Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Subdifferentials and penalty approximations of the obstacle problem

Amal Alphonse[1], Gerd Wachsmuth[2]

submitted: January 6, 2025

[1] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: amal.alphonse@wias-berlin.de

[2] Institute of Mathematics
Brandenburgische Technische Universität Cottbus-Senftenberg
03046 Cottbus
Germany
E-Mail: gerd.wachsmuth@b-tu.de

No. 3159
Berlin 2025

2020 *Mathematics Subject Classification.* 49J40, 65K15, 49K21, 49M41, 47B92, 28A12.

Key words and phrases. Variational inequalities, obstacle problem, optimal control, weak operator topology, measures.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Subdifferentials and penalty approximations of the obstacle problem

Amal Alphonse, Gerd Wachsmuth

Abstract

We consider a framework for approximating the obstacle problem through a penalty approach by nonlinear PDEs. By using tools from capacity theory, we show that derivatives of the solution maps of the penalised problems converge in the weak operator topology to an element of the strong-weak Bouligand subdifferential. We are able to treat smooth penalty terms as well as nonsmooth ones involving for example the positive part function $\max(0,\cdot)$. Our abstract framework applies to several specific choices of penalty functions which are omnipresent in the literature. We conclude with consequences to the theory of optimal control of the obstacle problem.

1 Introduction

A ubiquitous method to approximate solutions of the classical obstacle problem

$$u \in H^1_0(\Omega),\ u \le \psi : \langle -\Delta u - f, u - v\rangle \le 0 \qquad \forall v \in H^1_0(\Omega),\ v \le \psi \tag{1}$$

is by penalisation through a nonlinear PDE

$$-\Delta u_\rho + \Lambda_\rho(u_\rho - \psi) = f_\rho. \tag{2}$$

The equation (2) approximates the variational inequality (1) in the sense that its solutions satisfy $u_\rho \to u$ as $\rho \to 0$ provided $f_\rho \to f$. If we define the source-to-solution map $f_\rho \mapsto u_\rho$ of (2) by

$$S_\rho \colon H^{-1}(\Omega) \to H^1_0(\Omega)$$

and consider its derivative at f_ρ in a direction d denoted by $\alpha_\rho := S_\rho'(f_\rho)(d)$, then we know that it satisfies the linearised equation

$$-\Delta \alpha_\rho + \Lambda_\rho'(u_\rho - \psi)(\alpha_\rho) = d. \tag{3}$$

A natural question arises concerning the convergence of the derivatives $\alpha_\rho = S_\rho'(f_\rho)(d)$ in the limit $\rho \to 0$, or more generally, the convergence of the operators $S_\rho'(f_\rho)\colon H^{-1}(\Omega) \to H^1_0(\Omega)$. The possibility that immediately comes to mind is that they converge to the directional derivative of the VI solution map $S\colon u \mapsto f$, i.e., to $S'(f)$. If this were to be the case, the limit $\alpha(d)$ would satisfy

$$\alpha(d) \in \mathcal{K} : \langle -\Delta\alpha(d) - d, \alpha(d) - v\rangle \le 0 \quad \forall v \in \mathcal{K}$$

where $\mathcal{K} := \{v \in H^1_0(\Omega) : v \le 0 \text{ q.e. on } \{u = \psi\}, \langle -\Delta u - f, v\rangle = 0\}$ is the critical cone at $u = S(f)$. In general, this cannot happen since $d \mapsto \alpha(d)$ is nonlinear (unless $\mathcal{K}$ is a subspace), whereas $d \mapsto \alpha_\rho(d)$ is often linear for all ρ (e.g., if Λ_ρ is Fréchet differentiable) and linearity would be preserved in the limit.

DOI 10.20347/WIAS.PREPRINT.3159

In this paper, we consider limits of the operator $S'_\rho(f_\rho)\colon d \mapsto \alpha_\rho$ for different choices of the penalty function Λ_ρ and prove that (under weak conditions) they converge in the weak operator topology to elements of the strong-weak subdifferential of S, defined as

$$\partial_B^{sw} S(f) := \left\{ \begin{aligned} L \in \mathcal{L}(H^{-1}(\Omega), H_0^1(\Omega)) : \exists \{f_n\} \subset F_S : f_n \to f \in H^{-1}(\Omega), \\ S'(f_n) \overset{\text{WOT}}{\to} L \text{ in } \mathcal{L}(H^{-1}(\Omega), H_0^1(\Omega)) \end{aligned} \right\},$$

where F_S is the set of all points in $H^{-1}(\Omega)$ at which S is Gâteaux differentiable and $\overset{\text{WOT}}{\to}$ refers to convergence in the weak operator topology, which we recall now.

Definition 1.1. A sequence $\{L_n\} \subset \mathcal{L}(X, Y)$ of bounded linear operators between Banach spaces X and Y converges to a bounded linear operator $L \in \mathcal{L}(X, Y)$ in the weak operator topology if and only if $L_n x \rightharpoonup Lx$ in Y for all $x \in X$. We write this as $L_n \overset{\text{WOT}}{\to} L$.

Our method of proof relies on a recent characterisation of $\partial_B^{sw} S(f)$ from [18] (see (37) below) involving so-called capacitary measures, which we will introduce in Section 4.2. In order to fulfil one of the conditions to utilise that characterisation (see also Remark 1.4), we assume the following.

Assumption 1.2 (Standing assumption on regularity)**.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set in dimension $n \geq 2$. For the obstacle, we assume $\psi \in C(\bar{\Omega}) \cap H^1(\Omega)$ with either $\psi \in H_0^1(\Omega)$ or $\psi > 0$ on $\partial\Omega$.*

Inspired by tradition and existing literature, we treat in particular the following specific examples[1]:

$$\Lambda_\rho^{\mathrm{m}}(u) = \frac{1}{\rho} u^+ \qquad \text{and} \qquad \Lambda_\rho^{\mathrm{c}}(u) = \frac{1}{\rho}(\rho\bar{\lambda} + u)^+ \tag{4}$$

where $\bar{\lambda} \in L^\infty(\Omega)$, $\bar{\lambda} \geq 0$ is given, and the corresponding smooth versions

$$\Lambda_\rho^{\mathrm{sm}}(u) = \frac{1}{\rho} m_\rho(u) \qquad \text{and} \qquad \Lambda_\rho^{\mathrm{sc}}(u) = \frac{1}{\rho} m_\rho(\rho\bar{\lambda} + u) \tag{5}$$

where m_ρ is a regularisation (satisfying Assumption 2.4) that smooths out the positive part function $(\cdot)^+$, see Lemma 2.7 for some concrete examples. In particular, we have in mind the commonly used regularisations from [9] and [13] respectively, see (13) and (14) for their definitions. We will denote solution maps as well as other maps that depend on the specific choice of Λ with the superscript $\mathrm{m}, \mathrm{c}, \mathrm{sm}$ or sc as appropriate. Note that the choice $\bar{\lambda} \equiv 0$ yields $\Lambda_\rho^{\mathrm{m}} = \Lambda_\rho^{\mathrm{c}}$ and $\Lambda_\rho^{\mathrm{sm}} = \Lambda_\rho^{\mathrm{sc}}$. A typical choice of $\bar{\lambda}$ is $\bar{\lambda} := (f + \Delta\psi)^+$, see [10, Theorem 3.2].

Our penalty term $\Lambda_\rho^{\mathrm{sc}}$ covers also the penalisation

$$\Lambda_\rho^{\widetilde{\mathrm{sc}}}(u) = \tilde{m}_\rho(\bar{\lambda} + (1/\rho)u), \tag{6}$$

(given a regularisation $\tilde{m}_\rho$) which is used frequently in the literature, see, e.g., [14, 13, 20, 12, 10]. Indeed, $\Lambda_\rho^{\widetilde{\mathrm{sc}}}$ is of the form (5) using

$$m_\rho(r) := \rho \tilde{m}_\rho(r/\rho), \tag{7}$$

see Remark 2.6 for more details.

Let us give the main results of this work. We start with the smooth case (5). Below, we use the notation $C_0(\Omega)$ for the set of functions $v \in C(\bar{\Omega})$ with $v = 0$ on $\partial\Omega$.

[1] We have chosen the superscripts m and c in (4) to stand for **m**ax and **c**omplementarity respectively; the reason for the former is clear and the latter is due to the fact that $\Lambda_\rho^{\mathrm{c}}$ is obtained from writing the VI as a complementarity system. The superscripts sm and sc in (5) are supposed to denote that these are **s**moothed versions of m and c respectively.

Theorem 1.3. *For every $f \in H^{-1}(\Omega)$ with $S(f) \in C_0(\Omega)$, if $f_\rho \to f$ in $H^{-1}(\Omega)$, then there exist maps $L^{\mathrm{sm}}, L^{\mathrm{sc}} \in \partial_B^{sw} S(f)$ such that for a subsequence (that we relabel),*

$$(S_\rho^{\mathrm{sm}})'(f_\rho) \overset{\mathrm{WOT}}{\to} L^{\mathrm{sm}}, \qquad \text{and} \qquad (S_\rho^{\mathrm{sc}})'(f_\rho) \overset{\mathrm{WOT}}{\to} L^{\mathrm{sc}}.$$

Theorem 1.3 is a special case of Theorem 4.6 below.

Remark 1.4 (On the assumption $S(f) \in C_0(\Omega)$)**.** *Note that asking for the solution $u = S(f)$ of* (1) *to satisfy $u \in C_0(\Omega)$ (which is needed for the characterisation of [18, Lemma 4.3]) is not too restrictive. For example, when Ω is Lipschitz, [19, Theorem 2.7, §5] guarantees $u \in C(\bar{\Omega}) \cap H_0^1(\Omega)$ if $\psi \in C(\bar{\Omega})$ and $f \in W^{-1,p}(\Omega)$ for $p > n$. An alternative is if Ω satisfies a uniform exterior cone condition and $n \leq 3$, $\psi \in H^1(\Omega)$ with $\Delta\psi \in L^2(\Omega)$ and $f \in L^2(\Omega)$: then u is even Hölder continuous, see [16, Theorem 2.7].*

In the nonsmooth case (4), we additionally have to assume Gâteaux differentiability of the approximations.

