ON M-STATIONARY POINTS FOR MATHEMATICAL PROGRAMS WITH EQUILIBIRUM CONSTRAINTS

Michael L. Flegel and Christian Kanzow

University of Würzburg

Institute of Applied Mathematics and Statistics

Am Hubland

97074 Würzburg

Germany

e-mail: flegel@mathematik.uni-wuerzburg.de
          kanzow@mathematik.uni-wuerzburg.de

March 30, 2004

Abstract. Mathematical programs with equilibrium constraints are optimization problems which violate most of the standard constraint qualifications. Hence the usual Karush-Kuhn-Tucker conditions cannot be viewed as first order optimality conditions unless relatively strong assumptions are satisfied. This observation has lead to a number of weaker first order conditions, with M-stationarity being the strongest among these weaker conditions. Here we show that M-stationarity is a first order optimality condition under a very weak Abadie-type constraint qualification. This result has recently been established by Jane Ye, but her proof contains a gap, so we try to give a complete and self-contained presentation here.

1 Introduction

We consider the following program, known across the literature as a *mathematical program with complementarity*—or often also *equilibrium*—*constraints*, *MPEC* for short:

$$
\begin{aligned}
\min \quad & f(z) \\
\text{s.t.} \quad & g(z) \le 0, \quad h(z) = 0, \\
& G(z) \ge 0, \quad H(z) \ge 0, \quad G(z)^T H(z) = 0,
\end{aligned} \tag{1}
$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$, $h : \mathbb{R}^n \to \mathbb{R}^p$, $G : \mathbb{R}^n \to \mathbb{R}^l$, and $H : \mathbb{R}^n \to \mathbb{R}^l$ are continuously differentiable.

It is easily verified that the standard Mangasarian-Fromovitz constraint qualification is violated at every feasible point of the program (1), see, e.g. [2]. The weaker Abadie constraint qualification can be shown to only hold in restrictive circumstances, see [15, 4]. A still weaker CQ, the Guignard CQ, has a chance of holding, see [4]. Any of the classic CQs imply that a Karush-Kuhn-Tucker point (called a strongly stationary point by the MPEC community) is a necessary first order condition.

However, because only the weakest constraint qualifications have a chance of holding, new constraint qualifications tailored to MPECs, and with it new stationarity concepts, have arisen, see, e.g., [9, 18, 15, 13, 14, 6, 22].

One of the stronger stationarity concepts introduced is M-stationarity [13] (see (5)). It is second only to strong stationarity. Weaker stationarity concepts like A- and C-stationarity have also been introduced [5, 18], but it is commonly held that these are too weak since such points allow for trivial descent directions to exist.

M-stationary points also play an important role for some classes of algorithms for the solution of MPECs. For example, Scholtes [19] has introduced an algorithm which, under certain assumptions to the MPEC (1), converges to an M-stationary point, but not in general to a strongly stationary point. Later, Hu and Ralph [7] proved a generalization of this result by showing that a limit point of a whole class of algorithms is an M-stationary point of the MPEC (1).

Hence it is of some importance to know when an M-stationary point is in fact a first order condition. This paper is dedicated to answering that question. We will show M-stationarity to be a necessary first order condition under MPEC-ACQ, an MPEC variant of the classic Abadie CQ, see [6]. It is the weakest of the MPEC constraint qualifications known to date.

The path we take in order to prove this was described by Ye [21, 22]. We introduce MPEC variants of calmness and a local error bound, and show M-stationarity to be a necessary first order condition under these constraint qualifications. Then we proceed to show that a mathematical program with *affine* equilibrium constraints (see (43)) satisfies these, finally using [22, Theorem 3.1] which proves our result.

Unfortunately, the proof of [21, Theorem 3.2] is erroneous (as the author agrees [23]), which invalidates [22, Theorem 3.1] since it relies on that result. The aim of this paper is therefore to close this gap and to present a complete proof for the fact that a local minimizer of the MPEC (1) is an M-stationary point under an Abadie-type constraint qualification.

To this end, we transfer several definitions and results known for OPVICs (optimization problems with variational inequality constraints) to the MPEC setting, yielding slightly different, more general, conditions. The proofs of our results are inspired by those of Ye [21].

We chose to use a result by Treiman [20, Corollary 4.2] to replace [21, Theorem 3.2]. This has two predominant advantages. For one, the Treiman result is in a format more compatible with our MPEC (1), as opposed to an OPVIC. Secondly, we avoid the Mordukhovich coderivative. Both of these facts facilitate the notation of our proofs considerably, in particular that of Theorem 3.2. For some background on coderivatives in general, see [17, Chapter 8.G], and we refer to [11] for the Mordukhovich coderivative in particular.

The organization of this paper is as follows. In Section 2 we introduce necessary concepts and collect some useful results about normal cones and subgradients. Section 3 starts off with introducing a Fritz John-type M-stationarity result which is proved using the Treiman result. We then define MPEC variants of calmness and error bounds, and also detail the relationship between MPEC-calmness and exact penalization. These results will play an important role in order to establish the main result of this paper.

A word on notation. Given two vectors x and y, we use $(x, y) := (x^T, y^T)^T$ for ease of notation. By $B_\varepsilon(x)$ we denote the open ball around x with diameter ε with respect to the Euclidian norm; the dimension is given by the dimension of x. Differential operators such as ∂ and ∇ are always applied to all arguments of the function following it. If a differential operator is applied to only a part of the arguments, this is denoted by an appropriate subscript, i.e. we have $\partial f(x, y) = (\partial_x f(x, y), \partial_y f(x, y))$. Note that this is contrary to the common use of the subscript. Functions such as max and min are understood componentwise when applied to vectors. Similarly, comparisons such as $\leq$ and $\geq$ are also understood componentwise. The Euclidian distance between a point x and a closed set X is denoted by $\text{dist}(x, X)$ while a (not necessarily unique) projection of x onto X is denoted by $\Pi_X(x)$. Given a matrix $A \in \mathbb{R}^{m \times n}$, the i-th row is denoted by A_i. Similarly, a_i denotes the i-th component of a vector $a \in \mathbb{R}^n$. Given a set $\delta \subseteq \{1, \ldots, n\}$ we denote by $x_\delta \in \mathbb{R}^{|\delta|}$ that vector which consists of those components of $x \in \mathbb{R}^n$ which correspond to the indices in δ. By $\mathbb{R}_+^l := \{x \in \mathbb{R}^l \mid x \geq 0\}$ we mean the nonnegative orthant of $\mathbb{R}^l$. Finally, by $\text{cl}(\cdot)$ and $\text{conv}(\cdot)$ we denote the closure and convex hull of a given set, respectively.

2 Preliminaries

We will now introduce some notation and concepts in the context of MPECs which we will need for the remainder of this paper. The reader is referred to [6] for more detail.

