A FRITZ JOHN APPROACH TO FIRST ORDER OPTIMALITY CONDITIONS FOR MATHEMATICAL PROGRAMS WITH EQUILIBRIUM 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

October 24, 2002

Abstract. Mathematical programs with equilibrium constraints (MPECs) are nonlinear programs which do not satisfy any of the common constraint qualifications. In order to obtain first order optimality conditions, constraint qualifications tailored to MPECs have been developed and researched in the past. This has been done by falling back on technical proofs or results from nonsmooth analysis. In this paper, we use a completely dirfferent approach and show how the standard Fritz John conditions may be used in order to obtain the most important optimality conditions for MPECs. In this way, we obtain relatively short and elementary proofs for some known results in the MPEC field.

Key Words. mathematical programs with equilibrium constraints, constraint qualification, Fritz John conditions, first order conditions.

1 Introduction

Consider the constrained optimization problem

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

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 functions. Due to the complementarity term in the constraints, programs of this type are sometimes referred to as *mathematical programs with complementarity constraints*. More commonly, however, they are called *mathematical programs with equilibrium constraints*. This also yields the more pronouncable acronym MPEC, by which we will refer to the program (1) in the following. For more detail the reader is referred to, for example, the two monographs [7, 14].

It is well-known (see, e.g., [2, 21]) and easily verified that the MPEC (1) does not satisfy most of the common constraint qualification known from standard nonlinear programming at any feasible point. (One exception to this is the Guignard constraint qualification, see [4] for details.) Consequently, the usual Karush-Kuhn-Tucker conditions associated with the program (1) can, in general, not be viewed as first order optimality conditions for (1).

It has therefore been the subject of intensive research during the last few years to find suitable MPEC constraint qualifications under which a local minimizer of the problem (1) satisfies some first order optimality conditions. In fact, several first order optimality conditions have been derived under different sets of assumptions. The derivation of these optimality conditions, however, is usually either lengthy and technical or based on results from nonsmooth analysis.

For example, some authors reformulate the MPEC (1) as a nonsmooth program by rewriting the complementarity constraints as a nonsmooth equation. Results from nonsmooth analysis are then applied to this reformulated problem, see, e.g., [3, 16]. Other authors derive optimality conditions for (1) by using an exact penalty function for (1) and applying optimality conditions to this exact penalty function, which, however, is again nonsmooth in general, see, e.g., [7, 8, 22, 17]. Also, the tangent cone approach has been used by, e.g., [7, 15, 4]. Other approaches (like the implicit programming technique) are possible if one assumes that the MPEC (1) has a specific structure, see [11, 14, 19, 20, 22], for example.

In this paper, we present a straightforward and elementary approach to the most standard first order optimality conditions for (1). The basic idea is quite simple: Since the MPEC (1) does not satisfy any of the usual constraint qualifications known for constrained optimization problems, we apply the Fritz John conditions which hold without any regularity assumptions. However, the Fritz John conditions themselves provide relatively weak optimality conditions. Moreover, a direct application of the Fritz John conditions to the MPEC (1) does not lead to any meaningful results, so further thought is required in order to obtain suitable first order conditions.

Using this Fritz John approach, we reobtain existing results from the literature (see, in particular, [16], and to a certain extent [15]) by using a completely different technique. We feel that our derivation is much more elementary than existing approaches because our proofs are relatively short and we only need the standard Fritz John conditions for smooth optimization problems as prerequisites. As a by-product, we also obtain a new optimality condition under a Mangasarian-Fromovitz-type condition.

In Section 2 we review some existing constraint qualifications together with some stationarity concepts related to the MPEC (1), while Section 3 deals with our Fritz John approach.

The notation used in this paper is rather standard: $\mathbb{R}^n$ denotes the n-dimensional Euclidean space. For $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$, we simply write (x, y) for the $(n+m)$-dimensional column vector $(x^T, y^T)^T$. Given $x \in \mathbb{R}^n$ and a subset $\delta \subseteq \{1, \ldots, n\}$, we denote by x_δ the subvector in $\mathbb{R}^{|\delta|}$ consisting of all components x_i with $i \in \delta$. Finally, inequalities $x \geq 0$ with $x \in \mathbb{R}^n$ are defined componentwise.

2 Constraint Qualifications and Stationarity Concepts

We now commence by recalling some constraint qualifications for the MPEC (1) as well as some first order optimality conditions.

