OPTIMAL CONTROL OF STATIC ELASTOPLASTICITY IN PRIMAL FORMULATION

J. C. DE LOS REYES[‡], R. HERZOG[§], AND C. MEYER[¶]

Abstract. An optimal control problem of static plasticity with linear kinematic hardening and von Mises yield condition is studied. The problem is treated in its primal formulation, where the state system is a variational inequality of the second kind. First-order necessary optimality conditions are obtained by means of an approximation by a family of control problems with state system regularized by Huber-type smoothing, and a subsequent limit analysis. The equivalence of the optimality conditions with the C-stationarity system Herzog et al. [2012] for the equivalent dual formulation of the problem is proved. Numerical experiments are presented, which demonstrate the viability of the Huber-type smoothing approach.

Key words. optimal control, first-order necessary optimality conditions, mathematical program with equilibrium constraints (MPEC), variational inequality of the second kind, elastoplasticity

1. Introduction. We consider an optimal control problem for static small-strain elastoplasticity in its primal formulation. This model describes the deformation of a solid body under high loads, such that the yield stress is reached and permanent deformation ensues. Since elastoplastic deformation is the basis of many industrial production processes, its optimization is of significant importance. The static VI has only limited physical meaning, but can be regarded as time discretization of a corresponding quasi-static counterpart. The latter models elastoplastic deformation processes and thus appears in various industrial applications. When an instantaneous control strategy is applied to optimize or control such processes, then the static optimal control problem considered in this paper will arise.

We consider here a linear kinematic hardening model with von Mises yield condition and under the assumption of small strains. The description of the forward problem follows [Han and Reddy, 1999a, Chapter 12.2] where the quasistatic case is considered, see also Alberty et al. [1999]. The solid body $\Omega \subset \mathbb{R}^3$ is clamped on a non-vanishing Dirichlet part Γ_D of its boundary Γ, and it is subject to boundary loads on the remaining Neumann part Γ_N. The variables of the problem are the *displacement*

$$\boldsymbol{u} \in V := H^1_D(\Omega;\mathbb{R}^3) = \{\boldsymbol{u} \in H^1(\Omega;\mathbb{R}^3) : \boldsymbol{u} = \mathbf{0} \text{ on } \Gamma_D\}$$

and the *plastic strain*

$$\boldsymbol{p} \in Q := L^2(\Omega;\mathbb{Q}) = \{\boldsymbol{p} \in S : \operatorname{trace}(\boldsymbol{p}) = 0\}$$

where $S := L^2(\Omega;\mathbb{S})$ and $\mathbb{S} = \mathbb{R}^{3\times 3}_{\text{sym}}$ are spaces of symmetric matrices. By $\mathbb{Q} := \{\boldsymbol{q} \in \mathbb{S} : \operatorname{trace}(\boldsymbol{q}) = 0\}$ we denote the subspace of trace-free (deviatoric) symmetric matrices. All spaces are endowed with their natural inner products and norms. For $\mathbb{S}$ and $\mathbb{Q}$, this is the Frobenius norm, denoted by $|\boldsymbol{p}|$, and corresponding inner product $\boldsymbol{p} : \boldsymbol{q}$.

[‡]Research Center on Math. Modelling ModeMat, EPN Quito, Quito, Ecuador
[§]Chemnitz University of Technology, Department of Mathematics, Chemnitz, Germany.
[¶]TU Dortmund, Faculty of Mathematics, Dortmund, Germany.

Given $\ell \in V'$, the forward problem is to find $\boldsymbol{W} = (\boldsymbol{u}, \boldsymbol{p}) \in Z := V \times Q$ which satisfies the following variational inequality (VI) of the second kind.

$$a(\boldsymbol{W}, \boldsymbol{Y} - \boldsymbol{W}) + j(\boldsymbol{q}) - j(\boldsymbol{p}) \geq \langle \ell, \boldsymbol{v} - \boldsymbol{u} \rangle \quad \text{for all } \boldsymbol{Y} = (\boldsymbol{v}, \boldsymbol{q}) \in Z. \tag{1.1}$$

The forms a, j and ℓ in (1.1) are defined as follows:

$$a(\boldsymbol{W}, \boldsymbol{Y}) = \int_\Omega \left[(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{p}) : \mathbb{C}\,(\boldsymbol{\varepsilon}(\boldsymbol{v}) - \boldsymbol{q})\right] \mathrm{d}x + \int_\Omega \boldsymbol{p} : \mathbb{H}\,\boldsymbol{q}\,\mathrm{d}x \tag{1.2a}$$

$$j(\boldsymbol{p}) = \widetilde{\sigma}_0 \int_\Omega |\boldsymbol{p}|\,\mathrm{d}x \tag{1.2b}$$

$$\langle \ell, \boldsymbol{v} \rangle = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}\,\mathrm{d}x + \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{v}\,\mathrm{d}s. \tag{1.2c}$$

In (1.2a), $\mathbb{C}$ represents the material's fourth-order elasticity tensor, see e.g. [Han and Reddy, 1999a, Chapter 2.3], and $\mathbb{H}$ is the hardening modulus. The constant $\widetilde{\sigma}_0 > 0$ in (1.2b) denotes the material's yield stress. The data $\boldsymbol{f}$ and $\boldsymbol{g}$ in (1.2c) are volume and boundary loads, respectively. We remark that the second term of the energy form $a(\cdot, \cdot)$ and the specific choice of $j(\cdot)$ are characteristic for linear kinematic hardening.

We consider the following optimal control problem with control variables $(\boldsymbol{f}, \boldsymbol{g}) \in L^2(\Omega; \mathbb{R}^3) \times L^2(\Gamma_N; \mathbb{R}^3)$ and state variables $(\boldsymbol{u}, \boldsymbol{p}) \in V \times Q$.

$$\begin{aligned} \text{Minimize} \quad & J(\boldsymbol{u}, \boldsymbol{f}, \boldsymbol{g}) := \frac{1}{2}\|\boldsymbol{u} - \boldsymbol{u}_d\|^2_{L^2(\Omega;\mathbb{R}^3)} + \frac{\nu_1}{2}\|\boldsymbol{f}\|^2_{L^2(\Omega;\mathbb{R}^3)} + \frac{\nu_2}{2}\|\boldsymbol{g}\|^2_{L^2(\Gamma_N;\mathbb{R}^3)} \\ \text{s.t.} \quad & \text{the variational inequality (1.1).} \end{aligned} \tag{1.3}$$

One can prove by standard methods that there exists at least one global optimal solution of (1.3) so we do not discuss this in detail. The derivation of optimality conditions, however, is by no means standard since (1.3) constitutes a generalized mathematical program with equilibrium constraints (MPEC) in function space. Under assumptions made precise in Section 2, our main result is the following first-order necessary optimality system.

THEOREM 1.1. *Let $(\boldsymbol{f}, \boldsymbol{g}) \in L^2(\Omega; \mathbb{R}^3) \times L^2(\Gamma_N; \mathbb{R}^3)$ be a locally optimal solution for (1.3) with associated state $\boldsymbol{W} = (\boldsymbol{u}, \boldsymbol{p}) \in Z$. Then there exists an adjoint state $\boldsymbol{Z} = (\boldsymbol{w}, \boldsymbol{r}) \in Z$ and multipliers $\boldsymbol{\varrho} \in Q$, $\boldsymbol{\pi} \in Q$ and $\vartheta \in L^2(\Omega)$ such that the following optimality system is satisfied:*

$$a(\boldsymbol{W}, \boldsymbol{Y}) + \int_\Omega \boldsymbol{\varrho} : \boldsymbol{q}\,\mathrm{d}x = \langle \ell, \boldsymbol{v} \rangle \quad \textit{for all } \boldsymbol{Y} = (\boldsymbol{v}, \boldsymbol{q}) \in Z \tag{1.4a}$$

$$\boldsymbol{\varrho} : \boldsymbol{p} = \widetilde{\sigma}_0\,|\boldsymbol{p}|, \quad |\boldsymbol{\varrho}| \leq \widetilde{\sigma}_0 \quad \textit{a.e. in } \Omega \tag{1.4b}$$

$$a(\boldsymbol{Y}, \boldsymbol{Z}) + \int_\Omega \boldsymbol{\pi} : \boldsymbol{q}\,\mathrm{d}x = -\int_\Omega (\boldsymbol{u} - \boldsymbol{u}_d) \cdot \boldsymbol{v}\,\mathrm{d}x \quad \textit{for all } \boldsymbol{Y} = (\boldsymbol{v}, \boldsymbol{q}) \in Z \tag{1.4c}$$

$$\nu_1\,\boldsymbol{f} - \boldsymbol{w} = \boldsymbol{0} \quad \textit{a.e. in } \Omega, \quad \nu_2\,\boldsymbol{g} - \boldsymbol{w} = \boldsymbol{0} \quad \textit{a.e. in } \Gamma_N \tag{1.4d}$$

$$\boldsymbol{\pi} : \boldsymbol{p} = 0 \qquad \textit{a.e. in } \Omega \tag{1.4e}$$

$$\widetilde{\sigma}_0\,\boldsymbol{r} = |\boldsymbol{p}|\,\boldsymbol{\pi} + \vartheta\boldsymbol{\varrho} \qquad \textit{a.e. in } \Omega \tag{1.4f}$$

$$\vartheta\boldsymbol{\varrho} : \boldsymbol{\pi} \geq 0 \qquad \textit{a.e. in } \Omega \tag{1.4g}$$

$$\vartheta = 0 \qquad \textit{a.e. in } \mathcal{I} = \{x \in \Omega : |\boldsymbol{\varrho}(x)| < \widetilde{\sigma}_0\}. \tag{1.4h}$$

	space	state variables	test functions	adjoint variables
displacement	V	$\boldsymbol{u}$	$\boldsymbol{v}$	$\boldsymbol{w}$
plastic strain	Q	$\boldsymbol{p}$	$\boldsymbol{q}$	$\boldsymbol{r}$
joint variables	$Z = V \times Q$	$\boldsymbol{W} = (\boldsymbol{u}, \boldsymbol{p})$	$\boldsymbol{Y} = (\boldsymbol{v}, \boldsymbol{q})$	$\boldsymbol{Z} = (\boldsymbol{w}, \boldsymbol{r})$
		dual variables		
lower level	Q	$\boldsymbol{\varrho} \in \widetilde{\sigma}_0 \, \partial \lvert \boldsymbol{p} \rvert$		
upper level	Q	$\boldsymbol{\pi}$		
upper level	$L^2(\Omega)$	ϑ		
		control variables		
volume force	$L^2(\Omega; \mathbb{R}^3)$	$\boldsymbol{f}$		
traction force	$L^2(\Gamma_N; \mathbb{R}^3)$	$\boldsymbol{g}$		
		constant		
yield stress	$\mathbb{R}$	$\widetilde{\sigma}_0$		

TABLE 1.1
Variables in the primal optimal control problem (1.3) *and optimality system* (1.4).

Note that (1.4b) can be equivalently expressed as the pointwise relation $\boldsymbol{\varrho} \in \widetilde{\sigma}_0 \, \partial \lvert \boldsymbol{p} \rvert$.

This set of optimality conditions is proved by considering a family of regularized problems obtained by a Huber-type smoothing of the functional j in (1.2b). The individual conditions in (1.4) will be detailed later on in the paper, along with the proof. Due to the wealth of notation involved in (1.4), we give an overview over all variables in Table 1.1.

Let us make some general comments on the result of Theorem 1.1. First-order necessary optimality conditions such as (1.4) are of significant importance in practice. They are, among other uses, key to efficient solution algorithms for (1.3), as well as error estimates. Despite their importance, there does not seem to be a classification scheme for optimality conditions pertaining to MPECs which involve variational inequalities of the second kind. There are rather few problems treated in the literature, and each comes with its own specific set of optimality conditions, see see Wenbin and Rubio [1991], Bonnans and Tiba [1991], Bonnans and Casas [1995], de los Reyes [2011, 2012]. A comparison of the strengths of these conditions between problems seems to be lacking.

The picture is much more complete for problems involving VIs of the first kind, which can be written in terms of an equivalent complementarity system. Such optimization problems are also refered to as MPCCs (mathematical programs with complementarity constraints). The concept of *strong stationarity* is the most rigorous among first-order optimality conditions, and it can be expected to hold at local minima under suitable constraint qualifications. We refer to Flegel and Kanzow [2003] for a treatment in the finite dimensional case. Under less restrictive assumptions, weaker stationarity conditions can be shown. We refer to Scheel and Scholtes [2000], Hoheisel et al. [2013] for an overview.

Among the various notions of stationarity for optimization problems with VIs of the first kind, *C-stationarity* plays a prominent role. This concept of intermediate strength is generically obtained when a limit process for a sequence of regularized problems is considered. We refer to Hintermüller and Kopacka [2009], Herzog et al. [2012] as examples for infinite dimensional problems.

The problem of elastoplasticity considered here is special in the sense that equivalent formulations of the forward problem exist, either as a VI of the first kind, or of the second kind. Optimal control problems with the former, also called dual or stress-based formulation, were considered in Herzog et al. [2012], and optimality conditions of C-stationary and other types were derived there.

As a second result of this paper, we show that the optimality system (1.4) is precisely equivalent to the C-stationarity conditions for the problem obtained by replacing in (1.3) the primal (strain-based) formulation (1.1) by the corresponding dual (stress-based) VI. To our best knowledge, this is the first time that an optimality system for an optimization problem of a VI of the second kind is being classified in this sense.

To put our contributions into perspective, we wish to point out that the investigation of optimal control of primal elastoplasticity is not only of interest due to the comparison with the dual formulation. The primal formulation also suggests an alternative way of smoothing (the Huber-type regularization already mentioned), which appears to be competitive compared to the regularization approaches used in case of the dual formulation. This is demonstrated by preliminary numerical experiments presented in Section 8. Moreover, there are numerous constitutive laws that can be only formulated through primal variables, as for instance in case of thermoplasticity, see e.g. Bartels and Roubíček [2008]. Our work thus lays the foundations for the derivation of optimality conditions for such cases.

The outline of the paper is as follows. Sections 2 and 3 collect our standing assumptions, as well as some known facts concerning the forward problem (1.1). Sections 4 and 5 are devoted to the study of regularized optimal control problems, which are obtained by approximating the VI (1.1) by an equation. A Huber-type smoothing of $j(\cdot)$ is used for this purpose. We point out that an improved integrability result for $(\boldsymbol{u}, \boldsymbol{p})$ based on Herzog et al. [2011a] plays an essential role in the Fréchet differentiability of the regularized forward problem, see Theorem 5.2. The first-order optimality system for the regularized case is given in Theorem 6.1. In Section 6 we pass to the limit to prove Theorem 1.1. The equivalence of the optimality conditions (1.4) with the system of C-stationarity of the problem involving the corresponding dual formulation is shown in Section 7. Finally, Section 8 reports on some numerical experiments based on the proposed regularization of (1.3).

Notation. We shall use the short hand notation $(\cdot,\cdot)_\Omega$ to denote the standard L^2 inner product in spaces such as $L^2(\Omega)$, $L^2(\Omega;\mathbb{R}^3)$ and $L^2(\Omega;\mathbb{S})$. Similarly, $(\cdot,\cdot)_{\Gamma_N}$ denotes the L^2 inner product of functions defined on Γ_N. Besides the space $H^1_D(\Omega;\mathbb{R}^3)$ already defined, we will also use the more general Sobolev spaces

$$W^{1,p}_D(\Omega;\mathbb{R}^3) = \{\boldsymbol{u} \in W^{1,p}(\Omega;\mathbb{R}^3) : \boldsymbol{u} = \boldsymbol{0} \text{ on } \Gamma_D\}$$

for values of $p \in [1,\infty]$. We will denote the conjugate exponent of p by p' and write $W^{-1,p}_D(\Omega;\mathbb{R}^3)$ to denote the dual space of $W^{1,p'}_D(\Omega;\mathbb{R}^3)$. The dual of a normed linear space X is denoted as X' and the duality pairing by $\langle\cdot,\cdot\rangle$.

2. Standing assumptions.

Our first assumption concerns the domain Ω.

Assumption 2.1 (Domain and its boundary).

(a) *The boundary Γ of the domain $\Omega \subset \mathbb{R}^3$ is Lipschitz, i.e., the boundary consists of a finite number of local graphs of Lipschitz maps, see, e.g., [Grisvard, 1985, Definition 1.2.1.1].*
(b) *Moreover, the boundary is assumed to consist of two disjoint measurable parts Γ_N and Γ_D such that $\Gamma = \Gamma_N \cup \Gamma_D$. While Γ_N is relatively open, Γ_D is a relatively closed subset of Γ. Furthermore Γ_D is assumed to have positive measure.*
(c) *In addition, the set $\Omega \cup \Gamma_N$ is regular in the sense of Gröger [1989].*

The class of domains fulfilling Assumption 2.1 covers a wide range of geometries. In particular, a characterization of Gröger regular domains can be found in [Haller-Dintelmann et al., 2009, Section 5]. We make this assumption in order to apply the integrability results in Herzog et al. [2011a] pertaining to systems of nonlinear elasticity, which leads to Theorem 5.2.

Assumption 2.2 (Elasticity and hardening tensors). *The tensor-valued functions $\mathbb{C}$ and $\mathbb{H}$ are elements of $L^\infty(\Omega; \mathcal{L}(\mathbb{S}))$, where $\mathcal{L}(\mathbb{S})$ denotes the space of linear operators $\mathbb{S} \to \mathbb{S}$. Both $\mathbb{C}(x)$ and $\mathbb{H}(x)$ are assumed to be uniformly bounded and coercive with coercivity constants $\underline{c} > 0$ and $\underline{h} > 0$, respectively. That is, for all $\boldsymbol{\varepsilon} \in \mathbb{S}$ and almost all $x \in \Omega$ there holds*

$$\boldsymbol{\varepsilon} : \mathbb{C}(x)\, \boldsymbol{\varepsilon} \geq \underline{c}\, |\boldsymbol{\varepsilon}|^2 \quad \textit{and} \quad \boldsymbol{p} : \mathbb{H}(x)\, \boldsymbol{p} \geq \underline{h}\, |\boldsymbol{p}|^2.$$

Moreover, we assume as usual that $\mathbb{C}$ and $\mathbb{H}$ are symmetric, i.e.,

$$\mathbb{C}_{ijkl} = \mathbb{C}_{jikl} = \mathbb{C}_{klij}. \tag{2.1}$$

and similarly for $\mathbb{H}$.

In what follows, we abbreviate

$$\|\mathbb{C}\|_{L^\infty(\Omega;\mathcal{L}(\mathbb{S}))} =: \overline{c} \quad \text{and} \quad \|\mathbb{H}\|_{L^\infty(\Omega;\mathcal{L}(\mathbb{S}))} =: \overline{h}.$$

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

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

with Lamé constants satisfying $\mu_L > 0$ and $3\,\lambda_L + 2\,\mu_L > 0$. In this case, we have $\underline{c} = \min(3\,\lambda_L + 2\,\mu_L, 2\,\mu_L)$. A standard example for the hardening modulus is $\mathbb{H} = k_1\,\mathbb{I}$, with the fourth-order identity tensor $\mathbb{I}$ and a constant $k_1 > 0$, hence $\underline{h} = k_1$ holds.

As in Herzog and Meyer [2011], we introduce the linear and compact operator $R : L^2(\Omega; \mathbb{R}^3) \times L^2(\Gamma_N; \mathbb{R}^3) \to V'$ by

$$\langle R(\boldsymbol{f}, \boldsymbol{g}),\, \boldsymbol{v}\rangle := \langle \ell,\, \boldsymbol{v}\rangle.$$

3. Known results. First we address the solvability of (1.1). Note that $a(\cdot,\cdot)$ is clearly bounded and also coercive on Z, since Young's inequality implies

$$\begin{aligned} a(\boldsymbol{W}, \boldsymbol{W}) &= \int_\Omega \big[(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{p}) : \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{p}) + \boldsymbol{p} : \mathbb{H}\,\boldsymbol{p}\big]\,\mathrm{d}x \\ &\geq \underline{c}\,\|\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{p}\|_S^2 + \underline{h}\,\|\boldsymbol{p}\|_S^2 \\ &\geq \underline{c}\,(1-\kappa)\,\|\boldsymbol{\varepsilon}(\boldsymbol{u})\|_S^2 + \big(\underline{c}\,(1 - 1/\kappa) + \underline{h}\big)\|\boldsymbol{p}\|_S^2 \end{aligned} \tag{3.1}$$

for any $\kappa > 0$. Any choice of κ subject to $\underline{c}/(\underline{c}+\underline{h}) < \kappa < 1$, together with Korn's inequality ([Temam, 1983, Proposition 1.1]) then gives the ellipticity of a:

$$a(\boldsymbol{W}, \boldsymbol{W}) \geq \underline{a}\left(\|\boldsymbol{u}\|_V^2 + \|\boldsymbol{p}\|_S^2\right) \tag{3.2}$$

for some positive constant $\underline{a}$. Moreover, in the considered case of linear kinematic hardening, j is convex and finite on all of $Z = V \times Q$. Therefore, existence and uniqueness for (1.1) follow by standard arguments, see e.g. [Han and Reddy, 1999b, Theorem 6.6]:

Lemma 3.1. *For every $\ell \in V'$, there is a unique solution $(\boldsymbol{u}, \boldsymbol{p}) \in V \times Q$ of (1.1).*

While we treat primarily the primal formulation of static plasticity in this paper, it is useful to recall the dual formulation as well. In place of the plastic strain $\boldsymbol{p}$, the dual formulation uses the stress and backstress $(\boldsymbol{\sigma}, \boldsymbol{\chi})$, which are confined to a feasible set $\mathcal{K}$. In our case of the von Mises yield condition, $\mathcal{K}$ is defined as

$$\begin{aligned} &\mathcal{K} := \{(\boldsymbol{\tau}, \boldsymbol{\mu}) \in S \times S : (\boldsymbol{\tau}(x), \boldsymbol{\mu}(x)) \in K \text{ a.e. in } \Omega\} \\ \text{and} \quad &K := \{(\boldsymbol{\tau}, \boldsymbol{\mu}) \in \mathbb{S} \times \mathbb{S} : |\boldsymbol{\tau}^D + \boldsymbol{\mu}^D| \leq \widetilde{\sigma}_0\}. \end{aligned} \tag{3.3}$$

Here and throughout,

$$\boldsymbol{\tau}^D = \boldsymbol{\tau} - (1/3)\,\text{trace}(\boldsymbol{\tau})\, I \tag{3.4}$$

denotes the deviatoric (trace-free) part of the matrix $\boldsymbol{\tau} \in \mathbb{R}^{3\times 3}$, and I is the identity matrix.

Lemma 3.2. *Problem (1.1) is equivalent to the dual problem of the following VI of first kind in mixed form: given $\ell \in V'$, find $\boldsymbol{u} \in V$ and $(\boldsymbol{\sigma}, \boldsymbol{\chi}) \in \mathcal{K}$ such that*

$$\begin{aligned} \int_\Omega \boldsymbol{\sigma} : \mathbb{C}^{-1}(\boldsymbol{\tau} - \boldsymbol{\sigma})\,\mathrm{d}x + \int_\Omega \boldsymbol{\chi} : \mathbb{H}^{-1}(\boldsymbol{\mu} - \boldsymbol{\chi})\,\mathrm{d}x \\ - \int_\Omega \boldsymbol{\varepsilon}(\boldsymbol{u}) : (\boldsymbol{\tau} - \boldsymbol{\sigma})\,\mathrm{d}x \geq 0 \quad \textit{for all } (\boldsymbol{\tau}, \boldsymbol{\mu}) \in \mathcal{K} \end{aligned} \tag{3.5a}$$

$$\int_\Omega \boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\sigma}\,\mathrm{d}x = \langle \ell,\, \boldsymbol{v}\rangle \quad \textit{for all } \boldsymbol{v} \in V. \tag{3.5b}$$

The equivalence holds in the following sense: Problem (3.5) admits a unique solution $(\boldsymbol{\sigma}, \boldsymbol{\chi}, \boldsymbol{u}) \in S \times S \times V$. It is related to the unique solution $(\boldsymbol{u}, \boldsymbol{p})$ of (1.1) via

$$\boldsymbol{\sigma} = \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{p}) \quad \textit{and} \quad \boldsymbol{\chi} = -\mathbb{H}\,\boldsymbol{p}.$$

The proof of Lemma 3.2 is based on classical arguments, cf. e.g. [Han and Reddy, 1999b, Section 8.1] and it is given in Section A in the appendix.

In the sequel, we will frequently invoke the Browder-Minty theorem on monotone operators. We recall here a version for *strongly* monotone operators, which will be sufficient for our purposes.

