A DIRECT PROOF FOR M-STATIONARITY UNDER MPEC-ACQ FOR MATHEMATICAL PROGRAMS WITH EQUILIBRIUM CONSTRAINTS

Michael L. Flegel and Christian Kanzow

Preprint 254 May 2004

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

May 4, 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. Improving on her approach, we present a different, much more direct and shorter, approach.

Key Words. Mathematical programs with equilibrium constraints, M-stationarity, Abadie constraint qualification.

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:

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

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 [14, 3]. A still weaker CQ, the Guignard CQ, has a chance of holding, see [3]. 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., [10, 17, 14, 12, 13, 6, 21].

One of the stronger stationarity concepts introduced is M-stationarity [12] (see (5)). It is second only to strong stationarity. Weaker stationarity concepts like A- and C-stationarity have also been introduced [4, 17], 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 [18] 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 [8] 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]. This result has previously been established by Ye [21].

However, there has been some controversy about the correctness of her proof. It relies on [20, Theorem 3.2], the proof of which is erroneous, as the author agrees [22]. It is possible to fix this gap in her proof by using, for example, a result by Treiman [19]. We do not take this route in this paper. The interested reader is referred to [5] for an exhaustive discussion of this approach. Instead, we present here a very direct and short proof.

The organization of this paper is as follows: In Section 2 we introduce some concepts and results necessary for proving our main result. This is done in Section 3, referring to Section 2 and introducing additional concepts as needed.

A word on notation. Given two vectors x and y, we use $(x, y) := (x^T, y^T)^T$ for ease of notation. Comparisons such as $\le$ and $\ge$ are understood componentwise. Given a vector

$a \in \mathbb{R}^n$, a_i denotes the i-th component of that vector. Given a set $\nu \subseteq \{1, \dots, n\}$ we denote by $x_\nu \in \mathbb{R}^{|\nu|}$ that vector which consists of those components of $x \in \mathbb{R}^n$ which correspond to the indices in ν. Furthermore, we denote the set of all partitions of ν by $\mathcal{P}(\nu) := \{(\nu_1, \nu_2) \mid \nu_1 \cup \nu_2 = \nu,\ \nu_1 \cap \nu_2 = \emptyset\}$. By $\mathbb{R}^l_+ := \{x \in \mathbb{R}^l \mid x \geq 0\}$ we mean the nonnegative orthant of $\mathbb{R}^l$. Finally, the graph of a multifunction (set-valued function) $\Phi : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is defined as $\operatorname{gph} \Phi := \{(v, w) \in \mathbb{R}^{m+n} \mid w \in \Phi(v)\}$.

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.

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

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

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 [3]). Instead, the authors proposed the *MPEC-linearized* tangent cone in [6],

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

where $\mathcal{I}_g := \{i \mid g_i(z^*) = 0\}$ is the set of active inequality constraints at z^*. 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

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

(see [6]). 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*

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

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 [17, 14, 4] for more detail.

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

$$\begin{aligned}
&0 = \nabla f(z^*) + \sum_{i=1}^{m} \lambda_i^g \nabla g_i(z^*) + \sum_{i=1}^{p} \lambda_i^h \nabla h_i(z^*) - \sum_{i=1}^{l} \left[\lambda_i^G \nabla G_i(z^*) + \lambda_i^H \nabla H_i(z^*)\right],\\
&\begin{array}{lll} \lambda_\alpha^G \ \text{free}, & & \lambda_\gamma^G = 0, \\ & (\lambda_i^G > 0 \wedge \lambda_i^H > 0) \ \vee \ \lambda_i^G \lambda_i^H = 0 \quad \forall i \in \beta & \\ \lambda_\gamma^H \ \text{free}, & & \lambda_\alpha^H = 0, \end{array}\\
&g(z^*) \leq 0, \qquad \lambda^g \geq 0, \qquad g(z^*)^T \lambda^g = 0.
\end{aligned} \tag{5}$$

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 [17, 14]: $\lambda_i^G \geq 0 \wedge \lambda_i^H \geq 0 \ \forall i \in \beta$;
- M-stationarity [12]: $(\lambda_i^G > 0 \wedge \lambda_i^H > 0) \ \vee \ \lambda_i^G \lambda_i^H = 0 \ \forall i \in \beta$;
- C-stationarity [17]: $\lambda_i^G \lambda_i^H \geq 0 \ \forall i \in \beta$;
- A-stationarity [4]: $\lambda_i^G \geq 0 \vee \lambda_i^H \geq 0 \ \forall i \in \beta$.