Theorem 1.5. *For every sequence $\{f_\rho\} \subset H^{-1}(\Omega)$ such that S_ρ^{m} (S_ρ^{c}) is Gâteaux differentiable at f_ρ and $f_\rho \to f$ with $S(f) \in C_0(\Omega)$, there exists a map $L^{\mathrm{m}} \in \partial_B^{sw} S(f)$ ($L^{\mathrm{c}} \in \partial_B^{sw} S(f)$) such that for a subsequence (that we relabel),*

$$(S_\rho^{\mathrm{m}})'(f_\rho) \overset{\mathrm{WOT}}{\to} L^{\mathrm{m}}, \qquad ((S_\rho^{\mathrm{c}})'(f_\rho) \overset{\mathrm{WOT}}{\to} L^{\mathrm{c}}).$$

Theorem 1.5 is also a special case of Theorem 4.6 below.

The nonsmooth result Theorem 1.5 contains the assumption that each S_ρ is Gâteaux differentiable at f_ρ, which can be justified in the following sense. We rely on the key observation that there exists a dense set $F \subset H^{-1}(\Omega)$ such that S_ρ^{m} and S_ρ^{c} are Gâteaux differentiable from F into $H_0^1(\Omega)$ for every ρ taken from a countable set (see Lemma 4.2). Then the above theorem can always be applied for every constant sequence $f_{\rho_n} \equiv f \in F$ such that $S(f) \in C_0(\Omega)$.

In order to cover these cases without repetition and to generalise the structure of the penalty term as much as possible, we consider an abstract problem formulation, as described in the next subsection. In fact, as mentioned, Theorem 1.3 and Theorem 1.5 are consequences of our more general result Theorem 4.6. Finally, we also mention Theorem 5.1 where we obtain a first-order stationarity condition for an optimal control problem with a VI constraint.

Remark 1.6 (Generalisation to other elliptic operators)**.** *In this paper we consider the elliptic operator in* (1) *and related problems to be the Laplacian because the characterisation of $\partial_B^{sw} S(f)$ from [18] was shown for the Laplacian. Our results up to (but not including) the proof of Theorem 4.6 work in a more general setting where $-\Delta$ is replaced by a linear, bounded, coercive and T-monotone operator $A\colon H_0^1(\Omega) \to H^{-1}(\Omega)$. The estimates below should be adjusted to include the coercivity and boundedness constants.*

2 Abstract setup and properties of the penalised problem

Throughout, we equip the Sobolev space $H_0^1(\Omega)$ with the inner product

$$(u, v)_{H_0^1(\Omega)} := \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x.$$

2.1 Setup

For the results of this section, it suffices to take an obstacle $\psi \in H^1(\Omega)$ with $\psi|_{\partial\Omega} \geq 0$ in the sense that $\min(0, \psi) \in H_0^1(\Omega)$.

Let us now formulate an abstract penalty term. For each $\rho > 0$, we work with a general mapping $\Lambda_\rho \colon H_0^1(\Omega) \to H^{-1}(\Omega)$ defined as a Nemytskii map of a function $\lambda_\rho \colon \Omega \times \mathbb{R} \to \mathbb{R}$, i.e.,

$$\Lambda_\rho(u)(x) := \lambda_\rho(x, u(x)).$$

We make the following standing assumption on λ_ρ.

Assumption 2.1. *We assume that*

(i) for all $\rho \in (0, \infty)$, $\lambda_\rho \colon \Omega \times \mathbb{R} \to \mathbb{R}$ is a Carathéodory function,

(ii) for all $\rho \in (0, \infty)$, $\lambda_\rho(x, \cdot) \colon \mathbb{R} \to \mathbb{R}$ is increasing and convex for a.a. $x \in \Omega$,

(iii) for all $\rho \in (0, \infty)$, there exist $k_\rho, K_\rho \in L^\infty(\Omega)$ and $j_\rho \in L^2(\Omega)$ with $k_\rho \leq K_\rho$ such that

$$\lambda_\rho(x, r) = \begin{cases} 0 & \text{if } r \leq k_\rho(x), \\ \frac{r + j_\rho(x)}{\rho} & \text{if } r \geq K_\rho(x), \end{cases} \tag{8}$$

and, for some $C > 0$,

$$k_\rho \to 0 \text{ in } L^\infty(\Omega) \text{ as } \rho \to 0, \tag{9}$$

$$\|k_\rho\|_{L^2(\Omega)} + \|K_\rho\|_{L^2(\Omega)} \leq C\rho \quad \forall \rho \in (0, 1], \tag{10}$$

(iv) $\lambda_\rho(x, \cdot) \in C^1(\mathbb{R})$ for all $\rho \in (0, \infty)$ and a.a. $x \in \Omega$ or $k_\rho \equiv K_\rho$.

Throughout this work, bearing in mind Assumption 2.1 (iv), we refer to the case that $\lambda_\rho(x, \cdot) \in C^1(\mathbb{R})$ for a.a. $x \in \Omega$ as the "smooth case" and to $k_\rho \equiv K_\rho$ as the "nonsmooth case".

Remark 2.2.

(i) The smoothness $\lambda_\rho(x, \cdot) \in C^1(\mathbb{R})$ is equivalent to $k_\rho(x) < K_\rho(x)$.

(ii) Since we know that $\lambda_\rho(x, \cdot) \geq 0$, we have $\lambda_\rho(x, K_\rho(x)) \geq 0$, which yields

$$K_\rho + j_\rho \geq 0. \tag{11}$$

Lemma 2.3. *Assumption 2.1 is satisfied for the nonsmooth choices in* (4)*.*

Proof. Observe that for $\lambda_\rho^{\mathrm{m}}(r) = \frac{1}{\rho} r^+$,

$$k_\rho = K_\rho = j_\rho = 0,$$

whereas for $\lambda_\rho^{\mathrm{c}}(x, r) = (1/\rho)(\rho\bar{\lambda}(x) + r)^+$, we have

$$k_\rho(x) = K_\rho(x) = -\rho\bar{\lambda}(x), \quad j_\rho(x) = \rho\bar{\lambda}(x).$$

From this information we can verify the claim without difficulty. □

Regarding the smooth cases, let us first give sufficient conditions on the structure of m_ρ that will enable us to verify Assumption 2.1 more easily.

Assumption 2.4. *Let $m_\rho \in C^1(\mathbb{R})$ be given for all $\rho > 0$, such that there exist $\theta_\rho, \Theta_\rho, l_\rho \in \mathbb{R}$ with*

$$m_\rho(r) = 0 \quad \text{for all } r \le \theta_\rho, \qquad m_\rho(r) = r + l_\rho \quad \text{for all } r \ge \Theta_\rho,$$

for all $\rho > 0$ and

$$|\theta_\rho| + |\Theta_\rho| \le C\rho \quad \forall \rho \in (0, 1] \tag{12}$$

for some constant $C > 0$.

Lemma 2.5. *If a function m_ρ satisfies Assumption 2.4, then Assumption 2.1 is satisfied by both $\Lambda_\rho^{\mathrm{sm}}$ and $\Lambda_\rho^{\mathrm{sc}}$ defined as in* (5)*.*

Proof. The first part of Assumption 2.4 directly implies that Assumption 2.1 (i), (ii) and (iv) are satisfied. The remaining part follows from observing that in the $\Lambda_\rho^{\mathrm{sm}}$ case, we have $k_\rho = \theta_\rho$, $K_\rho = \Theta_\rho$ and $j_\rho = l_\rho$, and in the $\Lambda_\rho^{\mathrm{sc}}$ case, we have $k_\rho = \theta_\rho - \rho\bar{\lambda}$, $K_\rho = \Theta_\rho - \rho\bar{\lambda}$ and $j_\rho = \rho\bar{\lambda} + l_\rho$.

□

Remark 2.6. *The penalisation $\Lambda_\rho^{\widetilde{\mathrm{sc}}}$ from* (6) *can be handled too and in fact we can in this situation weaken the last condition of Assumption 2.4: we would need the parameters of $\tilde{m}_\rho$ to satisfy only*

$$|\theta_\rho| + |\Theta_\rho| \le C \quad \forall \rho \in (0, 1].$$

Indeed, if $\tilde{m}_\rho$ is a regularisation with the structure presented in Assumption 2.4 we find that m_ρ defined as in (7) *has associated parameters $\hat{\theta}_\rho := \rho\theta_\rho$, $\hat{\Theta}_\rho := \rho\Theta_\rho$ and $\hat{l}_\rho := \rho l_\rho$, which all contain a helpful factor of ρ.*

Let us now look at some examples of m_ρ that satisfy Assumption 2.4.

Lemma 2.7. *The following choices of m_ρ satisfy Assumption 2.4.*

(i) The global regularisation used, e.g., in [9]:

$$m_\rho(r) := \begin{cases} 0 & \text{if } r \le 0, \\ \frac{r^2}{2\rho} & \text{if } 0 < r < \rho, \\ r - \frac{\rho}{2} & \text{if } r \ge \rho, \end{cases} \tag{13}$$

with $\theta_\rho = 0$, $\Theta_\rho = \rho$, $l_\rho = -\rho/2$.

(ii) The regularisation from [13]:

$$\tilde{m}_\rho(r) := \begin{cases} 0 & \text{if } r \le -\frac{\rho}{2}, \\ \frac{1}{2\rho^3}\left(r + \frac{\rho}{2}\right)^3\left(\frac{3\rho}{2} - r\right) & \text{if } -\frac{\rho}{2} < r < \frac{\rho}{2}, \\ r & \text{if } r \ge \frac{\rho}{2}, \end{cases} \tag{14}$$

with $\theta_\rho = -\frac{\rho}{2}$, $\Theta_\rho = \frac{\rho}{2}$, $l_\rho = 0$.

(iii) The local regularisation from [9]:

$$m^l_\rho(r) := \begin{cases} 0 & \text{if } r \leq -\rho, \\ \frac{r^2}{4\rho} + \frac{r}{2} + \frac{\rho}{4} & \text{if } -\rho < r < \rho, \\ r & \text{if } r \geq \rho, \end{cases}$$

with $\theta_\rho = -\rho$, $\Theta_\rho = \rho$, $l_\rho = 0$.

