A NOTE ON THE RELATION BETWEEN STRONG AND M-STATIONARITY FOR A CLASS OF MATHEMATICAL PROGRAMS WITH EQUILIBRIUM CONSTRAINTS

René Henrion, Jiří Outrata and Thomas Surowiec

In this paper, we deal with strong stationarity conditions for mathematical programs with equilibrium constraints (MPEC). The main task in deriving these conditions consists in calculating the Fréchet normal cone to the graph of the solution mapping associated with the underlying generalized equation of the MPEC. We derive an inner approximation to this cone, which is exact under an additional assumption. Even if the latter fails to hold, the inner approximation can be used to check strong stationarity via the weaker (but easier to calculate) concept of M-stationarity.

Keywords: mathematical programs with equilibrium constraints, S-stationary points, M-stationary points, Fréchet normal cone, limiting normal cone

Classification: 90C30, 49J53

1. INTRODUCTION

Over the last twenty years, both researchers as well as practitioners have paid a considerable amount of attention to so-called *mathematical programs with equilibrium constraints* (MPECs). This class of problems offers a very suitable modeling framework for a number of practical problems (cf. the monographs [8, 13] and the references therein). In addition, the study of MPECs provides a fruitful area for application of various notions and tools from modern variational analysis ([11, Chapter 5]). Because of the intrinsic nonsmoothness arising in every MPEC setting, various stationarity concepts have been developed (cf. [18] for the case of equilibria governed by complementarity problems). Among these conditions, a distinguished role is played by so-called M(ordukhovich)-stationarity, which leads to rather sharp optimality conditions under none too restrictive qualification conditions, and by S(trong)-stationarity, where one pays for increased sharpness by the necessity to impose a rather strong qualification condition. This strong qualification condition, which apparently cannot be fulfilled in numerous MPECs, has been thoroughly analyzed in the case of equilibria given by complementarity problems (see ([12, 14]). On the other hand, very little is known about S-stationarity provided the equilibria are governed by variational inequalities (VIs).

The aim of this paper is to analyze the S-stationarity of MPEC solutions, where the equilibria are governed by strongly regular VIs ([16]), and there are no other constraints. We assume first that the constraint set of the considered VI is a convex polyhedron and compute an inner approximation of the Fréchet normal cone to the feasible set. Under a new qualification condition in terms of the problem data this inclusion becomes an equality, which in turn ensures S-stationarity of the point in question. Next, we assume that the constraint set of the VI is given by smooth convex inequalities and rewrite the VI into the "enhanced" Lagrangian formulation. On the basis of the previous result and the respective M-stationarity conditions, we are then able to obtain a qualification condition for S-stationarity also in this situation.

A standard notation is employed, where $\operatorname{gph}\Phi$ stands for the graph of a multifunction Φ, K^- is the negative polar cone to a cone K and $f'(x;d)$ denotes the directional derivative of f at x in the direction d and by Id we denote the identity operator. Throughout the paper, VIs will be written down in the form of *generalized equations* ([16]), which is well-suited for this type of analysis.

2. PROBLEM FORMULATION AND PRELIMINARIES

In this paper, we will consider MPECs of the type:

$$\min_{x,z} \{ f(x,z) \,|\, z \in S(x) \} . \tag{1}$$

Here, $f : \mathbb{R}^s \times \mathbb{R}^t \to \mathbb{R}$ is continuously differentiable and $S : \mathbb{R}^s \rightrightarrows \mathbb{R}^t$ is the solution mapping to the generalized equation

$$0 \in F(x,z) + N_C(z), \tag{2}$$

where $F : \mathbb{R}^s \times \mathbb{R}^t \to \mathbb{R}^t$ is continuously differentiable and

$$C := \left\{ z \in \mathbb{R}^t \,|\, A(z) \le 0 \right\},$$

with $A : \mathbb{R}^t \to \mathbb{R}^p$ twice continuously differentiable and having convex components A_j for all $j = 1,\dots,p$. Note that the convexity assumption implies $N_C(z)$ is the standard normal cone from convex analysis.

We begin by defining the notions of variational analysis important for our study. We refer the reader to [17] and [10] for a detailed introduction to the objects introduced in the following. For a closed set $C \subseteq \mathbb{R}^t$, we recall the definition of the *contingent cone* to C at some point $\bar{z} \in C$:

$$T_C(\bar{z}) := \left\{ d \in \mathbb{R}^t \,|\, \exists\, \tau_k \searrow 0, \exists\, d_k \to d : \forall\, k, \bar{z} + \tau_k d_k \in C \right\}.$$

In the case C is convex, this reduces to the standard tangent cone from convex analysis. Using the contingent cone, the *Fréchet normal cone* is defined as $\widehat{N}_C(\bar{z}) := [T_C(\bar{z})]^-$. By taking the outer limit of Fréchet normal cones in the following way

$$N_C(\bar{z}) := \{ z^* \mid \exists\, (z_n, z_n^*) \to (\bar{z}, z^*) : (z_n, z_n^*) \in C \times \widehat{N}_C(z_n)\, \forall\, n \},$$

one arrives at the *Mordukhovich* or *limiting normal cone* to C at $\bar{z}$. Note that if C is convex, then $\widehat{N}_C = N_C$ and the cones coincide with the classical notion of normal cone from convex analysis. Given a multifunction $\Phi : \mathbb{R}^s \rightrightarrows \mathbb{R}^t$, the *Mordukhovich coderivative* of Φ at $(\bar{x}, \bar{z}) \in \operatorname{gph}\Phi$ in (dual) direction $z^* \in \mathbb{R}^t$ is defined as

$$D^*\Phi(\bar{x},\bar{z})(z^*) := \{x^* \in \mathbb{R}^s \,|(x^*, -z^*) \in N_{\operatorname{gph}\Phi}(\bar{x},\bar{z})\}.$$

