DFG Deutsche Forschungsgemeinschaft
Priority Programme 1962

Elliptic Quasi-Variational Inequalities under a Smallness Assumption: Uniqueness, Differential Stability and Optimal Control

Gerd Wachsmuth

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

Preprint Number SPP1962-118

received on September 9, 2019

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

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

Preprint, Faculty 1
Institute of Mathematics

Brandenburg University of Technology Cottbus - Senftenberg

Elliptic quasi-variational inequalities under a smallness assumption: Uniqueness, differential stability and optimal control

Gerd Wachsmuth

September 6, 2019
Chair of Optimal Control, BTU Cottbus-Senftenberg

We consider a quasi-variational inequality governed by a moving set. We employ the assumption that the movement of the set has a small Lipschitz constant. Under this requirement, we show that the quasi-variational inequality has a unique solution which depends Lipschitz-continuously on the source term. If the data of the problem is (directionally) differentiable, the solution map is directionally differentiable as well. We also study the optimal control of the quasi-variational inequality and provide necessary optimality conditions of strongly stationary type.

Keywords: quasi-variational inequality, uniqueness, directional differentiability, strong stationarity

MSC: 47J20, 49K21, 35J87

1 Introduction

We consider the quasi-variational inequality (QVI)

$$\text{Find } y \in Q(y) \quad \text{such that} \quad \langle A(y) - f, v - y\rangle \ge 0 \qquad \forall v \in Q(y). \tag{1.1}$$

Here, V is a Hilbert space, $A\colon V \to V^\star$ is a (possibly nonlinear) mapping, and $f \in V^\star$. We will not cover the general situation of a set-valued mapping $Q\colon V \rightrightarrows V$, but we restrict the treatment of (1.1) to the case in which $Q(y)$ is a *moving set*, i.e.,

$$Q(y) = K + \Phi(y) \tag{1.2}$$

for some non-empty, closed and convex subset $K \subset V$ and $\Phi \colon V \to V$. It is well-known that QVIs have many important real-world applications, we refer exemplarily to Bensoussan, Lions, 1987; Prigozhin, 1996; Barrett, Prigozhin, 2013; Alphonse, Hintermüller, Rautenberg, 2019 and the references therein.

The main contributions of this paper are the following.

- We prove existence and uniqueness of solutions to (1.1) under a smallness assumption on the mapping Φ, see Section 3.
- If, additionally, the functions A and Φ are differentiable and if K is polyhedric, we establish the directional differentiability of the solution mapping of (1.1), see Section 5.
- For the associated optimal control problem, we derive necessary optimality conditions of strongly stationary type, see Section 6.

In particular, our results are applicable if the Lipschitz constant of Φ is small. Let us put our work in perspective. In the following discussion, we will assume that A is μ_A-strongly monotone and L_A-Lipschitz and that Φ is L_Φ-Lipschitz. We refer to Section 2 for the definitions. We further define the condition number of A via $\gamma_A := L_A/\mu_A \ge 1$. An existence and uniqueness result for the general QVI (1.1) was given in Noor, Oettli, 1994, Theorem 9. This result can be applied to the moving set case (1.2) via Nesterov, Scrimali, 2011, Lemma 3.2. One obtains the unique solvability of (1.1) under the condition

$$L_\Phi < 1 - \sqrt{1 - 1/\gamma_A^2} = \frac{1}{\gamma_A \left(\gamma_A + \sqrt{\gamma_A^2 - 1}\right)}. \tag{1.3}$$

In the work Nesterov, Scrimali, 2011, Corollary 2 the requirement was relaxed to

$$L_\Phi < \frac{1}{\gamma_A}. \tag{1.4}$$

In this work, we shall show that

$$L_\Phi < \frac{2\sqrt{\gamma_A}}{1 + \gamma_A} \tag{1.5}$$

is sufficient for existence and uniqueness under the condition that A is the derivative of a convex function. Note that A is indeed a derivative of a convex function in many important applications. Moreover, the conditions (1.4) and (1.5) are necessary for uniqueness in the following sense: Whenever the constants $L_\Phi < 1$, $0 < \mu_A \le L_A$ violate (1.4) with $\gamma_A := L_A/\mu_A$, there exist bounded and linear operators $A \colon V \to V^\star$ and $\Phi \colon V \to V$ possessing these constants such that (1.1) does not have a unique solution for every $f \in V^\star$. If even (1.5) is violated, A can chosen to be symmetric. We refer to Theorems 3.6 and 3.7 below for the precise formulation of this result.

For a different approach to obtain uniqueness of solutions to (1.1), we refer to Dreves, 2015.

To our knowledge, Alphonse, Hintermüller, Rautenberg, 2019 is the only contribution concerning differentiability of the solution mapping of (1.1). Their approach is based on monotonicity considerations and only the differentiability into non-negative directions is obtained. In what follows, we are able to relax the assumption required for the differentiability and we also obtain differentiability in all directions, see Theorem 5.5.

Finally, we are not aware of any contribution in which stationarity conditions for the optimal control of (1.1) are obtained.

2 Notation and preliminaries

Throughout this work, V will denote a Hilbert space. Its dual space is denoted by $V^\star$. The radial cone, the tangent cone and the normal cone of a closed, convex set $K \subset V$ at $y \in K$ are given by

$$\mathcal{R}_K(y) := \operatorname{cone}(K - y) = \bigcup_{\alpha>0} \alpha\,(K - y), \qquad \mathcal{T}_K(y) := \operatorname{cl}\{\mathcal{R}_K(y)\},$$
$$\mathcal{T}_K(y)^\circ := \{\lambda \in V^\star \mid \langle \lambda, v - y\rangle \le 0 \;\forall v \in K\},$$

respectively. The critical cone of K w.r.t. $(y, \lambda) \in K \times \mathcal{T}_K(y)^\circ$ is given by

$$\mathcal{K}_K(y, \lambda) := \mathcal{T}_K(y) \cap \lambda^\perp = \{v \in \mathcal{T}_K(y) \mid \langle \lambda, v\rangle = 0\}.$$

The set K is called polyhedric at (y, λ) if $\mathcal{K}_K(y, \lambda) = \operatorname{cl}\{\mathcal{R}_K(y) \cap \lambda^\perp\}$. We refer to Wachsmuth, 2019 for a recent review of polyhedricity.

A mapping $B\colon V \to V^\star$ is called μ-strongly monotone if $\mu > 0$ satisfies

$$\langle B(y_1) - B(y_2), y_1 - y_2\rangle \ge \mu\,\|y_1 - y_2\|_V^2 \qquad \forall y_1, y_2 \in V.$$

If H is another Hilbert space, a mapping $C\colon V \to H$ is called L-Lipschitz for some $L \ge 0$ if

$$\|C(y_1) - C(y_2)\|_H \le L\,\|y_1 - y_2\|_V \qquad \forall y_1, y_2 \in V.$$

The monotonicity of an operator implies some weak lower semicontinuity.

