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

Preprint ISSN 2198-5855

On M-stationarity conditions in MPECs and the associated qualification conditions

Lukáš Adam[1], René Henrion[2], Jiří Outrata[1]

submitted: January 22, 2016

[1] ÚTIA, Czech Academy of Sciences
Pod Vodárenskou věží 4
18208 Prague
Czech Republic
E-Mail: adam@utia.cas.cz
outrata@utia.cas.cz

[2] Weierstrass Institute
Mohrenstrasse 39
10117 Berlin
Germany
E-Mail: rene.henrion@wias-berlin.de

No. 2215
Berlin 2016

2010 *Mathematics Subject Classification.* 65K10, 90C30, 90C31, 90C46.

Key words and phrases. Mathematical programs with equilibrium constraints, optimality conditions, constraint qualification, calmness, perturbation mapping.

This work was partially supported by the Grant Agency of the Czech Republic under grant 15-00735S (L. Adam) and in part by the Australian Research Council under grant DP-160100854 (R. Henrion, J. Outrata).

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

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

Abstract

Depending on whether a mathematical program with equilibrium constraints (MPEC) is considered in its original or its enhanced (via KKT conditions) form, the assumed constraint qualifications (CQs) as well as the derived necessary optimality conditions may differ significantly. In this paper, we study this issue when imposing one of the weakest possible CQs, namely the calmness of the perturbation mapping associated with the respective generalized equations in both forms of the MPEC. It is well known that the calmness property allows one to derive so-called M-stationarity conditions. The strength of assumptions and conclusions in the two forms of the MPEC is strongly related with the CQs on the 'lower level' imposed on the set whose normal cone appears in the generalized equation. For instance, under just the Mangasarian-Fromovitz CQ (a minimum assumption required for this set), the calmness properties of the original and the enhanced perturbation mapping are drastically different. They become identical in the case of a polyhedral set or when adding the Full Rank CQ. On the other hand, the resulting optimality conditions are affected too. If the considered set even satisfies the Linear Independence CQ, both the calmness assumption and the derived optimality conditions are fully equivalent for the original and the enhanced form of the MPEC. A compilation of practically relevant consequences of our analysis in the derivation of necessary optimality conditions is provided in the main Theorem 4.3. The obtained results are finally applied to MPECs with structured equilibria.

1 Introduction

Starting with [22], efficient necessary optimality conditions for various types of *mathematical programs with equilibrium constraints* (MPECs) have been developed on the basis of the generalized differential calculus of Mordukhovich, e.g. [13, 15, 16, 21]. Following [19], we speak about M-stationarity conditions. Let us consider an MPEC of the form

$$\begin{aligned} &\underset{x,y}{\text{minimize}}\ \varphi(x,y) \\ &\text{subject to}\ \ 0 \in F(x,y) + \hat{N}_\Gamma(y), \\ &\qquad\qquad\ \ x \in \omega, \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^n$ is the *control* variable, $y \in \mathbb{R}^m$ is the *state* variable, $\varphi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is the objective, $\omega \subset \mathbb{R}^n$ is the (closed) set of admissible controls, $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ is a continuously differentiable mapping, and the constraint set $\Gamma \subset \mathbb{R}^m$ is given by inequalities, i.e.,

$$\Gamma = \{y \in \mathbb{R}^m \mid q_i(y) \le 0,\ i = 1,\dots,s\} \tag{2}$$

with a twice continuously differentiable mapping $q : \mathbb{R}^m \to \mathbb{R}^s$. Further, $\hat{N}_\Gamma$ refers to the *regular (Fréchet) normal* cone (see Definition 2.1).

Let $(\bar{x},\bar{y})$ be a (local) solution of (1). Throughout this paper, we shall assume that Γ satisfies the *Mangasarian-Fromovitz Constraint Qualification* (MFCQ) at $\bar{y}$ (see Definition 2.4). As a consequence, one has the representation

$$\hat{N}_\Gamma(y) = N_\Gamma(\bar{y}) = (\nabla q(y))^\top N_{\mathbb{R}^s_-}(q(y))$$

on a neighborhood of $\bar{y}$ so that the following equivalence holds true for the *generalized equation* (GE) in (1):

$$0 \in F(x,y) + N_\Gamma(y) \Leftrightarrow \exists \lambda : 0 \in H(x,y,\lambda) + N_{\mathbb{R}^m \times \mathbb{R}^s_+}(y,\lambda), \tag{3}$$

provided y is close to $\bar{y}$ and $H(x,y,\lambda) := (F(x,y) + (\nabla q(y))^\top \lambda, -q(y))$. This relation suggests also to consider the *enhanced* MPEC

$$\begin{aligned} &\underset{x,y,\lambda}{\text{minimize}} \ \varphi(x,y) \\ &\text{subject to} \ \ 0 \in H(x,y,\lambda) + N_{\mathbb{R}^m \times \mathbb{R}^s_+}(y,\lambda), \\ &\qquad\qquad\quad x \in \omega \end{aligned} \tag{4}$$

in variables (x,y,λ). The GE in (4) has a substantially simpler constraint set than the GE in (1). As the price for it, one has to do with an additional variable λ, not arising in the objective. Let us introduce the multifunction $\Lambda : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^s$ by

$$\Lambda(p,x,y) := \left\{ \lambda \in \mathbb{R}^s \,\middle|\, p \in F(x,y) + (\nabla q(y))^\top \lambda,\ q(y) \in N_{\mathbb{R}^s_+}(\lambda) \right\}$$

so that $\Lambda(p,x,y)$ is the set of Lagrange multipliers associated with a pair (x,y), feasible with respect to the canonically perturbed GE from (1). It is well-known (and easy to prove) that under the posed constraint qualification we have that $\Lambda(0,\bar{x},\bar{y}) \neq \emptyset$ and $(\bar{x},\bar{y})$ is a local solution to problems (1) if and only if $(\bar{x},\bar{y}.\lambda)$ is a local solution to (4) for all $\lambda \in \Lambda(0,\bar{x},\bar{y})$. Likewise, it is known that for a local solution $(\bar{x},\bar{y},\bar{\lambda})$ of (4) the pair $(\bar{x},\bar{y})$ need not be a local solution of (1), see [2] in the context of bilevel programming. It follows that numerical methods computing M-stationary points of (4) may terminate at points which are not M-stationary with respect to the original (1). A complete analysis of this issue requires, however, to compare also the *qualification conditions* imposed in the course of derivation of the M-stationarity conditions for (1) and (4), respectively. As in [22] or [15], we will make use of the so-called calmness qualification conditions [10] which ensure a certain Lipschitzian behavior of the canonically perturbed constraint maps in (1) and (4), cf. Definition 2.3 and formulas (6) and (7). It turns out that, very often, the calmness qualification condition related to (1) is satisfied, whereas the qualification condition of (4) is for certain multipliers λ not fulfilled. The main aim of this paper is thus a thorough analysis of both these qualification conditions and their mutual relationship. Not surprisingly, in the achieved results an important role is played by the *constraint qualifications* (CQs) which Γ fulfills at $\bar{y}$. The choice between M-stationarity conditions of (1) and (4) depends, however, also on some other circumstances. In the first line it is the question of workable criteria for the considered calmness qualification conditions which are typically somewhat simpler in the case of (4). Further, one has to take into account also the possibility to express M-stationarity conditions of (1) in terms of problem data because otherwise the results do not have a practical value.

In the paper, all these aspects will be considered. To state our aims rigorously, one needs some basic notions from modern variational analysis. They are introduced at the beginning of the next section (Section 2.1). Section 2.2 is then devoted to a proper problem setting whereas in the last part (Section 2.3) we present several auxiliary results needed in the sequel.

Our notation is standard. For $f : \mathbb{R} \to \mathbb{R}$ by f' we understand its derivative. For a vector $x \in \mathbb{R}^n$ and a set $C \subset \mathbb{R}^n$, by $\|x\|$ we understand the (Euclidean) norm of x and by $d(x,C)$ the distance of x from C. By $o(h)$ we understand any function such that $\lim_{h \searrow 0} \frac{o(h)}{\|h\|} = 0$. Finally, by $\#S$ we mean the cardinality of a set S.

2 Problem setting and preliminaries

Throughout the whole paper we consider equilibria governed by the *generalized equation* (GE) from (1), where Γ is given in (2). With minor modifications, however, the whole theory applies also to the case when Γ is given by inequalities and *equalities*. For the sake of brevity we assume (without any loss of generality) that, at the considered point $\bar{y}$, all inequality constraints are active, i.e,

$$q_i(\bar{y}) = 0, \ i = 1, \ldots, s.$$

2.1 Background from variational analysis

Definition 2.1 *For closed set $A \subset \mathbb{R}^n$ and $\bar{x} \in A$ we define the Fréchet and limiting (Mordukhovich) normal cone to A at $\bar{x}$ by*

$$\hat{N}_A(\bar{x}) = \{x^* \mid \langle x^*, x - \bar{x}\rangle \leq o(\|x - \bar{x}\|) \text{ for all } x \in A\}$$
$$N_A(\bar{x}) = \operatorname*{Limsup}_{x \to \bar{x}} \hat{N}_A(x) := \left\{x^* \,\middle|\, \exists (x_k, x_k^*) : x_k^* \in \hat{N}_A(x_k),\ x_k \to \bar{x},\ x_k^* \to x^* \right\}.$$

If A happens to be convex, both normal cones coincide and are equal to the normal cone in the sense of convex analysis

$$\hat{N}_A(\bar{x}) = N_A(\bar{x}) = \{x^* \mid \langle x^*, x - \bar{x}\rangle \leq 0 \text{ for all } x \in A\}.$$

In such a case, we prefer to use notation $N_A(\bar{x})$.

It follows from [18, Exercise 10.26(d)] that under the standing assumptions $\hat{N}_\Gamma(y) = N_\Gamma(y)$ for all y from a neighborhood of $\bar{y}$ and therefore one can replace the regular normal cone in the GE (1) by the limiting one, having a better calculus.

Definition 2.2 *For a multifunction $M : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and for any $\bar{y} \in M(\bar{x})$ we define the (limiting) coderivative $D^*M(\bar{x}, \bar{y}) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ at this point as*

$$D^*M(\bar{x}, \bar{y})(y^*) = \left\{x^* \,\middle|\, (x^*, -y^*) \in N_{\operatorname{gph} M}(\bar{x}, \bar{y}) \right\},$$

where $\operatorname{gph} M$ stands for the graph of M.

If M is single-valued, we write only $D^*M(\bar{x})(y^*)$ instead of $D^*M(\bar{x}, M(\bar{x}))(y^*)$.

Definition 2.3 *We say that a multifunction $M : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ has the Aubin property at $(\bar{x}, \bar{y}) \in \operatorname{gph} M$ if there exist a nonnegative modulus L and neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that for all $x, x' \in U$ the following inclusion holds true*

$$M(x) \cap V \subset M(x') + L\|x - x'\|\mathbb{B},$$

where $\mathbb{B} \subset \mathbb{R}^m$ stands for the unit ball. Similarly, we say that M is calm at at $(\bar{x}, \bar{y}) \in \operatorname{gph} M$ if there exist a nonnegative modulus L and neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that for all $x \in U$ the following inclusion holds true

$$M(x) \cap V \subset M(\bar{x}) + L\|x - \bar{x}\|\mathbb{B}.$$

Note that the calmness may be significantly weaker than the Aubin property. For example any polyhedral mapping (mapping whose graph is a finite union of convex polyhedra) satisfies the calmness property at any point of its graph but may fail to have the Aubin property at the same time.

In our analysis we make use of some basic CQs from nonlinear programming. For the reader's convenience, we recall them in the next definition, where $I(y)$ denotes the set of active constraints, i.e.,

$$I(y) := \{i \in \{1,\dots,S\} |\, q_i(y) = 0\}.$$

Definition 2.4 *Consider a set* Γ *defined by inequalities* (2) *and some point* $\bar{y} \in \Gamma$*. We say that* Γ *satisfies LICQ (linear independence constraint qualification) at* $\bar{y}$ *if the gradients corresponding to all active constraints are linearly independent, hence*

$$\sum_{i\in I(\bar{y})} \mu_i \nabla q_i(\bar{y}) = 0 \implies \mu_i = 0 \text{ for all } i \in I(\bar{y}).$$

Similarly, we say that Γ *satisfies MFCQ (Mangasarian-Fromovitz constraint qualification) at* $\bar{y}$ *if the gradients corresponding to all active constraints are positively linearly independent, hence*

$$\sum_{i\in I(\bar{y})} \mu_i \nabla q_i(\bar{y}) = 0,\ \mu_i \geq 0 \implies \mu_i = 0 \text{ for all } i \in I(\bar{y}).$$

Next, Γ *satisfies CRCQ (constant rank constraint qualification) at* $\bar{y}$ *if there is a neighborhood* Y *of* $\bar{y}$ *such that for all subsets* I *of active indices* $I(\bar{y})$ *we have that* $\operatorname{rank}\{\nabla q_i(y) |\, i \in I\}$ *is a constant value for all* $y \in Y$*. Finally,* Γ *satisfies FRCQ (full rank constraint qualification) at* $\bar{y}$ *if for all subsets* I *of active indices* $I(\bar{y})$ *we have that*

$$\operatorname{rank}\{\nabla q_i(y) |\, i \in I\} = \min\{\#I, m\}.$$

The relationship among the above constraint qualifications is given by the implications

$$MFCQ \impliedby LICQ \implies FRCQ \implies CRCQ \tag{5}$$

which can be easily verified. Note that neither of these implications can be reversed in general.

2.2 Problem setting

The above defined notions enable us to state the investigated problem rigorously. The perturbation mappings associated with MPECs (1) and (4) attain the form

$$M(z) := \{(x,y) \,|\, x \in \omega, z \in F(x,y) + N_\Gamma(y)\}, \tag{6}$$

$$\begin{aligned} \tilde{M}(z_1,z_2) &:= \left\{(x,y,\lambda) \,\middle|\, x \in \omega,\, (z_1,z_2) \in H(x,y,\lambda) + N_{\mathbb{R}^m \times \mathbb{R}^s_+}(y,\lambda)\right\} \\ &= \left\{(x,y,\lambda) \,\middle|\, x \in \omega,\ z_1 = F(x,y) + (\nabla q(y))^\top \lambda,\ z_2 \in -q(y) + N_{\mathbb{R}^s_+}(\lambda)\right\}, \end{aligned} \tag{7}$$

respectively. By virtue of [22, Theorem 3.2] the M-stationarity conditions for (1) can be formulated as follows.

