Control and Cybernetics

vol. **38** (2009) No. 4B

On the implicit programming approach in a class of mathematical programs with equilibrium constraints*†

by

Jiří V. Outrata and Michal Červinka

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

Abstract: In the paper we analyze the influence of implicit programming hypothesis and presence of state constraints on first order optimality conditions to mathematical programs with equilibrium constraints. In the absence of state constraints, we derive sharp stationarity conditions, provided the strong regularity condition holds. In the second part of the paper we suggest an exact penalization of state constraints and test the behavior of standard bundle trust region algorithm on academic examples.

Keywords: mathematical problem with equilibrium constraint, state constraints, implicit programming, calmness, exact penalization.

1. Introduction

Consider the optimization problem

$$
\begin{aligned}
&\text{minimize } f(x,y)\\
&\text{subject to}\\
&\qquad y \in S(x)\\
&\qquad x \in \omega\\
&\qquad y \in \Xi,
\end{aligned}
\tag{1}
$$

where $x \in \mathbb{R}^n$ is the *control*, $y \in \mathbb{R}^m$ is the *state variable*, $f[\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}]$ is a locally Lipschitz objective, $\omega \subset \mathbb{R}^n$ and $\Xi \subset \mathbb{R}^m$ are nonempty closed sets of *admissible* controls and states, respectively, and $S[\mathbb{R}^n \rightrightarrows \mathbb{R}^m]$ is a closed-graph multifunction. This multifunction represents in (1) the so-called equilibrium

*This research was supported by the Grant Agency of the Czech Republic under the grant 201/09/1957.

†Submitted: March 2009; Accepted: September 2009.

constraint and is usually defined by a control-dependent variational inequality in variable y. Problems of the type (1) are called *mathematical programs with equilibrium constraints* (MPECs) and for S we use the term *solution map*. MPECs have have been investigated in the monographs of Luo, Pang and Ralph (1996), Outrata, Kočvara and Zowe (1998), Dempe (2002) and in numerous papers. In Ye and Ye (1997), Outrata (1999, 2000), Ye (2000) and some other papers the authors used the Mordukhovich theory of generalized differentiation (Mordukhovich, 1988, 2006) to derive efficient optimality conditions for (1). Optimality conditions for an infinite-dimensional version of (1) can be found, e.g., in Mordukhovich (2006), Chapter 5. In the approach applied in these works neither the multi-valuedness of the solution map S nor the presence of geometric constraints $x \in \omega, y \in \Xi$ cause substantial difficulties. One has to impose, however, specific *constraint qualifications* requiring either the Aubin property or calmness of special multifunctions, reflecting the structure of the overall constraint system.

In numerous important MPECs coming, e.g., from continuum mechanics, the solution map S is (at least locally) single-valued and locally Lipschitz, see Outrata, Kočvara and Zowe (1998). In this case, in agreement with Luo, Pang and Ralph (1996), we will say that the so-called *implicit programming hypothesis* (ImP-hypothesis) is fulfilled. Problem (1) becomes then a special optimal control problem, where S specifies the behavior of the controlled system. The ImP-hypothesis enables us to apply to the derivation of optimality conditions and to the numerical solution of (1) various techniques of the so-called *ImP-approach*, based essentially on the implicit function argument, see Luo, Pang and Ralph (1996), Outrata, Kočvara and Zowe (1998). They are particularly efficient in the absence of state constraints.

The aim of this paper is to analyze the impact of the ImP-hypothesis on the optimality conditions and to examine exact penalization of state constraints in the framework of a standard ImP numerical technique. Notice that besides the mechanical equilibria, mentioned above (where the ImP-hypothesis can be verified directly), the ImP-hypothesis is usually ensured by the Robinson's strong regularity or the strong second order sufficient condition, see Robinson (1980).

The plan is as follows. In Section 2 we first show that ImP-hypothesis has a positive influence only on the imposed constraint qualification and not on the stationarity condition itself. Then we replace the abstract equilibrium constraint $y \in S(x)$ by a parameterized variational inequality and ensure the ImP-hypothesis by the Robinson's strong regularity. In this setting, under some additional assumptions, we are able to derive new, sharp optimality conditions.