From the complementarity term in (1) it is clear that for a feasible point z^*, either $G_i(z^*)$, or $H_i(z^*)$, or both must be zero. To differentiate between these cases, we divide the indices of G and H into three sets:

$$\alpha := \alpha(z^*) := \{i \mid G_i(z^*) = 0, \ H_i(z^*) > 0\}, \tag{2a}$$

$$\beta := \beta(z^*) := \{i \mid G_i(z^*) = 0, \ H_i(z^*) = 0\}, \tag{2b}$$

$$\gamma := \gamma(z^*) := \{i \mid G_i(z^*) > 0,\ H_i(z^*) = 0\}. \tag{2c}$$

The set β is called the *degenerate set*.

The classic Abadie CQ is defined using the tangent cone of the feasible set of a mathematical program. The MPEC variant of the Abadie CQ (see Definition 2.1) also makes use of this tangent cone. If we denote the feasible set of (1) by $\mathcal{Z}$, the tangent cone of (1) in a feasible point z^* is defined by

$$\mathcal{T}(z^*) := \left\{d \in \mathbb{R}^n \mid \exists\{z^k\} \subset \mathcal{Z}, \exists t_k \searrow 0\ :\ z^k \to z^* \text{ and } \frac{z^k - z^*}{t_k} \to d\right\}. \tag{3}$$

Note that the tangent cone is closed, but in general not convex.

For the classic Abadie CQ, the constraints of the mathematical program are linearized. This makes less sense in the context of MPECs because information we keep for G and H, we throw away for the complementarity term (see also [4]). Instead, the authors proposed the *MPEC-linearized* tangent cone in [6],

$$
\begin{aligned}
\mathcal{T}_{\mathrm{MPEC}}^{lin}(z^*) := \{d \in \mathbb{R}^n \mid \nabla g_i(z^*)^T d \leq 0, & \qquad \forall i \in \mathcal{I}_g, \\
\nabla h_i(z^*)^T d = 0, & \qquad \forall i = 1,\ldots,p, \\
\nabla G_i(z^*)^T d = 0, & \qquad \forall i \in \alpha, \\
\nabla H_i(z^*)^T d = 0, & \qquad \forall i \in \gamma, \\
\nabla G_i(z^*)^T d \geq 0, & \qquad \forall i \in \beta, \\
\nabla H_i(z^*)^T d \geq 0, & \qquad \forall i \in \beta, \\
(\nabla G_i(z^*)^T d) \cdot (\nabla H_i(z^*)^T d) = 0, & \quad \forall i \in \beta\,\},
\end{aligned} \tag{4}
$$

where $\mathcal{I}_g := \{i \mid g_i(z^*) = 0\}$ is the set of indices such that the inequality constraint $g_i(z^*) \leq 0$ is active. Note that here, the component functions of the complementarity term have been linearized separately, so that we end up with a *quadratic term* in (4).

Similar to the classic case, it holds that

$$\mathcal{T}(z^*) \subseteq \mathcal{T}_{\mathrm{MPEC}}^{lin}(z^*).$$

This inspires the following variant of the Abadie CQ for MPECs.

Definition 2.1 *The MPEC (1) is said to satisfy* MPEC-Abadie CQ *(or* MPEC-ACQ *for short) in a feasible vector z^* if*

$$\mathcal{T}(z^*) = \mathcal{T}_{\mathrm{MPEC}}^{lin}(z^*)$$

holds.

We refer the reader to [6] for a rigorous discussion of MPEC-ACQ.

As mentioned in the introduction, various stationarity concepts have arisen for MPECs. Though we only need M-stationarity, we also state A-, C- and strong stationarity for completeness' sake, see [18, 15, 5] for more detail.

5

Let $z^* \in \mathcal{Z}$ be feasible for the MPEC (1). We call z^* *M-stationary* if there exists λ^g, λ^h, λ^G, and λ^H such that

$$0 = \nabla f(z^*) + \sum_{i=1}^m \lambda_i^g \nabla g_i(z^*) + \sum_{i=1}^p \lambda_i^h \nabla h_i(z^*) - \sum_{i=1}^l \left[\lambda_i^G \nabla G_i(z^*) + \lambda_i^H \nabla H_i(z^*) \right],$$

$$\begin{array}{llll}
\lambda_\alpha^G & \text{free}, & & \lambda_\gamma^G = 0, \\
& & (\lambda_i^G > 0 \wedge \lambda_i^H > 0) \vee \lambda_i^G \lambda_i^H = 0 \quad \forall i \in \beta & \\
\lambda_\gamma^H & \text{free}, & & \lambda_\alpha^H = 0,
\end{array} \tag{5}$$

$$g(z^*) \leq 0, \qquad \lambda^g \geq 0, \qquad g(z^*)^T \lambda^g = 0.$$

The other stationarity concepts differ from M-stationarity only in the restriction imposed upon λ_i^G and λ_i^H for $i \in \beta$, as detailed in the following list:

- strong stationarity [18, 15]: $\lambda_i^G \geq 0 \wedge \lambda_i^H \geq 0 \; \forall i \in \beta$;

- M-stationarity [13]: $(\lambda_i^G > 0 \wedge \lambda_i^H > 0) \vee \lambda_i^G \lambda_i^H = 0 \; \forall i \in \beta$;

- C-stationarity [18]: $\lambda_i^G \lambda_i^H \geq 0 \; \forall i \in \beta$;

- A-stationarity [5]: $\lambda_i^G \geq 0 \vee \lambda_i^H \geq 0 \; \forall i \in \beta$.

Note that the intersection of A- and C-stationarity yields M-stationarity and that strong stationarity implies M- and hence A- and C-stationarity. Also note that Pang and Fukushima [15] call a strongly stationary point a *primal-dual stationary point*. The "C" and "M" stand for Clarke and Mordukhovich, respectively, since they occur when applying the Clarke or Mordukhovich calculus to the MPEC (1). The "A" might stand for "alternative" because that descibes the properties of the Lagrange multipliers, or "Abadie" because it first occured when MPEC-ACQ was applied to the MPEC (1), see [6].

We will now introduce some normal cones, which we will later use to define various subgradients. For more detail on the normal cones we use here, see [3, 11, 8].

Definition 2.2 *Let $\Omega \subseteq \mathbb{R}^l$ be nonempty and closed, and $v \in \Omega$ be given. We call*

$$N^\pi(v, \Omega) := \{ w \in \mathbb{R}^l \mid \exists \mu > 0 : w^T(u - v) \leq \mu \|u - v\|^2 \; \forall u \in \Omega \} \tag{6}$$

the proximal normal cone *to Ω at v,*

$$N(v, \Omega) := \{ \lim_{k \to \infty} w^k \mid \exists \{v^k\} \subset \Omega : \lim_{k \to \infty} v^k = v, \; w^k \in N^\pi(v^k, \Omega) \} \tag{7}$$

the limiting normal cone *to Ω at v, and*

$$N^{\mathrm{Cl}}(v, \Omega) := \mathrm{cl\,conv\,} N(v, \Omega) \tag{8}$$

the Clarke normal cone *to Ω at v. If Ω additionally is convex, we call*

$$N^{\mathrm{conv}}(v, \Omega) := \{ w \in \mathbb{R}^l \mid w^T(u - v) \leq 0 \; \forall u \in \Omega \} \tag{9}$$

the standard normal cone *to* Ω *at* v.