Theorem 2.1 ([22], Theorem 3.2) *Let $(\bar{x},\bar{y})$ be a local solution to (1). If M is calm at $(0,\bar{x},\bar{y})$, then there exists an MPEC multiplier $v \in \mathbb{R}^m$ such that*

$$0 = \nabla_x \varphi(\bar{x},\bar{y}) + [\nabla_x F(\bar{x},\bar{y})]^\top v + N_\omega(\bar{x}) \tag{8a}$$

$$0 \in \nabla_y \varphi(\bar{x},\bar{y}) + [\nabla_y F(\bar{x},\bar{y})]^\top v + D^* N_\Gamma(\bar{y}, -F(\bar{x},\bar{y}))(v). \tag{8b}$$

Since MPEC (4) has exactly the same structure as MPEC (1), the respective M-stationarity condition can be derived in the same way and one arrives at

Theorem 2.2 *Let $(\bar{x},\bar{y},\bar{\lambda})$ be a local solution to (4). If $\tilde{M}$ is calm at $\left(0,0,\bar{x},\bar{y},\bar{\lambda}\right)$, then there exist some multipliers $v \in \mathbb{R}^m$ and $w \in \mathbb{R}^s$ such that*

$$\begin{aligned}
&0 = \nabla_x \varphi(\bar{x},\bar{y}) + [\nabla_x F(\bar{x},\bar{y})]^\top v + N_\omega(\bar{x}) \\
&0 = \nabla_y \varphi(\bar{x},\bar{y}) + [\nabla_y F(\bar{x},\bar{y})]^\top v + \sum_{i=1}^{s} \bar{\lambda}_i \nabla^2 q_i(\bar{y}) v - [\nabla q(\bar{y})]^\top w \\
&0 = \nabla q_i(\bar{y}) v && \forall i : \bar{\lambda}_i > 0 \\
&0 = w_i && \forall i : q_i(\bar{y}) < 0 \\
&0 \geq w_i, 0 \leq \nabla q_i(\bar{y}) v \quad \text{or} \quad 0 = w_i \quad \text{or} \quad 0 = \nabla q_i(\bar{y}) v && \forall i : \bar{\lambda}_i = q_i(\bar{y}) = 0.
\end{aligned} \tag{9}$$

Theorem 2.2 can be interpreted as a variant of Theorem 2.1 in a different disguise addressing the same topic of MPEC (1) with differing assumptions and differing stationarity conditions. By taking into account the relationships between local solutions to (1) and (4) mentioned above, the combination of both theorems immediately leads to the following result.

Corollary 2.1 *Let $(\bar{x},\bar{y})$ be a local solution to* (1). *Then there exist multipliers v and w such that* (9) *holds true for those $\bar{\lambda} \in \Lambda(0,\bar{x},\bar{y})$ for which $\tilde{M}$ is calm at $\left(0,0,\bar{x},\bar{y},\bar{\lambda}\right)$.*

We observe first that Theorem 2.1 requires the computation of a coderivative while Theorem 2.2 provides fully explicit stationarity conditions. Precise formulas for this coderivative in terms of the problem data are available provided that Γ is polyhedral ([9, Theorem 3.2]), under LICQ at $\bar{y}$ ([7, Theorem 3.1]) or under a relaxation of MFCQ combined with the so-called 2-regularity ([5, Theorem 3]). An upper estimate has been derived in [7, Theorem 3.3] and further worked out in the Section 3.2 (Corollary 3.2).

Corollary 2.1 enables us to circumvent the difficulties associated with the coderivative in (8b) and to benefit from the explicit stationary conditions (9). This gain in convenience is bought by the need to check a calmness condition for $\tilde{M}$ which may be much more restrictive than the calmness condition for M imposed in Theorem 2.1. Indeed, Example .2 in the Appendix shows that even under MFCQ at $\bar{y}$, M may be calm while $\tilde{M}$ fails to be so at any $\bar{\lambda} \in \Lambda(0,\bar{x},\bar{y})$. Given the impact of constraint qualifications for Γ on the coderivative calculus discussed above, we are naturally led also to investigate the relation between calmness of M and $\tilde{M}$ under these constraint qualifications (MFCQ, MFCQ+FRCQ, LICQ, Γ polyhedral). This will be the topic of Section 3. Another important issue is to find qualification conditions, i.e., verifiable algebraic relations in terms of the initial data of the problem ensuring the calmness of M or $\tilde{M}$. Section 4 then compiles the main results of this paper in a concise form illustrating the information we gained about the relations and combinations of Theorems 2.1 and 2.2. Finally Section 5 illustrates the application of our results to a structured family of MPECs or bilevel problems.

2.3 Auxiliary results

At several places of the paper we will make use of the following statement from [11] which ensures the calmness of the intersection of two independently perturbed multifunctions.

Theorem 2.3 ([11], Theorem 3.6) *Consider the following multifunctions* $S_1 : \mathbb{R}^{n_1} \rightrightarrows \mathbb{R}^m$ *and* $S_2 : \mathbb{R}^{n_2} \rightrightarrows \mathbb{R}^m$ *and a point* $\bar{u} \in S_1(0) \cap S_2(0)$. *Then* $S_1 \cap S_2$ *is calm at* $(0,0,\bar{u})$ *provided the following conditions are satisfied:*

1. S_1 *is calm at* $(0,\bar{u})$
2. S_2 *is calm at* $(0,\bar{u})$
3. S_1^{-1} *has the Aubin property at* $(\bar{u},0)$
4. $S_1 \cap S_2(0)$ *is calm at* $(0,\bar{u})$.

If S_1 is the inverse of a single-valued function $f : \mathbb{R}^n \to \mathbb{R}^m$ then, to ensure the validity of the first condition of Theorem 2.3, we will often use the following result.

Lemma 2.1 *Let* $f : \mathbb{R}^n \to \mathbb{R}^m$ *be a differentiable function. Then* f^{-1} *is calm at* $(f(\bar{x}),\bar{x})$ *if at least one of the following conditions holds:*

1. f *is piecewise linear*
2. $\nabla f(\bar{x})$ *has full row rank*
3. $\nabla f(\bar{x})$ *has full column rank*

Proof. Denote $\bar{y} := f(\bar{x})$. The first case is the classical result of Robinson [17, Proposition 1]. The second one implies the Aubin property by standard calculus rule and the third one the isolated calmness property of f^{-1} at $(\bar{y},\bar{x})$ by [3, Corollary 3I.11]. Since both these properties imply calmness, the proof is complete. ☐

In the text, we will several times make use of the following consequence of Theorem 2.3.

Lemma 2.2 *Consider* $u = (u_1,u_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} = \mathbb{R}^n$, *continuously differentiable mappings* $H_1 : \mathbb{R}^n \to \mathbb{R}^m$, $H_2 : \mathbb{R}^n \to \mathbb{R}^{n_2}$, *closed sets* $\Delta \subset \mathbb{R}^n$, $\Omega \subset \mathbb{R}^{n_2}$ *and the following multifunctions*

$$\begin{aligned} S_1(z_1) &:= \{u \mid H_1(u) - z_1 = 0\} \\ S_2(z_2) &:= \{u \in \Delta \mid H_2(u) - z_2 \in N_\Omega(u_2)\}. \end{aligned} \tag{10}$$

Consider further a point $\bar{u} \in S_1(0) \cap S_2(0)$ *with the following properties:* S_1 *is calm at* $(0,\bar{u})$, S_2 *is calm at* $(0,\bar{u})$ *and the following qualification condition holds:*

$$(\nabla H_1(\bar{u}))^\top \mu \in \begin{pmatrix} 0 & \nabla_{u_1} H_2(\bar{u})^\top \\ I & \nabla_{u_2} H_2(\bar{u})^\top \end{pmatrix} N_{\operatorname{gph} N_\Omega}(\bar{u}_2, H_2(\bar{u})) + N_\Delta(\bar{u}) \implies \mu = 0. \tag{11}$$

Then $\Sigma(z_1,z_2) := S_1(z_1) \cap S_2(z_2)$ *is calm at* $(0,0,\bar{u})$.

Proof. Imitating the proof of [20, Proposition 5.2], it can be shown that Σ is calm at $(0,0,\bar{u})$ if and only if $S_1 \cap \tilde{S}_2$ is calm at $(0,0,0,\bar{u})$ with

$$\tilde{S}_2(z_2,z_3) := \left\{ u \in \Delta \,\middle|\, \begin{pmatrix} u_2 - z_3 \\ H_2(u) - z_2 \end{pmatrix} \in \operatorname{gph} N_\Omega \right\}.$$

To finish the proof, we will now apply Theorem 2.3 to S_1 and $\tilde{S}_2$, hence we need to verify that our assumptions ensure properties 1 - 4 from this theorem. Again, due to [20, Proposition 5.2] the calmness of $\tilde{S}_2$ at $(0,0,\bar{u})$ is equivalent to the calmness of S_2 at $(0,\bar{u})$, which is satisfied by our assumptions. The multifunction $S_1^{-1} = H_1$ is single-valued and locally Lipschitz continuous, and thus satisfies the Aubin property everywhere. Calmness of S_1 at $(0,\bar{u})$ is satisfied due to the assumptions.

To show that $G(z) := S_1(z) \cap \tilde{S}_2(0,0)$ is calm at $(0,\bar{u})$, we claim that (11) implies even the Aubin property of G around $(0,\bar{u})$. Indeed, by virtue of the Mordukhovich criterion [18, Theorem 9.40] this is equivalent to the implication

$$\begin{pmatrix} \mu \\ 0 \end{pmatrix} \in N_{\operatorname{gph} G}(0,\bar{u}) \implies \mu = 0,$$

which due to [18, Theorem 6.14] and simple calculus is implied by

$$(\nabla H_1(\bar{u}))^\top \mu \in N_{\tilde{S}_2(0,0)}(\bar{u}) \implies \mu = 0. \tag{12}$$

Since $\tilde{S}_2$ is calm at $(0,0,\bar{z})$, we may use [6, Theorem 4.1] to deduce

$$N_{\tilde{S}_2(0,0)}(\bar{u}) \subset \begin{pmatrix} 0 & I \\ \nabla_{u_1} H_2(\bar{u}) & \nabla_{u_2} H_2(\bar{u}) \end{pmatrix}^\top N_{\operatorname{gph} N_\Omega}(\tilde{u}_2, H_2(\bar{u})) + N_\Delta(\bar{u}). \tag{13}$$

However, due to (13), it is clear that (11) implies (12) and hence G has the Aubin property around $(0,\bar{u})$, which means that Σ is indeed calm at $(0,0,\bar{u})$. □

The following auxiliary lemma will be used on several places in the text.

Lemma 2.3 *Consider a multifunction $\phi : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^p \times \mathbb{R}^t$ with the separable structure*

$$\phi(u,v) = \phi_1(u) \times \phi_2(v),$$

and assume that $(\bar{w},\bar{z}) \in \phi_1(\bar{u}) \times \phi_2(\bar{v})$, whereby ϕ_1 is calm at $(\bar{u},\bar{w})$ and ϕ_2 is calm at $(\bar{v},\bar{z})$. Then ϕ is calm at $((\bar{u},\bar{v}),(\bar{w},\bar{z}))$.

Proof. Let us equip the Cartesian product $\mathbb{R}^p \times \mathbb{R}^t$ with the sum norm. Then one has for all $w \in \phi_1(u)$ and $z \in \phi_1(v)$ that

$$d((w,z),\phi(\bar{u},\bar{v})) = d(w,\phi_1(\bar{u})) + d(z,\phi_2(\bar{v})) \le l_1\|u-\bar{u}\| + l_2\|v-\bar{v}\| \tag{14}$$

whenever (u,v) and (w,z) are sufficiently close to $(\bar{u},\bar{v})$ and $(\bar{w},\bar{z})$, respectively. In (14), l_1 and l_2 signify the calmness moduli of ϕ_1 and ϕ_2 at $(\bar{u},\bar{w})$ and $(\bar{v},\bar{z})$, respectively. If we now endow the Cartesian product of $\mathbb{R}^n \times \mathbb{R}^m$ with the sum norm as well, we immediately conclude that ϕ is calm at the respective point with modulus $L = \max\{l_1,l_2\}$ and we are done. □

3 Relations of calmness properties of M and $\tilde{M}$

This section is devoted to a study of the general relationship between the calmness properties of M and $\tilde{M}$ defined in (6) and (7), respectively. Since we do not make use of any result from second-order variational analysis, we may relax our original assumption and for this section require functions q_i to be only $\mathscr{C}^1$. It turns out (see Example .2 in the Appendix) that under these circumstances M can be calm at $(0,\bar{x},\bar{y})$ whereas $\tilde{M}$ is not calm at $(0,0,\bar{x},\bar{y},\lambda)$ for any $\lambda \in \Lambda(0,\bar{x},\bar{y})$.

3.1 Calmness under MFCQ

Before proving first result concerning the relation between the calmness properties of M and $\tilde{M}$, we will need the following auxiliary lemma.