In Section 3 we suggest to penalize the state constraints under appropriate assumptions on S by the nonsmooth composition $d_\Xi \circ S$, where d_Ξ denotes a suitable distance function. We show that the exactness of this penalty can be ensured by a constraint qualification, playing a crucial role already in optimality conditions. The behavior of the proposed method is tested by academic examples.

The following notation is employed. $\mathbb{B}$ denotes the unit ball, $d_\Omega(\cdot)$ is the distance function to a set Ω and, for a closed cone D with vertex at the origin, D° denotes its negative polar cone. By $x \xrightarrow{\Omega} \bar{x}$ we mean that $x \to \bar{x}$ with $x \in \Omega$. $T_\Omega(x)$ denotes the contingent (Bouligand-Severi) cone to Ω at x and $P_\Omega(x)$ is the metric projection of x onto the closure of Ω. For a real-valued function f we use the notation $\operatorname{epi} f$ and $\bar{\partial} f(x)$ to denote its epigraph and the Clarke subdifferential of f at x, see Clarke (1983), respectively. If f is vector-valued, $\bar{\partial} f(x)$ stands for the generalized Jacobian of Clarke at x.

For the readers' convenience we state now the definitions of several basic notions from modern variational analysis.

For a set Ω and a point $\bar{x} \in \mathrm{cl}\Omega$, the *Fréchet normal cone* to Ω at $\bar{x}$ is defined by

$$\widehat{N}_\Omega(\bar{x}) := \left\{ x^* \in \mathbb{R}^n \,\middle|\, \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle x^*, x - \bar{x} \rangle}{\| x - \bar{x} \|} \leq 0 \right\}.$$

The *limiting normal cone* to Ω at $\bar{x}$ is given by

$$N_\Omega(\bar{x}) = \operatorname*{Lim\,sup}_{x \xrightarrow{\Omega} \bar{x}} \widehat{N}_\Omega(x),$$

where the "Lim sup" stands for the Painlevé-Kuratowski upper (or outer) limit. This limit is defined for a set-valued mapping $F[\mathbb{R}^n \rightrightarrows \mathbb{R}^m]$ at a point $\bar{x}$ by

$$\operatorname*{Lim\,sup}_{x \to \bar{x}} F(x) := \{ y \in \mathbb{R}^m \mid \exists x_k \to x, \exists y_k \to y \text{ with } y_k \in F(x_k) \}.$$

For a convex set Ω, both normal cones N_Ω and $\widehat{N}_\Omega$ reduce to the normal cone of convex analysis, for which we use simply the notation N_Ω.

For a function $f[\mathbb{R}^n \to \mathbb{R}]$, and a point $\bar{x} \in \mathbb{R}^n$, the set

$$\partial f(\bar{x}) = \{ y \in \mathbb{R}^n \mid (y, -1) \in N_{\operatorname{epi} f}(\bar{x}, f(\bar{x})) \}$$

is the *limiting subdifferential* of f at $\bar{x}$.

Given a set-valued mapping $F[\mathbb{R}^n \rightrightarrows \mathbb{R}^m]$ and a point $(\bar{x}, \bar{y})$ from its graph

$$\mathrm{Gph} F := \{ (x, y) \in \mathbb{R}^n \times \mathbb{R}^m | y \in F(x) \},$$

the *Fréchet coderivative* $\widehat{D}^* F(\bar{x}, \bar{y})[\mathbb{R}^m \rightrightarrows \mathbb{R}^n]$ of F at $(\bar{x}, \bar{y})$ is defined by

$$\widehat{D}^* F(\bar{x}, \bar{y})(y^*) := \{ x^* \in \mathbb{R}^n | (x^*, -y^*) \in \widehat{N}_{\mathrm{Gph} F}(\bar{x}, \bar{y}) \},$$

and the *(limiting) coderivative* $D^* F(\bar{x}, \bar{y})[\mathbb{R}^m \rightrightarrows \mathbb{R}^n]$ of F at $(\bar{x}, \bar{y})$ is defined by

$$D^* F(\bar{x}, \bar{y})(y^*) := \{ x^* \in \mathbb{R}^n | (x^*, -y^*) \in N_{\mathrm{Gph} F}(\bar{x}, \bar{y}) \}.$$

When F is single-valued at $\bar{x}$, we omit $\bar{y}$ in the notation $\widehat{D}^* F(\bar{x}, \bar{y})$ or $D^* F(\bar{x}, \bar{y})$.

