Chapter 2
Calmness as a Constraint Qualification for M-Stationarity Conditions in MPECs

René Henrion

Abstract Mathematical programs with equilibrium constraints (MPECs) represent an important class of nonlinear optimization problems. Due to their constraint set being defined as the solution of some parameter-dependent generalized equation, the application of standard constraint qualifications (CQs) from nonlinear programming to MPECs is not straightforward. Rather than turning MPECs into mathematical programs with complementarity constraints (MPCCs) and applying specially adapted CQs, we want to present here a variational-analytic approach to dual stationarity conditions for MPECs on the basis of Lipschitzian properties of the perturbed generalized equation. The focus will be on the so-called calmness property, ensuring an appropriate calculus rule for the Mordukhovich normal cone.

2.1 Introduction

This chapter is devoted to a rather self-contained introduction to the *calmness* concept of multifunctions and its application as a constraint qualification to *Mathematical Programs with Equilibrium Constraints*, MPECs for short. Here, under a constraint qualification we understand a property ensuring the derivation of dual necessary optimality conditions. We shall follow a variational-analytic approach to this problem. For this purpose, we consider an MPEC as a special case of an abstract optimization problem

$$\min\{f(x)|G(x) \in C\} \quad f : \mathbb{R}^n \to \mathbb{R}; \;\; G : \mathbb{R}^n \to \mathbb{R}^p; \;\; C \subseteq \mathbb{R}^p, \qquad (2.1)$$

where the objective f and the constraint mapping G are continuous and the set C is supposed to be closed. An obvious instance of (2.1) is a conventional nonlinear optimization problem with equality and inequality constraints, which results upon

R. Henrion (✉)
Weierstrass Institute for Applied Analysis and Stochstics, Berlin, Germany
e-mail: henrion@wias-berlin.de

D. Aussel and C. S. Lalitha (eds.), *Generalized Nash Equilibrium Problems, Bilevel Programming and MPEC*, Forum for Interdisciplinary Mathematics, https://doi.org/10.1007/978-981-10-4774-9_2

putting $C := \{0\}_{p_1} \times \mathbb{R}^{p_2}_+$ with $p_1 + p_2 = p$ and f, G being continuously differentiable.

An MPEC is an optimization problem whose constraint is given by a parameter-dependent generalized equation:

$$\min\{\varphi(x, y) | 0 \in F(x, y) + N_\Gamma(y)\} \quad \varphi : \mathbb{R}^{n+m} \to \mathbb{R},\ F : \mathbb{R}^{n+m} \to \mathbb{R}^m, \tag{2.2}$$

Here, $\Gamma \subseteq \mathbb{R}^m$ is closed and 'N' refers to an appropriate normal cone (e.g. normal cone of convex analysis if Γ is convex or Fréchet normal cone). Such problems have a wide range of applications in mechanics or economy (e.g. in the description of equilibria in electricity spot markets [3]). In order to derive dual stationarity conditions for (2.2), we provide first some introduction to some necessary concepts of nonsmooth calculus and to Lipschitzian properties of set-valued mappings. The role of calmness as a constraint qualification is illustrated then before discussing several options to check this property. Finally, M-stationarity conditions are derived and made fully explicit.

2.2 Some Tools from Variational Analysis

We recall that a set-valued mapping $\Phi : X \rightrightarrows Y$ between topological spaces X, Y is a conventional (single-valued) mapping $\Phi : X \to 2^Y$ assigning to each $x \in X$ a subset $\Phi(x) \subseteq Y$. A set-valued mapping is uniquely defined by its graph

$$\operatorname{gr} \Phi := \{(x, y) \in X \times Y \mid y \in \Phi(x)\}.$$

The inverse of Φ is defined as

$$\Phi^{-1}(y) := \{x \in X \mid y \in \Phi(x)\}.$$

2.2.1 Elements of Nondifferentiable Calculus

We recall the definition of the contingent cone and the Fréchet normal cone to a closed set:

Definition 2.1 Let $C \subseteq \mathbb{R}^n$ be closed and $\bar{x} \in C$. The **contingent cone** and the **Fréchet normal cone**, respectively, to C at $\bar{x}$ are defined as

$$T_C(\bar{x}) := \{d \in \mathbb{R}^n \mid \exists t_n \downarrow 0\ \exists x_n \in C : t_n^{-1}(x_n - \bar{x}) \to d\}$$
$$\widehat{N}_C(\bar{x}) := \{x^* \in \mathbb{R}^n \mid \langle x^*, d\rangle \leq 0\ \forall d \in T_C(\bar{x})\}.$$

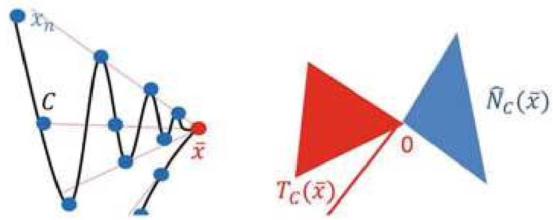


Fig. 2.1 Illustration of the contingent and the Fréchet normal cone to a closed set C at some $\bar{x} \in C$

For an illustration, see Fig. 2.1. Clearly, the contingent cone is nonconvex in general, whereas the Fréchet normal cone as its negative polar cone is always convex. In order to define $T_C, \widehat{N}_C$ as set-valued mappings, we formally put $T_C(\bar{x}) := \emptyset$ and $\widehat{N}_C(\bar{x}) := \emptyset$, whenever $\bar{x} \notin C$.

Exercise 2.1 Show the following statements:

1. If C is a closed cone, then $T_C(0) = C$.
2. $T_{\mathbb{R}^n_+}(x) = \{h \in \mathbb{R}^n \mid x_i > 0 \Rightarrow h_i \geq 0 \quad \forall i = 1, \ldots, n\}$.
3. $\widehat{N}_{\mathbb{R}^n_+}(x) = \{x^* \in \mathbb{R}^n_- \mid \langle x^*, x\rangle = 0\}$.

Since, by convention, $\widehat{N}_C(x) = \emptyset$ for $x \notin C$, 3. in Exercise 2.1 amounts to

$$\operatorname{gr} \widehat{N}_{\mathbb{R}^n_+} = \{(x, x^*) \in \mathbb{R}^n_+ \times \mathbb{R}^n_- \mid \langle x, x^*\rangle = 0\}. \tag{2.3}$$

If the closed set C happens to be convex, then the Fréchet normal cone to C coincides with the normal cone of convex analysis. The use of the Fréchet normal cone suffers from a lack of good calculus rules due to its graph being not closed in general. Therefore, it makes sense rather to consider a normal cone whose graph is the closure of the graph of the Fréchet normal cone [6]. Translating this into an explicit definition yields.

Definition 2.2 Let $C \subseteq \mathbb{R}^n$ be closed and $\bar{x} \in C$. The **Mordukhovich normal cone** to C at $\bar{x}$ is defined as

$$N_C(\bar{x}) := \{x^* | \exists (x_n, x_n^*) \rightarrow (\bar{x}, x^*) : x_n \in C, \ \ x_n^* \in \widehat{N}_C(x_n)\}$$

Figure 2.2 illustrates the computation of the Mordukhovich normal cone to some closed set C at one of its points $\bar{x}$. In a first step, Fréchet normal cones to C in a neighbourhood of $\bar{x}$ are computed. In the example, these are onedimensional linear subspaces for points $x \neq \bar{x}$ (normals to curves in the sense of classical analysis) and a solid polyhedral cone in $\bar{x}$ itself. In the second step, all limits of such Fréchet normals are aggregated according to Definition 2.2 to yield the Mordukhovich normal cone, which in contrast to the Fréchet normal cone may be nonconvex (see Fig. 2.2). Similar to the Fréchet normal cone, the Mordukhovich normal cone coincides with the one of convex analysis for convex sets.

