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

Preprint ISSN 2198-5855

On the convexity of optimal control problems involving non-linear PDEs or VIs and applications to Nash games (changed title: Vector-valued convexity of solution operators with application to optimal control problems)

Michael Hintermüller[1,2], Steven-Marian Stengl[1,2]

submitted: September 11, 2020 (revision: November 17, 2021)

[1] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: michael.hintermueller@wias-berlin.de
steven-marian.stengl@wias-berlin.de

[2] Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin
Germany
E-Mail: hint@math.hu-berlin.de
stengl@math.hu-berlin.de

No. 2759
Berlin 2020

2010 *Mathematics Subject Classification.* 06B99, 49K21, 47H04, 49J40.

Key words and phrases. PDE-constrained optimization, K-Convexity, set-valued analysis, subdifferential, semilinear elliptic PDEs, variational inequality.

This research was supported by the DFG under the grant HI 1466/10-1 associated to the project “Generalized Nash Equilibrium Problems with Partial Differential Operators: Theory, Algorithms, and Risk Aversion” within the priority programme SPP1962 “Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization”.

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

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

On the convexity of optimal control problems involving non-linear PDEs or VIs and applications to Nash games (changed title: Vector-valued convexity of solution operators with application to optimal control problems)

Michael Hintermüller, Steven-Marian Stengl

Abstract

In the present article, generalized convexity induced by a preorder-relation is investigated for solution operators of (generalized) equations. After the establishment of sufficient conditions, a generalized subdifferential is established. Along with mild conditions on the objective to ensure convexity of the optimization problem this concept can be used for the derivation of first-order optimality conditions. Throughout, examples illustrating the theoretical findings are given.

1 Introduction

Many optimization problems incorporating models from economics, engineering, or physics are formulated as mathematical programs with equilibrium constraints (abbr.: MPECs); see, e.g., [OKZ98, LLPR96] for a selection of examples in finite dimensions and [BdP84] for infinite dimensions. Often, the optimization problem is given in the form of an *optimal control problem*, distinguishing two classes of variables: controls and states, respectively. Both variable types are typically linked via, e.g., a system of partial differential equations (abbr.: PDEs) or variational inequalities (abbr.: VIs) appearing in the constraint set of the optimal control problem. For instances of such problems see the selected works [BdP84, TS10, LEG$^+$12, HK11, HK09, NST07] and the many references therein.
From an analytical and also numerical standpoint convexity of optimization problems is a very appealing property. This is in particular true as convex problems are well studied in the literature (see e.g. the monographs [ET76, BZ06]) and stand out by a number of desirable properties. These include that all local solutions are also global solutions, generalized differentiation concepts are available, there is a deep interlinkage between analytical and geometrical notions, necessary conditions are also sufficient, and the set of minimizers is convex – to name only a few examples. Returning to the class of optimal control problems alluded to above, if, for example, the constraint set links states and controls via a linear operator equation, then the optimal control problem or its control reduced form enjoy convexity properties, if the objective is convex in controls and states. This, however, does not need to be the case when the linear operator equation is replaced by non-linear PDEs or VIs.
Nevertheless, it is a priori not clear that the non-linearity in the constraints breaks the convexity of the associated optimization problem and thus the present work takes the last observation as a starting point for investigation. One of our central goals is to derive structural conditions to ensure convexity of optimal control problems in a reduced formulation, where the state is considered as dependent on the control. For this sake we utilize a preorder based vector-valued convexity notion imposed on the solution operator, the so-called control-to-state mapping (cf. [CLV13, Chapter 19] and [BS00, Section 2.3.5]). Exploiting its properties, we extend the theory of convex optimization to vector-valued operators with

regard to subdifferentials and also investigate calculus rules for the derivation of first-order optimality systems. The latter is of independent interest in *non-smooth optimization*, in general. Especially in view of MPECs involving VI constraints or bilevel optimization problems, differentiability of the control-to-state-map is often not guaranteed. Hence, novel techniques need to be developed to derive necessary optimality conditions.

The rest of this paper is organized as follows. In section 2 we introduce notation and preliminaries used in the rest of the work. In section 3 we introduce vector-valued convex operators. After investigating their properties, sufficient criteria for the K-convexity of solution operators of (generalized) equations are identified. Here, K stands for a non-empty closed convex cone. Such generalized equations will later appear as constraints in the optimization problems under investigation. Our results of this section are then applied to examples that are linked to optimal control problems discussed in the literature. In section 4 we draw our attention to the development of a subdifferential concept and its application to the aforementioned examples. In section 5 we use our combined results and ensure the convexity of a class of optimization problems containing a composition in the objective. Moreover, we derive calculus rules for the subdifferential of the (control reduced) objective aiming at the derivation of first-order optimality resp. stationarity conditions for the associated optimization problem. The work ends by an application of our combined findings to a problem on doping optimization in semiconductor physics.

2 Notation and preliminaries

In the following, let $(X, \|\cdot\|)$ be a Banach space and let X^* denote its *topological dual space*. The associated dual pairing $\langle\,\cdot\,,\,\cdot\,\rangle_{X^*,X} : X^* \times X \to \mathbb{R}$ is defined by $\langle x^*, x\rangle_{X^*,X} := x^*(x)$. Oftentimes, we simply denote $\langle\,\cdot\,,\,\cdot\,\rangle$, if the corresponding spaces are clear from the context. Two elements $x^* \in X^*$ and $x \in X$ are called *orthogonal*, if $\langle x^*, x\rangle = 0$ and we write $x^* \perp x$ or $x \perp x^*$. The *annihilator* of a subset $M \subseteq X$ is defined as

$$M^\perp = \{x^* \in X^* : \langle x^*, x\rangle = 0 \text{ for all } x \in M\}\,,$$

and analogously for a set $M^* \subseteq X^*$ as

$$M^{*\perp} = \{x \in X : \langle x^*, x\rangle = 0 \text{ for all } x^* \in M^*\}\,.$$

For a single element we may write $x^\perp := \{x\}^\perp$.

The closed unit ball of X is denoted by $\mathbb{B}_X := \{x \in X : \|x\| \leq 1\}$. The interior of a set $M \subseteq X$ is defined by

$$\operatorname{int}(M) := \{x \in M : \text{ there exists } \varepsilon > 0 : x + \varepsilon\mathbb{B}_X \subseteq M\}\,.$$

and its closure as

$$\operatorname{cl}(M) := \{x \in M : \text{ there exists } (x_n)_{n\in\mathbb{N}} \subseteq M \text{ with } x_n \to x\}\,.$$

A subset $C \subseteq X$ is called *convex*, if for all $t \in (0,1)$ and $x_0, x_1 \in C$ it holds that $tx_1 + (1-t)x_0 \in C$. A set $K \subseteq X$ is called a *cone*, if for all $t \in \mathbb{R}$, $t \geq 0$ and $x \in K$ also $tx \in K$ holds. The *tangential cone* of C in $x \in C$ is defined as

$$T_C(x) := \{d \in X : \text{ there exist } t_k \searrow 0, d_k \to d \text{ with } x + t_k d_k \in C\ \forall k \in N\}$$

and the normal cone is defined as

$$N_C(x) := \{x^* \in X^* : \langle x^*, x' - x\rangle_{X^*,X} \leq 0 \text{ for all } x' \in C\}\,.$$

The *core* (or *algebraic interior*) of a set $M \subseteq X$ is defined by

$$\operatorname{core}(M) := \{x \in M : \forall d \in X\ \exists\, \bar{t} > 0 : x + td \in M \text{ for all } |t| < \bar{t}\}\,.$$

We define the *indicator functional* $I_M : X \to \mathbb{R} \cup \{+\infty\}$ by

$$I_M(x) := \begin{cases} 0, & \text{if } x \in M, \\ +\infty, & \text{else.} \end{cases}$$

A functional $f : X \to \mathbb{R} \cup \{+\infty\}$ is called *proper*, if there exists an argument with finite value. It is called *lower semi-continuous*, if for all sequences $x_n \to x$ it holds that

$$f(x) \leq \liminf_{n \to \infty} f(x_n)$$

and it is called *weakly lower semi-continuous*, if the above holds even for weakly convergent $x_n \rightharpoonup x$. The functional is called *convex*, if

$$f(tx_1 + (1-t)x_0) \leq tf(x_1) + (1-t)(x_0) \text{ holds for all } t \in (0,1), x_0, x_1 \in X.$$

The subdifferential of f in $x \in X$ denoted $\partial f(x)$ is defined as the set of all directions $x^* \in X^*$ such that

$$f(x') \geq f(x) + \langle x^*, x' - x\rangle_{X^*,X} \text{ holds for all } x' \in X.$$

A subset $M \subseteq X$ is called *absorbing*, if for all $x \in X$ there exists $r > 0$ such that for all $|t| \leq r$ one has $tx \in M$.
A convex subset $C \subseteq X$ is called *cs-closed* (convex series closed) if for every sequence $(t_i)_{i\in\mathbb{N}}$ of non-negative numbers with $\sum_{i=1}^{\infty} t_i = 1$ and sequence $(x_i)_{i\in\mathbb{N}} \subseteq C$ such that $x := \sum_{i=1}^{\infty} t_i x_i$ exists, the inclusion $x \in C$ follows. Moreover, C is called *cs-compact* (convex series compact) if for all sequences $(t_i)_{i\in\mathbb{N}}$ of non-negative numbers with $\sum_{i=1}^{\infty} t_i = 1$ and an arbitrary sequence $(x_i)_{i\in\mathbb{N}} \subseteq C$ the limit $x := \sum_{i=1}^{\infty} t_i x_i$ exists and $x \in C$ holds.
A function $F : X \to \mathcal{P}(Y)$ is called a *set-valued operator* or *correspondence* and is denoted by $F : X \rightrightarrows Y$. Its *graph* is defined by

$$\operatorname{gph}(F) := \{(x, y) \in X \times Y : y \in F(x)\}$$

and its (effective) domain by

$$\mathcal{D}(F) := \{x \in X : F(x) \neq \emptyset\}\,.$$

In the scope of this work we make use of some terminology and aspects of order theory, which will be introduced next. For further references as well as details and additional information, the interested reader is referred to the monographs [Sch74, Bec08].
We equip X with a preorder relation (an order relation without antisymmetry) $\leq$. Then, X is called a preordered Banach space, if the preorder is compatible with the relation, i.e. for all $z \in X$ and $t \geq 0$ the implications $x_0 \leq x_1 \Rightarrow x_0 + z \leq x_1 + z$ and $x_0 \leq x_1 \Rightarrow tx_0 \leq tx_1$ hold true.
Let a subset $M \subseteq X$ be given. The infimum of M is an element $x \in X$ such that $x \leq x'$ for all $x' \in M$ and for every $y \in X$ with $y \leq x'$ for all $x' \in M$ one infers $y \leq x$. The supremum is defined analogously.
An ordered Banach space is called a *vector lattice*, if for two elements x_0, x_1 the *infimum* $\min(x_0, x_1) = x_0 \wedge x_1 := \inf\{x_0, x_1\}$ as well as the *supremum* $\max(x_0, x_1) = x_0 \vee x_1 := \sup\{x_0, x_1\}$ exist, respectively. For $x \in X$ we also abbreviate $x^+ := \max(x, 0)$.
Let $K := \{x \in X : x \geq 0\}$. If $(X, \leq)$ is a preordered vector space, then K is a non-empty, convex cone. On the other hand, let a vector space X and a non-empty, convex cone $K \subseteq X$ be given. Then,

K induces a preorder relation $\leq_K$ for all $x_0, x_1 \in X$ by $x_0 \leq_K x_1$, if $x_1 - x_0 \in K$. By definition $(X, \leq_K)$ is a preordered vector space and $\leq_K$ induces an order, if and only if $K \cap (-K) = \{0\}$. In this sense, it is possible to characterize the order equivalently by the cone of non-negative elements.
Let $d \in \mathbb{N}\backslash\{0\}$ and let $\Omega \subseteq \mathbb{R}^d$ be a bounded, open domain. Associated to this domain we denote the *Borel algebra* $\mathcal{B}(\Omega)$ as the smallest σ-algebra generated by the system of open subsets of Ω. The Lebesgue measure on the Borel-algebra is denoted by $\lambda^d : \mathcal{B}(\Omega) \to [0, \infty]$. For a set $E \in \mathcal{B}(\Omega)$ the *characteristic function* of E is given by

$$\mathbb{1}_E(x) := \begin{cases} 1, & x \in E, \\ 0, & \text{else.} \end{cases}$$

For $p \in [1, \infty)$ denote the Lebesgue space as

$$L^p(\Omega) := \left\{ u : \Omega \to \mathbb{R} \text{ measurable } : \int_\Omega |u|^p \mathrm{d}x < +\infty \right\}$$

with its elements only identified up to null sets, i.e. sets of Lebesgue measure zero. This space equipped with the norm $\|u\|_{L^p(\Omega)} := \left(\int_\Omega |u|^p \mathrm{d}x\right)^{\frac{1}{p}}$ is a Banach space for all $p \in [1, \infty)$ and a reflexive Banach space for $p \in (1, \infty)$. The *Sobolev spaces* $W^{1,p}(\Omega)$ are defined as

$$W^{1,p}(\Omega) := \left\{ u \in L^p(\Omega) : \nabla u \in L^p(\Omega; \mathbb{R}^d) \right\},$$

where ∇u denotes the distributional derivative of u. Equipped with the norm $\|u\|_{W^{1,p}(\Omega)} := \Big(\|u\|^p_{L^p(\Omega)} + \sum_{i=1}^d \|\partial_i u\|^p$ the space $W^{1,p}(\Omega)$ is a Banach space, and for $p \in (1, \infty)$ a reflexive Banach space. For $p = 2$ one also denotes $H^1(\Omega) = W^{1,2}(\Omega)$. For further details on Sobolev spaces the interested reader is referred to [AF03].

3 Vector-valued convex operators

Targeting solution maps of (generalized) equations, we introduce and study convexity for vector-valued operators in this section. This concept will later be utilized for PDEs and VIs. We start with basic facts on the polar, respectively dual cone associated with a subset $M \subseteq X$.
Let $M \subseteq X$. The *polar cone of* M is defined as

$$M^\circ := \{x^* \in X^* : \langle x^*, x \rangle \leq 0 \text{ for all } x \in M\}.$$

and the *dual cone* M^+ is defined as

$$M^+ := -M^\circ = \{x^* \in X^* : \langle x^*, x \rangle \geq 0 \text{ for all } x \in M\}. \tag{1}$$

Using the notation in (1) one observes for a closed, convex set C the relation $N_C(x) := (C - x)^\circ$ (cf. [Sch07, Definition 11.2.1 and Lemma 11.2.2]). Next, we establish some calculus rules for the dual cone. For the statements as well as their proofs we refer to [RW09, Corollary 11.25] (in finite dimensions) as well as to [BS00, Proposition 2.40]. To remain self-contained, we provide short proofs in the appendix.