Lemma 2.1. Let $A\colon V \to V^\star$ be a monotone operator. Suppose that $y_n \rightharpoonup y$ in V and $A(y_n) \rightharpoonup A(y)$ in $V^\star$. Then,

$$\liminf_{n\to\infty} \langle A(y_n), y_n\rangle \ge \langle A(y), y\rangle.$$

Proof. From the monotonicity of A we find

$$\langle A(y_n), y_n\rangle \geq \langle A(y_n), y\rangle + \langle A(y), y_n\rangle - \langle A(y), y\rangle.$$

The right-hand side converges towards $\langle A(y), y\rangle$ due to the weak convergences $y_n \rightharpoonup y$ and $A(y_n) \rightharpoonup A(y)$. This implies the claim.

In the case that A is additionally bounded and linear, the above claim can be obtained from the observation that $y \mapsto \langle A(y), y\rangle$ is convex. This convexity does not hold in the nonlinear setting: consider $A\colon \mathbb{R} \to \mathbb{R}$, $y \mapsto \max(-1, \min(1, y))$.

In order to obtain unique solvability of (1.1) via contraction-type arguments, one typically requires an inequality like

$$\|\operatorname{Proj}_{Q(x)}(z) - \operatorname{Proj}_{Q(y)}(z)\|_V \leq L_Q \, \|x - y\|_V \qquad \forall x, y, z \in V \tag{2.1}$$

for some $L_Q \geq 0$, see, e.g., Nesterov, Scrimali, 2011, Theorem 4.1. Note that this inequality is not related to the Lipschitz continuity of the projection since the arguments of the projections in (2.1) coincide. By means of an example, we show that (2.1) does not hold for obstacle-type problems if Q is not of the moving-set type. We consider the setting $\Omega = (0, 1)$, $V = H_0^1(\Omega)$ $f = 1$, $A = -\Delta$ and

$$K(h) := \{v \in V \mid v \leq h\}$$

for $\mathbb{R} \ni h \geq 0$. It is easy to check that the projection of $A^{-1}f$ onto the set $K(h)$ is given by

$$y_h(x) = \begin{cases} t_h \, x - \frac{1}{2} \, x^2 & \text{for } x \leq t_h \\ h & \text{for } t_h < x < 1 - t_h \\ -t_h \, x - \frac{1}{2} \, x^2 & \text{for } x \geq 1 - t_h \end{cases}$$

with $t_h := \sqrt{2\,h}$ for all $h \in [0, 1/8]$. Here, we used the norm $\|v\|_{H_0^1}^2 = \int_\Omega |\nabla v|^2 \, \mathrm{d}x$. Then, $\|y_h\|_{H_0^1(\Omega)} = C \, h^{3/4}$ for some $C > 0$. Since $y_0 = 0$, the mapping $h \mapsto y_h$ is not Lipschitz at $h = 0$. By choosing a suitable $\Psi\colon V \to \mathbb{R}$ it can be checked that $Q = K \circ \Psi$ violates (2.1).

3 Moving-set QVIs as VIs

In this section, we utilize the moving-set structure of $Q(y)$ to recast the QVI (1.1) as an equivalent variational inequality (VI). This is a classical approach, see also Alphonse, Hintermüller, Rautenberg, 2019, Section 2, Alphonse, Hintermüller, Rautenberg, 2018, Section 5.1. We start by defining the new solution variable $z := y - \Phi(y) \in K$. In order to not lose any information, we require that the function $I - \Phi\colon V \to V$ is a bijection. Hence, $y = (I - \Phi)^{-1}(z)$. Now, it is easy to check that (1.1) is equivalent to

$$\text{Find } z \in K \quad \text{such that} \quad \langle (A \circ (I - \Phi)^{-1})(z) - f, v - z\rangle \geq 0 \qquad \forall v \in K \tag{3.1}$$

and $y = (I - \Phi)^{-1}(z)$. This means the following: Under the assumption that $I - \Phi$ is a bijection, y is a solution of (1.1) if and only if $z = (I - \Phi)(y)$ is a solution of (3.1). In what follows, we are going to use the VI (3.1) in order to obtain information about the QVI (1.1). In the case in which both A and Φ are linear, such a strategy was suggested in Alphonse, Hintermüller, Rautenberg, 2019, Remark 7. We shall see that this is also viable in the fully nonlinear case.

In order to analyze (3.1) we will frequently make use of the following assumption.

Assumption 3.1. The operator $I - \Phi \colon V \to V$ is a bijection and the operator $B := A \circ (I - \Phi)^{-1}$ is strongly monotone and Lipschitz continuous.

Using the equivalence of (1.1) and (3.1) as well as the existence result Kinderlehrer, Stampacchia, 1980, Corollary III.1.8 for (3.1), we obtain the following existence and uniqueness result for (1.1).

Theorem 3.2. Under Assumption 3.1, the QVI (1.1) has a unique solution $y \in V$ for any $f \in V^\star$. Moreover, the mapping $f \mapsto y$ is Lipschitz continuous.

In the remainder of this section, we give some conditions implying Assumption 3.1.

Lemma 3.3. We assume that A is μ_A-strongly monotone and L_A-Lipschitz and that Φ is L_Φ-Lipschitz. We further assume that

$$L_\Phi < \frac{1}{\gamma_A}, \tag{3.2}$$

where $\gamma_A = L_A/\mu_A$. Then, the operator $B := A \circ (I - \Phi)^{-1}$ is μ_B-strongly monotone and L_B-Lipschitz with

$$\mu_B = \frac{\mu_a - L_A L_\Phi}{(1 + L_\Phi)^2} \qquad \text{and} \qquad L_B = \frac{L_A}{1 - L_\Phi}.$$

In particular, Assumption 3.1 is satisfied and (1.1) has a unique solution for every $f \in V^\star$.

Proof. First we remark that we have $L_\Phi < \gamma_A^{-1} \le 1$. Thus, Banach's fixed point theorem implies that $I - \Phi$ is a bijection. We claim that $(1 - L_\Phi)^{-1}$ is a Lipschitz constant of $(I - \Phi)^{-1}$. Indeed, let $y_1, y_2 \in V$ be arbitrary. We define $x_i := (I - \Phi)^{-1}(y_i)$ for $i = 1, 2$. Then, $x_i - y_i = \Phi(x_i)$, $i = 1, 2$ and this yields

$$\|x_1 - x_2\|_V - \|y_1 - y_2\|_V \le \|(x_1 - y_1) - (x_2 - y_2)\|_V \le L_\Phi \, \|x_1 - x_2\|_V.$$

This shows the claim concerning a Lipschitz constant of $(I - \Phi)^{-1}$. Moreover, this directly shows that L_B is a Lipschitz constant of B.