Lemma 3.1 *Fix any $(\bar{x},\bar{y}) \in M(0)$ and assume that MFCQ holds at $\bar{y} \in \Gamma$. Then there exist a constant L and a neighborhood $\mathscr{U}$ of $(0,0,\bar{x},\bar{y})$ such that $\|\lambda\| \le L$ for all $(z_1,z_2,x,y) \in \mathscr{U}$ and $(x,y,\lambda) \in \tilde{M}(z_1,z_2)$.*

Proof. If the statement was not true, then there would exist some sequences $(x_k,y_k,z_{k1},z_{k2}) \to (\bar{x},\bar{y},0,0)$ and λ_k with $\|\lambda_k\| \to \infty$ such that $x \in \omega$ and

$$z_{k1} = F(x_k,y_k) + (\nabla q(y_k))^\top \lambda_k, \qquad \lambda_k \in N_{\mathbb{R}^s_-}(q(y_k) - z_{k2}).$$

Since normal cone is a cone by definition, we obtain

$$(\nabla q(y_k))^\top \frac{\lambda_k}{\|\lambda_k\|} \to 0, \qquad \frac{\lambda_k}{\|\lambda_k\|} \in N_{\mathbb{R}^s_-}(q(y_k) - z_{k2}).$$

By using the outer semicontinuity of the normal cone mapping and passing to a subsequence if necessary, we might then conclude that there is a vector $\mu \in N_{\mathbb{R}^s_-}(q(\bar{y}))$ with $\|\mu\| = 1$ such that $(\nabla q(\bar{y}))^\top \mu = 0$, which contradicts MFCQ at $\bar{y}$. $\square$

Proposition 3.1 *Let MFCQ hold at $\bar{y} \in \Gamma$. Then the calmness of $\tilde{M}$ at $(0,0,\bar{x},\bar{y},\bar{\lambda})$ for all $\bar{\lambda} \in \Lambda(0,\bar{x},\bar{y})$ implies the calmness of M at $(0,\bar{x},\bar{y})$.*

Proof. Assume by contradiction that M is not calm at $(0,\bar{x},\bar{y})$, which means that there exist sequences $x_k \to \bar{x}$, $y_k \to \bar{y}$ and $p_k \to 0$ with $x_k \in \omega$ such that

$$p_k \in F(x_k,y_k) + N_\Gamma(y_k), \tag{15}$$
$$d((x_k,y_k),M(0)) > k\|p_k\|. \tag{16}$$

Since for k sufficiently large MFCQ holds at y_k, it follows from (15) and [18, Theorem 6.14] the existence of λ_k such that $(x_k,y_k,\lambda_k) \in \tilde{M}(p_k,0)$. From Lemma 3.1 we obtain that the sequence $\{\lambda_k\}$ is bounded and thus we may assume, by taking a subsequence if necessary, that $\{\lambda_k\}$ converges to some $\bar{\lambda} \in \Lambda(0,\bar{x},\bar{y})$. Since $M(0)$ is the canonical projection of $\tilde{M}(0,0)$ onto the space of the first two variables, one obtains from (16) and $(x_k,y_k,\lambda_k) \in \tilde{M}(p_k,0)$ that

$$d((x_k,y_k,\lambda_k),\tilde{M}(0,0)) \ge d((x_k,y_k),M(0)) > k\|p_k\|$$

and hence $\tilde{M}$ is not calm at $(0,0,\bar{x},\bar{y},\bar{\lambda})$. $\square$

We emphasize that the reverse implication of Proposition 3.1 cannot be expected to hold true as shown in the following example.

Example 3.1 *Consider the following data for* (1) *and* (2)

$$q(y_1,y_2) := \begin{pmatrix} y_1^2 - y_2 \\ -y_2 \end{pmatrix}, \quad F(x,y_1,y_2) := \begin{pmatrix} x \\ 1 \end{pmatrix}, \quad (\bar{x},\bar{y}_1,\bar{y}_2) := (0,0,0)$$

and $\omega = \mathbb{R}$. Note that MFCQ is satisfied for Γ. Some elementary calculus shows that, locally around $(0,0)$, we have

$$M(p_1,p_2) = \left\{ (x,y_1,y_2) \,\middle|\, y_1 = \frac{p_1 - x}{2(1-p_2)},\ y_2 = \frac{(p_1-x)^2}{4(1-p_2)^2} \right\}.$$

Since we can write $M(p_1,p_2) = \{(x,y_1,y_2) |\ G(p_1,p_2,x,y_1,y_2) = 0\}$ for a certain smooth mapping G with $\nabla_{x,y_1,y_2} G(0,0)$ having full row rank, we obtain that M has the Aubin property at $(0,0,0,0,0)$ due to [13, Corollary 4.42] and, hence, is calm there.

It can be easily computed that $\Lambda(0,\bar{x},\bar{y}) = \{\lambda \geq 0 |\ \lambda_1 + \lambda_2 = 1\}$. For $k \in \mathbb{N}$ we define

$$(z_{k1},z_{k2},z_{k3},z_{k4},x_k,y_{k1},y_{k2},\lambda_{k1},\lambda_{k2}) := (0,0,-k^{-2},0,0,k^{-1},0,0,1)$$

and observe that

$$(x_k,y_{k1},y_{k2},\lambda_{k1},\lambda_{k2}) \in \tilde{M}(z_{k1},z_{k2},z_{k3},z_{k4}).$$

Now, let $(\tilde{x},\tilde{y}_1,\tilde{y}_2,\tilde{\lambda}_1,\tilde{\lambda}_2) \in \tilde{M}(0,0,0,0)$ be arbitrarily given, where $(\tilde{\lambda}_1,\tilde{\lambda}_2)$ is close to $(0,1)$. By construction of the example, one has that $\tilde{x} = \tilde{y}_1 = \tilde{y}_2 = 0$. Consequently, one arrives at

$$\begin{aligned} d((x_k,y_{k1},y_{k2},\lambda_{k1},\lambda_{k2}),\tilde{M}(0,0,0,0)) &= \|(0,-k^{-1},0,0,1) - (0,0,0,0,1)\| \\ &= k^{-1} = k\|(z_{k1},z_{k2},z_{k3},z_{k4})\|, \end{aligned}$$

which implies that $\tilde{M}$ is not calm at $(0,0,0,0,\bar{x},\bar{y}_1,\bar{y}_2,\bar{\lambda}_1,\bar{\lambda}_2)$ with $\bar{\lambda} = (0,1)$. $\triangle$

One may easily check that in the previous example calmness of $\tilde{M}$ does hold true for the specific multiplier $\tilde{\lambda} = (1,0)$. From here, one may be tempted to deduce that calmness of M implies calmness of $\tilde{M}$ for at least **some** multiplier. However, in Example .2 we present a situation, for which M is calm at $(0,\bar{x},\bar{y})$ while $\tilde{M}$ is **not calm** at $(0,0,\bar{x},\bar{y},\lambda)$ for **any** $\lambda \in \Lambda(0,\bar{x},\bar{y})$. Since the construction of this counterexample is rather difficult, we have postponed it to the Appendix.

3.2 Calmness of perturbed complementarity constraints

In this section we collect auxiliary results that are important both for Section 3.3 devoted to the reversion of Proposition 3.1 as well as for our main results stated in Section 4. Moreover, estimate (24) represents an important amendment to the coderivative formulas from [14] and [7] and could be valuable also in a different context.

Define now the multifunction $T : \mathbb{R}^m \rightrightarrows \mathbb{R}^m \times \mathbb{R}^s$ by

$$T(p) := \left\{ (y,\lambda) \,|\, q(y) - p \in N_{\mathbb{R}^s_+}(\lambda) \right\}. \tag{17}$$

Moreover, for each arbitrary index set $I \subset \{1,\ldots,s\}$ define the multifunction $T_I : \mathbb{R}^s \rightrightarrows \mathbb{R}^m$ by

$$T_I(p) := \{y |\ q_i(y) = p_i\ (i \in I),\ q_i(y) \leq p_i\ (i \in I^c)\}. \tag{18}$$

Lemma 3.2 *Let $I \subset \{1,\dots,s\}$ and $\bar y \in q^{-1}(0)$ be arbitrary. Then, T_I is calm at $(0,\bar y)$ if and only if T is calm at any $\left(0,\bar y,\bar\lambda\right) \in \operatorname{gph} T$ such that $\bar\lambda_i > 0$ for $i \in I$ and $\bar\lambda_i = 0$ for $i \in I^c$.*

Proof. Assume first, that T_I is calm at $(0,\bar y)$ for any fixed $I \subset \{1,\dots,s\}$. This means that there are $L > 0$ and neighborhoods $\mathscr{V}$ and $\mathscr{W}$ of $\bar y$ and 0 such that for all $p \in \mathscr{W}$ and $y \in T(p) \cap \mathscr{V}$ we have

$$d(y, T_I(0)) \le L\|p\|. \tag{19}$$

Let $\bar\lambda$ be arbitrary such that $(\bar y,\bar\lambda) \in T(0)$ and $\bar\lambda_i > 0$ for $i \in I$ and $\bar\lambda_i = 0$ for $i \in I^c$. In order to verify the calmness of T at $(0,\bar y,\bar\lambda)$, let first be $\mathscr{X}$ a neighborhood of $\bar\lambda$ such that for all $\lambda \in \mathscr{X}$ one has that $\lambda_i > 0$ for $i \in I$. Finally choose a neighborhood $\mathscr{U}$ of $(0,\bar y,\bar\lambda)$ such that for all $(p,y,\lambda) \in \mathscr{U}$ one has $y \in \mathscr{V}$ and $\lambda \in \mathscr{X}$.

Select an arbitrary triple $(p,y,\lambda) \in \mathscr{U}$ such that $(y,\lambda) \in T(p)$. Since $\lambda_i > 0$ for $i \in I$, it follows that $q_i(y) - p_i = 0$ for $i \in I$ and $q_i(y) - p_i \le 0$ for $i \in I^c$. In other words, $y \in T_I(p)$, so that (19) yields $d(y,T_I(0)) \le L\|p\|$. Next, choose some $\tilde y \in T_I(0)$ such that $d(y,T_I(0)) = \|y - \tilde y\|$. Then, by definition of $T_I(0)$, we have that $q_i(\tilde y) = 0$ for $i \in I$ and $q_i(y) \le 0$ for $i \in I^c$. This entails that $(\tilde y,\lambda) \in T(0)$. Summarizing,

$$d((y,\lambda),T(0)) \le \|y-\tilde y\| = d(y,T_I(0)) \le L\|p\|$$

for all $(p,y,\lambda) \in \mathscr{U}$ such that $(y,\lambda) \in T(p)$. This implies the calmness of T at $(0,\bar y,\bar\lambda)$.

Conversely, assume that T_I fails to be calm at $(0,\bar y)$ for a fixed $I \subset \{1,\dots,s\}$. Accordingly, there exists a sequence $(p_k,y_k) \to (0,\bar y)$ such that for all k

$$q_i(y_k) = (p_k)_i \ (i \in I), \qquad q_i(y_k) \le (p_k)_i \ (i \in I^c) \tag{20}$$

and

$$d(y_k, T_I(0)) > k\|p_k\|. \tag{21}$$

Let $\bar\lambda$ be arbitrary such that $(\bar y,\bar\lambda) \in T(0)$ and $\bar\lambda_i > 0$ for $i \in I$ and $\bar\lambda_i = 0$ for $i \in I^c$. Our aim is to show that T fails to be calm at $(0,\bar y,\bar\lambda)$. We claim that, for k large enough, we have

$$d((y_k,\bar\lambda),T(0)) = d((y_k,\bar\lambda),T(0) \cap \{(y,\lambda) \,|\, \lambda_i > 0 \ (i \in I)\}). \tag{22}$$

Indeed, if this relation did not hold, then there would exist some $(\tilde y_k,\tilde\lambda_k) \in T(0)$ such that

$$\|(y_k,\bar\lambda) - (\tilde y_k,\tilde\lambda_k)\| = d((y_k,\bar\lambda),T(0)) < d((y_k,\bar\lambda),T(0) \cap \{(y,\lambda) \,|\, \lambda_i > 0 \ (i \in I)\}),$$

which implies that $(\tilde\lambda_k)_j = 0$ for some $j \in I$. On the other hand, $\bar\lambda_j > 0$ by assumption. Consequently, due to $(y_k,\bar\lambda) \to (\bar y,\bar\lambda) \in T(0)$, we end up at the contradiction

$$0 < \bar\lambda_j = |\bar\lambda_j - (\tilde\lambda_k)_j| \le \|(y_k,\bar\lambda) - (\tilde y_k,\tilde\lambda_k)\| = d((y_k,\bar\lambda),T(0)) \to 0.$$

By (22), there exists some $(\tilde y,\tilde\lambda) \in T(0)$ such that $\tilde\lambda_i > 0$ for all $i \in I$ and

$$d((y_k,\bar\lambda),T(0)) = \|(y_k,\bar\lambda) - (\tilde y,\tilde\lambda)\|. \tag{23}$$

Since $q(\tilde y) \in N_{\mathbb{R}^s_+}(\tilde\lambda)$, it follows that $q_i(\tilde y) = 0$ for all $i \in I$ and $q_i(\tilde y) \le 0$ for all $i \in I^c$. In other words, $\tilde y \in T_I(0)$. Now, (21) implies that $\|y_k - \tilde y\| > k\|p_k\|$. Combining this with (23) yields that

$$d((y_k,\bar\lambda),T(0)) > k\|p_k\|.$$

Moreover, we know that $(p_k,y_k,\bar\lambda) \to (0,\bar y,\bar\lambda)$. Finally, (20) along with $\bar\lambda_i > 0$ for $i \in I$ implies that $(y_k,\bar\lambda) \in T(p_k)$. Altogether, we have shown that T fails to be calm at $(0,\bar y,\bar\lambda)$ as well. □

Exploiting the above lemma for all index subsets $I \subset \{1,\dots,s\}$, we arrive at the following criterion for the calmness of T.

Corollary 3.1 *Let $\bar{y} \in q^{-1}(0)$ be arbitrary. Then, T is calm at any $(0,\bar{y},\bar{\lambda}) \in \operatorname{gph} T$ if and only if for all $I \subset \{1,\ldots,s\}$, T_I is calm at $(0,\bar{y})$. This holds in particular if CRCQ is satisfied at $\bar{y}$.*

Proof. The first part of the proof follows immediately from Lemma 3.2. To prove the second part, recall from [20, Proposition 5.2] that T is calm at $(0,\bar{y},\lambda)$ provided the following mapping

$$\tilde{T}(p,r) := \left\{ (y,\lambda) \,\middle|\, q(y) - p \in N_{\mathbb{R}^s_+}(\lambda + r) \right\}$$

is calm at $(0,0,\bar{y},\lambda)$. But from [7, Propositions 3.1 and 3.2] we obtain that this property holds provided for all index sets $I \subset \{1,\ldots,s\}$ the associated mappings $M_I : \mathbb{R}^{|I|} \rightrightarrows \mathbb{R}^m$, defined by

$$M_I(p) := \{y \,|\, q_i(y) = p_i \ (i \in I)\},$$

are calm at $(0,\bar{y})$. However, as shown in [12, Theorem 1], the calmness of a perturbed system of equalities and inequalities is implied by CRCQ at the respective point. Due to its definition, the imposed CRCQ is valid also for all subsystems generating the maps M_I, and so T is indeed calm at $(0,\bar{y},\bar{\lambda})$ for all $(\bar{y},\bar{\lambda}) \in T(0)$. □

In [7, 14] the authors computed (an upper estimate of) coderivative $D^*N_\Gamma(\bar{y}, -F(\bar{x},\bar{y}))$ under MFCQ at $\bar{y}$ and under the assumption that T is calm at $(0,\bar{y},\lambda)$ for all $\lambda \in \Lambda(0,\bar{x},\bar{y})$. By combining [7, Theorem 3.3] and Corollary 3.1, one arrives directly at the next statement.