Example 2.1 As an illustration, we compute the normal cone to different points of the set $C := \operatorname{gr} N_{\mathbb{R}_+} \subseteq \mathbb{R}^2$. By convexity of $\mathbb{R}_+$ and by (2.3), we have that

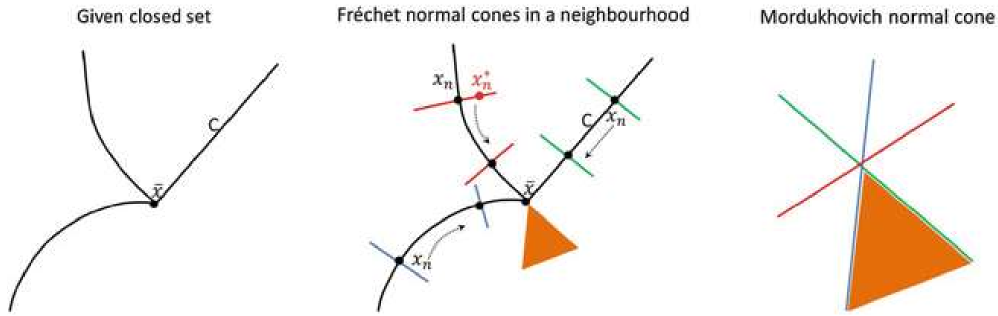
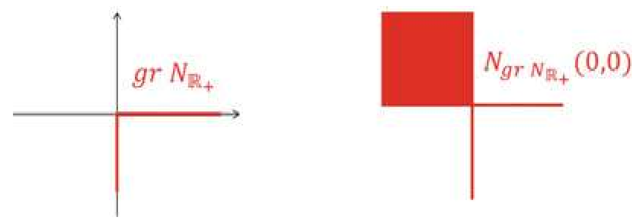


Fig. 2.2 Given closed set C and point $\bar{x} \in C$ (left), all Fréchet normal cones in a neighbourhood of $\bar{x}$ (middle) and Mordukhovich normal cone in $\bar{x}$ (right)

Fig. 2.3 Illustration of the sets $\mathrm{gr}\, N_{\mathbb{R}_+}$ (left) and $N_{\mathrm{gr}\, N_{\mathbb{R}_+}}(0,0)$ (right)



$$\mathrm{gr}\, N_{\mathbb{R}_+} = \mathrm{gr}\, \widehat{N}_{\mathbb{R}_+} = [\mathbb{R} \times \{0\}] \cup [\{0\} \times \mathbb{R}_-] \tag{2.4}$$

(see Fig. 2.3). Hence, there are three possibilities for a point $\bar{x}$ belonging to $\mathrm{gr}\, N_{\mathbb{R}_+}$: first, one may have that $\bar{x}_1 > 0$ and $\bar{x}_2 = 0$. Then, $T_{\mathrm{gr}\, N_{\mathbb{R}_+}}(\bar{x}) = \mathbb{R} \times \{0\}$ and, hence, $\widehat{N}_{\mathrm{gr}\, N_{\mathbb{R}_+}}(\bar{x}) = \{0\} \times \mathbb{R}$. Similarly, in the second case, $\bar{x}_1 = 0$ and $\bar{x}_2 > 0$, we derive that $\widehat{N}_{\mathrm{gr}\, N_{\mathbb{R}_+}}(\bar{x}) = \mathbb{R} \times \{0\}$. Finally, for the remaining third case, $\bar{x} = (0,0)$, the fact that $\mathrm{gr}\, N_{\mathbb{R}_+}$ is a closed cone implies via 1. of Exercise 2.1 that $T_{\mathrm{gr}\, N_{\mathbb{R}_+}}(0,0) = \mathrm{gr}\, N_{\mathbb{R}_+}$. Consequently,

$$\widehat{N}_{\mathrm{gr}\, N_{\mathbb{R}_+}}(0,0) = \{x^* \mid \langle x^*, h\rangle \le 0 \quad \forall h \in \mathrm{gr}\, N_{\mathbb{R}_+}\} = \mathbb{R}_- \times \mathbb{R}_+,$$

where the last equation follows from (2.4). Now, aggregating all limits of Fréchet normals in the neighbourhood of $\bar{x}$ in the sense of Definition 2.2 amounts in our example simply to collecting the union of Fréchet normal cones in the three discussed cases. Hence, at $\bar{x} = (0,0)$ we have that (see Fig. 2.3)

$$N_{\mathrm{gr}\, N_{\mathbb{R}_+}}(0,0) = [\mathbb{R}_- \times \mathbb{R}_+] \cup [\{0\} \times \mathbb{R}] \cup [\mathbb{R} \times \{0\}].$$

As the first and second cases considered above ($\bar{x} \neq (0,0)$) are stable (i.e. remain unchanged under a small perturbation of $\bar{x} \in \mathrm{gr}\, N_{\mathbb{R}_+}$), the Fréchet normal cones are locally constant around $\bar{x}$ and, consequently, coincide with the Mordukhovich normal cone.

An important property of the normal cone is that it commutes with the Cartesian product (see [6, Proposition 1.2]):

$$N_{C_1\times\cdots\times C_n}(\bar{x}_1,\ldots,\bar{x}_n) = N_{C_1}(\bar{x}_1)\times\cdots\times N_{C_n}(\bar{x}_n) \tag{2.5}$$

Exercise 2.2 Using the Cartesian product formula above, show that

$$\operatorname{gr} N_{\mathbb{R}^p_+} = L^{-1}(\Lambda),$$

where

$$L(x_1,\ldots,x_p,y_1,\ldots,y_p) := (x_1,y_1,\ldots,x_p,y_p),\quad \Lambda := \operatorname{gr} N_{\mathbb{R}_+}\times\cdots\times\operatorname{gr} N_{\mathbb{R}_+}.$$

As usual in nondifferentiable calculus, a normal cone induces a subdifferential of lower semicontinuous (possibly extended-valued) functions:

Definition 2.3 Let $f:\mathbb{R}^n\to\mathbb{R}\cup\{\infty\}$ be lower semicontinuous and define (closed) epigraph as $\operatorname{epi} f := \{(x,t)\in\mathbb{R}^{n+1} \mid t\geq f(x)\}$. Then, the **subdifferential** of f at $\bar{x}$ is defined as

$$\partial f(\bar{x}) := \{x^*\in\mathbb{R}^n | (x^*,-1)\in N_{\operatorname{epi} f}(\bar{x},f(\bar{x}))\}.$$

Analogous to the normal cone, the subdifferential is nonconvex in general, but coincides with the subdifferential of convex analysis for convex functions. If f happens to be continuously differentiable, then $\partial f(\bar{x}) = \nabla f(\bar{x})$.

Exercise 2.3 Show that for $f(x) := -|x|$ one has that $\partial f(0) = \{-1,1\}$. Hint: Verify that

$$\widehat{N}_{\operatorname{epi} f}(x,t) = \begin{cases} (0,0) & \text{if } t > f(x) \text{ or } x = t = 0\\ \mathbb{R}_+(1,-1)) & \text{if } t = f(x) \text{ and } x < 0\\ \mathbb{R}_+(-1,-1)) & \text{if } t = f(x) \text{ and } x > 0 \end{cases}$$

Using this, aggregate Fréchet normals for $(x,t)\in\operatorname{epi} f$ in a neighbourhood of $(\bar{x},f(\bar{x}))$ in order to derive that $N_{\operatorname{epi} f}(\bar{x},f(\bar{x})) = \operatorname{gr} f$. Apply Definition 2.3.

The subdifferential satisfies the following important sum rule

Theorem 2.1 (see [6], Theorem 2.33) *Let $f_1:\mathbb{R}^n\to\mathbb{R}$ be locally Lipschitzian, and let $f_2:\mathbb{R}^n\to\mathbb{R}$ be lower semicontinuous. Then,*

$$\partial(f_1+f_2)(\bar{x}) \subseteq \partial f_1(\bar{x}) + \partial f_2(\bar{x}).$$

The normal cone and the subdifferential introduced so far can be employed in order to state the following necessary optimality conditions for an abstract optimization problem:

Theorem 2.2 (see [8], Theorem 8.15) *Let $f:\mathbb{R}^n\to\mathbb{R}$ be locally Lipschitzian, and let $\bar{x}$ be a local solution of the optimization problem*

$$\min\{f(x) | x\in C\} \quad (C\subseteq\mathbb{R}^n \mathit{closed})$$

Then, $0\in\partial f(\bar{x}) + N_C(\bar{x})$.

We note that a similar formula would not be valid for Fréchet normal cones.