By convention, we set $N^\pi(v, \Omega) = N(v, \Omega) = N^{\mathrm{Cl}}(v, \Omega) = N^{\mathrm{conv}}(v, \Omega) := \emptyset$ *if* $v \notin \Omega$. *By* $N_\Omega^\times : \mathbb{R}^l \rightrightarrows \mathbb{R}^l$ *we denote the multifunction that maps* $v \mapsto N^\times(v, \Omega)$, *where* $\times$ *is a placeholder for one of the normal cones defined above.*

Since the limiting normal cone is the most important one in our subsequent analysis, we have not given it an index, to simplify notation.

Note that the Clarke normal cone defined in (8) is in fact the well-known Clarke normal cone [3]. The limiting normal cone is also called the *Mordukhovich normal cone* [20]. Additionally, it may also be defined using a limit of what Rockafellar and Wets [17] call the *regular normal cone* and others [11, 12] call the *Fréchet normal cone*. Since we do not need it in our analysis, we don't define it here. See in particular [11, Proposition 2.2] in this context.

We will now show that the normal cones defined in Definition 2.2 coincide if Ω is convex.

Proposition 2.3 *Let* $\Omega \subseteq \mathbb{R}^l$ *be nonempty, closed, and convex. Then the normal cones defined in Definition 2.2 coincide for every* $v \in \Omega$, *i.e. we have*

$$N^\pi(v, \Omega) = N(v, \Omega) = N^{\mathrm{Cl}}(v, \Omega) = N^{\mathrm{conv}}(v, \Omega) \tag{10}$$

for all $v \in \Omega$.

Proof. For an arbitrary $v \in \Omega$ it obviously holds that

$$N^{\mathrm{conv}}(v, \Omega) \subseteq N^\pi(v, \Omega) \subseteq N(v, \Omega) \subseteq N^{\mathrm{Cl}}(v, \Omega). \tag{11}$$

By [3, Proposition 2.4.4], it further holds that

$$N^{\mathrm{Cl}}(v, \Omega) = N^{\mathrm{conv}}(v, \Omega)$$

if Ω is convex. Together with (11) this yields the statement of the proposition. $\qquad\square$

Given that all normal cones coincide when Ω is convex, we will, in the following, simply say *normal cone* and denote it by $N(\cdot, \Omega)$ if the set Ω in question is convex.

It will become useful to know the normal cones to some specific sets. We start off with the normal to the nonnegative orthant.

Proposition 2.4 *Let* $\Omega = \mathbb{R}_+^l$ *be the nonnegative orthant in* $\mathbb{R}^l$. *Then the normal cone to* $\mathbb{R}_+^l$ *in* $v \in \mathbb{R}^l$ *is given by*

$$N(v, \mathbb{R}_+^l) = N(v_1, [0, \infty)) \times \cdots \times N(v_l, [0, \infty)) \tag{12}$$

with

$$N(v_i, [0, \infty)) = \begin{cases} \emptyset & : v_i < 0, \\ (-\infty, 0] & : v_i = 0, \\ \{0\} & : v_i > 0. \end{cases} \tag{13}$$

Proof. Observing that $\mathbb{R}_+^l$ is closed and convex, this is given by [17, Theorem 6.9 & Example 6.10]. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

A direct consequence of this proposition is the following lemma, which will play a central role when we try to cope with the complementarity constraints in (1).

Lemma 2.5 *Let $a, b \in \mathbb{R}^l$ be given. Then the following are equivalent:*

(a) $a \geq 0$, $b \geq 0$, $a^T b = 0$;

(b) $0 \in b + N(a, \mathbb{R}_+^l)$;

(c) $(a, -b) \in \operatorname{gph} N_{\mathbb{R}_+^l}$, *where*

$$\operatorname{gph} N_{\mathbb{R}_+^l} := \left\{ (v, u) \,\middle|\, u \in N(v, \mathbb{R}_+^l) \right\}$$

denotes the graph of the multifunction $N_{\mathbb{R}_+^l}$.

The following proposition, concerning the limiting normal cone to $\operatorname{gph} N_{\mathbb{R}_+^l}$, is due to Ye [21, Proposition 3.7].

Proposition 2.6 *For any $(x, y) \in \operatorname{gph} N_{\mathbb{R}_+^l}$, define*

$$\mathcal{I}_x := \{i \mid x_i > 0, y_i = 0\}, \quad \mathcal{I}_y := \{i \mid x_i = 0, y_i < 0\}, \quad \mathcal{I}_0 := \{i \mid x_i = 0, y_i = 0\}.$$

Then

$$N((x, y), \operatorname{gph} N_{\mathbb{R}_+^l}) = \{(a, -b) \in \mathbb{R}^{2l} \mid a_{\mathcal{I}_x} = 0, b_{\mathcal{I}_y} = 0, (a_i < 0 \wedge b_i < 0) \vee a_i b_i = 0 \,\forall i \in \mathcal{I}_0\}.$$

Just as the Clarke subgradient can be defined in terms of the Clarke normal cone [3, Corollary to Theorem 2.4.9], alternate subgradients can be defined using the normal cones introduced in Definition 2.2.

Definition 2.7 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz continuous. Then the* proximal subgradient *of f in $z \in \mathbb{R}^n$ is given by*

$$\partial^\pi f(z) := \{\xi \mid (\xi, -1) \in N^\pi((z, f(z)), \operatorname{epi} f)\}, \tag{14}$$

the limiting subgradient *of f in $z \in \mathbb{R}^n$ by*

$$\partial f(z) := \{\xi \mid (\xi, -1) \in N((z, f(z)), \operatorname{epi} f)\}, \tag{15}$$

and the Clarke subgradient *of f in $z \in \mathbb{R}^n$ by*

$$\partial^{\mathrm{Cl}} f(z) := \{\xi \mid (\xi, -1) \in N^{\mathrm{Cl}}((z, f(z)), \operatorname{epi} f)\}. \tag{16}$$

8

If f is convex, the convex subgradient *of f in $z \in \mathbb{R}^n$ is given by*

$$\partial^{\text{conv}} f(z) := \{\xi \mid (\xi, -1) \in N^{\text{conv}}((z, f(z)), \text{epi} f)\}. \tag{17}$$

