Deutsche
Forschungsgemeinschaft
Priority Programme 1962

First-order Conditions for the Optimal Control of the Obstacle Problem with State Constraints

Ira Neitzel, Gerd Wachsmuth

SPP 1962
Non-smooth and Complementarity-based
Distributed Parameter Systems:
Simulation and Hierarchical Optimization

Preprint Number SPP1962-154

received on December 30, 2020

Edited by
SPP1962 at Weierstrass Institute for Applied Analysis and Stochastics (WIAS)
Leibniz Institute in the Forschungsverbund Berlin e.V.
Mohrenstraße 39, 10117 Berlin, Germany
E-Mail: spp1962@wias-berlin.de

World Wide Web: `http://spp1962.wias-berlin.de/`

First-order conditions for the optimal control of the obstacle problem with state constraints*

Ira Neitzel† Gerd Wachsmuth‡

December 30, 2020

We consider an optimal control problem in which the state is governed by an unilateral obstacle problem (with obstacle from below) and restricted by a pointwise state constraint (from above). In the presence of control constraints, we prove, via regularization of the state constraints, that a system of C-stationarity is necessary for optimality. In the absence of control constraints, we show that local minimizers are even strongly stationary by a careful discussion of the primal first-order conditions of B-stationary type.

Keywords: Obstacle problem, state constraints, C-stationarity, strong stationarity
MSC (2020): 49K21, 35J86

1 Introduction

In this paper, we analyse an optimal control problem subject to an obstacle (from below) and subject to an additional pointwise state constraint (from above). More precisely, we are interested in the problem

$$\begin{aligned} &\text{Minimize} && J(y,u) \\ &\text{with respect to} && (y,u) \in H_0^1(\Omega) \times L^2(\Omega) \\ &\text{such that} && y \in K,\ \langle \mathcal{A}y - u, v - y\rangle \geq 0 \quad \forall v \in K, \\ &&& u \in U_{\text{ad}} \\ &\text{and} && y \leq y_b \quad \text{a.e. in } \Omega. \end{aligned} \tag{P}$$

*This research was supported by the German Research Foundation (DFG) within the priority program "Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization" (SPP 1962) under grant numbers NE 1941/1-2 and WA 3636/4-2.

†Rheinische Friedrich Wilhelms Universität Bonn, Institute for Numerical Simulation, Friedrich-Hirzebruch-Allee 7, 53115 Bonn, Germany, https://ins.uni-bonn.de/staff/neitzel
‡Brandenburgische Technische Universität Cottbus-Senftenberg, Institute of Mathematics, 03046 Cottbus, Germany, https://www.b-tu.de/fg-optimale-steuerung

In our notation, y is the state and u is the control, and the set K is defined by the lower obstacle y_a, i.e.,

$$K := \{v \in H_0^1(\Omega) \mid v \geq y_a \text{ a.e. in } \Omega\}.$$

For the precise assumptions we refer to Assumption 2.13. We just mention that we work with minimal regularity, i.e., the coefficients in the differential operator $\mathcal{A}$ are just assumed to be measurable and bounded, and we only require the so-called "uniform exterior cone condition" of the bounded domain $\Omega \subset \mathbb{R}^n$, $n \in \{2, 3\}$.

Since the solution operator S of the obstacle is not differentiable, its optimal control is a challenging problem. The first contribution is the seminal work [Mignot, 1976], which demonstrates the directional differentiability of S and in which optimality conditions of strongly stationary type were derived. Another classical work is [Barbu, 1984], in which regularization methods are used to provide stationarity conditions. It turns out that regularization techniques are applicable to a wider range of problems, in particular to problems including control constraints. However, the resulting optimality conditions are weaker than strong stationarity. It seems that the so-called system of C-stationarity is the best system which can be derived in this way. For some more recent contributions to the optimal control of the obstacle problem, we refer to [Hintermüller and Surowiec, 2011; Kunisch and D. Wachsmuth, 2011; Schiela and D. Wachsmuth, 2013; G. Wachsmuth, 2014; 2016; Harder and G. Wachsmuth, 2018a].

PDE-constrained optimal control problems with pointwise state constraints have also been known as a challenging problem class with respect to optimality conditions for quite some time. The theory is typically based on the so called Slater condition, for which continuity of the states is usually required. Lagrange multipliers in the first-order optimality system are then obtained in the space of regular Borel measures, see [Casas, 1986], which in turn leads to low regularity of the adjoint state. From the meanwhile very rich literature on different aspects of purely state-constrained problems we refer only to [Casas, 1993], where the results of [Casas, 1986] have also been extended to boundary control of semilinear elliptic PDEs, to [Raymond and Zidani, 1998; 1999] for some earlier results on (semilinear) parabolic problems, as well as to related problems with constraints of bottleneck type, [Bergounioux and Tröltzsch, 1999]. We also mention [Hintermüller and Kunisch, 2009] for problems with control, state and gradient constraints. The structure of Lagrange multipliers has for instance been discussed in [Bergounioux and Kunisch, 2002]. More recently, in [Casas et al., 2014], the authors showed improved regularity for the Lagrange multiplier, inspired by a result for sparse optimal controls, see [Pieper and Vexler, 2013]. Sparse control became a field of active research rather recently. These problems resemble state-constrained problems in the way that if for instance measures as control variables are considered, this leads to regularity difficulties in the state equation, instead of the adjoint equation in case of pointwise state constraints.

The challenges and regularity issues associated with the Lagrange multiplier also influence all further analysis, such as second-order sufficient conditions (SSC) for nonconvex problems, analysis of solution algorithms, or numerical analysis of such problems. For an introductory overview on general aspects of second-order sufficient conditions and finite element error analysis for PDE-constrained optimization, not restricted to pointwise state constraints, we mention [Casas and Tröltzsch, 2015] or [Hinze and Tröltzsch, 2010], respectively.

There are a number of well-established regularization techniques that help to avoid theoretical and numerical difficulties associated with pointwise state constraints. We mention the rather classical Moreau-Yosida regularization approach from [Ito and Kunisch, 2003] that we will pursue in this paper, a Lavrentiev-regularization technique presented originally in [Meyer et al., 2006], as well as the virtual control regularization approach from [Krumbiegel and Rösch, 2009]. Moreover, barrier methods are in use; we refer for instance to [Schiela, 2009] or [M. Ulbrich and S. Ulbrich, 2009]. As the literature on regularization of pointwise state constraints is meanwhile also rather rich, let us only refer to [Bergounioux et al., 2000; Hintermüller et al., 2003; 2008] as examples of a few publications related to the Moreau-Yosida penalization.

We also would like to point out [Rösch and Tröltzsch, 2007] and the references therein, where elliptic and parabolic problems with mixed control-state constraints have been considered. Lagrange multipliers are shown to exist in L^p-spaces. While these constraints exhibit better regularity properties, the analysis begins with existence of multipliers in the dual space of L^∞, which is even less regular than the space of regular Borel measures. The regularity of the multipliers and also the optimal control is subsequently improved. For a problem with bilateral control and mixed control-state constraints, a separation condition for the active sets allows to prove L^1-regularity of the multipliers and further regularity improvements via bootstrapping arguments. Such a separation condition has also been used in e.g. [Alt et al., 2010] for stability analysis of linear-quadratic elliptic problems with mixed constraints, for convergence analysis of the SQP method for nonlinear problems in [Griesse et al., 2008], and in [Neitzel and Tröltzsch, 2009] for Lavrentiev regularization of pointwise state constraints in parabolic problems.

In our analysis, we will rely on separate supports of the multipliers associated with the obstacle and the state constraints. In essence, this condition allows to apply typical cut-off-type arguments. If the support of a low-regularity Lagrange-multiplier is clearly separated from another point or rather area of interest, the smoothing properties of the solution operators can be used to prove higher regularity on appropriate subdomains. In our case, this means that the adjoint state admits H_0^1-regularity away from the support of the state-constraint multiplier. Cut-off arguments are a typical strategy to prove known higher interior regularity results for PDEs. For PDE-constrained optimization, such techniques have for example also been used to consider Dirichlet boundary control of Poisson's equation with pointwise state constraints in the interior, analyzed in [Mateos and Neitzel, 2016] even though the states only admit $H^{\frac{1}{2}}(\Omega)$ regularity on the whole domain.

The literature concerning optimal control of the obstacle problem with additional state constraints is rather scarce. We are only aware of three publications.

In [He, 1987], the author considers a problem which is more general than (**P**). In [He, 1987, Theorem 6.2], a system of C-stationarity is derived which includes a multiplier $\nu \in L^2(\Omega)$ for the state constraint. This unusual high regularity seems to be related to the requirement (6.5) therein. It is assumed that perturbations $z \in L^2(\Omega)$ in the state constraint lead to perturbations of the optimal value which can be bounded from below (up to first order) by the $L^2(\Omega)$-norm of z. It is not clear whether this assumption can be verified for a large class of problems. We do not expect that the multiplier of the state constraint considered in our paper belongs to $L^2(\Omega)$.

In the contribution [Bergounioux, 1998], a more general variational inequality is considered. This variational inequality is regularized and optimality conditions for the regularized problem (subject to the state constraints) are derived. The passage to the limit in this optimality system is not addressed.

[Bergounioux and Tiba, 1998, Section 3] addresses a problem very similar to (**P**), but the state constraint is defined by a closed convex set in $H_0^1(\Omega)$. Again, the obstacle problem is regularized. This contribution also addresses the passage to the limit in the optimality system. The resulting optimality system uses a limiting object defined in [Bergounioux and Tiba, 1998, Definition 3.1]. It is not clear how much information is carried by this object. We just mention that optimality systems defined by the so-called limiting normal cone seems to be of limited use in infinite dimensions, see [Mehlitz and G. Wachsmuth, 2019; Harder and G. Wachsmuth, 2018b].

In this work, we present three optimality systems. First, in Theorem 3.5 we derive a primal optimality system (i.e., it does not involve dual multipliers), which is typically called B-stationarity. To this end, we heavily utilize that the pointwise convexity of the solution operator of the obstacle problem (see Lemma 2.11) renders the state constraint convex w.r.t. the control, i.e., the set $U_{\text{state}} := \{u \in L^2(\Omega) \mid S(u) \le y_b\}$ is convex, where S is the solution operator of the obstacle problem, see Lemma 3.1.

Second, we use a regularization approach to derive optimality conditions of C-stationary type, see Theorem 4.1. To this end, we just regularize the state constraints and apply the C-stationarity conditions from [G. Wachsmuth, 2016] to the regularized problems. The passage to the limit in the optimality system requires some care due to the multiplier of the state constraint, which is a measure and appears on the right-hand side of the adjoint equation.

Finally, we show that the system of B-stationarity is equivalent to strong stationarity in the absence of control constraints, see Theorem 5.7. To this end, we utilize classical results by [Mignot, 1976]. Due to the state constraints, this is much more difficult than in the classical setting. We mention that this result also allows to characterize the normal cone to the convex set U_{state}.

The paper is structured as follows. In Section 2.1 we fix some notation. Linear PDEs governed by the differential operator $\mathcal{A}$ and its adjoint $\mathcal{A}^\star$ are discussed in Section 2.2. Of particular importance are Lemmas 2.5 and 2.6. In these results, we discuss PDEs with irregular but localized right-hand sides and show that the solution enjoys higher regularity if we neglect a neighborhood of the support of the right-hand side. Afterwards, the obstacle problem is discussed in Section 2.3. Using the regularity results of the previous section, we can provide uniform estimates for the directional derivative S' away from the active set $\{y = y_a\}$, see Lemma 2.10. In Section 2.4, we collect all the assumptions, see Assumption 2.13, and give some basic properties of (**P**). The optimality conditions are discussed in Sections 3 to 5, as described above.

2 Preliminaries and technical results

We start by setting up the notation in Section 2.1. Afterwards, we discuss the solution mapping of the differential operator in Section 2.2. The variational inequality will be addressed in Section 2.3, and finally, we give some basic properties of the optimal control problem in Section 2.4.

2.1 Notation

Let us fix some notation. The positive numbers are denoted by $\mathbb{R}^+ := (0, \infty)$. For a (real) Banach space X, we denote by $X^\star$ the (topological) dual space of X. The corresponding duality pairing is denoted by $\langle \cdot, \cdot \rangle_X : X^\star \times X \to \mathbb{R}$. If the space X is clear from the context, we may omit the index X. The inner product in $L^2(\Omega)$ is denoted by $(\cdot, \cdot)$.

For a subset set $C \subset X$, we define the polar cone and the annihilator by

$$C^\circ := \{\xi \in X^\star \mid \forall x \in C : \langle \xi, x \rangle_X \le 0\},$$
$$C^\perp := \{\xi \in X^\star \mid \forall x \in C : \langle \xi, x \rangle_X = 0\},$$

respectively. Analogously, we define $D^\circ, D^\perp \subset X$ for $D \subset X^\star$. In particular, the annihilator of $\xi \in X^\star$ is defined via

$$\xi^\perp := \{x \in X \mid \langle \xi, x \rangle_X = 0\}.$$

Now, let $C \subset X$ be closed and convex. For all $x \in C$, we define the radial cone, the tangent cone, and the normal cone to C at x via

$$\mathcal{R}_C(x) := \bigcup_{\lambda > 0} (C - x), \qquad \mathcal{T}_C(x) := \operatorname{cl}(\mathcal{R}_C(x)), \qquad \mathcal{N}_C(x) := \mathcal{T}_C(x)^\circ = (C - x)^\circ,$$

respectively. In case $x \in X \setminus C$, we define all these cones to be the empty set. In case $X = L^2(\Omega)$, where $\Omega \subset \mathbb{R}^n$ is measurable, we identify $X^\star$ with X in the canonical way and, therefore, interpret $\mathcal{N}_C(x)$ as a subset of X.

Further, we need some basic concepts of capacity theory. For a summary, we refer to [G. Wachsmuth, 2014, Section 2], [G. Wachsmuth, 2016, Section 1.2] and [Harder and G. Wachsmuth, 2018a, Section 3]. In particular, we require the definition of the quasi-support q-supp(ξ) for measures $\xi \in H^{-1}(\Omega)^-$, see [G. Wachsmuth, 2014, Lemma 3.1] and [Harder and G. Wachsmuth, 2018a, Lemma 3.7, Definition 3.8] (called "fine support" therein).

If $y : \Omega \to \mathbb{R}$ is some function, we define

$$\{y > 0\} := \{x \in \Omega \mid y(x) > 0\}.$$

Note that this set is defined up to sets of measure zero if $y \in L^r(\Omega)$, $r \in [1, \infty]$, and up to sets of capacity zero if $y \in H^1(\Omega)$. The same notation will be used for other relations and also with more than one function, e.g., $\{y_1 > y_2\}$ with the obvious meaning.

2.2 Solution operators of differential equations

In this section, we specify the properties of the differential operator $\mathcal{A}$ appearing in (**P**). To this end, let us define $\mathcal{A}\colon H^1(\Omega) \to H^{-1}(\Omega)$ via

$$\begin{aligned}\langle \mathcal{A}y, v\rangle &:= \int_\Omega \sum_{i,j=1}^n a_{ij}\,\partial_j y\,\partial_i v + \sum_{i=1}^n b_i\,y\,\partial_i v + \sum_{i=1}^n c_i\,\partial_i y\,v + d\,y\,v\,\mathrm{d}x \\ &= \int_\Omega \nabla v^\top A\,\nabla y + y\,(b^\top \nabla v) + v\,(c^\top \nabla y) + d\,y\,v\,\mathrm{d}x \qquad \forall y \in H^1(\Omega), v \in H_0^1(\Omega).\end{aligned} \tag{2.1}$$

For the coefficients appearing in (2.1), we assume measurability and boundedness, i.e.,

$$a_{ij}, b_i, c_i, d \in L^\infty(\Omega) \tag{2.2}$$

on a domain $\Omega \subset \mathbb{R}^n$, $n \in \{2, 3\}$, satisfying a uniform exterior cone condition. Note that this is precisely the definition of the operator [Gilbarg and Trudinger, 2001, (8.1)] and the regularity condition of Ω described on page 205 therein. As usual, we will denote by $\Gamma := \partial\Omega$ the boundary of Ω. For our analysis, we assume that $\mathcal{A}$ is strictly elliptic, i.e., there exists $\gamma_0 > 0$ such that

$$\sum_{i,j=1}^n a_{ij}(x)\,w_i\,w_j = w^\top A(x)\,w \ge \gamma_0\,\|w\|^2_{\mathbb{R}^n} \qquad \forall w \in \mathbb{R}^n \text{ and a.a. } x \in \Omega. \tag{2.3}$$

Further, $\mathcal{A}$ should be coercive on $H_0^1(\Omega)$, i.e., there exists $\gamma_1 > 0$ such that

$$\langle \mathcal{A}y, y\rangle \ge \gamma_1\,\|y\|^2_{H_0^1(\Omega)} \qquad \forall y \in H_0^1(\Omega). \tag{2.4}$$

These assumptions on $\mathcal{A}$ hold throughout the paper. In the remainder of this section, we are concerned with existence and regularity results for the differential equation

$$y \in H_0^1(\Omega), \qquad \mathcal{A}y = f \quad \text{in } H^{-1}(\Omega) \tag{2.5}$$

and the associated adjoint equation. The following is a standard existence and regularity result and the starting point of our analysis.