Theorem 3.3 (Browder-Minty Theorem for strongly monotone operators). *Suppose that X is a separable Hilbert space and that $\mathcal{M} : X \to X'$ is strongly monotone, i.e.,*

$$\langle \mathcal{M}\,x - \mathcal{M}\,y,\, x - y\rangle \geq \underline{m}\,\|x - y\|_X^2 \tag{3.6}$$

with some $\underline{m} > 0$, and hemicontinuous, i.e., $[0,1] \ni t \mapsto \langle \mathcal{M}(x+t\,y),\, z\rangle$ is continuous. Then $\mathcal{M}$ is invertible, and $\mathcal{M}^{-1}$ is Lipschitz continuous with constant $1/\underline{m}$, i.e.,

$$\|\mathcal{M}^{-1} F - \mathcal{M}^{-1} G\|_X \le \frac{1}{\underline{m}} \|F - G\|_{X'}. \tag{3.7}$$

If in addition $\mathcal{M}$ is Lipschitz with constant $L_{\mathcal{M}}$, then $\mathcal{M}^{-1}$ is also strongly monotone, viz.

$$\langle \mathcal{M}^{-1} F - \mathcal{M}^{-1} G,\, F - G\rangle \ge \frac{\underline{m}}{L_{\mathcal{M}}^2}\, \|F - G\|_{X'}^2. \tag{3.8}$$

The theorem and its proof can be found in [Zeidler, 1990, Theorem 26.A]. The strong monotonicity of $\mathcal{M}^{-1}$ follows easily from

$$\|F - G\|_{X'}^2 = \|\mathcal{M}\mathcal{M}^{-1}F - \mathcal{M}\mathcal{M}^{-1}G\|^2 \le L_{\mathcal{M}}^2 \|\mathcal{M}^{-1}F - \mathcal{M}^{-1}G\|_X^2$$

and the subsequent application of (3.6). In all our applications of Theorem 3.3, the hemicontinuity will be superseded by continuity of $\mathcal{M}$.

4. Regularized control problems.

Following the road map in de los Reyes [2011] we consider the following family of regularized optimal control problems:

$$\text{Minimize} \quad \frac{1}{2}\|\boldsymbol{u} - \boldsymbol{u}_d\|_{L^2(\Omega;\mathbb{R}^3)}^2 + \frac{\nu_1}{2}\|\boldsymbol{f}\|_{L^2(\Omega;\mathbb{R}^3)}^2 + \frac{\nu_2}{2}\|\boldsymbol{g}\|_{L^2(\Gamma_N;\mathbb{R}^3)}^2 \tag{4.1a}$$

$$\text{s.t.} \quad a(\boldsymbol{W},\boldsymbol{Y}) + \int_\Omega h_\gamma(\boldsymbol{p}) : \boldsymbol{q}\,\mathrm{d}x = \langle \ell,\, \boldsymbol{v}\rangle \quad \text{for all } \boldsymbol{Y} = (\boldsymbol{v},\boldsymbol{q}) \in Z. \tag{4.1b}$$

The function

$$h_\gamma(\boldsymbol{p}) := \widetilde{\sigma}_0\, \gamma\, \frac{\boldsymbol{p}}{m_\gamma(|\boldsymbol{p}|)} \tag{4.2}$$

is the derivative of a Huber-type regularization (see [Huber, 1981, eq. (7.14)] or [Huber, 1973, eq. (1.6)]) of j with the following local smoothing of the function $\mathbb{R} \ni p \mapsto \max(0,p) \in \mathbb{R}$, parametrized by $\gamma > 0$:

$$m_\gamma(p) := \begin{cases} \gamma\, p, & \text{if } \gamma\, p \ge \widetilde{\sigma}_0 + \frac{1}{2\gamma} \\ \widetilde{\sigma}_0 + \frac{\gamma}{2}\big(\gamma\, p - \widetilde{\sigma}_0 + \frac{1}{2\gamma}\big)^2, & \text{if } |\gamma\, p - \widetilde{\sigma}_0| \le \frac{1}{2\gamma} \\ \widetilde{\sigma}_0, & \text{if } \gamma\, p \le \widetilde{\sigma}_0 - \frac{1}{2\gamma}, \end{cases} \tag{4.3}$$

The functions m_γ and h_γ are displayed in Figure 4.1.

In this section, we next address some properties of m_γ and h_γ in Lemma 4.1. We then deduce the global Lipschitz continuity of the solution map for the regularized state equation (4.1b) in Theorem 4.2. Theorem 4.3 is devoted to the convergence of the solution of (4.1b) to that of the original VI (1.1) as $\gamma \to \infty$. Finally, Theorem 4.4 addresses the convergence of (global) minimizers of the regularized control problem (4.1) to those of the original problem (1.3).

Lemma 4.1. *For all $p \in \mathbb{R}$ and $\gamma > 0$ there holds*

$$\frac{\widetilde{\sigma}_0}{m_\gamma(p)} \le 1, \quad \frac{\gamma\, p}{m_\gamma(p)} \le 1, \tag{4.4}$$

$$\text{and} \quad 0 \le m_\gamma(p) - \max(\widetilde{\sigma}_0, \gamma\, p) \le \frac{1}{2\gamma}. \tag{4.5}$$

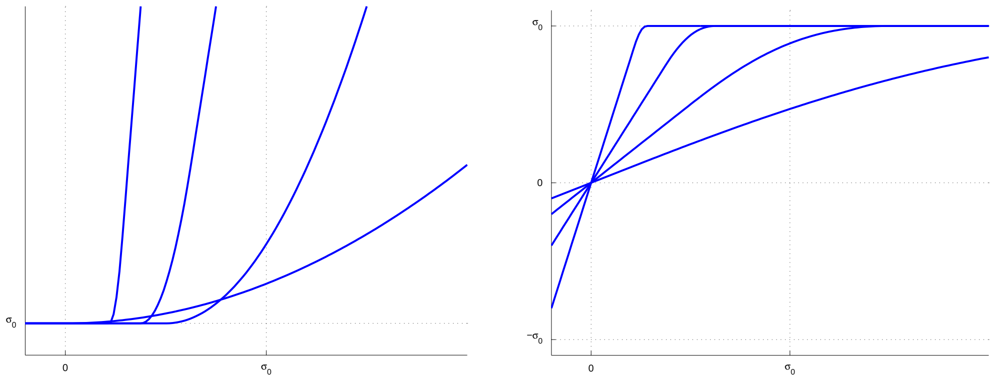

FIG. 4.1. *The function* $p \mapsto m_\gamma(p)$ *for* $\gamma \in \{0.5, 1, 2, 4\}$ *(left), and* $|\boldsymbol{p}| \mapsto |h_\gamma(\boldsymbol{p})|$ *(right).*

Moreover, $h_\gamma : \mathbb{S} \to \mathbb{S}$ *is bounded and continuously differentiable with*

$$h'_\gamma(\boldsymbol{p}) = \widetilde{\sigma}_0\, \gamma \left(\frac{1}{m_\gamma(|\boldsymbol{p}|)}\, \mathbb{I} - \frac{m'_\gamma(|\boldsymbol{p}|)}{m_\gamma(|\boldsymbol{p}|)^2\, |\boldsymbol{p}|} \boldsymbol{p} \otimes \boldsymbol{p} \right), \tag{4.6}$$

where $\mathbb{I} : \mathbb{S} \to \mathbb{S}$ *denotes the fourth-order identity tensor and* $\otimes : \mathbb{S} \times \mathbb{S} \to \mathcal{L}(\mathbb{S})$ *is the dyadic product, i.e.,* $(\boldsymbol{p} \otimes \boldsymbol{q})_{ijkl} = \boldsymbol{q}_{ij}\, \boldsymbol{p}_{kl}$. *The derivative of* m_γ *is given by*

$$m'_\gamma(p) = \begin{cases} \gamma, & \text{if } \gamma\, p \geq \widetilde{\sigma}_0 + \frac{1}{2\gamma}, \\ \gamma^2 \big(\gamma\, p - \widetilde{\sigma}_0 + \frac{1}{2\gamma}\big), & \text{if } |\gamma\, p - \widetilde{\sigma}_0| \leq \frac{1}{2\gamma}, \\ 0, & \text{if } \gamma\, p \leq \widetilde{\sigma}_0 - \frac{1}{2\gamma}. \end{cases} \tag{4.7}$$

Furthermore, there holds

$$0 \leq m'_\gamma(p) \leq \gamma \quad \text{for all } p \in \mathbb{R}, \tag{4.8}$$

$$\|h'_\gamma(\boldsymbol{p})\|_{\mathcal{L}(\mathbb{S})} \leq 2\, \gamma \quad \text{for all } \boldsymbol{p} \in \mathbb{S}, \tag{4.9}$$

$$\boldsymbol{q} : h'_\gamma(\boldsymbol{p})\, \boldsymbol{q} \geq 0 \quad \text{for all } \boldsymbol{p}, \boldsymbol{q} \in \mathbb{S}, \tag{4.10}$$

which shows that $h_\gamma : \mathbb{S} \to \mathbb{S}$ *is a globally Lipschitz and monotone operator.*

Proof. For values of p such that $\gamma\, p \leq \widetilde{\sigma}_0 - 1/(2\gamma)$ or $\gamma\, p \geq \widetilde{\sigma}_0 + 1/(2\gamma)$, (4.4) is evident by the definition (4.3) of m_γ. For all p satisfying $\widetilde{\sigma}_0 - 1/(2\gamma) < \gamma\, p < \widetilde{\sigma}_0 + 1/(2\gamma)$, (4.3) implies

$$m_\gamma(p) = \widetilde{\sigma}_0 + \underbrace{\frac{\gamma}{2} \Big(\gamma\, p - \widetilde{\sigma}_0 + \frac{1}{2\gamma}\Big)^2}_{\geq 0} \geq \widetilde{\sigma}_0.$$

Moreover, for the right hand boundary $\gamma\, p = \widetilde{\sigma}_0 + 1/(2\gamma)$, we get $m_\gamma(p) = \gamma\, p$. Since the derivative

$$m'_\gamma(p) = \gamma^2 \left(\gamma\, p - \widetilde{\sigma}_0 + \frac{1}{2\gamma}\right)$$

is easily seen to be $\leq \gamma$ on the interval under consideration, we can conclude $m_\gamma(p) \geq \gamma\, p$, i.e., (4.4) holds.

Inequality (4.5) holds trivially for p such that $\gamma\, p \le \widetilde{\sigma}_0 - 1/(2\gamma)$ or $\gamma\, p \ge \widetilde{\sigma}_0 + 1/(2\gamma)$ since the middle term is then identically zero. In case $\widetilde{\sigma}_0 - 1/(2\gamma) < \gamma\, p < \widetilde{\sigma}_0 + 1/(2\gamma)$, the first inequality in (4.5) follows from (4.4). For the second inequality, we estimate

$$m_\gamma(p) - \max(\widetilde{\sigma}_0, \gamma\, p) = \underbrace{\widetilde{\sigma}_0 - \max(\widetilde{\sigma}_0, \gamma\, p)}_{\le 0} + \frac{\gamma}{2}\Big(\gamma\, p - \widetilde{\sigma}_0 + \frac{1}{2\gamma}\Big)^2 \le \frac{1}{2\gamma},$$

which completes the proof of (4.5).

In view of (4.4) we have $|h_\gamma(\boldsymbol{p})| = \widetilde{\sigma}_0\, \gamma\, \frac{|\boldsymbol{p}|}{m_\gamma(|\boldsymbol{p}|)} \le \widetilde{\sigma}_0$ for all $\boldsymbol{p} \in \mathbb{S}$ so that h_γ is indeed bounded. The continuous differentiability of h_γ and m_γ are easily verified, i.e., (4.6) and (4.7) hold. To prove (4.8) we observe that, in case $|\gamma\, p - \widetilde{\sigma}_0| \le \frac{1}{2\gamma}$,

$$0 \le m_\gamma'(p) = \gamma^2\Big(\gamma\, p - \widetilde{\sigma}_0 + \frac{1}{2\gamma}\Big) \le \gamma \tag{4.11}$$

holds. If $\big|\gamma\, p - \widetilde{\sigma}_0\big| > \frac{1}{2\gamma}$, (4.8) is trivially fulfilled. Moreover, (4.9) follows directly from (4.8), (4.4), and (4.6):

$$\begin{aligned} h_\gamma'(\boldsymbol{p}) &= \widetilde{\sigma}_0\, \gamma\, \left(\frac{1}{m_\gamma(|\boldsymbol{p}|)}\, \mathbb{I} - \frac{m_\gamma'(|\boldsymbol{p}|)}{m_\gamma(|\boldsymbol{p}|)^2\, |\boldsymbol{p}|} \boldsymbol{p} \otimes \boldsymbol{p}\right) \\ \Rightarrow \quad \|h_\gamma'(\boldsymbol{p})\| &\le \widetilde{\sigma}_0\, \gamma\, \left(\frac{1}{\widetilde{\sigma}_0} + \frac{\gamma}{\widetilde{\sigma}_0\, \gamma\, |\boldsymbol{p}|^2} |\boldsymbol{p}|^2\right) = 2\,\gamma. \end{aligned}$$

It remains to prove (4.10). In case $\gamma\, |\boldsymbol{p}| \le \widetilde{\sigma}_0 - 1/(2\gamma)$, one obtains

$$\boldsymbol{q} : h_\gamma'(\boldsymbol{p})\, \boldsymbol{q} = \widetilde{\sigma}_0\, \gamma\, \frac{1}{m_\gamma(|\boldsymbol{p}|)} |\boldsymbol{q}|^2 = \gamma\, |\boldsymbol{q}|^2 \ge 0 \quad \text{for all } \boldsymbol{q} \in \mathbb{S}.$$

If $\gamma\, |\boldsymbol{p}| \ge \widetilde{\sigma}_0 + 1/(2\gamma)$, then (4.4) yields

$$\begin{aligned} \boldsymbol{q} : h_\gamma'(\boldsymbol{p})\, \boldsymbol{q} &= \widetilde{\sigma}_0\, \gamma\, \left(\frac{1}{m_\gamma(|\boldsymbol{p}|)} |\boldsymbol{q}|^2 - \gamma\, \frac{(\boldsymbol{p} : \boldsymbol{q})^2}{m_\gamma(|\boldsymbol{p}|)^2}\, \frac{1}{|\boldsymbol{p}|}\right) \\ &\ge \widetilde{\sigma}_0\, \gamma\, \left(\frac{1}{m_\gamma(|\boldsymbol{p}|)} - \frac{1}{m_\gamma(|\boldsymbol{p}|)}\, \frac{\gamma\, |\boldsymbol{p}|}{m_\gamma(|\boldsymbol{p}|)}\right) |\boldsymbol{q}|^2 = 0 \quad \text{for all } \boldsymbol{q} \in \mathbb{S}. \end{aligned}$$

In view of (4.11), a similar argument as above implies $\boldsymbol{q} : h_\gamma'(\boldsymbol{p})\, \boldsymbol{q} \ge 0$ whenever $\widetilde{\sigma}_0 - 1/(2\gamma) \le \gamma\, |\boldsymbol{p}| \le \widetilde{\sigma}_0 + 1/(2\gamma)$. ☐

The definition of m_γ implies that the Nemyzki operators associated with m_γ and m_γ' map $L^p(\Omega)$ into $L^p(\Omega)$ for every $p \in [1, \infty]$. Furthermore, since $h_\gamma : \mathbb{S} \to \mathbb{S}$ is bounded, it follows immediately that the associated Nemyzki operator maps $S = L^2(\Omega; \mathbb{S})$ into $L^\infty(\Omega; \mathbb{S})$. To simplify notation, we denote this Nemyzki operator by the same symbol. By standard arguments, this Nemyzki operator inherits the monotonicity and Lipschitz properties from $h_\gamma : \mathbb{S} \to \mathbb{S}$. We refer to Goldberg et al. [1992] or [Tröltzsch, 2010, Section 4.3]. An immediate consequence is the following result.

Theorem 4.2. *For each $\ell \in V'$ the regularized equation (4.1b) admits a unique solution $\boldsymbol{W}_\gamma = (\boldsymbol{u}_\gamma, \boldsymbol{p}_\gamma) \in Z$. The associated solution operator $G_\gamma : V' \ni \ell \mapsto \boldsymbol{W}_\gamma \in Z$ is Lipschitz continuous with a Lipschitz constant independent of γ.*

The proof based on the theory of monotone operators is given in Section B in the appendix. The following theorem uses arguments similar to [de los Reyes, 2012, Theorem 3.3]. Its proof is given in Section C in the appendix.

THEOREM 4.3. *Let $\ell \in V'$ be given and denote by $\boldsymbol{W}_\gamma = (\boldsymbol{u}_\gamma, \boldsymbol{p}_\gamma) \in Z$ the unique solution of (4.1b). The sequence $\{\boldsymbol{W}_\gamma\}_{\gamma>0}$ converges strongly in Z to the unique solution $\boldsymbol{W}$ of (1.1) as $\gamma \to \infty$.*

We can now address the convergence of minimizers of the regularized control problem (4.1) to those of the original problem (1.3). The proof follows along the lines of [de los Reyes, 2012, Theorem 3.5]. Nevertheless we briefly recall the arguments adapted to the present setting in the following theorem.

THEOREM 4.4.

(a) For each $\gamma > 0$, there exists a globally optimal solution for problem (4.1). Moreover, every sequence $\{(\boldsymbol{f}_\gamma, \boldsymbol{g}_\gamma)\}_{\gamma>0}$ of global solutions to (4.1) contains a weakly convergent subsequence. Any weak accumulation point is a globally optimal solution for (1.3).

(b) In case the global solution $(\boldsymbol{f}^*, \boldsymbol{g}^*) \in L^2(\Omega;\mathbb{R}^3) \times L^2(\Gamma_N;\mathbb{R}^3)$ of (1.3) is unique*, then $(\boldsymbol{f}_\gamma, \boldsymbol{g}_\gamma) \to (\boldsymbol{f}^*, \boldsymbol{g}^*)$ strongly in $L^2(\Omega;\mathbb{R}^3) \times L^2(\Gamma_N;\mathbb{R}^3)$ as $\gamma \to \infty$.*

Proof. Let $\gamma > 0$ be arbitrary but fixed. It follows from the compactness of $R : L^2(\Omega;\mathbb{R}^3) \times L^2(\Gamma_N;\mathbb{R}^3) \to V'$ and the (Lipschitz) continuity of $G_\gamma : V' \to V \times Q$ that G_γ is weakly-strongly continuous. Due to the quadratic structure of the objective functional, the existence of a global minimizer for (4.1) is a standard result.

To prove convergence, consider a sequence $\{\boldsymbol{W}_\gamma, \boldsymbol{f}_\gamma, \boldsymbol{g}_\gamma\} \in Z \times L^2(\Omega;\mathbb{R}^3) \times L^2(\Gamma_N;\mathbb{R}^3)$ of global minimizers to (4.1) for $\gamma \to \infty$. Note first that $(\boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0})$ is feasible for each regularized VI (4.1b). It then follows that

$$J(\boldsymbol{W}_\gamma, \boldsymbol{f}_\gamma, \boldsymbol{g}_\gamma) \le J(\boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}) \quad \text{for all } \gamma > 0.$$

Thanks to the structure of the cost functional, this implies that $\{(\boldsymbol{f}_\gamma, \boldsymbol{g}_\gamma)\}$ is bounded in $L^2(\Omega;\mathbb{R}^3) \times L^2(\Gamma_N;\mathbb{R}^3)$. Consequently, there exists a subsequence (denoted in the same way) and a weak accumulation point $(\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{g}})$ such that

$$(\boldsymbol{f}_\gamma, \boldsymbol{g}_\gamma) \rightharpoonup (\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{g}}) \quad \text{weakly in } L^2(\Omega;\mathbb{R}^3) \times L^2(\Gamma_N;\mathbb{R}^3).$$

And thus, again by compactness, $R(\boldsymbol{f}_\gamma, \boldsymbol{g}_\gamma) \to R(\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{g}})$ strongly in V'. Theorems 4.2 and 4.3 imply for the associated states $\boldsymbol{W}_\gamma$ and $\widehat{\boldsymbol{W}}$

$$\begin{aligned}
\|\boldsymbol{W}_\gamma - \widehat{\boldsymbol{W}}\|_Z &= \|G_\gamma(\boldsymbol{f}_\gamma, \boldsymbol{g}_\gamma) - G(\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{g}})\|_Z \\
&\le \|G_\gamma(\boldsymbol{f}_\gamma, \boldsymbol{g}_\gamma) - G_\gamma(\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{g}})\|_Z + \|G_\gamma(\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{g}}) - G(\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{g}})\|_Z \\
&\le L\,\|(\boldsymbol{f}_\gamma, \boldsymbol{g}_\gamma) - (\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{g}})\|_{L^2(\Omega;\mathbb{R}^3)\times L^2(\Gamma_N;\mathbb{R}^3)} \\
&\quad + \|G_\gamma(\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{g}}) - G(\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{g}})\|_Z \longrightarrow 0 \quad \text{as } \gamma \to \infty.
\end{aligned}$$

We recall that $L > 0$ is the Lipschitz constant of G_γ, which is independent of γ by Theorem 4.2. Now let $(\boldsymbol{W}^*, \boldsymbol{f}^*, \boldsymbol{g}^*)$ be an arbitrary global minimizer for (1.3). From the weak lower semicontinuity of the objective functional, we finally obtain that

$$J(\widehat{\boldsymbol{W}}, \widehat{\boldsymbol{f}}, \widehat{\boldsymbol{g}}) \le \liminf_{\gamma\to\infty} J(\boldsymbol{W}_\gamma, \boldsymbol{f}_\gamma, \boldsymbol{g}_\gamma) \le \liminf_{\gamma\to\infty} J(G_\gamma(\boldsymbol{f}^*, \boldsymbol{g}^*), \boldsymbol{f}^*, \boldsymbol{g}^*) = J(\boldsymbol{W}^*, \boldsymbol{f}^*, \boldsymbol{g}^*),$$

so (a) is proved. If $(\boldsymbol{f}^*, \boldsymbol{g}^*)$ is unique, the strong convergence in part (b) follows from weak convergence together with convergence of the norm. ☐

Remark 4.5. *The question arises, which minima of the original problem (1.3) can be approximated by minima of the regularized problems (4.1)? So far, we merely know from Theorem 4.4 that this is true for* one of the global minimizers *of (1.3). However, it is well known that with a slight modification of the regularized problems (4.1), this result can be sharpened so that* every local minimum *of (1.3) can be approximated.*

To be more precise, let $(\boldsymbol{f}^*, \boldsymbol{g}^*)$ be an arbitrary local minimum of (1.3). When the term

$$\frac{r}{2}\,\|\boldsymbol{f}-\boldsymbol{f}^*\|^2_{L^2(\Omega;\mathbb{R}^3)} + \frac{r}{2}\,\|\boldsymbol{g}-\boldsymbol{g}^*\|^2_{L^2(\Gamma_N;\mathbb{R}^3)} \tag{4.12}$$

with sufficiently large $r > 0$ is added to the objective in (4.1a), then it can be shown that a sequence of solutions of these modified regularized problems exists which converges strongly *in $L^2(\Omega;\mathbb{R}^3) \times L^2(\Gamma_N;\mathbb{R}^3)$ to $(\boldsymbol{f}^*, \boldsymbol{g}^*)$. This technique goes back to Barbu [1984] and Mignot and Puel [1984] and was used in [Herzog et al., 2012, Corollary 3.5] in the context of an optimal control problem for the* dual *formulation of the static elastoplasticity system, see (7.1).*

5. Differentiability of the regularized solution operator.

This section addresses the Fréchet differentiability of the solution map of (4.1b). In fact, we prove this result in a slightly more general setting of the form

$$a(\boldsymbol{W},\boldsymbol{Y}) + \int_\Omega h(\boldsymbol{p}) : \boldsymbol{q}\,\mathrm{d}x = \langle \ell, \boldsymbol{v}\rangle \quad \text{for all } \boldsymbol{Y} = (\boldsymbol{v},\boldsymbol{q}) \in Z, \tag{5.1}$$

with unknown $\boldsymbol{W} = (\boldsymbol{u},\boldsymbol{p}) \in Z$ and a general nonlinear Nemyzki operator h. Assumptions on h are given below, and they admit h_γ from (4.2) as a special case. Equation (5.1) with its general nonlinearity is of independent interest, since it comprises, for example, models of static *viscoplasticity* if $h : \mathbb{S} \to \mathbb{S}$ is properly chosen.