In this note, we will call upon certain stability properties of multifunctions, which we outline below. Let $\Phi : \mathbb{R}^s \rightrightarrows \mathbb{R}^t$ be a multifunction and $(\bar{x}, \bar{z}) \in \operatorname{gph}\Phi$. Then we say that Φ has the *Aubin property* at $(\bar{x}, \bar{z})$, if there exist neighborhoods $\mathcal{U}$ of $\bar{x}$ and $\mathcal{V}$ of $\bar{z}$ and a constant $\kappa > 0$ such that the following relation holds

$$d(z, \Phi(x')) \le \kappa||x - x'||, \forall\, z \in \mathcal{V} \cap \Phi(x), \forall\, x, x' \in \mathcal{U}.$$

By fixing $x' = \bar{x}$, we say that Φ is *calm* at $(\bar{x}, \bar{z})$. Evidently, the Aubin property implies calmness.

Let now $(\bar{x}, \bar{z})$ be a solution to (2) and define the multifunction $\Sigma : \mathbb{R}^t \rightrightarrows \mathbb{R}^t$ via a local partial linearization with respect to z

$$\Sigma(\xi) := \left\{z \in \mathbb{R}^t \,|\xi \in F(\bar{x},\bar{z}) + \nabla_z F(\bar{x},\bar{z})(z-\bar{z}) + N_C(z)\right\}.$$

If there exist neighborhoods $\mathcal{W}$ of $0 \in \mathbb{R}^t$ and $\mathcal{V}$ of $\bar{z}$ such that the map $\xi \mapsto \Sigma(\xi) \cap \mathcal{V}$ is *single-valued and Lipschitz continuous* on $\mathcal{W}$ with modulus κ, then the generalized equation is is called *strongly regular* at $(\bar{x}, \bar{z})$ in the sense of Robinson (see [16]). Strong regularity has the implication that for any $\varepsilon > 0$ there exist neighborhoods $\mathcal{U}_\varepsilon$ of $\bar{x}$ and $\mathcal{V}_\varepsilon$ of $\bar{z}$ such that the mapping $x \mapsto \sigma(x) := S(x) \cap \mathcal{V}_\varepsilon$ (with S from (1)) is single-valued and Lipschitz on $\mathcal{U}_\varepsilon$ with Lipschitz modulus $(\kappa+\varepsilon)L$, where L is the uniform Lipschitz modulus of $F(\cdot, z)$ on $\mathcal{U}_\varepsilon$ for all $z \in \mathcal{V}_\varepsilon$ (see [16], Theorem 2.1).

Given $(\bar{x}, \bar{z}) \in \operatorname{gph} S$, we say that the *strong second-order sufficient condition* (SSOSC) holds at $(\bar{x}, \bar{z})$, if $\nabla_z \mathcal{L}(\bar{x}, \bar{z}, \bar{\lambda})$ is positive definite on $\ker(\nabla A_{I_+(\bar{z},\bar{\lambda})}(\bar{z}))$ for all $\bar{\lambda} \in N_{\mathbb{R}^p_-}(A(\bar{z}))$ with

$$\nabla^T A(\bar{z})\bar{\lambda} = -F(\bar{x},\bar{z}),$$

where $\mathcal{L}(x, z, \lambda) = F(x, z) + \nabla^T A(z)\lambda$ and $I_+(\bar{z}, \bar{\lambda}) := \left\{j \in \{1, \dots, p\} | \bar{\lambda}_j > 0\right\}$.

Clearly, (1) can be converted to the mathematical-programming form

$$\min_{x,z}\left\{f(x,z)\,\middle|\,\begin{bmatrix} z \\ -F(x,z)\end{bmatrix} \in \operatorname{gph} N_C\right\}. \tag{3}$$

A feasible point $(\hat{x}, \hat{z})$ to (1) is called *strongly* or *S-stationary* if there is a pair of multipliers $(u^*, b^*) \in \widehat{N}_{\operatorname{gph} N_C}(\hat{z}, -F(\hat{x}, \hat{z}))$ such that

$$\begin{aligned} 0 &= \nabla_x^T f(\hat{x},\hat{z}) - \nabla_x^T F(\hat{x},\hat{z})b^* \\ 0 &= \nabla_z^T f(\hat{x},\hat{z}) - \nabla_z^T F(\hat{x},\hat{z})b^* + u^*. \end{aligned} \tag{4}$$

If $(u^*, b^*) \in N_{\operatorname{gph} N_C}(\hat{z}, -F(\hat{x}, \hat{z}))$, one speaks about *Mordukhovich* or *M-stationarity*. Equivalently, with $v^* = -b^*$, M-stationarity amounts to the existence of a multiplier v^* such that

$$0 = \nabla_x^T f(\hat{x},\hat{z}) + \nabla_x^T F(\hat{x},\hat{z})v^* \tag{5}$$

$$0 \in \nabla_z^T f(\hat{x},\hat{z}) + \nabla_z^T F(\hat{x},\hat{z})v^* + D^* N_C(\hat{z}, -F(\hat{x},\hat{z}))(v^*). \tag{6}$$

From [3, Theorem 4.1] we know that

$$N_{\mathrm{gph}S}(\hat{x},\hat{z}) \subseteq \begin{bmatrix} 0 & -\nabla_x^T F(\hat{x},\hat{z}) \\ \mathrm{Id} & -\nabla_z^T F(\hat{x},\hat{z}) \end{bmatrix} N_{\mathrm{gph}N_C}(\hat{z}, -F(\hat{x},\hat{z})), \tag{7}$$

whenever the multifunction

$$\Psi(u) := \{(x,z) \in \mathbb{R}^s \times \mathbb{R}^t \begin{bmatrix} z \\ -F(x,z) \end{bmatrix} - u \in \mathrm{gph}N_C\} \tag{8}$$

is calm at $(0,\hat{x},\hat{z})$. On the basis of [9] we arrive in this way at the following statement.