For arbitrary $y_1, y_2 \in V$ we again use $x_i := (I - \Phi)^{-1}(y_i)$ for $i = 1, 2$. Then,

$$\begin{aligned}\langle B(y_1) - B(y_2), y_1 - y_2\rangle &= \langle A(x_1) - A(x_2), (I - \Phi)(x_1) - (I - \Phi)(x_2)\rangle \\ &\ge (\mu_A - L_A\, L_\Phi)\, \|x_1 - x_2\|_V^2.\end{aligned}$$

The estimate

$$\|y_1 - y_2\|_V = \|x_1 - x_2 - (\Phi(x_1) - \Phi(x_2))\|_V \le (1 + L_\Phi)\, \|x_1 - x_2\|_V$$

yields the assertion concerning the strong monotonicity of B. The final claim follows from Theorem 3.2.

We recall that the condition (3.2) was used in Nesterov, Scrimali, 2011, Corollary 2 to obtain existence and uniqueness for solutions of (3.1). The above analysis shows, that this condition even implies Assumption 3.1.

Next, we show that the estimate (3.2) can be significantly relaxed if A is the derivative of a convex function. To this end, we need to recall an important inequality for convex functions. This inequality is well-known in the finite-dimensional case, see, e.g., Nesterov, 2004, Theorem 2.1.12 or Bubeck, 2015, Lemma 3.10, and the proof carries over to arbitrary Hilbert spaces. We are, however, not aware of a reference in the infinite-dimensional case.

Lemma 3.4. Let $g\colon V \to \mathbb{R}$ be a Fréchet differentiable convex function such that the derivative $g'\colon V \to V^\star$ is μ_g-strongly monotone and L_g-Lipschitz. Then,

$$\langle g'(x_1) - g'(x_2), x_1 - x_2\rangle \ge \frac{\mu_g\, L_g}{\mu_g + L_g}\, \|x_1 - x_2\|_V^2 + \frac{1}{\mu_g + L_g}\, \|g'(x_1) - g'(x_2)\|_{V^\star}^2$$

holds for all $x_1, x_2 \in V$.

Proof. One can transfer the proofs of Nesterov, 2004, Theorem 2.1.12 or Bubeck, 2015, Lemma 3.10 to the infinite-dimensional case by using Bauschke, Combettes, 2011, Theorem 18.15.

Lemma 3.5. We assume that A is μ_A-strongly monotone and L_A-Lipschitz and that Φ is L_Φ-Lipschitz. We further assume that there exists a Fréchet differentiable convex function $g\colon V \to \mathbb{R}$ such that $A = g'$ and

$$L_\Phi < \frac{2\sqrt{\gamma_A}}{1 + \gamma_A} = \frac{2\sqrt{\mu_A\, L_A}}{\mu_A + L_A} \tag{3.3}$$

where $\gamma_A = L_A/\mu_A$. Then, the operator $B := A\,(I - \Phi)^{-1}$ is μ_B-strongly monotone and

L_B-Lipschitz with

$$\mu_B = \frac{4\,\mu_a\,L_A - L_\Phi^2\,(\mu_A + L_A)^2}{4\,(\mu_A + L_A)\,(1 + L_\Phi)^2} \qquad \text{and} \qquad L_B = \frac{L_A}{1 - L_\Phi}.$$

In particular, Assumption 3.1 is satisfied and (1.1) has a unique solution for every $f \in V^\star$.

Proof. By arguing as in the proof of Lemma 3.3, we obtain that $I - \Phi$ is invertible and the value of the Lipschitz constant L_B follows.

Now, let $y_1, y_2 \in V$ be arbitrary and we set $x_i := (I - \Phi)^{-1}(y_i)$, $i = 1, 2$. Then, we apply Lemma 3.4 to obtain

$$\begin{aligned}\langle B(y_1) - B(y_2), y_1 - y_2\rangle &= \langle A(x_1) - A(x_2), (I - \Phi)(x_1) - (I - \Phi)(x_2)\rangle \\ &\ge \frac{\mu_A\,L_A}{\mu_A + L_A}\,\|x_1 - x_2\|_V^2 + \frac{1}{\mu_A + L_A}\,\|A(x_1) - A(x_2)\|_{V^\star}^2 \\ &\quad - L_\Phi\,\|x_1 - x_2\|_V\,\|A(x_1) - A(x_2)\|_{V^\star}.\end{aligned}$$

Next, we employ Young's inequality

$$\begin{aligned}L_\Phi\,\|x_1 - x_2\|_V\,\|A(x_1) - A(x_2)\|_{V^\star} \le{}& \frac{L_\Phi^2\,(\mu_A + L_A)}{4}\,\|x_1 - x_2\|_V^2 \\ &+ \frac{1}{\mu_A + L_A}\,\|A(x_1) - A(x_2)\|_{V^\star}^2\end{aligned}$$

and get

$$\begin{aligned}\langle B(y_1) - B(y_2), y_1 - y_2\rangle &\ge \Big(\frac{\mu_A\,L_A}{\mu_A + L_A} - \frac{L_\Phi^2\,(\mu_A + L_A)}{4}\Big)\,\|x_1 - x_2\|_V^2 \\ &= \frac{4\,\mu_A\,L_A - L_\Phi^2\,(\mu_A + L_A)^2}{4\,(\mu_A + L_A)}\,\|x_1 - x_2\|_V^2.\end{aligned}$$

From the proof of Lemma 3.3 we find $\|y_1 - y_2\|_V \le (1 + L_\Phi)\,\|x_1 - x_2\|_V$ and this yields the monotonicity. The final claim follows from Theorem 3.2.

Note that the inequality (3.3) is weaker than (3.2), unless $\gamma_A = 1$. Lemma 3.5 is an improvement of the corresponding results in the literature, e.g., Nesterov, Scrimali, 2011, Corollary 2, in the case that A is the derivative of a convex function. It is well known that A is a derivative of a convex function if and only if A is *maximally cyclically monotone*, see, e.g., Bauschke, Combettes, 2011, Theorem 22.14.

Finally, we demonstrate by the mean of two examples that the assumptions (3.2) and (3.3) are sharp, even in the case of linear operators. These examples are found by constructing operators for which the estimates in the proofs of Lemmas 3.3 and 3.5 are sharp. First, we validate the sharpness of (3.3).

Theorem 3.6. Let the constants $0 < \mu_A < L_A$ be given. We define $L_\Phi := \mu_A / L_A < 1$, i.e., (3.2) is violated. Then, there exist linear operators A and Φ on $V = \mathbb{R}^2$ (equipped with the Euclidean inner product), such that A is μ_A-strongly monotone and L_A-Lipschitz, Φ is L_Φ-Lipschitz and $A\,(I-\Phi)^{-1}$ is not coercive. Moreover, there exists a one-dimensional subspace $K \subset \mathbb{R}^2$ such that (3.1) and (1.1) are not uniquely solvable for all $f \in V^\star$.