Lemma 1. *Let X be a topological vector space and let the subsets $M, M_1, M_2 \subseteq X$ be given. Then the following assertions hold true.*

(i) If $M_1 \subseteq M_2$, then $M_2^+ \subseteq M_1^+$.

(ii) $M^+ = (\mathrm{cl}\,(M))^+$.

(iii) If M is a non-empty, closed, convex cone, then $M^{++} = M$.

(iv) If $0 \in M_1 \cap M_2$, then we have

$$(M_1 + M_2)^+ = M_1^+ \cap M_2^+.$$

(v) Let M_j be closed, convex cones, then it holds that

$$(M_1 \cap M_2)^+ = \mathrm{cl}\left(M_1^+ + M_2^+\right).$$

Next, we introduce convexity in a vector-valued setting, see [BS00]. Since the notion therein is related to a preorder cone and we work with different cones in different spaces simultanously, we use the term K-convexity to keep track of the cone involved.

Definition 2. Consider Banach spaces U and Y. Let a non-empty closed, convex cone $K \subseteq Y$ inducing a preorder relation $\geq_K$ on Y be given. A set-valued mapping $A : U \rightrightarrows Y$ is called *K-convex*, if for all $t \in (0,1)$ and $u_0, u_1 \in U$ the relation

$$tA(u_1) + (1-t)A(u_0) \subseteq A(tu_1 + (1-t)u_0) + K$$

holds true, and A is called K-concave, if it is ($-K$)-convex, i.e., for all $u_0, u_1 \in U$ and $t \in (0,1)$ holds

$$tA(u_1) + (1-t)A(u_0) \subseteq A(tu_1 + (1-t)u_0) - K.$$

As mentioned, we use the term K-convexity over just convexity to keep track of the involved cone and is not to be confused with the one in [Sca59].
We note here that $A : U \rightrightarrows Y$ is K-convex, if and only if the epigraph $\mathrm{epi}_K\,(A)$ of A with respect to K, defined by

$$\mathrm{epi}_K\,(A) := \{(x,y) : y \in A(x) + K\},$$

is a convex subset of $U \times Y$. By defining the set-valued mapping $A_K : U \rightrightarrows Y$ via $A_K(x) := A(x) + K$ one can rewrite

$$\mathrm{epi}_K\,(A) = \mathrm{gph}(A_K).$$

A special instance is the case of a single-valued operator $S : U \to Y$. Then, the K-convexity reads

$$S(tu_1 + (1-t)u_0) \leq_K tS(u_1) + (1-t)S(u_0)$$

for all $u_0, u_1 \in X$ and $t \in (0,1)$. It is noteworthy that for a convex set $\overline{C} \subseteq Y$ with $\overline{C} - K \subseteq \overline{C}$ its preimage under S is convex. To see this, take $u_0, u_1 \in S^{-1}(\overline{C})$ and $t \in (0,1)$. By the K-convexity of S we obtain

$$S(tu_1 + (1-t)u_0) \in tS(u_1) + (1-t)S(u_0) - K \in \overline{C} - K \subseteq \overline{C},$$

and thus $tu_1 + (1-t)u_0 \in S^{-1}(\overline{C})$. In this setting we can establish the following characterization.

Lemma 3. *Let U, Y be Banach spaces and let $K \subseteq Y$ be a non-empty, closed, convex cone inducing the preorder relation $\leq_K$ on Y. Consider an operator $S : U \to Y$. If S is (twice) Fréchet-differentiable, let DS (D^2S) denote its (second) Fréchet-derivative. Then, the following statements are equivalent:*

(i) S *is* K*-convex.*

(ii) *For all* $y^* \in K^+$ *the functional* $u \mapsto \langle y^*, S(u)\rangle$ *is convex.*

(iii) *If* S *is continuously Fréchet-differentiable, then for all* $u_1, u_0 \in U$ *it holds that*

$$DS(u_0)(u_1 - u_0) + S(u_0) \leq_K S(u_1).$$

(iv) *If* S *is continuously differentiable, then for all* $u_1, u_0 \in U$ *it holds that*

$$(DS(u_1) - DS(u_0))(u_1 - u_0) \geq_K 0.$$

(v) *If, moreover,* S *is twice continuously differentiable, then for all* $u \in U$ *and* $d \in U$ *it holds that*

$$D^2S(u)(d, d) \geq_K 0.$$

Proof. Consider $K^+ = \{y^* \in Y^* : \langle y^*, y\rangle \geq 0 \text{ for all } y \in K\}$. By *(iii)* in Lemma 1 it holds that $y \in K$, if and only if $y \in K^{++}$. The latter is equivalent to

$$\langle y^*, y\rangle \geq 0 \text{ for all } y^* \in K^+.$$

Hence, S is K-convex if and only if for all $y^* \in K^+$ the functionals $u \mapsto \langle y^*, S(u)\rangle$ are convex, which proves the equivalence $(i) \Leftrightarrow (ii)$.
For the C^1- and C^2-case we can hence utilize the characterization of convex functionals and obtain the equivalence of the remaining statements. □

3.1 Vector-valued convexity of solution operators of inverse problems

In this subsection we investigate conditions guaranteeing the solution operator of a generalized equation to be K-convex. This problem class covers a variety of problems including PDEs and VIs. For this sake consider a set-valued operator $A : Y \rightrightarrows W$ and an arbitrary $w \in W$. We are interested in the following *generalized equation*:
Seek $y \in Y$ such that

$$w \in A(y). \tag{GE}$$

In the following theorem we derive conditions on the operator A that guarantee the convexity of the solution mapping associated with (GE).

Theorem 4. *Let* Y, W *be Banach spaces both equipped with non-empty closed, convex cones* $K \subseteq Y$ *and* $K_W \subseteq W$*, respectively. Let* $A : Y \rightrightarrows W$ *be a set-valued operator and assume:*

(i) A *is* K_W*-concave.*

(ii) *The mapping* $A^{-1} : W \rightrightarrows Y$ *is single-valued, its effective domain is* W *and it is* K_W*-*K*-isotone, i.e. for* $w_1, w_0 \in W$ *with* $w_1 \geq_{K_W} w_0$ *it holds that* $A^{-1}(w_1) \geq_K A^{-1}(w_0)$.

Then, the mapping $A^{-1} : W \rightrightarrows Y$ *is* K*-convex.*

Proof. Let $t \in (0,1)$ and $w_0, w_1 \in W$. We denote by $y_j \in Y$ the unique solution of $w_j \in A(y_j)$ for $j = 0,1$. Let $y \in Y$ be the solution of $tw_1 + (1-t)w_0 \in A(y)$. Then we obtain by the assumed K_W-concavity

$$tw_1 + (1-t)w_0 \in tA(y_1) + (1-t)A(y_0) \subseteq A(ty_1 + (1-t)y_0) - K_W.$$

Hence, there exists $k_W \in K_W$ with $tw_1+(1-t)w_0+k_W \in A(ty_1+(1-t)y_0)$ and by the assumed isotonicity of the inverse A^{-1} we obtain

$$y = A^{-1}(tw_1 + (1-t)w_0) \leq_K A^{-1}(tw_1 + (1-t)w_0 + k_W) = ty_1 + (1-t)y_0,$$

which proves the K-convexity of A^{-1}. □

After establishing our core result for this subsection, we derive a variation of Theorem 4 aiming at formally more complicated generalized equations involving two components with different roles being assigned to them.
Given $u \in U, w \in W$, find $y \in Y$ such that

$$w \in A(u,y). \tag{2}$$

Corollary 5. *Consider the Banach spaces U, Y and W, the latter two equipped with non-empty, closed, convex cones $K \subseteq Y$ and $K_W \subseteq W$, respectively. Let $A : U \times Y \rightrightarrows W$ be a set-valued operator and assume:*

(i) The mapping A is K_W-concave.

(ii) For every fixed $u \in U$, the mapping $A(u,\cdot)^{-1} : W \rightrightarrows Y$ is single-valued, K_W-K-isotone and its domain is W.

Then the solution mapping $S : W \times U \to Y$ of (2) *is K-convex.*

Proof. In order to apply Theorem 4, we define $\bar{A} : U \times Y \rightrightarrows W \times U$ by

$$\bar{A}(v,y) := A(v,y) \times \{v\}$$

and equip the product spaces with the non-empty, closed, convex cones

$$\bar{K} := \{0\} \times K \subseteq U \times Y \text{ and } \bar{K}_W := K_W \times \{0\} \subseteq W \times U.$$

We check the conditions of Theorem 4:
The $\bar{K}_W$-concavity is immediately clear from the definition of $\bar{A}$. Considering the inverse, we see $\bar{A}^{-1}(w,u) = (u, A^{-1}(u,\cdot)(w))$ and obtain $(w_1,u_1) \geq_{\bar{K}_W} (w_2,u_2)$ if and only if $u_1 = u_2 =: u$ and $w_1 \geq_{K_W} w_2$. By our assumption it holds that $A^{-1}(u,\cdot)(w_1) \geq_K A^{-1}(u,\cdot)(w_2)$ and we deduce $\bar{A}^{-1}(w_1,u_1) \geq_{\bar{K}} \bar{A}^{-1}(w_2,u_2)$, which proves the isotonicity and by Theorem 4, the $\bar{K}$-convexity. Hence, we see that $(w,u) \mapsto (u, S(w,u)) = \bar{A}^{-1}(w,u)$ is $\bar{K}$-convex, which is equivalent to S being K-convex. □

The following corollary addresses a yet more specific case of (GE).

Corollary 6. *Consider the Banach spaces U, Y and W, the latter two equipped with the non-empty, closed, convex cones K and K_W, respectively. Let $A : Y \to W$ be invertible and K_W-concave and $B : U \to W$ be K_W-convex. Assume that A, B are Fréchet-differentiable and the operator $DA(y) \in \mathcal{L}(Y, W)$ has a K-K_W-isotone inverse. Then, the solution mapping $S : U \to Y$ of the equation*

$$A(y) = B(u)$$

is K-convex.

Proof. Consider the mapping $\widetilde{A}(u, y) := A(y) - B(u)$. Evidently, the mapping is K_W-concave and $\widetilde{A}(u, \cdot)^{-1}$ is a singleton and defined on all of W. Moreover, it is K_W-K-isotone. Let therefore $w_1 \geq_{K_W} w_0$, and we see, writing $A(y_t) = tw_1 + (1-t)w_0$, that

$$\begin{aligned} A^{-1}(w_1) - A^{-1}(w_0) &= \int_0^1 D(A^{-1})(tw_1 + (1-t)w_0)(w_1 - w_0)\mathrm{d}t \\ &= \int_0^1 DA(y_t)^{-1}(w_1 - w_0)\mathrm{d}t \geq_K 0. \end{aligned}$$

The assertion follows by Corollary 5. □

Next, we illustrate these results by several practically relevant examples.

3.2 Applications

In the upcoming subsection we apply our results to a selection of examples. Precisely, we show the K-convexity of the solution operators associated to a partial differential equation (abbr.: PDE) and a variational inequality (abbr.: VI). These problem classes enjoy an active research interest also with respect to optimal control problems as an underlying constraint in the context of MPECs (mathematical programming with equilibrium constraints, see e.g. [OKZ98, LLPR96]). We will revisit these examples in Section 4.1 and Section 4.2, respectively when we calculate their (generalized) subdifferentials. Now, we show the K-convexity of the solution operators.

3.2.1 Application to semilinear elliptic PDEs

As first application we propose conditions to a type of semilinear elliptic PDEs with homogeneous Dirichlet boundary condition. For an open, bounded domain $\Omega \subseteq \mathbb{R}^d$ with Lipschitz boundary and $d \in \mathbb{N}\backslash\{0\}$, take $Y := H_0^1(\Omega)$ and consider the following partial differential equation:
Given $w \in H^{-1}(\Omega)$, seek $y \in H_0^1(\Omega)$ such that

$$\begin{aligned} -\Delta y + \Phi(y) &= w \text{ in } \Omega, \\ y &= 0 \text{ on } \partial\Omega, \end{aligned} \tag{PDE}$$

where $\Phi : \mathbb{R} \to \mathbb{R}$ is a continuous, non-decreasing and concave function inducing a continuous superposition operator $\Phi : L^2(\Omega) \to L^2(\Omega)$. Take as operator associated to (PDE) $A : H_0^1(\Omega) \to H^{-1}(\Omega)$ defined by $\langle A(y), v\rangle = (\nabla y, \nabla v)_{L^2(\Omega;\mathbb{R}^d)} + (\Phi(y), v)_{L^2(\Omega)}$ for arbitrary $v \in H_0^1(\Omega)$.
Then, the associated solution operator $S : H^{-1}(\Omega) \to H_0^1(\Omega)$ is K-convex with respect to $K :=$

$\{v \in H_0^1(\Omega) : v \geq 0$ a.e. on $\Omega\}$. To show this, we apply Theorem 4. For this purpose, let $W = H^{-1}(\Omega)$ and

$$K_W := K^+ = \left\{\xi \in H^{-1}(\Omega) : \langle \xi, v\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} \geq 0 \text{ for all } v \in H_0^1(\Omega)\right\}.$$

By the assumed concavity of $\Phi : \mathbb{R} \to \mathbb{R}$ we obtain for an arbitrary test function $v \in K$ that

$$\begin{aligned}\langle A(ty_1 + (1-t)y_0), v\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} &= t(\nabla y_1, \nabla v)_{L^2(\Omega)} + (1-t)(\nabla y_0, \nabla v)_{L^2(\Omega)} \\ &+ (\Phi(ty_1 + (1-t)y_0), v)_{L^2(\Omega)} \geq t\left((\nabla y_1, \nabla v)_{L^2(\Omega)} + (\Phi(y_1), v)_{L^2(\Omega)}\right) \\ &+ (1-t)\left((\nabla y_0, \nabla v)_{L^2(\Omega)} + (\Phi(y_0), v)_{L^2(\Omega)}\right) \\ &= t\langle A(y_1), v\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} + (1-t)\langle A(y_0), v\rangle_{H^{-1}(\Omega),H_0^1(\Omega)}\end{aligned}$$

and hence the K_W-concavity of A. By the monotonicity of Φ the operator A is strongly monotone and moreover, it is continuous by the assumed continuity of Φ and hence we obtain its invertibility by the Browder-Minty Theorem, see [Cia13, Theorem 9.14-1]. Using Theorem 4 yields the claimed K-convexity. In fact, a closer look at the above arguments shows that the arguments can be used for other semilinear elliptic PDEs, e.g. involving Neumann boundary conditions as well. We will come back to this in the last section of this work. As a practically relevant case, choose $\Phi(y) := -(-y)^+$, which is equivalent to the setting of [CMWC18].