(iv) The regularisation from [14]:

$$\hat{m}_\rho(r) := \begin{cases} 0 & \text{if } r \leq -\frac{\rho}{2}, \\ \frac{1}{2\rho}\left(r + \frac{\rho}{2}\right)^2 & \text{if } -\frac{\rho}{2} < r < \frac{\rho}{2}, \\ r & \text{if } r \geq \frac{\rho}{2}, \end{cases}$$

with $\theta_\rho = -\frac{\rho}{2}$, $\Theta_\rho = \frac{\rho}{2}$, $l_\rho = 0$.

Hence Λ^{sm}_ρ and Λ^{sc}_ρ as defined in (5) associated to each of the above regularisations satisfy Assumption 2.1.

Proof. We can verify the claim using the information presented below each choice. □

2.2 First properties

Lemma 2.8. *For all $\rho \in (0,\infty)$ and a.a. $x \in \Omega$ the following holds.*

(i) The map $\lambda_\rho(x,\cdot)$ is directionally differentiable. In the nonsmooth case (with $k_\rho \equiv K_\rho$), the directional derivative is given by

$$\lambda'_\rho(x,r)(h) = \frac{1}{\rho}\chi_{\{r=k_\rho\}}(x)h^+ + \frac{1}{\rho}\chi_{\{r>k_\rho\}}(x)h \qquad \forall r,h \in \mathbb{R}. \tag{15}$$

(ii) In the smooth case we have

$$0 \leq \lambda'_\rho(x,r) \leq \frac{1}{\rho} \qquad \forall r \in \mathbb{R}. \tag{16}$$

(iii) We have

$$\lambda'_\rho(x,k_\rho(x))(\alpha)\alpha \geq 0, \quad \lambda'_\rho(x,k_\rho(x))(\alpha)\alpha^+ \geq 0 \qquad \forall \alpha \in \mathbb{R}, \tag{17}$$

$$|\lambda'_\rho(x,r)(h)| \leq \frac{1}{\rho}|h| \qquad \forall r,h \in \mathbb{R}. \tag{18}$$

(iv) The function $\lambda_\rho(x,\cdot)$ is Lipschitz continuous uniformly in x, i.e.,

$$|\lambda_\rho(x,u) - \lambda_\rho(x,v)| \leq \frac{1}{\rho}|u-v| \qquad \forall u,v \in \mathbb{R}. \tag{19}$$

(v) If $\rho \in (0,1]$, we have the growth condition (with the constant C from (10))

$$\|j_\rho\|_{L^2(\Omega)} \leq 2C\rho. \tag{20}$$

(vi) *The map* $\Lambda_\rho \colon L^2(\Omega) \to L^2(\Omega)$ *is well defined, Lipschitz continuous and directionally differentiable with the derivative given by*

$$\Lambda'_\rho(u)(h)(x) = \lambda'_\rho(x, u(x))(h(x)). \tag{21}$$

(vii) *The map* $\Lambda_\rho \colon L^2(\Omega) \to L^2(\Omega)$ *is monotone.*

(viii) *For all* $r \in \mathbb{R}$ *we have*

$$\lambda_\rho(x, r)r^+ \geq -\frac{K_\rho(x)}{\rho} r^+ + \frac{1}{\rho}|r^+|^2. \tag{22}$$

(ix) *If* $\rho \in (0, 1]$*, we have (with the constant* C *from* (10)*)*

$$\|\Lambda_\rho(v - \psi)\|_{L^2(\Omega)} \leq \frac{1}{\rho} \left\|(v - \psi)^+\right\|_{L^2(\Omega)} + C \qquad \forall v \in H^1_0(\Omega). \tag{23}$$

Proof. (i) In the nonsmooth case, we have $k_\rho \equiv K_\rho$. Consequently, the formula for $\lambda_\rho(x, \cdot)$ yields the differentiability everywhere on points other than at k_ρ, where it is directionally differentiable; it is standard to see that (15) is the expression for the derivative.

(ii) The fact that $\lambda_\rho(x, \cdot)$ is convex and differentiable implies that its derivative is increasing, so this follows from the structure of λ_ρ given in Assumption 2.1 (iii).

(iii) Regarding (17), for the nonnegativity, this is clear in the smooth case due to the linearity with respect to the direction and the nonnegativity of the derivative. In the nonsmooth case this follows from the expression (15) for the derivative above.

For the upper bound (18), in the smooth case this follows again by (16). In the nonsmooth case this follows by the expression for the directional derivative above and the fact that the sets appearing in the expression are disjoint.

(iv) Using the mean value theorem for directional derivatives [17, Proposition 2.29] and (18), we obtain the claim.

(v) We note that by (19),

$$\frac{1}{\rho}|K_\rho + j_\rho| = |\lambda_\rho(\cdot, K_\rho) - \lambda_\rho(\cdot, k_\rho)| \leq \frac{1}{\rho}|K_\rho - k_\rho|$$

which implies $|j_\rho| \leq |K_\rho - k_\rho| + |k_\rho|$ whence (20) follows by (10).

(vi) The above Lipschitz property implies

$$|\lambda_\rho(x, u)| = |\lambda_\rho(x, u) - \lambda_\rho(x, k_\rho(x))| \leq \frac{1}{\rho}|u - k_\rho(x)| \tag{24}$$

and thus Λ_ρ maps $L^2(\Omega)$ to $L^2(\Omega)$. Lipschitz continuity also follows easily by the above.

Due to the directional differentiability of $\lambda_\rho(x, \cdot)$ and a simple dominated convergence theorem argument, using the fact that the first derivative is bounded by (18), we obtain directional differentiability.

(vii) This follows from the fact that $\lambda_\rho(x, \cdot)$ is increasing.

(viii) From the basic properties of λ_ρ and (11) we get

$$\lambda_\rho(x, r) \geq \frac{r - K_\rho(x)}{\rho}.$$

Indeed, for $r < K_\rho$, the right-hand side is negative and for $r \geq K_\rho$ this follows from (8) and (11). Now, (22) easily follows.

(ix) Using the monotonicity of $\lambda_\rho(x, \cdot)$, $\lambda_\rho(x, k_\rho(x)) = 0$ and the Lipschitz estimate (19), we get

$$\lambda_\rho(x, v(x) - \psi(x)) \leq \lambda_\rho(x, (v(x) - \psi(x))^+) \leq \frac{1}{\rho}(v(x) - \psi(x))^+ + \frac{1}{\rho}|k_\rho(x)|.$$

Taking the $L^2(\Omega)$-norm and using (10), the inequality (23) follows. □

Next, we address the well posedness of the PDE (2) satisfied by u_ρ.

Lemma 2.9. *For all $\rho > 0$, the solution map $S_\rho \colon H^{-1}(\Omega) \to H^1_0(\Omega)$ of the PDE* (2) *is well defined. If $\rho \in (0, 1]$, we have*

$$\|S_\rho(f_\rho)\|_{H^1_0(\Omega)} \leq \|f_\rho\|_{H^{-1}(\Omega)} + 2\,\|\min(0, \psi)\|_{H^1_0(\Omega)} + C_P C \qquad \forall f_\rho \in H^{-1}(\Omega),$$

where C_P is the constant from Poincaré's inequality and C is from (10)*. Furthermore, $S_\rho \colon H^{-1}(\Omega) \to H^1_0(\Omega)$ is Lipschitz continuous with constant 1.*

Proof. We define the operator $T_\rho \colon H^1_0(\Omega) \to H^1_0(\Omega)$ via

$$T_\rho(u) := (-\Delta)^{-1}\Lambda_\rho(u - \psi) \qquad \forall u \in H^1_0(\Omega).$$

By applying $(-\Delta)^{-1} \colon H^{-1}(\Omega) \to H^1_0(\Omega)$ to (2), we get

$$(\operatorname{Id} + T_\rho)(u_\rho) = (-\Delta)^{-1} f_\rho.$$

Now, it is easy to check that the operator T_ρ is continuous and monotone. Consequently, the operator T_ρ is maximally monotone, see [2, Proposition 20.27]. Thus, we can apply Minty's theorem, see [2, Theorem 21.1 and Proposition 23.8], to obtain that $\operatorname{Id} + T_\rho$ bijective and the inverse has Lipschitz constant 1. Since $-\Delta \colon H^1_0(\Omega) \to H^{-1}(\Omega)$ is an isometric isomorphism, this shows that S_ρ is well defined and has Lipschitz constant 1.

It remains to provide an estimate for $S_\rho(f_\rho)$. We define $\psi_0 := \min(0, \psi)$ and recall that $\psi_0 \in H^1_0(\Omega)$ by our assumption on ψ. We observe $S_\rho(g_\rho) = \psi_0$ for $g_\rho := -\Delta\psi_0 + \Lambda_\rho(\psi_0 - \psi)$. Consequently, the Lipschitzness of S_ρ yields

$$\begin{aligned} \|S_\rho(f_\rho)\|_{H^1_0(\Omega)} &\leq \|S_\rho(f_\rho) - S_\rho(g_\rho)\|_{H^1_0(\Omega)} + \|S_\rho(g_\rho)\|_{H^1_0(\Omega)} \\ &\leq \|f_\rho + \Delta\psi_0 - \Lambda_\rho(\psi_0 - \psi)\|_{H^{-1}(\Omega)} + \|\psi_0\|_{H^1_0(\Omega)} \\ &\leq \|f_\rho\|_{H^{-1}(\Omega)} + 2\,\|\psi_0\|_{H^1_0(\Omega)} + C_P\,\|\Lambda_\rho(\psi_0 - \psi)\|_{L^2(\Omega)}. \end{aligned}$$

The last addend is estimated via (23) by using $(\psi_0 - \psi)^+ = 0$. □

Let us now address the differentiability of S_ρ.

Lemma 2.10. *Given $f_\rho, d \in H^{-1}(\Omega)$, the directional derivative $\alpha_\rho = S'_\rho(f_\rho)(d)$ exists and is the unique solution of the PDE* (3)*, i.e.,*

$$-\Delta\alpha_\rho + \Lambda'_\rho(u_\rho - \psi)(\alpha_\rho) = d$$

and satisfies the bound

$$\|\alpha_\rho\|_{H_0^1(\Omega)} \leq \|d\|_{H^{-1}(\Omega)} .$$

Proof. The directional differentiability and satisfaction of (3) follows by the directional differentiability and monotonicity of Λ_ρ, see e.g. [1, Lemma 6.1] or [4, Theorem 2.2].