Proof. We define the constant $c_A := \sqrt{L_A^2 - \mu_A^2} > 0$ and the operators

$$A := \begin{pmatrix} \mu_A & -c_A \\ c_A & \mu_A \end{pmatrix}, \qquad \Phi := \frac{L_\Phi}{L_A}\, y\, x^\top$$

where

$$x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad y = \begin{pmatrix} \mu_A \\ c_A \end{pmatrix} = A\,x.$$

Since A is the combination of a rotation and a scaling by L_A, it is easy to check that $z^\top A\, z = \mu_A \|z\|^2$ and $\|A\, z\| = L_A \|z\|$ hold for all $z \in \mathbb{R}^2$. Moreover, the Lipschitz constant of Φ is L_Φ. However,

$$z^\top A\,(I-\Phi)^{-1}\, z = 0, \qquad \text{where} \qquad z = (I-\Phi)\, x \neq 0.$$

Hence, $A\,(I-\Phi)^{-1}$ is not coercive. Moreover, if we set $K = \operatorname{span}\{z\}$ it is clear that (3.1) is not uniquely solvable for all $f \in V^\star = \mathbb{R}^2$. Since $I - \Phi$ is a bijection, this implies that (1.1) is not uniquely solvable for all $f \in V^\star = \mathbb{R}^2$.

The next result shows that (3.3) is sharp.

Theorem 3.7. Let $0 < \mu_A < L_A$ be given. We define $L_\Phi := 2\sqrt{\mu_A\, L_A}/(\mu_A + L_A) < 1$, i.e., (3.3) is violated. Then, there exists a linear symmetric operator A on $V = \mathbb{R}^2$ (equipped with the Euclidean inner product) and a linear operator Φ in $\mathbb{R}^2$, such that A is μ_A-strongly monotone and L_A-Lipschitz, Φ is L_Φ-Lipschitz and $A\,(I-\Phi)^{-1}$ is not coercive. Moreover, there exists a one-dimensional subspace $K \subset \mathbb{R}^2$ such that (3.1) and (1.1) cannot be uniquely solvable for all $f \in V^\star$.

Proof. We define

$$A := \begin{pmatrix} \mu_A & 0 \\ 0 & L_A \end{pmatrix}.$$

It is clear that the operator A is μ_A-strongly monotone and L_A-Lipschitz. We further set

$$x := \begin{pmatrix} \sqrt{L_A/(\mu_A + L_A)} \\ \sqrt{\mu_A/(\mu_A + L_A)} \end{pmatrix}, \qquad \Phi := \frac{2}{(\mu_A + L_A)^2} \begin{pmatrix} \mu_A^2\, L_A & \mu_A \sqrt{\mu_A\, L_A} \\ L_A \sqrt{\mu_A\, L_A} & \mu_A^2\, L_A \end{pmatrix}.$$

It can be checked that Φ is L_Φ-Lipschitz and $\|x\| = 1$. However,

$$x^\top A\,(I - \Phi)\,x = 0 \qquad \text{and} \qquad y^\top A\,(I - \Phi)^{-1}\,y = 0$$

where $y = (I - \Phi)\,x \neq 0$. Hence, $A\,(I - \Phi)$ and $A\,(I - \Phi)^{-1}$ are not coercive. Moreover, if we set $K = \operatorname{span}\{y\}$ it is clear that (3.1) cannot be uniquely solvable for all $f \in V^\star = \mathbb{R}^2$. Since $I - \Phi$ is a bijection, this implies that (1.1) is not uniquely solvable for all $f \in V^\star = \mathbb{R}^2$.

We further mention that it is also possible to obtain Assumption 3.1 in situations in which Φ is "not small", e.g., if $\Phi = \lambda\, I$ with some $\lambda < 1$, Assumption 3.1 follows automatically if A is strongly monotone and Lipschitz since $(I - \Phi)^{-1} = (1 - \lambda)^{-1}\, I$ in this case. Moreover, it is possible to analyze the situation in which A is a small perturbation of the derivative of a convex function by combining the ideas of Lemmas 3.3 and 3.5.

The combination of Theorem 3.2 and Lemma 3.3 yields a well-known result: under the assumption (3.2), the QVI (1.1) has a unique solution. Such a result is typically shown via contraction-type arguments, see, e.g., Nesterov, Scrimali, 2011 or Alphonse, Hintermüller, Rautenberg, 2018, Section 3.1.1. Thus, the approach of this section is able to reproduce this classical result. However, the combination of Theorem 3.2 and Lemma 3.5 yields a new result in case that A has a convex potential in which the condition (3.2) on the Lipschitz constant L_Φ of Φ is relaxed to (3.3).

4 Localization of the smallness assumption

We localize the assumptions concerning the Lipschitz constant of Φ.

Assumption 4.1. We assume that $A \colon V \to V^\star$ is (globally) μ_A-strongly monotone and L_A-Lipschitz. Further, let $\bar f \in V^\star$ be given and let $\bar y$ be a solution of (1.1). We suppose that there is a closed, convex neighborhood $Y \subset V$ of $\bar y$ such that Φ is L_Φ-Lipschitz continuous on Y. Finally,

(i) inequality (3.2) holds or

(ii) inequality (3.3) holds and A is the Fréchet derivative of a convex function.

Theorem 4.2. Suppose that Assumption 4.1 is satisfied. There is a neighborhood $F \subset V^\star$ of $\bar f$ such that (1.1) has exactly one solution in Y for all $f \in F$. Moreover, this solution depends Lipschitz-continuously on f.

Note that we do not claim that (1.1) is uniquely solvable for all $f \in F$ and (1.1) might have further solutions in $V \setminus Y$.

Proof. We define $\tilde{\Phi}\colon V \to V$ via

$$\tilde{\Phi}(y) := \Phi(\operatorname{Proj}_Y(y)).$$

Since projections are 1-Lipschitz, $\tilde{\Phi}$ is L_Φ-Lipschitz. Now, we consider the modified QVI

$$\text{Find } y \in \tilde{Q}(y) \quad \text{such that} \quad \langle A(y) - f, v - y \rangle \ge 0 \qquad \forall v \in \tilde{Q}(y) \tag{4.1}$$

with

$$\tilde{Q}(y) = K + \tilde{\Phi}(y).$$

From Assumption 4.1, Lemmas 3.3 and 3.5, and Theorem 3.2 it follows that (4.1) has a unique solution $y = \tilde{S}(f)$ for every $f \in F$ and the solution operator $S\colon V^\star \to V$ is Lipschitz continuous. Hence, we can choose a neighborhood $F \subset V^\star$ of $\bar{f}$, such that $\tilde{S}(f) \in Y$ for all $f \in F$.