Corollary 3.2 *Assume that $q \in \mathscr{C}^2$ and both MFCQ as well as CRCQ are fulfilled at $\bar{y}$. Then one has with $y^* := -F(\bar{x},\bar{y})$ for all $v^* \in \mathbb{R}^m$ the estimate*

$$D^*N_\Gamma(\bar{y},y^*)(v^*) \subset \bigcup_{\lambda \in \Lambda(0,\bar{x},\bar{y})} \left\{ \left(\sum_{i=1}^{s} \lambda_i \nabla^2 q_i(\bar{y}) \right) v^* + (\nabla q(\bar{y}))^\top D^*N_{\mathbb{R}^s_-}(0,\lambda)(\nabla q(\bar{y})v^*) \right\}. \tag{24}$$

3.3 Calmness under MFCQ and FRCQ

The main result of this section concerns the reversion of Proposition 3.1. This is already everything but evident even in the case when LICQ is fulfilled at $\bar{y}$. Much less it is clear under a weaker constraint qualification. The respective statement (Theorem 3.1), the proof of which requires a considerable effort, will play an important role in Section 4, where we compare and combine the optimality conditions stated in Theorems 2.1 and 2.2. Unfortunately, we have succeeded to prove this reverse implication only when, instead of $q \in \mathscr{C}^1$, one assumes that $q \in \mathscr{C}^{1,1}$ meaning that ∇q is also Lipschitz continuous around $\bar{y}$.

For the main theorem, we will define two auxiliary multifunctions which will be of use when partitioning $\tilde{M}$

$$\begin{aligned} S_1(z_1) &:= \{(x,y,\lambda) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^s \,|\, F(x,y) + (\nabla q(y))^\top \lambda - z_1 = 0\} \\ S_2(z_2) &:= \left\{ (x,y,\lambda) \in \omega \times \mathbb{R}^m \times \mathbb{R}^s \,\middle|\, \begin{pmatrix} \lambda \\ q(y) - z_2 \end{pmatrix} \in \operatorname{gph} N_{\mathbb{R}^s_+} \right\}. \end{aligned} \tag{25}$$

Further we consider the following family of mappings parameterized by $y \in \mathbb{R}^m$

$$W_y(p_1,p_2) := \left\{ \lambda \,\middle|\, (\nabla q(y))^\top \lambda = p_1,\, p_2 \in N_{\mathbb{R}^s_+}(\lambda) \right\}. \tag{26}$$

Theorem 3.1 *Let q be of class $\mathscr{C}^{1,1}$, i.e., q is differentiable with locally Lipschitz continuous gradient. Fix any $(\bar{x},\bar{y})$ and assume at least one of the following two conditions:*

1 *MFCQ is satisfied at $\bar{y}$ and q is affine linear;*

2 *MFCQ and FRCQ are satisfied at $\bar{y}$.*

Then the calmness of M at $(0,\bar{x},\bar{y})$ is equivalent to the calmness of $\tilde{M}$ at $(0,0,\bar{x},\bar{y},\bar{\lambda})$ for any $\bar{\lambda} \in \Lambda(0,\bar{x},\bar{y})$.

Proof. One implication follows directly from Proposition 3.1. Hence, it suffices to show that the calmness of M at $(0,\bar{x},\bar{y})$ implies the calmness of $\tilde{M}$ at $(0,0,\bar{x},\bar{y},\bar{\lambda})$ for any $\bar{\lambda} \in \Lambda(0,\bar{x},\bar{y})$. To this aim fix an arbitrary such $\bar{\lambda}$. We will show that under each of the two required assumptions, there are constants $\kappa \geq 0$ and $\varepsilon_1 > 0$ such that for all $(z_1,z_2,x',y',\lambda') \in \operatorname{gph}\tilde{M} \cap \mathbb{B}_{\varepsilon_1}(0,0,\bar{x},\bar{y},\bar{\lambda})$ we have

$$d((x',y',\lambda'),\tilde{M}(0,0)) \leq \kappa\|(z_1,z_2)\|. \tag{27}$$

Without loss of generality, we will work with the maximum norm throughout this proof. First we collect all information that is at our disposal in the following relations, where $\varepsilon, L > 0$ are certain positive constants which may be assumed to have common values in all of these relations:

$$\begin{aligned}
\|F(x_1,y_1) - F(x_2,y_2)\| &\leq L\|(x_1,y_1)-(x_2,y_2)\| \quad \forall (x_1,y_1),(x_2,y_2) \in \mathbb{B}_\varepsilon((\bar{x},\bar{y})) && \text{(28a)}\\
\|F(x,y)\| &\leq L \quad \forall (x,y) \in \mathbb{B}_\varepsilon((\bar{x},\bar{y})) && \text{(28b)}\\
\|q(y_1)-q(y_2)\| &\leq L\|y_1-y_2\| \quad \forall y_1,y_2 \in \mathbb{B}_\varepsilon(\bar{y}) && \text{(28c)}\\
\|\nabla q(y_1)-\nabla q(y_2)\| &\leq L\|y_1-y_2\| \quad \forall y_1,y_2 \in \mathbb{B}_\varepsilon(\bar{y}) && \text{(28d)}\\
\|\nabla q(y)\| &\leq L \quad \forall y \in \mathbb{B}_\varepsilon(\bar{y}) && \text{(28e)}\\
d((x,y),M(0)) &\leq L\|z\| \quad \forall (x,y,z) \in \mathbb{B}_\varepsilon(\bar{x},\bar{y},0) : (x,y) \in M(z) && \text{(28f)}\\
d((x,y,\lambda),S_2(0)) &\leq L\|z\| \quad \forall (x,y,\lambda,z) \in \mathbb{B}_\varepsilon(\bar{x},\bar{y},\bar{\lambda},0) : (x,y,\lambda) \in S_2(z). && \text{(28g)}\\
\|\lambda\| &\leq L \quad \forall \lambda\ \forall (x,y,z_1,z_2) \in \mathbb{B}_\varepsilon(\bar{x},\bar{y},0,0) : (x,y,\lambda) \in \tilde{M}(z_1,z_2) && \text{(28h)}
\end{aligned}$$

Here, (28a)-(28e) follow from the differentiability assumptions we have made, (28f) corresponds to the assumed calmness of M at $(0,\bar{x},\bar{y})$. (28g) means the calmness at $(0,\bar{x},\bar{y},\bar{\lambda})$ of the mapping S_2 defined in (10). This follows from Corollary 3.1 and Lemma 2.3 upon observing that both our assumptions imply MFCQ and CRCQ: indeed, CRCQ is always satisfied for affine linear inequality systems and it follows from FRCQ via (5). Formula (28h) is a consequence of Lemma 3.1.

Now, in order to verify the asserted calmness of $\tilde{M}$ at $(0,0,\bar{x},\bar{y},\bar{\lambda})$, define

$$\varepsilon_1 := \min\left\{\frac{\varepsilon}{2}, \frac{\varepsilon}{2L}, \frac{\varepsilon}{1+2L^2+L^3}, \frac{\varepsilon}{1+2L+2L^3+L^4}\right\} \tag{29}$$

and consider an arbitrary triple $(x',y',\lambda') \in \tilde{M}(z_1,z_2)$ with $(z_1,z_2,x',y',\lambda') \in \mathbb{B}_{\varepsilon_1}\left(0,0,\bar{x},\bar{y},\bar{\lambda}\right)$. Since $\tilde{M}(z_1,z_2) = S_1(z_1) \cap S_2(z_2)$, we may use (28g) to obtain the existence of some $(\tilde{x},\tilde{y},\tilde{\lambda}) \in S_2(0)$ such that

$$\max\left\{\|x'-\tilde{x}\|, \|y'-\tilde{y}\|, \|\lambda'-\tilde{\lambda}\|\right\} \leq L\|z_2\|. \tag{30}$$

By definition of S_2, relation $(\tilde{x},\tilde{y},\tilde{\lambda}) \in S_2(0)$ implies that $q(\tilde{y}) \in N_{\mathbb{R}^s_+}(\tilde{\lambda})$, which further means that $(\tilde{x},\tilde{y},\tilde{\lambda}) \in \tilde{M}(a,0)$ and thus $(\tilde{x},\tilde{y}) \in M(a)$ with

$$a = F(\tilde{x},\tilde{y}) + [\nabla q(\tilde{y})]^\top \tilde{\lambda}. \tag{31}$$

Moreover, since $(x',y',\lambda')\in S_1(z_1)$, we obtain

$$\begin{aligned}\|a\| &= \|F(\tilde{x},\tilde{y})+[\nabla q(\tilde{y})]^\top\tilde{\lambda}+z_1-F(x',y')-[\nabla q(y')]^\top\lambda'\|\\ &\le \|z_1\|+\|F(\tilde{x},\tilde{y})-F(x',y')\|+\|[\nabla q(\tilde{y})]^\top\tilde{\lambda}-\left[\nabla q(y')\right]^\top\lambda'\|\\ &\le \|z_1\|+\left\|F(\tilde{x},\tilde{y})-F(x',y')\right\|+\left\|\lambda'\right\|\left\|\nabla q(\tilde{y})-\nabla q(y')\right\|+\|\lambda'-\tilde{\lambda}\|\,\|\nabla q(\tilde{y})\|.\end{aligned}\tag{32}$$

Next, the relation $(x',y',\lambda')\in\mathbb{B}_{\varepsilon_1}(\bar{x},\bar{y},\bar{\lambda})$ and (29, first case) imply that

$$(x',y',\lambda')\in\mathbb{B}_{\varepsilon/2}(\bar{x},\bar{y},\bar{\lambda}).$$

Combining (30) with (29, second case) and recalling that $z_2\in\mathbb{B}_{\varepsilon_1}(0)$ yields

$$(\tilde{x},\tilde{y},\tilde{\lambda})\in\mathbb{B}_{L\|z_2\|}\left(x',y',\lambda'\right)\subset\mathbb{B}_{\varepsilon/2}\left(x',y',\lambda'\right)\subset\mathbb{B}_{\varepsilon}(\bar{x},\bar{y},\bar{\lambda}).\tag{33}$$

Now, relations (28a), (28d), (28e), (28h), and (29, third case) together with (30) allow us to continue our estimation from (32) and to obtain

$$\begin{aligned}\|a\| &\le \|z_1\|+L^2\|z_2\|+\left\|\lambda'\right\|L^2\|z_2\|+L^2\|z_2\|\le\|z_1\|+\left(2L^2+L^3\right)\|z_2\|\\ &\le\left(1+2L^2+L^3\right)\|(z_1,z_2)\|\le\varepsilon.\end{aligned}\tag{34}$$

Therefore, we are now allowed to apply (28f) and make use of the fact that $(\tilde{x},\tilde{y})\in M(a)$ implies the existence of some $(x^*,y^*)\in M(0)$ such that

$$\max\left\{\|x^*-\tilde{x}\|,\|y^*-\tilde{y}\|\right\}\le L\|a\|.\tag{35}$$

Note that (35) along with (34), (30), (29, fourth case) and the initial assumption $(x',y',z_1,z_2)\in\mathbb{B}_{\varepsilon_1}(\bar{x},\bar{y},0,0)$ leads to

$$\max\{\|x^*-\tilde{x}\|,\|y^*-\tilde{y}\|\}\le L\left(1+2L^2+L^3\right)\|(z_1,z_2)\|\tag{36a}$$

$$\max\{\left\|x^*-x'\right\|,\left\|y^*-y'\right\|\}\le L\left(2+2L^2+L^3\right)\|(z_1,z_2)\|\tag{36b}$$

$$\max\{\|x^*-\bar{x}\|,\|y^*-\bar{y}\|\}\le\left(1+2L+2L^3+L^4\right)\varepsilon_1\le\varepsilon.\tag{36c}$$

Referring back to the definition (26), $(x^*,y^*)\in M(0)$ implies that $W_{y^*}(-F(x^*,y^*),q(y^*))\neq\emptyset$. Similarly, $(\tilde{x},\tilde{y},\tilde{\lambda})\in S_2(0)$, yields

$$\tilde{\lambda}\in W_{y^*}\left([\nabla q(y^*)]^\top\tilde{\lambda},q(\tilde{y})\right).$$

Assumptions 1 and 2 of this theorem allow us to apply the respective estimates (58) and (59) from the Appendix. Putting

$$y:=y^*,\ \bar{p}_1:=-F(x^*,y^*),\ \bar{p}_2:=q(y^*),\ p_1:=[\nabla q(y^*)]^\top\tilde{\lambda},\ p_2:=q(\tilde{y}),\ \lambda:=\tilde{\lambda},$$

we derive in both cases the existence of some $\lambda^*\in W_{y^*}\left(-F(x^*,y^*),q(y^*)\right)$ such that

$$\|\lambda^*-\tilde{\lambda}\|\le L'\max\left\{\|[\nabla q(y^*)]^\top\tilde{\lambda}+F(x^*,y^*)\|,\|q(\tilde{y})-q(y^*)\|\right\}.\tag{37}$$

for some $L'>0$. To estimate both terms on the right-hand side of (37), we realize first that (33) and (36c) allow us to use estimates (28) for all terms on the right-hand side of (37). For estimation of the first norm on the right-hand side of (37), we use (31), (34), (28h) coupled with $(\tilde{x},\tilde{y},\tilde{\lambda})\in\tilde{M}(a,0)$, (28d), (28a) and (36a) while for the second one we use (28c) and (36a) to obtain some constant c_1 such that

$$\begin{aligned}\left\|[\nabla q(y^*)]^\top \tilde{\lambda}+F(x^*,y^*)\right\| &= \left\|a+\left([\nabla q(y^*)]^\top-[\nabla q(\tilde{y})]^\top\right)\tilde{\lambda}+F(x^*,y^*)-F(\tilde{x},\tilde{y})\right\| \\ &\le \|a\|+\|\tilde{\lambda}\|\,\|[\nabla q(y^*)]-[\nabla q(\tilde{y})]\|+\|F(x^*,y^*)-F(\tilde{x},\tilde{y})\| \\ &\le c_1\|(z_1,z_2)\|, \\ \|q(\tilde{y})-q(y^*)\| &\le L\|\tilde{y}-y^*\|\le c_1\|(z_1,z_2)\|.\end{aligned} \tag{38}$$

Estimates (30), (37) and (38) yield

$$\|\lambda^*-\lambda'\|\le\|\lambda^*-\tilde{\lambda}\|+\|\tilde{\lambda}-\lambda'\|\le L'c_1\,\|(z_1,z_2)\|+L\|z_2\|.$$

Adding this to (36b), we arrive at existence of some c such that

$$\left\|(x',y',\lambda')-(x^*,y^*,\lambda^*)\right\|\le c\,\|(z_1,z_2)\| \tag{39}$$

On the other hand, the already obtained relation $\lambda^*\in W_{y^*}(-F(x^*,y^*),q(y^*))$ amounts, by definition, to $(x^*,y^*,\lambda^*)\in\tilde{M}(0,0)$ and thus we have shown (27) with $\kappa=c$. This finishes the proof. $\square$

Corollary 3.3 *In the setting of Theorem 3.1, let LICQ be satisfied at $\bar{y}$. Then the calmness of M at $(0,\bar{x},\bar{y})$ is equivalent to the calmness of $\tilde{M}$ at $(0,0,\bar{x},\bar{y},\bar{\lambda})$ for the unique $\bar{\lambda}\in\Lambda(0,\bar{x},\bar{y})$.*

Proof. By (5), assumption 2 of Theorem 3.1 is satisfied. $\square$

In the next example we present an instance, where FRCQ and MFCQ hold true but LICQ is violated. Note that none of the inequalities is redundant and that the mapping q is not affine. This means that assumption 2 of Theorem 3.1 is not only strictly weaker than LICQ but also independent of assumption 1 of the same theorem.