Theorem 2.1. Let $(\bar{x},\bar{z})$ be a local solution to (1) and let the multifunction (8) be calm at $(0,\bar{x},\bar{z})$. Then $(\bar{x},\bar{z})$ is M-stationary, i. e., there exists a multiplier $v^* \in \mathbb{R}^t$ such that relations (5), (6) are fulfilled (with $(\hat{x},\hat{z})$ replaced by $(\bar{x},\bar{z})$).

Recent developments in the study of explicit formulae for coderivatives of the type displayed in (6) allow us to ultimately rewrite (5) and (6) even more explicitly, i. e., without any ambiguous terms, provided C enjoys certain regularity properties. In the following proposition we assume without loss of generality, that $A(\bar{z}) = 0$ at the considered point $\bar{z} \in C$. This is justified, because the derived necessary optimality conditions just depend on the problem data in a neighbourhood of the considered local solution. Hence, inequality constraints which are non-binding at this solution can be ignored.

Proposition 2.2. Let $(\bar{x},\bar{z})$ be a local solution to (1) and, without loss of generality, assume that $A(\bar{z}) = 0$. Suppose that

1. $\nabla A(\bar{z})$ is surjective;
2. The perturbation mapping (8) is calm at $(0,\bar{x},\bar{z})$.

Then there exists a unique $\bar{\lambda} \in \mathbb{R}^p_+$ and vectors $(v^*, w^*) \in \mathbb{R}^t \times \mathbb{R}^p$ such that

$$-\nabla_x^T f(\bar{x},\bar{z}) = \nabla_x^T F(\bar{x},\bar{z}) v^* \tag{9}$$

$$-\nabla_z^T f(\bar{x},\bar{z}) = \nabla_z^T F(\bar{x},\bar{z}) v^* + \left(\sum_{j=1}^p \bar{\lambda}_j \nabla^2 A_j(\bar{z})\right) v^* + \nabla^T A(\bar{z}) w^* \tag{10}$$

$$\nabla A_j(\bar{z}) v^* = 0 \quad \forall\, j : \bar{\lambda}_j > 0 \tag{11}$$

$$w_j^* = 0 \quad \forall\, j : \bar{\lambda}_j = 0,\ \nabla A_j(\bar{z}) v^* < 0 \tag{12}$$

$$w_j^* \geq 0 \quad \forall\, j : \bar{\lambda}_j = 0,\ \nabla A_j(\bar{z}) v^* > 0 \tag{13}$$

$$F(\bar{x},\bar{z}) = -\nabla^T A(\bar{z}) \bar{\lambda}. \tag{14}$$

Proof. Given assumption 2., there exists $v^* \in \mathbb{R}^t$ such that (5) and (6) hold, from which we immediately obtain (9). Turning now to the coderivative in (6), assumption 1. allows us to invoke Theorem 3.1 in [5], which states

$$D^* N_C(\bar{z}, -F(\bar{x},\bar{z}))(v^*) = \left(\sum_{j=1}^p \bar{\lambda}_j \nabla^2 A_j(\bar{z})\right) v^* + \nabla^T A(\bar{z}) D^* N_{\mathbb{R}^p_-}(A(\bar{z}), \bar{\lambda})(\nabla A(\bar{z}) v^*),$$

where $\bar{\lambda} \in \mathbb{R}^p_+$ is uniquely defined by (14). Furthermore, Corollary 3.1 in [5] states that the last term in the formula above amounts to $\nabla^T A(\bar{z})w^*$, where v^* and w^* satisfy (11) – (13). $\square$

3. APPROXIMATING $\widehat{N}_{\mathrm{GPH}\,S}$

3.1. Polyhedral feasible sets

In this section, we provide results detailing how one could possibly calculate, or at the very least, characterize, a large subset contained within $\hat{N}_{\mathrm{gph}\,S}$. These results will allow us to characterize S-stationarity under certain assumptions similar to those which were used for characterizing M-stationarity. In addition, the results will later be used to intimate how the gap between M-stationarity and S-stationarity might be bridged. To begin, we associate the solution mapping S with the following generalized equation

$$0 \in \mathcal{F}(x, w) + N_{\mathcal{C}}(w), \tag{15}$$

where $x \in \mathbb{R}^s$, $w \in \mathbb{R}^r$, $\mathcal{F} : \mathbb{R}^s \times \mathbb{R}^r \to \mathbb{R}^r$ is continuously differentiable, and $\mathcal{C} \subseteq \mathbb{R}^r$ is a convex polyhedron. Let us denote by $K(\bar{x}, \bar{w})$ the critical cone to $\mathcal{C}$ corresponding to $(\bar{w}, \mathcal{F}(\bar{x}, \bar{w}))$, i. e., the set $K(\bar{x}, \bar{w}) = T_{\mathcal{C}}(\bar{w}) \cap \{\mathcal{F}(\bar{x}, \bar{w})\}^{\perp}$.