Theorem 2.1. *Let $f \in H^{-1}(\Omega)$. Then there exists a unique weak solution $y \in H_0^1(\Omega)$ of* (2.5)*, satisfying*

$$\|y\|_{H_0^1(\Omega)} \le C\|f\|_{H^{-1}(\Omega)}. \tag{2.6}$$

Proof. Under our assumptions on $\mathcal{A}$, this follows by the well-known Lax-Milgram theory. □

If we can rely on more regularity of the data, we also obtain better regularity properties of the solution. However, we are limited by the low regularity (2.2) of the coefficients. The following theorem, that particularly includes the case of right-hand sides in $L^2(\Omega)$ for spatial dimensions up to $n = 3$, guarantees Hölder regularity. In the sequel, this will allow to rely on a so-called Slater condition as constraint qualification in order to obtain first-order necessary optimality conditions for controls in $L^2(\Omega)$.

Theorem 2.2. *For every $q' > n$ and $f \in W^{-1,q'}(\Omega)$, the unique solution $y \in H_0^1(\Omega)$ of equation (2.5) enjoys the additional regularity $y \in C_0^{0,\alpha}(\Omega)$ for some $\alpha > 0$ and we have the estimate*

$$\|y\|_{H_0^1(\Omega)} + \|y\|_{C_0^{0,\alpha}(\Omega)} \le C\|f\|_{W^{-1,q'}(\Omega)}. \tag{2.7}$$

Here, C and α do not depend on f.

Proof. The additional Hölder regularity follows from [Gilbarg and Trudinger, 2001, Theorem 8.29]. Together with [Gilbarg and Trudinger, 2001, Theorem 8.15], the norm estimate follows. □

Using this regularity, we can define the solution operator $T\colon W^{-1,q'}(\Omega) \to C_0(\Omega)$ of (2.5). Note that T is compact, since the Hölder space $C_0^{0,\alpha}(\Omega)$ is compactly embedded in $C_0(\Omega)$ by the Arzelà-Ascoli theorem. Its adjoint operator $T^\star\colon \mathcal{M}(\Omega) \to W_0^{1,q}(\Omega)$ is related to the adjoint equation of (2.5).

Theorem 2.3. *Let $q \in (1, n/(n-1))$ be given. For every $\nu \in \mathcal{M}(\Omega)$, the equation*

$$\mathcal{A}^\star p = \nu$$

admits a unique very weak solution $p \in W_0^{1,q}(\Omega)$, i.e.

$$\langle \mathcal{A}z, p\rangle_{W_0^{1,q}(\Omega)} = \int_\Omega z \,\mathrm{d}\nu \qquad \forall z \in Z, \tag{2.8}$$

where

$$Z := \{z \in H_0^1(\Omega) \mid \mathcal{A}z \in W^{-1,q'}(\Omega)\}. \tag{2.9}$$

Here, $q' \in (n, \infty)$ is the conjugate exponent of q. This solution fulfills the estimate

$$\|p\|_{W_0^{1,q}(\Omega)} \le C\|\nu\|_{\mathcal{M}(\Omega)}, \tag{2.10}$$

with C independent of ν. Finally, if $\nu_k \overset{\star}{\rightharpoonup} \nu$, we have $p_k \to p$ in $W_0^{1,q}(\Omega)$ for the corresponding solutions.

Note that $Z \subset C_0(\Omega)$ due to Theorem 2.2, hence, the right-hand side in (2.8) is well defined.

Proof. For $\nu \in \mathcal{M}(\Omega)$, we define the function $p := T^\star\nu \in W_0^{1,q}(\Omega)$. It is uniquely determined by the properties of the adjoint operator,

$$\langle f, p\rangle_{W_0^{1,q}(\Omega)} = \langle \nu, Tf\rangle_{\mathcal{M}(\Omega),C_0(\Omega)} = \int_\Omega (Tf)\,\mathrm{d}\nu \qquad \forall f \in W^{-1,q'}(\Omega). \tag{2.11}$$

From this, we observe that (2.10) holds. Note that (2.11) is equivalent to the very weak formulation (2.8).

In order to verify the compactness property, we use that $T^\star$ maps weak-$\star$ convergent sequences to weak-$\star$ convergent sequences (since it is an adjoint operator) and it maps bounded sequences to sequences possessing a strong accumulation point (since it is compact by Schauder's theorem). Hence, if $\nu_k \overset{\star}{\rightharpoonup} \nu$ in $\mathcal{M}(\Omega)$, every subsequence of $(T^\star \nu_k)_{k\in\mathbb{N}}$ possesses a strong accumulation point which has to coincide with $T^\star\nu$ due to $T^\star \nu_k \overset{\star}{\rightharpoonup} T^\star \nu$ in $W_0^{1,q}(\Omega)$. Hence, a subsequence-subsequence argument shows $T^\star\nu_k \to T^\star\nu$ in $W_0^{1,q}(\Omega)$. □

We point out that using the density of $C_0(\Omega)$ in $W^{-1,q'}(\Omega)$, it is also possible to use

$$\hat{Z} := \{z \in H_0^1(\Omega) \mid \mathcal{A}z \in C_0(\Omega)\},$$

as it was done in [Casas et al., 2014]. In particular, this shows that the very weak solution of (2.8) does not depend on the choice of the regularity exponent q.

Note that $\mathcal{A}z \in W^{-1,q'}(\Omega)$ in the definition of Z means that

$$|\langle \mathcal{A}z, v\rangle| \le C\|v\|_{W_0^{1,q}(\Omega)} \qquad \forall v \in H_0^1(\Omega).$$

Therefore, the functional $\mathcal{A}z \in H^{-1}(\Omega)$ can be extended continuously to a functional from $W^{-1,q'}(\Omega)$. Note that we cannot use integration by parts on the left-hand side of (2.8), since this would require $p \in H_0^1(\Omega)$ or $z \in W_0^{1,q'}(\Omega)$. This, however, may not hold under the low regularity (2.2). For a thorough discussion of the interpretation of the adjoint equation in the case of coefficients with low regularity, we refer to [Meyer et al., 2011].

Of course, it is also possible to discuss the adjoint equation with right-hand sides from $H^{-1}(\Omega)$, i.e.,

$$p \in H_0^1(\Omega), \qquad \mathcal{A}^\star p = \mu \quad \text{in } H^{-1}(\Omega).$$

Existence and uniqueness follows from Lax-Milgram. Due to $Z \subset H_0^1(\Omega) \cap C_0(\Omega)$, both notions of solutions coincide if $\mu \in H^{-1}(\Omega) \cap \mathcal{M}(\Omega)$. Indeed, for every $z \in Z$ and $p \in H_0^1(\Omega)$, we have

$$\langle \mathcal{A}z, p\rangle_{W_0^{1,q}(\Omega)} = \langle \mathcal{A}z, p\rangle_{H_0^1(\Omega)} \quad \text{and} \quad \int_\Omega z \,\mathrm{d}\mu = \langle \mu, z\rangle_{H_0^1(\Omega)}.$$

Thus, we can define a very weak solution $p \in W_0^{1,q}(\Omega)$ of

$$\mathcal{A}^\star p = \mu + \nu$$

for $\mu \in H^{-1}(\Omega)$ and $\nu \in \mathcal{M}(\Omega)$ via

$$\langle \mathcal{A}z, p\rangle_{W_0^{1,q}(\Omega)} = \langle \mu, z\rangle_{H_0^1(\Omega)} + \int_\Omega z \,\mathrm{d}\nu \qquad \forall z \in Z, \tag{2.12}$$

and this solution does not depend on the precise splitting of $\mu + \nu$ into $\mu \in H^{-1}(\Omega)$ and $\nu \in \mathcal{M}(\Omega)$.

Let us consider further properties and auxiliary results, starting with smooth multipliers for the space Z.

Lemma 2.4. *Let $z \in H_0^1(\Omega)$ be given such that $\mathcal{A}z \in W^{-1,q'}(\Omega)$ for some $q' > n$ with $q' < \infty$ in case $n = 2$ and $q' \le 6$ in case $n = 3$. Then, for all $\psi \in C^\infty(\Omega)$ we have $\psi z \in H_0^1(\Omega)$ and $\mathcal{A}(\psi z) \in W^{-1,q'}(\Omega)$.*

Proof. Theorem 2.2 implies $z \in L^\infty(\Omega)$. Further, we already know that the linear functional $v \mapsto \langle \mathcal{A}z, \psi v\rangle$ belongs to $W^{-1,q'}(\Omega)$. Now we consider

$$\begin{aligned}
&\langle \mathcal{A}(\psi z), v\rangle - \langle \mathcal{A}z, \psi v\rangle \\
&\quad = \int_\Omega \nabla v^\top A \nabla(\psi z) + (\psi z)\,(b^\top \nabla v) + v\,(c^\top \nabla(\psi z)) + d\,(\psi z) v \,\mathrm{d}x \\
&\qquad\quad - \int_\Omega \nabla(\psi v)^\top A \nabla z + z\,(b^\top \nabla(\psi v)) + (\psi v)\,(c^\top \nabla z) + dz\,(\psi v)\,\mathrm{d}x \\
&\quad = \int_\Omega z \nabla v^\top A \nabla \psi - v \nabla \psi^\top A \nabla z - zv\,(b^\top \nabla \psi) + vz\,(c^\top \nabla \psi)\,\mathrm{d}x.
\end{aligned}$$

This yields the bound

$$\left|\langle \mathcal{A}(\psi z), v\rangle - \langle \mathcal{A}z, \psi v\rangle\right| \le C\big(\|\nabla v\|_{L^1(\Omega)} + \|v\|_{L^2(\Omega)} + \|v\|_{L^1(\Omega)} + \|v\|_{L^1(\Omega)}\big).$$

The Sobolev embedding theorem yields $W_0^{1,s}(\Omega) \hookrightarrow L^2(\Omega)$ for $1/s = 1/2 + 1/n$, i.e., $s = 2n/(n+2)$. Thus, $s = 6/5$ for $n = 3$. Due to the assumption on q', we have $q' \le s'$, and therefore $\mathcal{A}(\psi z) \in W^{-1,q'}(\Omega)$. $\square$

Lemma 2.5. *Let subsets $U, V \subset \Omega$ be given, such that U is compact, V is open and $U \subset V$. Let $f \in H^{-1}(\Omega)$ and $\varphi \in C^\infty(\mathbb{R}^n)$ with $\varphi_{|V} = 0$. Then the weak solution $y \in H_0^1(\Omega)$ of*

$$\mathcal{A}y = \varphi f$$

is continuous in a neighborhood of U and fulfills

$$\|y\|_{C(U)} \le C\|f\|_{H^{-1}(\Omega)}. \tag{2.13}$$

Moreover, the mapping $H^{-1}(\Omega) \ni f \mapsto y \in C(U)$ is compact.

Proof. Choose $\psi \in C_c^\infty(\Omega)$ with $\psi_{|U} = 1$, and $\psi_{|\Omega\setminus V} = 0$. Then the product of φ and ψ vanishes, and we observe

$$\langle \mathcal{A}(\psi y), v\rangle = \langle \mathcal{A}(\psi y), v\rangle - \langle \varphi y, \psi v\rangle = \langle \mathcal{A}(\psi y), v\rangle - \langle \mathcal{A}y, \psi v\rangle.$$

Similar to the proof of Lemma 2.4 we therefore obtain

$$\begin{aligned}
\langle \mathcal{A}(\psi y), v\rangle &= \int_\Omega y \nabla v^\top A \nabla \psi - v \nabla \psi^\top A \nabla y - yv\,(b^\top \nabla \psi) + vy\,(c^\top \nabla \psi)\,\mathrm{d}x \\
&\le C\|y\|_{L^6(\Omega)}\|\nabla v\|_{L^{6/5}(\Omega)} + C\|\nabla y\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)} + C\|y\|_{L^6(\Omega)}\|v\|_{L^{6/5}(\Omega)}, \\
&\le C\|y\|_{H_0^1(\Omega)}\|v\|_{W^{1,6/5}(\Omega)},
\end{aligned}$$

where we have used the embeddings $H_0^1(\Omega) \hookrightarrow L^6(\Omega)$ for y as well as $W_0^{1,6/5}(\Omega) \hookrightarrow L^2(\Omega)$ in the case $n \le 3$ as before. With Theorem 2.1 we deduce

$$\langle \mathcal{A}(\psi y), v\rangle \le C\|f\|_{H^{-1}(\Omega)}\|v\|_{W^{1,6/5}(\Omega)},$$

hence $\|\mathcal{A}(\psi y)\|_{W^{-1,6}(\Omega)} \le C\|f\|_{H^{-1}(\Omega)}$. Finally, this yields

$$\|\psi y\|_{C_0^{0,\alpha}(\Omega)} \le C\|f\|_{H^{-1}(\Omega)}$$

by Theorem 2.2. Note that due to the compact embedding of $C_0^{0,\alpha}(\Omega) \hookrightarrow C_0(\Omega)$ the mapping $f \mapsto \psi y$ is compact from $H^{-1}(\Omega)$ into $C_0(\Omega)$, which concludes the proof. □

By a duality argument, we obtain a regularity result for the adjoint equation.

Lemma 2.6. *Let subsets $U, V \subset \Omega$ be given, such that U is compact, V is open and $U \subset V$. Let $\mu \in \mathcal{M}(\Omega)$ with $\operatorname{supp}(\mu) \subset U$ and $\varphi \in C^\infty(\mathbb{R}^n)$ with and $\varphi_{|V} = 0$. Then the very weak solution $p \in W_0^{1,q}(\Omega)$ of*

$$\mathcal{A}^\star p = \mu$$

fulfills

$$\|\varphi p\|_{H_0^1(\Omega)} \le C\|\mu\|_{\mathcal{M}(\Omega)}. \tag{2.14}$$

Moreover, if $\mu_k \overset{\star}{\rightharpoonup} \mu$, then $\varphi p_k \to \varphi p$ in $H_0^1(\Omega)$, where p_k is the very weak solution for the right-hand side μ_k.

Proof. We test the very weak formulation with the solution $y \in H_0^1(\Omega) \cap C_0^{0,\alpha}(\Omega)$ of $\mathcal{A}y = \varphi f$ for $f \in L^2(\Omega)$, and obtain

$$\int_\Omega p\varphi f \,\mathrm{d}x = \int_\Omega p\,(\mathcal{A}y)\,\mathrm{d}x = \int_\Omega y\,\mathrm{d}\mu \le \|y\|_{C(U)}\|\mu\|_{\mathcal{M}(\Omega)}.$$

Note that the last estimate uses $\operatorname{supp}(\mu) \subset U$. Applying Lemma 2.5 yields

$$\int_\Omega p\varphi f \,\mathrm{d}x \le C\|f\|_{H^{-1}(\Omega)}\|\mu\|_{\mathcal{M}(\Omega)}.$$

Since $f \in L^2(\Omega)$ was arbitrary, this yields $\|\varphi p\|_{H_0^1(\Omega)} \le C\|\mu\|_{\mathcal{M}(\Omega)}$.

