Deutsche Forschungsgemeinschaft
Priority Programme 1962

From Resolvents to Generalized Equations and Quasi-variational Inequalities: Existence and Differentiability

Gerd Wachsmuth

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

Preprint Number SPP1962-177

received on September 30, 2021

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

From resolvents to generalized equations and quasi-variational inequalities: existence and differentiability

Gerd Wachsmuth

September 28, 2021
Chair of Optimal Control, BTU Cottbus-Senftenberg

arXiv:2109.13569v1 [math.OC] 28 Sep 2021

We consider a generalized equation governed by a strongly monotone and Lipschitz single-valued mapping and a maximally monotone set-valued mapping in a Hilbert space. We are interested in the sensitivity of solutions w.r.t. perturbations of both mappings. We demonstrate that the directional differentiability of the solution map can be verified by using the directional differentiability of the single-valued operator and of the resolvent of the set-valued mapping. The result is applied to quasi-generalized equations in which we have an additional dependence of the solution within the set-valued part of the equation.

Keywords: generalized equations, variational inclusion, directional differentiability, resolvent operator, proto-derivative, quasi-variational inequality

MSC (2020): 49J53, 47J22, 49J50

1 Introduction

We consider the (local) solution mapping $S\colon U \to Y$, $u \mapsto y$, of the generalized equation

$$0 \in A(y, u) + B(y, u). \tag{1}$$

Here, Y is a (real) Hilbert space, U is a Banach space, $A\colon Y \times U \to Y^\star$ is (locally) strongly monotone and Lipschitz w.r.t. its first argument and the set-valued map $B\colon Y \times U \rightrightarrows Y^\star$ is assumed to be maximally monotone w.r.t. its first argument.

The equation (1) can be used to model many real-world phenomena. In the case where $B(\cdot, u)$ coincides with the subdifferential ∂j_u of a proper, lower-semicontinuous and convex function $j_u \colon Y \to (-\infty, \infty]$, (1) is a variational inequality (VI) of the second kind. If further, j_u is the indicator function $\delta_{C_u} \colon Y \to \{0, \infty\}$ of a non-empty, closed and convex set $C_u \subset Y$, it is a VI of the first kind.

We show by some simple arguments that the existence of solutions and directional differentiability of S follows from the properties of A and of the resolvent J_B of B and from the directional differentiability of A and J_B, respectively. Here, $J_B \colon Y \times U \to Y$, $(q, u) \mapsto y$, is the solution map of

$$0 \in R(y - q) + B(y, u) \tag{2}$$

where $R \colon Y \to Y^\star$ is the Riesz map of the Hilbert space Y. Similarly, we treat the equation

$$0 \in A(y, u) + B(y - \Phi(y, u), u), \tag{3}$$

where $\Phi \colon Y \times U \to Y$ is assumed to have a small Lipschitz constant w.r.t. the first variable. Note that (3) is an inclusion which generalizes the setting of so-called quasi-variational inequalities (QVIs).

Let us put our work in perspective. The existence of solutions to (1) is well understood, we refer to, e.g., Bauschke, Combettes, 2011, Section 23.4. The first contributions which study differentiability of problems similar to (1) are Mignot, 1976; Haraux, 1977 in which the case $B = \partial \delta_C$ is considered under the assumption that $C \subset Y$ is polyhedric, see also Wachsmuth, 2019. The case of non-linear A was treated in Levy, Rockafellar, 1994; Levy, 1999. Later, theory for the differentiability of J_B with $B = \partial j$ was set up in Do, 1992. It was shown that the differentiability of $J_{\partial j}$ is equivalent to the so-called twice epi-differentiability of j, see also Christof, Wachsmuth, 2020. Finally, Adly, Rockafellar, 2020, Theorem 1 study (1) with a real parameter $u \geq 0$. However, since we are mainly interested in directional differentiability, this is not a restriction.

Contributions corresponding to the QVI case (3) are rather new and are currently restricted to the special case

$$\text{Find } y \in K(y) \quad \text{s.t.} \quad \langle A(y) - u, v - y \rangle \geq 0 \quad \forall v \in K(y)$$

with $K(y) = K - \Phi(y)$ for some polyhedric set $K \subset Y$. Important parameters for the study of this problem are the constant μ_A of strong monotonicity of A and the Lipschitz constants L_A and L_Φ of A and Φ, respectively. The first contribution in this direction is Alphonse, Hintermüller, Rautenberg, 2019, Theorem 1. Therein, the authors showed directional differentiability into non-negative directions under monotonicity assumptions on A and $L_\Phi < \mu_A/(\mu_A + L_A)$. Afterwards, Wachsmuth, 2020, Theorem 5.5 proved the directional differentiability into arbitrary directions and the smallness assumption on Φ was relaxed to $L_\Phi < \mu_A/L_A$ (and an even weaker inequality suffices if A is the derivative of a convex function). However, this result needs that Φ is Fréchet differentiable (or, at least Bouligand differentiable). Later, Alphonse, Hintermüller, Rautenberg, 2021, Theorem 3.2 showed that it is sufficient to have a directionally differentiable Φ at the price of having again the stricter requirement $L_\Phi < \mu_A/(\mu_A + L_A)$. A different approach, which is based on concavity properties and which is not restricted to the special case above, is given in Christof, Wachsmuth, 2021.

The main contributions of our paper are the following.

(i) Theorem 12 shows that the solution map of (1) is directionally differentiable. This result is very similar to Adly, Rockafellar, 2020, Theorem 1. However, our assumptions are localized around a solution and we require the directional differentiability of the resolvent J_B instead of the proto-differentiability of B. This results in a much easier proof.

(ii) Lemma 16 shows that the directional differentiability of J_B is equivalent to B being proto-differentiable with a maximally monotone proto-derivative. Thus, the differentiability assumptions on B in Theorem 12 and Adly, Rockafellar, 2020, Theorem 1 coincide.

(iii) In Section 4 we show that the approaches of Alphonse, Hintermüller, Rautenberg, 2019, Theorem 1 and Wachsmuth, 2020, Theorem 5.5 can be generalized to deal with the solution map of (3). Moreover, we only need directional differentiability of the data functions A, Φ and J_B as in Alphonse, Hintermüller, Rautenberg, 2019 and only the weaker requirements on L_Φ from Wachsmuth, 2020.

The paper is structured as follows. In Section 2 we fix some notation and state Theorem 2 concerning convex functions with a strongly monotone and Lipschitz continuous derivative. Section 3 contains the differentiability result for (1), whereas Section 4 is concerned with (3). Some applications are presented in Section 5.

2 Notation

Let U and Y be a Banach space and a Hilbert space, respectively. We denote by $R\colon Y \to Y^\star$ the Riesz map of Y. For $\varepsilon > 0$ and $y^* \in Y$, $B_\varepsilon(y^*)$ $(U_\varepsilon(y^*))$ denotes the closed (open) ball in Y with center y^* and radius ε, respectively.

If $B\colon Y \times U \rightrightarrows Y^\star$ and $\mathcal{U} \subset U$ are given such that $B(\cdot, u)\colon Y \rightrightarrows Y^\star$ is maximally monotone for all $u \in \mathcal{U}$, we say that B is a parametrized maximally monotone operator and we define its resolvent $J_B\colon Y \times \mathcal{U} \to Y$ via $J_B(\cdot, u) := J_{B(\cdot,u)}$, i.e., for $(q, u) \in Y \times \mathcal{U}$ the point $y = J_B(q, u) = (R + B(\cdot, u))^{-1}$ is the unique solution of

$$0 \in R(y - q) + B(y, u).$$

For a closed, convex subset $K \subset Y$, we denote by $T_K\colon Y \rightrightarrows Y$ and $N_K\colon Y \rightrightarrows Y^\star$ the tangent-cone and normal-cone map. Moreover, by

$$K^\circ := \{\mu \in Y^\star \mid \langle \mu, v\rangle \le 0 \;\forall v \in K\}, \qquad \mu^\perp := \{v \in Y \mid \langle \mu, v\rangle = 0\}$$

we denote the polar cone of K and the annihilator of $\mu \in Y^\star$, respectively.

We need a characterization of convex functions defined on subsets of Y with a Lipschitz continuous derivative.

Theorem 1. Let $\mathcal{Y} \subset Y$ be nonempty, open and convex. Further, let $f\colon \mathcal{Y} \to \mathbb{R}$ be convex and let $L_f \in (0, \infty)$ be given. Then, the following assertions are equivalent.

(i) f is Gâteaux differentiable on $\mathcal{Y}$ and $f'\colon \mathcal{Y} \to Y^\star$ is L_f-Lipschitz continuous on $\mathcal{Y}$.

(ii) f is Gâteaux differentiable on $\mathcal{Y}$ and $f'\colon \mathcal{Y} \to Y^\star$ is $1/L_f$-cocoercive, i.e.,

$$\langle f'(y_2) - f'(y_1), y_2 - y_1 \rangle \ge \frac{1}{L_f} \|f'(y_2) - f'(y_1)\|^2 \qquad \forall y_1, y_2 \in \mathcal{Y}.$$

(iii) The map $\frac{L_f}{2}\|\cdot\|^2 - f$ is convex on $\mathcal{Y}$.

For the proof, we refer to Pérez-Aros, Vilches, 2019, Theorem 3.1.

Next, we give an important inequality for convex functions. This inequality is well-known in the finite-dimensional case if the convex function is defined on the entire space, see, e.g., Nesterov, 2004, Theorem 2.1.12 or Bubeck, 2015, Lemma 3.10. The infinite dimensional version (on the entire space) was given in Wachsmuth, 2020, Lemma 3.4. By using Theorem 1, we can adopt the proof to the situation at hand.

Theorem 2. Let $\mathcal{Y} \subset Y$ be nonempty, open and convex. Further, let $f\colon \mathcal{Y} \to \mathbb{R}$ be convex such that $f'\colon \mathcal{Y} \to Y^\star$ is strongly monotone with constant $\mu_f \in (0, \infty)$ and Lipschitz continuous with constant $L_f \in (0, \infty)$. Then,

$$\langle f'(y_2) - f'(y_1), y_2 - y_1 \rangle \ge \frac{\mu_f\, L_f}{\mu_f + L_f} \|y_2 - y_1\|^2 + \frac{1}{\mu_f + L_f} \|f'(y_2) - f'(y_1)\|^2$$

for all $y_1, y_2 \in \mathcal{Y}$.