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 [14] 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 will become important in our subsequent analysis. For more detail on the normal cones we use here, see [11, 9, 16].

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

$$\hat{N}(v, \Omega) := \{w \in \mathbb{R}^l \mid \limsup_{\substack{v^k \to v \\ \{v^k\} \subset \Omega \setminus \{v\}}} w^T(v^k - v) / \left\| v^k - v \right\| \leq 0\} \tag{6}$$

the Fréchet normal cone *or* regular normal cone *[16] to Ω at v,*

$$N^\pi(v, \Omega) := \{w \in \mathbb{R}^l \mid \exists \mu > 0 : w^T(u - v) \leq \mu \left\| u - v \right\|^2 \;\; \forall u \in \Omega\} \tag{7}$$

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

$$N(v, \Omega) := \{\lim_{k \to \infty} w^k \mid \exists \{v^k\} \subset \Omega : \lim_{k \to \infty} v^k = v, \; w^k \in N^\pi(v^k, \Omega)\} \tag{8}$$

the limiting normal cone *to Ω at v.*

By convention, we set $\hat{N}(v, \Omega) = N^\pi(v, \Omega) = N(v, \Omega) := \emptyset$ 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.

Note that, in the definition of the limiting normal cone $N(v, \Omega)$, $N^\pi(v^k, \Omega)$ may be replaced by $\hat{N}(v^k, \Omega)$, see [11, Proposition 2.2]. Since this is particularly important, we state this in its own proposition.

Proposition 2.3 *Let $\Omega \subseteq \mathbb{R}^l$ be nonempty and closed. Then it holds that*

$$\begin{aligned} N(v, \Omega) &= \{\lim_{k \to \infty} w^k \mid \exists \{v^k\} \subset \Omega : \lim_{k \to \infty} v^k = v, \; w^k \in N^\pi(v^k, \Omega)\} \\ &= \{\lim_{k \to \infty} w^k \mid \exists \{v^k\} \subset \Omega : \lim_{k \to \infty} v^k = v, \; w^k \in \hat{N}(v^k, \Omega)\}. \end{aligned} \tag{9}$$

In particular, it holds that $\hat{N}(v, \Omega) \subseteq N(v, \Omega)$.

Note also that if v is in the interior of Ω, any of the above normal cones reduces to $\{0\}$, as is well known.

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

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

$$N(v, \mathbb{R}^l_+) = N(v_1, [0, \infty)) \times \cdots \times N(v_l, [0, \infty)) \tag{10}$$

with

$$N(v_i, [0, \infty)) = \begin{cases} \emptyset & : \ v_i < 0, \\ (-\infty, 0] & : \ v_i = 0, \\ \{0\} & : \ v_i > 0. \end{cases} \tag{11}$$

Proof. Observing that $\mathbb{R}^l_+$ is closed and convex, this is given by [16, Theorem 6.9 & Example 6.10]. □

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) $(a, -b) \in \operatorname{gph} N_{\mathbb{R}^l_+}$.

The following proposition, concerning the limiting normal cone to $\operatorname{gph} N_{\mathbb{R}^l_+}$, is due to Outrata [12, Lemma 2.2], see also [20, 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

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

Another important set is a polyhedral convex set, whose limiting normal cone at the origin will be needed in our subsequent analysis and is therefore given in the following lemma.

Lemma 2.7 *Let the convex set*

$$\begin{aligned} \mathcal{D} := \{d \in \mathbb{R}^n \mid \ & a_i^T d \leq 0, \quad \forall i = 1, \ldots, k, \\ & b_j^T d = 0, \quad \forall j = 1, \ldots, l\} \end{aligned} \tag{12}$$

be given. Then the limiting normal cone of $\mathcal{D}$ at 0 is given by

$$\begin{aligned} N(0, \mathcal{D}) = \{v \in \mathbb{R}^n \mid \ & v = \sum_{i=1}^{k} \alpha_i a_i + \sum_{j=1}^{l} \beta_j b_j \\ & \alpha_i \geq 0, \quad \forall i = 1, \ldots, k\}, \end{aligned} \tag{13}$$

Proof. Since $\mathcal{D}$ is convex, [16, Theorem 6.9] may be invoked and the statement of this lemma is given by Theorem 3.2.2 and its proof in [1]. $\square$

3 M-Stationarity

We start off this section by stating our main result. The remainder of the paper is dedicated to proving this result.

