ÖAW AUSTRIAN ACADEMY OF SCIENCES

**Johann Radon Institute for
Computational and Applied Mathematics
Austrian Academy of Sciences (ÖAW)**

RICAM
JOHANN·RADON·INSTITUTE
FOR COMPUTATIONAL AND APPLIED MATHEMATICS

Optimal Control of an Abstract Evolution Variational Inequality with Application to Homogenized Plasticity

H. Meinlschmidt, C. Meyer, S. Walter

OPTIMAL CONTROL OF AN ABSTRACT EVOLUTION VARIATIONAL INEQUALITY WITH APPLICATION TO HOMOGENIZED PLASTICITY*

HANNES MEINLSCHMIDT†, CHRISTIAN MEYER‡, AND STEPHAN WALTHER‡

Abstract. The paper is concerned with an optimal control problem governed by a state equation in form of a generalized abstract operator differential equation involving a maximal monotone operator. The state equation is uniquely solvable, but the associated solution operator is in general not Gâteaux-differentiable. In order to derive optimality conditions, we therefore regularize the state equation and its solution operator, respectively, by means of a (smoothed) Yosida approximation. We show convergence of global minimizers for regularization parameter tending to zero and derive necessary and sufficient optimality conditions for the regularized problems. The paper ends with an application of the abstract theory to optimal control of homogenized quasi-static elastoplasticity.

Key words. Optimal control of operator differential equations, Yosida approximation, necessary and sufficient optimality conditions, homogenized plasticity, evolution variational inequality

AMS subject classifications. 49J27, 49J20, 49K27, 49K20, 74C05

1. Introduction. This paper is concerned with an optimal control problem of the following form, governed by an operator differential equation:

$$
\text{(P)} \quad
\begin{cases}
\min & J(z, \ell) := \Psi(z, \ell) + \Phi(\ell), \\
\text{s.t.} & \dot{z} \in A(R\ell - Qz), \qquad z(0) = z_0, \\
& (z, \ell) \in H^1(0, T; \mathcal{H}) \times \left(H^1(0, T; \mathcal{X}_c) \cap \mathcal{U}(z_0; M) \right).
\end{cases}
$$

Herein, A is a maximal monotone operator, while R and Q are linear and bounded operators in a Hilbert space $\mathcal{H}$. The control variable is denoted by ℓ, whereas z is the state of the system. The precise assumptions on the data are given in section 2 below.

The particular feature of the problem under consideration is the set-valued mapping A. Due to its maximal monotony, one can show that there is a well-defined single-valued control-to-state mapping $\ell \mapsto z$ (in suitable function spaces), but this mapping is in general *not Gâteaux-differentiable*. We are thus faced with a non-smooth optimal control problem, for which the derivation of necessary and sufficient optimality conditions is a particular challenge.

Depending on the precise choice of A, R, and Q, problem (P) covers various application problems. For instance, quasi-static elastoplasticity is frequently modeled in this way. Here, R is the solution operator associated with the equations of linear elasticity for given load distribution ℓ. Moreover, Q is the sum of the solution operator of linear elasticity for given stress distribution, the elasticity tensor, and a coercive

---

1

operator modeling hardening effects. Finally, A is the convex subdifferential of the indicator functional of the closed and convex set of feasible stresses defined by a suitable yield condition. Details on models in elastoplasticity can be found in [25]. Another model which is covered by the state equation in (P) is the system of homogenized elastoplasticity, which we will study in detail in section 7 below.

Let us put our work into perspective. Assume for a moment that A is the convex subdifferential of a proper, convex, and lower semicontinuous functional ϕ and that Q is self adjoint. Then, by convex duality, the state equation is equivalent to

$$(1.1) \qquad\qquad 0 \in \partial\phi^*(\dot{z}) + \mathcal{E}'(z), \quad z(0) = z_0,$$

where $\mathcal{E}$ is the quadratic energy functional given by

$$\mathcal{E}(z) := \tfrac{1}{2}(Qz, z)_{\mathcal{H}} - \langle R\ell, z \rangle.$$

Systems of this type have been intensively studied concerning existence of solutions and their numerical approximation, and we only refer to [29] and the references therein. In contrast to this, the literature on optimization problems governed by (1.1) is rather scarce. The research on optimal control of equations of type (1.1) probably started with the sweeping process, where $\phi = I_{-C(t)}$ is the indicator functional of a moving convex set $C(t)$, see [30]. In the optimal control setting, $C(t)$ is most frequently set to $C(t) = \ell(t) - Z$ with a convex set Z and a driving force ℓ. This fits into the setting of (1.1) by defining $\phi := I_Z$ and $Q = R = \mathrm{id}$ (identity). Optimal control problems of this type are investigated in [1, 2, 9–11, 15–17], where the underlying Hilbert space is mostly finite dimensional. Problems in an infinite dimensional Hilbert space are investigated in [21, 37]. To be more precise, in these contributions, $\mathcal{H}$ is the Sobolev space $H_0^1(\Omega)$ and $\mathcal{E}$ is the Dirichlet energy. Moreover, ϕ^* is set to $\phi^*(z) = \|z\|_{L^1(\Omega)}$ and its viscous regularization, respectively, i.e., $\phi_\delta^*(z) = \|z\|_{L^1(\Omega)} + \frac{\delta}{2}\|z\|^2_{H_0^1(\Omega)}$. Optimal control problems governed by quasi-static elastoplasticity with linear kinematic hardening and von Mises yield conditions are treated in [40, 42, 43]. As already indicated above, ϕ is the indicator functional of the convex set of feasible stresses in this case. All mentioned problems fit into our framework and can thus be seen as special cases of our non-smooth evolution. Our analysis therefore represents a generalization of existing results on optimal control of non-smooth evolution problems and can also be applied to application problems that were not treated in the literature so far such as optimal control of homogenized plasticity, which is investigated in section 7. We emphasize that problems with non-convex energies such as damage evolution are not covered by our analysis. Optimal control problems governed by (1.1) with non-convex energy are investigated in [31, 32].

Our strategy to analyze (P) is as follows: After showing well-posedness of the state equation and the optimal control problem, we employ the Yosida-regularization with an additional smoothing to obtain a smooth (i.e., Fréchet-differentiable) control-to-state map. We will prove that accumulation points of global minimizers of the regularized optimal control problems for vanishing regularization are solutions of the original non-smooth problem (P). Moreover, first-order necessary and second-order sufficient optimality conditions for the regularized problems are derived. The passage to the limit to establish optimality conditions for the original problem goes beyond the scope of this paper and is subject to future work. The results of [37, 43] indicate that the optimality conditions obtained in this way are rather weak and we expect the same all the more for our general setting. Let us underline that regularization is

a widely used approach to treat optimal control problems governed by non-smooth evolutions. We only refer to [10, 21, 42] and the references therein.

The paper is organized as follows. After stating our standing assumptions in section 2, we investigate the state equation and its regularization in section 3. In section 4, we then turn to the optimal control problem and show that it admits an optimal solution under our standing assumptions. Moreover, we establish the convergence result for vanishing regularization indicated above. Section 5 is devoted to first-order necessary optimality conditions for the regularized problems in form of a KKT-system involving an adjoint equation. In section 6, we address second-order sufficient conditions for the regularized problem. The paper ends with the adaptation of our general results to a concrete application problem, namely the optimal control of homogenized elastoplasticity.

2. Notation and Standing Assumptions. We start with a short introduction in the notation used throughout the paper.

Notation. Given two vector spaces $\mathcal{X}$ and $\mathcal{Y}$, we denote the space of linear and continuous functions from $\mathcal{X}$ into $\mathcal{Y}$ by $L(\mathcal{X}, \mathcal{Y})$. If $\mathcal{X} = \mathcal{Y}$, we simply write $L(\mathcal{X})$. The dual space of $\mathcal{X}$ is denoted by $\mathcal{X}^* = L(\mathcal{X}, \mathbb{R})$. If $\mathcal{H}$ is a Hilbert space, we denote its inner product by $(\cdot, \cdot)_{\mathcal{H}}$. For the whole paper, we fix the final time $T > 0$. We denote the Bochner space of square-integrable functions by $L^2(0, T; \mathcal{X})$ and the Bochner-Sobolev space by $H^1(0, T; \mathcal{X})$. Given a coercive operator $G : \mathcal{H} \to \mathcal{H}$ in a Hilbert space $\mathcal{H}$, we denote its coercivity constant by γ_G, i.e., $(Gh, h)_{\mathcal{H}} \geq \gamma_G \|h\|_{\mathcal{H}}^2$ for all $h \in \mathcal{H}$. Finally, $c > 0$ and $C > 0$ denote generic constants.

Standing Assumptions. The following standing assumptions are tacitly assumed for the rest of paper without mentioning them every time.

Spaces. Throughout the paper, $\mathcal{X}, \mathcal{X}_c, \mathcal{Y}, \mathcal{Z}, \mathcal{W}$ are real Banach spaces. Moreover, $\mathcal{X}_c$ reflexive and $\mathcal{H}$ is a separable Hilbert space. The space $\mathcal{X}_c$ is compactly embedded into $\mathcal{X}$ and the embeddings $\mathcal{Y} \hookrightarrow \mathcal{Z} \hookrightarrow \mathcal{H} \hookrightarrow \mathcal{W}$ are continuous.

Operators. The operator $A \colon \mathcal{H} \to 2^{\mathcal{H}}$ is maximal monotone, its domain $D(A)$ is closed and we define

$$(2.1) \qquad A^0 \colon D(A) \to \mathcal{H}, \quad h \mapsto \arg\min_{v \in A(h)} \|v\|_{\mathcal{H}}.$$

Furthermore, by $A_\lambda \colon \mathcal{H} \to \mathcal{H}$, $\lambda > 0$, we denote the Yosida-approximation of A and by $R_\lambda = (I + \lambda A)^{-1}$ the resolvent of A, so that $A_\lambda = \frac{1}{\lambda}(I - R_\lambda)$. We assume that the operator $A^0 \colon D(A) \to \mathcal{H}$ is bounded on bounded sets. For further reference on maximal monotone operators, we refer to [8], [44, Ch. 32], [7, Ch. 55], and [35, Ch. 55]. Furthermore, $R \in L(\mathcal{X}; \mathcal{Y})$ and $Q \in L(\mathcal{W}; \mathcal{W})$, are linear and bounded operators, and the restriction of Q to $\mathcal{H}$, $\mathcal{Z}$, or $\mathcal{Y}$ maps into these spaces and is again linear and bounded. To ease notation, we denote this restriction by the same symbol. Moreover, $Q \colon \mathcal{H} \to \mathcal{H}$ is coercive and self-adjoint.

Optimization Problem. By $J \colon H^1(0, T; \mathcal{W}) \times H^1(0, T; \mathcal{X}_c) \to \mathbb{R}$ we denote the objective function. We assume that both $\Psi \colon H^1(0, T; \mathcal{W}) \times H^1(0, T; \mathcal{X}_c) \to \mathbb{R}$ and $\Phi \colon H^1(0, T; \mathcal{X}_c) \to \mathbb{R}$ are weakly lower semicontinuous. Moreover, Ψ is bounded from below, while Φ is coercive. The set M is a nonempty and closed subset of $D(A)$ and $z_0 \in \mathcal{Y}$ is a given initial state.

Remark 2.1. We emphasize that not all of the above assumptions are always needed. For instance, in the next two sections, Q and R are only considered as operators with values in $\mathcal{H}$ and the spaces $\mathcal{Y}$ and $\mathcal{Z}$ are not needed, before the investigation

of optimality conditions starts in Sections 5 and 6. However, to keep the discussion concise, we present the standing assumption in the present form.

3. State Equation.

3.1. Existence and Uniqueness. We start the investigation of (P) with the discussion of the state equation, i.e.,

$$(3.1) \qquad\qquad \dot{z} \in A(R\ell - Qz), \qquad z(0) = z_0.$$

DEFINITION 3.1. *Let $\ell \in H^1(0, T; \mathcal{X})$ and $z_0 \in \mathcal{H}$. Then $z \in H^1(0, T; \mathcal{H})$ is called solution of* (3.1), *if $z(0) = z_0$ and $\dot{z}(t) \in A(Rl(t) - Qz(t))$ holds for almost all $t \in [0, T]$.*

In order to obtain the existence of a solution to (3.1), the data have to fulfill a certain compatibility condition. For this reason, we introduce the following

DEFINITION 3.2. *For $z_0 \in \mathcal{H}$ and $M \subset D(A)$, we define the set*

$$\mathcal{U}(z_0, M) := \left\{ \ell \in H^1(0, T; \mathcal{X}) \colon R\ell(0) - Qz_0 \in M \right\}$$

of admissible loads.

THEOREM 3.3 (Existence result for the state equation). *Let $z_0 \in \mathcal{H}$ and $\ell \in \mathcal{U}(z_0, D(A))$. Then there exists a unique solution $z \in H^1(0, T; \mathcal{H})$ of* (3.1)*. Furthermore, there exists a constant C, independent of z_0 and ℓ, such that*

$$(3.2) \qquad \|z\|_{C([0,T];\mathcal{H})} \le C\big(1 + \|z_0\|_{\mathcal{H}} + \|\ell\|_{C([0,T];\mathcal{X})} + \|\dot{\ell}\|_{L^1(0,T;\mathcal{X})}\big),$$

$$(3.3) \qquad \|\dot{z}\|_{L^2(0,T;\mathcal{H})} \le C\big(\|\dot{\ell}\|_{L^2(0,T;\mathcal{X})} + \sup_{\tau \in [0,T]} \|A^0(R\ell(\tau) - Qz(\tau))\|_{\mathcal{H}}\big),$$

where A^0 is as defined in (2.1)*.*

Proof. The proof essentially follows the lines of [23, Theorem 4.1] and [8, Proposition 3.4]. For convenience of the reader, we sketch the main arguments. At first, one employs the transformation $H^1(0, T; \mathcal{H}) \ni z \mapsto q := R\ell - Qz \in H^1(0, T; \mathcal{H})$ with its inverse $H^1(0, T; \mathcal{H}) \ni q \mapsto z := Q^{-1}(R\ell - q) \in H^1(0, T; \mathcal{H})$ to see that (3.1) is equivalent to

$$(3.4) \qquad\qquad \dot{q} + QA(q) \ni R\dot{\ell}, \qquad q(0) = R\ell(0) - Qz_0.$$

Since Q is coercive the operator, $\tilde{A} \colon \mathcal{H} \to 2^{\mathcal{H}}, h \mapsto QA(h)$ is maximal monotone with respect to the scalar product

$$(h_1, h_2)_{\mathcal{H}, Q^{-1}} := \big(Q^{-1}h_1, h_2\big)_{\mathcal{H}}, \quad h_1, h_2 \in \mathcal{H}.$$

Therefore, [8, Proposition 3.4] yields the existence of a unique solution $q \in H^1(0, T; \mathcal{H})$ of (3.4). To verify the estimate in (3.2), we employ [8, Lemme 3.1], which gives

$$\|q(t) - \tilde{q}(t)\|_{\mathcal{H}, Q^{-1}} \le \|R\ell(0) - Qz_0 - a\|_{\mathcal{H}, Q^{-1}} + \int_0^t \|R\dot{\ell}(\tau)\|_{\mathcal{H}, Q^{-1}} \, d\tau,$$

where $\tilde{q}$ is the unique solution of

$$\dot{\tilde{q}} + QA(\tilde{q}) \ni 0, \qquad q(0) = a$$

with an arbitrary element $a \in D(A)$. This gives the desired first inequality.

To prove the second inequality, we deduce from [8, Proposition 3.4] and the associated proof that

$$\|q(t) - q(s)\|_{\mathcal{H},Q^{-1}} \leq \int_s^t \|R\dot{\ell}(\tau)\|_{\mathcal{H},Q^{-1}} \, d\tau + \sup_{\tau \in [0,T]} \|\tilde{A}^0(q(\tau))\|_{\mathcal{H},Q^{-1}}(t-s).$$

Dividing this inequality by $(t-s)$ and letting $t \to s$ yields

$$(3.5) \qquad \|\dot{q}(s)\|_{\mathcal{H},Q^{-1}} \leq \|R\dot{\ell}(s)\|_{\mathcal{H},Q^{-1}} + \sup_{\tau \in [0,T]} \|\tilde{A}^0(q(\tau))\|_{\mathcal{H},Q^{-1}}$$

for almost all $s \in [0,T]$. From the definition of $\tilde{A}^0$ (with respect to $(\cdot,\cdot)_{\mathcal{H},Q^{-1}}$) we see that $\|\tilde{A}^0(h)\|_{\mathcal{H},Q^{-1}} \leq \|Qv\|_{\mathcal{H},Q^{-1}}$ for all $v \in A(h)$. This holds in particular for $v = A^0(h)$ so that $z = Q^{-1}(R\ell - q)$ and (3.5) imply

$$\|\dot{z}(s)\|_{\mathcal{H}} \leq C\big(\|\dot{\ell}(s)\|_{\mathcal{X}} + \|\dot{q}(s)\|_{\mathcal{H},Q^{-1}}\big)$$
$$\leq C\Big(\|\dot{\ell}(s)\|_{\mathcal{X}} + \sup_{\tau \in [0,T]} \|A^0(R\ell(\tau) - Qz(\tau))\|_{\mathcal{H}}\Big),$$

which gives the second inequality. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

Remark 3.4. In order to prove Theorem 3.3, it is sufficient to require that A^0 is bounded on compact subsets (in addition to the closedness of $D(A)$), cf. [8, Proposition 3.4]. However, the boundedness on bounded sets of A^0 is needed to prove Theorem 3.11 below and therefore, we impose it as a standing assumption.

Based on Theorem 3.3, we may introduce the solution operator associated with (3.1) and reduce (P) to an optimization problem in the control variable only, see Definition 4.1 below. Due to the set-valued operator A, this solution operator will in general be *non-smooth*, which complicates the derivation of first- and second-order optimality conditions. A prominent way to overcome this issue is to regularize the state equation in order to obtain a smooth solution operator. This is frequently done by means of the Yosida-approximation, see e.g. [5], and we will pursue the same approach. For this purpose, we will investigate the Yosida-approximation and its convergence properties in the next subsection.

3.2. Regularization and Convergence Results. For the rest of this section, we fix $z_0 \in \mathcal{H}$ and $\ell \in \mathcal{U}(z_0, D(A))$ and denote the unique solution of (3.1) by z. We start with a convergence result of the Yosida-approximation for fixed data z_0 and ℓ and then turn to perturbation of the data.

PROPOSITION 3.5 (Convergence of the Yosida-approximation for fixed data). *Let $z_\lambda \in H^1(0,T;\mathcal{H})$ be the solution of*

$$(3.6) \qquad\qquad \dot{z}_\lambda = A_\lambda(R\ell - Qz_\lambda), \qquad z_\lambda(0) = z_0$$

for all $\lambda > 0$. Then $z_\lambda \to z$ in $H^1(0,T;\mathcal{H})$ as $\lambda \searrow 0$ and the following inequality holds

$$(3.7) \quad \|z_\lambda - z\|_{C([0,T];\mathcal{H})}^2 + \frac{\lambda}{\gamma_Q}\|\dot{z}_\lambda\|_{L^2(0,T;\mathcal{H})}^2 + \frac{\lambda}{\gamma_Q}\|\dot{z}_\lambda - \dot{z}\|_{L^2(0,T;\mathcal{H})}^2 \leq \frac{\lambda}{\gamma_Q}\|\dot{z}\|_{L^2(0,T;\mathcal{H})}^2.$$

Proof. The proof in principle follows the lines of [8, Proposition 3.11], since our assumptions and assertions however are slightly different, we provide the arguments in detail.

First of all, since $z \mapsto A_\lambda(R\ell - Qz)$ is Lipschitz-continuous by [7, Proposition 55.2(b)], the existence of a unique solution of (3.6) follows from Banach's contraction principle by standard arguments, cf. e.g. [19]. Moreover, [7, Proposition 55.2(a)] and the definition of A_λ give

$$
\begin{aligned}
\frac{d}{dt} \left(Q(z_\lambda(t) - z(t)), z_\lambda(t) - z(t) \right)_\mathcal{H} &= 2 \left(\dot{z}_\lambda(t) - \dot{z}(t), Q(z_\lambda(t) - z(t)) \right)_\mathcal{H} \\
&= -2 \left(\dot{z}_\lambda(t) - \dot{z}(t), R_\lambda \left[R\ell(t) - Qz_\lambda(t) \right] - \left[R\ell(t) - Qz(t) \right] \right)_\mathcal{H} \\
&\quad - 2 \left(\dot{z}_\lambda(t) - \dot{z}(t), R\ell(t) - Qz_\lambda(t) - R_\lambda \left[R\ell(t) - Qz_\lambda(t) \right] \right)_\mathcal{H} \\
&\le -2\lambda \left(\dot{z}_\lambda(t) - \dot{z}(t), \dot{z}_\lambda(t) \right)_\mathcal{H} = \lambda \left(\|\dot{z}(t)\|_\mathcal{H}^2 - \|\dot{z}_\lambda(t)\|_\mathcal{H}^2 - \|\dot{z}_\lambda(t) - \dot{z}(t)\|_\mathcal{H}^2 \right).
\end{aligned}
$$

By integrating this inequality and using the coercivity of Q, we obtain the desired inequality.

In order to prove the strong convergence of z_λ to z in $H^1(0, T; \mathcal{H})$, we note that $z_\lambda \to z$ in $C([0, T]; \mathcal{H})$ and $\|\dot{z}_\lambda\|_{L^2(0,T;\mathcal{H})} \le \|\dot{z}\|_{L^2(0,T;\mathcal{H})}$ follow from the gained inequality. Hence, $z_\lambda \rightharpoonup z$ in $H^1(0, T; \mathcal{H})$ and the desired strong convergence follows from Lemma C.1 in the appendix. $\square$

Remark 3.6. The above proof shows that the inequality in (3.7) is by no means restricted to the specific setting in (3.1), i.e., whenever $\zeta \in H^1(0, T; \mathcal{H})$ and $\zeta_\lambda \in H^1(0, T; \mathcal{H})$ solve

$$
\begin{aligned}
\dot{\zeta} &= \mathcal{A}(Rg - Q\zeta), & \zeta(0) &= \zeta_0, \\
\dot{\zeta}_\lambda &= \mathcal{A}_\lambda(Rg - Q\zeta_\lambda), & \zeta_\lambda(0) &= \zeta_0,
\end{aligned}
$$

where $\mathcal{A} \colon \mathcal{H} \to 2^\mathcal{H}$ is a maximal monotone operator, $\mathcal{A}_\lambda \colon \mathcal{H} \to \mathcal{H}$ its Yosida-approximation and $g \in L^2(0, T; \mathcal{X})$ and $\zeta_0 \in \mathcal{H}$ are given, then an inequality analogue to (3.7) holds (with ζ and ζ_λ instead of z and z_λ).

Since we are concerned with an optimal control problem with the external loads as control variable, the continuity of the solution operator of (3.1) and its regularization w.r.t. variations in the external loads is of particular interest, for instance when it comes to the existence of optimal controls, see section 4 below. Since we aim to have a less restrictive control space in order to allow for as many control functions as possible, the topology for the variations of the loads needed for our continuity results should be as weak as possible. In particular, we aim to avoid *strong convergence* of (time-)derivatives of the loads. The underlying idea is similar to [8, Theorem 3.16] and leads to the following

LEMMA 3.7. *Let $\{z_{n,0}\}_{n\in\mathbb{N}} \subset \mathcal{H}$ and $\{\ell_n\}_{n\in\mathbb{N}} \subset L^2(0, T; \mathcal{X})$ be sequences such that $z_{n,0} \to z_0$ in $\mathcal{H}$ and $\ell_n \to \ell$ in $L^1(0, T; \mathcal{X})$. Assume further that $\{A_n\}_{n\in\mathbb{N}}$ is a sequence of maximal monotone operators such that*

$$
\tag{3.8} A_{n,\lambda}(h) \to A_\lambda(h)
$$

for all $\lambda > 0$ and all $h \in (R\ell - Qz_\lambda)([0, T])$, as $n \to \infty$, where z_λ is the solution of (3.6) and $A_{n,\lambda}$ denotes the Yosida-approximation of A_n. Then, if a sequence $\{z_n\}_{n\in\mathbb{N}} \subset H^1(0, T; \mathcal{H})$ satisfies

$$
\tag{3.9} \dot{z}_n \in A_n(R\ell_n - Qz_n), \qquad z_n(0) = z_{n,0}.
$$

and the derivatives $\dot{z}_n$ are bounded in $L^2(0, T; \mathcal{H})$, then $z_n \rightharpoonup z$ in $H^1(0, T; \mathcal{H})$ and $z_n \to z$ in $C([0, T]; \mathcal{H})$.