The main step stone in proving the differentiability of the solution map $\ell \mapsto \boldsymbol{W}$ of (5.1) is to establish an integrability result for the solution $\boldsymbol{W}$, i.e., to show that $(\boldsymbol{u},\boldsymbol{p}) \in W^{1,p}_D(\Omega;\mathbb{R}^3) \times L^p(\Omega;\mathbb{Q})$. This is achieved in Theorem 5.2 and Corollary 5.3. The main result of this section in the abstract setting is Theorem 5.5. Lemma 5.6 shows its applicability to the regularized state equation (4.1b), and Corollary 5.7 summarizes the differentiability result for the regularized state equation (4.1b).

In this section, we work with the following assumption for the nonlinearity h in order to achieve the higher integrability result. An additional assumption will be added later (Assumption 5.4).

Assumption 5.1 (Nonlinearity h). *The function $h : \mathbb{S} \to \mathbb{S}$ is continuously differentiable and its derivative satisfies*

$$\|h'(\boldsymbol{p})\|_{\mathcal{L}(\mathbb{S})} \le L_h \quad \textit{for all } \boldsymbol{p} \in \mathbb{S}, \tag{5.2}$$

$$\boldsymbol{q} : h'(\boldsymbol{p})\,\boldsymbol{q} \ge 0 \quad \textit{for all } \boldsymbol{p},\boldsymbol{q} \in \mathbb{S} \tag{5.3}$$

with a constant $L_h > 0$. For every $\boldsymbol{p} \in \mathbb{S}$, the operator $h'(\boldsymbol{p}) \in \mathcal{L}(\mathbb{S})$ is self-adjoint, i.e. $h'(\boldsymbol{p})$ is a symmetric fourth-order tensor in the sense of (2.1). For simplicity, we also assume $h(\boldsymbol{0}) = \boldsymbol{0}$.

The conditions in Assumption 5.1 clearly imply that $h : \mathbb{S} \to \mathbb{S}$ is monotone and globally Lipschitz continuous with Lipschitz constant L_h. These properties carry over to the Nemyzki operator associated with h, which maps $L^p(\Omega;\mathbb{S})$ into $L^p(\Omega;\mathbb{S})$, for any $p \in [1,\infty]$. For simplicity, we denote this Nemyzki operator by the same symbol.

Just as in the proof of Theorem 4.2 (Appendix B), we define the nonlinear map $\mathcal{N} : Z \to Z'$ by the left hand side of (5.1), i.e.,

$$\langle \mathcal{N}\boldsymbol{W},\, \boldsymbol{Y}\rangle = a(\boldsymbol{W},\boldsymbol{Y}) + (h(\boldsymbol{p}),\boldsymbol{q})_\Omega\,. \tag{5.4}$$

We know from (3.2) and the monotonicity of h that $\mathcal{N}$ is strongly monotone, i.e.,

$$\langle \mathcal{N}(\boldsymbol{u}_1-\boldsymbol{u}_2,\boldsymbol{p}_1-\boldsymbol{p}_2),\,(\boldsymbol{u}_1-\boldsymbol{u}_2,\boldsymbol{p}_1-\boldsymbol{p}_2)\rangle \geq \underline{a}\left(\|\boldsymbol{u}_1-\boldsymbol{u}_2\|_V^2 + \|\boldsymbol{p}_1-\boldsymbol{p}_2\|_S^2\right).$$

Moreover $\mathcal{N}$ is clearly continuous. Just as in Theorem 4.2, it thus follows by the Browder-Minty Theorem 3.3 that (5.1) admits a unique solution $\boldsymbol{W} = (\boldsymbol{u},\boldsymbol{p}) \in V \times Q$ for every $\ell \in V'$.

We now transform (5.1) into an equation in the displacement $\boldsymbol{u}$ only. Let us briefly sketch the approach. We borrow ideas from linear saddle-point systems and apply them to

$$\begin{bmatrix}\boldsymbol{\varepsilon}(\boldsymbol{v})\\ \boldsymbol{q}\end{bmatrix}^\top \begin{bmatrix}\mathbb{C}(x) & -\mathbb{C}(x)\,\Pi^D\\ -\Pi^D\,\mathbb{C}(x) & \Pi^D\big[\mathbb{C}(x)+\mathbb{H}(x)+h(\cdot)\big]\,\Pi^D\end{bmatrix}\begin{bmatrix}\boldsymbol{\varepsilon}(\boldsymbol{u})\\ \boldsymbol{p}\end{bmatrix},$$

which is a pointwise representation of the terms appearing on the left hand side of (5.1). The symbol $\Pi^D : \mathbb{S} \to \mathbb{Q}$ denotes the orthogonal projection to the deviatoric part,

$$\Pi^D\boldsymbol{\tau} := \boldsymbol{\tau}^D, \quad \boldsymbol{\tau} \in \mathbb{S},$$

see (3.4). Note that the (2,2) block is nonlinear.

To achieve the reduction to $\boldsymbol{u}$, we test (5.1) with $\boldsymbol{v} = \boldsymbol{0}$ and $\boldsymbol{q} \in Q$ arbitrary and arrive at

$$\Pi^D\big[\mathbb{C}(x)\,\boldsymbol{p}(x) + \mathbb{H}(x)\,\boldsymbol{p}(x) + h(\boldsymbol{p}(x))\big] = \Pi^D\,\mathbb{C}(x)\,\boldsymbol{\varepsilon}(\boldsymbol{u}(x)) \quad \text{a.e. in } \Omega. \tag{5.5}$$

Note that Q consists of all elements of S with vanishing trace so that only the deviatoric parts show up in (5.5). Let us denote the left hand side in (5.5) by $F : \mathbb{Q} \to \mathbb{Q}$, i.e. $F(\boldsymbol{p}) := \Pi^D\big[\mathbb{C}\,\boldsymbol{p} + \mathbb{H}\,\boldsymbol{p} + h(\boldsymbol{p})\big]$. Of course F depends on x, since $\mathbb{C}$ and $\mathbb{H}$ need not to be constant, but we suppress this dependency in the following for the sake of readability. Thanks to $\Pi^D\boldsymbol{\tau} : \boldsymbol{q} = \boldsymbol{\tau} : \boldsymbol{q}$ for all $\boldsymbol{\tau} \in \mathbb{S}$ and $\boldsymbol{q} \in \mathbb{Q}$, the monotonicity of h yields

$$\big(F(\boldsymbol{q}_1) - F(\boldsymbol{q}_2)\big) : (\boldsymbol{q}_1 - \boldsymbol{q}_2) \geq (\underline{c} + \underline{h})\,|\boldsymbol{q}_1 - \boldsymbol{q}_1|^2 \quad \text{for all } \boldsymbol{q}_1, \boldsymbol{q}_2 \in \mathbb{Q}, \tag{5.6}$$

where $\underline{c}$ and $\underline{h}$ are the coercivity constants from Assumption 2.2. Moreover, as Π^D is linear and bounded with constant one, Assumptions 2.2 and 5.1 imply

$$|F(\boldsymbol{q}_1) - F(\boldsymbol{q}_2)| \leq \big(\overline{c} + \overline{h} + L_h\big)\,|\boldsymbol{q}_1 - \boldsymbol{q}_2| \quad \text{for all } \boldsymbol{q}_1, \boldsymbol{q}_2 \in \mathbb{Q}. \tag{5.7}$$

As the constants in (5.6) and (5.7) are independent of x, Theorem 3.3 implies that $F(x,\cdot) : \mathbb{Q} \to \mathbb{Q}$ is continuously invertible for almost every $x \in \Omega$ such that (5.5) gives

$$\boldsymbol{p}(x) = F^{-1}\big(x, \Pi^D\mathbb{C}(x)\,\boldsymbol{\varepsilon}(\boldsymbol{u}(x))\big) \quad \text{a.e. in } \Omega. \tag{5.8}$$

Moreover, by Theorem 3.3, the pointwise inverse $F^{-1}(x,\cdot) : \mathbb{Q} \to \mathbb{Q}$ is strongly monotone and Lipschitz continuous with constants independent of x. Arguing as in [Betz and Meyer, 2012, Theorem 2.4], one can show in addition that $x \mapsto F^{-1}(x, \boldsymbol{q})$ is measurable for every $\boldsymbol{q} \in \mathbb{Q}$. Thus the Nemyzki operator associated with F^{-1} satisfies the Carathéodory condition and it maps $L^p(\Omega;\mathbb{Q})$ into $L^p(\Omega;\mathbb{Q})$ for every $p \in [1,\infty]$. We will denote this operator by the same symbol. Moreover, according to Goldberg et al. [1992] or [Tröltzsch, 2010, Section 4.3], the Lipschitz continuity of $F^{-1}(x,\cdot)$ carries over to its Nemyzki operator.

We now test (5.1) with $\boldsymbol{q} = \boldsymbol{0}$ and $\boldsymbol{v} \in V$ and eliminate $\boldsymbol{p}$ by (5.8). This yields the desired reduced formulation:

$$\int_\Omega b(x, \boldsymbol{\varepsilon}(\boldsymbol{u})) : \boldsymbol{\varepsilon}(\boldsymbol{v}) \,\mathrm{d}x = \langle \ell,\, \boldsymbol{v} \rangle \quad \text{for all } \boldsymbol{v} \in V, \tag{5.9}$$

where we are using the abbreviation

$$b(x,\cdot) : \mathbb{S} \ni \boldsymbol{\varepsilon} \mapsto \mathbb{C}(x)\Big(\boldsymbol{\varepsilon} - F^{-1}\big(x, \Pi^D \mathbb{C}(x)\, \boldsymbol{\varepsilon}\big)\Big) \in \mathbb{S}. \tag{5.10}$$

The above derivation shows that, if $(\boldsymbol{u}, \boldsymbol{p}) \in V \times Q$ solves (5.1), then $\boldsymbol{u}$ is a solution of (5.9). On the other hand it is easily seen that, if $\boldsymbol{u} \in V$ solves (5.9), then $(\boldsymbol{u}, \boldsymbol{p})$ with $\boldsymbol{p}$ as defined in (5.8) is a solution of (5.1). Thus (5.1) and (5.9) are indeed equivalent. The unique solvability of (5.1) implies that, for every $\ell \in V'$, (5.9) admits a unique solution $\boldsymbol{u} \in V$.

As indicated at the beginning of this section, we wish to prove higher integrability of $\boldsymbol{u}$ provided that ℓ is more regular than just an element of V'. To this end we aim to apply [Herzog et al., 2011a, Theorem 1.1] to the reduced problem (5.9). This requires us to verify that $b(x,\cdot)$ is strongly monotone and Lipschitz continuous with constants independent of x. The uniform Lipschitz continuity of b follows immediately from the uniform boundedness of $\mathbb{C}$ and the uniform Lipschitz continuity of $F^{-1}(x,\cdot)$. To prove the strong monotonicity of b, let us define, for any $x \in \Omega$, the pointwise map $\mathcal{M}(x,\cdot,\cdot) : \mathbb{S} \times \mathbb{Q} \to \mathbb{S} \times \mathbb{Q}$ by

$$\mathcal{M}(x, \boldsymbol{\varepsilon}, \boldsymbol{p}) := \begin{bmatrix} \mathbb{C}(x)(\boldsymbol{\varepsilon} - \boldsymbol{p}) \\ F(x, \boldsymbol{p}) - \Pi^D \mathbb{C}(x)\, \boldsymbol{\varepsilon} \end{bmatrix}. \tag{5.11}$$

Arguing as in (3.1) and using again the monotonicity of h, we infer that $\mathcal{M}(x,\cdot,\cdot)$ is strongly monotone:

$$\big(\mathcal{M}(x, \boldsymbol{\varepsilon}_1, \boldsymbol{p}_1) - \mathcal{M}(x, \boldsymbol{\varepsilon}_2, \boldsymbol{p}_2)\big) : (\boldsymbol{\varepsilon}_1 - \boldsymbol{\varepsilon}_2, \boldsymbol{p}_1 - \boldsymbol{p}_2) \geq \underline{m}\left(|\boldsymbol{\varepsilon}_1 - \boldsymbol{\varepsilon}_2|^2 + |\boldsymbol{p}_1 - \boldsymbol{p}_2|^2\right) \tag{5.12}$$

with some $\underline{m} > 0$. Due to the uniform coercivity of $\mathbb{C}$ and $\mathbb{H}$, the constant $\underline{m}$ is independent of x. Now let $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2 \in \mathbb{S}$ be arbitrary and let $\boldsymbol{p}_1 = F^{-1}\big(x, \Pi^D \mathbb{C}(x)(\boldsymbol{\varepsilon}_1 - \boldsymbol{\varepsilon}_2)\big)$ and $\boldsymbol{p}_2 = \boldsymbol{0}$. Inserting these into (5.12) yields

$$\begin{aligned}
\underline{m}\, |\boldsymbol{\varepsilon}_1 - \boldsymbol{\varepsilon}_2|^2 &\leq \underline{m}\left(|\boldsymbol{\varepsilon}_1 - \boldsymbol{\varepsilon}_2|^2 + |\boldsymbol{p}_1|^2\right) \\
&\leq \big(\mathcal{M}(\boldsymbol{\varepsilon}_1, \boldsymbol{p}_1) - \mathcal{M}(\boldsymbol{\varepsilon}_2, \boldsymbol{0})\big) : (\boldsymbol{\varepsilon}_1 - \boldsymbol{\varepsilon}_2, \boldsymbol{p}_1) \\
&= \Big(\mathbb{C}(\boldsymbol{\varepsilon}_1 - \boldsymbol{\varepsilon}_2) - \mathbb{C}F^{-1}(\Pi^D \mathbb{C}(\boldsymbol{\varepsilon}_1 - \boldsymbol{\varepsilon}_2))\Big) : (\boldsymbol{\varepsilon}_1 - \boldsymbol{\varepsilon}_2) - F(\boldsymbol{0}) : \boldsymbol{p}_1 \\
&= b(\cdot, \boldsymbol{\varepsilon}_1 - \boldsymbol{\varepsilon}_2) : (\boldsymbol{\varepsilon}_1 - \boldsymbol{\varepsilon}_2) - F(\boldsymbol{0}) : \boldsymbol{p}_1,
\end{aligned}$$

where again we suppress the dependency on x. Because of Assumption 5.1 we have $F(\mathbf{0}) = \mathbf{0}$ so that (5.10) gives

$$\underline{m}\,|\boldsymbol{\varepsilon}_1 - \boldsymbol{\varepsilon}_2|^2 \le b(x, \boldsymbol{\varepsilon}_1 - \boldsymbol{\varepsilon}_2) : (\boldsymbol{\varepsilon}_1 - \boldsymbol{\varepsilon}_2),$$

uniformly for $x \in \Omega$. In addition, because of $h(\mathbf{0}) = \mathbf{0}$, we have $b(\mathbf{0}) = \mathbf{0} \in L^\infty(\Omega; \mathbb{S})$. Since $x \mapsto b(x, \boldsymbol{\varepsilon})$ is also measurable for every $\boldsymbol{\varepsilon} \in \mathbb{S}$ — thanks to the measurability of F^{-1} mentioned above — the operator b satisfies the conditions in [Herzog et al., 2011a, Assumption 1.5(2)]. Taking into account Assumption 2.1, [Herzog et al., 2011a, Theorem 1.1] is applicable and it yields the following higher integrability result:

Theorem 5.2. *There exists an index $p > 2$ such that for every $\ell \in W_D^{-1,p}(\Omega; \mathbb{R}^3) = W_D^{-1,p'}(\Omega; \mathbb{R}^3)$, the equation (5.9) admits a unique solution $\boldsymbol{u} \in W_D^{1,p}(\Omega; \mathbb{R}^3)$. Moreover, the associated solution mapping is globally Lipschitz, i.e., there exists $L > 0$ such that*

$$\|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_{W^{1,p}(\Omega;\mathbb{R}^3)} \le L\,\|\ell_1 - \ell_2\|_{W^{-1,p}(\Omega;\mathbb{R}^3)}$$

holds for all $\ell_1, \ell_2 \in W^{-1,p}(\Omega; \mathbb{R}^3)$.

Due to the trace theorem and Sobolev embeddings, an inhomogeneity of the form

$$\langle \ell, \boldsymbol{v}\rangle = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}\,\mathrm{d}x + \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{v}\,\mathrm{d}s = \langle R(\boldsymbol{f}, \boldsymbol{g}), \boldsymbol{v}\rangle \tag{5.13}$$

with $\boldsymbol{f} \in L^2(\Omega; \mathbb{R}^3)$ and $\boldsymbol{g} \in L^2(\Gamma_N; \mathbb{R}^3)$ does represent an element of $W^{-1,p}(\Omega; \mathbb{R}^3)$ for every $p < 4$.

In order to transfer the result to the original problem (5.1), we exploit that we can recover the plastic strain $\boldsymbol{p}$ from $\boldsymbol{u}$ by the pointwise relation (5.8). Since F^{-1} is globally Lipschitz from $L^p(\Omega; \mathbb{S})$ to $L^p(\Omega; \mathbb{Q})$ as seen above, Assumption 2.2 implies the same for $F^{-1} \circ \Pi^D \mathbb{C}$ and we can conclude the following result.

Corollary 5.3. *There exists an index $p \in (2, 4)$ such that for every $\boldsymbol{f} \in L^2(\Omega; \mathbb{R}^3)$ and $\boldsymbol{g} \in L^2(\Gamma_N; \mathbb{R}^3)$, the equation (5.1) with ℓ as in (5.13) admits a unique solution $(\boldsymbol{u}, \boldsymbol{p}) \in W_D^{1,p}(\Omega; \mathbb{R}^3) \times L^p(\Omega; \mathbb{Q})$. Moreover, the associated solution mapping $\mathcal{G}$: $L^2(\Omega; \mathbb{R}^3) \times L^2(\Gamma_N; \mathbb{R}^3) \to W_D^{1,p}(\Omega; \mathbb{R}^3) \times L^p(\Omega; \mathbb{Q})$ is globally Lipschitz continuous, i.e., there exists $L > 0$ such that*

$$\|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_{W^{1,p}(\Omega;\mathbb{R}^3)} + \|\boldsymbol{p}_1 - \boldsymbol{p}_2\|_{L^p(\Omega;\mathbb{Q})} \le L\left(\|\boldsymbol{f}_1 - \boldsymbol{f}_2\|_{L^2(\Omega;\mathbb{R}^3)} + \|\boldsymbol{g}_1 - \boldsymbol{g}_2\|_{L^2(\Gamma_N;\mathbb{R}^3)}\right)$$

holds for all $\boldsymbol{f}_1, \boldsymbol{f}_2 \in L^2(\Omega; \mathbb{R}^3)$ and $\boldsymbol{g}_1, \boldsymbol{g}_2 \in L^2(\Gamma_N; \mathbb{R}^3)$.

Based on this integrability result, we are now in the position to prove the differentiability of $\mathcal{G}$. An additional assumption is needed.

Assumption 5.4. *Assume that the Nemyzki operator associated with h is Fréchet differentiable from $L^p(\Omega; \mathbb{S})$ to $L^2(\Omega; \mathbb{S})$ with $p > 2$ as in Corollary 5.3.*

This assumption will be verified for the particular nonlinearity h_γ from (4.2) in Lemma 5.6 below.

Theorem 5.5. *Under Assumptions 5.1 and 5.4, $\mathcal{G}$ is Fréchet differentiable from $L^2(\Omega; \mathbb{R}^3) \times L^2(\Gamma_N; \mathbb{R}^3)$ to Z, and the derivative $\delta\boldsymbol{W} = (\delta\boldsymbol{u}, \delta\boldsymbol{p}) = \mathcal{G}'(\boldsymbol{f}, \boldsymbol{g})(\delta\boldsymbol{f}, \delta\boldsymbol{g})$ at $(\boldsymbol{f}, \boldsymbol{g})$ in the direction $(\delta\boldsymbol{f}, \delta\boldsymbol{g})$ is given by the unique solution $\delta\boldsymbol{W} = (\delta\boldsymbol{u}, \delta\boldsymbol{p}) \in Z$ of the linearized equation*

$$a(\delta\boldsymbol{W}, \boldsymbol{Y}) + \int_\Omega h'(\boldsymbol{p})\,\delta\boldsymbol{p} : \boldsymbol{q}\,\mathrm{d}x = \int_\Omega \delta\boldsymbol{f} \cdot \boldsymbol{v}\,\mathrm{d}x + \int_{\Gamma_N} \delta\boldsymbol{g} \cdot \boldsymbol{v}\,\mathrm{d}s \quad \textit{for all } \boldsymbol{Y} = (\boldsymbol{v}, \boldsymbol{q}) \in Z, \tag{5.14}$$

where $(\boldsymbol{u},\boldsymbol{p}) = \mathcal{G}(\boldsymbol{f},\boldsymbol{g})$.

Proof. Thanks to (5.3) and (5.2) the bilinear form on the left hand side of (5.14) is bounded and coercive so that (5.14) admits a unique solution $\delta\boldsymbol{W} = (\delta\boldsymbol{u},\delta\boldsymbol{p})$. Note that due to (5.2) the operator $h'(\boldsymbol{p})$ can be extended to an operator from S to S, which we denote by the same symbol. Next let us introduce the remainder

$$\boldsymbol{R} = (\boldsymbol{r}_u,\boldsymbol{r}_p) := \underbrace{\mathcal{G}(\boldsymbol{f}+\delta\boldsymbol{f},\boldsymbol{g}+\delta\boldsymbol{g})}_{=:(\widehat{\boldsymbol{u}},\widehat{\boldsymbol{p}})} - \mathcal{G}(\boldsymbol{f},\boldsymbol{g}) - \delta\boldsymbol{W}$$

The remainder $\boldsymbol{R}$ solves

$$a(\boldsymbol{R},\boldsymbol{Y}) + \int_\Omega h'(\boldsymbol{p})\,\boldsymbol{r}_p : \boldsymbol{q}\,\mathrm{d}x = -\int_\Omega \big(h(\widehat{\boldsymbol{p}}) - h(\boldsymbol{p}) - h'(\boldsymbol{p})(\widehat{\boldsymbol{p}}-\boldsymbol{p})\big) : \boldsymbol{q}\,\mathrm{d}x$$

for all $\boldsymbol{Y} = (\boldsymbol{v},\boldsymbol{q}) \in Z$. Due to the coercivity of the bilinear form in the linearized equation (5.14) and the differentiability assumption on h we find

$$\begin{aligned}
\|\boldsymbol{R}\|_Z &\le c\,\|h(\widehat{\boldsymbol{p}}) - h(\boldsymbol{p}) - h'(\boldsymbol{p})(\widehat{\boldsymbol{p}}-\boldsymbol{p})\|_S \\
&= o\big(\|\widehat{\boldsymbol{p}}-\boldsymbol{p}\|_{L^p(\Omega;\mathbb{S})}\big) \\
&= o\big(\|(\delta\boldsymbol{f},\delta\boldsymbol{g})\|_{L^2(\Omega;\mathbb{R}^3)\times L^2(\Gamma_N;\mathbb{R}^3)}\big),
\end{aligned}$$

where the Lipschitz continuity of $\mathcal{G}$ by Corollary 5.3 was used in the previous estimate. Thus $\mathcal{G}$ is Fréchet differentiable with derivative $\mathcal{G}'(\boldsymbol{f},\boldsymbol{g})$. ☐

Lemma 5.6. *For every $\gamma > 0$, the function h_γ defined in (4.2) satisfies the conditions in Assumption 5.1. Moreover, the associated Nemyzki operator is Fréchet differentiable from $L^p(\Omega;\mathbb{S})$ to $L^2(\Omega;\mathbb{S})$ for every $p > 2$, so that Assumption 5.4 holds as well. The derivative of h_γ at $\boldsymbol{p} \in L^p(\Omega;\mathbb{S})$ in the direction $\delta\boldsymbol{p} \in L^p(\Omega;\mathbb{S})$ is given by*

$$\big(h'_\gamma(\boldsymbol{p})\,\delta\boldsymbol{p}\big)(x) = \widetilde{\sigma}_0\,\gamma\left(\frac{\delta\boldsymbol{p}(x)}{m_\gamma(|\boldsymbol{p}(x)|)} - m'_\gamma(|\boldsymbol{p}(x)|)\,\frac{\boldsymbol{p}(x):\delta\boldsymbol{p}(x)}{m_\gamma(|\boldsymbol{p}(x)|)^2}\,\frac{\boldsymbol{p}(x)}{|\boldsymbol{p}(x)|}\right) \tag{5.15}$$

with $m_\gamma : \mathbb{R}\to\mathbb{R}$ as defined in (4.3) and m'_γ as given in (4.7).