3.2.2 Application to variational inequalities

Let Y be a reflexive and preordered Banach space as well as a vector lattice with order cone K and consider a K^+-concave, demicontinuous (i.e. for all sequences $(y_n)_{n\in\mathbb{N}}$ with $y_n \to y$ in Y it holds $A(y_n) \rightharpoonup A(y)$ in Y^*, cf. [Rou05, Definition 2.3]) and strongly monotone (cf. [Rou05, Definition 2.1 (iii)]) operator $A : Y \to Y^*$, that is moreover strictly T-monotone (cf. [Rod87, Equation (5.7)]), i.e.,

$$\langle A(y+z) - A(y), (-z)^+\rangle_{Y^*,Y} < 0 \text{ for } z \in Y \text{ with } (-z)^+ \neq 0.$$

Furthermore, let $C : U \rightrightarrows Y$ be a set-valued operator with a convex graph and values *with lower bound*, i.e. for all $u \in U$ it holds that $C(u) + K \subseteq C(u)$ and $y_0, y_1 \in C(u)$ implies $\min(y_0, y_1) \in C(u)$ (see [Wac16, Definition 5.4.9]). Moreover, let $w \in Y^*$ be given. We consider the following variational inequality problem (VI):
Seek $y \in C(u)$ such that

$$w \in A(y) + N_{C(u)}(y), \tag{VI}$$

where $N_{C(u)}(\,\cdot\,)$ denotes the normal cone mapping (see [AF90]). We have $N_{C(u)}(y) = (C(u) - y)^\circ$. Then, the solution operator $S : Y^* \times U \to Y$ is K-convex:
Setting $\bar{A}(u, y) := A(y) + N_{C(u)}(y)$, we check the conditions of Corollary 5. The strong monotonicity and the existence theory for VIs (cf. [KS80]) yield the single-valuedness of $\bar{A}(u, \cdot\,)^{-1}$.
To prove the isotonicity condition take $w_1 \geq_{K_W} w_0$ and set $y_j = S(w_j, u)$ for $j = 0, 1$. By testing with $z_1 := \max(y_0, y_1) = y_1 + (y_0 - y_1)^+ \in C(u)$ and $z_0 := \min(y_0, y_1) = y_0 - (y_0 - y_1)^+ \in C(u)$ we obtain

$$\begin{aligned}\langle A(y_1), z_1 - y_1\rangle_{Y^*,Y} &= \langle A(y_1), (y_0 - y_1)^+\rangle_{Y^*,Y} \geq \langle w_1, (y_0 - y_1)^+\rangle_{Y^*,Y} \\ &\geq \langle w_0, (y_0 - y_1)^+\rangle_{Y^*,Y} \geq \langle A(y_0), (y_0 - y_1)^+\rangle_{Y^*,Y} \\ &= \langle A(y_0), y_0 - z_0\rangle_{Y^*,Y}\end{aligned}$$

and hence $\langle A(y_1) - A(y_0), (y_0 - y_1)^+\rangle \geq 0$, which implies $y_1 \geq_K y_0$ by the strict T-monotonicity. To show the K^+-concavity of $\bar{A}$ we use the K^+-concavity of A. Since $C(u) + K \subseteq C(u)$ we can easily show that $N_{C(u)}(y) \subseteq -K^+$ and since $0 \in N_{C(u)}(y)$ for $y \in C(u)$ we have $N_{C(u)}(y) - K^+ = -K^+$. Letting $(u_j, y_j) \in \operatorname{gph}(C)$ for $j = 0, 1$ we obtain by the convexity of the graph $ty_1 + (1-t)y_0 \in C(tu_1 + (1-t)u_0)$, which implies $N_{C(tu_1+(1-t)u_0)}(ty_1 + (1-t)y_0) \neq \emptyset$. Hence, the concavity condition reads as

$$\begin{aligned} tN_{C(u_1)}(y_1) + (1-t)N_{C(u_0)}(y_0) &\subseteq N_{C(tu_1+(1-t)u_0)}(ty_1 + (1-t)y_0) - K^+ \\ &= -K^+, \end{aligned}$$

which is fulfilled since $N_{C(u_j)}(y_j) \subseteq -K^+$. Hence, we have verified the assumptions in Corollary 5 and deduce the K-convexity of the solution operator.

4 Subdifferential of vector-valued convex operators

In the previous section, we devoted our attention to the vector-valued convexity of solution operators. After deriving results on the abstract level we applied them to a selection of applications. We want to proceed in this manner and return to optimal control problems. A central task is the derivation of first-order optimality conditions. The latter is challenging, if the underlying solution operator is non-smooth as it is the case for VIs. On the one hand, we are interested in extending the preceding generalization of convexity to (generalized) subdifferentials for vector-valued convex operators. On the other hand, the presence of possibly non-smooth solution operators necessitates a weakened differentiation concept, not only in view of optimal control problems. Regarding the latter, the establishment of calculus rules incorporating compositions of convex functionals and K-convex operators is of interest. As a starting point, we prove some basic properties on S, respectively $S_K(\,\cdot\,) := S(\,\cdot\,) + K$ with $S_K : U \rightrightarrows Y$, where $K \subseteq Y$ is again a non-empty, closed, convex cone in the Banach space Y.

Theorem 7. *Let $S : U \to Y$ be a locally bounded, K-convex operator. Then, the normal cone of the graph of S_K is characterized as*

$$\begin{aligned} N_{\operatorname{gph}(S_K)}(u, y) = \{(h^*, d^*) \in U^* \times Y^* : \;& d^* \in N_K(y - S(u)), \\ & h^* \in \partial\langle -d^*, S(\,\cdot\,)\rangle(u)\}. \end{aligned}$$

Proof. Defining the set

$$\mathcal{N} := \{(h^*, d^*) \in U^* \times Y^* : d^* \in N_K(y - S(u)), h^* \in \partial\langle -d^*, S(\,\cdot\,)\rangle(u)\}$$

we have to prove $N_{\operatorname{gph}(S_K)}(u, y) = \mathcal{N}$:
Step 1: $N_{\operatorname{gph}(S_K)}(u, y) \subseteq \mathcal{N}$.
For $u \in U$ and $y \in S_K(u)$ the relation $(h^*, d^*) \in N_{\operatorname{gph}(S_K)}(u, y)$ holds if and only if

$$\langle h^*, v - u\rangle_{U^*,U} + \langle d^*, z - y\rangle_{Y^*,Y} \leq 0 \text{ for all } (v, z) \in \operatorname{gph}(S_K).$$

By $S_K(\,\cdot\,) = S(\,\cdot\,) + K$ we can write $y = S(u) + k$ and $z = S(v) + \tilde{k}$, with $k, \tilde{k} \in K$ respectively. Taking $v = u$ we obtain

$$\langle d^*, z - y\rangle_{Y^*,Y} = \langle d^*, \tilde{k} - k\rangle_{Y^*,Y} \leq 0 \text{ for all } \tilde{k} \in K,$$

which yields $d^* \in N_K(k) = N_K(y - S(u))$ $(\subseteq -K^+)$. Hence, we obtain with $z = S(v)$ and $v \in U$ that

$$\begin{aligned} \langle h^*, v - u\rangle_{U^*,U} &\leq \langle -d^*, S(v) - S(u)\rangle_{Y^*,Y} + \langle d^*, k\rangle_{Y^*,Y} \\ &\leq \langle -d^*, S(v)\rangle_{Y^*,Y} - \langle -d^*, S(u)\rangle_{Y^*,Y} \end{aligned}$$

for all $v \in U$. Since $-d^* \in K^+$, the functional $\langle -d^*, S(\,\cdot\,)\rangle : U \to \mathbb{R}$ is convex and the above inequality characterizes

$$h^* \in \partial\left(\langle -d^*, S(\,\cdot\,)\rangle\right)(u).$$

So we obtain $N_{\mathrm{gph}(S_K)}(u,y) \subseteq \mathcal{N}$.
Step 2: $\mathcal{N} \subseteq N_{\mathrm{gph}(S_K)}(u,y)$.
Take on the other hand $(h^*, d^*) \in \mathcal{N}$. Then we get for arbitrary $v \in U$ and $z = (S(v) + \tilde{k}) \in S_K(v)$ with some $\tilde{k} \in K$ that

$$\begin{aligned}\langle h^*, v-u\rangle_{U^*,U} + \langle d^*, z-y\rangle_{Y^*,Y} &= \langle h^*, v-u\rangle_{U^*,U} + \langle d^*, S(v)-S(u)\rangle_{Y^*,Y} \\ &\quad + \langle d^*, \tilde{k}-k\rangle_{Y^*,Y} \le 0 + 0 = 0,\end{aligned}$$

which proves the equality. □

From the above lemma we are able to formulate the coderivative of the multifunction $S_K : U \rightrightarrows Y$ in $(u,y) \in \mathrm{gph}(S_K)$, see [Mor06, Definition 1.32] for the general definition of the coderivative of a mapping. In fact, we have

$$\begin{aligned} D^*S_K(u,y)(y^*) &= \{u^* \in U^* : (u^*, -y^*) \in N_{\mathrm{gph}(S_K)}(u,y)\} \\ &= \begin{cases} \partial\langle y^*, S(\cdot)\rangle(u), & \text{if } -y^* \in N_K(y - S(u)), \\ \emptyset, & \text{else.} \end{cases} \end{aligned} \tag{3}$$

Based on (3) we formulate our definition of the subdifferential of a vector-valued convex operator

Definition 8. Let Banach spaces U, Y be given, the latter equipped with a closed, convex cone $K \subseteq Y$. The subdifferential $D^*S(u) : K^+ \rightrightarrows Y^*$ is for $u \in U$ and $y^* \in K^+$ defined by

$$D^*S(u)(y^*) = D^*S_K(u, S(u))(y^*) = \partial\langle y^*, S(\,\cdot\,)\rangle(u).$$

Using the standard sum rule for subdifferentials it is straightforward to deduce the linearity relation

$$D^*S(u)(\lambda y_1^* + y_2^*) = \lambda D^*S(u)(y_1^*) + D^*S(u)(y_2^*)$$

for all $y_1^*, y_2^* \in K^+$ and $\lambda \ge 0$.
In the light of our previously discussed results for solution operators on generalized equations, we derive a characterization of the subdifferential for a solution operator of (GE) in the following inversion formula.

Theorem 9. *Let $y^* \in K^+$. Then, it holds that*

$$w^* \in D^*S(w)(y^*) \text{ if and only if } (-y^*, w^*) \in N_{\mathrm{gph}(A_{-K_W})}(S(w), w)$$

with $A_{-K_W}(y) = A(y) - K_W$ for all $y \in Y$.

Proof. Let $w^* \in D^*S(w)(y^*)$. Then $w^* \in \partial\langle y^*, S(\,\cdot\,)\rangle(u)$ holds or in other words

$$\langle w^*, w' - w\rangle_{W^*,W} + \langle -y^*, S(w') - S(w)\rangle_{W^*,W} \le 0 \text{ for all } w' \in W.$$

Taking now $w' = w - k_W$ with $k_W \in K_W$ and $w' \in A(y')$, which is the same as $y' = S(w')$, we find by the isotonicity of the solution operator $y' \le_K S(w)$ and hence

$$\langle w^*, -k_W\rangle = \langle w^*, w' - w\rangle \le \langle y^*, y' - S(w)\rangle \le 0,$$

which yields $w^* \in K_W^+$. Hence, for $y' \in Y$ and $w' = \bar{w} - k_W$ with $\bar{w} \in A(y')$ and $k_W \in K_W$ we have

$$\begin{aligned}\langle w^*, w' - w\rangle_{W^*,W} + \langle -y^*, y' - S(w)\rangle_{Y^*,Y} &= \langle w^*, -k_W\rangle_{W^*,W} + \langle w^*, \bar{w} - w\rangle_{W^*,W} \\ &\quad + \langle -y^*, S(w') - S(w)\rangle_{Y^*,Y} \leq 0,\end{aligned}$$

which means $(w^*, -y^*) \in N_{\mathrm{gph}(A_{-K_W})}(w, S(w))$.
For the other direction we assume the latter. Then, take $w' = w - k_W \in A(S(w)) - K_W$ for an arbitrary $k_W \in K_W$. We have $\langle w^*, -k_W\rangle_{W^*,W} \leq 0$ and hence $w^* \in K_W^+$. For $w' \in A(y')$ the assumption yields

$$\langle w^*, w' - w\rangle_{W^*,W} + \langle -y^*, S(w') - S(w)\rangle_{Y^*,Y} \leq 0.$$

Since by assumption $y^* \in K^+$, the map $w \mapsto \langle y^*, S(w)\rangle$ is convex and hence the above reads $w^* \in \partial\langle y^*, S(\,\cdot\,)\rangle(w) = D^*S(w)(y^*)$. □

We continue our investigation with $A : W \to Y$ being single-valued. In this case, we obtain the following result.

Corollary 10. *Let $S : W \to Y$ denote the solution operator of the equation $w = A(y)$ for an operator $A : W \to Y$ being K_W-concave with an isotone inverse. Then, for $y^* \in K^+$ it holds that*

$$w^* \in D^*S(w)(y^*) \text{ if and only if } -y^* \in D^*(-A)(S(w))(w^*).$$

Proof. By Theorem 9 we have $w^* \in D^*S(w)(y^*)$ if and only if $(-y^*, w^*) \in N_{\mathrm{gph}(A_{-K_W})}(y, w)$ with $y = S(w)$. The latter is equivalent to

$$\langle -y^*, y' - y\rangle_{Y^*,Y} + \langle w^*, w' - w\rangle_{W^*,W} \leq 0 \text{ for all } y' \in Y \text{ and } w' \in A(y') - K_W.$$

Setting $y' = y$ and $w' = w - k_W$ for an arbitrary $k_W \in K_W$ yields again $w^* \in K_W^+$. Since A is K_W-concave, the mapping $y \mapsto \langle w^*, -A(y)\rangle_{W^*,W}$ is a convex functional. Testing with arbitrary $y' \in Y$ and $w' = A(y')$, we obtain

$$\langle -y^*, y' - y\rangle_{Y^*,Y} + \langle w^*, -A(y)\rangle_{W^*,W} \leq \langle w^*, -A(y')\rangle_{W^*,W},$$

which yields $-y^* \in D^*(-A)(y)(w^*)$. The other direction follows as in the proof of Theorem 9. □

With these results at hand, we return to the applications represented in the Subsections 3.2.1 and 3.2.2 and calculate the subdifferentials of the respective solution operators.