Proof. Let $\lambda > 0$ be fixed, but arbitrary and define $z_{n,\lambda} \in H^1(0, T; \mathcal{H})$ as solution of

$$\dot{z}_{n,\lambda} = A_{n,\lambda}(R\ell_n - Qz_{n,\lambda}), \quad z_{n,\lambda}(0) = z_{n,0},$$

whose existence and uniqueness can again be shown by Banach's contraction principle as in case of (3.6). Owing to [7, Proposition 55.2(b)], we obtain

$$
\begin{aligned}
\|\dot{z}_\lambda(t) - \dot{z}_{n,\lambda}(t)\|_{\mathcal{H}} &\leq \|A_\lambda(R\ell(t) - Qz_\lambda(t)) - A_{n,\lambda}(R\ell(t) - Qz_\lambda(t))\|_{\mathcal{H}} \\
&\quad + \|A_{n,\lambda}(R\ell(t) - Qz_\lambda(t)) - A_{n,\lambda}(R\ell_n(t) - Qz_{n,\lambda}(t))\|_{\mathcal{H}} \\
&\leq \|A_\lambda(R\ell(t) - Qz_\lambda(t)) - A_{n,\lambda}(R\ell(t) - Qz_\lambda(t))\|_{\mathcal{H}} \\
&\quad + \frac{\|Q\|_{L(\mathcal{H};\mathcal{H})}}{\lambda}\|z_\lambda(t) - z_{n,\lambda}(t)\|_{\mathcal{H}} + \frac{\|R\|_{L(\mathcal{X};\mathcal{H})}}{\lambda}\|\ell(t) - \ell_n(t)\|_{\mathcal{X}},
\end{aligned}
$$

and therefore, Gronwall's inequality implies

$$
\begin{aligned}
\|z_\lambda - z_{n,\lambda}\|_{C([0,T];\mathcal{H})} &\leq C(\lambda)\Big(\|z_0 - z_{n,0}\|_{\mathcal{H}} + \|\ell - \ell_n\|_{L^1(0,T;\mathcal{X})} \\
&\quad + \|A_\lambda(R\ell - Qz_\lambda) - A_{n,\lambda}(R\ell - Qz_\lambda)\|_{L^1(0,T;\mathcal{H})}\Big).
\end{aligned}
$$

The operators $A_{n,\lambda}$ are uniformly Lipschitz continuous with Lipschitz constant λ^{-1}. Thus, thanks to assumption (3.8), we can apply Lemma C.2 with $\mathcal{M} := (R\ell - Qz_\lambda)[0, T]$, $\mathcal{N} := \mathcal{H}$, $G_n := A_{n,\lambda}$ and $G := A_\lambda$. Together with the assumptions on ℓ_n and $z_{n,0}$ this gives that the right side in the inequality above converges to zero as $n \to \infty$. Using this, Proposition 3.5, and Remark 3.6 (with $\mathcal{A} = A_n$), we conclude

$$
\begin{aligned}
(3.10) \quad \limsup_{n \to \infty} \|z - z_n\|_{C([0,T];\mathcal{H})} &\leq \|z - z_\lambda\|_{C([0,T];\mathcal{H})} + \limsup_{n \to \infty} \|z_{n,\lambda} - z_n\|_{C([0,T];\mathcal{H})} \\
&\leq \sqrt{\frac{\lambda}{\gamma_Q}}\big(\|\dot{z}\|_{L^2(0,T;\mathcal{H})} + \sup_{n \in \mathbb{N}} \|\dot{z}_n\|_{L^2(0,T;\mathcal{H})}\big).
\end{aligned}
$$

Now, since λ was arbitrary, (3.10) holds for every $\lambda > 0$. Therefore, as $\dot{z}_n$ is bounded in $L^2(0, T; \mathcal{H})$ by assumption, we obtain $z_n \to z$ in $C([0, T]; \mathcal{H})$. Moreover, again due to the boundedness assumption on $\dot{z}_n$, there is a weakly converging subsequence in $H^1(0, T; \mathcal{H})$. Due to $z_n \to z$ in $C([0, T]; \mathcal{H})$, the weak limit is unique and hence, the whole sequence z_n converges weakly to z in $H^1(0, T; \mathcal{H})$. $\quad\square$

LEMMA 3.8. *Let $\{\lambda_n\}_{n \in \mathbb{N}} \subset (0, \infty)$ be a sequence converging towards zero. Then the sequence $A_n := A_{\lambda_n}$, $n \in \mathbb{N}$, of maximal monotone operators fulfills (3.8) for all $\lambda > 0$ and all $h \in \mathcal{H}$.*

Proof. At first we prove that, for all $h \in \mathcal{H}$ and $2\lambda > \mu > 0$, the following inequality holds

$$(3.11) \qquad \|R_\lambda(h) - R_{\lambda+\mu}(h)\|_{\mathcal{H}} \leq \sqrt{\frac{\mu}{2\lambda - \mu}}\,\|h - R_\lambda(h)\|_{\mathcal{H}}.$$

For this purpose, let $h \in \mathcal{H}$ be arbitrary and set $y_1 := R_\lambda(h)$ and $y_2 := R_{\lambda+\mu}(h)$. Then we have $h \in y_1 + \lambda A(y_1)$, hence, $\frac{h-y_1}{\lambda} \in A(y_1)$ and analogously $\frac{h-y_2}{\lambda+\mu} \in A(y_2)$.

The monotonicity of A thus implies

$$0 \le \left(\frac{\lambda + \mu}{\lambda}(h - y_1) - (h - y_2), y_1 - y_2 \right)_{\mathcal{H}} \le \left(\frac{\mu}{2\lambda} - 1 \right) \|y_1 - y_2\|_{\mathcal{H}}^2 + \frac{\mu}{2\lambda} \|h - y_1\|_{\mathcal{H}}^2,$$

hence,

$$\|y_1 - y_2\|_{\mathcal{H}}^2 \le \frac{\mu}{2\lambda - \mu} \|h - y_1\|_{\mathcal{H}}^2,$$

which yields (3.11). With this inequality and [7, Proposition 55.2 (d)] at hand, we obtain

$$(A_{\lambda_n})_\lambda(h) = A_{\lambda_n + \lambda}(h) = \frac{1}{\lambda_n + \lambda}(h - R_{\lambda_n + \lambda}(h)) \to \frac{1}{\lambda}(h - R_\lambda(h)) = A_\lambda(h),$$

which completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Now, we are in the position to state our main convergence results in Theorem 3.10 and Theorem 3.11, where the loads and initial data are no longer fixed. The first theorem addresses the continuity properties of the solution operator to the original equation (3.1), whereas Theorem 3.11 deals with the Yosida-approximation. In order to sharpen these convergence results and prove strong convergence in $H^1(0, T; \mathcal{H})$, we additionally need the following

ASSUMPTION 3.9. *The maximal monotone operator A is given as a subdifferential of a proper, convex and lower semicontinuous function $\phi \colon \mathcal{H} \to (-\infty, \infty]$, that is, $A = \partial\phi$.*

THEOREM 3.10 (Continuity of the solution operator). *Let $\{z_{n,0}\}_{n\in\mathbb{N}} \subset \mathcal{H}$ and $\{\ell_n\}_{n\in\mathbb{N}} \subset \mathcal{U}(z_{n,0}, D(A))$ be sequences such that $z_{n,0} \to z_0$ in $\mathcal{H}$, $\ell_n \rightharpoonup \ell$ in $H^1(0, T; \mathcal{X})$ and $\ell_n \to \ell$ in $L^1(0, T; \mathcal{X})$. Moreover, denote the solution of*

$$\dot{z}_n \in A(R\ell_n - Qz_n), \qquad z_n(0) = z_{n,0}$$

by $z_n \in H^1(0, T; \mathcal{H})$ (whose existence is guaranteed by Theorem 3.3). Then $z_n \rightharpoonup z$ in $H^1(0, T; \mathcal{H})$ and $z_n \to z$ in $C([0, T]; \mathcal{H})$.

If additionally $\ell_n \to \ell$ in $H^1(0, T; \mathcal{X})$, A fulfills Assumption 3.9, and $\phi(R\ell_n(0) - Qz_{n,0}) \to \phi(R\ell(0) - Qz_0)$, then $z_n \to z$ in $H^1(0, T; \mathcal{H})$.

Proof. Thanks to Theorem 3.3, to be more precise (3.2), $\{z_n\}$ is bounded in $C([0, T]; \mathcal{H})$. Since A^0 is bounded on bounded sets by assumption, (3.3) then gives that $\{\dot{z}_n\}$ is bounded in $L^2(0, T; \mathcal{H})$. Therefore, we can apply Lemma 3.7 with $A_n := A$ for all $n \in \mathbb{N}$ to obtain $z_n \rightharpoonup z$ in $H^1(0, T; \mathcal{H})$ and $z_n \to z$ in $C([0, T]; H)$.

If additionally $\ell_n \to \ell$ in $H^1(0, T; \mathcal{X})$, A fulfills Assumption 3.9, and $\phi(R\ell_n(0) - Qz_{n,0}) \to \phi(R\ell(0) - Qz_0)$, we can follow the lines of [23, Theorem 4.2 step 3)] to get

$$\limsup_{n \to \infty} \int_0^T (Q\dot{z}_n, \dot{z}_n)_{\mathcal{H}} \, dt$$

$$= \limsup_{n \to \infty} -(R\dot{\ell}_n - Q\dot{z}_n, \dot{z}_n)_{L^2(0,T;\mathcal{H})} + (R\dot{\ell}, \dot{z})_{L^2(0,T;\mathcal{H})}$$

$$= \limsup_{n \to \infty} \phi(R\ell_n(0) - Qz_{n,0}) - \phi(R\ell_n(T) - Qz_n(T)) + (R\dot{\ell}, \dot{z})_{L^2(0,T;\mathcal{H})}$$

$$\le \phi(R\ell(0) - Qz_0) - \phi(R\ell(T) - Qz(T)) + (R\dot{\ell}, \dot{z})_{L^2(0,T;\mathcal{H})}$$

$$= \int_0^T (Q\dot{z}, \dot{z})_{\mathcal{H}} \, dt$$

where the second and last equation follows from [8, Lemme 3.3]. Hence, by equipping $\mathcal{H}$ with the equivalent norm $\sqrt{(Q\cdot,\cdot)_{\mathcal{H}}}$, the strong convergence $z_n \to z$ in $H^1(0,T;\mathcal{H})$ follows from Lemma C.1 with $\mathcal{H} = L^2(0,T;\mathcal{H})$, $x_n = \dot{z}_n$ and $a_n = \sup_{k \geq n} \|\dot{z}_k\|_{L^2(0,T;\mathcal{H})}$. □

THEOREM 3.11 (Convergence of the Yosida-approximation). *Consider again sequences* $\{z_{n,0}\}_{n\in\mathbb{N}} \subset \mathcal{H}$, *and* $\{\ell_n\}_{n\in\mathbb{N}} \subset \mathcal{U}(z_{n,0}, D(A))$ *such that* $z_{n,0} \to z_0$ *in* $\mathcal{H}$, $\ell_n \rightharpoonup \ell$ *in* $H^1(0,T;\mathcal{X})$ *and* $\ell_n \to \ell$ *in* $L^1(0,T;\mathcal{X})$. *Furthermore, let* $\{\lambda_n\}_{n\in\mathbb{N}} \subset (0,\infty)$ *be a sequence converging to zero and* z_n *be, for every* $n \in \mathbb{N}$, *the solution of*

$$(3.12) \qquad \dot{z}_n = A_{\lambda_n}(R\ell_n - Qz_n), \qquad z_n(0) = z_{n,0}$$

Then $z_n \rightharpoonup z$ *in* $H^1(0,T;\mathcal{H})$ *and* $z_n \to z$ *in* $C([0,T];\mathcal{H})$.

If in addition, $\ell_n \to \ell$ *in* $H^1(0,T;\mathcal{X})$, A *fulfills Assumption 3.9, and* $\phi(R\ell_n(0) - Qz_{n,0}) \to \phi(R\ell(0) - Qz_0)$, *then* $z_n \to z$ *in* $H^1(0,T;\mathcal{H})$.

Proof. According to Lemma 3.8, the sequence of maximal monotone operators $A_n := A_{\lambda_n}$ fulfills (3.8) so that it only remains to prove that $\dot{z}_n$ is bounded in $L^2(0,T;\mathcal{H})$ to apply again Lemma 3.7. To this end, let $v_n \in H^1(0,T;\mathcal{H})$ be the solution of

$$\dot{v}_n \in A(R\ell_n - Qv_n), \qquad v_n(0) = z_{n,0},$$

whose existence is guaranteed by Theorem 3.3 (note that $\ell_n \in \mathcal{U}(z_{n,0}, D(A))$ by assumption). Thanks to Theorem 3.10, it holds $v_n \rightharpoonup z$ in $H^1(0,T;\mathcal{H})$. From Proposition 3.5, it follows $\|\dot{z}_n\|_{L^2(0,T;\mathcal{H})} \leq \|\dot{v}_n\|_{L^2(0,T;\mathcal{H})}$ and consequently, $\dot{z}_n$ is bounded in $L^2(0,T;\mathcal{H})$. Thus, Lemma 3.7 yields $z_n \rightharpoonup z$ in $H^1(0,T;\mathcal{H})$ and $z_n \to z$ in $C([0,T];\mathcal{H})$, as claimed.

If additionally $\ell_n \to \ell$ in $H^1(0,T;\mathcal{X})$, A fulfills Assumption 3.9, and $\phi(R\ell_n(0) - Qz_{n,0}) \to \phi(R\ell(0) - Qz_0)$, then Theorem 3.10 implies $v_n \to z$ in $H^1(0,T;\mathcal{H})$ so that Lemma C.1 gives the strong convergence $z_n \to z$ in $H^1(0,T;\mathcal{H})$ because of $\|\dot{z}_n\|_{L^2(0,T;\mathcal{H})} \leq \|\dot{v}_n\|_{L^2(0,T;\mathcal{H})}$ as seen above. □

Remark 3.12. The assertions of Theorem 3.10 and Theorem 3.11 are remarkable due to the following: As a first approach to prove the (strong) convergence of the states in $H^1(0,T;\mathcal{H})$, one is tempted to follow the lines of the proofs of [8, Lemme 3.1] and Proposition 3.5, respectively. This would however require the strong convergence of the *derivatives* of the given loads, which we want to avoid in order to enable less regular controls. The detour via the Yosida-regularization in Lemma 3.7 allows to overcome this issue.

Remark 3.13. If we would allow for more regular controls, then we could weaken the assumptions on the maximal monotone operator A. For instance, if $\ell \in H^2(0,T;\mathcal{X})$, then we can drop the assumptions that $D(A)$ is closed and A^0 is bounded on bounded sets. In this case, one can use [7, Theorem 55.A] instead of [8, Proposition 3.4] in the proof of Theorem 3.3. The proof of [7, Theorem 55.A] also gives the boundedness of $\dot{z}_n$ in $L^\infty(0,T;\mathcal{H})$ in this case. Thus, Lemma 3.7 is again applicable and we can argue similar as we did in the proof of Theorem 3.11 to verify the previous convergence results. In this setting, we would not have any restrictions on A (except monotonicity), but would need more regular loads, which is not favorable, as the latter serve as control variables in our optimization problem. Moreover, the boundedness assumption on A is fulfilled for our concrete application problem in subsection 7.4. Therefore, we decided to choose the present setting and to impose the additional boundedness assumption on A.

Remark 3.14. It is to be noted that most of the above results can also be shown in more general Bochner-Sobolev spaces, that is, when loads are contained in $W^{1,r}(0,T;\mathcal{X})$ and states in $W^{1,r}(0,T;\mathcal{H})$ for some $r \in [1,\infty)$. However, since a Hilbert space setting is advantageous when it comes to the derivation of optimality conditions, we focus on the case $r = 2$.

Unfortunately, the Yosida-approximation is frequently not sufficient for the derivation of optimality conditions by means of the standard adjoint approach, since the solution operator associated with (3.6) is in general still not Gâteaux-differentiable. Therefore, we apply a second regularization turning the Yosida-approximation of A into a smooth operator. The properties needed to ensure convergence of this second regularization are investigated in the following

LEMMA 3.15 (Convergence of the Regularized Yosida-Approximation). *Consider a sequence $\{\lambda_n\}_{n\in\mathbb{N}} \subset (0,\infty)$ and a sequence of Lipschitz continuous operators $A_n \colon \mathcal{H} \to \mathcal{H}$, $n \in \mathbb{N}$, such that*

$$(3.13) \qquad \lambda_n \searrow 0 \quad \text{and} \quad \frac{1}{\lambda_n}\exp\left(\frac{T\|Q\|_{L(\mathcal{H};\mathcal{H})}}{\lambda_n}\right)\sup_{h\in\mathcal{H}}\|A_n(h)-A_{\lambda_n}(h)\|_{\mathcal{H}} \to 0.$$

Let moreover $\{\ell_n\}_{n\in\mathbb{N}} \subset C([0,T];\mathcal{X})$ be given and denote by $z_n, z_{\lambda_n} \in C^1([0,T];\mathcal{H})$ the solutions of

$$(3.14) \qquad\qquad\qquad \dot{z}_n = A_n(R\ell_n - Qz_n), \qquad z_n(0) = z_0,$$

$$(3.15) \qquad\qquad \text{and} \quad \dot{z}_{\lambda_n} = A_{\lambda_n}(R\ell_n - Qz_{\lambda_n}), \quad z_{\lambda_n}(0) = z_0.$$

Then $\|z_n - z_{\lambda_n}\|_{C^1([0,T];\mathcal{H})} \to 0$.

Proof. Again, thanks to the Lipschitz continuity of A_n and A_{λ_n}, the existence and uniqueness of z_n and $z_{n,\lambda}$ follows from Banach's contraction principle by classical arguments. Moreover, the continuity of ℓ_n carries over to the continuity of $\dot{z}_n$ and $\dot{z}_{n,\lambda}$. Let us abbreviate $c_n := \sup_{h\in\mathcal{H}}\|A_n(h) - A_{\lambda_n}(h)\|_{\mathcal{H}}$. Then, in light of (3.14) and (3.15), we find

$$\|\dot{z}_n(t) - \dot{z}_{\lambda_n}(t)\|_{\mathcal{H}} \leq c_n + \frac{\|Q\|_{L(\mathcal{H};\mathcal{H})}}{\lambda_n}\|z_n(t) - z_{\lambda_n}(t)\|_{\mathcal{H}} \quad \forall\, t \in [0,T]$$

so that Gronwall's inequality yields

$$\|\dot{z}_n(t) - \dot{z}_{\lambda_n}(t)\|_{\mathcal{H}} \leq \frac{\|Q\|_{L(\mathcal{H};\mathcal{H})}}{\lambda_n}\left(T\exp\left(\frac{\|Q\|_{L(\mathcal{H};\mathcal{H})}}{\lambda_n}T\right) + 1\right)c_n \quad \forall\, t \in [0,T],$$

which completes the proof. $\square$

4. Existence and Approximation of Optimal Controls. Now we turn to the optimal control problem (P). We first address the existence of optimal solutions and afterwards discuss the approximation of (P) in subsection 4.2.

4.1. Existence of Optimal Controls. Based on Theorem 3.3, we reduce the optimal control problem (P) into a problem in the control variable only. Recall that the control space $\mathcal{X}_c$ embeds compactly in $\mathcal{X}$.

DEFINITION 4.1. *Let $z_0 \in \mathcal{H}$ and $M \subset D(A)$. Due to Theorem 3.3, there exists for every $\ell \in H^1(0,T;\mathcal{X}_c) \cap \mathcal{U}(z_0;M)$ a solution $z \in H^1(0,T;\mathcal{H})$ of the state equation in (3.1). Consequently, we may define the solution operator*

$$\mathcal{S} \colon H^1(0,T;\mathcal{X}_c) \cap \mathcal{U}(z_0;M) \ni \ell \to z \in H^1(0,T;\mathcal{H}).$$

This operator will be frequently called control-to-state map.

With the definition above, problem (P) is equivalent to the *reduced problem*:

$$(\mathrm{P}) \quad \Longleftrightarrow \quad \begin{cases} \min & J(\mathcal{S}(\ell), \ell), \\ \text{s.t.} & \ell \in H^1(0, T; \mathcal{X}_c) \cap \mathcal{U}(z_0; M). \end{cases}$$

Recall our standing assumptions on the objective, namely that $J(z, \ell) = \Psi(z, \ell) + \Phi(\ell)$, where $\Psi \colon H^1(0, T; \mathcal{W}) \times H^1(0, T; \mathcal{X}_c) \to \mathbb{R}$ is weakly lower semicontinuous and bounded from below and $\Phi \colon H^1(0, T; \mathcal{X}_c) \to \mathbb{R}$ is weakly lower semicontinuous and coercive. These assumptions allow us to show the existence of (globally) optimal solutions:

THEOREM 4.2 (Existence of Optimal Solutions). *Let $z_0 \in \mathcal{H}$ and M be a closed subset of $D(A)$. Then there exists a global solution of* (P).

Proof. Based on Theorem 3.10, the proof follows the standard direct method of the calculus of variations. First of all, since Ψ is bounded from below and Φ is coercive, every infimal sequence of controls is bounded in $H^1(0, T; \mathcal{X}_c)$ and thus admits a weakly converging subsequence. Due to the compact embedding of $\mathcal{X}_c$ in $\mathcal{X}$, this sequence converges strongly in $C([0, T]; \mathcal{X})$ so that the weak limit belongs to $\mathcal{U}(z_0; M)$, due the closeness of M. Moreover, thanks to weak convergence in $H^1(0, T; \mathcal{X})$ and strong convergence in $C([0, T]; \mathcal{X})$, Theorem 3.10 gives weak convergence of the associated states in $H^1(0, T; \mathcal{H})$. The weak lower semicontinuity of Ψ and Φ together with $\mathcal{H} \hookrightarrow \mathcal{W}$ then implies the optimality of the weak limit. □

Clearly, in view of the nonlinear state equation, one cannot expect the optimal solution to be unique. Note that, since $D(A)$ is closed by our standing assumptions, the choice $M = D(A)$ is feasible.

4.2. Convergence of Global Minimizers. While the existence of optimal solutions for (P) can be shown by well-established techniques as seen above, the derivation of optimality conditions is all but standard because of the lack of differentiability of the control-to-state map. We therefore apply a regularization of A built upon the Yosida-approximation in order to obtain a smooth control-to-state mapping. In view of Lemma 3.15, this regularization is assumed to satisfy the following

ASSUMPTION 4.3. *Let $\{A_n\}_{n \in \mathbb{N}}$ be a sequence of Lipschitz continuous operators from $\mathcal{H}$ to $\mathcal{H}$ such that, together with a sequence $\{\lambda_n\}_{n \in \mathbb{N}} \subset (0, \infty)$, it holds*

$$(4.1) \qquad \lambda_n \searrow 0 \quad and \quad \frac{1}{\lambda_n} \exp\left(\frac{T \|Q\|_{L(\mathcal{H}; \mathcal{H})}}{\lambda_n}\right) \sup_{h \in \mathcal{H}} \|A_n(h) - A_{\lambda_n}(h)\|_{\mathcal{H}} \to 0,$$

i.e., the requirements in Lemma 3.15 are fulfilled.

In subsection 7.4, we show how to construct such a regularization for a concrete application problem. Given the regularization of A, we define the corresponding optimal control problem:

$$(\mathrm{P}_n) \quad \begin{cases} \min & J(z, \ell), \\ \text{s.t.} & \dot{z} = A_n(R\ell - Qz), \qquad z(0) = z_0, \\ & (z, \ell) \in H^1(0, T; \mathcal{H}) \times \left(H^1(0, T; \mathcal{X}_c) \cap \mathcal{U}(z_0; M)\right). \end{cases}$$

Since A_n is Lipschitz continuous, the equation

$$\dot{z} = A_n(R\ell - Qz), \quad z(0) = z_0$$

admits a unique solution for every $z_0 \in \mathcal{H}$ and every $\ell \in L^2(0,T;\mathcal{X})$. Similar to Definition 4.1, we denote the associated solution operator by

$$\mathcal{S}_n \colon L^2(0,T;\mathcal{X}) \to H^1(0,T;\mathcal{H}).$$

Moreover, the solution operator associated with the Yosida-approximation, i.e., the solution operator of $\dot{z} = A_{\lambda_n}(R\ell - Qz)$, $z(0) = z_0$, is denoted by

$$\mathcal{S}_{\lambda_n} \colon L^2(0,T;\mathcal{X}) \to H^1(0,T;\mathcal{H}).$$

PROPOSITION 4.4 (Existence of Optimal Solutions of the Regularized Problems). *Let $n \in \mathbb{N}$, $z_0 \in \mathcal{H}$, and M a closed subset of $D(A)$. Then, under Assumption 4.3, there exists a global solution of* (P_n).

Proof. Let $\ell_1, \ell_2 \in L^2(0,T;\mathcal{X})$ be arbitrary and define $z_i \coloneqq \mathcal{S}_n(\ell_i)$, $i = 1, 2$. Then, due to the Lipschitz continuity of A_n, we have for almost all $t \in [0,T]$

(4.2)
$$\begin{aligned}\|\dot{z}_1(t) - \dot{z}_2(t)\|_{\mathcal{H}} &= \|A_n(R\ell_1(t) - Qz_1(t)) - A_n(R\ell_2(t) - Qz_2(t))\|_{\mathcal{H}} \\ &\leq c\big(\|l_1(t) - l_2(t)\|_{\mathcal{X}} + \|z_1(t) - z_2(t)\|_{\mathcal{H}}\big),\end{aligned}$$

which yields, thanks to Gronwall's inequality , the Lipschitz continuity of $\mathcal{S}_n$. Using this together with the fact that $\mathcal{X}_c$ is compactly embedded into $\mathcal{X}$, one can argue as in the proof of Theorem 4.2 to obtain the existence of a global solution of (P_n) for all $n \in \mathbb{N}$. □

THEOREM 4.5 (Weak Approximation of Global Minimizers). *Let $z_0 \in \mathcal{H}$ and M be a closed subset of $D(A)$. Suppose moreover that Assumption 4.3 holds and let $\{\bar{\ell}_n\}_{n\in\mathbb{N}}$ be a sequence of globally optimal controls of* (P_n), *$n \in \mathbb{N}$. Then there exists a weak accumulation point and every weak accumulation point is a global solution of* (P).