Before we begin, we need to introduce some notation. Given a feasible vector z^* of the MPEC (1), we define the following sets of indices:

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

The set β is known as the *degenerate* set. If it is empty, the vector z^* is said to fulfill *strict complementarity*. As we shall see, it will become convenient to split β into its partitions, which are defined as follows:

$$\mathcal{P}(\beta) := \{(\beta_1, \beta_2) \mid \beta_1 \cup \beta_2 = \beta, \beta_1 \cap \beta_2 = \emptyset\}. \tag{3}$$

To define altered constraint qualifications, we introduce the following program, dependent on z^*, and called the *tightened nonlinear program* *TNLP* $:=$ *TNLP*(z^*):

$$\begin{array}{lll} \min & f(z) & \\ \text{s.t.} & g(z) \leq 0, & h(z) = 0, \\ & G_{\alpha\cup\beta}(z) = 0, & G_\gamma(z) \geq 0, \\ & H_\alpha(z) \geq 0, & H_{\gamma\cup\beta}(z) = 0. \end{array} \tag{4}$$

The above nonlinear program is called *tightened* since the feasible region is a subset of the feasible region of the MPEC (1). This implies that if z^* is a local minimizer of the MPEC (1), then it is also a local minimizer of the corresponding tightened nonlinear program TNLP(z^*).

We will need the Karush-Kuhn-Tucker (KKT) conditions of the nonlinear program (4) throughout the remainder of this paper. We therefore write them down for a vector z and Lagrange multiplier $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in \mathbb{R}^{m+p+2l}$:

$$\begin{aligned} 0 = \nabla f(z) + \sum_{i=1}^{m} \lambda_i^g \nabla g_i(z) + \sum_{i=1}^{p} \lambda_i^h \nabla h_i(z) - \sum_{i=1}^{l} \left[\lambda_i^G \nabla G_i(z) + \lambda_i^H H_i(z)\right], \\ h(z) = 0, \quad g(z) \leq 0, \quad \lambda^g \geq 0, \quad (\lambda^g)^T g(z) = 0, \\ G_{\alpha\cup\beta}(z) = 0, \quad G_\gamma(z) \geq 0, \quad \lambda_\gamma^G \geq 0, \quad (\lambda_\gamma^G)^T G_\gamma(z) = 0, \\ H_{\gamma\cup\beta}(z) = 0, \quad H_\alpha(z) \geq 0, \quad \lambda_\alpha^H \geq 0, \quad (\lambda_\alpha^H)^T H_\alpha(z) = 0. \end{aligned} \tag{5}$$

The vector (z, λ) is said to be a *KKT point* of the tightened nonlinear program (4) if the conditions (5) hold.

The TNLP (4) can now be used to define suitable MPEC variants of the standard linear independence, Mangasarian-Fromovitz- and strict Mangasarian-Fromovitz constraint qualifications (LICQ, MFCQ, and SMFCQ for short).

Definition 2.1 *The MPEC (1) is said to satisfy the* MPEC-LICQ (MPEC-MFCQ, MPEC-SMFCQ) *in a feasible vector* z^* *if the corresponding TNLP(*z^**) satisfies the LICQ (MFCQ, SMFCQ) in that vector* z^**.*

Since we will need them in Section 3, we shall explicitly write down the constraint qualifications from Definition 2.1. The MPEC-LICQ expands to the condition that the gradient vectors

$$\begin{array}{ll} \nabla g_i(z^*) & \forall i \in \mathcal{I}_g := \{i \mid g_i(z^*) = 0\}, \\ \nabla h_i(z^*) & \forall i = 1, \ldots, p, \\ \nabla G_i(z^*) & \forall i \in \alpha \cup \beta, \\ \nabla H_i(z^*) & \forall i \in \gamma \cup \beta \end{array} \tag{6}$$

must be linearly independent. The MPEC-LICQ can also be defined using the so-called *relaxed* nonlinear program, which we shall not elaborate upon here. The resulting definition, however, is the same, see, e.g., [15].

Similarly, the MPEC-MFCQ expands to the following set of conditions: The gradient vectors

$$\begin{array}{ll} \nabla h_i(z^*) & \forall i = 1, \ldots, p, \\ \nabla G_i(z^*) & \forall i \in \alpha \cup \beta, \\ \nabla H_i(z^*) & \forall i \in \gamma \cup \beta \end{array} \tag{7a}$$

are linearly independent, and there exists a vector $d \in \mathbb{R}^n$ such that

$$\begin{array}{ll} \nabla h_i(z^*)^T d = 0 & \forall i = 1, \ldots, p, \\ \nabla G_i(z^*)^T d = 0 & \forall i \in \alpha \cup \beta, \\ \nabla H_i(z^*)^T d = 0 & \forall i \in \gamma \cup \beta, \\ \nabla g_i(z^*)^T d < 0 & \forall i \in \mathcal{I}_g. \end{array} \tag{7b}$$