4.1 Application to semilinear elliptic PDEs

As a first application of the results in Section 4, we return to the class of semilinear elliptic PDEs discussed in Subsection 3.2.1. The characterization of the subdifferential of the solution operator is presented in the following theorem.

Theorem 11. *Let $S : H^{-1}(\Omega) \to H_0^1(\Omega)$ denote the solution operator of the following elliptic PDE problem:*
Given $w \in H^{-1}(\Omega)$ seek $y \in H_0^1(\Omega)$ such that

$$\begin{aligned} -\Delta y + \Phi(y) &= w \text{ in } \Omega, \\ y &= 0 \text{ on } \partial\Omega \end{aligned} \tag{4}$$

holds. Assume $\Phi : \mathbb{R} \to \mathbb{R}$ to be a continuous, non-decreasing and concave function inducing a continuous superposition operator $\Phi : L^2(\Omega) \to L^2(\Omega)$.
Let $y^* \in K^+$ with $K := \{z \in H_0^1(\Omega) : z \geq 0 \text{ a.e. on } \Omega\}$. Then $w^* \in D^*S(w)(y^*)$ holds if and only if there exists a measurable function $m : \Omega \to \mathbb{R}$ with $-m(x) \in \partial(-\Phi)(y(x))$ a.e. such that the following PDE is satisfied

$$\begin{aligned} -\Delta w^* + m w^* &= y^* \text{ in } \Omega, \\ w^* &= 0 \text{ on } \partial\Omega. \end{aligned} \tag{5}$$

Proof. First let $w^* \in D^*S(w)(y^*)$. Define the operator $A : H_0^1(\Omega) \to H^{-1}(\Omega)$

$$\langle A(y), z\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} = (\nabla y, \nabla z)_{L^2(\Omega;\mathbb{R}^d)} + (\Phi(y), z)_{L^2(\Omega)}.$$

By Corollary 10 this is equivalent to $-y^* \in D^*(-A)(y)(w^*)$, where it is also proven that $w^* \in K_W^+ = K$. Hence, $w^* \geq 0$ a.e. on Ω. For arbitrary $z \in H_0^1(\Omega)$ we obtain for y^* the following inequality:

$$\begin{aligned} \langle -y^*, z\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} &\leq -\langle -\Delta z + \Phi(y+z) - \Phi(y), w^*\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} \\ &= -\langle \Phi(y+z) - \Phi(y), w^*\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} - \langle -\Delta w^*, z\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} \end{aligned} \tag{6}$$

and hence $\langle \Delta w^* + y^*, z\rangle \geq \langle \Phi(y+z) - \Phi(y), w^*\rangle$. Testing now with $z \in K$ yields $\langle \Delta w^* + y^*, z\rangle \geq 0$ by the non-decreasing nature of $\Phi : \mathbb{R} \to \mathbb{R}$. Hence, we can identify the distribution $\Delta w^* + y^*$ with a Borel measure μ. Let $E \in \mathcal{B}(\Omega)$ be a Borel set. Since $C_0^\infty(\Omega) \subseteq L^2(\Omega)$ is dense, there exists a sequence $\widetilde{\varphi}_n \in C_0^\infty(\Omega)$ with $\widetilde{\varphi}_n \to \mathbb{1}_E$ in $L^2(\Omega)$, where $\mathbb{1}_E$ denotes the characteristic function of E. Taking a subsequence we also obtain the pointwise convergence (Fischer–Riesz) and by setting $\varphi_n := \min(\max(\widetilde{\varphi}_n, 0), 1)$ we have a non-negative sequence in $H_0^1(\Omega)$ pointwise bounded by 1 from above and converging pointwise and in $L^2(\Omega)$ to $\mathbb{1}_E$. Using Fatou's Lemma we obtain

$$\begin{aligned} 0 \leq \mu(E) &= \int_\Omega \mathbb{1}_E \mathrm{d}\mu = \int_\Omega \liminf_{n\to\infty} \varphi_n \mathrm{d}\mu \leq \liminf_{n\to\infty} \int_\Omega \varphi_n \mathrm{d}\mu \\ &\leq \liminf_{n\to\infty} \langle \Phi(y) - \Phi(y - \varphi_n), w^*\rangle = \int_E (\Phi(y) - \Phi(y-1)) w^* \mathrm{d}x < \infty, \end{aligned}$$

where we used $L^2(\Omega) \ni \Phi(y) - \Phi(y-1) \geq \Phi(y) - \Phi(y - \varphi_n) \geq 0$ a.e. on Ω as well as the continuity of $\Phi : \mathbb{R} \to \mathbb{R}$, which gives by dominated convergence the last equality. If $\lambda^d(E) = 0$, then we obtain $\mu(E) = 0$ and hence we infer that the measure μ is absolutely continuous with respect to the Lebesgue measure. Thus, by the Radon–Nikodym theorem there exists a non-negative function $\rho \in L^1(\Omega)$ with $\mu(E) = \int_E \rho \,\mathrm{d}x$ for all $E \in \mathcal{B}(\Omega)$. Testing with $E \subseteq \{w^* = 0\}$ yields as well $\mu(E) = 0$ and $\rho = 0$ on $\{w^* = 0\}$ and we rewrite $\rho = m w^*$ for a measurable function $m : \Omega \to \mathbb{R}$. Using the characterization in equation (6) we get

$$\int_\Omega \Big((-\Phi)(y+z) - (-\Phi)(y) + mz\Big) w^* \mathrm{d}x \geq 0$$

for all $z \in H_0^1(\Omega)$. Using the same density argument as before with $mw^* \in L^1(\Omega)$, we can as well test with $z = t\mathbb{1}_E \in L^\infty(\Omega)$ for $t \in \mathbb{R}$ and E again a Borel set. Hence, we find on $\{w^* > 0\}$ that for all $t \in \mathbb{R}$ it holds that

$$(-\Phi)(y+t) - (-\Phi)(y) \geq -mt \text{ a.e. on } \{w^* > 0\},$$

which, due to the convexity of $-\Phi$, implies $-m(x) \in \partial(-\Phi)(y(x))$ for almost all $x \in \{w^* > 0\}$. Since the values of m on $\{w^* = 0\}$ do not matter, we can deduce without loss of generality $-m(x) \in \partial(-\Phi)(y(x))$ on the entire domain Ω. Hence, we conclude that for all $z \in H_0^1(\Omega)$ it holds that

$$0 = \langle -\Delta w^* - y^*, z\rangle + \int_\Omega z \,\mathrm{d}\mu = \int_\Omega (\nabla w^* \cdot \nabla z + mw^* z)\,\mathrm{d}x - \langle y^*, z\rangle,$$

which is the weak formulation of the PDE in the assertion.
For the other direction let now m be a measurable function with $-m(x) \in \partial(-\Phi)(y(x))$ a.e. on Ω and let $w^* \in H_0^1(\Omega)$ be the solution of (5). For an arbitrary function $z \in H_0^1(\Omega)$ we find

$$\begin{aligned} (-\Phi)(y+z) - (-\Phi)(y) &\geq -mz \quad \text{a.e. on } \Omega \text{ as well as} \\ (-\Phi)(y-z) - (-\Phi)(y) &\geq mz \quad \text{a.e. on } \Omega. \end{aligned}$$

Together, we get

$$\Phi(y+z) - \Phi(y) \leq mz \leq \Phi(y) - \Phi(y-z)$$

and since by assumption $\Phi : H_0^1(\Omega) \to L^2(\Omega)$ is well defined, we obtain $mz \in L^2(\Omega)$. Since Φ is non-decreasing $m \geq 0$ a.e. on Ω holds, so testing (5) with $z = (-w^*)^+$ yields

$$\begin{aligned} 0 &\geq -\|\nabla(-w^*)^+\|_{L^2(\Omega;\mathbb{R}^d)}{}^2 \geq -\|\nabla(-w^*)^+\|_{L^2(\Omega;\mathbb{R}^d)}{}^2 - \int_\Omega m\left((-w^*)^+\right)^2 \mathrm{d}x \\ &= (\nabla w^*, \nabla(-w^*)^+)_{L^2(\Omega;\mathbb{R}^d)} + \int_\Omega mw^*(-w^*)^+ \mathrm{d}x \\ &= \langle y^*, (-w)^+\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} \geq 0, \end{aligned}$$

from which we deduce $w^* \geq 0$. Multiplying (5) by the solution of (4) yields

$$\begin{aligned} \langle -y^*, z\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} &= \langle \Delta w^* - mw^*, z\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} = \langle \Delta z, w^*\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} + \int_\Omega (-mz)w^* \mathrm{d}x \\ &\leq \langle \Delta z, w^*\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} - \langle \Phi(y+z) - \Phi(y), w^*\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} \\ &= -\langle -\Delta z + \Phi(y+z) - \Phi(y), w^*\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} \end{aligned}$$

for all $z \in H_0^1(\Omega)$, which proves $y^* \in D^*A(y)(-w^*)$ and equivalently $w^* \in D^*S(w)(y^*)$. □

Theorem 11 is related to the results in [CMWC18] as follows:
Consider the case $\Phi(z) := -(-z)^+$ yielding the PDE

$$\begin{aligned} -\Delta y - (-y)^+ &= w \quad \text{in } \Omega, \\ y &= 0 \quad \text{on } \partial\Omega \end{aligned}$$

for given $w \in H^{-1}(\Omega)$. Then, for $y^* \in H^{-1}(\Omega)$ our subdifferential reads as $w^* \in D^*S(w)(y^*)$, if and only if there exists a function $m \in L^\infty(\Omega)$ with

$$\begin{aligned} 0 \leq m &\leq 1 \text{ a.e. on } \{y = 0\}, \\ m &= 1 \text{ a.e. on } \{y < 0\}, \text{ and} \\ m &= 0 \text{ a.e. on } \{y > 0\}, \end{aligned}$$

such that

$$\begin{aligned} -\Delta w^* + m w^* &= y^* \text{ on } \Omega, \\ w^* &= 0 \text{ on } \partial\Omega. \end{aligned}$$

This corresponds to the *strong-weak Bouligand subdifferential* calculated in [CMWC18]. For the sake of brevity we do not introduce the details of Bouligand subdifferentials here. In the next subsection, we return to variational inequalities, where we apply our results on an abstract level as well as for the obstacle problem (cf. [KS80, Rod87]).

4.2 Applications to VI solution operators

Returning to the setup in Subsection 3.2.2, we draw our attention to (VI), where we additionally assume the operator $A : Y \to Y^*$ to be linear, bounded and coercive.
There, we provided the K-convexity of the solution operator $S : U \times Y^* \to Y$ of (VI). For the calculation of its subdifferential we utilize the inversion formula given in Theorem 9 and obtain the following result.

Theorem 12. *Let $S : Y^* \times U \to Y$ denote the solution operator of* (VI) *and take $y^* \in K^+$. The subdifferential of S in (w, u) reads with $y := S(w, u)$ as*

$$\begin{aligned} D^*S(w,u)(y^*) = \{(w^*, u^*) \in Y \times U^* : w^* \in K \cap \{Ay - w\}^\perp \text{ and} \\ (u^*, A^*w^* - y^*) \in N_{\mathrm{gph}(C)}(u, y)\}. \end{aligned}$$

Proof. Our aim is the use of Theorem 9. For this purpose we introduce — as in the proof of Corollary 5 — the mapping $\bar{A} : U \times Y \to Y^* \times U$ defined by $\bar{A}(u, y) := (Ay + N_{C(u)}(y)) \times \{u\}$ and obtain as solution mapping $(w, u) \mapsto \bar{S}(w, u) := (u, S(w, u))$. From the inversion formula in Theorem 9, we infer for $y^* \in K^+$ that

$$(w^*, u^*) \in D^*S(w,u)(y^*) \text{ if and only if } (w^*, u^*) \in D^*\bar{S}(w,u)(0, y^*),$$

which is equivalent to

$$(0, -y^*, w^*, u^*) \in N_{\mathrm{gph}(\bar{A}_{-K^+ \times \{0\}})}(u, S(w,u), w, u).$$

Hence, it is left to calculate the normal cone of the graph of $\bar{A} - (K^+ \times \{0\})$. For this sake let $(u, y, w, u) \in \mathrm{gph}(\bar{A}_{-K^+ \times \{0\}})$ and $(-v^*, -y^*, w^*, u^*) \in N_{\mathrm{gph}(\bar{A}_{-K^+ \times \{0\}})}(u, y, w, u)$. Since for $u' \in U$ it holds that $C(u') + K \subseteq C(u')$ we obtain for all $y' \in C(u')$ that $N_{C(u')}(y') \subseteq -K^+$, which yields $\bar{A}(u', y') - K^+ \times \{0\} = (Ay' - K^+) \times \{u'\}$. With $\xi := Ay - w \in K^+$ we obtain for all $(w', u') = (Ay' - \xi', u') \in (Ay' - K^+) \times \{u'\}$ that

$$\begin{aligned} 0 &\geq \langle w' - w, w^* \rangle_{Y^*,Y} + \langle u^*, u' - u \rangle_{U^*,U} + \langle -v^*, u' - u \rangle_{U^*,U} + \langle -y^*, y' - y \rangle_{Y^*,Y} \\ &= -\langle \xi' - \xi, w^* \rangle_{Y^*,Y} + \langle A^*w^* - y^*, y' - y \rangle_{Y^*,Y} + \langle u^* - v^*, u' - u \rangle_{U^*,U} \end{aligned} \tag{$*$}$$

for all $\xi' \in K^+, u' \in U, y' \in C(u')$.
First, we test $(*)$ with $y' = y$ and $u' = u$. Then, we get $\langle \xi' - \xi, -w^* \rangle_{Y^*,Y} \leq 0$ for all $\xi' \in K^+$. By setting $\xi' = \xi + k^+$ for a $k^+ \in K^+$ we see $\langle k^+, w^* \rangle_{Y^*,Y} \geq 0$ and using $k^+ = \xi$ especially $\langle \xi, w^* \rangle_{Y^*,Y} \geq 0$. Setting $\xi' = 0$ yields $\langle \xi, w^* \rangle_{Y^*,Y} \leq 0$ and thus $w^* \in K \cap \{Ay - w\}^\perp$. By testing with an arbitrary $u' \in U$ with $y' \in C(u')$ and $\xi' = 0$ we get $(u^* - v^*, A^*w^* - y^*) \in N_{\mathrm{gph}(C)}(u, y)$.