Proof. We define $g := f - \frac{\mu_f}{2}\|\cdot\|^2$. The strong monotonicity of f' implies that g is convex. From Theorem 1 we infer that $\frac{L_f}{2}\|\cdot\|^2 - f = \frac{L_f - \mu_f}{2}\|\cdot\|^2 - g$ is convex. Applying Theorem 1 again shows that g' is $1/(L_f - \mu_f)$-cocoercive, i.e., for arbitrary $y_1, y_2 \in \mathcal{Y}$ we have

$$\begin{aligned}
&(L_f - \mu_f)\Big(\langle f'(y_2) - f'(y_1), y_2 - y_1 \rangle - \mu_f \|y_2 - y_1\|^2\Big) \\
&\qquad = (L_f - \mu_f)\langle g'(y_2) - g'(y_1), y_2 - y_1 \rangle \\
&\qquad \ge \|g'(y_2) - g'(y_1)\|^2 \\
&\qquad = \|f'(y_2) - f'(y_1)\|^2 - 2\mu_f \langle f'(y_2) - f'(y_1), y_2 - y_1 \rangle + \mu_f^2 \|y_2 - y_1\|^2.
\end{aligned}$$

Rearranging terms yields the claim.

3 Generalized equations

We consider the solution mapping of the generalized equation (1). We show that solutions are locally stable and directionally differentiable under suitable assumptions.

3.1 Local solvability

We set up the standing assumptions which allow to prove that (1) is uniquely solvable around a given reference solution (y^*, u^*).

Assumption 3 (Standing assumptions)**.** Let $(y^*, u^*) \in Y \times U$ be given and let $\mathcal{U} \subset U$ be a neighborhood of u^*.

(i) For all $u \in \mathcal{U}$, $A(\cdot, u)$ is locally (uniformly) strongly monotone and Lipschitz in a neighborhood of y^*. That is, there exists constants $\varepsilon, \mu_A, L_A > 0$ such that

$$\begin{aligned} \langle A(y_2, u) - A(y_1, u), y_2 - y_1 \rangle &\ge \mu_A \|y_2 - y_1\|^2 \\ \|A(y_2, u) - A(y_1, u)\| &\le L_A \|y_2 - y_1\| \end{aligned}$$

holds for all $y_1, y_2 \in B_\varepsilon(y^*)$ and for all $u \in \mathcal{U}$.

(ii) $B(\cdot, u) \colon Y \rightrightarrows Y^\star$ is maximally monotone for all $u \in \mathcal{U}$.

Here and in the sequel, it is sufficient that A is defined only on $B_\varepsilon(y^*) \times \mathcal{U}$ and that B is defined on $Y \times \mathcal{U}$.

We check that the VI has a unique solution in a neighborhood of y^* for certain "small" perturbations u of u^*.

Theorem 4. Let Assumption 3 be satisfied by a solution (y^*, u^*) of (1). We select $\rho \in (0, 2\mu_A/L_A^2)$, $r \in (0, \varepsilon]$ and set $c := \sqrt{1 - 2\rho\mu_A + \rho^2 L_A^2} \in (0, 1)$. Suppose that $\zeta \in Y^\star$ and $u \in \mathcal{U}$ satisfy

$$\|J_{\rho B}(q_\rho^*, u) - J_{\rho B}(q_\rho^*, u^*)\| + \rho\|A(y^*, u) - A(y^*, u^*)\| + \rho\|\zeta\| \le (1 - c)r, \tag{4}$$

where $q_\rho^* = y^* - \rho R^{-1} A(y^*, u^*)$. Then, there exist unique solutions $y, \tilde{y}$ of

$$0 \in \zeta + A(y, u) + B(y, u), \quad y \in B_\varepsilon(y^*) \tag{5}$$

and

$$0 \in A(\tilde{y}, u) + B(\tilde{y}, u), \quad \tilde{y} \in B_\varepsilon(y^*). \tag{6}$$

These solutions satisfy $y, \tilde{y} \in B_r(y^*)$ and $\|y - \tilde{y}\| \le \frac{\rho}{1-c}\|\zeta\|$.

Proof. The proof is inspired by the proof of Brezis, 2011, Theorem 5.6 which treats the special case of a variational inequality.
We define $T\colon Y \to Y$, $y \mapsto z$ as the solution operator of

$$0 \in R(z - (y - \rho R^{-1} A(y,u) - \rho R^{-1}\zeta)) + \rho B(z,u),$$

i.e.,

$$T(y) = J_{\rho B}(y - \rho R^{-1} A(y,u) - \rho R^{-1}\zeta, u).$$

In order to apply the Banach fixed-point theorem, we check that the mapping $Q\colon B_\varepsilon(y^*) \to Y$, $x \mapsto x - \rho R^{-1} A(x,u)$ is a contraction. Indeed,

$$\begin{aligned}\|x - \rho R^{-1}A(x,u) - y + \rho R^{-1}A(y,u)\|^2 &= \|x-y\|^2 + \|\rho R^{-1}(A(x,u) - A(y,u))\|^2 \\ &\qquad - 2\rho\langle A(x,u) - A(y,u), x - y\rangle \\ &\le \Big(1 - 2\rho\mu_A + \rho^2 L_A^2\Big)\|x-y\|^2\end{aligned}$$

holds for all $x, y \in B_\varepsilon(y^*)$ due to Assumption 3(i). Hence, Q is Lipschitz with constant $c \in (0,1)$. Since $J_{\rho B}(\cdot,u)$ is Lipschitz with constant 1, $T\colon B_\varepsilon(y^*) \to Y$ is a contraction. It remains to check that T maps $B_r(y^*)$ onto itself. To this end, we denote the three terms on the left-hand side of (4) by κ_1, κ_2 and κ_3, respectively. By using $y^* = J_{\rho B}(q_\rho^*, u^*)$, we have for an arbitrary $y \in B_r(y^*)$

$$\begin{aligned}\|T(y) - y^*\| &= \|J_{\rho B}(Q(y) - \rho R^{-1}\zeta, u) - J_{\rho B}(q_\rho^*, u^*)\| \\ &\le \|J_{\rho B}(Q(y) - \rho R^{-1}\zeta, u) - J_{\rho B}(q_\rho^*, u)\| + \|J_{\rho B}(q_\rho^*, u) - J_{\rho B}(q_\rho^*, u^*)\| \\ &\le \|y - \rho R^{-1}A(y,u) - \rho R^{-1}\zeta - (y^* - \rho R^{-1}A(y^*,u^*))\| + \kappa_1 \\ &\le \|y - \rho R^{-1}A(y,u) - (y^* - \rho R^{-1}A(y^*,u))\| + \rho\|A(y^*,u) - A(y^*,u^*)\| \\ &\qquad + \rho\|\zeta\| + \kappa_1 \\ &= \|Q(y) - Q(y^*)\| + \kappa_1 + \kappa_2 + \kappa_3 \le r.\end{aligned}$$

Hence, we have shown that $T\colon B_r(y^*) \to B_r(y^*)$ is a contraction.
The Banach fixed-point theorem yields a unique fixed point $y \in B_r(y^*)$ and this is a solution of (5). Similarly, the solvability of (6) is obtained by using the same arguments with $\zeta = 0$. By repeating the same argument with $r = \varepsilon$, we establish the uniqueness of solutions in $B_\varepsilon(y^*)$.
For the final estimate, we note

$$\tilde y = J_{\rho B}(\tilde y - \rho R^{-1}A(\tilde y, u), u), \qquad y = J_{\rho B}(y - \rho R^{-1}A(y,u) - \rho R^{-1}\zeta, u).$$

Thus,

$$\|y - \tilde y\| \le \|Q(y) - Q(\tilde y)\| + \rho\|\zeta\| \le c\|y - \tilde y\| + \rho\|\zeta\|$$

and this gives the desired estimate.

The second term on the left-hand side of (4) becomes small if we assume some continuity of A and u is close to u^*. In order to control the first term, we use a famous result by Attouch.

Lemma 5 (Attouch, 1984, Proposition 3.60)**.** Let $(B_n)_{n\in\mathbb{N}}$ be a sequence of maximal monotone operators on Y and let $B_0 \colon Y \rightrightarrows Y^\star$ be maximally monotone. Then, the following are equivalent.

(i) There exists $\rho_0 > 0$ such that for all $y \in Y$ we have $J_{\rho_0 B_n}(y) \to J_{\rho_0 B_0}(y)$.

(ii) For all $\rho > 0$ and all $y \in Y$ we have $J_{\rho B_n}(y) \to J_{\rho B_0}(y)$.

Corollary 6. Suppose that for every $v \in Y$, the mapping $J_B(v, \cdot) \colon \mathcal{U} \to Y$ is continuous at u^*. Then, for all $\rho > 0$ and $v \in Y$, $J_{\rho B}(v, \cdot) \colon \mathcal{U} \to Y$ is continuous at u^*.

Proof. We have to show that $J_{\rho B}(v, u_n) \to J_{\rho B}(v, u^*)$ for all sequences $(u_n)_{n\in\mathbb{N}} \subset \mathcal{U}$ with $u_n \to u^*$. This follows directly from Lemma 5 by setting $B_n := B(\cdot, u_n)$ and $B_0 := B(\cdot, u^*)$.

Hence, if Assumption 3 and the assumptions of Corollary 6 are satisfied and if the map $A(y^*, \cdot) \colon \mathcal{U} \to Y$ is continuous at u^*, the estimate (4) holds if ζ is small enough and if u is sufficiently close to u^*. For later reference, we also give a directional version of this statement.

Lemma 7. Let Assumption 3 be satisfied by a solution (y^*, u^*) of (1) and choose ρ, r, c as in Theorem 4. Further, let $h \in U$ be arbitrary. We assume that $t \mapsto A(y^*, u^* + th) \in Y^\star$ and $t \mapsto J_B(v, u^* + th) \in Y$ are right-continuous at $t = 0$ for all $v \in Y$. Then, the estimate (4) is satisfied by $u = u^* + th$ and $\zeta \in Y^\star$ if $t > 0$ and $\|\zeta\|$ are small enough.

3.2 Directional differentiability

In order to prove directional differentiability of the solution mapping, we need some differentiability assumptions concerning the mappings A and B.

Assumption 8 (Differentiability assumptions)**.** In addition to Assumption 3, we suppose the following.

(i) A is directionally differentiable at (y^*, u^*).

(ii) J_B is directionally differentiable at (q^*, u^*) with $q^* = y^* - R^{-1}A(y^*, u^*)$.

Interestingly, the next result shows that Assumption 8(ii) already implies that the directional derivative of J_B is again a resolvent of a parametrized maximally monotone operator. In the setting that B is a subdifferential and independent of u, this result follows from Do, 1992, Theorems 3.9, 4.3.