It remains to verify the compactness property. As in the beginning of the proof, we have

$$\langle f, \varphi\, p\rangle_{H_0^1(\Omega)} = \langle \mu, y\rangle_{C(U)}$$

for all $f \in L^2(\Omega)$, $\mu \in \mathcal{M}(U)$, where $\mathcal{A}y = \varphi f$ and $\mathcal{A}^\star p = \mu$ (in the very weak sense). Using the density of $L^2(\Omega)$ in $H^{-1}(\Omega)$ and Lemma 2.5, this equation extends to all $f \in H^{-1}(\Omega)$. Therefore, the mapping $\mathcal{M}(U) \ni \mu \mapsto \varphi\, p \in H_0^1(\Omega)$ is the adjoint of the mapping $H^{-1}(\Omega) \ni f \mapsto y \in C(U)$ from Lemma 2.5. Now, we can argue as in the proof of Theorem 2.3. □

2.3 Solution operator of the obstacle problem

In this section, we give some properties of the solution operator of the variational inequality (VI)

$$\text{Find } y \in K \quad \text{such that } \langle \mathcal{A}y - u, v - y\rangle \ge 0 \quad \forall v \in K \tag{\textbf{VI}}$$

which appears as a constraint in (**P**). Here

$$K := \{v \in H_0^1(\Omega) \mid v \geq y_a \quad \text{a.e. in } \Omega\}.$$

We assume the same regularity of $\mathcal{A}$ and Ω as in the previous section. We further suppose that $y_a \leq 0$ on Γ in the sense of $H^1(\Omega)$ and this implies $K \neq \emptyset$. First of all, it is well-known that this VI admits a unique solution $y \in H_0^1(\Omega)$ for each $u \in H^{-1}(\Omega)$, see [Kinderlehrer and Stampacchia, 1980, Theorem II.2.1] or [Troianiello, 1987, Theorem 4.4]. The solution operator is denoted by

$$S\colon H^{-1}(\Omega) \to H_0^1(\Omega), \quad u \mapsto y.$$

It is known that (**VI**) is equivalent to the existence of $\xi \in H^{-1}(\Omega)$ such that

$$\mathcal{A}y = u - \xi, \quad \xi \in \mathcal{N}_K(y).$$

Here, $\mathcal{N}_K(y)$ is the normal cone of the convex set K.

Next, we address Hölder regularity of the solutions.

Theorem 2.7. *We assume that the obstacle y_a satisfies $y_a \in H^1(\Omega)$ and $\mathcal{A}y_a \in L^2(\Omega)$. Then, for any $u \in L^2(\Omega)$, we have $\mathcal{A}y, \xi \in L^2(\Omega)$, where $y := S(u)$, $\xi := u - \mathcal{A}y$. Moreover, $S\colon L^2(\Omega) \to C_0^{0,\alpha}(\Omega)$ is continuous for some $\alpha \in (0,1)$.*

Proof. We can apply [Troianiello, 1987, Theorem 4.32] and obtain the pointwise a.e. inequality $u \leq \mathcal{A}y \leq \max\{\mathcal{A}y_a, u\}$, where $u \in L^2(\Omega)$ is arbitrary and $y = S(u)$. This implies

$$\|\mathcal{A}y\|_{L^2(\Omega)} \leq \|u\|_{L^2(\Omega)} + \|\mathcal{A}y_a\|_{L^2(\Omega)},$$

i.e., $u \mapsto \mathcal{A}y$ is bounded from $L^2(\Omega)$ to $L^2(\Omega)$.

Now, let a sequence with $u_k \to u$ in $L^2(\Omega)$ be given. We set $y_k := S(u_k)$. Since $\mathcal{A}y_k$ is bounded in $L^2(\Omega)$ and converges in $H^{-1}(\Omega)$ to $\mathcal{A}y$, we get $\mathcal{A}y_k \rightharpoonup \mathcal{A}y$ in $L^2(\Omega)$.

Since the embedding from $W_0^{1,5/4}(\Omega)$ into $L^2(\Omega)$ is compact ([Gilbarg and Trudinger, 2001, Theorem 7.22]), the adjoint embedding from $L^2(\Omega)$ into $W^{-1,5}(\Omega) = W_0^{1,5/4}(\Omega)^\star$ is compact as well (by Schauder's theorem). Hence, $\mathcal{A}y_k \to \mathcal{A}y$ in $W^{-1,5}(\Omega)$. Finally, $\mathcal{A}^{-1}$ is continuous from $W^{-1,5}(\Omega)$ to $C_0^{0,\alpha}(\Omega)$, see Theorem 2.2. □

Using the standard truncation idea due to Stampacchia, we can also show the Lipschitz continuity of S w.r.t. a weaker norm on the forces u.

Lemma 2.8. *Let $q' > n$ be given. Then, there exists $C > 0$ such that*

$$\|S(u_1) - S(u_2)\|_{L^\infty(\Omega)} \leq C\|u_1 - u_2\|_{W^{-1,q'}(\Omega)} \qquad \forall u_1, u_2 \in W^{-1,q'}(\Omega).$$

Proof. Let $u_1, u_2 \in W^{-1,q'}(\Omega)$ be given and set $y_j := S(u_j)$ for $j = 1, 2$. For $k > 0$, we define $\hat{y}_k := (y_1 - y_2 - k)^+$. Due to $y_1 - \hat{y}_k = \min(y_2 + k, y_1) \geq y_a$, we can test the VIs for y_1 and y_2 with $y_1 - \hat{y}_k$ and $y_2 + \hat{y}_k$, respectively. Adding the resulting inequalities leads to

$$\langle \mathcal{A}(y_1 - y_2), \hat{y}_k\rangle \leq \langle u_1 - u_2, \hat{y}_k\rangle \qquad \forall k > 0.$$

Now, we can use the arguments of [Troianiello, 1987, Lemma 2.8] (see also the remark following this lemma) to obtain

$$\operatorname{ess\,sup}(y_1 - y_2) \le C\|u_1 - u_2\|_{W^{-1,p'}(\Omega)}.$$

Note that we can avoid the $L^2(\Omega)$-norm of $y_1 - y_2$ on the right-hand side of this estimate due to the Lipschitz-continuity of S from $H^{-1}(\Omega)$ to $H_0^1(\Omega)$. By interchanging the roles of y_1 and y_2, we arrive at the claimed estimate. □

An important property is the monotonicity

$$u_1 \le u_2 \quad \Rightarrow \quad S(u_1) \le S(u_2), \tag{2.15}$$

see, e.g., [Troianiello, 1987, Corollary, p. 242].

From the seminal work of Mignot, we get the directional differentiability of the mapping $S\colon H^{-1}(\Omega) \to H_0^1(\Omega)$, see [Mignot, 1976, Théorème 3.3].

Theorem 2.9. *The solution operator $S\colon H^{-1}(\Omega) \to H_0^1(\Omega)$ is directionally differentiable at all points $\bar u \in H^{-1}(\Omega)$. The directional derivative $z := S'(\bar u; h) \in H_0^1(\Omega)$ in direction $h \in H^{-1}(\Omega)$ is given by the unique solution of the VI*

$$z \in \mathcal{K}(\bar u), \qquad \langle \mathcal{A}z - h, v - z\rangle \ge 0 \quad \forall v \in \mathcal{K}(\bar u). \tag{2.16}$$

Here,

$$\mathcal{K}(\bar u) := \mathcal{T}_K(\bar y) \cap \bar\xi^\perp,$$

where $\bar y = S(\bar u)$ and $\bar\xi := \bar u - \mathcal{A}\bar y$ are the associated state and multiplier, respectively. For the critical cone $\mathcal{K}(\bar u)$, we have the representation

$$\mathcal{K}(\bar u) = \{v \in H_0^1(\Omega) \mid v \ge 0 \text{ q.e. on } \{\bar y = y_a\} \text{ and } v = 0 \text{ q.e. on } \operatorname{q-supp}(\bar\xi)\}. \tag{2.17}$$

The formula for the critical cone involving the quasi-support of $\bar\xi$ can be found in [G. Wachsmuth, 2014, Lemma 3.1].

Since $S\colon H^{-1}(\Omega) \to H_0^1(\Omega)$ is Lipschitz continuous, we obtain that S is even Hadamard differentiable, i.e., $(S(\bar u + t_k h_k) - S(\bar u))/t_k \to S'(\bar u; h)$ if $h_k \to h$ in $H^{-1}(\Omega)$ and $t_k \searrow 0$.

We note that the monotonicity (2.15) implies

$$h_1 \le h_2 \quad \Rightarrow \quad S'(u; h_1) \le S'(u; h_2) \tag{2.18}$$

for the directional derivative.

The next lemma shows that the difference quotients converge uniformly on the set where $\bar y$ has a positive distance from the lower bound y_a.

Lemma 2.10. *Let the assumptions of Theorem 2.7 and, additionally, $y_a \in C(\bar\Omega)$ be satisfied. For $\bar u \in L^2(\Omega)$, we define the state $\bar y := S(\bar u)$ and the set*

$$\hat\Omega := \{\bar y \ge y_a + \sigma\},$$

where $\sigma > 0$ is arbitrary and we use the continuous representatives of $\bar y$ and y_a.

(a) For an arbitrary $q' > n$, there exists a constant $C > 0$, such that

$$\|S'(\bar{u};h_1) - S'(\bar{u};h_2)\|_{L^\infty(\hat{\Omega})} \le C\|h_1 - h_2\|_{W^{-1,q'}(\Omega)} \qquad \forall h_1, h_2 \in W^{-1,q'}(\Omega).$$

(b) Let $\tilde{\varphi} \in C^\infty(\mathbb{R}^n)$ be given such that $\tilde{\varphi}$ vanishes on a neighborhood of $\hat{\Omega}$. Then,

$$\|S'(\bar{u};\tilde{\varphi}h_1) - S'(\bar{u};\tilde{\varphi}h_2)\|_{L^\infty(\hat{\Omega})} \le C\|h_1 - h_2\|_{H^{-1}(\Omega)} \qquad \forall h_1, h_2 \in H^{-1}(\Omega).$$

(c) Let sequences $(h_k)_{k\in\mathbb{N}} \subset L^2(\Omega)$ and $(t_k)_{k\in\mathbb{N}} \subset \mathbb{R}^+$ be given such that $h_k \to h$ in $W^{-1,q'}(\Omega)$ and $t_k \searrow 0$. We define the perturbed states $y_k := S(\bar{u} + t_k h_k)$. Then, the difference quotients $(y_k - \bar{y})/t_k$ converge towards $S'(\bar{u};h)$ uniformly on the set $\hat{\Omega}$ as $k \to \infty$. In particular, $S'(\bar{u};h)$ is continuous on $\hat{\Omega}$.

Proof. Note that we get continuity of $\bar{y}$ from Theorem 2.7 and y_a is continuous by assumption. Thus, the sets $\hat{\Omega}$ and

$$\hat{\Omega}_2 := \{\bar{y} \le y_a + \sigma/2\}$$

are closed. In the sequel, we are going to apply the regularity result Lemma 2.5 with $U = \hat{\Omega}$. Therefore, we fix a function $\varphi \in C^\infty(\mathbb{R}^n)$, such that $0 \le \varphi \le 1$ on $\mathbb{R}^n$, $\varphi = 0$ on a neighborhood V of $U = \hat{\Omega}$ and $\varphi = 1$ on $\hat{\Omega}_2$. Note that this is possible, since the sets $\hat{\Omega}, \hat{\Omega}_2$ have a positive distance.

We start with (a). Let $h_1, h_2 \in W^{-1,q'}(\Omega)$ be given. We define the functional

$$\hat{\xi} := (h_1 - \mathcal{A}S'(\bar{u};h_1)) - (h_2 - \mathcal{A}S'(\bar{u};h_2)) \in \mathcal{K}(\bar{u})^\circ,$$

see (2.16). Note that

$$\|\hat{\xi}\|_{H^{-1}(\Omega)} \le \|h_1 - h_2\|_{H^{-1}(\Omega)}.$$

For arbitrary $v \in H^1_0(\Omega)$, we have $(1-\varphi)v = 0$ q.e. on $\hat{\Omega}_2 \supset \{\bar{y} = y_a\}$. Thus, $\pm(1-\varphi)v \in \mathcal{K}(\bar{u})$. Thus, $\langle (1-\varphi)\hat{\xi}, v\rangle = 0$, i.e., $\hat{\xi} = \varphi\hat{\xi}$. By Theorem 2.2 and Lemma 2.5, we get

$$\begin{aligned}\|S'(\bar{u};h_1) - S'(\bar{u};h_2)\|_{L^\infty(\hat{\Omega})} &= \|\mathcal{A}^{-1}(\varphi\hat{\xi} - h_1 + h_2)\|_{L^\infty(\hat{\Omega})} \\ &\le C\Big(\|\hat{\xi}\|_{H^{-1}(\Omega)} + \|h_1 - h_2\|_{W^{-1,q'}(\Omega)}\Big) \\ &\le C\Big(\|h_1 - h_2\|_{H^{-1}(\Omega)} + \|h_1 - h_2\|_{W^{-1,q'}(\Omega)}\Big)\end{aligned}$$

and the assertion follows.

The proof of (b) is very similar. One just has to replace Theorem 2.2 by another application of Lemma 2.5.

Next, we show (c). Since $S\colon W^{-1,q'}(\Omega) \to L^\infty(\Omega)$ is Lipschitz by Lemma 2.8, there exists $N \in \mathbb{N}$ such that

$$y_k(x) > y_a(x) \qquad \forall x \in \Omega \setminus \hat{\Omega}_2, k \ge N.$$

Hence, the associated multiplier $\xi_k := (\bar{u} + t_k h_k) - \mathcal{A}y_k \in L^2(\Omega)$ is supported on the set $\hat{\Omega}_2$. Now, we consider the difference quotient of the multipliers $\hat{\xi}_k := (\xi_k - \bar{\xi})/t_k$, where $\bar{\xi} := \bar{u} - \mathcal{A}\bar{y}$. This difference quotient is supported on $\hat{\Omega}_2$ and converges in $H^{-1}(\Omega)$ towards $\hat{\xi} := h - \mathcal{A}S'(\bar{u}; h)$, see Theorem 2.9. This implies $\hat{\xi}_k = \varphi\hat{\xi}_k \to \varphi\hat{\xi} = \hat{\xi}$.

Thus, we can apply Lemma 2.5 and obtain

$$\left\|\mathcal{A}^{-1}\hat{\xi}_k - \mathcal{A}^{-1}\hat{\xi}\right\|_{L^\infty(\hat{\Omega})} = \|\mathcal{A}^{-1}\Big(\varphi(\hat{\xi}_k - \hat{\xi})\Big)\|_{L^\infty(\hat{\Omega})} \le \|\hat{\xi}_k - \hat{\xi}\|_{H^{-1}(\Omega)} \to 0.$$

Together with

$$\frac{y_k - \bar{y}}{t_k} = \mathcal{A}^{-1}h + \mathcal{A}^{-1}\hat{\xi}_k \qquad \text{and} \qquad S'(\bar{u}; h) = \mathcal{A}^{-1}h + \mathcal{A}^{-1}\hat{\xi},$$

this shows the claim. □

A well-known property of the obstacle problem with a lower bound is the pointwise convexity of the solution operator. This renders the state constraint convex w.r.t. the control u and will become important for our analysis of the optimal control problem.

Lemma 2.11. *Let $u_1, u_2 \in H^{-1}(\Omega)$ and $\alpha \in (0, 1)$ be given. Then,*

$$S(\alpha u_1 + (1-\alpha)u_2) \le \alpha S(u_1) + (1-\alpha)S(u_2) \qquad \text{a.e. in } \Omega.$$

Proof. For convenience, we reproduce the short proof from [Mignot, 1976, Lemme 4.1].