Here, epi *denotes the epigraph of a function $f : \mathbb{R}^n \to \mathbb{R}$:*

$$\text{epi} f := \{(z, \zeta) \in \mathbb{R}^{n+1} \mid \zeta \geq f(z)\}.$$

Remark. It immediately follows from (11) and the definition of the various subgradients that the string of inclusions

$$\partial^\pi f(z) \subseteq \partial f(z) \subseteq \partial^{\text{Cl}} f(z) \tag{18}$$

holds for all $f : \mathbb{R}^n \to \mathbb{R}$ locally Lipschitz continuous in $z \in \mathbb{R}^n$. If f is convex (and, therefore, locally Lipschitz) in $z \in \mathbb{R}^n$, it follows from Proposition 2.3 that

$$\partial^\pi f(z) = \partial f(z) = \partial^{\text{Cl}} f(z) = \partial^{\text{conv}} f(z). \tag{19}$$

The following proposition collects some useful properties and calculus rules of the limiting subgradient. The nonnegative scalar multiplication property is easily verified, while the rest may be found, e.g., in their respective order, in Remark 4C.3, Proposition 5A.4, and Theorem 5A.8 in [8], or in the discussion after Definition 2.9, and Corollaries 4.6 and 5.8 in [11].

Proposition 2.8 *(a) Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz continuous in z and $\alpha \geq 0$. Then it holds that*

$$\partial(\alpha f)(z) = \alpha \partial f(z). \tag{20}$$

(b) Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable in z. Then the limiting subgradient reduces to a singleton:

$$\partial f(z) = \{\nabla f(z)\}. \tag{21}$$

(c) Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz continuous in z. Then the following sum rule holds:

$$\partial(f + g)(z) \subseteq \partial f(z) + \partial g(z). \tag{22}$$

(d) Let $F : \mathbb{R}^n \to \mathbb{R}$ and $f : \mathbb{R} \to \mathbb{R}$ be locally Lipschitz continuous in z and $F(z)$, respectively. Then the following chain rule holds for $\Phi : \mathbb{R}^n \to \mathbb{R}$ defined by $\Phi(z) := (f \circ F)(z)$:

$$\partial\Phi(z) \subseteq \{\partial(\eta F)(z) \mid \eta \in \partial f(F(z))\}. \tag{23}$$

For more detail on the calculus of the limiting subgradient, the interested reader is referred to the works of Loewen [8] and Mordukhovich [11]. Also of interest in this context is a comparison to the closely related calculus for the Clarke subgradient, see [3, Section 2.3].

3 M-stationarity

We now state a theorem by Treiman [20, Corollary 4.2] which we will use to prove a Fritz John-type condition in Theorem 3.2. A result similar to Treiman's may be found in [12, Theorem 4.4].

Theorem 3.1 *Let the program*

$$
\begin{aligned}
\min \quad & f(z) \\
s.t. \quad & g(z) \le 0, \quad h(z) = 0, \\
& z \in U
\end{aligned}
\tag{24}
$$

be given, where $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$, and $h : \mathbb{R}^n \to \mathbb{R}^p$ are locally Lipschitz continuous functions and $U \subseteq \mathbb{R}^n$ is closed. Further, let z^* be a local minimizer of (24). Then there exist $r \ge 0$, $\lambda^g \ge 0$, and λ^h, not all zero, such that

$$
0 \in r\partial f(z^*) + \sum_{i=1}^{m} \lambda_i^g \partial g(z^*) + \sum_{i=1}^{p} \partial(\lambda_i^h h_i)(z^*) + N(z^*, U),
$$

$$
g(z^*)^T \lambda^g = 0.
\tag{25}
$$

We will now apply Theorem 3.1 to our problem (1). Note, however, that we use a slightly weaker smoothness assumption on f since this result will later be applied to a specific MPEC whose objective function is locally Lipschitz but not differentiable in general.

Theorem 3.2 *(Fritz John-type M-stationarity condition)*
Let z^* be a local minimizer of the MPEC (1), where the objective function f is locally Lipschitz and all other functions are continuously differentiable. Then there exists $r \ge 0$, $\lambda^g \in \mathbb{R}^m$, $\lambda^h \in \mathbb{R}^p$, $\lambda^G \in \mathbb{R}^l$, $\lambda^H \in \mathbb{R}^l$, not all zero, such that

$$
0 \in r\partial f(z^*) + \sum_{i=1}^{m} \lambda_i^g \nabla g_i(z^*) + \sum_{i=1}^{p} \lambda_i^h \nabla h_i(z^*) - \sum_{i=1}^{l} \left[\lambda_i^G \nabla G_i(z^*) + \lambda_i^H \nabla H_i(z^*) \right],
$$

$$
\begin{array}{llll}
\lambda_\alpha^G & \text{free}, & & \lambda_\gamma^G = 0, \\
& & (\lambda_i^G > 0 \wedge \lambda_i^H > 0) \vee \lambda_i^G \lambda_i^H = 0 \quad \forall i \in \beta & \\
\lambda_\gamma^H & \text{free}, & & \lambda_\alpha^H = 0,
\end{array}
\tag{26}
$$

$$
g(z^*) \le 0, \qquad \lambda^g \ge 0, \qquad g(z^*)^T \lambda^g = 0.
$$

Proof. We will prove this result by applying Theorem 3.1 to our MPEC (1). In order to do so, we must formulate the complementarity constraints as a set constraint. To facilitate the proof, we also introduce slack variables, ξ and η, in the following equivalent reformulation of the MPEC (1):

$$
\begin{aligned}
\min_{z,\xi,\eta} \quad & f(z) \\
s.t. \quad & g(z) \le 0, \quad h(z) = 0, \\
& \Gamma(z,\xi,\eta) := \begin{pmatrix} G(z) - \xi \\ H(z) + \eta \end{pmatrix} = 0 \\
& \xi \ge 0, \quad \eta \le 0, \quad \xi^T \eta = 0.
\end{aligned}
\tag{27}
$$

In view of Lemma 2.5, this in turn is equivalent to

$$
\begin{aligned}
\min_{z,\xi,\eta} \quad & f(z) \\
\text{s.t.} \quad & g(z) \le 0, \quad h(z) = 0, \quad \Gamma(z,\xi,\eta) = 0, \\
& (z,\xi,\eta) \in \mathbb{R}^n \times \operatorname{gph} N_{\mathbb{R}^l_+}.
\end{aligned}
\tag{28}
$$

We now apply Theorem 3.1 to the program (28). This is possible because, in particular, all functions involved are locally Lipschitz continuous and $\mathbb{R}^n \times \operatorname{gph} N_{\mathbb{R}^l_+}$ is obviously closed.