Example 3.2 *Define $\bar{y}:=(0,0,0)$ and*

$$q_1(y):=y_1+y_1^2-y_3,\ q_2(y):=-y_1+y_1^2-y_3,\ q_3(y):=y_2+y_2^2-y_3,\ q_4(y):=-y_2+y_2^2-y_3.$$

Then it is not difficult to show that these data satisfy the properties stated above. $\triangle$

To conclude this section, we show that Theorem 3.1 does not hold if we assume only $q\in\mathscr{C}^1$.

Example 3.3 *Consider function $q:\mathbb{R}\to\mathbb{R}$ defined as*

$$q(y)=\begin{cases} y+y^{3/2} & \text{if } y\ge 0\\ y-|y|^{3/2} & \text{if } y<0.\end{cases}$$

Further define $F(x,y)=-1$, $\omega=\mathbb{R}$ and consider the reference point $(\bar{x},\bar{y},\bar{\lambda})=(0,0,1)$. Since $q'(0)=1$, LICQ is satisfied around $\bar{y}$. Moreover, it is clear that $\Gamma=(-\infty,0]$ and that q' is continuous at 0 but it is not Lipschitz continuous there. For all p close to 0 it holds true that

$$M(p):=\{(x,y)\,|\,p+1\in N_\Gamma(y)\}=\mathbb{R}\times\{0\}$$

and thus M is calm at $(0,\bar{x},\bar{y})$. Since $\bar{\lambda}=1$, we may find a neighborhood of the reference point such that

$$\tilde{M}(z_1,z_2) := \{(x,y,\lambda) \,|\, z_1+1 = q'(y)\lambda,\, q(y) = -z_2\}$$

and thus, due to Lemma 2.3, the calmness of $\tilde{M}$ at $(0,0,\bar{x},\bar{y},\bar{\lambda})$ is equivalent to the calmness of $\hat{M}$ at $(0,0,\bar{y},\bar{\lambda})$ with

$$\hat{M}(z_1,z_2) := \{(y,\lambda) \,|\, z_1+1 = q'(y)\lambda,\, q(y) = -z_2\}.$$

Since q is continuously differentiable and $q'(0)\neq 0$, we may use the inverse function theorem to obtain there there exists a continuously differentiable function h with such that on some neighborhood of 0, relation $-q(y)=z_2$ is equivalent to $h(z_2)=y$. Further we have

$$h'(z_2) = -1/q'(h(z_2)).$$

Plugging this into the first first equation defining $\hat{M}$ and performing simple algebraic operations, we obtain the following system of equations

$$\lambda = -h'(z_2)(z_1+1), \quad y = h(z_2).$$

This means that $\hat{M}$ is single-valued and to show that $\hat{M}$ is not calm at $(0,0,\bar{y},\bar{\lambda})$ it is sufficient to show that $p\mapsto h'(p)$ is not calm at 0. Since h' is continuous, we do not have to consider a neighborhood in the range from the definition of calmness. It is simple to see that

$$\frac{|h'(p)-h'(0)|}{|p-0|} = \frac{1}{|q'(h(p))q'(h(0))|}\frac{|q'(h(0))-q'(h(p))|}{|p-0|} \geq \frac{|q'(h(0))-q'(h(p))|}{2|h(p)-h(0)|} \overset{p\to 0}{\longrightarrow} \infty$$

because q' is not calm at 0. In the inequality we have used the estimate

$$\frac{1}{|q'(h(p))q'(h(0))|}\frac{|h(p)-h(0)|}{|p-0|} \geq \frac{1}{2},$$

for all p sufficiently close to zero as $q'(0)=1$ and $h'(0)=-1/q'(0)=-1$ and both q and h are continuously differentiable at 0. Thus, we have managed to find and example, in which LICQ holds, M is calm at $(0,\bar{x},\bar{y})$ but $\tilde{M}$ is not calm at $(0,0,\bar{x},\bar{y},\bar{\lambda})$. $\triangle$

4 Main results

In the first part of this section we address the question how the calmness property of M and $\tilde{M}$ can be ensured by suitable pointbased conditions. Concerning the calmness of M, we present here only a standard result in which one enforces in fact even the (substantially more restrictive) Aubin property. In [18] and [13], exclusively this type of qualification conditions is used. We are aware about the possibility to employ to this purpose some less restrictive calmness criteria from, e.g., [4, 10] but this goes beyond the aim of this paper. Throughout this section it is assumed that $q\in\mathscr{C}^2$.

Theorem 4.1 *Assume that the implication*

$$\left.\begin{aligned} (\nabla_x F(\bar{x},\bar{y}))^\top a &\in -N_\omega(\bar{x}) \\ -(\nabla_y F(\bar{x},\bar{y}))^\top a &\in D^* N_\Gamma(\bar{y},-F(\bar{x},\bar{y}))(a) \end{aligned}\right\} \implies a = 0 \tag{40}$$

is fulfilled. Then M has the Aubin property around $(0,\bar{x},\bar{y})$ and hence it is also calm at this point.

Proof. The assertion follows immediately from the Mordukhovich criterion and the standard first-order calculus. □

For the verification of the calmness of $\tilde{M}$, however, we present here a new condition based on Lemma 3.2. To this aim, we define the Lagrangian as

$$\mathscr{L}(x,y,\lambda) := F(x,y) + (\nabla q(y))^\top \lambda. \tag{41}$$

Note that in the statement of the theorem, T_I is automatically calm at $(0,\bar{y})$ if CRCQ is satisfied at $\bar{y}$. This follows directly from Corollary 3.1.

Theorem 4.2 *Assume that* $(\bar{x},\bar{y},\bar{\lambda}) \in \tilde{M}(0,0)$ *and that the implication*

$$\left.\begin{array}{lll} (\nabla_x F(\bar{x},\bar{y}))^\top a \in -N_\omega(\bar{x}) & & \\ (\nabla_y \mathscr{L}(\bar{x},\bar{y},\bar{\lambda}))^\top a + (\nabla q(\bar{y}))^\top c = 0 & & \\ 0 = \nabla q_i(\bar{y})a & \forall i : \bar{\lambda}_i > 0 & \\ 0 = c_i & \forall i : q_i(\bar{y}) < 0 & \\ 0 \le c_i, 0 \le \nabla q_i(\bar{y})a \quad \textit{or} \quad 0 = c_i \quad \textit{or} \quad 0 = \nabla q_i(\bar{y})a & \forall i : \bar{\lambda}_i = q_i(\bar{y}) = 0. & \end{array}\right\} \Longrightarrow a = 0. \tag{42}$$

holds true. Define $I := \{i | \bar{\lambda}_i > 0\}$ *and assume that the mapping* $T_I : \mathbb{R}^s \to \mathbb{R}^m$ *defined by* (18) *is calm at* $(0,\bar{y})$. *Then* $\tilde{M}$ *is calm at* $(0,0,\bar{x},\bar{y},\bar{\lambda})$.

Proof. Taking into account that $\tilde{M}(z_1,z_2) = S_1(z_1) \cap S_2(z_2)$ with S_1 and S_2 defined in (25), to obtain the calmness of $\tilde{M}$ at $(0,0,\bar{x},\bar{y},\bar{\lambda})$ it suffices to verify the assumptions of Lemma 2.2 for the following data: $u_1 = (x,y)$, $u_2 = \lambda$, $H_1(u) = \mathscr{L}(x,y,\lambda)$, $H_2(u) = q(y)$, $\Delta = \omega \times \mathbb{R}^m \times \mathbb{R}^s$, $\mu = a$ and $\Omega = \mathbb{R}^s_+$. It is not difficult to show that condition (11) takes form (42) and so it remains to show that S_1 and S_2 are calm at $(0,\bar{x},\bar{y},\bar{\lambda})$.

To show that S_1 has this property, we will apply Lemma 2.1 according to which it is sufficient to show that $\nabla \mathscr{L}(\bar{x},\bar{y},\bar{\lambda})$ has full row rank. Hence consider any a such that $\nabla \mathscr{L}(\bar{x},\bar{y},\bar{\lambda})^\top a = 0$. But then $(a,0)$ satisfies the relations on the left-hand side of (42) and thus $a = 0$, implying that S_1 is indeed calm at $(0,\bar{x},\bar{y},\bar{\lambda})$. But since by Lemma 3.2 the calmness of S_2 at $(0,\bar{x},\bar{y},\bar{\lambda})$ is equivalent with the calmness of T_I at $(0,\bar{y})$, the remaining assumption is ensured by our assumptions due to Lemma 2.3. The last three lines of (42) provide an equivalent representation of the resulting coderivative. □

Note that if ω is a convex set, then N_ω is the standard normal cone in the sense of convex analysis. Moreover, if $\omega = \mathbb{R}^n$, then $N_\omega(\bar{x}) = \{0\}$ and the inclusion reduces to an equality.

In the MPEC literature, one finds under various names (GMFCQ, NNAMCQ) a qualification condition similar to (42) with the difference that $a = c = 0$ is required instead of only $a = 0$. Clearly, under LICQ at $\bar{y}$, both these conditions coincide. However, if we impose only MFCQ and CRCQ at $\bar{y}$, (42) is strictly better (less restrictive) than GMFCQ.

The next example illustrates the possible applications and limitations of Theorem 4.2.

Example 4.1 *We consider the data of Example 3.1 with the only exception that now* $F(x,y_1,y_2) := (0,1)^\top$. *Similar to Example 3.1, locally around* $(0,0)$, *we have*

$$M(p_1,p_2) = \left\{ (x,y_1,y_2) \,\middle|\, y_1 = \frac{p_1}{2(1-p_2)},\ y_2 = \frac{p_1^2}{4(1-p_2)^2} \right\}$$

and that M has the Aubin property at $(0,0,0,0,0)$ and, hence, is calm there. Moreover, the multiplier set still has the description

$$\Lambda(0,\bar{x},\bar{y}) = \{\lambda \geq 0 | \lambda_1 + \lambda_2 = 1\}.$$

Now we want to check calmness of $\tilde{M}$ for different multipliers λ.

Case 1: $\bar{\lambda} = (1,0)$. It is easy to see that the associated mapping $T_{\{1\}}$ is calm at $(0,0,\bar{y})$. Now we check system (42)*. Since*

$$\mathscr{L}(x,y,\lambda) = \begin{pmatrix} 2y_1\lambda_1 \\ 1-\lambda_1-\lambda_2 \end{pmatrix},$$

the second equation implies $a_1 = 0$. Since $\bar{\lambda}_2 > 0$, we obtain $0 = \nabla q_2(\bar{y})a = -a_2$. Hence, all assumptions of Theorem 4.2 has been verified for $\bar{\lambda} = (1,0)$ and hence $\tilde{M}$ is calm at $(0,0,\bar{x},\bar{y},\bar{\lambda})$.

Case 2: $\bar{\lambda} = (0,1)$: In this case $\tilde{M}$ is not calm at $(0,0,\bar{x},\bar{y},\bar{\lambda})$, which can be verified in exactly the same way as in Example 3.1. This naturally implies that the assumptions of Theorem 4.2 are not satisfied.

Case 3: $\bar{\lambda}_1 > 0$, $\bar{\lambda}_2 > 0$: Finally consider any $\bar{\lambda} \in \Lambda(0,\bar{x},\bar{y})$ with $\bar{\lambda}_1, \bar{\lambda}_2 > 0$ and choose any sequence

$$(x_k, y_{k1}, y_{k2}, \lambda_{k1}, \lambda_{k2}) \in \tilde{M}(z_{k1}, z_{k2}, z_{k3}, z_{k4}).$$

Since we are interested in a local property, we can consider purely $\lambda_{k1} \geq \varepsilon$ for some $\varepsilon > 0$, which implies $|y_{k1}| \leq \frac{1}{2\varepsilon}|z_{k1}|$. Then we have

$$d((x_k,y_{k1},y_{k2},\lambda_{k1},\lambda_{k2}),\tilde{M}(0,0,0,0)) \leq |y_{k1}| + |y_{k2}| + |1-\lambda_{k1}-\lambda_{k2}| \leq \frac{1}{2\varepsilon}|z_{k1}| + |z_{k4}| + |z_{k2}|$$

and hence $\tilde{M}$ is calm at $(0,0,\bar{x},\bar{y},\bar{\lambda})$. Note that this holds true even though the assumptions of Theorem 4.2 are not satisfied because $T_{\{1,2\}}$ is not calm at $(0,\bar{y})$. △

In the remainder of this section we will state the main result of the paper. It comprises in a concise form the information which we have gained in the course of our analysis about the relationship between Theorems 2.1 and 2.2. It leads to several useful conclusions important in deriving workable M-stationarity conditions for MPEC (1).