Since $Q(y) = \tilde{Q}(y)$ for all $y \in Y$, it is clear that $y \in Y$ is a solution of (1.1) if and only if $y \in Y$ solves (4.1). Hence, (1.1) has a unique solution in Y for all $f \in F$.

5 Differential stability

In this section, we consider the situation of Assumption 4.1. However, we do not need Assumption 4.1 directly, but the assertion of Theorem 4.2 is enough.

Assumption 5.1. We suppose that the following assumptions are satisfied.

(i) We assume the existence of sets $F \subset V^\star$, $Y \subset V$ such that for every $f \in F$, (1.1) has a unique solution y in Y and the solution map $S\colon F \to Y$, $f \mapsto y$ is Lipschitz continuous. For fixed $\bar{f} \in F$, we set $\bar{y} := S(\bar{f})$. The sets F, Y are assumed to be neighborhoods of $\bar{f}$, $\bar{y}$, respectively.

(ii) The operator $\Phi\colon V \to V$ is Lipschitz on Y, i.e., there exists $L_\Phi > 0$ with

$$\|\Phi(y_1) - \Phi(y_2)\|_V \le L_\Phi \|y_1 - y_2\|_V \qquad \forall y_1, y_2 \in Y.$$

We suppose that $I - \Phi\colon Y \to Z$ is bijective with a Lipschitz continuous inverse, where $Z := (I - \Phi)(Y)$. Further, Φ is Fréchet differentiable at $\bar{y}$ and the bounded linear operator $I - \Phi'(\bar{y})$ is bijective.

(iii) The operator A is Fréchet differentiable at $\bar{y}$ and the bounded linear operator

$$A'(\bar{y})\,(I - \Phi'(\bar{y}))^{-1} \tag{5.1}$$

is assumed to be coercive.

(iv) The set K is polyhedric at $(\bar{z}, \bar{f} - A(\bar{y}))$, where $\bar{z} = (I - \Phi)(\bar{y})$.

Due to (1.2), the last assumption is equivalent to the polyhedricity of $Q(y)$ at $(\bar y, \bar f - A(\bar y))$.

First, we show that Assumption 5.1 follows from Assumption 4.1 and from the differentiability of Φ and A.

Theorem 5.2. Suppose that Assumption 4.1 is satisfied. Then, Assumption 5.1 (i) holds. If Φ is Fréchet differentiable at $\bar y$, then Assumption 5.1 (ii) holds. If, additionally, A is Fréchet differentiable at $\bar y$, then Assumption 5.1 (iii) is satisfied.

Proof. Assumption 5.1 (i) follows from Theorem 4.2.

Since $L_\Phi < 1$, Banach's fixed point theorem implies that $I - \Phi$ is bijective with a Lipschitz continuous inverse. The invertibility of $I - \Phi'(\bar y)$ follows from the Neumann series since $\|\Phi'(\bar y)\| \le L_\Phi < 1$.

If A is Fréchet differentiable at $\bar y$, Assumption 4.1 implies that $A'(\bar y)$ is μ_A-strongly monotone and L_A-Lipschitz. In case that (3.2) is satisfied, we can invoke Lemma 3.3 to obtain Assumption 5.1 (iii). Otherwise, A is the Fréchet derivative of a convex function. Hence, $A'(\bar y)$ is symmetric since it is a second Fréchet derivative, see Cartan, 1967, Theorem 5.1.1. Thus, $A'(\bar y)$ is the derivative of the convex function $v \mapsto \langle A'(\bar y)\, v, v\rangle/2$. Therefore, we can invoke Lemma 3.5 to obtain Assumption 5.1 (iii).

Lemma 5.3. Let us assume that Assumption 5.1 (i)–(ii) is satisfied. Then, $(I - \Phi)^{-1}$ is Fréchet differentiable at $\bar z := (I - \Phi)(\bar y)$ and $((I - \Phi)^{-1})'(\bar z) = (I - \Phi'(\bar y))^{-1}$.

Proof. For arbitrary $h \in V$ we have

$$h = (I - \Phi)\big[(I - \Phi)^{-1}(\bar z + h) - \bar y + \bar y\big] - \bar z.$$

Using the Fréchet differentiability of Φ at $\bar y$ implies

$$h = (I - \Phi'(\bar y))\big[(I - \Phi)^{-1}(\bar z + h) - \bar y\big] + o(\|(I - \Phi)^{-1}(\bar z + h) - \bar y\|_V)$$

as $\|(I - \Phi)^{-1}(\bar z + h) - \bar y\|_V \to 0$. Finally, using the fact that $(I - \Phi)^{-1}$ is Lipschitz implies

$$(I - \Phi'(\bar y))^{-1} h = (I - \Phi)^{-1}(\bar z + h) - (I - \Phi)^{-1}(\bar z) + o(\|h\|_V) \quad \text{as} \quad \|h\|_V \to 0.$$

This shows the claim.

Lemma 5.4. Let us assume that Assumption 5.1 (i)–(iii) is satisfied. The operator $B := A \circ (I - \Phi)^{-1}$ is Fréchet differentiable at $\bar z$ and its Fréchet derivative is given by $B'(\bar z) = A'(\bar y)\,(I - \Phi'(\bar y))^{-1}$.

Proof. Follows from Lemma 5.3 together with a chain rule.

Theorem 5.5. Let us assume that Assumption 5.1 is satisfied. Then, the solution map S is directionally differentiable at $\bar f$ and the directional derivative $x := S'(\bar f; h)$ in direction $h \in V^\star$ is given by the unique solution of the QVI

$$\text{Find } x \in Q^{\bar y}(x) \quad \text{such that} \quad \langle A'(\bar y)\, x - h, v - x\rangle \ge 0 \qquad \forall v \in Q^{\bar y}(x), \tag{5.2}$$

where the set-valued mapping $Q^{\bar y} \colon V \rightrightarrows V$ is given by

$$Q^{\bar y}(x) := \mathcal{K}_K(\bar z, \bar f - A(\bar y)) + \Phi'(\bar y)\, x.$$

Note that we have

$$\mathcal{K}_K(\bar z, \bar f - A(\bar y)) = \mathcal{K}_{Q(\bar y)}(\bar y, \bar f - A(\bar y))$$

due to (1.2).