Finally, again based on the definition of the Mordukhovich normal cone, we introduce a concept for the derivative of a general set-valued mapping:

Definition 2.4 Let $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ have a closed graph. Fix any $(\bar{x}, \bar{y}) \in \operatorname{gr} \Phi$. Then, the **coderivative** of Φ at $(\bar{x}, \bar{y})$ is defined as a multifunction $D^*\Phi(\bar{x}, \bar{y}) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ such that

$$D^*\Phi(\bar{x}, \bar{y})(y^*) := \{x^* | (x^*, -y^*) \in N_{\operatorname{gr} \Phi}(\bar{x}, \bar{y})\}$$

One should observe that the coderivative is not just defined at an argument $\bar{x}$ of the preimage space but also needs the specification of a point $\bar{y} \in \Phi(\bar{x})$ in the image $\Phi(\bar{x})$. Indeed, the coderivative is generally a different mapping for different $\bar{y} \in \Phi(\bar{x})$ even for foxed $\bar{x}$. In case of single-valued Φ, one necessarily has $\bar{y} = \Phi(\bar{x})$, so the specification of $\bar{y}$ is omitted and one simply writes $D^*\Phi(\bar{x})$ rather than $D^*\Phi(\bar{x}, f(\bar{x}))$. It can be shown that for single-valued, continuously differentiable mappings Φ the coderivative of Φ at $(\bar{x})$ reduces to its adjoint Jacobian $D^T\Phi(\bar{x})$.

Exercise 2.4 Let $(\bar{x}, \bar{y}) \in \operatorname{gr} N_{\mathbb{R}_+}$. Using Example 2.1 for computing $N_{\operatorname{gr} N_{\mathbb{R}_+}}(\bar{x}, \bar{y})$, show that

$$\text{If } (\bar{x}, \bar{y}) = (0, 0) \Longrightarrow D^*N_{\mathbb{R}_+}(\bar{x}, \bar{y})(y^*) = \begin{cases} \mathbb{R} & \text{if } y^* = 0 \\ \{0\} & \text{if } y^* > 0 \\ \mathbb{R}_- & \text{if } y^* < 0 \end{cases}$$

$$\text{If } \bar{x} > 0,\ \bar{y} = 0 \Longrightarrow D^*N_{\mathbb{R}_+}(\bar{x}, \bar{y})(y^*) = \{0\}$$

$$\text{If } \bar{x} = 0,\ \bar{y} < 0 \Longrightarrow D^*N_{\mathbb{R}_+}(\bar{x}, \bar{y})(y^*) = \begin{cases} \mathbb{R} & \text{if } y^* = 0 \\ \emptyset & \text{if } y^* \neq 0 \end{cases}$$

The following important scalarization formula for coderivatives of single-valued mappings and subdifferentials of their components holds true:

Proposition 2.1 (see [6], Theorem 1.90) *If $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ is locally Lipschitzian, then $D^*\Phi(\bar{x})(y^*) = \partial\langle y^*, \Phi\rangle(\bar{x})$ for all $y^* \in \mathbb{R}^m$.*

2.2.2 Lipschitz Properties of Set-Valued Mappings

In this section, we consider a set-valued mapping $F : X \rightrightarrows Y$ between metric spaces, i.e. a mapping assigning to each $x \in X$ an image set $F(x) \subseteq Y$. We want to introduce two generalizations of Lipschitz properties from single-valued to set-valued mappings. We recall that a single-valued mapping $f : X \to Y$ is locally Lipschitzian at some $\bar{x} \in X$ if there exists some $L \geq 0$ such that $d(f(x_1), f(x_2)) \leq Ld(x_1, x_2)$ for all x_1, x_2 in a neighbourhood of $\bar{x}$. A strictly weaker, yet related to Lipschitzian behaviour property results from fixing one of the two arguments in the previous definition. More precisely, f is *calm* at $\bar{x}$, if $d(f(x), f(\bar{x})) \leq Ld(x, \bar{x})$ for all x in a neighbourhood of $\bar{x}$. A function which is calm but fails to be locally Lipschitz (due to unbounded slopes for pairs of points close to the fixed point) is illustrated in Fig. 2.4.

When transferring these concepts to set-valued mappings, one has to take into account first that images $F(x)$ are sets now. A straightforward generalization would be obtained by considering the Hausdorff distance between subsets of the image space (see Fig. 2.4):

$$d_H(A, B) := \max\{\sup_{a \in A} d(a, B), \sup_{b \in B} d(b, A)\} \quad \forall A, B \subseteq Y.$$

Then, for instance, local Lipschitz continuity of F at some $\bar{x} \in X$ would amount to the relation $d_H(F(x_1), F(x_2)) \leq Ld(x_1, x_2)$ for all x_1, x_2 in a neighbourhood of $\bar{x}$. However, the use of the Hausdorff distance in variational analysis has several drawbacks: first, if the considered sets are unbounded, then convergence of sets may not be well reflected (see Fig. 2.4, where 'intuitively' sets A_n converge to A while $d_H(A_n, A) = \infty$); second, the Hausdorff distance is a global measure and may even for bounded sets not well describe the local convergence of sets around a fixed point in the limit set (see Fig. 2.4).

In order to circumvent the mentioned inconveniences of the use of the Hausdorff distance, the following definitions have proven to be useful for describing the (locally)

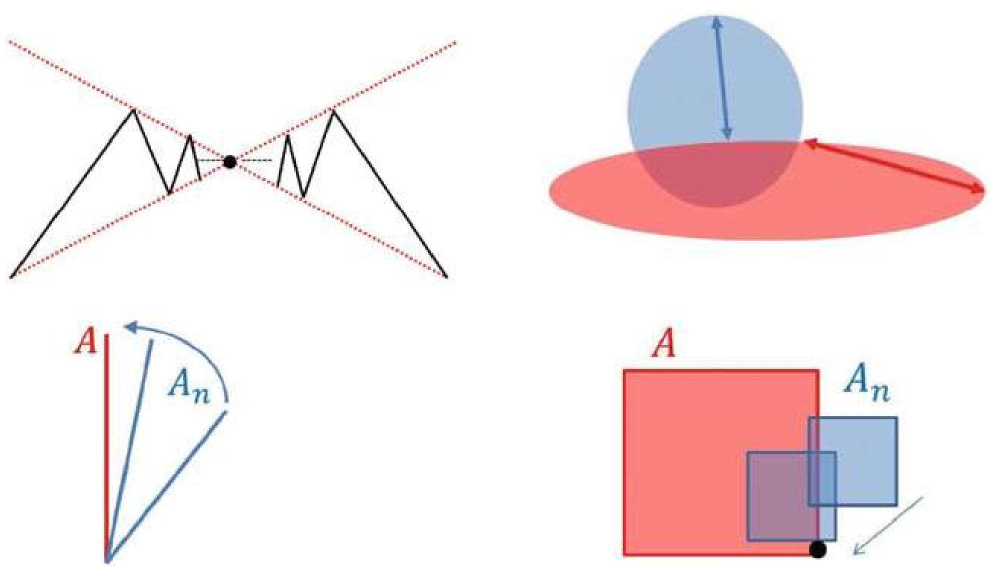


Fig. 2.4 Top left: example for a calm function not being locally Lipschitzian; top right: illustration of the Hausdorff distance as maximum of the excesses of one set over the other; bottom left: 'convergence' of sets having Hausdorff distance ∞ to the limit set; bottom right: the sequence of sets does not converge to the set A in a global sense, although it does so locally around the fixed point

Lipschitzian behaviour of a set-valued mapping $F : X \rightrightarrows Y$ between metric spaces (see, e.g. [8]):

Definition 2.5 F has the **Aubin property** at $(\bar{x}, \bar{y}) \in \operatorname{gr} F$ if there are $L, \delta > 0$ with

$$d\,(y, F\,(x_1)) \leq Ld\,(x_1, x_2) \quad \forall x_1, x_2 \in \mathbb{B}_\delta(\bar{x}),\ \ \forall y \in \left[F\,(x_2) \cap \mathbb{B}_\delta(\bar{y})\right].$$

F is said to be **calm** at $(\bar{x}, \bar{y}) \in \operatorname{gr} F$ if there are $L, \delta > 0$ such that

$$d\,(y, F\,(\bar{x})) \leq Ld\,(x, \bar{x}) \quad \forall x \in \mathbb{B}_\delta(\bar{x}),\ \ \forall y \in \left[F\,(x) \cap \mathbb{B}_\delta(\bar{y})\right].$$

Here, $\mathbb{B}_\delta(z)$ refers to the closed ball around z with radius δ.

One immediately verifies that the Aubin property and calmness presented in Definition 2.5 reduce in the case of single-valued functions to local Lipschitz continuity and calmness as introduced in the beginning of this section. Therefore, it is clear that for set-valued mappings too, calmness is strictly weaker than the Aubin property. Observe also that, in contrast to single-valued mappings, we now have not only to fix some argument $\bar{x} \in X$ but also a point $\bar{y} \in F(\bar{x})$ in the image set, because the local behaviour of F at $(\bar{x}, y)$ may be different for different $y \in F(\bar{x})$.

