# First Order Optimality Conditions for Elliptic Mathematical Programs with Equilibrium Constraints via Variational Analysis

M. Hintermüller* and T. Surowiec†

**Abstract**

Mathematical programs in which the constraint set is partially defined by the solutions of an elliptic variational inequality, so-called "elliptic MPECs", are formulated in reflexive Banach spaces. With the goal of deriving explicit first order optimality conditions amenable to the development of numerical procedures, variational analytic concepts are both applied and further developed. The paper is split into two main parts. The first part concerns the derivation of conditions in which the state constraints are assumed to be polyhedric sets. This part is then completed by two examples, the latter of which involves pointwise bilateral bounds on the gradient of the state. The second part begins with the derivation of a formula for the second order (Mosco) epiderivative of the indicator function of a general convex set. This result is then used to derive analogous conditions to those which are presented in the first part. Finally, an elliptic MPEC is considered important to the study of elasto-plasticity in which the pointwise Euclidean norm of the gradient of the state is bounded. Explicit strong stationarity conditions are provided for this problem.

## 1 Introduction

The mathematical modeling of real-world phenomena often leads to infinite dimensional, i.e., function space, problem formulations containing variational inequalities. For example, certain problems in elasticity [25], elasto-plasticity [17, 24], and mathematical finance [1] all lead to models in which a variational inequality arises. In addition, minimization problems involving certain classes of non-smooth functionals result in variational inequalities via Fenchel-Legendre dualization and associated Euler-Lagrange conditions [14]. Due to their practical relevance, many research efforts have been devoted to the study of variational inequalities and their numerical solution since their conception, see e.g., [15, 16, 26] and the references therein.

Frequently, one is often interested in controlling the solution of a variational inequality in order to achieve a desired state or to minimize a target quantity. On an abstract level this leads to minimization problems of the type

$$\begin{aligned} &\text{minimize} && J(u,y) \quad \text{over } (u,y) \in \mathcal{U} \times Y \\ &\text{subject to (s.t.)} && y \in S(u), \end{aligned} \tag{1}$$

where $(u, y)$ denotes the associated control-state pair with respective control space $\mathcal{U}$ and state space $Y$, $J$ is a sufficiently smooth objective function, and $S : \mathcal{U} \to Y$ represents the solution operator of the underlying variational inequality. Problems of type (1) are sometimes called mathematical

*Dept. of Mathematics, Humboldt University of Berlin, E-Mail:hint@math.hu-berlin.de and Dept. of Mathematics and Scientific Computing, Karl-Franzens-University of Graz.

†Dept. of Mathematics, Humboldt University of Berlin, E-Mail:surowiec@math.hu-berlin.de

programs with equilibrium constraints (MPECs), as the variational inequality often represents an equilibrium condition, e.g., first-order optimality conditions of a convex optimization problem. Although the literature on finite dimensional MPECs has reached a certain level of sophistication, as evidenced by the monographs [28, 32, 34] and the many references therein, far less is known about MPECs in function spaces. With respect to the latter, we mention [6, 33] and the selected papers [7, 8, 9, 20, 23, 30]. We would also like to mention that parameter identification problems for variational inequalities lead to problems of type (1); see, e.g., [18, 19] and references therein.

From a mathematical optimization point of view the difficulties associated with (1) result from a lack of constraint regularity, which in turn prevents the application of well-known results for mathematical programs in Banach space; see, e.g., [41]. Moreover, upon reducing (1) to a problem in $u$ by considering $y = y(u) = S(u)$, following the so-called implicit programming approach, the problem typically becomes a non-smooth and non-convex problem, which is then hard to tackle analytically as well as numerically. In particular, the explicit representation of first order optimality conditions suitable for numerical realization remains an issue.

In a recent work motivated by similar results in finite dimensions as found in [39], an attempt at systematizing stationarity conditions for function-space-based problems of the type (1) was undertaken in [20]. Remarkably, versions of weak, C, and strong stationarity were derived that paralleled the concepts in finite dimensions; and whereas many approaches applied in the past relied on penalization techniques, the method of [20] utilizes a relaxation approach yielding stronger stationarity conditions than those resulting from penalty techniques. Penalization or relaxation techniques have the advantage that they readily facilitate the application of the well-established theory on mathematical programs in Banach space for the existence of Lagrange multipliers. Moreover, these techniques may be turned into algorithmic frameworks by closely following the derivation of first order optimality systems for the MPEC. Using these techniques, as was demonstrated in [20, 21], makes the problems amenable to the application of fast solvers such as semismooth Newton and multigrid methods.

Despite the appeal of penalization and regularization methods, variational analysis (see e.g., [5, 38, 31]) provides a different set of analytical tools able to directly derive sharp, i.e., strong, stationarity conditions without needing to pass to the limit with respect to certain parameters that arise in the relaxation/penalty approaches. Oftentimes one needs only to verify that the data of the underlying problem satisfy a certain constraint qualification in order to derive first order optimality conditions, thereby providing a means for avoiding relaxation/penalization techniques and limit processes in a problem-dependent fashion.

In this respect, the aim of this work is two-fold: (i) We utilize and extend tools from variational analysis in order to derive an abstract first order optimality system (in the sense of strong stationarity) for a rather broad class of control problems of elliptic variational inequalities. (ii) We treat systems involving (pointwise) gradient constraints of the type

$$M := \left\{ y \in H_0^1(\Omega) \,|\, |\nabla y| \le \psi, \text{almost everywhere (a.e.) on } \Omega \right\}.$$

Gradient constraints in function space MPECs have yet to be treated in the literature, despite having been mentioned as an important (open) problem class in [30]. For control of obstacle-type problems (i.e., with pointwise unilateral constraints on the state, rather than its gradient) it turns out that we recover the strong stationarity conditions derived earlier by Mignot and Puel in [30], who applied a mix of penalization techniques and conical derivatives.

The rest of the paper is organized as follows. In section 2, we introduce notation and basic concepts. The remaining sections subdivide our investigation into a polyhedric setting (Sections 3–5) and a more general setting (Sections 6–8) with respect to the constraints on the state. Section 3 is devoted to studying differentiability properties of the control-to-state mapping, i.e. the solution operator of the underlying variational inequality. Our results extend those obtained by Mignot in his fundamental paper [29] on the conical derivative of the solution operator associated with the obstacle problem. In section 4, these results are applied to derive strong stationarity conditions for the MPEC, and in section 5 two case studies are performed yielding first the well-known stationarity result of Mignot and Puel [30] (here in the sense of a validation of our technique) and then explicit strong stationarity conditions in the presence of pointwise constraints on the gradient of the state. The extension to the general case is more delicate as the variational arguments in the polyhedric case do not immediately apply. Hence, in section 6, we establish a new result on the second-order Mosco epiderivative of the indicator function of a closed convex set. This result enables us to derive strong stationarity in the general case in section 7. Finally, in section 8 we use our abstract results to calculate explicit strong stationarity conditions for the case of pointwise constraints on the 2-norm of the gradient of the state.

# 2 Notation and Basic Concepts from Variational Analysis

Throughout the text we make significant use of certain objects that are more or less standard in the literature. New or lesser known concepts are introduced throughout the text so that they may be better understood in context.

## 2.1 Assumptions and Notation

Within the entirety of this paper, we will only consider real Banach spaces and we make the additional assumption that the topologies of some Banach space $X$ along with its topological dual $X^*$ are compatible. If $X$ is in addition reflexive, then the strong topologies on both spaces are considered, otherwise we assume $X^*$ is equipped with the weak$^*$-topology, so that the dual of $X^*$ is isometric to $X$. For more on this subject, the reader is referred to any standard reference on Functional Analysis, e.g., [40]. We denote the dual pairing between $X$ and $X^*$ by $\langle\cdot,\cdot\rangle_{X^*,X}$ and denote strong convergence in any space, e.g., $X$, via the symbol "$\rightarrow_X$" and weak convergence by "$\rightharpoonup_X$". The embedding of a space $X$ into $Y$ is denoted $X \hookrightarrow Y$. If $X$ is an inner product space, then the inner product will be denoted by $(\cdot,\cdot)_X$ and the norm defining the topology on $X$ is denoted by $||\cdot||_X$. We denote the closure in the topology on $X$ by $\mathrm{cl}\{\cdot\}_X$. In all cases, we leave off the subscript "$X$" if it is clear in context. Finally, if $x, y \in \mathbb{R}^l$, then $x \cdot y$ represents their scalar product and for any subset $A \subseteq \mathbb{R}^l$, we use "a.e. $A$" to represent "almost everywhere on $A$".

## 2.2 A Few Important Function Spaces

At some points in this paper, we provide examples in which certain function spaces are present. We always assume that the subset $\Omega \subseteq \mathbb{R}^l$ is a bounded open subset with Lipschitz boundary $\partial\Omega$ and let $l \geq 1$. We denote the standard Lebesgue space of square integrable functions/vector fields by $L^2(\Omega)^l$, leaving off the "$l$"-subscript if $l = 1$; and we denote the space of all infinitely differentiable functions whose (compact) support is contained in $\Omega$ by $C_0^\infty(\Omega)$. We then define the Sobolev space $H_0^1(\Omega)$ as the completion of the space $C_0^\infty(\Omega)$ with respect to the norm

$$||x||_{H_0^1(\Omega)} := ||\nabla x||_{L^2(\Omega)^l}.$$

The usage of this norm for defining $H_0^1(\Omega)$ follows from the boundedness of $\Omega$. Here, the gradient of $x$ is understood in a weak sense. Finally, we denote the *pointwise* $\infty$-norm and 2-norm of the gradients of $H_0^1$-functions for any $\omega \in \Omega$ by

$$|\nabla y(\omega)|_\infty := \max_{1\le i\le l} |\nabla y_i(\omega)| \quad \text{and} \quad |\nabla y(\omega)|_2 := \left(\sum_{i=1}^{l} \nabla y_i(\omega)^2\right)^{1/2},$$

respectively. When it is clear in context, we leave off the arguments "$\omega$". For more on Sobolev spaces we refer the reader to [2].

### 2.3 Variational Analytic Concepts

Throughout this subsection we assume that $X$ is some arbitrary Banach space paired with its dual $X^*$ and let $C \subseteq X$ be a nonempty closed convex set. In addition, we define the indicator function of $C$

$$I_C(x) := \begin{cases} 0, & y \in C, \\ +\infty, & y \notin C. \end{cases}$$

The *tangent cone* to $C$ at $x \in C$ is defined by

$$T_C(x) := \left\{ d \in X \,\middle|\, \exists t_k \to 0^+, \exists d_k \to_X d : x + t_k d_k \in C,\ \forall k \right\}.$$

In the event that $C$ is only closed but not convex, then we refer to this cone as the *contingent cone*. As we will see in a moment, the tangent cone can be derived via calculating the polar cone to another variational object. This however, is not true when $C$ is not convex.

Given another arbitrary Banach space $Y$, we refer to any mapping $F$ from $X$ into the set of subsets of $Y$ as a *multifunction* or *set-valued mapping*. We use the notation $F : X \rightrightarrows Y$ to denote that $F$ is a multifunction. The *graph* of a multifunction is defined by

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

Clearly, $\operatorname{gph} F \subset X \times Y$. Though multifunctions are very different from single-valued mappings, we can still define (generalized) derivatives using contingent cones. Accordingly, we define the *contingent derivative* of $F$ at a point $(x, y) \in \operatorname{gph} F$ to be the mapping $DF[(x, y)] : X \rightrightarrows Y$ whose graph equals $T_{\operatorname{gph} F}(x, y)$, i.e.,

$$w \in DF[(x, y)](u) \Leftrightarrow (u, w) \in T_{\operatorname{gph} F}(x, y) \Leftrightarrow \begin{cases} \exists t_k \to 0^+, \exists u_k \to_X u, \exists w_k \to_Y w : \\ y + t_k w_k \in F(x + t_k u_k). \end{cases}$$

For more on these and related concepts, see e.g., [5]. Another important object in our study is the so-called *normal cone*. The normal cone to a closed convex set $C \subseteq X$ at some point $x \in C$ is defined by

$$N_C(x) := \left\{ x^* \in X^* \,\middle|\, \langle x^*, x' - x\rangle_{X^*,X} \le 0, \forall x' \in C \right\}.$$

In addition, $N_C(x)$ can also be (equivalently) defined by

$$N_C(x) := \{x^* \in X^* \,|\, \langle x^*, d\rangle_{X^*,X} \le 0, \forall d \in T_C(x)\} = [T_C(x)]^-,$$

i.e., as the negative polar/dual cone of the tangent cone. Given that $X$ and $X^*$ are paired spaces, it holds that

$$[N_C(x)]^- = \left[[T_C(x)]^-\right]^- = T_C(x),$$

as $T_C$ itself is a closed convex cone. When $C$ is merely closed and not convex, we can define another type of normal cone known as the *Fréchet normal cone*

$$\widehat{N}_C(x) := \left\{ x^* \in X^* \left| \limsup_{x' \to_X x} \frac{\langle x^*, x' - x \rangle}{||x' - x||_X} \leq 0 \right. \right\}.$$

Note that the lim sup in the previous definition must hold for *all* sequences $x' \to_X x$. This often makes the Fréchet normal cone extremely difficult to calculate explicitely. Nevertheless, Theorem 1.10 in [31] provides the following upper approximation

$$\widehat{N}_C(y) \subset [T_C(y)]^- ,$$

where equality holds if $X$ is reflexive and one considers weak limits in the definition of $T_C$ or $X$ is finite dimensional.

Finally, we define the *subdifferential* of a convex lower-semicontinuous function $f : X \to \bar{\mathbb{R}}$ at $x \in \operatorname{dom} f$ by

$$\partial f(x) := \left\{ x^* \in X^* \left| \langle x^*, x' - x \rangle + f(x) \leq f(x'),\ \forall x' \in X \right. \right\}.$$

Note that for any non-empty closed convex set $C \subseteq X$, $\partial I_C(x) = N_C(x)$. We direct the reader to [31] for more information concerning the dual variational objects.

# 3 Generalized Differentiation of the Solution: Polyhedric Case

In this section, we demonstrate the application of some known theoretical results and obtain a general formula for the contingent derivative of the solution of the variational inequality in the case where the set of state constraints is polyhedric.

Throughout this section, we define $S$ to be the solution mapping of the variational inequality

$$u \in Ay + N_M(y),$$

where $A$ is a coercive bounded linear operator from $Y$ into $Y^*$, i.e., there exists a $\xi \in \mathbb{R}_+ \setminus \{0\}$ such that

$$\langle Ay, y \rangle_{Y^*,Y} \geq \xi ||y||_Y^2, \forall y \in Y.$$

In addition, we define $M$ to be a closed convex subset of the reflexive Banach space $Y$ and $u \in Y^*$. By referring to certain classical results, e.g., Chapter 3 in [26], we see that $S$ is in fact a single-valued and locally Lipschitz function of $u$.

In order to provide a formula for the contingent derivative of $S$, we need a way of characterizing the contingent derivative of $N_M$. We begin with the case of so-called *polyhedric* sets $M$.