Proof. Let $h \in V^\star$ be given. There exists $T > 0$ such that $\bar f + t\, h \in F$ for all $t \in [0, T)$. For $t \in (0, T)$ we define

$$y_t := S(\bar f + t\, h), \qquad x_t := \frac{y_t - \bar y}{t}$$
$$z_t := (I - \Phi)(y_t) \qquad w_t := \frac{z_t - \bar z}{t}.$$

Since S is assumed to be Lipschitz continuous on F, the difference quotients $\{x_t \mid t \in (0, T)\}$ are bounded in V. The Lipschitz continuity of Φ implies the boundedness of $\{w_t \mid t \in (0, T)\}$ in V.

Since z_t solves the VI (3.1), i.e.,

$$\text{Find } z \in K \quad \text{such that} \quad \langle B(z) - f, v - z\rangle \ge 0 \qquad \forall v \in K$$

with $f := \bar f + t\, h$, we can apply Christof, Wachsmuth, 2019, Theorem 2.13 to obtain the convergence of the difference quotients w_t. Let us check that the assumptions of Christof, Wachsmuth, 2019, Theorem 2.13 are satisfied. The standing assumption Christof, Wachsmuth, 2019, Assumption 2.1 is satisfied in our Hilbert space setting with $j = \delta_K$ being the indicator function (in the sense of convex analysis) of the set K. The validity of Christof, Wachsmuth, 2019, Assumption 2.2 follows from the Taylor expansion

$$B(z_t) = B(\bar z + t\, w_t) = B(\bar z) + t\, B'(\bar z)\, w_t + r(t)$$

with $r(t) = o(\|z_t - \bar z\|_V) = o(t)$, see Lemma 5.4. It remains to check that the assumption Christof, Wachsmuth, 2019, Theorem 2.13 (ii) holds:

- Since K is assumed to be polyhedric at $\bar z$ w.r.t. $\bar f - A(\bar y)$, its indicator function δ_K is twice epi-differentiable at $(\bar z, \bar f - A(\bar y))$, see Christof, Wachsmuth, 2019, Corollary 3.3. Moreover, its second subderivative is the indicator function of the critical cone $\mathcal{K}_K(\bar z, \bar f - A(\bar y)) := \mathcal{T}_K(\bar z) \cap (\bar f - A(\bar y))^\perp$.

- The weak convergence $w_n \rightharpoonup w$ in V implies $B'(\bar z)\, w_n \rightharpoonup B'(\bar z)\, w$ in $V^\star$ and $\liminf_{n\to\infty} \langle B'(\bar z) w_n, w_n\rangle \ge \langle B'(\bar z)\, w, w\rangle$ follows from the coercivity of the linear operator $B'(\bar z) = A'(\bar y)\,(I - \Phi'(\bar y))^{-1}$, see Lemma 2.1, Assumption 5.1 (iii) and Lemma 5.4.

Thus, the application of Christof, Wachsmuth, 2019, Theorem 2.13 yields that all accumulation points w of w_t for $t \searrow 0$ are solutions of the linearized VI

$$\text{Find } w \in \mathcal{K}_K(\bar z, \bar f - A(\bar y)) \quad \text{such that} \quad \langle B'(\bar z)\, w - h, v - w\rangle \ge 0 \qquad \forall v \in \mathcal{K}_K(\bar z, \bar f - A(\bar y)). \tag{5.3}$$

Since $B'(\bar z)$ is coercive, this linearized VI has a unique solution. Hence, the last part of Christof, Wachsmuth, 2019, Theorem 2.13 implies $w_t \to w$ as $t \searrow 0$.

It remains to prove the convergence of x_t towards the solution of (5.2). Using the differentiability of $(I - \Phi)^{-1}$, we find

$$\begin{aligned} x_t = \frac{y_t - \bar y}{t} &= \frac{(I-\Phi)^{-1}(z_t) - (I-\Phi)^{-1}(\bar z)}{t} \\ &= (I - \Phi'(\bar y))^{-1} \frac{z_t - \bar z}{t} + \frac{o(\|z_t - \bar z\|_V)}{t} \\ &\to (I - \Phi'(\bar y))^{-1}\, w =: x. \end{aligned}$$

The change of variables $w = (I - \Phi'(\bar y))\, x$ shows the equivalence of (5.2) and (5.3). Thus, x is the unique solution of (5.2).

Some remarks concerning Theorem 5.5 are in order.

Remark 5.6.

(i) The polyhedricity assumption Assumption 5.1 (iv) can be replaced by the strong twice epi-differentiability of the indicator function δ_K in the sense of Christof, Wachsmuth, 2019, Definition 2.9. Under this generalized assumption, the second epi-derivative of δ_K appears as a curvature term in the linearized inequalities (5.2) and (5.3). Note that the indicator function of the critical cone $\mathcal{K}_K(\bar z, \bar f - A(\bar y))$, which appears implicitly in (5.2) and (5.3), is just the second epi-derivative of δ_K in the case of K being polyhedric.

(ii) We have derived the differentiability result under the assumption that Φ is Fréchet differentiable at $\bar y$. In the notation of Christof, Wachsmuth, 2019, this translates to linearity of the operator A_x. However, the inspection of the proof of Christof, Wachsmuth, 2019, Theorem 2.13 entails that it is possible to replace the Fréchet differentiability of Φ by the following set of assumptions:

(a) Φ is Bouligand differentiable at $\bar y$, i.e., there exists an operator $\Phi'(\bar y; \cdot)\colon V \to V$

such that

$$\|\Phi(\bar y + h) - \Phi(\bar y) - \Phi'(\bar y; h)\|_V = o(\|h\|_V) \qquad \text{as } \|h\|_V \to 0.$$

(b) For every sequence $w_n \rightharpoonup w$ in V, we assume

$$(I - \Phi'(\bar y; \cdot))^{-1}(w_n) \rightharpoonup (I - \Phi'(\bar y; \cdot))^{-1}(w), \tag{5.4a}$$

$$\liminf_{n\to\infty} \langle A'(\bar y)\,(I - \Phi'(\bar y; \cdot))^{-1}(w_n), w_n\rangle \ge \langle A'(\bar y)\,(I - \Phi'(\bar y; \cdot))^{-1}(w), w\rangle. \tag{5.4b}$$

Note that (a) implies that $\Phi'(\bar y; \cdot)$ is Lipschitz on V with constant L_Φ. Hence, $(I - \Phi'(\bar y; \cdot))$ is invertible and Lemmas 3.3 and 3.5 can be used to obtain the strong monotonicity of $A'(\bar y)\,(I - \Phi'(\bar y; \cdot))^{-1}$.

Property (5.4a) can be verified by assuming, e.g., weak continuity of $\Phi'(\bar y; \cdot)$. Indeed, the sequence $z_n := (I - \Phi'(\bar y; \cdot))^{-1}(w_n)$ is bounded, hence, $z_n \rightharpoonup z$ along a subsequence. Now, weak continuity implies $w_n = z_n - \Phi'(\bar y; z_n) \rightharpoonup z - \Phi'(\bar y; z)$ and $w_n \rightharpoonup w$ implies $z = (I - \Phi'(\bar y; \cdot))^{-1}(w)$, i.e. $z_n \rightharpoonup (I - \Phi'(\bar y; \cdot))^{-1}(w)$ along a subsequence. The uniqueness of the limit point implies the convergence of the entire sequence.