Theorem 3.1 (polyhedral feasible sets). Let $(\bar{x}, \bar{w}) \in \mathrm{gph}\, S$. Then

$$\hat{N}_{\mathrm{gph}\,S}(\bar{x}, \bar{w}) \supseteq \left\{ \left[\begin{array}{c} -\nabla_x^T \mathcal{F}(\bar{x}, \bar{w}) v^* \\ u^* - \nabla_w^T \mathcal{F}(\bar{x}, \bar{w}) v^* \end{array} \right] \middle| u^* \in K^-(\bar{x}, \bar{w}), v^* \in K(\bar{x}, \bar{w}) \right\}. \tag{16}$$

If, in addition, (15) is strongly regular at $(\bar{x}, \bar{w})$ and

$$\nabla_w \mathcal{F}(\bar{x}, \bar{w}) K(\bar{x}, \bar{w}) \subseteq \mathrm{Im}\, \nabla_x \mathcal{F}(\bar{x}, \bar{w}), \tag{17}$$

then equality holds true in (16).

Proof. By definition, $S(x)$ consists of those w satisfying (15). Equivalently,

$$\mathrm{gph}\, S = \{(x, w) | G(x, w) \in \mathrm{gph} N_{\mathcal{C}}\} \quad G(x, w) := \left(\begin{array}{c} w \\ -F(x, w) \end{array} \right). \tag{18}$$

By [17, Theorem 6.14], we infer that

$$\hat{N}_{\mathrm{gph}S}(\bar{x}, \bar{w}) \supseteq \nabla^T G(\bar{x}, \bar{w}) \hat{N}_{\mathrm{gph}N_{\mathcal{C}}} G(\bar{x}, \bar{w}).$$

Recalling that, in accordance with our definition of $K(\bar{x}, \bar{w})$,

$$\hat{N}_{\mathrm{gph}N_{\mathcal{C}}}(\bar{w}, -F(\bar{x}, \bar{w})) = K^-(\bar{x}, \bar{w}) \times K(\bar{x}, \bar{w}) \tag{19}$$

(see [1, proof of Theorem 2]), we arrive at inclusion (16).

Suppose now, that (15) is strongly regular at $(\bar{x}, \bar{w})$ and (17) holds true. For readability, we leave off the arguments of the critical cone. As $\mathcal{C}$ is a polyhedron, the strong regularity assumption implies that the Lipschitz localization of S, denoted by $\sigma(x)$, is directionally differentiable at $\bar{x}$ for each $d \in \mathbb{R}^s$. Moreover, one has $\sigma'(\bar{x}; d) = v$, where v is the unique solution of the generalized equation

$$0 \in \nabla_x \mathcal{F}(\bar{x}, \bar{w})d + \nabla_w \mathcal{F}(\bar{x}, \bar{w})v + N_K(v)$$

(see e. g., Theorem 6.3 [13]). We calculate first the contingent cone to $\operatorname{gph} S$:

$$\begin{aligned} & T_{\operatorname{gph} S}(\bar{x}, \bar{w}) \\ = \;& \{(d, v) \in \mathbb{R}^s \times \mathbb{R}^r \,|\, \exists\, \tau_i \searrow 0, (d_i, v_i) \to (d, v) : \bar{w} + \tau_i v_i = \sigma(\bar{x} + \tau_i d_i)\, \forall\, i\} \\ = \;& \{(d, v) \in \mathbb{R}^s \times \mathbb{R}^r \,|\, v = \sigma'(\bar{x}; d)\}, \end{aligned}$$

where the last equality follows from the Lipschitz continuity of σ. Hence,

$$\begin{aligned} & T_{\operatorname{gph} S}(\bar{x}, \bar{w}) \\ = \;& \{(d, v) \in \mathbb{R}^s \times \mathbb{R}^r \,|\, 0 \in \nabla_x \mathcal{F}(\bar{x}, \bar{w})d + \nabla_w \mathcal{F}(\bar{x}, \bar{w})v + N_K(v)\} \\ = \;& \left\{(d, v) \in \mathbb{R}^s \times \mathbb{R}^r \,\middle|\, \begin{bmatrix} v \\ -\nabla_x \mathcal{F}(\bar{x}, \bar{w})d - \nabla_w \mathcal{F}(\bar{x}, \bar{w})v \end{bmatrix} \in \operatorname{gph} N_K \right\}. \end{aligned}$$

Moreover, given K is a convex cone, it is easy to see that

$$\operatorname{gph} N_K = \left\{(v, u) \in K \times K^- \,|\, \langle v, u \rangle = 0\right\}.$$

As a consequence,

$$T_{\operatorname{gph} S}(\bar{x}, \bar{w}) = \left\{(d, v) \in \mathbb{R}^s \times \mathbb{R}^r \,\middle|\, \begin{array}{c} v \in K \\ -\nabla_x \mathcal{F}(\bar{x}, \bar{w})d - \nabla_w \mathcal{F}(\bar{x}, \bar{w})v \in K^- \\ \langle v, \nabla_x \mathcal{F}(\bar{x}, \bar{w})d + \nabla_w \mathcal{F}(\bar{x}, \bar{w})v \rangle = 0 \end{array}\right\}. \quad (20)$$

Now, in order to show the reverse inclusion to (16), let $(d^*, v^*) \in \hat{N}_{\operatorname{gph} S}(\bar{x}, \bar{w}) = [T_{\operatorname{gph} S}(\bar{x}, \bar{w})]^-$ be arbitrary. Then, $\langle d^*, d \rangle + \langle v^*, v \rangle \leq 0$ for all $(d, v) \in T_{\operatorname{gph} S}(\bar{x}, \bar{w})$. Setting $v := 0 \in K$, (20) implies that $\langle d^*, d \rangle \leq 0$ for all d such that $-\nabla_x \mathcal{F}(\bar{x}, \bar{w})d \in K^-$. In other words,

$$d^* \in [[-\nabla_x \mathcal{F}(\bar{x}, \bar{w})]^{-1}(K^-)]^- = -\nabla_x^T \mathcal{F}(\bar{x}, \bar{w})(K).$$