For the estimate, we test (3) with α_ρ and obtain

$$\|\alpha_\rho\|^2_{H_0^1(\Omega)} + \int_\Omega \lambda'_\rho(\cdot, u_\rho - \psi)(\alpha_\rho)\alpha_\rho \,\mathrm{d}x = \langle d, \alpha_\rho\rangle.$$

Using the nonnegativity by (17) we get the result.

Uniqueness follows by the monotonicity of $\Lambda'_\rho(u_\rho - \psi)\colon H_0^1(\Omega) \to H^{-1}(\Omega)$, which is a consequence of the fact that the derivative derivative of a monotone map at a particular point is also monotone. □

In what follows, we show that $\Lambda'_\rho(u_\rho - \psi)$ can be represented by a function in $L^\infty(\Omega)$ under the assumption that the directional derivative $S'_\rho(f_\rho)$ is linear.

Lemma 2.11. *If $f_\rho \in H^{-1}(\Omega)$ is given such that $S'_\rho(f_\rho)$ is linear, i.e., $S'_\rho(f_\rho) \in \mathcal{L}(H^{-1}(\Omega), H_0^1(\Omega))$, then for a.a. $x \in \Omega$, $\lambda'_\rho(x,\cdot)$ is differentiable at the point $S_\rho(f_\rho)(x) - \psi(x)$ and the derivative belongs to $[0, 1/\rho]$. Consequently, the operator $\Lambda'_\rho(S_\rho(f_\rho) - \psi)\colon L^2(\Omega) \to L^2(\Omega)$ is linear and can be identified with a function in $L^\infty(\Omega)$.*

Proof. The operator

$$S'_\rho(f_\rho)\colon H^{-1}(\Omega) \to H_0^1(\Omega)$$

is the inverse of

$$-\Delta + \Lambda'_\rho(S_\rho(f_\rho) - \psi)\colon H_0^1(\Omega) \to H^{-1}(\Omega).$$

Since the former operator is linear by assumption, the latter operator is linear as well. Consequently, $\Lambda'_\rho(S_\rho(f_\rho) - \psi)$ is a linear operator. From (21) we infer the differentiability of $\lambda'_\rho(x,\cdot)$ at $S_\rho(f_\rho)(x) - \psi(x)$ and then (15), (16) imply $\lambda'_\rho(\cdot, S_\rho(f_\rho) - \psi) \in L^\infty(\Omega)$. □

In this linear case, we can define a measure μ_ρ via

$$\mu_\rho(B) := \int_B \Lambda'_\rho(S_\rho(f_\rho) - \psi)\,\mathrm{d}x \tag{25}$$

and (3) becomes

$$-\Delta\alpha_\rho + \alpha_\rho\mu_\rho = d, \tag{26}$$

with an appropriate interpretation of the term $\alpha_\rho\mu_\rho$. We will use the notion of capacitary measures to pass to the limit in this equation, see Section 4.2.

2.3 Convergence results

We now consider the limit $\rho \to 0$. The convergence result

$$u_\rho \to u := S(f) \qquad \text{in } H_0^1(\Omega)$$

has been shown for various choices of the penalty function in, e.g., [1, Lemma 3.3 and Lemma 3.5], [12, Theorem 3.1], [10, Theorem 4.1], [11, Theorem 2.1] and [20, Theorem 2.10]. We demonstrate that this convergence already holds under our general setting. In particular, our results are more general than currently available in the literature and are phrased in the natural function spaces for the problem (e.g., the source terms are only assumed to converge in the dual space). We begin with a preparatory lemma.

Lemma 2.12. *For all $f_\rho \in H^{-1}(\Omega)$ and $\rho \in (0,1]$, we have*

$$\left\|(S_\rho(f_\rho) - \psi)^+\right\|_{L^2(\Omega)} \le \frac{\sqrt{\rho}}{2}\Big(\|f_\rho + \Delta\psi\|_{H^{-1}(\Omega)} + C_P C\Big), \tag{27}$$

$$\|\Lambda_\rho(S_\rho(f_\rho) - \psi)\|_{L^2(\Omega)} \le \frac{1}{2\sqrt{\rho}}\Big(\|f_\rho + \Delta\psi\|_{H^{-1}(\Omega)} + C_P C\Big) + C, \tag{28}$$

where C is the constant from (10).

Proof. Denoting $u_\rho := S_\rho(f_\rho)$, we test the equation

$$-\Delta(u_\rho - \psi) + \Lambda_\rho(u_\rho - \psi) = f_\rho + \Delta\psi$$

with $(u_\rho - \psi)^+ \in H_0^1(\Omega)$. One has (by T-monotonicity of the Laplacian)

$$\langle -\Delta(u_\rho - \psi), (u_\rho - \psi)^+\rangle \ge \left\|(u_\rho - \psi)^+\right\|^2_{H^1(\Omega)}$$

and using (22) we obtain

$$\int_\Omega \Lambda_\rho(u_\rho - \psi)(u_\rho - \psi)^+ \,\mathrm{d}x \ge \frac{1}{\rho}\left\|(u_\rho - \psi)^+\right\|^2_{L^2(\Omega)} - \frac{1}{\rho}\int_\Omega K_\rho (u_\rho - \psi)^+ \,\mathrm{d}x,$$

which leads to

$$\begin{aligned}\left\|(u_\rho - \psi)^+\right\|^2_{H_0^1(\Omega)} + \frac{1}{\rho}\left\|(u_\rho - \psi)^+\right\|^2_{L^2(\Omega)} &\le \Big\langle f_\rho + \Delta\psi + \frac{K_\rho}{\rho}, (u_\rho - \psi)^+\Big\rangle \\ &\le \frac{1}{4}\left\|f_\rho + \Delta\psi + \frac{K_\rho}{\rho}\right\|^2_{H^{-1}(\Omega)} + \left\|(u_\rho - \psi)^+\right\|^2_{H_0^1(\Omega)},\end{aligned} \tag{29}$$

where we used Young's inequality. This yields

$$\left\|(u_\rho - \psi)^+\right\|_{L^2(\Omega)} \le \frac{\sqrt{\rho}}{2}\left\|f_\rho + \Delta\psi + \frac{K_\rho}{\rho}\right\|_{H^{-1}(\Omega)}$$

and (27) follows from Poincaré's inequality. Inequality (28) follows with (23). □

Remark 2.13. *By arguing as in the proof of [20, Lemma 2.3] we can obtain a better rate than above if we have additional regularity. Indeed, we can show that if $f_\rho + \Delta\psi + K_\rho/\rho \in L^2(\Omega)$ and $\rho \in (0,1]$, we have that*

$$\left\|(S_\rho(f_\rho) - \psi)^+\right\|_{L^2(\Omega)} \le \rho \left\| f_\rho + \Delta\psi + \frac{K_\rho}{\rho} \right\|_{L^2(\Omega)} \tag{30}$$

$$\|\Lambda_\rho(S_\rho(f_\rho) - \psi)\|_{L^2(\Omega)} \le \left\| f_\rho + \Delta\psi + \frac{K_\rho}{\rho} \right\|_{L^2(\Omega)} + C \tag{31}$$

where C is the constant from (10)*. To see this, we can adapt* (29) *as follows, making use of the assumed L^2 regularity:*

$$\begin{aligned}\left\|(u_\rho - \psi)^+\right\|^2_{H^1_0(\Omega)} + \frac{1}{\rho}\left\|(u_\rho - \psi)^+\right\|^2_{L^2(\Omega)} &\le \left\langle f_\rho + \Delta\psi + \frac{K_\rho}{\rho}, (u_\rho - \psi)^+ \right\rangle \\ &\le \left\| f_\rho + \Delta\psi + \frac{K_\rho}{\rho} \right\|_{L^2(\Omega)} \left\|(u_\rho - \psi)^+\right\|_{L^2(\Omega)},\end{aligned}$$

whence (30)*. The second estimate* (31) *follows as before.*

Proposition 2.14. *Let $f_\rho \to f$ in $H^{-1}(\Omega)$. We have $u_\rho = S_\rho(f_\rho) \to S(f) =: u$ in $H^1_0(\Omega)$, i.e., the limit u solves the VI* (1)*.*

Proof. The estimate in Lemma 2.9 implies that $\{u_\rho\}$ is uniformly bounded in $H^1_0(\Omega)$, hence there exists a $v \in H^1_0(\Omega)$ such that for a subsequence (that we relabel), we have

$$u_\rho \rightharpoonup v \quad \text{in } H^1_0(\Omega).$$

In what follows, we argue that the convergence is strong and that the limit v equals u. Hence, the entire sequence converges towards $u := S(f)$ in $H^1_0(\Omega)$.

We first note that (27) implies the feasibility of v, i.e., $v \le \psi$. Consequently, we can use v in the VI (1) and test the PDE for u_ρ with $u_\rho - u$. Adding the resulting expressions gives

$$\|u - u_\rho\|^2_{H^1_0(\Omega)} \le \langle f - f_\rho, u - u_\rho\rangle + \langle -\Delta u - f, v - u_\rho\rangle - \langle \Lambda_\rho(u_\rho - \psi), u_\rho - u\rangle.$$

It remains to check that the last addend on the right-hand side converges to 0. From $v - \psi + k_\rho \le k_\rho$ we get $\Lambda_\rho(v - \psi + k_\rho) = 0$, see (8). Together with the monotonicity of Λ_ρ, we obtain

$$\begin{aligned}\langle \Lambda_\rho(u_\rho - \psi), u_\rho - v\rangle &= \langle \Lambda_\rho(u_\rho - \psi) - \Lambda_\rho(v - \psi + k_\rho), u_\rho - (v + k_\rho) + k_\rho\rangle \\ &\ge \langle \Lambda_\rho(u_\rho - \psi), k_\rho\rangle.\end{aligned}$$

Combining (28) with (10) yields that the right-hand side converges to 0. This can be used in the above estimate and the claim follows. □

Now, we can revisit the estimate in Lemma 2.12 and improve the rate implied by (27). This will be crucial later on.