**Definition 3.1** (**Polyhedric Sets**). *A closed convex set $C$ of a Banach space $X$ is called polyhedric if for all $x \in C$*

$$T_C(x) \cap \{v\}^\perp = \operatorname{cl} \left\{ R_C(y) \cap \{v\}^\perp \right\}_X ,$$

*where $R_C(y)$ represents the so-called radial cone and is defined by*

$$R_C(y) := \{ h \in X \,| \exists \tau^* > 0 : \forall \tau \in [0, \tau^*],\, y + \tau h \in C \}$$

*and $v \in N_C(y)$.*

Note that in general, the tangent cone to $M$ contains the radial cone, in fact, it can also be defined as the closure of $R_M$. We will later provide two (non-trivial) examples containing polyhedric sets, but first we state the following important result due to Levy [27].

**Theorem 3.1** (**[27], Theorem 3.1**). *Let $M$ be a polyhedric subset of some reflexive Banach space $Y$ and $v \in N_M(y)$. Then for any $(y, v) \in \operatorname{gph} N_M$, the following are equivalent*

1. $w \in DN_M[(y, v)](d)$.
2. $w \in N_{\mathcal{K}(y,v)}(d)$.
3. $(d, w) \in \mathcal{K}(y, v) \times [\mathcal{K}(y, v)]^- : \langle w, d \rangle = 0$.

*Here, $\mathcal{K}(y, v) := T_M(y) \cap \{v\}^\perp$, i.e., the critical cone.*

We note that the original theorem was stated in terms of so-called proto-derivatives of $N_M$, see Definition 6.4 in Section 6, however, we later argue that this object coincides with the contingent derivative. We would also like to bring the reader's attention to the fact that Theorem 3.1 can be can be easily derived by using Equation (2.13) in Example 2.10 together with Theorem 3.9 in [13].

By appealing to a special calculus rule for the contingent derivatives of certain classes of multifunctions, we prove the next important corollary.

**Corollary 3.1** (**The Contingent Derivative of $S$**). *Let $M$ be a polyhedric subset of a reflexive Banach space $Y$ and $S$ be as above. If $(u, y) \in \operatorname{gph} S$, then the following are equivalent:*

1. $d \in DS[(u, y)](w)$.
2. $w \in Ad + N_{\mathcal{K}(y,v)}(d)$.
3. $w - Ad \in [\mathcal{K}(y, v)]^-$, $d \in \mathcal{K}(y, v)$, $\langle w - Ad, d \rangle = 0$.

*or equivalently*

$$T_{\operatorname{gph} S}(u, y) = \left\{(w, d) \in Y^* \times Y \,\middle|\, w - Ad \in [\mathcal{K}(y, v)]^-,\ d \in \mathcal{K}(y, v),\ \langle w - Ad, d \rangle = 0 \right\}.$$

*Here, $v := u - Ay$.*

*Proof.* Define

$$S^{-1}(y) := \{u \in Y^* \,|\, u \in Ay + N_M(y)\}.$$

It can be easily derived from the definition of the contingent derivative ( see e.g., Chapter 4 in [5]) that

$$d \in DS[(u, y)](w) \Leftrightarrow w \in DS^{-1}[(y, u)](d).$$

Then since $A$ is Fréchet differentiable from $Y$ to $Y^*$, we can refer to Proposition 5.1.2 in [5], which states

$$w \in DS^{-1}[(y, u)](d) = Ad + DN_M[(y, u - Ay)](d).$$

The rest follows from Theorem 3.1. ☐

**Remark 3.1** (**Contingent Derivatives vs. Conical Derivatives**). In his seminal 1976 paper [29], Mignot introduces a type of one-sided directional derivative for continuous mappings between Banach spaces called the *conical derivative*. Perhaps the most stunning result concerning these derivatives is found in Theorem 3.3 ([29]), where solutions to a specific class of variational inequalities are shown to admit a conical derivative for every perturbation parameter (similar to $w$ in Corollary 3.1). In turn, the conical derivative is the solution of a variational inequality of the type found in Corollary 3.1. However, Mignot does so only for a restricted class of function spaces and choice of $M$. In this sense, Corollary 3.1 extends Mignot's result to reflexive Banach spaces for all state constraints $M$, provided the sets are polyhedric. Indeed, 2. in Corollary 3.1 can be viewed as the necessary and sufficient optimality conditions to the optimization problem

$$\min_{d \in Y} \left\{ \frac{1}{2}\langle Ad, d\rangle - \langle w, d\rangle + I_{\mathcal{K}(y,v)}(d) \right\}.$$

Since this objective function is strictly convex, coercive and lower semicontinuous, the optimization problem always has a unique solution, see. e.g., Theorem 3.3.4 in [4]; this holds regardless of the choice of $w \in Y^*$. Moreover, it is easy to ascertain the Lipschitz continuity of these solutions as functions of $w$ using the coercivity of the operator $A$. Hence, for any fixed $(u, y) \in \operatorname{gph} S$, i.e., a solution $y$ of the original variational inequality for a given $u$, we see that at every point $w \in Y^*$, this solution admits a contingent derivative $DS[(u, y)](w) = d$, where $d$ solves the variational inequality:

$$\text{Find } d \in \mathcal{K}(y, v) : \langle Ad, d' - d\rangle \geq \langle w, d' - d\rangle, \ \forall d' \in \mathcal{K}(y, v).$$

In Corollary 6.2, we fully generalize this result to hold for *all* closed convex sets of state constraints $M$ in reflexive Banach spaces.

Continuing, we have the following important result, which incidentally can be derived via arguments in Chapter 5 of [5] (Proposition 5.1.3 with "$f(x) := 0$"). Nevertheless, we provide an alternate proof for completeness. Keep note of how the result follows without the requirement that $M$ is polyhedric.

**Proposition 3.1** (**Domain-Constrained Solution Mappings**). *Assume that $Y$ is a reflexive Banach space and $\mathcal{U}$ a Hilbert space such that $Y \hookrightarrow \mathcal{U} \hookrightarrow Y^*$, where the embedding of $\mathcal{U}$ (identified with its dual $\mathcal{U}^*$) in $Y^*$ is dense. If $M \subset Y$ is nonempty, closed, and convex, then for any $(u, y) \in [\mathcal{U} \times Y] \cap \operatorname{gph} S$ it holds that*

$$T_{[\mathcal{U}\times Y]\cap \operatorname{gph} S}(u, y) = T_{\operatorname{gph} S}(u, y).$$

*Proof.* Throughout the proof, we identify $\mathcal{U}$ with its dual, making it a proper subspace of $Y^*$. Therefore, $\mathcal{U}$ is formally a convex set. Moreover, since the subspace $\mathcal{U}$ is dense in $Y^*$ and the convergence of $w_k$ is taken in $Y^*$, we see that $T_{\mathcal{U}}(u) = Y^*$.

Begin by letting $(w, d) \in T_{[\mathcal{U}\times Y]\cap \operatorname{gph} S}(u, y)$. Then by definition there exist sequences $t_k \to 0^+$, $w_k \to_{Y^*} w$, and $d_k \to_Y d$ such that

$$u + t_k w_k \in \mathcal{U} \wedge y + t_k d_k \in S(u + t_k w_k), \ \forall k.$$

It follows that $(w, d) \in [T_{\mathcal{U}}(u) \times Y] \cap T_{\operatorname{gph} S}(u, y) = T_{\operatorname{gph} S}(u, y)$.

Consider now that since $S$ is locally Lipschitz near $u$, there exists a scalar $L > 0$ and a neighborhood $\mathcal{W}$ of $u$ such that

$$S(u') \subset S(u'') + L||u' - u''||_{Y^*}\mathbb{B}, \ \forall u', u'' \in \mathcal{W}.$$

Here, $\mathbb{B}$ represents the open unit ball in $Y$. Now let $(w, d) \in [T_{\mathcal{U}}(u) \times Y] \cap T_{\operatorname{gph} S}(u, y)$. Then there exist sequences $t_k \to 0^+$, $w_k \to_{Y^*} w$, $w'_k \to_{Y^*} w$, and $d'_k \to_Y d$ such that

$$u + t_k w_k \in \mathcal{U} \wedge y + t_k d'_k \in S(u + t_k w'_k), \ \forall k.$$

For large $k$ we have from the Lipschitz continuity of $S$ that

$$y + t_k d'_k \in S(u + t_k w'_k) \subset S(u + t_k w_k) + t_k L||w'_k - w_k||_{Y^*}\mathbb{B}$$

Due to the single-valuedness of the solution mapping $S$ near $u$, we obtain a bounded sequence $b_k \in \mathbb{B}$ such that for $k$ large

$$y + t_k d'_k = S(u + t_k w_k) + t_k L||w'_k - w_k||_{Y^*} b_k.$$

Then by defining $d_k := d'_k - t_k L||w'_k - w_k||_{Y^*} b_k$, we obtain a sequence $d_k \to d$ such that for large enough $k$

$$u + t_k w_k \in \mathcal{U} \wedge y + t_k d_k \in S(u + t_k w_k),$$

Hence, $(w, d) \in T_{[\mathcal{U} \times Y] \cap \operatorname{gph} S}(u, y)$. The assertion follows. □

The previous result has wide-ranging consequences that will become clear as soon as we present the examples in the coming sections. We close this section with an additional result needed for the derivation of dual first order optimality conditions for the class of elliptic MPECs with polyhedric state constraints and subspace constraints on the control.

**Proposition 3.2 (Approximation of the Fréchet Normal Cone).** *Given the setting from Proposition 3.1 with $(u, y) \in [\mathcal{U} \times Y] \cap \operatorname{gph} S$, the following holds*

$$\widehat{N}_{[\mathcal{U} \times Y] \cap \operatorname{gph} S}(u, y) \subseteq \left\{ (p^*, q^*) \in Y \times Y^* \,\middle|\, p^* \in \mathcal{K}(y, v),\ q^* \in -A^* p^* + [\mathcal{K}(y, v)]^- \right\}.$$

*Here, $v := u - Ay$.*

*Proof.* From Theorem 1.10 in [31] and Proposition 3.1, we know that

$$\widehat{N}_{[\mathcal{U} \times Y] \cap \operatorname{gph} S}(u, y) \subseteq [T_{[\mathcal{U} \times Y] \cap \operatorname{gph} S}(u, y)]^- = [T_{\operatorname{gph} S}(u, y)]^-.$$

Passing now to the definition of the polar cone we see that

$$[T_{\operatorname{gph} S}(u, y)]^- = \{ (p^*, q^*) \in Y \times Y^* \,|\, \langle w, p^* \rangle + \langle q^*, d \rangle \le 0, \forall (w, d) \in T_{\operatorname{gph} S}(u, y) \}.$$

Using Corollary 3.1, we can proceed in a manner similar to the proof of Lemma 3.1 in [35].

$$\begin{aligned}
[T_{\operatorname{gph} S}(u, y)]^- &= \{ (p^*, q^*) \in Y \times Y^* \,|\, \langle w, p^* \rangle + \langle q^*, d \rangle \le 0, \forall (w, d) \in T_{\operatorname{gph} S}(u, y) \} \\
&= \left\{ (p^*, q^*) \in Y \times Y^* \,\middle|\, \langle Ad + h, p^* \rangle + \langle q^*, d \rangle \le 0, \forall (d, h) \in \operatorname{gph} N_{\mathcal{K}(y, v)} \right\} \\
&= \left\{ (p^*, q^*) \in Y \times Y^* \,\middle|\, \langle A^* p^* + q^*, d \rangle + \langle h, p^* \rangle \le 0, \forall (d, h) \in \operatorname{gph} N_{\mathcal{K}(y, v)} \right\}
\end{aligned}$$

Recall from Theorem 3.1 that

$$\operatorname{gph} N_{K(y, v)} = \left\{ (d, h) \in Y \times Y^* \,\middle|\, d \in \mathcal{K}(y, v),\ h \in [\mathcal{K}(y, v)]^-,\ \langle h, d \rangle = 0 \right\}.$$

Then by ignoring the complementarity relation $\langle h, d \rangle = 0$, we observe that

$$[T_{\operatorname{gph} S}(u, y)]^- \supset \left\{ (p^*, q^*) \in Y \times Y^* \,\middle|\, A^* p^* + q^* \in [\mathcal{K}(y, v)]^-,\ p^* \in \mathcal{K}(y, v) \right\}.$$

Conversely, since $\mathcal{K}(y, v) \times \{0\}$ and $\{0\} \times [\mathcal{K}(y, v)]^-$ are subsets of $\operatorname{gph} N_{K(y, v)}$ we obtain the reverse inclusion, i.e.,

$$[T_{\operatorname{gph} S}(u, y)]^- \subset \left\{ (p^*, q^*) \in Y \times Y^* \,\middle|\, A^* p^* + q^* \in [\mathcal{K}(y, v)]^-,\ p^* \in \mathcal{K}(y, v) \right\},$$

as was to be shown. □

## 4 Optimality Conditions for the Elliptic MPEC: Polyhedric Case

Using the results of the previous section, we derive what amount to so-called *strong* or *S-stationarity conditions* (cf. [39]) for elliptic MPECs. The choice of nomenclature is done so in an effort to unify terminology with the classical MPEC literature. Though the abstract conditions below do not directly reflect the sign conditions normally associated with S-stationarity conditions, the examples in the following sections do in fact exhibit this behavior.

Consider the following class of elliptic MPECs

$$\begin{cases} \min J(u,y) \\ \text{s.t. } u \in \mathcal{U},\ y \in S(u). \end{cases} \tag{2}$$

Here, we assume as before that $S$ is the solution mapping of some variational inequality of the form

$$u \in Ay + N_M(y),$$

where $A : Y \to Y^*$ is a coercive bounded linear operator and $M \subseteq Y$ is polyhedric. In addition, we assume that $Y$ is a subspace of $\mathcal{U}$, where $\mathcal{U}$ is some Hilbert space densely embedded in $Y^*$ (upon identification with its dual $\mathcal{U}^*$. Finally, we let $J$ be Fréchet differentiable from $\mathcal{U} \times Y$ to $\mathbb{R}$.

**Theorem 4.1** (**Existence of a Solution to the Elliptic MPEC**). *In addition to the assumptions above, we assume that the objective functional $J$ satisfies the following additional criteria.*

1. $\forall u_k \rightharpoonup_{\mathcal{U}} u$, $\forall y_k \rightharpoonup_Y y$, $\liminf_k J(u_k, y_k) \geq J(u,y)$ *(weak lower-semicontinuity).*
2. $\exists c_1, c_2 \in \mathbb{R} : \forall y \in Y,\ J(u,y) \geq c_1 ||u||_{\mathcal{U}} + c_2$. *(partial coercivity).*
3. $\exists K > 0 : J(u,y) \geq K > -\infty\ \forall u, y \in \mathcal{U} \times Y$ *(boundedness from below).*

*Then the elliptic MPEC (2) possesses at least one solution $(\bar{u}, \bar{y})$, provided the embedding "$\mathcal{U} \hookrightarrow Y^*$" is compact.*

*Proof.* The proof is standard, however, in the interest of completeness, we provide it in the appendix. □

**Remark 4.1** (**Strength of the Assumptions**). *In the proof of Theorem 4.1, we do note require $M$ to be polyhedric. In fact, it is enough that $M$ be closed and convex. Furthermore, we would require only a slight change of proof if we were to restrict $u$ to a closed convex subset of $\mathcal{U}$. Thus, the result provides an existence proof for a large class of elliptic MPECs.*

*One example of an objective functional that satisfies the needed requirements is*

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

*where $\alpha \geq 0$ and $z_d \in L^2(\Omega)$ with $Y = H^1_0(\Omega)$, $Y^* = H^{-1}(\Omega)$, and $\mathcal{U} = L^2(\Omega)$. Since $\Omega$ is bounded by assumption, $H^1_0(\Omega)$ is densely embedded into $L^2(\Omega)$, which in turn implies that $L^2(\Omega)$ is densely embedded into $H^{-1}(\Omega)$, as $L^2(\Omega)$ is a Hilbert space. Moreover, the boundedness also implies the compactness of the embedding $L^2(\Omega) \hookrightarrow H^{-1}(\Omega)$.*

We now present our first main result providing dual optimality conditions for (2).