We now set $\xi^* := G(z^*)$ and $\eta^* := -H(z^*)$. Theorem 3.1 then states that there exist $r \ge 0$, $\lambda^g \ge 0$, λ^h, and $\lambda^\Gamma =: (-\lambda^G, -\lambda^H)$, not all zero, such that

$$
\begin{aligned}
0 \in & \begin{pmatrix} r\partial f(z^*) \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \sum_{i=1}^m \lambda_i^g \partial g_i(z^*) \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \sum_{i=1}^p \partial(\lambda_i^h h_i)(z^*) \\ 0 \\ 0 \end{pmatrix} \\
& + \sum_{i=1}^{2l} \partial(\lambda_i^\Gamma \Gamma_i)(z^*, \xi^*, \eta^*) + N\big((z^*, \xi^*, \eta^*), \mathbb{R}^n \times \operatorname{gph} N_{\mathbb{R}^l_+}\big) \\
= & \begin{pmatrix} r\partial f(z^*) \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \sum_{i=1}^m \lambda_i^g \nabla g_i(z^*) \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \sum_{i=1}^p \lambda_i^h \nabla h_i(z^*) \\ 0 \\ 0 \end{pmatrix} \\
& - \begin{pmatrix} \sum_{i=1}^l \big(\lambda_i^G \nabla G_i(z^*) + \lambda_i^H \nabla H_i(z^*)\big) \\ -\lambda^G \\ \lambda^H \end{pmatrix} + \begin{pmatrix} 0 \\ N\big((\xi^*, \eta^*), \operatorname{gph} N_{\mathbb{R}^l_+}\big) \end{pmatrix},
\end{aligned}
\tag{29}
$$

where we used that all functions but f are continuously differentiable, and that $N((a,b), A \times B) = N(a, A) \times N(b, B)$ for A, B closed (see, e.g., [11, (2.2)]).

We now take a closer look at those components pertaining to ξ and η in (29),

$$
(-\lambda^G, \lambda^H) \in N\big((\xi^*, \eta^*), \operatorname{gph} N_{\mathbb{R}^l_+}\big).
\tag{30}
$$

Invoking Proposition 2.6, we obtain the following rules for the components of λ^G and λ^H:

$$
(\lambda_i^G, \lambda_i^H) \in \begin{cases} \{(a,b) \mid a \text{ free}, b = 0\} & : \xi_i^* = 0, \eta_i^* < 0, \\ \{(a,b) \mid (a > 0 \wedge b > 0) \vee ab = 0\} & : \xi_i^* = \eta_i^* = 0, \\ \{(a,b) \mid a = 0, b \text{ free}\} & : \xi_i^* > 0, \eta_i^* = 0. \end{cases}
\tag{31}
$$

Taking into account that $\xi_i^* = G_i(z^*)$, that $\eta_i^* = -H_i(z^*)$, the definitions of α, β, and γ (see (2)), (31), (29), and the final condition in (25) yield the statement of this theorem. $\square$

Note that a version of Theorem 3.2 employing stronger smoothness assumptions appears in [22, Theorem 2.1].

In standard nonlinear programming, *calmness* is often used as a constraint qualification in order to guarantee that $r > 0$ in a Fritz John-setting, yielding a KKT point as a necessary

first order condition. Applying this reasoning to MPECs yields a KKT point (which is equivalent to a strongly stationary point, see [4]) under calmness.

Our current goal, however, is M-stationarity, weaker than strong stationarity. Therefore we will define an MPEC variant of calmness and show M-stationarity to be a necessary first order condition under it. The following definition of MPEC-calmness is inspired by the definition of calmness for OPVICs, as proposed by Ye [21].

Definition 3.3 *The MPEC (1) is said to satisfy* MPEC-calmness *in z^* (or, alternatively, the MPEC is* MPEC-calm *at z^*) if there exist $\varepsilon > 0$ and $\mu > 0$ such that for all $(p, q, r, s) \in B_\varepsilon(0) \subset \mathbb{R}^{m+p+2l}$ and all $z \in B_\varepsilon(z^*)$ with*

$$
\begin{aligned}
g(z) + p \leq 0, \quad &h(z) + q = 0, \\
G(z) + r \geq 0, \quad H(z) + s \geq 0, \quad &(G(z) + r)^T (H(z) + s) = 0
\end{aligned}
\tag{32}
$$

it holds that

$$
f(z^*) \leq f(z) + \mu \, \|(p, q, r, s)\| \,.
\tag{33}
$$

Since all norms are equivalent in a finite-dimensional setting, MPEC-calmness (just as standard calmness) is independent of the norm used.

Note that a straightforward application of OPVIC-calmness [21] yields a slightly degenerated version of the above, where s (or r) is omitted.

Also of interest is that, once more, we treated the factors of the complementarity term independently, not the complementarity term as a whole. This is reminiscent of the way the MPEC-linearized tangent cone $\mathcal{T}^{lin}_{\mathrm{MPEC}}(z^*)$ is defined (see (4)).

Just as in standard nonlinear programming, MPEC-calmness is closely linked to exact penalization (see, e.g., [1]), which we will examine in the following proposition.

Proposition 3.4 *Let z^* be a local minimizer of the MPEC (1). Then the following are equivalent:*

(a) *the MPEC is MPEC-calm at z^*;*

(b) *there exists a $\rho_0 > 0$ such that for all $\rho \geq \rho_0$ the vector $(z^*, 0, 0) \in \mathbb{R}^{n+2l}$ is a local minimizer of*

$$
\begin{aligned}
\min_{(z,r,s)} \quad & f(z) + \rho(\|\max\{0, g(z)\}\| + \|h(z)\| + \|(r, s)\|) \\
s.t. \quad & G(z) + r \geq 0, \quad H(z) + s \geq 0, \\
& (G(z) + r)^T (H(z) + s) = 0.
\end{aligned}
\tag{34}
$$

Proof. (a) $\Rightarrow$ (b). Then, for some $\mu > 0$, condition (33) is satisfied for all (z, p, q, r, s) in a neighborhood of $(z^*, 0, 0, 0, 0)$ satisfying (32). Consequently, condition (33) also holds for (z, r, s) in a neighborhood of $(z^*, 0, 0)$ satisfying the second line of (32) (which is identical to

the constraints of (34)), and for the specific choice $p_i := -\max\{0, g_i(z)\}$ for $i = 1, \ldots, m$, and $q := -h(z)$ (feasible for (32)), we obtain

$$
\begin{aligned}
f(z^*) &\leq f(z) + \mu \, \|(p, q, r, s)\| \\
&\leq f(z) + \mu \big(\|p\| + \|q\| + \|(r, s)\| \big) \\
&= f(z) + \mu(\|\max\{0, g(z)\}\| + \|h(z)\| + \|(r, s)\|) \\
&\leq f(z) + \rho(\|\max\{0, g(z)\}\| + \|h(z)\| + \|(r, s)\|),
\end{aligned}
\tag{35}
$$

for all $\rho \geq \mu > 0$. Together with the arguments above, this implies that $(z^*, 0, 0)$ is a local minimizer of the program (34) for all $\rho \geq \rho_0 := \mu$.