Finally, (5.4b) can be obtained via (5.4a) and Lemma 2.1.

In the next remark, we compare our differentiability result with Alphonse, Hintermüller, Rautenberg, 2019, Theorem 1.

Remark 5.7. In Alphonse, Hintermüller, Rautenberg, 2019, Theorem 1 a similar differentiability result is obtained in a more restrictive setting:

(i) Therein, the leading operator A has to be linear and T-monotone (w.r.t. a vector space order on V). Our approach also allows for non-linear operators and we do not need any order structure on V. Similarly, we do not need any monotonicity assumptions on Φ.

(ii) They require the complete continuity of $\Phi'(\bar y)$, which is not needed in Theorem 5.5.

(iii) One of their most restrictive assumptions is the assumption (**A5**). Via Cartan, 1967, Theorem 3.1.2, this assumption is equivalent to Φ being L_Φ-Lipschitz with

$$L_\Phi < \frac{1}{1 + \gamma_A}. \tag{5.5}$$

This inequality is much stronger than (3.2). Thus, their assumption (**A5**) implies that the solutions to the QVI (1.1) are unique.

Moreover, they obtained the differentiability only for non-negative directions whereas our approach is applicable to arbitrary perturbations of the right-hand side.

One assumption in Alphonse, Hintermüller, Rautenberg, 2019 is weaker: they only need Hadamard differentiability of Φ. We need Fréchet differentiability (or Bouligand differentiability, see Remark 5.6).

6 Optimal control

In this section, we consider the optimal control problem

$$\begin{aligned} \text{Minimize} \quad & J(y, u) \\ \text{w.r.t.} \quad & y \in V, u \in U \\ \text{s.t.} \quad & y \in Q(y) \quad \text{and} \quad \langle A(y) - (B\,u + f), v - y\rangle \ge 0 \quad \forall v \in Q(y). \end{aligned} \tag{6.1}$$

Here, $f \in V^\star$ is fixed, U is a Hilbert space and the bounded, linear operator $B\colon U \to V^\star$ is assumed to have a dense range. Moreover, the objective $J\colon V \times U \to \mathbb{R}$ is Fréchet differentiable.

Theorem 6.1. Suppose that $(\bar y, \bar u)$ is locally optimal for (6.1). In addition to the assumptions on B and J, we assume that Assumption 5.1 is satisfied at $\bar f := B\,\bar u + f$. Then, there exist unique multipliers $p \in V$, $\mu \in V^\star$ such that the system

$$J_y(\cdot) + A'(\bar y)^\star p + (I - \Phi'(\bar y))^\star \mu = 0, \tag{6.2a}$$

$$J_u(\cdot) - B^\star p = 0, \tag{6.2b}$$

$$p \in -\mathcal{K}_K(\bar z, \bar\lambda), \tag{6.2c}$$

$$\mu \in \mathcal{K}_K(\bar z, \bar\lambda)^\circ \tag{6.2d}$$

is satisfied. Here,

$$\bar z = (I - \Phi)(\bar y) \in K \qquad \text{and} \qquad \bar\lambda = B\,\bar u + f - A(\bar y) \in \mathcal{T}_K(\bar z)^\circ, \tag{6.3}$$

and $J_y(\cdot) \in V^\star$ and $J_u(\cdot) \in U^\star$ are the partial Fréchet derivatives of J at $(\bar u, \bar y)$.

Proof. We use classical arguments dating back to Mignot, 1976, Proposition 4.1, see also Wachsmuth, 2019, Theorem 5.3.

Due to Assumption 5.1 we can invoke Theorem 5.5 to obtain the directional differentiability of the control-to-state map. Combined with the local optimality of $(\bar y, \bar u)$, this implies

$$\langle J_y(\cdot), S'(B\,\bar u + f; B\,h)\rangle_{V^\star,V} + \langle J_u(\cdot), h\rangle_{U^\star,U} \ge 0 \qquad \forall h \in U. \tag{6.4}$$

Due to the Lipschitz estimate $\|S'(B\,\bar u + f; B\,h)\|_V \le C\,\|B\,h\|_{V^\star}$, the above inequality implies

$$|\langle J_u(\cdot), h\rangle_{U^\star,U}| \le C\,\|B\,h\|_{V^\star} \qquad \forall h \in U.$$

Hence, there is $p \in V^{\star\star} \cong V$ (by defining it as in the next line on the dense subspace image$(B) \subset V^\star$ and extending it by continuity on the whole space $V^\star$) such that

$$\langle J_u(\cdot), h\rangle_{U^\star,U} = \langle p, B\,h\rangle_{V,V^\star} \qquad \forall h \in U.$$

This yields (6.2b) and

$$\langle J_y(\cdot), S'(B\,\bar u + f; B\,h)\rangle_{V^\star,V} + \langle p, B\,h\rangle_{V,V^\star} \ge 0 \qquad \forall h \in U.$$

Using the density of image(B) in $V^\star$ we get

$$\langle J_y(\cdot), S'(B\,\bar u + f; h)\rangle_{V^\star,V} + \langle p, h\rangle_{V,V^\star} \ge 0 \qquad \forall h \in V^\star. \tag{$*$}$$

In what follows, we set $\mathcal{K} := \mathcal{K}_K(\bar z, \bar\lambda)$ for convenience. We recall that $S'(B\,\bar u + f; h)$ is the unique solution of

$$\text{Find } x \in Q^{\bar y}(x) \quad \text{such that} \quad \langle A'(\bar y)\,x - h, v - x\rangle \ge 0 \qquad \forall v \in Q^{\bar y}(x), \tag{$**$}$$

where the set-valued mapping $Q^{\bar y} \colon V \rightrightarrows V$ is given by

$$Q^{\bar y}(x) = \mathcal{K} + \Phi'(\bar y)\,x.$$

We choose $h \in \mathcal{K}^\circ$ in ($*$). Then, ($**$) shows $S'(\bar u + f; h) = 0$ and, thus,

$$\langle p, h\rangle_{V,V^\star} \ge 0 \qquad \forall h \in \mathcal{K}^\circ,$$

i.e., $p \in -\mathcal{K}$, which shows (6.2c).