Exercise 2.5 Show that the mapping $F(t) := \{x \mid x^2 \geq t\}$ is calm but fails to have the Aubin property at the point (0, 0) of its graph.

2.3 Calmness and Aubin Property in Optimization Problems

2.3.1 Calmness as a Constraint Qualification for Abstract Optimization Problems

In this section, we want to derive dual necessary optimality conditions for the abstract optimization problem (2.1). Observe first that (2.1) can be compactly rewritten as $\min\{f(x)|x \in G^{-1}(C)\}$. Therefore, we may apply Theorem 2.2 to derive the following

Corollary 2.1 *Let $\bar{x}$ be a local solution of problem (2.1), where we assume that f is locally Lipschitzian, G is continuous, and C is closed. Then,* $0 \in \partial f(\bar{x}) + N_{G^{-1}(C)}(\bar{x})$.

The necessary optimality condition obtained in the Corollary is also referred to as an *abstract M-stationarity condition* because it is based on the Mordukhovich normal cone and the associated subdifferential. As the name suggests, the condition is abstract in that it does not provide yet an expression for the normal cone in terms of the data G and C of problem (2.1). Any condition providing such resolution of normal cones in terms of constraint data is usually called a *constraint qualification*. A key constraint qualification in the context of M-stationarity is calmness as introduced in Definition 2.5. This is explained by the following *preimage formula*:

Theorem 2.3 (see [2], Theorem 4.1) *Let $G : \mathbb{R}^n \to \mathbb{R}^m$ be locally Lipschitz, and let $C \subseteq \mathbb{R}^m$ be closed. If, for some $\bar{x} \in G^{-1}(C)$ the multifunction $\Psi(y) := \{x | G(x) + y \in C\}$ is calm at $(0, \bar{x})$, then*

$$N_{G^{-1}(C)}(\bar{x}) \subseteq D^*G(\bar{x})\left[N_C(G(\bar{x}))\right].$$

Combining this preimage formula with Corollary 2.1 leads immediately to the following *resolved M-stationarity conditions* for problem (2.1):

Corollary 2.2 *Let $\bar{x}$ be a local solution of problem (2.1), where we assume that f and G are locally Lipschitzian and C is closed. Under the calmness assumption for Ψ in Theorem 2.3, there exists some $v^* \in N_C(G(\bar{x}))$ such that $0 \in \partial f(\bar{x}) + D^*G(\bar{x})(v^*)$.*

In the special case that f and G are continuously differentiable and that $C = \mathbb{R}^p_+$, the last corollary yields the classical Karush–Kuhn–Tucker conditions of classical nonlinear optimization under inequality constraints. Indeed, recalling that in this smooth case $\partial f(\bar{x}) = \nabla f(\bar{x})$ and the coderivative coincides with the adjoint Jacobian $D^T G(\bar{x})$, the inclusion from Corollary 2.2 reduces to the equation $0 = \partial f(\bar{x}) + D^T G(\bar{x})v^*$. On the other hand, $v^* \in N_{\mathbb{R}^p_+} G(\bar{x}) = \widehat{N}_{\mathbb{R}^p_+} G(\bar{x})$ by convexity of $\mathbb{R}^p_+$. Now, (2.3) entails the complementarity relations

$$G(\bar{x}) \geq 0, \quad v^* \leq 0, \quad \sum_{i=1}^{p} v_i^* G_i(\bar{x}) = 0.$$

These statements suggest to compare the constraint qualification (CQ) via calmness in Theorem 2.3 with known CQs in nonlinear programming. One can show that calmness implies the *Abadie CQ* but is implied by the *Mangasarian–Fromovitz CQ* (MFCQ) which in turn is equivalent to the stronger Aubin property of Ψ in Theorem 2.3 (see Exercise 2.6). This observation from nonlinear programming already provides some idea of how calmness could work as a strictly weaker CQ in MPECs than the (easier to characterize algebraically) Aubin property.

2.3.2 Verification of Calmness and Aubin Property

As far as the Aubin property of set-valued mapping is concerned, the coderivative introduced in Definition 2.4 provides a powerful equivalent characterization via the celebrated *Mordukhovich criterion*:

Theorem 2.4 (see [8], Theorem 9.40) *Let $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ have a closed graph. Then, F has the Aubin property at $(\bar{x}, \bar{y}) \in \mathrm{gr}F$ if and only if*

$$D^*F(\bar{x},\bar{y})(0) = \{0\}$$

This criterion may be used, in order to derive an easy-to-verify algebraic characterization of the Aubin property for smooth constraint systems:

Proposition 2.2 *Consider the set-valued mapping given by the perturbation*

$$F(p) := \{x \in \mathbb{R}^n | G(x) - p \in C\}$$

of the constraint $G(x) \in C$ in the abstract optimization problem (2.1). Here, we assume that G is continuously differentiable and C is closed. Let $\bar{x}$ be a feasible point of the unperturbed problem, i.e. $G(\bar{x}) \in C$. Then,

$$F \textit{ has the Aubin property at } (0,\bar{x}) \Longleftrightarrow \mathrm{Ker}[\nabla G(\bar{x})]^T \cap N_C(G(\bar{x})) = \{0\}$$

Proof Define $\widetilde{G}(p,x) := G(x) - p$. Then, $\mathrm{gr}F = \widetilde{G}^{-1}(C)$. Clearly, the Jacobian $\nabla\widetilde{G}(0,\bar{x}) = (-I, \nabla G(\bar{x}))$ is surjective. This allows us to invoke the following preimage formula

$$N_{\widetilde{G}^{-1}(C)}(0,\bar{x}) = [\nabla\widetilde{G}(0,\bar{x})]^T N_C(\widetilde{G}(0,\bar{x}))$$

(see [8], [Exercise 6.7]), which in our special case with smooth mappings corresponds to the inclusion of Theorem 2.3 being actually satisfied as an equality. Hence,

$$N_{\mathrm{gr}F}(0,\bar{x}) = N_{\widetilde{G}^{-1}(C)}(0,\bar{x}) = \begin{pmatrix} -I \\ [\nabla G(\bar{x})]^T \end{pmatrix} N_C(\widetilde{G}(0,\bar{x})).$$

Observing that $\widetilde{G}(0,\bar{x}) = G(\bar{x})$, we derive from this last relation that for all p^* the following holds true:

$$\begin{aligned}(p^*,0) \in N_{\mathrm{gr}F}(0,\bar{x}) &\Longleftrightarrow \exists z^* \in N_C(G(\bar{x})) :\; p^* = -z^*,\; [\nabla G(\bar{x})]^T z^* = 0 \\ &\Longleftrightarrow -p^* \in \mathrm{Ker}[\nabla G(\bar{x})]^T \cap N_C(G(\bar{x})).\end{aligned}$$

By Definition 2.4, this amounts to

$$D^*F(0,\bar{x})(0) = \{p^* | (p^*,0) \in N_{\mathrm{gr}F}(0,\bar{x})\} = -\left\{\mathrm{Ker}[\nabla G(\bar{x})]^T \cap N_\Theta(G(\bar{x}))\right\}.$$

The result now follows from Theorem 2.4.

Exercise 2.6 For a smooth inequality system $G(x) \in \mathbb{R}^p_-$ the Aubin property of the perturbation mapping $F(p) := \{x \in \mathbb{R}^n | G(x) - p \in \mathbb{R}^p_+\}$ at some feasible point $\bar{x}$ is equivalent to the validity of the Mangasarian–Fromovitz CQ at $\bar{x}$, i.e. with the existence of some d such that

$$\langle \nabla G_i(\bar{x}), d\rangle < 0 \quad \forall i : G_i(\bar{x}) = 0 \tag{2.6}$$

Hint: using Proposition 2.2 and (2.3) show that F has the Aubin property at $(0, \bar{x})$ if and only if the following relation holds true:

$$[\nabla G(\bar{x})]^T \lambda = 0,\ \lambda \geq 0,\ \ \lambda_i = 0\ \forall i:\ G_i(\bar{x}) < 0 \Rightarrow \lambda = 0,$$

which by Motzkin's Theorem of the alternative is equivalent to (2.6).