Consequently, there exists a $\bar{u} \in K$ such that

$$d^* = -\nabla_x^T \mathcal{F}(\bar{x}, \bar{w})\bar{u}. \quad (21)$$

Plugging in this information yields that $\langle -\nabla_x^T \mathcal{F}(\bar{x}, \bar{w})\bar{u}, d \rangle + \langle v^*, v \rangle \leq 0$ for all $(d, v) \in T_{\operatorname{gph} S}(\bar{x}, \bar{w})$. Let a $v \in K$ be arbitrarily fixed. By the additional assumption of the Theorem, to this v there exists a d_v such that

$$-\nabla_x \mathcal{F}(\bar{x}, \bar{w})d_v = \nabla_w \mathcal{F}(\bar{x}, \bar{w})v.$$

Therefore, $(d_v, v) \in T_{\operatorname{gph} S}(\bar{x}, \bar{w})$ and the previous inequality yields

$$\langle \bar{u}, \nabla_w \mathcal{F}(\bar{x}, \bar{w})v \rangle + \langle v^*, v \rangle \leq 0.$$

As $v \in K$ is arbitrary, we have $\bar{v} := v^* + \nabla_w^T \mathcal{F}(\bar{x}, \bar{w})\bar{u} \in K^-$. Then along with (21), we get

$$(d^*, v^*) = \begin{bmatrix} 0 & -\nabla_x^T \mathcal{F}(\bar{x}, \bar{w}) \\ I & -\nabla_w^T \mathcal{F}(\bar{x}, \bar{w}) \end{bmatrix} \begin{pmatrix} \bar{v} \\ \bar{u} \end{pmatrix} \in \nabla^T G(\bar{x}, \bar{w})[K^- \times K], \tag{22}$$

where G is defined in (18). Since $(d^*, v^*) \in \hat{N}_{\mathrm{gph}\, S}(\bar{x}, \bar{w})$ was arbitrary, this proves the reverse inclusion to (16). □

We note that the set on the right-hand side of (16) can be easily made fully explicit (similarly to what was done in Prop. 2.2) by exploiting representations of the cartesian product of the critical cone with its polar as provided in [4] or [7]. From Theorem 3.1 we can also immediately derive a setting under which an exact formula for the studied Fréchet normal cone is available:

Corollary 3.2. Under the assumptions of Theorem 3.1, (16) holds as an equality whenever $\nabla_x \mathcal{F}(\bar{x}, \bar{w})$ is surjective, i. e., the controls provide an ample parametrization of (2) ([2]).

On the basis of Theorem 3.1 we arrive now at the following sharp optimality condition for the MPEC

$$\min_{x,w} \{\varphi(x, w) | w \in S(x)\}, \tag{23}$$

where $\varphi[\mathbb{R}^s \times \mathbb{R}^r \to \mathbb{R}]$ is continuously differentiable and S is given by (15).

Theorem 3.3. Let $(\bar{x}, \bar{w})$ be a local solution to (23). Assume that the generalized equation (15) is strongly regular at $(\bar{x}, \bar{w})$ and inclusion (17) is fulfilled. Then $(\bar{x}, \bar{w})$ is S-stationary.

Proof. Clearly, $0 \in \nabla^T \varphi(\bar{x}, \bar{w}) + \hat{N}_{gphS}(\bar{x}, \bar{w})$, where $\hat{N}_{gphS}(\bar{x}, \bar{w})$ amounts to the set on the right-hand side of (16). It remains to recall (19) in order to check that (4) is fulfilled at $(\hat{x}, \hat{z}) := (\bar{x}, \bar{w})$. □

3.2. Nonpolyhedral convex feasible sets

We now use Theorem 3.1 to obtain a similar statement for the solution mapping S to the generalized equation (2), i. e., for settings in which a non-polyhedral convex feasible set is considered. We start by writing the so-called *enhanced generalized equation* associated with (2)

$$0 \in \begin{bmatrix} \mathcal{L}(x, z, \lambda) \\ -A(z) \end{bmatrix} + N_{\mathbb{R}^t \times \mathbb{R}^p_+}(z, \lambda), \tag{24}$$

where, as before,

$$\mathcal{L}(x, z, \lambda) = F(x, z) + \nabla^T A(z)\lambda$$

and λ is a vector of Lagrange multipliers associated with the constraint mapping A. For the enhanced generalized equation, we introduce the *enhanced solution mapping*

$$S^e(x) := \left\{(z, \lambda) \in \mathbb{R}^t \times \mathbb{R}^p \,|(24) \text{ is fulfilled}\right\}. \tag{25}$$

Clearly (24) is of the form (15), where

$$w := (z, \lambda)\,, \quad \mathcal{F}\,(x, w) := \begin{bmatrix} \mathcal{L}(x, z, \lambda) \\ -A(z) \end{bmatrix}, \quad \mathcal{C} := \mathbb{R}^t \times \mathbb{R}^p_+.$$

On the basis of Theorem 3.1 we arrive now at the following statement.