Lemma 2.15. *Let $f_\rho \to f$ in $H^{-1}(\Omega)$. Then,*

$$\rho^{-1/2}\left\|(S_\rho(f_\rho) - \psi)^+\right\|_{L^2(\Omega)} \to 0.$$

Proof. Thanks to Proposition 2.14, we have $(u_\rho - \psi)^+ \to 0$ in $H^1_0(\Omega)$ with $u_\rho := S_\rho(f_\rho)$. In the inequality (29), i.e.,

$$\left\|(u_\rho - \psi)^+\right\|^2_{H^1_0(\Omega)} + \frac{1}{\rho}\left\|(u_\rho - \psi)^+\right\|^2_{L^2(\Omega)} \le \left\langle f_\rho + \Delta\psi + \frac{K_\rho}{\rho}, (u_\rho - \psi)^+ \right\rangle,$$

the duality product on the right-hand side clearly goes to zero, see (10). This shows the assertion. □

3 Pointwise characterisation of the limit of the derivatives

In this section, we study the convergence of the (directional) derivatives from Lemma 2.10 and provide properties for the limits.

Proposition 3.1. *There exists an $\alpha \in H^1_0(\Omega)$ such that, for a subsequence (that we relabel) the derivative satisfies*

$$S'_\rho(f_\rho)(d) \rightharpoonup \alpha \quad \text{in } H^1_0(\Omega).$$

Proof. This follows directly from the bound in Lemma 2.10. □

We want to know what properties the limit α satisfies.

3.1 A complementarity condition on α

We start with a simple lemma.

Lemma 3.2. *If $\sigma\colon \mathbb{R} \to \mathbb{R}$ is C^1 and convex, and satisfies $\sigma = 0$ on $(-\infty, r_0]$, we have*

$$\sigma(r) \leq (r - r_0)\sigma'(r) \quad \forall r \in \mathbb{R}.$$

Proof. As σ is a differentiable convex function, it satisfies $\sigma(r) \leq \sigma(s) + (r-s)\sigma'(r)$ for all $r, s \in \mathbb{R}$. The result follows from the fact that $\sigma(r_0) = 0$. □

In the smooth setting where $\lambda_\rho(x,\cdot)$ is C^1, we can take in Lemma 3.2 $\sigma(\cdot) = \lambda_\rho(x,\cdot)$ for fixed x and obtain

$$\lambda_\rho(x,r) \leq (r - k_\rho(x,\rho))\lambda'_\rho(x,r) \quad \forall r \in [k_\rho(x), K_\rho(x)]. \tag{32}$$

Note that this inequality also holds in the nonsmooth case, since this implies $k_\rho \equiv K_\rho$.

Proposition 3.3. *Let $f_\rho \to f$ in $H^{-1}(\Omega)$. With $\xi := f + \Delta u$ and α denoting the weak limit of (a subsequence of) $S'_\rho(f_\rho)(d)$ from Proposition 3.1, we have*

$$\langle \xi, \alpha \rangle = \langle \xi, \alpha^+ \rangle = \langle \xi, \alpha^- \rangle = 0.$$

For a characterisation of this result in terms of the strictly active set, see (39).

Proof. Set $u_\rho := S_\rho(f_\rho)$ and define the sets

$$M_0 = \{u_\rho - \psi \leq k_\rho\}, \quad M_1 = \{k_\rho < u_\rho - \psi \leq K_\rho\}, \quad M_2 = \{K_\rho \leq u_\rho - \psi\}.$$

Note that $M_1 = \emptyset$ in the nonsmooth case. Further, (11) implies

$$u_\rho - \psi + j_\rho \geq 0 \text{ on } M_2. \tag{33}$$

Define $\xi_\rho := f_\rho + \Delta u_\rho$. We have

$$\begin{aligned}
\langle \xi_\rho, \alpha_\rho \rangle &= \int_\Omega \Lambda_\rho(u_\rho - \psi)\alpha_\rho \,\mathrm{d}x \\
&= \int_{M_1} \Lambda_\rho(u_\rho - \psi)\alpha_\rho \,\mathrm{d}x + \int_{M_2} \frac{u_\rho - \psi + j_\rho}{\rho}\alpha_\rho \,\mathrm{d}x \\
&= \int_{M_1} \Lambda_\rho(u_\rho - \psi)\alpha_\rho \,\mathrm{d}x + \int_{M_2} \frac{(u_\rho - \psi + j_\rho)^+}{\rho}\alpha_\rho \,\mathrm{d}x
\end{aligned}$$

DOI 10.20347/WIAS.PREPRINT.3159

with no integral over M_0 because $\lambda_\rho(x,\cdot)$ vanishes on $(-\infty, k_\rho(x)]$ and the positive part is introduced due to (33).

Regarding the first term, using the estimate (32), we have

$$\begin{aligned}\left|\int_{M_1} \Lambda_\rho(u_\rho-\psi)\alpha_\rho\right| \mathrm{d}x &\le \int_{M_1} \lambda_\rho(\cdot, u_\rho-\psi)|\alpha_\rho| \,\mathrm{d}x \\ &\le \int_{M_1} (u_\rho-\psi-k_\rho)\lambda_\rho'(\cdot,u_\rho-\psi)|\alpha_\rho|\,\mathrm{d}x \\ &\le \frac{\|K_\rho-k_\rho\|_{L^2(\Omega)}}{\sqrt{\rho}} \left\|\sqrt{\lambda_\rho'(\cdot,u_\rho-\psi)}\alpha_\rho\chi_{M_1}\right\|_{L^2(\Omega)}\end{aligned}$$

where we used the bounds on λ_ρ' given in (16) to justify the last estimate. As for the second integral, we obtain

$$\left|\int_{M_2} \frac{(u_\rho-\psi+j_\rho)^+}{\rho}\alpha_\rho\,\mathrm{d}x\right| \le \frac{\|(u_\rho-\psi+j_\rho)^+\chi_{M_2}\|_{L^2(\Omega)}}{\sqrt{\rho}} \frac{\|\alpha_\rho\chi_{M_2}\|_{L^2(\Omega)}}{\sqrt{\rho}}.$$

Thus

$$\begin{aligned}|\langle\xi_\rho,\alpha_\rho\rangle| \le& \frac{\|K_\rho-k_\rho\|_{L^2(\Omega)}}{\sqrt{\rho}} \left\|\sqrt{\lambda_\rho'(\cdot,u_\rho-\psi)}\alpha_\rho\chi_{M_1}\right\|_{L^2(\Omega)} \\ &+ \frac{\|(u_\rho-\psi+j_\rho)^+\chi_{M_2}\|_{L^2(\Omega)}}{\sqrt{\rho}} \frac{\|\alpha_\rho\chi_{M_2}\|_{L^2(\Omega)}}{\sqrt{\rho}}.\end{aligned} \tag{34}$$

By (10), the first multiplicand in the first term above vanishes as $\rho \to 0$. Let us show that the first multiplicand in the second term also vanishes. By using the triangle inequality and the fact that $(a+b)^+ \le a^+ + b^+$ for all $a, b \in \mathbb{R}$,

$$\frac{\|(u_\rho-\psi+j_\rho)^+\chi_{M_2}\|_{L^2(\Omega)}}{\sqrt{\rho}} \le \frac{\|(u_\rho-\psi)^+\|_{L^2(\Omega)}}{\sqrt{\rho}} + \frac{\|j_\rho\|_{L^2(\Omega)}}{\sqrt{\rho}}$$

and both terms on the right-hand side vanish by Lemma 2.15 and (20) respectively.

Now we simply need to show that second multiplicand in the two terms on the right-hand side of (34) are bounded. Since α_ρ is bounded, so is $d + \Delta\alpha_\rho$. Hence,

$$\begin{aligned}C \ge |\langle d+\Delta\alpha_\rho, \alpha_\rho\rangle| &= \int_\Omega \Lambda_\rho'(u_\rho-\psi)(\alpha_\rho)\alpha_\rho\,\mathrm{d}x \\ &\ge \int_{M_1} \lambda_\rho'(\cdot,u_\rho-\psi)(\alpha_\rho)\alpha_\rho\,\mathrm{d}x + \int_{M_2} \lambda_\rho'(\cdot,u_\rho-\psi)(\alpha_\rho)\alpha_\rho\,\mathrm{d}x,\end{aligned}$$

where we used that the integral over M_0 is nonnegative due to (17). Now, on M_1, the derivative is linear w.r.t. the direction and on M_2, the derivative is $1/\rho$. Thus, we continue with

$$\begin{aligned}C &\ge \int_{M_1} \lambda_\rho'(\cdot,u_\rho-\psi)\alpha_\rho^2 + \frac{1}{\rho}\int_{M_2} \alpha_\rho^2\,\mathrm{d}x \\ &= \left\|\sqrt{\lambda_\rho'(\cdot,u_\rho-\psi)}\alpha_\rho\chi_{M_1}\right\|_{L^2(\Omega)}^2 + \frac{1}{\rho}\|\alpha_\rho\chi_{M_2}\|_{L^2(\Omega)}^2.\end{aligned}$$

Plugging this information back into (34), we find that $\langle\xi_\rho,\alpha_\rho\rangle \to 0$ and hence, since $\alpha_\rho \rightharpoonup \alpha$ in $H_0^1(\Omega)$ and $\xi_\rho \to \xi$ in $H^{-1}(\Omega)$,

$$\langle\xi,\alpha\rangle = 0.$$

In order to produce the same identity for α^+, one can repeat exactly the same calculation with α_ρ^+ instead of α_ρ. In the final step, we can use that $\alpha_\rho \rightharpoonup \alpha$ in $H_0^1(\Omega)$ implies $\alpha_\rho^+ \rightharpoonup \alpha^+$ in $H_0^1(\Omega)$. This yields the claim. □

3.2 An orthogonality condition on $-\Delta\alpha - d$

In this section we will show that $-\Delta\alpha - d$ satisfies an orthogonality condition involving the coincidence set.