Proof. Due to $M \subset D(A)$, Proposition 3.5 gives $\mathcal{S}_{\lambda_n}(\bar{\ell}_1) \to \mathcal{S}(\bar{\ell}_1)$ in $H^1(0,T;\mathcal{H})$ so that Lemma 3.15 yields $\mathcal{S}_n(\bar{\ell}_1) \to \mathcal{S}(\bar{\ell}_1)$ in $H^1(0,T;\mathcal{H})$ and thus

$$\begin{aligned}\limsup_{n\to\infty} \Psi(\mathcal{S}_n(\bar{\ell}_n),\bar{\ell}_n) + \Phi(\bar{\ell}_n)& \\ = \limsup_{n\to\infty} J(\mathcal{S}_n(\bar{\ell}_n),\bar{\ell}_n) &\leq \limsup_{n\to\infty} J(\mathcal{S}_n(\bar{\ell}_1),\bar{\ell}_1) = J(\mathcal{S}(\bar{\ell}_1),\bar{\ell}_1).\end{aligned}$$

Hence, by virtue of the boundedness of Ψ from below and the radial unboundedness of Φ, $\{\bar{\ell}_n\}$ is bounded and therefore admits a weak accumulation point in $H^1(0,T;\mathcal{X}_c)$.

Let us now assume that a given subsequence of $\{\bar{\ell}_n\}_{n\in\mathbb{N}}$, denoted by the same symbol for simplicity, converges weakly to $\tilde{\ell}$ in $H^1(0,T;\mathcal{X}_c)$. Since $\mathcal{X}_c$ is compactly embedded in $\mathcal{X}$, we obtain $\bar{\ell}_n \to \tilde{\ell}$ in $C([0,T];\mathcal{X})$ and consequently, $\tilde{\ell} \in \mathcal{U}(z_0;M)$. In addition, the strong convergence in $C([0,T];\mathcal{X})$ in combination with Theorem 3.11 and Lemma 3.15 yields weak convergence of the states, i.e., $\mathcal{S}_n(\bar{\ell}_n) \rightharpoonup \mathcal{S}(\tilde{\ell})$ in $H^1(0,T;\mathcal{H})$ and thus also in $H^1(0,T;\mathcal{W})$. Now, let $\bar{\ell}$ be a global solution of (P). We can again use Proposition 3.5 and Lemma 3.15 to obtain $\mathcal{S}_n(\bar{\ell}) \to \mathcal{S}(\bar{\ell})$ in $H^1(0,T;\mathcal{H})$. This, together with the weak lower semicontinuity of Ψ and Φ, implies

(4.3)
$$\begin{aligned}J(\mathcal{S}(\tilde{\ell}),\tilde{\ell}) &= \Psi(\mathcal{S}(\tilde{\ell}),\tilde{\ell}) + \Phi(\tilde{\ell}) \\ &\leq \liminf_{n\to\infty} \Psi(\mathcal{S}_n(\bar{\ell}_n),\bar{\ell}_n) + \Phi(\bar{\ell}_n) \\ &\leq \limsup_{n\to\infty} J(\mathcal{S}_n(\bar{\ell}_n),\bar{\ell}_n) \leq \limsup_{n\to\infty} J(\mathcal{S}_n(\bar{\ell}),\bar{\ell}) = J(\mathcal{S}(\bar{\ell}),\bar{\ell}),\end{aligned}$$

giving in turn the optimality of the weak limit. □

COROLLARY 4.6 (Strong Approximation of Global Minimizers). *In addition to Assumption* 4.3, *assume that* $\Phi\colon H^1(0,T;\mathcal{X}_c) \to \mathbb{R}$ *is such that, if a sequence* $\{\ell_n\}_{n\in\mathbb{N}}$ *satisfies* $\ell_n \rightharpoonup \ell$ *in* $H^1(0,T;\mathcal{X}_c)$ *and* $\Phi(\ell_n) \to \Phi(\ell)$, *then* $\ell_n \to \ell$ *in* $H^1(0,T;\mathcal{X}_c)$. *Then every weak accumulation point of a sequence of globally optimal controls of* (P_n) *is also a strong one.*

Moreover, if in addition, at least one of the following holds

- *Assumption* 3.9 *is satisfied, that is* $A = \partial\phi$, *and* ϕ *is continuous on* M *or*
- $\Psi\colon H^1(0,T;\mathcal{W}) \times H^1(0,T;\mathcal{X}_c) \to \mathbb{R}$ *is such that, if sequences* $\{z_n\}_{n\in\mathbb{N}}$ *and* $\{\ell_n\}_{n\in\mathbb{N}}$ *satisfy* $z_n \rightharpoonup z$ *in* $H^1(0,T;\mathcal{H})$ *and* $\ell_n \to \ell$ *in* $H^1(0,T;\mathcal{X}_c)$ *and* $\Psi(z_n,\ell_n) \to \Psi(z,\ell)$, *then* $z_n \to z$ *in* $H^1(0,T;\mathcal{H})$,

then the associated sequence of state also converges strongly in $H^1(0,T;\mathcal{H})$.

Proof. Consider an arbitrary accumulation point $\tilde{\ell}$ of a sequence of global minimizers of (P_n), i.e., $\bar{\ell}_n \rightharpoonup \tilde{\ell}$ in $H^1(0,T;\mathcal{X}_c)$. From the previous proof, we know that then (4.3) holds, giving in turn

$$\Psi(\mathcal{S}_n(\bar{\ell}_n),\bar{\ell}_n) + \Phi(\bar{\ell}_n) \to \Psi(\mathcal{S}(\tilde{\ell}),\tilde{\ell}) + \Phi(\tilde{\ell}).$$

Since $\mathcal{S}_n(\bar{\ell}_n) \rightharpoonup \mathcal{S}(\tilde{\ell})$, as seen in the previous proof, and both, Ψ and Φ, are weakly lower semicontinuous by assumption, this implies $\Phi(\bar{\ell}_n) \to \Phi(\tilde{\ell})$ and $\Psi(\mathcal{S}_n(\bar{\ell}_n),\bar{\ell}_n) \to \Psi(\mathcal{S}(\tilde{\ell}),\tilde{\ell})$. The hypothesis on Φ thus yields $\bar{\ell}_n \to \tilde{\ell}$ in $H^1(0,T;\mathcal{X}_c)$ so that $\tilde{\ell}$ is indeed a strong accumulation point as claimed.

Due to $\mathcal{X}_c \hookrightarrow \mathcal{X}$, the strong convergence carries over to $H^1(0,T;\mathcal{X})$ and therefore, we deduce from Theorem 3.11 that $\mathcal{S}_{\lambda_n}(\bar{\ell}_n) \to \mathcal{S}(\tilde{\ell})$ in $H^1(0,T;\mathcal{H})$, provided that Assumption 3.9 is fulfilled and $\phi(R\bar{\ell}_n(0) - Qz_0) \to \phi(R\tilde{\ell}(0) - Qz_0)$ holds. If the additional requirements on Ψ are fulfilled, we also obtain the strong convergence $\mathcal{S}_{\lambda_n}(\bar{\ell}_n) \to \mathcal{S}(\tilde{\ell})$ in $H^1(0,T;\mathcal{H})$, since we already showed $\bar{\ell}_n \to \tilde{\ell}$ in $H^1(0,T;\mathcal{X}_c)$. Thus, in both cases, Lemma 3.15 gives $\mathcal{S}_n(\bar{\ell}_n) \to \mathcal{S}(\tilde{\ell})$ in $H^1(0,T;\mathcal{H})$, which is the second assertion. □

EXAMPLE 4.7. *Let us assume that* $\mathcal{X}_c$ *is a Hilbert space. Then a possible objective functional fulfilling the requirements on* Φ *in Corollary* 4.6 *reads as follows:*

$$J(z,\ell) = \Psi(z,\ell) + \frac{\alpha}{2}\,\|\ell\|^2_{H^1(0,T;\mathcal{X}_c)},$$

i.e., $\Phi(\ell) \coloneqq \alpha/2\,\|\ell\|^2_{H^1(0,T;\mathcal{X}_c)}$. *Herein,* $\Psi\colon H^1(0,T;\mathcal{H}) \times H^1(0,T;\mathcal{X}_c) \to \mathbb{R}$ *is again lower semicontinuous and bounded from below and* $\alpha > 0$ *is a given constant. Since* $H^1(0,T;\mathcal{X}_c)$ *is a Hilbert space, too, weak convergence and norm convergence give strong convergence and consequently, this specific choice of g fulfills the condition in Corollary* 4.6.

Remark 4.8 (Approximation of Local Minimizers). By standard localization arguments, the above convergence analysis can be adapted to approximate local minimizers. Following the lines of, for instance, [12], one can show that, under the assumptions of Corollary 4.6, every strict local minimum of (P) can be approximated by a sequence of local minima of (P_n). A local minimizer $\bar{\ell}$ of (P), which is not necessarily strict, can be approximated by replacing the objective in (P_n) by $\overline{J}(z,l) \coloneqq J(z,l) + \|\ell - \bar{\ell}\|_{H^1(0,T;\mathcal{X}_c)}$, which is of course only of theoretical interest, cf. e.g. [4]. Since these results and their proofs are standard, we omitted them.

Now that we answered the question of approximation of optimal controls via regularization, we turn to the regularized problems and derive optimality conditions for these in the next two sections.

5. First-Order Optimality Conditions. In the following, we consider a single element of the sequence of regularized problems. The associated regularized operator is denoted by A_s so that the regularized optimal control problems reads as follows:

$$(\mathrm{P}_s) \qquad \begin{cases} \min & J(z, \ell), \\ \text{s.t.} & \dot{z} = A_s(R\ell - Qz), \qquad z(0) = z_0, \\ & (z, \ell) \in H^1(0, T; \mathcal{H}) \times \big(H^1(0, T; \mathcal{X}_c) \cap \mathcal{U}(z_0; M)\big). \end{cases}$$

Beside our standing assumption and the Lipschitz continuity required in Assumption 4.3, we need the following additional assumptions for the derivation of first-order necessary optimality conditions for (P_s). Recall the continuous embeddings $\mathcal{Y} \hookrightarrow \mathcal{Z} \hookrightarrow \mathcal{H}$ from section 2.

ASSUMPTION 5.1.
  (i) $J \colon H^1(0, T; \mathcal{W}) \times H^1(0, T; \mathcal{X}_c) \to \mathbb{R}$ *is Fréchet differentiable.*
  (ii) $A_s \colon \mathcal{Y} \to \mathcal{Y}$ *is Lipschitz continuous and Fréchet differentiable from $\mathcal{Y}$ to $\mathcal{Z}$. Moreover, $A_s'(y)$ can be extended to elements of $L(\mathcal{Z}; \mathcal{Z})$ and $L(\mathcal{H}; \mathcal{H})$, respectively, denoted by the same symbol. There exists a constant C such that these extensions satisfy $\|A_s'(y)z\|_\mathcal{Z} \leq C \|z\|_\mathcal{Z}$ and $\|A_s'(y)h\|_\mathcal{H} \leq C\|h\|_\mathcal{H}$ for all $y \in \mathcal{Y}$, $z \in \mathcal{Z}$, and $h \in \mathcal{H}$.*

Remark 5.2. It is well known that a norm gap is often indispensable to ensure Fréchet-differentiability. This is also the case in our application example in subsection 7.4. This is the reason for considering two different spaces $\mathcal{Y}$ and $\mathcal{Z}$ in context of the Fréchet-differentiability of A_s in Assumption 5.1.

We start the derivation of optimality conditions for (P_s) with the Fréchet-derivative of the associated control-to-state mapping.

5.1. Differentiability of the Regularized Control-to-State Mapping. As $A_s \colon \mathcal{Y} \to \mathcal{Y}$ is supposed to be Lipschitz continuous and R and Q are not only linear and continuous as operators with values in $\mathcal{H}$, but also in $\mathcal{Y}$ according to our standing assumptions, Banach's fixed point theorem immediately implies that the state equation in (P_s), i.e.,

$$(5.1) \qquad\qquad \dot{z} = A_s(R\ell - Qz), \quad z(0) = z_0,$$

admits a unique solution $z \in H^1(0, T; \mathcal{Y})$ for every right hand side $\ell \in L^2(0, T; \mathcal{X})$, provided that $z_0 \in \mathcal{Y}$. Therefore, similar to above, we can define the associated solution operator $\mathcal{S}_s \colon L^2(0, T; \mathcal{X}) \to H^1(0, T; \mathcal{Y})$ (for fixed $z_0 \in \mathcal{Y}$). We will frequently consider this operator with different domains, e.g. $H^1(0, T; \mathcal{X})$, and ranges, in particular $H^1(0, T; \mathcal{Z})$. With a little abuse of notation, these operators are denoted by the same symbol.

LEMMA 5.3 (Lipschitz Continuity of $\mathcal{S}_s$). *The solution operator $\mathcal{S}_s$ is globally Lipschitz continuous from $L^2(0, T; \mathcal{X})$ to $H^1(0, T; \mathcal{Y})$.*

Proof. This can be proven completely analogously to the Lipschitz continuity of $\mathcal{S}_n$ from $L^2(0, T; \mathcal{X})$ to $H^1(0, T; \mathcal{Y})$ in Proposition 4.4.   □

LEMMA 5.4. *Assume that Assumption 5.1(ii) is fulfilled and let $y \in L^2(0, T; \mathcal{Y})$ and $w \in L^2(0, T; \mathcal{Z})$ be given. Then there exists a unique solution $\eta \in H^1(0, T; \mathcal{Z})$ of*

$$(5.2) \qquad\qquad \dot{\eta} = A_s'(y)(w - Q\eta), \quad \eta(0) = 0.$$

Proof. Let us define

$$B\colon [0,T] \times \mathcal{Z} \to \mathcal{Z}, \quad (t,\eta) \mapsto A_s'(y(t))(w(t) - Q\eta)$$

so that (5.2) becomes $\dot{\eta}(t) = B(t,\eta(t))$ a.e. in $(0,T)$, $\eta(0) = 0$. Now, given $\eta \in L^2(0,T;\mathcal{Z})$, $[0,T] \ni t \mapsto B(t,\eta(t)) \in \mathcal{Z}$ is Bochner measurable as a pointwise limit of Bochner measurable functions. Furthermore, Assumption 5.1(ii) implies for almost all $t \in [0,T]$ and all $\eta_1, \eta_2 \in \mathcal{Z}$ that $\|B(t,\eta_1) - B(t,\eta_2)\|_{\mathcal{Z}} \le C \|\eta_1 - \eta_2\|_{\mathcal{Z}}$. Therefore, we can apply Banach's fixed point argument to the integral equation associated with (5.2), which gives the assertion. □

THEOREM 5.5 (Fréchet-Differentiability of the Regularized Solution Operator). *Under Assumption 5.1(ii), the solution operator $\mathcal{S}_s$ is Fréchet differentiable from $H^1(0,T;\mathcal{X})$ to $H^1(0,T;\mathcal{Z})$. Its directional derivative at $\ell \in H^1(0,T;\mathcal{X})$ in direction $h \in H^1(0,T;\mathcal{X})$ is given by the unique solution of*

$$(5.3) \qquad \dot{\eta} = A_s'(R\ell - Qz)(Rh - Q\eta), \quad \eta(0) = 0,$$

where $z := \mathcal{S}_s(\ell) \in H^1(0,T;\mathcal{Y})$. Moreover, there exists a constant C such that $\|\mathcal{S}_s'(\ell)h\|_{H^1(0,T;\mathcal{Z})} \le C\|h\|_{L^2(0,T;\mathcal{X})}$ holds for all $\ell, h \in H^1(0,T;\mathcal{X})$.

Proof. Let $\ell, h \in H^1(0,T;\mathcal{X})$ be arbitrary and abbreviate $z_h := \mathcal{S}_s(\ell+h)$. Thanks to Lemma 5.4, there exists a unique solution $\eta \in H^1(0,T;\mathcal{Z})$ of (5.3). Clearly, the solution operator of (5.3) is linear with respect to h. Moreover, Assumption 5.1(ii) implies for almost all $t \in [0,T]$ that

$$\|\dot{\eta}(t)\|_{\mathcal{Z}} \le C\big(\|h(t)\|_{\mathcal{X}} + \|\eta(t)\|_{\mathcal{Z}}\big),$$

so that Gronwall's inequality gives $\|\eta\|_{H^1(0,T;\mathcal{Z})} \le C\|h\|_{L^2(0,T;\mathcal{X})}$, i.e., the continuity of the solution operator of (5.3). This also proves the asserted inequality (after having proved that $\eta = \mathcal{S}_s'(\ell)h$, which we do next).

It remains to verify the remainder term property. For this purpose, let us denote the remainder term of A_s by r_1, i.e.,

$$A_s(y+\zeta) = A_s(y) + A_s'(y)\zeta + r_1(y;\zeta) \quad \text{with} \quad \frac{\|r_1(y;\zeta)\|_{\mathcal{Z}}}{\|\zeta\|_{\mathcal{Y}}} \to 0 \text{ as } \zeta \to 0 \text{ in } \mathcal{Y}.$$

Moreover, we abbreviate

$$y := R\ell - Qz \in H^1(0,T;\mathcal{Y}) \quad \text{and} \quad \zeta := Rh - Q(z_h - z) \in H^1(0,T;\mathcal{Y}).$$

Then, in view of the definition of z, z_h, and η (as solution of (5.3)), we find for almost all $t \in [0,T]$

$$\|\dot{z}_h(t) - \dot{z}(t) - \dot{\eta}(t)\|_{\mathcal{Z}}$$
$$= \|A_s(y(t) + \zeta(t)) - A_s(y(t)) - A_s'(y(t))(\zeta(t) + Q(z_h(t) - z(t) - \eta(t)))\|_{\mathcal{Z}}$$
$$\le \|A_s'(y(t))Q(z_h(t) - z(t) - \eta(t))\|_{\mathcal{Z}} + \|r_1(y(t);\zeta(t))\|_{\mathcal{Z}}.$$

Hence, Assumption 5.1(ii) and Gronwall's inequality yield

$$(5.4) \qquad \|z_h - z - \eta\|_{H^1(0,T;\mathcal{Z})} \le C \|r_1(y;\zeta)\|_{L^2(0,T;\mathcal{Z})}.$$

(note that $r_1(y;\zeta) \in L^2(0,T;\mathcal{Z})$ by its definition as remainder term). Furthermore, thanks to Lemma 5.3 and the definition of ζ, we obtain

$$(5.5) \qquad \|\zeta\|_{H^1(0,T;\mathcal{Y})} \le C \|h\|_{H^1(0,T;\mathcal{X})}$$

such that $h \to 0$ in $H^1(0, T; \mathcal{X})$ implies $\zeta \to 0$ in $H^1(0, T; \mathcal{Y})$. The continuous embedding $H^1(0, T; \mathcal{Y}) \hookrightarrow C([0, T]; \mathcal{Y})$ and the remainder term property of r_1 thus give for almost all $t \in (0, T)$ that

$$
(5.6) \qquad \frac{\|r_1(y(t); \zeta(t))\|_{\mathcal{Z}}}{\|h\|_{H^1(0,T;\mathcal{X})}} \leq C \frac{\|r_1(y(t); \zeta(t))\|_{\mathcal{Z}}}{\|\zeta(t)\|_{\mathcal{Y}}} \frac{\|\zeta\|_{H^1(0,T;\mathcal{Y})}}{\|h\|_{H^1(0,T;\mathcal{X})}} \to 0
$$

as $h \to 0$ in $H^1(0, T; \mathcal{X})$. Moreover, the Lipschitz continuity of $A_s \colon \mathcal{Y} \to \mathcal{Y}$ together with 5.1(ii), $\mathcal{Y} \hookrightarrow \mathcal{Z}$, and (5.5) yield for almost all $t \in (0, T)$ that

$$
\frac{\|r_1(y(t); \zeta(t))\|_{\mathcal{Z}}}{\|h\|_{H^1(0,T;\mathcal{X})}} = \frac{\|(A_s(y + \zeta) - A_s(y) - A_s'(y)\zeta)(t)\|_{\mathcal{Z}}}{\|h\|_{H^1(0,T;\mathcal{X})}} \leq C \frac{\|\zeta(t)\|_{\mathcal{Y}}}{\|h\|_{H^1(0,T;\mathcal{X})}} \leq C.
$$

In combination with (5.6) and Lebesgue's dominated convergence theorem, this yields

$$
\frac{\|r_1(y; \zeta)\|_{L^2(0,T;\mathcal{Z})}}{\|h\|_{H^1(0,T;\mathcal{X})}} \to 0
$$

as $h \to 0$ in $H^1(0, T; \mathcal{X})$, which, in view of (5.4) finishes the proof.   □

Remark 5.6. It is to be noted that we did not employ the implicit function theorem to show the differentiability of $\mathcal{S}_s$. The reason is that $H \colon z \mapsto \dot{z} - A_s(R\ell - Qz)$ is Fréchet-differentiable from $H^1(0, T; \mathcal{Y})$ to $L^2(0, T; \mathcal{Z})$, but the derivative $H'(z)$ is not continuously invertible in these spaces, cf. Lemma 5.4. On the other hand, H is not differentiable from $H^1(0, T; \mathcal{Y})$ to $L^2(0, T; \mathcal{Y})$ (due to the differentiability properties of A_s, see Remark 5.2), which would be the right spaces for the existence result from Lemma 5.4. The same observation for a more abstract setting was already made in [41].

5.2. Adjoint Equation. Now that we know that the (regularized) control-to-state map is Gâteaux-differentiable, we can apply the standard adjoint approach to derive first-order necessary optimality conditions in form of a Karush-Kuhn-Tucker (KKT) system. To keep the discussion concise, we restrict our analysis to the case without further control constraints. To be more precise, we require the following:

ASSUMPTION 5.7. *Let $z_0 \in \mathcal{Y}$ such that $-Qz_0 \in D(A)$. The set M in the definition of the set of admissible controls is given by the singleton $M = \{-Qz_0\}$ such that*

$$
\mathcal{U} := \mathcal{U}(z_0; \{-Qz_0\}) = \{\ell \in H^1(0, T; \mathcal{X}) \colon \ell(0) \in \ker R\}.
$$

Note that $\mathcal{U}$ is a linear subspace of $H^1(0, T; \mathcal{X})$.

Remark 5.8 (Additional Control Constraints). One could allow for additional control constraints in our analysis, even more complex ones than the ones covered by $\mathcal{U}(z_0; M)$ such as for instance box constraints over the whole time interval or vanishing initial and final loading, i.e., $\ell(0) = \ell(T) = 0$, which is certainly meaningful for many practically relevant problems. However, since the differentiability of the control-to-state map is the essential issue in the derivation of optimality conditions and additional (convex and closed) control constraints can be incorporated by standard argument, we decided to leave them out in order to keep the discussion as concise as possible.

However, without any further assumptions, the existence of solutions to the unregularized state equation (3.1) cannot be guaranteed. To be more precise, one needs

that $R\ell(0) - Qz_0 \in D(A)$, see Theorem 3.3, which holds in case of $\mathcal{U}$, provided that $-Qz_0 \in D(A)$. This is the reason for considering the set $H^1(0,T;\mathcal{X}_C) \cap \mathcal{U}$ as set of admissible controls in the rest of the paper.

Note moreover that, if the operator R is injective (which is the case in section 7), then $\mathcal{U} = \{\ell \in H^1(0,T;\mathcal{X}) : \ell(0) = 0\}$.

The chain rule immediately gives that the reduced objective defined by

$$(5.7) \qquad F \colon H^1(0,T;\mathcal{X}_c) \to \mathbb{R}, \quad \ell \mapsto J(\mathcal{S}_s(\ell), \ell)$$

is Fréchet-differentiable, too. Thus, by standard arguments, one derives the following

LEMMA 5.9 (Purely Primal Necessary Optimality Conditions). *Let Assumption 5.1 and Assumption 5.7 hold. Then, if a control $\bar{\ell} \in H^1(0,T;\mathcal{X}_c) \cap \mathcal{U}$ with associated state $\bar{z} = \mathcal{S}_s(\bar{\ell})$ is locally optimal for (P_s), then*

$$(5.8) \qquad F'(\bar{\ell})h = J'_z(\bar{z},\bar{\ell})\mathcal{S}'_s(\bar{\ell})h + J'_\ell(\bar{z},\bar{\ell})h = 0,$$

for all $h \in H^1(0,T;\mathcal{X}_c) \cap \mathcal{U}$.