(b) $\Rightarrow$ (a). The following arguments apply to the Euclidian norm $\|\cdot\| = \|\cdot\|_2$. The extension to an arbitrary norm is trivial since all norms are equivalent in finite dimensions. Only $\tilde{\mu}$ and μ below would be adjusted.

Let ε be given such that $(z^*, 0, 0)$ is a global minimizer of (34) with $\rho = \rho_0$ in the ball $B_{2\varepsilon}(z^*, 0, 0)$ (note the radius is 2ε).

Now, let (z, p, q, r, s) be chosen arbitrarily such that (z, p, q, r, s) satisfies (32), $z \in B_\varepsilon(z^*)$, and $(p, q, r, s) \in B_\varepsilon(0)$. In this case, the vector (z, r, s) is, in particular, feasible for (34) and is within the region where $(z^*, 0, 0)$ is a global minimizer. Therefore, it holds for $\rho = \rho_0$ that

$$
\begin{aligned}
f(z^*) &\leq f(z) + \rho_0(\|\max\{0, g(z)\}\| + \|h(z)\| + \|(r, s)\|) \\
&= f(z) + \rho_0(\|\max\{0, g(z) + p - p\}\| + \|h(z) + q - q\| + \|(r, s)\|) \\
&\leq f(z) + \tilde{\mu}(\|\max\{0, g(z) + p - p\}\|_1 + \|h(z) + q - q\|_1 + \|(r, s)\|_1) \\
&\leq f(z) + \tilde{\mu}(\|\max\{0, \underbrace{g(z) + p}_{\leq 0}\}\|_1 + \|p\| + \|\underbrace{h(z) + q}_{=0}\|_1 + \|q\|_1 + \|(r, s)\|_1 \\
&= f(z) + \tilde{\mu}(\|p\|_1 + \|q\|_1 + \|(r, s)\|_1) \\
&= f(z) + \tilde{\mu} \, \|(p, q, r, s)\|_1 \\
&\leq f(z) + \mu \, \|(p, q, r, s)\|,
\end{aligned}
$$

for some $\tilde{\mu}, \mu > 0$, due to the equivalence of norms. This is exactly the definition of MPEC-calmness. $\qquad \square$

We can now use this proposition to prove that M-stationarity is a necessary first order condition under MPEC-calmness.

Theorem 3.5 *Let z^* be a local minimizer of the MPEC (1) at which the MPEC is MPEC-calm. Then there exists $\lambda^g \in \mathbb{R}^m$, $\lambda^h \in \mathbb{R}^p$, $\lambda^G \in \mathbb{R}^l$, $\lambda^H \in \mathbb{R}^l$ such that*

$$
0 = \nabla f(z^*) + \sum_{i=1}^{m} \lambda_i^g \nabla g_i(z^*) + \sum_{i=1}^{p} \lambda_i^h \nabla h_i(z^*) - \sum_{i=1}^{l} \left[\lambda_i^G \nabla G_i(z^*) + \lambda_i^H \nabla H_i(z^*) \right],
$$

$$
\begin{array}{llll}
\lambda_\alpha^G \;\; \text{free}, & & & \lambda_\gamma^G = 0, \\
\lambda_\gamma^H \;\; \text{free}, & (\lambda_i^G > 0 \wedge \lambda_i^H > 0) \vee \lambda_i^G \lambda_i^H = 0 \quad \forall i \in \beta & & \lambda_\alpha^H = 0, \\
\end{array}
\tag{36}
$$

$$
g(z^*) \leq 0, \qquad \lambda^g \geq 0, \qquad g(z^*)^T \lambda^g = 0,
$$

13

i.e. z^* *is M-stationary.*

Proof. The idea and proof of this theorem is based on the same principles as [21, Theorem 3.6].

Since the MPEC (1) satisfies MPEC-calmness at z^*, $(z^*, 0, 0)$ is a local minimizer of (34) according to Proposition 3.4 for $\rho = \rho_0$.

Applying Theorem 3.2 to (34) yields the existence of $(r, \lambda^G, \lambda^H) \neq (0, 0, 0)$ such that $r \geq 0$, λ^G and λ^H satisfy the sign conditions from (26), and

$$
0 \in r \partial \tilde{f}(z^*, 0, 0) - \sum_{i=1}^{l} \Big[\lambda_i^G \begin{pmatrix} \nabla G_i(z^*) \\ e_i \\ 0 \end{pmatrix} + \lambda_i^H \begin{pmatrix} \nabla H_i(z^*) \\ 0 \\ e_i \end{pmatrix} \Big], \tag{37}
$$

where $\tilde{f}(z, r, s) := f(z) + \rho_0(\|\max\{0, g(z)\}\|_1 + \|h(z)\|_1 + \|(r, s)\|_1)$, and $e_i \in \mathbb{R}^l$ is the i-th unit vector. Note that we choose the 1-norm specifically. Also note that Theorem 3.2 may be applied since $\tilde{f}$ is locally Lipschitz continuous.

If we assume $r = 0$, it follows immediately from (37) that $\lambda^G = \lambda^H = 0$, a contradiction to $(r, \lambda^G, \lambda^H) \neq (0, 0, 0)$. Therefore, $r = 1$ can be assumed without loss of generality.

Now, let us consider $\partial_z \tilde{f}(z^*, 0, 0)$. Using the properties for limiting subgradients from Proposition 2.8 (note that $\rho_0 \geq 0$), we get

$$
\partial_z \tilde{f}(z^*, 0, 0) \subseteq \nabla f(z^*) + \rho_0 \Big[\sum_{i=1}^{m} \partial \max\{0, g_i(z^*)\} + \sum_{i=1}^{p} \partial |h_i(z^*)| \Big]
$$

$$
\subseteq \nabla f(z^*) + \rho_0 \Big[\sum_{i=1}^{m} \{\eta_i^g \nabla g(z^*) \mid \eta_i^g \in \partial \max\{0, g_i(z^*)\}\} \tag{38}
$$

$$
+ \sum_{i=1}^{p} \{\eta_i^h \nabla h(z^*) \mid \eta_i^h \in \partial |h_i(z^*)|\} \Big].
$$

Both $\max\{0, \cdot\}$ and $|\cdot|$ are convex functions, and in view of (19), their limiting subgradients are equal to their convex subgradients, which are well known. Since z^* is, in particular, feasible, it follows that $h(z^*) = 0$ and $g(z^*) \leq 0$. Hence we state the limiting subgradients of $\max\{0, \cdot\}$ and $|\cdot|$ only for the relevant part of their domain:

$$
\partial \max\{0, x\} = \begin{cases} 0 & : \ x < 0, \\ [0, 1] & : \ x = 0, \end{cases} \qquad \partial |x| = [-1, 1] \ \text{ for } x{=}0.
$$

Incorporating this into (38) and substituting $\partial_z \tilde{f}(z^*, 0, 0)$ into (37) with $r = 1$ yields the first part of (36).