We set $u_3 := \alpha u_1 + (1-\alpha)u_2$ and $y_i := S(u_i)$ for $i = 1, \ldots, 3$. We have to show that $w := (y_3 - \alpha y_1 - (1-\alpha)y_2)^+$ is zero. This definition directly implies $w \ge 0$ and $y_3 - w = \min\{y_3, \alpha y_1 + (1-\alpha)y_2\} \ge y_a$. Thus, we obtain

$$\begin{aligned}
\langle \mathcal{A}y_1 - u_1, w\rangle &= \langle \mathcal{A}y_1 - u_1, y_1 + w - y_1\rangle \ge 0,\\
\langle \mathcal{A}y_2 - u_2, w\rangle &= \langle \mathcal{A}y_2 - u_2, y_2 + w - y_2\rangle \ge 0,\\
\langle \mathcal{A}y_3 - u_3, -w\rangle &= \langle \mathcal{A}y_3 - u_3, y_3 - w - y_3\rangle \ge 0.
\end{aligned}$$

We multiply the first two inequalities by α and $(1-\alpha)$, respectively. Adding the three resulting inequalities yields

$$\langle \mathcal{A}(y_3 - \alpha y_1 - (1-\alpha)y_2), w\rangle \le 0.$$

The structure of the differential operator gives

$$\langle \mathcal{A}v^+, v^+\rangle = \langle \mathcal{A}v, v^+\rangle \qquad \forall v \in H_0^1(\Omega)$$

and together with the coercivity of $\mathcal{A}$ we obtain $w = 0$. □

From this pointwise convexity, we obtain two inequalities for the directional derivative by the usual arguments.

Corollary 2.12. *Let $u, h, h_2 \in H^{-1}(\Omega)$ be given. Then,*

$$\begin{aligned}
S(u) + S'(u; h) &\le S(u + h) \qquad \text{a.e. in } \Omega,\\
S'(u; h + h_2) &\le S'(u; h) + S'(u; h_2) \qquad \text{a.e. in } \Omega.
\end{aligned}$$

Proof. For any $t \in (0,1)$ we have

$$S(u+th) = S((1-t)u + t(u+h)) \leq (1-t)S(u) + tS(u+h).$$

Now, we subtract $S(u)$, divide by $t > 0$ and pass to the limit $t \searrow 0$ to arrive at the first assertion. The second assertion follows similarly by considering

$$2\left(S(u + t\,(h+h_2)) - S(u)\right) \leq S(u + t\,(2h)) - S(u) + S(u + t\,(2h_2)) - S(u),$$

dividing by $t > 0$, passing to the limit $t \searrow 0$ and using the positive homogeneity of $S'(u;\cdot)$. □

2.4 Optimal control problem

We will now discuss the optimal control problem (**P**). To this end, let us collect all the assumptions which have been made in the previous preliminary results. Additionally, we make further assumptions concerning the optimal control problem, in particular we will assume the existence of a Slater point, from which we will eventually deduce existence of a Lagrange multiplier associated with the state constraints.

Assumption 2.13.

(i) *The domain $\Omega \subset \mathbb{R}^n$, $n \in \{2,3\}$, is bounded and satisfies the uniform exterior cone condition, see [Gilbarg and Trudinger, 2001, p. 205].*

(ii) *The differential operator $\mathcal{A}$ is given as in (2.1), such that (2.2)–(2.4) are satisfied.*

(iii) *The obstacle $y_a \in H^1(\Omega) \cap C(\bar{\Omega})$ satisfies $y_a \leq 0$ on Γ in the sense $\max\{y_a, 0\} \in H^1_0(\Omega)$ and $\mathcal{A}y_a \in L^2(\Omega)$.*

(iv) *The state constraint has the regularity $y_b \in C(\bar{\Omega})$ and satisfies $y_b > 0$ on Γ.*

(v) *The control set $U_{\mathrm{ad}} \subset L^2(\Omega)$ is convex, closed, and non-empty.*

(vi) *The objective $J\colon H^1_0(\Omega) \times L^2(\Omega) \to \mathbb{R}$ is assumed to be continuously Fréchet-differentiable and bounded from below. We require that J is sequentially lower semi-continuous w.r.t. to the strong topology in $H^1_0(\Omega)$ and the weak topology in $L^2(\Omega)$, that is $J(y,u) \leq \liminf_{k\to\infty} J(y_k,u_k)$ for all sequences $\left((y_k,u_k)\right)_{k\in\mathbb{N}} \subset H^1_0(\Omega) \times L^2(\Omega)$ satisfying $y_k \to y$ in $H^1_0(\Omega)$ and $u_k \rightharpoonup u$ in $L^2(\Omega)$. Finally, we assume that J is coercive w.r.t. the second variable on the feasible set U_{ad}, that is the boundedness of $(u_k)_{k\in\mathbb{N}}$ in $L^2(\Omega)$ follows from the boundedness of $(J(y_k,u_k))_{k\in\mathbb{N}}$ for all sequences $\left((y_k,u_k)\right)_{k\in\mathbb{N}} \subset H^1_0(\Omega) \times U_{\mathrm{ad}}$.*

(vii) *There exists a Slater point $\hat{u} \in U_{\mathrm{ad}}$ with $S(\hat{u}) \leq y_b - \tau$ on Ω for some $\tau > 0$.*

Due to the Slater condition, we have $y_a \leq y_b - \tau$ on Ω. In the case without control constraints, this apparently weaker condition already implies the Slater condition.

Lemma 2.14. *We assume Assumption 2.13 (i)–(iv). If, additionally, $y_a \leq y_b - \tau$ on Ω for some $\tau > 0$, $y_a < 0$ on Γ, and $U_{\text{ad}} = L^2(\Omega)$, then Assumption 2.13 (vii) follows.*

Proof. Since y_a, y_b are continuous, $y_b > 0 > y_a$ on Γ, and $y_b - y_a \geq \tau > 0$, it is possible to construct an arbitrarily smooth function $\tilde{y} \geq y_a$ with $\tilde{y} = 0$ on Γ and positive distance to y_b. Now we define $\tilde{u} := \mathcal{A}\tilde{y} \in W^{-1,q'}(\Omega)$, $q' > n$. By a density argument, we can smooth $\tilde{u}$ and construct $\hat{u} \in L^2(\Omega)$ such that, with $\hat{y} = S(\hat{u})$, Lemma 2.8 guarantees $\|\tilde{y} - \hat{y}\|_{L^\infty(\Omega)} \leq \varepsilon$ for any fixed, positive ε. Therefore, choosing $\hat{u}$ corresponding to $\varepsilon > 0$ small enough, $\hat{y}$ has positive distance to y_b, meaning that $\hat{u}$ fulfills the Slater point property. □

Note that one cannot expect $\hat{u} \in U_{\text{ad}}$ if $U_{\text{ad}} \neq L^2(\Omega)$.

If the admissible set U_{ad} has a minimal point, i.e., $u_b \in U_{\text{ad}}$ with $u \geq u_b$ for all $u \in U_{\text{ad}}$, then there exists a Slater point if and only if u_b is a Slater point. This follows easily from the monotonicity of S, see (2.15).

From now on, we will always assume that Assumption 2.13 is satisfied.

The existence of the Slater point $\hat{u}$ will not only be used to show optimality conditions. As a side effect, it also guarantees the existence of solutions.

Theorem 2.15. *There exists at least one globally optimal control $\bar{u} \in U_{\text{ad}}$ to (**P**) with associated optimal state $\bar{y} \in H_0^1(\Omega)$.*

Proof. Due to Assumption 2.13 (vii), there exists a feasible pair $(\hat{y}, \hat{u})$. Then, we infer the existence of an optimal solution $(\bar{y}, \bar{u})$ to Problem (**P**) by standard arguments. □

Note that due to the nonlinearity of the solution operator S, one cannot show uniqueness of the solution.

3 Primal optimality conditions

In this section, we address necessary optimality conditions for (**P**) which do not involve dual quantities. We start by masking the state constraint as a convex control constraint via Lemma 2.11.

Lemma 3.1. *We define*

$$\begin{aligned} U_{\text{state}} &:= \{u \in L^2(\Omega) \mid S(u) \leq y_b \text{ in } \Omega\}, \\ U_{\text{eff}} &:= U_{\text{ad}} \cap U_{\text{state}} = \{u \in U_{\text{ad}} \mid S(u) \leq y_b \text{ in } \Omega\}. \end{aligned}$$

The sets $U_{\text{state}}, U_{\text{eff}} \subset L^2(\Omega)$ are closed and convex.

Proof. The convexity follows from Lemma 2.11 and the closedness from the continuity of $S\colon L^2(\Omega) \to C_0^{0,\alpha}(\Omega)$, see Theorem 2.7. □

Using this result, we can reformulate ($\mathbf{P}$) and obtain the equivalent problem

$$\begin{aligned} \text{Minimize} \quad & J(S(u),u) \\ \text{with respect to} \quad & (y,u) \in H_0^1(\Omega) \times L^2(\Omega) \\ \text{such that} \quad & u \in U_{\text{eff}}. \end{aligned} \tag{3.1}$$

Since the set U_{eff} is closed and convex, we could apply [G. Wachsmuth, 2016, Theorem 1.1] to obtain a system of C-stationarity. This system contains the normal cone to U_{eff} in $L^2(\Omega)$ and it is not immediately clear how to evaluate this normal cone. After we have characterized the normal cone in Lemma 5.9, we comment on this approach at the end of Section 5.

Another possibility is to use the directional differentiability of S to arrive at a primal optimality system.

Lemma 3.2. *Let $\bar{u}$ be locally optimal for ($\mathbf{P}$) with associated state $\bar{y} = S(\bar{u})$. Then,*

$$\langle J_y(\bar{y},\bar{u}), S'(\bar{u};h)\rangle + (J_u(\bar{y},\bar{u}),h) \geq 0 \qquad \forall h \in \mathcal{T}_{U_{\text{eff}}}(\bar{u})$$

is necessary for the optimality of $\bar{u}$.

Here, $J_y(\bar{y},\bar{u}) \in H^{-1}(\Omega)$ and $J_u(\bar{y},\bar{u}) \in L^2(\Omega)$ are the partial derivatives of J w.r.t. y and u evaluated at $(\bar{y},\bar{u})$.

Proof. For any $h \in \mathcal{R}_{U_{\text{eff}}}(\bar{u})$, we have $u + th \in U_{\text{eff}}$ for $t > 0$ small enough. Thus,

$$J(S(u+th), u+th) - J(S(u),u) \geq 0.$$

Dividing by $t > 0$ and passing to the limit $t \searrow 0$, we obtain

$$\langle J_y(\bar{y},\bar{u}), S'(\bar{u};h)\rangle + (J_u(\bar{y},\bar{u}),h) \geq 0 \qquad \forall h \in \mathcal{R}_{U_{\text{eff}}}(\bar{u}).$$

Since $\mathcal{R}_{U_{\text{eff}}}(\bar{u})$ is dense in $\mathcal{T}_{U_{\text{eff}}}(\bar{u})$ and since the left-hand side of the inequality is continuous w.r.t. $u \in L^2(\Omega)$, the claim follows. $\square$

Our next goal is the characterization of the tangent cone of U_{eff}. We start by the investigation of the tangent cone of U_{state}.

Theorem 3.3. *Let $\bar{u} \in U_{\text{state}}$ be given. Then,*

$$\mathcal{T}_{U_{\text{state}}}(\bar{u}) = \{h \in L^2(\Omega) \mid S'(\bar{u};h) \leq 0 \text{ everywhere on } \Omega_b\}, \tag{3.2}$$

where $\Omega_b := \{\bar{y} = y_b\}$ with $\bar{y} = S(\bar{u})$.

Note that $S'(\bar{u};h)$ is continuous in the neighborhood $\{\bar{y} \geq y_a + \zeta/2\}$ of Ω_b via Lemma 2.10. Thus, the inequality can be understood in an "everywhere"-sense.

Proof. "$\subset$": Let $h \in \mathcal{T}_{U_{\mathrm{state}}}(\bar u)$ be given. Then, there are sequences $(u_k)_{k\in\mathbb{N}} \subset U_{\mathrm{state}}$, $(t_k)_{k\in\mathbb{N}} \subset \mathbb{R}^+$ such that $u_k \to \bar u$, $t_k \searrow 0$ and $h_k := (u_k - \bar u)/t_k \to h$ in $L^2(\Omega)$. We define $y_k := S(u_k) = S(\bar u + t_k h_k)$. Then, $0 \ge \frac{y_k - \bar y}{t_k}$ holds everywhere on Ω_b. Due to Lemma 2.10 (c), the right-hand side converges uniformly on Ω_b towards $S'(\bar u; h)$. Hence, $S'(\bar u; h) \le 0$ on Ω_b.

"$\supset$": Let $h \in L^2(\Omega)$ with $S'(\bar u; h) \le 0$ on Ω_b be given. In case $\Omega_b = \emptyset$, the continuous functions $\bar y$ and y_b have a positive distance. Therefore, the claim follows from the continuity of $S\colon L^2(\Omega) \to C_0^{0,\alpha}(\Omega)$, see Theorem 2.7.

Otherwise, let $(t_k)_{k\in\mathbb{N}} \subset \mathbb{R}^+$ with $t_k \searrow 0$ be given. Due to the continuity of S from $L^2(\Omega)$ to $C_0^{0,\alpha}(\Omega)$, there exists a sequence $(s_k)_{k\in\mathbb{N}} \subset \mathbb{R}^+$, $s_k \searrow 0$, such that

$$S(\bar u + t_k h) \le S(\bar u) + s_k \qquad \text{on } \Omega.$$

W.l.o.g. we assume $s_k \le \zeta/3$. Therefore, the sets

$$\Omega_k := \{\bar y \ge y_b - s_k\}, \quad \hat\Omega := \{\bar y \ge y_a + \zeta/3\}$$

satisfy $\Omega_b \subset \Omega_k \subset \hat\Omega$ for all $k \in \mathbb{N}$. We define the scalar sequence

$$d_k := \sup\{S'(\bar u; h)(x) \mid x \in \Omega_k\} \ge 0.$$

We claim that $d_k \searrow 0$. Indeed, otherwise we would get a sequence $(x_k)_{k\in\mathbb{N}}$ with $x_k \in \Omega_k$ and $S'(\bar u; h)(x_k) \ge \varepsilon > 0$. This sequence has accumulation points and due to continuity, all accumulation points $\bar x$ satisfy $S'(\bar u; h)(\bar x) \ge \varepsilon > 0$ and $\bar y(\bar x) \ge y_b(\bar x)$, i.e., $\bar x \in \Omega_b$. This is a contradiction to $S'(\bar u; h) \le 0$ on Ω_b.

Due to Lemma 2.10 (c),

$$r_k := \left\| \frac{S(\bar u + t_k h) - \bar y}{t_k} - S'(\bar u; h) \right\|_{C(\hat\Omega)} \searrow 0.$$

Now we have

$$\begin{aligned}
S(\bar u + t_k h) &\le S(\bar u) + t_k S'(\bar u; h) + t_k r_k \le y_b + t_k (d_k + r_k) && \text{on } \Omega_k \subset \hat\Omega,\\
S(\bar u + t_k h) &\le S(\bar u) + s_k \le y_b && \text{on } \Omega \setminus \Omega_k.
\end{aligned}$$

Next, we use the Slater point $\hat u \in L^2(\Omega)$, i.e., $S(\hat u) \le y_b - \tau$ for some $\tau > 0$. We set

$$h_k := (1 - \alpha_k) h + \frac{\alpha_k}{t_k} (\hat u - \bar u), \qquad \alpha_k := \frac{d_k + r_k}{\tau} t_k.$$

From $\alpha_k / t_k \to 0$ we get $h_k \to h$ in $L^2(\Omega)$. Moreover, for k large enough we have $\alpha_k \in (0, 1)$ and via Lemma 2.11 we obtain

$$\begin{aligned}
S(\bar u + t_k h_k) &= S((1 - \alpha_k)(\bar u + t_k h) + \alpha_k \hat u)\\
&\le (1 - \alpha_k) S(\bar u + t_k h) + \alpha_k S(\hat u)\\
&\le (1 - \alpha_k)(y_b + t_k (d_k + r_k)) + \alpha_k (y_b - \tau)\\
&= y_b + (1 - \alpha_k) t_k (d_k + r_k) - \alpha_k \tau\\
&\le y_b + t_k (d_k + r_k) - \alpha_k \tau = y_b \qquad \text{on } \Omega.
\end{aligned}$$

This shows $\bar u + t_k h_k \in U_{\mathrm{state}}$. Together with $h_k \to h$ in $L^2(\Omega)$ we get $h \in \mathcal{T}_{U_{\mathrm{state}}}(\bar u)$. $\square$

Using the Slater point again, we can characterize the tangent cone and normal cone to $U_{\text{eff}} = U_{\text{ad}} \cap U_{\text{state}}$.

