COMPUTATION OF A BOULIGAND GENERALIZED DERIVATIVE FOR THE SOLUTION OPERATOR OF THE OBSTACLE PROBLEM*

ANNE-THERESE RAULS† AND STEFAN ULBRICH†

Abstract. The non-differentiability of the solution operator of the obstacle problem is the main challenge in tackling optimization problems with an obstacle problem as a constraint. Therefore, the structure of the Bouligand generalized differential of this operator is interesting from a theoretical and from a numerical point of view. The goal of this article is to characterize and compute a specific element of the Bouligand generalized differential for a wide class of obstacle problems. We allow right-hand sides that can be relatively sparse in $H^{-1}(\Omega)$ and do not need to be distributed among all of $H^{-1}(\Omega)$. Under assumptions on the involved order structures, we investigate the relevant set-valued maps and characterize the limit of a sequence of Gâteaux derivatives. With the help of generalizations of Rademacher's theorem, this limit is shown to be in the considered Bouligand generalized differential. The resulting generalized derivative is the solution operator of a Dirichlet problem on a quasi-open domain.

Key words. optimal control, non-smooth optimization, Bouligand generalized differential, generalized derivatives, obstacle problem, variational inequalities, Mosco convergence

AMS subject classifications. 49J40, 35J86, 47H04, 49J52, 58C06, 58E35

1. Introduction.

This article is concerned with the derivation of a generalized gradient for the reduced objective associated with the optimal control of an obstacle problem

$$\begin{aligned} &\min_{y,u} \quad J(y,u) \\ &\text{subject to} \quad y \in K_\psi, \\ &\qquad\qquad \langle Ly - f(u), z - y\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} \geq 0 \quad \text{for all } z \in K_\psi. \end{aligned} \tag{1.1}$$

Here, $\Omega \subset \mathbb{R}^d$ is open and bounded, $J\colon H_0^1(\Omega)\times U \to \mathbb{R}$ is a continuously differentiable objective function and $L \in \mathcal{L}\left(H_0^1(\Omega), H^{-1}(\Omega)\right)$ is a coercive and strictly T-monotone operator. Furthermore, $f\colon U \to H^{-1}(\Omega)$ is a Lipschitz continuous, continuously differentiable and monotone operator on a partially ordered Banach space U. On $H^{-1}(\Omega)$ we use the partial ordering induced by the ordering on the predual space $H_0^1(\Omega)$. We say that $H^{-1}(\Omega) \ni \Lambda \geq 0$ if and only if $\Lambda(v) \geq 0$ for all $v \in H_0^1(\Omega)$ with $v \geq 0$ a.e. on Ω. We will specify the precise assumptions on U later on, but the class of Banach spaces U we consider includes the important examples $H^{-1}(\Omega)$, $L^2(\Omega)$ and $\mathbb{R}^n$. The closed convex set K_ψ is of the form

$$K_\psi := \left\{z \in H_0^1(\Omega) : z \geq \psi\right\}$$

and the quasi upper-semicontinuous obstacle ψ, see Definition 2.1, is chosen such that K_ψ is nonempty. In additon to the case $\psi \in H^1(\Omega)$ with $\psi \leq 0$ on $\partial\Omega$ this allows also for thin obstacles. The inequality "$z \geq \psi$" is to be understood pointwise quasi-everywhere (q.e.) in Ω. We note that for $\psi \in H^1(\Omega)$ with $\psi \leq 0$ on $\partial\Omega$

*Submitted to the editors June 14, 2019.

Funding: This work was supported by the DFG under grant UL158/10-1 within the Priority Program SPP 1962 Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization.

†Department of Mathematics, Technische Universität Darmstadt, Dolivostr. 15, 64293 Darmstadt, Germany (rauls@mathematik.tu-darmstadt.de, ulbrich@mathematik.tu-darmstadt.de).

this is equivalent to $z \geq \psi$ pointwise almost everywhere (a.e.) in Ω. We denote by $S\colon U \to H_0^1(\Omega)$ the solution operator of (1.1), by $A(u) := \{\omega \in \Omega : S(u)(\omega) = \psi(\omega)\}$ the active set and by $A_s(u) = \text{f-supp}(\mu) \subset A(u)$ the strictly active set. Here, μ is the regular Borel measure associated with the multiplier $LS(u) - f(u) \in H^{-1}(\Omega)^+$ and $\text{f-supp}(\mu)$ denotes the fine support of this measure, see also [42, App. A]. Details of the variational inequality in (1.1) will be discussed in section 3.

As a model problem, one can think of the optimal control problem

$$
\begin{aligned}
&\min_{y,u} \quad J(y,u) = \frac{1}{2}\int_\Omega |y - y_d|^2\,dx + \frac{\alpha}{2}\|u\|_U^2 \\
&\text{subject to} \quad y \in K_\psi, \\
&\qquad\qquad \langle -\Delta y - f(u), z - y\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} \geq 0 \quad \text{for all } z \in K_\psi.
\end{aligned} \tag{1.2}
$$

Instead of the least squares term in the first part of J other continuously differentiable functionals $y \in H_0^1(\Omega) \to \mathbb{R}$ are possible.

We give some examples of operators f and control spaces U that can enter the optimal control problem (1.2). For example, the operator f can realize controls given as $L^2(\Omega)$ functions with support in a measurable subset $\tilde{\Omega}$ of Ω by choosing $U = L^2(\tilde{\Omega})$ and f the embedding of $L^2(\tilde{\Omega})$ into $H^{-1}(\Omega)$. Next, the range of f can consist of weighted sums $\sum_{i=1}^n u_i f_i$ with fixed nonnegative functionals $f_i \in H^{-1}(\Omega)$, $i = 1,\dots,n$ by setting $U = \mathbb{R}^n$ and $f(u_1,\dots,u_n) := \sum_{i=1}^n u_i f_i$. Moreover, U can be a closed linear subspace of $H^{-1}(\Omega)$ satisfying Assumption 5.4, where f is the embedding of U into $H^{-1}(\Omega)$. Hence, in particular also different types of sparse controls are possible.

In particular, the assumptions on f include also the choices $U = H^{-1}(\Omega)$, $f(u) = \mathrm{id}(u) = u$ and $U = L^2(\Omega)$, $f\colon L^2(\Omega) \hookrightarrow H^{-1}(\Omega)$, which are interesting from a theoretical point of view. The class of obstacles we consider contains also thin obstacles which live on sets with positive capacity but Lebesgue measure zero, e.g., a line segment in the two-dimensional case.

In points $u \in U$ where the Lipschitz solution operator $S\colon U \to H_0^1(\Omega)$ is Gâteaux differentiable, the Gâteaux derivative $S'(u)$ is the solution operator to a Dirichlet problem of the form

$$
\begin{aligned}
&\text{Find } \xi \in H_0^1(D(u)), \\
&\langle L\xi - f'(u;h), z\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} = 0 \quad \text{for all } z \in H_0^1(D(u)).
\end{aligned} \tag{1.3}
$$

Here, any quasi-open set $D(u)$ with $\Omega \setminus A(u) \subset D(u) \subset \Omega \setminus A_s(u)$ up to a set of zero capacity is valid and yields the same solution operator equal to $S'(u)$. This variational equation for the Gâteaux derivative is a consequence of the classical result by Mignot [30], see also Haraux [19], which gives the directional derivative as the solution to a related variational inequality. For sensitivity analysis of variational inequalities, see also [2, 11, 13, 14] and the references therein.

Note that the strict complementarity condition, which states that $A_s(u) = A(u)$ holds up to a set of zero capacity and which was introduced, e.g., in [9, Def. 6.59], is not automatically implied by Gâteaux differentiability for general f other than the identity map on $H^{-1}(\Omega)$. Therefore, there can really be a gap between the sets $A_s(u)$ and $A(u)$, even when S is Gâteaux differentiable in u.

Since, in general, S is not Gâteaux differentiable in all points of U, we are interested in computing at least one element of a suitable generalized differential in such points. In order to obtain an element of the Bouligand generalized differential $\partial_B S(u)$ of S in a point $u \in U$, we construct and characterize limits of solution operators of

(1.3) for points u_n with $u_n \to u$ in U. The Bouligand generalized differential of S in u, which will be introduced in the subsequent subsection, is exactly defined as such a set of limits of Gâteaux derivative operators in points u_n converging to u with respect to a certain topology. The construction and characterization of suitable limits of solution operators respective to (1.3) requires a careful study of the set-valued maps $u \mapsto H_0^1(D(u))$ and we will focus on the special case $D(u) = \Omega \setminus A(u)$. The spaces $H_0^1(\Omega \setminus A(u_n))$ are shown to converge in the sense of Mosco towards $H_0^1(\Omega \setminus A(u))$ in the case where $u_n \uparrow u$ in U.

We impose assumptions on U to guarantee that the order structure in U is rich enough to construct a sequence of Gâteaux points $(u_n)_{n\in\mathbb{N}}$ with $u_n \uparrow u$ for arbitrary $u \in U$, see Assumption 5.4. In particular, these assumptions on U apply for the special cases $U = \mathbb{R}^n$, $U = L^2(\Omega)$ and $U = H^{-1}(\Omega)$.

Altogether, this leads to Theorem 5.6, which the following is a simplified version of.

Theorem 1.1. *Assume that $U = \mathbb{R}^n$, $U = L^2(\Omega)$ or $U = H^{-1}(\Omega)$ and let $f\colon U \to H^{-1}(\Omega)$ be a Lipschitz continuous, continuously differentiable and monotone operator. Let $u, h \in U$. Denote the unique solution to the variational equation*

$$\begin{aligned} &\text{Find } \xi \in H_0^1(\Omega \setminus A(u)),\\ &\langle L\xi - f'(u;h), z\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} = 0 \qquad \text{for all } z \in H_0^1(\Omega \setminus A(u)), \end{aligned} \tag{1.4}$$

by $\Xi\,(u;h)$. Then

$$\Xi(u;\cdot) \in \partial_B S(u).$$

Furthermore, for each increasing sequence $(u_n)_{n\in\mathbb{N}}$ with $u_n \uparrow u$ it holds

$$\Xi(u_n;h) \to \Xi(u;h)$$

for each $h \in U$.

If u is a point where S is Gâteaux differentiable, then the equality $S'(u) = \Xi(u;\cdot)$ holds, where $\Xi(u;\cdot)$ is determined by the variational equation (1.4). Thus, an element of the Bouligand generalized differential is obtained independently from differentiability. This explains why algorithms that assume not to strike non-differentiability points while working with formulas for the Gâteaux derivative will actually compute a Bouligand generalized derivative, also in points of non-Gâteaux differentiability.

In numerical realization it might be difficult to solve the system (1.4) because of the structure of the space $H_0^1(\Omega \setminus A(u))$. This is due to the possibly bad behaved form of the set $A(u)$. Our approach suggests that, instead of solving (1.4), it is also possible to solve the system for some $\tilde{u} < u$ to obtain a better behaved active set $A(\tilde{u})$ and an inexact subgradient that is a good approximation for $\tilde{u}$ close enough to u.

We conjecture that similar results to the statement in the theorem could be obtained by considering the sets $\Omega \setminus A_s(u)$ and sequences $(u_n)_{n\in\mathbb{N}}$ with opposite monotonicity properties. The generalized derivatives obtained by considering the sets $A(u)$ and $A_s(u)$ are somehow the ‘extremal’ elements in the Bouligand generalized differential, whereas the remaining elements are related to sets between them, see also Lemma 3.7 and Remark 3.8.

Once such an analogous result is established also for a quasi-open set $D(u)$ with

$$H_0^1(\Omega \setminus A(u)) \subset H_0^1(D(u)) \subset H_0^1(\Omega \setminus A_s(u)),$$

one can also switch to a better behaved set $D(u)$ to avoid difficulties with bad structure and to obtain a (possibly different) element of the Bouligand generalized differential.