Proof. The first assertion has already been proved in Lemma 4.1, cf. (4.10) and (4.9). To show the differentiability property of the Nemyzki operator associated with h_γ, let $p > 2$ be given and define $r = 2p/(p-2) < \infty$. Thanks to (4.9) we have that $S \ni \boldsymbol{p} \mapsto h'_\gamma(\boldsymbol{p}) \in L^\infty(\Omega;\mathcal{L}(\mathbb{S})) \hookrightarrow L^r(\Omega;\mathcal{L}(\mathbb{S}))$. Since moreover h_γ maps S and thus also $L^p(\Omega;\mathbb{S})$ into $L^\infty(\Omega;\mathbb{S}) \hookrightarrow S$, the desired Fréchet differentiability follows from abstract results for Nemyzki operators, see Goldberg et al. [1992] or [Tröltzsch, 2010, Section 4.3]. ☐

Corollary 5.7. *For any $\gamma > 0$, there exists $p \in (2,4)$ such that the solution operator G_γ of (4.1b), with $\ell = R(\boldsymbol{f},\boldsymbol{g})$ as in (5.13), maps $L^2(\Omega;\mathbb{R}^3)\times L^2(\Gamma_N;\mathbb{R}^3)$ into $W^{1,p}_D(\Omega;\mathbb{R}^3)\times L^p(\Omega;\mathbb{Q})$, and it is Fréchet differentiable from $L^2(\Omega;\mathbb{R}^3)\times L^2(\Gamma_N;\mathbb{R}^3)$ to Z. The derivative $\delta\boldsymbol{W} = (\delta\boldsymbol{u},\delta\boldsymbol{p}) = G'_\gamma(\boldsymbol{f},\boldsymbol{g})(\delta\boldsymbol{f},\delta\boldsymbol{g})$ at $(\boldsymbol{f},\boldsymbol{g})$ in the direction $(\delta\boldsymbol{f},\delta\boldsymbol{g})$ is given by the unique solution $\delta\boldsymbol{W} = (\delta\boldsymbol{u},\delta\boldsymbol{p}) \in Z$ of the linearized equation (5.14) with h_γ in place of h.*

Proof. Lemma 5.6 shows that the nonlinearity h_γ from (4.2) verifies Assumptions 5.1 and 5.4. Consequently, Corollary 5.3 and Theorem 5.5 apply for this particular choice of h. Thus the solution operator G_γ of (4.1b) maps $L^2(\Omega;\mathbb{R}^3)\times L^2(\Gamma_N;\mathbb{R}^3)$ to $W^{1,p}_D(\Omega;\mathbb{R}^3)\times L^p(\Omega;\mathbb{Q})$ for some $p \in (2,4)$. The Fréchet derivative is given by (5.14) with h'_γ in place of h'. ☐

6. Optimality system. Based on the differentiability result established in Theorem 5.5, we can now derive first-order necessary optimality conditions for the regularized control problem (4.1). Later on we pass to the limit $\gamma \to \infty$ to obtain the optimality system for the original problem (1.3) stated in Theorem 1.1.

THEOREM 6.1 (Regularized optimality system). *Let $(\boldsymbol{f}_\gamma, \boldsymbol{g}_\gamma) \in L^2(\Omega;\mathbb{R}^3) \times L^2(\Gamma_N;\mathbb{R}^3)$ be a locally optimal solution of the regularized problem (4.1) with associated state $\boldsymbol{W}_\gamma = (\boldsymbol{u}_\gamma, \boldsymbol{p}_\gamma) \in Z$. Then there exists an adjoint state $\boldsymbol{Z}_\gamma = (\boldsymbol{w}_\gamma, \boldsymbol{r}_\gamma) \in Z$ and multipliers $\boldsymbol{\varrho}_\gamma \in Q$ and $\boldsymbol{\pi}_\gamma \in Q$ such that the following optimality system is satisfied:*

$$a(\boldsymbol{W}_\gamma, \boldsymbol{Y}) + \int_\Omega \boldsymbol{\varrho}_\gamma : \boldsymbol{q}\,\mathrm{d}x = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}\,\mathrm{d}x + \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{v}\,\mathrm{d}s \quad \textit{for all } \boldsymbol{Y} = (\boldsymbol{v}, \boldsymbol{q}) \in Z \tag{6.1a}$$

$$\boldsymbol{\varrho}_\gamma = \widetilde{\sigma}_0 \frac{\gamma\,\boldsymbol{p}_\gamma}{m_\gamma(|\boldsymbol{p}_\gamma|)} \quad \textit{a.e. in } \Omega \tag{6.1b}$$

$$a(\boldsymbol{Y}, \boldsymbol{Z}_\gamma) + \int_\Omega \boldsymbol{\pi}_\gamma : \boldsymbol{q}\,\mathrm{d}x = -\int_\Omega (\boldsymbol{u}_\gamma - \boldsymbol{u}_d) \cdot \boldsymbol{v}\,\mathrm{d}x \quad \textit{for all } \boldsymbol{Y} = (\boldsymbol{v}, \boldsymbol{q}) \in Z \tag{6.1c}$$

$$\boldsymbol{\pi}_\gamma = \widetilde{\sigma}_0\,\gamma \left(\frac{\boldsymbol{r}_\gamma}{m_\gamma(|\boldsymbol{p}_\gamma|)} - m_\gamma'(|\boldsymbol{p}_\gamma|)\, \frac{\boldsymbol{r}_\gamma : \boldsymbol{p}_\gamma}{m_\gamma(|\boldsymbol{p}_\gamma|)^2} \frac{\boldsymbol{p}_\gamma}{|\boldsymbol{p}_\gamma|} \right) \quad \textit{a.e. in } \Omega \tag{6.1d}$$

$$\nu_1\,\boldsymbol{f}_\gamma - \boldsymbol{w}_\gamma = \boldsymbol{0} \quad \textit{a.e. in } \Omega \tag{6.1e}$$

$$\nu_2\,\boldsymbol{g}_\gamma - \boldsymbol{w}_\gamma = \boldsymbol{0} \quad \textit{a.e. in } \Gamma_N. \tag{6.1f}$$

We note that (6.1a)–(6.1b) represents the state equation (4.1). We introduced the term $\boldsymbol{\varrho}_\gamma$ through (6.1b) in order to facilitate the passage to the limit in Proposition 6.2 below. Equations (6.1c)–(6.1d) are the adjoint equation, while (6.1e)–(6.1f) represent the stationarity condition w.r.t. the controls.

Proof. We first eliminate the multiplier-like terms $\boldsymbol{\varrho}_\gamma$ and $\boldsymbol{\pi}_\gamma$ from (6.1) by plugging in (6.1b) into (6.1a) and (6.1d) into (6.1d). Then (6.1a) becomes the regularized state equation (4.1b), whose adjoint equation for $\boldsymbol{Z}_\gamma = (\boldsymbol{w}_\gamma, \boldsymbol{r}_\gamma)$ reads

$$a(\boldsymbol{Y}, \boldsymbol{Z}_\gamma) + (h_\gamma'(\boldsymbol{p}_\gamma)\,\boldsymbol{r}_\gamma, \boldsymbol{q})_\Omega = -(\boldsymbol{u}_\gamma - \boldsymbol{u}_d, \boldsymbol{v})_\Omega \quad \text{for all } \boldsymbol{Y} \in Z. \tag{6.2}$$

Here we used that $\boldsymbol{\pi}_\gamma$ in (6.1d) is equal to $h_\gamma'(\boldsymbol{p}_\gamma)\,\boldsymbol{r}_\gamma$, cf. (4.6). Since the bilinear form $a(\cdot,\cdot)$ is symmetric (due to the symmetry properties of $\mathbb{C}$ and $\mathbb{H}$), the left hand side of the adjoint equation thus coincides with the one of the linearized equation so that there exists a unique solution $\boldsymbol{Z}_\gamma \in Z$ to (6.2).

Next we consider the reduced cost functional $\mathcal{J}(\boldsymbol{f}, \boldsymbol{g}) := J(G_\gamma^{\boldsymbol{u}}(\boldsymbol{f}, \boldsymbol{g}), (\boldsymbol{f}, \boldsymbol{g}))$. Here $G_\gamma = (G_\gamma^{\boldsymbol{u}}, G_\gamma^{\boldsymbol{p}})$ denotes the splitting of the (regularized) solution operator into displacement and plastic strain components. Since J is of quadratic type and G_γ is Fréchet differentiable from $L^2(\Omega;\mathbb{R}^3) \times L^2(\Gamma_N;\mathbb{R}^3)$ to Z by Corollary 5.7. It follows from the local optimality that

$$\mathcal{J}'(\boldsymbol{f}_\gamma, \boldsymbol{g}_\gamma)(\delta\boldsymbol{f}, \delta\boldsymbol{g}) = 0 \quad \text{for all } (\delta\boldsymbol{f}, \delta\boldsymbol{g}) \in L^2(\Omega;\mathbb{R}^3) \times L^2(\Gamma_N;\mathbb{R}^3). \tag{6.3}$$

Let $\delta\boldsymbol{W} = (\delta\boldsymbol{p}, \delta\boldsymbol{u})$ denote the unique solution to the linearized equation (5.14) with h_γ' in place of h'. Using the chain rule we obtain that

$$\begin{aligned}
&\mathcal{J}'(\boldsymbol{f}_\gamma, \boldsymbol{g}_\gamma)(\delta\boldsymbol{f}, \delta\boldsymbol{g}) \\
&\quad = (\boldsymbol{u}_\gamma - \boldsymbol{u}_d, \delta\boldsymbol{u})_\Omega + \nu_1(\boldsymbol{f}, \delta\boldsymbol{f})_\Omega + \nu_2(\boldsymbol{g}, \delta\boldsymbol{g})_{\Gamma_N} \\
&\quad = -a(\delta\boldsymbol{W}, \boldsymbol{Z}_\gamma) - (h_\gamma'(\boldsymbol{p}_\gamma)\,\delta\boldsymbol{p}, \boldsymbol{r}_\gamma)_\Omega + \nu_1(\boldsymbol{f}, \delta\boldsymbol{f})_\Omega + \nu_2(\boldsymbol{g}, \delta\boldsymbol{g})_{\Gamma_N} \\
&\quad = -(\boldsymbol{w}_\gamma, \delta\boldsymbol{f})_\Omega - (\boldsymbol{w}_\gamma, \delta\boldsymbol{g})_{\Gamma_N} + \nu_1(\boldsymbol{f}, \delta\boldsymbol{f})_\Omega + \nu_2(\boldsymbol{g}, \delta\boldsymbol{g})_{\Gamma_N}.
\end{aligned}$$

Together with (6.3) this implies (6.1e) and (6.1f). Finally, in view of (4.4) and (4.9), it is easy to check that indeed $\boldsymbol{\varrho}_\gamma \in Q$ and $\boldsymbol{\pi}_\gamma \in Q$ as claimed. □

It will be convenient to refer to (6.1) as the regularized optimality system (rather than the optimality system of the regularized control problem), and to (6.1c)–(6.1d) as the regularized adjoint equation, etc. These slightly imprecise terms will not give rise to confusion.

We now pass to the limit in the regularized optimality system (6.1). As an intermediate step in the proof of Theorem 1.1, we obtain a preliminary version of first-order necessary optimality conditions for the original problem (1.3). This result will be refined in Lemma 6.4 below, which subsequently leads to the proof of Theorem 1.1.

PROPOSITION 6.2 (Preliminary optimality system). *Let $(\boldsymbol{f}, \boldsymbol{g})$ be a locally optimal solution for (1.3) with associated state $(\boldsymbol{u}, \boldsymbol{p}) \in Z$. Then there exists an adjoint state $\boldsymbol{Z} = (\boldsymbol{w}, \boldsymbol{r}) \in Z$ and multipliers $\boldsymbol{\varrho} \in Q$ and $\boldsymbol{\pi} \in Q$ such that the following optimality system is satisfied:*

$$a(\boldsymbol{W}, \boldsymbol{Y}) + \int_\Omega \boldsymbol{\varrho} : \boldsymbol{q} \,\mathrm{d}x = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} \,\mathrm{d}x + \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{v} \,\mathrm{d}s \quad \text{for all } \boldsymbol{Y} = (\boldsymbol{v}, \boldsymbol{q}) \in Z, \tag{6.4a}$$

$$\boldsymbol{\varrho} : \boldsymbol{p} = \widetilde{\sigma}_0 \, |\boldsymbol{p}| \quad \text{a.e. in } \Omega, \tag{6.4b}$$

$$|\boldsymbol{\varrho}| \le \widetilde{\sigma}_0 \quad \text{a.e. in } \Omega, \tag{6.4c}$$

$$a(\boldsymbol{Y}, \boldsymbol{Z}) + \int_\Omega \boldsymbol{\pi} : \boldsymbol{q} \,\mathrm{d}x = -\int_\Omega (\boldsymbol{u} - \boldsymbol{u}_d) \cdot \boldsymbol{v} \,\mathrm{d}x \quad \text{for all } \boldsymbol{Y} = (\boldsymbol{v}, \boldsymbol{q}) \in Z, \tag{6.4d}$$

$$\nu_1 \, \boldsymbol{f} - \boldsymbol{w} = \boldsymbol{0} \quad \text{a.e. in } \Omega \tag{6.4e}$$

$$\nu_2 \, \boldsymbol{g} - \boldsymbol{w} = \boldsymbol{0} \quad \text{a.e. in } \Gamma_N, \tag{6.4f}$$

as well as

$$\boldsymbol{\pi} : \boldsymbol{p} = 0 \qquad \text{a.e. in } \Omega, \tag{6.4g}$$

$$\boldsymbol{\pi} : \boldsymbol{r} \ge 0 \qquad \text{a.e. in } \Omega, \tag{6.4h}$$

$$\boldsymbol{r} = \boldsymbol{0} \qquad \text{a.e. in } \mathcal{I} = \{x \in \Omega : |\boldsymbol{\varrho}(x)| < \widetilde{\sigma}_0\}. \tag{6.4i}$$

Note that (6.4b)–(6.4c) are equivalent to $\boldsymbol{\varrho} \in \widetilde{\sigma}_0 \, \partial |\boldsymbol{p}|$, the subdifferential of the Frobenius norm.

Proof. The proof principally follows the lines of [de los Reyes, 2011, Theorem 5.1]. Nevertheless, since substantial parts of the proof have to be modified due to the special structure of static elastoplasticity, we present the proof in detail.

We elaborate on the arguments under the assumption that $(\boldsymbol{f}, \boldsymbol{g}) \in L^2(\Omega; \mathbb{R}^3) \times L^2(\Gamma_N; \mathbb{R}^3)$ is a (local or global) optimal control of (1.3) which can be approximated by a strongly convergent sequence of (local or global) solutions $(\boldsymbol{f}_\gamma, \boldsymbol{g}_\gamma)$ of (4.1). As was mentioned in Remark 4.5, this may not be the case for all local minimizers of (1.3). We briefly come back to the necessary modifications to the arguments at the end of the proof.

Let us denote the state associated to $(\boldsymbol{f}, \boldsymbol{g})$ by $(\boldsymbol{u}, \boldsymbol{p}) \in V \times Q$, i.e., $(\boldsymbol{u}, \boldsymbol{p})$ solves (1.1) with ℓ in the right hand side given by (1.2c).

Step 1. The state system (6.4a)–(6.4c):
The state system is simply an alternative formulation of (1.1), so it can be checked without a limit argument. We first define

$$\boldsymbol{\varrho} := \left[\mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{p}) - \mathbb{H}\, \boldsymbol{p}\right]^D.$$

We multiply this equation by $\boldsymbol{q} \in Q$ and integrate over Ω to obtain

$$\int_\Omega \left[(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{p}) : \mathbb{C}\,(\boldsymbol{\varepsilon}(\boldsymbol{0}) - \boldsymbol{q})\right] \mathrm{d}x + \int_\Omega \boldsymbol{p} : \mathbb{H}\, \boldsymbol{q} \,\mathrm{d}x + \int_\Omega \boldsymbol{\varrho} : \boldsymbol{q} \,\mathrm{d}x = 0 \quad \text{for all } \boldsymbol{q} \in Q.$$

On the other hand, we test (1.1) with $\boldsymbol{Y} = (\boldsymbol{v} + \boldsymbol{u}, \boldsymbol{p})$ with $\boldsymbol{v} \in V$ arbitrary to get

$$\int_\Omega \left[(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{p}) : \mathbb{C}\,(\boldsymbol{\varepsilon}(\boldsymbol{v}) - \boldsymbol{0})\right] \mathrm{d}x \geq \langle \ell, \boldsymbol{v} \rangle \quad \text{for all } \boldsymbol{v} \in V.$$

Since $\boldsymbol{v} \in V$ is arbitrary, we have indeed an equality. When added to the previous equality, (6.4a) follows.

Next we set $\boldsymbol{Y} = (\boldsymbol{u}, \boldsymbol{q})$ in (1.1) so that we arrive at

$$\int_\Omega \left[(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{p}) : \mathbb{C}\,(\boldsymbol{\varepsilon}(\boldsymbol{0}) - (\boldsymbol{q} - \boldsymbol{p}))\right] \mathrm{d}x + \int_\Omega \boldsymbol{p} : \mathbb{H}\,(\boldsymbol{q} - \boldsymbol{p}) \,\mathrm{d}x + j(\boldsymbol{q}) - j(\boldsymbol{p}) \geq 0$$

for all $\boldsymbol{q} \in Q$. Using the definition of $\boldsymbol{\varrho}$ and since $\boldsymbol{q} \in Q$ is arbitrary, this can be written as

$$\int_\Omega \Big(\widetilde{\sigma}_0\big(|\boldsymbol{p}| - |\boldsymbol{q}|\big) - \boldsymbol{\varrho} : (\boldsymbol{p} - \boldsymbol{q})\Big) \,\mathrm{d}x = 0 \quad \text{for all } \boldsymbol{q} \in Q. \tag{6.5}$$

From this we will deduce (6.4c) and (6.4b) by contradiction. Assume first that there is a subset $E_1 \subset \Omega$ of positive measure with $|\boldsymbol{\varrho}| > \widetilde{\sigma}_0$ a.e. in E_1. If we insert $\boldsymbol{q} = \chi_{E_1}(\boldsymbol{p} - \boldsymbol{\varrho}) + \chi_{\Omega \setminus E_1} \boldsymbol{p} \in Q$ into (6.5), then we get (using $|\boldsymbol{p}| \leq |\boldsymbol{p} - \boldsymbol{\varrho}| + |\boldsymbol{\varrho}|$)

$$0 = \int_{E_1} \Big(\widetilde{\sigma}_0\big(|\boldsymbol{p}| - |\boldsymbol{p} - \boldsymbol{\varrho}|\big) - |\boldsymbol{\varrho}|^2\Big) \,\mathrm{d}x \leq \int_{E_1} |\boldsymbol{\varrho}|\big(\widetilde{\sigma}_0 - |\boldsymbol{\varrho}|\big) \,\mathrm{d}x < 0,$$

which yields a contradiction. Consequently, (6.4c) is proved. To verify (6.4b), we assume to the contrary that there is a subset $E_2 \subset \Omega$ of positive measure with the property $\boldsymbol{p} : \boldsymbol{\varrho} < \widetilde{\sigma}_0 \,|\boldsymbol{p}|$ a.e. in E_2. (Note that (6.4c) forbids $\boldsymbol{p} : \boldsymbol{\varrho} > \widetilde{\sigma}_0 \,|\boldsymbol{p}|$.) By choosing $\boldsymbol{q} = \chi_{\Omega \setminus E_2} \boldsymbol{p} \in Q$ in (6.5), we obtain

$$0 = \int_{E_2} \big(\widetilde{\sigma}_0 \,|\boldsymbol{p}| - \boldsymbol{\varrho} : \boldsymbol{p}\big) \,\mathrm{d}x < 0,$$

again a contradiction, which shows that (6.4b) must hold.

Step 2. The adjoint equation (6.4d):
By definition of $\boldsymbol{\pi}_\gamma$, see (6.1d), we have a.e. in Ω

$$\begin{aligned}
\boldsymbol{\pi}_\gamma : \boldsymbol{r}_\gamma &= \widetilde{\sigma}_0\, \gamma \left(\frac{|\boldsymbol{r}_\gamma|^2}{m_\gamma(|\boldsymbol{p}_\gamma|)} - m_\gamma'(|\boldsymbol{p}_\gamma|)\, \frac{(\boldsymbol{r}_\gamma : \boldsymbol{p}_\gamma)^2}{m_\gamma(|\boldsymbol{p}_\gamma|)^2 |\boldsymbol{p}_\gamma|} \right) \\
&\geq \widetilde{\sigma}_0\, \gamma \left(\frac{|\boldsymbol{r}_\gamma|^2}{m_\gamma(|\boldsymbol{p}_\gamma|)} - m_\gamma'(|\boldsymbol{p}_\gamma|)\, \frac{|\boldsymbol{r}_\gamma|^2 |\boldsymbol{p}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)^2} \right) \\
&\geq \frac{\widetilde{\sigma}_0\, \gamma}{m_\gamma(|\boldsymbol{p}_\gamma|)} \left(|\boldsymbol{r}_\gamma|^2 - \frac{m_\gamma'(|\boldsymbol{p}_\gamma|)}{\gamma}\, |\boldsymbol{r}_\gamma|^2 \right) \geq 0 \quad \text{by (4.4) and (4.8).}
\end{aligned} \tag{6.6}$$

Testing the regularized adjoint equation (6.1c) with its solution $\boldsymbol{Z}_\gamma = (\boldsymbol{w}_\gamma, \boldsymbol{r}_\gamma)$ it thus follows that

$$-(\boldsymbol{u}_\gamma - \boldsymbol{u}_d, \boldsymbol{w}_\gamma)_\Omega = a(\boldsymbol{Z}_\gamma, \boldsymbol{Z}_\gamma) + (\boldsymbol{\pi}_\gamma, \boldsymbol{r}_\gamma)_\Omega \geq a(\boldsymbol{Z}_\gamma, \boldsymbol{Z}_\gamma) \geq \underline{a}\, \|\boldsymbol{Z}_\gamma\|^2.$$

Therefore,

$$\underline{a}\,\|\boldsymbol{Z}_\gamma\|_Z^2 \le \|\boldsymbol{u}_\gamma - \boldsymbol{u}_d\|_V\,\|\boldsymbol{w}_\gamma\|_V \le K\|\boldsymbol{Z}_\gamma\|_Z$$

holds, which shows the boundedness of the sequence of adjoint states. Thus there exists a subsequence (denoted in the same way) such that $(\boldsymbol{w}_\gamma, \boldsymbol{r}_\gamma) = \boldsymbol{Z}_\gamma \rightharpoonup \boldsymbol{Z} = (\boldsymbol{w}, \boldsymbol{r})$ in Z. We need to show that $\boldsymbol{Z}$ satisfies (6.4d).

When we insert $\boldsymbol{Y} = (\boldsymbol{v}, \boldsymbol{0})$ with arbitrary $\boldsymbol{v} \in V$ into (6.1c) and take the limit $\gamma \to \infty$, then

$$\int_\Omega \big(\boldsymbol{\varepsilon}(\boldsymbol{w}) - \boldsymbol{r}\big) : \mathbb{C}\,\boldsymbol{\varepsilon}(\boldsymbol{v})\,\mathrm{d}x = -\int_\Omega (\boldsymbol{u} - \boldsymbol{u}_d)\cdot\boldsymbol{v}\,\mathrm{d}x \quad \text{for all } \boldsymbol{v} \in V \tag{6.7}$$

follows by weak convergence. Testing (6.1c) with $\boldsymbol{Y} = (\boldsymbol{0}, \boldsymbol{q})$ yields

$$\boldsymbol{\pi}_\gamma = \big[\mathbb{C}\,(\boldsymbol{\varepsilon}(\boldsymbol{w}_\gamma) - \boldsymbol{r}_\gamma) - \mathbb{H}\,\boldsymbol{r}_\gamma\big]^D \quad \rightharpoonup \quad \big[\mathbb{C}\,(\boldsymbol{\varepsilon}(\boldsymbol{w}) - \boldsymbol{r}) - \mathbb{H}\,\boldsymbol{r}\big]^D =: \boldsymbol{\pi} \in Q.$$

Multiplying the previous equation with $\boldsymbol{q} \in Q$, integrating over Ω and subtraction the arising equation from (6.7) then verifies the adjoint equation (6.4d).