Theorem 3.4. *Let $\bar u \in U_{\text{eff}}$ be given. Then,*

$$\mathcal{T}_{U_{\text{eff}}}(\bar u) = \mathcal{T}_{U_{\text{ad}}}(\bar u) \cap \mathcal{T}_{U_{\text{state}}}(\bar u), \qquad \mathcal{N}_{U_{\text{eff}}}(\bar u) = \mathcal{N}_{U_{\text{ad}}}(\bar u) + \mathcal{N}_{U_{\text{state}}}(\bar u).$$

Proof. Due to the continuity of $S\colon L^2(\Omega) \to C_0^{0,\alpha}(\Omega)$, see Theorem 2.7, the Slater point $\hat u$ is an interior point of U_{state} and belongs to U_{ad}. Hence, we can apply the sum rule of convex analysis [Bauschke and Combettes, 2011, Corollary 16.38] to the indicator function $\delta_{U_{\text{eff}}} = \delta_{U_{\text{ad}}} \cap \delta_{U_{\text{state}}}$ and obtain

$$\mathcal{N}_{U_{\text{eff}}}(\bar u) = \partial\delta_{U_{\text{eff}}}(\bar u) = \partial\delta_{U_{\text{ad}}}(\bar u) + \partial\delta_{U_{\text{state}}}(\bar u) = \mathcal{N}_{U_{\text{ad}}}(\bar u) + \mathcal{N}_{U_{\text{state}}}(\bar u).$$

The tangent cone can be obtained by polarization via the bipolar theorem. □

Together with Lemma 3.2, we obtain the following optimality condition.

Theorem 3.5. *Every locally optimal solution $\bar u$ of* (**P**) *satisfies*

$$\langle J_y(\bar y, \bar u), S'(\bar u; h)\rangle + (J_u(\bar y, \bar u), h) \ge 0 \qquad \forall h \in \mathcal{T}_{U_{\text{ad}}}(\bar u), S'(\bar u; h) \le 0 \text{ on } \Omega_b, \tag{3.3}$$

where $\Omega_b := \{\bar y = y_b\}$ and $\bar y := S(\bar u)$.

Although we have derived a characterization of the tangent cone of U_{state}, see Theorem 3.3, this cannot be employed to obtain an expression for the normal cone, due to the nonlinearity of $S'(\bar u; \cdot)$. Even if an explicit formula for this normal cone would be available, the primal optimality condition (3.3) cannot be turned directly into a dual optimality condition, since the left-hand side in (3.3) depends nonlinearly on h. We mention that a formula for $\mathcal{N}_{U_{\text{state}}}(\bar u)$ will be given in Lemma 5.9 below.

4 Dual optimality conditions via regularization

In this section, we are going to derive optimality conditions which include multipliers via a regularization procedure. We will prove the following theorem.

Theorem 4.1. *Every local solution $(\bar y, \bar u)$ of* (**P**) *is C-stationary, i.e., there exist multipliers $p \in W_0^{1,q}(\Omega)$, $\mu \in H^{-1}(\Omega)$, $\nu \in \mathcal{M}(\Omega)^+$, $\lambda \in L^2(\Omega)$ such that $p \in H^1(\hat\Omega_a)$ for some open $\hat\Omega_a \supset \{\bar y = y_a\}$ and such that the system*

$$\mathcal{A}^\star p + J_y(\bar y, \bar u) + \nu + \mu = 0, \tag{4.1a}$$

$$J_u(\bar y, \bar u) + \lambda - p = 0, \tag{4.1b}$$

$$p = 0 \text{ q.e. on } \operatorname{q-supp}(\bar\xi), \tag{4.1c}$$

$$\langle \mu, v\rangle_{H_0^1(\Omega)} = 0 \quad \forall v \in H_0^1(\Omega), v = 0 \text{ q.e. on } \{\bar y = y_a\}, \tag{4.1d}$$

$$\langle \mu, \Phi p\rangle_{H_0^1(\Omega)} \ge 0 \quad \forall \Phi \in W^{1,\infty}(\Omega)^+,\ \Phi_{|_{\Omega\setminus\hat\Omega_a}} = 0, \tag{4.1e}$$

$$\operatorname{supp}(\nu) \subset \Omega_b, \tag{4.1f}$$

$$\lambda \in \mathcal{N}_{U_{\text{ad}}}(\bar u) \tag{4.1g}$$

is satisfied. Here, the adjoint equation is to be understood in the very weak sense, see (2.12)*, and* $q \in (1, n/(n-1))$ *can be chosen arbitrarily. Note that* $\Phi p \in H_0^1(\Omega)$ *even though* p *itself is only in* $W_0^{1,p}(\Omega)$*.*

The proof of this theorem is divided into several steps, which will be addressed in the remaining part of this section:

- Section 4.1: Existence of solutions and optimality condition for regularized problems.
- Section 4.2: Boundedness of the multipliers of the regularized optimality system.
- Section 4.3: Passage to the limit in the optimality system.

Throughout the remaining part of this section, we fix a local solution $(\bar{y}, \bar{u})$ of (**P**).

4.1 Regularized problems

In order to derive an optimality condition for problem (**P**), we consider a regularization of the state constraint by penalization of any violation of the constraints, see [Ito and Kunisch, 2003]. Clearly, other regularization approaches would be viable as well, i.e., a regularization of the obstacle problem. For a regularization parameter $\gamma > 0$, define the regularized problem

$$\begin{aligned} \text{Minimize} \quad & J(y,u) + \frac{\gamma}{2}\|\max\{0, y - y_b\}\|^2_{L^2(\Omega)} + \frac{1}{2}\|u - \bar{u}\|^2_{L^2(\Omega)} \\ \text{with respect to} \quad & (y,u) \in H_0^1(\Omega) \times L^2(\Omega) \\ \text{such that} \quad & y \in K,\ \langle \mathcal{A}y - u, v - y\rangle \geq 0 \quad \forall v \in K, \\ \text{and} \quad & u \in U_{\text{ad}}. \end{aligned} \tag{$\mathbf{P}_\gamma$}$$

Note that the term $\frac{1}{2}\|u - \bar{u}\|^2_{L^2(\Omega)}$ in the regularized objective functional is necessary to prove convergence if $(\bar{y}, \bar{u})$ is not a strict local minimizer.

We proceed by proving that the minimizer $(\bar{y}, \bar{u})$ can be approximated by local solutions of the regularized problem ($\mathbf{P}_\gamma$).

Lemma 4.2. *There exists a sequence* $(\gamma_k)_{k\in\mathbb{N}}$ *with* $\gamma_k \to \infty$*, such that there exists a local solution* (y_k, u_k) *of* ($\mathbf{P}_\gamma$) *with* $\gamma = \gamma_k$ *for each* $k \in \mathbb{N}$ *and* $u_k \to \bar{u}$ *in* $L^2(\Omega)$*. Thus,* $y_k \to \bar{y}$ *in* $H_0^1(\Omega)$ *and* $\xi_k := u_k - \mathcal{A}y_k \to \bar{u} - \mathcal{A}\bar{y} =: \bar{\xi}$ *in* $H^{-1}(\Omega)$*.*

Proof. We use a meanwhile classical localization argument. Let $\delta > 0$ denote the radius of optimality of $\bar{u}$. We introduce the auxiliary problems

$$\begin{aligned} \text{Minimize} \quad & J(y,u) + \frac{\gamma}{2}\|\max\{0, y - y_b\}\|^2_{L^2(\Omega)} + \frac{1}{2}\|u - \bar{u}\|^2_{L^2(\Omega)} \\ \text{with respect to} \quad & (y,u) \in H_0^1(\Omega) \times L^2(\Omega) \\ \text{such that} \quad & y \in K,\ \langle \mathcal{A}y - u, v - y\rangle \geq 0 \quad \forall v \in K \\ \text{and} \quad & u \in U_{\text{ad}} \cap B_\delta(\bar{u}). \end{aligned} \tag{$\mathbf{P}_{\gamma,\delta}$}$$

By the usual arguments, these problems possess global solutions. Let $(\gamma_k)_{k\in\mathbb{N}}$ be an arbitrary sequence of positive numbers with $\gamma_k \to \infty$. We denote by (y_k, u_k) a global solution of problem ($\mathbf{P}_{\gamma,\delta}$) with $\gamma = \gamma_k$.

Since the sequence $(u_k)_{k\in\mathbb{N}}$ is bounded in $L^2(\Omega)$, we can extract a weakly convergent subsequence (denoted by the same symbol). We denote by $\tilde{u}$ the weak limit and by compact embedding, we have $y_k \to \tilde{y} := S(\tilde{u})$ in $H_0^1(\Omega)$. Since $(\bar{y}, \bar{u})$ is feasible for ($\mathbf{P}_{\gamma,\delta}$), we find

$$J(\bar{y}, \bar{u}) \geq J(y_k, u_k) + \frac{\gamma_k}{2}\|\max\{0, y_k - y_b\}\|^2_{L^2(\Omega)} + \frac{1}{2}\|u_k - \bar{u}\|^2_{L^2(\Omega)}. \tag{4.2}$$

From this inequality we infer that $\tilde{y} \leq y_b$. Hence, $(\tilde{y}, \tilde{u})$ is a feasible point for ($\mathbf{P}$) and by lower semicontinuity of J, we find

$$J(\bar{y}, \bar{u}) \geq J(\tilde{y}, \tilde{u}) + \frac{1}{2}\limsup_{k\to\infty}\|u_k - \bar{u}\|^2_{L^2(\Omega)} \geq J(\tilde{y}, \tilde{u}) + \frac{1}{2}\|\tilde{u} - \bar{u}\|^2_{L^2(\Omega)}.$$

Since $\tilde{u} \in B_\delta(\bar{u})$, we obtain

$$(\tilde{y}, \tilde{u}) = (\bar{y}, \bar{u}) \quad \text{and} \quad \|u_k - \bar{u}\|_{L^2(\Omega)} \to 0,$$

i.e., $u_k \to \bar{u}$ in $L^2(\Omega)$. Hence, the constraint $u_k \in B_\delta(\bar{u})$ is not active for large k and the result follows. □

The regularized problem ($\mathbf{P}_\gamma$) is a standard optimal control problem of the obstacle problem with control constraints and a differentiable objective function. Thus, we obtain a primal optimality condition similar to Lemma 3.2, i.e.,

$$0 \leq \langle J_y(y_k, u_k), S'(u_k, h)\rangle + \gamma\,(\max\{0, y_k - y_b\}, S'(u_k; h)) + (J_u(y_k, u_k), h) + (u_k - \bar{u}, h) \tag{4.3}$$

folds for all $h \in \mathcal{T}_{U_{\text{ad}}}(u_k)$. On the other hand, local solutions satisfy a system of C-stationarity, see [G. Wachsmuth, 2016, Theorem 1.1] and (under higher regularity assumptions on the data) [Schiela and D. Wachsmuth, 2013, Propositions 3.5–3.8]. This yields the following result.

Lemma 4.3. *Let (y_k, u_k) be locally optimal for ($\mathbf{P}_\gamma$). Then, there exist $\mu_k \in H^{-1}(\Omega)$, $\lambda_k \in L^2(\Omega)$ and $p_k \in H_0^1(\Omega)$ such that the system*

$$\mathcal{A}^\star p_k + J_y(y_k, u_k) + \gamma_k \max\{0, y_k - y_b\} + \mu_k = 0 \quad \textit{in } H^{-1}(\Omega), \tag{4.4a}$$

$$J_u(y_k, u_k) + (u_k - \bar{u}) + \lambda_k - p_k = 0 \quad \textit{in } L^2(\Omega), \tag{4.4b}$$

$$p_k = 0 \quad \textit{q.e. on } \Omega_{s,k}, \tag{4.4c}$$

$$\langle \mu_k, v\rangle_{H_0^1(\Omega)} = 0 \quad \forall v \in H_0^1(\Omega), v = 0 \textit{ q.e. on } \Omega_{a,k}, \tag{4.4d}$$

$$\langle \mu_k, \Phi p_k\rangle_{H_0^1(\Omega)} \geq 0 \quad \forall \Phi \in W^{1,\infty}(\Omega)^+, \tag{4.4e}$$

$$\lambda_k \in \mathcal{N}_{U_{\text{ad}}}(u_k) \tag{4.4f}$$

is satisfied. Here,

$$\Omega_{a,k} := \{y_k = y_a\}, \qquad\qquad \Omega_{s,k} := \text{q-supp}\,\xi_k$$

are the active and strictly active set for the obstacle problem at (y_k, u_k)*, respectively, and* $\xi_k := u_k - \mathcal{A}y_k$ *is the corresponding multiplier. Note that both sets are defined up to sets of capacity zero.*

4.2 Boundedness of the multipliers

From now on, we will not only fix $(\bar{y}, \bar{u})$ (with associated multiplier $\bar{\xi}$), but also sequences $(\gamma_k)_{k\in\mathbb{N}}$ and $((y_k, u_k))_{k\in\mathbb{N}}$ as in Lemma 4.2 and the corresponding sequences of multipliers from Lemma 4.3. Recall that Lemma 4.2 already implies the convergence results for the primal quantities u_k, y_k, and ξ_k. We check that this implies bounds on the dual variables, in order to pass to the limit in the optimality system (4.4) in Lemma 4.3.

For brevity, we introduce the regularized counterpart to the multiplier ν for the state constraint via

$$\nu_k := \gamma_k \max\{0, y_k - y_b\}, \tag{4.5}$$

as well as the set on which the state constraint is violated or active, i.e.,

$$\Omega_{b,k} := \{y_k \geq y_b\}.$$

Note that ν_k is an approximation of a Lagrange multiplier for the pointwise state constraint $\bar{y} \leq y_b$ in the unregularized problem (**P**) with support contained in $\Omega_{b,k}$. Our first goal is to bound ν_k, μ_k, and p_k in appropriate spaces. To this end, we observe that the supports of ν_k and μ_k are uniformly separated.

Lemma 4.4. *There exists a constant* $\rho > 0$ *such that*

$$\operatorname{dist}(\Omega_{a,k}, \Omega_{b,k}) \geq \rho$$

holds for all k*.*

Proof. By Assumption 2.13 (i) the controls u_k are bounded in $L^2(\Omega)$, and the associated states y_k are bounded in $H^2(\Omega)$ due to the mapping properties of S. Hence, their Hölder-norm is uniformly bounded and the result follows from $y_a \leq y_b - \tau$. □

As a consequence, we obtain the following auxiliary result:

Lemma 4.5. *There exist open sets* $\hat{\Omega}_a \supset \Omega_a$ *and* $\hat{\Omega}_b \supset \Omega_b$ *such that* $\Omega_{a,k} \subset \hat{\Omega}_a$, $\Omega_{b,k} \subset \hat{\Omega}_b$ *for all* k *sufficiently large as well as* $\rho > 0$ *such that*

$$\operatorname{dist}(\hat{\Omega}_a, \hat{\Omega}_b) > \rho.$$

Proof. We define

$$\hat{\Omega}_a := \Big\{\bar{y} < y_a + \frac{\tau}{4}\Big\}, \qquad\qquad \hat{\Omega}_b := \Big\{\bar{y} > y_b - \frac{\tau}{4}\Big\}.$$

For $x \in \Omega_{a,k}$, we observe that $\bar{y}(x) = \bar{y}(x) - y_k(x) + y_k(x) \le \|\bar{y} - y_k\|_{L^\infty(\Omega)} + y_a(x)$. For k large enough, uniform convergence of y_k towards $\bar{y}$ yields $x \in \hat{\Omega}_a$. The set $\hat{\Omega}_b$ can be treated analogously. From $y_b - \frac{\tau}{4} - y_a - \frac{\tau}{4} < \tau - \frac{\tau}{2} = \frac{\tau}{2}$ and the Hölder continuity of $\bar{y}$ the result follows. □

The boundedness of the multiplier approximations ν_k is a simple consequence of the Slater point property.