**Theorem 4.2 (Strong Stationarity Conditions I).** *Let $(\bar{u}, \bar{y})$ be a locally optimal solution of (2), then there exist multipliers $p^* \in \mathcal{K}(\bar{y}, \bar{v})$, $r^* \in [\mathcal{K}(\bar{y}, \bar{v})]^-$, and $\bar{v} \in N_M(\bar{y})$ such that*

$$0 = \nabla_u J(\bar{u}, \bar{y}) + p^*, \tag{3}$$

$$0 = \nabla_y J(\bar{u}, \bar{y}) + r^* - A^* p^*, \tag{4}$$

$$0 = A\bar{y} - \bar{u} + \bar{v}. \tag{5}$$

*Proof.* The result follows from Proposition 5.1 in [32] given the Fréchet differentiability of $J$ on $\mathcal{U} \times Y$ and using the estimate from Proposition 3.2. □

# 5 Two Examples with Polyhedric State Constraints $M$

In this section, we provide examples in which the conditions (3) and (4) from Theorem 4.2 are made explicit. We begin with a classical example for illustration, following which we derive new conditions for an important example from the study of elasto-plasticity (cf. [24, 25]).

## 5.1 Optimal Control of the Obstacle Problem

In [29], Mignot demonstrates the polyhedricity of constraint sets of the type

$$\left\{ y \in H_0^1(\Omega) \,|\, \varphi \leq y \leq \psi, \ a.e.\, \Omega \right\},$$

where $\psi, \varphi \in H^1(\Omega)$ are appropriately chosen. Furthermore, many other obstacle-problem-type constraints sets, i.e., box-constraints, in $L^p$-spaces are shown to be polyhedric in Chapter 6 of [10]. Using the optimality conditions derived in the previous section, along with the well-known characterizations for both the associated critical cone and its dual, we quickly rederive the well-known conditions of Mignot and Puel [30]. As their conditions are considered to be the best possible for the optimal control of the obstacle problem, we consider this brief derivation as a type of validation for the optimality of our conditions.

In the following,

- $Y := H_0^1(\Omega)$,
- $\mathcal{U} := L^2(\Omega)$,
- $M := \left\{ y \in H_0^1(\Omega) \,|\, y \geq 0, \ a.e.\, \Omega \right\}$,
- $J : L^2(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$, Fréchet differentiable.

For the definitions and assumptions on these spaces, we refer the reader to Section 2.

By Theorem 4.2, if $(\bar{u}, \bar{y})$ is a locally optimal solution of the associated elliptic MPEC, then there exist $(p^*, r^*) \in H_0^1(\Omega) \times H^{-1}(\Omega)$ such that

$$\begin{aligned} 0 &= \nabla_u J(\bar{u}, \bar{y}) + p^* \\ 0 &= \nabla_y J(\bar{u}, \bar{y}) + r^* - A^* p^*, \end{aligned}$$

where $p^* \in T_M(\bar{y}) \cap \{\bar{u} - A\bar{y}\}^\perp$ and $r^* \in \left[T_M(\bar{y}) \cap \{\bar{u} - A\bar{y}\}^\perp\right]^-$. Then by referring to Lemma 3.2 in [35] (along with the discussion following Lemma 2.2), it holds that

$$\begin{aligned} p^* &\geq 0, \ a.e.\, \mathcal{A}, \\ \langle \bar{u} - A\bar{y}, p^* \rangle &= 0, \\ \langle r^*, \varphi \rangle &= 0, \ \forall \varphi \in H_0^1(\Omega) : \varphi = 0, \ a.e.\, \mathcal{A}, \\ \langle r^*, \varphi \rangle &\leq 0, \ \forall \varphi \in M : \langle \bar{u} - A\bar{y}, \varphi \rangle = 0. \end{aligned}$$

By carefully comparing these conditions to Theorem 3.3 in [30], we see that our conditions yield those of Mignot and Puel. Note that in [30], $J(u,y) := \frac{1}{2}||y - z_d||^2_{L^2(\Omega)} + \frac{\alpha}{2}||u||^2_{L^2(\Omega)}$, with $\alpha > 0$. As expected in this setting, we see that the regularity of the optimal control $\bar{u}$ is better than $L^2(\Omega)$, in fact, $\bar{u} \in H^1_0(\Omega)$.

## 5.2 Pointwise Constraints on the Gradient of the State using the $\infty$-Norm

Many important problems in the study of elasto-plasticity require the pointwise bounding of the gradient of the displacement, i.e., the stress on an isotropic body at each point in the presence of a given force. The optimal control problem then results in an elliptic MPEC in which the gradients of the state $y$ are pointwise bounded (almost everywhere). In the following example, we consider a setting in which the gradients of the state are pointwise bounded using the $\infty$-norm on vectors in $\mathbb{R}^l$. In Section 8, we consider the 2-norm instead, which does not allow a simple reformulation to a bilateral setting.

In the following,

- $Y := H^1_0(\Omega)$,
- $\mathcal{U} := L^2(\Omega)$,
- $M := \left\{ y \in H^1_0(\Omega) \,|\, |\nabla y|_\infty \leq \psi, \ a.e.\,\Omega \right\}$, $\psi \in L^\infty(\Omega)$ and $\exists \underline{\psi} \in \mathbb{R}_+ \setminus \{0\} : \psi \geq \underline{\psi} > 0, \ a.e.\,\Omega$,
- $J : L^2(\Omega) \times H^1_0(\Omega) \to \mathbb{R}$, Fréchet differentiable.

We use $\nabla y$ to represent the full gradient and $\nabla y_i$ for its components. This simple rule will be applied throughout this example for all vectors and their components. It is easy to see that $M$ can be equivalently defined by

$$M = \left\{ y \in H^1_0(\Omega) \,|\, -\psi \leq \nabla y_i \leq \psi, \ a.e.\,\Omega, 1 \leq i \leq l \right\}.$$

In addition to these basic assumptions, we reduce the space of the gradient used in the state constraints, i.e., more specifically:

- $\nabla : H^1_0(\Omega) \to G(\Omega)$, where $G(\Omega) := \nabla(H^1_0(\Omega))$, i.e., the image space of the gradient.

In shrinking the image space of the gradient, we obtain a surjective bounded linear operator. This leads to the new formulation of the constraint set $M$:

$$M = \left\{ y \in H^1_0(\Omega) \,|\, \nabla y \in B_\psi \right\},$$

where

$$B_\psi := \{ \mathbf{z} \in G(\Omega) \,|\, -\psi \leq z_i \leq \psi, \ a.e.\,\Omega, \ 1 \leq i \leq l \}.$$

The image space is not merely chosen for its convenience. In fact, as shown in Proposition 1 (Chapter IX, Section 1) in [12], $L^2(\Omega)^l$ can be written as the orthogonal direct sum

$$L^2(\Omega)^l = G(\Omega) \oplus H(\operatorname{div} 0, \Omega),$$

where

$$H(\operatorname{div} 0, \Omega) := \left\{ \mathbf{x} \in L^2(\Omega)^l \,|\, \operatorname{div} \mathbf{x} = 0 \right\}.$$

This then allows us to use the $L^2(\Omega)^l$-norm on $G(\Omega)$. Not only is $G(\Omega)$ closed with respect to the $L^2(\Omega)^l$-norm, but it is also a Hilbert space with inner product $(\cdot,\cdot)_{L^2(\Omega)^l}$. Using these facts, it is easy to see the $B_\psi$ is closed and convex in $G(\Omega)$. We now show, using essentially the same argument as in the proof of Proposition 6.33 in [10], that $B_\psi$ is polyhedric in $G(\Omega)$. In order to continue, we will need the following definitions.

**Definition 5.1 (The Active and Inactive Sets).** *For some $y \in M$, we define the upper active set for the i-th component of $\nabla y$, $\mathcal{A}_i^+(\nabla y) \subseteq \Omega$, such that*

$$\mathcal{A}_i^+(\nabla y) := \{\omega \in \Omega \,|\, \nabla y_i(\omega) = \psi(\omega)\}.$$

*and the lower active set for the i-th component of $\nabla y$, $\mathcal{A}_i^-(\nabla y) \subseteq \Omega$, such that*

$$\mathcal{A}_i^-(\nabla y) := \{\omega \in \Omega \,|\, \nabla y_i(\omega) = -\psi(\omega)\}.$$

*The i-th inactive set, $\mathcal{I}_i(\nabla y)$, is therefore defined by*

$$\mathcal{I}_i(\nabla y) := \Omega \setminus \mathcal{A}_i^+(\nabla y) \cup \mathcal{A}_i^-(\nabla y).$$

Note that we can analogously define the active and inactive sets for any element of $B_\psi$.

**Proposition 5.1 (The Tangent and Normal Cones to $B_\psi$).** *Let the set $B_\psi$ be defined as above and $\mathbf{z} \in B_\psi$. Then*

$$T_{B_\psi}(\mathbf{z}) = \left\{\mathbf{h} \in G(\Omega) \,\middle|\, h_i \le 0,\ a.e.\, \mathcal{A}_i^+(\mathbf{z}),\ h_i \ge 0,\ a.e.\, \mathcal{A}_i^-(\mathbf{z}), 1 \le i \le l\right\}.$$

*Moreover, the properties of $G(\Omega)$ allow the elements of the normal cone $N_{B_\psi}(\mathbf{z})$ to be identified with elements of $G(\Omega)$ so that*

$$N_{B_\psi}(\mathbf{z}) = \left\{\boldsymbol{\lambda} \in G(\Omega) \,\middle|\, \lambda_i \ge 0,\ a.e.\, \mathcal{A}_i^+(\mathbf{z}),\ \lambda_i \le 0,\ a.e.\, \mathcal{A}_i^-(\mathbf{z}),\ \lambda_i = 0,\ a.e.\, \mathcal{I}_i(\mathbf{z}),\ 1 \le i \le l\right\}.$$

*Therefore, the critical cone to $B_\psi$ at $\mathbf{z}$ for some $\boldsymbol{\lambda} \in N_{B_\psi}(\mathbf{z})$ is characterized as*

$$\mathcal{K}(\mathbf{z}, \boldsymbol{\lambda}) = T_{B_\psi}(\mathbf{z}) \cap \{\boldsymbol{\lambda}\}^\perp = \left\{ \mathbf{h} \in G(\Omega) \,\middle|\, \begin{array}{ll} h_i \le 0 & a.e.\, \mathcal{A}_i^+(\mathbf{z}) : \lambda_i = 0 \\ h_i = 0 & a.e.\, \mathcal{A}_i^+(\mathbf{z}) : \lambda_i > 0 \\ h_i \ge 0 & a.e.\, \mathcal{A}_i^-(\mathbf{z}) : \lambda_i = 0 \\ h_i = 0 & a.e.\, \mathcal{A}_i^-(\mathbf{z}) : \lambda_i < 0 \end{array}, 1 \le i \le l \right\},$$

*where*

$$\{\boldsymbol{\lambda}\}^\perp := \left\{\mathbf{h} \in G(\Omega) \,\middle|\, (\boldsymbol{\lambda}, \mathbf{h})_{L^2(\Omega)^l} = 0\right\}.$$

*Here, conditions of the type "$a.e.\, \mathcal{A}_i^+(\mathbf{z}) : \lambda_i = 0$" are to be understood: "almost everywhere on the set $\left\{\omega \in \mathcal{A}_i^+(\mathbf{z}) \,|\, \lambda_i(\omega) = 0\right\}$".*

*Proof.* Let $\mathbf{h} \in T_{B_\psi}(\mathbf{z})$. Then by definition, there exist sequences $t_k \to 0^+$ and $\mathbf{h}_k \to \mathbf{h}$ in $G(\Omega)$ such that

$$\mathbf{z} + t_k \mathbf{h}_k \in B_\psi,\ \forall k.$$

Then for any $i$ such that $1 \le i \le l$, it holds that

$$-\psi \le z_i + t_k h_k^i \le \psi,\ a.e.\, \Omega,\ \forall k.$$

Thus,

$$0 \le h_k^i \le \frac{\psi - z_i}{t_k},\ a.e.\, \mathcal{A}_i^-(\mathbf{z}),\ \forall k$$

and

$$\frac{-\psi - z_i}{t_k} \le h_k^i \le 0,\ a.e.\, \mathcal{A}_i^+(\mathbf{z}),\ \forall k.$$

Here, we use $h_k^i$ to represent the $i$-th component of the $\mathbf{h}_k$. Hence,

$$T_{B_\psi}(\mathbf{z}) \subseteq \left\{\mathbf{h} \in G(\Omega) \,\middle|\, h_i \leq 0,\ a.e.\, \mathcal{A}_i^+(\mathbf{z}),\ h_i \geq 0,\ a.e.\, \mathcal{A}_i^-(\mathbf{z}), 1 \leq i \leq l \right\}. \tag{6}$$

Now let $\mathbf{h} \in G(\Omega)$ satisfy the righthand side of (6) and define the following class of vector fields indexed by $\tau > 0$

$$r_\tau := \frac{\Pi_{B_\psi}(\mathbf{z} + \tau\mathbf{h}) - \mathbf{z}}{\tau}.$$

Here, $\Pi_{B_\psi}$ represents the metric projection onto $B_\psi$. Since $\mathbf{z} + \tau r_\tau = \Pi_{B_\psi}(\mathbf{z} + \tau\mathbf{h})$, $r_\tau \in R_{B_\psi}(\mathbf{z})$ for all $\tau > 0$.

Consider now that for almost every $\omega \in \Omega$, $r_\tau(\omega) \to \mathbf{h}(\omega)$. Indeed, pointwise, we can always find $\tau > 0$ small enough such that $\mathbf{z}(\omega) + \tau\mathbf{h}(\omega) \in B_\psi(\omega)$. Moreover, since $G(\Omega)$ is a Hilbert space and $B_\psi(\omega)$ is closed and convex for almost every $\omega \in \Omega$, the metric projection is single-valued and Lipschitz continuous with modulus 1 (non-expansive). Therefore, it holds that $|r_\tau(\omega)| \leq |h(\omega)|$ for almost every $\omega \in \Omega$. Then given $G(\Omega)$ is a closed subspace of $L^2(\Omega)^l$, we can apply Lebesgue's Dominating Convergence Theorem, which yields $r_\tau \to \mathbf{h}$ in $G(\Omega)$. As the set of all $r_\tau$ is contained in $R_{B_\psi}(\mathbf{z})$ and $\mathrm{cl}\{R_{B_\psi}(\mathbf{z})\}_{G(\Omega)} = T_{B_\psi}(\mathbf{z})$, (6) holds as an equality.