Step 3. The sign condition (6.4h):
Let $\varphi \in C_0^\infty(\Omega)$ with $\varphi \ge 0$ be arbitrary. If one tests the regularized adjoint equation (6.1c) with $\varphi\boldsymbol{Z}_\gamma = (\varphi\,\boldsymbol{w}_\gamma, \varphi\,\boldsymbol{r}_\gamma)$, which clearly belongs to $V \times Q$, then one obtains

$$\begin{aligned}
0 &= \int_\Omega \big(\boldsymbol{\varepsilon}(\boldsymbol{w}_\gamma) - \boldsymbol{r}_\gamma\big) : \mathbb{C}\big(\boldsymbol{\varepsilon}(\varphi\,\boldsymbol{w}_\gamma) - \varphi\,\boldsymbol{r}_\gamma\big)\,\mathrm{d}x \\
&\qquad + \int_\Omega \varphi\,\boldsymbol{r}_\gamma : \mathbb{H}\,\boldsymbol{r}_\gamma\,\mathrm{d}x + \int_\Omega \varphi\,\boldsymbol{\pi}_\gamma : \boldsymbol{r}_\gamma\,\mathrm{d}x + \int_\Omega \varphi\,(\boldsymbol{u}_\gamma - \boldsymbol{u}_d)\cdot\boldsymbol{w}_\gamma\,\mathrm{d}x \\
&\ge \int_\Omega \varphi\,\Big(\big(\boldsymbol{\varepsilon}(\boldsymbol{w}_\gamma) - \boldsymbol{r}_\gamma\big) : \mathbb{C}\big(\boldsymbol{\varepsilon}(\boldsymbol{w}_\gamma) - \boldsymbol{r}_\gamma\big) + \boldsymbol{r}_\gamma : \mathbb{H}\,\boldsymbol{r}_\gamma\Big)\,\mathrm{d}x \\
&\qquad + \int_\Omega \big(\boldsymbol{\varepsilon}(\boldsymbol{w}_\gamma) - \boldsymbol{r}_\gamma\big) : \mathbb{C}\,\frac{1}{2}\big(\nabla\varphi\,\boldsymbol{w}_\gamma^\top + \boldsymbol{w}_\gamma\nabla\varphi^\top\big)\,\mathrm{d}x + \int_\Omega \varphi\,(\boldsymbol{u}_\gamma - \boldsymbol{u}_d)\cdot\boldsymbol{w}_\gamma\,\mathrm{d}x,
\end{aligned}$$

where we used $\boldsymbol{\pi}_\gamma : \boldsymbol{r}_\gamma \ge 0$ from (6.6) for the preceding estimate. Due to the coercivity of $\mathbb{C}$ and $\mathbb{H}$ and $\varphi \ge 0$, the first addend is convex and continuous w.r.t. $(\boldsymbol{w}_\gamma, \boldsymbol{r}_\gamma)$, thus weakly lower semicontinuous in $V \times Q$. Moreover, thanks to $\boldsymbol{w}_\gamma \rightharpoonup \boldsymbol{w}$ in V, we obtain $\boldsymbol{w}_\gamma \to \boldsymbol{w}$ in $L^2(\Omega;\mathbb{R}^3)$ by compact embedding. Since also $\boldsymbol{u}_\gamma \to \boldsymbol{u}$ in $L^2(\Omega;\mathbb{R}^3)$, we arrive at

$$\begin{aligned}
0 &\ge \liminf_{\gamma\to\infty} \int_\Omega \varphi\,\Big(\big(\boldsymbol{\varepsilon}(\boldsymbol{w}_\gamma) - \boldsymbol{r}_\gamma\big) : \mathbb{C}\big(\boldsymbol{\varepsilon}(\boldsymbol{w}_\gamma) - \boldsymbol{r}_\gamma\big) + \boldsymbol{r}_\gamma : \mathbb{H}\,\boldsymbol{r}_\gamma\Big)\,\mathrm{d}x \\
&\qquad + \lim_{\gamma\to\infty} \int_\Omega \big(\boldsymbol{\varepsilon}(\boldsymbol{w}_\gamma) - \boldsymbol{r}_\gamma\big) : \mathbb{C}\,\frac{1}{2}\big(\nabla\varphi\,\boldsymbol{w}_\gamma^\top + \boldsymbol{w}_\gamma\nabla\varphi^\top\big)\,\mathrm{d}x \\
&\qquad + \lim_{\gamma\to\infty} \int_\Omega \varphi\,(\boldsymbol{u}_\gamma - \boldsymbol{u}_d)\cdot\boldsymbol{w}_\gamma\,\mathrm{d}x \\
&\ge \int_\Omega \varphi\,\Big(\big(\boldsymbol{\varepsilon}(\boldsymbol{w}) - \boldsymbol{r}\big) : \mathbb{C}\big(\boldsymbol{\varepsilon}(\boldsymbol{w}) - \boldsymbol{r}\big) + \boldsymbol{r} : \mathbb{H}\,\boldsymbol{r}\Big)\,\mathrm{d}x \\
&\qquad + \int_\Omega \big(\boldsymbol{\varepsilon}(\boldsymbol{w}) - \boldsymbol{r}\big) : \mathbb{C}\,\frac{1}{2}\big(\nabla\varphi\,\boldsymbol{w}^\top + \boldsymbol{w}\nabla\varphi^\top\big)\,\mathrm{d}x + \int_\Omega \varphi\,(\boldsymbol{u} - \boldsymbol{u}_d)\cdot\boldsymbol{w}\,\mathrm{d}x \\
&= a(\boldsymbol{Z}, \varphi\boldsymbol{Z}) + \int_\Omega \varphi\,(\boldsymbol{u} - \boldsymbol{u}_d)\cdot\boldsymbol{w}\,\mathrm{d}x = -\int_\Omega \varphi\,\boldsymbol{\pi} : \boldsymbol{r}\,\mathrm{d}x,
\end{aligned}$$

where we used the definition of $a(\cdot,\cdot)$ and the adjoint equation (6.4d) for the final two equalities. Since $\varphi \ge 0$ was arbitrary, this implies $\boldsymbol{\pi} : \boldsymbol{r} \ge 0$ a.e. in Ω, which is (6.4h).

Step 4. The complementarity relation (6.4g):
Let us first notice that the definition of $\boldsymbol{\pi}_\gamma$ in (6.1d) implies

$$\int_\Omega \widetilde{\sigma}_0 \left(\frac{|\boldsymbol{r}_\gamma|^2}{m_\gamma(|\boldsymbol{p}_\gamma|)} - m_\gamma'(|\boldsymbol{p}_\gamma|) \frac{(\boldsymbol{r}_\gamma : \boldsymbol{p}_\gamma)^2}{m_\gamma(|\boldsymbol{p}_\gamma|)^2 |\boldsymbol{p}_\gamma|} \right) \mathrm{d}x \\ = \frac{1}{\gamma} \int_\Omega \boldsymbol{\pi}_\gamma : \boldsymbol{r}_\gamma \to 0 \quad \text{as } \gamma \to \infty, \tag{6.8}$$

since $\boldsymbol{\pi}_\gamma$ and $\boldsymbol{r}_\gamma$ are weakly convergent and thus bounded in Q. With this result at hand, (6.4g) follows similarly as in [de los Reyes, 2011, Theorem 5.1]. Nevertheless we include the proof, since, in contrast to [de los Reyes, 2011, Theorem 5.1], we derive here a pointwise equation.

Using the definition of $\boldsymbol{\pi}_\gamma$ in (6.1d), we obtain

$$\begin{aligned} |\boldsymbol{\pi}_\gamma : \boldsymbol{p}_\gamma| &= \widetilde{\sigma}_0\, \gamma \left| \frac{\boldsymbol{r}_\gamma : \boldsymbol{p}_\gamma}{m_\gamma(|\boldsymbol{p}_\gamma|)} - m_\gamma'(|\boldsymbol{p}_\gamma|) \frac{(\boldsymbol{r}_\gamma : \boldsymbol{p}_\gamma)|\boldsymbol{p}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)^2} \right| \\ &\le \widetilde{\sigma}_0 |\boldsymbol{r}_\gamma| \frac{\gamma\, |\boldsymbol{p}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)} \left| 1 - m_\gamma'(|\boldsymbol{p}_\gamma|) \frac{|\boldsymbol{p}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)} \right| && \text{since } m_\gamma(|\boldsymbol{p}_\gamma|) > 0 \\ &\le \widetilde{\sigma}_0 \left| |\boldsymbol{r}_\gamma| - m_\gamma'(|\boldsymbol{p}_\gamma|) \frac{|\boldsymbol{r}_\gamma|\, |\boldsymbol{p}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)} \right| && \text{by } (4.4)_2. \end{aligned}$$

Therefore, Hölder's inequality implies

$$\|\boldsymbol{\pi}_\gamma : \boldsymbol{p}_\gamma\|_{L^1(\Omega)}^2 \le \widetilde{\sigma}_0^2\, |\Omega| \int_\Omega \left(|\boldsymbol{r}_\gamma| - m_\gamma'(|\boldsymbol{p}_\gamma|) \frac{|\boldsymbol{r}_\gamma||\boldsymbol{p}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)} \right)^2 \mathrm{d}x. \tag{6.9}$$

To estimate the above integral, let us define—up to sets of zero measure—the following subsets of Ω:

$$\mathcal{A}_\gamma = \{x \in \Omega : \gamma\, |\boldsymbol{p}_\gamma| \ge \widetilde{\sigma}_0 + \tfrac{1}{2\gamma}\}, \tag{6.10a}$$
$$\mathcal{S}_\gamma = \{x \in \Omega : \big|\gamma\, |\boldsymbol{p}_\gamma| - \widetilde{\sigma}_0\big| \le \tfrac{1}{2\gamma}\} \tag{6.10b}$$
$$\mathcal{I}_\gamma = \Omega \setminus (\mathcal{A}_\gamma \cup \mathcal{S}_\gamma) = \{x \in \Omega : \gamma\, |\boldsymbol{p}_\gamma| \le \widetilde{\sigma}_0 - \tfrac{1}{2\gamma}\}. \tag{6.10c}$$

By the formulas for m_γ and m_γ', see (4.3) and (4.7), we have

$$\begin{aligned} &m_\gamma(|\boldsymbol{p}_\gamma|) = \gamma\, |\boldsymbol{p}_\gamma| \quad \text{and} \quad m_\gamma'(|\boldsymbol{p}_\gamma|) = \gamma \quad \text{a.e. in } \mathcal{A}_\gamma \\ &m_\gamma(|\boldsymbol{p}_\gamma|) = \widetilde{\sigma}_0 \qquad \text{and} \quad m_\gamma'(|\boldsymbol{p}_\gamma|) = 0 \quad \text{a.e. in } \mathcal{I}_\gamma. \end{aligned} \tag{6.11}$$

Thus we find

$$|\boldsymbol{r}_\gamma| - m_\gamma'(|\boldsymbol{p}_\gamma|) \frac{|\boldsymbol{r}_\gamma|\, |\boldsymbol{p}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)} = \left\{ \begin{matrix} 0 \\ |\boldsymbol{r}_\gamma| \end{matrix} \right\} \le \begin{cases} 0 & \text{a.e. in } \mathcal{A}_\gamma \\ \widetilde{\sigma}_0 \dfrac{|\boldsymbol{r}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)} & \text{a.e. in } \mathcal{I}_\gamma \end{cases} \tag{6.12}$$

by $(4.4)_1$. On $\mathcal{S}_\gamma$ one obtains by inserting (4.3) and (4.7) that

$$|\boldsymbol{r}_\gamma| - m_\gamma'(|\boldsymbol{p}_\gamma|)\,\frac{|\boldsymbol{r}_\gamma|\,|\boldsymbol{p}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)} = \frac{|\boldsymbol{r}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)}\big(m_\gamma(|\boldsymbol{p}_\gamma|) - m_\gamma'(|\boldsymbol{p}_\gamma|)|\boldsymbol{p}_\gamma|\big)$$
$$= \frac{|\boldsymbol{r}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)}\left(\widetilde{\sigma}_0 + \frac{\gamma}{2}\Big[\Big(\widetilde{\sigma}_0 - \frac{1}{2\gamma}\Big)^2 - \big(\gamma\,|\boldsymbol{p}_\gamma|\big)^2\Big]\right).$$

On the subset of $\mathcal{S}_\gamma$ where $\gamma\,|\boldsymbol{p}_\gamma| < \widetilde{\sigma}_0$ holds, the definition of $\mathcal{S}_\gamma$ in (6.10b) yields $\gamma\,|\boldsymbol{p}_\gamma| \geq \widetilde{\sigma}_0 - 1/(2\gamma)$. On the remaining subset of $\mathcal{S}_\gamma$, $\gamma\,|\boldsymbol{p}_\gamma| \geq \widetilde{\sigma}_0 > \widetilde{\sigma}_0 - 1/(2\gamma)$ follows. So in any case we have

$$\left(\widetilde{\sigma}_0 - \frac{1}{2\gamma}\right)^2 - \big(\gamma\,|\boldsymbol{p}_\gamma|\big)^2 \leq 0$$

a.e. in $\mathcal{S}_\gamma$, provided that $\widetilde{\sigma}_0 - 1/(2\gamma) \geq 0$, i.e. $\gamma \geq 1/(2\widetilde{\sigma}_0)$. This is not a restriction, as we will pass to the limit $\gamma \to \infty$ below, so we conclude

$$|\boldsymbol{r}_\gamma| - m_\gamma'(|\boldsymbol{p}_\gamma|)\,\frac{|\boldsymbol{r}_\gamma|\,|\boldsymbol{p}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)} \leq \widetilde{\sigma}_0\frac{|\boldsymbol{r}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)} \quad \text{a.e. in } \mathcal{S}_\gamma. \tag{6.13}$$

By combining the estimates (6.13) and (6.12), we find

$$|\boldsymbol{r}_\gamma| - m_\gamma'(|\boldsymbol{p}_\gamma|)\,\frac{|\boldsymbol{r}_\gamma|\,|\boldsymbol{p}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)} \leq \widetilde{\sigma}_0\,\frac{|\boldsymbol{r}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)} \quad \text{a.e. in } \Omega \tag{6.14}$$

for the integrand in (6.9). Furthermore the estimates $(4.4)_2$ and (4.8) yield

$$|\boldsymbol{r}_\gamma| - m_\gamma'(|\boldsymbol{p}_\gamma|)\,\frac{|\boldsymbol{r}_\gamma|\,|\boldsymbol{p}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)} \geq 0 \quad \text{a.e. in } \Omega. \tag{6.15}$$

Using these upper and lower bounds for the integrand in (6.9), together with the Cauchy-Schwarz inequality, we finally arrive at

$$\begin{aligned}
&\|\boldsymbol{\pi}_\gamma : \boldsymbol{p}_\gamma\|^2_{L^1(\Omega)} \\
&\quad\leq \widetilde{\sigma}_0^2\,|\Omega| \int_\Omega \left(|\boldsymbol{r}_\gamma| - m_\gamma'(|\boldsymbol{p}_\gamma|)\,\frac{|\boldsymbol{r}_\gamma||\boldsymbol{p}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)}\right)\widetilde{\sigma}_0\,\frac{|\boldsymbol{r}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)}\,\mathrm{d}x \\
&\quad\leq \widetilde{\sigma}_0^3\,|\Omega| \int_\Omega \left(\frac{|\boldsymbol{r}_\gamma|^2}{m_\gamma(|\boldsymbol{p}_\gamma|)} - m_\gamma'(|\boldsymbol{p}_\gamma|)\,\frac{(\boldsymbol{r}_\gamma : \boldsymbol{p}_\gamma)^2}{m_\gamma(|\boldsymbol{p}_\gamma|)^2|\boldsymbol{p}_\gamma|}\right)\mathrm{d}x.
\end{aligned} \tag{6.16}$$

Consequently, (6.8) implies

$$\|\boldsymbol{\pi}_\gamma : \boldsymbol{p}_\gamma\|_{L^1(\Omega)} \to 0 \quad \text{as } \gamma \to \infty.$$

Moreover, since $\boldsymbol{\pi}_\gamma \rightharpoonup \boldsymbol{\pi}$ in Q and $\boldsymbol{p}_\gamma \to \boldsymbol{p}$ in Q, we conclude $\boldsymbol{\pi}_\gamma : \boldsymbol{p}_\gamma \rightharpoonup \boldsymbol{\pi} : \boldsymbol{p}$ in $L^1(\Omega)$. The uniqueness of the weak limit then confirms that $\boldsymbol{\pi} : \boldsymbol{p} = 0$ must hold a.e. in Ω, which is (6.4g).

Step 5. The complementarity relation (6.4i):
By choosing $\boldsymbol{v} = \boldsymbol{0}$ in (6.1a) we obtain $\boldsymbol{\varrho}_\gamma = \big[\mathbb{C}\,(\boldsymbol{\varepsilon}(\boldsymbol{u}_\gamma) - \boldsymbol{p}_\gamma) - \mathbb{H}\,\boldsymbol{p}_\gamma\big]^D$. Since

$(\boldsymbol{u}_\gamma, \boldsymbol{p}_\gamma) \to (\boldsymbol{u}, \boldsymbol{p})$ strongly in Z, this shows $\boldsymbol{\varrho}_\gamma \to \boldsymbol{\varrho}$ strongly in Q. Therefore the weak convergence of $\boldsymbol{r}_\gamma$ in Q (see step 2) implies

$$\int_\Omega |\boldsymbol{r}_\gamma|(|\boldsymbol{\varrho}_\gamma| - |\boldsymbol{\varrho}|)\,\mathrm{d}x \to 0 \quad \text{as } \gamma \to \infty. \tag{6.17}$$

Due to (6.4c) the mapping $Q \ni \boldsymbol{r} \mapsto \int_\Omega |\boldsymbol{r}|(\widetilde{\sigma}_0 - |\boldsymbol{\rho}|)\,\mathrm{d}x \in \mathbb{R}$ is convex and continuous and thus weakly lower semicontinuous. Consequently we obtain

$$\begin{aligned}
0 &\le \int_\Omega |\boldsymbol{r}|(\widetilde{\sigma}_0 - |\boldsymbol{\varrho}|)\,\mathrm{d}x \\
&\le \liminf_{\gamma\to\infty} \int_\Omega |\boldsymbol{r}_\gamma|(\widetilde{\sigma}_0 - |\boldsymbol{\varrho}|)\,\mathrm{d}x \\
&\le \limsup_{\gamma\to\infty} \int_\Omega |\boldsymbol{r}_\gamma|(\widetilde{\sigma}_0 - |\boldsymbol{\varrho}_\gamma|)\,\mathrm{d}x + \limsup_{\gamma\to\infty} \int_\Omega |\boldsymbol{r}_\gamma|(|\boldsymbol{\varrho}_\gamma| - |\boldsymbol{\varrho}|)\,\mathrm{d}x \\
&= \limsup_{\gamma\to\infty} \int_\Omega \left(\widetilde{\sigma}_0\,|\boldsymbol{r}_\gamma| - |\boldsymbol{r}_\gamma||\boldsymbol{\varrho}_\gamma|\right)\mathrm{d}x && \text{by (6.17)} \\
&= \limsup_{\gamma\to\infty} \widetilde{\sigma}_0 \int_\Omega \left(|\boldsymbol{r}_\gamma| - \gamma \frac{|\boldsymbol{r}_\gamma||\boldsymbol{p}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)}\right)\mathrm{d}x && \text{by (6.1b)} \\
&\le \limsup_{\gamma\to\infty} \widetilde{\sigma}_0 \int_\Omega \left(|\boldsymbol{r}_\gamma| - m_\gamma'(|\boldsymbol{p}_\gamma|) \frac{|\boldsymbol{r}_\gamma||\boldsymbol{p}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)}\right)\mathrm{d}x && \text{by (4.8)} \\
&\le \limsup_{\gamma\to\infty} \widetilde{\sigma}_0\,|\Omega|^{1/2} \left(\int_\Omega \left(|\boldsymbol{r}_\gamma| - m_\gamma'(|\boldsymbol{p}_\gamma|) \frac{|\boldsymbol{r}_\gamma||\boldsymbol{p}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)}\right)^2 \mathrm{d}x\right)^{1/2},
\end{aligned}$$

where we again applied Hölder's inequality for the last estimate. So we end up with the expression known from (6.9), which can be estimated by (6.16) as seen before, and we conclude

$$\int_\Omega |\boldsymbol{r}|(\widetilde{\sigma}_0 - |\boldsymbol{\varrho}|)\,\mathrm{d}x = 0.$$

The non-negativity of the integrand now shows the pointwise complementarity relation $|\boldsymbol{r}|(\widetilde{\sigma}_0 - |\boldsymbol{\varrho}|) = 0$ a.e. in Ω. Together with $|\boldsymbol{\varrho}| \le \widetilde{\sigma}_0$ from (6.4c), this finally implies (6.4i).

Step 6. The gradient equations (6.4e) and (6.4f):
By passing to the limit in (6.1e) and (6.1f) we finally obtain (6.4e) and (6.4f).

This concludes the proof under the assumption initially made, namely that $(\boldsymbol{f}, \boldsymbol{g}) \in L^2(\Omega; \mathbb{R}^3) \times L^2(\Gamma_N; \mathbb{R}^3)$ is being approximated by the sequence $(\boldsymbol{f}_\gamma, \boldsymbol{g}_\gamma)$. As was mentioned in Remark 4.5, to achieve the same in the general case, the regularized problems (4.1) need to be modified by the additional term (4.12) in the objective. The optimality systems for problems undergo the obvious change that (6.4e) and (6.4f) are replaced by $\nu_1\,\boldsymbol{f}_\gamma + r\,(\boldsymbol{f}_\gamma - \boldsymbol{f}^*) - \boldsymbol{w}_\gamma = \boldsymbol{0}$ and $\nu_2\,\boldsymbol{g}_\gamma + r\,(\boldsymbol{g}_\gamma - \boldsymbol{g}^*) - \boldsymbol{w}_\gamma = \boldsymbol{0}$. In the limit, the conditions are again (6.4e) and (6.4f), respectively. □

For easy reference, we summarize in Table 6.1 the variables associated with (4.1) and their convergence as $\gamma \to \infty$.

Remark 6.3. *We would like to point out some differences to optimality conditions for other optimal control problems governed by VIs of second kind obtained in [de los*

variable	definition	convergence
$\boldsymbol{f}_\gamma \in L^2(\Omega;\mathbb{R}^3)$	local solution to (4.1)	strongly in $L^2(\Omega;\mathbb{R}^3)$
$\boldsymbol{g}_\gamma \in L^2(\Gamma_N;\mathbb{R}^3)$	local solution to (4.1)	strongly in $L^2(\Gamma_N;\mathbb{R}^3)$
$\boldsymbol{u}_\gamma \in V$	local solution to (4.1)	strongly in V
$\boldsymbol{p}_\gamma \in Q$	local solution to (4.1)	strongly in Q
$\boldsymbol{w}_\gamma \in V$	Theorem 6.1	weakly in V
$\boldsymbol{r}_\gamma \in Q$	Theorem 6.1	weakly in Q
$\boldsymbol{\pi}_\gamma \in Q$	Theorem 6.1	weakly in Q
$\boldsymbol{\varrho}_\gamma \in Q$	Theorem 6.1	strongly in Q
$\vartheta_\gamma \in L^2(\Omega)$	Lemma 6.4	weakly in $L^2(\Omega)$

Table 6.1

Variables associated with the regularized control problem (4.1) *and their convergence when* $\gamma \to \infty$*, as proved in Proposition 6.2 and Lemma 6.4.*

Reyes, 2011, Theorem 5.1] and [de los Reyes, 2012, Theorem 4.1]. First of all, notice that the complementarity relations in (6.4h)–(6.4i) *hold in a* pointwise *sense in contrast to the results mentioned above. Apart from that, further structural differences to the above arise when we further exploit the specific setting of our problem in the sequel. It turns out that we can refine the result of Proposition 6.2 by introducing a new scalar valued multiplier* ϑ*, which gives more structure to the adjoint plastic strain* $\boldsymbol{r}$ *and leads to a strengthening of conditions* (6.4h) *and* (6.4i)*, so that we finally arrive at the optimality system presented in Theorem 1.1, eq.* (1.4)*. This step is essential in proving the equivalence of* (1.4) *to the C-stationarity system for the equivalent dual formulation of the problem in Section 7.*

Lemma 6.4. *Under the conditions of Proposition 6.2, there exists a multiplier* $\vartheta \in L^2(\Omega)$ *such that the adjoint plastic strain can be decomposed as follows:*

$$\boldsymbol{r} = \frac{1}{\widetilde{\sigma}_0}\big(|\boldsymbol{p}|\boldsymbol{\pi} + \vartheta\boldsymbol{\varrho}\big). \tag{6.18}$$

Proof. From (6.1b) and (6.1d) it follows that $\boldsymbol{r}_\gamma$ satisfies

$$\boldsymbol{r}_\gamma = \frac{1}{\widetilde{\sigma}_0}\left(\frac{m_\gamma(|\boldsymbol{p}_\gamma|)}{\gamma}\,\boldsymbol{\pi}_\gamma + \frac{m'_\gamma(\boldsymbol{p}_\gamma)}{\gamma}\,\frac{\boldsymbol{r}_\gamma : \boldsymbol{p}_\gamma}{|\boldsymbol{p}_\gamma|}\,\boldsymbol{\varrho}_\gamma\right).$$

By introducing the approximate multiplier

$$\vartheta_\gamma := \frac{m'_\gamma(\boldsymbol{p}_\gamma)}{\gamma}\,\frac{\boldsymbol{r}_\gamma : \boldsymbol{p}_\gamma}{|\boldsymbol{p}_\gamma|} \tag{6.19}$$

it follows from (4.8) and the boundedness of $\boldsymbol{r}_\gamma$ (see after (6.8)) that

$$\|\vartheta_\gamma\|_{L^2(\Omega)} \le \|\boldsymbol{r}_\gamma\|_S \le K < \infty.$$

Consequently, up to a subsequence, $\vartheta_\gamma \rightharpoonup \vartheta$ in $L^2(\Omega)$ for some $\vartheta \in L^2(\Omega)$.