Theorem 3.1 *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).*

The fundamental idea of the proof is due to Ye, see [21, Theorem 3.1]. It is based on the fact that, under MPEC-ACQ, the tangent cone is described by some linear equations and inequalities, and a linear complementarity problem. Using the tangent cone as the feasible set of a mathematical program yields an MPEC. We are then able to glean the conditions for M-stationarity for the original MPEC (1) from this "affine" MPEC. In the following, we make this idea more precise.

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

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

and since MPEC-ACQ holds, this is equivalent to

$$\nabla f(z^*)^T d \geq 0 \qquad \forall d \in \mathcal{T}^{lin}_{\text{MPEC}}(z^*). \tag{14}$$

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

$$\begin{array}{ll} \min\limits_{d} & \nabla f(z^*)^T d \\ \text{s.t.} & d \in \mathcal{T}^{lin}_{\text{MPEC}}(z^*). \end{array} \tag{15}$$

This is a degenerated form of a *mathematical program with affine equilibrium constraints*, or MPAEC. It easily verified that $d^* = 0$ being a minimizer of (15) is equivalent to $(d^*, \xi^*, \eta^*) = (0, 0, 0)$ being a minimizer of

$$\begin{array}{ll} \min\limits_{(d,\xi,\eta)} & \nabla f(z^*)^T d \\ \text{s.t.} & (d, \xi, \eta) \in \mathcal{D} := \mathcal{D}_1 \cap \mathcal{D}_2 \end{array} \tag{16}$$

with

$$\begin{array}{lll} \mathcal{D}_1 := \{(d, \xi, \eta) \mid & \nabla g_i(z^*)^T d \leq 0, & \forall i \in \mathcal{I}_g, \\ & \nabla h_i(z^*)^T d = 0, & \forall i = 1, \dots, p, \\ & \nabla G_i(z^*)^T d = 0, & \forall i \in \alpha, \\ & \nabla H_i(z^*)^T d = 0, & \forall i \in \gamma, \\ & \nabla G_i(z^*)^T d - \xi_i = 0, & \forall i \in \beta, \\ & \nabla H_i(z^*)^T d + \eta_i = 0, & \forall i \in \beta\,\} \end{array} \tag{17}$$

and

$$\mathcal{D}_2 := \{(d, \xi, \eta) \mid \xi \geq 0,\ \eta \leq 0,\ \xi^T \eta = 0\}. \tag{18}$$

Once more, since $(0,0,0)$ is a minimizer of (16), B-stationarity holds, which in this case means that

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

where $\mathcal{T}((0,0,0), \mathcal{D})$ denotes the tangent cone to the set $\mathcal{D}$ in the point $(0,0,0)$. By virtue of [16, Proposition 6.5], this is equivalent to

$$(-\nabla f(z^*), 0, 0) \in \hat{N}((0,0,0), \mathcal{D}) \subseteq N((0,0,0), \mathcal{D}). \tag{19}$$

Note, once again, that the limiting normal cone $N(\cdot,\cdot)$ is equal to the limit of the Fréchet normal cone $\hat{N}(\cdot,\cdot)$ (see Proposition 2.3).

In order to calculate $N((0,0,0), \mathcal{D})$ in a fashion conducive to our goal, we need to consider the normal cones $\mathcal{D}_1$ and $\mathcal{D}_2$ separately. To be able to do this, we need some auxiliary results. We start off with the definition of a polyhedral multifunction (see [15]).

Definition 3.2 *We say that a multifunction $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is a* polyhedral multifunction *if its graph is the union of finitely many polyhedral convex sets.*

We now show that a certain multifunction, which is defined using $\mathcal{D}_1$ and $\mathcal{D}_2$, is a polyhedral multifunction. We will need this to apply a result by Henrion, Jourani and Outrata [7].

Lemma 3.3 *Let the multifunction $\Phi : \mathbb{R}^{n+2|\beta|} \rightrightarrows \mathbb{R}^{n+2|\beta|}$ be given by*

$$\Phi(v) := \{w \in \mathcal{D}_1 \mid v + w \in \mathcal{D}_2\}. \tag{20}$$

Then Φ is a polyhedral multifunction.