Lemma 9. Let Assumption 8 be satisfied by a solution (y^*, u^*) of (1). We further set $\xi^* := -A(y^*, u^*) \in B(y^*, u^*)$. Then, the operator $DB(y^*, u^* \mid \xi^*)\colon Y \times U \rightrightarrows Y^\star$,

$$DB(y^*, u^* \mid \xi^*)(\delta, h) := \{R(k - \delta) \mid k \in Y, \delta = J_B'(q^*, u^*; k, h)\} \qquad \forall (\delta, h) \in Y \times U, \tag{7}$$

is a parametrized maximally monotone operator and we have

$$J_B'(q^*, u^*; \cdot) = J_{DB(y^*, u^* \mid \xi^*)}, \tag{8}$$

i.e., $\delta = J_B'(q^*, u^*; k, h) = J_{DB(y^*, u^* \mid \xi^*)}(k, h)$ if and only if δ solves

$$0 \in R(\delta - k) + DB(y^*, u^* \mid \xi^*)(\delta, h).$$

We have chosen such a complicated name for the linearization of B, since it will turn out that this linearization coincides with the so-called proto-derivative of B, see Section 3.3.

Proof. Let us check the monotonicity w.r.t. the parameter δ. For fixed $h \in U$ and arbitrary $k_1, k_2 \in Y$, we have

$$\begin{aligned}
&\|J_B'(q^*, u^*; k_1, h) - J_B'(q^*, u^*; k_2, h)\|^2 \\
&\qquad \leftarrow \left\| \frac{J_B(q^* + tk_1, u^* + th) - J_B(x)}{t} - \frac{J_B(q^* + tk_2, u^* + th) - J_B(x)}{t} \right\|^2 \\
&\qquad = \frac{1}{t^2} \|J_B(q^* + tk_1, u^* + th) - J_B(q^* + tk_2, u^* + th)\|^2 \\
&\qquad \le \frac{1}{t^2} (J_B(q^* + tk_1, u^* + th) - J_B(q^* + tk_2, u^* + th), t\,(k_1 - k_2)) \\
&\qquad = \left(\frac{J_B(q^* + tk_1, u^* + th) - J_B(x)}{t} - \frac{J_B(q^* + tk_2, u^* + th) - J_B(x)}{t}, k_1 - k_2 \right) \\
&\qquad \to (J_B'(q^*, u^*; k_1, h) - J_B'(q^*, u^*; k_2, h), k_1 - k_2).
\end{aligned}$$

Here, the convergences are as $t \searrow 0$ and the inequality follows from the fact that resolvents of maximally monotone operators are firmly non-expansive. Hence,

$$(J_B'(q^*, u^*; k_1, h) - J_B'(q^*, u^*; k_2, h), k_1 - J_B'(q^*, u^*; k_1) - (k_2 - J_B'(q^*, u^*; k_2))) \ge 0,$$

i.e.,

$$\langle R(k_1 - \delta_1) - R(k_2 - \delta_2), \delta_1 - \delta_2 \rangle \ge 0,$$

where $\delta_i := J_B'(q^*, u^*; k_i, h)$. This shows monotonicity of $DB(y^*, u^* \mid \xi^*)(\cdot, h)$.
Minty's theorem, see Bauschke, Combettes, 2011, Theorem 21.1, implies that the operator $DB(y^*, u^* \mid \xi^*)(\cdot, h)$ is maximally monotone if and only if $R + DB(y^*, u^* \mid \xi^*)(\cdot, h)$ is surjective. This, however, is obvious: for an arbitrary $\zeta \in Y^\star$, we can set $\delta = J_B'(q^*, u^*; R^{-1}\zeta, h)$ and have $\zeta \in R\delta + DB(y^*, u^* \mid \xi^*)(\delta, h)$.
Finally, (8) follows from a straightforward calculation.

Remark 10. Minimal changes to the proof show that the assertion of Lemma 9 remains true if we only assume that J_B is weakly directionally differentiable at (q^*, u^*), i.e., if

$$\frac{J_B(q^* + tk, u^* + th) - J(q^*, u^*)}{t} \rightharpoonup J_B'(q^*, u^*; k, h) \qquad \text{in } Y \text{ as } t \searrow 0$$

for all $(k, h) \in Y \times U$.

Next, we apply Lemma 9 to the normal cone mapping of a polyhedric set.

Proposition 11. Suppose that $K \subset Y$ is polyhedric and $B(\cdot, u) := N_K$ is the normal cone mapping to K (independent of u). Then, Assumption 8(ii) is satisfied and we have

$$DB(y^* \mid \xi^*)(\delta) = N_{\mathcal{K}}(\delta) = \begin{cases} \mathcal{K}^\circ \cap \delta^\perp & \text{if } \delta \in \mathcal{K}, \\ \emptyset & \text{if } \delta \notin \mathcal{K}, \end{cases}$$

where $\mathcal{K} = T_K(y^*) \cap (\xi^*)^\perp$ denotes the critical cone. That is, $DB(y^* \mid \xi^*)$ is the normal cone mapping to the critical cone $\mathcal{K}$. Here, we have suppressed the arguments u^* and h.

Proof. From Mignot, 1976; Haraux, 1977 we get the directional differentiability of $J_B = \operatorname{Proj}_K$ and

$$\operatorname{Proj}_K'(q^*; k) = \operatorname{Proj}_{\mathcal{K}}(k).$$

Using (7), we have

$$\begin{aligned} DB(y^* \mid \xi^*)(\delta) &:= \{R(k - \delta) \mid k \in Y, \delta = \operatorname{Proj}_{\mathcal{K}}(k)\} \\ &= \{R(k - \delta) \mid k \in Y, \delta \in \mathcal{K}, R(k - \delta) \in \mathcal{K}^\circ, \langle R(k - \delta), \delta\rangle = 0\} \end{aligned}$$

and the claim follows.

Theorem 12. Suppose that Assumption 8 is satisfied by a solution (y^*, u^*) of (1). We denote by S the local solution mapping of (1), cf. Theorem 4 and Lemma 7. Then, S is directionally differentiable at u^*. For $h \in U$, the derivative $\delta = S'(u^*; h)$ is the unique solution of

$$0 \in A'(y^*, u^*; \delta, h) + DB(y^*, u^* \mid \xi^*)(\delta, h), \tag{9}$$

where $\xi^* = -A(y^*, u^*)$.

Proof. The proof is inspired by Levy, 1999. Since $A'(y^*, u^*; \cdot, h) \colon Y \to Y^\star$ is again strongly monotone and Lipschitz, and since the operator $DB(y^*, u^* \mid \xi^*)(\cdot, h) \colon Y \rightrightarrows Y^\star$ is maximally monotone by Lemma 9, it is clear that the linearized equation possesses a unique solution $\delta \in Y$ for an arbitrary $h \in U$. It remains to check that δ is the directional derivative of S. We set

$$q^* := y^* - R^{-1}A(y^*, u^*), \qquad k := \delta - R^{-1}A'(y^*, u^*; \delta, h).$$

This implies $y^* = J_B(q^*, u^*)$ and $\delta = J_B'(q^*, u^*; k, h)$, cf. (8).

Next, we take an arbitrary sequence $(t_n)_{n\in\mathbb{N}} \subset (0, \infty)$ with $t_n \searrow 0$ and define

$$y_n := J_B(q^* + t_n k, u^* + t_n h), \qquad \text{i.e.,} \quad 0 \in R(y_n - q^* - t_n k) + B(y_n, u^* + t_n h)$$

for n large enough. Then, Assumption 8(ii) implies

$$\frac{y_n - y^*}{t_n} \to J_B'(q^*, u^*; k, h) = \delta.$$

Now, we define

$$\zeta_n := \frac{R(y_n - q^* - t_n k) - A(y_n, u^* + t_n h)}{t_n}.$$

By using the definition of q^* we get

$$\begin{aligned}\zeta_n &= \frac{R(y_n - y^* - t_n k) + A(y^*, u^*) - A(y_n, u^* + t_n h)}{t_n} \\ &= R\Big(\frac{y_n - y^*}{t_n} - k\Big) + \frac{A(y^*, u^*) - A(y^* + t_n\delta, u^* + t_n h)}{t_n} \\ &\qquad + \frac{A(y^* + t_n\delta, u^* + t_n h) - A(y_n, u^* + t_n h)}{t_n}.\end{aligned}$$

Due to Assumption 3(i), the last addend can be bounded by $L_A\|(y_n - y^*)/t_n - \delta\| \to 0$. Thus, the directional differentiability of A implies $\zeta_n \to R(\delta - k) - A'(y^*, u^*, \delta, h) = 0$. Next, we note that the definition of ζ_n yields

$$0 \in t_n\zeta_n + A(y_n, u^* + t_n h) + B(y_n, u^* + t_n h).$$

Due to Lemma 7, we can apply Theorem 4 for n large enough. This yields the existence of a unique solution $\tilde{y}_n \in B_\varepsilon(y^*)$ of

$$0 \in A(\tilde{y}_n, u^* + t_n h) + B(\tilde{y}_n, u^* + t_n h)$$

and this solution satisfies $\|y_n - \tilde{y}_n\| \le C\|t_n\zeta_n\|$. Note that $\tilde{y}_n = S(u^* + t_n h)$. Hence,

$$\frac{S(u^* + t_n h) - S(u^*)}{t_n} = \frac{\tilde{y}_n - y_n}{t_n} + \frac{y_n - y^*}{t_n} \to 0 + \delta.$$

This shows the claim.

We mention that a similar result has been given in Adly, Rockafellar, 2020, Theorem 1. Therein, global assumptions on A are used, i.e., Assumption 3(i) is required to hold for $\varepsilon = \infty$. Moreover, this contribution uses the concept of proto-differentiability, which is not utilized in our proof. In particular, instead of Assumption 8(ii), Adly, Rockafellar, 2020 require that B is proto-differentiable with a maximally monotone proto-derivative.

In our opinion, it is often easier and more natural to study the differentiability properties of J_B instead of checking whether B is proto-differentiable. Indeed, in the case that B

is the normal cone mapping of a polyhedric set $C \subset Y$, the directional differentiability of $J_B = \operatorname{Proj}_C$ was already shown in Mignot, 1976; Haraux, 1977, whereas the proto-differentiability of B was proved later in Do, 1992, Example 4.6, see also Levy, 1999. Moreover, the former proofs are rather elementary, whereas the latter proof utilizes Do, 1992, Theorem 3.9 which is based on Attouch's theorem linking the Mosco-convergence of convex functions with the graphical convergence of their subdifferentials.