We put this paper into perspective. There are several contributions dealing with optimal control of variational inequalities. Different techniques are applied to deal with the difficulties resulting from the variational inequality constraints and the resulting non-smoothness. Usually, the variational inequality is approximated by a sequence of penalized, relaxed or regularized problems. Exemplary, we mention [8, 5, 31, 25, 24, 28, 37, 29]. The provision of a generalized derivative for the solution operator of the obstacle problem and hence a subgradient for the reduced objective function of an optimal control problem subject to obstacle problem constraints naturally suggests to consider numerical optimization methods based on subgradient techniques. In the recent contribution [23], a bundle method for nonconvex nonsmooth optimization in infinite dimensional Hilbert spaces is developed. This method is appropriate for the numerical optimization of optimal control problems with obstacle constraints and the subgradient we construct can be used. For finite dimensional obstacle problems, a characterization of the whole Clarke subdifferential (of the reduced objective function) was obtained in [21]. To achieve this result, the authors impose the local surjectivity assumption $Df(u) = \mathbb{R}^n$ on $f\colon \mathbb{R}^m \to \mathbb{R}^n$. With this assumption, each index set between $A_s(u)$ and $A(u)$ can be associated with an element of the subdifferential. In [12], the Bouligand generalized differentials of the solution operator to a nonsmooth semilinear equation relative to several combinations of topologies are characterized. This is the first contribution dealing with generalized derivativess for nonsmooth mappings between infinite dimensional spaces. In the recent [35], these characterizations are obtained for the infinite dimensional obstacle problem, except for the one involving only weak topologies. The authors consider right-hand sides distributed among all of $H^{-1}(\Omega)$, i.e., f is the identity on $H^{-1}(\Omega)$. In this setting, for a given $u \in H^{-1}(\Omega)$ and an arbitrarily chosen quasi-open set $D(u)$ between $\Omega \setminus A(u)$ and $\Omega \setminus A_s(u)$, it is possible to construct a sequence $(u_n)_{n\in\mathbb{N}}$ of right-hand sides in $H^{-1}(\Omega)$, which converges to u and which fulfills $A(u_n) = A_s(u_n) = \Omega \setminus D(u)$ up to a set of capacity zero. This approach cannot be translated to our setting, since, in general, $(u_n)_{n\in\mathbb{N}}$ is not contained in the range of f. The characterization for the generalized differential involving the weak operator topology is obtained by considering capacitary measures and the corresponding notion of γ-convergence.

We will proceed as follows: In the remainder of this section we formally define the Bouligand generalized differential. Afterwards, in section 2, we give a short introduction to capacity theory, quasi-open sets, and Sobolev spaces on quasi-open sets. In section 3, we collect features of the variational inequality (1.1). For our purpose, especially results on monotonicity, differentiability and convergence are important. We proceed in section 4 with a precise analysis of the set-valued map connected to the sets that appear in the variational equation for the Gâteaux derivative and show a result on monotone continuity, see Corollary 4.2. This will help us to prove the Mosco convergence of the corresponding sets, provided the sequence $(u_n)_{n\in\mathbb{N}}$ converging to u has suitable monotonicity properties. Establishing these convergence results, identifying the limit and infering the main result, using density of Gâteaux differentiability points, will be done in section 5. An adjoint equation for a Clarke subgradient of the reduced objective will also be provided. Moreover, we sketch how our techniques can be applied to handle also shape optimization problems for the obstacle problem.

The Bouligand generalized differential. In the following, we will formally define the Bouligand generalized differential which we consider for the solution operator of the obstacle problem (1.1). It is not clear how to extend the definition of the Bouligand generalized differential from finite to infinite dimensions, see [12], since the respective weak and strong topologies in infinite dimensions are not equivalent anymore. Therefore, our definition is only one possible out of several generalizations to infinite dimensions. We choose one that can also be found in [12]. See also [40, Sect. 2] for the concepts, although they do not carry the specific name Bouligand generalized differential.

For the finite dimensional case, see, e.g., [33, Def. 2.12], [17, Def. 4.6.2].

Definition 1.2. *Let X be a separable Banach space and Y a separable Hilbert space. Furthermore, assume that the operator $T\colon X \to Y$ is Lipschitz continuous. We will denote the subset of X on which T is Gâteaux differentiable by D_T.*

The Bouligand generalized differential of T in u is defined as

$$\partial_B T(u) := \Big\{ \Sigma \in \mathcal{L}(X,Y) : T'(u_n) \to \Sigma \text{ in the weak operator topology for some sequence } (u_n)_{n\in\mathbb{N}} \subset D_T \text{ with } \lim_{n\to\infty} u_n = u \Big\}.$$

Here, T' denotes the Gâteaux derivative of T.

Remark 1.3.
1. We recall that $T'(u_n) \to \Sigma$ in the weak operator topology if $T'(u_n)\,x \rightharpoonup \Sigma x$ weakly in Y for all $x \in X$. Moreover, $T'(u_n) \to \Sigma$ in the strong operator topology if $T'(u_n)\,x \to \Sigma x$ strongly in Y for all $x \in X$.
2. Since D_T is dense in X, each $u \in X$ can be approximated by a sequence $(u_n)_{n\in\mathbb{N}} \subset D_T \subset X$, see Theorem 5.3 or [7, Thm. 6.42]. Now, since T is Lipschitz continuous, the Gâteaux derivatives $T'(u_n)$ are bounded by the Lipschitz constant in $\mathcal{L}(X,Y)$. The sequential compactness of the unit ball in $\mathcal{L}(X,Y)$ with respect to the weak operator topology, which follows from a generalization of the Banach-Alaoglu theorem to the weak operator topology, yields the existence of an accumulation point of the sequence $(T'(u_n))_{n\in\mathbb{N}}$. Thus, $\partial_B T(u)$ is nonempty.
3. Let $X = U$, $Y = H_0^1(\Omega)$ and let $\hat{J}(u) = J(S(u),u)$ be the reduced objective respective to a continuously differentiable objective function $J\colon H_0^1(\Omega) \times U \to \mathbb{R}$. Here, $S\colon U \to H_0^1(\Omega)$ is, as usual, the solution operator to the obstacle problem (1.1). Then

$$\{\Sigma^* J_y(S(u),u) + J_u(S(u),u) : \Sigma \in \partial_B S(u)\} \subset \partial_B \hat{J}(u) \subset \partial_C \hat{J}(u),$$

   where $\partial_C \hat{J}(u)$ is Clarke's generalized gradient, see [15].
   The first inclusion follows from the chain rule for the mixed Fréchet/ Gâteaux derivative, since J is Fréchet differentiable. The second inclusion is implied by weak*-closedness of Clarke's generalized gradient, see [15, Prop. 2.1.5b]. These arguments are also used in [12]. Moreover, by [40, Prop. 2.2], $\partial_C \hat{J}(u)$ is the closed convex hull of $\partial_B \hat{J}(u)$ in the weak operator topology.
4. From Definition 1.2, it directly follows that if T is Gâteaux differentiable in u with Gâteaux derivative $T'(u)$, then $T'(u)$ belongs to $\partial_B T(u)$.
5. In our derivation of elements in $\partial_B S(u)$, we will construct limits of sequences $(S'(u_n))_{n\in\mathbb{N}}$ in the strong operator topology, so we could have taken the strong operator topology in Definition 1.2 as well.

2. Fundamentals of capacity theory and Sobolev spaces on quasi-open domains.

2.1. Capacity theory. We define a capacity that is appropriate for our purposes when dealing with the space $H_0^1(\Omega)$. Note that it is possible to define analoguous capacities also for other Sobolev spaces, see, e.g., [1], [22]. In addition, the following definitions can be found, e.g., in [4, Sect. 5.8.2, Sect. 5.8.3], [16, Def. 6.2].

DEFINITION 2.1. *a) Let $E \subset \Omega$ be a set. Then we define the capacity of E by*

$$\operatorname{cap}(E) := \inf \left\{ \int_\Omega |\nabla v|^2 d\lambda^d : v \in H_0^1(\Omega), v \geq 1 \text{ a.e. in a neighborhood of } E \right\}.$$

b) A set $O \subset \Omega$ is called quasi-open if for all $\varepsilon > 0$ there exists an open set $\Omega_\varepsilon \subset \Omega$ with $\operatorname{cap}(\Omega_\varepsilon) < \varepsilon$ such that $O \cup \Omega_\varepsilon$ is open. The complement of a quasi-open set is called quasi-closed.

c) A function $v\colon \Omega \to \mathbb{R}$ is called quasi-continuous (quasi upper-semicontinuous, quasi lower-semicontinuous, respectively) if for all $\varepsilon > 0$ there exists an open set $\Omega_\varepsilon \subset \Omega$ with $\operatorname{cap}(\Omega_\varepsilon) < \varepsilon$ and such that $v|_{\Omega\setminus\Omega_\varepsilon}$ is continuous (upper-semicontinuous, lower-semicontinuous, respectively).

We use the prefix 'quasi-' to express that a property holds except on a set of zero capacity.

Remark 2.2. Countable unions of quasi-open sets are quasi-open and countable intersections of quasi-closed sets are quasi-closed. Nevertheless, the family of quasi-open subsets does not define a topology on Ω, since arbitrary unions of quasi-open sets do not have to be quasi-open, see also [10, Chap. 4].

The next lemma is taken from [10, Thm. 4.1.6] and [4, Thm. 5.8.6].

LEMMA 2.3. *Suppose $v\colon \Omega \to \mathbb{R}$ is a function. The following assertions are equivalent:*

a) v is quasi lower-semicontinuous,

b) the sets $\{v > c\}$ are quasi-open for all $c \in \mathbb{R}$,

c) $-v$ is quasi upper-semicontinuous.

The following result on quasi-continuous representatives can be found, e.g., in [16, Chap. 6, Thm. 6.1].

LEMMA 2.4. *Each element of $H^1(\Omega)$ has a quasi-continuous representative and this quasi-continuous representative is uniquely defined up to values on subsets of Ω with zero capacity.*

When we speak about function values of elements $v \in H_0^1(\Omega)$, then we usually mean the function values of a quasi-continuous representative. In this sense, for a quasi upper-semicontinuous ψ, the sets

$$\{\omega \in \Omega : v(\omega) = \psi(\omega)\}$$

or

$$\{\omega \in \Omega : v(\omega) \geq 0\}$$

have to be understood. They are defined up to a set of zero capacity. Furthermore, these sets are quasi-closed, see Lemma 2.3.

Let $v, w \in H_0^1(\Omega)$. We say that $v \geq w$ if

$$\operatorname{cap}(\{\omega \in \Omega : v(\omega) < w(\omega)\}) = 0,$$

i.e., if $v \geq w$ q.e. on Ω. Equivalently, we could demand $v \geq w$ a.e. on Ω, see [42, Lem. 2.3].

Together with the relation '$\geq$', the space $H_0^1(\Omega)$ becomes a vector lattice, see [36]. In particular, $\sup(v, w) \in H_0^1(\Omega)$ exists for all $v, w \in H_0^1(\Omega)$.

The following lemma on convergence in $H_0^1(\Omega)$ is taken out of [9, Lem. 6.52].

Lemma 2.5. *Suppose that $y_n \to y$ in $H_0^1(\Omega)$. Denote by $(\tilde{y}_n)_{n\in\mathbb{N}}$ and $\tilde{y}$ fixed quasi-continuous representatives of $(y_n)_{n\in\mathbb{N}}$ and y, respectively. Then there is a subsequence of $(\tilde{y}_n)_{n\in\mathbb{N}}$ that converges to $\tilde{y}$ pointwise quasi-everywhere.*

2.2. The space $H_0^1(O)$ with O quasi-open. The goal of this subsection is to define the spaces $H_0^1(O)$ for quasi-open sets O as generalizations of the spaces $H_0^1(\Omega)$ on open domains $\Omega \subset \mathbb{R}^d$. These spaces are interesting for our purposes since closed linear subspaces of $H_0^1(\Omega)$ are of this form and since they naturally appear in the Dirichlet problems characterizing the Gâteaux derivative of the solution operator, see subsection 3.1.

The capacity we have introduced in Definition 2.1 can be evaluated on arbitrary subsets of Ω. In the following subsection, we need to formally define the capacity of arbitrary subsets of $\mathbb{R}^d$, quasi-open sets in $\mathbb{R}^d$ and quasi-continuous representatives of equivalence classes in $H^1(\mathbb{R}^d)$. Therefore, we introduce the following notion of capacity which is also called 'Sobolev capacity', see also [22, Sect. 2.35], [4, Sect. 5.8.2]:

Definition 2.6. *a) For a set $E \subset \mathbb{R}^d$ we define*

$$\operatorname{Cap}(E) := \inf \left\{ \int_{\mathbb{R}^d} \left(|v|^2 + |\nabla v|^2\right) \, d\lambda^d : \right. \\ \left. v \in H^1(\mathbb{R}^d), v \geq 1 \text{ a.e. in a neighborhood of } E \right\}.$$

b) A set $O \subset \mathbb{R}^d$ is called $\mathbb{R}^d$-quasi-open if for all $\varepsilon > 0$ there exists an open set $\Omega_\varepsilon \subset \mathbb{R}^d$ such that $O \cup \Omega_\varepsilon$ is open and such that $\operatorname{Cap}(\Omega_\varepsilon) < \varepsilon$. The complement of a $\mathbb{R}^d$-quasi-open set in $\mathbb{R}^d$ is called $\mathbb{R}^d$-quasi-closed.

c) The function $v \colon \mathbb{R}^d \to \mathbb{R}$ is called quasi-continuous if for all $\varepsilon > 0$ there exists an open set $\Omega_\varepsilon \subset \mathbb{R}^d$ such that $\operatorname{Cap}(\Omega_\varepsilon) < \varepsilon$ and such that $v|_{\mathbb{R}^d \setminus \Omega_\varepsilon}$ is continuous.