Let us now define the auxiliary quantity $\alpha_\gamma \in L^2(\Omega)$ by

$$\alpha_\gamma = \frac{m_\gamma(|\boldsymbol{p}_\gamma|)}{\gamma} = \frac{\widetilde{\sigma}_0}{\gamma} + \begin{cases} |\boldsymbol{p}_\gamma| - \frac{\widetilde{\sigma}_0}{\gamma} & \text{a.e. in } \mathcal{A}_\gamma, \\ \frac{1}{2}(\gamma\,|\boldsymbol{p}_\gamma| - \widetilde{\sigma}_0 + \frac{1}{2\gamma})^2 & \text{a.e. in } \mathcal{S}_\gamma, \\ 0 & \text{a.e. in } \mathcal{I}_\gamma, \end{cases} \tag{6.20}$$

where $\mathcal{A}_\gamma$, $\mathcal{S}_\gamma$, and $\mathcal{I}_\gamma$ are as defined in (6.10). The definition of m_γ in (4.3) was used here.

We now verify that $\alpha_\gamma \to |\boldsymbol{p}|$ strongly in $L^2(\Omega)$. First Young's inequality yields

$$\begin{aligned}\int_\Omega (\alpha_\gamma - |\boldsymbol{p}|)^2 \,\mathrm{d}x &= \int_\Omega \left(\frac{m_\gamma(|\boldsymbol{p}_\gamma|)}{\gamma} - |\boldsymbol{p}_\gamma| + |\boldsymbol{p}_\gamma| - |\boldsymbol{p}| \right)^2 \mathrm{d}x \\ &\le 2 \int_\Omega \left(\frac{m_\gamma(|\boldsymbol{p}_\gamma|)}{\gamma} - |\boldsymbol{p}_\gamma| \right)^2 \mathrm{d}x + 2 \big\| |\boldsymbol{p}_\gamma| - |\boldsymbol{p}| \big\|^2_{L^2(\Omega)}. \end{aligned} \tag{6.21}$$

From the representation (6.20) we get

$$\frac{m_\gamma(|\boldsymbol{p}_\gamma|)}{\gamma} - |\boldsymbol{p}_\gamma| = \begin{cases} 0, & \text{a.e. in } \mathcal{A}_\gamma \\ \frac{\widetilde{\sigma}_0}{\gamma} - |\boldsymbol{p}_\gamma|, & \text{a.e. in } \mathcal{I}_\gamma. \end{cases} \tag{6.22}$$

On $\mathcal{S}_\gamma$ we have by (6.20) and Young's inequality

$$\left(\frac{m_\gamma(|\boldsymbol{p}_\gamma|)}{\gamma} - |\boldsymbol{p}_\gamma| \right)^2 \le 2 \left(\frac{\widetilde{\sigma}_0}{\gamma} - |\boldsymbol{p}_\gamma| \right)^2 + \frac{1}{2} \left(\gamma\, |\boldsymbol{p}_\gamma| - \widetilde{\sigma}_0 + \frac{1}{2\gamma} \right)^4 \quad \text{a.e. in } \mathcal{S}_\gamma. \tag{6.23}$$

Moreover the definitions of $\mathcal{S}_\gamma$ and $\mathcal{I}_\gamma$, respectively, in (6.10) immediately yield

$$\left| \gamma\, |\boldsymbol{p}_\gamma| - \widetilde{\sigma}_0 + \frac{1}{2\gamma} \right| \le \left| \gamma\, |\boldsymbol{p}_\gamma| - \widetilde{\sigma}_0 \right| + \frac{1}{2\gamma} \le \frac{1}{\gamma} \quad \text{a.e. in } \mathcal{S}_\gamma \tag{6.24a}$$

and

$$\left| \frac{\widetilde{\sigma}_0}{\gamma} - |\boldsymbol{p}_\gamma| \right| \le \frac{\widetilde{\sigma}_0}{\gamma} + |\boldsymbol{p}_\gamma| \le \frac{2\,\widetilde{\sigma}_0}{\gamma} - \frac{1}{2\gamma^2} \le \frac{2\,\widetilde{\sigma}_0}{\gamma} \quad \text{a.e. in } \mathcal{I}_\gamma. \tag{6.24b}$$

By inserting the estimates (6.22) and (6.23) into (6.21) we continue with

$$\begin{aligned} &\int_\Omega (\alpha_\gamma - |\boldsymbol{p}|)^2 \,\mathrm{d}x \\ &\qquad \le 4 \int_{S_\gamma} \left(\frac{\widetilde{\sigma}_0}{\gamma} - |\boldsymbol{p}_\gamma| \right)^2 \mathrm{d}x + \int_{S_\gamma} \left(\gamma\, |\boldsymbol{p}_\gamma| - \widetilde{\sigma}_0 + \frac{1}{2\gamma} \right)^4 \mathrm{d}x \\ &\qquad\quad + 2 \int_{I_\gamma} \left(\frac{\widetilde{\sigma}_0}{\gamma} - |\boldsymbol{p}_\gamma| \right)^2 \mathrm{d}x + 2\, \|\boldsymbol{p}_\gamma - \boldsymbol{p}\|_S^2 \\ &\qquad \le 4\, |\mathcal{S}_\gamma|\, \frac{1}{4\gamma^4} + |\mathcal{S}_\gamma|\, \frac{1}{\gamma^4} + 2\, |\mathcal{I}_\gamma|\, \frac{4\, \widetilde{\sigma}_0^2}{\gamma^2} + 2\, \|\boldsymbol{p}_\gamma - \boldsymbol{p}\|_S^2 \quad \text{by (6.24a) and (6.24b)} \\ &\qquad \le c\, |\Omega| \left(\frac{1}{\gamma^4} + \frac{1}{\gamma^2} \right) + 2\, \|\boldsymbol{p}_\gamma - \boldsymbol{p}\|_S^2 \to 0 \quad \text{as } \gamma \to \infty. \end{aligned}$$

Since $\boldsymbol{\varrho}_\gamma \to \boldsymbol{\varrho}$ strongly in Q (see Step 5 in the proof of Proposition 6.2), $\vartheta_\gamma \rightharpoonup \vartheta$ weakly in $L^2(\Omega)$ as shown above, $\boldsymbol{\pi}_\gamma \rightharpoonup \boldsymbol{\pi}$ weakly in S (see end of Step 4 in the proof of Proposition 6.2), and $\alpha_\gamma \to |\boldsymbol{p}|$ strongly in $L^2(\Omega)$ as just shown, we finally obtain

$$\widetilde{\sigma}_0\, \boldsymbol{r}_\gamma = \alpha_\gamma\, \boldsymbol{\pi}_\gamma + \vartheta_\gamma\, \boldsymbol{\varrho}_\gamma \quad \rightharpoonup \quad |\boldsymbol{p}|\, \boldsymbol{\pi} + \vartheta \boldsymbol{\varrho} \quad \text{in } L^1(\Omega; \mathbb{S})$$

so that $\widetilde{\sigma}_0\, \boldsymbol{r} = |\boldsymbol{p}|\, \pi + \vartheta \boldsymbol{\varrho}$ holds as claimed. □

With the help of the above lemma, we can now finalize the proof of Theorem 1.1.

Proof of Theorem 1.1. *Proof.* Let $(\boldsymbol{f}, \boldsymbol{g}) \in L^2(\Omega;\mathbb{R}^3) \times L^2(\Gamma_N;\mathbb{R}^3)$ be a locally optimal solution for (1.3) with associated state $\boldsymbol{W} = (\boldsymbol{u}, \boldsymbol{p}) \in Z$. In view of Proposition 6.2, the preliminary optimality system (6.4) holds. A comparison between (6.4) and (1.4) shows that we only have to verify (1.4f)–(1.4h).

First we note that (6.4b) and (6.4c) imply

$$\boldsymbol{\varrho}(x) : \boldsymbol{p}(x) = |\boldsymbol{\varrho}(x)||\boldsymbol{p}(x)|, \tag{6.25}$$

i.e., the alignment of $\boldsymbol{\varrho}$ and $\boldsymbol{p}$, which will be useful in the course of the proof.

Lemma 6.4 gives the existence of $\widetilde{\vartheta} \in L^2(\Omega)$ such that $\widetilde{\sigma}_0\, \boldsymbol{r} = |\boldsymbol{p}|\boldsymbol{\pi} + \widetilde{\vartheta}\boldsymbol{\varrho}$ holds. If we define

$$\vartheta(x) := \begin{cases} \widetilde{\vartheta}(x), & \text{if } \boldsymbol{\varrho}(x) \neq \boldsymbol{0}, \\ 0, & \text{if } \boldsymbol{\varrho}(x) = \boldsymbol{0}, \end{cases} \tag{6.26}$$

then we obtain $\vartheta \in L^2(\Omega)$ and still $\widetilde{\sigma}_0\, \boldsymbol{r} = |\boldsymbol{p}|\boldsymbol{\pi} + \vartheta\boldsymbol{\varrho}$ holds, so (1.4f) is shown.

Next we show that $\vartheta = 0$ holds on the set $\mathcal{I}$ (where $|\boldsymbol{\varrho}| < \widetilde{\sigma}_0$), which is (1.4h). Equation (6.4b) implies $|\boldsymbol{p}| = 0$ a.e. in $\mathcal{I}$, while (6.4i) shows $\boldsymbol{r} = \boldsymbol{0}$ on $\mathcal{I}$. Now (1.4f) implies $\vartheta\boldsymbol{\varrho} = \boldsymbol{0}$ a.e. in $\mathcal{I}$. Due to (6.26) this gives (1.4h).

It remains to prove (1.4g), i.e., $\vartheta\boldsymbol{\varrho} : \boldsymbol{\pi} \geq 0$. If $\boldsymbol{p}(x) \neq \boldsymbol{0}$, then (6.25) gives that $\boldsymbol{\varrho}(x) = \kappa\, \boldsymbol{p}(x)$ with some $\kappa \in \mathbb{R}$ and we obtain from (6.4g) that $\boldsymbol{\pi}(x) = \boldsymbol{0}$ and thus in particular

$$\vartheta(x)\, \boldsymbol{\varrho}(x) : \boldsymbol{\pi}(x) = \vartheta(x)\, \kappa\, \boldsymbol{p}(x) : \boldsymbol{\pi}(x) = 0$$

holds. If, on the other hand, $\boldsymbol{p}(x) = \boldsymbol{0}$ holds, then (1.4f) and (6.4h) imply

$$\vartheta(x)\, \boldsymbol{\varrho}(x) : \boldsymbol{\pi}(x) = \widetilde{\sigma}_0\, \boldsymbol{r}(x) : \boldsymbol{\pi}(x) \geq 0.$$

Thus in any case (1.4g) holds, which concludes the proof. □

Remark 6.5. *Assume that* (1.4) *holds. Then* (1.4b) *implies* $\boldsymbol{p} = \boldsymbol{0}$ *a.e. in* $\mathcal{I}$. *Hence* (1.4h) *gives*

$$\boldsymbol{r} = \frac{1}{\widetilde{\sigma}_0}\big(|\boldsymbol{p}|\boldsymbol{\pi} + \vartheta\boldsymbol{\varrho}\big) = \boldsymbol{0} \quad \textit{a.e. in } \mathcal{I}, \tag{6.27}$$

i.e. (6.4i). *Moreover, by* (1.4g) *we obtain*

$$\boldsymbol{\pi} : \boldsymbol{r} = \frac{1}{\widetilde{\sigma}_0}\big(|\boldsymbol{p}|\, |\boldsymbol{\pi}|^2 + \vartheta\boldsymbol{\varrho} : \boldsymbol{\pi}\big) \geq 0,$$

i.e. (6.4h). *Thus* (1.4) *implies* (6.4).

7. Equivalence to C-stationarity. As was mentioned in the introduction, the strength of the optimality system obtained in Theorem 1.1 is not obvious to gauge. This is due to the lack of a classification scheme for optimality conditions of MPECs which involve variational inequalities of the second kind. We prove in this section that the result of Theorem 1.1 is in fact equivalent to C-stationarity of another optimal control problem closely related and equivalent to (1.3). That problem is obtained when the primal formulation of elastoplasticity (1.1) is replaced by the equivalent dual formulation (3.5), a variational inequality of first kind, see Lemma 3.2. To our

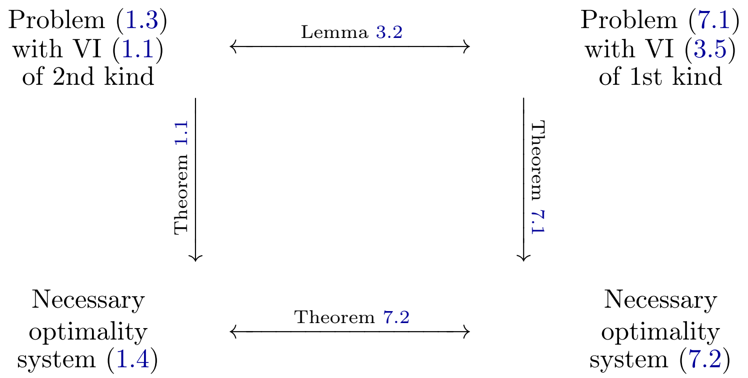

Fig. 7.1. *Relation between optimal control problems* (1.3) *and* (7.1) *and their first-order optimality systems.*

best knowledge, this is the first time that an optimality system for an MPEC with a VI of the second kind is being classified in this sense. Figure 7.1 illustrates the situation.

By the replacement of the primal by the dual VI, we obtain the following optimal control problem

$$\text{Minimize} \quad J(\boldsymbol{u}, \boldsymbol{f}, \boldsymbol{g}) \quad \text{s.t.} \quad (3.5), \tag{7.1}$$

which is equivalent to (1.3). Since (3.5) is a VI of first kind, it can be reformulated by means of a complementarity system involving a Lagrange multiplier $\lambda \in L^2(\Omega)$, the so-called plastic multiplier, see (7.2a)–(7.2c) below. For a rigorous proof of the existence and uniqueness of λ, we refer to [Herzog et al., 2011b, Theorem 1.4 and Corollary 1.2]. Problem (7.1) can thus be considered an MPCC (mathematical programs with complementarity constraints). The following first-order optimality conditions for (7.1) are proved in [Herzog et al., 2012, Theorem 3.16] by means of a Moreau-Yosida based regularization of (3.5) and a subsequent limit analysis:

Theorem 7.1 (C-stationarity). *Let* $(\boldsymbol{u}, \boldsymbol{\sigma}, \boldsymbol{\chi}, \lambda, \boldsymbol{f}, \boldsymbol{g}) \in V \times S \times S \times L^2(\Omega) \times L^2(\Omega;\mathbb{R}^3) \times L^2(\Gamma_N;\mathbb{R}^3)$ *be a locally optimal solution of* (7.1)*. Then there exist adjoint stresses* $(\boldsymbol{\zeta}, \boldsymbol{\psi}) \in S \times S$ *and displacement* $\boldsymbol{w} \in V$*, and a multiplier* $\theta \in L^2(\Omega)$ *such that the following optimality system is satisfied:*

$$\left.\begin{aligned} &\int_\Omega \boldsymbol{\sigma} : \mathbb{C}^{-1}\boldsymbol{\tau}\,\mathrm{d}x + \int_\Omega \boldsymbol{\chi} : \mathbb{H}^{-1}\boldsymbol{\mu}\,\mathrm{d}x - \int_\Omega \boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\tau}\,\mathrm{d}x \\ &\qquad + \int_\Omega \lambda\,(\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D) : (\boldsymbol{\tau}^D + \boldsymbol{\mu}^D)\,\mathrm{d}x = 0 \quad \textit{for all } \boldsymbol{\tau}, \boldsymbol{\mu} \in S \end{aligned}\right\} \tag{7.2a}$$

$$\int_\Omega \boldsymbol{\sigma} : \boldsymbol{\varepsilon}(\boldsymbol{v})\,\mathrm{d}x = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}\,\mathrm{d}x \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{v}\,\mathrm{d}s \quad \textit{for all } \boldsymbol{v} \in V \tag{7.2b}$$

$$0 \le \lambda \quad \perp \quad |\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D|^2 - \widetilde{\sigma}_0^2 \le 0 \quad \textit{a.e. in } \Omega \tag{7.2c}$$

$$
\left.
\begin{aligned}
&\int_\Omega \boldsymbol{\zeta} : \mathbb{C}^{-1}\boldsymbol{\tau}\,\mathrm{d}x + \int_\Omega \boldsymbol{\psi} : \mathbb{H}^{-1}\boldsymbol{\mu}\,\mathrm{d}x - \int_\Omega \boldsymbol{\varepsilon}(\boldsymbol{w}) : \boldsymbol{\tau}\,\mathrm{d}x \\
&\qquad + \int_\Omega \lambda\,(\boldsymbol{\zeta}^D + \boldsymbol{\psi}^D) : (\boldsymbol{\tau}^D + \boldsymbol{\mu}^D)\,\mathrm{d}x \\
&\qquad + \int_\Omega \theta\,(\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D) : (\boldsymbol{\tau}^D + \boldsymbol{\mu}^D)\,\mathrm{d}x = 0 \quad \textit{for all } \boldsymbol{\tau}, \boldsymbol{\mu} \in S
\end{aligned}
\right\} \tag{7.2d}
$$

$$
-\int_\Omega \boldsymbol{\zeta} : \boldsymbol{\varepsilon}(\boldsymbol{v})\,\mathrm{d}x = +\int_\Omega (\boldsymbol{u} - \boldsymbol{u}_d)\cdot \boldsymbol{v}\,\mathrm{d}x \quad \textit{for all } \boldsymbol{v} \in V \tag{7.2e}
$$

$$
\nu_1\,\boldsymbol{f} - \boldsymbol{w} = \boldsymbol{0} \quad \textit{a.e. in } \Omega, \qquad \nu_2\,\boldsymbol{g} - \boldsymbol{w} = \boldsymbol{0} \quad \textit{a.e. on } \Gamma_N \tag{7.2f}
$$

$$
\lambda\,(\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D) : (\boldsymbol{\zeta}^D + \boldsymbol{\psi}^D) = 0 \quad \textit{a.e. in } \Omega \tag{7.2g}
$$

$$
\theta\,\big(|\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D|^2 - \widetilde{\sigma}_0^2\big) = 0 \quad \textit{a.e. in } \Omega \tag{7.2h}
$$

$$
\theta\,(\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D) : (\boldsymbol{\zeta}^D + \boldsymbol{\psi}^D) \ge 0 \quad \textit{a.e. in } \Omega. \tag{7.2i}
$$

For convenience, we list in Table 7.1 the variables pertaining to the dual formulation. The notation is the same as in Herzog et al. [2012] with the minor exception that here, the multiplier μ associated with the non-negativity constraint for the plastic multiplier is not introduced as an extra variable. Instead, it has been replaced by its governing equation, $\mu = (\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D) : (\boldsymbol{\zeta}^D + \boldsymbol{\psi}^D)$. Moreover, in order to comply with the sign of the adjoint displacement $\boldsymbol{w}$ in (1.4), we needed to change the sign of all adjoint states and multipliers ($\boldsymbol{\zeta}$, $\boldsymbol{\psi}$, $\boldsymbol{w}$ and θ) appearing in [Herzog et al., 2012, Theorem 3.16].

	space	state variables	test functions	adjoint variables
displacement	V	$\boldsymbol{u}$	$\boldsymbol{v}$	$\boldsymbol{w}$
stress	S	$\boldsymbol{\sigma}$	$\boldsymbol{\tau}$	$\boldsymbol{\zeta}$
backstress	S	$\boldsymbol{\chi}$	$\boldsymbol{\mu}$	$\boldsymbol{\psi}$
generalized stresses	$S \times S$	$\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi})$	$\boldsymbol{T} = (\boldsymbol{\tau}, \boldsymbol{\mu})$	$\boldsymbol{\Upsilon} = (\boldsymbol{\zeta}, \boldsymbol{\psi})$
		dual variables		
plastic multiplier	$L^2(\Omega)$	$\lambda \ge 0$		μ
yield condition		$\phi(\boldsymbol{\Sigma}) = \big(|\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D|^2 - \widetilde{\sigma}_0^2\big)/2 \le 0$		θ

Table 7.1
Variables in the dual formulation.

The optimality system in (7.2) is known as the system of C-stationarity. Note that, as usual for MPCCs, the optimality system does not involve a multiplier for the complementarity relation in (7.2c). Moreover, it is characteristic for C-stationarity that a sign condition is only known for the product in (7.2i), and not for each term individually. In the following we will show that the optimality systems (1.4) and (7.2) are indeed equivalent.

Theorem 7.2. *Let $(\boldsymbol{u}, \boldsymbol{p}, \boldsymbol{f}, \boldsymbol{g}) \in V \times Q \times L^2(\Omega;\mathbb{R}^3) \times L^2(\Gamma_N;\mathbb{R}^3)$ satisfy the primal optimality system (1.4) with multipliers $(\boldsymbol{w}, \boldsymbol{r}, \boldsymbol{\varrho}, \boldsymbol{\pi}, \vartheta) \in V \times Q \times Q \times Q \times L^2(\Omega)$.*

Define

$$\boldsymbol{\sigma} = \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{p}), \quad \boldsymbol{\chi} = -\mathbb{H}\,\boldsymbol{p}, \quad \lambda = \frac{|\boldsymbol{p}|}{\widetilde{\sigma}_0}, \tag{7.3a}$$

$$\boldsymbol{\zeta} = \mathbb{C}\,(\boldsymbol{\varepsilon}(\boldsymbol{w}) - \boldsymbol{r}), \quad \boldsymbol{\psi} = -\mathbb{H}\,\boldsymbol{r}, \quad \theta = \frac{\vartheta}{\widetilde{\sigma}_0}. \tag{7.3b}$$

Then $(\boldsymbol{u}, \boldsymbol{\sigma}, \boldsymbol{\chi}, \lambda, \boldsymbol{f}, \boldsymbol{g}, \boldsymbol{w}, \boldsymbol{\zeta}, \boldsymbol{\psi}, \theta)$ *fulfill the C-stationarity conditions in* (7.2)*.*

Let, on the other hand, $(\boldsymbol{u}, \boldsymbol{\sigma}, \boldsymbol{\chi}, \lambda, \boldsymbol{f}, \boldsymbol{g}) \in V \times S \times S \times L^2(\Omega) \times L^2(\Omega;\mathbb{R}^3) \times L^2(\Gamma_N;\mathbb{R}^3)$ *fulfill* (7.2) *with multipliers* $(\boldsymbol{w}, \boldsymbol{\zeta}, \boldsymbol{\psi}, \theta) \in V \times S \times S \times L^2(\Omega)$*. If we define*

$$\boldsymbol{p} = \boldsymbol{\varepsilon}(\boldsymbol{u}) - \mathbb{C}^{-1}\boldsymbol{\sigma}, \quad \boldsymbol{\varrho} = \boldsymbol{\sigma}^D + \boldsymbol{\chi}^D, \tag{7.4a}$$

$$\boldsymbol{r} = \boldsymbol{\varepsilon}(\boldsymbol{w}) - \mathbb{C}^{-1}\boldsymbol{\zeta}, \quad \boldsymbol{\pi} = \boldsymbol{\zeta}^D + \boldsymbol{\psi}^D, \quad \vartheta = \widetilde{\sigma}_0\,\theta, \tag{7.4b}$$

then $(\boldsymbol{u}, \boldsymbol{p}, \boldsymbol{f}, \boldsymbol{g}, \boldsymbol{w}, \boldsymbol{r}, \boldsymbol{\varrho}, \boldsymbol{\pi}, \vartheta)$ *satisfies* (1.4)*.*

Proof. We first assume that (1.4) holds and show (7.2).