If it comes to check the calmness property of a set-valued mapping, then one could of course keep using the criterion of Theorem 2.4 for the stronger Aubin property. But in this way, one might loose the potential of strictly weakening the assumptions needed, for instance, for the derivation of necessary optimality conditions. Refining the criterion of Theorem 2.4 towards calmness seems to be possible only in special cases (see [1, Theorem 3.1], [2, Theorem 3.2]). An instance, where calmness (but not necessarily the Aubin property) may always be taken for granted without further assumptions, is *polyhedral mappings*, i.e. set-valued mappings whose graph is a finite union of convex polyhedra. The following result is a slight reduction of a theorem by Robinson:

Proposition 2.3 ([7], Proposition 1) *A polyhedral set-valued mapping is calm at any point of its graph.*

Note that the graph of a polyhedral mapping need not be convex. A prototype example is the set $\operatorname{gr} N_{\mathbb{R}^p_+}$:

Example 2.2 From (2.4), we know that $\operatorname{gr} N_{\mathbb{R}^p_+}$ is the union of two polyhedra:

$$\operatorname{gr} N_{\mathbb{R}_+} = \underbrace{\left[\mathbb{R}_+ \times \{0\}\right]}_{A_0} \cup \underbrace{\left[\{0\} \times \mathbb{R}_-\right]}_{A_1}$$

so that the set Λ from Exercise 2.2 may be represented as a finite union of polyhedra because the Cartesian product of polyhedra is a polyhedron again:

$$\Lambda = \bigcup_{(i_1,\ldots,i_p)\in\{0,1\}^p} \underbrace{A_{i_1} \times \cdots \times A_{i_p}}_{\text{polyhedron}}.$$

Now, owing to Exercise 2.2 and the mapping L defined there, we arrive at

$$\operatorname{gr} N_{\mathbb{R}^p_+} = \bigcup_{(i_1,\ldots,i_p)\in\{0,1\}^p} \underbrace{L^{-1}(A_{i_1} \times \cdots \times A_{i_p})}_{\text{polyhedron}}.$$

showing that $\operatorname{gr} N_{\mathbb{R}^p_+}$ is a polyhedron as the preimage of a polyhedron under a linear mapping. Consequently, the normal cone mapping $x \mapsto N_{\mathbb{R}^p_+}$ is polyhedral.

In many applications, set-valued mappings are neither polyhedral nor satisfy the Aubin property, so that the previous approaches for verifying calmness would not apply. On the other hand, more than often a structure is present which is partially polyhedral and partially 'Aubin-like'. In such cases, the following useful characterization of structured calmness can be useful:

Theorem 2.5 ([5], Theorem 3.6) *Let $T_1 : X_1 \rightrightarrows X$ and $T_2 : X_2 \rightrightarrows X$ be multifunctions between metric spaces X_1, X_2, X. If*

1. *T_1 is calm at $(x_1, x) \in \operatorname{gr} T_1$*
2. *T_2 is calm at $(x_2, x) \in \operatorname{gr} T_2$*
3. *T_2^{-1} has the Aubin property at (x, x_2)*
4. *$T_1(x_1) \cap T_2(\cdot)$ is calm at (x_2, x),*

then the multifunction $(T_1 \cap T_2)(x_1, x_2) := T_1(x_1) \cap T_2(x_2)$ is calm at (x_1, x_2, x).

Exercise 2.7 Provide an example for two set-valued mappings being calm but without their pointwise intersection being calm likewise.

The previous exercise illustrates, why the first two conditions of Theorem 2.5 alone are not sufficient to yield the calmness of the intersection mapping.

2.4 M-Stationarity Conditions for MPECs

We consider the MPEC introduced in (2.2) and specify now that the normal cone N appearing there refers to the Mordukhovich normal cone, so for convex sets Γ it coincides with the normal cone of convex analysis. On the other hand, the use of the Mordukhovich normal cone allows an application to general closed sets and, in contrast to the Fréchet normal cone, its graph will be always closed (see Sect. 2.2.1). Observe that by passing to the concept of the graph of a multifunction, one may equivalently rewrite it as

$$\min\{\varphi(x, y) | \underbrace{(y, -F(x, y))}_{H(x,y)} \in \operatorname{gr} N_\Gamma\}, \tag{2.7}$$

which is exactly of the form of (2.1) with $f := \varphi$, $G := H$ [(as defined in (2.7)] and $C := \operatorname{gr} N_\Gamma$. This being done, we may immediately apply Corollary 2.2 on M-stationarity conditions for abstract optimization problems in order to specify them in the case of MPECs:

Proposition 2.4 *Let $(\bar{x}, \bar{y})$ be a local solution of the MPEC (2.2), where we assume that φ and F are locally Lipschitzian and Γ is closed. Then, if the mapping*

$$\Psi(p_1, p_2) := \{(x, y) | p_2 \in F(x, y) + N_\Gamma(y + p_1)\} \tag{2.8}$$

is calm at $(0, 0, \bar{x}, \bar{y})$, then there exists an MPEC multiplier $v^* \in N_{\operatorname{gr} N_\Gamma}(H(\bar{x}, \bar{y}))$ (with H as introduced in (2.7)) such that

$$0 \in \partial\varphi(\bar{x}, \bar{y}) + D^*(H(\bar{x}, \bar{y}))(v^*). \tag{2.9}$$

Proof We consider the MPEC (2.2) in its equivalent form (2.7). Since F is locally Lipschitzian, the same holds true for H. Moreover, as mentioned above, the set $C := \operatorname{gr} N_\Gamma$ is closed. Finally, we observe that the calmness assumption in our proposition amounts in graphical form to the calmness of the mapping

$$\Psi(p_1, p_2) := \{(x, y) | (p_1 + y, p_2 - F(x, y)) \in \operatorname{gr} N_\Gamma\}$$

at $(0, 0, \bar{x}, \bar{y})$. Recalling the definition of H and putting $p := (p_1, p_2)$ this can be rephrased as the calmness of the mapping $\Psi(p) := \{(x, y) | (p + H(x, y)) \in \operatorname{gr} N_\Gamma\}$ occurring in Theorem 2.3 and needed in Corollary 2.2. Summarizing, all assumptions of Corollary 2.2 are satisfied for problem (2.7) and the assertion follows.

The necessary optimality condition (2.9) is not fully efficient, yet in that it is formulated in terms of the intermediary mapping H rather than the input mapping F of (2.7). Moreover, one can simplify the calmness condition according to the following statement.

Lemma 2.1 (see [9], Proposition 5.2) *The full perturbation mapping Ψ in (2.8) is calm at $(0, 0, \bar{x}, \bar{y})$ if and only if the associated reduced perturbation mapping $\widetilde{\Psi}(p) := \{(x, y) | p \in F(x, y) + N_\Gamma(y)\}$ is calm at $(0, \bar{x}, \bar{y})$.*

Next, we develop a more handy expression for the coderivative of $H(x, y) = (y, -F(x, y))$ in (2.9). Put $H := (H_1, H_2)$. The scalarization formula of Proposition 2.1 and the sum rule of Theorem 2.1 yield that

$$\begin{aligned} D^*H(\bar{x}, \bar{y})(u^*, v^*) &= \partial\langle(u^*, v^*), (H_1, H_2)\rangle(\bar{x}, \bar{y}) \\ &\subseteq \partial\langle u^*, H_1\rangle(\bar{x}, \bar{y}) + \partial\langle v^*, H_2\rangle(\bar{x}, \bar{y}) \\ &= \{(0, u^*)\} + D^*H_2(\bar{x}, \bar{y})(v^*) \\ &= \{(0, u^*)\} + D^*(-F)(\bar{x}, \bar{y})(v^*). \end{aligned} \tag{2.10}$$

Here, we made use of the fact that $\langle u^*, H_1\rangle = \langle u^*, y\rangle$ is a linear function of (x, y) and, hence, the subdifferential reduces to its gradient. Combining (2.10) with Lemma 2.1, Proposition 2.4 yields the following M-stationarity conditions for the MPEC (2.2) with locally Lipschitzian mappings completely in terms of the input data of the problem:

Theorem 2.6 *Let $(\bar{x}, \bar{y})$ be a local solution of the MPEC (2.2), where φ and F are locally Lipschitzian and Γ is closed. Then, if the mapping $\widetilde{\Psi}$ in Lemma 2.1 is calm at $(0, \bar{x}, \bar{y})$, then there exist MPEC multipliers $(u^*, v^*) \in N_{\operatorname{gr} N_\Gamma}(\bar{y}, -F(\bar{x}, \bar{y}))$ such that*

$$0 \in \partial\varphi(\bar{x}, \bar{y}) + \{(0, u^*)\} + D^*(-F)(\bar{x}, \bar{y})(v^*).$$

In the following, we are going to specify the M-stationarity conditions of Theorem 2.6 to the case of smooth input data for the MPEC (2.2):