To show the other direction, let $w^* \in K \cap \{Ay - w\}^\perp$ such that $(u^* - v^*, A^* w^* - y^*) \in N_{\mathrm{gph}(C)}(y)$ and write again $\xi = Ay - u$. Then we get

$$\begin{aligned} 0 &\geq \langle \xi' - \xi, -w^* \rangle_{Y^*,Y} + \langle A^* w^* - y^*, y' - y \rangle_{Y^*,Y} + \langle u^* - v^*, u' - u \rangle_{U^*,U} \\ &= \langle (Ay' - \xi') - (Ay - \xi), w^* \rangle_{Y^*,Y} + \langle u^*, u' - u \rangle_{U^*,U} \\ &+ \langle -v^*, u' - u \rangle_{U^*,U} + \langle -y^*, y' - y \rangle_{Y^*,Y} \end{aligned}$$

for all $\xi' \in K^+, u' \in U$ and $y' \in C(u')$. This implies $(-v^*, -y^*, w^*, u^*) \in N_{\mathrm{gph}(\bar{A}_{-K^+ \times \{0\}})}(u, y, w, u)$. Summarizing, we obtain for the operator S

$$\begin{aligned} D^* S(w, u)(y^*) &= D^* \bar{S}(w, u)(0, y^*) \\ &= \{(w^*, u^*) \in Y \times U^* : w^* \in K \cap \{Ay - w\}^\perp \text{ and} \\ &\qquad (u^*, A^* w^* - y^*) \in N_{\mathrm{gph}(C)}(u, y)\}, \end{aligned}$$

which yields the assertion. □

An important subclass of VIs is associated with $C(u) = C$ for all $u \in U$ with $C \subseteq Y$ a non-empty, closed, convex set. In this case, it holds that $N_{\mathrm{gph}(C)}(u, y) = N_{U \times C}(u, y) = \{0\} \times N_C(y)$ for all $y \in C$ and $u \in U$. For the characterization of the subdifferential some aspects and results from *Capacity Theory* (cf. [BS00]) are provided in the appendix and are used in the following theorem.

Theorem 13. *Let $S : Y^* \to Y, w \mapsto y$, denote in the setting of Theorem 12 the solution operator of the following VI:*
Seek $y \in Y$ such that

$$w \in Ay + N_C(y) \text{ in } Y^*. \tag{7}$$

Then, we obtain for $y := S(w)$ that

$$D^* S(w)(y^*) = \left\{ w^* \in Y : w^* \in K \cap \{Ay - w\}^\perp \text{ and } A^* w^* - y^* \in N_C(y) \right\}.$$

Proof. We apply Theorem 12 using $C(u) = C$ for all $u \in U$. Further, one observes

$$w^* \in D^* S(w)(y^*) \Leftrightarrow (w^*, 0) \in D^* \bar{S}(w, u)(y^*),$$

where $\bar{S}$ is the solution operator defined in the proof of Theorem 12. This yields $-w^* \in K \cap \{Ay - w\}^\perp$ and $(0, A^* w^* - y^*) \in N_{\mathrm{gph}(C)}(u, \bar{S}(w, u)) = \{0\} \times N_C(S(w))$ and thus the assertion. □

Next, we study the special case of the obstacle problem. For this, we let $Y = H_0^1(\Omega)$ be equipped with the order cone $K := \{z \in H_0^1(\Omega) : z \geq 0 \text{ a.e. on } \Omega\}$. For the VI we assume $A = -\Delta$ and

$$C := \{z \in H_0^1(\Omega) : z \geq \psi \text{ a.e. on } \Omega\}$$

with $\psi \in H^1(\Omega), \psi \leq 0$ on $\partial\Omega$. For $w \in H^{-1}(\Omega)$ we set $y = S(w)$ and define the inactive set $\mathcal{I}(y) := \{x \in \Omega : y(x) > \psi(x)\}$, the active set $\mathcal{A}(y) := \Omega \backslash \mathcal{I}(y)$ and the strictly active set as $\mathcal{A}_s(y) := \text{f-supp}(w + \Delta y)$, where f-supp denotes the *fine support*, see Lemma 21 in the appendix. Then it can be shown, that the tangential cone of C in $y \in C$ reads

$$T_C(y) = \{z \in H_0^1(\Omega) : z \geq 0 \text{ q.e. on } \mathcal{A}(y)\},$$

where 'q.e.' stands for 'quasi-everywhere' (cf. [BS00] for more details). By the techniques involving capacitary measures from [RW19] and the references therein we deduce the following characterization of the subdifferential of S:

Theorem 14. *Let $w \in H^{-1}(\Omega)$ with $y = S(w)$ and $y^* \in K^+$. Then, $w^* \in D^*S(w)(y^*)$ if and only if there exists a capacitary measure $m \in \mathcal{M}_0(\Omega)$ such that*

$$m(\mathcal{I}(y)) = 0 \text{ and } m = +\infty \text{ on } \mathcal{A}_s(y),$$

and $w^* \in H_0^1(\Omega) \cap L_m^2(\Omega)$ with $L_m^2(\Omega)$ defined as in (13) *in the appendix, solves the system*

$$\begin{aligned} -\Delta w^* + mw^* &= y^* \text{ in } \Omega, \\ w^* &= 0 \text{ on } \partial\Omega, \text{ i.e.} \end{aligned} \tag{8}$$

for all $v \in H_0^1(\Omega) \cap L_m^2(\Omega)$ it holds that

$$(\nabla w^*, \nabla v)_{L^2(\Omega)} + \int_\Omega w^* v \, \mathrm{d}m = \langle y^*, v\rangle_{H^{-1}(\Omega), H_0^1(\Omega)}.$$

Proof. We use the characterization of the subdifferential given in Theorem 13. Let first $w^* \in H_0^1(\Omega)$ be given with $w^* \in K \cap \{-\Delta y - w\}^\perp$ and $-\Delta w^* - y^* \in N_C(y)$.
The latter implies $\Delta w^* + y^* \in K^+$ and according to Lemma 21 we can identify the functional with a non-negative Borel measure. Let $E \in \mathcal{B}(\Omega)$ be an arbitrary Borel set. We define the measure m as follows

$$m(E) := \begin{cases} \int_E \frac{1}{w^*} \mathrm{d}(y^* + \Delta w^*), & \text{if } \operatorname{cap}(E \cap \{w^* = 0\}) = 0, \\ +\infty, & \text{else.} \end{cases}$$

Since $\langle w + \Delta y, w^*\rangle_{H^{-1}(\Omega), H_0^1(\Omega)} = 0$ and $w^* \geq 0$ a.e. on Ω, we obtain that $\{w^* = 0\}$ q.e. on $\mathcal{A}_s(y)$ and hence $m = +\infty$ on $\mathcal{A}_s(y)$.
Since $\langle y^* + \Delta w^*, v\rangle_{H^{-1}(\Omega), H_0^1(\Omega)} = 0$ for all $v \in H_0^1(\mathcal{I}(y))$ we see as well $\operatorname{cap}(\text{f-supp}(y^* + \Delta w^*) \cap \mathcal{I}(y)) = 0$ and hence $m(\mathcal{I}(y)) = 0$. It is left to show, that the system is fulfilled. At first we see that $w^* \in L_m^2(\Omega)$:

$$\begin{aligned} \int_\Omega w^{*2} \, \mathrm{d}m = \int_{\{w^* \neq 0\}} w^{*2} \, \mathrm{d}m = \int_{\{w^* \neq 0\}} w^* \mathrm{d}(y^* + \Delta w^*) = \int_\Omega w^* \mathrm{d}(y^* + \Delta w^*) \\ = \langle y^* + \Delta w^*, w^*\rangle_{H^{-1}(\Omega), H_0^1(\Omega)} < \infty. \end{aligned}$$

Take now $v \in H_0^1(\Omega) \cap L_m^2(\Omega)$. Then $v = 0$ q.e. on $\{w^* = 0\}$ by the construction of m, and we obtain

$$\begin{aligned} \int_\Omega w^* v \, \mathrm{d}m &= \int_{\{w^* \neq 0\}} w^* v \, \mathrm{d}m = \int_{\{w^* \neq 0\}} v \, \mathrm{d}(y^* + \Delta w^*) \\ &= \int_\Omega v \, \mathrm{d}(y^* + \Delta w^*) = \langle y^* + \Delta w^*, v\rangle_{H^{-1}(\Omega), H_0^1(\Omega)}, \end{aligned}$$

which proves the assertion.
To prove the other direction let now $m \in \mathcal{M}_0(\Omega)$ be a capacitary measure with $m(\mathcal{I}(y)) = 0$ and $m = +\infty$ on $\mathcal{A}_s(y)$. Let $w^* \in H_0^1(\Omega) \cap L_m^2(\Omega)$ denote (8). Then, we see that $w^* = 0$ q.e. on $\mathcal{A}_s(y)$, and since $y^* \in K^+$ we deduce by testing with $v = (-w^*)^+$ that

$$0 \geq -\|\nabla(-w^*)^+\|_{L^2(\Omega)}^2 - \int_{\{w^* < 0\}} (-w^*)^2 \, \mathrm{d}m = \langle y^*, (-w^*)^+\rangle_{H^{-1}(\Omega), H_0^1(\Omega)} \geq 0,$$

and hence $w^* \geq 0$ a.e. on Ω, which proves $w^* \in K \cap \{Ay - w\}^\perp$. Let now $v \in T_C(w)$ and define similar to the proof of [DMG94, Proposition 2.6] $v_n := \min\left(\frac{1}{n} v, w^*\right)$. Then we see $0 \leq v_n \leq w^*$ q.e. on $\mathcal{A}(y)$ and $v_n = 0$ q.e. on $\mathcal{A}_s(y)$. Since $m(\mathcal{I}(y)) = 0$ we obtain

$$\int_\Omega v_n^2 \, \mathrm{d}m = \int_{\mathcal{A}(y) \setminus \mathcal{A}_s(y)} v_n^2 \, \mathrm{d}m \leq \int_{\mathcal{A}(y) \setminus \mathcal{A}_s(y)} w^{*2} \, \mathrm{d}m < \infty,$$

and hence $v_n \in H_0^1(\Omega) \cap L_m^2(\Omega)$. Testing (8) with v_n we obtain similarly to before

$$\begin{aligned}\frac{1}{n}\langle y^*, v\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} &\geq \langle y^*, v_n\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} = (\nabla w^*, \nabla v_n)_{L^2(\Omega)} + \int_\Omega w^* v_n \,\mathrm{d}m \\ &= \int_{\{nw^*\leq v\}} |\nabla w^*|^2 \mathrm{d}x + \frac{1}{n}\int_{\{nw^*>v\}} \nabla w^* \cdot \nabla v \mathrm{d}x + \int_{A(y)} w^* v_n \,\mathrm{d}m \\ &\geq \frac{1}{n}\int_{\{nw^*>v\}} \nabla w^* \cdot \nabla v \mathrm{d}x.\end{aligned}$$

We multiply by n, let $n \to \infty$ and obtain using $\nabla w^* = 0$ on $\{w^* = 0\}$ that $\langle y^* + \Delta w^*, v\rangle \geq 0$. Hence, we deduce $-\Delta w^* - y^* \in N_C(y)$. □

As mentioned before, also in this case the involved operator corresponds to the *strong-weak Bouligand subdifferential* (cf. [RW19]).

5 Convex optimization problems involving vector-valued convex operators

In the previous sections, a preorder-related notion of vector-valued convex operators and an extended subdifferential concept for these operators have been introduced. Aiming at optimal control problems we might be confronted with the treatment of optimization problems of the following type:

$$\text{minimize} \quad f(u) + g(S(u)) \text{ over } u \in U. \tag{9}$$

The minimization problem in (9) matches the reduced formulation of an optimal control problem, where S represents the solution operator of a partial differential equation or a variational inequality as discussed in the previous subsections. The derivation of necessary optimality conditions is a central task. Especially in the context of VIs this is non-trivial and requires the use of generalized differentiation concepts. In view of the latter the previously derived results for the subdifferential will be analyzed next for their use for the derivation of first-order conditions. For this sake, the upcoming section is devoted to the derivation of calculus rules addressing compositions of a K-convex operators with convex functionals which are compatible with the preorder relation. We start with the derivation of a chain rule.