At this point, it is important to note that under MPEC-MFCQ, a local minimizer z^* of the MPEC (1) implies the existence of a Lagrange multiplier λ^* such that (z^*, λ^*) satisfies the KKT conditions (5) (see, e.g., [10]). Therefore, if we assume that MPEC-MFCQ holds for a local minimizer z^* of the MPEC (1), we can use any Lagrange multiplier λ^* (which we now know exists) to define the MPEC-SMFCQ, i.e., taking (z^*, λ^*), we require the following to hold: The gradient vectors

$$\begin{array}{ll} \nabla h_i(z^*) & \forall i = 1, \dots, p, \\ \nabla G_i(z^*) & \forall i \in \alpha \cup \beta, \\ \nabla H_i(z^*) & \forall i \in \gamma \cup \beta, \\ \nabla g_i(z^*) & \forall i \in \mathcal{J}_g := \{i \in \mathcal{I}_g \mid (\lambda_i^g)^* > 0\} \end{array} \tag{8a}$$

are linearly independent, and there exists a vector $d \in \mathbb{R}^n$ such that

$$\begin{array}{ll} \nabla h_i(z^*)^T d = 0 & \forall i = 1, \dots, p, \\ \nabla G_i(z^*)^T d = 0 & \forall i \in \alpha \cup \beta, \\ \nabla H_i(z^*)^T d = 0 & \forall i \in \gamma \cup \beta, \\ \nabla g_i(z^*)^T d = 0 & \forall i \in \mathcal{J}_g, \\ \nabla g_i(z^*)^T d < 0 & \forall i \in \mathcal{K}_g := \{i \in \mathcal{I}_g \mid (\lambda_i^g)^* = 0\}. \end{array} \tag{8b}$$

Note that the above assumption that MPEC-MFCQ holds in order to define MPEC-SMFCQ is no restriction since MPEC-SMFCQ implies MPEC-MFCQ.

As mentioned earlier, classic KKT conditions are not appropriate in the context of MPECs. We therefore introduce two stationarity conditions used in [15, 16].

A feasible point z of the MPEC (1) is called *weakly stationary* [16] if there exists a Lagrange multiplier $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H)$ such that the following conditions hold:

$$\begin{array}{l} 0 = \nabla f(z) + \sum_{i=1}^{m} \lambda_i^g \nabla g_i(z) + \sum_{i=1}^{p} \lambda_i^h \nabla h_i(z) - \sum_{i=1}^{l} \left[\lambda_i^G \nabla G_i(z) + \lambda_i^H H_i(z)\right], \\ \lambda_\alpha^G \quad \text{free}, \qquad \lambda_\beta^G \quad \text{free}, \qquad \lambda_\gamma^G = 0, \\ \lambda_\gamma^H \quad \text{free}, \qquad \lambda_\beta^H \quad \text{free}, \qquad \lambda_\alpha^H = 0, \\ g(z) \le 0, \qquad \lambda^g \ge 0, \qquad (\lambda^g)^T g(z) = 0. \end{array} \tag{9}$$

A feasible point z of the MPEC (1) is called *strongly stationary* [16] or *primal-dual stationary* [15] if there exists a Lagrange multiplier $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H)$ such that the following conditions hold:

$$\begin{array}{l} 0 = \nabla f(z) + \sum_{i=1}^{m} \lambda_i^g \nabla g_i(z) + \sum_{i=1}^{p} \lambda_i^h \nabla h_i(z) - \sum_{i=1}^{l} \left[\lambda_i^G \nabla G_i(z) + \lambda_i^H H_i(z)\right], \\ \lambda_\alpha^G \quad \text{free}, \qquad \lambda_\beta^G \ge 0, \qquad \lambda_\gamma^G = 0, \\ \lambda_\gamma^H \quad \text{free}, \qquad \lambda_\beta^H \ge 0, \qquad \lambda_\alpha^H = 0, \\ g(z) \le 0, \qquad \lambda^g \ge 0, \qquad (\lambda^g)^T g(z) = 0. \end{array} \tag{10}$$

Note that the difference between the two stationarity conditions is the sign-restriction imposed on λ_β^G and λ_β^H in the case of strong stationarity. It is easily verified that strong stationarity coincides with the KKT conditions of the MPEC (1) (see, e.g., [4]). Furthermore, in the nondegenerate case, i.e. if $\beta = \emptyset$, strong stationarity is identical to weak stationarity.