Remark 13. We comment on some extensions and limitations of Theorem 12.

(i) The strong monotonicity of A can be replaced by requiring that the linearized equation (9) possesses solutions for all $h \in U$ and by assuming that the assertion of Theorem 4 holds.

(ii) It is not possible to adapt the proof to the situation of Remark 10. Indeed, if we only assume weak directional differentiability of J_B, we only get $(y_n - y^*)/t_n \rightharpoonup \delta$ and, thus, only $\zeta_n \rightharpoonup 0$ (if A is Bouligand differentiable). This, however, is not enough to obtain $\|y_n - \tilde{y}_n\| = o(t_n)$ in the last step of the proof.

(iii) Another approach for proving Theorem 12 is to directly consider $\tilde{y}_n := S(u^* + t_n h)$. Due to the Lipschitz continuity of S, one gets $(\tilde{y}_n - y^*)/t_n \rightharpoonup \delta$ along a subsequence for some $\delta \in Y$. The next step would be to perform a Taylor expansion of $A(\tilde{y}_n, u^* + t_n h) - A(y^*, u^*)$, but this needs stronger differentiability assumptions for A, e.g., Bouligand/Fréchet differentiability w.r.t. its first argument, cf. Christof, Wachsmuth, 2020, Remark 2.3(iii), Theorem 2.14. In the above proof, this is avoided via the construction of y_n.

Finally, we mention that the directional differentiability of $J_{\rho_0 B}$ for some $\rho_0 > 0$ implies the directional differentiability for all $\rho > 0$.

Corollary 14. Let $B \colon Y \times \mathcal{U} \rightrightarrows Y^\star$ as in Assumption 3(ii) be given and fix $(q^*, u^*) \in Y \times \mathcal{U}$. With $y^* = J_B(q^*, u^*)$, the following are equivalent.

(i) There exists $\rho_0 > 0$ such that the resolvent $J_{\rho_0 B}$ is directionally differentiable at $(y^* + \rho_0(q^* - y^*), u^*)$.

(ii) For all $\rho > 0$, $J_{\rho B}$ is directionally differentiable at $(y^* + \rho(q^* - y^*), u^*)$.

Proof. This follows from Theorem 12, since $J_{\rho B} \colon (x, u) \mapsto y$ is the solution mapping of $0 \in \frac{\rho_0}{\rho} R(y - x) + \rho_0 B(y, u)$ and $y^* = J_{\rho B}(y^* + \rho(q^* - y^*), u^*)$.

Similarly, an application of Theorem 12 shows that the directional differentiability of the resolvent J_B is actually independent of the Riesz isomorphism R of Y and, thus, independent of the inner product in Y.

3.3 Relation to proto-differentiability

The purpose of this section is to shed some light on the relation of directional differentiability of the resolvent J_B and the proto-differentiability of B.

We first fix the notion of proto-differentiability of the parametrized set-valued map B.

Definition 15. Let $(y^*, u^*) \in Y \times U$ be given, such that Assumption 3(ii) is satisfied. For some $\xi^* \in B(y^*, u^*)$ and $(\delta, h) \in Y \times U$, we define

$$\Delta_t B(y^*, u^* \mid \xi^*)(\delta, h) := \frac{B(y^* + t\delta, u^* + th) - \xi^*}{t} \tag{10}$$

for $t > 0$ small enough. We say that B is proto-differentiable at (y^*, u^*) relative to $\xi^* \in B(y^*, u^*)$ if the graph of $\Delta_t B(y^*, u^* \mid \xi^*)(\cdot, h) \colon Y \rightrightarrows Y^\star$ converges as $t \searrow 0$ in the sense of Painlevé–Kuratowski, see Adly, Rockafellar, 2020, Definition 2, for all $h \in U$. In this case, we define its proto-derivative $DB(y^*, u^* \mid \xi^*) \colon Y \times U \rightrightarrows Y^\star$ via

$$\operatorname{graph} DB(y^*, u^* \mid \xi^*)(\cdot, h) := \operatorname{graph} \lim_{t \searrow 0} \Delta_t B(y^*, u^* \mid \xi^*)(\cdot, h)$$

for all $h \in U$.

Although the operator $\Delta_t B(y^*, u^* \mid \xi^*)(\cdot, h)$ defined in (10) is maximally monotone, cf. Adly, Rockafellar, 2020, Lemma 1, its graphical limit $DB(y^*, u^* \mid \xi^*)(\cdot, h)$ might fail to be maximally monotone, even if it exists, see Wachsmuth, 2021, Theorem 2.

Lemma 16. Let $(y^*, u^*) \in Y \times U$ be given, such that Assumption 3(ii) is satisfied. Further, let $q^* \in Y$ be given such that $y^* = J_B(q^*, u^*)$ and we set $\xi^* := R(q^* - y^*) \in B(y^*, u^*)$. Then, the following are equivalent.

(i) The mapping J_B is directionally differentiable at (q^*, u^*).

(ii) The mapping B is proto-differentiable at (y^*, u^*) relative to ξ^* and its proto-derivative $DB(y^*, u^* \mid \xi^*)(\cdot, h)$ is maximally monotone for all $h \in U$.

Proof. By defining $A(y, u) := R(y - q^*)$ it can be easily checked that (y^*, u^*) is a solution of (1) and that Assumption 8 is satisfied.
In order to apply the results from Adly, Rockafellar, 2020, we fix $h \in U$ and $t_0 > 0$ such that $u^* + th \in \mathcal{U}$ for all $t \in [0, t_0)$. We define

$$\tilde{B}_h(t, y) := B(y, u^* + th) \qquad \forall t \in [0, t_0), y \in Y.$$

Now, it can be checked that $\Delta_t B(y^*, u^* \mid \xi^*)(\cdot, h)$ coincides with $\Delta_t \tilde{B}_h(y^* \mid \xi^*)(\cdot)$ as defined in Adly, Rockafellar, 2020, (9). Thus, B is proto-differentiable at (y^*, u^*) relative to ξ^* if and only if $\tilde{B}_h$ is proto-differentiable at y^* relative to ξ^* for all $h \in U$ and we have the formula

$$DB(y^*, u^* \mid \xi^*)(\cdot, h) = D\tilde{B}_h(y^* \mid \xi^*)(\cdot)$$

for the corresponding proto-derivatives.
"⇒": Due to $J_B(q, u^* + th) = J_{\tilde{B}_h}(t, q)$, we get that $J_{\tilde{B}_h}$ is directionally differentiable. Further, $J_{\tilde{B}_h}(t, q) \in Y$ is Lipschitz continuous w.r.t. $q \in Y$ and, thus, we get the semi-differentiability (see Adly, Rockafellar, 2020, Definition 1) of $J_{\tilde{B}_h}$ at q^*. Now, we can apply Adly, Rockafellar, 2020, Remark 5 to obtain that $J_{\tilde{B}_h}$ is proto-differentiable at q^* relative to $y^* = J_{\tilde{B}_h}(0, q^*)$. Now, Adly, Rockafellar, 2020, Lemma 2 yields that $\tilde{B}_h$ is proto-differentiable at y^* relative to ξ^*. As explained above, this yields the desired proto-differentiability of B. Moreover, from these arguments, we can distill the formula

$$\begin{aligned} J_B'(q^*, u^*; \cdot, h) = J_{\tilde{B}_h}'(q^*; \cdot) = DJ_{\tilde{B}_h}(q^* \mid y^*) = (R + D\tilde{B}_h(y^* \mid \xi^*))^{-1} \\ = (R + DB(y^*, u^* \mid \xi^*)(\cdot, h))^{-1}. \end{aligned}$$

Since a maximal monotone mapping is uniquely determined by its resolvent, this shows that $DB(y^*, u^* \mid \xi^*)$ coincides with the mapping defined in Lemma 9 and, therefore, is maximally monotone.
"⇐": Using similar arguments as above, this follows from Adly, Rockafellar, 2020, Theorem 2.

In particular, the normal cone mapping to a polyhedric set is proto-differentiable and the proto-derivative is given as in Proposition 11. As already mentioned, this result is known from Do, 1992; Levy, 1999.

4 Quasi-generalized equations

In this section, we treat the generalization

$$0 \in A(y, u) + B(y - \Phi(y, u), u) \tag{3}$$

of (1). Here, $\Phi\colon Y \times U \to Y$ is an additional mapping. We will follow two approaches to investigate the solution mapping of (3). In the first approach, we reformulate (3) in the form (1) by introducing a new variable $z = y - \Phi(y, u)$. This idea was successfully used in Wachsmuth, 2020 to prove the directional differentiability of QVIs. The second approach, which was pioneered in Alphonse, Hintermüller, Rautenberg, 2019, uses a iteration approach, i.e., it builds a sequence $(y_{u,n})_{n\in\mathbb{N}}$, in which $y_{u,n}$ solves

$$0 \in A(y_{u,n}, u) + B(y_{u,n} - \Phi(y_{u,n-1}, u), u).$$

We shall see that both approaches will use a similar set of assumptions.

Assumption 17. We assume that $(y^*, u^*) \in Y \times U$ is a solution of (3) such that the operators A and B satisfy Assumption 3. In addition, $\Phi\colon Y \times U \to Y$ is continuous at (y^*, u^*) and $\Phi(\cdot, u)$ is locally a uniform contraction, i.e., there exists a Lipschitz constant $L_\Phi \in [0, 1)$ such that

$$\|\Phi(y_2, u) - \Phi(y_1, u)\| \le L_\Phi \|y_2 - y_1\|$$

for all $y_1, y_2 \in B_\varepsilon(y^*)$ and for all $u \in \mathcal{U}$.

Assumption 18. In addition to Assumption 17, we require

(a) $L_\Phi < \gamma_A^{-1}$, or

(b) $L_\Phi < 2\sqrt{\gamma_A}/(1 + \gamma_A)$ and there exists a function $g\colon U_\varepsilon(y^*) \times \mathcal{U} \to \mathbb{R}$ such that A is the Fréchet derivative of g w.r.t. the first variable,

where $\gamma_A = L_A/\mu_A \ge 1$ is the local condition number of A.

Assumption 19. In addition to Assumption 17, we assume that

(i) A is directionally differentiable at (y^*, u^*).

(ii) J_B is directionally differentiable at $(q^* - \phi^*, u^*)$ with $q^* = y^* - R^{-1}A(y^*, u^*)$ and $\phi^* = \Phi(y^*, u^*)$.

(iii) Φ is directionally differentiable at (y^*, u^*).

As a preparation, we state a consequence of Assumption 18.