Step 1. The state system (7.2a)–(7.2c):
Employing the definition of $\boldsymbol{\sigma}$ in (7.3a) and testing (1.4a) with $(\boldsymbol{0}, \boldsymbol{v})$, $\boldsymbol{v} \in V$ arbitrary, immediately gives (7.2b). Taking $\boldsymbol{v} = \boldsymbol{0}$ and $\boldsymbol{q} \in Q$ arbitrary in (1.4a), the definition of $\boldsymbol{\sigma}$ and $\boldsymbol{\chi}$ in (7.3a) yields

$$\int_\Omega (\boldsymbol{\varrho} - \boldsymbol{\sigma} - \boldsymbol{\chi}) : \boldsymbol{q}\,\mathrm{d}x = 0 \quad \text{for all } \boldsymbol{q} \in Q,$$

which implies

$$\boldsymbol{\varrho} = \boldsymbol{\varrho}^D = \boldsymbol{\sigma}^D + \boldsymbol{\chi}^D, \tag{7.5}$$

since Q consists of all trace-free (purely deviatoric) tensor functions in S. Thus (1.4b) yields $|\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D| \le \widetilde{\sigma}_0$ a.e. in Ω. Next we show

$$\lambda\,(\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D) = \boldsymbol{p} \quad \text{a.e. in } \Omega. \tag{7.6}$$

Thanks to the definition of λ in (7.3a) this is obviously true for $\boldsymbol{p}(x) = \boldsymbol{0}$. To show the relation where $\boldsymbol{p}(x) \neq 0$, employ again the definition of λ and (7.5), which give

$$\lambda\,(\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D) = \frac{|\boldsymbol{p}|}{\widetilde{\sigma}_0}\,\boldsymbol{\varrho}. \tag{7.7}$$

Since $|\boldsymbol{\varrho}(x)| = \widetilde{\sigma}_0$ holds a.e. where $\boldsymbol{p}(x) \neq \boldsymbol{0}$, (6.25) implies

$$\boldsymbol{p}(x) = \frac{|\boldsymbol{p}(x)|}{|\boldsymbol{\varrho}(x)|}\,\boldsymbol{\varrho}(x) = \frac{|\boldsymbol{p}(x)|}{\widetilde{\sigma}_0}\,\boldsymbol{\varrho}(x) \quad \text{a.e. in } \{x \in \Omega : \boldsymbol{p}(x) \neq \boldsymbol{0}\}.$$

Together with (7.7), this yields (7.6). Using (7.6) and the definition of λ immediately gives $\lambda\left(|\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D| - \widetilde{\sigma}_0\right) = 0$ so that the complementarity system in (7.2c) is verified. Note that the sign condition on λ follows from its definition in (7.3a).

By solving the definitions of $\boldsymbol{\sigma}$ and $\boldsymbol{\chi}$ in (7.3a) for $\boldsymbol{p}$, we find $\boldsymbol{p} = \boldsymbol{\varepsilon}(\boldsymbol{u}) - \mathbb{C}^{-1}\boldsymbol{\sigma}$ and $\boldsymbol{p} = -\mathbb{H}^{-1}\boldsymbol{\chi}$. By multiplying the first equation with an arbitrary $\boldsymbol{\tau} \in S$ and the second with an arbitrary $\boldsymbol{\mu} \in S$, integrating over Ω, and using that $\boldsymbol{p}$ is trace-free, we arrive at

$$\int_\Omega \boldsymbol{\sigma} : \mathbb{C}^{-1}\boldsymbol{\tau}\,\mathrm{d}x + \int_\Omega \boldsymbol{\chi} : \mathbb{H}^{-1}\boldsymbol{\mu}\,\mathrm{d}x - \int_\Omega \boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\tau}\,\mathrm{d}x + \int_\Omega \boldsymbol{p} : (\boldsymbol{\tau}^D + \boldsymbol{\mu}^D)\,\mathrm{d}x = 0$$

for all $\boldsymbol{\tau}, \boldsymbol{\mu} \in S$. Inserting (7.6) then yields (7.2a).

Step 2. The adjoint equation (7.2d) and (7.2e):
Similarly to above, we choose $\boldsymbol{q} = \mathbf{0}$ in (1.4c) so that the definition of $\boldsymbol{\zeta}$ in (7.3b) yields (7.2e). Choosing $\boldsymbol{v} = \mathbf{0}$ in (1.4c), we obtain completely analogously to (7.5) that

$$\boldsymbol{\pi} = \boldsymbol{\zeta}^D + \boldsymbol{\psi}^D. \tag{7.8}$$

If we solve the definitions of $\boldsymbol{\zeta}$ and $\boldsymbol{\psi}$ for $\boldsymbol{r}$ and test the arising equations with $\boldsymbol{\tau}, \boldsymbol{\mu} \in S$, then we arrive at

$$\begin{aligned}\int_\Omega \boldsymbol{\zeta} : \mathbb{C}^{-1}\boldsymbol{\tau}\,\mathrm{d}x + \int_\Omega \boldsymbol{\psi} : \mathbb{H}^{-1}\boldsymbol{\mu}\,\mathrm{d}x - \int_\Omega \boldsymbol{\varepsilon}(\boldsymbol{w}) : \boldsymbol{\tau}\,\mathrm{d}x \\ + \int_\Omega \boldsymbol{r} : (\boldsymbol{\tau}^D + \boldsymbol{\mu}^D)\,\mathrm{d}x = 0 \quad \text{for all } \boldsymbol{\tau}, \boldsymbol{\mu} \in S.\end{aligned} \tag{7.9}$$

From (1.4f), (7.8), (7.5), and the definitions of λ and θ it follows that

$$\begin{aligned}\int_\Omega \boldsymbol{r} : (\boldsymbol{\tau}^D + \boldsymbol{\mu}^D)\,\mathrm{d}x &= \int_\Omega \frac{1}{\widetilde{\sigma}_0}(|\boldsymbol{p}|\boldsymbol{\pi} + \vartheta\boldsymbol{\varrho}) : (\boldsymbol{\tau}^D + \boldsymbol{\mu}^D)\,\mathrm{d}x \\ &= \int_\Omega \big(\lambda\,(\boldsymbol{\zeta}^D + \boldsymbol{\psi}^D) + \theta\,(\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D)\big) : (\boldsymbol{\tau}^D + \boldsymbol{\mu}^D)\,\mathrm{d}x.\end{aligned}$$

Inserting this into (7.9) results in (7.2d).

Step 3. The complementarity relations (7.2g)–(7.2i):
Thanks to (7.6) and (7.8), (7.2g) follows immediately from (1.4e). Similarly, (7.2i) is a direct consequence of (1.4g) together with (7.5), (7.8), and the definition of θ in (7.3b). The complementarity relation in (7.2h) follows from (1.4h) and the definition of θ which imply that

$$\theta = 0 \quad \text{a.e. in } \{x \in \Omega : |\boldsymbol{\sigma}^D(x) + \boldsymbol{\chi}^D(x)| < \widetilde{\sigma}_0\}.$$

Since the gradient equations in (7.2f) coincide with these in (1.4d), this ends the first part of the proof.

To show the reverse direction, assume that (7.2) holds.

Step 1. The state system (1.4a) and (1.4b):
If one tests (7.2a) with $(\boldsymbol{\tau}, -\boldsymbol{\tau})$ with $\boldsymbol{\tau} \in S$ arbitrary, then $\boldsymbol{\varepsilon}(\boldsymbol{u}) - \mathbb{C}^{-1}\boldsymbol{\sigma} + \mathbb{H}^{-1}\boldsymbol{\chi} = \mathbf{0}$ is obtained. The definition of $\boldsymbol{p}$ in (7.4a) thus yields

$$\boldsymbol{p} = -\mathbb{H}^{-1}\boldsymbol{\chi}. \tag{7.10}$$

Consequently, (7.2a) implies

$$\int_\Omega \big(\lambda\,(\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D) - \boldsymbol{p}\big) : (\boldsymbol{\tau} + \boldsymbol{\mu})\,\mathrm{d}x = 0 \quad \text{for all } \boldsymbol{\tau}, \boldsymbol{\mu} \in S,$$

where we used that

$$A^D : B^D = A^D : B \quad \text{for all } A, B \in \mathbb{R}^{3\times 3}. \tag{7.11}$$

Therefore

$$\boldsymbol{p} = \lambda\,(\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D), \tag{7.12}$$

which in particular implies that $\operatorname{trace}(\boldsymbol{p}) = 0$ a.e. in Ω such that $\boldsymbol{p} \in Q$. In view of the definition of $\boldsymbol{p}$ in (7.4a) and (7.10) we find $\boldsymbol{\varrho} = \big[\mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{p})\big]^D - \big[\mathbb{H}\,\boldsymbol{p}\big]^D$ and thus

$$\int_\Omega \boldsymbol{\varrho} : \boldsymbol{q} \,\mathrm{d}x + \int_\Omega \boldsymbol{q} : \mathbb{H}\,\boldsymbol{p} \,\mathrm{d}x - \int_\Omega \mathbb{C}\big(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{p}\big) : \boldsymbol{q} \,\mathrm{d}x = 0 \quad \text{for all } \boldsymbol{q} \in Q, \tag{7.13}$$

where we again used (7.11). Solving the definition of $\boldsymbol{p}$ for $\boldsymbol{\sigma}$ and inserting this expression into (7.2b) gives

$$\int_\Omega \mathbb{C}\big(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{p}\big) : \boldsymbol{\varepsilon}(\boldsymbol{v}) \,\mathrm{d}x = \langle \ell,\, \boldsymbol{v}\rangle \quad \text{for all } \boldsymbol{v} \in V.$$

Adding this equation to (7.13) yields (1.4a). By the definition of $\boldsymbol{\varrho}$ in (7.4a) and (7.2c) we immediately obtain $|\boldsymbol{\varrho}| \leq \widetilde{\sigma}_0$ a.e. in Ω. Moreover, from (7.12), the definition of $\boldsymbol{\varrho}$, and the complementarity relation in (7.2c) it follows

$$\boldsymbol{\varrho} : \boldsymbol{p} = \lambda \left|\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D\right|^2 = \left|\lambda \left(\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D\right)\right| \, \widetilde{\sigma}_0 = \widetilde{\sigma}_0 \, |\boldsymbol{p}|,$$

i.e. the remaining condition in (1.4b).

Step 2. The adjoint equation (1.4c):
Analogously to the derivation of (7.10) we test (7.2d) with $(\boldsymbol{\tau}, -\boldsymbol{\tau})$ with arbitrary $\boldsymbol{\tau} \in S$, which gives

$$\boldsymbol{r} = -\mathbb{H}^{-1}\boldsymbol{\psi}. \tag{7.14}$$

By solving this equation for $\boldsymbol{\psi}$, the definition of $\boldsymbol{r}$ and $\boldsymbol{\pi}$ in (7.4a) and (7.4b), respectively imply $\boldsymbol{\pi} = [\boldsymbol{\zeta} + \boldsymbol{\psi}]^D = \big[\mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{w}) - \boldsymbol{r}) - \mathbb{H}\,\boldsymbol{r}\big]^D$ so that

$$\int_\Omega \Big(\mathbb{H}\,\boldsymbol{r} : \boldsymbol{q} - \mathbb{C}\big(\boldsymbol{\varepsilon}(\boldsymbol{w}) - \boldsymbol{r}\big) : \boldsymbol{q} + \boldsymbol{\pi} : \boldsymbol{q}\Big) \,\mathrm{d}x = 0 \quad \text{for all } \boldsymbol{q} \in Q, \tag{7.15}$$

where we again used (7.11). Now, we can argue as in the case of the state equation. The definition of $\boldsymbol{r}$ gives $\boldsymbol{\zeta} = \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{w}) - \boldsymbol{r})$. If we insert this into (7.2e) and add the arising equation to (7.15), then (1.4c) is obtained.

Step 3. The complementarity relations (1.4e)–(1.4h):
If ones inserts (7.14) together with the definition of $\boldsymbol{\varrho}$ and $\boldsymbol{\pi}$ into (7.4a) and (7.4b), then (7.11) implies similarly to (7.12) that

$$\boldsymbol{r} = \lambda\, \boldsymbol{\pi} + \theta\, \boldsymbol{\varrho}. \tag{7.16}$$

Since $\operatorname{trace}(\boldsymbol{\varrho}) = \operatorname{trace}(\boldsymbol{\pi}) = 0$ a.e. in Ω by definition, this yields $\boldsymbol{r} \in Q$. By (7.12) we have $\boldsymbol{p}(x) = \boldsymbol{0}$ a.e. where $\lambda\,(x) = 0$ holds. If $\lambda\,(x) \neq 0$, then (7.2c) implies $|\boldsymbol{\sigma}^D(x) + \boldsymbol{\chi}^D(x)| = \widetilde{\sigma}_0$. Therefore, in any case, (7.12) gives

$$|\boldsymbol{p}| = \widetilde{\sigma}_0\, \lambda.$$

Together with (7.16) and the definition of ϑ in (7.4b) this yields (1.4f).

In view of (7.12) and the definition of $\boldsymbol{\pi}$ in (7.4b), (1.4e) follows directly from (7.2g). Moreover, using the definitions of $\boldsymbol{\varrho}$, $\boldsymbol{\pi}$, and ϑ in (7.4a) and (7.4b), we obtain (1.4g) immediately from (7.2i). In addition, the complementarity relation in (7.2h) shows $\theta(x) = 0$ a.e. where $|\boldsymbol{\varrho}(x)| = |\boldsymbol{\sigma}^D(x) + \boldsymbol{\chi}^D(x)| < \widetilde{\sigma}_0$, which gives in turn (1.4h). Since the gradient equations in (7.2f) and (1.4d) are the same, this ends the proof. □

8. Numerical Experiments. We consider the numerical solution of a particular example of the *regularized* problem (4.1) for fixed $\gamma > 0$. In fact, we apply some minor changes to the original problem setting in an effort to make the problem more realistic from a practical point of view. These modifications do not affect our theory, and the changes to the regularized optimality system (6.1) as well as the optimality system (1.4) for the original problem are going to be obvious.

In our model problem, we restrict the discussion to boundary loads $\boldsymbol{g}$ as controls. These controls act only on a part Γ_C of the Neumann boundary Γ_N, and $\boldsymbol{g} = \boldsymbol{0}$ is fixed on $\Gamma_N \setminus \Gamma_C$. We also slightly modify the first term in the objective so that only the first two components of the displacement field are observed, and the desired state $\boldsymbol{u}_d$ has only two components. Finally the observation takes place on part of the boundary $\Gamma_O \subset \Gamma$, instead of inside the domain. We arrive at the following problem.

$$\begin{aligned} &\text{Minimize} \quad J(\boldsymbol{u},\boldsymbol{g}) := \frac{\beta}{2}\|\boldsymbol{u}^{(1,2)} - \boldsymbol{u}_d\|^2_{L^2(\Gamma_O;\mathbb{R}^2)} + \frac{\nu_2}{2}\|\boldsymbol{g}\|^2_{L^2(\Gamma_C;\mathbb{R}^3)} \\ &\text{s.t.} \quad a(\boldsymbol{W},\boldsymbol{Y}) + \int_\Omega h_\gamma(\boldsymbol{p}) : \boldsymbol{q}\,\mathrm{d}x = \int_{\Gamma_C} \boldsymbol{g}\cdot\boldsymbol{v}\,\mathrm{d}s \quad \text{for all } \boldsymbol{Y} = (\boldsymbol{v},\boldsymbol{q}) \in Z, \end{aligned} \tag{8.1}$$

with h_γ given in (4.2). For notational convenience, we denote the variables by $\boldsymbol{u}$ instead of $\boldsymbol{u}_\gamma$ etc.

Our computational domain is the scaled Fichera corner $\Omega = (-L, L)^3 \setminus (0, L)^3$ with $L = 100$ $[mm]$. The control boundary Γ_C is the upper boundary at $z = L$, and it coincides with the observation boundary Γ_O. The Dirichlet boundary Γ_D is located at the opposite face, i.e., at $z = -L$, see also Figure 8.1.

We take as material data the relevant parameters from [Neff and Wieners, 2003, Table 2], i.e.,

elasticity modulus	$E = 206\ 900\ [N/mm^2]$
Poisson ratio	$\nu = 0.29$
shear modulus	$\mu_L = \frac{E}{2\,(1+\nu)} \approx 80\ 194\ [N/mm^2]$
dilation modulus	$\lambda_L = \frac{E\,\nu}{(1+\nu)(1-2\,\nu)} \approx 110\ 744\ [N/mm^2]$
yield stress	$\widetilde{\sigma}_0 = 450\sqrt{2/3}\ [N/mm^2] \approx 367.42\ [N/mm^2]$

and in addition we choose the

hardening parameter	$k_1 = 100\ 000\ [N/mm^2]$.

Moreover, we use the following data related to the objective in problem (8.1):

desired state	$\boldsymbol{u}_d = (15, 0)^\top\ [mm]$
coefficient	$\beta = 1\ [1/mm^4]$
coefficient	$\nu_2 = 10^{-10}\ [mm^2/N^2]$.

The boundary traction $\boldsymbol{g}$ has $[N/mm^2]$ as its unit. The units of the coefficients β and ν_2 are chosen in such a way that the objective $J(\boldsymbol{u},\boldsymbol{g})$ becomes dimensionless. In our calculations, we choose

$$\gamma = 10^4$$

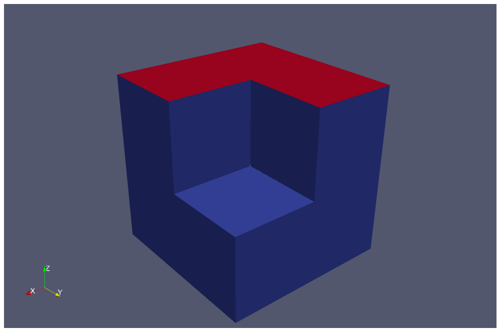

FIG. 8.1. *Computational domain with the control and observation boundary $\Gamma_C = \Gamma_O$ at $z = L$ highlighted. The body is clamped at the opposite boundary $z = -L$.*

as the regularization parameter. We use this rather large value of γ in place of a path-following strategy in this first numerical experiment.

This numerical exercise was solved within the finite element framework FEniCS, see Logg et al. [2012a]. We formulated the Lagrangian pertaining to (8.1) in the Unified Form Language UFL (Alnæs [2012]) and exploited the automatic differentiation capabilities of the form compiler FFC (Logg et al. [2012b]) in order to automatically generate the first-order optimality system, which is similar to (6.1). Minor changes to (6.1) are necessary due to the changes which led from (4.1) to (8.1).

The optimality system was discretized using vector-valued continuous $\mathbb{P}_1$ elements for the primal and adjoint displacements $\boldsymbol{u}$ and $\boldsymbol{w}$ as well as the control $\boldsymbol{g}$, while symmetric matrix-valued and trace free discontinuous $\mathbb{P}_0$ elements were used to discretize the primal and adjoint plastic strains $\boldsymbol{p}$ and $\boldsymbol{r}$. The quantities $\boldsymbol{\varrho}$ and $\boldsymbol{\pi}$ were not introduced as extra variables, cf. (6.1b) and (6.1d). We then used Newton's method (in the form of the `solve` method for a FEniCS nonlinear variational problem) with automatically generated second-order derivatives to solve the nonlinear optimality system. The overall algorithm is thus a basic sequential quadratic programming (SQP) approach.

For this study, sparse direct linear algebra was used to solve the arising linear systems. The uniform tetrahedral mesh has approximately 7 000 nodes, 35 000 cells and the total number of unknowns is roughly 415 000. Three Newton steps (without globalization efforts) were required to solve the system for $\gamma = 10^4$ to reasonable accuracy, starting from an all-zeros initial guess for $\boldsymbol{u}$, $\boldsymbol{p}$, $\boldsymbol{g}$, $\boldsymbol{w}$ and $\boldsymbol{r}$. To accelerate the solution, the built-in MPI-based parallel assembly and solution capabilities of FEniCS were used to distribute the problem onto 24 cores. The overall wall clock time was approximately 321 seconds.

Figure 8.2 shows the displacement $\boldsymbol{u}$ obtained at the solution with regularization parameter $\gamma = 10^4$. The desired state $\boldsymbol{u}_d$ is achieved rather closely. Figure 8.2 also shows the control $\boldsymbol{g}$, acting on the upper surface. The boundary stresses are in the range $|\boldsymbol{g}| \in [0, 2\,918]$ $[N/mm^2]$. Moreover, we show in Figure 8.3 the Frobenius norm of the deviator of the combined stress $\boldsymbol{\sigma} + \boldsymbol{\chi} = \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{p}) - \mathbb{H}\,\boldsymbol{p}$, compare (7.3a). We point out that $|\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D|$ nearly reaches the upper bound of $\widetilde{\sigma}_0 = 367.4235$ $[N/mm^2]$. Figure 8.3 also shows the Frobenius norm of the plastic strain $|\boldsymbol{p}|$. As expected, nonzero plastic strains are concentrated in the areas where the stresses are at the yield stress limit. It is also as expected that these large stresses occur in areas around the edges leading to the re-entrant corner. All visualizations were done in ParaView.

Table 8.1 gives some insight into the behavior of the Huber-type regularization. We see that below $\gamma = 10^3$, the plastic behavior is effectively suppressed, which leads to a fast convergence of Newton's method (always from an all-zeros initial guess) within two steps. Between $\gamma = 10^3$ and $\gamma = 10^4$, we observe the sudden onset of a pronounced elastoplastic behavior, which is reflected by the behavior of Newton's method due to increasing nonlinear effects.

γ	Newton	deviator	distance
10^0	2	$6.3374 \cdot 10^{-2}$	$3.6736 \cdot 10^2$
10^1	2	$6.3371 \cdot 10^{-1}$	$3.6679 \cdot 10^2$
10^2	2	$6.3350 \cdot 10^0$	$3.6109 \cdot 10^2$
10^3	2	$6.3132 \cdot 10^1$	$3.0429 \cdot 10^2$
10^4	3	$3.6742 \cdot 10^2$	$4.7411 \cdot 10^{-6}$
10^5		did not converge	

TABLE 8.1

Dependence of number of Newton steps, the value of the deviator sum $\|\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D\|_{L^\infty(\Omega;\mathbb{Q})}$, *and the distance to the dual feasibility constraint* $\widetilde{\sigma}_0 - \|\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D\|_{L^\infty(\Omega;\mathbb{Q})} \geq 0$ *on the choice of the regularization parameter* γ *at fixed discretization level. At* $\gamma = 10^5$, *Newton's method no longer converges from an all-zero initial guess.*

It is interesting that the Huber-type regularization of the primal formulation of elastoplasticity has the effect that the combined stresses $\boldsymbol{\sigma} + \boldsymbol{\chi}$ in the dual formulation actually stay *below* the yield threshold. When one starts with the dual formulation (7.1), it is natural to employ a penalty-type regularization, which leads to a convergence of the combined stresses from *above*. We point out that the penalized dual formulation can be interpreted as an elasto*visco*plastic model. A similar interpretation seems to be lacking in the primal formulation and deserves further investigation. Finally, we mention that more efficient numerical implementations can be achieved by adaptively coupling the regularization and discretization parameters, and by employing preconditioned iterative solvers for the Newton systems. This is postponed to future work.

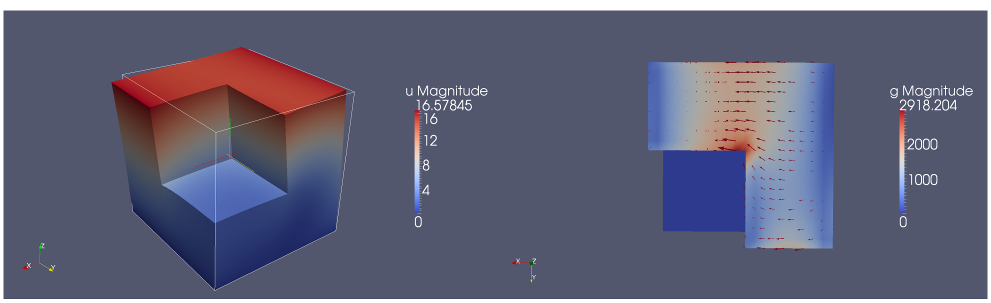

FIG. 8.2. *Displacement magnitude* $|\boldsymbol{u}|$, *and control* $\boldsymbol{g}$ *and its magnitude* $|\boldsymbol{g}|$ *at the solution.*

Appendix A. Proof of Lemma 3.2.