Lemma 15. *Let U, Y be Banach spaces, the latter one equipped with a non-empty, closed, convex cone K. Let $g : Y \to \mathbb{R} \cup \{+\infty\}$ be a convex, lower semi-continuous, proper and K-isotone functional. Suppose $S : U \to Y$ is a locally bounded, K-convex operator. Then, $g \circ S : U \to \mathbb{R} \cup \{+\infty\}$ is convex as well.*
Moreover, consider $u \in U$ with $S(u) \in \mathcal{D}(\partial g)$ and let the following constraint qualification hold

$$0 \in \operatorname{core}\left(S(U) - \operatorname{dom}(g)\right).$$

Then, for the subdifferential it holds that

$$\partial(g \circ S)(u) = D^*S(u)\left(\partial g(S(u))\right) = \bigcup_{y^* \in \partial g(S(u))} \partial\langle y^*, S(\,\cdot\,)\rangle(u).$$

Proof. Let $u \in U$ be as above and define $M := \bigcup_{y^* \in \partial g(S(u))} \partial \langle y^*, S(\,\cdot\,)\rangle(u)$. By the assumption on u we get $\partial g(S(u)) \neq \emptyset$. Let $y^* \in \partial g(y)$ for some $y \in \mathcal{D}(\partial g)$. Then we obtain for $k \in K$, that

$$0 \geq g(y-k) - g(y) \geq \langle y^*, y-k-y\rangle_{Y^*,Y} = -\langle y^*, k\rangle_{Y^*,Y}$$

and hence we obtain $\partial g(y) \subseteq K^+$ and further the convexity of $u \mapsto \langle y^*, S(u)\rangle$. By the local boundedness of S we obtain local boundedness of $u \mapsto \langle y^*, S(u)\rangle$ as well and by [ET76, Lemma 2.1] also its continuity on all of U. So the set M is well defined in the sense of the convex subdifferential and non-empty.
Take $u^* \in M$. Then, there exists $y^* \in \partial g(S(u))$ with $u^* \in \partial\langle y^*, S(\,\cdot\,)\rangle(u)$ such that

$$g(S(v)) \geq g(S(u)) + \langle y^*, S(v) - S(u)\rangle_{Y^*,Y} \geq g(S(u)) + \langle u^*, v-u\rangle_{U^*,U}$$

and hence $M \subseteq \partial(g \circ S)(u)$. Since $M \neq \emptyset$ we can now take $u^* \in \partial(g\circ S)(u)$ and obtain by the Fenchel-Legendre identity, that

$$\langle u^*, u\rangle_{U^*,U} = (g\circ S)(u) + (g \circ S)^*(u^*).$$

Hence, we know that $u \in \operatorname{argmin}_{v\in U}(g(S(v)) - \langle u^*, v\rangle)$. Using the K-isotonicity this is equivalent to

$$(u, S(u)) \in \operatorname{argmin}_{v \in U, y\in Y}\left(g(y) - \langle u^*, v\rangle + I_{\operatorname{gph}(S_K)}(v,y)\right),$$

with I_M Hence, the first-order condition holds at $(u, S(u))$, i.e.,

$$0 \in \partial\left(g(\operatorname{pr}_Y(\,\cdot\,)) - \langle u^*, \operatorname{pr}_U(\,\cdot\,)\rangle + i_{\operatorname{gph}(S_K)}(\,\cdot\,)\right)(u, S(u)), \tag{10}$$

where pr_U and pr_Y denote the projection operator on the component in U and Y, respectively. By our constraint qualification we know that for every $z \in Y$ there exists a $t > 0$ such that $z \in t\left(S(U) - \operatorname{dom}(g)\right)$. Thus, there exists a pair $(u_1, y_2) \in U \times \operatorname{dom}(g)$ with $z = t(S(u_1) - y_2)$. For an arbitrary $v \in U$ choose $u_2 = u_1 - \frac{1}{t}v$ and $y_1 = S(u_1)$. Then, we obtain $v = t(u_1 - u_2)$ and $z = t(y_1 - y_2)$, which means that $(v, z) \in t\left(\operatorname{gph}(S_K) - U \times \operatorname{dom}(g)\right)$. This yields the constraint qualification $0 \in \operatorname{core}\left(\operatorname{gph}(S_K) - U \times \operatorname{dom}(g)\right)$ and allows us to use the sum rule in the inclusion (10), which yields

$$0 \in \{-u^*\} \times \partial g(S(u)) + N_{\operatorname{gph}(S_K)}(u, S(u)).$$

Utilizing Theorem 7 we deduce the existence of $y^* \in \partial g(S(u))$ together with $d^* \in N_K(S(u) - S(u)) = -K^+$ as well as $h^* \in \partial\langle -d^*, S(\,\cdot\,)\rangle(u)$ such that

$$\begin{aligned} 0 &= y^* + d^*,\\ 0 &= -u^* + h^*, \end{aligned}$$

which yields $u^* = h^* \in \partial\langle -d^*, S(\,\cdot\,)\rangle(u) = \partial\langle y^*, S(\,\cdot\,)\rangle(u)$ and eventually $u^* \in M$. □

Next, we derive a differentiation rule combining a sum as well as a composition.

Theorem 16. *Let U, Y be Banach spaces, the latter one equipped with a closed, convex cone K. Let $f : U \to \mathbb{R} \cup \{+\infty\}$ and $g : Y \to \mathbb{R}\cup\{+\infty\}$ be convex, proper, lower semi-continuous functionals, and moreover let g be K-isotone. Consider $S : U \to Y$ a locally bounded, K-convex operator. Then, the functional $f + g\circ S : U \to \mathbb{R}\cup\{+\infty\}$ is convex. Moreover, consider $u \in \mathcal{D}(\partial f)$ with $S(u) \in \mathcal{D}(\partial g)$ and let the following constraint qualification hold:*

$$0 \in \operatorname{core}\left(\operatorname{dom}(f) \times \operatorname{dom}(g) - \operatorname{gph}(S)\right).$$

Then, the subdifferential reads as

$$\partial(f + g\circ S)(u) = \partial f(u) + D^*S(u)\left(\partial g(S(u))\right).$$

Proof. Consider the functional $h(u,y) := f(u) + g(y)$ together with the convex, closed cone $\bar{K} := \{0\} \times K$ and the operator $T : U \to U \times Y$ defined by $T(u) := (u, S(u))$. Then the operator T is $\bar{K}$-convex and locally bounded, and the functional h is convex, proper, lower semi-continuous and $\bar{K}$-isotone. By assumption on u we have $(u, S(u)) = T(u) \in \mathcal{D}(\partial h) = \mathcal{D}(\partial f) \times \mathcal{D}(\partial g)$ and the constraint qualification reads as $0 \in \operatorname{core}(\operatorname{dom}(h) - T(U))$. Hence, we are in the position to use Lemma 15 and obtain with

$$\begin{aligned} D^*T(u)(u^*, y^*) &= \partial\langle (u^*, y^*), T(\,\cdot\,)\rangle(u) = \partial\langle u^*, \cdot\rangle_{U^*,U}(u) + \partial\langle y^*, S(\,\cdot\,)\rangle_{Y^*,Y}(u) \\ &= u^* + D^*S(u)(y^*) \end{aligned}$$

finally for the subdifferential that

$$\begin{aligned} \partial(f + g\circ S)(u) = \partial(h \circ T)(u) &= D^*T(u)\left(\partial h(T(u))\right) \\ &= D^*T(u)\left(\partial f(u) \times \partial g(S(u))\right) \\ &= \partial f(u) + D^*S(u)\left(\partial g(S(u))\right). \end{aligned}$$

□

Next, we propose a variant of the previous result using a different constraint qualification. For this purpose we need a generalization of the Moreau-Rockafellar theorem suitable for our framework. For this sake we adapt the techniques in [BZ06, Section 4.3].

Proposition 17. *Let U, Y be Banach spaces, the latter one equipped with a closed, convex cone K. Let $f : U \to \mathbb{R} \cup \{+\infty\}$ and $g : Y \to \mathbb{R} \cup \{+\infty\}$ be convex, proper, lower semi-continuous functionals and moreover let g be K-isotone. Consider $S : U \to Y$ a* demi-continuous*, K-convex operator. Suppose the following constraint qualification to be satisfied*

$$0 \in \operatorname{core}\left(\operatorname{dom}(g) - S(\operatorname{dom}(f))\right).$$

Then, there exists $y^* \in Y^*$ such that for all $u \in U$ and $y \in Y$ it holds that

$$\inf_{u\in U}\left(f(u) + g(S(u))\right) \le \left(f(u) + \langle y^*, S(u)\rangle_{Y^*,Y}\right) + \left(g(y) - \langle y^*, y\rangle_{Y^*,Y}\right).$$

Proof. The lemma and the proof are strongly based on the one of [BZ06, Lemma 4.3.1].
Define the functional $h : Y \to [-\infty, +\infty]$ by

$$h(y) := \inf_{u\in U}\left(f(u) + g(S(u) + y)\right).$$

Then h is a convex functional with $\operatorname{dom}(h) = \operatorname{dom}(g) - S(\operatorname{dom}(f))$. We show that $0 \in \operatorname{int}(\operatorname{dom}(h))$. Without loss of generality we assume $f(0) = g(S(0)) = 0$ (else take $\bar{u} \in \operatorname{dom}(f)$, $\bar{y} \in \operatorname{dom}(g)$ and consider $\bar{f}(u) := f(u + \bar{u}) - f(\bar{u})$ and $\bar{g}(y) := g(y + \bar{y} - S(0)) - g(\bar{y})$). Define the set

$$M := \bigcup_{u\in\mathbb{B}_U}\{y \in Y : f(u) + g(S(u) + y) \le 1\}.$$

It is straightforward to argue the convexity of M. We show that M is *absorbing* and *cs-closed*.
We start by showing the former. For this purpose let $y \in Y$. We need to prove, that there exists $r > 0$ such that $\lambda y \in M$ for $|\lambda| \le r$. By the assumed constraint qualification $0 \in \operatorname{core}(\operatorname{dom}(g) - S(\operatorname{dom}(f)))$,

there exists $\bar{t} > 0$ with $ty \in \operatorname{dom}(g) - S(\operatorname{dom}(f))$ for all $|t| \le \bar{t}$. Hence, there exist $u_\pm \in \operatorname{dom}(f)$ with $S(u_\pm) \pm ty \in \operatorname{dom}(g)$ and we define

$$\alpha := \max(f(u_\pm) + g(S(u_\pm) \pm \bar{t}y), 1) < \infty.$$

Then we see for $|t| \le \bar{t}$ with $u_t := |t|\,\bar{t}^{-1} u_{\operatorname{sign}(t)}$ and $u_t \in \mathbb{B}_U$ that

$$\begin{aligned} f(u_t) + g(S(u_t) + ty) &= f\left(\frac{|t|}{\bar{t}} u_{\operatorname{sign}(t)}\right) + g\left(S\left(\frac{|t|}{\bar{t}} u_{\operatorname{sign}(t)}\right) + \frac{|t|}{\bar{t}} \operatorname{sign}(t)\, \bar{t}y\right) \\ &\le \frac{|t|}{\bar{t}} \left(f(u_{\operatorname{sign}(t)}) + g(S(u_{\operatorname{sign}(t)}) + \operatorname{sign}(t)\, \bar{t}y)\right) \le \frac{|t|}{\bar{t}}\alpha \le \alpha. \end{aligned}$$

Choose now $m := \max(\|u_\pm\|, \alpha, 1)$. Then we see that $\frac{u_t}{m} \in \mathbb{B}_U$ and further

$$f\left(\frac{u_t}{m}\right) + g\left(S\left(\frac{u_t}{m}\right) + \frac{t}{m}y\right) \le \frac{1}{m}\left(f(u_t) + g(S(u_t) + ty)\right) \le \frac{\alpha}{m} \le 1.$$

Hence, we can choose $r = \frac{\bar{t}}{m}$ to obtain $\lambda y \in M$ for all $\lambda \le r$. To prove the cs-closedness take $y = \sum_{k=1}^\infty \lambda_k y_k$ where $\lambda_k \ge 0$, $\sum_{k=1}^\infty \lambda_k = 1$, and $(y_k)_{k\in\mathbb{N}}$ is a sequence in M. By the definition of M there exist $(u_k)_{k\in\mathbb{N}} \subseteq \mathbb{B}_U$ with

$$f(u_k) + g(S(u_k) + y_k) \le 1 \text{ for all } k \in \mathbb{N}.$$

Since $\mathbb{B}_U$ is bounded and closed it is cs-closed and hence also cs-compact (cf. [Jam74, Theorem 22.2]). By this we set $u := \sum_{k=1}^\infty \lambda_k u_k$. As the operator S is assumed to be demi-continuous and g is convex, lower semi-continuous and hence weakly lower semi-continuous we obtain

$$f(u) + g(S(u) + y) \le 1,$$

which yields $y \in M$.
Due to the cs-closedness we obtain $\operatorname{core}(M) = \operatorname{int}(M)$ by [Sch07, Proposition 1.2.3] and since M is absorbent $0 \in \operatorname{core}(M)$, which implies $0 \in \operatorname{int}(\operatorname{dom}(h))$. From this we see that $\partial h(0) \neq \emptyset$ and take $y^* \in \partial h(0)$. Eventually, we observe for all $u \in U$ and $y \in Y$ that

$$\begin{aligned} \inf_{u\in U} \Big(f(u) + g(S(u))\Big) &= h(0) \le h(y - S(u)) - \langle y^*, y - S(u)\rangle \\ &\le f(u) + g(S(u) + y - S(u)) - \langle y^*, y - S(u)\rangle \\ &\le \Big(f(u) + \langle y^*, S(u)\rangle\Big) + \Big(g(y) - \langle y^*, y\rangle\Big), \end{aligned}$$

which proves the assertion. □

We are now ready to state another version of the differentiation rule given in Theorem 16.

Theorem 18. *Let U, Y be Banach spaces, the latter one equipped with a closed, convex cone K. Let $f : U \to \mathbb{R} \cup \{+\infty\}$ and $g : Y \to \mathbb{R} \cup \{+\infty\}$ be convex, proper, lower semi-continuous functionals and moreover let g be K-isotone. Suppose $S : U \to Y$ to be a demi-continuous, K-convex operator. Then the functional $f + g \circ S : U \to \mathbb{R} \cup \{+\infty\}$ is convex. Moreover, consider $u \in \mathcal{D}(\partial f)$ with $S(u) \in \mathcal{D}(\partial g)$ and let the following constraint qualification hold*

$$0 \in \operatorname{core}\left(S\left(\operatorname{dom}(f)\right) - \operatorname{dom}(g)\right).$$

Then, the subdifferential reads as

$$\partial(f + g \circ S)(u) = \partial f(u) + D^* S(u)\Big(\partial g(S(u))\Big).$$

Proof. The inclusion $\partial f(u) + D^*S(u)\left(\partial g(S(u))\right) \subseteq \partial(f + g \circ S)(u)$ is straightforward and its proof will therefore be omitted here. To show the reverse direction let $u^* \in \partial(f + g \circ S)(u)$. Then we obtain by the Fenchel-Legendre identity the relation

$$f(u) + g(S(u)) + (f + g \circ S)^*(u^*) = \langle u^*, u\rangle.$$

Applying Lemma 17 to $f - \langle u^*, \cdot\rangle$ (instead of f) we deduce the existence of $y^* \in Y^*$ such that for all $v \in U$ and $y \in Y$ it holds that

$$\begin{aligned} f(u) + g(S(u)) - \langle u^*, u\rangle &= -(f + g \circ S)^*(u^*) \\ &= \inf_{w\in U} \Big(f(w) - \langle u^*, w\rangle + g(S(w)) \Big) \\ &\leq f(v) - \langle u^*, v\rangle + \langle y^*, S(v)\rangle + g(y) - \langle y^*, y\rangle. \end{aligned}$$

On the one hand, setting $v = u$ implies

$$g(S(u)) + \langle y^*, y - S(u)\rangle \leq g(y) \text{ for all } y \in Y,$$

which yields $y^* \in \partial g(S(u))$. Since g is assumed to be K-isotone it holds that $y^* \in K^+$.
On the other hand, setting $y = S(u)$ implies

$$f(u) + \langle y^*, S(u)\rangle + \langle u^*, v - u\rangle \leq f(v) + \langle y^*, S(v)\rangle \text{ for all } v \in U.$$

Hence, we see $u^* \in \partial\left(f + \langle y^*, S(\,\cdot\,)\rangle\right)(u)$. Since S is defined on all of U, the second function has a domain equal to the entire space. Hence, we can apply the usual sum rule to deduce

$$u^* \in \partial f(u) + D^*S(u)(y^*) \subseteq \partial f(u) + D^*S(u)\Big(\partial g(S(u))\Big),$$

which proves the assertion. □

A closer comparison of the two versions of chain rules formulated in Theorem 16 and Theorem 18 shows, that the additional requirement of S being demicontinuous is traded with a weaker constraint qualification. The difference between these two does not occur for linear operators and is thus of interest. We briefly address the relation between these conditions in the following theorem.