Proof. Since the graph of Φ may be expressed as

$$\begin{aligned}
\operatorname{gph}\Phi = \{(d^v, \xi^v, \eta^v, d^w, \xi^w, \eta^w) \in \mathbb{R}^{2(n+2|\beta|)} \mid \\
&\nabla g_i(z^*)^T d^w \leq 0, && \forall i \in \mathcal{I}_g, \\
&\nabla h_i(z^*)^T d^w = 0, && \forall i = 1, \dots, p, \\
&\nabla G_i(z^*)^T d^w = 0, && \forall i \in \alpha, \\
&\nabla H_i(z^*)^T d^w = 0, && \forall i \in \gamma, \\
&\nabla G_i(z^*)^T d^w - \xi_i^w = 0, && \forall i \in \beta, \\
&\nabla H_i(z^*)^T d^w + \eta_i^w = 0, && \forall i \in \beta, \\
&\xi^v + \xi^w \geq 0,\ \eta^v + \eta^w \leq 0,\ (\xi^v + \xi^w)^T (\eta^v + \eta^w) = 0\,\}
\end{aligned}$$

$$
\begin{aligned}
= \bigcup_{(\nu_1,\nu_2)\in\mathcal{P}(\{1,\dots,|\beta|\})} \{(d^v,\xi^v,\eta^v,d^w,\xi^w,\eta^w) \in \mathbb{R}^{2(n+2|\beta|)} \mid \\
&\nabla g_i(z^*)^T d^w \le 0, && \forall i \in \mathcal{I}_g, \\
&\nabla h_i(z^*)^T d^w = 0, && \forall i = 1,\dots,p, \\
&\nabla G_i(z^*)^T d^w = 0, && \forall i \in \alpha, \\
&\nabla H_i(z^*)^T d^w = 0, && \forall i \in \gamma, \\
&\nabla G_i(z^*)^T d^w - \xi_i^w = 0, && \forall i \in \beta, \\
&\nabla H_i(z^*)^T d^w + \eta_i^w = 0, && \forall i \in \beta, \\
&\xi_{\nu_1}^v + \xi_{\nu_1}^w = 0, \quad \xi_{\nu_2}^v + \xi_{\nu_2}^w \ge 0, \\
&\eta_{\nu_1}^v + \eta_{\nu_1}^w \le 0, \quad \eta_{\nu_2}^v + \eta_{\nu_2}^w = 0 \},
\end{aligned}
$$

it is obviously the union of finitely many polyhedral convex sets. □

Since Φ defined in (20) is a polyhedral multifunction, [15, Proposition 1] may be invoked to show that Φ is locally upper Lipschitz at every point $v \in \mathbb{R}^{n+2|\beta|}$. It is therefore in particular calm at every $(v,w) \in \operatorname{gph}\Phi$ in the sense of [7]. By invoking [7, Corollary 4.2] we see that (19) implies

$$
(-\nabla f(z^*), 0, 0) \in N((0,0,0),\mathcal{D}_1) + N((0,0,0),\mathcal{D}_2). \tag{21}
$$

Now, the limiting normal cone of $\mathcal{D}_1$ is given by Lemma 2.7. This yields that there exist λ^g, λ^h, λ^G and λ^H with $\lambda^g_{\mathcal{I}_g} \ge 0$ such that

$$
\begin{aligned}
\begin{pmatrix} -\nabla f(z^*) \\ 0 \\ 0 \end{pmatrix} \in & \sum_{i\in\mathcal{I}_g} \lambda_i^g \begin{pmatrix} \nabla g_i(z^*) \\ 0 \\ 0 \end{pmatrix} + \sum_{i=1}^p \lambda_i^h \begin{pmatrix} \nabla h_i(z^*) \\ 0 \\ 0 \end{pmatrix} \\
& - \sum_{i\in\alpha} \lambda_i^G \begin{pmatrix} \nabla G_i(z^*) \\ 0 \\ 0 \end{pmatrix} - \sum_{i\in\gamma} \lambda_i^H \begin{pmatrix} \nabla H_i(z^*) \\ 0 \\ 0 \end{pmatrix} \\
& - \sum_{i\in\beta} \left[\lambda_i^G \begin{pmatrix} \nabla G_i(z^*) \\ -e^i \\ 0 \end{pmatrix} + \lambda_i^H \begin{pmatrix} \nabla H_i(z^*) \\ 0 \\ e^i \end{pmatrix} \right] \\
& + N((0,0,0),\mathcal{D}_2).
\end{aligned} \tag{22}
$$

where e^i denotes that unit vector in $\mathbb{R}^{|\beta|}$ which corresponds to the position of i in β. Note that since the signs in the second and third lines of (22) are arbitrary, they were chosen to facilitate the notation of the proof.