Lemma 4.6. *There exists $C > 0$ such that $\|\nu_k\|_{L^1(\Omega)} \le C$.*

Proof. We start with the B-stationarity (4.3) with $h = \hat{u} - u_k$, i.e.,

$$\begin{aligned} 0 \le \langle J_y(y_k, u_k), S'(u_k; \hat{u} - u_k)\rangle + \gamma_k \left(\max\{0, y_k - y_b\}, S'(u_k; \hat{u} - u_k)\right) \\ + \left(J_u(y_k, u_k), \hat{u} - u_k\right) + (u_k - \bar{u}, \hat{u} - u_k). \end{aligned}$$

Due to the convergence properties of y_k and u_k, the first, third and fourth addend can be bounded by a constant. Thus,

$$\gamma_k \left(\max\{0, y_k - y_b\}, S'(u_k; \hat{u} - u_k)\right) \ge -C.$$

Due to the convexity of the solution operator S, Corollary 2.12 can be used to obtain a linearized Slater condition for the local solutions u_k of ($\mathbf{P}_\gamma$) from the Slater point $\hat{u}$. Indeed, for all $k > 0$ we have

$$y_k + S'(u_k; \hat{u} - u_k) \le S(\hat{u}) \le y_b - \tau.$$

Combining the last two inequalities yields

$$(\nu_k, y_b - y_k - \tau) = \gamma_k \left(\max\{0, y_k - y_b\}, y_b - y_k - \tau\right) \ge -C.$$

Since $(\nu_k, y_b - y_k) \le 0$ by definition of ν_k, we obtain $\|\nu_k\|_{L^1(\Omega)} \le C\tau^{-1}$. □

Next, we show the boundedness of the adjoint state p and of the multiplier μ.

Lemma 4.7. *For every $q \in (1, n/(n-1))$, there exists $C > 0$ such that*

$$\|p_k\|_{W_0^{1,q}(\Omega)} + \|p_k\|_{H^1(\hat{\Omega}_a)} + \|\mu_k\|_{H^{-1}(\Omega)} \le C,$$

where $\hat{\Omega}_a$ is defined in Lemma 4.5.

Proof. We split the adjoints p_k into the sum of $p_k^y, p_k^\mu, p_k^\nu \in H_0^1(\Omega)$, defined via

$$\begin{aligned} \mathcal{A}^\star p_k^y + J_y(y_k, u_k) &= 0 \quad \text{in } H^{-1}(\Omega), \\ \mathcal{A}^\star p_k^\mu + \mu_k &= 0 \quad \text{in } H^{-1}(\Omega), \\ \mathcal{A}^\star p_k^\nu + \gamma_k \max\{0, y_k - y_b\} &= 0 \quad \text{in } H^{-1}(\Omega). \end{aligned}$$

Due to $(y_k, u_k) \to (\bar{y}, \bar{u})$, the term $J_y(y_k, u_k)$ is bounded in $H^{-1}(\Omega)$, see Assumption 2.13 (vi). This implies the boundedness of p_k^y, i.e.,

$$\|p_k^y\|_{H_0^1(\Omega)} \leq C.$$

Next, we are going to bound p_k^ν. To this end, let $\varphi \in C_c^\infty(\Omega)$ be given, such that $\varphi = 1$ on $\hat{\Omega}_a$ and $\varphi = 0$ on $\hat{\Omega}_b$. Since the equation for p_k^ν can be understood in the very weak sense, see (2.12), we can apply Theorem 2.3 and Lemma 2.6 in combination with Lemma 4.6 to obtain

$$\|p_k^\nu\|_{W_0^{1,q}(\Omega)} + \|\varphi p_k^\nu\|_{H_0^1(\Omega)} \leq C.$$

To obtain a uniform bound for p_k^μ, we write

$$\gamma_1 \|p_k^\mu\|^2_{H_0^1(\Omega)} \leq \langle \mathcal{A}^\star p_k^\mu, p_k^\mu \rangle = -\langle \mu_k, p_k^\mu \rangle = \langle \mu_k, p_k^\nu \rangle + \langle \mu_k, p_k^y \rangle - \langle \mu_k, p_k \rangle.$$

In order to bound the first term, we use

$$\langle \mu_k, (1 - \varphi) v \rangle = 0 \qquad \forall v \in H_0^1(\Omega)$$

due to (4.4d). For the third term, we can apply (4.4e) with $\Phi = 1$ and obtain $-\langle \mu_k, p_k \rangle \leq 0$. Now, the above inequality yields

$$\begin{aligned} \gamma_1 \|p_k^\mu\|^2_{H_0^1(\Omega)} &\leq \langle \mu_k, \varphi p_k^\nu \rangle + \langle \mu_k, p_k^y \rangle \\ &\leq C \|\mu_k\|_{H^{-1}(\Omega)} \|\varphi p_k^\nu\|_{H_0^1(\Omega)} + C \|\mu_k\|_{H^{-1}(\Omega)} \|p_k^y\|_{H_0^1(\Omega)}. \end{aligned}$$

Together with

$$C^{-1} \|p_k^\mu\|_{H_0^1(\Omega)} \leq \|\mu_k\|_{H^{-1}(\Omega)} \leq C \|p_k^\mu\|_{H_0^1(\Omega)}$$

which follows from the coercivity of $\mathcal{A}^\star$, we obtain the claim. □

4.3 Passage to the limit in the optimality system

From the boundedness results in Lemma 4.6 and Lemma 4.7 we conclude that there exist weakly convergent subsequences, denoted by the same index k, satisfying

$$p_k \rightharpoonup p \quad \text{in } W_0^{1,q}(\Omega), \qquad \nu_k \overset{\star}{\rightharpoonup} \nu \quad \text{in } \mathcal{M}(\Omega), \qquad \mu_k \rightharpoonup \mu \quad \text{in } H^{-1}(\Omega).$$

In the following steps, we will prove that the limits satisfy the optimality system of Theorem 4.1. To this end, we recall the strong convergences

$$u_k \to \bar{u} \quad \text{in } L^2(\Omega), \qquad y_k \to \bar{y} \quad \text{in } H_0^1(\Omega)$$

from Lemma 4.2, and note that the strong convergence

$$\lambda_k \to \lambda \quad \text{in } L^2(\Omega)$$

with $\lambda \in \mathcal{N}_{U_{\mathrm{ad}}}(\bar{u})$ is then a simple consequence of the gradient equation (4.4b) in Lemma 4.3 and the closedness of the graph of the normal cone. This proves (4.1b) and (4.1g). Also, (4.1a) is immediately clear. It remains to prove the complementarity condition for ν, as well as the properties (4.1c)-(4.1e). First, let us show that the weak limit ν fulfills the complementarity condition (4.1f) for the unregularized problem (**P**).

Lemma 4.8. *The weak limit ν fulfills $\nu \geq 0$ as well as $\operatorname{supp}(\nu) \subset \Omega_b$, see (4.1f).*

Proof. Nonnegativity of ν is an immediate consequence of $\nu_k \geq 0$ for all k. Moreover, from (4.2), we observe

$$\int_\Omega \nu_k \, (y_k - y_b) \, \mathrm{d}x = \gamma_k \int_\Omega \max\{0, y_k - y_b\}^2 \, \mathrm{d}x \to 0.$$

The mapping properties of S guarantee $y_k \to \bar{y}$ in $C_0(\Omega)$, and hence

$$\langle \nu, \bar{y} - y_b \rangle_{C_0(\Omega)} = 0.$$

Feasibility of $\bar{y}$, i.e. $\bar{y} - y_b \leq 0$ concludes the proof. □

The conditions (4.1c) and (4.1d) on p and μ follow from results in [G. Wachsmuth, 2016]:

Lemma 4.9. *The weak limit p of (p_k) satisfies $p = 0$ q.e. on $\operatorname{q-supp}(\bar{\xi})$, see (4.1c).*

Proof. This follows from (4.4c) via [G. Wachsmuth, 2016, Lemma 4.2]. □

Lemma 4.10. *The weak limit μ of (μ_k) satisfies*

$$\mu \in \{v \in H_0^1(\Omega) \mid v = 0 \textit{ q.e. on } \{\bar{y} = y_a\}\}^\perp,$$

see (4.1d).

Proof. This follows from (4.4d) via [G. Wachsmuth, 2016, Lemma 4.3]. □

Finally, we prove (4.1e).

Lemma 4.11. *The weak limits p and μ fulfill $\langle \mu, \Phi p \rangle_{H_0^1(\Omega)} \geq 0$ for all $\Phi \in W^{1,\infty}(\Omega)^+$ that satisfy $\Phi_{|_{\Omega \setminus \hat{\Omega}_a}} = 0$, see (4.1e).*

Proof. From (4.4e) in the C-stationarity system of Lemma 4.3, we know

$$\langle \mu_k, p_k \Phi \rangle \geq 0 \qquad \forall \Phi \in W^{1,\infty}(\Omega)^+.$$

Using the separation of the adjoint state into p_k^y, p_k^μ, and p_k^ν, we therefore observe

$$0 \leq \langle \mu_k, p_k \Phi \rangle = \langle \mu_k, (p_k^y + p_k^\mu + p_k^\nu)\Phi \rangle = \langle \mu_k, p_k^y \Phi \rangle + \langle -\mathcal{A}^\star p_k^\mu, p_k^\mu \Phi \rangle + \langle \mu_k, p_k^\nu \Phi \rangle. \tag{4.6}$$

For the first term on the right-hand-side of (4.6), we note that p_k^y converges strongly in $H_0^1(\Omega)$ due to the mapping properties of $T^\star$, cf. Theorem 2.1 which is applicable to the adjoint equation. This yields $\langle \mu_k, p_k^y \Phi \rangle \to \langle \mu, p^y \Phi \rangle$. The arguments in [G. Wachsmuth, 2016, Proof of Lemma 4.5, (4.2)] applied to the second term yield

$$\limsup_{k \to \infty} \langle -\mathcal{A}^\star p_k^\mu, p_k^\mu \Phi \rangle \leq \langle -\mathcal{A}^\star p^\mu, p^\mu \Phi \rangle.$$

Finally, for the third term $\langle \mu_k, p_k^\nu \Phi\rangle$ we apply the separation of sets from Lemma 4.5, and point out that $\Phi p \in H_0^1(\Omega)$. Note that the supports of ν and all ν_k are contained in $\hat\Omega_b$. We apply Lemma 2.6 with $U = \hat\Omega_b$ and V an open set containing $\hat\Omega_b$ with positive distance to $\hat\Omega_a$, $\varphi \geq 0$, $\varphi = 1$ on $\hat\Omega_a$, and $\varphi = 0$ on $\hat\Omega_b$. Hence, $\mu_k = \varphi\mu_k$ converges weakly towards $\mu = \varphi\mu$ in $H^{-1}(\Omega)$ and φp_k^ν converges strongly to φp^ν in $H_0^1(\Omega)$. Thus we obtain

$$\langle \mu_k, p_k^\nu \Phi\rangle = \langle \mu_k, \varphi p_k^\nu \Phi\rangle \to \langle \mu, \varphi p^\nu \Phi\rangle = \langle \mu, p^\nu \Phi\rangle.$$

Collecting all arguments yields the assertion. $\square$

The proof of Theorem 4.1 is complete.

5 Strong stationarity without control constraints

In this section, we consider the problem (**P**) in the case without control constraints, i.e., $U_{\text{ad}} = L^2(\Omega)$. We will show that in this case local solutions are strongly stationary. As a byproduct, we will obtain a characterization of the normal cone of U_{state}.

In the following, we will follow the approach by [Mignot, 1976, Théorème 4.3] and show the system of strong stationarity by employing Theorem 3.5. On the one hand, this has the advantage of showing the equivalency of B-stationarity and strong stationarity. On the other hand, it enables us to derive the announced characterization of the normal cone of U_{state}.

Before we dive into the proofs, let us state the system of strong stationarity. Note that we state the system without assuming $U_{\text{ad}} = L^2(\Omega)$. However, we only show that it is a necessary optimality condition in case $U_{\text{ad}} = L^2(\Omega)$.

Definition 5.1. *Let an admissible control $\bar u \in U_{\text{eff}}$ be given. We denote by $\bar y = S(\bar u)$ and $\bar\xi = \bar u - \mathcal{A}\bar y$ the associated state and multiplier. We say that $\bar u$ is strongly stationary if there exist $p \in W_0^{1,q}(\Omega)$, $\mu \in H^{-1}(\Omega)$, $\nu \in \mathcal{M}(\Omega)^+$, $\lambda \in L^2(\Omega)$ such that $p \in H^1(\hat\Omega_a)$ for some open $\hat\Omega_a \supset \{\bar y = y_a\}$ and such that the system*

$$\mathcal{A}^\star p + J_y(\bar y, \bar u) + \nu + \mu = 0 \tag{5.1a}$$
$$J_u(\bar y, \bar u) + \lambda - p = 0 \tag{5.1b}$$
$$p = 0 \text{ q.e. on } \operatorname{q-supp}(\bar\xi), \quad p \leq 0 \text{ q.e. on } \{\bar y = y_a\} \tag{5.1c}$$
$$\mu \in \mathcal{K}(\bar u)^\circ \tag{5.1d}$$
$$\operatorname{supp}(\nu) \subset \Omega_b \tag{5.1e}$$
$$\lambda \in \mathcal{N}_{U_{\text{ad}}}(\bar u) \tag{5.1f}$$

is satisfied. Here, $\Omega_s := \operatorname{q-supp}(\bar\xi)$, $\Omega_a := \{\bar y = y_a\}$, $\Omega_b = \{\bar y = y_b\}$, and the adjoint equation is to be understood in the very weak sense, see (2.12).

Note that (5.1f) implies $\lambda = 0$ in case $U_{\text{ad}} = L^2(\Omega)$.

In the sequel of this section, we will use two smooth test functions. These functions have the properties

$$\psi \in C_c^\infty(\Omega), \quad 0 \le \psi \le 1 \text{ in } \Omega, \quad \psi = 0 \text{ in nbhd. of } \Omega_a, \quad \psi = 1 \text{ in nbhd. of } \Omega_b, \tag{5.2a}$$
$$\varphi \in C^\infty(\mathbb{R}^n) \quad 0 \le \varphi \le 1 \text{ in } \Omega, \quad \varphi = 1 \text{ in } \{\psi < 1\}, \qquad \varphi = 0 \text{ in nbhd. of } \Omega_b. \tag{5.2b}$$

We fix φ and ψ throughout this section. Note that such a choice of φ and ψ is possible since Ω_a and Ω_b have a positive distance and since Ω_b has a positive distance to the boundary.

Further, we argue that the regularity $p \in H^1(\hat{\Omega}_a)$ is enough to write down the condition (5.1c). Indeed, it implies that $\varphi p \in H_0^1(\Omega)$ has a quasi-continuous representative. Since $\varphi = 1$ in a neighborhood of Ω_a, p is quasi-continuous on Ω_a and therefore it makes sense to state (5.1c). In the case that the adjoint state has additionally the regularity $p \in H_0^1(\Omega)$, one can formulate (5.1c) as $p \in -\mathcal{K}(\bar{u})$.

Now, we start with the B-stationarity system Theorem 3.5. In order to satisfy (5.1b) and (5.1f), we set $p = J_u(\bar{y}, \bar{u})$ in case $U_{\text{ad}} = L^2(\Omega)$. The differentiability properties of J only yield the low regularity $p \in L^2(\Omega)$. In the next two results, we show that p enjoys some increased regularity and that the first-order condition (3.3) can be extended to a larger test spaces.