Now, we choose $v \in (I - \Phi'(\bar y))^{-1}\mathcal{K}$ and set $h = A'(\bar y)\,v$. It can be checked that ($**$) implies $S'(B\,\bar u + f; h) = v$. With this choice, ($*$) implies

$$\langle J_y(\cdot) + A'(\bar y)^\star\,p, v\rangle_{V^\star,V} \ge 0 \qquad \forall v \in (I - \Phi'(\bar y))^{-1}\mathcal{K}.$$

We define $\mu := -(I - \Phi'(\bar y))^{-\star}(J_y(\cdot) + A^\star\,p)$ and get (6.2a) and

$$\langle (I - \Phi'(\bar y))^\star \mu, v\rangle_{V^\star,V} \le 0 \qquad \forall v \in (I - \Phi'(\bar y))^{-1}\mathcal{K}.$$

Since $I - \Phi'(\bar y)$ is a bijection, this is equivalent to (6.2d).

The uniqueness of p and μ follows from the injectivity of $B^\star$ and the bijectivity of $(I - \Phi'(\bar y))^\star$.

The approach of Christof, 2018, Section 6.1 can be used to provide strong stationarity systems under less restrictive assumptions on K, i.e., the polyhedricity assumption can be replaced by the twice epi-differentiability of the indicator function δ_K.

In the case that Φ is merely Bouligand differentiable, cf. Remark 5.6, conditions (6.2a) and (6.2d) could be rewritten as

$$J_y(\cdot) + A'(\bar y)^\star p + \hat\mu = 0, \qquad \hat\mu \in \big[(I - \Phi'(\bar y;\cdot))^{-1}\mathcal{K}_K(\bar z, \bar\lambda)\big]^\circ, \tag{6.5}$$

see the proof of Theorem 6.1.

Finally, we show that the system of strong stationarity is of reasonable strength, i.e., it implies the B-stationarity (6.4).

Lemma 6.2. Let $(\bar y, \bar u)$ be a feasible point of (6.1) such that Assumption 5.1 is satisfied at $\bar f := B\,\bar u + f$. Moreover, suppose that J is Fréchet differentiable. If there exist multipliers $p \in V$, $\mu \in V^\star$ satisfying (6.2), then (6.4) holds.

Proof. For an arbitrary $h \in U$ we define $x := S'(B\,\bar u + f; B\,h)$. Then

$$\begin{aligned}\langle J_y(\cdot), x\rangle_{V^\star,V} + \langle J_u(\cdot), h\rangle_{U^\star,U} &= \langle -A'(\bar y)^\star p - (I - \Phi'(\bar y))^\star\mu, x\rangle_{V^\star,V} + \langle B^\star p, h\rangle_{U^\star,U} \\ &= -\langle p, A'(\bar y)\,x - B\,h\rangle_{V^\star,V} - \langle \mu, (I - \Phi'(\bar y))\,x\rangle_{V^\star,V}.\end{aligned}$$

From the linearized QVI (5.2) and the strong stationarity system (6.2), we have

$$\begin{aligned}(I - \Phi'(\bar y))\,x &\in \mathcal{K}_K(\bar z, \bar\lambda), & A'(\bar y)\,x - B\,h &\in -\mathcal{K}_K(\bar z, \bar\lambda)^\circ, \\ p &\in -\mathcal{K}_K(\bar z, \bar\lambda), & \mu &\in \mathcal{K}_K(\bar z, \bar\lambda)^\circ,\end{aligned}$$

where we used (6.3). Thus, (6.4) follows.

Acknowledgement

This work is supported by the DFG grant *Analysis and Solution Methods for Bilevel Optimal Control Problems* (grant number WA3636/4-1) within the Priority Program SPP 1962 (Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization).

The simple proof of Lemma 2.1 was pointed out by Felix Harder.

References

Alphonse, Amal; Hintermüller, Michael; Rautenberg, Carlos N. (2018). “Recent trends and views on elliptic quasi-variational inequalities”. Preprint. DOI: 10.20347/wias.preprint.2518.

Alphonse, Amal; Hintermüller, Michael; Rautenberg, Carlos N. (2019). “Directional differentiability for elliptic quasi-variational inequalities of obstacle type”. *Calculus of Variations and Partial Differential Equations* 58.1. DOI: 10.1007/s00526-018-1473-0.

Barrett, John W.; Prigozhin, Leonid (2013). “Sandpiles and superconductors: nonconforming linear finite element approximations for mixed formulations of quasi-variational inequalities”. *IMA Journal of Numerical Analysis* 35.1, pp. 1–38. DOI: 10.1093/imanum/drt062.

Bauschke, Heinz H.; Combettes, Patrick L. (2011). *Convex analysis and monotone operator theory in Hilbert spaces*. Berlin: Springer, pp. xvi + 468. DOI: 10.1007/978-1-4419-9467-7.

Bensoussan, Alain; Lions, Jacques-Louis (1987). *Impulse Control and Quasi Variational Inequalities*. Modern Applied Mathematics Series. John Wiley & Sons Canada, Limited.

Bubeck, Sébastien (2015). “Convex Optimization: Algorithms and Complexity”. *Foundations and Trends® in Machine Learning* 8.3-4, pp. 231–357. DOI: 10.1561/2200000050.

Cartan, Henri (1967). *Calcul différentiel*. Paris: Hermann, p. 178.

Christof, Constantin (2018). “Sensitivity analysis of elliptic variational inequalities of the first and the second kind”. PhD thesis. DOI: 10.17877/de290r-19056.

Christof, Constantin; Wachsmuth, Gerd (2019). “Differential Sensitivity Analysis of Variational Inequalities with Locally Lipschitz Continuous Solution Operators”. *Applied Mathematics & Optimization*. To appear. DOI: 10.1007/s00245-018-09553-y. arXiv: 1711.02720.

Dreves, Axel (2015). “Uniqueness for Quasi-variational Inequalities”. *Set-Valued and Variational Analysis* 24.2, pp. 285–297. DOI: 10.1007/s11228-015-0339-2.

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

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

Nesterov, Yurii (2004). *Introductory Lectures on Convex Optimization*. Springer US. DOI: 10.1007/978-1-4419-8853-9.

Nesterov, Yurii; Scrimali, Laura (2011). “Solving strongly monotone variational and quasi-variational inequalities”. *Discrete and Continuous Dynamical Systems* 31.4, pp. 1383–1396. DOI: 10.3934/dcds.2011.31.1383.

Noor, M Aslam; Oettli, Werner (1994). “On general nonlinear complementarity problems and quasi-equilibria”. *Le Matematiche* 49.2, pp. 313–331.

Prigozhin, Leonid (1996). “Variational model of sandpile growth”. *European Journal of Applied Mathematics* 7.03. DOI: 10.1017/s0956792500002321.

Wachsmuth, Gerd (2019). “A guided tour of polyhedric sets”. *Journal of Convex Analysis* 26.1, pp. 153–188. URL: http://www.heldermann.de/JCA/JCA26/JCA261/jca26010.htm.