Lemma 20. Let Assumption 18 be satisfied. Then, there are constants $C > 0$ and $\tilde{c} \in (0, 1)$, such that for all $y_1, y_2, z_1, z_2 \in B_\varepsilon(y^*)$ we have

$$\langle A(z_2, u) - A(z_1, u), z_2 - \Phi(y_2, u) - z_1 + \Phi(y_1, u)\rangle \ge C\Big(\|z_2 - z_1\|^2 - \tilde{c}^2\|y_2 - y_1\|^2\Big).$$

Proof. We denote the left-hand side of the inequality by $M \in \mathbb{R}$.
In case that Assumption 18(a) holds, we have

$$M \ge \mu_A\|z_2 - z_1\|^2 - L_A L_\Phi \|z_2 - z_1\|\|y_2 - y_1\| \ge \frac{\mu_A}{2}\|z_2 - z_1\|^2 - \frac{L_A^2 L_\Phi^2}{2\mu_A}\|y_2 - y_1\|^2,$$

i.e., we can use $C = \mu_A/2$ and $\tilde{c} = L_A L_\Phi/\mu_A = \gamma_A L_\Phi < 1$.
Under Assumption 18(b), we first consider $z_1, z_2 \in U_\varepsilon(y^*)$. Theorem 2 yields

$$\begin{aligned} M &\ge \frac{\mu_A L_A}{\mu_A + L_A}\|z_2 - z_1\|^2 + \frac{1}{\mu_A + L_A}\|A(z_2, u) - A(z_1, u)\|^2 \\ &\quad - L_\Phi\|A(z_2, u) - A(z_1, u)\|\|y_2 - y_1\| \\ &\ge \frac{\mu_A L_A}{\mu_A + L_A}\|z_2 - z_1\|^2 - \frac{(\mu_A + L_A)L_\Phi^2}{4}\|y_2 - y_1\|^2. \end{aligned}$$

Here, we have used Young's inequality. Since everything is continuous w.r.t. $z_1, z_2 \in Y$, the inequality carries over to $z_1, z_2 \in B_\varepsilon(y^*)$. Now, we can choose $C = \mu_A L_A/(\mu_A + L_A)$ and $\tilde{c} = (\mu_A + L_A)L_\Phi/(2\sqrt{\mu_A L_A}) = (1 + \gamma_A)L_\Phi/(2\sqrt{\gamma_A}) < 1$.

Interestingly, this already shows that (locally) (3) has at most one solution.

Lemma 21. Let Assumption 18 be satisfied. Then, for each $u \in \mathcal{U}$, (3) has at most one solution in $B_\varepsilon(y^*)$.

Proof. Let $y_1, y_2 \in B_\varepsilon(y^*)$ be solutions of (3). Due to $-A(y_i, u) \in B(y_i - \Phi(y_i), u)$ we get

$$\langle A(y_2, u) - A(y_1, u), y_1 - \Phi(y_1, u) - y_2 + \Phi(y_2, u)\rangle \geq 0$$

and Lemma 20 shows $C(1 - \tilde{c}^2)\|y_2 - y_1\|^2 \leq 0$, i.e., $y_1 = y_2$.

4.1 Via a reformulation as GE

In this first approach, we reformulate (3) via the new variable

$$z = y - \Phi(y, u). \tag{11}$$

Let us first check that this transformation is locally well defined.

Lemma 22. Let Assumption 17 be satisfied. There exists neighborhoods $\mathcal{Z} \subset Y$ and $\hat{\mathcal{U}} \subset \mathcal{U}$ of $z^* := y^* - \Phi(y^*, u^*)$ and of u^*, respectively, such that the equation (11) has a unique solution $y \in B_\varepsilon(y^*)$ for all $(z, u) \in \mathcal{Z} \times \hat{\mathcal{U}}$.
Moreover, the mapping

$$(\operatorname{id} - \Phi(\cdot, u))^{-1} \colon \mathcal{Z} \to B_\varepsilon(y^*)$$

is Lipschitz with constant $(1 - L_\Phi)^{-1}$ for all $u \in \hat{\mathcal{U}}$.

Proof. We define the mappings $\mathcal{A} \colon Y \times (Y \times U) \to Y^\star$ and $\mathcal{B} \colon Y \times (Y \times U) \rightrightarrows Y^\star$ via

$$\mathcal{A}(y, (z, u)) := R(y - z - \Phi(y, u)), \qquad \mathcal{B}(y, (z, u)) := \{0\}.$$

Now, our transformation (11) is equivalent to the generalized equation

$$0 \in \mathcal{A}(y, (z, u)) - \mathcal{B}(y, (z, u))$$

and the first part of the assertion follows from Theorem 4.
To estimate the Lipschitz constant, we take $z_1, z_2 \in \mathcal{Z}$ and set $y_i = (\operatorname{id} - \Phi(\cdot, u))^{-1}(z_i)$ for $i = 1, 2$. Then,

$$\|y_2 - y_1\| = \|z_2 + \Phi(y_2, u) - z_1 - \Phi(y_1, u)\| \leq \|z_2 - z_1\| + L_\Phi\|y_2 - y_1\|$$

and this shows the claim.

Using the result of Lemma 22, we transform (3) into

$$0 \in \tilde{A}(z,u) + B(z,u), \tag{12}$$

where $\tilde{A}\colon \mathcal{Z} \times \hat{\mathcal{U}} \to Y$ is defined via

$$\tilde{A}(z,u) := A((\operatorname{id} - \Phi(\cdot,u))^{-1}(z), u).$$

By inserting the definitions, we verify that (3) and (12) are locally equivalent.

Lemma 23. Let Assumption 17 be satisfied.

(a) If $(y,u) \in B_\varepsilon(y^*) \times \hat{\mathcal{U}}$ is a solution of (3) with $z := y - \Phi(y,u) \in \mathcal{Z}$, then z is a solution of (12).

(b) If $(z,u) \in \mathcal{Z} \times \hat{\mathcal{U}}$ is a solution of (12), then $y := (\operatorname{id} - \Phi(\cdot,u))^{-1}(z)$ is a solution of (3).

Note that $z \in \mathcal{Z}$ in (a) is satisfied if (y,u) is sufficiently close to (y^*,u^*).

By using the ideas of Wachsmuth, 2020, Lemmas 3.3, 3.5, we check that the analysis from Section 3 applies to (12). In what follows, we use $z^* := y^* - \Phi(y^*,u^*)$.

Lemma 24. Let Assumption 18 be satisfied. Then, the operator $\tilde{A}\colon \mathcal{Z} \times \hat{\mathcal{U}} \to Y$ satisfies Assumption 3 at (z^*,u^*).

Proof. Let $\eta > 0$ be given such that $B_\eta(z^*) \subset \mathcal{Z}$. We have to show the existence of $\mu_{\tilde{A}}, L_{\tilde{A}} > 0$ such that

$$\begin{aligned}\langle \tilde{A}(z_2,u) - \tilde{A}(z_1,u), z_2 - z_1\rangle &\ge \mu_{\tilde{A}} \|z_2 - z_1\|^2 \\ \|\tilde{A}(z_2,u) - \tilde{A}(z_1,u)\| &\le L_{\tilde{A}} \|z_2 - z_1\|\end{aligned}$$

holds for all $z_1, z_2 \in B_\eta(z^*)$ and for all $u \in \hat{\mathcal{U}}$. The Lipschitz property follows directly from Lemma 22 and it remains to prove the strong monotonicity.
Let $z_1, z_2 \in B_\eta(z^*)$ and $u \in \hat{\mathcal{U}}$ be arbitrary and set $y_i := (\operatorname{id} - \Phi(\cdot,u))^{-1}(z_i) \in B_\varepsilon(y^*)$ for $i = 1,2$. Then, Lemma 20 implies

$$\begin{aligned}\langle \tilde{A}(z_2,u) - \tilde{A}(z_1,u), z_2 - z_1\rangle &= \langle A(y_2,u) - A(y_1,u), y_2 - y_1 - \Phi(y_2,u) + \Phi(y_1,u)\rangle \\ &\ge C(1 - \tilde{c})\|y_2 - y_1\|^2.\end{aligned}$$

In combination with

$$\|z_2 - z_1\| = \|y_2 - y_1 - \Phi(y_2,u) + \Phi(y_1,u)\| \le (1 + L_\Phi)\|y_2 - y_1\|,$$

this shows the uniform strong monotonicity of $\tilde{A}$.

In order to apply Theorem 12, we have to check that $\tilde{A}$ is directionally differentiable. To this end, we verify that $(z, u) \mapsto (\mathrm{id} - \Phi(\cdot, u))^{-1}(z)$ is directionally differentiable.

Lemma 25. Let Assumption 19 be satisfied. Then, $\Psi\colon \mathcal{Z} \times \hat{\mathcal{U}} \to B_\varepsilon(y^*)$, defined via $\Psi(z, u) := (\mathrm{id} - \Phi(\cdot, u))^{-1}(z)$, is directionally differentiable at (z^*, u^*) and the directional derivative $\delta = \Psi'(z^*, u^*; k, h)$ in direction $(k, h) \in Y \times U$ is the unique solution of

$$\delta - \Phi'(y^*, u^*; \delta, h) = k.$$

Proof. We define $\mathcal{A}$ and $\mathcal{B}$ as in the proof of Lemma 22. Then, Ψ is the solution mapping $(y, u) \mapsto z$ of

$$0 \in \mathcal{A}(y, (z, u)) - \mathcal{B}(y, (z, u))$$

and the assertion follows from Theorem 12.

Corollary 26. Let Assumption 19 be satisfied. Then, the operator $\tilde{A}$ is directionally differentiable at (z^*, u^*) and we have

$$\tilde{A}'(z^*, u^*; k, h) = A'(y^*, u^*; \Psi'(z^*, u^*; k, h), h).$$

Proof. We have

$$\tilde{A}(z, u) = A(\Psi(z, u), u)$$

with the operator Ψ from Lemma 25. Let $(k, h) \in Y \times U$ be arbitrary. For $t > 0$ we have

$$\begin{aligned} \frac{\tilde{A}(z^* + tk, u^* + th) - \tilde{A}(z^*, u^*)}{t} &= \frac{A(\Psi(z^* + tk, u^* + th), u^* + th) - A(\Psi(z^*, u^*), u^*)}{t} \\ &= \frac{A(\Psi(z^* + tk, u^* + th), u^* + th) - A(\Psi(z^*, u^*) + t\Psi'(z^*, u^*; k, h), u^* + th)}{t} \\ &\quad + \frac{A(\Psi(z^*, u^*) + t\Psi'(z^*, u^*; k, h), u^* + th) - A(\Psi(z^*, u^*), u^*)}{t} =: I_1 + I_2. \end{aligned}$$

Due to the Lipschitz continuity of A, the first term on the right-hand side is bounded by

$$\|I_1\| \le \frac{L_A}{t} \|\Psi(z^* + tk, u^* + th) - \Psi(z^*, u^*) - t\Psi'(z^*, u^*; k, h)\| \to 0,$$

where we used the directional differentiability of Ψ, see Lemma 25. Since A is directionally differentiable, we have

$$\lim_{t \searrow 0} I_2 = A'(y^*, u^*; \Psi'(z^*, u^*; k, h), h)$$

and this shows the claim.