Remark 2.7.

1. Compared to Definition 2.1, the domain Ω is replaced by $\mathbb{R}^d$ and $H_0^1(\Omega)$ by $H^1(\mathbb{R}^d)$. The definition of the capacity Cap works with the natural norm on $H^1(\mathbb{R}^d)$. Due to the Poincaré inequality, we could also have taken this norm (with $\mathbb{R}^d$ replaced by Ω) in Definition 2.1 to obtain an equivalent capacity on Ω.
2. The capacity Cap is bounded from above by the capacity cap, i.e., there is a constant $c > 0$ such that $\operatorname{Cap}(E) \leq c \operatorname{cap}(E)$ holds for all subsets $E \subset \Omega$, see [42, Lem. A1]. Furthermore, for $E \subset \Omega$, it holds $\operatorname{cap}(E) = 0$ if and only if $\operatorname{Cap}(E) = 0$, see, e.g., [22, Cor. 2.39]. In particular, quasi-open sets are $\mathbb{R}^d$-quasi-open and quasi-closed sets are $\mathbb{R}^d$-quasi-closed. When we say that a property holds true quasi-everywhere, then this statement is independent of the capacity used.
   In the rest of the paper, we will leave out the prefix $\mathbb{R}^d$ and just say quasi-open also for subsets of $\mathbb{R}^d$ and quasi-continuous also for functions on $\mathbb{R}^d$.
3. Analogous to the result in Lemma 2.4, each $v \in H^1(\mathbb{R}^d)$ has a quasi-continuous representative, see [1, Prop. 6.1.2].

Remark 2.8. From the definition it can be seen that for a quasi-open set $O \subset \mathbb{R}^d$ there exists a decreasing sequence of open sets $\Omega_n \supset O$ such that

$$\lim_{n\to\infty} \operatorname{Cap}(\Omega_n \setminus O) = 0.$$

Remark 2.8 can be seen as a motivation for the following definition, which is taken from [26], as well as many other definitions and results in this subsection.

DEFINITION 2.9. *For a quasi-open set $O \subset \mathbb{R}^d$ we define the space*

$$H_0^1(O) := \bigcap \left\{ H_0^1(\Omega) : \Omega \text{ open}, \Omega \supset O \right\}$$

equipped with the norm of $H^1\left(\mathbb{R}^d\right)$.

Remark 2.10. Let $\Omega \subset \mathbb{R}^d$ be open and let $v \in H^1(\Omega)$. Then it holds $v \in H_0^1(\Omega)$ if and only if there is a quasi-continuous function $\tilde{v}$ on $\mathbb{R}^d$ such that $\tilde{v} = v$ a.e. in Ω and $\tilde{v} = 0$ q.e. on $\mathbb{R}^d \setminus \Omega$, see [22, Thm. 4.5], [16, Lem. 6.1] or [4, Thm. 5.8.5].

Definition 2.9 offers a natural extension of the spaces $H_0^1(\Omega)$ for open sets Ω, which are a subclass of quasi-open sets. Remark 2.10 explains how we can interpret the intersection over function spaces with different domains as in Definition 2.9. We identify $v \in H_0^1(\Omega)$ with its quasi-continuous representative on $\mathbb{R}^d$.

The spaces $H_0^1(O)$ are closed subspaces of $H^1\left(\mathbb{R}^d\right)$ and thus Hilbert spaces.

Formally, elements in $H_0^1(O)$ are defined on all of $\mathbb{R}^d$. The next lemma shows that we can also take a different perspective and derive an analogue to Remark 2.10 for quasi-open sets.

LEMMA 2.11. *Let O be quasi-open. It holds $v \in H_0^1(O)$ if and only if $v \in H^1\left(\mathbb{R}^d\right)$ such that $v = 0$ q.e. on $\mathbb{R}^d \setminus O$.*

Proof. We have to show that

$$\bigcap \left\{ H_0^1(\Omega) : \Omega \text{ open}, \Omega \supset O \right\} = \left\{ v \in H^1\big(\mathbb{R}^d\big) : v = 0 \text{ q.e. on } \mathbb{R}^d \setminus O \right\}.$$

Since for open sets Ω it holds

$$H_0^1(\Omega) = \left\{ v \in H^1\big(\mathbb{R}^d\big) : v = 0 \text{ q.e. on } \mathbb{R}^d \setminus \Omega \right\},$$

the inclusion

$$\left\{ v \in H^1\big(\mathbb{R}^d\big) : v = 0 \text{ q.e. on } \mathbb{R}^d \setminus O \right\} \subset \bigcap \left\{ H_0^1(\Omega) : \Omega \text{ open}, \Omega \supset O \right\}$$

follows.

By Remark 2.8, there exists a sequence of open sets $(\Omega_n)_{n\in\mathbb{N}}$ with $O \subset \Omega_n$ such that

$$\lim_{n\to\infty} \operatorname{Cap}(\Omega_n \setminus O) = 0. \tag{2.1}$$

So if $v \in \bigcap \left\{ H_0^1(\Omega) : \Omega \text{ open}, \Omega \supset O \right\}$, then, in particular, $v \in \bigcap_{n\in\mathbb{N}} H_0^1(\Omega_n)$. This implies $v = 0$ q.e. on $\bigcup_{n\in\mathbb{N}} \left(\mathbb{R}^d \setminus \Omega_n\right)$.

Now, by (2.1) and by monotonicity of the capacity, it follows

$$\operatorname{Cap}\left(\bigcap_{n\in\mathbb{N}} \Omega_n \setminus O \right) = 0$$

and it holds

$$\mathbb{R}^d \setminus O = \left(\bigcup_{n \in \mathbb{N}} \left(\mathbb{R}^d \setminus \Omega_n \right) \right) \cup \left(\bigcap_{n \in \mathbb{N}} \Omega_n \setminus O \right),$$

so we have shown that $v = 0$ q.e. on $\mathbb{R}^d \setminus O$. ∎

Lemma 2.11 can be interpreted as follows: A function v defined on O belongs to $H_0^1(O)$ provided its zero extension belongs to $H^1\left(\mathbb{R}^d\right)$. The above result also shows that if O_1, O_2 are quasi-open sets with $\operatorname{Cap}(O_1 \triangle O_2) = 0$, then $H_0^1(O_1) = H_0^1(O_2)$. Here, $O_1 \triangle O_2 = (O_1 \setminus O_2) \cup (O_2 \setminus O_1)$ denotes the symmetric difference of the sets O_1 and O_2.

We need the following definition from [26]:

Definition 2.12. *We call a collection $\mathcal{O}$ of quasi-open sets a quasi-covering of a set E if there is a countable subfamily of $\mathcal{O}$ such that its union P satisfies* $\operatorname{Cap}(E \setminus P) = 0$.

The following lemma gives an application of quasi-coverings. It is a direct consequence of [26, Lem. 2.4] and [26, Thm. 2.10].

Lemma 2.13. *Let O be quasi-open and $v \in H_0^1(O)$. Suppose that the family $\mathcal{O} = \{O_i\}_{i \in I}$ of quasi-open subsets of O is a quasi-covering of O. Then there is a sequence $(v_n)_{n \in \mathbb{N}} \subset H_0^1(O)$ converging to v such that each v_n is a finite sum of functions in $\bigcup_{i \in I} H_0^1(O_i)$.*

3. Features of the obstacle problem. In this section, we recall important properties of the variational inequality characterizing the obstacle problem.

Let $L \in \mathcal{L}\left(H_0^1(\Omega), H^{-1}(\Omega)\right)$ be a coercive operator, i.e., for some positive constant $\alpha > 0$ it holds

$$\langle Ly, y \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} \geq \alpha \|y\|_{H_0^1(\Omega)}^2.$$

Furthermore, let L be strictly T-monotone, i.e., let L satisfy

$$\langle L(y-z), (y-z)_+ \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} > 0 \tag{3.1}$$

for all $y, z \in H_0^1(\Omega)$ with $(y-z)_+ \neq 0$, see [36, p. 105]. Here, $(y-z)_+ = \sup(0, y-z) \in H_0^1(\Omega)$ is defined q.e. on Ω.

In the rest of this article, we usually write $\langle \cdot, \cdot \rangle$ instead of $\langle \cdot, \cdot \rangle_{H^{-1}(\Omega), H_0^1(\Omega)}$ for the dual pairing between $H_0^1(\Omega)$ and $H^{-1}(\Omega)$. If we mean another pairing with different spaces involved, then we will explicitly write it down.

Let U be a separable partially ordered space and let $f\colon U \to H^{-1}(\Omega)$ be Lipschitz continuous, continuously differentiable and monotone. Let ψ be a quasi upper-semicontinuous function such that the closed convex set

$$K_\psi := \left\{ z \in H_0^1(\Omega) : z \geq \psi \right\}$$

is nonempty. Consider $u \in U$. We deal with the variational inequality

$$\begin{aligned} &\text{Find } y \in K_\psi, \\ &\langle Ly - f(u), z - y \rangle \geq 0 \quad \text{for all } z \in K_\psi. \end{aligned} \tag{3.2}$$

It is well known that this obstacle problem has a unique solution and that the solution operator

$$S\colon U \to H_0^1(\Omega)$$

that assigns the solution of the variational inequality to a given $u \in U$ is Lipschitz continuous, see, e.g., [5], [27], [18].

The next lemma on monotonicity properties of solutions to the variational inequality can be found in [18, Prob. 3, p. 30] and [36, Thm. 5.1]. The result is of crucial importance in our approach.

LEMMA 3.1. *The solution operator $S\colon U \to H_0^1(\Omega)$ for the obstacle problem (3.2) is increasing: Suppose that $u_1, u_2 \in U$ such that $u_1 \geq u_2$. Then it holds $S(u_1) \geq S(u_2)$ a.e. and q.e. in Ω.*

3.1. Differentiability. Let $\mathrm{id}\colon H^{-1}(\Omega) \to H^{-1}(\Omega)$ be the identity mapping on $H^{-1}(\Omega)$. In order to distinguish the solution operator respective to a general f from the solution operator respective to id, we denote the solution operator relative to id by S_{id}.

It is well known that the solution operator S_{id} is directionally differentiable and that the directional derivative in a point u and in direction h, denoted $S'_{\mathrm{id}}(u;h)$, is given as the unique solution to the variational inequality

$$\begin{aligned} &\text{Find } \xi \in \mathcal{K}_{K_\psi}(u), \\ &\langle L\xi - h, z - \xi\rangle \geq 0 \quad \text{for all } z \in \mathcal{K}_{K_\psi}(u). \end{aligned} \tag{3.3}$$

Here,

$$\mathcal{K}_{K_\psi}(u) := \mathcal{T}_{K_\psi}(S_{id}(u)) \cap \mu^\perp \tag{3.4}$$

is called the critical cone. In (3.4), $\mathcal{T}_{K_\psi}(S_{id}(u))$ denotes the tangent cone of K_ψ at $S_{id}(u) \in K_\psi$, and $\mu^\perp = \{z \in H_0^1(\Omega) : \langle \mu, z\rangle = 0\}$ is the annihilator with respect to the functional $\mu = LS_{\mathrm{id}}(u) - u \in H^{-1}(\Omega)$. With the help of capacity theory, one can find the following characterization of the critical cone

$$\begin{aligned} \mathcal{K}_{K_\psi}(u) &= \left\{z \in H_0^1(\Omega) : z \geq 0 \text{ q.e. on } A^{\mathrm{id}}(u), \langle \mu, z\rangle = 0\right\} \\ &= \left\{z \in H_0^1(\Omega) : z \geq 0 \text{ q.e. on } A^{\mathrm{id}}(u), z = 0 \text{ q.e. on } A_s^{\mathrm{id}}(u)\right\}. \end{aligned} \tag{3.5}$$

The active set

$$A^{\mathrm{id}}(u) := \{\omega \in \Omega : S_{\mathrm{id}}(u)(\omega) = \psi(\omega)\}$$

as well as the strictly active set $A_s^{\mathrm{id}}(u) \subset A^{\mathrm{id}}(u)$ are quasi-closed subsets of Ω that are defined up to a set of zero capacity. The second characterization in (3.5) gives an implicit representation of the strictly active set, while it can also be defined explicitly as the fine support of the regular Borel measure associated with $\mu = LS_{\mathrm{id}}(u) - u \in H^{-1}(\Omega)^+$. For details we refer to [30], [9, Sect. 6.4], [42, App. A].

In order to obtain the directional derivative for the composite mapping $S = S_{\mathrm{id}} \circ f$ for an operator $f\colon U \to H^{-1}(\Omega)$ as specified before, we will apply a chain rule for the directional derivatives.