Proposition 3.4 (an inner approximation of $\widehat{N}_{\operatorname{gph} S^e}$). Consider a reference point $(\bar{x}, \bar{z}, \bar{\lambda}) \in \operatorname{gph} S^e$ and assume w.l.o.g that $A(\bar{z}) = 0$ and that (24) is strongly regular at $(\bar{x}, \bar{z}, \bar{\lambda})$. Then,

$$\begin{aligned} \widehat{N}_{\operatorname{gph} S^e}(\bar{x}, \bar{z}, \bar{\lambda}) \quad \supseteq \quad & \{(a, b, c) \in \mathbb{R}^s \times \mathbb{R}^t \times \mathbb{R}^p \mid \exists\, v \in \mathbb{R}^t, u \in \mathbb{R}^{a_+} \times \mathbb{R}^{a_0}_+, \\ & u' \in \{0\} \times \mathbb{R}^{a_0}_- : \\ & a = -\nabla^T_x F(\bar{x}, \bar{z}) v \\ & b = -\nabla^T_z \mathcal{L}(\bar{x}, \bar{z}, \bar{\lambda}) v + \nabla^T A_{I_+}(\bar{z}) u_{I_+} + \nabla^T A_{I_0}(\bar{z}) u_{I_0} \\ & c_{I_+} = -\nabla A_{I_+}(\bar{z}) v \\ & c_{I_0} = u'_{I_0} - \nabla A_{I_0}(\bar{z}) v\} \end{aligned}$$

where $\nabla^T_z \mathcal{L}(\bar{x}, \bar{z}, \bar{\lambda}) v = [\nabla^T_z F(\bar{x}, \bar{z}) + \sum_{i=1}^p \bar{\lambda}_i \nabla^2 A_i(\bar{z})] v$ and

$$\begin{aligned} I_+ \quad &:= \quad \left\{ j \in \{1, \ldots, p\} \,\middle|\, \bar{\lambda}_j > 0 \right\}, \;\; I_0 := \left\{ j \in \{1, \ldots, p\} \,\middle|\, \bar{\lambda}_j = 0 \right\} \\ a_+ \quad &:= \quad |I_+|, \;\; a_0 := |I_0|. \end{aligned}$$

P r o o f. Letting $w = (z, \lambda)$ in Theorem 3.1, it suffices to compute

$$\begin{aligned} K(\bar{x}, \bar{z}, \bar{\lambda}) \quad &= \quad T_{\mathbb{R}^t \times \mathbb{R}^p_+}(\bar{z}, \bar{\lambda}) \cap \begin{bmatrix} \mathcal{L}(\bar{x}, \bar{z}, \bar{\lambda}) \\ -A(\bar{z}) \end{bmatrix}^\perp \\ &= \quad \left\{ (v, u) \in \mathbb{R}^t \times \mathbb{R}^p \,|u_{I_0} \geq 0 \right\} \cap \begin{bmatrix} 0 \\ -A(\bar{z}) \end{bmatrix}^\perp \\ &= \quad \left\{ (v, u) \in \mathbb{R}^t \times \mathbb{R}^p \,|{-A(\bar{z})^T u} = 0, \;\; u_{I_0} \geq 0 \right\} \\ &= \quad \left\{ (v, u) \in \mathbb{R}^t \times \mathbb{R}^p \,|u_{I_0} \geq 0 \right\} \end{aligned}$$

and

$$K^-(\bar{x}, \bar{z}, \bar{\lambda}) = \left\{ (v', u') \in \mathbb{R}^m \times \mathbb{R}^p \,\middle|\, v' = 0, u'_{I_+} = 0, u'_{I_0} \leq 0 \right\}$$

and apply Theorem 3.1 with

$$\begin{aligned} \nabla_x \mathcal{F}\,(\bar{x}, \bar{w}) \quad &= \quad \begin{bmatrix} \nabla_x F\,(\bar{x}, \bar{z}) \\ 0 \end{bmatrix} \\ \nabla_w \mathcal{F}\,(\bar{x}, \bar{w}) \quad &= \quad \begin{bmatrix} \nabla_z F\,(\bar{x}, \bar{z}) + \sum_{i=1}^p \bar{\lambda}_i \nabla^2 A_i(\bar{z}) & \nabla^T A(\bar{z}) \\ -\nabla A(\bar{z}) & 0 \end{bmatrix}. \end{aligned}$$

$\square$

Based on the structure provided by Proposition 3.4, we next compute a similar inner approximation for $\widehat{N}_{\mathrm{gph}\, S}(\bar{x}, \bar{z})$, where S is the solution map associated with (2).

Proposition 3.5 (an inner approximation of $\widehat{N}_{\mathrm{gph}\, S}$). Under the assumptions of Proposition 3.4, we have that

$$\widehat{N}_{\mathrm{gph}\, S}(\bar{x}, \bar{z}) \supseteq \left\{ \begin{bmatrix} \nabla_x^T F(\bar{x}, \bar{z}) v^* \\ \nabla_z^T \mathcal{L}(\bar{x}, \bar{z}, \bar{\lambda}) v^* + \nabla^T A(\bar{z}) w^* \end{bmatrix} \,\middle|\, \begin{array}{r} \nabla A_{I_+}(\bar{z}) v^* = 0 \\ \nabla A_{I_0}(\bar{z}) v^* \geq 0 \\ w_{I_0}^* \geq 0 \end{array} \right\}. \tag{26}$$

P r o o f. Note that due to strong regularity, $\bar{\lambda}$, used to define I_+ and I_0, is the unique multiplier vector associated with the pair $(\bar{x}, \bar{z})$. We claim that

$$\widehat{N}_{\mathrm{gph}\, S}(\bar{x}, \bar{z}) = \left\{ (a, b) \in \mathbb{R}^s \times \mathbb{R}^t \,\middle|\, (a, b, 0) \in \widehat{N}_{\mathrm{gph}\, S^e}(\bar{x}, \bar{z}, \bar{\lambda}) \right\}. \tag{27}$$