Theorem 27. Suppose that Assumptions 18 and 19 are satisfied. Then, for each

$h \in U$, there is $t_0 > 0$ such that

$$0 \in A(y_t, u + th) + B(y_t - \Phi(y_t, u + th), u + th)$$

possesses a unique solution $y_t \in B_\varepsilon(y^*)$ for all $t \in [0, t_0]$. Moreover, $(y_t - y^*)/t \to \delta$ as $t \searrow 0$, where $\delta \in Y$ is the unique solution of

$$0 \in A'(y^*, u^*; \delta, h) + DB(y^* - \Phi(y^*, u^*), u^* \mid \xi^*)(\delta - \Phi'(y^*, u^*; \delta, h), h), \tag{13}$$

where $\xi^* = -A(y^*, u^*)$.

Proof. Due to Lemma 23, (3) locally equivalent to (12). Owing to the previous results, we can apply Theorem 12 to (12). In particular, Assumption 8(i) follows from Corollary 26, whereas Assumption 8(ii) requires directional differentiability of J_B at $(z^* - R^{-1}\tilde{A}(z^*, u^*), u^*) = (q^* - \phi^*, u^*)$ and this is ensured by Assumption 19(ii). If we denote by $z_t \in \mathcal{Z}$ the local solution of (12) w.r.t. $u = u^* + th$, then $(z_t - z^*)/t \to k$, where k is the unique solution to

$$\begin{aligned} 0 &\in \tilde{A}'(z^*, u^*; k, h) + DB(z^*, u^* \mid \xi^*)(k, h) \\ &= A'(y^*, u^*; \Psi'(z^*, u^*; k, h), h) + DB(z^*, u^* \mid \xi^*)(k, h). \end{aligned}$$

It is clear that this equation is equivalent to (13) via the transformation

$$k = \delta - \Phi'(y^*, u^*; \delta, h), \qquad \text{i.e.,} \qquad \delta = \Psi'(z^*, u^*; k, h).$$

Finally,

$$\begin{aligned} \frac{y_t - y^*}{t} &= \frac{\Psi(z_t, u + th) - \Psi(z^*, u^*)}{t} \\ &= \frac{\Psi(z_t, u + th) - \Psi(z^* + tk, u + th)}{t} + \frac{\Psi(z^* + tk, u + th) - \Psi(z^*, u^*)}{t} \\ &\to \Psi'(z^*, u^*; k, h) = \delta, \end{aligned}$$

where we used that $\Psi(\cdot, u + th)$ is Lipschitz continuous, see Lemma 22. The uniqueness of y_t in $B_\varepsilon(y^*)$ follows from Lemma 20.

We mention that Assumption 18 is mainly used (see Lemma 24) to show that the operator $\tilde{A} \colon \mathcal{Z} \times \hat{\mathcal{U}} \to Y$ is locally (uniformly) strongly monotone in a neighborhood of (z^*, u^*).

4.2 Via an iteration approach

Here, we use a different approach to tackle

$$0 \in A(y, u) + B(y - \Phi(y, u), u). \tag{3}$$

For $u \in \mathcal{U}$ sufficiently close to u^*, we consider the sequence $(y_{u,n})_{n\in\mathbb{N}} \subset Y$ defined via

$$y_{u,0} := y^*, \tag{14a}$$
$$0 \in A(y_{u,n}, u) + B(y_{u,n} - \Phi(y_{u,n-1}, u), u). \tag{14b}$$

We will see that this iteration is well defined in the sense that (14b) has a unique solution $y_{u,n} \in B_\varepsilon(y^*)$ under appropriate assumptions. This idea was used in Alphonse, Hintermüller, Rautenberg, 2019 to show the directional differentiability of QVIs. We demonstrate that this idea can also be applied to (3).

In order to study (14b) with the methods from Section 3, we introduce the operators $\mathcal{A}\colon Y \times (Y \times U) \to Y^\star$, $\mathcal{B}\colon Y \times (Y \times U) \rightrightarrows Y^\star$ via

$$\mathcal{A}(y, (\phi, u)) := A(y, u), \qquad \mathcal{B}(y, (\phi, u)) := B(y - \phi, u). \tag{15}$$

Moreover, we set $\phi^* := \Phi(y^*, u^*)$. Then, under Assumption 17, it is clear that Assumption 3 is satisfied by $(\mathcal{A}, \mathcal{B})$ at $(y^*, (y^* - \phi^*, u^*))$.

Moreover, for arbitrary $(q, (\phi, u)) \in Y \times (Y \times \mathcal{U})$ and $\rho > 0$, the point $y = J_{\rho\mathcal{B}}(q, (\phi, u))$ solves

$$0 \in R(y - q) + \rho\mathcal{B}(y, (\phi, u)) = R((y - \phi) - (q - \phi)) + \rho B(y - \phi, u),$$

i.e., we have the relation

$$J_{\rho\mathcal{B}}(q, (\phi, u)) = J_{\rho B}(q - \phi, u) + \phi \tag{16}$$

between the resolvents of B and $\mathcal{B}$.

Next, we address the local solvability of (14b).

Lemma 28. Let Assumption 17 be satisfied and fix ρ, c and q_ρ^* as in Theorem 4. Then, for all $r \in (0, \varepsilon]$, $(\phi, u) \in Y \times \mathcal{U}$ with

$$2\|\phi - \phi^*\| + \|J_{\rho B}(q_\rho^* - \phi^*, u) - J_{\rho B}(q_\rho^* - \phi^*, u^*)\| + \rho\|A(y^*, u) - A(y^*, u^*)\| \le (1 - c)r$$

the equation

$$0 \in A(z, u) + B(z - \phi, u)$$

has a unique solution $z \in B_r(y^*)$.

Proof. The equation can be recast as

$$0 \in \mathcal{A}(y, (\phi, u)) + \mathcal{B}(y, (\phi, u))$$

and we are going to apply Theorem 4 with $\zeta = 0$. It is clear that Assumption 3 is satisfied by $(\mathcal{A}, \mathcal{B})$ and the operator $\mathcal{A}$ possesses the same constants as A. Thus, it remains to show that

$$\|J_{\rho\mathcal{B}}(q_\rho^*, (\phi, u)) - J_{\rho\mathcal{B}}(q_\rho^*, (\phi^*, u^*))\| + \rho\|\mathcal{A}(y^*, (\phi, u)) - \mathcal{A}(y^*, (\phi^*, u^*))\| \le (1 - c)r$$

is satisfied. This, however, follows from the estimate

$$
\begin{aligned}
\|J_{\rho\mathcal{B}}(q_\rho^*, (\phi, u)) - J_{\rho\mathcal{B}}(q_\rho^*, (\phi^*, u^*))\| &= \|J_{\rho B}(q_\rho^* - \phi, u) + \phi - J_{\rho B}(q_\rho^* - \phi^*, u^*) - \phi^*\| \\
&\le \|J_{\rho B}(q_\rho^* - \phi, u) - J_{\rho B}(q_\rho^* - \phi^*, u)\| + \|J_{\rho B}(q_\rho^* - \phi^*, u) - J_{\rho B}(q_\rho^* - \phi^*, u^*)\| \\
&\qquad + \|\phi - \phi^*\| \\
&\le 2\|\phi - \phi^*\| + \|J_{\rho B}(q_\rho^* - \phi^*, u) - J_{\rho B}(q_\rho^* - \phi^*, u^*)\|.
\end{aligned}
$$

Lemma 29. Let Assumption 17 be satisfied and fix ρ, c and q_ρ^* as in Theorem 4. Then, there exists a constant $\lambda \in (0, \varepsilon]$, such that for all $y \in B_\lambda(y^*)$ and all $u \in \mathcal{U}$ with

$$
\begin{aligned}
\mathcal{C}_\rho(u) := 2\|\Phi(y^*, u) - \Phi(y^*, u^*)\| + \|J_{\rho B}(q_\rho^* - \phi^*, u) - J_{\rho B}(q_\rho^* - \phi^*, u^*)\| + \\
\rho\|A(y^*, u) - A(y^*, u^*)\| \le \frac{1-c}{2}\varepsilon
\end{aligned}
\tag{17}
$$

the equation

$$0 \in A(z, u) + B(z - \Phi(y, u), u)$$

has a unique solution $z := T_u(y) \in B_\varepsilon(y^*)$.

Proof. Using

$$
\begin{aligned}
\|\Phi(y, u) - \phi^*\| &\le \|\Phi(y, u) - \Phi(y^*, u)\| + \|\Phi(y^*, u) - \Phi(y^*, u^*)\| \\
&\le L_\Phi\|y - y^*\| + \|\Phi(y^*, u) - \Phi(y^*, u^*)\| \le L_\Phi\lambda + \|\Phi(y^*, u) - \Phi(y^*, u^*)\|
\end{aligned}
$$

the assertion follows from Lemma 28 with $\lambda = (1 - c)\varepsilon/(4L_\Phi)$.

In the next lemma, we apply the Banach fixed-point theorem to T_u in order to show the convergence of (14).

Lemma 30. Let Assumption 18 be satisfied and fix ρ, c, q_ρ^* as in Theorem 4 and choose λ according to Lemma 29.

(a) There is a constant $\tilde{c} \in (0, 1)$, such that $T_u\colon B_\lambda(y^*) \to B_\varepsilon(y^*)$ is Lipschitz continuous with modulus $\tilde{c}$ for all $u \in \mathcal{U}$.

(b) If $u \in \mathcal{U}$ is chosen such that $\mathcal{C}_\rho(u) \le (1-c)\min\{(1-\tilde{c})\lambda, \varepsilon\}$, then T_u maps $B_\lambda(y^*)$ to $B_\lambda(y^*)$. Moreover, the sequence $(y_{u,n})_{n\in\mathbb{N}}$ given by the iteration (14) satisfies

$$\|y_{u,n} - y_u\| \le \frac{\tilde{c}^n}{(1-c)(1-\tilde{c})}\mathcal{C}_\rho(u),$$

where $y_u \in B_\lambda(y^*)$ is the solution of (3).