For general directionally differentiable mappings the chain rule might not hold. A stronger form of differentiability is needed. The following definition is taken from [9, Def. 2.45].

Definition 3.2. *Let X, Y be Banach spaces and consider a mapping $T\colon X \to Y$. Then T is directionally differentiable at $x \in X$ in the Hadamard sense (or Hadamard directionally differentiable) if the directional derivative $T'(x;h)$ exists for all $h \in X$ and fulfills*

$$T'(x;h) = \lim_{n\to\infty} \frac{T(x + t_n h_n) - T(x)}{t_n}$$

for all sequences $(t_n)_{n\in\mathbb{N}}$ and $(h_n)_{n\in\mathbb{N}}$ with $t_n \downarrow 0$ and $h_n \to h$.

The following proposition can be found in [9, Prop. 2.49].

Proposition 3.3. *Suppose that $T\colon X \to Y$ is directionally differentiable in x and Lipschitz continuous in a neighborhood of x. Then T is directionally differentiable at x in the Hadamard sense.*

Corollary 3.4. *The solution operator $S_{\mathrm{id}}\colon H^{-1}(\Omega) \to H_0^1(\Omega)$ is directionally differentiable in the Hadamard sense.*

Remark 3.5. Hadamard directional differentiability according Definition 3.2 is equivalent to compact directional differentiability, see [38, Prop. 3.2 and 3.3], which means that the asymptotic expansion is uniform on any compact set of directions. We note that Mignot [30] shows directly compact directional differentiability of the solution operator S_{id}, but in fact it follows also from directional differentiability by Corollary 3.4 and the equivalence of Hadamard directional differentiability and compact directional differentiability.

The following chain rule for directional derivatives holds, see [9, Prop. 2.47].

Lemma 3.6. *Let X, Y, Z be Banach spaces and assume that $T\colon X \to Y$ is directionally differentiable at x and that $R\colon Y \to Z$ is Hadamard directionally differentiable at $T(x)$. Then the composite mapping $R \circ T$ is directionally differentiable at x and the following chain rule holds*

$$(R \circ T)'(x;h) = R'(T(x); T'(x;h)).$$

This means the directional derivative $S'(u;h)$ of $S = S_{\mathrm{id}} \circ f$ in direction $h \in U$ is given by the solution of the variational inequality

$$\begin{aligned}&\text{Find } \xi \in \mathcal{K}_{K_\psi}(f(u))\\ &\langle L\xi - f'(u;h), z - \xi\rangle \geq 0 \quad \text{ for all } z \in \mathcal{K}_{K_\psi}(f(u)),\end{aligned}$$

where

$$\mathcal{K}_{K_\psi}(f(u)) = \left\{z \in H_0^1(\Omega) : z \geq 0 \text{ q.e. on } A(u), z = 0 \text{ q.e. on } A_s(u)\right\}.$$

Here,

$$A(u) := A^{\mathrm{id}}(f(u)) = \{\omega \in \Omega : S(u)(\omega) = \psi(\omega)\}$$

is the active set and $A_s(u) := A_s^{\mathrm{id}}(f(u))$ the strictly active set.

We now specify the behavior of S' in points where S is Gâteaux differentiable. We consider the largest linear subset of the critical cone $\mathcal{K}_{K_\psi}(f(u))$, the set

$$H_0^1(\Omega \setminus A(u)) = \left\{z \in H_0^1(\Omega) : z = 0 \text{ q.e. on } A(u)\right\} \tag{3.6}$$

and the smallest linear superset of $\mathcal{K}_{K_\psi}(f(u))$, the linear hull of $\mathcal{K}_{K_\psi}(f(u))$,

$$H_0^1(\Omega \setminus A_s(u)) = \left\{z \in H_0^1(\Omega) : z = 0 \text{ q.e. on } A_s(u)\right\}. \tag{3.7}$$

Recall that the sets on the right-hand sides of (3.6) and (3.7) coincide with the spaces on the left hand side, since $\Omega \setminus A(u)$ and $\Omega \setminus A_s(u)$ are quasi-open sets, see Definition 2.9 and Lemma 2.11.

Observe that whenever $A(u) = A_s(u)$ holds up to a set of zero capacity, then these sets coincide and S is Gâteaux differentiable in u. This is also known as the strict complementarity condition, see, e.g., [9, Sect. 6.4.4].

The strict complementarity condition is also necessary for Gâteaux differentiability if we consider the identity id on $H^{-1}(\Omega)$ and the corresponding solution operator S_{id}. This does not hold anymore when the range of f in $H^{-1}(\Omega)$ is smaller. Nevertheless, the following lemma holds.

Lemma 3.7. *Suppose that S is Gâteaux differentiable in $u \in U$ and let $h \in U$ be arbitrary. Then the directional derivative $S'(u;h)$ is determined by the solution to the variational equation*

$$\begin{aligned} &\textit{Find } \xi \in H_0^1(D(u)),\\ &\langle L\xi - f'(u;h), z\rangle = 0 \quad \textit{for all } z \in H_0^1(D(u)). \end{aligned} \tag{3.8}$$

Here, $D(u) = \Omega \setminus A(u)$ and $D(u) = \Omega \setminus A_s(u)$ are admissible sets that provide the same solution ξ.

Proof. The assumption that u is a point where S is Gâteaux differentiable implies that for all $h \in U$ the elements $S'(u;h)$ lie in a linear subspace of $\mathcal{K}_{K_\psi}(f(u))$, which means they lie in $H_0^1(\Omega \setminus A(u))$. Thus, for all $h \in U$ it holds

$$\begin{aligned} &S'(u;h) \in H_0^1(\Omega \setminus A(u)),\\ &\langle LS'(u;h) - f'(u;h), z - S'(u;h)\rangle \geq 0 \quad \text{for all } z \in H_0^1(\Omega \setminus A(u)) \subset \mathcal{K}_{K_\psi}(f(u)). \end{aligned}$$

Since $H_0^1(\Omega \setminus A(u))$ is a linear subspace, the variational inequality becomes a variational equation and thus $S'(u;h)$ is determined by the unique solution to the variational equation

$$\begin{aligned} &\text{Find } \xi \in H_0^1(\Omega \setminus A(u)),\\ &\langle L\xi - f'(u;h), z\rangle = 0 \quad \text{for all } z \in H_0^1(\Omega \setminus A(u)). \end{aligned}$$

On the other hand, for all $h \in U$, the element $S'(u;h)$ is clearly contained in $H_0^1(\Omega \setminus A_s(u))$, the linear hull of $\mathcal{K}_{K_\psi}(f(u))$.

We argue that the inequality

$$\langle LS'(u;h) - f'(u;h), z - S'(u;h)\rangle \geq 0$$

is fulfilled for all test functions z from $H_0^1(\Omega \setminus A_s(u))$, and not only from $\mathcal{K}_{K_\psi}(f(u))$.

Fix $z \in \mathcal{K}_{K_\psi}(f(u))$ and take an arbitrary $h \in U$. Then the test function z fulfills the variational inequality for the direction $-h$, namely

$$\langle LS'(u;-h) - f'(u;-h), z - S'(u;-h)\rangle \geq 0$$

or, equivalently,

$$\langle LS'(u;h) - f'(u;h), -z - S'(u;h)\rangle \geq 0.$$

This shows that $-z$ is also an admissible test function for the direction h.

Now, consider an arbitrary $z \in H_0^1(\Omega \setminus A_s(u))$, i.e., $z \in H_0^1(\Omega)$ with $z = 0$ q.e. on $A_s(u)$. Then we can write z as $z = z_+ - z_-$, with $z_+, z_- \in \mathcal{K}_{K_\psi}(f(u))$. Since $\mathcal{K}_{K_\psi}(f(u))$ is a cone, $2z_+$, respectively $2z_-$, are in $\mathcal{K}_{K_\psi}(f(u))$ and it holds

$$\langle LS'(u;h) - f'(u;h), 2z_+ - S'(u;h)\rangle \geq 0$$

and

$$\langle LS'(u;h) - f'(u;h), -2z_- - S'(u;h)\rangle \geq 0$$

for all $h \in U$. Adding up both inequalities and dividing by 2 yields

$$\langle LS'(u;h) - f'(u;h), z - S'(u;h)\rangle \geq 0.$$

Therefore, each $z \in H_0^1(\Omega \setminus A_s(u))$ is a valid test function and $S'(u;h)$ is the unique solution of the variational equation

$$\begin{aligned} &\text{Find}\xi \in H_0^1(\Omega \setminus A_s(u)),\\ &\langle L\xi - f'(u;h), z\rangle = 0 \quad \text{ for all } z \in H_0^1(\Omega \setminus A_s(u)). \end{aligned}$$

In summary, in a Gâteaux point u the unique solutions of the variational equations

$$\begin{aligned} &\text{Find } \xi \in H_0^1(D(u)),\\ &\langle L\xi - f'(u;h), z\rangle = 0 \quad \text{ for all } z \in H_0^1(D(u)) \end{aligned}$$

with $D(u) = \Omega \setminus A(u)$ and $D(u) = \Omega \setminus A_s(u)$ coincide and determine the Gâteaux differential of S in u. □

Remark 3.8. In particular, all linear subsets $H_0^1(D(u))$ with a quasi-open set $D(u)$ satisfying

$$\Omega \setminus A(u) \subset D(u) \subset \Omega \setminus A_s(u)$$

up to a set of capacity zero are admissible. Up to disagreement on a set of capacity zero, there is only one set $D(u)$ satisfying these properties in cases when the strict complementarity condition is fulfilled in u, i.e., in cases where $A(u) = A_s(u)$ holds up to a set of zero capacity.

In the following analysis, we will focus on the set $H_0^1(\Omega \setminus A(u))$. Although the sets $D(u)$ in Lemma 3.7 and Remark 3.8 can be replaced without effect in the system (3.8) for the Gâteaux derivative in respective points, they can have a different Mosco limit when considering the sequences $\left(H_0^1(D(u_n))\right)_{n\in\mathbb{N}}$ for $(u_n)_{n\in\mathbb{N}} \subset D_S$ converging to u. In this regard, the gap between $A_s(u)$ and $A(u)$ is responsible for the richness of the generalized differential in certain points u. We will encounter the notion of Mosco convergence and see how it is connected to convergence of solutions to variational inequalities in the following subsection.

3.2. A stability result. In Lemma 3.7 we have seen that the Gâteaux derivatives of the solution operator S in differentiability points evaluated in a direction h solve a variational equation. Since the Bouligand generalized differential contains limits of Gâteaux derivatives, we thus need to study the convergence behaviour of solutions to variational equations or more generally variational inequalities.

A closer look at the variational equation (3.8) reveals that for a fixed direction $h \in U$ and a sequence of elements $(u_n)_{n\in\mathbb{N}} \subset D_S$ with $\lim_{n\to\infty} u_n = u$, the corresponding variational equations for $S'(u_n; h)$ differ only in the closed convex sets $H_0^1(D(u_n))$. We will see that we can use a well suited convergence result for such systems based on the notion of Mosco convergence for closed convex sets.

The following definition goes back to Umberto Mosco, see also [32].

Definition 3.9 (Mosco convergence). *Let X be a Banach space and denote by $(C_n)_{n\in\mathbb{N}}$ a sequence of nonempty subsets of X. If the sets*

$$\tilde{C}_1 := \{x \in X : x_n \to x \text{ for a sequence } (x_n)_{n\in\mathbb{N}} \text{ satisfying } x_n \in C_n \text{ for every } n \in \mathbb{N}\}$$

and

$$\tilde{C}_2 := \{x \in X : x_k \rightharpoonup x \text{ for a sequence } (x_k)_{k\in\mathbb{N}} \text{ with } x_k \in C_{n_k} \text{ for a subsequence } (C_{n_k})_{k\in\mathbb{N}} \text{ of } (C_n)_{n\in\mathbb{N}}\}$$

coincide, then we say that the sequence $(C_n)_{n\in\mathbb{N}}$ converges to $C := \tilde{C}_1 = \tilde{C}_2$ in the sense of Mosco.

The next result on Mosco limits of monotone sequences can be found in [32, Lem. 1.2, 1.3]:

Lemma 3.10. *Let $(C_n)_{n\in\mathbb{N}}$ be a sequence of nonempty closed convex subsets of a Banach space X.*

i) If $(C_n)_{n\in\mathbb{N}}$ is an increasing sequence of sets, it follows

$$C_n \to \overline{\bigcup_{n\in\mathbb{N}} C_n}.$$

ii) If $(C_n)_{n\in\mathbb{N}}$ is decreasing, it follows

$$C_n \to \bigcap_{n\in\mathbb{N}} C_n.$$

Remark 3.11.