Other stationary conditions are derived and examined elsewhere, among which are C-stationarity [16] (see (18)) and M-stationarity [12, 13]. Both lie between the weak and strong stationarity conditions (9) and (10), respectively. Hence they all coincide in the nondegenerate case, whereas, in general, differences occur in the properties of the multipliers λ_β^G and λ_β^H.

3 Fritz John Approach to Optimality Conditions

Before we are able to present the main results of this paper, we need to define another nonlinear program derived from the MPEC (1). The notation we use here is borrowed from [15].

Given a partition $(\beta_1, \beta_2) \in \mathcal{P}(\beta)$, let $\mathrm{NLP}_*(\beta_1, \beta_2)$ denote the following nonlinear program:

$$\begin{array}{lll} \min & f(z) & \\ \text{s.t.} & g(z) \le 0, & h(z) = 0, \\ & G_{\alpha\cup\beta_1}(z) = 0, & H_{\alpha\cup\beta_1}(z) \ge 0, \\ & G_{\gamma\cup\beta_2}(z) \ge 0, & H_{\gamma\cup\beta_2}(z) = 0. \end{array} \tag{11}$$

Note that the program $\mathrm{NLP}_*(\beta_1, \beta_2)$ depends on the vector z^*.

Note that a local minimizer z^* of the MPEC (1) is a local minimizer of the $\mathrm{NLP}_*(\beta_1, \beta_2)$ since z^* is feasible for the latter program and its feasible region is a subset of the feasible region of the MPEC (1).

As they are needed in the following, we shall write down the KKT conditions for a feasible point z with Lagrange multiplier λ of the $\mathrm{NLP}_*(\beta_1, \beta_2)$:

$$\begin{array}{c} 0 = \nabla f(z) + \sum_{i=1}^m \lambda_i^g \nabla g_i(z) + \sum_{i=1}^p \lambda_i^h \nabla h_i(z) - \sum_{i=1}^l \left[\lambda_i^G \nabla G_i(z) + \lambda_i^H H_i(z)\right], \\ \begin{array}{cccc} h(z) = 0, & g(z) \le 0, & \lambda^g \ge 0, & (\lambda^g)^T g(z) = 0, \\ G_{\alpha\cup\beta_1}(z) = 0, & G_{\gamma\cup\beta_2}(z) \ge 0, & \lambda^G_{\gamma\cup\beta_2} \ge 0, & (\lambda^G_{\gamma\cup\beta_2})^T G_{\gamma\cup\beta_2}(z) = 0, \\ H_{\gamma\cup\beta_2}(z) = 0, & H_{\alpha\cup\beta_1}(z) \ge 0, & \lambda^H_{\alpha\cup\beta_1} \ge 0, & (\lambda^H_{\alpha\cup\beta_1})^T H_{\alpha\cup\beta_1}(z) = 0. \end{array} \end{array} \tag{12}$$

We are now able to state the first main result of this paper.

Theorem 3.1 *Let $z^* \in \mathbb{R}^n$ be a local minimizer of the MPEC (1). If MPEC-LICQ holds in z^*, then there exists a unique Lagrange multiplier λ^* such that (z^*, λ^*) is strongly stationary.*

Proof. We shall consider two programs (the reason for this will become clear as the proof unfolds): The $\mathrm{NLP}_*(\beta_1, \beta_2)$ and its (in a sense) complementary program $\mathrm{NLP}_*(\beta_2, \beta_1)$ (note the inverted positions of β_1 and β_2). The vector z^* is a local minimum of these two programs since in both cases it is feasible and the feasible region of the corresponding program is a subset of the feasible region of the original MPEC (1).

Now let us first consider the nonlinear program $\mathrm{NLP}_*(\beta_1, \beta_2)$. Well-known results (see, e.g., [1, Proposition 3.3.5]) yield the existence of a nonzero vector $(r, \mu) = (r, \mu^g, \mu^h, \mu^G, \mu^H) \in \mathbb{R}^{1+m+p+2l}$, such that (r, z^*, μ) is a Fritz John point of the $\mathrm{NLP}_*(\beta_1, \beta_2)$, i.e., the following conditions hold:

$$
\begin{aligned}
0 = r\nabla f(z^*) + \sum_{i=1}^{m} \mu_i^g \nabla g_i(z^*) + \sum_{i=1}^{p} \mu_i^h \nabla h_i(z^*) - \sum_{i=1}^{l} \left[\mu_i^G \nabla G_i(z^*) + \mu_i^H \nabla H_i(z^*)\right]&, \\
r \geq 0&, \\
h(z^*) = 0, \qquad g(z^*) \leq 0, \qquad \mu^g \geq 0, \qquad (\mu^g)^T g(z^*) = 0&, \\
G_{\alpha\cup\beta_1}(z^*) = 0, \quad G_{\gamma\cup\beta_2}(z^*) \geq 0, \quad \mu^G_{\gamma\cup\beta_2} \geq 0, \quad (\mu^G_{\gamma\cup\beta_2})^T G_{\gamma\cup\beta_2}(z^*) = 0&, \\
H_{\gamma\cup\beta_2}(z^*) = 0, \quad H_{\alpha\cup\beta_1}(z^*) \geq 0, \quad \mu^H_{\alpha\cup\beta_1} \geq 0, \quad (\mu^H_{\alpha\cup\beta_1})^T H_{\alpha\cup\beta_1}(z^*) = 0&.
\end{aligned} \tag{13}
$$

It follows from the latter two lines of (13) and by the definition of the sets α and γ ((2a) and (2c)) that $\mu^G_\gamma = 0$ and $\mu^H_\alpha = 0$. Furthermore, we know that $\mu^g_i = 0$ for all $i \notin \mathcal{I}_g$, where $\mathcal{I}_g$ is defined as in (6).

We want to show that $r > 0$. To do this, we assume that $r = 0$ and examine the first line of (13), taking into account the special structure of μ:

$$
0 = \sum_{i\in\mathcal{I}_g} \mu_i^g \nabla g_i(z^*) + \sum_{i=1}^{p} \mu_i^h \nabla h_i(z^*) - \sum_{i\in\alpha\cup\beta} \mu_i^G \nabla G_i(z^*) - \sum_{i\in\gamma\cup\beta} \mu_i^H \nabla H_i(z^*). \tag{14}
$$

Since by MPEC-LICQ, all terms in the sum (14) are linearly independent, it follows that $\mu = 0$, which is a contradiction to the assumption that $(r, \mu) \neq (0, 0)$. We therefore have $r > 0$ and can, without loss of generality, set $r = 1$.

Hence, we have the existence of a Lagrange multiplier $\tilde{\lambda}$ (with certain properties, inherited from μ) which satisfies the KKT conditions (12) of the $\mathrm{NLP}_*(\beta_1, \beta_2)$:

$$
\begin{aligned}
0 = \nabla f(z^*) + \sum_{i=1}^{m} \tilde{\lambda}_i^g \nabla g_i(z^*) + \sum_{i=1}^{p} \tilde{\lambda}_i^h \nabla h_i(z^*) - \sum_{i=1}^{l} \left[\tilde{\lambda}_i^G \nabla G_i(z^*) + \tilde{\lambda}_i^H H_i(z^*)\right]&, \\
h(z^*) = 0, \qquad g(z^*) \leq 0, \qquad \tilde{\lambda}^g \geq 0, \qquad (\tilde{\lambda}^g)^T g(z^*) = 0&, \\
G_{\alpha\cup\beta_1}(z^*) = 0, \quad G_{\gamma\cup\beta_2}(z^*) \geq 0, \quad \tilde{\lambda}^G_{\beta_2} \geq 0, \ \tilde{\lambda}^G_\gamma = 0, \quad (\tilde{\lambda}^G_{\beta_2})^T G_{\beta_2}(z^*) = 0&, \\
H_{\gamma\cup\beta_2}(z^*) = 0, \quad H_{\alpha\cup\beta_1}(z^*) \geq 0, \quad \tilde{\lambda}^H_{\beta_1} \geq 0, \ \tilde{\lambda}^H_\alpha = 0, \quad (\tilde{\lambda}^H_{\beta_1})^T H_{\beta_1}(z^*) = 0&.
\end{aligned} \tag{15}
$$

Similarly, taking the complementary $\mathrm{NLP}_*(\beta_2, \beta_1)$, we can show the existence of a Lagrange multiplier $\hat{\lambda}$ which satisfies the KKT conditions for the program NLP$_*(\beta_2, \beta_1)$:

$$
\begin{aligned}
0 = \nabla f(z^*) + \sum_{i=1}^{m} \hat{\lambda}_i^g \nabla g_i(z^*) + \sum_{i=1}^{p} \hat{\lambda}_i^h \nabla h_i(z^*) - \sum_{i=1}^{l} \left[\hat{\lambda}_i^G \nabla G_i(z^*) + \hat{\lambda}_i^H H_i(z^*)\right], \\
h(z^*) = 0, \qquad g(z^*) \leq 0, \qquad \hat{\lambda}^g \geq 0, \qquad (\hat{\lambda}^g)^T g(z^*) = 0, \\
G_{\alpha\cup\beta_2}(z^*) = 0, \quad G_{\gamma\cup\beta_1}(z^*) \geq 0, \quad \hat{\lambda}_{\beta_1}^G \geq 0, \; \hat{\lambda}_\gamma^G = 0, \quad (\hat{\lambda}_{\beta_1}^G)^T G_{\beta_1}(z^*) = 0, \\
H_{\gamma\cup\beta_1}(z^*) = 0, \quad H_{\alpha\cup\beta_2}(z^*) \geq 0, \quad \hat{\lambda}_{\beta_2}^H \geq 0, \; \hat{\lambda}_\alpha^H = 0, \quad (\hat{\lambda}_{\beta_2}^H)^T H_{\beta_2}(z^*) = 0.
\end{aligned} \tag{16}
$$

We now equate the first equations of (15) and (16), taking into account the special structure of $\tilde{\lambda}$ and $\hat{\lambda}$ inherited from the multiplier μ:

$$
\begin{aligned}
0 &= \sum_{i=1}^{m} (\hat{\lambda}_i^g - \tilde{\lambda}_i^g) \nabla g_i(z^*) + \sum_{i=1}^{p} (\hat{\lambda}_i^h - \tilde{\lambda}_i^h) \nabla h_i(z^*) \\
&\quad - \sum_{i=1}^{l} \left[(\hat{\lambda}_i^G - \tilde{\lambda}_i^G) \nabla G_i(z^*) + (\hat{\lambda}_i^H - \tilde{\lambda}_i^H) \nabla H_i(z^*)\right] \\
&= \sum_{i\in\mathcal{I}_g} (\hat{\lambda}_i^g - \tilde{\lambda}_i^g) \nabla g_i(z^*) + \sum_{i=1}^{p} (\hat{\lambda}_i^h - \tilde{\lambda}_i^h) \nabla h_i(z^*) \\
&\quad - \sum_{i\in\alpha\cup\beta} (\hat{\lambda}_i^G - \tilde{\lambda}_i^G) \nabla G_i(z^*) - \sum_{i\in\gamma\cup\beta} (\hat{\lambda}_i^H - \tilde{\lambda}_i^H) \nabla H_i(z^*).
\end{aligned}
$$

Again using the linear independence of all terms involved, we have $\tilde{\lambda} = \hat{\lambda}$. By setting $\lambda^* := \tilde{\lambda} = \hat{\lambda}$, we see that λ^* has the combined characteristics of the Lagrange multipliers in both (15) and (16), in particular

$$
\begin{aligned}
(\lambda_\beta^G)^* \geq 0, \qquad (\lambda_\gamma^G)^* = 0, \\
(\lambda_\beta^H)^* \geq 0, \qquad (\lambda_\alpha^H)^* = 0,
\end{aligned}
$$

satisfying the conditions for strong stationarity.

Uniqueness of the Lagrange multiplier follows immediately from (10). This concludes the proof. □

Theorem 3.1 is essentially one direction of [15, Theorem 3] and is also a direct consequence of [16, Theorem 2, (2)], which was proved using different techniques.

Although the above theorem is sufficient for most practical purposes (see [18] for a discussion), it is of interest whether the same result (or perhaps a weaker one) still holds if the MPEC-LICQ is relaxed. To this end, we examine the MPEC-MFCQ and the MPEC-SMFCQ in the following two theorems.

Theorem 3.2 *Let $z^* \in \mathbb{R}^n$ be a local minimizer of the MPEC (1). If MPEC-MFCQ holds in z^* then there exists a Lagrange multiplier λ^* such that (z^*, λ^*) satisfies the following stationarity conditions:*

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

$$\begin{array}{lll} (\lambda_\alpha^G)^* \quad \textit{free}, & & (\lambda_\gamma^G)^* = 0, \\ & (\lambda_i^G)^* \geq 0 \vee (\lambda_i^H)^* \geq 0 \quad \forall i \in \beta & \\ (\lambda_\gamma^H)^* \quad \textit{free}, & & (\lambda_\alpha^H)^* = 0, \end{array}$$

$$g(z^*) \leq 0, \qquad (\lambda^g)^* \geq 0, \qquad g(z^*)^T (\lambda^g)^* = 0. \tag{17}$$

In particular, (z^*, λ^*) is weakly stationary.