Indeed, by Theorem 6.11, [17], one has $(a, b) \in \widehat{N}_{\mathrm{gph}\, S}(\bar{x}, \bar{z})$ if and only if there is a smooth function h that achieves its local maximum relative to $\mathrm{gph}\, S$ at $(\bar{x}, \bar{z})$ and $\nabla h(\bar{x}, \bar{z}) = (a, b)$. Then clearly $(\bar{x}, \bar{z}, \bar{\lambda})$ is a local maximum of the function $\widetilde{h}$ on $\mathrm{gph}\, S^e$, where

$$\widetilde{h}(x, z, \lambda) = h(x, z) \text{ for all } \lambda.$$

Consequently, $(a, b, 0) \in \widehat{N}_{\mathrm{gph}\, S^e}(\bar{x}, \bar{z}, \bar{\lambda})$. For the reverse direction, we appeal to the equivalent definition of the Fréchet normal cone (see e. g., Definition 1.1 [10]), which states

$$(a, b, 0) \in \widehat{N}_{\mathrm{gph}\, S^e}(\bar{x}, \bar{z}, \bar{\lambda}) \Leftrightarrow \limsup_{\substack{(x,z,\lambda) \to (\bar{x}, \bar{z}, \bar{\lambda}) \\ (x,z,\lambda) \in \mathrm{gph}\, S^e \\ (x,z,\lambda) \neq (\bar{x}, \bar{z}, \bar{\lambda})}} \frac{\langle a, x - \bar{x} \rangle + \langle b, z - \bar{z} \rangle + 0}{||(x, z, \lambda) - (\bar{x}, \bar{z}, \bar{\lambda})||} \leq 0.$$

We claim now that this implies $(a, b) \in \widehat{N}_{\mathrm{gph}\, S}(\bar{x}, \bar{z})$. Indeed, due to the strong regularity assumption, both z and λ are single-valued locally Lipschitz functions of x near $\bar{x}$. In particular, it is easy to argue that locally around $(\bar{x}, \bar{z}, \bar{\lambda})$ one has that

$$(x, z) \in \mathrm{gph}\, S \Leftrightarrow (x, z, \lambda(x)) \in \mathrm{gph}\, S^e.$$

Then we may continue the inequality given above as follows

$$0 \geq \limsup_{\substack{(x,z) \to (\bar{x}, \bar{z}) \\ (x,z) \in \mathrm{gph}\, S \\ (x,z) \neq (\bar{x}, \bar{z})}} \frac{\langle a, x - \bar{x} \rangle + \langle b, z - \bar{z} \rangle}{||(x, z, \lambda(x)) - (\bar{x}, \bar{z}, \bar{\lambda})||} \geq \frac{1}{L+1} \limsup_{\substack{(x,z) \to (\bar{x}, \bar{z}) \\ (x,z) \in \mathrm{gph}\, S \\ (x,z) \neq (\bar{x}, \bar{z})}} \frac{\langle a, x - \bar{x} \rangle + \langle b, z - \bar{z} \rangle}{||x - \bar{x}|| + ||z - \bar{z}||}$$

where L is the Lipschitz modulus of $\lambda(x)$. Thus by definition, $(a, b) \in \widehat{N}_{\mathrm{gph}\, S}(\bar{x}, \bar{z})$, which proves (27). Then the asserted formula follows immediately from Proposition 3.4. □

4. USING M-STATIONARITY CONDITIONS TO OBTAIN STRONG STATIONARITY CONDITIONS

We can now use the results of the previous section to obtain a new condition ensuring that a local minimizer is S-stationary for (1).

Let the assumptions of Theorem 3.1 be fulfilled. Then one could adapt the notion of S-stationarity with respect to the structure of C as follows.

A feasible point $(\bar{x}, \bar{z})$ to the MPEC (1) is S-stationary, assuming $A(\bar{z}) = 0$, if there exist $\bar{\lambda} \in \mathbb{R}^p_+$ and $(v^*, w^*) \in \mathbb{R}^t \times \mathbb{R}^p$ such that

$$-\nabla_x^T f(\bar{x}, \bar{z}) = \nabla_x^T F(\bar{x}, \bar{z}) v^* \tag{28}$$

$$-\nabla_z^T f(\bar{x}, \bar{z}) = \nabla_z^T F(\bar{x}, \bar{z}) v^* + \left(\sum_{j=1}^{p} \bar{\lambda}_j \nabla^2 A_j(\bar{z}) \right) v^* + \nabla^T A(\bar{z}) w^* \tag{29}$$

$$\nabla A_{I_+}(\bar{z}) v^* = 0 \tag{30}$$

$$\nabla A_{I_0}(\bar{z}) v^* \geq 0 \tag{31}$$

$$w^*_{I_0} \geq 0 \tag{32}$$

$$F(\bar{x}, \bar{z}) = -\nabla^T A(\bar{z}) \bar{\lambda}. \tag{33}$$

It is clear that a solution $(\bar{x}, \bar{z})$ to (1) need not be S-stationary, provided (26) holds as a proper inclusion. Nevertheless, under an additional condition, $(\bar{x}, \bar{z})$ will be S-stationary as shown in the next statement via comparison of the M-stationarity conditions in Proposition 2.1 with inclusion (26).

Theorem 4.1 (using M-stationarity to obtain S-stationarity). Let $(\bar{x}, \bar{z})$ be a local solution to (1) and assume without loss of generality that $A(\bar{z}) = 0$. Assume that

1. $\nabla A(\bar{z})$ is surjective;
2. SSOSC holds at $(\bar{x}, \bar{z})$;
3. there is a v^* satisfying (9), (10) such that $\nabla A_{I_0}(\bar{z}) v^* > 0$.