Lemma 5.2. *We assume $U_{\text{ad}} = L^2(\Omega)$. Let $\bar{u} \in U_{\text{eff}}$ satisfy (3.3). Then, the function $p := J_u(\bar{y}, \bar{u}) \in L^2(\Omega)$ satisfies $p \in W^{1,q}(\Omega)$ for all $q < n/(n-1)$. Moreover,*

$$\langle J_y(\bar{y}, \bar{u}), S'(\bar{u}; h)\rangle + \langle h, p\rangle \ge 0 \qquad \forall h \in W^{-1,q'}(\Omega), S'(\bar{u}; h) \le 0 \text{ on } \Omega_b, \tag{5.3}$$

where $\Omega_b := \{\bar{y} = y_b\}$ with $\bar{y} = S(\bar{u})$.

Proof. Let $q \in (1, n/(n-1))$ be given, i.e., $q' \in (n, \infty)$. For all $h \in L^2(\Omega)$, we have $z_h := S'(\bar{u}; h)$ satisfies $\mathcal{A}z = h - \upsilon_h$ with $\upsilon_h \in \mathcal{K}(\bar{u})^\circ$ and $\|\upsilon_h\|_{H^{-1}(\Omega)} \le \|h\|_{H^{-1}(\Omega)}$ Now, from Lemma 2.10 (a) we get the estimate

$$\|S'(\bar{u}; h)\|_{L^\infty(\Omega_b)} = \|S'(\bar{u}; h) - S'(\bar{u}; 0)\|_{L^\infty(\Omega_b)} \le C\|h\|_{W^{-1,q'}(\Omega)},$$

where $C > 0$ is independent of h. We set $c := \zeta/C$. Then, for all $h \in L^2(\Omega)$ with $\|h\|_{W^{-1,q'}(\Omega)} \le c$ we have

$$\begin{aligned} S'(\bar{u}; h + (\hat{u} - \bar{u})) &\le S'(\bar{u}; h) + S'(\bar{u}; \hat{u} - \bar{u}) \\ &\le S'(\bar{u}; h) + S(\hat{u}) - S(\bar{u}) \le \zeta + (y_b - \zeta) - y_b = 0 \end{aligned}$$

on Ω_b, see Corollary 2.12. Therefore, we can use $h + (\hat{u} - \bar{u})$ as a test function in Theorem 3.5 and obtain

$$\langle J_y(\bar{y}, \bar{u}), S'(\bar{u}; h + (\hat{u} - \bar{u}))\rangle + (J_u(\bar{y}, \bar{u}), h + (\hat{u} - \bar{u})) \ge 0.$$

Thus, there is $K > 0$ such that

$$(J_u(\bar{y}, \bar{u}), h) \ge -\langle J_y(\bar{y}, \bar{u}), S'(\bar{u}; h + (\hat{u} - \bar{u}))\rangle - (J_u(\bar{y}, \bar{u}), \hat{u} - \bar{u}) \ge -K$$

holds for all $h \in L^2(\Omega)$ with $\|h\|_{W^{-1,q}(\Omega)} \le c$. Since we can replace h by $-h$, we infer

$$|(J_u(\bar{y}, \bar{u}), h)| \le K \qquad \forall h \in L^2(\Omega), \|h\|_{W^{-1,q}(\Omega)} \le C.$$

By scaling we get

$$|(J_u(\bar{y}, \bar{u}), h)| \le \frac{K}{C} \|h\|_{W^{-1,q}(\Omega)} \qquad \forall h \in L^2(\Omega).$$

This and the density of $L^2(\Omega)$ in $W^{-1,q}(\Omega)$, imply $p = J_u(\bar{y}, \bar{u}) \in W_0^{1,q}(\Omega)$.

It remains to show (5.3). Let $h \in W^{-1,q'}(\Omega)$ with $S'(\bar{u}; h) \le 0$ on Ω_b be given. Then, there is a sequence $(h_k)_{k \in \mathbb{N}} \subset L^2(\Omega)$ with $h_k \to h$ in $W^{-1,q'}(\Omega)$. Then, Lemma 2.10 (a) implies

$$r_k := \|S'(\bar{u}; h_k) - S'(\bar{u}; h)\|_{L^\infty(\Omega_b)} \to 0.$$

Thus,

$$S'(\bar{u}; h_k + r_k \zeta^{-1} (\hat{u} - \bar{u})) \le 0 \qquad \text{a.e. on } \Omega_b.$$

Hence,

$$\langle J_y(\bar{y}, \bar{u}), S'(\bar{u}; h_k + r_k \zeta^{-1} (\hat{u} - \bar{u})) \rangle + (J_u(\bar{y}, \bar{u}), h_k + r_k \zeta^{-1} (\hat{u} - \bar{u})) \ge 0.$$

Due to $J_u(\bar{y}, \bar{u}) = p \in W_0^{1,q}(\Omega)$, we can pass to the limit $k \to \infty$ and obtain (5.3). □

Using similar arguments, we get that p has H^1-regularity if we stay away from the active set Ω_b.

Lemma 5.3. *We assume $U_{\text{ad}} = L^2(\Omega)$. Let $\bar{u} \in U_{\text{eff}}$ satisfy (3.3) and $p := J_u(\bar{y}, \bar{u}) \in L^2(\Omega)$. We have $\varphi p \in H_0^1(\Omega)$ and*

$$\langle J_y(\bar{y}, \bar{u}), S'(\bar{u}; \varphi h) \rangle + \langle h, \varphi p \rangle \ge 0 \qquad \forall h \in H^{-1}(\Omega), S'(\bar{u}; \varphi h) \le 0 \text{ on } \Omega_b. \tag{5.4}$$

Proof. The proof is very similar to the proof of Lemma 5.2. We mainly have to replace the regularity result Lemma 2.10 (a) by Lemma 2.10 (b). This yields

$$\|S'(\bar{u}; \varphi h)\|_{L^\infty(\Omega_b)} \le C \|h\|_{H^{-1}(\Omega)}$$

for all $h \in L^2(\Omega)$, where $C > 0$ is independent of h. Now we can argue along the lines of the proof of Lemma 5.2. □

Using this extended stationarity condition, we can show the sign conditions on p.

Lemma 5.4. *We assume $U_{\text{ad}} = L^2(\Omega)$. Let $\bar{u} \in U_{\text{eff}}$ satisfy (3.3) and $p := J_u(\bar{y}, \bar{u}) \in L^2(\Omega)$. Then, $p \in H^1(\hat{\Omega}_a)$ for some open $\hat{\Omega}_a \supset \Omega_a := \{S(\bar{u}) = y_a\}$ and (5.1c) holds.*

Proof. We choose $\hat{\Omega}_a$ with a positive distance to $\Omega_b := \{S(\bar{u}) = y_b\}$. Then, there exist $\varphi \in C_c^\infty(\Omega)$ such that the support of φ does not intersect Ω_b, $\varphi \ge 0$ and $\varphi = 1$ on $\hat{\Omega}_a$. Now, Lemma 5.3 implies $p \in H^1(\hat{\Omega}_a)$.

Next, let $h \in \mathcal{K}(\bar{u})^\circ$ be arbitrary. Since $\varphi v \in \mathcal{K}(\bar{u})$ for all $v \in \mathcal{K}(\bar{u})$, we have $\varphi h \in \mathcal{K}(\bar{u})^\circ$ as well. Thus, $S'(\bar{u}; \varphi h) = 0$, see Theorem 2.9. Thus, (5.4) implies

$$\langle h, \varphi p \rangle \ge 0 \qquad \forall h \in \mathcal{K}(\bar{u})^\circ.$$

Hence, $\varphi p \in -\mathcal{K}(\bar{u})$. Since $\varphi = 1$ in a neighborhood of Ω_a, this shows (5.1c). □

It remains to verify the adjoint equation and the sign conditions on μ and ν. First, we consider the adjoint equation in a neighborhood of Ω_b.

Lemma 5.5. *We assume $U_{\mathrm{ad}} = L^2(\Omega)$. Let $\bar u \in U_{\mathrm{eff}}$ satisfy (3.3) and $p := J_u(\bar y, \bar u) \in W_0^{1,q}(\Omega)$ with arbitrary $q \in (2n/(n+2), n/(n-1))$. Then, there is $\nu \in \mathcal{M}(\Omega)$ such that (5.1e) and*

$$\langle \mathcal{A}(\psi z), p\rangle + \langle J_y(\bar y, \bar u), \psi z\rangle + \int_\Omega z \,\mathrm{d}\nu = 0 \qquad \forall z \in Z, \tag{5.5}$$

where Z is defined in (2.9).

Proof. We have $1/q \in (1-1/n, 1/2+1/n)$ and $1/q' \in (1/2-1/n, 1/n)$. Thus, Lemma 2.4 implies $\mathcal{A}(\psi z) \in W^{-1,q'}(\Omega)$, hence the first term in (5.5) is well defined.

Due to the properties of $S'(\bar u; \cdot)$, we have $S'(\bar u; \mathcal{A}(\psi z)) = \psi z$ for all $z \in Z$. If, additionally, $z \le 0$ on Ω_b, we have by (5.3)

$$\langle J_y(\bar y, \bar u), \psi z\rangle + \langle \mathcal{A}(\psi z), p\rangle \ge 0.$$

Hence, the left-hand side defines a negative functional w.r.t. $z \in Z$. Moreover, if $M := \|z\|_{C_0(\Omega)}$, we have $\varphi\,(z + M) \ge 0$ on Ω_b, thus

$$\langle J_y(\bar y, \bar u), \psi\,(z + M)\rangle + \langle \mathcal{A}(\psi\,(z + M)), p\rangle \ge 0,$$

i.e.,

$$\langle J_y(\bar y, \bar u), \psi z\rangle + \langle \mathcal{A}(\psi z), p\rangle \ge -\langle J_y(\bar y, \bar u), M\psi\rangle - \langle \mathcal{A}(M\psi), p\rangle =: -CM.$$

Similarly, by considering $\varphi\,(z - M)$, one can show that the left-hand side is bounded from above by CM. Since $C_c^\infty(\Omega)^+ \subset Z$ is dense in $C_0(\Omega)^+$, the first two addends in (5.5) define a negative Borel measure $-\nu \in \mathcal{M}(\Omega)$. This shows (5.5). Moreover, by considering $z \in C_c^\infty(\Omega)$ with $z = 0$ on Ω_b is arbitrary, we get $\int_\Omega z \,\mathrm{d}\nu = 0$. Hence, (5.1e) follows. □

Next, we argue in the neighborhood of Ω_a. To this end, we use the test function φ from (5.2b).

Lemma 5.6. *We assume $U_{\mathrm{ad}} = L^2(\Omega)$. Let $\bar u \in U_{\mathrm{eff}}$ satisfy (3.3) and $p := J_u(\bar y \bar u)$. We define $\mu \in H^{-1}(\Omega)$ via*

$$\langle \mu, v\rangle := -\big[\langle J_y(\bar y, \bar u), (1-\psi)v\rangle + \langle \mathcal{A}((1-\psi)v), p\rangle\big] \qquad \forall v \in H_0^1(\Omega). \tag{5.6}$$

Then, (5.1d) is satisfied.

Proof. Since $\varphi p \in H_0^1(\Omega)$ by Lemma 5.3, we have $p \in H^1(\{\psi < 1\})$. Thus, the definition (5.6) implies the regularity $\mu \in H^{-1}(\Omega)$.

In order to check (5.1d), we take an arbitrary $v \in \mathcal{K}(\bar u)$. Then, $(1-\psi)v \in \mathcal{K}(\bar u)$ as well and, consequently, $S'(\bar u; h) = (1-\psi)v$ for $h = \mathcal{A}((1-\psi)v)$. Note that $h = \varphi h$ due to the construction of φ and ψ, i.e., $S'(\bar u; \varphi h) = (1-\psi)v$ as well. Since $(1-\psi)v = 0$ on Ω_b, we can use h in (5.4) and obtain

$$\begin{aligned} -\langle \mu, v\rangle &= \langle J_y(\bar y, \bar u), (1-\psi)v\rangle + \langle \mathcal{A}((1-\psi)v), p\rangle \\ &= \langle J_y(\bar y, \bar u), (1-\psi)v\rangle + \langle \mathcal{A}((1-\psi)v), \varphi p\rangle \ge 0. \end{aligned}$$

Since $v \in \mathcal{K}(\bar u)$ was arbitrary, this shows $\mu \in \mathcal{K}(\bar u)^\circ$. □

By collection of the results of Lemmas 5.2 to 5.6, we can show that the system of strong stationarity is equivalent to the B-stationarity from Theorem 3.5.

Theorem 5.7. *We assume $U_{\mathrm{ad}} = L^2(\Omega)$ and let $\bar u \in U_{\mathrm{state}}$ be given. Then, $\bar u$ is strongly stationary if and only if the B-stationarity (3.3) is satisfied.*

Proof. "⇐": By using the results from the previous lemmas, it remains to show that the adjoint PDE (5.1a) is satisfied. To this end, let $z \in Z$ be arbitrary. Using (5.5) and (5.6), we have

$$\begin{aligned}\langle J_y(\bar y, \bar u), z\rangle + \langle \mathcal{A}z, p\rangle &= \langle J_y(\bar y, \bar u), (1-\psi)z\rangle + \langle \mathcal{A}((1-\psi)z), p\rangle \\ &\quad + \langle J_y(\bar y, \bar u), \psi z\rangle + \langle \mathcal{A}(\psi z), p\rangle \\ &= -\langle \mu, \nu\rangle - \int_\Omega z \,\mathrm{d}\nu.\end{aligned}$$

Hence, the adjoint PDE is satisfied.

"⇒": Now assume that the system of strong stationarity (5.1) is satisfied. Let $h \in L^2(\Omega)$ with $v := S'(\bar u; h) \le 0$ on Ω_b be given. We set $\xi_h := h - \mathcal{A}v \in \mathcal{K}(\bar u)^\circ \subset H^{-1}(\Omega)$. Note that, in general, $v \notin Z$, therefore we cannot use v directly as a test function in the adjoint PDE.

Due to the differentiability result Theorem 2.9, we have $\hat\xi := h - \mathcal{A}v \in \mathcal{K}(\bar u)^\circ$. This implies $\varphi\hat\xi = \hat\xi$. In order to test the adjoint PDE, we approximate $\hat\xi \in H^{-1}(\Omega)$ by a sequence $(\xi_k)_{k\in\mathbb{N}} \subset L^2(\Omega)$ such that $\xi_k \to \hat\xi$ in $H^{-1}(\Omega)$. This implies $\varphi\xi_k \to \varphi\hat\xi = \hat\xi$ in $H^{-1}(\Omega)$. Now, we can test the adjoint PDE by $z_k := \mathcal{A}^{-1}(h - \varphi\xi_k) \in Z$ and obtain

$$\int_\Omega p\,(h - \varphi\xi_k)\,\mathrm{d}x + \langle J_y(\bar y, \bar u) + \mu, z_k\rangle + \int_\Omega z_k \,\mathrm{d}\nu.$$

Now, we have $z_k \to v$ in $H_0^1(\Omega)$ and $z_k \to v$ in $C(\Omega_b)$ due to Theorem 2.1 and Lemma 2.5. Since the measure ν is supported on Ω_b, we can pass to the limit $k \to \infty$ and obtain

$$\int_\Omega ph\,\mathrm{d}x - \langle \hat\xi, \varphi p\rangle_{H_0^1(\Omega)} + \langle J_y(\bar y, \bar u) + \mu, v\rangle_{H_0^1(\Omega)} + \int_\Omega v\,\mathrm{d}\nu = 0.$$

Now, we can use the sign conditions from the adjoint system and from v and $\hat\xi$. This results in

$$\int_\Omega ph\,\mathrm{d}x + \langle J_y(\bar y, \bar u), v\rangle_{H_0^1(\Omega)} \ge 0.$$

Since h was arbitrary (as above) and $p = J_u(\bar y, \bar u)$, this shows (3.3). □

Note that the second part of the proof also works in case $U_{\mathrm{ad}} \ne L^2(\Omega)$.

Remark 5.8.