Corollary 2.3 *Let* $(\bar{x}, \bar{y})$ *be a local solution of the MPEC* (2.2)*, where* φ *and* F *are continuously differentiable and* Γ *is closed. Then, if the mapping* $\widetilde{\Psi}$ *in Lemma* 2.1 *is calm at* $(0, \bar{x}, \bar{y})$*, then there exists an MPEC multiplier* v^* *such that*

$$0 = \nabla_x \varphi(\bar{x}, \bar{y}) + \left[\nabla_x F(\bar{x}, \bar{y})\right]^T v^*$$
$$0 \in \nabla_y \varphi(\bar{x}, \bar{y}) + \left[\nabla_y F(\bar{x}, \bar{y})\right]^T v^* + D^* N_\Gamma(\bar{y}, -F(\bar{x}, \bar{y}))(v^*).$$

Proof In the case of smooth data, the subdifferential and coderivative, respectively, reduce to

$$\partial \varphi(\bar{x}, \bar{y}) = (\nabla_x \varphi(\bar{x}, \bar{y}), \nabla_y \varphi(\bar{x}, \bar{y}))$$
$$D^*(-F)(\bar{x}, \bar{y})(v^*) = (-\left[\nabla_x F(\bar{x}, \bar{y})\right]^T v^*, -\left[\nabla_y F(\bar{x}, \bar{y})\right]^T v^*).$$

Now, Theorem 2.6 guarantees the existence of $(u^*, v^*) \in N_{\mathrm{gr} N_\Gamma}(\bar{y}, -F(\bar{x}, \bar{y}))$ such that

$$0 = \nabla_x \varphi(\bar{x}, \bar{y}) - \left[\nabla_x F(\bar{x}, \bar{y})\right]^T v^*$$
$$0 = \nabla_y \varphi(\bar{x}, \bar{y}) - \left[\nabla_y F(\bar{x}, \bar{y})\right]^T v^* + u^*$$

Since $(u^*, v^*) \in N_{\mathrm{gr} N_\Gamma}(\bar{y}, -F(\bar{x}, \bar{y}))$ if and only if $u^* \in D^* N_\Gamma(\bar{y}, -F(\bar{x}, \bar{y}))(-v^*)$ (see Definition 2.4), we can substitute for the second multiplier u^* by turning the second equation above into an inclusion.

We note that the preceding Corollary has been proven first in [10, Theorem 3.2] using a different way of reasoning. Looking at Corollary 2.3, there remain two issues to be clarified for an efficient application of the obtained M-stationarity conditions: first, the calmness of the perturbation mapping $\widetilde{\Psi}$ has to be verified, and second, explicit formulae for the coderivative $D^* N_\Gamma$ have to be found. This will be the object of the following two sections.

2.5 Verification of Calmness for the Perturbation Mapping

2.5.1 Using the Aubin Property

We start by providing an algebraic condition ensuring the calmness of the mapping $\widetilde{\Psi}$ introduced in Lemma 2.1 via checking the stronger Aubin property for the perturbation mapping (2.8). Observe first the general relation $\widetilde{\Psi}(p_2) = \Psi(0, p_2)$ between both mappings. This means that the perturbation of Ψ is richer than that of $\widetilde{\Psi}$ while $\widetilde{\Psi}(0) = \Psi(0, 0)$. As a consequence, Ψ having the Aubin property at $(0, 0, \bar{x}, \bar{y})$ would imply $\widetilde{\Psi}$ having the Aubin property at $(0, \bar{x}, \bar{y})$ which in turn would imply

$\widetilde{\Psi}$ being calm at $(0, \bar{x}, \bar{y})$. Therefore, the following proposition yields a sufficient algebraic condition for the calmness of $\widetilde{\Psi}$ as required in Corollary 2.3:

Proposition 2.5 *Let F be continuously differentiable, let Γ be closed, and let $(\bar{x}, \bar{y})$ be such that $0 \in F(\bar{x}, \bar{y}) + N_\Gamma(\bar{y})\}$. Then, the perturbation mapping Ψ defined in (2.8) has the Aubin property at $(0, 0, \bar{x}, \bar{y})$ if and only if the following implication holds true:*

$$\left[\nabla_x F(\bar{x}, \bar{y})\right]^T v^* = 0, \;\left[\nabla_y F(\bar{x}, \bar{y})\right]^T v^* \in D^* N_\Gamma(\bar{y}, -F(\bar{x}, \bar{y}))(-v^*) \Longrightarrow v^* = 0$$

As a consequence, this implication guarantees the calmness of $\widetilde{\Psi}$ at $(0, \bar{x}, \bar{y})$ as required in Corollary 2.3.

Proof By (2.8), Ψ may be rewritten in graphical form as

$$\Psi(p_1, p_1) := \{(x, y) | H(x, y) - (p_1, p_1) \in \operatorname{gr} N_\Gamma\},$$

where H is defined in (2.7). Now, by Proposition 2.2, Ψ has the Aubin property at $(0, 0, \bar{x}, \bar{y})$ if and only if

$$\operatorname{Ker} \begin{pmatrix} 0 - \left[\nabla_x F(\bar{x}, \bar{y})\right]^T \\ I \; - \left[\nabla_y F(\bar{x}, \bar{y})\right]^T \end{pmatrix} \cap N_{\operatorname{gr} N_\Gamma}(\bar{y}, -F(\bar{x}, \bar{y})) = \{0\}.$$

which is equivalent to the implication

$$\begin{array}{r} \left[\nabla_x F(\bar{x}, \bar{y})\right]^T v^* = 0, \quad u^* - \left[\nabla_y F(\bar{x}, \bar{y})\right]^T v^* = 0, \quad u^* \in D^* N_\Gamma(\bar{y}, -F(\bar{x}, \bar{y}))(-v^*) \\ \Longrightarrow u^* = 0, \; v^* = 0. \end{array}$$

This yields the assertion.

Of course, the application of Proposition 2.5 hinges on concrete formulae for the coderivative $D^* N_\Gamma$. Possibilities to do so will be discussed in Sect. 2.6. Alternatively, one could try to check the Aubin property of $\widetilde{\Psi}$ directly using the definition in order to deduce its calmness. This is illustrated in the following example:

Example 2.3 Let $F : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$ be given by $F(x, y_1, y_2) := (0, 1)$ and

$$\Gamma := \{y \in \mathbb{R}^2 | y_2 \geq 0, \; y_2 \geq y_1^2\}.$$

Then, $\widetilde{\Psi}$ introduced in Lemma 2.1 takes the form

$$\widetilde{\Psi}(p_1, p_2) = \{(x, y_1, y_2) | (p_1, p_2 - 1) \in N_\Gamma(y)\}$$

We verify the Aubin property of $\widetilde{\Psi}$ at the point $(\bar{p}_1, \bar{p}_2, \bar{x}, \bar{y}_1, \bar{y}_2) := (0, 0, 0, 0, 0)$ which belongs to $\operatorname{gr} \widetilde{\Psi}$ due to $(0, -1) \in N_\Gamma(y)$. If $y \in \operatorname{int} \Gamma$, then $N_\Gamma(y_1, y_2) =$

$\{(0,0)\}$. Therefore, $(y_1, y_2) \in \mathrm{bd}\,\Gamma$ for all $(x, y_1, y_2) \in \widetilde{\Psi}(p_1, p_2)$ and (p_1, p_2) close to $(\bar{p}_1, \bar{p}_2)$. In particular, $y_2 = y_1^2$. The first inequality in the definition of Γ has been arranged to be redundant. Hence, $N_\Gamma(y) = \mathbb{R}_+\{(2y_1, -1)\}$ for all $(y_1, y_2) \in \mathrm{bd}\,\Gamma$. This implies that

$$(p_1, p_2 - 1) = \lambda(y)(2y_1, -1)$$

with some function $\lambda(y) \geq 0$ for all (p_1, p_2) close to $(\bar{p}_1, \bar{p}_2)$ and $(x, y_1, y_2) \in \widetilde{\Psi}(p_1, p_2)$. A comparison of components along with $y_2 = y_1^2$ yields the relations

$$\lambda(y) = 1 - p_2, \quad y_1 = p_1/2(1 - p_2), \quad y_2 = (p_1/2(1 - p_2))^2.$$

Consequently, for p close to $\bar{p}$ we have that

$$\widetilde{\Psi}(p) = \{(x, y) | y_1 = p_1/2(1 - p_2),\ y_2 = (p_1/2(1 - p_2))^2\}.$$

Clearly, the images of $\widetilde{\Psi}$ do not involve x. Moreover, the y-components are locally Lipschitzian functions of (p_1, p_2) in a neighbourhood of $(\bar{p}_1, \bar{p}_2)$. Therefore, $\widetilde{\Psi}$ has the Aubin property at $(\bar{p}_1, \bar{p}_2, \bar{x}, \bar{y}_1, \bar{y}_2)$.