Proof. The arguments are classical, cf. e.g. [Han and Reddy, 1999b, Section 8.1]. Nevertheless, we shortly sketch the proof for convenience of the reader. The existence and uniqueness of solutions for (3.5) is for instance proved in Herzog and Meyer [2011]

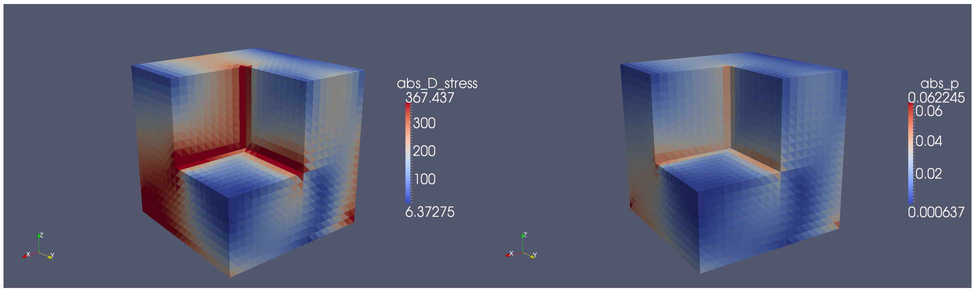

FIG. 8.3. *Sum of stress deviators* $|\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D|$ *and plastic strain* $|\boldsymbol{p}|$ *at the solution.*

by standard direct methods of variational calculus. Moreover it is shown in [Herzog and Meyer, 2011, Lemma A.2] by means of Lagrange duality that the dual problem of (3.5) is given by

$$\int_\Omega \big(\boldsymbol{\varepsilon}(\widehat{\boldsymbol{u}}) - \widehat{\boldsymbol{p}}\big) : \mathbb{C}\big(\boldsymbol{\varepsilon}(\boldsymbol{v} - \widehat{\boldsymbol{u}}) - (\boldsymbol{q} - \widehat{\boldsymbol{p}})\big)\,\mathrm{d}x + \int_\Omega \widehat{\boldsymbol{\xi}} : \mathbb{H}(\boldsymbol{\eta} - \widehat{\boldsymbol{\xi}})\,\mathrm{d}x \\ + \int_\Omega \sup_{(\boldsymbol{\tau},\boldsymbol{\mu})\in K} (\boldsymbol{\tau} : \boldsymbol{q} + \boldsymbol{\mu} : \boldsymbol{\eta})\,\mathrm{d}x - \int_\Omega \sup_{(\boldsymbol{\tau},\boldsymbol{\mu})\in K} (\boldsymbol{\tau} : \widehat{\boldsymbol{p}} + \boldsymbol{\mu} : \widehat{\boldsymbol{\xi}})\,\mathrm{d}x \\ \geq \langle \ell,\, \boldsymbol{v} - \boldsymbol{u}\rangle \quad \text{for all } (\boldsymbol{v}, \boldsymbol{q}, \boldsymbol{\eta}) \in V \times S \times S, \qquad \text{(A.1)}$$

where

$$\widehat{\boldsymbol{u}} = \boldsymbol{u}, \quad \widehat{\boldsymbol{p}} = \boldsymbol{\varepsilon}(\boldsymbol{u}) - \mathbb{C}^{-1}\boldsymbol{\sigma}, \quad \text{and} \quad \widehat{\boldsymbol{\xi}} = -\mathbb{H}^{-1}\boldsymbol{\chi}. \qquad \text{(A.2)}$$

Note that the result of [Herzog and Meyer, 2011, Lemma A.2] applies to more general flow rules with problem (3.5) as a special case. Moreover, we point out that ℓ was defined in Herzog and Meyer [2011] with the opposite sign compared to (1.2c), as is customary for the dual formulation, cf. [Han and Reddy, 1999b, Chapter 8]. When we test (3.5a) with $(\boldsymbol{\tau}, \boldsymbol{\mu}) = (\boldsymbol{\omega} + \boldsymbol{\sigma}, \boldsymbol{\chi} - \boldsymbol{\omega})$ and $\boldsymbol{\omega} \in S$ arbitrary, which is feasible due to the structure of $\mathcal{K}$, then

$$\mathbb{H}^{-1}\boldsymbol{\chi} = \mathbb{C}^{-1}\boldsymbol{\sigma} - \boldsymbol{\varepsilon}(\boldsymbol{u})$$

is obtained, which in turn implies $\widehat{\boldsymbol{p}} = \widehat{\boldsymbol{\xi}}$. Hence, by choosing $\boldsymbol{\eta} = \boldsymbol{q}$ in (A.1), we arrive at

$$a(\widehat{\boldsymbol{W}}, \boldsymbol{Y} - \widehat{\boldsymbol{W}}) + \int_\Omega \sup_{(\boldsymbol{\tau},\boldsymbol{\mu})\in K} (\boldsymbol{\tau} + \boldsymbol{\mu}) : \boldsymbol{q}\,\mathrm{d}x - \int_\Omega \sup_{(\boldsymbol{\tau},\boldsymbol{\mu})\in K} (\boldsymbol{\tau} + \boldsymbol{\mu}) : \widehat{\boldsymbol{p}}\,\mathrm{d}x \\ \geq 0 \quad \text{for all } \boldsymbol{Y} = (\boldsymbol{v}, \boldsymbol{q}) \in V \times S \qquad \text{(A.3)}$$

with $\widehat{\boldsymbol{W}} = (\widehat{\boldsymbol{u}}, \widehat{\boldsymbol{p}})$. Since K only involves restrictions on the deviatoric parts of $\boldsymbol{\tau}$ and $\boldsymbol{\mu}$, we find

$$\sup_{(\boldsymbol{\tau},\boldsymbol{\mu})\in K} (\boldsymbol{\tau} + \boldsymbol{\mu}) : \boldsymbol{q} = \infty \quad \text{for all } \boldsymbol{q} \in \mathbb{S} \text{ with } \operatorname{trace}(\boldsymbol{q}) \neq 0. \qquad \text{(A.4)}$$

On the other hand, for all $\boldsymbol{q} \in \mathbb{S}$ with vanishing trace one obtains

$$\sup_{(\boldsymbol{\tau},\boldsymbol{\mu})\in K} (\boldsymbol{\tau} + \boldsymbol{\mu}) : \boldsymbol{q} = \sup_{|\boldsymbol{\tau}^D + \boldsymbol{\mu}^D| \leq \widetilde{\sigma}_0} (\boldsymbol{\tau}^D + \boldsymbol{\mu}^D) : \boldsymbol{q} = \widetilde{\sigma}_0\,|\boldsymbol{q}|. \qquad \text{(A.5)}$$

Now, (A.3)–(A.5) imply that $\operatorname{trace}(\widehat{\boldsymbol{p}}) = 0$ holds a.e. in Ω and that $\widehat{\boldsymbol{W}}$ solves (1.1). Since the solution of (1.1) is unique, the assertion is proved. ◻

Appendix B. Proof of Theorem 4.2.

Proof. Let us define the nonlinear map $\mathcal{N}_\gamma : Z \to Z'$ by the left hand side of (4.1b), i.e.,

$$\langle \mathcal{N}_\gamma \boldsymbol{W},\, \boldsymbol{Y} \rangle := a(\boldsymbol{W}, \boldsymbol{Y}) + \left(h_\gamma(\boldsymbol{p}), \boldsymbol{q}\right)_\Omega .$$

We know from (3.2) and the monotonicity of h that $\mathcal{N}_\gamma$ is strongly monotone, i.e.,

$$\langle \mathcal{N}_\gamma(\boldsymbol{u}_1 - \boldsymbol{u}_2, \boldsymbol{p}_1 - \boldsymbol{p}_2),\, (\boldsymbol{u}_1 - \boldsymbol{u}_2, \boldsymbol{p}_1 - \boldsymbol{p}_2)\rangle \ge \underline{a}\left(\|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_V^2 + \|\boldsymbol{p}_1 - \boldsymbol{p}_2\|_S^2\right).$$

Moreover $\mathcal{N}_\gamma$ is clearly continuous. The Browder-Minty Theorem 3.3 thus shows that (4.1b) admits a unique solution $\boldsymbol{W} = (\boldsymbol{u}, \boldsymbol{p}) \in V \times Q$ for every $\ell \in V'$. The inverse $\mathcal{N}_\gamma^{-1}$ has Lipschitz constant $1/\underline{a}$, independent of γ. ◻

Appendix C. Proof of Theorem 4.3.

Proof. Let $\ell \in V'$ be given and denote by $\boldsymbol{W}_\gamma = (\boldsymbol{u}_\gamma, \boldsymbol{p}_\gamma) \in Z$ the unique solution of (4.1b) and by $\boldsymbol{W}$ the unique solution of (1.1). The idea is to estimate

$$\|\boldsymbol{W} - \boldsymbol{W}_\gamma\|_Z \le \|\boldsymbol{W} - \widetilde{\boldsymbol{W}}_\gamma\|_Z + \|\widetilde{\boldsymbol{W}}_\gamma - \boldsymbol{W}_\gamma\|_Z \tag{C.1}$$

with a suitable intermediate element $\widetilde{\boldsymbol{W}}_\gamma$. We choose $\widetilde{\boldsymbol{W}}_\gamma = (\widetilde{\boldsymbol{u}}_\gamma, \widetilde{\boldsymbol{p}}_\gamma)$ as the unique solution to the auxiliary problem

$$a(\widetilde{\boldsymbol{W}}_\gamma, \boldsymbol{Y}) + \widetilde{\sigma}_0\, \gamma \int_\Omega \frac{\widetilde{\boldsymbol{p}}_\gamma : \boldsymbol{q}}{\max(\widetilde{\sigma}_0, \gamma\, |\widetilde{\boldsymbol{p}}_\gamma|)}\, \mathrm{d}x = \langle \ell,\, \boldsymbol{v} \rangle \quad \text{for all } \boldsymbol{Y} = (\boldsymbol{v}, \boldsymbol{q}) \in Z. \tag{C.2}$$

The second term above corresponds to the derivative of the convex Huber function $j_\gamma : S \to S$ given by

$$j_\gamma(\boldsymbol{p}) := \widetilde{\sigma}_0 \int_\Omega g_\gamma\big(|\boldsymbol{p}|\big)\, \mathrm{d}x$$

with

$$g_\gamma : \mathbb{R} \to \mathbb{R}, \quad g_\gamma(p) := \begin{cases} |p| - \frac{\widetilde{\sigma}_0}{2\gamma}, & \text{if } \gamma\,|p| \ge \widetilde{\sigma}_0 \\ \frac{\gamma}{2\,\widetilde{\sigma}_0}\, |p|^2, & \text{if } \gamma\,|p| \le \widetilde{\sigma}_0. \end{cases}$$

Existence and uniqueness of a solution to (C.2) can be obtained by the Browder-Minty Theorem 3.3, using the strong monotonicity of the underlying operator.

Equivalently, equation (C.2) can be written as the following variational inequality

$$a(\widetilde{\boldsymbol{W}}_\gamma, \boldsymbol{Y} - \widetilde{\boldsymbol{W}}_\gamma) + j_\gamma(\boldsymbol{q}) - j_\gamma(\widetilde{\boldsymbol{p}}_\gamma) \ge \langle \ell,\, \boldsymbol{Y} - \widetilde{\boldsymbol{W}}_\gamma \rangle \quad \text{for all } \boldsymbol{Y} = (\boldsymbol{v}, \boldsymbol{q}) \in Z. \tag{C.3}$$

To estimate the first term on the right hand side of (C.1), we plug in $\boldsymbol{Y} = \boldsymbol{W}$ into (C.3). Then we plug in $\boldsymbol{Y} = \widetilde{\boldsymbol{W}}_\gamma$ into (1.1) and subtract the results to obtain

$$a(\boldsymbol{W} - \widetilde{\boldsymbol{W}}_\gamma, \boldsymbol{W} - \widetilde{\boldsymbol{W}}_\gamma) \le j(\widetilde{\boldsymbol{p}}_\gamma) - j(\boldsymbol{p}) + j_\gamma(\boldsymbol{p}) - j_\gamma(\widetilde{\boldsymbol{p}}_\gamma).$$

$$0 \le j(\boldsymbol{q}) - j_\gamma(\boldsymbol{q}) = \widetilde{\sigma}_0 \int_\Omega \big(|\boldsymbol{q}| - g_\gamma(|\boldsymbol{q}|)\big) \, \mathrm{d}x \le |\Omega| \, \frac{\widetilde{\sigma}_0^2}{\gamma}$$

holds for all $\boldsymbol{q} \in S$, it follows that

$$a(\boldsymbol{W} - \widetilde{\boldsymbol{W}}_\gamma, \boldsymbol{W} - \widetilde{\boldsymbol{W}}_\gamma) \le j(\widetilde{\boldsymbol{p}}_\gamma) - j_\gamma(\widetilde{\boldsymbol{p}}_\gamma) \le |\Omega| \, \frac{\widetilde{\sigma}_0^2}{\gamma}$$

and, consequently, the coercivity of a (see (3.2)) yields

$$\|\boldsymbol{W} - \widetilde{\boldsymbol{W}}_\gamma\|_Z \le \frac{c}{\sqrt{\gamma}} \to 0 \quad \text{as } \gamma \to \infty. \tag{C.4}$$

To estimate the second term on the right hand side of (C.1) we test both (4.1b) and (C.2) with $\boldsymbol{Y} = \boldsymbol{W}_\gamma - \widetilde{\boldsymbol{W}}_\gamma$ and take their difference. Exploiting again the coercivity of a we deduce

$$c \, \|\boldsymbol{W}_\gamma - \widetilde{\boldsymbol{W}}_\gamma\|_Z^2 \le -\widetilde{\sigma}_0 \, \gamma \int_\Omega \left(\frac{\boldsymbol{p}_\gamma}{m_\gamma(|\boldsymbol{p}_\gamma|)} - \frac{\widetilde{\boldsymbol{p}}_\gamma}{\max(\widetilde{\sigma}_0, \gamma \, |\widetilde{\boldsymbol{p}}_\gamma|)} \right) : (\boldsymbol{p}_\gamma - \widetilde{\boldsymbol{p}}_\gamma) \, \mathrm{d}x.$$

The monotonicity of $\mathbb{S} \ni \boldsymbol{p} \mapsto \frac{\boldsymbol{p}}{\max(\widetilde{\sigma}_0, \gamma \, |\boldsymbol{p}|)} \in \mathbb{S}$ implies for the integrand:

$$\begin{aligned}
&\left(\frac{\boldsymbol{p}_\gamma}{m_\gamma(|\boldsymbol{p}_\gamma|)} - \frac{\widetilde{\boldsymbol{p}}_\gamma}{\max(\widetilde{\sigma}_0, \gamma \, |\widetilde{\boldsymbol{p}}_\gamma|)} \right) : (\boldsymbol{p}_\gamma - \widetilde{\boldsymbol{p}}_\gamma) \\
&= \left(\frac{1}{m_\gamma(|\boldsymbol{p}_\gamma|)} - \frac{1}{\max(\widetilde{\sigma}_0, \gamma \, |\boldsymbol{p}_\gamma|)} \right) \boldsymbol{p}_\gamma : (\boldsymbol{p}_\gamma - \widetilde{\boldsymbol{p}}_\gamma) \\
&\quad + \left(\frac{\boldsymbol{p}_\gamma}{\max(\widetilde{\sigma}_0, \gamma \, |\boldsymbol{p}_\gamma|)} - \frac{\widetilde{\boldsymbol{p}}_\gamma}{\max(\widetilde{\sigma}_0, \gamma \, |\widetilde{\boldsymbol{p}}_\gamma|)} \right) : (\boldsymbol{p}_\gamma - \widetilde{\boldsymbol{p}}_\gamma) \\
&\ge \left(\frac{1}{m_\gamma(|\boldsymbol{p}_\gamma|)} - \frac{1}{\max(\widetilde{\sigma}_0, \gamma \, |\boldsymbol{p}_\gamma|)} \right) \boldsymbol{p}_\gamma : (\boldsymbol{p}_\gamma - \widetilde{\boldsymbol{p}}_\gamma) \\
&= \frac{\max(\widetilde{\sigma}_0, \gamma \, |\boldsymbol{p}_\gamma|) - m_\gamma(|\boldsymbol{p}_\gamma|)}{m_\gamma(|\boldsymbol{p}_\gamma|) \max(\widetilde{\sigma}_0, \gamma \, |\boldsymbol{p}_\gamma|)} \, \boldsymbol{p}_\gamma : (\boldsymbol{p}_\gamma - \widetilde{\boldsymbol{p}}_\gamma) \\
&\ge -\frac{1}{2\gamma} \, \frac{1}{\max(\widetilde{\sigma}_0, \gamma \, |\boldsymbol{p}_\gamma|)} \, \frac{|\boldsymbol{p}_\gamma|}{m_\gamma(|\boldsymbol{p}_\gamma|)} \, |\boldsymbol{p}_\gamma - \widetilde{\boldsymbol{p}}_\gamma| && \text{by (4.5)} \\
&\ge -\frac{1}{2\gamma} \, \frac{1}{\widetilde{\sigma}_0} \, \frac{1}{\gamma} \, |\boldsymbol{p}_\gamma - \widetilde{\boldsymbol{p}}_\gamma| && \text{by (4.4).}
\end{aligned}$$

Therefore,

$$c \, \|\boldsymbol{W}_\gamma - \widetilde{\boldsymbol{W}}_\gamma\|_Z^2 \le \frac{1}{2\gamma} \int_\Omega |\boldsymbol{p}_\gamma - \widetilde{\boldsymbol{p}}_\gamma| \, \mathrm{d}x \le \frac{|\Omega|^{1/2}}{2\gamma} \, \|\boldsymbol{W}_\gamma - \widetilde{\boldsymbol{W}}_\gamma\|_Z. \tag{C.5}$$

From (C.4) and (C.5) the result follows. □

Acknowledgment. This work was supported by a DFG grant within the Priority Program SPP 1253 (*Optimization with Partial Differential Equations*), which is gratefully acknowledged. Roland Herzog would like to thank Gerd Wachsmuth for advice on the FEniCS implementation.

References.

J. Alberty, C. Carstensen, and D. Zarrabi. Adaptive numerical analysis in primal elastoplasticity with hardening. *Computer Methods in Applied Mechanics and Engineering*, 171:175–204, 1999. doi: 10.1016/S0045-7825(98)00210-2.

Martin S. Alnæs. UFL: a finite element form language. In Anders Logg, Kent-Andre Mardal, and Garth N. Wells, editors, *Automated Solution of Differential Equations by the Finite Element Method, Volume 84 of Lecture Notes in Computational Science and Engineering*, chapter 17. Springer, 2012.

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

S. Bartels and T. Roubíček. Thermoviscoplasticity at small strains. *Zeitschrift für Angewandte Mathematik und Mechanik*, 88(9):735—-754, 2008.

T. Betz and C. Meyer. Second-order sufficient conditions for optimal control of static elastoplasticity with hardening. submitted, 2012.

F. Bonnans and E. Casas. An extension of Pontryagin's principle for state-constrained optimal control of semilinear elliptic equations and variational inequalities. *SIAM Journal on Control and Optimization*, 33(1):274–298, 1995. doi: 10.1137/S0363012992237777.

F. Bonnans and D. Tiba. Pontryagin's principle in the control of semilinear elliptic variational inequalities. *Applied Mathematics and Optimization*, 23:299–312, 1991. doi: 10.1007/BF01442403.

J.C. de los Reyes. Optimal control of a class of variational inequalities of the second kind. *SIAM Journal on Control and Optimization*, 49(4):1629–1658, 2011. doi: 10.1137/090764438.

J.C. de los Reyes. Optimization of mixed variational inequalities arising in flow of viscoplastic materials. *Computational Optimization and Applications*, 52:757–784, 2012. doi: 10.1007/s10589-011-9435-x.

Michael L. Flegel and Christian Kanzow. A Fritz John approach to first order optimality conditions for mathematical programs with equilibrium constraints. *Optimization. A Journal of Mathematical Programming and Operations Research*, 52 (3):277–286, 2003. ISSN 0233-1934. doi: 10.1080/0233193031000120020.

H. Goldberg, W. Kampowsky, and F. Tröltzsch. On Nemytskij operators in L_p-spaces of abstract functions. *Mathematische Nachrichten*, 155:127–140, 1992. ISSN 0025-584X. doi: 10.1002/mana.19921550110.

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. doi: 10.1007/BF01442860.

R. Haller-Dintelmann, C. Meyer, J. Rehberg, and A. Schiela. Hölder continuity and optimal control for nonsmooth elliptic problems. *Applied Mathematics and Optimization*, 60(3):397–428, 2009. doi: 10.1007/s00245-009-9077-x.

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

W. Han and B. D. Reddy. Convergence analysis of discrete approximations of problems in hardening plasticity. *Computer Methods in Applied Mechanics and Engineering*, 171(3-4):327–340, 1999b. doi: 10.1016/S0045-7825(98)00214-X.

R. Herzog and C. Meyer. Optimal control of static plasticity with linear kinematic hardening. *Journal of Applied Mathematics and Mechanics*, 91(10):777–794, 2011. doi: 10.1002/zamm.200900378.

R. Herzog, C. Meyer, and G. Wachsmuth. Integrability of displacement and stresses in linear and nonlinear elasticity with mixed boundary conditions. *Journal of Mathematical Analysis and Applications*, 382(2):802–813, 2011a. doi: 10.1016/j.jmaa.2011.04.074.

R. Herzog, C. Meyer, and G. Wachsmuth. Existence and regularity of the plastic multiplier in static and quasistatic plasticity. *GAMM Reports*, 34(1):39–44, 2011b. doi: 10.1002/gamm.201110006.

R. Herzog, C. Meyer, and G. Wachsmuth. C-stationarity for optimal control of static plasticity with linear kinematic hardening. *SIAM Journal on Control and Optimization*, 50(5):3052–3082, 2012. doi: 10.1137/100809325.

M. Hintermüller and I. Kopacka. Mathematical programs with complementarity constraints in function space: C- and strong stationarity and a path-following algorithm. *SIAM Journal on Optimization*, 20(2):868–902, 2009. ISSN 1052-6234. doi: 10.1137/080720681.

Tim Hoheisel, Christian Kanzow, and Alexandra Schwartz. Theoretical and numerical comparison of relaxation methods for mathematial programs with complementarity constraints. *Mathematical Programming*, 137(1–2):257–288, 2013. doi: 10.1007/s10107-011-0488-5.

Peter J. Huber. Robust regression: asymptotics, conjectures and Monte Carlo. *The Annals of Statistics*, 1:799–821, 1973. ISSN 0090-5364.

Peter J. Huber. *Robust statistics*. John Wiley & Sons Inc., New York, 1981. ISBN 0-471-41805-6. Wiley Series in Probability and Mathematical Statistics.

Anders Logg, Kent-Andre Mardal, Garth N. Wells, et al. *Automated Solution of Differential Equations by the Finite Element Method*. Springer, 2012a. ISBN 978-3-642-23098-1. doi: 10.1007/978-3-642-23099-8.

Anders Logg, Kristian B. Ølgaard, Marie E. Rognes, and Garth N. Wells. Ffc: the fenics form compiler. In Anders Logg, Kent-Andre Mardal, and Garth N. Wells, editors, *Automated Solution of Differential Equations by the Finite Element Method, Volume 84 of Lecture Notes in Computational Science and Engineering*, chapter 11. Springer, 2012b.

F. Mignot and J.-P. Puel. Optimal control in some variational inequalities. *SIAM Journal on Control and Optimization*, 22(3):466–476, 1984. doi: 10.1137/0322028.

P. Neff and C. Wieners. Comparison of models for finite plasticity: A numerical study. *Computing and Visualization in Science*, 6(1):23–35, 2003. doi: 10.1007/s00791-003-0104-1.

H. Scheel and S. Scholtes. Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity. *Mathematics of Operations Research*, 25 (1):1–22, 2000. doi: 10.1287/moor.25.1.1.15213.

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

F. Tröltzsch. *Optimal Control of Partial Differential Equations*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, 2010. Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels.

Liu Wenbin and J. E. Rubio. Maximum principles for optimal controls for elliptic variational inequalities of the second kind. *IMA Journal of Mathematical Control and Information*, 8(3):211–230, 1991. ISSN 0265-0754. doi: 10.1093/imamci/8.3.211.

E. Zeidler. *Nonlinear Functional Analysis and its Applications*, volume II/B. Springer, New York, 1990.