Then $(\bar{x}, \bar{z})$ is S-stationary. In particular, there exists a unique $\bar{\lambda} \in \mathbb{R}^p_+$ and $(v^*, w^*) \in \mathbb{R}^t \times \mathbb{R}^p$ such that

$$-\nabla_x^T f(\bar{x}, \bar{z}) = \nabla_x^T F(\bar{x}, \bar{z}) v^*$$

$$-\nabla_z^T f(\bar{x}, \bar{z}) = \nabla_z^T F(\bar{x}, \bar{z}) v^* + \left(\sum_{j=1}^{p} \bar{\lambda}_j \nabla^2 A_j(\bar{z}) \right) v^* + \nabla^T A(\bar{z}) w^*$$

$$\nabla A_{I_+}(\bar{z}) v^* = 0, \quad \nabla A_{I_0}(\bar{z}) v^* > 0, \quad w^*_{I_0} \geq 0, \quad F(\bar{x}, \bar{z}) = -\nabla^T A(\bar{z}) \bar{\lambda}.$$

Proof. Given assumptions 1. and 2., we know (24) is strongly regular at $(\bar{x}, \bar{z}, \bar{\lambda})$ (see Theorem 4.1 [16]). Thus, the inclusion (26) holds and we can write down the

S-stationarity conditions in the form (28) – (33). In addition, assumptions 1. and 2. also imply that the perturbation mapping associated with (2) has the Aubin property at $(\bar{x}, \bar{z})$ (see Corollary 5.1 in [19] in general or Proposition 5.1 in [6] for a special case). Then by Proposition 2.2, there exists a unique $\bar{\lambda} \in \mathbb{R}^p_+$ and multipliers $(v^*, w^*) \in \mathbb{R}^t \times \mathbb{R}^p$ such that (9) – (14) hold. At this point, it is clear that relations (9) – (11),(14) are equivalent to (28) – (30),(33). Finally, by assumption 3., we observe that (12) vanishes and (13) becomes equivalent to $w^*_{I_0} \geq 0$ and $\nabla A_{I_0}(\bar{z})v^* > 0$. Therefore, the pair (v^*, w^*) fulfills (together with $\bar{x}, \bar{z}, \bar{\lambda}$) the S-stationarity conditions and we are done. □

Note that if strict complementarity were to hold, then the S-stationarity conditions would always coincide with the M-stationarity conditions, in which case M- and S-stationarity become equivalent conditions, as would be expected.

The interested reader can find an example in which all three conditions of the previous theorem hold in [19] (Example 8.6).

ACKNOWLEDGEMENT

The work of the first author was supported by the DFG Research Center Matheon "Mathematics for key technologies" in Berlin. The work of the second author was supported by the Grant Agency of the Czech Academy of Sciences (Grant No. IAA 100750802).

(Received April 14, 2010)

REFERENCES

[1] A. L. Dontchev and R. T. Rockafellar: Characterizations of strong regularity for variational inequalities over polyhedral convex sets. SIAM J. Optim. *6* (1996), 1087–1105.

[2] A. L. Dontchev and R. T. Rockafellar: Ample parameterization of variational inclusions. SIAM J. Optim. *12* (2001), 170–187.

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

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

[5] R. Henrion, J. V. Outrata, and T. Surowiec: On the coderivative of normal cone mappings to inequality systems. Nonlinear Anal. *71* (2009), 1213–1226.

[6] R. Henrion, J. V. Outrata, and T. Surowiec: Analysis of M-stationary points to an EPEC modeling oligopolistic competition in an electricity spot market. Weierstraß-Institute of Applied Analysis and Stochastics, Preprint No. *1433* (2009) and submitted.

[7] R. Henrion, B. S. Mordukhovich, and N. M. Nam: Second-order analysis of polyhedral systems in finite and infinite dimensions with applications to robust stability of variational inequalities. SIAM J. Optim., to appear.

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

[9] B. S. Mordukhovich: Approximation Methods in Problems of Optimization and Control (in Russian). Nauka, Moscow 1988.

[10] B. S. Mordukhovich: Variational Analysis and Generalized Differentiation. Vol. 1: Basic Theory. Springer, Berlin 2006.

[11] B. S. Mordukhovich: Variational Analysis and Generalized Differentiation. Vol. 2: Applications. Springer, Berlin 2006.

[12] M. L. Flegel, C. Kanzow, and J. V. Outrata: Optimality conditions for disjunctive programs with application to mathematical programs with equilibrium constraints. Set-Valued Anal. *15* (2007), 139–162.

[13] J. V. Outrata, M. Kočvara, and J. Zowe: Nonsmooth Approach to Optimization Problems with Equilibrium Constraints. Kluwer Academic Publishers, Dordrecht 1998.

[14] J.-S. Pang and M. Fukushima: Complementarity constraint qualifications and simplified B-stationarity conditions for mathematical programs with equilibrium constraints. Comput. Optim. Appl. *13* (1999), 111–136.

[15] S. M. Robinson: Some continuity properties of polyhedral multifunctions. Math. Program. Studies *14* (1976), 206–214.

[16] S. M. Robinson: Strongly regular generalized equations. Math. Oper. Res. *5* (1980), 43–62.

[17] R. T. Rockafellar and R. J.-B. Wets: Variational Analysis. Springer, Berlin 1998.

[18] H. Scheel and S. Scholtes: Mathematical programs with complementarity constraints: Stationarity, optimality and sensitivity. Math. Oper. Res. *25* (2000), 1–22.

[19] T. Surowiec: Explicit Stationarity Conditions and Solution Characterization for Equilibrium Problems with Equilibrium Constraints. Doctoral Thesis, Humboldt University, Berlin 2009.

René Henrion, Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstr. 39, 10117 Berlin. Germany.
e-mail: henrion@wias-berlin.de

Jiří Outrata, Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Praha 8. Czech Republic.
e-mail: outrata@utia.cas.cz

Thomas Surowiec, Humboldt University Berlin, Unter den Linden 6, 10099 Berlin. Germany.
e-mail: surowiec@math.hu-berlin.de