2.5.2 *Using Polyhedrality or Structured Calmness*

If the MPEC (2.2) is governed by a linear generalized equation (i.e. F is affine linear and Γ is a convex polyhedron), then calmness of the perturbation mapping $\widetilde{\Psi}$ introduced in Lemma 2.1 comes for free and, hence, the M-stationarity conditions of Corollary 2.3 can be derived without any further assumptions. More precisely, we have the following result:

Proposition 2.6 *If in the MPEC (2.2) (or (2.7), respectively) Γ is a polyhedron and $F(x, y) = Ax + By + c$ is an affine linear mapping, then the perturbation mapping $\widetilde{\Psi}$ introduced in Lemma 2.1 is calm at all points of its graph.*

Proof By definition of $\widetilde{\Psi}$, it holds that

$$\mathrm{gr}\,\widetilde{\Psi} = \{(p, x, y) | \underbrace{(y, p - Ax - By - c)}_{H(p,x,y)} \in \mathrm{gr}\,N_\Gamma\} = H^{-1}(\mathrm{gr}\,N_\Gamma)$$

Since H is an affine linear mapping, it will be sufficient to verify that $\mathrm{gr}\,N_\Gamma$ is a finite union of polyhedra, because then so is $\mathrm{gr}\,\widetilde{\Psi}$ and the result follows from Proposition 2.3.

In order to carry out this verification, we describe the polyhedron Γ explicitly as the solution of a finite inequality system: $\Gamma = \{y|Cy \le d\}$. It is well known that the normal cone to a polyhedron of such description calculates as

$$N_\Gamma = C^T N_{\mathbb{R}^p_-}(Cy - d).$$

It follows that

$$(y, z) \in \operatorname{gr} N_\Gamma \Leftrightarrow z \in N_\Gamma(y) \Leftrightarrow \exists \lambda \in N_{\mathbb{R}^p_-}(Cy - d) : \, z = C^T \lambda.$$

Therefore, $\operatorname{gr} N_\Gamma = P(\Theta)$ with $P(y, z, \lambda) := (y, z)$ and

$$\Theta := \{(y, z, \lambda)|z = C^T \lambda, \;\; (\lambda, Cy - d) \in \operatorname{gr} N_{\mathbb{R}^p_+}\}.$$

Here, we exploited that

$$x^* \in N_{\mathbb{R}^p_-}(x) \Leftrightarrow x \in N_{\mathbb{R}^p_+}(x^*). \tag{2.11}$$

Defining

$$H(y, z, \lambda) := (z - C^T \lambda, \lambda, Cy - d),$$

we then have that $\Theta = H^{-1}\left(\{0\} \times \operatorname{gr} N_{\mathbb{R}^p_+}\right)$. Thus, by Example 2.2, Θ is the preimage of a finite union of polyhedra under a linear mapping and as such is a finite union $\Theta = \cup_{i=1}^q A_i$ of certain polyhedra A_i. But then,

$$\operatorname{gr} N_\Gamma = P(\cup_{i=1}^q A_i) = \cup_{i=1}^q P(A_i).$$

As the projection of a polyhedron is a polyhedron again, Γ is a polyhedral map.

Finally, we give an idea about how to employ structured calmness by formulating without proof a result which can be derived from Theorem 2.5 along the lines of [3, Theorem 7.1]:

Theorem 2.7 *For the mapping $\widetilde{\Psi}(p)$ defined in Lemma 2.1 fix any $(\bar{x}, \bar{y}) \in \widetilde{\Psi}(0)$. Assume that Γ is polyhedral and*

$$F(x, y) = \begin{pmatrix} F_1(x, y) \\ F_2(y) \end{pmatrix} \quad \textit{with } F_2 \textit{ affine linear and } \nabla_x F_1(\bar{x}, \bar{y}) \textit{ surjective.}$$

Then, $\widetilde{\Psi}$ is calm at $(0, \bar{x}, \bar{y})$.

Note that in this theorem, neither the stronger Aubin property (or its equivalent characterization via Proposition 2.5) nor the (full) affine linearity of the mapping F as in Proposition 2.6 is required.

2.6 Coderivative Formulae and Fully Explicit M-Stationarity Conditions

After providing various possibilities of verifying the calmness property of the perturbation mapping, the missing link for an efficient application of Corollary 2.3 is the computation of the coderivative to the normal cone mapping of the set Γ. Several results in this direction are known. We content ourselves here with the case of Γ being described by a finite system of smooth inequalities, i.e.

$$\Gamma := \{y \in \mathbb{R}^m | g_i(y) \leq 0 \ (i = 1, \ldots, p)\} \tag{2.12}$$

where the $g = (g_i)$ is a twice continuously differentiable. As unbinding constraints are of no interest for a local analysis, we assume without loss of generality that $g(\bar{y}) = 0$ at some fixed point of interest $\bar{y} \in \Gamma$. The following result is a consequence of a chain rule for second-order subdifferentials presented in [6, Theorem 1.127]:

Theorem 2.8 *In (2.12), fix any $\bar{y} \in \Gamma$ and $\bar{v} \in N_\Gamma(\bar{y})$. Assume that $g(\bar{y}) = 0$. If $\nabla g(\bar{y})$ is surjective, then*

$$D^* N_\Gamma(\bar{y}, \bar{v})(v^*) = \left(\sum_{i=1}^{p} \bar{\lambda}_i \nabla^2 g_i(\bar{y}) \right) v^* + \left[\nabla g\,(\bar{y}) \right]^T D^* N_{\mathbb{R}^p_-} \left(g(\bar{y}), \bar{\lambda} \right) \left(\nabla g\,(\bar{y})\, v^* \right).$$

Here, $\bar{\lambda}$ is the unique multiplier satisfying $\bar{v} = \nabla^T g(\bar{y})\lambda$.

Given Theorem 2.8, the last step needed for calculating the coderivative of N_Γ consists in specifying the coderivative of $N_{\mathbb{R}^p_-}$. We will only need the formula evaluated at points $(0, \bar{y})$ here.

Lemma 2.2 *Let $\bar{y} \in N_{\mathbb{R}^p_-}(0)$. Then,*

$$D^* N_{\mathbb{R}^p_-}(0, \bar{y})(y^*) = \emptyset \quad \textit{if there exists some } i \textit{ such that } \bar{y}_i > 0,\ y_i^* \neq 0$$

Otherwise:

$$D^* N_{\mathbb{R}^p_-}(0, \bar{y})(y^*) = \left\{ x^* \,\middle|\, \begin{array}{l} x_i^* = 0 \textit{ if } \bar{y}_i = 0,\ y_i^* < 0 \\ x_i^* \geq 0 \textit{ if } \bar{y}_i = 0,\ y_i^* > 0 \end{array} \right\}.$$

Proof From (2.11) and Exercise 2.2, we conclude that $\mathrm{gr}\, N_{\mathbb{R}^p_-} = \widetilde{L}^{-1}(\Lambda)$, where

$$\widetilde{L}(x_1, \ldots, x_p, y_1, \ldots, y_p) := (y_1, x_1, \ldots, y_p, x_p), \quad \Lambda := \mathrm{gr}\, N_{\mathbb{R}_+} \times \cdots \times \mathrm{gr}\, N_{\mathbb{R}_+}.$$

Clearly, $\widetilde{L}$ is surjective (actually regular) and $\widetilde{L} = \widetilde{L}^{-1} = \widetilde{L}^T$. Therefore, as in the proof of Proposition 2.2, we are allowed to invoke the preimage formula from [8, Exercise 6.7] in order to verify that

$$N_{\mathrm{gr}\, N_{\mathbb{R}^p_-}}(x, y) = N_{\widetilde{L}^{-1}(\Lambda)}(x, y) = \widetilde{L} N_\Lambda(\widetilde{L}(x, y)).$$