Next, we reformulate (5.8) in terms of a KKT-system by introducing an adjoint equation, which formally reads

$$(5.9) \qquad \dot{\varphi} = QA'_s(y)^*\varphi + v, \quad \varphi(T) = 0.$$

Depending on the regularity of the right hand side v, we define different notions of solutions:

DEFINITION 5.10. *Let $y \in L^2(0,T;\mathcal{Y})$ and $v \in H^1(0,T;\mathcal{H})^*$ be given. A function $\varphi \in L^2(0,T;\mathcal{H})$ is called weak solution of (5.9), if*

$$(5.10) \qquad -\left(\varphi, \dot{\eta}\right)_{L^2(0,T;\mathcal{H})} = (\varphi, A'_s(y)Q\eta)_{L^2(0,T;\mathcal{H})} + v(\eta)$$

holds for all $\eta \in H^1(0,T;\mathcal{H})$ with $\eta(0) = 0$.
If v takes the form

$$(5.11) \qquad v(\eta) = (v_1, \eta)_{L^2(0,T;\mathcal{H})} + (v_2, \eta(T))_{\mathcal{H}}$$

with some $v_1 \in L^2(0,T;\mathcal{H})$ and $v_2 \in \mathcal{H}$, then we call $\varphi \in H^1(0,T;\mathcal{H})$ strong solution of (5.9), if, for almost all $t \in (0,T)$,

$$(5.12) \qquad \dot{\varphi}(t) = \left(QA'_s(y)^*\varphi\right)(t) + v_1(t) \quad \text{in } \mathcal{H}, \quad \varphi(T) = -v_2 \quad \text{in } \mathcal{H}.$$

In the following, we will—as usual—identify $v \in L^2(0,T;\mathcal{H})$ with an element of $H^1(0,T;\mathcal{H})^*$ via $(v, \cdot\,)_{L^2(0,T;\mathcal{H})}$ and denote this element with a slight abuse of notation by the same symbol.

LEMMA 5.11. *Let $y \in L^2(0,T;\mathcal{Y})$ and $v \in H^1(0,T;\mathcal{H})^*$. Then there is a unique weak solution of (5.9), which is given by $\varphi := -v \circ \mathcal{S}_y \in L^2(0,T;\mathcal{H})^* = L^2(0,T;\mathcal{H})$, where $\mathcal{S}_y \colon L^2(0,T;\mathcal{H}) \to H^1(0,T;\mathcal{H})$ is the solution operator of*

$$(5.13) \qquad \dot{\eta} = -A'_s(y)Q\eta + w, \quad \eta(0) = 0,$$

that is, $\mathcal{S}_y(w) = \eta$.
Moreover, if v is of the form (5.11), then there exists a unique strong solution of (5.9), and the weak and the strong solution coincide.

Proof. At first note that the existence of a solution of (5.13) can be proven exactly as in Lemma 5.4. Let $\eta \in H^1(0,T;\mathcal{H})$ with $\eta(0) = 0$ be arbitrary and define $w := \dot{\eta} + A'_s(y)Q\eta \in L^2(0,T;\mathcal{H})$, hence, $\eta = \mathcal{S}_y(w)$. By the definition of w and φ, it follows that

$$(\dot{\eta} + A'_s(y)Q\eta, \varphi)_{L^2(0,T;\mathcal{H})} = (\varphi, w)_{L^2(0,T;\mathcal{H})} = -v(\mathcal{S}_y(w)) = -v(\eta),$$

i.e., (5.10) holds. Since η was arbitrary, we see that φ is a weak solution of (5.9).

To prove uniqueness, let $\tilde{\varphi} \in L^2(0,T;\mathcal{H})$ be another weak solution. Then, we choose an arbitrary $w \in L^2(0,T;\mathcal{H})$ and set $\eta := \mathcal{S}_y(w)$ to see that

$$(\varphi, w)_{L^2(0,T;\mathcal{H})} = -v(\eta) = \left(\tilde{\varphi}, \dot{\eta} + A'_s(y)Q\eta\right)_{L^2(0,T;\mathcal{H})} = (\tilde{\varphi}, w)_{L^2(0,T;\mathcal{H})},$$

and therefore $\varphi = \tilde{\varphi}$.

Now we turn to the strong solution and suppose that v is as given in (5.11). Existence and uniqueness of a strong solution can again be shown by means of Banach's fixed point theorem. To this end, let us consider the affine-linear operator

$$B \colon [0,T] \times \mathcal{Z} \to \mathcal{Z}, \quad B(t,\varphi) = QA'_s(y(t))^*\varphi + v_1(t).$$

Since $\mathcal{H}$ is separable by our standing assumptions, we can apply [20, Chap. IV, Thm. 1.4] to obtain that, for every $\varphi \in L^2(0,T;\mathcal{H})$, the mapping $(0,T) \mapsto B(t,\varphi(t))$ is Bochner measureable. Moreover, since $\|A'_s(y)^*\|_{L(\mathcal{H};\mathcal{H})} = \|A'_s(y)\|_{L(\mathcal{H};\mathcal{H})}$, Assumption 5.1(ii) yields that B is also Lipschitz continuous w.r.t. the second variable for almost all $t \in (0,T)$. Therefore, similarly to the proof of Lemma 5.4, one can apply Banach's fixed point theorem to the integral equation associated with (5.12) to establish the existence of a unique strong solution.

Finally, if we test (5.12) with an arbitrary $\eta \in H^1(0,T;\mathcal{H})$ with $\eta(0) = 0$ and integrate by parts, then we see that every strong solution is also a weak solution. Since the latter one is unique, as seen above, we deduce that weak and strong solution coincide. $\qquad\square$

THEOREM 5.12 (KKT-Conditions for (P_s)). *Assume that Assumption 5.1 and Assumption 5.7 hold and let $\bar{\ell} \in H^1(0,T;\mathcal{X}_c) \cap \mathcal{U}$ be a locally optimal control for (P_s) with associated state $\bar{z} = \mathcal{S}_s(\bar{\ell})$. Then there exists a unique adjoint state $\varphi \in L^2(0,T;\mathcal{H})$ such that the following optimality system is fulfilled*

$$(5.14a) \quad \dot{\bar{z}} = A_s(R\bar{\ell} - Q\bar{z}), \qquad\qquad\qquad \bar{z}(0) = z_0,$$

$$(5.14b) \quad \begin{cases} -\left(\varphi, \dot{\eta}\right)_{L^2(0,T;\mathcal{H})} & \forall\, \eta \in H^1(0,T;\mathcal{H}): \\ = \left(\varphi, A'_s(R\bar{\ell} - Q\bar{z})Q\eta\right)_{L^2(0,T;\mathcal{H})} + J'_z(\bar{z},\bar{\ell})\eta & \eta(0) = 0 \end{cases}$$

$$(5.14c) \quad \left(\varphi, A'_s(R\bar{\ell} - Q\bar{z})Rh\right)_{L^2(0,T;\mathcal{H})} = J'_\ell(\bar{z},\bar{\ell})h \qquad \forall\, h \in H^1(0,T;\mathcal{X}_c) \cap \mathcal{U}.$$

If J enjoys extra regularity, namely

$$(5.15) \qquad\qquad J(z,\ell) = \Psi_1(z,\ell) + \Psi_2(z(T),\ell(T)) + \Phi(\ell)$$

with two Fréchet-differentiable functionals $\Psi_1 \colon L^2(0,T;\mathcal{H}) \times H^1(0,T;\mathcal{X}_c) \to \mathbb{R}$ and $\Psi_2 \colon \mathcal{H} \times \mathcal{X}_c \to \mathbb{R}$, then $\varphi \in H^1(0,T;\mathcal{H})$ is a strong solution of

$$(5.16) \quad \dot{\varphi}(t) = \left(QA'_s(R\bar{\ell} - Q\bar{z})^*\varphi\right)(t) + \frac{\partial\Psi_1}{\partial z}(\bar{z},\bar{\ell}), \qquad \varphi(T) = -\frac{\partial\Psi_2}{\partial z}(\bar{z}(T),\bar{\ell}(T)).$$

Remark 5.13. The exemplary objective functionals in section 7 are precisely of the form in (5.15).

Proof of Theorem 5.12. Since $J'_z(\overline{z}, \overline{\ell}) \in H^1(0,T;\mathcal{W})^* \hookrightarrow H^1(0,T;\mathcal{H})^*$, Lemma 5.11 gives the existence of a unique solution of (5.14b). Now, let $h \in H^1(0,T;\mathcal{X}_c) \cap \mathcal{U}$ be arbitrary and define $\eta := \mathcal{S}'_s(\overline{\ell})h \in H^1(0,T;\mathcal{Z}) \subset H^1(0,T;\mathcal{H})$. The weak form of the adjoint equation then implies

(5.17)
$$\begin{aligned}
\left(\varphi, A'_s(R\overline{\ell} - Q\overline{z})Rh\right)_{L^2(0,T;\mathcal{H})} &= \left(\varphi, \dot{\eta} + A'_s(R\overline{\ell} - Q\overline{z})Q\eta\right)_{L^2(0,T;\mathcal{H})} \\
&= -J'_z(\mathcal{S}_s(\overline{\ell}), \overline{\ell})\eta.
\end{aligned}$$

This together with Lemma 5.9 shows that $(\overline{z}, \overline{\ell}, \varphi)$ fulfills the optimality system (5.14). If J is of the form in (5.15), then Lemma 5.11 implies that the weak solution of the adjoint equation is in fact a strong solution and solves (5.16). □

COROLLARY 5.14. *Let Assumption 5.1 and Assumption 5.7 hold. Then $\overline{\ell} \in H^1(0,T;\mathcal{X}_c) \cap \mathcal{U}$ with associated state $\overline{z} = \mathcal{S}_s(\overline{\ell})$ fulfills (5.8) if and only if there exists an adjoint state $\varphi \in L^2(0,T;\mathcal{H})$ such that $(\overline{z}, \overline{\ell}, \varphi)$ satisfies the optimality system (5.14).*

Proof. The proof of Theorem 5.12 already shows that (5.8) implies the optimality system in (5.14).

To prove the reverse implication, assume that $(\overline{z}, \overline{\ell}, \varphi)$ fulfills the optimality system (5.14). Then choose an arbitrary $h \in H^1(0,T;\mathcal{H})$, define $\eta := \mathcal{S}'_s(\overline{\ell})h$, and use the fact that φ is the weak solution of (5.14b) to obtain (5.17). This together with (5.14c) finally give (5.8). □

EXAMPLE 5.15. *Under suitable additional assumptions, it is possible to further simplify the gradient equation (5.14c). For this purpose assume that R is injective (so that $\mathcal{U} = \{\ell \in H^1(0,T,\mathcal{X}) : \ell(0) = 0\}$), $\mathcal{X}_c$ is a Hilbert space, and*

(5.18)
$$J(z, \ell) = \Psi_1(z, \ell) + \Psi_2(z(T), \ell(T)) + \frac{\gamma}{2}\|\dot{\ell}\|^2_{L^2(0,T;\mathcal{X}_c)},$$

where $\Psi_1 \colon H^1(0,T;\mathcal{W}) \times L^2(0,T;\mathcal{X}_c) \to \mathbb{R}$ and $\Psi_2 \colon \mathcal{W} \times \mathcal{X}_c \to \mathbb{R}$ are Fréchet-differentiable and $\gamma > 0$. This type of objective will also appear in the application problem in section 7. Then (5.14c) becomes

(5.19)
$$\begin{aligned}
\gamma(\partial_t\overline{\ell}, \partial_t\overline{h})_{L^2(0,T;\mathcal{X}_c)} &- \int_0^T \langle R^* A'_s(R\overline{\ell} - Q\overline{z})^*\varphi, h\rangle_{\mathcal{X}^*, \mathcal{X}} \, dt \\
&+ \int_0^T \frac{\partial\Psi_1}{\partial\ell}(\overline{z}, \overline{\ell})h \, dt + \frac{\partial\Psi_2}{\partial\ell}(\overline{z}(T), \overline{\ell}(T))h(T) = 0 \\
&\qquad\qquad \forall \, h \in H^1(0,T;\mathcal{X}_c) \text{ with } h(0) = 0,
\end{aligned}$$

where we identified $\partial_\ell\Psi_1(\overline{z}, \overline{\ell}) \in (L^2(0,T;\mathcal{X}_c))^* = L^2(0,T;\mathcal{X}_c)$. Note that $\mathcal{X}_c$ as a Hilbert space satisfies the Radon-Nikodým-property. Since $\mathcal{X}_c \hookrightarrow \mathcal{X}$, we may identify $R^* A'_s(R\overline{\ell} - Q\overline{z})^*\varphi$ with an element of $L^2(0,T;\mathcal{X}_c)$, too, which we denote by the same symbol. Then, if we choose $h(t) = \psi(t)\xi$ with $\psi \in C_c^\infty(0,T)$ and $\xi \in \mathcal{X}_c$ arbitrary, we obtain

$$\left(-\int_0^T \left[\gamma\,\partial_t\psi\,\partial_t\overline{\ell} + R^* A'_s(R\overline{\ell} - Q\overline{z})^*\varphi\,\psi - \frac{\partial\Psi_1}{\partial\ell}(\overline{z}, \overline{\ell})\psi\right] dt, \xi\right)_{\mathcal{X}_c} = 0.$$

Now, since $\xi \in \mathcal{X}_c$ was arbitrary, we find that the second distributional time derivative of $\overline{\ell}$ is a regular distribution in $L^2(0, T; \mathcal{X}_c)$, i.e., $\overline{\ell} \in H^2(0, T; \mathcal{X}_c)$, satisfying for almost all $t \in (0, T)$

$$(5.20) \qquad \gamma \, \partial_t^2 \overline{\ell}(t) + R^* A_s'\big(R\overline{\ell}(t) - Q\overline{z}(t)\big)^* \varphi(t) = \frac{\partial \Psi_1}{\partial \ell}(\overline{z}, \overline{\ell})(t) \quad in \ \mathcal{X}_c.$$

Since $\mathcal{X}_c$ is supposed to be a Hilbert space, we can apply integration by parts to (5.19). Together with $\overline{\ell} \in \mathcal{U} = \{\ell \in H^1(0, T; \mathcal{X}) : \ell(0) = 0\}$ and (5.20), this implies the following boundary conditions:

$$(5.21) \qquad \overline{\ell}(0) = 0, \quad \gamma \, \partial_t \overline{\ell}(T) = -\frac{\partial \Psi_2}{\partial \ell}(\overline{z}(T), \overline{\ell}(T)),$$

where we again identified $\partial_\ell \Psi_2(\overline{z}(T), \overline{\ell}(T)) \in \mathcal{X}_c^*$ with its Riesz representative. In summary, we have thus seen that the gradient equation in (5.14c) becomes an operator boundary value problem in $\mathcal{X}_c$, namely (5.20)–(5.21).

6. Second-Order Sufficient Conditions. The next section is devoted to the derivation of second-order sufficient optimality conditions for the regularized problem (P_s). As it was the case for the first-order conditions, the main part concerns the differentiability properties of the control-to-state map $\mathcal{S}_s$ and the reduced objective, to be more precise to show that these are twice continuously Fréchet-differentiable. For this purpose, we need the following sharpened assumptions on the objective and the regularized operator A_s:

ASSUMPTION 6.1.
 (i) $J \colon H^1(0, T; \mathcal{W}) \times H^1(0, T; \mathcal{X}_c) \to \mathbb{R}$ *is twice continuously Fréchet differentiable.*
 (ii) *The Fréchet-derivative A_s' is Lipschitz continuous from $\mathcal{Y}$ to $L(\mathcal{Z}; \mathcal{Z})$. Moreover, for every $y \in \mathcal{Y}$, $A_s'(y)$ can be extended to an element of $L(\mathcal{W}; \mathcal{W})$. The mapping arising in this way is Lipschitz continuous from $\mathcal{Y}$ to $L(\mathcal{W}; \mathcal{W})$. Furthermore, there is a constant $C > 0$ such that $\|A_s'(y)w\|_{\mathcal{W}} \leq C\|w\|_{\mathcal{W}}$ hold for all $y \in \mathcal{Y}$ and all $w \in \mathcal{W}$.*
 (iii) *A_s' is Fréchet-differentiable from $\mathcal{Y}$ to $L(\mathcal{Z}; \mathcal{W})$. For all $y \in \mathcal{Y}$, its derivative $A_s''(y)$ can be extended to an element of $L(\mathcal{Z}; L(\mathcal{Z}; \mathcal{W}))$ and the mapping $y \mapsto A_s''(y)$ is continuous in these spaces. Moreover, there exists a constant C such that $\|A_s''(y)[z_1, z_2]\|_{\mathcal{W}} \leq C\|z_1\|_{\mathcal{Z}}\|z_2\|_{\mathcal{Z}}$ for all $y \in \mathcal{Y}$ and all $z_1, z_2 \in \mathcal{Z}$.*

Remark 6.2. We point out that a second norm gap arises in Assumption 6.1, since A_s' is only Fréchet-differentiable as an operator with values in $\mathcal{W}$ and not in $\mathcal{Z} \hookrightarrow \mathcal{W}$. This assumption is again motivated by the application problem in section 7. The example given there demonstrates that such as second norm gap is indeed necessary in general, since, given a concrete application, one cannot expect A_s to be twice Fréchet-differentiable in $\mathcal{Y}$, and even not as an operator from $\mathcal{Y}$ to $\mathcal{Z}$.

The following proposition addresses the second derivative of the solution operator under the above assumptions. Its proof is in principle completely along the lines of the proof of Theorem 5.5 on the first derivative of $\mathcal{S}$. We therefore postpone it to Appendix A.

PROPOSITION 6.3 (Second Derivative of the Solution Operator). *Under Assumption 5.1(ii) and Assumption 6.1(ii) & (iii), the solution operator $\mathcal{S}_s \colon H^1(0, T; \mathcal{X}) \to H^1(0, T; \mathcal{W})$ is twice Fréchet differentiable. Given $\ell, h_1, h_2 \in H^1(0, T; \mathcal{X})$, its second*

derivative $\mathcal{S}_s''(\ell)[h_1, h_2] \in H^1(0,T;\mathcal{W})$ *is given by the unique solution of*

$$(6.1) \qquad \dot{\xi} = A_s''(R\ell - Qz)[Rh_1 - Q\eta_1, Rh_2 - Q\eta_2] - A_s'(R\ell - Qz)Q\xi, \quad \xi(0) = 0,$$

where $z := \mathcal{S}_s(\ell) \in H^1(0,T;\mathcal{Y})$ *and* $\eta_i := \mathcal{S}_s'(\ell)h_i \in H^1(0,T;\mathcal{Z})$, $i = 1, 2$.
    Moreover, there exists a constant C *such that*

$$(6.2) \qquad \|\mathcal{S}_s''(\ell)[h_1,h_2]\|_{H^1(0,T;\mathcal{W})} \le C\|h_1\|_{H^1(0,T;\mathcal{X})}\|h_2\|_{H^1(0,T;\mathcal{X})}$$

for all $\ell, h_1, h_2 \in H^1(0,T;\mathcal{X})$.

LEMMA 6.4. *Assume that Assumption 5.1 (ii) and Assumption 6.1 (ii) and (iii) are fulfilled. Then there exists a constant* C *such that*

$$\|\mathcal{S}_s''(\ell_1) - \mathcal{S}_s''(\ell_2)\|_{L(H^1(0,T;\mathcal{X});L(H^1(0,T;\mathcal{X});H^1(0,T;\mathcal{W})))}$$
$$\le C\big(\|A_s''(R\ell_1 - Qz_1) - A_s''(R\ell_2 - Qz_2)\|_{L^2(0,T;L(\mathcal{Z};L(\mathcal{Z};\mathcal{W})))} + \|\ell_1 - \ell_2\|_{H^1(0,T;\mathcal{X})}\big)$$

holds for all $\ell_1, \ell_2 \in H^1(0,T;\mathcal{X})$, *where* $z_i := \mathcal{S}_s(\ell_i)$, $i = 1, 2$.

Proof. Let $\ell_1, \ell_2, h_1, h_2 \in H^1(0,T;\mathcal{X})$ be arbitrary. We again abbreviate $z_i := \mathcal{S}_s(\ell_i)$, $\eta_{i,j} := \mathcal{S}_s'(\ell_i)h_j$, $\xi_i := \mathcal{S}_s''(\ell_i)[h_1, h_2]$, and $y_i := R\ell_i - Qz_i$ for $i, j \in \{1, 2\}$. By the equation for $\mathcal{S}_s''$, we obtain for almost all $t \in [0, T]$

$$\dot{\xi}_1 - \dot{\xi}_2 = A_s''(y_1)[Rh_1 - Q\eta_{1,1}, Rh_2 - Q\eta_{1,2}] - A_s'(y_1)Q\xi_1$$
$$\quad - A_s''(y_2)[Rh_1 - Q\eta_{2,1}, Rh_2 - Q\eta_{2,2}] - A_s'(y_2)Q\xi_2$$
$$= \big(A_s''(y_1)(Rh_1 - Q\eta_{1,1}) - A_s''(y_2)(Rh_1 - Q\eta_{2,1})\big)(Rh_2 - Q\eta_{1,2})$$
$$\quad + A_s''(y_2)[Rh_1 - Q\eta_{2,1}, Q(\eta_{2,2} - \eta_{1,2})]$$
$$\quad + \big(A_s'(y_2) - A_s'(y_1)\big)Q\xi_1 + A_s'(y_2)Q(\xi_2 - \xi_1).$$

With the help of

$$A_s''(y_1)(Rh_1 - Q\eta_{1,1}) - A_s''(y_2)(Rh_1 - Q\eta_{2,1})$$
$$= \big(A_s''(y_1) - A_s''(y_2)\big)(Rh_1 - Q\eta_{1,1}) + A_s''(y_2)Q(\eta_{2,1} - \eta_{1,1}),$$

and Gronwall's inequality, we thus arrive at

$$\|\xi_1 - \xi_2\|_{H^1(0,T;\mathcal{W})}$$
$$\le C\Big[\|Rh_1 - Q\eta_{2,1}\|_{H^1(0,T;\mathcal{Z})}\|\eta_{1,2} - \eta_{2,2}\|_{H^1(0,T;\mathcal{Z})}$$
$$\quad + \|y_1 - y_2\|_{H^1(0,T;\mathcal{Y})}\|\xi_1\|_{H^1(0,T;\mathcal{W})}$$
$$\quad + \big(\|A_s''(y_1) - A_s''(y_2)\|_{L^2(0,T;L(\mathcal{Z};L(\mathcal{Z};\mathcal{W})))}\|Rh_1 - Q\eta_{1,1}\|_{H^1(0,T;\mathcal{Z})}$$
$$\quad\quad\quad\quad\quad\quad + \|\eta_{1,1} - \eta_{2,1}\|_{H^1(0,T;\mathcal{Z})}\big)\|Rh_2 - Q\eta_{1,2}\|_{H^1(0,T;\mathcal{Z})}\Big]$$
$$\le C\Big(\|A_s''(y_1) - A_s''(y_2)\|_{L^2(0,T;L(\mathcal{Z};L(\mathcal{Z};\mathcal{W})))}$$
$$\quad\quad\quad\quad\quad\quad + \|\ell_1 - \ell_2\|_{H^1(0,T;\mathcal{X})}\Big)\|h_1\|_{H^1(0,T;\mathcal{X})}\|h_2\|_{H^1(0,T;\mathcal{X})},$$

where we used the estimate in Theorem 5.5, (6.2), and the Lipschitz continuity of $\mathcal{S}_s'$ by Lemma A.1 in the Appendix.   □

If A_s'' were Lipschitz continuous from $\mathcal{Y}$ to $L(\mathcal{Z}; L(\mathcal{Z}; \mathcal{W}))$, then Lemma 6.4 would immediately imply the Lipschitz continuity of $\mathcal{S}_s''$. However, to obtain the continuity of the second derivative, this additional assumption is not necessary as the following theorem shows:

THEOREM 6.5 (Second-Order Continuous Fréchet-Differentiability of the Solution Operator). *Suppose that Assumption 5.1(ii) and Assumption 6.1(ii) & (iii) are fulfilled. Then $\mathcal{S}_s \colon H^1(0,T;\mathcal{X}) \to H^1(0,T;\mathcal{W})$ is twice continuously Fréchet-differentiable. Its second derivative at $\ell \in H^1(0,T;\mathcal{X})$ in directions $h_1, h_2 \in H^1(0,T;\mathcal{X})$ is given by the unique solution of* (6.1).

Proof. Thanks to Proposition 6.3, we only have to show that $\mathcal{S}_s'' \colon H^1(0,T;\mathcal{X}) \to L(H^1(0,T;\mathcal{X}); L(H^1(0,T;\mathcal{X}); H^1(0,T;\mathcal{W})))$ is continuous. For this let $\{\ell_n\}_{n\in\mathbb{N}} \subset H^1(0,T;\mathcal{X})$ and $\ell \in H^1(0,T;\mathcal{X})$ be given such that $\ell_n \to \ell$ in $H^1(0,T;\mathcal{X})$ so that in particular $\ell_n \to \ell$ in $C([0,T];\mathcal{X})$. Then, Lemma 5.3 implies $z_n := \mathcal{S}_s(\ell_n) \to \mathcal{S}_s(\ell) =: z$ in $C([0,T];\mathcal{Y})$. With this convergence results at hand, we can apply Lemma C.3 with $M = [0,T]$, $N = \mathcal{Y}$, $G_n = R\ell_n - Qz_n$ and $G = R\ell - Qz$ to see that

$$U := \Big(\bigcup_{n=1}^{\infty} (R\ell_n - Qz_n)([0,T]) \Big) \cup \Big((R\ell - Qz)([0,T]) \Big)$$

is compact. Therefore, thanks to the continuity assumption in Assumption 6.1(iii), $A_s'' \colon \mathcal{Y} \to L(\mathcal{Z}; L(\mathcal{Z}; \mathcal{W}))$ is uniformly continuous on U. Consequently, $A_s''(R\ell_n - Qz_n)$ converges to $A_s''(R\ell - Qz)$ in $C([0,T]; L(\mathcal{Z}; L(\mathcal{Z}; \mathcal{W})))$, which, together with Lemma 6.4, yields the assertion. $\qquad\square$

Remark 6.6. It is to be noted that the regularized state equation (5.1) and the equations corresponding to the derivatives of $\mathcal{S}_s$, i.e., (5.3) and (6.1), provide more regular solutions under the hypotheses of Assumption 5.1(ii) and Assumption 6.1(ii) & (iii). Indeed, if $\ell, h_1, h_2 \in H^1(0,T;\mathcal{X})$, then the solutions of all three equations can be shown to be continuously differentiable in time with values in the respective spaces ($\mathcal{Y}$, $\mathcal{Z}$, and $\mathcal{W}$, respectively). Moreover, the time derivatives of z and η are absolutely continuous and the same would hold for ξ, if A_s'' were Lipschitz continuous. However, we did not exploit this additional regularity, since the original unregularized problem (P) does not provide this property in general.