Proof. Up to and including the sum (14), this proof is identical to the proof of Theorem 3.1. Using MPEC-MFCQ and Motzkin's theorem of the alternative (cf. [9, Theorem 2.4.2]), we obtain $\mu = 0$ from (14). Since this is a contradiction to the existence of a nonzero vector (r, μ) we have shown, without loss of generality, the existence of a Lagrange multiplier $\tilde{\lambda}$ satisfying (15). In particular, we are able to conclude that

$$\tilde{\lambda}_{\beta_2}^G \geq 0, \qquad \tilde{\lambda}_{\beta_1}^H \geq 0.$$

Setting $\lambda^* := \tilde{\lambda}$ proves the result. □

Note that the proof of Theorem 3.2 holds for an arbitrary partition (β_1, β_2) of the index set β. Hence we can choose, a priori, such a partition and obtain corresponding Lagrange multipliers $(\lambda^G)^*$ and $(\lambda^H)^*$ such that $(\lambda_i^G)^* \geq 0$ for all $i \in \beta_1$ and $(\lambda_i^H)^* \geq 0$ for all $i \in \beta_2$.

Motivated by Theorem 3.2, we call a weakly stationary point z^* of the MPEC (1) *A-stationary* if there exists a corresponding Lagrange multiplier λ^* such that

$$(\lambda_i^G)^* \geq 0 \quad \text{or} \quad (\lambda_i^H)^* \geq 0 \quad \forall i \in \beta,$$

i.e., z^* is A-stationary if and only if (17) holds for some multiplier λ^*. Here, the letter 'A' may stand for 'alternative' since, for each $i \in \beta$, we have the alternative that either $(\lambda_i^G)^* \geq 0$ or $(\lambda_i^H)^* \geq 0$ (or both) hold. However, the letter 'A' may also be interpreted as an abbreviation for 'Abadie' since, as is discussed in [5], an Abadie-type constraint qualification also implies A-stationarity.

By the above remark that a partition (β_1, β_2) may be chosen a priori, the following corollary is an immediate consequence of Theorem 3.2.

Corollary 3.3 *Let $z^* \in \mathbb{R}^n$ be a local minimizer of the MPEC (1). If MPEC-MFCQ holds in z^* then for every partition $(\beta_1, \beta_2) \in \mathcal{P}(\beta)$ there exists a Lagrange multiplier λ^* such that (z^*, λ^*) is A-stationary.*

The first order condition in Corollary 3.3 is called a *primal-dual first-order condition* in [7] and this particular one can be found in Theorem 3.3.6 of the same reference.

A slightly different theorem under MPEC-MFCQ is stated in [16], where the existence of Lagrange multipliers $(\lambda^G)^*$ and $(\lambda^H)^*$ satisfying

$$(\lambda_i^G)^*(\lambda_i^H)^* \geq 0 \qquad \forall i \in \beta \tag{18}$$

is shown. A weakly stationary point satisfying this condition is called *Clarke-* or simply *C-stationary* in [16].

Note that either theorem merely show the existence of Lagrange multipliers with certain characteristics, but not the exclusion of other Lagrange multipliers. In fact, as the following example (taken from [16]) demonstrates, the respective conditions are not, in general, satisfied by the same set of Lagrange multipliers:

$$\begin{array}{ll} \min & f(z) := z_1 + z_2 - z_3 \\ \text{s.t.} & g(z) := \begin{pmatrix} -4z_1 + z_3 \\ -4z_2 + z_3 \end{pmatrix} \leq 0, \\ & G(z) := z_1 \geq 0, \\ & H(z) := z_2 \geq 0, \\ & G(z)^T H(z) = z_1 z_2 = 0. \end{array}$$

The origin is the unique solution of this program, and it satisfies the MPEC-MFCQ. It is easily verified that the corresponding Lagrange multipliers $\lambda^G := (\lambda_1^G)^*$ and $\lambda^H := (\lambda_1^H)^*$ are subject to the following restrictions:

$$\begin{array}{l} \lambda^H \in [-3;1], \\ \lambda^G = -\lambda^H - 2. \end{array}$$

The Lagrange multipliers satisfying A-stationarity (17) are $\{(\lambda^G,\lambda^H) \mid \lambda^H \in [-3;-2] \cup [0;1],\ \lambda^G = -\lambda^H - 2\}$, while the conditions (18) for C-stationarity are satisfied by the multipliers $\{(\lambda^G,\lambda^H) \mid \lambda^H \in [-2;0],\ \lambda^G = -\lambda^H - 2\}$. Only for $(\lambda^G,\lambda^H) = (0,-2)$ and $(\lambda^G,\lambda^H) = (-2,0)$ are both conditions (17) and (18) satisfied simultaneously.

The following result can also be found in [16], where it is derived using different techniques.