1. *In the case that $\bar u$ is even a locally optimal control, one can use the results of Section 4 to skip some parts of the proofs, since the system of C-stationarity already includes the regularity of p and the adjoint equation. However, one still needs to extend the B-stationarity condition via density to (5.3) and (5.4) to show the signs of p and μ.*

2. *In the case $U_{\mathrm{ad}} = L^2(\Omega)$, one can obtain the uniqueness of the multipliers. First, we infer $\lambda = 0$ and, thus, p is unique via (5.1b). Let us argue that μ and ν are unique by using (5.1a) and the fact that the supports of μ and ν are disjoint. To this end, let $\varphi \in C_c^\infty(\Omega)$ be given such that $\varphi = 0$ on $\{\bar y = y_a\}$. Thus, $\pm\varphi \in \mathcal{K}(\bar u)$ and (5.1d) implies $\langle \mu, \varphi\rangle = 0$. Hence, (5.1a) implies*

$$\langle \nu, \varphi\rangle = -\langle \mathcal{A}^\star p + J_y(\bar y, \bar u), \varphi\rangle \qquad \forall \varphi \in C_c^\infty, \varphi = 0 \text{ on } \{\bar y = y_a\}.$$

Since the support of ν is contained in $\{\bar y = y_b\}$, the measure ν is uniquely determined by the values of $\langle \nu, \varphi\rangle$ for these test functions φ. Since p is unique, the uniqueness of ν follows. Consequently, the uniqueness of μ follows from (5.1a).

The next result addresses the question of characterizing the normal cone to U_{state}, which was left open in Section 3.

Lemma 5.9. *Let $\bar u, \tau \in L^2(\Omega)$ be given and set $\bar y := S(\bar u)$. Then, $\tau \in \mathcal{N}_{U_{\mathrm{state}}}(\bar u)$ is equivalent to the existence of $p \in W_0^{1,q}(\Omega)$, $\mu \in H^{-1}(\Omega)$, $\nu \in \mathcal{M}(\Omega)^+$ such that $p \in H^1(\hat\Omega_a)$ for some open $\hat\Omega_a \supset \{\bar y = y_a\}$, $p = -\tau$,*

$$\mathcal{A}^\star p + \nu + \mu = 0$$

and the sign conditions (5.1c), (5.1d), (5.1e) are satisfied.

Proof. Let $\tau \in L^2(\Omega)$ be arbitrary. We consider the auxiliary problem

$$\begin{aligned} \text{Minimize} \quad & -(\tau, u) \\ \text{such that} \quad & u \in U_{\mathrm{state}}. \end{aligned}$$

Note that this is a special case of problem (**P**) with $U_{\mathrm{ad}} = L^2(\Omega)$ and $J(y,u) := -(\tau, u)$. Now, $\tau \in \mathcal{N}_{U_{\mathrm{state}}}(\bar u)$ is equivalent to

$$-(\tau, h) \ge 0 \qquad \forall h \in \mathcal{T}_{U_{\mathrm{state}}}(\bar u).$$

Using the characterization (3.2) of the tangent cone, this is equivalent to the B-stationarity (3.3) of $\bar u$ for the auxiliary problem. Now, the assertion follows from the equivalency of B-stationarity and strong stationarity in Theorem 5.7. $\square$

Using the characterization of the normal cone to U_{state}, we can use directly the optimality system from [G. Wachsmuth, 2016, Theorem 1.1], i.e., with U_{eff} as control constraints. Let us check that this does not yield a system of C-stationarity in the case $U_{\mathrm{ad}} \subset L^2(\Omega)$. From the referenced optimality system, we get the existence of $\mu_1 \in H^{-1}(\Omega)$, $\hat\lambda \in L^2(\Omega)$ and $p_1 \in H_0^1(\Omega)$ such that the system

$$\begin{aligned} \mathcal{A}^\star p_1 + J_y(\bar y, \bar u) + \mu_1 &= 0 \quad \text{in } H^{-1}(\Omega), \\ J_u(y,u) + \hat\lambda - p_1 &= 0 \quad \text{in } L^2(\Omega), \\ p_1 &= 0 \quad \text{q.e. on q-supp}(\bar\xi), \\ \langle \mu_1, v\rangle_{H^{-1}, H_0^1} &= 0 \quad \forall v \in H_0^1(\Omega), v = 0 \text{ q.e. on } \Omega_a, \\ \langle \mu_1, \Phi p_1\rangle &\ge 0 \quad \forall \Phi \in W^{1,\infty}(\Omega)^+, \\ \hat\lambda &\in \mathcal{N}_{U_{\mathrm{eff}}}(u) \end{aligned}$$

is satisfied. Next we use Theorem 3.4 and Lemma 5.9 to evaluate $\hat{\lambda} \in \mathcal{N}_{U_{\text{eff}}}(\bar{u})$. This yields the existence of $\lambda \in \mathcal{N}_{U_{\text{ad}}}(\bar{u})$, $p_2 \in W_0^{1,q}(\Omega)$, $\mu_2 \in H^{-1}(\Omega)$, $\nu \in \mathcal{M}(\Omega)^+$ such that $\hat{\lambda} = \lambda - p_2$, $p_2 \in H^1(\hat{\Omega}_a)$ for some open $\hat{\Omega}_a \supset \{\bar{y} = y_a\}$,

$$\mathcal{A}^\star p_2 + \nu + \mu_2 = 0$$

and the sign conditions (5.1c), (5.1d), (5.1e) are satisfied by p_2, μ_2 and ν, respectively.

By defining $p = p_1 + p_2$ and $\mu = \mu_1 + \mu_2$, we arrive at (4.1a) and (4.1b). The conditions (4.1f) and (4.1g) on ν and λ follow. Moreover, it is straightforward to see that (4.1c), (4.1d) are satisfied. However, the sign condition (4.1e) will, in general, not be valid.

References

Alt, Walter, Griesse, Roland, Metla, Nataliya, and Rösch, Arnd (2010). "Lipschitz Stability for Elliptic Optimal Control Problems with Mixed Control-State Constraints". In: *Optimization* 59, pp. 833–849. DOI: 10.1080/02331930902863749.

Barbu, Viorel (1984). *Optimal Control of Variational Inequalities*. Vol. 100. Research Notes in Mathematics. Boston: Pitman.

Bauschke, Heinz H. and Combettes, Patrick L. (2011). *Convex analysis and monotone operator theory in Hilbert spaces*. Berlin: Springer. ISBN: 978-1-4419-9466-0/hbk; 978-1-4419-9467-7/ebook. DOI: 10.1007/978-1-4419-9467-7.

Bergounioux, Maïtine (1998). "Optimal Control of Problems Governed by Abstract Elliptic Variational Inequalities with State Constraints". In: *SIAM Journal on Control and Optimization* 36.1, pp. 273–289. DOI: 10.1137/s0363012996302615.

Bergounioux, Maïtine, Haddou, Mounir, Hintermüller, Michael, and Kunisch, Karl (2000). "A comparison of a Moreau-Yosida-based active set strategy and interior point methods for constrained optimal control problems". In: *SIAM J. Optimization* 11, pp. 495–521. DOI: 10.1137/s1052623498343131.

Bergounioux, Maïtine and Kunisch, Karl (2002). "On the structure of the Lagrange multipliers for state-constrained optimal control problems". In: *Systems and Control Letters* 48, pp. 16–176. DOI: 10.1016/s0167-6911(02)00262-1.

Bergounioux, Maïtine and Tiba, Dan (1998). "Optimal control for the obstacle problem with state constraints". In: *ESAIM: Proceedings* 4, pp. 7–19. DOI: 10.1051/proc:1998018.

Bergounioux, Maïtine and Tröltzsch, Fredi (1999). "Optimal Control of Semilinear Parabolic Equations with State-Constraints of Bottleneck Type". In: *ESAIM COCV* 4, pp. 595–608. DOI: 10.1051/cocv:1999124.

Casas, Eduardo (1986). "Control of an elliptic problem with pointwise state constraints". In: *SIAM J. Control and Optimization* 4, pp. 1309–1322. DOI: 10.1137/0324078.

Casas, Eduardo (1993). "Boundary control of semilinear elliptic equations with pointwise state constraints". In: *SIAM J. Control and Optimization* 31, pp. 993–1006. DOI: 10.1137/0331044.

Casas, Eduardo, Mateos, Mariano, and Vexler, Boris (2014). "New regularity results and improved error estimates for optimal control problems with state constraints". In: *ESAIM: Control, Optimisation and Calculus of Variations* 20.3, pp. 803–822. DOI: 10.1051/cocv/2013084.

Casas, Eduardo and Tröltzsch, Fredi (2015). "Second order optimality conditions and their role in PDE control". In: *Jahresber. Dtsch. Math.-Ver.* 117.1, pp. 3–44. DOI: 10.1365/s13291-014-0109-3.

Gilbarg, David and Trudinger, Neil S. (2001). *Elliptic Partial Differential Equations of Second Order*. Springer Berlin Heidelberg. DOI: 10.1007/978-3-642-61798-0.

Griesse, Roland, Metla, Nataliya, and Rösch, Arnd (2008). "Convergence analysis of the SQP method for nonlinear mixed-constrained elliptic optimal control problems". In: *ZAMM* 88.10, pp. 776–792. DOI: 10.1002/zamm.200800036.

Harder, Felix and Wachsmuth, Gerd (2018a). "Comparison of Optimality Systems for the Optimal Control of the Obstacle Problem". In: *GAMM-Mitteilungen* 40.4, pp. 312–338. DOI: 10.1002/gamm.201740004.

Harder, Felix and Wachsmuth, Gerd (2018b). "The limiting normal cone of a complementarity set in Sobolev spaces". In: *Optimization* 67.10, pp. 1579–1603. DOI: 10.1080/02331934.2018.1484467.

He, Zheng-Xu (1987). "State Constrained Control Problems Governed by Variational Inequalities". In: *SIAM Journal on Control and Optimization* 25.5, pp. 1119–1144. DOI: 10.1137/0325061.

Hintermüller, Michael, Ito, Kazufumi, and Kunisch, Karl (2003). "The primal-dual active set strategy as a semismooth Newton method". In: *SIAM J. Optim.* 13, pp. 865–888. DOI: 10.1137/s1052623401383558.

Hintermüller, Michael and Kunisch, Karl (2009). "PDE-constrained optimization subject to pointwise constraints on the control, the state, and its derivative". In: *SIAM J. Optim.* 20.3, pp. 1133–1156. ISSN: 1052-6234. DOI: 10.1137/080737265. URL: http://dx.doi.org/10.1137/080737265.

Hintermüller, Michael and Surowiec, Thomas (2011). "First-order optimality conditions for elliptic mathematical programs with equilibrium constraints via variational analysis". In: *SIAM Journal on Optimization* 21.4, pp. 1561–1593. ISSN: 1052-6234. DOI: 10.1137/100802396.

Hintermüller, Michael, Tröltzsch, Fredi, and Yousept, Irwin (2008). "Mesh-independence of semismooth Newton methods for Lavrentiev-regularized state constrained nonlinear optimal control problems". In: *Numerische Mathematik* 108.4, pp. 571–603. DOI: 10.1007/s10589-007-9114-0.

Hinze, Michael and Tröltzsch, Fredi (2010). "Discrete concepts versus error analysis in PDE-constrained optimization". In: *GAMM-Mitt.* 33, pp. 148–166. DOI: 10.1002/gamm.201010012.

Ito, Kazufumi and Kunisch, Karl (2003). "Semi-Smooth Newton Methods for State-Constrained Optimal Control Problems". In: *Systems and Control Letters* 50, pp. 221–228. DOI: 10.1016/s0167-6911(03)00156-7.

Kinderlehrer, David and Stampacchia, Guido (1980). *An Introduction to Variational Inequalities and Their Applications*. New York: Academic Press.

Krumbiegel, Klaus and Rösch, Arnd (2009). “A virtual control concept for state constrained optimal control problems”. In: *COAP* 43.2, pp. 213–233. DOI: 10.1007/s10589-007-9130-0.

Kunisch, Karl and Wachsmuth, Daniel (2011). “Sufficient optimality conditions and semi-smooth newton methods for optimal control of stationary variational inequalities”. In: *ESAIM: Control, Optimisation and Calculus of Variations* 18.2, pp. 520–547. DOI: 10.1051/cocv/2011105.

Mateos, Mariano and Neitzel, Ira (2016). “Dirichlet control of elliptic state constrained problems”. In: *Comput Optim Appl* 63, pp. 825–853. DOI: 10.1007/s10589-015-9784-y.

Mehlitz, Patrick and Wachsmuth, Gerd (2019). “The weak sequential closure of decomposable sets in Lebesgue spaces and its application to variational geometry”. In: *Set-Valued and Variational Analysis* 27.1, pp. 265–294. DOI: 10.1007/s11228-017-0464-1.

Meyer, Christian, Panizzi, Lucia, and Schiela, Anton (2011). “Uniqueness Criteria for the Adjoint Equation in State-Constrained Elliptic Optimal Control”. In: *Numerical Functional Analysis and Optimization* 32.9, pp. 983–1007. DOI: 10.1080/01630563.2011.587074.

Meyer, Christian, Rösch, Arnd, and Tröltzsch, Fredi (2006). “Optimal control of PDEs with regularized pointwise state constraints”. In: *Computational Optimization and Applications* 33.2003-14, pp. 209–228. DOI: 10.1007/s10589-005-3056-1.

Mignot, Fulbert (1976). “Contrôle dans les inéquations variationelles elliptiques”. In: *Journal of Functional Analysis* 22.2, pp. 130–185. DOI: 10.1016/0022-1236(76)90017-3.

Neitzel, Ira and Tröltzsch, Fredi (2009). “On regularization methods for the numerical solution of parabolic control problems with pointwise state constraints”. In: *ESAIM Control, Optimisation and Calculus of Variations* 15.2, pp. 426–453. DOI: 10.1051/cocv:2008038.

Pieper, Konstantin and Vexler, Boris (2013). “A Priori Error Analysis for Discretization of Sparse Elliptic Optimal Control Problems in Measure Space”. In: *SIAM J. Control and Optimization* 51.4, pp. 2788–2808. DOI: 10.1137/120889137.

Raymond, Jean-Pierre and Zidani, Hasnaa (1998). “Pontryagin’s principle for state-constrained control problems governed by parabolic equations with unbounded controls”. In: *SIAM J. Control and Optimization* 36, pp. 1853–1879. DOI: 10.1137/s0363012996302470.

Raymond, Jean-Pierre and Zidani, Hasnaa (1999). “Hamiltonian Pontryagin’s Principles for control Problems governed by Semilinear Parabolic Equations”. In: *Applied Mathematics and Optimization* 39, pp. 143–177. DOI: 10.1007/s002459900102.

Rösch, Arnd and Tröltzsch, Fredi (2007). “On regularity of solutions and Lagrange multipliers of optimal control problems for semilinear equations with mixed pointwise control-state constraints”. In: *SIAM J. Control and Optimization* 46.3, pp. 1098–1115. DOI: 10.1137/060671565.

Schiela, Anton (2009). “Barrier Methods for Optimal Control Problems with State Constraints”. In: *SIAM J. on Optimization* 20 (2), pp. 1002–1031. DOI: 10.1137/070692789.

Schiela, Anton and Wachsmuth, Daniel (2013). "Convergence analysis of smoothing methods for optimal control of stationary variational inequalities with control constraints". In: *ESAIM Math. Model. Numer. Anal.* 47.3, pp. 771–787. DOI: 10.1051/m2an/2012049.

Troianiello, Giovanni Maria (1987). *Elliptic differential equations and obstacle problems.* The University Series in Mathematics. New York: Plenum Press, pp. xiv+353. ISBN: 0-306-42448-7.

Ulbrich, Michael and Ulbrich, Stefan (2009). "Primal-dual interior-point methods for PDE-constrained optimization". In: *Math. Program.* 117.1-2, Ser. B, pp. 435–485. ISSN: 0025-5610. DOI: 10.1007/s10107-007-0168-7. URL: http://dx.doi.org/10.1007/s10107-007-0168-7.

Wachsmuth, Gerd (2014). "Strong Stationarity for Optimal Control of the Obstacle Problem with Control Constraints". In: *SIAM Journal on Optimization* 24.4, pp. 1914–1932. DOI: 10.1137/130925827.

Wachsmuth, Gerd (2016). "Towards M-stationarity for optimal control of the obstacle problem with control constraints". In: *SIAM Journal on Control and Optimization* 54.2, pp. 964–986. DOI: 10.1137/140980582.