1. Let H be a Hilbert space and denote by $\mathcal{P}_C$ the projection onto the set C. Then the convergence of a sequence $(C_n)_{n\in\mathbb{N}}$ of nonempty closed convex sets in the sense of Definition 3.9 is equivalent to the convergence of the sequence $(\mathcal{P}_{C_n}x)_{n\in\mathbb{N}}$ to $\mathcal{P}_C x$ in the norm of X and for all $x \in X$, i.e., the convergence of $(\mathcal{P}_{C_n})_{n\in\mathbb{N}}$ with respect to the strong operator topology on $\mathcal{L}(H)$, see [36, Thm. 4.3]. The book [3] gives an extensive overview of Mosco convergence and relations to other notions of convergence. In non-reflexive spaces, the slice topology, see [6], is an extension of Mosco convergence.
2. Let $(O_n)_{n\in\mathbb{N}}$ be a sequence of quasi-open subsets of Ω. Then the convergence of $\left(H_0^1(O_n)\right)_{n\in\mathbb{N}}$ to $H_0^1(O)$ in the sense of Mosco is equivalent to the γ-convergence of the sequence $(O_n)_{n\in\mathbb{N}}$ to O, see [10, Prop. 4.5.3 and Rem. 4.5.4] This notion of convergence is also used in [35] and appropriate when investigating the full Bouligand generalized differential involving the weak operator topology, where capacitary measures play a role.

The following proposition is taken from [36, Thm. 4.1]. The result will be our tool for studying convergence of solutions to variational inequalities and thus Gâteaux derivatives in differentiability points.

PROPOSITION 3.12. *Let $L \in \mathcal{L}\left(H_0^1(\Omega), H^{-1}(\Omega)\right)$ be coercive and let $(C_n)_{n\in\mathbb{N}}, C$ be closed convex subsets of $H_0^1(\Omega)$. Assume that $C_n \to C$ in the sense of Mosco and $h_n \to h$ in $H^{-1}(\Omega)$, then the unique solutions of*

$$\begin{aligned}&\text{Find } \xi_n \in C_n,\\ &\langle L\xi_n - h_n, z - \xi_n\rangle \geq 0 \quad \text{ for all } z \in C_n\end{aligned}$$

converge to the solution of the limit problem

$$\begin{aligned}&\text{Find } \xi \in C,\\ &\langle L\xi - h, z - \xi\rangle \geq 0 \text{ for all } z \in C.\end{aligned}$$

4. The set-valued map $u \mapsto H_0^1(\Omega \setminus A(u))$. In this section, we analyze the set-valued map $u \mapsto H_0^1(\Omega \setminus A(u))$, as a preparation to show Mosco convergence of the sets $H_0^1(\Omega \setminus A(u_n))$ for suitable sequences $(u_n)_{n\in\mathbb{N}}$ converging to u.

4.1. A monotone continuity result for the set-valued map. We now establish continuity results for the set-valued map $u \mapsto H_0^1(\Omega \setminus A(u))$.

PROPOSITION 4.1. *Let $u \in U$ be arbitrary and let $(u_n)_{n\in\mathbb{N}}$ be an increasing sequence such that $u_n \uparrow u$. Then for each $v \in H_0^1(\Omega \setminus A(u))$ there is a sequence $(v_n)_{n\in\mathbb{N}}$ satisfying $v_n \in \bigcap_{j=J(n)}^{\infty} H_0^1(\Omega \setminus A(u_j))$ for some natural number $J(n) \in \mathbb{N}$ as well as $v_n \to v$ in $H_0^1(\Omega \setminus A(u))$.*

Proof. We want to show that the sequence $(\Omega \setminus A(u_n))_{n\in\mathbb{N}}$ of quasi-open sets is a quasi-covering of $\Omega \setminus A(u)$, i.e., we have to show that

$$\operatorname{Cap}\left((\Omega \setminus A(u)) \setminus \bigcup_{n\in\mathbb{N}} (\Omega \setminus A(u_n))\right) = \operatorname{Cap}\left((\Omega \setminus A(u)) \cap \bigcap_{n\in\mathbb{N}} A(u_n)\right) = 0. \tag{4.1}$$

Since $u_n \to u$, it follows $S(u_n) \to S(u)$ in $H_0^1(\Omega)$ and therefore, pointwise quasi-everywhere for a subsequence. This means for quasi all $x \in \bigcap_{n\in\mathbb{N}} A(u_n)$ it holds $S(u)(x) = \psi(x)$, i.e., x belongs to $A(u)$ and (4.1) follows. Therefore, the family $(\Omega \setminus A(u_n))_{n\in\mathbb{N}}$ is a quasi-covering of $\Omega \setminus A(u)$.

Let $v \in H_0^1(\Omega \setminus A(u))$. By Lemma 2.13, there exists a sequence $(v_n)_{n\in\mathbb{N}}$ converging to v such that each v_n is a finite sum of functions in $\bigcup_{i\in I} H_0^1(\Omega \setminus A(u_i))$. This means for each $n \in \mathbb{N}$ there is $N(n) \in \mathbb{N}$ and a map $\pi_n \colon \{1, \dots, N(n)\} \to \mathbb{N}$ such that

$$v_n = \sum_{i=1}^{N(n)} g_{\pi_n(i)}^{(n)}$$

with $g_j^{(n)} \in H_0^1(\Omega \setminus A(u_j))$. It holds $v_n \in H_0^1\left(\Omega \setminus \bigcap_{i=1}^{N(n)} A\left(u_{\pi_n(i)}\right)\right)$. We define $J(n) := \max_{i=1,\dots,N(n)} \pi_n(i)$. Since $(A(u_n))_{n\in\mathbb{N}}$ is decreasing, we conclude

$$v_n = 0 \text{ q.e. on } A(u_{J(n)})$$

and therefore also

$$v_n = 0 \text{ q.e. on } \bigcup_{j=J(n)}^{\infty} A(u_j).$$

It follows that

$$v_n \in \bigcap_{j=J(n)}^{\infty} H_0^1(\Omega \setminus A(u_j)),$$

see Lemma 2.11. ∎

Corollary 4.2. *Let $u \in U$ be arbitrary and let $(u_n)_{n\in\mathbb{N}}$ be an increasing sequence such that $u_n \uparrow u$. Then for each $v \in H_0^1(\Omega \setminus A(u))$ there is a sequence $(w_n)_{n\in\mathbb{N}}$ satisfying $w_n \in H_0^1(\Omega \setminus A(u_n))$ for each $n \in \mathbb{N}$ as well as $w_n \to v$ in $H_0^1(\Omega)$.*

Proof. Let $v \in H_0^1(\Omega \setminus A(u))$. Proposition 4.1 yields a sequence $(v_n)_{n\in\mathbb{N}}$ such that, with some natural number $J(n) \in \mathbb{N}$,

$$v_n \in H_0^1(\Omega \setminus A(u_j)) \quad \text{for all } j \geq J(n)$$

holds as well as $\lim_{n\to\infty} v_n = v$. Set

$$j_0 := 0$$

and for $n \geq 1$ set

$$j_n := \max\left(\max_{1\leq n'\leq n} J(n'), j_{n-1} + 1\right).$$

Then $(j_n)_{n\in\mathbb{N}}$ is strictly increasing and it holds

$$v_n \in H_0^1(\Omega \setminus A(u_j)) \quad \text{for all } j \geq j_n.$$

Now, for $i = 1, \ldots, j_1 - 1$, choose

$$w_i \in H_0^1(\Omega \setminus A(u_i)).$$

If $i \geq j_1$, then there is exactly one $n \in \mathbb{N}$ such that $i \in \{j_n, \ldots, j_{n+1} - 1\}$ and we set

$$w_i := v_n$$

in this case. By this process, we find a sequence $(w_n)_{n\in\mathbb{N}}$ with $w_n \in H_0^1(\Omega \setminus A(u_n))$ for all $n \in \mathbb{N}$ and $\lim_{n\to\infty} w_n = v$. ∎

Remark 4.3. It is easy to see that the result from Corollary 4.2 holds also when $u_n \downarrow u$ since in this case the inclusions $H_0^1(\Omega \setminus A(u)) \subset H_0^1(\Omega \setminus A(u_n))$ imply the statement.

If the result held for all sequences $(u_n)_{n\in\mathbb{N}}$ converging to u, we would obtain lower semicontinuity of the set-valued map $u \mapsto H_0^1(\Omega \setminus A(u))$. In this regard, the property in Corollary 4.2 could be called monotone lower semicontinuity.

5. An element of the Bouligand generalized differential. We show the convergence of $(S'(u_n; h))_{n\in\mathbb{N}} \subset H_0^1(\Omega)$ for Gâteaux differentiability points u_n of S converging from below towards u and identify the limit, even when S is not Gâteaux differentiable in u. This will give us an element of the Bouligand generalized differential.

The influence of monotonicity of the sequence $(u_n)_{n\in\mathbb{N}}$ on Mosco convergence of the sequences $\left(H_0^1(D(u_n))\right)_{n\in\mathbb{N}}$, in particular, for the sets $D(u_n) = \Omega \setminus A(u_n)$ and

$D(u_n) = \Omega \setminus A_s(u_n)$, is visualized in Figure 1. The top of Figure 1 shows solutions of the obstacle problem for $u = 0$ and for some values $u_i < 0$ and $u_i > 0$ close to 0. The respective sets $A(u_i)$, $A_s(u_i)$ are shown in the middle and at the bottom of Figure 1.

In $u = 0$, the strict complementarity condition does not hold, the isolated point in $A(0)$ belongs to the set $A(0)$, but is not contained in $A_s(0)$. Note that this point has capacity strictly positive in this one-dimensional case. Therefore, $u = 0$ is a point where the respective solution operator is potentially non-Gâteaux differentiable. In the values shown for $u_i > 0$ and $u_i < 0$, the strict complementarity condition holds and the solution operator is Gâteaux differentiable in u_i. For $u_i > 0$, the solutions $S(u_i)$ lose contact to the obstacle in this point, so the sets $A(u_i)$ have a jump when approaching $A(0)$. Vice versa, the sets $A_s(u_i)$ for $u_i < 0$ do all contain a neighborhood of the isolated point, whereas in the limit case, $A_s(0)$ does not contain this point anymore.

These observations stress that the Mosco convergence of the sets $\left(H_0^1(D(u_n))\right)_{n\in\mathbb{N}}$ is connected to monotonicity properties of the sequences $(u_n)_{n\in\mathbb{N}}$.

Example 5.1. When dealing, e.g., with the sets $(H_0^1(\Omega\setminus A(u_n)))_{n\in\mathbb{N}}$, it can be seen that the Mosco limit will not be $H_0^1(\Omega \setminus A(u))$ for a decreasing sequence $(u_n)_{n\in\mathbb{N}}$ in cases as in Figure 1. One can choose an element $v \in H^1(\mathbb{R}^d)$ with $\{v > 0\} = \Omega \setminus A_s(0)$ up to a set of zero capacity, see [41, Prop. 2.3.14] or [20, Lem. 3.6], and define $v_n := v$ for all $n \in \mathbb{N}$. Then, it holds $v_n \in H_0^1(\Omega \setminus A(u_n))$ for all $n \in \mathbb{N}$ as well as $v_n \to v$. Nevertheless, v is not an element of $H_0^1(\Omega \setminus A(0))$. Therefore, the Mosco limit of the sequence $(H_0^1(\Omega \setminus A(u_n)))_{n\in\mathbb{N}}$ is not $H_0^1(\Omega \setminus A(0))$ (but rather $H_0^1(\Omega \setminus A_s(0))$).

The idea to consider increasing sequences of controls $(u_n)_{n\in\mathbb{N}}$ approximating u in order to deduce Mosco convergence of the sets $\left(H_0^1(\Omega \setminus A(u_n))\right)_{n\in\mathbb{N}}$ towards $H_0^1(\Omega \setminus A(u))$ becomes apparent. This idea is formalized in the following theorem.