We now move on to the derivation of the normal cone. By definition

$$N_{B_\psi}(\mathbf{z}) = \left[T_{B_\psi}(\mathbf{z})\right]^- = \left\{\boldsymbol{\lambda} \in G(\Omega)^* \,\middle|\, \langle \boldsymbol{\lambda}, \mathbf{h}\rangle_{G(\Omega)^*, G(\Omega)} \leq 0,\ \forall h \in T_{B_\psi}(\mathbf{z}) \right\}.$$

By virtue of the Riesz Representation Theorem, there exists a unique $\tilde{\boldsymbol{\lambda}} \in G(\Omega)$ for each $\boldsymbol{\lambda} \in G(\Omega)^*$ such that

$$\langle \boldsymbol{\lambda}, \mathbf{h}\rangle_{G(\Omega)^*, G(\Omega)} = \left(\tilde{\boldsymbol{\lambda}}, \mathbf{h}\right)_{L^2(\Omega)^l}.$$

Hence, we identify all $\boldsymbol{\lambda} \in N_{B_\psi}(\mathbf{z})$ with their $G(\Omega)$ counterparts, so that

$$\langle \boldsymbol{\lambda}, \mathbf{h}\rangle_{G(\Omega)^*, G(\Omega)} \leq 0,\ \forall h \in T_{B_\psi}(\mathbf{z}) \Leftrightarrow \left(\tilde{\boldsymbol{\lambda}}, \mathbf{h}\right)_{L^2(\Omega)^l} \leq 0,\ \forall h \in T_{B_\psi}(\mathbf{z}).$$

Defining $\mathcal{A}^+(\mathbf{z}) := \mathcal{A}_1^+(\mathbf{z}) \times \cdots \times \mathcal{A}_l^+(\mathbf{z})$, $\mathcal{A}^-(\mathbf{z}) := \mathcal{A}_1^-(\mathbf{z}) \times \cdots \times \mathcal{A}_l^-(\mathbf{z})$, and $\mathcal{I}(\mathbf{z}) := \mathcal{I}_1(\mathbf{z}) \times \cdots \times \mathcal{I}_l(\mathbf{z})$, the polarity inequality becomes

$$\left(\tilde{\boldsymbol{\lambda}}, \mathbf{h}\right)_{L^2(\Omega)^l} = \int_{\mathcal{A}^+(\mathbf{z})} \tilde{\boldsymbol{\lambda}} \cdot \mathbf{h} d\omega + \int_{\mathcal{A}^-(\mathbf{z})} \tilde{\boldsymbol{\lambda}} \cdot \mathbf{h} d\omega + \int_{\mathcal{I}(\mathbf{z})} \tilde{\boldsymbol{\lambda}} \cdot \mathbf{h} d\omega \leq 0,\ \ \forall h \in T_{B_\psi}(\mathbf{z}).$$

Referring to the above, we see that if $\mathbf{h} \in G(\Omega)$ such that $h_i = 0$ a.e. on $\mathcal{A}_i^+(\mathbf{z}) \cup \mathcal{A}_i^-(\mathbf{z})$ and free on $\mathcal{I}_i(\mathbf{z})$ for all $i = 1, \ldots, l$, then $\mathbf{h} \in T_{B_\psi}(\mathbf{z})$. Therefore, $\tilde{\boldsymbol{\lambda}}$ must equal zero a.e. on $\mathcal{I}(\mathbf{z})$. Then since the components of $\mathbf{h}$ are almost everywhere non-positive on $\mathcal{A}^+(\mathbf{z})$ and almost everywhere non-negative on $\mathcal{A}^-(\mathbf{z})$ for all $\mathbf{h} \in T_{B_\psi}(\mathbf{z})$, $\tilde{\boldsymbol{\lambda}}$ must always have the opposite signs (a.e.) on these sets. By identifying the $\tilde{\boldsymbol{\lambda}}$ with the $\boldsymbol{\lambda}$, the asserted formula for the normal cone holds.

Given the formulae for the tangent and normal cones, the characterization for the critical cone follows trivially. □

**Corollary 5.1 (Polyhedricity of $B_\psi$).** *The set $B_\psi$ as defined above is polyhedric in $G(\Omega)$, i.e., for any $\boldsymbol{\lambda} \in N_{B_\psi}(\mathbf{z})$, it holds that*

$$T_{B_\psi}(\mathbf{z}) \cap \{\boldsymbol{\lambda}\}^\perp = \mathrm{cl}\left\{R_{B_\psi}(\mathbf{z}) \cap \{\boldsymbol{\lambda}\}^\perp\right\}_{G(\Omega)}.$$

*Proof.* The argument follows analogously to the derivation of the tangent cone and mirrors the proof of Proposition 6.33 in [10]. □

This leads to our next result.

**Proposition 5.2** (**The Tangent and Normal Cones to** $M$)**.** *Let $y \in M$, where $M$ is defined as above. Then*

$$T_M(y) = \left\{ d \in H_0^1(\Omega) \,\middle|\, \nabla d \in T_{B_\psi}(\nabla y) \right\}$$

*and*

$$N_M(y) = \left\{ -\operatorname{div} \boldsymbol{\lambda} \in H^{-1}(\Omega) \,\middle|\, \boldsymbol{\lambda} \in G(\Omega) : \boldsymbol{\lambda} \in N_{B_\psi}(\nabla y) \right\},$$

*Here, the associated critical cone to $M$ at $y \in M$ for some $v \in N_M(y)$ is characterized by*

$$\mathcal{K}(y,v) = T_M(y) \cap \{v\}^\perp = \left\{ d \in H_0^1(\Omega) \,\middle|\, \begin{array}{ll} \nabla d_i \le 0 & a.e.\, \mathcal{A}_i^+(\nabla y) : \lambda_i = 0 \\ \nabla d_i = 0 & a.e.\, \mathcal{A}_i^+(\nabla y) : \lambda_i > 0 \\ \nabla d_i \ge 0 & a.e.\, \mathcal{A}_i^-(\nabla y) : \lambda_i = 0 \\ \nabla d_i = 0 & a.e.\, \mathcal{A}_i^-(\nabla y) : \lambda_i < 0 \end{array}, 1 \le i \le l \right\}.$$

*Here, $\boldsymbol{\lambda} \in N_{B_\psi}(\nabla y)$ such that $v = -\operatorname{div} \boldsymbol{\lambda}$.*

*Proof.* Due to the assumption on the range space of $\nabla$, the classical generalized Slater condition, i.e.,

$$0 \in \operatorname{int} \left\{ \nabla(H_0^1(\Omega)) - B_\psi \right\}$$

automatically holds. Indeed, since $\nabla(H_0^1(\Omega)) = G(\Omega)$ and $B_\psi \subset G(\Omega)$, there always exists an $\varepsilon > 0$ such that $\mathbb{B}_\varepsilon(0) \subset \nabla(H_0^1(\Omega)) - B_\psi$. Thus, the assertions hold for $T_M(y)$ and $N_M(y)$ (cf., e.g., [5], Chapter 4.2). Note that the adjoint of $\nabla$ is $-\operatorname{div}$, understood in a weak sense.

Due to Proposition 5.1, for each $v \in N_M(y)$, there exists a $\boldsymbol{\lambda} \in N_{B_\psi}(\nabla y)$ such that $v = -\operatorname{div} \boldsymbol{\lambda}$. Therefore, taking any arbitrary $d \in \mathcal{K}(y,v)$ requires

$$\langle -\operatorname{div} \boldsymbol{\lambda}, d \rangle_{H^{-1}, H_0^1} = (\boldsymbol{\lambda}, \nabla d)_{L^2(\Omega)^l} = 0.$$

Continuing the previous relation further yields

$$(\boldsymbol{\lambda}, \nabla d)_{L^2(\Omega)^l} = \int_{\mathcal{A}^+(\nabla y)} \boldsymbol{\lambda} \cdot \nabla d d\omega + \int_{\mathcal{A}^-(\nabla y)} \boldsymbol{\lambda} \cdot \nabla d d\omega + \int_{\mathcal{I}(\nabla y)} \boldsymbol{\lambda} \cdot \nabla d d\omega = \int_{\mathcal{A}^+(\nabla y)} \boldsymbol{\lambda} \cdot \nabla d d\omega + \int_{\mathcal{A}^-(\nabla y)} \boldsymbol{\lambda} \cdot \nabla d d\omega = 0.$$

The rest follows from the fact that $\nabla d_i$ and $\lambda_i$ always have opposite signs. The reverse inclusion is trivial. ☐

Given the explicit formula for the critical cone associated with $M$ provided by the previous proposition, we now demonstrate that $M$ is in fact polyhedric.

**Proposition 5.3** (**Polyhedricity of** $M$)**.** *Given $M$ as above, $y \in M$, and $v \in N_M(y)$, it holds that*

$$\mathcal{K}(y,v) = T_M(y) \cap \{v\}^\perp = \operatorname{cl} \left\{ R_M(y) \cap \{v\}^\perp \right\}_{H_0^1(\Omega)},$$

*i.e., $M$ is polyhedric.*

*Proof.* Since $\nabla$ is onto and $B_\psi$ is polyhedric (Propostion 5.1), Proposition 3.54 in [10] implies

$$\operatorname{cl} \left\{ d \in \mathcal{K}(y,v) \,\middle|\, \nabla d \in R_{B_\psi}(\nabla y) \right\}_{H_0^1(\Omega)} = \mathcal{K}(y,v).$$

Given

$$T_M(y) \cap \{v\}^\perp \supseteq \mathrm{cl}\left\{R_M(y) \cap \{v\}^\perp\right\}_{H_0^1(\Omega)},$$

it suffices to show that

$$\left\{d \in \mathcal{K}(y,v) \,\middle|\, \nabla d \in R_{B_\psi}(\nabla y)\right\} \subset R_M(y) \cap \{v\}^\perp.$$

By definition $\nabla d \in R_{B_\psi}(\nabla y)$ implies the existence of a $\tau > 0$ such that $\nabla y + \tau \nabla d \in B_\psi$. Hence, $d \in R_M(y)$. Then since $d \in \mathcal{K}(y,v)$ implies $\langle v, d\rangle_{H^{-1},H_0^1} = 0$, it holds that $d \in R_M(y) \cap \{v\}^\perp$. $\square$

Given the previous results, we need one last component in order to provide the explicit stationarity conditions for the elliptic MPEC.

**Proposition 5.4** (**The Polar Cone** $[\mathcal{K}(y,v)]^-$)**.** *Given $M$ as above, $y \in M$, and $v \in N_M(y)$, it holds that*

$$[\mathcal{K}(y,v)]^- = \left\{-\mathrm{div}\,\boldsymbol{\mu} \in H^{-1}(\Omega) \,\middle|\, \boldsymbol{\mu} \in G(\Omega) : \begin{array}{ll} \mu_i \geq 0 & a.e. \mathcal{A}_i^+(\nabla y) : \lambda_i = 0 \\ \mu_i \leq 0 & a.e. \mathcal{A}_i^-(\nabla y) : \lambda_i = 0 \\ \mu_i = 0 & a.e. \mathcal{I}_i(\nabla y) \end{array}, 1 \leq i \leq l \right\}$$

*Proof.* By definition,

$$[\mathcal{K}(y,v)]^- = \left\{d^* \in H^{-1}(\Omega) \,\middle|\, \langle d^*, d\rangle_{H^{-1},H_0^1} \leq 0, \, \forall d \in \mathcal{K}(y,v)\right\}.$$

Let $\boldsymbol{\mu}$ satisfy the requirements for the righthand side of the asserted result. Then by Proposition 5.2, for any $d \in \mathcal{K}(y,v)$, we have

$$\langle -\mathrm{div}\,\boldsymbol{\mu}, d\rangle_{H^{-1},H_0^1} = (\boldsymbol{\mu}, \nabla d)_{L^2} \leq 0.$$

Therefore, the inclusion "$\supseteq$" holds. For the reverse direction, define

$$L := \left\{\boldsymbol{\mu} \in G(\Omega) \,\middle|\, \begin{array}{ll} \mu_i \geq 0 & a.e. \mathcal{A}_i^+(\nabla y) : \lambda_i = 0 \\ \mu_i \leq 0 & a.e. \mathcal{A}_i^-(\nabla y) : \lambda_i = 0 \\ \mu_i = 0 & a.e. \mathcal{I}_i(\nabla y) \end{array}, 1 \leq i \leq l \right\}.$$

It is easy to show that $L$ is closed and convex in $G(\Omega)$. Assume now that there exists some $d^* \in [\mathcal{K}(y,v)]^-$ such that $d^* \notin -\mathrm{div}\,(L)$. Then there must exist some $\delta \in H_0^1(\Omega)$ strongly separating $d^*$ from $-\mathrm{div}\,(L)$, see e.g., II.38 Prop. 4 [11], i.e.,

$$\langle d^*, \delta\rangle_{H^{-1},H_0^1} > 0, \quad \langle -\mathrm{div}\,\boldsymbol{\mu}, \delta\rangle_{H^{-1},H_0^1} \leq 0, \ \forall \boldsymbol{\mu} \in L.$$

Then $\delta$ cannot be in $\mathcal{K}(y,v)$. However, for any arbitrary $\boldsymbol{\mu} \in L$, it holds that

$$0 \geq \langle -\mathrm{div}\,\boldsymbol{\mu}, \delta\rangle_{H^{-1},H_0^1} = (\boldsymbol{\mu}, \nabla \boldsymbol{\delta})_{L^2(\Omega)^l}.$$

Since the previous must hold for all $\boldsymbol{\mu} \in L$, we deduce that $\delta \in \mathcal{K}(y,v)$, a contradiction. The assertion follows. $\square$

Now that we have all the necessary characterizations, we can provide explicit strong stationarity conditions for the elliptic MPEC via Theorem 4.2.

**Proposition 5.5** (**Explicit Strong Stationarity Conditions**). *Under the given data assumptions, let $(\bar{u}, \bar{y})$ be a (locally) optimal solution to corresponding MPEC. Then there exist multipliers $p \in H_0^1(\Omega)$, $\boldsymbol{\lambda} \in G(\Omega)$, and $\boldsymbol{\mu} \in G(\Omega)$ such that*

$$0 = \nabla_u J(\bar{u}, \bar{y}) + p, \tag{7}$$

$$0 = \nabla_y J(\bar{u}, \bar{y}) - \operatorname{div} \boldsymbol{\mu} - A^* p, \tag{8}$$

$$0 = A\bar{y} - \bar{u} - \operatorname{div} \boldsymbol{\lambda}, \tag{9}$$

$$\tag{10}$$

*where for all $i = 1, \ldots, l$*

$$\left\{ \begin{array}{ll} \nabla p_i \leq 0, & a.e.\, \mathcal{A}_i^+(\nabla y) : \lambda_i = 0 \\ \nabla p_i = 0, & a.e.\, \mathcal{A}_i^+(\nabla y) : \lambda_i > 0 \\ \nabla p_i \geq 0, & a.e.\, \mathcal{A}_i^-(\nabla y) : \lambda_i = 0 \\ \nabla p_i = 0, & a.e.\, \mathcal{A}_i^-(\nabla y) : \lambda_i < 0 \end{array} \right| \begin{array}{ll} \mu_i \geq 0, & a.e.\, \mathcal{A}_i^+(\nabla y) : \lambda_i = 0 \\ \mu_i \leq 0, & a.e.\, \mathcal{A}_i^-(\nabla y) : \lambda_i = 0 \\ \mu_i = 0, & a.e.\, \mathcal{I}_i(\nabla y) \end{array} \left| \begin{array}{ll} \lambda_i \geq 0, & a.e.\, \mathcal{A}_i^+(\nabla y) \\ \lambda_i \leq 0, & a.e.\, \mathcal{A}_i^-(\nabla y) \\ \lambda_i = 0, & a.e.\, \mathcal{I}_i(\nabla y) \end{array} \right\}$$

*and*

$$\begin{aligned} \mathcal{A}_i^+(\nabla y) &= \{\omega \in \Omega \,|\, \nabla y_i(\omega) = \psi(\omega)\}, \\ \mathcal{A}_i^-(\nabla y) &= \{\omega \in \Omega \,|\, \nabla y_i(\omega) = -\psi(\omega)\}, \\ \mathcal{I}_i(\nabla y) &= \Omega \setminus \mathcal{A}_i^+(\nabla y) \cup \mathcal{A}_i^-(\nabla y). \end{aligned}$$

*Proof.* The result follows from Theorem 4.2 via Propositions 5.2, 5.3, and 5.4. □

As in the simple obstacle case, we see that if the tracking functional

$$J(u, y) = \frac{1}{2} ||y - z_d||_{L^2}^2 + \frac{\alpha}{2} ||u||_{L^2(\Omega)}^2,$$

is chosen with $\alpha > 0$, then the optimal control $\bar{u} \in H_0^1(\Omega)$.