Theorem 4.3 *Let $(\bar{x},\bar{y})$ be a local solution to* (1) *and assume that MFCQ holds at $\bar{y} \in \Gamma$.*

1 *If CRCQ holds at $\bar{y}$, then for those $\lambda \in \Lambda(0,\bar{x},\bar{y})$ satisfying the qualification condition* (42)*, there exist v and w fulfilling the stationarity conditions* (9)*.*

2 *If CRCQ holds at $\bar{y}$ and M is calm at $(0,\bar{x},\bar{y})$, then there exist $\lambda \in \Lambda(0,\bar{x},\bar{y})$, v and w fulfilling the stationarity conditions* (9)*.*

3 *Let M be calm at $(0,\bar{x},\bar{y})$. If FRCQ holds at $\bar{y}$ or q is affine linear, then, for* ***all*** *$\lambda \in \Lambda(0,\bar{x},\bar{y})$ there exist v and w fulfilling the stationarity conditions* (9)*.*

4 *If even LICQ holds at $\bar{y}$, then Theorems 2.1 and 2.2 are completely equivalent in their assumptions and their results. In particular, if* (42) *is satisfied for the unique $\lambda \in \Lambda(0,\bar{x},\bar{y})$, then there exist v and w fulfilling the stationarity conditions* (9) *with this λ.*

Before proving this Theorem, we include some comments on the statements 1-3. The big progress of statement 1 over Theorems 2.1 and 2.2 or Corollary 2.1 is that under MFCQ and CRCQ it completely frees us from the necessity of checking any calmness condition or computing the complicated coderivative $D^*N_\Gamma(\bar{y},-F(\bar{x},\bar{y}))$. It just relies on checking the explicit qualification condition (42) and provides explicit stationarity conditions (9). For instance, in order to exclude $(\bar{x},\bar{y})$ from being a local solution to (1), it will be sufficient to find some $\lambda\in\Lambda(0,\bar{x},\bar{y})$ satisfying (42) and violating (9) for any v and w. Unfortunately, it is not excluded that the set of $\lambda\in\Lambda(0,\bar{x},\bar{y})$ satisfying (42) is empty so that statement 1 cannot be applied. But even then, one might be successful in checking calmness of M and thus apply statement 2. Excluding $(\bar{x},\bar{y})$ from being a local solution to (1) would then amount to verifying that (9) is violated for any $\lambda\in\Lambda(0,\bar{x},\bar{y})$ and any v and w. Statement 3 provides two instances under which we do not have to care about specific $\lambda\in\Lambda(0,\bar{x},\bar{y})$. This facilitates the task of excluding $(\bar{x},\bar{y})$ from being a local solution to (1) in the sense that we just have to find **some** $\lambda\in\Lambda(0,\bar{x},\bar{y})$ such that (9) is violated for any v and w. *Proof.* [of Theorem 4.3] First recall that under the joint assumptions $(\bar{x},\bar{y},\lambda)$ is a local solution of MPEC (4) for all $\lambda\in\Lambda(0,\bar{x},\bar{y})$. Concerning statement 1, observe that under CRCQ at $\bar{y}$ we have that $\tilde{M}$ is calm at all points $(0,0,\bar{x},\bar{y},\lambda)$ with $\lambda\in\Lambda(0,\bar{x},\bar{y})$ satisfying (42) by virtue of Theorem 4.2 and Corollary 3.1. Statement 1 thus follows from Theorem 2.2.

Statement 2 is a direct consequence of Theorem 2.1 and inclusion (24), where one needs just to express the coderivative $D^*N_{\mathbb{R}_{+}^{-s}}(q(\bar{y}),\lambda)$ in terms of $q(\bar{y})$ and λ. To prove statement 3, it suffices to combine Theorem 2.2 with Theorem 3.1, according to which under the posed assumptions the calmness of $(0,\bar{x},\bar{y})$ implies the calmness of $\tilde{M}$ at $(0,0,\bar{x},\bar{y},\lambda)$ for all $\lambda\in\Lambda(0,\bar{x},\bar{y})$. Finally, in statement 4, the equivalence of Theorems 2.1 and 2.2 follows from Theorem 3.1 and [7, Theorem 3.1]. The second assertion follows from the fact that under LICQ at $\bar{y}$, condition (24) ensures both the calmness of M at $(0,\bar{x},\bar{y})$ as well as the calmness of $\tilde{M}$ at $(0,\bar{x},\bar{y})$, where $\bar{\lambda}$ is the unique multipliers from $\Lambda(0,\bar{x},\bar{y})$. □

5 MPECs with structured equilibria

Some of the tools and/or results from the preceding part of the paper can be utilized in deriving stationarity conditions for MPECs with equilibria governed by GEs having a special structure. In Section 5.1 we illustrate this fact by such an equilibrium with a polyhedral constraint set. In Section 5.2 we then apply a result from Section 5.1 to a class of bilevel programming problems arising in models of electricity spot markets.

5.1 Structured equilibria with polyhedral constraint sets

Let us consider a generalized equation of the considered type where

$$F(x,y)=\begin{pmatrix}F_1(x,y)\\F_2(x,y)\end{pmatrix},\ q(y)=Ay-b \tag{43}$$

with $F_1:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^{m_1}$, $F_2:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^{m_2}$, $A=(A_1,A_2)$ and $y=(y_1,y_2)\in\mathbb{R}^{m_1}\times\mathbb{R}^{m_2}$. Even though there is no structural difference between F_1 and F_2 yet, we will impose different assumptions on them later in the text. Structure (43) with $F_2(x,y)\equiv F_2(y)$ arises typically in a hierarchical bilevel multileader game where one looks for a Nash equilibrium on the upper level. In this case we obtain a finite number of MPECs in which the equilibria on the lower level are governed by generalized equation having the special structure (43), see e.g. [8].

In this case it may be reasonable to define the mappings S_1, S_2, employed in Section 3, in a different way, namely

$$\begin{aligned} S_1(z_1) &:= \left\{ (x,y,\lambda) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^s \,\middle|\, z_1 = F_1(x,y) + A_1^\top \lambda \right\} \\ S_2(z_2,z_3) &:= \left\{ (x,y,\lambda) \in \omega \times \mathbb{R}^m \times \mathbb{R}^s \,\middle|\, z_2 = F_2(x,y) + A_2^\top \lambda,\ q(y) - z_3 \in N_{\mathbb{R}^s_+}(\lambda) \right\}. \end{aligned} \tag{44}$$

Theorem 5.1 *In the setting of* (43) *fix some* $(\bar{x},\bar{y}) \in M(0)$ *and* $\bar{\lambda} \in \Lambda(0,\bar{x},\bar{y})$. *Assume that* $\omega = \mathbb{R}^n$, $F_2(x,y) \equiv F_2(y)$ *is affine linear and that* $\nabla_x F_1(\bar{x},\bar{y})$ *is surjective. Then* $\tilde{M}$ *is calm at* $(\bar{z}_1,\bar{z}_2,\bar{z}_3,\bar{x},\bar{y},\lambda)$ *for* $(\bar{z}_1,\bar{z}_2,\bar{z}_3) = (0,0,0)$ *and all* $\lambda \in \Lambda(0,\bar{x},\bar{y})$. *If in addition* Γ *has nonempty interior, then* M *is calm at* $(0,\bar{x},\bar{y})$.

Proof. Clearly $\tilde{M}(z_1,z_2,z_3) = S_1(z_1) \cap S_2(z_2,z_3)$. We will apply Lemma 2.2. Due to Lemma 2.1 and the assumed surjectivity of $\nabla_x F_1(\bar{x},\bar{y})$ we obtain that S_1 is calm at $(0,\bar{x},\bar{y},\bar{\lambda})$. As S_2 is generated by affine linear functions, it is calm at every point of its graph and it remains to verify condition (11), which takes the form

$$\left.\begin{aligned} (\nabla_x F_1(\bar{x},\bar{y}))^\top a &\in -N_\omega(\bar{x}) \\ (\nabla_y F_1(\bar{x},\bar{y}))^\top a + (\nabla_y F_2(\bar{y}))^\top d + A^\top c &= 0 \\ -A_1 a - A_2 d &\in D^* N_{\mathbb{R}^s_+}(\bar{\lambda}, A\bar{y} - b)(-c) \end{aligned}\right\} \Longrightarrow a = 0.$$

However, we easily conclude that this condition is fulfilled by virtue of $\omega = \mathbb{R}^n$ and the surjectivity of $\nabla_x F_1(\bar{x},\bar{y})$. The last statement follows directly from Proposition 3.1 and the equivalence of nonempty interior and MFCQ for polyhedral sets. ☐

By relaxing the assumptions imposed on F_2, we obtain the following weaker statement.

Theorem 5.2 *In the setting of* (43) *fix some* $(\bar{x},\bar{y}) \in M(0)$ *and* $\bar{\lambda} \in \Lambda(0,\bar{x},\bar{y})$. *Assume first that function*

$$(x,y,\lambda) \mapsto F_1(x,y) + A_1^\top \lambda \tag{45}$$

satisfies the assumptions of Lemma 2.1 and that the following system is satisfied

$$\left.\begin{aligned} (\nabla_x F_1(\bar{x},\bar{y}))^\top a + (\nabla_x F_2(\bar{x},\bar{y}))^\top d &\in -N_\omega(\bar{x}) \\ (\nabla_y F_1(\bar{x},\bar{y}))^\top a + (\nabla_y F_2(\bar{x},\bar{y}))^\top d + A^\top c &= 0 \\ -A_1 a - A_2 d &\in D^* N_{\mathbb{R}^s_+}(\bar{\lambda}, A\bar{y} - b)(-c) \end{aligned}\right\} \Longrightarrow a = 0. \tag{46}$$

Moreover, assume that at least one of the three following assumptions is satisfied:

1 F_2 *is affine linear;*

2 $\nabla_x F_2(\bar{x},\bar{y})$ *has full row rank;*

3 *MFCQ holds at* $\bar{y}$ *and denoting*

$$G := \left\{ c \in \mathbb{R}^{m_2} \,\middle|\, (\nabla_x F_2(\bar{x},\bar{y}))^\top c = 0 \right\}, \tag{47}$$

then for all $c \in G \setminus \{0\}$ *we have*

$$c^\top \nabla_{y_2} F_2(\bar{x},\bar{y}) c > 0. \tag{48}$$

Then $\tilde{M}$ is calm at $(0,0,0,\bar{x},\bar{y},\bar{\lambda})$.

Proof. Again we will employ Lemma 2.2 with the same partition of $\tilde{M}$ into S_1 and S_2 as in Theorem 5.1. Since (11) takes form (46), to finish the proof it remain to verify the calmness of S_2 at $(0,0,\bar{x},\bar{y},\bar{\lambda})$. It is simple to see that this property holds provided F_2 is affine or $\nabla_x F_2(\bar{x},\bar{y})$ has full row rank. Hence, we assume that (48) holds. First, we will pass to a different mapping $\hat{S}_2$ whose calmness ensures the calmness of S_2 and then we will verify even the Aubin property of $\hat{S}_2$. This will then finish the whole proof.

Define the following mapping

$$\tilde{S}_2(z_1,z_2,z_3) := \left\{ (x,y,\lambda,v) \,\middle|\, \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} v \\ F_2(x,y) \end{pmatrix} + \begin{pmatrix} A_1^\top \\ A_2^\top \end{pmatrix} \lambda,\ q(y) - z_3 \in N_{\mathbb{R}^s_+}(\lambda) \right\}$$

and show that if $\tilde{S}_2$ is calm at $(0,0,0,\bar{x},\bar{y},\bar{\lambda},-A_1^\top\bar{\lambda})$, then S_2 is calm at $(0,0,\bar{x},\bar{y},\bar{\lambda})$. Indeed, consider any $(x,y,\lambda) \in S_2(z_2,z_3)$ close to $(\bar{x},\bar{y},\bar{\lambda})$ and $(0,0)$, respectively. Then we have $(x,y,\lambda,-A_1^\top\lambda) \in \tilde{S}_2(0,z_2,z_3)$ and due to the calmness of $\tilde{S}_2$ we obtain

$$d((x,y,\lambda,-A_1^\top\lambda),\tilde{S}_2(0,0,0)) \le L\|(z_2,z_3)\|. \tag{49}$$

Find any $(\tilde{x},\tilde{y},\tilde{\lambda},\tilde{v}) \in \tilde{S}_2(0,0,0)$ minimizing the distance on the left-hand side of (49). Then we have

$$\begin{aligned} d((x,y,\lambda,-A_1^\top\lambda),\tilde{S}_2(0,0,0)) &= \|(x,y,\lambda,-A_1^\top\lambda) - (\tilde{x},\tilde{y},\tilde{\lambda},\tilde{v})\| \\ &\ge \|(x,y,\lambda) - (\tilde{x},\tilde{y},\tilde{\lambda})\| \ge d((x,y,\lambda),S_2(0,0)) \end{aligned} \tag{50}$$

because $(\tilde{x},\tilde{y},\tilde{\lambda}) \in S_2(0,0)$. Combining (49) and (50) yields that the calmness of $\tilde{S}_2$ implies the calmness of S_2.

If we apply Theorem 3.1 to $\tilde{S}_2$, where we consider the partition of (x,y,v) into (x,v) and y, then we obtain that the desired calmness of $\tilde{S}_2$ at $(0,0,0,\bar{x},\bar{y},\bar{\lambda},-A_1^\top\bar{\lambda})$ is equivalent to the calmness of

$$\hat{S}_2(z_1,z_2) := \left\{ (x,y,v) \,\middle|\, \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} v \\ F_2(x,y) \end{pmatrix} + N_\Gamma(y) \right\}$$

at $(0,0,\bar{x},\bar{y},-A_1\bar{\lambda}) = (0,0,\bar{x},\bar{y},F_1(\bar{x},\bar{y}))$. The Aubin property of $\hat{S}_2$ around this point is due to the Mordukhovich criterion equivalent to the following implication

$$\left. \begin{aligned} &(\nabla_x F_2(\bar{x},\bar{y}))^\top c = 0 \\ &\begin{pmatrix} (\nabla_y F_2(\bar{x},\bar{y}))^\top c \\ 0 \\ c \end{pmatrix} \in N_{\operatorname{gph} N_\Gamma}(\bar{y},-F_1(\bar{x},\bar{y}),-F_2(\bar{x},\bar{y})) \end{aligned} \right\} \implies c = 0. \tag{51}$$

If c satisfies the left-hand side of (51), then [9, Proposition 3.2] tells us that

$$0 \ge c^\top \nabla_y F_2(\bar{x},\bar{y}) \begin{pmatrix} 0 \\ c \end{pmatrix} = c^\top (\nabla_{y_1} F_2(\bar{x},\bar{y}), \nabla_{y_2} F_2(\bar{x},\bar{y})) \begin{pmatrix} 0 \\ c \end{pmatrix} = c^\top \nabla_{y_2} F_2(\bar{x},\bar{y}) c.$$

Due to the third assumption this implies that $c = 0$. Thus, formula (51) indeed holds and the statement has been proved. □

We are aware that by using [9, Proposition 3.2], we could obtain the same result even for a smaller set G, thus weakening the assumptions of the theorem. However, for the presentational simplicity, we prefer to keep it in the current form.