Proof. We start by proving (a). Let $y_1, y_2 \in B_\lambda(y^*)$ be given and set $z_i := T_u(y_i) \in B_\varepsilon(y^*)$, $i = 1, 2$. Then, $-A(z_i, u) \in B(z_i - \Phi(y_i, u), u)$ and, thus, the monotonicity of

B yields
$$\langle A(z_2, u) - A(z_1, u), z_1 - \Phi(y_1, u) - z_2 + \Phi(y_2, u)\rangle \ge 0.$$
Consequently, Lemma 20 implies
$$0 \ge C(\|z_2 - z_1\|^2 - \tilde{c}^2\|y_2 - y_1\|^2),$$
i.e., $\|z_2 - z_1\|^2 \le \tilde{c}\|y_2 - y_1\|$.
Now, let $u \in \mathcal{U}$ be chosen as in (b). This enables us to apply Lemma 28 with the choices $r = (1-c)^{-1}\mathcal{C}_\rho(u) \le \varepsilon$ and $\phi = \Phi(y^*, u)$. This shows that $T_u(y^*)$, which is the solution of $0 \in A(z, u) + B(z - \Phi(y^*, u), u)$, satisfies $\|T_u(y^*) - y^*\| \le r \le (1-\tilde{c})\lambda$. Consequently, every $y \in B_\lambda(y^*)$ satisfies
$$\|T_u(y) - y^*\| \le \|T_u(y) - T_u(y^*)\| + \|T_u(y^*) - y^*\| \le \tilde{c}\lambda + (1-\tilde{c})\lambda = \lambda.$$
Thus, we can apply the Banach fixed-point theorem to obtain the existence of $y_u \in B_\lambda(y^*)$. Due to $y_{0,n} = y^*$ and $y_{1,n} = T_u(y^*)$, this also yields the a-priori estimate
$$\|y_{u,n} - y_u\| \le \frac{\tilde{c}^n}{1-\tilde{c}}\|y_{u,1} - y_{u,0}\| \le \frac{\tilde{c}^n}{(1-c)(1-\tilde{c})}\mathcal{C}_\rho(u).$$

The next lemma helps us to control the term $\mathcal{C}_\rho(u^* + th)$.

Lemma 31. Let Assumption 19 be satisfied and fix $h \in U$. Then, for any $\rho > 0$, there exist constants $C \ge 0$ and $t_0 > 0$ such that
$$\mathcal{C}_\rho(u^* + th) \le Ct \qquad \forall t \in [0, t_0].$$

Proof. With $q_\rho^* = y^* - \rho R^{-1}A(y^*, u^*)$ and $\phi^* = \Phi(y^*, u^*)$ we have $y^* - \phi^* = J_{\rho B}(q_\rho^* - \phi^*, u^*)$. Owing to Corollary 14, the directional differentiability of $J_{\rho B}$ at $(q_\rho^* - \phi^*, u^*)$ follows from the directional differentiability of J_B at $(q^* - \phi^*, u^*)$ and this is guaranteed by Assumption 19. Hence, for all terms appearing in the definition (17) of $\mathcal{C}_\rho(u^* + th)$, we can utilize the directional differentiabilities of the involved operators and this yields the desired estimate.

Theorem 32. Let Assumptions 18 and 19 be satisfied. For all $h \in U$ there exists $t_0 > 0$, such that for all $t \in [0, t_0]$, the equation
$$0 \in A(y_t, u + th) + B(y_t - \Phi(y_t, u + th), u + th)$$
has a unique solution $y_t \in B_\lambda(y^*)$, where λ is chosen as in Lemma 29. Moreover, the difference quotient $(y_t - y^*)/t$ converges strongly in Y towards $\delta \in Y$ which is the unique solution of the linearized equation
$$0 \in A'(y^*, u^*; \delta, h) + DB(y^* - \Phi(y^*, u^*), u^* \mid \xi^*)(\delta - \Phi'(y^*, u^*; \delta, h), h), \tag{13}$$
where $\xi^* = -A(y^*, u^*)$.

Proof. We fix $h \in U$. Then, Lemmas 30 and 31 imply the existence of $t_0 > 0$, such that for all $t \in [0, t_0]$ the sequence $(y_{t,n})_{n\in\mathbb{N}}$ defined via $y_{t,0} := y^*$ and each $y_{t,n} \in B_\varepsilon(y^*)$ solves

$$0 \in A(y_{t,n}, u^* + th) + B(y_{t,n} - \Phi(y_{t,n-1}, u^* + th), u^* + th),$$

satisfies

$$\|y_{t,n} - y_t\| \le Ct\tilde{c}^n \tag{18}$$

where $\tilde{c} \in (0, 1)$ is as in Lemma 30 and $C > 0$ is a constant.
Next, we study the differentiability of $y_{t,n}$ w.r.t. $t > 0$. We claim that for all $n \ge 0$, we have the directional differentiabilities

$$\frac{y_{t,n} - y_t}{t} \to \delta_n, \qquad \frac{\Phi(y_{t,n}, u^* + th) - \phi^*}{t} \to \Phi'(y^*, u^*; \delta_n, h), \tag{19}$$

where $\delta_0 = 0$ and for $n \ge 1$ the point $\delta_n \in Y$ solves

$$0 \in A'(y^*, u^*; \delta_n, h) + DB(y^* - \phi^*, u^* \mid \xi^*)(\delta_n - \Phi'(y^*, u^*; \delta_{n-1}, h), h). \tag{20}$$

We argue by induction over n. The base case $n = 0$ is clear since $y_{t,0} = y^* = y_0$. Assume that the assertion holds for $n - 1$. We abbreviate $\phi_{t,n} := \Phi(y_{t,n}, u^* + th)$. Using the operators from (15), we can recast the equation for $y_{t,n}$ as

$$0 \in \mathcal{A}(y_{t,n}, (\phi_{t,n-1}, u^* + th)) + \mathcal{B}(y_{t,n}, (\phi_{t,n-1}, u^* + th)).$$

Next, we apply Lemma 28 for $t > 0$ small enough (depending on n) with $\phi := \phi^* + t\psi_{t,n}$, $\psi_{t,n} = \Phi'(y^*, u^*; \delta_{n-1}, h)$ to obtain a solution $\tilde{y}_{t,n} \in B_\varepsilon(y^*)$ of

$$0 \in A(\tilde{y}_{t,n}, u^* + th) + B(\tilde{y}_{t,n} - \phi^* - t\Phi'(y^*, u^*; \delta_{n-1}, h), u^* + th)$$

or, equivalently,

$$0 \in \mathcal{A}(\tilde{y}_{t,n}, (\phi^* + t\psi_{t,n}, u^* + th)) + \mathcal{B}(\tilde{y}_{t,n}, (\phi^* + t\psi_{t,n}, u^* + th)).$$

Now, we are in position to apply Theorem 12 and this yields $(\tilde{y}_{t,n} - y^*)/t \to \delta_n$ in Y as $t \searrow 0$, where $\delta_n \in Y$ solves

$$0 \in \mathcal{A}'(y^*, (\phi^*, u^*); \delta_n, (\psi_{t,n}, h)) + D\mathcal{B}(y^*, (\phi^*, u^*); \delta_n, (\psi_{t,n}, h)).$$

Using Lemma 9 and (16), we can relate DB and $D\mathcal{B}$ as follows

$$\begin{aligned}
&D\mathcal{B}(y^*, (\phi^*, u^*) \mid \xi^*)(\delta, (\psi, h)) \\
&\qquad = \{R(k - \delta) \mid k \in Y, \delta = J'_{\mathcal{B}}(q^*, (\phi^*, u^*); k, (\psi, h))\} \\
&\qquad = \{R(k - (\delta - \psi)) \mid k \in Y, \delta - \psi = J'_B(q^* - \phi^*, u^*; k - \psi, h)\} \\
&\qquad = DB(y^* - \phi^*, u^* \mid \xi^*)(\delta - \psi, h).
\end{aligned}$$

This results in the equation (20). By using the equations satisfied by $y_{t,n}$ and $\tilde y_{t,n}$, we get

$$\begin{aligned}&\langle A(\tilde y_{t,n}, u^* + th) - A(y_{t,n}, u^* + th), \tilde y_{t,n} - y_{t,n}\rangle \\ &\quad\le \langle A(\tilde y_{t,n}, u^* + th) - A(y_{t,n}, u^* + th), \phi^* + t\Phi'(y^*, u^*; \delta_{n-1}, h) - \Phi(y_{t,n-1}, u^* + th)\rangle.\end{aligned}$$

Consequently,

$$\frac{1}{t}\|\tilde y_{t,n} - y_{t,n}\| \le \frac{L_A}{\mu_A t}\|\phi^* + t\Phi'(y^*, u^*; \delta_{n-1}, h) - \Phi(y_{t,n-1}, u^* + th)\| \to 0 \quad \text{as } t \searrow 0.$$

In combination with $(\tilde y_{t,n} - y^*)/t \to \delta_n$, this yields the first convergence in (19). The second convergence in (19) follows since Φ is directionally differentiable and Lipschitz w.r.t. its first argument. Consequently, (19) holds for all $n \ge 0$.
Thus, we have shown

$$\lim_{n\to\infty} \frac{y_{t,n} - y^*}{t} = \frac{y_t - y^*}{t} \quad \text{and} \quad \lim_{t\searrow 0} \frac{y_{t,n} - y^*}{t} = \delta_n, \tag{21}$$

where both limits exist (strongly) in Y. Moreover,

$$\left\|\frac{y_{t,n} - y^*}{t} - \frac{y_t - y^*}{t}\right\| = \left\|\frac{y_{t,n} - y_t}{t}\right\| \le C\tilde c^n,$$

cf. (18). This shows that the limit $n \to \infty$ in (21) is uniform in $t \in (0, t_0]$. Hence, the classical theorem on the existence and equality of iterated limits ensures

$$\delta := \lim_{t\searrow 0} \frac{y_t - y^*}{t} = \lim_{t\searrow 0}\lim_{n\to\infty} \frac{y_{t,n} - y^*}{t} = \lim_{n\to\infty}\lim_{t\searrow 0} \frac{y_{t,n} - y^*}{t} = \lim_{n\to\infty} \delta_n \quad \text{in } Y.$$

Finally, passing to the limit $n \to \infty$ in (20) yields the equation for δ.

Note that Assumption 18 is only used in Lemma 30. It can be replaced by requiring that $T_u \colon B_\lambda(y^*) \to B_\varepsilon(y^*)$ is a contraction uniformly in $u \in \mathcal{U}$.