# 6 Generalized Differentiation of the Solution via Variational Convergence and Epiderivatives

In the following section, we derive a new formula for the contingent derivative of the normal cone mapping that is valid for all closed convex sets in reflexive Banach spaces. This important result allows us to provide optimality conditions for a much wider array of problems and extends the characterizations of the polyhedric setting. We begin by introducing an important type of variational convergence.

**Definition 6.1** (**Mosco Epiconvergence**). *Let $\{\varphi_t\}$ be a family of functions from a Banach space $X$ into the extended reals $\bar{\mathbb{R}}$ parameterized by $t > 0$ and $\varphi : X \to \bar{\mathbb{R}}$. Then the family $\varphi_t$ is said to Mosco epiconverge to $\varphi$ as $t \to 0^+$ if for all sequences $t_n \to 0^+$ and every $x \in X$, the following two conditions hold*

$$\forall x_n \rightharpoonup x, \quad \varphi(x) \leq \liminf_n \varphi_{t_n}(x_n), \tag{11}$$

$$\exists x_n \to x, \quad \varphi(x) \geq \limsup_n \varphi_{t_n}(x_n). \tag{12}$$

For more on this and related types of variational convergence, we refer the reader to [3]. In the next definition, we will use second-order differential quotients associated with some proper convex lower semicontinuous function $f : X \to \bar{\mathbb{R}}$, where $X$ is again some arbitrary Banach space. We assume $f$ is finite at $x \in X$, $x^* \in X^*$, and $h \in X$ arbitrary. The so-called *second-order difference quotient* associated with $f$ are then defined by

$$(\Delta_t^2 f)_{x,x^*}(h) := \frac{f(x+th) - f(x) - t\langle x^*, x\rangle}{\frac{1}{2}t^2}.$$

**Definition 6.2** (**Second-Order Mosco Epiderivatives**). *Let $f : X \to \bar{\mathbb{R}}$ be any proper convex lower semicontinuous function, $X$ an arbitrary Banach space and $x^* \in X^*$. If the family of associated second-order difference quotients $(\Delta_t^2 f)_{x,x^*}$ Mosco epiconverges to some function $\varphi$ as $t \to 0^+$ with $\varphi(0) \neq -\infty$, then $f$ is said to be twice Mosco epidifferentiable at $x$ relative to $x^*$. Here, $\varphi$ represents the second-order Mosco epiderivative of $f$ at $x$ relative to $z$, which we denote by $f''_{x,x^*}$.*

Second-order Mosco epiderivatives were introduced by Rockafellar for extended real-valued functionals from $\mathbb{R}^n$ in [36] and there is a compendium of results for finite dimensional objects in [38]. Some important references for the infinite dimensional setting include, but are by no means limited to, [13, 22, 27].

Before introducing the next type of generalized derivative, we recall two notions of set limits, see e.g., [5].

**Definition 6.3** (**Painlevé-Kuratowski Upper/Lower-Limits**). *Let $C_t \subseteq X$ be a family of subsets indexed by $t > 0$ and $X$ is some Banach space. We define*

- *the Painlevé-Kuratowski Upper Limit of $\{C_t\}_{t>0}$*

$$\operatorname{Lim\,sup}_{t\to 0^+} C_t := \left\{ x \in X \,\middle|\, \exists t_n \to 0^+ \exists x_n \to x : x_n \in C_{t_n} \right\};$$

- *the Painlevé-Kuratowski Lower Limit of $\{C_t\}_{t>0}$*

$$\operatorname{Lim\,inf}_{t\to 0^+} C_t := \left\{ x \in X \,\middle|\, \forall t_n \to 0^+ \forall x_n \to x : x_n \in C_{t_n} \right\}.$$

As mentioned in a remark following Theorem 3.1, the original calculus rules of Rockafellar (finite dimensions) and Do (infinite dimensions) were stated in terms of so-called *proto-derivatives.* These objects were introduced by Rockafellar in [37] and were also developed by A.Levy and R.Poliquin. We direct the reader to [38] for a comprehensive listing of results and a proper bibliography of important works.

**Definition 6.4** (**Proto-Derivatives**). *Let $F : X \rightrightarrows Y$ be a multifunction and both $X$ and $Y$ be Banach spaces. Then $F : X \rightrightarrows Y$ is said to be proto-differentiable at $x \in X$ relative to $z \in Y$ if and only if*

$$\operatorname{Lim\,sup}_{t\to 0^+} \frac{\operatorname{gph} F - (x,z)}{t} = \operatorname{Lim\,inf}_{t\to 0^+} \frac{\operatorname{gph} F - (x,z)}{t},$$

*with the common cone being the graph of the proto-derivative. We denote the proto-derivative of $F$ at $x$ relative to $z$ by $PF[(x,z)]$.*

Note that since

$$\operatorname{gph} DF[(x,z)] = T_{\operatorname{gph} F}(x,z) = \operatorname{Lim\,sup}_{t\to 0^+} \frac{\operatorname{gph} F - (x,z)}{t},$$

the proto-differentiability of a multifunction $F$ immediately implies the contingent-differentiability of $F$ and the two derivatives coincide. We can now state the following important calculus rule, which forms the basis for our interest in second-order Mosco epiderivatives.

**Theorem 6.1** (Do **[13], Theorem 3.9**). *Let $f : X \to \bar{\mathbb{R}}$ be a proper convex lower-semicontinuous function, $X$ a reflexive Banach space, and $f(x)$ be finite. Then the following statements are equivalent:*

- *$f$ is twice Mosco epidifferentiable at $x \in X$ relative to $x^* \in X^*$.*
- *The subdifferential $\partial f$ is proto-differentiable at $(x, x^*) \in \operatorname{gph} \partial f$.*

*In addition, it holds that*

$$\partial(\frac{1}{2} f''_{x,x^*})(h) = P\partial f[(x, x^*)](h), \quad h \in X. \tag{13}$$

This brings us to our main result.

**Theorem 6.2** (**The Second-Order Mosco Epiderivative of $I_M$**). *Let $M$ be a nonempty closed convex set in a reflexive Banach space $Y$ and $v \in \partial I_M(y) = N_M(y)$. Then the indicator function $I_M : Y \to \bar{\mathbb{R}}$ is twice Mosco epidifferentiable at $y$ relative to $v$ and the second order Mosco epiderivative is characterized as follows:*

$$(I''_M)_{y,v}(d) = \begin{cases} 0 & d \in T_M(y) \cap \{v\}^{\perp}, \\ \infty & \textit{otherwise}. \end{cases} \tag{14}$$

*In other words,*

$$(I''_M)_{y,v}(d) = I_{\mathcal{K}(y,v)}(d),$$

*where*

$$\mathcal{K}(y, v) := T_M(y) \cap \{v\}^{\perp}.$$

*Proof.* Clearly the indicator function of a nonempty closed convex set is proper, convex, and lower semicontinuous and for any $y \in M$, $I_M(y) = 0$, i.e., $I_M$ is finite. Begin by letting $t_n \to 0^+$ be an arbitrary sequence of scalars converging to zero from above and let $d_n \rightharpoonup_Y d$, for an arbitrary $d \in Y$. By the definition of the indicator function, the lower limit

$$\liminf_n \frac{I_M(y + t_n d_n) - I_M(y) - t_n \langle v, d_n \rangle}{t_n^2/2}$$

will be equal to infinity unless there exists a subsequence of $d_n$ or some large $N_0 \in \mathbb{N}$ such that $y + t_n d_n \in M$ for all $n \geq N_0$. Therefore, suppose such a sequence exists. Given any $v' \in N_M(y)$, the closure and convexity of $M$ imply that $\langle v', y' - y \rangle \leq 0$ for all $y' \in M$, in which case it follows from the assumption that $\langle v', d_n \rangle \leq 0$. Since $d_n \rightharpoonup_Y d$, we observe that $\langle v', d \rangle \leq 0$ for all $v' \in N_M(y)$. Then from the convexity of $M$, we deduce: $d \in T_M(y)$. In addition, we see that the second-order difference quotients are all non-negative and in fact reduce to

$$\frac{-2\langle v, d_n \rangle}{t_n} \geq 0.$$

Hence, if $\langle v, d_n \rangle$ does not converge to zero, $t_n \to 0^+$ implies that the lower limit tends to infinity, which leads to the following inequality.

$$\liminf_n \frac{I_M(y + t_n d_n) - I_M(y) - t_n \langle v, d_n \rangle}{t_n^2/2} \geq \begin{cases} 0 & d \in T_M(y) \cap \{v\}^{\perp}, \\ \infty & \text{otherwise}, \end{cases} \tag{15}$$

Since $t_n$ was arbitrarily chosen, (15) holds for all sequences $t_n \to 0^+$ and $d_n \rightharpoonup d$. Thus, the righthand side of (15) is a good candidate for the second-order Mosco epiderivative. Recalling

Definition 6.2, we see that in order to complete the proof, we need to show for all $t_n \to 0^+$ that there exists a strongly converging sequence $d_n \to d$ such that

$$\limsup_n \frac{I_M(y+t_nd_n) - I_M(y) - t_n\langle v, d_n\rangle}{t_n^2/2} \leq \begin{cases} 0 & d \in T_M(y) \cap \{v\}^\perp\,, \\ \infty & \text{otherwise}\,. \end{cases}$$

In what follows, let $t_n \to 0^+$ be arbitrary. Clearly, if $d \notin T_M(y) \cap \{v\}^\perp$, then the inequality will always hold. Therefore, we only need to construct sequences for $d \in T_M(y) \cap \{v\}^\perp$. Since both $T_M(y)$ and $\{v\}^\perp$ are strongly closed in Y, their intersection is as well. Therefore, for all $d \in T_M(y) \cap \{v\}^\perp$ there exists a strongly convergent sequence $\delta_n \to d$ with $\delta_n \in T_M(y) \cap \{v\}^\perp$ for all $n$, such that $\langle v, \delta_n\rangle = 0$ and, by the definition of $T_M(y)$, sequences $\tau_k^n \to 0^+$ and $\delta_k^n \to \delta_n$ such that $y + \tau_k^n\delta_k^n \in M$ for all $k$ and each $n$. From the convexity of $M$, we infer that for any $t \in [0, \tau_k^n]$, we have

$$y + t\delta_k^n = (1 - \frac{t}{\tau_k^n})y + \frac{t}{\tau_k^n}(y + \tau_k^n\delta_k^n) \in M, \ \forall k, \forall n. \tag{16}$$

Moreover, we deduce that $\langle v, \delta_k^n\rangle \to 0$ as $k \to \infty$ and similar to before, we see that $\langle v, \delta_k^n\rangle \leq 0$. Hence, for some $\varepsilon > 0$, there exists $K_n \in \mathbb{N}$ for each $n$ such that

1. $|2\langle v, \delta_k^n\rangle| \leq t_{n_{min}}^{1+\varepsilon}$ for all $k \geq K_n$, where $n_{min} := \operatorname{argmin}_n \{t_1, \ldots, t_n\}$.
2. $||\delta_n - \delta_k^n|| \leq t_n$ for all $k \geq K_n$.
3. $\tau_k^n \leq f_n$ for all $k \geq K_n$, with $f_n$ arbitrary such that $f_n \to 0^+$ monotonely.

We now build our strongly converging sequence $d_n$. Begin by fixing $m_1 \in \mathbb{N}$ and define

$$\varepsilon_{m_1} := \min_{1 \leq i \leq m_1} \tau_{K_i}^i.$$

Given $t_n \to 0^+$, there exists an $N(\varepsilon_{m_1}) \in \mathbb{N}$ such that for all $n \geq N(\varepsilon_{m_1})$

$$t_n \leq \varepsilon_{m_1}.$$

By the definition of $\varepsilon_{m_1}$, it also holds that

$$t_n \leq \tau_{K_{m_1}}^{m_1}.$$

For all $n < N(\varepsilon_{m_1})$, set $d_n := \delta_{K_1}^1$. Now define $j_1$ to be the smallest index such that

$$t_{N(\varepsilon_{m_1})} > \tau_{K_{m_1+j_1}}^{m_1+j_1}$$

and define $l_1 \geq 1$ to be the first index such that

$$t_{N(\varepsilon_{m_1})+l_1} \leq \tau_{K_{m_1+j_1}}^{m_1+j_1}.$$

By the third assumption on $K_n$, $j_1$ exists since $\tau_{K_n}^n \to 0^+$, with $n$; and $l_1$ exists since $t_n \to 0^+$. Then using these indices, we set $d_n = \delta_{K_{m_1}}^{m_1}$ for all $n = N(\varepsilon_{m_1}), \ldots, N(\varepsilon_{m_1}) + l_1 - 1$.

Given for all $n = N(\varepsilon_{m_1}) + i$ with $i = 0, \ldots, l_1 - 1$, $t_n \in [0, \tau_{K_{m_1}}^{m_1}]$ and $d_n = \delta_{K_{m_1}}^{m_1}$, it holds that $y + t_nd_n \in M$ (cf. (16)). Thus, $I_M(y + t_nd_n) = 0$.

At this point we define $m_2 := m_1 + j_1$ and repeat the process described above with $m_2$ in place of $m_1$. Clearly, $\varepsilon_{m_2} \leq \varepsilon_{m_1}$ and $N(\varepsilon_{m_2}) \geq N(\varepsilon_{m_1})$. If $N(\varepsilon_{m_1}) + l_1 - 1 < N(\varepsilon_{m_2})$, then set $d_n = \delta_{K_{m_1}}^{m_1}$ for all $n$ such that $N(\varepsilon_{m_1}) + l_1 - 1 \leq n < N(\varepsilon_{m_2})$. As $l_1 \geq 1$, $t_n \leq \varepsilon_{m_1}$ so that $t \in [0, \tau_{K_{m_1}}^{m_1}]$ and

$y + t_n d_n \in M$ still holds. Also note that the case "$N(\varepsilon_{m_2}) < N(\varepsilon_{m_1}) + l_1 - 1$" cannot happen. Indeed, this would require, for all $n = N(\varepsilon_{m_2}), \ldots, N(\varepsilon_{m_1}) + l_1 - 1$, that $\tau^{m_2}_{K_{m_2}} < t_n \leq \varepsilon_{m_2}$, but $\varepsilon_{m_2} \leq \tau^{m_2}_{K_{m_2}}$ by definition, a contradiction. The rest continues as before.