With the above differentiability result at hand, it is now standard to derive the following:

THEOREM 6.7 (Second-Order Sufficient Optimality Conditions for (P_s)). *Assume that Assumption 5.1, Assumption 5.7, and Assumption 6.1 hold. Let $(\bar{z},\bar{\ell},\varphi) \in H^1(0,T;\mathcal{Y}) \times (H^1(0,T;\mathcal{X}_c) \cap \mathcal{U}) \times L^2(0,T;\mathcal{H})$ be a solution of the optimality system* (5.14). *Moreover, suppose that there is a $\delta > 0$ such that*

(SSC) $$F''(\bar{\ell})h^2 \geq \delta \|h\|_{H^1(0,T;\mathcal{X}_c)}^2$$

for all $h \in H^1(0,T;\mathcal{X}_c) \cap \mathcal{U}$, where F is the reduced objective from (5.7). *Then $(\bar{z},\bar{\ell})$ is locally optimal for (P_s) and there exist $\varepsilon > 0$ and $\tau > 0$ such that the following quadratic growth condition*

(6.3) $$F(\ell) \geq F(\bar{\ell}) + \tau \|\ell - \bar{\ell}\|_{H^1(0,T;\mathcal{X}_c)}^2$$

holds for all $\ell \in H^1(0,T;\mathcal{X}_c) \cap \mathcal{U}$ with $\|\ell - \bar{\ell}\|_{H^1(0,T;\mathcal{X}_c)} \leq \varepsilon$.

Proof. Thanks to the assumptions on J and Theorem 6.5, the chain rule implies that the reduced objective function $F(\cdot) = J(\mathcal{S}_s(\cdot), \cdot) \colon H^1(0, T; \mathcal{X}_c) \to \mathbb{R}$ is twice continuously Fréchet-differentiable and, according to Corollary 5.14, the equation in (5.8) holds for all $h \in H^1(0, T; \mathcal{X}_c) \cap \mathcal{U}$. Since $\mathcal{U}$ is a linear subspace, the claim then follows from standard arguments, see e.g. [39, Satz 4.23]. □

Remark 6.8. As already mentioned in Remark 5.8, one could also account for additional control constraints. In this case, a critical cone would arise in the second-order conditions, cf. e.g. the survey article [13].

Using the adjoint equation, the second derivative of the reduced objective in (SSC) can be reformulated as follows:

COROLLARY 6.9. *Assume in addition to the hypotheses of Assumption* 6.1*(iii) that* $\|A_s''(y)[z_1, z_2]\|_{\mathcal{H}} \leq C \|z_1\|_{\mathcal{Z}} \|z_2\|_{\mathcal{Z}}$ *for all* $y \in \mathcal{Y}$ *and* $z_1, z_2 \in \mathcal{Z}$*, i.e., the last inequality in Assumption* 6.1 *holds in* $\mathcal{H}$ *instead of the weaker space* $\mathcal{W}$*. Then it holds for all* $\ell, h \in H^1(0, T; \mathcal{H})$ *that*

$$F''(\ell)h^2 = \Psi''(z, \ell)(\eta, h)^2 + \Phi''(\ell)h^2 - \big(\varphi, A_s''(R\ell - Qz)(Rh - Q\eta)^2\big)_{L^2(0,T;\mathcal{H})},$$

where $z = \mathcal{S}_s(\ell)$*,* $\eta = \mathcal{S}_s'(\ell)h$*, and* φ *solves the adjoint equation in* (5.14b)*.*

Proof. Let us again abbreviate $y = R\ell - Qz$. According to the chain rule, the second derivative of the reduced objective is given by

$$F''(\ell)h^2 = \tfrac{\partial^2}{\partial \ell^2} J(z, \ell)h^2 + \tfrac{\partial^2}{\partial z^2} J(z, \ell)\eta^2 + 2 \tfrac{\partial^2}{\partial \ell \partial z} J(z, \ell)[h, \eta] + \tfrac{\partial}{\partial z} J(z, \ell)\xi$$
$$= \Psi''(z, \ell)(\eta, h)^2 + \Phi''(\ell)h^2 + \tfrac{\partial}{\partial z} J(z, \ell)\xi$$

with $z = \mathcal{S}_s(\ell)$, $\eta = \mathcal{S}_s'(\ell)h$, and $\xi = \mathcal{S}_s''(\ell)h^2$. Now, since $A_s''(y)$ is a bilinear form on $\mathcal{H}$ by assumption, we obtain that $\xi \in H^1(0, T; \mathcal{H})$. Therefore, we are allowed to test the adjoint equation in (5.14b) (in its weak form) with ξ, which results in

$$\tfrac{\partial}{\partial z} J(z, \ell)\xi = -(\varphi, \dot{\xi} + A_s'(y)Q\xi)_{L^2(0,T;\mathcal{H})} = -(\varphi, A_s''(y)(Rh - Q\eta)^2)_{L^2(0,T;\mathcal{H})},$$

where we used the precise form of $\mathcal{S}_s''(\ell)$ in (6.1) for the last identity. □

7. Application to Optimal Control of Homogenized Elastoplasticity. In the upcoming sections, we apply the analysis from the previous sections to an optimal control problem governed by a system of equations that arise as homogenization limit in elastoplasticity and was derived in [33, Theorem. 2.2]. It describes the evolution of plastic deformation in a material with periodic microstructure and formally (i.e., in its strong form) reads as follows:

$$
\begin{align}
(7.1a) && -\nabla_x \cdot \pi\Sigma &= f && \text{in } \Omega, \\
(7.1b) && \Sigma &= \mathbb{C}(\nabla_x^s u + \nabla_y^s v - Bz) && \text{in } \Omega \times Y, \\
(7.1c) && -\nabla_y \cdot \Sigma &= 0 && \text{in } \Omega \times Y, \\
(7.1d) && \dot{z} &\in A(B^\top \Sigma - \mathbb{B}z) && \text{in } \Omega \times Y, \\
(7.1e) && u &= 0 && \text{on } \Gamma_D, \\
(7.1f) && \nu \cdot \pi\Sigma &= g && \text{on } \Gamma_N, \\
(7.1g) && z(0) &= z_0 && \text{in } \Omega \times Y.
\end{align}
$$

Herein, $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, is a given domain occupied by the body under consideration, while $Y = [0, 1]^d$ is the unit cell. The boundary of Ω consists of two disjoint parts, the

Dirichlet boundary Γ_D and the Neumann boundary Γ_N. Furthermore, $u\colon (0,T)\times\Omega \to \mathbb{R}^d$ is the displacement on the macro level, while $v\colon (0,T)\times\Omega\times Y \to \mathbb{R}^d$ is the displacement reflecting the micro structure. The stress tensor is denoted by $\Sigma\colon (0,T)\times\Omega\times Y \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ and $z\colon (0,T)\times\Omega\times Y \to \mathbb{V}$ is the internal variable describing changes in the material behavior under plastic deformation (such as hardening), where $\mathbb{V}$ is a finite dimensional Banach space. Moreover, $\nabla^s_x := \frac{1}{2}(\nabla_x + \nabla^\top_x)$ is the linearized strain in Ω and ∇^s_y is defined analogously. The elasticity tensor $\mathbb{C}\colon \Omega\times Y \to \mathcal{L}(\mathbb{R}^{d\times d}_{\mathrm{sym}})$ and the hardening parameter $\mathbb{B}\colon \Omega\times Y \to \mathcal{L}(\mathbb{V})$ are given linear and coercive mappings and, by $B\colon \Omega\times Y \to \mathcal{L}(\mathbb{V};\mathbb{R}^{d\times d}_{\mathrm{sym}})$, one recovers the plastic strain from the internal variables z. The evolution of the internal variables is determined by a maximal monotone operator $A\colon \mathbb{V} \to 2^{\mathbb{V}}$. In subsection 7.4 below, we present a concrete example for such an operator, namely the case of linear kinematic hardening with von Mises yield condition. Finally, z_0 is a given initial state and π is the averaging over the unit cell, i.e.,

$$(7.2) \qquad \pi\colon \Sigma \mapsto \fint_Y \Sigma(\cdot,y)\, dy := \frac{1}{|Y|}\int_Y \Sigma(\cdot,y)\, dy.$$

The precise assumptions on these data as well as the precise notion of solutions to (7.1) are given below.

The volume force $f\colon (0,T)\times\Omega \to \mathbb{R}^d$ and the boundary loads $g\colon (0,T)\times\Gamma_N \to \mathbb{R}^d$, serve as control variables. In the following, we will frequently write ℓ for the tuple (f,g). Possible objectives could include a desired displacement or stress distribution at end time, i.e.,

$$J(u,\Sigma,\ell) := \frac{\alpha}{2}\int_\Omega |u(T) - u_d|^2\, dx + \frac{\beta}{2}\int_\Omega |(\pi\Sigma)(T) - \sigma_d|^2\, dx + \Phi(\ell),$$

where $u_d\colon \Omega \to \mathbb{R}^d$ and $\sigma_d\colon \Omega \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ are given desired displacement and stress field, respectively, $\alpha,\beta \geq 0$, and Φ is a regularization term depending on the choice of the control space that will be specified below, see Remark 7.14.

7.1. Homogenized Plasticity – Notation and Standing Assumptions. Before discussing the optimal control problem, we first have to introduce the precise notion of solution for homogenized elastoplasticity system in (7.1). For this purpose, we need several assumptions and definitions. We start with the following

ASSUMPTION 7.1 (Hypotheses on the data in (7.1)).
- Regularity of the domain: *The domain $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, is bounded with Lipschitz boundary Γ. The boundary consists of two disjoint measurable parts Γ_N and Γ_D such that $\Gamma = \Gamma_N \cup \Gamma_D$. While Γ_N is a relatively open subset, Γ_D is a relatively closed subset of Γ with positive measure. In addition, the set $\Omega \cup \Gamma_N$ is regular in the sense of Gröger, cf. [24].*
- Assumptions on the coefficients: *The elasticity tensor and the hardening parameter satisfy $\mathbb{C} \in L^\infty(\Omega\times Y;\mathcal{L}(\mathbb{R}^{d\times d}_{\mathrm{sym}}))$ and $\mathbb{B} \in L^\infty(\Omega\times Y;\mathcal{L}(\mathbb{V}))$ and are symmetric and uniformly coercive, i.e., there exist constants $\underline{c} > 0$ and $\underline{b} > 0$ such that*

$$\mathbb{C}(x,y)\sigma\colon \sigma \geq \underline{c}\,\|\sigma\|^2_{\mathbb{R}^{d\times d}} \quad \forall\,\sigma \in \mathbb{R}^{d\times d}_{\mathrm{sym}}, \quad f.a.a.\ (x,y) \in \Omega\times Y,$$
$$\mathbb{B}(x,y)\zeta\colon \zeta \geq \underline{b}\,\|\zeta\|^2_{\mathbb{V}} \qquad \forall\,\zeta \in \mathbb{V}, \qquad f.a.a.\ (x,y) \in \Omega\times Y.$$

In addition, $B \in L^\infty(\Omega\times Y;\mathcal{L}(\mathbb{V};\mathbb{R}^{d\times d}_{\mathrm{sym}}))$ is a given linear mapping.

Next, we define the function spaces for the various variables in (7.1):

DEFINITION 7.2 (Function spaces). *Let $s \in [1, \infty)$. For the quantities in* (7.1), *we define the following spaces:*

- *space for the macro displacement u:*

$$U^s := W_D^{1,s}(\Omega; \mathbb{R}^d) := \overline{\left\{\psi|_\Omega \colon \psi \in C_0^\infty(\mathbb{R}^d; \mathbb{R}^d),\ \operatorname{supp}(\psi) \cap \Gamma_D = \emptyset\right\}}^{W^{1,s}(\Omega;\mathbb{R}^d)}$$

- *space for the internal variable z:*

$$Z^s := L^s\big(\Omega \times Y; \mathbb{V}\big)$$

- *stress space for Σ:*

$$S^s := L^s\big(\Omega \times Y; \mathbb{R}_{\mathrm{sym}}^{d \times d}\big).$$

- *space for the micro displacement v:*

$$V^s := L^s\big(\Omega; W_{\mathrm{per},\perp}^{1,s}(Y; \mathbb{R}^d)\big).$$

For the latter, we denote by $C_{\mathrm{per}}^\infty(Y; \mathbb{R}^d)$ the space of $C^\infty(\mathbb{R}^d; \mathbb{R}^d)$ functions which are Y-periodic, identified with their restriction on Y, and define $W_{\mathrm{per}}^{1,s}(Y; \mathbb{R}^d)$ to be the closure of $C_{\mathrm{per}}^\infty(Y; \mathbb{R}^d)$ with respect to the $W^{1,s}(Y; \mathbb{R}^d)$ norm. Further, $W_\perp^{1,s}(Y; \mathbb{R}^d)$ is the closed subspace of $W^{1,s}(Y; \mathbb{R}^d)$ consisting of functions of mean 0, and

$$W_{\mathrm{per},\perp}^{1,s}(Y; \mathbb{R}^d) = W_{\mathrm{per}}^{1,s}(Y; \mathbb{R}^d) \cap W_\perp^{1,s}(Y; \mathbb{R}^d).$$

We set the norm on V^s to be

$$\|v\|_{V^s} := \|v\|_{L^s(\Omega \times Y; \mathbb{R}^d)} + \|\nabla_y^s v\|_{L^s(\Omega \times Y; \mathbb{R}^{d \times d})},$$

with which V^s becomes a Banach space and for the case $s = 2$ a Hilbert space with the obvious scalar product.

ASSUMPTION 7.3 (Maximal monotone operator).   *The maximal monotone operator A from the evolution law in* (7.1d) *is a set-valued map in the Hilbert space $\mathcal{H} = Z^2$, i.e., $A \colon Z^2 \to 2^{Z^2}$. It is assumed to satisfy our standing assumptions from section 2. Moreover, we assume that there is a sequence of operators $\{A_n\}$ from Z^2 to Z^2 satisfying Assumption 4.3.*

In subsection 7.4 below, we will investigate the maximal monotone operator arising in the case of linear kinematic hardening with von Mises yield condition and show how to construct the approximating sequence of smooth operators for this particular case. With the above definitions at hand, we are now in the position to define our precise notion of solutions to (7.1):

DEFINITION 7.4 (Weak solutions).   *Let $\ell \in H^1(0, T; (U^2)^*)$ and $z_0 \in Z^2$. Then we say that a tuple*

$$(u, v, z, \Sigma) \in H^1(0, T; U^s) \times H^1(0, T; V^2) \times H^1(0, T; Z^2) \times H^1(0, T; S^2)$$

is a solution of (7.1)*, if, for almost all* $t \in (0, T)$*, there holds*

$$(7.3\text{a}) \qquad \int_{\Omega} (\pi \Sigma(t))(x) \cdot \nabla_x^s \varphi(x) \, dx = \langle \ell(t), \varphi \rangle \qquad\qquad \forall \varphi \in U^2,$$

$$(7.3\text{b}) \qquad\qquad\qquad \Sigma(t) = \mathbb{C}(\pi_r^{-1} \nabla_x^s u(t) + \nabla_y^s v(t) - Bz(t)) \quad in \; S^2,$$

$$(7.3\text{c}) \quad \int_{\Omega \times Y} \Sigma(t, x, y) \cdot \nabla_y^s \psi(x, y) \, dy = 0 \qquad\qquad \forall \psi \in V^2,$$

$$(7.3\text{d}) \qquad\qquad\qquad \dot{z}(t) \in A(B^\top \Sigma(t) - \mathbb{B}z(t)) \qquad\qquad in \; Z^2,$$

$$(7.3\text{e}) \qquad\qquad\qquad z(0) = z_0 \qquad\qquad\qquad\qquad in \; Z^2,$$

where $\pi \colon S^2 \to L^2(\Omega; \mathbb{R}_{\mathrm{sym}}^{d \times d})$ *is the average mapping from* (7.2) *and*

$$\pi_r^{-1} \colon L^2(\Omega; \mathbb{R}_{\mathrm{sym}}^{d \times d}) \ni \varepsilon \mapsto \left(\Omega \times Y \ni (x, y) \mapsto \varepsilon(x) \in \mathbb{R}_{\mathrm{sym}}^{d \times d} \right) \in S^2.$$

In the following, we will frequently consider π_r^{-1} *in different domains and ranges, for simplicity denoted by the same symbol.*

7.2. Reduction of the System. In the following, we reduce the system (7.3) to an equation in the internal variable z only and it will turn out that this equation has exactly the form of our general equation (3.1). To this end we proceed analog to [23, Chapter 4]. For this purpose, let us define the following operators:

DEFINITION 7.5. *Let* $s \in [1, \infty)$. *Then we define*

$$\nabla_{(x,y)}^s \colon U^s \times V^s \to S^s, \quad \nabla_{(x,y)}^s(u, v) := \pi_r^{-1} \nabla_x^s u + \nabla_y^s v.$$

For its adjoint, we write

$$\mathrm{Div}_{(x,y)} \colon S^{s'} \to (U^s)^* \times (V^s)^*,$$
$$\langle \mathrm{Div}_{(x,y)} \sigma, (\varphi, \psi) \rangle := -\langle \nabla_{(x,y)}^s {}^* \sigma, (\varphi, \psi) \rangle$$
$$= - \int_{\Omega \times Y} \sigma(x, y) \colon (\nabla_x^s \varphi(x) + \nabla_y^s \psi(x, y)) \, d(x, y).$$

With a slight abuse of notation, we denote these operators for different values of s always by the same symbol.

LEMMA 7.6. *Let* Assumption 7.1 *be fulfilled. Then there is an index* $\bar{s} > 2$ *such that, for every* $s \in [\bar{s}', \bar{s}]$ *and every* $(\mathfrak{f}, \mathfrak{g}) \in (U^{s'})^* \times (V^{s'})^*$*, there exists a unique solution* $(u, v) \in U^s \times V^s$ *of*

$$(7.4) \qquad\qquad - \mathrm{Div}_{(x,y)} \left(\mathbb{C} \nabla_{(x,y)}^s(u, v) \right) = (\mathfrak{f}, \mathfrak{g}) \quad in \; (U^{s'})^* \times (V^{s'})^*$$

and there is a constant $C_s > 0$*, independent of* $\mathfrak{f}$ *and* $\mathfrak{g}$*, such that*

$$\|(u, v)\|_{U^s \times V^s} \le C_s \big(\|\mathfrak{f}\|_{(U^{s'})^*} + \|\mathfrak{g}\|_{(V^{s'})^*} \big).$$

Proof. The claim is equivalent to $-\mathrm{Div}_{(x,y)} \mathbb{C} \nabla_{(x,y)}^s$ being a topological isomorphism between $U^s \times V^s$ and its dual space $(U^s)^* \times (V^s)^*$ for $s \in [\bar{s}', \bar{s}]$. We start with the case $s = 2$. For this, the left hand side of (7.4) gives rise to a bilinear form $\mathfrak{b}$ on the Hilbert space $U^2 \times V^2$:

$$\mathfrak{b}\big((u, v), (\varphi, \psi)\big) := \big(\mathbb{C} \nabla_{(x,y)}^s(u, v), \nabla_{(x,y)}^s(\varphi, \psi) \big)_{S^2}$$

Clearly, $\mathfrak{b}$ is bounded. Due to Poincaré's inequality for functions with zero mean value, which implies

$$\mathfrak{b}\big((u,v),(u,v)\big) \geq \underline{c}\,|Y| \int_\Omega |\nabla_x^s u(x)|^2\, dx + \underline{c} \int_{\Omega \times Y} |\nabla_y^s v(x,y)|^2\, d(x,y)$$

$$\geq C\big(\|u\|_{H_D^1(\Omega;\mathbb{R}^d)} + \|v\|_{V^2(\Omega \times Y;\mathbb{R}^d)}\big)^2 \quad \forall\, (u,v) \in U^2 \times V^2,$$

it is also coercive so that the claim and isomorphism property for $-\operatorname{Div}_{(x,y)} \mathbb{C}\nabla_{(x,y)}^s$ for $s = 2$ follows from the Lax-Milgram lemma.

We next extrapolate this isomorphism property to $U^s \times V^s$ for s around 2 using the fundamental stability theorem by Šneĭberg [36]. (See also the more accessible and extensive [3, Appendix A].) More precisely, we show that the spaces $U^s \times V^s$ and their duals form complex interpolation scales in s. Then the stability theorem shows that the set of scale parameters s such that $-\operatorname{Div}_{(x,y)} \mathbb{C}\nabla_{(x,y)}^s$ is a topological isomorphism between $U^s \times V^s$ and its dual space is open. Since the set includes 2, as seen above, this then implies the claim.

To establish the interpolation scale, it is enough to consider the primal case, since the dual interpolation scale is inherited from the primal one by duality properties of the complex interpolation functor [38, Theorem 1.11.3]. So, we show that

$$U^{s_\theta} \times V^{s_\theta} = \big[U^{s_0} \times V^{s_0}, U^{s_1} \times V^{s_1}\big]_\theta \quad \text{for} \quad \frac{1}{s} = \frac{1-\theta}{s_0} + \frac{\theta}{s_1}$$

for all $s_0, s_1 \in (1, \infty)$ and $\theta \in (0,1)$. It is moreover sufficient to consider each component in the interpolation separately.

For the $U^s = W_D^{1,s}(\Omega;\mathbb{R}^d)$ spaces, the interpolation scale property is well known by now in the setting of Assumption 7.1 and even much more general ones; we refer to [6]. The result for V^s is proven by reducing the problem to the $W_{\mathrm{per},\perp}^{1,s}(Y;\mathbb{R}^d)$ spaces and showing that these are complemented subspaces of $W^{1,s}(Y;\mathbb{R}^d)$ and thus inherit the latter's interpolation properties. This is done in the appendix, Theorem B.3, and finishes the proof. $\qquad\square$

Remark 7.7. In general, one cannot expect $\bar{s}$ to be significantly larger than 2, due to both the irregular coefficient tensors and the mixed boundary conditions, see e.g. [18, 28, 34]. This issue will become crucial in the discussion of second-order necessary optimality conditions in 7.3 below.

Now we are in the position to reduce (7.3) to an equation in the variable z only. For this purpose, we need the following

DEFINITION 7.8 (Q and R for the case of homogenized plasticity). *Let $s \in [\bar{s}', \bar{s}]$ be given. By Lemma 7.6, the solution operator associated with (7.4), denoted by*

$$\mathcal{G} := \big(-\operatorname{Div}_{(x,y)} \mathbb{C}\nabla_{(x,y)}^s\big)^{-1} \colon (U^{s'})^* \times (V^{s'})^* \to U^s \times V^s,$$

is well defined, linear and bounded. The components of $\mathcal{G}$ are abbreviated by

$$\mathcal{G}_u := (1,0)\,\mathcal{G} \colon (U^{s'})^* \times (V^{s'})^* \to U^s, \quad \mathcal{G}_v := (0,1)\,\mathcal{G} \colon (U^{s'})^* \times (V^{s'})^* \to V^s.$$

Based on this solution operator, we moreover define

$$T \colon Z^s \ni z \mapsto B^\top \mathbb{C}\nabla_{(x,y)}^s \mathcal{G}(-\operatorname{Div}_{(x,y)}(\mathbb{C}Bz)) \in Z^s.$$

Now, we have everything at hand to define the mappings R and Q from our general equation (3.1) *for the special case of homogenized plasticity:*

$$(7.5) \qquad\qquad R: (U^{s'})^* \ni \ell \mapsto B^\top \mathbb{C} \nabla^s_{(x,y)} \mathcal{G}(\ell, 0) \in Z^s,$$

$$(7.6) \qquad\qquad Q: Z^s \ni z \mapsto (B^\top \mathbb{C} B + \mathbb{B} - T)z \in Z^s.$$

Again, with a slight abuse of notation, we denote all of the above operators for different values of $s \in [\bar{s}', \bar{s}]$ always by the same symbol.

The reason for defining the operators Q and R in the way we did in Definition 7.8 is the following: Owing to Lemma 7.6, given $z \in Z^2$, one can solve (7.3a)–(7.3c) for u, v, and Σ so that the tuple $(u, v, \Sigma) \in U^2 \times V^2 \times S^2$ is uniquely determined by z. Even more, using the operators from Definition 7.8, we see that the solution of (7.3a)–(7.3c) for given z is

$$(7.7) \qquad\qquad (u, v) = \mathcal{G}(-\operatorname{Div}_{(x,y)}(\mathbb{C}Bz)) + (\ell, 0)),$$

$$(7.8) \qquad\qquad \Sigma = \mathbb{C}\left[\nabla^s_{(x,y)}\mathcal{G}\left(-\operatorname{Div}_{(x,y)}(\mathbb{C}Bz)) + (\ell, 0)\right) - Bz\right].$$

Inserting the last equation in (7.3d) and employing the definition of Q and R in (7.6) and (7.5) then yields

$$\dot{z} \in A(B^\top \Sigma - \mathbb{B}z) = A(R\ell - Qz),$$

i.e., exactly an evolution equation of the general form in (3.1). This shows that the system (7.3) of homogenized elastoplasticity can equivalently be rewritten as an abstract operator evolution equation of the form (3.1).