Hence, by (2.5),

$$
\begin{aligned}
& (x^*, y^*) \in N_{\operatorname{gr} N_{\mathbb{R}^p_-}}(x, y) \\
\Longleftrightarrow\ & \widetilde{L}^{-1}(x^*, y^*) \in N_\Lambda(\widetilde{L}(x, y)) \\
\Longleftrightarrow\ & (y_1^*, x_1^*, \ldots, y_p^*, x_p^*) \in N_{\operatorname{gr} N_{\mathbb{R}_+}}(y_1, x_1) \times \cdots \times N_{\operatorname{gr} N_{\mathbb{R}_+}}(y_p, x_p).
\end{aligned}
$$

In other words, $x^* \in D^*N_{\mathbb{R}^p_-}(0, \bar{y})(y^*)$ if and only if $-y_i^* \in D^*N_{\mathbb{R}_+}(\bar{y}_i, 0)(-x_i^*)$ for all $i = 1, \ldots p$. Now, using Exercise 2.4, one arrives at the asserted formula: Indeed, to see it for example for the first statement, let there exist some i such that $\bar{y}_i > 0$, $y_i^* \neq 0$. If there was some $x^* \in D^*N_{\mathbb{R}^p_-}(0, \bar{y})(y^*)$, then $-y_i^* \in D^*N_{\mathbb{R}_+}(\bar{y}_i, 0)(-x_i^*)$. Then, the second case in Exercise 2.4 yields the contradiction $y_i^* = 0$. Hence, we infer the desired statement $D^*N_{\mathbb{R}^p_-}(0, \bar{y})(y^*) = \emptyset$. The second asserted statement follows similarly.

The finally obtained coderivative formula allows us to combine Theorem 2.8 with Lemma 2.2 in order to make the M-stationarity conditions of Corollary 2.3 fully explicit:

Theorem 2.9 *Let $(\bar{x}, \bar{y})$ be a local solution to the MPEC with smooth data*

$$\min\{\varphi(x, y) | 0 \in F(x, y) + N_\Gamma(y)\}, \quad \Gamma := \{y \in \mathbb{R}^p | g_i(y) \le 0\ (i = 1, \ldots, p)\},$$

where φ, F are once and g is twice continuously differentiable. Assume that $g(\bar{y}) = 0$, that $\nabla g(\bar{y})$ is surjective and that the perturbation mapping $\widetilde{\Psi}$ from Lemma 2.1 is calm at $(0, \bar{x}, \bar{y})$. Then, there exist MPEC multipliers u^*, v^* such that

$$
\begin{aligned}
0 &= \nabla_x \varphi(\bar{x}, \bar{y}) + \left[\nabla_x F(\bar{x}, \bar{y})\right]^T v^* \\
0 &= \nabla_y \varphi(\bar{x}, \bar{y}) + \left(\left[\nabla_y F(\bar{x}, \bar{y})\right]^T + \sum_{i=1}^p \bar{\lambda}_i \nabla^2 g_i(\bar{y})\right) v^* + \left[\nabla g\,(\bar{y})\right]^T u^* \\
0 &= \nabla g_i(\bar{y}) v^* \quad \forall i : \bar{\lambda}_i > 0 \\
0 &= u_i^* \quad \forall i : \bar{\lambda}_i = 0,\ \nabla g_i(\bar{y}) v^* < 0 \\
0 &\le u_i^* \quad \forall i : \bar{\lambda}_i = 0,\ \nabla g_i(\bar{y}) v^* > 0
\end{aligned}
$$

Here, $\bar{\lambda}$ is the unique solution of $F(\bar{x}, \bar{y}) = \left[\nabla g(\bar{y})\right]^T \bar{\lambda}$.

Proof By Corollary 2.3, there exist multipliers w^*, v^* such that the first of our asserted equations and

$$0 = \nabla_y \varphi(\bar{x}, \bar{y}) - \left[\nabla_y F(\bar{x}, \bar{y})\right]^T v^* + w^*, \quad w^* \in D^*N_\Gamma(\bar{y}, -F(\bar{x}, \bar{y}))(v^*)$$

hold true. By Theorem, 2.8, there exists some $u^* \in D^*N_{\mathbb{R}^p_-}\left(g(\bar{y}), \bar{\lambda})\right)(\nabla g(\bar{y})\, v^*)$ with

$$w^* = \left(\sum_{i=1}^{p} \bar{\lambda}_i \nabla^2 g_i(\bar{y})\right) v^* + \left[\nabla g(\bar{y})\right]^T u^*$$

yielding the second of the asserted equations. Now, Lemma 2.2 provides the last asserted relations of Theorem upon recalling that $g(\bar{y}) = 0$ and that

$$u^* \in D^*N_{\mathbb{R}^p_-}\left(g(\bar{y}), \bar{\lambda})\right)\left(\nabla g(\bar{y})\, v^*\right) \neq \emptyset.$$

In the special case of a polyhedral set Γ, the surjectivity condition $\nabla g(\bar{y})$ in Theorem 2.8 can be dispensed with and a precise coderivative formula is available too. More precisely, let $\Gamma := \{x \in \mathbb{R}^n | Ax \leq b\}$ for some (q, n)-matrix A. Denote the rows of A by a_i. Fix $\bar{x} \in \Gamma$ and $\bar{v} \in N_\Gamma(\bar{x})$, i.e. $\bar{v} = A^T\lambda$ for some $\lambda \in \mathbb{R}^q_+$. For each $x \in \Gamma$ let $I(x) := \{i | a_i x = b_i\}$. Define the family of active index sets as

$$\mathscr{I} := \{I \subseteq \{1, \ldots, q\} | \, \exists x \in \Gamma : I = I(x)\}$$

Then, the following coderivative formula holds true for N_Γ:

Theorem 2.10 ([4], Proposition 3.2)

$$D^*N_\Gamma(\bar{x}, \bar{v})(v^*) = \left\{x^* \,\middle|\, \left(x^*, -v^*\right) \in \bigcup_{J \subseteq I_1 \subseteq I_2 \subseteq I(\bar{x})} P_{I_1,I_2} \times Q_{I_1,I_2}\right\},$$

Here,

$$\begin{aligned} P_{I_1,I_2} &= \text{con}\,\{a_i | i \in \chi(I_2) \backslash I_1\} + \text{span}\,\{a_i | i \in I_1\} \\ Q_{I_1,I_2} &= \{h \in \mathbb{R}^n | \, \langle a_i, h\rangle = 0 \quad (i \in I_1), \quad \langle a_i, h\rangle \leq 0 \quad (i \in \chi(I_2) \backslash I_1)\} \end{aligned}$$

and 'con' and 'span' refer to the convex conic and linear hulls, respectively. Moreover,

$$J := \{j \in I | \lambda_j > 0\} \quad \text{and} \quad \chi(I') := \cap\{J \in \mathscr{I} | I' \subseteq J\} \quad \forall I' \subseteq \{1, \ldots, q\}.$$

This last theorem can be combined, for instance, with Proposition 2.6 in order to derive fully explicit M-stationarity conditions in the spirit of Theorem 2.9 without any further assumptions (with the perturbation mapping $\widetilde{\Psi}$ being automatically calm thanks to Proposition 2.6).

References

1. Henrion, R., Outrata, J.: A subdifferential criterion for calmness of multifunctions. J. Math. Anal. Appl. **258**, 110–130 (2001)
2. Henrion, R., Jourani, A., Outrata, J.V.: On the calmness of a class of multifunctions. SIAM J. Optim. **13**, 603–618 (2002)
3. Henrion, R., Outrata, J., Surowiec, T.: Analysis of M-stationary points to an EPEC modeling oligopolistic competition in an electricity spot market. ESAIM Control Optim. Calc. Var. **18**, 295–317 (2012)
4. Henrion, R., Römisch, W.: On M-stationary points for a stochastic equilibrium problem under equilibrium constraints in electricity spot market modeling. Appl. Math. **52**, 473–494 (2007)
5. Klatte, D., Kummer, B.: Constrained minima and Lipschitzian penalties in metric mpaces. SIAM J. Optim. **13**, 619–633 (2002)
6. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation. Springer, Berlin (2006)
7. Robinson, S.M.: Some continuity properties of polyhedral multifunctions. Math. Program. Stud. **14**, 206–214 (1976)
8. Rockafellar, R.T., Wets, J.-B.: Variational Analysis. Springer, Berlin (1998)
9. Surowiec, T.: Explicit stationarity conditions and solution characterization for equilibrium problems with equilibrium constraints. Ph.D. thesis, Humboldt University Berlin (2010)
10. Ye, J.J., Ye, X.Y.: Necessary optimality conditions for optimization problems with variational inequality constraints. Math. Oper. Res. **22**, 977–997 (1997)