Theorem 19. *Let U, Y be Banach spaces the latter one equipped with a closed, convex cone K. Then the following assertions hold:*

(i) If S is demi-continuous, then it is locally bounded.

(ii) Let $S : U \to Y$ be a K-convex operator. If S is locally bounded and K is an order cone (i.e.: $K \cap (-K) = \{0\}$), then S is demi-continuous.

Proof. ad (i): If S is not locally bounded, then there exists a point $u \in U$ such that for all $n \in \mathbb{N}$ there exists $u_n \in u + \frac{1}{n}\mathbb{B}_U$ with $\|S(u_n)\|_Y \geq n$. Then this holds $u_n \to u$ in U and by the demi-continuity $S(u_n) \rightharpoonup S(u)$ in Y implying the boundedness of $(S(u_n))_{n\in\mathbb{N}}$ — a contradiction.
ad (ii): We consider first $y^* \in K^+$. Then the mapping $u \mapsto \langle y^*, S(u)\rangle$ is convex and locally bounded from above in every point and hence continuous by [ET76, Lemma 2.1]. Then we deduce the continuity of the functional also for $y^* \in K^+ - K^+$. Let now $y^* \in Y^*$ be arbitrary. By the calculus rules of the dual cone in Lemma 1 we see that

$$\operatorname{cl}\left(K^+ - K^+\right) = \operatorname{cl}\left(K^+ + (-K)^+\right) = (K \cap (-K))^+ = \{0\}^+ = Y^*.$$

So for every $\varepsilon > 0$ we find $y^*_\varepsilon \in K^+ - K^+$ such that $\|y^* - y^*_\varepsilon\|_{Y^*} < \varepsilon$. Taking now a convergent sequence $u_n \to u$ we get by assumption the boundedness of $S(u_n)$ by some constant B. This yields

$$\begin{aligned}|\langle y^*, S(u_n)\rangle - \langle y^*, S(u)\rangle| &\le |\langle y^*_\varepsilon, S(u_n)\rangle - \langle y^*_\varepsilon, S(u)\rangle| \\ &\quad + |\langle y^* - y^*_\varepsilon, S(u_n) - S(u)\rangle| \\ &\le |\langle y^*_\varepsilon, S(u_n)\rangle - \langle y^*_\varepsilon, S(u)\rangle| + 2B\varepsilon.\end{aligned}$$

Using the continuity of $\langle y^*_\varepsilon, S(\,\cdot\,)\rangle$ the first term tends to zero as $n \to \infty$, and we finally see that

$$0 \le \limsup_{n\to\infty} |\langle y^*, S(u_n)\rangle - \langle y^*, S(u)\rangle| \le 2B\varepsilon.$$

Since the choice of ε was arbitrary we deduce the desired continuity of $u \mapsto \langle y^*, S(u)\rangle$ and hence the demi-continuity of S. □

Interestingly, Theorem 19 can also be interpreted as a generalization of [Har77, Theorem 3, Part (a)]. This has the following consequence: Having a vector lattice Y with order cone K, we obtain that the mapping $y \mapsto y^+ = \max(0, y)$ is demi-continuous, if and only if it is locally bounded (see also [Har77, Proposition 1]).

5.1 Application to doping optimization

In the last part we draw our attention to the following optimization problem, that gained some interest recently in [KS20, Section 5.2]. Here, we want to discuss in our notation the deterministic counterpart of the problem therein. Therefore consider an open, bounded domain $\Omega \subseteq \mathbb{R}^d$ with $d \in \{1, 2\}$ and Lipschitz boundary $\partial\Omega$ as well as $\Omega_o \subseteq \Omega$ an open subset.

$$\begin{aligned}\min_{u\in U_{\mathrm{ad}}} \frac{1}{2}\int_{\Omega_o} (1-z)^{2+}\mathrm{d}x &+ \frac{\alpha}{2}\int_\Omega u^2\mathrm{d}x,\\ -\operatorname{div}(\kappa\nabla z) + c\sinh z &= Bu + b \text{ in } \Omega,\\ \kappa\frac{\partial z}{\partial\nu} &= 0 \text{ on } \partial\Omega,\\ &\text{with } B \text{ as the solution operator of}\\ -r\Delta d + d &= z \text{ in } \Omega,\\ r\frac{\partial d}{\partial\nu} &= 0 \text{ on } \partial\Omega.\end{aligned} \tag{11}$$

Within the scope of this example we assume $U_{\mathrm{ad}} \subseteq L^2_+(\Omega)$ and $b \in L^2_+(\Omega)$. First, we introduce the variable $y = -z$. As the hyperbolic sine is an odd function, we rewrite (11) as

$$\begin{aligned}\min_{z\in U_{\mathrm{ad}}} \frac{1}{2}\int_{\Omega_o} (y+1)^{2+}\mathrm{d}x &+ \frac{\alpha}{2}\int_\Omega z^2\mathrm{d}x,\\ -\operatorname{div}(\kappa\nabla y) + c\sinh y &= -Bu - b \text{ in } \Omega,\\ \kappa\frac{\partial y}{\partial\nu} &= 0 \text{ on } \partial\Omega.\\ &\text{with } B \text{ as above.}\end{aligned}$$

First, we need to discuss the exponential non-linearity involved in the hyperbolic sine. For this sake, we use the version of the Trudinger–Moser inequality given in [LR08]. First, the well-definedness of the expression $\sinh : H^1(\Omega) \to L^2(\Omega)$ is ensured. Given $u \in H^1(\Omega)$, there exists $n \in \mathbb{N}$ such that $\|u - u_n\|_{H^1(\Omega)} \leq 1$ with $u_n := \min(\max(u, -n), n)$. Then, we obtain by multiple use of the Young inequality the relation

$$\begin{aligned}\exp(|u|) &\leq \exp\left(\frac{1}{2\alpha_d}\right)\exp\left(\frac{\alpha_d}{2}u^2\right) = \exp\left(\frac{1}{2\alpha_d}\right)\exp\left(\frac{\alpha_d}{2}(u_n + (u - u_n))^2\right)\\ &\leq \exp\left(\frac{1}{2\alpha_d}\right)\exp(\alpha_d n^2)\exp(\alpha_d(u - u_n)^2),\end{aligned}$$

where $\alpha_d > 0$ is as in [LR08] and thus $\int_\Omega \exp(|u|)\mathrm{d}x < \infty$ for all $u \in H^1(\Omega)$. By using $|\sinh(u)| \leq \exp(|u|)$ we obtain the well definedness.

As we aim for the application of our combined results, we first ensure the K-convexity of the solution operator $S : u \mapsto y$ with $K := \{y \in H^1(\Omega) : y \geq 0 \text{ a.e. on } \Omega\}$. For this sake, we use the arguments in Subsection 3.2.1 with homogenous Neumann boundary conditions instead. First, we observe, that the right hand side of the state equation is pointwise non-positive, as b is non-negative and B is a sign-preserving solution operator. Thus, y is non-positive, too. Next, introduce the concave, monotone superposition operator $\Phi(y) := (-\sinh(y))^+$ and formulate the modified equation

$$\begin{aligned}-\Delta y + \Phi(y) &= w \text{ in } \Omega,\\ \frac{\partial y}{\partial \nu} &= 0 \text{ on } \partial\Omega.\end{aligned}$$

By the above sign argument it is evident, that the solution does not change, when $\sinh$ is substituted with Φ. In order to show continuity and differentiability it is sufficient to show the Fréchet-differentiability of $\sinh : H^1(\Omega) \to L^2(\Omega)$ as the proof for Φ is analogous. Next, we will show, that the operator is Fréchet differentiable with first derivative $D\sinh(u)h = \cosh(u)h$. For this sake, consider the Taylor expansion for a sequence $h_n \to 0$ in $H^1(\Omega)$ reading as

$$\|\Phi(u + h_n) - \Phi(u) - D\Phi(u)h_n\|^2_{L^2} = \int_\Omega (\sinh(u + h_n) - \sinh(u) - \cosh(u)h_n)^2 \mathrm{d}x.$$

Define

$$\xi_n := \sinh(u + h_n) - \sinh(u) - \cosh(u)h_n,$$

then we rewrite ξ_n as

$$\xi_n = \sinh(u)(\cosh(h_n) - 1) + \cosh(u)(\sinh(h_n) - 1)$$

and obtain

$$\xi_n^2 \leq 2\sinh(u)^2(\cosh(h_n) - 1)^2 + 2\cosh(u)^2(\sinh(h_n) - h_n)^2.$$

For an arbitrary $t \in \mathbb{R}$ we derive the estimates

$$\begin{aligned}&0 \leq \cosh(t) - 1 \leq \int_0^t \sinh(s)\mathrm{d}s \leq |\sinh(t)||t|,\\ &0 \leq |\sinh(t) - t| \leq \left|\int_0^t (\cosh(s) - 1)\mathrm{d}s\right| \leq |\cosh(t) - 1||t| \text{ and}\\ &0 \leq \cosh(t) \leq \exp(|t|) \text{ and } |\sinh(t)| \leq \exp(|t|).\end{aligned}$$

Thus we get

$$\xi_n^2 \leq 2\sinh(u)^2\sinh(h_n)^2 h_n^2 + 2\sinh(u)^2\sinh(h_n)^2 h_n^4$$

and using the embeddings $H^1(\Omega) \hookrightarrow L^6(\Omega)$ and $H^1(\Omega) \hookrightarrow L^{12}(\Omega)$ with constants $C_6, C_{12} > 0$ we obtain

$$\begin{aligned}\int_\Omega \xi_n^2 \mathrm{d}x &\leq 2\left(\int_\Omega \sinh(u)^6 \mathrm{d}x\right)^{\frac{1}{3}} \left(\int_\Omega \sinh(h_n)^6 \mathrm{d}x\right)^{\frac{1}{3}} \left(\int_\Omega h_n^6 \mathrm{d}x\right)^{\frac{1}{3}} \\ &\quad + 2\left(\int_\Omega \sinh(u)^6 \mathrm{d}x\right)^{\frac{1}{3}} \left(\int_\Omega \sinh(h_n)^6 \mathrm{d}x\right)^{\frac{1}{3}} \left(\int_\Omega h_n^{12} \mathrm{d}x\right)^{\frac{1}{3}} \\ &\leq C(u)\left(\|h_n\|^2_{L^6(\Omega)} + \|h_n\|^4_{L^{12}(\Omega)}\right) \\ &\leq C(u)\left(C_6^2\|h_n\|^2_{H^1(\Omega)} + C_{12}^4\|h_n\|^4_{H^1(\Omega)}\right) \to 0 \text{ as } h_n \to 0 \text{ in } H^1(\Omega).\end{aligned}$$

This proves the Fréchet-differentiability.
Thus, we can deduce from the arguments in Subsection 3.2.1 the K-convexity with $Y = H^1(\Omega)$ equipped with $K := \{v \in H^1(\Omega) : v \geq 0 \text{ a.e. on } \Omega\}$ and $w \in H^1(\Omega)^*$ with $K_W = K^+$. Setting $w = -Bz - b$ yields the same solution as the original state equation, thus on $U_{\mathrm{ad}} \subseteq L^2_+(\Omega)$ both solution mappings S and the one induced by Φ coincide. Thus, the K-convexity is proven.
It is straightforward to see, that the mapping $\xi \to (\xi + 1)^{2+}$ is increasing and convex. Thus, defining $f(u) := \frac{\alpha}{2}\int_\Omega u^2 \mathrm{d}x$ as well as $g(y) := \int_{\Omega_o} (y+1)^{2+} \mathrm{d}x$ and $S : L^2(\Omega) \to H^1(\Omega)$ the solution operator $z \mapsto (-\Delta + \Phi)^{-1}(-Bz - b)$ we see the conditions discussed in Theorem 16 and Theorem 18 fulfilled. Thus, the above optimization problem is indeed convex. In fact, the application of one of our chain rules is not necessary and can be accomplished by standard analytical results and reads as

$$\begin{aligned} z - \operatorname{Proj}_{Z_{\mathrm{ad}}}\left(-\frac{1}{\alpha}B^* p\right) &= 0 && \text{in } \Omega, \\ -\operatorname{div}(\kappa \nabla y) + c\sinh(y) &= -Bz - b && \text{in } \Omega, \\ -\operatorname{div}(\kappa \nabla p) + c\cosh(p) &= (y+1)^+ \mathbb{1}_{\Omega_o} && \text{in } \Omega, \\ \frac{\partial y}{\partial \nu} = \frac{\partial p}{\partial \nu} &= 0 && \text{on } \partial\Omega. \end{aligned}$$

6 Conclusion

In this paper, we investigated a class of operators fulfilling a generalized, order-based convexity concept and their properties with regard to convex analysis and optimization theory. As part of we utilized and generalized methods from non-smooth and set-valued analysis and illustrated the applicability of these concepts to a selection of operator equations and variational inequalities closely related to the types of problems discussed in the recent literature.

A Introduction to capacity theory and capacitary measures

For the sake of selfcontainment of the present work, we collect some basic definitions and results regarding capacity theory. Our exposition is strongly inspired by the one in [RW19]. For more details besides references mentioned below we refer to [BS00], [EG15].

Definition 20. (cf. [BS00, Definition 6.4.7], [BMA06, Section 5.8.2, Section 5.8.3], [DZ11, Definition 6.4])

(i) For a subset $A \subseteq \Omega$ its *capacity* in the sense of $H_0^1(\Omega)$ is defined by

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

(ii) A subset $\hat{\Omega} \subseteq \Omega$ is called *quasi-open* if for all $\varepsilon > 0$ there exists and open set $O_\varepsilon \subseteq \Omega$ such that $\hat{\Omega} \cup O_\varepsilon$ is open and $\operatorname{cap}(O_\varepsilon) < \varepsilon$ holds.