For the differentiability properties needed in sections 5 and 6, a norm gap is required such that it is no longer sufficient to consider just the Hilbert space $\mathcal{H} = Z^2$. In accordance with the definitions of R and Q, we therefore define the spaces in the abstract setting in our concrete application problem as follows:

DEFINITION 7.9 (Spaces in case of homogenized plasticity). *The spaces $\mathcal{Y}$, $\mathcal{Z}$, $\mathcal{H}$, and $\mathcal{W}$ from section 2 are set to*

$$(7.9) \qquad\qquad \mathcal{Y} := Z^{s_1} \hookrightarrow \mathcal{Z} := Z^{s_2} \hookrightarrow \mathcal{H} = Z^2 \hookrightarrow \mathcal{W} := Z^{s_3}$$

with $s_1 \geq s_2 \geq 2 \geq s_3$. The integrability indices s_1, s_2, and s_3 depend crucially on the differentiability properties of the regularized version of A and will be specified for a concrete realization of A in subsection 7.4 below. Moreover, we choose

$$\mathcal{X} := (U^{s_1'})^* = W_D^{1,s_1'}(\Omega; \mathbb{R}^d)^* =: W_D^{-1,s_1}(\Omega; \mathbb{R}^d).$$

Furthermore, the control space is given by

$$(7.10) \qquad \mathcal{X}_c := L^p(\Omega; \mathbb{R}^d) \times L^r(\Gamma_N; \mathbb{R}^d) \quad \text{with} \quad p > \frac{ds_1}{d+s_1} \quad \text{and} \quad r > \frac{(d-1)s_1}{d}.$$

Due to $s_1 \geq 2$, $\mathcal{X}_c$ is reflexive and embeds compactly in $\mathcal{X}$ by Sobolev embedding and trace theorems. Therefore, all our standing assumptions on the spaces in section 2 are fulfilled.

Of course, elements in $\mathcal{X}_c$ are identified with those in $\mathcal{X}$ by

$$\langle (f, g), u \rangle_{(U^{s_1})^*, U^{s_1}} := \int_\Omega f\, u\, dx + \int_{\Gamma_N} g\, u\, ds, \quad (f, g) \in \mathcal{X}_c,\, u \in U^{s_1}.$$

In order to apply our general theory to the present setting, we need the following assumption on the regularity of the linear equation (7.4). As we will see in subsections 7.3 and 7.4 below, this assumption may become fairly restrictive, if one aims to establish second-order sufficient optimality conditions, since, in this case, s_1 and the conjugate index s_3' may be rather large.

ASSUMPTION 7.10 (Critical regularity condition). *The index $\bar{s}$ from Lemma 7.6 satisfies $\bar{s} \geq \max\{s_1, s_3'\}$, where s_1 and s_3 are the integrability indices from (7.9).*

PROPOSITION 7.11. *Under Assumption 7.10 and with the spaces defined in Definition 7.9, the operators R and Q from (7.5) and (7.6), respectively, satisfy the standing assumptions from section 2, that is, R is linear and bounded from $(U^{s_1'})^*$ to Z^{s_1} and Q is a linear and bounded operator from Z^s to Z^s for all $s \in [s_3, s_1]$ and, considered as an operator in Z^2, coercive and self-adjoint.*

Proof. The required mapping properties of Q and R directly follow from their construction in Definition 7.8 in combination with Lemma 7.6 and Assumption 7.10, respectively. It remains to show that Q is coercive and self-adjoint. Since $\mathbb{B}$ is symmetric and coercive according to Assumption 7.1, it is sufficient to prove that the operator $B^\top \mathbb{C} B - T \colon Z^2 \to Z^2$ is symmetric and positive. To prove the symmetry, first observe that $B^\top \mathbb{C} B$ is symmetric by the symmetry of $\mathbb{C}$. The symmetry of $\mathbb{C}$ moreover implies that $\mathcal{G} \colon (U^2)^* \times (V^2)^* \to U^2 \times V^2$, i.e., the solution operator of (7.4), is self-adjoint. Therefore, the construction of T in Definition 7.8 implies for all $z_1, z_2 \in Z^2$ that

$$
\begin{aligned}
(Tz_1, z_2)_{Z^2} &= \langle -\operatorname{Div}_{(x,y)}(\mathbb{C}Bz_2), \mathcal{G}(-\operatorname{Div}_{(x,y)}(\mathbb{C}Bz_1)) \rangle \\
&= \langle \mathcal{G}(-\operatorname{Div}_{(x,y)}(\mathbb{C}Bz_2)), -\operatorname{Div}_{(x,y)}(\mathbb{C}Bz_1) \rangle = (z_1, Tz_2)_{Z^2}
\end{aligned}
$$

so that T is also symmetric. To show the positivity of $B^\top \mathbb{C} B - T$, let $z \in Z^2$ be arbitrary. To shorten the notation, we abbreviate $(u_z, v_z) := \mathcal{G}(-\operatorname{Div}_{(x,y)}(\mathbb{C}Bz))$. Then, by testing the equation for (u_z, v_z), i.e., (7.4) with $(\mathfrak{f}, \mathfrak{g}) = -\operatorname{Div}_{(x,y)}(\mathbb{C}Bz)$, with $(-u_z, -v_z)$, we arrive at

$$
\left(\mathbb{C}(Bz - \nabla^s_{(x,y)}(u_z, v_z)), -\nabla^s_{(x,y)}(u_z, v_z) \right)_{S^2} = 0.
$$

Since, by construction, $Tz = B^\top \mathbb{C} \nabla^s_{(x,y)}(u_z, v_z)$, the coercivity of $\mathbb{C}$ therefore implies

$$
\begin{aligned}
\left((B^\top \mathbb{C} B - T)z, z \right)_{Z^2} &= \left(\mathbb{C}(Bz - \nabla^s_{(x,y)}(u_z, v_z)), Bz \right)_{S^2} \\
&= \left(\mathbb{C}(Bz - \nabla^s_{(x,y)}(u_z, v_z)), Bz - \nabla^s_{(x,y)}(u_z, v_z) \right)_{S^2} \geq 0.
\end{aligned}
$$

As z was arbitrary, this proves the positivity. □

We point out that the whole analysis in sections 3 and 4 is carried out in the Hilbert space $\mathcal{H} = Z^2$. Therefore, for the mere existence and approximation results from these two sections, the critical regularity condition in Assumption 7.10 is *automatically fulfilled* by setting $s_1 = s_2 = s_3 = 2$ (so that $\mathcal{Y} = \mathcal{Z} = \mathcal{W} = \mathcal{H} = Z^2$). Note that, in this case, the Lax-Milgram lemma guarantees the assertion of Assumption 7.10 without any further regularity assumptions, see the proof of Lemma 7.6. The additional crucial regularity assumption only comes into play, when first- and second-order optimality conditions are investigated, see Remark 2.1. In subsection 7.4 below, we will elaborate in detail, where the critical Assumption 7.10 is needed to ensure the required differentiability properties of the regularized control-to-state map for the example of a specific yield condition.

We collect our findings so far in the following

THEOREM 7.12 (Homogenized plasticity as abstract evolution VI). *Under the Assumptions 7.1 and 7.3, the system of homogenized elastoplasticty in its weak form in* (7.3) *is equivalent to an abstract operator differential equation of the form*

$$(7.11) \qquad\qquad \dot{z} \in A(R\ell - Qz), \qquad z(0) = z_0,$$

with Q and R as defined in (7.5) *and* (7.6) *in the following sense: If (u, v, z, Σ) solves* (7.3), *then z is a solution* (7.11), *and vice versa, if z solves* (7.11), *then z together with (u, v) and Σ as defined in* (7.7) *and* (7.8), *respectively, form a solution of* (7.3).

In addition, Q and R satisfy the standing assumptions from section 2 (provided that the function spaces are chosen according to Definition 7.9). Therefore, the existence and approximation results of Sections 3 and 4 hold for (7.11), *in particular:*

- *For every $\ell \in \mathcal{U}(z_0, D(A))$, there is a unique solution (u, v, z, Σ) of the weak system of homogenized plasticity in* (7.3), *cf. Theorem 3.3.*
- *Optimal control problems governed by the weak system of homogenized elastoplasticity admit globally optimal solutions, provided that the standing assumptions on the objective are fulfilled, cf. Theorem 4.2.*
- *The approximation results of Theorem 4.5 and Corollary 4.6 apply in case of homogenized elastoplasticity.*

Remark 7.13. Existence and uniqueness of solutions to (7.3) was already established in [33].

Remark 7.14. The example for objective functionals mentioned above, i.e.,

$$(7.12) \qquad \begin{aligned} &J(u, \Sigma, f, g) := \\ &\tfrac{\alpha}{2} \int_{\Omega} |u(T) - u_d|^2 \, dx + \tfrac{\beta}{2} \int_{\Omega} |(\pi\Sigma)(T) - \sigma_d|^2 \, dx + \tfrac{\gamma}{2} \|(f, g)\|_{H^1(0, T; \mathcal{X}_c)}^2 \end{aligned}$$

with $u_D \in L^2(\Omega; \mathbb{R}^d)$ and $\sigma_D \in L^2(\Omega; \mathbb{R}_{\text{sym}}^{d \times d})$ and $\alpha, \beta \geq 0$, $\gamma > 0$, satisfies the standing assumptions on the objective functional, as we will see in the following. In this case, the functional $\Psi \colon (z, \ell) \to \mathbb{R}$ in the general setting consists of the two integrals at end point T. Let us consider the first one containing the displacement u. Since the latter is given by the first component of the solution operator of (7.7), which maps $H^1(0, T; Z^2) \times H^1(0, T; \mathcal{X}_c)$ to $H^1(0, T; U^2) \hookrightarrow C([0, T]; L^2(\Omega; \mathbb{R}^d))$, this integral is well defined. (Note that the operators in (7.7) just act pointwise in time and the time regularity of z and ℓ carries over to u and v.) Clearly, this solution operator is linear and bounded. In case of the second integral involving Σ, one can argue completely analogously based on the solution operator of (7.8). Thus, Ψ is convex and continuous, hence weakly lower semicontinuous, and in addition, bounded from below by zero. Moreover, the purely control part of the objective is given by $\Phi(f, g) = \tfrac{\gamma}{2} \|(f, g)\|_{H^1(0, T; \mathcal{X}_c)}^2$ and therefore clearly weakly lower semicontinuous and coercive as required. Thus all standing assumptions are fulfilled as claimed. Of course, various other objective functionals are possible as well, such as tracking type objectives over the whole space-time-cylinder, but to keep the discussion concise, we just mention the example above.

7.3. Optimality System. In the following section, we establish necessary and sufficient optimality conditions for the optimal control of regularized homogenized elastoplasticity. To be more precise, we consider a single element of the sequence

of regularizations of the maximal monotone operator A from Assumption 7.3, which we again denote by A_s, and apply the general theory from section 5 and section 6. In view of the norm gap needed for the differentiability of A_s, we will consider A_s in different domains and ranges (denoted by the same symbol) and assume that A_s maps $\mathcal{Y} = Z^{s_1}$ to itself. Accordingly, we treat the regularized version of the state equation (with A_s instead of A) in the same manner, i.e., with integrability index s_1 instead of 2, see (7.13) below. To keep the discussion concise, we moreover assume in all what follows that $s_1 \geq 2$ is such that $p = r = 2$ satisfy the conditions in (7.10). For $d = \dim(\Omega) = 3$, this implies $s_1 < 3$ and, in case without boundary control, i.e., $g \equiv 0$, $s_1 < 6$ is sufficient. This will become important in the discussion of second-order sufficient conditions, as we will see below. Motivated by (7.12), we consider an optimal control problem of the form

$$(7.13) \quad \begin{cases} \min & J(u, \Sigma, f, g) := F_1(u, \Sigma) + F_2(u(T), \Sigma(T)) \\ & \qquad\qquad + \dfrac{\gamma}{2}\Big(\|\dot{f}\|^2_{L^2(0,T;L^2(\Omega;\mathbb{R}^d))} + \|\dot{g}\|^2_{L^2(0,T;L^2(\Gamma_N;\mathbb{R}^d))}\Big) \\ \text{s.t.} & (f, g) \in H^1(0, T; L^2(\Omega; \mathbb{R}^d) \times L^2(\Gamma_N; \mathbb{R}^d)), \\ & (u, v, z, \Sigma) \in H^1(0, T; U^{s_1} \times V^{s_1} \times Z^{s_1} \times S^{s_1}), \\ \text{and} & -\operatorname{Div}_{(x,y)} \Sigma = ((f, g), 0), \\ & \Sigma = \mathbb{C}\big(\nabla^s_{(x,y)}(u, v) - Bz\big), \\ & \dot{z} = A_s(B^\top \Sigma - \mathbb{B}z), \quad z(0) = z_0, \\ & \ell(0) = 0. \end{cases}$$

Since, in many applications, displacement and stress on the macro level are of special interest, especially at end time, we focus on objectives with this particular structure with continuously Fréchet-differentiable mappings

$$(7.14) \qquad F_1 \colon L^2(0, T; U^2) \times L^2(0, T; S^2) \to \mathbb{R}, \quad F_2 \colon U^2 \times S^2 \to \mathbb{R}.$$

In order to apply our general theory, we not only have to reduce the state system to an equation of the form (7.11), but also have to reduce the objective. For this purpose, let us denote the solution operators of (7.7) and (7.8) by $\mathfrak{u} \colon (\ell, z) \mapsto u$ and $\mathfrak{S} \colon (\ell, z) \mapsto \Sigma$. To shorten the notation, we will consider $\mathfrak{u}$ and $\mathfrak{S}$ with different domains and ranges, e.g. $\mathfrak{u} \colon (U^{s'})^* \times Z^s \to U^s$ and $\mathfrak{u} \colon L^2(0, T; (U^{s'})^*) \times L^2(0, T; Z^s) \to L^2(0, T; U^s)$ with $s \in [\bar{s}', \bar{s}]$ and analogously for $\mathfrak{S}$. Note again that the time regularity of z and ℓ directly carries over to the time regularity of u and Σ. Given these operators, we define

$$\begin{aligned} \Psi_1 &\colon L^2(0, T; Z^2) \times L^2(0, T; (U^2)^*) \to \mathbb{R}, & \Psi_1(z, \ell) &:= F_1(\mathfrak{u}(z, \ell), \mathfrak{S}(z, \ell)), \\ \Psi_2 &\colon Z^2 \times (U^2)^* \to \mathbb{R}, & \Psi_2(z, \ell) &:= F_2(\mathfrak{u}(z, \ell), \mathfrak{S}(z, \ell)), \end{aligned}$$

so that the objective in (7.13) becomes

$$(7.15) \qquad J(z, \ell) = \Psi_1(z, \ell) + \Psi_2(z(T), \ell(T)) + \tfrac{\gamma}{2}\|(\dot{f}, \dot{g})\|^2_{L^2(0,T;\mathcal{X}_c)},$$

i.e., exactly an objective of the form in (5.15) and (5.18), respectively. Since $\mathfrak{u}$ and $\mathfrak{S}$ are linear and bounded and F_1 and F_2 are assumed to be continuously Fréchet-differentiable, the chain rule implies the differentiability of Ψ_1 and Ψ_2 so that Assumption 5.1(i) is met, if we set $s_3 = 2$ so that $\mathcal{W} = Z^2$. To apply the results of section 5 in order to establish an optimality system for (7.13), we additionally need that A_s satisfies Assumption 5.1(ii), which is ensured by the following

ASSUMPTION 7.15. *We set $s_2 = s_3 = 2$, i.e., $\mathcal{W} = \mathcal{Z} = Z^2$, and assume that A_s fulfills Assumption 5.1(ii) with $\mathcal{Y} = Z^{s_1}$, $s_1 \geq 2$, i.e., in particular that A_s is Fréchet-differentiable from $\mathcal{Z}^{s_1}$ to $\mathcal{Z}^2$.*

In light of Lemma 7.6, Assumption 7.15 does not impose any restriction for practical realizations of A_s, as we will see in subsection 7.4 below. Given this assumption, Theorem 5.12 and Example 5.15 imply for a locally optimal solution $(\bar{\ell}) = (\bar{f}, \bar{g})$ with associated optimal internal variable $\bar{z}$:

$$(7.16a) \quad \dot{\bar{z}}(t) = A_s(R\bar{\ell} - Q\bar{z})(t), \quad \bar{z}(0) = z_0,$$

$$(7.16b) \quad \dot{\varphi}(t) = \left(QA'_s(R\bar{\ell} - Q\bar{z})^*\varphi\right)(t) + \frac{\partial \Psi_1}{\partial z}(\bar{z}, \bar{\ell})(t), \quad \varphi(T) = -\frac{\partial \Psi_2}{\partial z}(\bar{z}(T), \bar{\ell}(T)),$$

$$(7.16c) \quad \begin{cases} \gamma\, \partial_t^2 \bar{\ell}(t) + R^* A'_s\left(R\bar{\ell}(t) - Q\bar{z}(t)\right)^* \varphi(t) = \dfrac{\partial \Psi_1}{\partial \ell}(\bar{z}, \bar{\ell})(t), \\[2mm] \qquad\qquad \bar{\ell}(0) = 0, \quad \partial_t \bar{\ell}(T) = -\dfrac{\partial \Psi_2}{\partial \ell}(\bar{z}(T), \bar{\ell}(T)). \end{cases}$$

Then, owing to the precise structure of R, Q, Ψ_1, and Ψ_2, this leads us to the following

THEOREM 7.16 (KKT-system for optimal control of homogenized plasticity). *Let Assumption 7.1 be satisfied and assume that A_s fulfills Assumption 7.15. Suppose moreover that the regularity condition in Assumption 7.10 is satisfied, i.e., $\bar{s} \geq s_1$. Then, if $(\bar{f}, \bar{g}) \in H^1(0,T;L^2(\Omega)) \times H^1(0,T;L^2(\Gamma_N))$ is locally optimal for (7.13) with associated state $(\bar{u}, \bar{v}, \bar{z}, \bar{\Sigma}) \in H^1(0,T;U^{s_1}) \times H^1(0,T;V^{s_1}) \times H^1(0,T;Z^{s_1}) \times H^1(0,T;S^{s_1})$, then there exists an adjoint state*

$$(w, q, \varphi, \Upsilon) \in H^1(0,T;U^2) \times H^1(0,T;V^2) \times H^1(0,T;Z^2) \times H^1(0,T;S^2),$$
$$(w_T, q_T, \Upsilon_T) \in U^2 \times V^2 \times S^2$$

such that the following optimality system is satisfied:

State equation:

$$(7.17a) \quad -\operatorname{Div}_{(x,y)} \bar{\Sigma} = ((\bar{f}, \bar{g}), 0),$$
$$(7.17b) \quad \bar{\Sigma} = \mathbb{C}\left(\nabla^s_{(x,y)}(\bar{u}, \bar{v}) - B\bar{z}\right),$$
$$(7.17c) \quad \dot{\bar{z}} = A_s(B^\top \bar{\Sigma} - \mathbb{B}\bar{z}), \quad \bar{z}(0) = z_0,$$

Adjoint equation:

$$(7.17d) \quad -\operatorname{Div}_{(x,y)} \Upsilon = \left(\tfrac{\partial}{\partial u} F_1(\bar{u}, \bar{\Sigma}), 0\right) - \operatorname{Div}_{(x,y)}\left(\mathbb{C}BA'_s(B^\top \bar{\Sigma} - \mathbb{B}\bar{z})^*\varphi\right),$$
$$(7.17e) \quad \Upsilon = \mathbb{C}\left(\nabla^s_{(x,y)}(w, q) - \tfrac{\partial}{\partial \Sigma} F_1(\bar{u}, \bar{\Sigma})\right),$$
$$(7.17f) \quad \dot{\varphi} = (B^\top \mathbb{C}B + \mathbb{B})A'_s(B^\top \bar{\Sigma} - \mathbb{B}\bar{z})^*\varphi + B^\top \Upsilon, \quad \varphi(T) = -B^\top \Upsilon_T,$$
$$(7.17g) \quad -\operatorname{Div}_{(x,y)} \Upsilon_T = \left(\tfrac{\partial}{\partial u} F_2(\bar{u}(T), \bar{\Sigma}(T)), 0\right),$$
$$(7.17h) \quad \Upsilon_T = \mathbb{C}\left(\nabla^s_{(x,y)}(w_T, q_T) - \tfrac{\partial}{\partial \Sigma} F_2(\bar{u}(T), \bar{\Sigma}(T))\right),$$

Gradient equation:

$$(7.17i) \quad \gamma\, \partial_t^2 \bar{f} + w = 0, \quad \bar{f}(0) = 0, \quad \gamma\, \partial_t \bar{f}(T) + w_T = 0,$$
$$(7.17j) \quad \gamma\, \partial_t^2 \bar{g} + w = 0, \quad \bar{g}(0) = 0, \quad \gamma\, \partial_t \bar{g}(T) + w_T = 0.$$

Remark 7.17. A passage to the limit w.r.t. the regularization in order to obtain an optimality system for the original optimal control problem involving the maximal monotone operator A would of course be of particular interest. The results of [43] however indicate that the optimality conditions obtained in this way are in general rather weak. In [43], an optimal control problem governed by quasi-static elastoplasticity (without homogenization) is considered, which provides substantial similarities to (7.13). This system could also be treated by means of a reduction to the internal variable similar to our procedure for (7.1). In [43] however, a time discretization followed by a regularization was employed for the derivation of first-order optimality conditions. The reason for the comparatively weak optimality conditions obtained for the original (non-smooth) problem is the poor regularity of the dual variables in the limit, in particular the adjoint state. We expect a similar behavior in case of (7.17), when the regularization is driven to zero. This however is subject to future research.

Next, we turn to second-order sufficient optimality conditions. Now, Ψ_1, Ψ_2, and A_s have to fulfill Assumption 6.1. For this purpose, we require the following

ASSUMPTION 7.18. *The mappings F_1 and F_2 from (7.14) are twice Fréchet-differentiable. Moreover, A_s satisfies Assumption 5.1 (ii) and Assumption 6.1(ii) and (iii) with $\mathcal{W} = \mathcal{H} = Z^2$ (i.e., $s_3 = 2$), $\mathcal{Z} = Z^{s_2}$, $s_2 \geq 2$, and $\mathcal{Y} = Z^{s_1}$, $s_1 \geq s_2$.*

In contrast to Assumption 7.15, this assumption is very restrictive. If we assume that A_s arises as a Nemyzki operator from a nonlinear function $A_s \colon \mathbb{V} \to \mathbb{V}$ (for simplicity denoted by the same symbol), then the last condition in Assumption 6.1, i.e.,

$$(7.18) \qquad \|A_s''(z)[z_1, z_2]\|_{\mathcal{W}} \leq C \, \|z_1\|_{Z^{s_2}} \|z_2\|_{Z^{s_2}} \quad \forall \, z \in Z^{s_1}, \; z_1, z_2 \in Z^{s_2}$$

with $\mathcal{W} = Z^2$, may only hold—even in case $A_s''(z) \in L^\infty(\Omega \times Y; \mathcal{L}(\mathbb{V}, \mathbb{V}))$—provided that $s_2 \geq 2 \, s_3 = 4$. In order to have that A_s is Fréchet-differentiable from Z^{s_1} to Z^{s_2}, we therefore need $s_1 > 4$ such that Assumption 7.10 indeed becomes very restrictive, see Remark 7.7 and Remark 7.28 below. Moreover, as described at the beginning of this subsection, if one sets $r = 2$, i.e., considers to boundary loads in $H^1(0, T; L^2(\Gamma_N; \mathbb{R}^d))$, then, in view of (7.10), $s_1 < 3$ has to hold (at least in three spatial dimensions) so that our second-order analysis cannot be applied in case of boundary controls (at least if $r = 2$ and $d = 3$). Therefore, we restrict to volume forces, i.e., controls in $H^1(0, T; L^2(\Omega; \mathbb{R}^d))$ in what follows.