Setting $\lambda_i^g := 0$ for all $i \notin \mathcal{I}_g$ finally yields the last line of (36), thereby completing the proof. $\qquad \square$

Before we continue, we will introduce a set which will facilitate the statements and proofs of the following results. Taking (32) from the definition of MPEC-calmness, we collect all those z satisfying a specific choice (p, q, r, s):

$$
\begin{aligned}
\mathcal{Z}(p, q, r, s) := \{ z \in \mathbb{R}^n \mid\ & g(z) + p \leq 0,\ h(z) + q = 0, \\
& G(z) + r \geq 0,\ H(z) + s \geq 0, \\
& (G(z) + r)^T (H(z) + s) = 0 \}.
\end{aligned}
\tag{39}
$$

Note that $\mathcal{Z} = \mathcal{Z}(0, 0, 0, 0)$ is the feasible set of the MPEC (1).

We now define an MPEC-variant of a local error bound, which we will show to imply MPEC-calmness. Again, the following definition is inspired by Ye [21, Definition 4.1].

Definition 3.6 *The system*

$$
\begin{aligned}
g(z) \leq 0, \quad & h(z) = 0, \\
G(z) \geq 0, \quad & H(z) \geq 0, \quad G(z)^T H(z) = 0
\end{aligned}
\tag{40}
$$

is said to have a local MPEC-error bound *at z^* if there exist $\nu, \delta > 0$ such that*

$$
\operatorname{dist}(z, \mathcal{Z}) \leq \nu \, \|(p, q, r, s)\|
\tag{41}
$$

for all $(p, q, r, s) \in B_\delta(0)$ and all $z \in \mathcal{Z}(p, q, r, s) \cap B_\delta(z^*)$.

Note that there exists a distinct difference to a standard local error bound. A standard local error bound would allow *all $z \in B_\delta(z^*)$*, and we could hence replace $\|(p, q, r, s)\|$ with something like $\|(\max\{0, g(z)\}, h(z), \min\{G(z), H(z)\})\|$.

Note that the system (40) is the constraint system of the MPEC (1), i.e. that it defines $\mathcal{Z}$.

We now show that a local MPEC-error bound implies MPEC-calmness. This is also inspired by an analogous result for OPVICs by Ye [21, Proposition 4.2].

Proposition 3.7 *Let z^* be a local minimizer of the MPEC (1), and let the constraint system (40) of the MPEC have a local MPEC-error bound. Then the MPEC is MPEC-calm at z^*.*

Proof. Let $\tilde{\delta}$ be chosen such that z^* is a global minimizer of f in $\mathcal{Z} \cap B_{2\tilde{\delta}}(z^*)$. Furthermore, let δ be the quantity from the definition of the local MPEC-error bound (see Definition 3.6). Setting $\varepsilon := \min\{\tilde{\delta}, \delta\}$, we have that $(p, q, r, s) \in B_\varepsilon(0)$ and $z \in \mathcal{Z}(p, q, r, s) \cap B_\varepsilon(z^*)$ satisfy (41). By our choice of z and the fact that $z^* \in \mathcal{Z}$ is feasible, it holds that

$$
\|\Pi_{\mathcal{Z}}(z) - z^*\| \leq \|\Pi_{\mathcal{Z}}(z) - z\| + \|z - z^*\| \leq 2\varepsilon.
$$

Using this, and the facts that $2\varepsilon \leq 2\tilde{\delta}$ and f is continuously differentiable and hence locally Lipschitz continuous, it follows that

$$
\begin{aligned}
f(z^*) &\leq f(\Pi_{\mathcal{Z}}(z)) \\
&= f(z) + (f(\Pi_{\mathcal{Z}}(z)) - f(z)) \\
&\leq f(z) + L\,\|\Pi_{\mathcal{Z}}(z) - z\| \\
&= f(z) + L\mathrm{dist}(z, \mathcal{Z}) \\
&\overset{(41)}{\leq} f(z) + L\nu\,\|(p, q, r, s)\|\,.
\end{aligned}
\tag{42}
$$

Setting $\mu := L\nu$ yields the conditions for MPEC-calmness at z^*. This concludes the proof. $\square$

We next consider mathematical programs with equilibrium constraints where the constraints are characterized by affine functions, i.e. we consider the following *mathematical program with affine equilibrium constraints*, or *MPAEC*

$$
\begin{aligned}
\min \quad & f(z) \\
\text{s.t.} \quad & Az + a \leq 0, \quad Bz + b = 0, \\
& Cz + c \geq 0, \quad Dz + d \geq 0, \quad (Cz+c)^T(Dz+d) = 0,
\end{aligned}
\tag{43}
$$

where $f : \mathbb{R}^n \to \mathbb{R}$, not necessarily affine, $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{p \times n}$, $C, D \in \mathbb{R}^{l \times n}$ are given matrices, and $a \in \mathbb{R}^m$, $b \in \mathbb{R}^p$, $c, d \in \mathbb{R}^l$ are given vectors.

The following result states that an MPAEC has a local MPEC-error bound. This, though important by itself, will be instrumental when we prove M-stationarity to be a first order optimality condition under MPEC-ACQ in Theorem 3.9.

Proposition 3.8 *Let z^* be feasible for the MPAEC (43). Then the constraint system of the MPAEC has a local MPEC-error bound at z^*.*

A proof of this result may be obtained by following the techniques of the proof for [10, Theorem 2.3]. Alternatively, Proposition 3.8 may be derived using a result by Robinson [16], where polyhedral multifunctions are shown to be locally upper Lipschitz. See [16, 10, 21] for proofs using this approach.

We are finally able to present our main result, whose proof is entirely due to Ye [22, Theorem 3.1].

Theorem 3.9 *Let z^* be a local minimizer of the MPEC (1) at which MPEC-ACQ holds. Then there exists a Lagrange multiplier λ^* such that (z^*, λ^*) satisfies the conditions for M-stationarity (5).*

Proof. Since z^* is a local minimum of (1), B-stationarity holds, i.e.

$$
\nabla f(z^*)^T d \geq 0 \qquad \forall d \in \mathcal{T}(z^*),
$$

16

and since MPEC-ACQ holds, this is equivalent to

$$\nabla f(z^*)^T d \geq 0 \qquad \forall d \in \mathcal{T}_{\text{MPEC}}^{lin}(z^*). \tag{44}$$

This, in turn, is equivalent to $d^* = 0$ being a minimum of

$$\begin{aligned}
\min_{d} \quad & \nabla f(z^*)^T d \\
\text{s.t.} \quad & d \in \mathcal{T}_{\text{MPEC}}^{lin}(z^*),
\end{aligned} \tag{45}$$

which is an MPAEC. By Proposition 3.8, the constraint system of (45) has a local MPEC-error bound, which in turn implies that (45) is MPEC-calm by Proposition 3.7.