To see that the process continues indefinitely, let $p \geq 1$ and consider that the convergence of $\tau^n_{K_n} \to 0^+$ ensures the existence of a $j_p \in \mathbb{N}$ such that $t_{N(\varepsilon_{m_p})} > \tau^{m_p+j_p}_{K_{m_p+j_p}}$. With $j_p$ fixed, we now look to increase the number $n$ larger than $N(\varepsilon_{m_p})$. We do this by checking if $t_{N(\varepsilon_{m_p})+i} > \tau^{m_p+j_p}_{K_{m_p+j_p}}$ for $i = 1, 2, \ldots$. As $j_p$ is fixed, so is $\tau^{m_p+j_p}_{K_{m_p+j_p}}$. Therefore, the convergence of $t_n \to 0^+$ implies the existence of some $l_p$ such that $t_{N(\varepsilon_{m_p})+l_p} \leq \tau^{m_p+j_p}_{K_{m_p+j_p}}$, by definition, we define $m_{p+1}$ and continue as before. This ensures that the process is perpetual.

Summarizing, we describe the construction via the following diagram

$$d_n = \underbrace{\delta^1_{K_1}, \ldots, \delta^1_{K_1}}_{1 \leq n < N(\varepsilon_{m_1})} \mid \underbrace{\delta^{m_1}_{K_{m_1}}, \ldots, \delta^{m_1}_{K_{m_1}}}_{N(\varepsilon_{m_1}) \leq n < N(\varepsilon_{m_2})} \mid \underbrace{\delta^{m_2}_{K_{m_2}}, \ldots, \delta^{m_2}_{K_{m_2}}}_{N(\varepsilon_{m_2}) \leq n < N(\varepsilon_{m_3})} \mid \ldots,$$

that is, $d_n$ is a subsequence of $\delta^n_{K_n}$. Furthermore, we have for all $\mu > 0$ and large $n$ that

$$||d - \delta^n_{K_n}|| \leq ||d - \delta_n|| + ||\delta_n - \delta^n_{K_n}|| \leq \mu + t_n \leq 2\mu,$$

so that $\delta^n_{K_n} \to d$ as $n \to \infty$ and thus, $d_n$ as well. Finally, since

$$0 \leq \frac{-2\langle v, d_n \rangle}{t_n} \leq \frac{t^{1+\varepsilon}_{n_{min}}}{t_n} \leq \frac{t^{1+\varepsilon}_n}{t_n} = t^\varepsilon_n \text{ and } t^\varepsilon_n \to 0,$$

we see that the true limit of the difference quotients exists so that the limit superior equals zero. Lastly, since $0 \in T_M(y) \cap \{v\}^\perp$ the function

$$\varphi(d) := \begin{cases} 0 & d \in T_M(y) \cap \{v\}^\perp, \\ \infty & \text{otherwise}, \end{cases}$$

is the Mosco epilimit of the second-order difference quotients and therefore amounts to the second-order Mosco epiderivative of $I_M$, as was to be shown. □

We immediately obtain the next two corollaries.

**Corollary 6.1 (The Contingent Derivative of $N_M$).** *Let $M$ be a nonempty closed convex set of a reflexive Banach space $Y$ and $v \in \partial I_M(y) = N_M(y)$. Then for any $(y, v) \in \operatorname{gph} N_M$, the following are equivalent:*

1. $w \in DN_M[(y, v)](d)$.
2. $w \in N_{\mathcal{K}(y,v)}(d)$.
3. $(d, w) \in \mathcal{K}(y, v) \times [\mathcal{K}(y, v)]^- : \langle w, d \rangle = 0$.

*Here, $\mathcal{K}(y, v) := T_M(y) \cap \{v\}^\perp$, i.e., the critical cone.*

*Proof.* The result follows from (13) via the formula (14) in Theorem 6.2. Indeed, $\partial I_{\mathcal{K}(y,v)}(d) = N_{\mathcal{K}(y,v)}(d)$ due to convexity. Furthermore, since $\mathcal{K}(y, v)$ is a cone, we have from Lemma 4.2.5 in [5] that

$$w \in N_{\mathcal{K}(y,v)}(d) \Longleftrightarrow (d, w) \in \mathcal{K}(y, v) \times [\mathcal{K}(y, v)]^- : \langle w, d \rangle = 0.$$

The assertion then follows. □

**Corollary 6.2** (**The Contingent Derivative of** $S$)**.** *Let $M$ be a nonempty closed convex set of a reflexive Banach space $Y$ and $S$ be as above. If $(u, y) \in \operatorname{gph} S$, then the following are equivalent:*

1. $d \in DS[(u, y)](w)$.
2. $w \in Ad + N_{\mathcal{K}(y,v)}(d)$.
3. $w - Ad \in [\mathcal{K}(y, v)]^-$ , $d \in \mathcal{K}(y, v)$, $\langle w - Ad, d\rangle = 0$.

*or equivalently*

$$T_{\operatorname{gph} S}(u, y) = \left\{(w, d) \in Y^* \times Y \,\middle|\, w - Ad \in [\mathcal{K}(y, v)]^- ,\ d \in \mathcal{K}(y, v),\ \langle w - Ad, d\rangle = 0\right\}.$$

*Here, $v := u - Ay$.*

*Proof.* The proof is analogous to that of Theorem 4.2. The rest follows from Corollary 6.1. □

Corollary 6.2 once again demonstrates the power of the variational analytic method, as we now have a formula characterizing the contingent derivative of the solution of *any* linear elliptic variational inequality in reflexive Banach spaces. Furthermore, as mentioned in Remark 3.1, we have now fully generalized Mignot's classical result.

# 7 Optimality Conditions for the Elliptic MPEC: Nonlinear Case

We are now in a position to derive abstract optimality conditions for the more general elliptic MPEC (2), which we restate here for convenience.

$$\begin{cases} \min J(u, y) \\ \text{s.t. } u \in \mathcal{U},\ y \in S(u). \end{cases} \tag{17}$$

We assume as always that $S$ is the solution mapping of some variational inequality of the form

$$u \in Ay + N_M(y),$$

where $A : Y \to Y^*$ is a coercive bounded linear operator and $M \subseteq Y$ is any non-empty closed convex set. As before, we assume that $Y \hookrightarrow \mathcal{U} \hookrightarrow Y^*$, with $\mathcal{U}$ a Hilbert space, is densely embedded in $Y^*$, and we assume $J$ is Fréchet differentiable from $\mathcal{U} \times Y$ to $\mathbb{R}$.

**Theorem 7.1** (**Strong Stationarity Conditions II**)**.** *Let $(\bar{u}, \bar{y})$ be a locally optimal solution of (17), then there exist multipliers $p^* \in \mathcal{K}(\bar{y}, \bar{v})$, $r^* \in [\mathcal{K}(\bar{y}, \bar{v})]^-$, and $\bar{v} \in N_M(\bar{y})$ such that*

$$0 = \nabla_u J(\bar{u}, \bar{y}) + p^*, \tag{18}$$

$$0 = \nabla_y J(\bar{u}, \bar{y}) + r^* - A^* p^*, \tag{19}$$

$$0 = A\bar{y} - \bar{u} + \bar{v}. \tag{20}$$

*Proof.* The result follows from Proposition 5.1 in [32] given the Fréchet differentiability of $J$ on $\mathcal{U} \times Y$ and using the estimate from Proposition 3.2, which holds via the characterization of $DS[(\bar{u}, \bar{y})]$ in Corollary 6.2. □

# 8 Pointwise Constraints on the Gradient of the State using the 2-Norm

Earlier we motivated the need for the non-linear non-smooth case with

$$M := \left\{ y \in H_0^1(\Omega) \,|\, |\nabla y|_2 \le \psi,\ a.e.\,\Omega \right\}.$$

where $\nabla : H_0^1(\Omega) \to G(\Omega)$ with $G(\Omega) := \nabla(H_0^1(\Omega)) \subset L^2(\Omega)^l$, a closed (Hilbert) subspace of $L^2(\Omega)^l$. As we are about to see, the derivation of the objects needed for the stationarity conditions mirrors that of the polyhedric case. Once again, we have the following data assumptions, with $M$ as above.

- $Y := H_0^1(\Omega)$,
- $\mathcal{U} := L^2(\Omega)$,
- $J : L^2(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$, Fréchet differentiable,
- $\psi$ has a lower bound, $\underline{\psi} \in \mathbb{R}_+ \setminus \{0\}$, i.e., $\psi \ge \underline{\psi} > 0$.
- $\nabla : H_0^1(\Omega) \to G(\Omega)$, where $G(\Omega) := \nabla(H_0^1(\Omega))$, i.e., the image space of the gradient.

By defining $K_\psi := \{\mathbf{z} \in G(\Omega) \,||\mathbf{z}|_2 \le \psi,\ a.e.\,\Omega\}$, we reformulate $M$ as

$$M = \left\{ y \in H_0^1(\Omega) \,|\nabla y \in K_\psi \right\}.$$

Since $\nabla$, as defined here, is a surjective bounded linear operator, we only need to calculate $T_{K_\psi}$ and $N_{K_\psi}$ in order to derive formulae for $T_M$ and $N_M$, for which we have the next result.

**Proposition 8.1 (The Tangent and Normal Cones to $K_\psi$).** *Let $K_\psi$ be defined as above and assume $\mathbf{z} \in K_\psi$. Then*

$$T_{K_\psi}(\mathbf{z}) = \{\mathbf{h} \in G(\Omega) \,|\, \mathbf{z}\cdot\mathbf{h} \le 0,\ a.e.\,\mathcal{A}(\mathbf{z})\}$$

*and*

$$N_{K_\psi}(\mathbf{z}) = \left\{ \lambda\mathbf{q} \in G(\Omega) \;\middle|\; \begin{array}{l} \lambda \in L^2(\Omega) : \lambda \ge 0,\ a.e.\,\mathcal{A}(\mathbf{z}), \lambda = 0,\ a.e.\,\mathcal{I}(\mathbf{z}) \\ \mathbf{q}(\omega) \in \left\{ \begin{array}{ll} \{\mathbf{z}(\omega)\,/|\mathbf{z}(\omega)|_2\} & |\mathbf{z}(\omega)| \ne 0 \\ \mathbb{B}_1(0) & |\mathbf{z}(\omega)| = 0 \end{array} \right. \end{array} \right\}$$

*Here, $\mathbb{B}_1(0)$ represents the closed unit ball in $G(\Omega)$.*

*Proof.* Let $\mathbf{h} \in T_{K_\psi}(\mathbf{z})$. Then there exist sequences $t_k \to 0^+$ and $\mathbf{h}_k \to \mathbf{h}$ in $G(\Omega)$ such that $|\mathbf{z} + t_k\mathbf{h}_k|_2 \le \psi$ a.e. $\Omega$. By squaring both sides and rearranging terms, we observe that $\mathbf{z}\cdot\mathbf{h} \le 0$ a.e. $\mathcal{A}(\mathbf{z})$. Thus, the inclusion "$\subseteq$" holds. For the reverse direction, we again us an argument based on Lebesgue's Dominating Convergence Theorem. Indeed, by defining the family of $G(\Omega)$-vector fields

$$\mathbf{p}_\tau := \frac{\Pi_{K_\psi}(\mathbf{z} + t_k\mathbf{h}) - \mathbf{z}}{\tau},$$

it holds that for all $\tau > 0$, $\mathbf{p}_\tau \in R_{K_\psi}(\mathbf{z})$. Moreover, for almost every $\omega$, $\mathbf{p}_\tau(\omega) \to \mathbf{h}(\omega)$ and since $G(\Omega)$ is a Hilbert space and $K_\psi$ is closed and convex, the metric projection is non-expansive, i.e., $|\mathbf{p}_\tau(\omega)| \le |\mathbf{h}(\omega)|$. As $G(\Omega)$ is a closed subspace of $L^2(\Omega)^l$ in the $L^2(\Omega)^l$-norm, it holds that $\mathbf{p}_\tau \to \mathbf{h}$ in $G(\Omega)$. Therefore, the reverse inclusion holds.

We now characterize the normal cone. Formally,

$$N_{K_\psi}(\mathbf{z}) = \left\{ \mu^* \in G(\Omega)^* \,\middle|\, \langle \mu^*, \mathbf{h}\rangle_{G^*,G} \le 0, \forall \mathbf{h} \in T_{K_\psi}(\mathbf{z}) \right\}.$$

Nevertheless, since $G(\Omega)$ is a Hilbert space with $L^2(\Omega)^l$ inner product, we can associate with every $\mu^* \in N_{K_\psi}(\mathbf{z})$ its $G(\Omega)$-counterpart $\mu$. Therefore, we view the normal cone as follows

$$N_{K_\psi}(\mathbf{z}) = \left\{\mu \in G(\Omega) \,\middle|\, (\mu, \mathbf{h})_{L^2(\Omega)^l} \leq 0, \forall \mathbf{h} \in T_{K_\psi}(\mathbf{z}) \right\}.$$

Suppose that $\lambda \in L^2(\Omega)$ satisfies $\lambda \geq 0$ *a.e.* $\mathcal{A}(\mathbf{z})$ and $\lambda = 0$ *a.e.* $\mathcal{I}(\mathbf{z})$ and $\mathbf{q} \in G(\Omega)$ such that

$$\mathbf{q}(\omega) \in \begin{cases} \{\mathbf{z}(\omega)\,/|\mathbf{z}(\omega)|_2\,\} & |\mathbf{z}(\omega)| \neq 0, \\ \mathbb{B}_1(0) & |\mathbf{z}(\omega)| = 0. \end{cases}$$

In general, $\lambda \mathbf{q} \in L^2(\Omega)^l$ as the components of $\mathbf{q}$ are all essentially bounded on $\Omega$ (and therefore elements of $L^\infty(\Omega)$). Then clearly

$$(\lambda \mathbf{q}, \mathbf{h})_{L^2(\Omega)^l} = \int_{\mathcal{A}(\mathbf{z})} \lambda \frac{\mathbf{z}}{|\mathbf{z}|_2} \cdot \mathbf{h} d\omega \leq 0, \;\; \forall \mathbf{h} \in T_{K_\psi}(\mathbf{z}).$$

Define the set

$$N^0(\mathbf{z}) := \left\{ \lambda \mathbf{q} \in G(\Omega) \,\middle|\, \begin{array}{l} \lambda \in L^2(\Omega) : \lambda \geq 0, \;\; a.e.\, \mathcal{A}(\mathbf{z}), \lambda = 0, \;\; a.e.\, \mathcal{I}(\mathbf{z}) \\ \mathbf{q}(\omega) \in \begin{cases} \{\mathbf{z}(\omega)\,/|\mathbf{z}(\omega)|_2\,\} & |\mathbf{z}(\omega)| \neq 0 \\ \mathbb{B}_1(0) & |\mathbf{z}(\omega)| = 0 \end{cases} \end{array} \right\}.$$

Clearly, $N_{K_\psi}(\mathbf{z}) \supseteq N^0(\mathbf{z})$. To see that $N^0(\mathbf{z})$ is convex in $G(\Omega)$, let $\alpha \in (0,1)$ and $\mathbf{v}_1, \mathbf{v}_2 \in N^0(\mathbf{z})$. Then since

$$\alpha \mathbf{v}_1(\omega) + (1-\alpha)\mathbf{v}_2(\omega) = \begin{cases} (\alpha\lambda_1(\omega) + (1-\alpha)\lambda_2(\omega))\frac{\mathbf{z}(\omega)}{|\mathbf{z}(\omega)|_2}, & |\mathbf{z}(\omega)| \neq 0, \\ 0, & |\mathbf{z}(\omega)| = 0, \end{cases}$$

and the set

$$\left\{\lambda \in L^2(\Omega) \,|\, \lambda \geq 0, \;\; a.e.\, \mathcal{A}(\mathbf{z}), \lambda = 0, \;\; a.e.\, \mathcal{I}(\mathbf{z}) \right\}$$

is convex, it holds that $\alpha \mathbf{v}_1 + (1-\alpha)\mathbf{v}_2 \in N^0(\mathbf{z})$.