5.2 Application to a class of bilevel programming problems

As an application of the results from the previous section we introduce a special class of bilevel programming problems automatically satisfying the calmness conditions required for deriving necessary optimality conditions according to Theorem 2.1:

$$\min_{x,y}\{\varphi(x,y)|y\in\operatorname{argmin}\{f(x,z)|z\in\Gamma\}\} \tag{52}$$

with $f(x,z):=\langle x^a,z^a\rangle+\delta(x^b,z^a)+\langle z^b,Cz^b\rangle+\langle c,z^b\rangle$. Here, $x=\left(x^a,x^b\right)$, $z=\left(z^a,z^b\right)$, C is a positive semi-definite matrix of appropriate size, Γ is a polyhedron described by the linear inequality system $\Gamma:=\{z|\,Az\leq b\}$ with nonempty interior, φ is continuously differentiable and δ is twice continuously differentiable and convex in the second variable. Before deriving the corresponding necessary optimality conditions, we provide an example for this class of problems:

Example 5.1 *A special instance of (52) occurs in Equilibrium Problems with Equilibrium Constraints (EPECs) as they arise in certain electricity spot markets (see, e.g., [1, 8]). In this model, each player (power producer) optimizes his own decision while fixing the decisions of his competitors. More precisely, each player $i\in\{1,\dots,N\}$ solves the bilevel problem*

$$\min_{\alpha_i,\beta_i}\left\{\alpha_i g_i+2\beta_i g_i^2|\,(g,t)\in\operatorname{argmin}\left\{\sum\nolimits_{j=1}^N\alpha_j\tilde{g}_j+\beta_j\tilde{g}_j^2|\,\tilde{g}+A\tilde{t}\geq d\right\}\right\}, \tag{53}$$

where $\left(\alpha_j,\beta_j\right)$ is the decision vector of player j (representing 2 coefficients of a quadratic bidding curve), $(\tilde{g},\tilde{t})$ is the vector of power produced at the nodes and transmitted along the arcs of some network and the linear inequality system describes the satisfaction of a given demand vector d in the network using its incidence matrix A. In this MPEC, the decisions of competitors $j\neq i$ are supposed to be fixed. Putting

$$x^a:=\alpha_i,\,x^b:=\beta_i,\,z^a:=\tilde{g}_i,\,z^b:=(\tilde{g}_{-i},\tilde{t}),$$

where the lower index $-i$ refers to the subvector in which index i is omitted, the lower level objective function in (53) can be rewritten as

$$x^az^a+x^b\,(z^a)^2+\langle\alpha_{-i},\tilde{g}_{-i}\rangle+\langle\tilde{g}_{-i},[\operatorname{diag}\beta_{-i}]\,\tilde{g}_{-i}\rangle.$$

Here, $[\operatorname{diag}\beta_{-i}]$ refers to the diagonal matrix built up from the (fixed) vector β_{-i}. Now, (53) can be recast in the form of (52) with

$$\delta(x^b,z^a):=x^b\,(z^a)^2,\,c:=(\alpha_{-i},0),\,C:=\begin{pmatrix}[\operatorname{diag}\beta_{-i}] & 0\\ 0 & 0\end{pmatrix},\,\Gamma:=\{z|Bz\geq d\},\,B:=\begin{pmatrix}A & 0\\ 0 & I\end{pmatrix}.$$

Since all quadratic bidding coefficients β_j in the problem are required to be nonnegative, it follows that C is positive semidefinite. △

Theorem 5.3 *Let $(\bar{x},\bar{y})$ be a solution to (52). Under the assumptions made (without any further constraint qualification), there exist multipliers v^*, $u^*=\left(u_1^*,u_2^*\right)$ such that:*

$$\begin{aligned}
0 &= \nabla_{y^a}\varphi(\bar{x},\bar{y})-\nabla^2_{y^ay^a}\delta(\bar{x}^b,\bar{y}^a)\nabla_{x^a}\varphi(\bar{x},\bar{y})+u_1^*\\
0 &= \nabla_{y^b}\varphi(\bar{x},\bar{y})+(C+C^T)v^*+u_2^*\\
0 &= \nabla_{x^b}\varphi(\bar{x},\bar{y})-\nabla^2_{y^a,x^b}\delta(\bar{x}^b,\bar{y}^a)\nabla_{x^a}\varphi(\bar{x},\bar{y})\\
u^* &\in D^*N_\Gamma(\bar{y},-F(\bar{x},\bar{y}))(-\nabla_{x^a}\varphi(\bar{x},\bar{y}),v^*).
\end{aligned}$$

Proof. Thanks to our assumptions, the lower level problem $\min\{f(x,z)|z\in\Gamma\}$ is convex. Hence (52) can be equivalently rewritten as the MPEC

$$\min\{\varphi(x,y)|0\in\nabla_y f(x,y)+N_\Gamma(y)\}. \tag{54}$$

Observing that

$$\nabla_y f(x,y)=\left(x^a+\nabla_{y^a}\delta(x^b,y^a),(C+C^T)y^b+c\right)=:F(x,y),$$

(54) fits to the setting (43) with $F_1(x,y):=x^a+\nabla_{y^a}\delta(x^b,y^a)$ and $F_2(y):=(C+C^T)y^b+c$. Clearly, $\nabla_x F_1(\bar{x},\bar{y})=\left(I|\nabla^2_{y^a,x^b}\delta(\bar{x}^b,\bar{y}^a)\right)$ is surjective and F_2 is affine linear. Finally, due to the assumption of Γ having nonempty interior, MFCQ is satisfied for the description $\Gamma:=\{z|Az\leq b\}$. This allows us to apply Theorem 5.1 and to derive the calmness of M at $(0,\bar{x},\bar{y})$ via Proposition 3.1. Now, Theorem 2.1 yields in our special setting the existence of some multiplier $v=(v_1,v_2)$ such that

$$\begin{aligned}
0 &= \nabla_{x^a}\varphi(\bar{x},\bar{y})+v_1\\
0 &= \nabla_{x^b}\varphi(\bar{x},\bar{y})+\left[\nabla^2_{x^b,y^a}\delta(\bar{x}^b,\bar{y}^a)\right]^T v_1\\
u^* &\in D^*N_\Gamma(\bar{y},-F(\bar{x},\bar{y}))(v_1,v_2),
\end{aligned}$$

where $u^*:=-\nabla_y\varphi(\bar{x},\bar{y})-[\nabla_y F(\bar{x},\bar{y})]^T v$. In particular, $u^*=(u_1^*,u_2^*)$, where

$$\begin{aligned}
u_1^* &= -\nabla_{y^a}\varphi(\bar{x},\bar{y})-\nabla^2_{y^a,y^a}\delta(\bar{x}^b,\bar{y}^a)v_1\\
u_2^* &= -\nabla_{y^b}\varphi(\bar{x},\bar{y})-(C+C^T)v_2.
\end{aligned}$$

Combining all the obtained relations upon substituting for v_1 and setting $v^*:=v_2$, one arrives at the necessary conditions asserted in the theorem. ☐

References

[1] Aussel, D., Correa, R., Marechal, M.: Electricity spot market with transmission losses. Journal of Industrial and Management Optimization **9**, 275–290 (2013)

[2] Dempe, S., Dutta, J.: Is bilevel programming a special case of a mathematical program with complementarity constraints? Math. Program. A **131**(1-2), 37–48 (2012)

[3] Dontchev, A., Rockafellar, R.T.: Implicit Functions and Solution Mappings. Springer (2009)

[4] Gfrerer, H.: First order and second order characterizations of metric subregularity and calmness of constraint set mappings. SIAM J. Optim. **21**(4), 1439–1474 (2011)

[5] Gfrerer, H., Outrata, J.: On computation of limiting coderivatives of the normal-cone mapping to inequality systems and their applications. Optimization (accepted)

[6] Henrion, R., Jourani, A., Outrata, J.: On the calmness of a class of multifunctions. SIAM J. Optim. **13**(2), 603–618 (2002)

[7] Henrion, R., Outrata, J., Surowiec, T.: On the co-derivative of normal cone mappings to inequality systems. Nonlinear Anal. Theory, Methods Appl. **71**(3-4), 1213–1226 (2009)

[8] 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**(2), 295–317 (2012)

[9] Henrion, R., Römisch, W.: On M-stationary points for a stochastic equilibrium problem under equilibrium constraints in electricity spot market modeling. Appl. Math. **52**(6), 473–494 (2007)

[10] Ioffe, A.D., Outrata, J.: On metric and calmness qualification conditions in subdifferential calculus. Set-Valued Anal. **16**(2-3), 199–227 (2008)

[11] Klatte, D., Kummer, B.: Constrained minima and Lipschitzian penalties in metric spaces. SIAM J. Optim. **13**(2), 619–633 (2002)

[12] Minchenko, L., Stakhovski, S.: Parametric nonlinear programming problems under the relaxed constant rank condition. SIAM J. Optim. **21**(1), 314–332 (2011)

[13] Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation I. Springer (2006)

[14] Mordukhovich, B.S., Outrata, J.: Coderivative analysis of quasi-variational inequalities with applications to stability and optimization. SIAM J. Optim. **18**(2), 389–412 (2007)

[15] Outrata, J.: A generalized mathematical program with equilibrium constraints. SIAM J. Control Optim. **38**(5), 1623–1638 (2000)

[16] Outrata, J.V.: Optimality conditions for a class of mathematical programs with equilibrium constraints. Mathematics of Operations Research **24**(3), 627–644 (1999)

[17] Robinson, S.M.: Some continuity properties of polyhedral multifunctions. In: H. König, B. Korte, K. Ritter (eds.) Mathematical Programming at Oberwolfach, *Mathematical Programming Studies*, vol. 14, pp. 206–214. Springer Berlin Heidelberg (1981)

[18] Rockafellar, R.T., Wets, R.J.B.: Variational Analysis. Springer (1998)

[19] Scholtes, S., Stöhr, M.: How stringent is the linear independence assumption for mathematical programs with complementarity constraints? Mathematics of Operations Research **26**(4), 851–863 (2001)

[20] Surowiec, T.: Explicit stationarity conditions and solution characterization for equilibrium problems with equilibrium constraints. PhD thesis, Humboldt University Berlin (2010)

[21] Ye, J.J.: Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints. SIAM J. Optim. **10**(4), 943–962 (2000)

[22] Ye, J.J., Ye, X.Y.: Necessary optimality conditions for optimization problems with variational inequality constraints. Math. Oper. Res. **22**(4), 977–997 (1997)

Appendix

Example .2 *In this example we construct a set Γ satisfying MFCQ at given $\bar{y}$ and a function F such that M is calm at $(0,\bar{x},\bar{y})$ while $\tilde{M}$ is not calm at $(0,0,\bar{x},\bar{y},\lambda)$ for any $\lambda \in \Lambda(0,\bar{x},\bar{y})$.*

Define first $\varphi_1,\varphi_2 : [-1,1] \to \mathbb{R}$ and $q_1,q_2 : [-1,1]\times\mathbb{R} \to \mathbb{R}$ as

$$\begin{aligned}
\varphi_1(t) &:= \begin{cases} (-1)^k\left(t-\frac{1}{k}\right)^3\left(t-\frac{1}{k+1}\right)^3 & \text{for } t \in \left[\frac{1}{k+1},\frac{1}{k}\right],\ k\in\mathbb{N} \\ 0 & \text{for } t\le 0 \end{cases}\\
\varphi_2(t) &:= \begin{cases} (-1)^k\left(t-\frac{1}{k}\right)^5\left(t-\frac{1}{k+1}\right)^5 & \text{for } t \in \left[\frac{1}{k+1},\frac{1}{k}\right],\ k\in\mathbb{N} \\ 0 & \text{for } t\le 0 \end{cases}\\
q_1(y) &:= \varphi_1(y_1)-y_2\\
q_2(y) &:= \varphi_2(y_1)-y_2,
\end{aligned}$$

put $\omega=\mathbb{R}$ and as the reference point take $(\bar{x},\bar{y}_1,\bar{y}_2)=(0,0,0)$. These functions are depicted in Figure 1. Note first that MFCQ is indeed satisfied for Γ. Moreover, it is easy to verify that φ_1 and φ_2

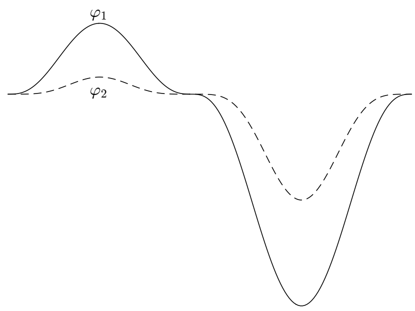

Figure 1: Segments of graphs φ_1 and $2.3\cdot 10^9\varphi_2$. The constant in front of φ_2 is used for graphical purposes.

are twice continuously differentiable. Define further

$$\phi(t) := \max\{\varphi_1(t),\varphi_2(t)\}.$$

The twice continuous differentiability of ϕ is obvious apart from 0. At 0 we compute

$$|\phi'(0)| = \lim_{t\to 0} t^{-1}|\phi(t)-\phi(0)| = \lim_{t\to 0} t^{-1}|\varphi_1(t)| = |\varphi_1'(0)| = 0$$

and similarly we obtain $\phi''(0)=0$ and hence ϕ is twice continuously differentiable. Finally, we define $F(x,y):=(-\phi'(y_1),\ 1)$. By construction of ϕ, we obtain that F is continuously differentiable. Since $\Gamma=\operatorname{epi}\phi$ we have that

$$M(0) = \left\{(x,y)\,\middle|\, \begin{pmatrix}\phi'(y_1)\\ -1\end{pmatrix} \in N_\Gamma(y)\right\} = \mathbb{R}\times\operatorname{gph}\phi.$$

As $M(p)\subset M(0)$ for all p small enough, we obtain that M is calm at $(0,\bar{x},\bar{y})$. It is easy to see that $\Lambda(0,\bar{x},\bar{y}) = \{\lambda\ge 0 \,|\, \lambda_1+\lambda_2=1\}$. We will show now that $\tilde{M}$ is not calm at $(0,0,\bar{x},\bar{y},\lambda)$ for any

$\lambda \in \Lambda(0,\bar{x},\bar{y})$. Define

$$\begin{aligned}
\Omega_1 &:= \{t \in [0,1] |\, \varphi_1(t) = \varphi_2(t)\} \\
\Omega_2 &:= \{t \in [0,1] |\, \varphi_1(t) \neq \varphi_2(t),\ \varphi_1'(t) = \varphi_2'(t)\} \\
\Omega_3 &:= [0,1] \setminus (\Omega_1 \cup \Omega_2)
\end{aligned}$$

and note that for all $t \in \Omega_2 \cup \Omega_3$ small enough it holds that $|\varphi_2(t)| < |\varphi_1(t)|$ and for all $t \in \Omega_3$ small enough we have $|\varphi_2'(t)| < |\varphi_1'(t)|$.