Lemma 3.4. *We have*

$$\langle -\Delta\alpha_\rho - d, v\rangle = 0 \quad \forall v \in H^1_0(\Omega), \quad v = 0 \text{ q.e. on } \{(u_\rho - \psi + \|k_\rho\|_{L^\infty(\Omega)})^- = 0\}.$$

Note that the set $\{(u_\rho - \psi + \|k_\rho\|_{L^\infty(\Omega)})^- = 0\}$ is defined up to sets of capacity zero, since $u_\rho - \psi + \|k_\rho\|_{L^\infty(\Omega)} \in H^1(\Omega)$ is quasicontinuous, see [7, Theorem 6.1, §8].

Proof. Taking $v \in H^1_0(\Omega)$ such that $v = 0$ q.e. on $\{(u_\rho - \psi + \|k_\rho\|_{L^\infty(\Omega)})^- = 0\}$, we have

$$\langle -\Delta\alpha_\rho - d, v\rangle = -\int_{\{u_\rho - \psi \geq k_\rho\}} \Lambda'_\rho(u_\rho - \psi)(\alpha_\rho)v \,\mathrm{d}x \tag{35}$$

because $\Lambda'_\rho(u_\rho - \psi)$ vanishes when $u_\rho - \psi \leq k_\rho$. By definition, we have that $v = 0$ q.e. on the set $\{u_\rho - \psi + \|k_\rho\|_{L^\infty(\Omega)} \geq 0\}$. Note that $u_\rho - \psi + \|k_\rho\|_{L^\infty(\Omega)} \geq u_\rho - \psi - k_\rho$ a.e. and so we have the inclusion

$$\{u_\rho - \psi - k_\rho \geq 0\} \subset \{u_\rho - \psi + \|k_\rho\|_{L^\infty(\Omega)} \geq 0\}.$$

It follows that $v = 0$ a.e. on $\{u_\rho - \psi - k_\rho \geq 0\}$ too. Using this in (35) we obtain the desired statement. □

Proposition 3.5. *We have*

$$\langle -\Delta\alpha - d, v\rangle = 0 \qquad \forall v \in H^1_0(\Omega) : v = 0 \text{ q.e. on } \{u = \psi\}.$$

Proof. By (9) it follows that $u_\rho - \psi + \|k_\rho\|_{L^\infty(\Omega)} \to u - \psi$ in $H^1(\Omega)$. We want to show that $s_\rho := (u_\rho - \psi + \|k_\rho\|_{L^\infty(\Omega)})^-$ converges to $s := (u - \psi)^- = \psi - u$ in capacity.

From Proposition 2.14 and [21, Lemma 2.2], we get that u_ρ converges towards u in capacity, i.e.,

$$\operatorname{cap}(\{|u_\rho - u| \geq \varepsilon\}) \to 0 \quad \text{as } \rho \to 0$$

for all $\varepsilon > 0$. Let $\varepsilon > 0$ be arbitrary. By (9), there exists $\rho_0 > 0$ such that $\rho \leq \rho_0$ implies $\|k_\rho\|_{L^\infty(\Omega)} \leq \frac{\varepsilon}{2}$. Restricting to such sufficiently small ρ, we have

$$|s_\rho - s| \leq |u_\rho + \|k_\rho\|_{L^\infty(\Omega)} - u| \leq |u_\rho - u| + \frac{\varepsilon}{2}.$$

Consequently,

$$\operatorname{cap}(\{|s_\rho - s| \geq \varepsilon\}) \leq \operatorname{cap}(\{|u_\rho - u| \geq \varepsilon/2\}) \to 0 \quad \text{as } \rho \to 0.$$

So we have shown that $s_\rho \to s$ in capacity as desired.

Now, the result of Lemma 3.4 implies

$$\pm(-\Delta\alpha_\rho - d) \in \{v \in H^1_0(\Omega) : v \geq 0 \text{ q.e. on } \Omega \text{ and } v = 0 \text{ q.e. on } \{s_\rho = 0\}\}^\circ$$

with $s_\rho = (u_\rho - \psi + \|k_\rho\|_{L^\infty(\Omega)})^-$ as above. Applying [21, Lemma 2.6] we find

$$\pm(-\Delta\alpha - d) \in \{v \in H^1_0(\Omega) : v \geq 0 \text{ q.e. on } \Omega \text{ and } v = 0 \text{ q.e. on } \{s = 0\}\}^\circ$$

where $s = (u - \psi)^- = \psi - u$ is the limit (w.r.t. convergence in capacity) of s_ρ. This yields the desired statement after using the decomposition $v = v^+ - v^-$. □

Remark 3.6. *The proof of Lemma 3.4 even shows*

$$\langle -\Delta\alpha_\rho - d, v\rangle = 0 \quad \forall v \in H^1_0(\Omega), \quad v = 0 \text{ a.e. on } \{(u_\rho - \psi + \|k_\rho\|_{L^\infty(\Omega)})^- = 0\}.$$

However, one cannot generalise (the proof of) Proposition 3.5 to obtain

$$\langle -\Delta\alpha - d, v\rangle = 0 \qquad \forall v \in H^1_0(\Omega) : v = 0 \text{ a.e. on } \{u = \psi\}.$$

The problem is the following. The set $\{u = \psi\}$ could have measure zero but positive capacity. In this case, the assumption $v = 0$ a.e. on $\{u = \psi\}$ is void. Consequently, v could be constant 1 in a neighbourhood of $\{u = \psi\}$. If further $v_\rho \to v$ in $H^1_0(\Omega)$ is given (this is needed for passing to the limit with $\langle -\Delta\alpha_\rho - d, v\rangle$), we have $v_\rho \to v$ in capacity. Further, $\{u_\rho - \psi + \|k_\rho\|_{L^\infty} \geq 0\}$ could contain an open neighbourhood of $\{u = \psi\}$. Therefore, $v_\rho = 0$ a.e. on $\{u_\rho - \psi + \|k_\rho\|_{L^\infty} \geq 0\}$ implies $v_\rho = 0$ q.e. on $\{u = \psi\}$. This contradicts $v_\rho \to v$ in capacity.

If $\psi \in H^1_0(\Omega)$, the argument above can be simplified and we could also alternatively have argued in a similar way to [1, §7].

4 Generalised derivatives as limits

In the convergence results of the previous section, the direction d was fixed. We need something stronger than this: we would like limiting statements for the derivative when seen as a linear operator, and furthermore, we would like the subsequence that converges to be independent of d.

4.1 Existence of limiting elements

First, we prove that the Gâteaux derivatives converge in the weak operator topology.

Proposition 4.1. *Let a sequence $\{f_\rho\} \subset H^{-1}(\Omega)$ be given such that $f_\rho \to f$ in $H^{-1}(\Omega)$ and such that S_ρ is Gâteaux differentiable at f_ρ for all $\rho > 0$. Then, there exists a map $L \in \mathcal{L}(H^{-1}(\Omega), H^1_0(\Omega))$ such that for a subsequence (that we relabel),*

$$S'_\rho(f_\rho) \overset{\text{WOT}}{\to} L.$$

Proof. The bound of Lemma 2.10 shows that

$$\left\|S'_\rho(f_\rho)\right\|_{\mathcal{L}(H^{-1}(\Omega), H^1_0(\Omega))} \leq 1$$

for all $\rho > 0$. By the separability of $H^1_0(\Omega)$, one can show that the unit ball is sequentially compact in the WOT, see also [4, Remark 3.1]. The result follows. □

Note that the Gâteaux differentiability assumption of the previous result holds automatically in the smooth case; this follows from (21). The question remains whether it can also hold in the nonsmooth case. We can in fact use the generalised Rademacher's theorem of Mignot [15, Theorem 1.2], which tells us that for each ρ, there exists a dense set on which S_ρ is Gâteaux differentiable. However, this dense set depends on ρ, and because we wish to consider limits it would be ideal to have *one* dense set on which S_ρ is Gâteaux for all ρ (or at least for a sequence $\{\rho_n\}$). We cannot simply take the intersection because an intersection of dense sets may not be dense. Thus we need another argument where we essentially embed $\{S_\rho\}$ into an infinite dimensional space, which we give next.

DOI 10.20347/WIAS.PREPRINT.3159

Lemma 4.2. *Let a sequence $\{\rho_n\} \subset (0,1]$ with $\rho_n \to 0$ be given and, for brevity, set $S_n := S_{\rho_n}$. Then, there exists a dense set $F \subset H^{-1}(\Omega)$ such that S_n is Gâteaux differentiable on F for all $n \in \mathbb{N}$.*

We remark that this result also follows from the theory of Aronszajn null sets. Indeed, because S_ρ is Lipschitz, S_ρ is Gâteaux differentiable outside an Aronszajn null set [3, Theorem 6.42] and the countable union of Aronszajn null sets remains Aronszajn null (see the paragraph after Definition 6.23 of [3]). This implies the result. For convenience, however, we give a different and direct proof.

Proof. We denote by $\mathcal{V} := \ell^2(\mathbb{N}; H^1_0(\Omega))$ the set of ℓ^2-summable $H^1_0(\Omega)$-valued sequences. Consider the map $H\colon H^{-1}(\Omega) \to \mathcal{V}$ defined by

$$H(f) := \big(2^{-1/2} S_1(f), 2^{-2/2} S_2(f), \ldots\big),$$

i.e., the nth component is $[H(f)]_n = 2^{-n/2} S_n(f)$. Note that

$$\|H(f)\|^2_{\mathcal{V}} = \sum_{n=1}^{\infty} \left\|2^{-n/2} S_n(f)\right\|^2_{H^1_0(\Omega)} \le C^2 \sum_{n=1}^{\infty} 2^{-n} = C^2 < \infty,$$

using the fact that $\|S_n(f)\|_{H^1_0(\Omega)} \le C$ for a constant C independent of n, see Lemma 2.9. Further, H is Lipschitz, since

$$\|H(f) - H(g)\|^2_{\mathcal{V}} \le \sum_{n=1}^{\infty} 2^{-n} \|(S_n(f) - S_n(g))\|^2_{H^1_0(\Omega)} \le \|f - g\|^2_{H^{-1}(\Omega)},$$

where we used the fact that each S_n is Lipschitz with constant 1, see Lemma 2.9. Since $\mathcal{V}$ is a Hilbert space and $H^{-1}(\Omega)$ is a separable Hilbert space, it follows [15, Theorem 1.2] that H is Gâteaux differentiable on a dense set $F \subset H^{-1}(\Omega)$. Now, it is straightforward to check that the Gâteaux differentiability of H on F implies that S_n is Gâteaux differentiable on F for all $n \in \mathbb{N}$. □

Observe that this result holds for a countable sequence and does not and cannot hold for the original uncountable family $\{S_\rho\}$. This is because the differentiability holds up to a set of measure zero and countable unions (of the exceptional sets) still have measure zero, which of course, does not apply for the uncountable case. A simple real-valued example is $S_\rho(x) := |x - \rho|$. This function is not differentiable at ρ, so if one considers all of these functions, the common differentiability set is just the empty set.