Next, we demonstrate that $N^0(\mathbf{z})$ is closed in $G(\Omega)$. Let $\mathbf{v}_n \in N^0(\mathbf{z})$ such that $\mathbf{v}_n \to \mathbf{v}$ in $L^2(\Omega)^l$. By definition, there exist $\lambda_n \in L^2(\Omega)$, where $\lambda_n \geq 0$ almost everywhere on $\mathcal{A}(\mathbf{z})$ and $\lambda_n = 0$ almost everywhere on $\mathcal{I}(\mathbf{z})$, and $\mathbf{q}_n \in G(\Omega) \subset L^2(\Omega)^l$ with the pointwise description for almost every $\omega \in \Omega$:

$$\mathbf{q}_n(\omega) \in \begin{cases} \{\mathbf{z}(\omega)/|\mathbf{z}(\omega)|\} & |\mathbf{z}(\omega)| \neq 0, \\ \mathbb{B}_1(0) & |\mathbf{z}(\omega)| = 0, \end{cases}$$

such that $\mathbf{v}_n = \lambda_n \mathbf{q}_n$. Clearly, $\mathbf{q}_n \in L^\infty(\Omega)^l$ for all $n$.

Given $\mathbf{v}_n \to \mathbf{v}$ in $L^2(\Omega)^l$, there exists a positive constant $C$ such that $||\mathbf{v}_n||^2_{L^2(\Omega)^l} \leq C$. Hence,

$$\int_\Omega \lambda_n \mathbf{q}_n \cdot \lambda_n \mathbf{q}_n d\omega = \int_\Omega \lambda_n^2 \mathbf{q}_n \cdot \mathbf{q}_n d\omega = \int_{\mathcal{A}(\mathbf{z})} \lambda_n^2 d\omega = ||\lambda_n||^2_{L^2(\Omega)} \leq C.$$

Therefore, there exists $\hat{\lambda} \in L^2(\Omega)$ and a subsequence $\{n_k\}_{k=1}^\infty \subset \{n\}_{n=1}^\infty$ such that $\lambda_{n_k} \rightharpoonup \hat{\lambda}$ in $L^2(\Omega)$. Moreover, since the set

$$U = \left\{\lambda \in L^2(\Omega) \,|\, \lambda \geq 0, \;\; a.e.\, \mathcal{A}(\mathbf{z}), \lambda = 0, \;\; a.e.\, \mathcal{I}(\mathbf{z}) \right\}$$

is closed and convex in $L^2(\Omega)$, and thus, weakly closed in $L^2(\Omega)$, $\hat{\lambda} \in U$ as well.

Continuing, we use the fact that $G(\Omega)$ is a Hilbert space with the $L^2(\Omega)^l$-inner-product to rewrite $\mathbf{v}$. Indeed, by letting $\alpha, \beta \in \mathbb{R} \setminus \{0\}$, we can write $\mathbf{v} = \alpha\hat{\lambda}\mathbf{q} + \beta\mathbf{w}$, where $\mathbf{q}$ is described analogously to the $\mathbf{q}_n$ and $\mathbf{w} \in \left\{\hat{\lambda}\mathbf{q}\right\}^{\perp}$. Consider then that

$$\begin{aligned}
||\mathbf{v}_n - \mathbf{v}||^2_{L^2(\Omega)^l} &= ||\mathbf{v}_n||^2_{L^2(\Omega)} - 2(\mathbf{v}_n, \mathbf{v})_{L^2(\Omega)^l} + ||\mathbf{v}||^2_{L^2(\Omega)^l} \\
&= ||\lambda_n\mathbf{q}_n||^2_{L^2(\Omega)} - 2(\lambda_n\mathbf{q}_n, \alpha\hat{\lambda}\mathbf{q} + \beta\mathbf{w})_{L^2(\Omega)^l} + ||\alpha\hat{\lambda}\mathbf{q} + \beta\mathbf{w}||^2_{L^2(\Omega)^l} \\
&= ||\lambda_n\mathbf{q}_n - \alpha\hat{\lambda}\mathbf{q}||^2_{L^2(\Omega)^l} - 2\beta(\lambda_n\mathbf{q}_n, \mathbf{w})_{L^2(\Omega)^l} + ||\beta\mathbf{w}||^2_{L^2(\Omega)^l}.
\end{aligned}$$

Then since $\lambda_{n_k} \rightharpoonup \hat{\lambda}$ in $L^2(\Omega)$ and $\mathbf{w}$ is orthogonal to $\hat{\lambda}\mathbf{q}$,

$$-2\beta(\lambda_{n_k}\mathbf{q}_{n_k}, \mathbf{w})_{L^2(\Omega)^l} = -2\beta\int_{\mathcal{A}(\mathbf{z})} \lambda_{n_k}\frac{\mathbf{z}}{|\mathbf{z}|_2}\cdot\mathbf{w}d\omega \to -2\beta\int_{\mathcal{A}(\mathbf{z})} \hat{\lambda}\frac{\mathbf{z}}{|\mathbf{z}|_2}\cdot\mathbf{w}d\omega = 0.$$

However, this implies that $||\lambda_{n_k}\mathbf{q}_{n_k} - \alpha\hat{\lambda}\mathbf{q}||^2_{L^2(\Omega)^l} + ||\beta\mathbf{w}||^2_{L^2(\Omega)^l} \to 0$, which can only hold when $||\beta\mathbf{w}||^2_{L^2(\Omega)^l} = 0$ and $||\lambda_{n_k}\mathbf{q}_{n_k} - \alpha\hat{\lambda}\mathbf{q}||^2_{L^2(\Omega)^l} \to 0$. Finally, we note that

$$||\lambda_{n_k}\mathbf{q}_{n_k} - \alpha\hat{\lambda}\mathbf{q}||^2_{L^2(\Omega)^l} = \\
\int_\Omega \left|\lambda_{n_k}\mathbf{q}_{n_k} - \alpha\hat{\lambda}\mathbf{q}\right|_2^2 d\omega = \int_{\mathcal{A}(\mathbf{z})}\left|(\lambda_{n_k} - \alpha\hat{\lambda})\frac{\mathbf{z}}{|\mathbf{z}|_2}\right|_2^2 d\omega = \int_{\mathcal{A}(\mathbf{z})}\left|(\lambda_{n_k} - \alpha\hat{\lambda})\right|_2^2 d\omega.$$

Hence, $\lambda_n \to \alpha\hat{\lambda}$ in $L^2(\Omega)$, $\alpha > 0$, and $\mathbf{v}$ has the form: $\tilde{\lambda}\mathbf{q}$, where $\tilde{\lambda} \in U$. Therefore, $N^0(\mathbf{z})$ is closed.

Finally, suppose there exists $\mathbf{v}^* \in N_{K_\psi}(\mathbf{z})$ such that $\mathbf{v}^* \notin N^0(\mathbf{z})$. Then then there exists a $\mathbf{u}^* \in G(\Omega)^*$ strongly separating $\mathbf{v}^*$ from $N^0(\mathbf{z})$ (see e.g., II.38 Prop. 4 [11]). That is,

$$\langle\mathbf{u}^*, \mathbf{v}^*\rangle_{G^*,G} > 0, \quad \langle\mathbf{u}^*, \mathbf{v}\rangle_{G^*,G} \le 0, \forall\mathbf{v} \in N^0(\mathbf{z}).$$

Using again the fact that $G(\Omega)$ is a Hilbert space, we immediately identify $\mathbf{u}^*$ with its counterpart $\mathbf{u} \in G(\Omega)$ so that the previous relation becomes

$$\exists\mathbf{u} \in G(\Omega) : (\mathbf{u}, \mathbf{v}^*)_{L^2(\Omega)^l} > 0, \quad (\mathbf{u}, \mathbf{v})_{L^2(\Omega)^l} \le 0, \forall\mathbf{v} \in N^0(\mathbf{z}).$$

The first inequality, being strict, implies that $\mathbf{u} \notin T_{K_\psi}(\mathbf{z})$. However,

$$(\mathbf{u}, \mathbf{v})_{L^2(\Omega)^l} \le 0, \forall\mathbf{v} \in N^0(\mathbf{z}) \Leftrightarrow \int_{\mathcal{A}(\mathbf{z})} \lambda\frac{\mathbf{z}}{|\mathbf{z}|_2}\cdot\mathbf{u}d\omega \le 0.$$

Then since this must hold for all $\lambda$ which are non-negative almost everywhere on $\mathcal{A}(\mathbf{z})$ it holds that $\mathbf{z}\cdot\mathbf{u} \le 0$ a.e. $\mathcal{A}(\mathbf{z})$, a contradiction. Therefore, $N_{K_\psi}(\mathbf{z}) \setminus N^0(\mathbf{z}) = \emptyset$, as was to be shown. □

We immediately obtain our next result.

**Proposition 8.2 (The Tangent and Normal Cones to $M$).** *Let $y \in M$, where $M$ is defined as above. Then*

$$T_M(y) = \left\{d \in H^1_0(\Omega)\,\middle|\,\nabla y\cdot\nabla d \in T_{K_\psi}(\nabla y)\right\}$$

*and*

$$N_M(y) = \left\{-\mathrm{div}\,(\lambda\mathbf{q}) \in H^{-1}(\Omega)\,\middle|\,\lambda \in L^2(\Omega) \wedge \mathbf{q} \in G(\Omega) : \lambda\mathbf{q} \in N_{K_\psi}(\nabla y)\right\},$$

*In addition, the critical cone to $M$ at $y$ for any normal $v \in N_M(y)$ becomes*

$$\mathcal{K}(y,v) = T_M(y) \cap \{v\}^{\perp} = \left\{ d \in H_0^1(\Omega) \,\middle|\, \nabla y \cdot \nabla d \leq 0, \ a.e.\, \mathcal{A}^0(\nabla y), \nabla y \cdot \nabla d = 0, \ a.e.\, \mathcal{A}^+(\nabla y) \right\},$$

*where*

$$\mathcal{A}^0(\nabla y) := \{\omega \in \mathcal{A}(\nabla y) \,|\, \lambda(\omega) = 0\}, \quad \mathcal{A}^+(\nabla y) := \{\omega \in \mathcal{A}(\nabla y) \,|\, \lambda(\omega) > 0\},$$

*i.e., the weakly and the strongly active sets, respectively.*

*Proof.* The same argument as used in the proof of Proposition 5.2 applies to the derivation of the normal and tangent cones. For the critical cone, let $d \in T_M(y) \cap \{v\}^{\perp}$. Then

$$\nabla y \cdot \nabla d \leq 0, \ a.e.\, \mathcal{A}(\nabla y) \wedge \langle v, d \rangle_{H^{-1}, H_0^1} = 0.$$

Using the characterization of the normals $v$, it holds for all $d$ in the critical cone that

$$\langle v, d \rangle_{H^{-1}, H_0^1} = \langle -\mathrm{div}\, \lambda \mathbf{q}, d \rangle_{H^{-1}, H_0^1} = (\lambda \mathbf{q}, \nabla d)_{L^2(\Omega)^l} = \int_{\mathcal{A}(\nabla y)} \lambda \frac{\nabla y}{|\nabla y|_2} \cdot \nabla d d\omega = \int_{\mathcal{A}^+(\nabla y)} \lambda \frac{\nabla y}{|\nabla y|_2} \cdot \nabla d d\omega = 0.$$

Hence, the inclusion "$\subseteq$" holds, whereas the reverse direction is trivial and follows via a direct verification. $\square$

This last corollary leads to the final variational object needed for the stationarity conditions.

**Proposition 8.3** (**The Polar Cone** $[\mathcal{K}(y,v)]^-$). *Let $y \in M$, where $M$ is defined as above. Then*

$$[\mathcal{K}(y,v)]^- = \left\{ w \in H^{-1}(\Omega) \,\middle|\, w = -\mathrm{div}\,(\mu \mathbf{q}) : \begin{array}{l} \mu \in L^2(\Omega) : \left\{ \begin{array}{ll} \mu \geq 0, & a.e.\, \mathcal{A}^0(\nabla y) \\ \mu = 0, & a.e.\, \mathcal{I}(\nabla y) \end{array} \right. \\ \\ \mathbf{q}(\omega) \in \left\{ \begin{array}{ll} \{\nabla \bar{y}(\omega) / |\nabla \bar{y}(\omega)|_2\} & |\nabla \bar{y}(\omega)| \neq 0 \\ \mathbb{B}_1(0) & |\nabla \bar{y}(\omega)| = 0 \end{array} \right. \end{array} \right\}.$$

*Proof.* We begin by demonstrating the inclusion "$\supseteq$" for the assertion and denote the righthand side of the equation by $K^0$. Let $w \in K^0$ and consider an arbitrary $d \in \mathcal{K}(y,v)$. Then

$$\langle w, d \rangle_{H^{-1}, H_0^1} = \langle -\mathrm{div}\,(\mu \mathbf{q}), d \rangle_{H^{-1}, H_0^1} = (\mu \mathbf{q}, \nabla d)_{L^2(\Omega)^l} = \int_{\mathcal{A}^0(\nabla y)} \mu \frac{\nabla y}{|\nabla y|_2} \cdot \nabla d d\omega + \int_{\mathcal{A}^+(\nabla y)} \mu \frac{\nabla y}{|\nabla y|_2} \cdot \nabla d d\omega.$$

Continuing, we recall the characterization of $d$ provided in Proposition 8.2, which provides us with

$$\int_{\mathcal{A}^0(\nabla y)} \mu \frac{\nabla y}{|\nabla y|_2} \cdot \nabla d d\omega + \int_{\mathcal{A}^+(\nabla y)} \mu \frac{\nabla y}{|\nabla y|_2} \cdot \nabla d d\omega = \int_{\mathcal{A}^0(\nabla y)} \mu \frac{\nabla y}{|\nabla y|_2} \cdot \nabla d d\omega \leq 0.$$

Hence, the inclusion holds.