First, we take a look at the second and third components in (22). To this end, we rewrite the normal cone to $\mathcal{D}_2$ in the following fashion:

$$
\begin{aligned}
N((0,0,0),\mathcal{D}_2) &= N(0,\mathbb{R}^n) \times N((0,0), \{(\xi,\eta) \mid \xi \ge 0, \eta \le 0, \xi^T\eta = 0\}) \\
&= \{0\} \times N((0,0), \operatorname{gph} N_{\mathbb{R}^l_+}).
\end{aligned} \tag{23}
$$

Here the first equality is due to the product rule (see, e.g., [11] or [16, Proposition 6.41]). The second equality uses that 0 is in the interior of $\mathbb{R}^n$ (and hence any normal cone reduces to $\{0\}$) as well as Lemma 2.5.

Substituting (23) into (22) yields

$$(-\lambda_\beta^G, \lambda_\beta^H) \in N((0,0), \operatorname{gph} N_{\mathbb{R}^l_+}).$$

By Proposition 2.6, we obtain that

$$(\lambda_i^G > 0 \ \wedge \ \lambda_i^H > 0) \quad \vee \quad \lambda_i^G \lambda_i^H = 0$$

for all $i \in \beta$. Note that since we need to determine the limiting normal cone in the point $(0,0)$, it holds that $\mathcal{I}_x = \mathcal{I}_y = \emptyset$ in Proposition 2.6.

Finally, we set $\lambda_\gamma^G := 0$, $\lambda_\alpha^H := 0$, and $\lambda_i^g := 0$ for all $i \notin \mathcal{I}_g$ and have thus acquired the conditions for M-stationarity (5) with $\lambda^* := (\lambda^g, \lambda^h, \lambda^G, \lambda^H)$, completing the proof of Theorem 3.1. Note that even though we derived our conditions using the MPAEC (15), we have in fact acquired the conditions for M-stationarity of our original MPEC (1).

Remark. We wish to draw attention to two fundamental ideas used in the proof of Theorem 3.1. The first is due to Ye and entails introducing an MPAEC (see (15) and the discussion preceeding it).

The second idea is that we can separate the benign constraints from the complementarity constraints (divided here into $\mathcal{D}_1$ and $\mathcal{D}_2$) and consider the two types of constraints separately. We are able to do this because

$$N((0,0,0), \mathcal{D}) \subseteq N((0,0,0), \mathcal{D}_1) + N((0,0,0), \mathcal{D}_2)$$

(see (19) and (21)) holds due to a result by Henrion, Jourani and Outrata (see [7, Corollary 4.2]). Note that this does not hold in general, but is a direct consequence of our MPAEC (15) having constraints characterized by affine functions.

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 [21]. 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 [17]. However, the authors of [17] were only able to prove C-stationarity to be a necessary first order condition 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 [17] 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 [21]. However, the proof given in [21] is somewhat incomplete. We therefore presented a complete and much shorter proof of this result.

5 Acknowledgements

The authors wish to thank Jiři Outrata for pointing out a mistake and possible improvements in the initial draft of our approach. Also, our gratitude goes out to Jane Ye, who was always helpful when questions pertaining to her approach arose.

References

[1] M. S. Bazaraa and C. M. Shetty, *Foundations of Optimization*, vol. 122 of Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, Berlin, Heidelberg, New York, 1976.

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

[3] 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.

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

[5] ——, *On M-stationarity for mathematical programs with equilibrium constraints*. Institute of Applied Mathematics and Statistics, University of Würzburg, Preprint 253, March 2004.

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

[7] R. Henrion, A. Jourani, and J. Outrata, *On the calmness of a class of multifunctions*, SIAM Journal on Optimization, 13 (2002), pp. 603–618.

[8] X. Hu and D. Ralph, *Convergence of a penalty method for mathematical programming with complementarity constraints*, tech. report, Judge Institute of Management Studies, University of Cambridge, England, October 2000.

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

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

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

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

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

[14] 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.

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

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

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

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

[19] 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.

[20] 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.

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

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