4.2 Capacitary measures and characterisation of limits

In what follows, we assume that the reader is familiar with the notions of capacity, quasi-open and quasi-closed sets and quasicontinuous representatives. For an introduction tailored to optimal control, we refer to [8, Section 3].

The measure μ_ρ defined in (25) has some additional regularity that we can exploit. We begin with recalling the notion of capacitary measures, see, e.g., [18, Definition 3.1].

Definition 4.3 (Capacitary measure)**.** A *capacitary measure* is a Borel measure μ such that

(i) $\mu(B) = 0$ for every Borel set $B \subset \Omega$ with $\operatorname{cap}(B) = 0$

(ii) μ is regular in the sense that every Borel set $B \subset \Omega$ satisfies

$$\mu(B) = \inf\{\mu(O) : O \text{ quasi-open and } \operatorname{cap}(B \setminus O) = 0\}.$$

We denote by $\mathcal{M}_0(\Omega)$ the set of all capacitary measures on Ω.

Note that a measure which is absolutely continuous w.r.t. the Lebesgue measure is a capacitary measure, see also [5, Remark 3.2 and Definition 2.1]. In particular, (25) defines a capacitary measure in the setting of Lemma 2.11.

From now on, given (an equivalence class) $v \in H_0^1(\Omega)$, we will always work with a Borel measurable and quasicontinuous representative. Note that such a representative is uniquely determined up to subsets of capacity zero which are μ-nullsets for every $\mu \in \mathcal{M}_0(\Omega)$.

We denote by $L^2_\mu(\Omega)$ the usual Lebesgue space w.r.t. a measure $\mu \in \mathcal{M}_0(\Omega)$.

For $\mu \in \mathcal{M}_0(\Omega)$ and $f \in H^{-1}(\Omega)$, one can check that there is a unique solution $y \in H_0^1(\Omega) \cap L^2_\mu(\Omega)$ of

$$\langle -\Delta y, v\rangle + \int_\Omega yv \, \mathrm{d}\mu = \langle f, v\rangle \qquad \forall v \in H_0^1(\Omega) \cap L^2_\mu(\Omega). \tag{36}$$

The solution map of this weak formulation is denoted by $L_\mu \colon H^{-1}(\Omega) \to H_0^1(\Omega)$, i.e., $L_\mu f = y$.

Associated to the VI (1), define the *inactive set* (or non-coincidence set) $I := \{u < \psi\}$ and the *active set* (or coincidence set) $A := \{u = \psi\}$. We denote by $A_s \subset A$ the *strictly active set* which can be defined via the quasi-support (or fine support) of the measure $\xi = f + \Delta u$; this is a quasi-closed subset. Note that all these sets are defined up to subsets of capacity zero.

Under Assumption 1.2 and if $S(f) \in C_0(\Omega)$, the strong-weak generalised derivative $\partial_B^{sw} S(f) \subset \mathcal{L}(H^{-1}(\Omega), H_0^1(\Omega))$ of S at f was characterised in [18, Theorem 5.6] and we have

$$\partial_B^{sw} S(f) = \{L_\mu : \mu \in \mathcal{M}_0(\Omega),\ \mu(I) = 0,\ \mu = \infty \text{ on } A_s\}, \tag{37}$$

where $\mu = \infty$ on A_s is to be understood in the sense that

$$v = 0 \quad \text{q.e. on } A_s \text{ for all } v \in H_0^1(\Omega) \cap L^2_\mu(\Omega)$$

(see [18, Lemma 5.2] for some equivalent characterisations). We will show that the limits of the derivatives from Proposition 4.1 belong to this set.

In the linear setting of Lemma 2.11, μ_ρ is a measure and the equation (26) for α_ρ is to be understood in the weak form

$$\langle -\Delta \alpha_\rho, z\rangle + \int_\Omega \alpha_\rho z \, \mathrm{d}\mu_\rho = \langle d, z\rangle \qquad \forall z \in H_0^1(\Omega).$$

Note that (25) implies

$$\int_\Omega \alpha_\rho z \, \mathrm{d}\mu_\rho = \int_\Omega \Lambda'_\rho(u_\rho - \psi)\alpha_\rho z \, \mathrm{d}x.$$

Since $\Lambda'_\rho(u_\rho - \psi)$ belongs to $L^\infty(\Omega)$, see Lemma 2.11, we have $H_0^1(\Omega) \subset L^2_{\mu_\rho}(\Omega)$. Consequently, the above equation is equivalent to $L_{\mu_\rho}(d) = \alpha_\rho$ and this implies

$$L_{\mu_\rho} \equiv S'_\rho(f_\rho). \tag{38}$$

To summarise, if the directional derivative $S'_\rho(f_\rho)$ is linear, we can find a capacitary measure μ_ρ such that (38) holds.

We now define a notion of convergence for measures related to convergence in the weak operator topology (see Definition 1.1) of the associated solution operators L_μ defined in (36).

Definition 4.4. Let $\{\mu_n\}$ and μ be capacitary measures. We say that $\mu_n \overset{\gamma}{\to} \mu$ if and only if $L_{\mu_n} \overset{\text{WOT}}{\to} L_\mu$.

The sequential compactness of the space $\mathcal{M}_0(\Omega)$, see, e.g., [6, Theorem 4.14], is a crucial property which implies the next result.

Proposition 4.5. *Let a sequence $\{f_\rho\}$ be given such that $S'_\rho(f_\rho)$ is linear for all ρ. We denote by μ_ρ the associated capacitary measure, see* (25)*. Then, there exists a capacitary measure $\mu \in \mathcal{M}_0(\Omega)$ such that $\mu_\rho \overset{\gamma}{\to} \mu$ (for a subsequence that we have relabelled). Equivalently, the derivatives satisfy $S'_\rho(f_\rho) \overset{\text{WOT}}{\to} L_\mu$.*

The key observation that allows us to prove the next theorem, which is the main result, is the following. Proposition 4.5 implies, under the stated assumptions, that there is a subsequence (which we shall relabel) such that for every $d \in H^{-1}(\Omega)$, $\alpha_\rho := S'_\rho(f_\rho)d \rightharpoonup L_\mu d =: \alpha$. We emphasise that the subsequence is independent of d. Thus, all of the results of Section 3 can be applied to obtain information on α and, consequently, on the limiting operator L_μ.

We come now to our main result, which concatenates the above results and characterises the limiting elements obtained in Proposition 4.1. We will use the fact that the result of Proposition 3.3 is equivalent [8, Lemma 3.7] to

$$\alpha = 0 \quad \text{q.e. on } A_s. \tag{39}$$

Theorem 4.6. *Let $f_\rho \to f$ in $H^{-1}(\Omega)$ with $S(f) \in C_0(\Omega)$ be given. We further assume that S_ρ is Gâteaux differentiable at f_ρ for all $\rho > 0$. Then there exists a map $L \in \partial_B^{sw} S(f)$ such that, for a subsequence (that we relabel),*

$$S'_\rho(f_\rho) \overset{\text{WOT}}{\to} L.$$

Proof. Recalling the characterisation

$$\partial_B^{sw} S(f) = \{L_\mu : \mu \in \mathcal{M}_0(\Omega),\ \mu(I) = 0,\ \mu = \infty \text{ on } A_s\},$$

we will show that the operator L_μ from Proposition 4.5 belongs to the set on the right-hand side.

We choose a nonnegative $v \in H_0^1(\Omega)$ such that $\{v > 0\} = I$. Further, set $d := -\Delta v$ and $\alpha := L_\mu d$, i.e.,

$$-\Delta\alpha + \mu\alpha = -\Delta v.$$

Testing with α we get $\|\nabla\alpha\|_{L^2} \le \|\nabla v\|_{L^2}$. Consequently,

$$\langle -\Delta(\alpha - v), \alpha - v\rangle \le 2\langle -\Delta v, v\rangle + 2\langle \Delta\alpha, v\rangle = 2\langle d, v\rangle + 2\langle \Delta\alpha, v\rangle = 0$$

with the final equality by Proposition 3.5. Thus, $\alpha = v$. Testing the equation for α by v, we get

$$\int_\Omega v^2 \, \mathrm{d}\mu = 0$$

and this means $\mu(I) = 0$.

Now, we choose $d := 1$ and $\alpha := L_\mu d$. Note that this α coincides with $w_\mu := L_\mu(1)$ defined below Theorem 3.8 of [18]. From the characterisation (39) of the result of Proposition 3.3, we get $\alpha = w_\mu = 0$ q.e. on A_s and by [18, Lemma 5.2] we get $\mu = +\infty$ on A_s. □

In the nonsmooth case, Theorem 4.6 and Lemma 4.2 imply that there exists a dense subset $F \subset H^{-1}(\Omega)$ such that if $f \in F$ and $S(f) \in C_0(\Omega)$, there exists a map $L \in \partial_B^{sw} S(f)$ such that, for a subsequence (that we relabel),

$$S_n'(f) \overset{\text{WOT}}{\to} L.$$

5 Optimal control of the obstacle problem

We apply our findings to the optimal control of the obstacle problem. Let $F_{ad} \subset L^2(\Omega)$ be a nonempty, closed and convex set satisfying

$$S(f) \in C_0(\Omega) \qquad \forall f \in F_{ad}$$

(see Remark 1.4) and let $J\colon H_0^1(\Omega) \times L^2(\Omega) \to \mathbb{R}$ be a given objective function satisfying

(i) J is continuously Fréchet differentiable with partial derivatives J_y and J_f,

(ii) If $y_n \to y$ in $H_0^1(\Omega)$ and $f_n \rightharpoonup f$ in $L^2(\Omega)$, then

$$J(y, f) \le \liminf_{n\to\infty} J(y_n, f_n).$$

An example of J satisfying the above conditions is

$$J(y, f) := \frac{1}{2} \|y - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2} \|f\|^2_{L^2(\Omega)},$$

where $y_d \in L^2(\Omega)$ is a given desired state and $\nu \ge 0$ is a constant.