Then, given all these assumptions, we can apply Theorem 6.7 and Corollary 6.9 to (7.13) to obtain the following

THEOREM 7.19 (Second-order sufficient conditions for optimal control of homogenized plasticity). *Let Assumptions 7.1 and 7.18 hold and suppose that $\bar{s} \geq s_1$ so that Assumption 7.10 is fulfilled. Assume moreover that $\overline{f} \in H^1(0, T; L^2(\Omega; \mathbb{R}^d))$ together with its associated state $(\overline{u}, \overline{v}, \overline{z}, \overline{\Sigma})$ and an adjoint state $(w, q, \varphi, \Upsilon, w_T, q_T, \Upsilon_T)$ satisfies the optimality system (7.17a)–(7.17i) (without (7.17j) because boundary controls are omitted) and, in addition, that there exists $\delta > 0$ such that*

$$\frac{\partial^2}{\partial u^2} F_1(\overline{u}, \overline{\Sigma}) \eta_u^2 + \frac{\partial^2}{\partial \Sigma^2} F_1(\overline{u}, \overline{\Sigma}) \eta_\Sigma^2$$

$$+ \frac{\partial^2}{\partial u^2} F_2(\overline{u}(T), \overline{\Sigma}(T)) \eta_u(T)^2 + \frac{\partial^2}{\partial \Sigma^2} F_2(\overline{u}(T), \overline{\Sigma}(T)) \eta_\Sigma(T)^2 + \gamma \, \|\dot{h}\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^d))}^2$$

$$+ \int_0^T \int_\Omega \left(\varphi, A_s''(B^\top \overline{\Sigma} - \mathbb{B}\overline{z})(B^\top \eta_\Sigma - \mathbb{B}\eta_z)^2 \right)_{\mathbb{V}} dx \, dt \geq \delta \, \|h\|_{H^1(0,T;L^2(\Omega;\mathbb{R}^d))}^2$$

holds for all $h \in H^1(0, T; L^2(\Omega; \mathbb{R}^d))$ with $h(0) = 0$, where $(\eta_u, \eta_v, \eta_z, \eta_\Sigma) \in H^1(0, T; U^2 \times$

$V^2 \times Z^2 \times S^2)$ *is the solution of the linearized state system associated with* h, *i.e.,*

$$
\begin{aligned}
-\operatorname{Div}_{(x,y)} \eta_\Sigma &= (h, 0), \\
\eta_\Sigma &= \mathbb{C}\big(\nabla^s_{(x,y)}(\eta_u, \eta_v) - B\eta_z\big), \\
\dot\eta_z &= A'_s(B^\top \overline\Sigma - \mathbb{B}\overline z)(B^\top \eta_\Sigma - \mathbb{B}\eta_z), \quad \eta_z(0) = 0.
\end{aligned}
$$

Then $\overline f$ *is a strict local minimizer of* (7.13) *and satisfies the quadratic growth condition* (6.3).

Remark 7.20. As indicated above, Assumption 7.18 in combination with Assumption 7.10 is very restrictive. One can however weaken these assumptions, if the objective provides certain properties. Let us for instance consider an objective of the form

$$
(7.19) \qquad\qquad J(u, f) := F_3(u) + \tfrac{\gamma}{2} \|\dot f\|^2_{L^2(0,T;L^2(\Omega;\mathbb{R}^d))}
$$

with a twice Fréchet-differentiable functional

$$
F_3 \colon H^1(0, T; L^2(\Omega; \mathbb{R}^d)) \to \mathbb{R}.
$$

In this case, it is sufficient to choose s_3 such that $\mathfrak{u} \colon (z, \ell) \mapsto u$ maps $\mathcal{W} \times \mathcal{X}_c$ with $\mathcal{W} = Z^{s_3}$ to $W^{1,p}(\Omega; \mathbb{R}^d) \hookrightarrow L^2(\Omega; \mathbb{R}^d)$, i.e., $p \geq 6/5$ for $d = 3$. Moreover, as seen above, in order to have that the Nemyzki operator A_s fulfills (7.18), we need $s_1 > 2s_3$. Thus, we are tempted to choose s_3 as small as possible. However, the crucial, limiting condition is the regularity assumption in Assumption 7.10 involving the conjugate exponent, i.e., $\bar s \geq \max\{s_1, s'_3\}$, and this leads to the following equilibration of s_1 and s_3 in the case $d = 3$: $s_1 > 3$, $s_3 = 3/2$ (such that $s'_3 = 3$). Then, in view of (7.7), $\mathfrak{u}$ maps Z^{s_3} to $W^{1,s_3}(\Omega; \mathbb{R}^d)$, which is continuously embedded in $L^2(\Omega; \mathbb{R}^d)$ for $d \leq 3$ as desired. In the next subsection, we will present an example for a Nemyzki operator A_s fulfilling all assumptions for s_1 arbitrarily close to 3. But nonetheless, $\bar s > 3$ in Assumption 7.10 is still a rather restrictive assumption and will not be satisfied in general (depending on the regularity of $\mathbb{C}$ and the boundary of Ω). This shows that the second-order analysis for problems of type (P) (resp. its regularized counterparts, to be precise) is in general a delicate issue, mainly due to the quasi-linear structure of the state equation.

7.4. A Concrete Flow Rule. In the following, we will discuss a concrete realization of the maximal monotone operator A and its regularization, respectively, in order to demonstrate how the Assumptions 7.3, 7.15, and 7.18 can be satisfied in practice and how restrictive they are, in particular Assumption 7.18.

We consider the case of linear kinematic hardening with von Mises yield condition, cf. [25] for a detailed description of this model. In this case, the finite dimensional space for the internal variable is given by $V = \mathbb{R}^{d \times d}_{\mathrm{sym}}$ and $B \colon \mathbb{R}^{d \times d}_{\mathrm{sym}} \to \mathbb{R}^{d \times d}_{\mathrm{sym}}$ is the identity so that $Z^2 = S^2$. Moreover, A is the convex subdifferential of the indicator functional $I_\mathcal{K}$ of following set of admissible stresses

$$
\mathcal{K} := \{\tau \in S^s \colon |\tau^D(x, y)| \leq \sigma_0 \quad \text{f.a.a. } (x, y) \in \Omega \times Y\},
$$

where $\tau^D := \tau - \frac{1}{d} \operatorname{tr}(\tau) I$ is the deviator of $\tau \in \mathbb{R}^{d \times d}_{\mathrm{sym}}$ and σ_0 denotes the initial uniaxial yield stress, a given material parameter. The domain of $A = \partial I_\mathcal{K}$ is trivially $\mathcal{K}$, which is closed and convex. Furthermore, it is easily seen that $A^0(\tau) = 0$ for all

$\tau \in D(A) = \mathcal{K}$ so that all of our standing assumptions are fulfilled in this case. Note moreover that A satisfies Assumption 3.9 so that the second approximation result on the convergence of the optimal states in Corollary 4.6 applies in this case. For the Yosida-approximation of $\partial I_{\mathcal{K}}$, one obtains

$$(7.20) \qquad A_\lambda = \frac{1}{\lambda}(I - \pi_{\mathcal{K}}) = \frac{1}{\lambda} \max\left\{0, 1 - \frac{\sigma_0}{|\tau^D|}\right\}\tau^D,$$

cf. [26], where $\pi_{\mathcal{K}}$ denotes the projection on $\mathcal{K}$ in Z^2, i.e., $\pi_{\mathcal{K}}(\sigma) := \arg\min_{\tau \in \mathcal{K}} \|\tau - \sigma\|_{Z^2}^2$. Herein, with a slight abuse of notation, we denote the Nemyzki operator in $L^\infty(\Omega \times Y)$ associated with the pointwise maximum, i.e., $\mathbb{R} \ni r \mapsto \max\{0, r\} \in \mathbb{R}$, by the same symbol. In addition, we set $\max\{0, 1 - \sigma_0/r\} := 0$, if $r = 0$.

The precise form of A_λ in (7.20) immediately suggests the following regularization of the Yosida approximation:

$$A_{\lambda,\epsilon}(\tau) := \frac{1}{\lambda} \max_\epsilon \left(1 - \frac{\sigma_0}{|\tau^D|}\right)\tau^D,$$

where $\max_\epsilon$ is a regularized version of the max-function, depending on a regularization parameter $\epsilon > 0$. To be more precise, $\max_\epsilon : L^\infty(\Omega \times Y) \to L^\infty(\Omega \times Y)$ is the Nemyzki operator associated with a real valued function (again denoted by the same symbol) with the following properties:

1. For every $\epsilon > 0$, there holds $\max_\epsilon \in C^2(\mathbb{R})$,
2. $\max_\epsilon(r) = \max\{0, r\}$ for $|r| \geq \frac{1}{2}$ and every $0 < \epsilon \leq 1/2$,
3. $|\max_\epsilon(r) - \max\{0, r\}| \leq \mathcal{O}(\epsilon)$ for all $r \in \mathbb{R}$.

EXAMPLE 7.21. *An example for a function satisfying the above conditions is*

$$\max_\epsilon(r) := \begin{cases} \max\{0, r\}, & |r| \geq \epsilon, \\ \frac{1}{16\epsilon^3}(r + \epsilon)^3(3\epsilon - r), & |r| < \epsilon. \end{cases}$$

One easily verifies that $\max_\varepsilon$ is twice continuously differentiable and that $|\max_\varepsilon(r) = \max\{0, r\}| \leq \frac{3}{16}\epsilon$.

LEMMA 7.22. *Let $\{\lambda_n\}_{n \in \mathbb{N}} \subset \mathbb{R}^+$ and $\{\epsilon_n\}_{n \in \mathbb{N}} \subset \mathbb{R}^+$ be null sequences satisfying*

$$(7.21) \qquad \epsilon_n = o\left(\lambda_n^2 \exp\left(-\frac{T\|Q\|}{\lambda_n}\right)\right),$$

and define $A_n := A_{\lambda_n, \epsilon_n} : Z^2 \to Z^2$. Then, the sequence $\{A_n\}_{n \in \mathbb{N}}$ satisfies Assumption 4.3. Thus, Assumption 7.3 is fulfilled in this case so that the approximation results from Theorem 7.12 apply.

Proof. Based on our assumptions on $\max_\epsilon$, we find for every $\tau \in Z^2$ and all $n \in \mathbb{N}$ such that $\epsilon_n \leq 1/2$ that

$$\|A_n(\tau) - A_{\lambda_n}(\tau)\|_{Z^2}^2$$
$$= \frac{1}{\lambda_n^2} \int_\Omega \left| \max_{\epsilon_n}\left(1 - \frac{\sigma_0}{|\tau^D|}\right) - \max\left\{0, 1 - \frac{\sigma_0}{|\tau^D|}\right\} \right|^2 |\tau^D|^2 \, dx \leq C \frac{\epsilon_n^2}{\lambda_n^2}.$$

The coupling of ϵ_n and λ_n in (7.21) then implies that (4.1) is fulfilled. $\quad\square$

Remark 7.23. We point out that we neither claim that the coupling of λ and ϵ in (7.21) is optimal nor that our regularization approach is the most efficient one for this specific flow rule.

Let us now fix $n \in \mathbb{N}$ and set $\lambda_s := \lambda_n$, $\max_s := \max_{\epsilon_n}$, and $A_s := A_n$. As before, we will denote the Nemyzki operators generated by $\max_s$ and its derivatives by the same symbol. The following result can be proven as in [27, Prop. 2.11] by using [22, Theorem 7]:

LEMMA 7.24. *Let $s > r \geq 1$ be arbitrary. Then A_s is continuously Fréchet-differentiable from Z^s to Z^r and its directional derivative at $\tau \in Z^s$ in direction $h \in Z^r$ is given by*

$$A_s'(\tau)h = \frac{1}{\lambda_s} \max_s' \left(1 - \frac{\sigma_0}{|\tau^D|}\right) \frac{\sigma_0}{|\tau^D|^3} (\tau^D : h^D)\tau^D + \frac{1}{\lambda_s} \max_s \left(1 - \frac{\sigma_0}{|\tau^D|}\right) h^D.$$

As a consequence of this result, we obtain the following

COROLLARY 7.25. *Assumption 7.15 is fulfilled by setting $s_1 := \bar{s}$, where $\bar{s} > 2$ is the exponent, whose existence is guaranteed by Lemma 7.6. Thus, in case of linear kinematic hardening with von Mises yield condition and the regularization introduced above, the optimality condition in Theorem 7.16 are indeed necessary for local optimality without any further assumptions (except our standing Assumption 7.1).*

Proof. We have to verify Assumption 5.1(ii) for $\mathcal{Y} = Z^{\bar{s}}$ and $\mathcal{Z} = \mathcal{H} = Z^2$. The Fréchet-differentiability from $Z^{\bar{s}}$ to Z^2 follows from Lemma 7.24. Moreover, the (global) Lipschitz continuity of A_s in $Z^{\bar{s}}$ can be deduced from the smoothness of $\max_s$ and the condition $\max_s(r) = \max\{0, r\}$ for all $|r| \geq 1/2$. The latter condition also guarantees that $\|A_s'(y)h\|_{Z^2} \leq C \|h\|_{Z^2}$ for all $y \in Z^{\bar{s}}$ and all $h \in Z^2$. This completes the proof. □

Furthermore, following the lines of [41] and [22, Theorem 9], one proves the following

LEMMA 7.26. *For every $s > 2$ and $1 \geq r < s/2$, A_s is twice Fréchet-differentiable and its second derivative at $\tau \in Z^s$ in directions $h_1, h_2 \in Z^r$ is given by*

$$
\begin{aligned}
&A_s''(\tau)[h_1, h_2] \\
&= \frac{\gamma}{\lambda_s |\tau^D|^3} \Big[\max_s'' \left(1 - \frac{\gamma}{|\tau^D|}\right) \frac{\gamma}{|\tau^D|^3} (\tau^D : h_1^D)(\tau^D : h_2^D)\tau^D \\
&\qquad + \max_s' \left(1 - \frac{\gamma}{|\tau^D|}\right) \left(-\frac{3}{|\tau^D|^2}(\tau^D : h_1^D)(\tau^D : h_2^D)\tau^D + (h_1^D : h_2^D)\tau^D \right. \\
&\qquad\qquad\qquad\qquad\qquad\qquad \left. + (\tau^D : h_1^D)h_2^D + (\tau^D : h_2^D)h_1^D \right) \Big].
\end{aligned}
$$

COROLLARY 7.27. *The conditions on A_s in Assumption 7.18 are satisfied, if the index $\bar{s}$ from Lemma 7.6 fulfills $\bar{s} > 4$.*

Proof. If we set $s_1 = \bar{s} > 4$, $s_2 \in]4, s_1[$, and $s_3 = 2$, then Lemma 7.26 implies the differentiability conditions in Assumption 6.1(iii) with $\mathcal{Y} = Z^{s_1}$, $\mathcal{Z} = Z^{s_2}$, and $\mathcal{W} = Z^{s_3}$. The Lipschitz continuity of A_s' from Z^{s_1} to $L(Z^{s_2})$ as well as the estimate $\|A_s'(y)w\|_{Z^2} \leq C \|w\|_{Z^2}$ follow from the condition $\max_s(r) = \max\{0, r\}$ for all $|r| \geq 1/2$. This condition also ensures that $\|A_s''(y)[z_1, z_2]\|_{Z^2} \leq C \|z_1\|_{Z^4}\|z_2\|_{Z^4}$, which in turn implies the last condition in Assumption 6.1(iii) thanks to $s_2 > 4$. □

Remark 7.28. As indicated in Remark 7.7, the assumption $\bar{s} > 4$ is very restrictive. However, if $\mathcal{W} = Z^2$, then any Nemyzki operator is only twice Fréchet-differentiable from $\mathcal{Y}$ to $\mathcal{W}$, if $\mathcal{Y} = Z^s$ with $s > 4$, see e.g. [22] and the references therein. In light of this observation, the above regularization is rather well-behaved. As explained in Remark 7.20, one can reduce the value of s_3, if only the L^2-norm

of the displacement appears in the objective. However, one still needs $\bar{s} > 3$ in this case, which is not guaranteed by Lemma 7.6 in general. But again, one can show that any $\bar{s} > 3$ is sufficient for our concrete example, no matter how close it is to 3. This concrete realization of A_s thus allows for the weakest possible regularity assumptions.

Appendix A. Second Derivative of the Solution Operator.

Before we are in the position to show that $\mathcal{S}$ is twice Fréchet-differentiable, we need the following result on the Lipschitz continuity of the first derivative, which is also needed in the proof of Lemma 6.4.

LEMMA A.1. *Assume that Assumption* 5.1*(ii) and Assumption* 6.1*(ii) are fulfilled. Then* $\mathcal{S}'_s$ *is Lipschitz continuous from* $H^1(0,T;\mathcal{X})$ *to* $L(H^1(0,T;\mathcal{X}); H^1(0,T;\mathcal{Z}))$.

Proof of Proposition 6.3. Let $\ell_1, \ell_2, h \in H^1(0,T;\mathcal{X})$ be arbitrary and abbreviate

$$z_i := \mathcal{S}_s(\ell_i), \quad \eta_i := \mathcal{S}'_s(\ell_i)h, \quad \text{and} \quad y_i := R\ell_i - Qz_i, \quad i = 1,2.$$

Using the first Lipschitz-assumption in Assumption 6.1(ii), we deduce for almost all $t \in [0,T]$ that

$$\|\dot{\eta}_1(t) - \dot{\eta}_2(t)\|_{\mathcal{Z}}$$
$$= \|\left(A'_s(y_1(t)) - A'_s(y_2(t))\right)(Rh(t) - Q\eta_1(t)) + A'_s(y_2(t))Q(\eta_1(t) - \eta_2(t))\|_{\mathcal{Z}}$$
$$\leq C\left(\|y_1(t) - y_2(t)\|_Y \|Rh(t) - Q\eta_1(t)\|_{\mathcal{Z}} + \|\eta_1(t) - \eta_2(t)\rangle\|_{\mathcal{Z}}\right).$$

Gronwall's inequality and the definition of y_1 and y_2 then yield

$$\|\eta_1 - \eta_2\|_{H^1(0,T;\mathcal{Z})} \leq C\|R(\ell_1 - \ell_2) - Q(z_1 - z_2)\|_{L^2(0,T;Y)}\|Rh - Q\eta_1\|_{H^1(0,T;\mathcal{Z})}$$
$$\leq C\|\ell_1 - \ell_2\|_{L^2(0,T;\mathcal{X})}\|h\|_{H^1(0,T;\mathcal{X})},$$

where we used Lemma 5.3 and the estimate in Theorem 5.5.   □

Now, we are ready to prove that the solution operator is twice Fréchet-differentiable. The proof is based on an estimate of the remainder term and thus similar to the one of Theorem 5.5.

Proof. The proof is similar to the one of Theorem 5.5. Let $\ell, h_1, h_2 \in H^1(0,T;\mathcal{X})$ be arbitrary and define $z := \mathcal{S}_s(\ell)$, $z_1 := \mathcal{S}_s(\ell + h_1)$, $\eta_i := \mathcal{S}'_s(\ell)h_i \in H^1(0,T;\mathcal{Z})$, $i \in \{1,2\}$, and $\eta_{1,2} := \mathcal{S}'_s(\ell + h_1)h_2$.

We first address the existence of solutions to (6.1). We argue similarly to Lemma 5.4 and set

$$w: [0,T] \to \mathcal{W}, \quad t \mapsto A''_s(R\ell(t) - Qz(t))[Rh_1(t) - Q\eta_1(t), Rh_2(t) - Q\eta_2(t)].$$

From the estimate in Assumption 6.1(iii) it follows that

$$\|w(t)\|_{\mathcal{W}} \leq C\|Rh_1(t) - Q\eta_1(t)\|_{\mathcal{Z}}\|Rh_2(t) - Q\eta_2(t)\|_{\mathcal{Z}},$$

and, since the limit of Bochner measurable functions is Bochner measurable, we obtain $w \in L^2(0,T;\mathcal{W})$. Since $A'_s(y)$ is assumed to be bounded in $\mathcal{W}$ by 6.1(ii), we can now follow the proof of Lemma 5.4 (with $\mathcal{W}$ instead of $\mathcal{Z}$) to deduce the existence of a unique solution $\xi \in H^1(0,T;\mathcal{W})$ of (6.1). The (bi-)linearity of the associated solution operator w.r.t. h_1 and h_2 is straightforward to see. For its continuity, we calculate

$$\|\dot{\xi}(t)\|_{\mathcal{W}} \leq C\|Rh_1(t) - Q\eta_1(t)\|_{\mathcal{Z}}\|Rh_2(t) - Q\eta_2(t)\|_{\mathcal{Z}} + C\|\xi(t)\|_{\mathcal{W}}$$

so that Gronwall's inequality and the estimate in Theorem 5.5 give

$$\|\xi\|_{H^1(0,T;\mathcal{W})} \le C \|h_1\|_{H^1(0,T;\mathcal{X})} \|h_2\|_{H^1(0,T;\mathcal{X})}.$$

This shows also (6.2) (after having proved that $\xi = \mathcal{S}_s''(\ell)[h_1, h_2]$).

It only remains to prove the remainder term property. To this end, we define

$$y := R\ell - Qz, \quad \zeta := Rh_1 - Q(z_1 - z).$$

Then, the equations for $\eta_{1,2}$, η_2, and ξ lead to

$$
\begin{aligned}
\dot\eta_{1,2} - \dot\eta_2 - \dot\xi &= \big(A_s'(y+\zeta) - A_s'(y)\big)(Rh_2 - Q\eta_{1,2}) \\
&\quad - A_s''(y)[Rh_1 - Q\eta_1, Rh_2 - Q\eta_2] - A_s'(y)Q(\eta_{1,2} - \eta_2 - \xi) \\
&= A_s''(y)[\zeta, Rh_2 - Q\eta_{1,2}] + r_2(y;\zeta)(Rh_2 - Q\eta_{1,2}) \\
&\quad - A_s''(y)[Rh_1 - Q\eta_1, Rh_2 - Q\eta_2] - A_s'(y)Q(\eta_{1,2} - \eta_2 - \xi) \\
&= A_s''(y)[\zeta, Q(\eta_2 - \eta_{1,2})] - A_s''(y)[Q(z_1 - z - \eta_1), Rh_2 - Q\eta_2] \\
&\quad + r_2(y;\zeta)(Rh_2 - Q\eta_{1,2}) - A_s'(y)Q(\eta_{1,2} - \eta_2 - \xi),
\end{aligned}
$$

where $r_2(y;\zeta) := A_s'(y+\zeta) - A_s'(y) - A_s''(y)\zeta \in L^2(0,T;L(\mathcal{Z};\mathcal{W}))$ denotes the corresponding remainder term. The estimate in Assumption 6.1(iii) thus implies

$$
\begin{aligned}
&\|\dot\eta_{1,2}(t) - \dot\eta_2(t) - \dot\xi(t)\|_{\mathcal{W}} \\
&\le C\big(\|\zeta(t)\|_{\mathcal{Z}}\|\eta_2(t) - \eta_{1,2}(t)\|_{\mathcal{Z}} + \|z_1(t) - z(t) - \eta_1(t)\|_{\mathcal{Z}}\|Rh_2(t) - Q\eta_2(t)\|_{\mathcal{Z}} \\
&\quad + \|r_2(y(t),\zeta(t))\|_{L(\mathcal{Z};\mathcal{W})}\|Rh_2(t) - Q\eta_{1,2}(t)\|_{\mathcal{Z}} + \|\eta_{1,2}(t) - \eta_2(t) - \xi(t)\|_{\mathcal{W}}\big)
\end{aligned}
$$

for almost all $t \in [0,T]$ such that Gronwall's inequality yields

$$
\begin{aligned}
&\|\eta_{1,2} - \eta_2 - \xi\|_{H^1(0,T;\mathcal{W})} \\
&\le C\big(\|Rh_1 - Q(z_1 - z)\|_{L^\infty(0,T;\mathcal{Z})}\|\eta_2 - \eta_{1,2}\|_{L^2(0,T;\mathcal{Z})} \\
&\quad + \|z_1 - z - \eta_1\|_{L^\infty(0,T;\mathcal{Z})}\|Rh_2 - Q\eta_2\|_{L^2(0,T;\mathcal{Z})} \\
&\quad + \|r_2(y;\zeta)\|_{L^2(0,T;L(\mathcal{Z};\mathcal{W}))}\|Rh_2 - Q\eta_{1,2}\|_{H^1(0,T;\mathcal{Z})}\big) \\
&\le C\|h_2\|_{H^1(0,T;\mathcal{X})}\big(\|h_1\|_{H^1(0,T;\mathcal{X})}^2 + \|z_1 - z - \eta_1\|_{H^1(0,T;\mathcal{Z})} + \|r_2(y;\zeta)\|_{L^2(0,T;L(\mathcal{Z};\mathcal{W}))}\big),
\end{aligned}
$$

where we used Lemma 5.3, Lemma A.1 and the estimate in Theorem 5.5. Denoting the solution operator of (6.1) already by $\mathcal{S}_s''(\ell)[h_1, h_2]$, we have thus shown

$$
\begin{aligned}
&\|\mathcal{S}_s'(\ell + h_1) - \mathcal{S}_s'(\ell) - \mathcal{S}_s''(\ell)h_1\|_{L(H^1(0,T;\mathcal{X});H^1(0,T;\mathcal{W}))} \\
&\qquad \le C\big(\|h_1\|_{H^1(0,T;\mathcal{X})}^2 + \|\mathcal{S}_s(\ell + h_1) - \mathcal{S}_s(\ell) - \mathcal{S}_s'(\ell)h_1\|_{H^1(0,T;\mathcal{Z})} \\
&\qquad\qquad\qquad\qquad\qquad\qquad + \|r_2(y;\zeta)\|_{L^2(0,T;L(\mathcal{Z};\mathcal{W}))}\big).
\end{aligned}
$$

Therefore, thanks to the Fréchet-differentiability of $\mathcal{S}_s \colon H^1(0,T;\mathcal{X}) \to H^1(0,T;\mathcal{Z})$, it only remains to show that

$$(A.1) \qquad \frac{\|r_2(y;\zeta)\|_{L^2(0,T;L(\mathcal{Z};\mathcal{W}))}}{\|h_1\|_{H^1(0,T;\mathcal{X})}} \to 0,$$

as $0 \ne h_1 \to 0$ in $H^1(0,T;\mathcal{X})$. To this end, we note that the embedding $H^1(0,T;\mathcal{Y}) \hookrightarrow C([0,T];\mathcal{Y})$ and Lemma 5.3 yield for all $t \in [0,T]$

$$(A.2) \qquad \frac{\|\zeta(t)\|_{\mathcal{Y}}}{\|h_1\|_{H^1(0,T;\mathcal{X})}} \le C \frac{\|\zeta\|_{H^1(0,T;\mathcal{Y})}}{\|h_1\|_{H^1(0,T;\mathcal{X})}} = C \frac{\|Rh_1 - Q(z_1 - z)\|_{H^1(0,T;\mathcal{Y})}}{\|h_1\|_{H^1(0,T;\mathcal{X})}} \le C$$

Hence, thanks to the Fréchet-differentiability of $A_s' \colon \mathcal{Y} \to L(\mathcal{Z};\mathcal{W})$, we have for almost all $t \in [0,T]$

$$\frac{\|r_2(y;\zeta)(t)\|_{L(\mathcal{Z};\mathcal{W})}}{\|h_1\|_{H^1(0,T;\mathcal{X})}} \leq C \frac{\|r_2(y;\zeta)(t)\|_{L(\mathcal{Z};\mathcal{W})}}{\|\zeta(t)\|_{\mathcal{Y}}} \to 0$$

as $0 \neq h_1 \to 0$ in $H^1(0,T;\mathcal{X})$. Furthermore, using the Lipschitz continuity of $A_s' \colon \mathcal{Y} \to L(\mathcal{Z};\mathcal{Z})$, the estimate for A_s'' in Assumption 6.1(iii) and again (A.2), we deduce

$$\frac{\|r_2(y;\zeta)(t)\|_{L(\mathcal{Z};\mathcal{W})}}{\|h_1\|_{H^1(0,T;\mathcal{X})}}$$
$$= \frac{\|A_s'(y(t)+\zeta(t)) - A_s'(y(t)) - A_s''(y(t))\zeta(t)\|_{L(\mathcal{Z};\mathcal{W})}}{\|h_1\|_{H^1(0,T;\mathcal{X})}} \leq C \frac{\|\zeta(t)\|_{\mathcal{Y}}}{\|h_1\|_{H^1(0,T;\mathcal{X})}} \leq C$$

for almost all $t \in [0,T]$. The convergence in (A.1) now follows from Lebesgue's dominated convergence theorem. □

Appendix B. Interpolation for the V^s spaces. We prove that the spaces $V^s = L^s(\Omega; W^{1,s}_{\mathrm{per},\perp}(Y;\mathbb{R}^d))$ defined in section 7 form a complex interpolation scale in s. This result is a cornerstone in the proof of Lemma 7.6.