It is quite interesting to see that the approaches from Sections 4.1 and 4.2 use the same assumptions and, actually, Theorems 27 and 32 coincide. If we look a little bit more carefully, we see that in Section 4.1, Lemma 20 is only applied in the special case $y_i = z_i$, $i = 1, 2$, see Lemmas 21 and 24, whereas Section 4.2 requires the application in the general case $y_i \ne z_i$, $i = 1, 2$, see Lemma 30.

Altogether, it seems to be possible to craft special situations in which only one of the approaches of Sections 4.1 and 4.2 is applicable. However, we think that (up to some exceptional boundary cases) the range of applicability of both approaches coincide.

5 Applications

5.1 Optimization with a parameter-dependent sparsity functional

As a first application, we consider the minimization problem

$$\text{Minimize} \quad F(y) + G(y, u) \quad \text{w.r.t. } y \in Y \tag{$P(u)$}$$

with a parameter $u \in U$. Here, (Ω, Σ, μ) is a measure space and $Y = U = L^2(\mu)$. Moreover, $F\colon Y \to \mathbb{R}$ is a given functional and $G\colon Y \times U \to \mathbb{R}$ is defined via

$$G(y, u) := \int_\Omega |uy| \, \mathrm{d}\mu.$$

Thus, $(P(u))$ models, e.g., optimal control problems which include a sparsity functional and we are interested in the sensitivity of the solution y w.r.t. the distributed sparsity parameter u.

Suppose that $u^* \in U$ is fixed and $y^* \in Y$ is a local minimizer of $(P(u))$. We assume that F is Fréchet differentiable in $B_\varepsilon(y^*)$ for some $\varepsilon > 0$, such that its derivative F' is (uniformly) strongly monotone and Lipschitz continuous on $B_\varepsilon(y^*)$, see Assumption 3(i).

As it is usually done, we identify the dual space of $Y = L^2(\mu)$ with itself.

Due to the convexity of F, a point $y \in Y$ with $\|y - y^*\| < \varepsilon$ is a local minimizer of $(P(u))$ with $u \in U$ if and only if

$$0 \in F'(y) + \partial_y G(y, u),$$

where

$$\partial_y G(y, u) = \{g \in L^2(\Omega) \mid G(v, u) \geq G(y, u) + \int_\Omega g(v - y) \, \mathrm{d}\mu \quad \forall v \in Y\}$$

is the subdifferential of G w.r.t. y. It is clear that Assumption 3 is satisfied with the setting

$$A(y, u) := F'(y), \qquad B(y, u) := \partial_y G(y, u).$$

In order to apply the results from Section 3, we have to study the properties of the resolvent $J_{\rho B}$, $\rho > 0$. It is clear that

$$J_{\rho B}(q, u) = \operatorname{prox}_{\rho G(\cdot, u)}(q) = \operatorname*{arg\,min}_{v \in Y} \int_\Omega \frac{1}{2}(v - q)^2 + \rho |uv| \, \mathrm{d}\mu.$$

Now, a pointwise discussion shows that the resolvent can be computed pointwise and is given by a soft-shrinkage with parameter $\rho|u|$, i.e.,

$$J_{\rho B}(q, u)(x) = \operatorname{shrink}_{\rho|u(x)|}(q(x)) := \max(|q(x)| - \rho|u(x)|, 0) \operatorname{sign}(q(x)).$$

Now, since

$$\mathbb{R}^2 \ni (q, u) \mapsto \operatorname{shrink}_{\rho|u|}(q) = \max(|q| - \rho|u|, 0) \operatorname{sign}(q) \in \mathbb{R}$$

is Lipschitz continuous and directionally differentiable, it is easy to check that also the associated Nemytskii operator $J_{\rho B}\colon Y \times U \to Y$ is Lipschitz continuous and directionally differentiable. If, additionally, $F'\colon Y \to Y$ is directionally differentiable, we are in position to apply Theorem 12 to obtain the directional differentiability of the (local) solution mapping of $(P(u))$. Using Lemma 9, it is also possible to characterize the directional derivative.

5.2 Quasi-linear QVIs

We demonstrate the applicability of our results to a QVI governed by a quasi-linear operator. To this end, let $\Omega \subset \mathbb{R}^d$ be open and bounded and with $Y = H_0^1(\Omega)$ and $U = L^2(\Omega)$ we define the quasi-linear operator $A\colon Y \times U \to Y^\star$ via

$$A(y, u) := -\operatorname{div} g(\nabla y, u) + f(u), \quad \text{i.e.} \quad \langle A(y, u), v\rangle = \int_\Omega g(\nabla y, u)\nabla v + f(u)v \, \mathrm{d}x,$$

where $g\colon \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$ is (uniformly) strongly monotone w.r.t. its first argument; and Lipschitz continuous and differentiable on $\mathbb{R}^d \times \mathbb{R}$. Moreover, $f\colon \mathbb{R} \to \mathbb{R}$ is differentiable and Lipschitz continuous. These conditions imply that A satisfies Assumption 3 globally on Y. Moreover, using the dominated convergence theorem, we can check that A is directionally differentiable with

$$\langle A'(y, u; \delta, h), v\rangle = \int_\Omega (g_y'(\nabla y, u)\nabla\delta + g_u'(\nabla y, u)h)\nabla v + f'(u)hv \, \mathrm{d}x,$$

see also Goldberg, Kampowsky, Tröltzsch, 1992, Theorem 8.

To define the operator B, let $K \subset Y$ be given such that Proj_K is directionally differentiable, e.g., we could choose a polyhedric K. We set $B(\cdot) := N_K(\cdot)$, where N_K is the normal cone mapping of K. Note that B is independent of the variable u.

With this setting, Assumptions 3 and 8 are satisfied. Next, we choose $\Phi\colon Y \times U \to Y$ such that there exists a Lipschitz constant $L_\Phi \in [0, 1)$ with

$$\|\Phi(y_2, u) - \Phi(y_1, u)\| \le L_\Phi \|y_2 - y_1\|$$

for all $y_1, y_2 \in Y$ and all $u \in U$. Further, we suppose that Φ is directionally differentiable and that Assumption 18 concerning the smallness of L_Φ is satisfied.

Since our assumptions on A, B and Φ are global, we obtain that for all $u \in U$, there exists a unique solution $y \in Y$ of the QVI

$$0 \in A(y, u) + B(y - \Phi(y, u)),$$

cf. Wachsmuth, 2020, Section 3. Since all the assumptions from Section 4 are satisfied, our differentiability theorems imply that the mapping $y \mapsto u$ is directionally differentiable.

For a fixed parameter $u^* \in U$ we denote the solution by $y^* \in Y$. Then, The directional derivative $\delta \in Y$ in the direction $h \in U$ is given by the solution of

$$0 \in A'(y^*, u^*; \delta, h) + DB(y^* - \Phi(y^*, u^*) \mid \xi^*)(\delta - \Phi'(y^*, u^*; \delta, h)),$$

where $\xi^* = -A(y^*, u^*)$. We mention that in the particular case that the set K is polyhedric, the set-valued mapping $DB(y^* - \Phi(y^*, u^*) \mid \xi^*)$ coincides with the normal cone mapping of the critical cone $\mathcal{K} = T_K(y^*) \cap (\xi^*)^\perp$, see Proposition 11.

It is also clear that the above assumptions and arguments can be localized if we already have a solution y^* of the QVI corresponding to the parameter u^*.

References

Adly, Samir; Rockafellar, Ralph Tyrrell (2020). “Sensitivity analysis of maximally monotone inclusions via the proto-differentiability of the resolvent operator”. *Mathematical Programming*. DOI: 10.1007/s10107-020-01515-z.

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.

Alphonse, Amal; Hintermüller, Michael; Rautenberg, Carlos N. (2021). *Optimal control and directional differentiability for elliptic quasi-variational inequalities*. arXiv: 2008.13586v2.

Attouch, Hedy (1984). *Variational Convergence for Functions and Operators*. Applicable Mathematics Series. Pitman Advanced Pub. Program.

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

Brezis, Haim (2011). *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York. DOI: 10.1007/978-0-387-70914-7.

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

Christof, Constantin; Wachsmuth, Gerd (2020). “Differential Sensitivity Analysis of Variational Inequalities with Locally Lipschitz Continuous Solution Operators”. *Applied Mathematics & Optimization* 81.1, pp. 23–62. DOI: 10.1007/s00245-018-09553-y.

Christof, Constantin; Wachsmuth, Gerd (2021). *Lipschitz Stability and Hadamard Directional Differentiability for Elliptic and Parabolic Obstacle-Type Quasi-Variational Inequalities*. arXiv: 2105.05895.

Do, Chi Ngoc (1992). “Generalized second-order derivatives of convex functions in reflexive Banach spaces”. *Transactions of the American Mathematical Society* 334.1, pp. 281–301. DOI: 10.2307/2153983.

Goldberg, Hyman; Kampowsky, Winfried; Tröltzsch, Fredi (1992). “On Nemytskij operators in L_p-spaces of abstract functions”. *Mathematische Nachrichten* 155, pp. 127–140. DOI: 10.1002/mana.19921550110.

Haraux, Alain (1977). “How to differentiate the projection on a convex set in Hilbert space. Some applications to variational inequalities”. *Journal of the Mathematical Society of Japan* 29.4, pp. 615–631. DOI: 10.2969/jmsj/02940615.

Levy, Adam B. (1999). “Sensitivity of Solutions to Variational Inequalities on Banach Spaces”. *SIAM Journal on Control and Optimization* 38.1, pp. 50–60. DOI: 10.1137/s036301299833985x.

Levy, Adam B.; Rockafellar, Ralph Tyrrell (1994). “Sensitivity analysis of solutions to generalized equations”. *Transactions of the American Mathematical Society* 345.2, pp. 661–671. DOI: 10.1090/s0002-9947-1994-1260203-5.

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.

Pérez-Aros, Pedro; Vilches, Emilio (2019). “An Enhanced Baillon extendashHaddad Theorem for Convex Functions Defined on Convex Sets”. *Applied Mathematics & Optimization* 83.3, pp. 2241–2252. DOI: 10.1007/s00245-019-09626-6.

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.

Wachsmuth, Gerd (2020). “Elliptic quasi-variational inequalities under a smallness assumption: Uniqueness, differential stability and optimal control”. *Calculus of Variations and Partial Differential Equations* 59.2. DOI: 10.1007/s00526-020-01743-3.

Wachsmuth, Gerd (2021). *Maximal monotone operators with non-maximal graphical limit*. arXiv: 2107.10055.