Keeping in mind that in (45), d is the variable, and that $d^* = 0$ solves (45), Theorem 3.5 then yields the existence of Lagrange multipliers $\lambda_{\mathcal{I}_g}^g$, λ^h, $\lambda_{\alpha \cup \beta}^G$, and $\lambda_{\gamma \cup \beta}^H$ such that, in particular

$$\begin{aligned}
0 = \nabla f(z^*) + \sum_{i \in \mathcal{I}_g} \lambda_i^g \nabla g_i(z^*) + \sum_{i=1}^p \lambda_i^h \nabla h_i(z^*) - \sum_{i \in \alpha} \lambda_i^G \nabla G_i(z^*) - \sum_{i \in \gamma} \lambda_i^H \nabla H_i(z^*) \\
- \sum_{i \in \beta} \left[\lambda_i^G \nabla G_i(z^*) + \lambda_i^H \nabla H_i(z^*) \right],
\end{aligned} \tag{46}$$

$$\lambda_\alpha^G \text{ free}, \qquad \lambda_\gamma^H \text{ free}, \qquad \lambda^h \text{ free}, \qquad \lambda_{\mathcal{I}_g}^g \geq 0,$$
$$(\lambda_i^G > 0 \wedge \lambda_i^H > 0) \vee \lambda_i^G \lambda_i^H = 0 \quad \forall i \in \beta.$$

The latter condition follows from the fact that since $d^* = 0$, it holds that $\nabla G_i(z^*)^T d^* = \nabla H_i(z^*)^T d^* = 0$ for all $i \in \beta$, and hence the whole set β is degenerate in the corresponding constraints of (45). Also note that, since λ_α^G and λ_γ^H have no sign restriction imposed upon them, their signs were chosen in such a manner as to facilitate the notation of the proof.

By setting $\lambda_i^g := 0$ for all $i \notin \mathcal{I}_g$, $\lambda_\gamma^G := 0$, and $\lambda_\alpha^H := 0$ we obtain the conditions (5) for M-stationarity with $\lambda^* = (\lambda^g, \lambda^h, \lambda^G, \lambda^H)$ from (46). This concludes the proof. $\qquad \square$

Note that the MPEC-Abadie constraint qualification is satisfied under many other conditions like the MPEC-MFCQ assumption or an MPEC-variant of a Slater-condition, see [6], as well as a number of other constraint qualifications, see [22]. Hence all these stronger constraint qualifications imply that M-stationarity is a necessary first order optimality condition. In particular, a local minimizer is an M-stationary point under the MPEC-MFCQ assumption used in [18]. However, the authors of [18] were only able to prove C-stationarity to be a necessary first order condtion under MPEC-MFCQ.

We also note that the MPEC-Abadie constraint qualification does not guarantee that a local minimizer is a strongly stationary point. This follows from the observation that even the stronger MPEC-MFCQ condition does not imply strong stationarity, see [18] for a counterexample.

4 Conclusion

We proved that a very weak assumption, the MPEC-Abadie constraint qualification, implies that a local minimum satisfies the relatively strong first order optimality condition, M-stationarity, in the framework of mathematical programs with equilibrium constraints. This result was established before in a very recent paper by Ye [22]. However, the proof given in [22] is somewhat incomplete. We therefore presented a complete and largely self-contained proof of this result.

References

[1] J. V. BURKE, *Calmness and exact penalization*, SIAM Journal on Control and Optimization, 29 (1991), pp. 493–497.

[2] Y. CHEN AND M. FLORIAN, *The nonlinear bilevel programming problem: Formulations, regularity and optimality conditions*, Optimization, 32 (1995), pp. 193–209.

[3] F. H. CLARKE, *Optimization and Nonsmooth Analysis*, John Wiley & Sons, New York, 1983.

[4] M. L. FLEGEL AND C. KANZOW, *On the Guignard constraint qualification for mathematical programs with equilibrium constraints*. Institute of Applied Mathematics and Statistics, University of Würzburg, Preprint 248, October 2002.

[5] ——, *A Fritz John approach to first order optimality conditions for mathematical programs with equilibrium constraints*, Optimization, 52 (2003), pp. 277–286.

[6] ——, *An Abadie-type constraint qualification for mathematical programs with equilibrium constraints*, Journal of Optimization Theory and Applications, to appear.

[7] X. HU AND D. RALPH, *Convergence of a penalty method for mathematical programming with complementarity constraints*, tech. rep., Judge Institute of Management Studies, University of Cambridge, England, October 2000.

[8] P. D. LOEWEN, *Optimal Control via Nonsmooth Analysis*, vol. 2 of CRM Proceedings & Ledcture Notes, American Mathematical Society, Providence, RI, 1993.

[9] Z.-Q. LUO, J.-S. PANG, AND D. RALPH, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, UK, 1996.

[10] Z.-Q. LUO AND P. TSENG, *Error bound and convergence analysis of matrix splitting algorithms for the affine variational inequality problem*, SIAM Journal on Optimization, 2 (1992), pp. 43–54.

[11] B. S. MORDUKHOVICH, *Generalized differential calculus for nonsmooth and set-valued mappings*, Journal of Mathematical Analysis and Applications, (1994), pp. 250–288.

[12] ——, *Necessary conditions in nonsmooth minimization via lower and upper subgradients*, Set-Valued Analalysis, to appear.

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

[14] ——, *A generalized mathematical program with equilibrium constraints*, SIAM Jounral of Control and Optimization, 38 (2000), pp. 1623–1638.

[15] J.-S. Pang and M. Fukushima, *Complementarity constraint qualifications and simplified B-stationarity conditions for mathematical programs with equilibrium constraints*, Computational Optimization and Applications, 13 (1999), pp. 111–136.

[16] S. M. Robinson, *Some continuity properties of polyhedral multifunctions*, Mathematical Programming Study, 14 (1981), pp. 206–214.

[17] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, vol. 317 of A Series of Comprehensive Studies in Mathematics, Springer, Berlin, Heidelberg, 1998.

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

[19] S. Scholtes, *Convergence properties of a regularization scheme for mathematical programs with complementarity constraints*, SIAM Journal on Optimization, 11 (2001), pp. 918–936.

[20] J. S. Treiman, *Lagrange multipliers for nonconvex generalized gradients with equality, inequality, and set constraints*, SIAM Journal on Control and Optimization, 37 (1999), pp. 1313–1329.

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

[22] ——, *Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints.* Preprint, October 2003.

[23] ——, *Personal communication.* By email, October 2003.