Consider the optimal control problem

$$\min_{f \in F_{ad}} J(y, f) \quad \text{such that} \quad y = S(f). \tag{40}$$

In the next result, we will derive a first-order optimality condition for this control problem. This rigorously shows the satisfaction of [18, (19)], which was derived only formally there.

Theorem 5.1. *For any local minimiser $(\bar{y}, \bar{f}) \in H_0^1(\Omega) \times F_{ad}$ of (40), there exists $L \in \partial_B^{sw} S(\bar{f})$ such that*

$$0 \in L^* J_y(\bar{y}, \bar{f}) + J_f(\bar{y}, \bar{f}) + N_{F_{ad}}(\bar{f})$$

is satisfied where $N_{F_{ad}}(\bar{f})$ is the normal cone to F_{ad} at $\bar{f}$.

Proof. We denote by $\varepsilon > 0$ the radius of optimality, i.e.,

$$J(S(f), f) \ge J(\bar{y}, \bar{f}) \qquad \forall f \in F_{ad} \cap B_\varepsilon(\bar{f}),$$

where $B_\varepsilon(\bar{f})$ is a closed ball in $L^2(\Omega)$.

We regularise the problem (40) and consider

$$\min_{f \in F_{ad} \cap B_\varepsilon(\bar{f})} J(S_\rho(f), f) + \frac{1}{2} \left\|f - \bar{f}\right\|^2_{L^2(\Omega)},$$

where we take S_ρ to be S_ρ^{sm} or S_ρ^{sc}. Denote by f_ρ a global minimiser of this problem and set $y_\rho := S_\rho(f_\rho)$.

Using standard arguments, one can show $f_\rho \to \bar{f}$ in $L^2(\Omega)$. Consequently, the constraint $f_\rho \in B_\varepsilon(\bar{f})$ is not binding for small enough ρ. By the standard minimisation principle, we get

$$0 \in \bar{J}'(f_\rho) + (f_\rho - \bar{f}) + N_{F_{ad}}(f_\rho)$$

where $\bar{J}(f) := J(S_\rho(f), f)$ defines the reduced functional. Using the chain rule and the fact that J is Fréchet, we can write this as

$$0 \in S_\rho'(f_\rho)^* J_y(y_\rho, f_\rho) + J_f(y_\rho, f_\rho) + (f_\rho - \bar{f}) + N_{F_{ad}}(f_\rho) \quad \text{in } L^2(\Omega).$$

Note that $J_y(y_\rho, f_\rho) \in H^{-1}(\Omega)$ and $S_\rho'(f_\rho)^* \colon H^{-1}(\Omega) \to H_0^1(\Omega) \subset L^2(\Omega)$.

It remains to pass to the limit with this optimality condition. From Proposition 2.14 we get $y_\rho \to \bar{y}$ in $H_0^1(\Omega)$. Further, Theorem 4.6 enables us to select a subsequence (which we relabel) such that $S_\rho'(f_\rho) \overset{\text{WOT}}{\to} L$ for some $L \in \partial_B^{sw} S(\bar{f})$. Further, we note that $S_\rho'(f_\rho)$ and L are self-adjoint, which yields $S_\rho'(f_\rho)^* \overset{\text{WOT}}{\to} L^*$ Together with the product rule [18, Lemma 2.9 (ii)] we get

$$S_\rho'(f_\rho)^* J_y(y_\rho, f_\rho) + J_f(y_\rho, f_\rho) + (f_\rho - \bar{f}) \rightharpoonup L^* J_y(\bar{y}, \bar{f}) + J_f(\bar{y}, \bar{f})$$

in $H_0^1(\Omega)$ and, consequently, strongly in $L^2(\Omega)$. Since the graph of the normal cone map is closed, this implies

$$0 \in L^* J_y(\bar{y}, \bar{f}) + J_f(\bar{y}, \bar{f}) + N_{F_{ad}}(\bar{f})$$

as claimed. □

Thanks to this result, [18, Lemma 7.2] implies the existence of

$$p \in H_0^1(\Omega \setminus A_s), \qquad \nu \in H^{-1}(\Omega), \qquad \lambda \in N_{F_{ad}}(\bar{f}),$$

such that

$$\begin{aligned} p + J_f(\bar{y}, \bar{f}) + \lambda &= 0, & \langle \nu, v \rangle &= 0 \quad \forall v \in H_0^1(\Omega \setminus A), \\ J_y(\bar{y}, \bar{f}) + \Delta p - \nu &= 0, & \langle \nu, p\varphi \rangle &\geq 0 \quad \forall \varphi \in W^{1,\infty}(\Omega)^+, \end{aligned}$$

which is a necessary condition satisfied by every local minimiser as shown in [20]. Here, A_s and A denote the strictly active set and the active set associated with $(\bar{y}, \bar{f})$, respectively.

References

[1] Amal Alphonse, Michael Hintermüller, Carlos N. Rautenberg, and Gerd Wachsmuth. *Minimal and maximal solution maps of elliptic QVIs: penalisation, Lipschitz stability, differentiability and optimal control*. 2023. arXiv: 2312.13879 [math.OC].

[2] Heinz H. Bauschke and Patrick L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. Berlin: Springer, 2011. DOI: 10.1007/978-1-4419-9467-7.

[3] Yoav Benyamini and Joram Lindenstrauss. *Geometric nonlinear functional analysis. Vol. 1*. Vol. 48. American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI, 2000, pp. xii+488. DOI: 10.1090/coll/048.

[4] Constantin Christof, Christian Meyer, Stephan Walther, and Christian Clason. "Optimal control of a non-smooth semilinear elliptic equation". In: *Math. Control Relat. Fields* 8.1 (2018), pp. 247–276. DOI: 10.3934/mcrf.2018011.

[5] Gianni Dal Maso. "Γ-convergence and μ-capacities". In: *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* 14.3 (1987), pp. 423–464. URL: http://www.numdam.org/item?id=ASNSP_1987_4_14_3_423_0.

[6] Gianni Dal Maso and Umberto Mosco. "Wiener's criterion and Γ-convergence". In: *Appl. Math. Optim.* 15.1 (1987), pp. 15–63. DOI: 10.1007/BF01442645.

[7] Michael Delfour and Jean-Paul Zolésio. *Shapes and geometries*. Second. Vol. 22. Advances in Design and Control. Metrics, analysis, differential calculus, and optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011, pp. xxiv+622. DOI: 10.1137/1.9780898719826.

[8] Felix Harder and Gerd Wachsmuth. "Comparison of optimality systems for the optimal control of the obstacle problem". In: *GAMM-Mitt.* 40.4 (2018), pp. 312–338. DOI: 10.1002/gamm.201740004.

[9] Michael Hintermüller and Ian Kopacka. "A smooth penalty approach and a nonlinear multigrid algorithm for elliptic MPECs". In: *Comput. Optim. Appl.* 50.1 (2011), pp. 111–145. DOI: 10.1007/s10589-009-9307-9.

[10] Kazufumi Ito and Karl Kunisch. "Optimal control of elliptic variational inequalities". In: *Appl. Math. Optim.* 41.3 (2000), pp. 343–364. DOI: 10.1007/s002459911017.

[11] Kazufumi Ito and Karl Kunisch. "Optimal control of obstacle problems by H^1-obstacles". In: *Appl. Math. Optim.* 56.1 (2007), pp. 1–17. DOI: 10.1007/s00245-007-0877-6.

[12] Kazufumi Ito and Karl Kunisch. "Semi-smooth Newton methods for variational inequalities of the first kind". In: *M2AN Math. Model. Numer. Anal.* 37.1 (2003), pp. 41–62. DOI: 10.1051/m2an:2003021.

[13] Karl Kunisch and Daniel Wachsmuth. "Path-following for optimal control of stationary variational inequalities". In: *Comput. Optim. Appl.* 51.3 (2012), pp. 1345–1373. DOI: 10.1007/s10589-011-9400-8.

[14] Karl Kunisch and Daniel Wachsmuth. "Sufficient optimality conditions and semi-smooth Newton methods for optimal control of stationary variational inequalities". In: *ESAIM Control Optim. Calc. Var.* 18.2 (2012), pp. 520–547. DOI: 10.1051/cocv/2011105.

[15] Fulbert Mignot. "Contrôle dans les inéquations variationelles elliptiques". In: *J. Functional Analysis* 22.2 (1976), pp. 130–185. DOI: 10.1016/0022-1236(76)90017-3.

[16] Ira Neitzel and Gerd Wachsmuth. “First-order conditions for the optimal control of the obstacle problem with state constraints”. In: *Pure Appl. Funct. Anal.* 7.5 (2022), pp. 1881–1911. URL: http://yokohamapublishers.jp/online2/oppafa/vol7/p1881.html.

[17] Jean-Paul Penot. *Calculus without derivatives*. Vol. 266. Graduate Texts in Mathematics. Springer, New York, 2013, pp. xx+524. DOI: 10.1007/978-1-4614-4538-8.

[18] Anne-Therese Rauls and Gerd Wachsmuth. “Generalized derivatives for the solution operator of the obstacle problem”. In: *Set-Valued Var. Anal.* 28.2 (2020), pp. 259–285. DOI: 10.1007/s11228-019-0506-y.

[19] José-Francisco Rodrigues. *Obstacle problems in mathematical physics*. Vol. 134. North-Holland Mathematics Studies. Notas de Matemática [Mathematical Notes], 114. North-Holland Publishing Co., Amsterdam, 1987, pp. xvi+352.

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

[21] Gerd Wachsmuth. “Towards M-stationarity for optimal control of the obstacle problem with control constraints”. In: *SIAM J. Control Optim.* 54.2 (2016), pp. 964–986. DOI: 10.1137/140980582.