Theorem 5.2. *Let $(u_n)_{n\in\mathbb{N}} \subset U$ be a convergent sequence of controls with limit $u \in U$ such that $u_n \uparrow u$. Then $H_0^1(\Omega \setminus A(u_n)) \to H_0^1(\Omega \setminus A(u))$ in the sense of Mosco.*

If, furthermore, S is Gâteaux differentiable in u_n for all $n \in \mathbb{N}$, then $(S'(u_n;\cdot))_{n\in\mathbb{N}}$ converges in the strong operator topology to $\Xi(u;\cdot)$, where, for a given $h \in U$ the element $\Xi(u;h)$ is given by the unique solution of

$$\begin{aligned}
&\textit{Find } \xi \in H_0^1(\Omega \setminus A(u)),\\
&\langle L\xi - f'(u;h), z\rangle = 0 \qquad \textit{for all } z \in H_0^1(\Omega \setminus A(u)).
\end{aligned}$$

Proof. The monotonicity of $(u_n)_{n\in\mathbb{N}}$ implies that $(S(u_n))_{n\in\mathbb{N}}$ is an increasing sequence in $H_0^1(\Omega)$ (compare Lemma 3.1), in particular, the sets $A(u_n)$ are decreasing and, therefore, the sequence $\left(H_0^1(\Omega \setminus A(u_n))\right)_{n\in\mathbb{N}}$ is increasing. Lemma 3.10 implies

$$\lim_{n\to\infty} H_0^1(\Omega \setminus A(u_n)) = \overline{\bigcup_{n\in\mathbb{N}} H_0^1(\Omega \setminus A(u_n))}.$$

We want to show that

$$H_0^1(\Omega \setminus A(u)) = \overline{\bigcup_{n\in\mathbb{N}} H_0^1(\Omega \setminus A(u_n))}.$$

Corollary 4.2 implies the inclusion

$$H_0^1(\Omega \setminus A(u)) \subset \overline{\bigcup_{n\in\mathbb{N}} H_0^1(\Omega \setminus A(u_n))}.$$

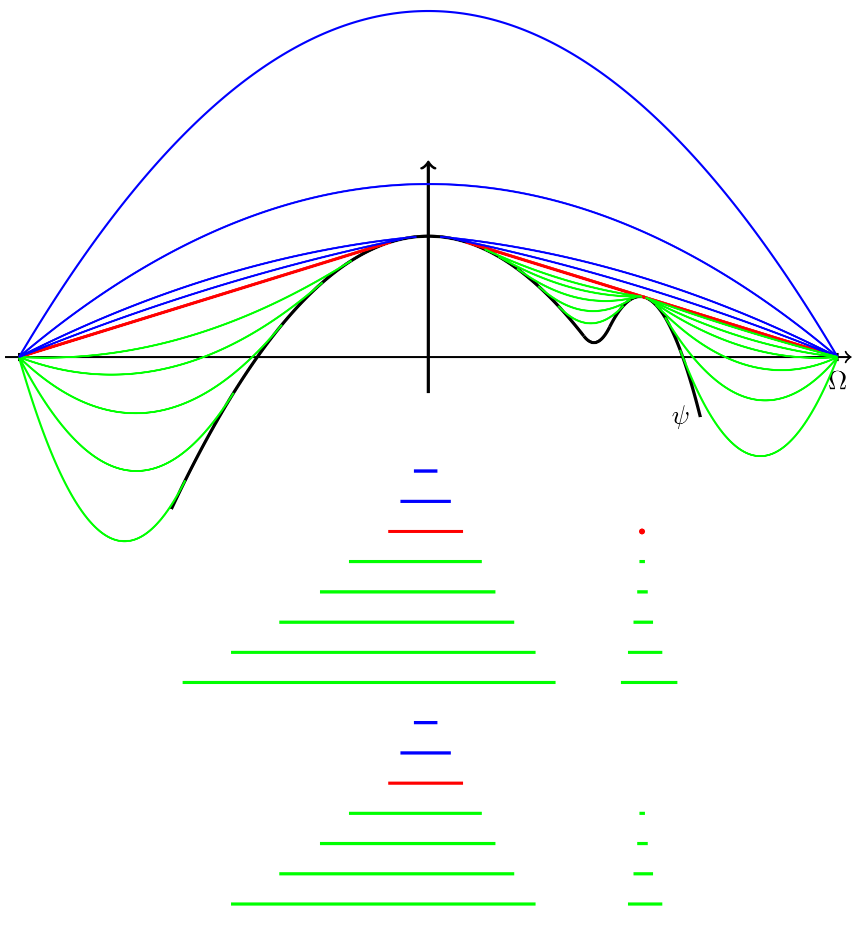

Fig. 1: Top: An instance of the obstacle problem for a piecewise quadratic obstacle ψ. The solution $S(0)$ is plotted in red, while solutions for $S(u)$ with different parameters for $u \leq 0$ are plotted in green and for $u \geq 0$ in blue. Middle: The corresponding active sets $A(u)$ for the different values of u. Bottom: The corresponding strictly active sets $A_s(u)$ for the different values of u.

Let v be an element of $\bigcup_{n\in\mathbb{N}} H_0^1(\Omega \setminus A(u_n))$. Then there is $m > 0$ such that $v \in H_0^1(\Omega \setminus A(u_m))$, i.e., $v = 0$ q.e. on $A(u_m) \supset A(u)$. Thus,

$$\bigcup_{n\in\mathbb{N}} H_0^1(\Omega \setminus A(u_n)) \subset H_0^1(\Omega \setminus A(u))$$

with Lemma 2.11. Since the set on the right-hand side is closed, also

$$\overline{\bigcup_{n\in\mathbb{N}} H_0^1(\Omega \setminus A(u_n))} \subset H_0^1(\Omega \setminus A(u))$$

holds.

All in all, we deduce the Mosco convergence of the sequence $\left(H_0^1(\Omega \setminus A(u_n))\right)_{n\in\mathbb{N}}$ towards $H_0^1(\Omega \setminus A(u))$.

The assertion that for given $h \in U$ the sequence $S'(u_n; h)$ converges to $\Xi(u; h)$ is implied by Proposition 3.12 and Lemma 3.7. By the Banach Steinhaus theorem, the operator $\Xi(u; \cdot)$ is an element of $\mathcal{L}\left(U, H_0^1(\Omega)\right)$ and the second statement of the theorem follows. ∎

5.1. Existence of appropriate Gâteaux differentiability points. In order to infer that the element $\Xi(u; \cdot) \in \mathcal{L}\left(U, H_0^1(\Omega)\right)$ constructed in Theorem 5.2 is in the generalized differential $\partial_B S(u)$ for all $u \in U$, we have to make sure that a sequence of Gâteaux differentiability points $(u_n)_{n\in\mathbb{N}}$ converging from below to a non-Gâteaux differentiability point u always exists. Therefore, we investigate certain subsets of U as exceptional sets of differentiability. Since we want to construct a monotone sequence, the sets we consider are particular subsets of order cones. We recall the following theorem:

Theorem 5.3. *Let X be a separable Banach space and let Y be a Hilbert space. Let S be a Lipschitz function from X into Y. Then the set of points in X in which S is Gâteaux differentiable is dense in X.*

For a proof we refer to [7, Thm. 6.42]. Here, it is shown that the set of points in X where S is not Gâteaux differentiable is Aronszajn null. It can be shown that the complements of Aronszajn null sets are dense in X, see [7, Chap. 6.1, p. 127].

If we assume that X is an Asplund space, it can be shown that even the set of Fréchet differentiability points is dense and that the mean value theorem holds, see [34, Thm. 2.5]. The requirement that Y is a Hilbert space can be weakened, but for simplicity and since it is sufficient for our purposes we state the theorem like that.

The following proposition shows that we can always find an approximating sequence of Gâteaux differentiability points converging from below to a given $u \in U$, so that Theorem 5.2 really yields an element of the Bouligand generalized differential of S in each point u. Therefore, we impose assumptions on the size of the positive cone in the space U, which are all satisfied for the examples $U = L^2(\Omega)$, $U = H^{-1}(\Omega)$ or $U = \mathbb{R}^n$.

Assumption 5.4. We assume that V is an ordered space such that the positive cone $\mathcal{P} := \{v \in V : v \geq 0\}$ has nonempty interior. Let U be a separable ordered space such that V is embedded into U. The embedding $\iota\colon V \to U$ is assumed to be continuous, dense and compatible with the order structures of V and U, i.e., if $v_1, v_2 \in V$ with $v_1 \leq v_2$ then $\iota(v_1) \leq \iota(v_2)$ in U.

Proposition 5.5. *Let Assumption 5.4 be satisfied and let $u \in U$ be arbitrary. Then there is a sequence $(u_n)_{n\in\mathbb{N}}$ such that S is Gâteaux differentiable in each u_n and $u_n \uparrow u$.*

Proof. Fix $u \in U$. Denote by $\mathcal{P}$ the positive cone in V and for $r > 0$ and $v \in V$ denote by $B_r(v)$ the closed ball of radius r around v in V. For short, we write also B_r instead of $B_r(0)$. Let v_0 be an interior point of $\mathcal{P}$. Note that this implies that for all $\lambda > 0$ the element λv_0 is an interior point of $\mathcal{P}$. Without loss of generality, assume that $\|v_0\|_V = 1$.

We define $\vartheta_n := -2^{-n} v_0$ for all $n \in \mathbb{N}$, then each ϑ_n is an element of $-\mathcal{P}$, i.e., $\vartheta_n \leq 0$. We define $\tilde{S}\colon V \to H_0^1(\Omega)$ by

$$\tilde{S}(v) := S(v + u).$$

Since V is continuously embedded into U, the operator $\tilde{S}$ is Lipschitz continuous on V. Therefore, by Theorem 5.3, the set of points in V in which $\tilde{S}$ is Gâteaux differentiable is dense in V. Thus, we find and fix $u_1 \in \vartheta_1 - (\mathcal{P} \cap B_{2^{-1}})$ where $\tilde{S}$ is Gâteaux differentiable.

It holds

$$\|u_1\|_V \leq \|\vartheta_1\|_V + \|u_1 - \vartheta_1\|_V \leq 2^{-1} + 2^{-1} = 1.$$

Now, for $n \geq 2$ assume that we have fixed $u_{n-1} \in \vartheta_{n-1} - \mathcal{P}$ with

$$\|u_{n-1}\|_V \leq 2^{-(n-2)}.$$

We argue that in V, the set

$$(\vartheta_n - \mathcal{P}) \cap (u_{n-1} + \mathcal{P}) \cap B_{2^{-n}}(\vartheta_n) \tag{5.1}$$

has nonempty interior:

Let $\delta_n > 0$ be such that $B_{\delta_n}(2^{-(n+1)}v_0) \subset \mathcal{P}$ holds and let y_n be an arbitrary element of B_{δ_n}. The following arguments show that the element $\vartheta_n - 2^{-(n+1)}v_0 + y_n$ is contained in all the sets that are intersected in (5.1).

It holds $\vartheta_n - (2^{-(n+1)}v_0 - y_n) \leq \vartheta_n$, i.e.,

$$\vartheta_n - 2^{-(n+1)}v_0 + y_n \in (\vartheta_n - \mathcal{P}) \tag{5.2}$$

as well as $2^{-(n+1)}v_0 + y_n \in \mathcal{P}$, i.e.,

$$y_n \geq -2^{-(n+1)}v_0. \tag{5.3}$$

From (5.3), the definition of ϑ_n and from $-u_{n-1} \in -\vartheta_{n-1} + \mathcal{P}$, i.e., $-u_{n-1} \geq -\vartheta_{n-1}$, we conclude that

$$\begin{aligned} \vartheta_n - 2^{-(n+1)}v_0 + y_n &= (u_{n-1} - u_{n-1}) + \vartheta_n - 2^{-(n+1)}v_0 + y_n \\ &\geq u_{n-1} - \vartheta_{n-1} + \vartheta_n - 2^{-(n+1)}v_0 - 2^{-(n+1)}v_0 \\ &= u_{n-1} + 2^{-(n-1)}v_0 - 2^{-n}v_0 - 2^{-n}v_0 \\ &= u_{n-1}. \end{aligned} \tag{5.4}$$

Furthermore, we estimate

$$\|\vartheta_n - 2^{-(n+1)}v_0 + y_n - \vartheta_n\|_V = \|2^{-(n+1)}v_0 - y_n\|_V \leq 2^{-(n+1)} + \delta_n \leq 2^{-n}. \tag{5.5}$$

Here, we have used that $B_{\delta_n}(2^{-(n+1)}v_0) \subset \mathcal{P}$ and $\|v_0\| = 1$ imply $\delta_n \leq 2^{-(n+1)}$.

The arguments in (5.2), (5.4) and (5.5) show that the set in (5.1) has nonempty interior. Thus, we can find a point in the intersection (5.1), denoted u_n, where $\tilde{S}$ is Gâteaux differentiable and which then fulfills

$$0 \geq \vartheta_n \geq u_n \geq u_{n-1}$$

as well as

$$\|u_n\|_V \leq \|\vartheta_n\|_V + \|\vartheta_n - u_n\|_V \leq 2^{-n} + 2^{-n} = 2^{-(n-1)}.$$

Next, we argue that S is Gâteaux differentiable in $u + u_n$ for all $n \in \mathbb{N}$. Since $\tilde{S}$ is just the restriction of S to V composed with a translation of u, the directional derivative $\tilde{S}'(v; h)$ for $v, h \in V$ is given by $S'(v+u; h)$. Let $n \in \mathbb{N}$. Since $\tilde{S}$ is Gâteaux differentiable in u_n, the operator $S'(u_n + u; \cdot)\colon V \to H_0^1(\Omega)$ is linear. Furthermore, $S'(u_n+u; \cdot)$ is continuous on U since it is given as the solution operator of a variational inequality and since $f'(u; \cdot)$ is continuous on U. Now, density of V in U implies that $S'(u_n + u; \cdot)$ is a bounded linear operator.