(iii) A subset $\hat{\Omega} \subseteq \Omega$ is called *quasi-closed* if its relative complement $\Omega \backslash A$ is quasi-open.

(iv) A function $v : \Omega \to [-\infty, +\infty]$ is called *quasi-continuous (quasi lower semi-continuous, quasi upper semi-continuous)* if for all $\varepsilon > 0$ there is an open set $O_\varepsilon \subseteq \Omega$ with $\operatorname{cap}(O_\varepsilon) < \varepsilon$ such that v is continuous (lower semi-continuous, upper semi-continuous) on $\Omega \backslash O_\varepsilon$.

In the same fashion as with the Lebesgue measure a pointwise property of a function on Ω is called to hold *quasi everywhere* if it holds on subsets that differ from the whole domain only by a set of capacity zero.
For two Borel sets $E_0, E_1 \in \mathcal{B}(\Omega)$ such that E_0 is a subset of E_1 up to a set of capacity zero, we also write $E_0 \subseteq_q E_1$. If both $E_0 \subseteq_q E_1$ and $E_1 \subseteq_q E_0$ hold, then we might also write $E_0 =_q E_1$.

Lemma 21. (cf. [BS00, p. 564, 565] with [Rud87, Theorem 2.18] for (i),(ii); [HW18, Lemmata 3.5, 3.7], [Wac14, Lemma A.4] for (iii)) *Let $\xi \in H^{-1}(\Omega)$ with $\langle \xi, v \rangle \geq 0$ for all $v \in H_0^1(\Omega)$ with $v \geq 0$ a.e. on Ω be given.*

(i) The functional ξ can be identified with a regular Borel measure on Ω which is finite on compact sets and which possesses the following property: For every Borel set $E \subseteq \Omega$ with $\operatorname{cap}(E) = 0$, we have $\xi(E) = 0$.

(ii) Every function $v \in H_0^1(\Omega)$ is ξ-integrable and it holds

$$\langle \xi, v \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} = \int_\Omega v \, \mathrm{d}\xi.$$

(iii) There exists a quasi-closed set $\operatorname{f-supp}(\xi) \subseteq \Omega$ with the property that for all $v \in H_0^1(\Omega)$ with $v \geq 0$ a.e. it holds that $\langle \xi, v \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} = 0$ if and only if $v = 0$ q.e. on $\operatorname{f-supp}(\xi)$. The set $\operatorname{f-supp}(\xi)$ is uniquely defined up to a set of zero capacity and is called the fine support *of ξ.*

One is able to extend the definition of Sobolev spaces to quasi-open subsets $\hat{\Omega} \subseteq \Omega$ by

$$H_0^1(\hat{\Omega}) = \left\{ v \in H_0^1(\Omega) : \ v = 0 \text{ q.e. on } \Omega \backslash \hat{\Omega} \right\} \tag{12}$$

Definition 22. (cf. [DM87, Definition 2.1, 3.1]) Let $\mathcal{M}_0(\Omega)$ be the set of all Borel measures μ on Ω such that $\mu(E) = 0$ for every Borel set $E \subseteq \Omega$ with $\operatorname{cap}(E) = 0$ and such that μ is regular in the sense that $\mu(E) = \inf \{\mu(O) : O \text{ quasi-open }, E \subseteq_q O\}$. The set $\mathcal{M}_0(\Omega)$ is called the set of capacitary measures on Ω.

For a given capacitary measure $m \in \mathcal{M}_0(\Omega)$ and for a quasi-continuous function $v : \Omega \to \mathbb{R}$ we define the space

$$L^2_m(\Omega) := \left\{ v : \Omega \to \mathbb{R} : \int_\Omega |v|^2 \mathrm{d}m < +\infty \right\}. \tag{13}$$

Let $T_m \in \mathcal{L}(H^{-1}(\Omega), H_0^1(\Omega))$ denote the solution operator which maps a given $f \in H^{-1}(\Omega)$ to the solution of the following equation:

$$\int_\Omega \nabla y \nabla z \mathrm{d}x + \int_\Omega yz \,\mathrm{d}m = \langle f, z\rangle_{H^{-1},H_0^1} \text{ for all } z \in H_0^1(\Omega).$$

Definition 23. (cf. [DM87, Section 5], [RW19, Definition 3.2, Lemma 3.4]) Let a sequence of capacitary measures $(m_n)_{n\in\mathbb{N}} \subseteq \mathcal{M}_0(\Omega)$ be given. We say that $(m_n)_{n\in\mathbb{N}}$ *γ-converges* towards $m \in \mathcal{M}_0(\Omega)$ if the sequence of operators (T_{m_n}) converges in the *weak operator topology* towards T_m, i.e., for all $h \in H^{-1}(\Omega)$ holds $T_{m_n}h \rightharpoonup T_m h$ in $H_0^1(\Omega)$. If $(m_n)_{n\in\mathbb{N}}$ γ-converges to m we write $m_n \xrightarrow{\gamma} m$.

Lemma 24. (cf. [RW19, Corollary 3.5]) *The γ-convergence on $\mathcal{M}_0(\Omega)$ is metrizable with the metric*

$$d_{\mathcal{M}_0}(m, m') := \|T_m(1) - T_{m'}(1)\|.$$

Moreover, $(\mathcal{M}_0(\Omega), d_{\mathcal{M}_0})$ is a complete metric space.

Theorem 25. (cf. [DMM87, Proposition 4.14]) *Let $(m_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{M}_0(\Omega)$. Then there exists a subsequence $(m_{n_k})_{k\in\mathbb{N}}$ and a measure $m \in \mathcal{M}_0(\Omega)$ such that $m_{n_k} \xrightarrow{\gamma} m$.*

B Appendix

Proof of Lemma 1*.* ad (i): Let $x^* \in M_2^+$. Then, $\langle x^*, x\rangle \geq 0$ for all $x \in M_2$ and hence especially for all $x \in M_1$. This yields $x^* \in M_1^+$.
ad (ii): Since it always holds, that $M \subseteq \mathrm{cl}\,(M)$ we deduce $(\mathrm{cl}\,(M))^+ \subseteq M^+$ by (i). Let now $x^* \in M^+$ and take $x \in \mathrm{cl}\,(M)$. Then there exists a sequence $x_n \to x$ with $x_n \in M$ and we obtain $\langle x^*, x\rangle = \lim_{n\to\infty}\langle x^*, x_n\rangle \geq 0$ and hence the equality.
ad (iii): see [BS00, Proposition 2.40].
ad (iv): Since $0 \in M_1 \cap M_2$ we have that $M_j \subseteq M_1 + M_2$ and hence $(M_1 + M_2)^+ \subseteq M_j^+$ for $j = 1, 2$. This yields the inclusion $(M_1 + M_2)^+ \subseteq M_1^+ \cap M_2^+$.
Let, on the other hand, $x^* \in M_1^+ \cap M_2^+$. Then we get for all $x_j \in M_j$ that $\langle x^*, x_1 + x_2\rangle = \langle x^*, x_1\rangle + \langle x^*, x_2\rangle \geq 0$, which gives $x^* \in (M_1 + M_2)^+$.
ad (v): Since M_j^+ are closed, convex cones the set $\mathrm{cl}\left(M_1^+ + M_2^+\right)$ is a closed, convex cone as well. Hence, by the application of (ii), (iii) and (iv) we obtain, that

$$\left(\mathrm{cl}\left(M_1^+ + M_2^+\right)\right)^+ = \left(M_1^+ + M_2^+\right)^+ = M_1^{++} \cap M_2^{++} = M_1 \cap M_2.$$

The subsequent application of (ii) yields

$$\mathrm{cl}\left(M_1^+ + M_2^+\right) = \left(\mathrm{cl}\left(M_1^+ + M_2^+\right)\right)^{++} = (M_1 \cap M_2)^+.$$

□

References

[AF90] J.P. Aubin and H. Frankowska. *Set-Valued Analysis*. Modern Birkhäuser Classics. Springer, 1990.

[AF03] R.A. Adams and J.J.F. Fournier. *Sobolev Spaces*. ISSN. Elsevier Science, 2003.

[BdP84] V. Barbu and G. da Prato. *Optimal Control of Variational Inequalities*. Research notes in mathematics. Pitman Advanced Pub. Program, 1984.

[Bec08] R. Becker. *Ordered Banach Spaces*. Collection Travaux en cours : mathématiques. Hermann, 2008.

[BMA06] G. Buttazzo, G. Michaille, and H. Attouch. *Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization*. MPS-SIAM series on optimization. Society for Industrial and Applied Mathematics, 2006.

[BS00] J.F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer Series in Operations Research and Financial Engineering. Springer New York, 2000.

[BZ06] J. Borwein and Q. Zhu. *Techniques of Variational Analysis*. CMS Books in Mathematics. Springer New York, 2006.

[Cia13] P.G. Ciarlet. *Linear and Nonlinear Functional Analysis with Applications*. Other Titles in Applied Mathematics. Society for Industrial and Applied Mathematics, 2013.

[CLV13] J.P. Crouzeix, J.E.M. Legaz, and M. Volle. *Generalized Convexity, Generalized Monotonicity: Recent Results: Recent Results*. Nonconvex Optimization and Its Applications. Springer US, 2013.

[CMWC18] C. Christof, C. Meyer, S. Walther, and C. Clason. Optimal control of a non-smooth semilinear elliptic equation. *Mathematical Control & Related Fields*, 8:247, 2018.

[DM87] G. Dal Maso. Γ-Convergence and μ-Capacities. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 14(3):423–464, 1987.

[DMG94] G. Dal Maso and A. Garroni. New results on the asymptotic behavior of Dirichlet problems in perforated domains Dirichlet problems in perforated domains. *Mathematical Models and Methods in Applied Sciences*, 4(03):373–407, 1994.

[DMM87] G. Dal Maso and U. Mosco. Wiener's criterion and γ-convergence. *Applied Mathematics and Optimization*, 15(1):15–63, 1987.

[DZ11] M.C. Delfour and J.P. Zolesio. *Shapes and Geometries: Metrics, Analysis, Differential Calculus, and Optimization*. Advances in Design and Control. Society for Industrial and Applied Mathematics, Second edition, 2011.

[EG15] L.C. Evans and R.F. Gariepy. *Measure Theory and Fine Properties of Functions, Revised Edition*. Textbooks in Mathematics. Taylor & Francis, 2015.

[ET76] I. Ekeland and R. Temam. *Convex Analysis and Variational Problems*. Studies in Mathematics and its Applications. Elsevier Science, 1976.

[Har77] A. Haraux. 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):615–631, 1977.

[HK09] M. Hintermüller and I. Kopacka. Mathematical programs with complementarity constraints in function space: C-and strong stationarity and a path-following algorithm. *SIAM Journal on Optimization*, 20(2):868–902, 2009.

[HK11] M. Hintermüller and I. Kopacka. A smooth penalty approach and a nonlinear multigrid algorithm for elliptic mpecs. *Computational Optimization and Applications*, 50(1):111–145, 2011.

[HW18] F. Harder and G. Wachsmuth. Comparison of optimality systems for the optimal control of the obstacle problem. *GAMM-Mitteilungen*, 40(4):312–338, 2018.

[Jam74] G.J.O. Jameson. *Topology and Normed Spaces*. Chapman and Hall Mathematics Series. Chapman and Hall, 1974.

[KS80] D. Kinderlehrer and G. Stampacchia. *An Introduction to Variational Inequalities and Their Applications*. ISSN. Elsevier Science, 1980.

[KS20] D. P. Kouri and T. M. Surowiec. Risk-averse optimal control of semilinear elliptic PDEs. *ESAIM: Control, Optimisation and Calculus of Variations*, 26:53, 2020.

[LEG+12] G. Leugering, S. Engell, A. Griewank, M. Hinze, R. Rannacher, V. Schulz, M. Ulbrich, and S. Ulbrich. *Constrained Optimization and Optimal Control for Partial Differential Equations*. International Series of Numerical Mathematics. Springer Basel, 2012.

[LLPR96] Z.Q. Luo, Z.Q. Luo, J.S. Pang, and D. Ralph. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, 1996.

[LR08] Y. Li and B. Ruf. A Sharp Trudinger-Moser Type Inequality for Unbounded Domains in $\mathbf{R}^n$. *Indiana University Mathematics Journal*, pages 451–480, 2008.

[Mor06] B.S. Mordukhovich. *Variational Analysis and Generalized Differentiation I: Basic Theory*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2006.

[NST07] P. Neittaanmaki, J. Sprekels, and D. Tiba. *Optimization of Elliptic Systems: Theory and Applications*. Springer Monographs in Mathematics. Springer New York, 2007.

[OKZ98] J. Outrata, M. Kocvara, and J. Zowe. *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints: Theory, Applications and Numerical Results*. Nonconvex Optimization and Its Applications. Springer US, 1998.

[Rod87] J.F. Rodrigues. *Obstacle Problems in Mathematical Physics*. ISSN. Elsevier Science, 1987.

[Rou05] T. Roubicek. *Nonlinear Partial Differential Equations with Applications*. International Series of Numerical Mathematics. Birkhäuser Basel, 2005.

[Rud87] W. Rudin. *Real and Complex Analysis*. Higher Mathematics Series. McGraw-Hill, 1987.

[RW09] R.T. Rockafellar and R.J.B. Wets. *Variational Analysis*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2009.

[RW19] A.-T. Rauls and G. Wachsmuth. Generalized Derivatives for the Solution Operator of the Obstacle Problem. *Set-Valued and Variational Analysis*, 2019.

[Sca59] H. Scarf. The Optimality of (S, s) Policies in the Dynamic Inventory Problem. *Optimal pricing, inflation, and the cost of price adjustment*, 1959.

[Sch74] H.H. Schäfer. *Banach Lattices and Positive Operators*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 1974.

[Sch07] W. Schirotzek. *Nonsmooth Analysis*. Universitext. Springer Berlin Heidelberg, 2007.

[TS10] F. Tröltzsch and J. Sprekels. *Optimal Control of Partial Differential Equations: Theory, Methods, and Applications*. Graduate studies in mathematics. American Mathematical Society, 2010.

[Wac14] G. Wachsmuth. Strong Stationarity for Optimal Control of the Obstacle Problem with Control Constraints. *SIAM Journal on Optimization*, 24(4):1914–1932, 2014.

[Wac16] G. Wachsmuth. Optimization problems with complementarity constraints in infinite-dimensional spaces, 2016. Habilitation Thesis.