LEMMA B.1. *Let $\theta \in (0,1)$ and $s_0, s_1 \in (1,\infty)$ and set $\frac{1}{s} = \frac{1-\theta}{s_0} + \frac{\theta}{s_1}$. Then*

(B.1)
$$\left[W^{1,s_0}_{\mathrm{per}}(Y;\mathbb{R}^d), W^{1,s_1}_{\mathrm{per}}(Y;\mathbb{R}^d)\right]_\theta = W^{1,s}_{\mathrm{per}}(Y;\mathbb{R}^d).$$

Proof. The proof relies on the complemented subspace interpolation theorem [38, Theorem 1.17.1] which essentially says that one can transfer interpolation properties to complemented subspaces provided there exists a common projection onto these subspaces on all involved spaces.

In this spirit, we first consider a larger regular domain $Y^\# \supset Y$ which includes a finite open covering of Y, and, for all $1 < r < \infty$, identify $W^{1,r}_{\mathrm{per}}(Y;\mathbb{R}^d)$ isomorphically with the closed subspace $W^{1,r}_{\mathrm{per},Y}(Y^\#;\mathbb{R}^d)$ of $W^{1,r}(Y^\#;\mathbb{R}^d)$ consisting of periodic extensions of $W^{1,r}_{\mathrm{per}}(Y;\mathbb{R}^d)$-functions. This is possible since the periodic extension of a $W^{1,r}_{\mathrm{per}}(Y;\mathbb{R}^d)$ function will preserve the Sobolev regularity [14, Proposition 3.50].

We next argue that there exists a projection P_{per} which projects $W^{1,r}(Y^\#;\mathbb{R}^d)$ onto $W^{1,r}_{\mathrm{per},Y}(Y^\#;\mathbb{R}^d)$. (We will not give a detailed proof of this since the details are somewhat tedious and lengthy.) To this end, we first wrap $u \in W^{1,r}(Y^\#;\mathbb{R}^d)$ around the torus $\mathbb{T}^d$ in a smooth manner using a fixed smooth partition of unity on $\mathbb{T}^d$ derived from the open covering of Y, and then pull it back. One checks that this indeed yields a function $P_{\mathrm{per}}u$ which is periodic on Y. Moreover, P_{per} is in fact a continuous linear operator on $W^{1,r}(Y^\#;\mathbb{R}^d)$, which in addition acts as the identity on the periodic extensions of $C^\infty_{\mathrm{per}}(Y;\mathbb{R}^d)$. This implies that P_{per} is indeed the searched-for projection of $W^{1,r}(Y^\#;\mathbb{R}^d)$ onto $W^{1,r}_{\mathrm{per},Y}(Y^\#;\mathbb{R}^d)$.

The complemented subspace interpolation theorem [38, Theorem 1.17.1] then al-

lows to argue as follows:

$$
\left[W^{1,s_0}_{\mathrm{per}}(Y;\mathbb{R}^d), W^{1,s_1}_{\mathrm{per}}(Y;\mathbb{R}^d)\right]_\theta = \left[W^{1,s_0}_{\mathrm{per},Y}(Y^\#;\mathbb{R}^d), W^{1,s_1}_{\mathrm{per},Y}(Y^\#;\mathbb{R}^d)\right]_\theta
$$
$$
= \left[W^{1,s_0}(Y^\#;\mathbb{R}^d)\cap W^{1,\max(s_0,s_1)}_{\mathrm{per},Y}(Y^\#;\mathbb{R}^d), W^{1,s_1}(Y^\#;\mathbb{R}^d)\cap W^{1,\max(s_0,s_1)}_{\mathrm{per},Y}(Y^\#;\mathbb{R}^d)\right]_\theta
$$
$$
= \left[W^{1,s_0}(Y^\#;\mathbb{R}^d), W^{1,s_1}(Y^\#;\mathbb{R}^d)\right]_\theta \cap W^{1,\max(s_0,s_1)}_{\mathrm{per},Y}(Y^\#;\mathbb{R}^d)
$$
$$
= W^{1,s}(Y^\#;\mathbb{R}^d)\cap W^{1,\max(s_0,s_1)}_{\mathrm{per},Y}(Y^\#;\mathbb{R}^d) = W^{1,s}_{\mathrm{per},Y}(Y^\#;\mathbb{R}^d) = W^{1,s}_{\mathrm{per}}(Y;\mathbb{R}^d).
$$

Here, the interpolation of $W^{1,r}(Y^\#;\mathbb{R}^d)$ is classical since we have assumed $Y^\#$ to be regular. Overall, this gives the claim. □

LEMMA B.2. *Let* $\theta \in (0,1)$ *and* $s_0, s_1 \in (1,\infty)$ *and set* $\frac{1}{s} = \frac{1-\theta}{s_0} + \frac{\theta}{s_1}$. *Then*

$$
\left[W^{1,s_0}_{\mathrm{per},\perp}(Y;\mathbb{R}^N), W^{1,s_1}_{\mathrm{per},\perp}(Y;\mathbb{R}^N)\right]_\theta = W^{1,s}_{\mathrm{per},\perp}(Y;\mathbb{R}^N).
$$

Proof. We again argue via the complemented subspace interpolation theorem. For every $1 < r < \infty$, the operator

$$
P_\perp u := u - \fint_Y u\,dy
$$

is a projection of $W^{1,r}(Y;\mathbb{R}^d)$ onto $W^{1,r}_\perp(Y;\mathbb{R}^d)$. Note that $P_\perp$ maps the closed subspace $W^{1,r}_{\mathrm{per}}(Y;\mathbb{R}^d)$ into itself, hence $W^{1,r}_{\mathrm{per},\perp}(Y;\mathbb{R}^d)$ is a complemented subspace of $W^{1,r}_{\mathrm{per}}(Y;\mathbb{R}^d)$ by means of the projection $P_\perp$. Using [38, Theorem 1.17.1] and Lemma B.1, we thus obtain

$$
\left[W^{1,s_0}_{\mathrm{per},\perp}(Y;\mathbb{R}^d), W^{1,s_1}_{\mathrm{per},\perp}(Y;\mathbb{R}^d)\right]_\theta
$$
$$
= \left[W^{1,s_0}_{\mathrm{per}}(Y;\mathbb{R}^d)\cap W^{1,\max(s_0,s_1)}_\perp(Y;\mathbb{R}^d), W^{1,s_1}_{\mathrm{per}}(Y;\mathbb{R}^d)\cap W^{1,\max(s_0,s_1)}_\perp(Y;\mathbb{R}^d)\right]_\theta
$$
$$
= \left[W^{1,s_0}_{\mathrm{per}}(Y;\mathbb{R}^d), W^{1,s_1}_{\mathrm{per}}(Y;\mathbb{R}^d)\right]_\theta \cap W^{1,\max(s_0,s_1)}_\perp(Y;\mathbb{R}^d)
$$
$$
= W^{1,s}_{\mathrm{per}}(Y;\mathbb{R}^d)\cap W^{1,\max(s_0,s_1)}_\perp(Y;\mathbb{R}^d) = W^{1,s}_{\mathrm{per},\perp}(Y;\mathbb{R}^d),
$$

as desired. □

THEOREM B.3. *Let* $\theta \in (0,1)$ *and* $s_0, s_1 \in (1,\infty)$ *and set* $\frac{1}{s} = \frac{1-\theta}{s_0} + \frac{\theta}{s_1}$. *Then*

$$
\left[V_{s_0}, V_{s_1}\right]_\theta = V_s.
$$

Proof. By [38, Thm. 1.18.4], we have

$$
\left[V_{s_0}, V_{s_1}\right]_\theta = \left[L^{s_0}\bigl(\Omega; W^{1,s_0}_{\mathrm{per},\perp}(Y;\mathbb{R}^d)\bigr), L^{s_1}\bigl(\Omega; W^{1,s_1}_{\mathrm{per},\perp}(Y;\mathbb{R}^d)\bigr)\right]_\theta
$$
$$
= L^s\bigl(\Omega; \left[W^{1,s_0}_{\mathrm{per},\perp}(Y;\mathbb{R}^d), W^{1,s_1}_{\mathrm{per},\perp}(Y;\mathbb{R}^d)\right]_\theta\bigr)
$$
$$
= L^s\bigl(\Omega; W^{1,s}_{\mathrm{per},\perp}(Y;\mathbb{R}^d)\bigr) = V_s,
$$

where the interpolation identity for $W^{1,s}_{\mathrm{per},\perp}(Y;\mathbb{R}^d)$ is a consequence of Lemma B.2. □

Appendix C. Auxiliary Results.

LEMMA C.1. *Let $\mathcal{H}$ be a Hilbert space, $\{x_n\}_{n\in\mathbb{N}} \subset \mathcal{H}$, $\{a_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$, $x \in \mathcal{H}$ and $a \in \mathbb{R}$. If $x_n \rightharpoonup x$ and $\|x_n\|_{\mathcal{H}} \leq a_n \to a \leq \|x\|_{\mathcal{H}}$, then $x_n \to x$.*

Proof. Since $\mathcal{H}$ is a Hilbert space and $x_n \rightharpoonup x$, it is sufficient to prove $\|x_n\|_{\mathcal{H}} \to \|x\|_{\mathcal{H}}$. To this end we argue by contradiction. Let us assume that $\|x_n\|_{\mathcal{H}}$ does not converge towards $\|x\|_{\mathcal{H}}$. Then there exists $\varepsilon > 0$ and a subsequence, again denoted by x_n, such that $\varepsilon \leq |\|x\|_{\mathcal{H}} - \|x_n\|_{\mathcal{H}}|$. Due to $\|x_n\|_{\mathcal{H}} \leq a_n \to a \leq \|x\|_{\mathcal{H}}$, we can assume that $\|x_n\|_{\mathcal{H}} \leq \|x\|_{\mathcal{H}} - \varepsilon$, that is, $x_n \in B := B_{\|x\|_{\mathcal{H}}-\varepsilon}^{\mathcal{H}}(0)$ for all $n \in \mathbb{N}$. Since x_n converges weakly towards x and B is weakly closed, we have $x \in B$, which is not true. $\qquad\square$

LEMMA C.2. *Let $\mathcal{M}$ be a compact metric space and $\mathcal{N}$ a metric space. Furthermore, let $\{G_n\}_{n\in\mathbb{N}} \subset C(\mathcal{M};\mathcal{N})$, $G \in C(\mathcal{M};\mathcal{N})$ with $G_n(x) \to G(x)$ for all $x \in \mathcal{M}$ and suppose that G_n is uniformly Lipschitz continuous, that is, there exists a constant L such that*

$$d_{\mathcal{N}}(G_n(x), G_n(y)) \leq L d_{\mathcal{M}}(x,y)$$

for all $n \in \mathbb{N}$, $x,y \in \mathcal{M}$.
Then $G_n \to G$ in $C(\mathcal{M};\mathcal{N})$.

Proof. We argue by contradiction. Assume that there exists $\varepsilon > 0$ and a strictly monotonically increasing function $n\colon \mathbb{N} \to \mathbb{N}$, such that for all $k \in \mathbb{N}$ there exists $x_k \in \mathcal{M}$ with

$$\varepsilon \leq d_{\mathcal{N}}(G_{n(k)}(x_k), G(x_k))$$

for all $k \in \mathbb{N}$. Since $\mathcal{M}$ is compact, we can extract a subsequence x_{k_j} of x_k such that $x_{k_j} \to x$ in $\mathcal{M}$, thus

$$d_{\mathcal{N}}(G_{n(k_j)}(x_{k_j}), G(x_{k_j})) \leq d_{\mathcal{N}}(G_{n(k_j)}(x_{k_j}), G_{n(k_j)}(x)) + d_{\mathcal{N}}(G_{n(k_j)}(x), G(x_{k_j}))$$
$$\leq L d_{\mathcal{M}}(x_{k_j}, x) + d_{\mathcal{N}}(G_{n(k_j)}(x), G(x_{k_j})) \to 0,$$

which gives the contradiction. $\qquad\square$

LEMMA C.3. *Let $\mathcal{M}$ be a compact metric space and $\mathcal{N}$ a metric space. Furthermore, let $\{G_n\}_{n\in\mathbb{N}} \subset C(\mathcal{M};\mathcal{N})$, $G \in C(\mathcal{M};\mathcal{N})$ with $G_n \to G$ in $C(\mathcal{M};\mathcal{N})$. Define $U_n := G_n(\mathcal{M})$ and $U_0 := G(\mathcal{M})$. Then the set $U := \cup_{n=0}^{\infty} U_n$ is compact.*

Proof. Let $\{y_k\}_{k\in\mathbb{N}} \subset U$. Since a finite union of compact sets is compact, we can assume that there exists a subsequence $\{y_{k_j}\}_{j\in\mathbb{N}}$ and a strictly monotonically increasing function $n\colon \mathbb{N} \to \mathbb{N}$, such that $y_{k_j} \in U_{n(j)}$. Then there exists a sequence $\{x_j\}_{j\in\mathbb{N}} \subset \mathcal{M}$, with $y_{k_j} = G_{n(j)}(x_j)$. Because $\mathcal{M}$ is compact we can select a subsequence, again denoted by x_j, and a limit $x \in \mathcal{M}$, such that $x_j \to x$, hence,

$$d_{\mathcal{N}}(y_{k_j}, G(x)) \leq d_{\mathcal{N}}(y_{k_j}, G(x_j)) + d_{\mathcal{N}}(G(x_j), G(x)) \to 0,$$

thus the proof is complete. $\qquad\square$

REFERENCES

[1] L. ADAM AND J. OUTRATA, *On optimal control of a sweeping process coupled with an ordinary differential equation*, Discrete Contin. Dyn. Syst. Ser. B, 19 (2014), pp. 2709–2738, https://doi.org/10.3934/dcdsb.2014.19.2709, https://doi.org/10.3934/dcdsb.2014.19.2709.

[2] C. E. Arroud and G. Colombo, *A maximum principle for the controlled sweeping process*, Set-Valued Var. Anal., 26 (2018), pp. 607–629, https://doi.org/10.1007/s11228-017-0400-4, https://doi.org/10.1007/s11228-017-0400-4.

[3] P. Auscher, S. Bortz, M. Egert, and O. Saari, *Nonlocal self-improving properties: a functional analytic approach*, Tunisian J. Math., 1 (2019), pp. 151–183.

[4] V. Barbu, *Optimal Control of Variational Inequalities*, vol. 100 of Research Notes in Mathematics, Pitman, Boston, 1984.

[5] V. Barbu, *Analysis and control of nonlinear infinite dimensional systems*, Academic Press Boston, 1993.

[6] S. Bechtel and M. Egert, *Interpolation theory for sobolev functions with partially vanishing trace on irregular open sets*, Journal of Fourier Analysis and Applications, (2019), https://doi.org/10.1007/s00041-019-09681-1.

[7] L. Boron and E. Zeidler, *Nonlinear Functional Analysis and its Applications: III: Variational Methods and Optimization*, Springer New York, 1984, https://books.google.de/books?id=jpcnAQAAIAAJ.

[8] H. Brézis, *Opérateurs maximaux monotones*, North-Holland, Amsterdam, 1973.

[9] M. Brokate, *Optimale Steuerung von gewöhnlichen Differentialgleichungen mit Nichtlinearitäten vom Hysteresis-Typ*, vol. 35 of Methoden und Verfahren der Mathematischen Physik [Methods and Procedures in Mathematical Physics], Verlag Peter D. Lang, Frankfurt am Main, 1987.

[10] M. Brokate and P. Krejčí, *Optimal control of ODE systems involving a rate independent variational inequality*, Discrete Contin. Dyn. Syst. Ser. B, 18 (2013), pp. 331–348, https://doi.org/10.3934/dcdsb.2013.18.331, https://doi.org/10.3934/dcdsb.2013.18.331.

[11] M. Brokate and P. Krejčí, *Weak differentiability of scalar hysteresis operators*, Discrete Contin. Dyn. Syst., 35 (2015), pp. 2405–2421, https://doi.org/10.3934/dcds.2015.35.2405, https://doi.org/10.3934/dcds.2015.35.2405.

[12] E. Casas and F. Tröltzsch, *Second order necessary and sufficient conditions for optimization problems and applications to control theory*, SIAM Journal on Optimization, 13 (2002), pp. 406–431.

[13] E. Casas and F. Tröltzsch, *Second order optimality conditions and their role in pde control*, Jahresbericht der Deutschen Mathematiker-Vereinigung, 117 (2015), pp. 3–44.

[14] D. Cioranescu and P. Donato, *An Introduction to Homogenization*, Oxford University Press, 1999.

[15] G. Colombo, R. Henrion, N. D. Hoang, and B. S. Mordukhovich, *Discrete approximations of a controlled sweeping process*, Set-Valued Var. Anal., 23 (2015), pp. 69–86, https://doi.org/10.1007/s11228-014-0299-y, https://doi.org/10.1007/s11228-014-0299-y.

[16] G. Colombo, R. Henrion, N. D. Hoang, and B. S. Mordukhovich, *Optimal control of the sweeping process over polyhedral controlled sets*, J. Differential Equations, 260 (2016), pp. 3397–3447, https://doi.org/10.1016/j.jde.2015.10.039, https://doi.org/10.1016/j.jde.2015.10.039.

[17] G. Colombo and M. Palladino, *The minimum time function for the controlled Moreau's sweeping process*, SIAM J. Control Optim., 54 (2016), pp. 2036–2062, https://doi.org/10.1137/15M1043364, https://doi.org/10.1137/15M1043364.

[18] J. Elschner, J. Rehberg, and G. Schmidt, *Optimal regularity for elliptic transmission problems including C^1 interfaces*, Interfaces Free Bound., 9 (2007), pp. 233–252, https://doi.org/10.4171/ifb/163.

[19] E. Emmrich, *Gewöhnliche und Operator-Differentialgleichungen*, Vieweg+Teubner Verlag, Wiesbaden, 2004.

[20] H. Gajewski, K. Gröger, and K. Zacharias, *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen*, Mathematische Nachrichten, 67 (1975), pp. iv–iv.

[21] T. Geiger and D. Wachsmuth, *Optimal control of an evolution equation with non-smooth dissipation.* arXiv:1801.04077, 2018.

[22] H. Goldberg, W. Kampowsky, and F. Tröltzsch, *On Nemytskij operators in L^p-spaces of abstract functions*, Mathematische Nachrichten, 155 (1992), pp. 127–140, https://doi.org/10.1002/mana.19921550110.

[23] K. Gröger, *Initial value problems for elastoplastic and elasto-viscoplastic systems*, BSB BG Teubner Verlagsgesellschaft, 1979.

[24] K. Gröger, *A $W^{1,p}$-estimate for solutions to mixed boundary value problems for second order elliptic differential equations*, Math. Ann., 283 (1989), pp. 679–687, https://doi.org/10.1007/bf01442860.

[25] W. Han and B. D. Reddy, *Plasticity*, vol. 9 of Interdisciplinary Applied Mathematics, Springer-Verlag, New York, 1999, https://doi.org/10.1007/978-1-4614-5940-8.

[26] R. Herzog and C. Meyer, *Optimal control of static plasticity with linear kinematic hardening*, Journal of Applied Mathematics and Mechanics (ZAMM), 91 (2011), pp. 777–794.

[27] 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 (2012), pp. 3052–3082.

[28] N. G. Meyers, *An L^p-estimate for the gradient of solutions of second order elliptic divergence equations*, Annali della Scuola Normale Superiore di Pisa - Classe di Scienze, 17 (1963), pp. 189–206, http://eudml.org/doc/83302.

[29] A. Mielke and T. Roubíček, *Rate-independent systems. Theory and application.*, vol. 193, New York, NY: Springer, 2015, https://doi.org/10.1007/978-1-4939-2706-7.

[30] J.-J. Moreau, *Problème d'évolution associé à un convexe mobile d'un espace Hilbertien*, C. R. Acad. Sci. Paris Sér. A-B, 276 (1973), pp. A791–A794.

[31] F. Rindler, *Optimal control for nonconvex rate-independent evolution processes*, SIAM J. Control Optim., 47 (2008), pp. 2773–2794, https://doi.org/10.1137/080718711, https://doi.org/10.1137/080718711.

[32] F. Rindler, *Approximation of rate-independent optimal control problems*, SIAM J. Numer. Anal., 47 (2009), pp. 3884–3909, https://doi.org/10.1137/080744050, https://doi.org/10.1137/080744050.

[33] B. Schweizer and M. Veneroni, *Homogenization of plasticity equations with two-scale convergence methods*, Applicable Analysis, 94 (2015), pp. 375–398.

[34] E. Shamir, *Regularization of mixed second-order elliptic problems*, Israel J. Math., 6 (1968), pp. 150–168, https://doi.org/10.1007/bf02760180.

[35] R. E. Showalter, *Monotone operators in Banach space and nonlinear partial differential equations*, vol. 49, American Mathematical Soc., 2013.

[36] I. J. Šneĭberg, *Spectral properties of linear operators in interpolation families of Banach spaces*, Mat. Issled., 9 (1974), pp. 214–229, 254–255.

[37] U. Stefanelli, D. Wachsmuth, and G. Wachsmuth, *Optimal control of a rate-independent evolution equation via viscous regularization*, Discrete Contin. Dyn. Syst. Ser. S, 10 (2017), pp. 1467–1485, https://doi.org/10.3934/dcdss.2017076, https://doi.org/10.3934/dcdss.2017076.

[38] H. Triebel, *Interpolation theory, function spaces, differential operators*, vol. 18 of North-Holland Mathematical Library, North-Holland Publishing Co., Amsterdam-New York, 1978.

[39] F. Tröltzsch, *Optimale Steuerung partieller Differentialgleichungen*, Springer, 2005.

[40] G. Wachsmuth, *Optimal control of quasi-static plasticity with linear kinematic hardening, Part I: Existence and discretization in time*, SIAM J. Control Optim., 50 (2012), pp. 2836–2861 + loose erratum, https://doi.org/10.1137/110839187, https://doi.org/10.1137/110839187.

[41] G. Wachsmuth, *Differentiability of implicit functions: Beyond the implicit function theorem*, Journal of Mathematical Analysis and Applications, 414 (2014), pp. 259–272.

[42] G. Wachsmuth, *Optimal control of quasistatic plasticity with linear kinematic hardening II: Regularization and differentiability*, Z. Anal. Anwend., 34 (2015), pp. 391–418, https://doi.org/10.4171/ZAA/1546, https://doi.org/10.4171/ZAA/1546.

[43] G. Wachsmuth, *Optimal control of quasistatic plasticity with linear kinematic hardening III: Optimality conditions*, Z. Anal. Anwend., 35 (2016), pp. 81–118, https://doi.org/10.4171/ZAA/1556, https://doi.org/10.4171/ZAA/1556.

[44] E. Zeidler and L. Boron, *Nonlinear Functional Analysis and Its Applications: II/ A: Linear Monotone Operators*, Monotone operators / transl. by the author and by Leo F. Boron, Springer New York, 1989, https://books.google.de/books?id=sk1M2Vw-Y24C.