By construction, $(u_n + u)_{n\in\mathbb{N}}$ is an increasing sequence of Gâteaux differentiability points in U that converges to u. ∎

5.2. Main result. We are now in the position to summarize our results and prove our main result.

THEOREM 5.6. *Assume that U satisfies Assumption 5.4 and let $u \in U$ be arbitrary. Then the operator $\Xi(u;\cdot) \in \mathcal{L}\left(U, H_0^1(\Omega)\right)$, where $\Xi(u;h)$ is given by the unique solution to the variational equation*

$$\begin{aligned} &\textit{Find } \xi \in H_0^1(\Omega \setminus A(u)), \\ &\langle L\xi - f'(u;h), z\rangle = 0 \qquad \textit{for all } z \in H_0^1(\Omega \setminus A(u)), \end{aligned} \tag{5.6}$$

is in the Bouligand generalized differential of S in u.

Proof. Proposition 5.5 yields the existence of a sequence $(u_n)_{n\in\mathbb{N}}$ such that S is Gâteaux differentiable in u_n for each $n \in \mathbb{N}$ with $u_n \uparrow u$ in U.

By Theorem 5.2, for all $h \in U$, the sequence $(S'(u_n;h))_{n\in\mathbb{N}}$ converges to the solution $\Xi(u;h)$ of (5.6). By definition of $\partial_B S(u)$, see Definition 1.2, it holds $\Xi(u;\cdot) \in \partial_B S(u)$. ∎

5.3. Adjoint representation of the subgradient. Let $J\colon H_0^1(\Omega)\times U \to \mathbb{R}$ be a continuously differentiable objective function. We consider an optimization problem with respect to this objective function, which is constrained by our variational inequality

$$\begin{aligned} &\min_{y,u} J(y,u) \\ &\text{subject to } \; y \in K_\psi, \\ &\qquad\qquad \langle Ly - f(u), z - y\rangle \geq 0 \qquad \text{for all } z \in K_\psi. \end{aligned}$$

At this point, let us also recall our model optimal control problem

$$\begin{aligned} &\min_{y,u} \quad J(y,u) = \frac{1}{2}\int_\Omega |y-y_d|^2\,dx + \frac{\alpha}{2}\|u\|_U^2 \\ &\text{subject to} \quad y \in K_\psi, \\ &\qquad\qquad \langle -\Delta y - f(u), z-y\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} \geq 0 \qquad \text{for all } z \in K_\psi \end{aligned}$$

mentioned in the introduction.

The question arises, how an element of Clarke's generalized gradient $\partial_C \hat{J}(u)$ respective to the reduced objective function

$$\hat{J}(u) := J\left(S(u), u\right)$$

can be computed. See [15] for details on Clarke's generalized gradient.

THEOREM 5.7. *Let q be the unique solution of the variational equation*

$$\begin{aligned} &\textit{Find } q \in H_0^1(\Omega \setminus A(u)), \\ &\langle L^* q, v\rangle = \langle J_y\left(S(u),u\right), v\rangle \qquad \textit{for all } v \in H_0^1(\Omega \setminus A(u)). \end{aligned} \tag{5.7}$$

Then it holds

$$f'(u)^* q + J_u\left(S(u),u\right) \in \partial_C \hat{J}(u).$$

Here, J_y and J_u denote the contiuous Fréchet derivatives of J with respect to y and u, respectively, $f'(u)^* \in \mathcal{L}\left(H_0^1(\Omega), U^*\right)$ is the (Banachian) adjoint operator of $f'(u) \in \mathcal{L}\left(U, H^{-1}(\Omega)\right)$ and $L^* \in \mathcal{L}\left(H_0^1(\Omega), H^{-1}(\Omega)\right)$ is the (Banachian) adjoint operator of $L \in \mathcal{L}\left(H_0^1(\Omega), H^{-1}(\Omega)\right)$.

Proof. The coercivity of L^* follows from the coercivity of L, thus the variational equation (5.7) has a unique solution. It holds

$$\partial_C \hat{J}(u) \ni \Sigma^* J_y\left(S(u), u\right) + J_u\left(S(u), u\right)$$

for all $\Sigma \in \partial_B S(u)$, see Remark 1.3, Item 3.

Let $q \in H_0^1(\Omega)$ be the solution to (5.7). We verify that $f'(u)^* q$ fulfills

$$\left\langle \Xi(u;\cdot)^* J_y\left(S(u), u\right), w\right\rangle_{U^*,U} = \left\langle f'(u)^* q, w\right\rangle_{U^*,U} \tag{5.8}$$

for all $w \in U$, where $\Xi\left(u;h\right)$ is given by the solution to

$$\begin{aligned} &\text{Find } \xi \in H_0^1(\Omega \setminus A(u)),\\ &\langle L^* z, \xi\rangle = \langle f'(u;h), z\rangle \quad \text{for all } z \in H_0^1(\Omega \setminus A(u)), \end{aligned} \tag{5.9}$$

see Theorem 5.6.

In fact, for all $w \in U$ it holds

$$\begin{aligned} \left\langle f'(u)^* q, w\right\rangle_{U^*,U} &= \langle f'(u;w), q\rangle \\ &\overset{(5.9)}{=} \langle L^* q, \Xi\left(u;w\right)\rangle \\ &\overset{(5.7)}{=} \left\langle J_y\left(S(u), u\right), \Xi\left(u;w\right)\right\rangle \\ &= \left\langle \Xi(u;\cdot)^* J_y\left(S(u), u\right), w\right\rangle_{U^*,U} \end{aligned}$$

and thus (5.8) holds. ∎

5.4. Extension to shape optimization. We briefly sketch how the presented approach can also be applied to shape optimization problems for the obstacle problem.

Let $\Omega \subset \mathbb{R}^d$ be a bounded, connected domain with C^1-boundary. Moreover, let $U = C^1(\mathbb{R}^d)^d$ be equipped with the norm

$$\|u\|_U = \|u\|_{L^\infty(\mathbb{R}^d)^d} + \|\nabla u\|_{L^\infty(\mathbb{R}^d)^{d,d}}.$$

Consider the set of transformations

$$T_\rho = \{\tau : \mathbb{R}^d \to \mathbb{R}^d : \tau = \mathrm{id}_{\mathbb{R}^d} + u,\ \|u\|_U < \rho\}, \quad U_\rho = \{u \in U :\ \|u\|_U < \rho\}.$$

It can be shown that for $0 < \rho < 1$ small enough all $\tau \in T_\rho$ are order preserving C^1-diffeomorphisms and that the family of domains

$$\Omega_\tau = \tau(\Omega)$$

are bounded domains with C^1-boundary.

Let $\psi \in H^1(\mathbb{R}^d)$ with $\operatorname{ess\,sup}_{\omega\in\Omega} \psi(\omega) > 0$ and such that there exists $\delta > 0$ with $\psi|_{G_\delta} < 0$ on $G_\delta = \{\omega \in \Omega :\ \mathrm{dist}(\omega, \partial\Omega) \le \delta\}$.

Now let $\rho < \delta$ and consider for $\tau \in T_\rho$ and $\Omega_\tau = \tau(\Omega)$ the obstacle problem

$$\begin{aligned} &\text{Find } y_\tau \in K_{\psi,\Omega_\tau},\\ &\langle -\Delta y_\tau, z - \psi\rangle_{H^{-1}(\Omega_\tau), H_0^1(\Omega\tau)} \ge 0 \quad \text{for all } z \in K_{\psi,\Omega_\tau}. \end{aligned} \tag{5.10}$$

where $K_{\psi,\Omega_\tau} = \left\{z \in H_0^1(\Omega_\tau) : z \ge \psi\right\}$.

Let $\nabla\tau = (\tau')^\top$ be the transposed jacobian. If we introduce

$$y^\tau = y_\tau \circ \tau \in H_0^1(\Omega)$$

and the bilinear form $a_\tau : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$,

$$a_\tau(y,z) = \int_\Omega \nabla y^\top \nabla\tau^{-\top}\nabla\tau^{-1}\nabla z \det(\nabla\tau)\, d\omega$$

then transformation of (5.10) to Ω yields the equivalent problem

$$\begin{aligned} &\text{Find } y^\tau \in K_{\psi^\tau},\\ &a_\tau(y^\tau, z-\psi^\tau) \geq 0 \quad \text{ for all } z \in K_{\psi^\tau}, \end{aligned} \tag{5.11}$$

where $\psi^\tau = \psi \circ \tau$ and $K_{\psi^\tau} = \{z \in H_0^1(\Omega) : z \geq \psi^\tau\}$.

Since $a_\tau(y,z)$ is linear in y, z and the coefficient matrix depends smoothly on $\tau, \nabla\tau$, it is easy to see that

$$(u,y,z) \in U_\rho \times H_0^1(\Omega) \times H_0^1(\Omega) \mapsto a_{\mathrm{id}+u}(y,z) \in \mathbb{R}$$

is smooth with derivative

$$\left\langle \frac{d}{d(u,y,z)} a_{\mathrm{id}+u}(y,z), (\delta u, \delta y, \delta z)\right\rangle = a_{\mathrm{id}+u}(\delta y, z) + a_{\mathrm{id}+u}(y,\delta z) + \langle a'_{\mathrm{id}+u}(y,z), \delta u\rangle,$$

where with $\tau = \mathrm{id} + u$

$$\begin{aligned}\langle a'_\tau(y,z), \delta u\rangle = \int_\Omega \nabla y^\top \big(&-\nabla\tau^{-\top}\nabla\delta u^\top \nabla\tau^{-\top}\nabla\tau^{-1} - \nabla\tau^{-\top}\nabla\tau^{-1}\nabla\delta u \nabla\tau^{-1}\\ &+\nabla\tau^{-\top}\nabla\tau^{-1}\mathrm{tr}(\nabla\delta u\nabla\tau^{-1})\big)\nabla z \det(\nabla\tau) d\omega\end{aligned}$$

Let $u \in U_\rho$ and denote by $S(u) = y^{\mathrm{id}+u}$ the solution y^τ of (5.11) for $\tau = \mathrm{id} + u$. Moreover, denote as before by

$$A(u) = \{\omega \in \Omega : S(u)(\omega) = \psi(\omega)\}$$

the active set.

Our aim is to compute an element in the Bouligand generalized differential of the solution operator $u \in U_\rho \mapsto S(u) \in H_0^1(\Omega)$ at $u = 0$. We will sketch how one can show that the solution operator $\Xi(0;\cdot) \in \mathcal{L}\big(U, H_0^1(\Omega)\big)$, where $\xi = \Xi(0;h)$ is given by the unique solution to the variational equation

$$\begin{aligned} &\text{Find } \xi \in H_0^1(\Omega \setminus A(0)),\\ &a_{\mathrm{id}}(\xi, z) + \langle a'_{\mathrm{id}}(S(0), z), h\rangle = 0 \quad \text{ for all } z \in H_0^1(\Omega\setminus A(0)), \end{aligned} \tag{5.12}$$

is an element of the Bouligand generalized differential $\partial S_B(0)$.

To this end, we can proceed as follows.

1. Since we consider the point $u = 0$, we can without restriction choose
$$u \in U_{\rho,0} = \{u \in U_\rho :\, \mathrm{supp}(u) \subset G_\delta\}.$$
2. By using the coercivity of a_{id} one easily deduces that $u \in U_{\rho,0} \mapsto S(u) \in H_0^1(\Omega)$ is Lipschitz-continuous.