Theorem 3.4 *Let $z^* \in \mathbb{R}^n$ be a local minimizer of the MPEC (1). If MPEC-SMFCQ holds in z^* then there exists a unique Lagrange multiplier λ^* such that (z^*,λ^*) is strongly stationary.*

Proof. Since MPEC-SMFCQ implies MPEC-MFCQ, the proof of Theorem 3.2 yields the existence of a Lagrange multiplier $\tilde{\lambda}$ which satisfies the KKT conditions (15). By the same arguments we have the existence of a Lagrange multiplier $\hat{\lambda}$ which satisfies the KKT conditions (16).

It is easily verified (taking into account that $G_\beta(z^*) = 0$ and $H_\beta(z^*) = 0$ by definition of the index set β) that both Lagrange multipliers $\tilde{\lambda}$ and $\hat{\lambda}$ also satisfy the KKT conditions

(5) of the TNLP (4). Since SMFCQ holds for the TNLP (4), the Lagrange multiplier is unique (cf. [6]) and hence $\tilde{\lambda} = \hat{\lambda}$. By the same arguments as in the proof of Theorem 3.1, we set $\lambda^* := \tilde{\lambda} = \hat{\lambda}$, which satisfies the combined characteristics of (15) and (16). This implies strong stationarity.

Since a Lagrange multiplier satisfying (10) also satisfies the KKT conditions (5) of the TNLP (4), and SMFCQ holds for this TNLP, the Lagrange multiplier is unique (cf., again, [6]). This concludes the proof. □

4 Conclusion

We have examined some of the more important constraint qualifitcations for MPECs (MPEC-LICQ, MPEC-SMFCQ, and MPEC-MFCQ) and have proven known first order conditions using an elementary approach. We were able to show that strong stationarity holds under both MPEC-LICQ and MPEC-SMFCQ using our technique. The same technique yielded a new stationarity concept, A-stationarity, under MPEC-MFCQ.

References

[1] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, Massachusetts, second ed., 1999.

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

[3] F. Facchinei, H. Jiang, and L. Qi, *A smoothing method for mathematical programs with equilibrium constraints*, Mathematical Programming, 85 (1999), pp. 107–134.

[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] ——, *Optimality conditions for mathematical programs with equilibrium contraints: Fritz John and Abadie-type approaches*. Institute of Applied Mathematics and Statistics, University of Würzburg, Preprint 245, May 2002.

[6] J. Kyparisis, *On uniqueness of Kuhn-Tucker multipliers in nonlinear programming*, Mathematical Programming, 32 (1985), pp. 242–246.

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

[8] Z.-Q. Luo, J.-S. Pang, D. Ralph, and S.-Q. Wu, *Exact penalization and stationary conditions of mathematical programs with equilibrium constraints*, Mathematical Programming, 75 (1996), pp. 19–76.

[9] O. L. Mangasarian, *Nonlinear Programming*, Classics in Applied Mathematics, SIAM, Philadelphia, 1994.

[10] O. L. Mangasarian and S. Fromovitz, *The Fritz John optimality necessary conditions in the presence of equality and inequality constraints*, Journal of Mathematical Analysis and Applications, 17 (1967), pp. 37–47.

[11] J. V. Outrata, *On optimization problems with variational inequality constraints*, SIAM Journal on Optimization, 4 (1994), pp. 340–357.

[12] ——, *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. V. Outrata, M. Kočvara, and J. Zowe, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*, Nonconvex Optimization and its Applications, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

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

[17] S. Scholtes and M. Stöhr, *Exact penalization of mathematical programs with equilibrium constraints*, SIAM Journal on Control and Optimization, 37 (1999), pp. 617–652.

[18] ——, *How stringent is the linear independence assumption for mathematical programs with complementarity constraints?*, Mathematics of Operations Research, 26 (2001), pp. 851–863.

[19] J. J. Ye, *Optimality conditions for optimization problems with complementarity constraints*, SIAM Journal on Optimization, 9 (1999), pp. 374–387.

[20] J. J. Ye and X. Y. Ye, *Necessary optimality conditions for optimization problems with variational inequality constraints*, Mathematics of Operations Research, 22 (1997), pp. 977–997.

[21] J. J. Ye and D. L. Zhu, *Optimality conditions for bilevel programming problems*, Optimization, 33 (1995), pp. 9–27.

[22] J. J. Ye, D. L. Zhu, and Q. J. Zhu, *Exact penalization and necessary optimality conditions for generalized bilevel programming problems*, SIAM Journal on Optimization, 7 (1997), pp. 481–507.