We will show first that $T_{\{1\}}$ defined in (18) *is not calm at $(0,\bar{y})$. From the definition we see that*

$$T_{\{1\}}(p) = \{y |\, \varphi_1(y_1) = y_2 + p_1, \varphi_2(y_1) \le y_2 + p_2\}.$$

and thus

$$T_{\{1\}}(0) = \{y |\, \varphi_1(y_1) = y_2, \varphi_2(y_1) \le y_2\} = \{(y_1, \varphi_1(y_1)) |\, \varphi_1(y_1) \ge 0\}.$$

Now pick any sequence $y_{k1} > 0$, $y_{k1} \to 0$ such that $y_{k1} \in \Omega_2$ and $\varphi_1(y_{k1}) < 0$ and define $p_{k1} := 0$, $y_{k2} := \varphi_1(y_{k1})$ and $p_{k2} := \varphi_2(y_{k1}) - y_{k2}$. Then $y_k \in T_{\{1\}}(p_k)$. Moreover, as φ_1 and φ_2 have the same signs

$$0 < p_{k2} = \varphi_2(y_{k1}) - y_{k2} = \varphi_2(y_{k1}) - \varphi_1(y_{k1}) \le |\varphi_1(y_{k1})|.$$

Consider now a point $\tilde{y}_{k1} \in \Omega_1$ at which $d(y_{k1},\Omega_1)$ is realized. Then we obtain

$$\frac{|d(y_k, T_{\{1\}}(0))|}{|p_k|} \ge \frac{|d(y_{k1},\Omega_1)|}{|\varphi_1(y_{k1})|} = \frac{|y_{k1} - \tilde{y}_{k1}|}{|\varphi_1(y_{k1}) - \varphi_1(\tilde{y}_{k1})|} = \frac{1}{\varphi_1'(\xi_k)},$$

where in the last equality we have used the mean value theorem to find some ξ_k which lies in the line segment connecting y_{k1} and $\tilde{y}_{k1}$. Since φ_1 is twice continuously differentiable with $\varphi_1'(0) = 0$, we have proved that $T_{\{1\}}$ is not calm at $(0,\bar{y})$. For $T_{\{2\}}$ we proceed with a similar construction. In this case we have

$$T_{\{2\}}(0) = \{y |\, \varphi_1(y_1) \le y_2, \varphi_2(y_1) = y_2\} = \{(y_1, \varphi_2(y_1)) |\, \varphi_1(y_1) \le 0\}$$

and for the contradicting sequence we choose some $y_{k1} > 0$, $y_{k1} \to 0$ such that $y_{k1} \in \Omega_2$ and $\varphi_1(y_{k1}) > 0$ and define again $p_{k1} := 0$, $y_{k2} := \varphi_1(y_{k1})$ and $p_{k2} := \varphi_2(y_{k1}) - y_{k2}$ and perform the estimates as in the previous case. Since for $T_{\{1,2\}}$ we have

$$T_{\{1,2\}}(0) = \{y |\, \varphi_1(y_1) = y_2, \varphi_2(y_1) = y_2\} = \{(y_1, \varphi_1(y_1)) |\, \varphi_1(y_1) = 0\}$$

either of the previous contradicting sequences can be chosen.

Fix now any $\bar{\lambda} \in \Lambda(0,\bar{x},\bar{y})$ and consider the corresponding index set $I = \{i |\, \bar{\lambda}_i > 0\}$. In the previous several paragraphs we have shown that T_I is not calm at $(0,\bar{y})$ and found a sequence $(\tilde{p}_k, \tilde{y}_k)$ violating the calmness property. By virtue of Lemma 3.2 we obtain that T is not calm at $(0,\bar{y},\bar{\lambda})$. Moreover, from the proof of this lemma we see that the sequence (p_k, y_k, λ_k), which violates the calmness of T at $(0,\bar{y},\bar{\lambda})$, can be taken in such a way that $p_k = \tilde{p}_k$, $y_k = \tilde{y}_k$ and $\lambda_k = \bar{\lambda}$ with $(\tilde{y}_k, \bar{\lambda}) \in T(\tilde{p}_k)$ and

$$d((\tilde{y}_k, \bar{\lambda}), T(0)) > k\|\tilde{p}_k\|. \tag{55}$$

Furthermore, in all the previous cases we have chosen $\tilde{y}_k$ in such a way that $\tilde{y}_{k1} \in \Omega_2$.

We will show that $\tilde{M}$ is not calm at $(0,0,\bar{x},\bar{y},\bar{\lambda})$. Consider sequence

$$(0,0,\tilde{p}_{k1},\tilde{p}_{k2},\bar{x},\tilde{y}_{k1},\tilde{y}_{k2},\bar{\lambda}_1,\bar{\lambda}_2) \to (0,0,0,0,\bar{x},0,0,\bar{\lambda}_1,\bar{\lambda}_2) \tag{56}$$

and show first that $(\bar{x},\tilde{y}_{k1},\tilde{y}_{k2},\bar{\lambda}_1,\bar{\lambda}_2)\in\tilde{M}(0,0,\tilde{p}_{k1},\tilde{p}_{k2})$, which amounts to showing

$$\begin{pmatrix}0\\0\end{pmatrix}=\begin{pmatrix}-\phi'(\tilde{y}_{k1})\\1\end{pmatrix}+\begin{pmatrix}\varphi_1'(\tilde{y}_{k1}) & \varphi_2'(\tilde{y}_{k1})\\-1 & -1\end{pmatrix}\begin{pmatrix}\bar{\lambda}_1\\\bar{\lambda}_2\end{pmatrix}$$
$$q(\tilde{y}_k)-\tilde{p}_k\in N_{\mathbb{R}^2_+}(\bar{\lambda}).$$

We know that $(\tilde{y}_k,\bar{\lambda})\in T(\tilde{p}_k)$ and hence the inclusion is satisfied. Moreover, as $\tilde{y}_{k1}\in\Omega_2$ by construction of this sequence and as $\bar{\lambda}_1+\bar{\lambda}_2=1$, we indeed obtain

$$(\bar{x},\tilde{y}_{k1},\tilde{y}_{k2},\bar{\lambda}_1,\bar{\lambda}_2)\in\tilde{M}(0,0,\tilde{p}_{k1},\tilde{p}_{k2}). \tag{57}$$

Because the relations defining $\tilde{M}$ do not depend on x and one of them defines also T, we infer that $\tilde{M}(z_1,z_2,z_3,z_4)\subset\mathbb{R}^n\times T(z_3,z_4)$ and consequently due to (55) *we obtain*

$$d((\bar{x},\tilde{y}_{k1},\tilde{y}_{k2},\bar{\lambda}_1,\bar{\lambda}_2),\tilde{M}(0,0,0,0))\geq d((\tilde{y}_{k1},\tilde{y}_{k2},\bar{\lambda}_1,\bar{\lambda}_2),T(0,0))>k\|\tilde{p}_k\|.$$

This together with (56) *and* (57) *implies that $\tilde{M}$ is indeed not calm at $(0,0,\bar{x},\bar{y},\bar{\lambda})$. Since $\bar{\lambda}$ was chosen arbitrarily from $\Lambda(0,\bar{x},\bar{y})$, the construction has been completed.* △

Lemma .1 *If FRCQ is satisfied at some $\bar{y}$ with $q(\bar{y})\leq 0$, then there is an $\varepsilon>0$ such that $\#I(y)\leq m-1$ for all $y\in\mathbb{B}_\varepsilon(\bar{y})\setminus\{\bar{y}\}$.*

Proof. Otherwise, there is a sequence $y^k\to\bar{y}$ with $y^k\neq\bar{y}$ and $\#I\left(y^k\right)\geq m$. Upon passing to a subsequence, we may assume without loss of generality that $I\left(y^k\right)\equiv I$ for some $I\subset\{1,\ldots,s\}$ and that $\#I\geq m$. From $q_i\left(y^k\right)=0$ for all $i\in I$ and all $k\in\mathbb{N}$ it follows that $q_i(\bar{y})=0$ for all $i\in I$, whence $I\subset I(\bar{y})$. Now, FRCQ implies that $\operatorname{rank}\{\nabla q_i(\bar{y})\}_{i\in I}=m$. By the classical inverse function theorem, $\bar{y}$ is the locally unique solution of the equation $q_i(y)=0\quad(i\in I)$. This, however, contradicts the fact that $q_i\left(y^k\right)=0\quad(i\in I)$ for a sequence $y^k\to\bar{y}$ with $y^k\neq\bar{y}$. □

For the following lemma recall that W_y has been defined in (26).

Lemma .2 *Let $\bar{y}$ with $q(\bar{y})=0$ be given. If q is affine linear, then, there exist $L>0$ such that for all $y\in\mathbb{R}^m$, for all $(p_1,p_2,\bar{p}_1,\bar{p}_2)$ with $W_y(\bar{p}_1,\bar{p}_2)\neq\emptyset$ and for all $\lambda\in W_y(p_1,p_2)$ the following estimate holds true:*

$$d\left(\lambda,W_y(\bar{p}_1,\bar{p}_2)\right)\leq L\max\left\{\|p_1-\bar{p}_1\|,\|p_2-\bar{p}_2\|\right\}. \tag{58}$$

Alternatively, if FRCQ is satisfied at $\bar{y}$, then there exist $\varepsilon,L>0$ such that for all $y\in\mathbb{B}_\varepsilon(\bar{y})$, for all $(p_1,p_2,\bar{p}_1)$ with $W_y(\bar{p}_1,q(y))\neq\emptyset$ and for all $\lambda\in W_y(p_1,p_2)$ the following estimate holds true:

$$d\left(\lambda,W_y(\bar{p}_1,q(y))\right)\leq L\|p_1-\bar{p}_1\|. \tag{59}$$

Proof. Assume first, that q is affine linear, i.e., $q(y)=Ay+b$. Then, for each fixed $y\in\mathbb{R}^m$,

$$\operatorname{gph}W_y=\left\{(p_1,p_2,\lambda)\,|A^\top\lambda=p_1,\,(\lambda,p_2)\in\operatorname{gph}N_{\mathbb{R}^s_+}\right\}=H^{-1}\left(\{0\}\times\operatorname{gph}N_{\mathbb{R}^s_+}\right),$$

where $H(p_1,p_2,\lambda)=\left(A^\top\lambda-p_1,\lambda,p_2\right)$. Observe, that $\operatorname{gph}W_y$ does actually not depend on y. Since $\{0\}\times\operatorname{gph}N_{\mathbb{R}^s_+}$ is a finite union of polyhedra, the same holds true for its preimage $\operatorname{gph}W_y$ under the linear mapping H. Hence W_y is a polyhedral mapping (not depending on y) and, by Robinson's

Theorem [3, Theorem 3D.1], there exists some $L > 0$ such that for all $y \in \mathbb{R}^m$, for all $(p_1, p_2, \bar{p}_1, \bar{p}_2)$ with $W_y(\bar{p}_1, \bar{p}_2) \neq \emptyset$ and for all $\lambda \in W_y(p_1, p_2)$ estimate (58) holds true.

For proving (59) assume that FRCQ is satisfied at $\bar{y}$ and define for any p_1 set

$$K(p_1) := \left\{ \lambda \,|\, (\nabla q(\bar{y}))^\top \lambda = p_1,\, \lambda \geq 0 \right\}.$$

By Hoffman's Lemma, there is some L_1 such that for any $\bar{p}_1$ with $K(\bar{p}_1) \neq \emptyset$ and any p_1 one has the estimate

$$d(\lambda, K(\bar{p}_1)) \leq L_1 \|p_1 - \bar{p}_1\| \quad \forall \lambda \in K(p_1). \tag{60}$$

Assume now that $y = \bar{y}$. Then, since $q(\bar{y}) = 0$, we have that $W_{\bar{y}}(\bar{p}_1, q(\bar{y})) = K(\bar{p}_1)$. Moreover, $\lambda \in W_y(p_1, p_2) = W_{\bar{y}}(p_1, p_2)$ implies that $\lambda \in K(p_1)$. Therefore, (60) yields

$$d(\lambda, W_y(\bar{p}_1, q(y))) = d(\lambda, W_{\bar{y}}(\bar{p}_1, q(\bar{y}))) = d(\lambda, K(\bar{p}_1)) \leq L_1 \|p_1 - \bar{p}_1\|$$

for any $\lambda \in W_y(p_1, p_2)$.

For the rest of the proof, assume that $y \neq \bar{y}$. Since FRCQ is satisfied at $\bar{y}$, we obtain that $\nabla q_I(\bar{y})$ is a surjective matrix for all $I \subset I(\bar{y})$ with $\#I \leq m$, where q_I stands for restriction of q on components I. In particular, the pseudo-inverse matrices to $\nabla q(\bar{y})^\top$

$$A_I(y) := \left(\nabla q_I(y) (\nabla q_I(y))^\top \right)^{-1} \nabla q_I(y) \quad (I \subset I(\bar{y}) : \#I \leq m), \tag{61}$$

exist and are continuous on some neighborhood $\mathscr{Y}$ of $\bar{y}$. Let $\varepsilon > 0$ be such that $\mathbb{B}_\varepsilon(\bar{y}) \subset \mathscr{Y}$ and $I(y) \subset I(\bar{y})$ for all $y \in \mathbb{B}_\varepsilon(\bar{y})$. Observing that the family of index sets $I \subset I(\bar{y})$ with $\#I \leq m$ is finite, the following quantity is well-defined and finite:

$$L_2 := \max\{\|A_I(y)\| \,| y \in \mathbb{B}_\varepsilon(\bar{y}), I \subset I(\bar{y}) : \#I \leq m\}.$$

Consider an arbitrary $y \in \mathbb{B}_\varepsilon(\bar{y})$ with $y \neq \bar{y}$ and arbitrary (p_1, p_2). Then, $I(y) \subset I(\bar{y})$ by definition of ε and $\#I(y) \leq m - 1$ by Lemma .1. Now, (61) yields that $\lambda = A_{I(y)}(y) p_1$ for any $\lambda \in W_y(p_1, p_2)$. Summarizing, we have that

$$W_y(p_1, p_2) = \{A_{I(y)}(y) p_1\} \tag{62}$$

for any $y \in \mathbb{B}_\varepsilon(\bar{y})$ with $y \neq \bar{y}$ and any (p_1, p_2) with $W_y(p_1, p_2) \neq \emptyset$. From (62) it follows that $W_y(\bar{p}_1, q(y)) = \{A_{I(y)}(y) \bar{p}_1\}$ which, together with (62), yields

$$d(\lambda, W_y(\bar{p}_1, q(y))) = \left\| A_{I(y)}(y) (p_1 - \bar{p}_1) \right\| \leq L_2 \|p_1 - \bar{p}_1\|.$$

By taking $L := \max\{L_1, L_2\}$, the proof is complete. □