3. By T-monotonicity and the properties of ψ one can show that $S(u) \geq 0$ and thus $(S(u)-\psi)|_{\Omega\cap G_\delta} > 0$ for all $u \in U_{\rho,0}$. Hence, the active sets $A(u)$ of $S(u)$ satisfy
$$A(u) \subset \Omega \setminus G_\delta \quad \forall\, u \in U_{\rho,0}$$
and thus $u|_{A(u)} = 0$ and $\psi^\tau|_{A(u)} = \psi|_{A(u)}$ with $\tau = \mathrm{id} + u$.
Hence, for all $u \in U_{\rho,0}$, (5.11) is equivalent to (5.11) with K_ψ instead of K_ψ^τ.
4. Again by T-monotonicity we deduce that
$$\begin{aligned} &u_1, u_2 \in U_{\rho,0} \text{ with } \Omega_{\mathrm{id}+u_1} \subset \Omega_{\mathrm{id}+u_2} \\ &\implies \quad (S(u_1) - S(u_2))|_{\Omega\setminus G_\delta} \leq 0 \text{ and thus } A(u_1) \supset A(u_2). \end{aligned}$$
5. One can show, see [39, Cor. 4.17], that $u \in U_{\rho,0} \mapsto S(u)$ is directionally differentiable and $\xi = S'(u;h)$ is analogously to (3.3) given by
$$\begin{aligned} &\text{Find } \xi \in \mathcal{K}_{K_\psi}(u), \\ &a_{\mathrm{id}+u}(\xi, z-\xi) + \langle a'_{\mathrm{id}+u}(S(u), z-\xi), h\rangle \geq 0 \quad \text{ for all } z \in \mathcal{K}_{K_\psi}(u). \end{aligned}$$
with the critical cone
$$\mathcal{K}_{K_\psi}(u) = \left\{ z \in H_0^1(\Omega) : z \geq 0 \text{ q.e. on } A(u), a_{\mathrm{id}+u}(S(u), z) = 0 \right\}.$$
6. If $u \in U_{\rho,0}$ is a point where S is Gâteaux differentiable then $\xi = S'(u;h)$ is by Lemma 3.7 given by the solution of the variational equation
$$\begin{aligned} &\text{Find } \xi \in H_0^1(\Omega \setminus A(u)), \\ &a_{\mathrm{id}+u}(\xi, z) + \langle a'_{\mathrm{id}+u}(S(u), z), h\rangle = 0 \quad \text{ for all } z \in H_0^1(\Omega \setminus A(u)). \end{aligned} \tag{5.13}$$
7. Now let $N : \partial\Omega \to \mathbb{R}^d$ be a C^1 vector field pointing in the interior of Ω. If Ω has C^2-boundary, we can choose N as the inward pointing unit normal field. The set
$$\{u \in U_{\rho,0} : \, -N(\omega)^\top u(\omega) > 0 \text{ for all } \omega \in \partial\Omega\}$$
has an strictly interior point and by using the ordering $u \leq v$ if $-N^\top u|_{\partial\Omega} \leq -N^\top v|_{\partial\Omega}$ we can argue similarly as in Proposition 5.5 that there exists a sequence $(u_n) \subset U_{\rho,0}$ of point, where S is Gâteaux differentiable, with $u_n \to 0$ in U such that $\Omega_{\mathrm{id}+u_n} \subset \Omega_{\mathrm{id}+u_{n+1}}$ for all n and thus $A(u_n) \supset A(u_{n+1})$ by 4.. Hence, the sequence $(H_0^1(\Omega \setminus A(u_n)))$ is increasing and we obtain as in Theorem 5.2 the Mosco convergence of $(H_0^1(\Omega \setminus A(u_n)))$ to $H_0^1(\Omega \setminus A(0))$. By using [36, Thm. 4.1] and the smooth dependence of $a_{\mathrm{id}+u}$ on u we obtain as in the proof of Theorem 5.2 that the solutions of (5.13) converge to the solution of (5.12).

Acknowledgement. We would like to thank the referees for their constructive comments.

REFERENCES

[1] D. R. Adams and L. I. Hedberg, *Function spaces and potential theory*, vol. 314 of Grundlehren Math. Wiss., Springer-Verlag, Berlin, 1996, https://doi.org/10.1007/978-3-662-03282-4.
[2] S. Adly and L. Bourdin, *Sensitivity analysis of variational inequalities via twice epi-differentiability and proto-differentiability of the proximity operator*, SIAM J. Optim., 28 (2018), pp. 1699–1725, https://doi.org/10.1137/17M1135013.

[3] H. Attouch, *Variational convergence for functions and operators*, Applicable Mathematics Series, Pitman (Advanced Publishing Program), Boston, MA, 1984.

[4] H. Attouch, G. Buttazzo, and G. Michaille, *Variational analysis in Sobolev and BV spaces*, vol. 17 of MOS-SIAM Ser. Optim., SIAM, Philadelphia, PA, second ed., 2014, https://doi.org/10.1137/1.9781611973488. Applications to PDEs and optimization.

[5] V. Barbu, *Optimal control of variational inequalities*, vol. 100 of Research Notes in Mathematics, Pitman, Boston, MA, 1984.

[6] G. Beer, *Topologies on closed and closed convex sets*, vol. 268 of Mathematics and its Applications, Kluwer Academic Publishers Group, Dordrecht, 1993, https://doi.org/10.1007/978-94-015-8149-3.

[7] Y. Benyamini and J. Lindenstrauss, *Geometric nonlinear functional analysis*, vol. 1 of Amer. Math. Soc. Colloq. Publ., Amer. Math. Soc., Providence, RI, 2000.

[8] M. Bergounioux, *Optimal control of an obstacle problem*, Applied Mathematics and Optimization, 36 (1997), pp. 147–172, https://doi.org/10.1007/BF02683341, https://doi.org/10.1007/BF02683341.

[9] J. F. Bonnans and A. Shapiro, *Perturbation analysis of optimization problems*, Springer Ser. Oper. Res., Springer-Verlag, New York, 2000, https://doi.org/10.1007/978-1-4612-1394-9.

[10] D. Bucur and G. Buttazzo, *Variational methods in shape optimization problems*, vol. 65 of Progr. Nonlinear Differential Equations Appl., Birkhäuser, Boston, MA, 2005.

[11] C. Christof, *Sensitivity analysis and optimal control of obstacle-type evolution variational inequalities*, SIAM J. Control Optim., 57 (2019), pp. 192–218, https://doi.org/10.1137/18M1183662, https://doi.org/10.1137/18M1183662.

[12] C. Christof, C. Clason, C. Meyer, and S. Walther, *Optimal control of a non-smooth semilinear elliptic equation*, Math. Control Relat. Fields, 8 (2018), pp. 247–276, https://doi.org/10.3934/mcrf.2018011.

[13] C. Christof and C. Meyer, *Sensitivity analysis for a class of h_0^1-elliptic variational inequalities of the second kind*, Set-Valued and Variational Analysis, (2018), https://doi.org/10.1007/s11228-018-0495-2, https://doi.org/10.1007/s11228-018-0495-2.

[14] C. Christof and G. Wachsmuth, *Differential sensitivity analysis of variational inequalities with locally lipschitz continuous solution operators*, Applied Mathematics & Optimization, (2019), https://doi.org/10.1007/s00245-018-09553-y, https://doi.org/10.1007/s00245-018-09553-y.

[15] F. H. Clarke, *Optimization and nonsmooth analysis*, Classics Appl. Math., SIAM, Philadelphia, PA, reprint ed., 1990, https://doi.org/10.1137/1.9781611971309.

[16] M. C. Delfour and J.-P. Zolésio, *Shapes and geometries. Metrics, analysis, differential calculus, and optimization*, Adv. Des. Control, SIAM, Philadelphia, PA, second ed., 2011, https://doi.org/10.1137/1.9780898719826.

[17] F. Facchinei and J.-S. Pang, *Finite-dimensional variational inequalities and complementarity problems*, vol. 1 of Springer Ser. Oper. Res., Springer-Verlag, New York, NY, 2003, https://doi.org/10.1007/b97543.

[18] A. Friedman, *Variational principles and free-boundary problems*, Robert E. Krieger, Malabar, FL, reprint ed., 1988.

[19] A. Haraux, *How to differentiate the projection on a convex set in Hilbert space. Some applications to variational inequalities*, J. Math. Soc. Japan, 29 (1977), pp. 615–631, https://doi.org/10.2969/jmsj/02940615.

[20] F. Harder and G. Wachsmuth, *Comparison of optimality systems for the optimal control of the obstacle problem*, GAMM-Mitt., 40 (2018), pp. 312–338, https://doi.org/10.1002/gamm.201740004.

[21] J. Haslinger and T. Roubíček, *Optimal control of variational inequalities. Approximation theory and numerical realization*, Appl. Math. Optim., 14 (1986), pp. 187–201, https://doi.org/10.1007/BF01442235.

[22] J. Heinonen, T. Kilpeläinen, and O. Martio, *Nonlinear potential theory of degenerate elliptic equations*, Oxford Math. Monogr., Oxford Univ. Press, New York, 1993.

[23] L. Hertlein and M. Ulbrich, *An inexact bundle algorithm for nonconvex nondifferentiable functions in hilbert space*, Preprint SPP1962-084, 2018, https://spp1962.wias-berlin.de/preprints/084.pdf.

[24] M. Hintermüller and I. Kopacka, *A smooth penalty approach and a nonlinear multigrid algorithm for elliptic MPECs*, Comput. Optim. Appl., 50 (2011), pp. 111–145, https://doi.org/10.1007/s10589-009-9307-9, https://doi.org/10.1007/s10589-009-9307-9.

[25] K. Ito and K. Kunisch, *Optimal control of elliptic variational inequalities*, Appl. Math. Optim., 41 (2000), pp. 343–364, https://doi.org/10.1007/s002459911017, https://doi.org/10.1007/s002459911017.

[26] T. Kilpeläinen and J. Malý, *Supersolutions to degenerate elliptic equations on quasi open sets*, Comm. Partial Differential Equations, 17 (1992), pp. 371–405, https://doi.org/10.1080/03605309208820847.

[27] D. Kinderlehrer and G. Stampacchia, *An introduction to variational inequalities and their applications*, Classics Appl. Math., SIAM, Philadelphia, PA, reprint of the 1980 original ed., 2000, https://doi.org/10.1137/1.9780898719451.

[28] K. Kunisch and D. Wachsmuth, *Path-following for optimal control of stationary variational inequalities*, Comput. Optim. Appl., 51 (2012), pp. 1345–1373, https://doi.org/10.1007/s10589-011-9400-8, https://doi.org/10.1007/s10589-011-9400-8.

[29] C. Meyer, A. Rademacher, and W. Wollner, *Adaptive optimal control of the obstacle problem*, SIAM J. Sci. Comput., 37 (2015), pp. A918–A945, https://doi.org/10.1137/140975863, https://doi.org/10.1137/140975863.

[30] F. Mignot, *Contrôle dans les inéquations variationelles elliptiques*, J. Funct. Anal., 22 (1976), pp. 130–185, https://doi.org/10.1016/0022-1236(76)90017-3.

[31] F. Mignot and J.-P. Puel, *Optimal control in some variational inequalities*, SIAM J. Control Optim., 22 (1984), pp. 466–476, https://doi.org/10.1137/0322028, https://doi.org/10.1137/0322028.

[32] U. Mosco, *Convergence of convex sets and of solutions of variational inequalities*, Adv. Math., 3 (1969), pp. 510–585, https://doi.org/10.1016/0001-8708(69)90009-7.

[33] J. Outrata, M. Kočvara, and J. Zowe, *Nonsmooth approach to optimization problems with equilibrium constraints. Theory, applications and numerical results.*, Nonconvex Optim. Appl., Kluwer Acad., Dordrecht, 1998, https://doi.org/10.1007/978-1-4757-2825-5.

[34] D. Preiss, *Differentiability of Lipschitz functions on Banach spaces*, J. Funct. Anal., 91 (1990), pp. 312–345, https://doi.org/10.1016/0022-1236(90)90147-D.

[35] A.-T. Rauls and G. Wachsmuth, *Generalized derivatives for the solution operator of the obstacle problem*, Set-Valued and Variational Analysis, (2019), https://doi.org/10.1007/s11228-019-0506-y, https://doi.org/10.1007/s11228-019-0506-y.

[36] J.-F. Rodrigues, *Obstacle problems in mathematical physics*, vol. 134, North-Holland, Amsterdam, 1987.

[37] A. Schiela and D. Wachsmuth, *Convergence analysis of smoothing methods for optimal control of stationary variational inequalities with control constraints*, ESAIM Math. Model. Numer. Anal., 47 (2013), pp. 771–787, https://doi.org/10.1051/m2an/2012049, https://doi.org/10.1051/m2an/2012049.

[38] A. Shapiro, *On concepts of directional differentiability*, J. Optim. Theory Appl., 66 (1990), pp. 477–487, https://doi.org/10.1007/BF00940933, https://doi.org/10.1007/BF00940933.

[39] J. Sokolowski and J.-P. Zolésio, *Introduction to shape optimization*, vol. 16 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1992, https://doi.org/10.1007/978-3-642-58106-9, https://doi.org/10.1007/978-3-642-58106-9. Shape sensitivity analysis.

[40] L. Thibault, *On generalized differentials and subdifferentials of Lipschitz vector- valued functions*, Nonlinear Anal., 6 (1982), pp. 1037–1053, https://doi.org/10.1016/0362-546X(82)90074-8.

[41] B. Velichkov, *Existence and regularity results for some shape optimization problems*, vol. 19 of Tesi Sc. Norm. Super. Pisa (N. S.), Ed. Norm., Pisa, 2015, https://doi.org/10.1007/978-88-7642-527-1.

[42] G. Wachsmuth, *Strong stationarity for optimal control of the obstacle problem with control constraints*, SIAM J. Optim., 24 (2014), pp. 1914–1932, https://doi.org/10.1137/130925827.