We now use an analogous argument as in the proof of Proposition 8.1 to demonstrate equality. We first need to argue that $K^0$ is closed and convex. Since the argument is identical to the one in Proposition 8.1 we only need to demonstrate closedness of $K^0$. First note that $K^0$ is the image of the set

$$L^0 := \left\{ \mu \mathbf{q} \in G(\Omega) \,\middle|\, \begin{array}{l} \mu \in L^2(\Omega) : \left\{ \begin{array}{ll} \mu \geq 0, & a.e.\, \mathcal{A}^0(\nabla y) \\ \mu = 0, & a.e.\, \mathcal{I}(\nabla y) \end{array} \right. \\ \\ \mathbf{q}(\omega) \in \left\{ \begin{array}{ll} \{\nabla \bar{y}(\omega) / |\nabla \bar{y}(\omega)|_2\} & |\nabla \bar{y}(\omega)| \neq 0 \\ \mathbb{B}_1(0) & |\nabla \bar{y}(\omega)| = 0 \end{array} \right. \end{array} \right\}.$$

under the divergence operator. Then since the data assumptions imply that $-\mathrm{div}$ is a bounded linear operator from $G(\Omega)$ into $H^{-1}(\Omega)$, it is enough to show that $L^0$ is closed. Let $\mathbf{w}_n \in L^0$ such that $\mathbf{w}_n \to \mathbf{w}$ in $L^2(\Omega)^l$. Then by the closedness of $G(\Omega)$, $\mathbf{w} \in G(\Omega)$. By definition, $\mathbf{w_n} \in L^0$ implies that there exist $\mu_n \in L^2(\Omega)$ and $\mathbf{q}_n \in G(\Omega)$ satisfying the requirements for $L^0$ such that $\mathbf{w}_n = \mu_n \mathbf{q}_n$. The rest follows analogously to the the closure argument for $N^0(\mathbf{z})$ found in the proof of Proposition 8.1. Hence, $L^0$ and, as argued above, $K^0$ are closed in their respective spaces.

Assume now there exists $w^* \in [\mathcal{K}(y,v)]^-$ such that $w^* \notin K^0$. Then there exists a $\delta \in H^1_0(\Omega)$ strongly separately the two sets, i.e.,

$$\langle w^*, \delta \rangle_{H^{-1},H^1_0} > 0,\ \langle w, \delta \rangle_{H^{-1},H^1_0} \leq 0,\ \forall w \in K^0.$$

Therefore, $\delta \notin \mathcal{K}(y,v)$. Conversely, using the definition of $K^0$ and the characterization of $\mathcal{K}(y,v)$ provided by Proposition 8.2, we obtain $\delta \in \mathcal{K}(y,v)$, a contradiction. Hence, equality holds. □

We may now provide the explicit strong stationarity conditions.

**Proposition 8.4** (**Explicit Strong Stationarity Conditions**). *Under the given data assumptions, let $(\bar{u},\bar{y})$ be a (locally) optimal solution to corresponding MPEC. Then there exist multipliers $p \in H^1_0(\Omega)$, $\mathbf{q} \in G(\Omega)$, $\mu \in L^2(\Omega)$, and $\lambda \in L^2(\Omega)$ such that*

$$0 = \nabla_u J(\bar{u},\bar{y}) + p, \tag{21}$$

$$0 = \nabla_y J(\bar{u},\bar{y}) - \mathrm{div}\,\mu\mathbf{q} - A^* p, \tag{22}$$

$$0 = A\bar{y} - \bar{u} - \mathrm{div}\,\lambda\mathbf{q}, \tag{23}$$

*where*

$$\left\{ \begin{array}{ll} \nabla\bar{y}\cdot\nabla p \leq 0 & a.e.\, \mathcal{A}^0(\nabla\bar{y}) \\ \nabla\bar{y}\cdot\nabla p = 0 & a.e.\, \mathcal{A}^+(\nabla\bar{y}) \end{array} \middle|\, \mathbf{q}(\omega) \in \left\{ \begin{array}{ll} \{\nabla\bar{y}(\omega)\,/|\nabla\bar{y}(\omega)|_2\} & |\nabla\bar{y}(\omega)| \neq 0 \\ \mathbb{B}_1(0) & |\nabla\bar{y}(\omega)| = 0 \end{array} \right. \right\},$$

$$\left\{ \begin{array}{ll} \mu \geq 0 & a.e.\, \mathcal{A}^0(\nabla\bar{y}) \\ \mu = 0 & a.e.\, \mathcal{I}(\nabla\bar{y}) \end{array} \middle|\, \begin{array}{ll} \lambda \geq 0 & a.e.\, \mathcal{A}(\nabla\bar{y}) \\ \lambda = 0 & a.e.\, \mathcal{I}(\nabla\bar{y}) \end{array} \right\}.$$

*Moreover,* $\mathcal{A}(\nabla\bar{y}) = \{\omega \in \Omega \,|\, |\nabla\bar{y}(\omega)|_2 = \psi(\omega)\}$,

$$\mathcal{A}^+(\nabla\bar{y}) = \{\omega \in \mathcal{A}(\nabla\bar{y}) \,|\, \lambda(\omega) > 0\},\ \text{and } \mathcal{A}^0(\nabla\bar{y}) = \{\omega \in \mathcal{A}(\nabla\bar{y}) \,|\, \lambda(\omega) = 0\}.$$

*Proof.* The result follows from Theorem 7.1 via Propositions 8.2, 8.1, and 8.3. □

As a final remark, we note that if the tracking functional $J(u,y)$ mentioned in Remark 4.1 were to be used in this example, then we again observe the increased regularity of the optimal control, i.e., $\bar{u} \in H^1_0(\Omega)$. Furthermore, using such a $J$ in the context of this example, we can guarantee that an optimal pair $(\bar{u},\bar{y})$ exists via Theorem 4.1.

# Appendix

### Proof of Theorem 4.1

*Proof.* We first derive that $S$ is a Lipschitz continuous function with respect to $u \in Y^*$. By Theorem 3.3.4 in [4], there exists a unique $y$ solving the variational inequality for each $u \in Y^*$. Let $(u_1, y_1)$

and $(u_2, y_2)$ be two arbitrary control-state pairs. Then by the convexity of $M$, the generalized equation used to describe $S$ can be formulated in variational form for each pair as follows:

$$\langle Ay_1 - u_1, y' - y_1\rangle_{Y^*,Y} \geq 0, \forall y' \in M, \tag{24}$$

$$\langle Ay_2 - u_2, y'' - y_2\rangle_{Y^*,Y} \geq 0, \forall y'' \in M. \tag{25}$$

Substituting $y' = y_2$ and $y'' = y_1$ into (24) and (25), respectively, and recognizing that

$$\langle Ay_2 - u_2, y'' - y_2\rangle = \langle u_2 - Ay_2, y_2 - y''\rangle_{Y^*.Y},$$

we add the two inequalities together and obtain

$$\langle Ay_1 - u_1 + u_2 - Ay_2, y_2 - y_1\rangle_{Y^*.Y} \geq 0.$$

By the coercivity of $A$, there exists a $\xi \in \mathbb{R}_+ \setminus \{0\}$ such that

$$\xi||y_2 - y_1||_Y^2 \leq \langle A(y_2 - y_1), y_2 - y_1\rangle_{Y^*.Y} \leq \langle u_2 - u_1, y_2 - y_1\rangle_{Y^*,Y} \leq ||u_2 - u_1||_{Y^*}||y_2 - y_1||_Y.$$

It follows that there exists $L = 1/\xi$ such that

$$||S(u_2) - S(u_1)||_Y \leq L||u_2 - u_1||_{Y^*}, \ \forall u_1, u_2 \in Y^*. \tag{26}$$

Therefore, we can rewrite (2) as follows:

$$\min\left\{\tilde{J}(u)\,|u \in \mathcal{U}\right\},$$

where $\tilde{J}(u) := J(u, S(u))$. Due to the continuity of $S$, $\tilde{J}$ remains coercive and bounded from below by some $K$. Therefore, the level sets of $\tilde{J}$, i.e., the sets defined by

$$\text{lev}_\gamma \tilde{J} := \left\{u \in \mathcal{U}\,\middle|\,\tilde{J}(u) \leq \gamma\right\}, \gamma \in \mathbb{R}$$

are bounded in $\mathcal{U}$ for all $\gamma \in \mathbb{R}$ (cf. Proposition 3.2.8. [4]). Now let $u_k$ be an infimizing sequence of $\tilde{J}$, that is,

$$\lim_k \tilde{J}(u_k) = \inf_{u\in\mathcal{U}} \tilde{J}(u).$$

Clearly, there exists some $\gamma_0 \in \mathbb{R}_+ \setminus \{0\}$ such that $u_k \in \text{lev}_{\gamma_0}\tilde{J}$ for all $k$ large. Since $\mathcal{U}$ is a Hilbert space and therefore reflexive, we can select a weakly converging subsequence of $\{u_k\}$, denoted by $\{u_{k_l}\}$, such that $u_{k_l} \rightharpoonup \bar{u}$. Moreover, the compactness of the embedding of $\mathcal{U}$ into $Y^*$, which itself is a Banach space, implies that there exists a further subsequence $u_{k_{l_n}} \to_{Y^*} \bar{u}$. Therefore, $y_{k_{l_n}} = S(u_{k_{l_n}}) \to S(\bar{u}) = \bar{y}$, via (26), so that the assumptions on $J$ imply

$$-\infty < K \leq \inf_{u\in\mathcal{U}} \tilde{J}(u) = \tilde{J}(\bar{u}) \leq \liminf_n \tilde{J}(u_{k_{l_n}}) = \lim_k \tilde{J}(u_k) = \inf_{u\in\mathcal{U}} \tilde{J}(u) \leq \gamma_0,$$

as was to be shown. ☐

## Acknowledgements

The authors would like to acknowledge the financial support by the DFG Research Center MATHEON Project C28, the SPP 1253 "Optimization with Partial Differential Equations", and the START Project Y 305 "Interfaces and Free Boundaries" funded by the Austrian Ministry of Science and Education and administered by the Austrian Science Fund FWF.

## References


[1] Achdou, Y., and Pironneau, O. *Computational Methods for Option Pricing*, vol. 30 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005.

[2] Adams, R. A., and Fournier, J. J.-F. *Sobolev Spaces*, second ed. Elsevier, Amsterdam, 2008.

[3] Attouch, H. *Variational Convergence for Functions and Operators*. Pitman Advanced Publishing Program, Boston, London, Melbourne, 1984.

[4] Attouch, H., Buttazzo, G., and Michaille, G. *Variational Analysis in Sobolev and BV Spaces*. MPS-SIAM Series on Optimization. SIAM and MPS, Philadelphia, 2006.

[5] Aubin, J.-P., and Frankowska, H. *Set-Valued Analysis*. Birkhäuser, Boston, 1990.

[6] Barbu, V. *Optimal Control of Variational Inequalities*, vol. 100 of *Research Notes in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1984.

[7] Bergounioux, M. Optimal control of an obstacle problem. *Appl. Math. Optim. 36*, 2 (1997), 147–172.

[8] Bergounioux, M. Optimal control of problems governed by abstract elliptic variational inequalities with state constraints. *SIAM J. Control Optim. 36*, 1 (1998), 273–289 (electronic).

[9] Bergounioux, M., and Lenhart, S. Optimal control of bilateral obstacle problems. *SIAM J. Control Optim. 43*, 1 (2004), 240–255 (electronic).

[10] Bonnans, J.-F., and Shapiro, A. *Perturbation Analysis of Optimization Problems*. Springer-Verlag, Berlin, 2000.

[11] Bourbaki, N. *Topological Vector Spaces: Chapters 1-5*. Springer-Verlag, Berlin, 2003.

[12] Dautray, R., and Lions, J.-L. *Mathematical Analysis and Numerical Methods for Science and Technology*, vol. 3. Springer-Verlag, Berlin Heidelberg, 1990.

[13] Do, C. Generalized second-order derivatives of convex functions in reflexive Banach spaces. *Transactions of the American Mathematical Society 334*, 1 (1992), 281–301.

[14] Ekeland, I., and Témam, R. *Convex Analysis and Variational Problems*, english ed., vol. 28 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999.

[15] Friedman, A. *Variational principles and Free-Boundary Problems*. John Wiley & Sons, Inc., New York, 1982.

[16] Glowinski, R. *Numerical Methods for Nonlinear Variational Problems*. Computational Physics. New York, 1984.

[17] Han, W., and Reddy, B. D. *Plasticity*, vol. 9 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, 1999.

[18] Hintermüller, M. Inverse coefficient problems for variational inequalities: optimality conditions and numerical realization. *M2AN Math. Model. Numer. Anal. 35*, 1 (2001), 129–152.

[19] Hintermüller, M. An active-set equality constrained Newton solver with feasibility restoration for inverse coefficient problems in elliptic variational inequalities. *Inverse Problems 24*, 3 (2008), 23pp.

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

[21] Hintermüller, M., and Kopacka, I. A smooth penalty approach and a nonlinear multigrid algorithm for elliptic MPECs. *Computational Optimization and Applications* (2009).

[22] Ioffe, A. Variational analysis of a composite function: A formula for the lower second order epi-derivative. *Journal of Mathematical Analysis and Applications 160* (1991), 379–405.

[23] Ito, K., and Kunisch, K. Optimal control of elliptic variational inequalities. *Appl. Math. Optim. 41*, 3 (2000), 343–364.

[24] Katchanov, L. *Élements de la Théorie de la Plasticité*. MIR. Moscow, 1975.

[25] Kikuchi, N., and Oden, J. T. *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods*, vol. 8 of *SIAM Studies in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1988.

[26] Kinderlehrer, D., and Stampacchia, G. *An Introduction to Variational Inequalities and Their Applications*. Academic Press, New York and London, 1980.

[27] Levy, A. B. Sensitivity of solutions to variational inequalities on Banach spaces. *SIAM J. Control Optim. 38*, 1 (1999), 50–60.

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

[29] Mignot, F. Contrôle dans les inéquations variationelles elliptiques. *J. Functional Analysis 22*, 2 (1976), 130–185.

[30] Mignot, F., and Puel, J.-P. Optimal control in some variational inequalities. *SIAM J. Control and Optimization 22*, 3 (1984), 466–476.

[31] Mordukhovich, B. S. *Variational Analysis and Generalized Differentiation. Vol. 1: Basic Theory*. Springer-Verlag, Berlin, 2006.

[32] Mordukhovich, B. S. *Variational Analysis and Generalized Differentiation. Vol. 2: Applications*. Springer-Verlag, Berlin, 2006.

[33] Neittaanmaki, P., Sprekels, J., and Tiba, D. *Optimization of Elliptic Systems*. Springer Monographs in Mathematics. Springer, New York, 2006.

[34] Outrata, J., Kočvara, M., and Zowe, J. *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*, vol. 28 of *Nonconvex Optimization and its Applications*. Kluwer Academic Publishers, Dordrecht, 1998.

[35] Outrata, J. V., Jařusek, J., and Stará, J. On optimality condtions in control of elliptic variational inequalities. Preprint (Submitted), 2008.

[36] Rockafellar, R. First and second-order epi-differentiability in nonlinear programming. *Trans. Amer. Math. Soc. 307* (1988), 75–108.

[37] Rockafellar, R. Proto-differentiability of set-valued mappings and its applications in optimization. *Analyse Non Linéaire* (1989), 449–482.

[38] Rockafellar, R. T., and Wets, R. J.-B. *Variational Analysis*. Springer-Verlag, Berlin, 1998.

[39] Scheel, H., and Scholtes, S. Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity. *Mathematics of Operations Research 25*, 1 (2000), 1–22.

[40] Yosida, K. *Functional Analysis*. Spinger-Verlag, Berlin, 1980.

[41] Zowe, J., and Kurcyusz, S. Regularity and stability for the mathematical programming problem in Banach spaces. *Appl. Math. Optim. 5*, 1 (1979), 49–62.