Finally, throughout the paper we use the notion of calmness. A set-valued mapping $F[\mathbb{R}^n \rightrightarrows \mathbb{R}^m]$ is said to be *calm* at $(\bar{x}, \bar{y}) \in \operatorname{Gph} F$ with modulus $L \geq 0$ if there are neighborhoods $\mathcal{U}$ of $\bar{x}$ and $\mathcal{V}$ of $\bar{y}$ such that

$$F(x) \cap \mathcal{V} \subset F(\bar{x}) + L\|x - \bar{x}\|\mathbb{B} \quad \text{for all } x \in \mathcal{U}.$$

2. Analysis of optimality conditions

Consider first the MPEC (1) without any additional assumptions. From Henrion, Jourani and Outrata (2002), Theorem 4.1, we get readily the following result.

THEOREM 1 *Let $(\hat{x}, \hat{y})$ be a (local) solution of (1) and assume that the* perturbation map $M[\mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n \times \mathbb{R}^m]$, *defined by*

$$M(p_1, p_2) := \{(x, y) \in \operatorname{Gph} S \mid x - p_1 \in \omega, y - p_2 \in \Xi\}, \tag{2}$$

is calm at $(0, 0, \hat{x}, \hat{y})$.

Then there is a subgradient $(\xi, \gamma) \in \partial f(\hat{x}, \hat{y})$ and a normal vector $\hat{\eta} \in N_\Xi(\hat{y})$ such that

$$0 \in \xi + D^*S(\hat{x}, \hat{y})(\gamma + \hat{\eta}) + N_\omega(\hat{x}). \tag{3}$$

Proof. It suffices to put $z := (x, y)$ and apply the mentioned result from Henrion, Jourani and Outrata (2002) to the computation of an upper estimate of the limiting normal cone $N_{\operatorname{Gph} S \cap (\omega \times \Xi)}$ at $\hat{z} = (\hat{x}, \hat{y})$. On the basis of Mordukhovich (1988), Theorem 7.1, we arrive at the optimality condition

$$0 \in \partial f(\hat{z}) + N_{\operatorname{Gph} S}(\hat{z}) + N_{\omega \times \Xi}(\hat{z}),$$

which immediately implies (3). ∎

The ImP-hypothesis amounts to the existence of neighborhoods $\mathcal{P}$ of $\hat{x}$, $\mathcal{Q}$ of $\hat{y}$ and a Lipschitz function $s[\mathcal{P} \to \mathbb{R}^m]$ such that $s(\hat{x}) = \hat{y}$ and

$$S(x) \cap \mathcal{Q} = \{s(x)\} \quad \text{for all } x \in \mathcal{P}.$$

In agreement with Rockafellar and Wets (1998), s will be called a Lipschitz localization of S around $(\hat{x}, \hat{y})$. So, the ImP-hypothesis enables us to replace (1) locally around $(\hat{x}, \hat{y})$ by a nonlinear program in variable x only:

$$\begin{aligned} &\text{minimize } \theta(x) \\ &\text{subject to} \\ &\qquad x \in \omega \\ &\qquad s(x) \in \Xi, \end{aligned} \tag{4}$$

where $\theta(x) := f(x, s(x))$. This reformulation leads to the following statement.

Theorem 2 *Let the ImP-hypothesis be fulfilled around* $(\hat{x},\hat{y})$*, where* $\hat{x}$ *is a (local) solution of (4) and* $\hat{y}=s(\hat{x})$*. Further, assume that the map* $\widetilde{M}[\mathbb{R}^m \rightrightarrows \mathbb{R}^n]$ *defined by*

$$\widetilde{M}(q) := \{x \in \omega \mid s(x) - q \in \Xi\} \tag{5}$$

is calm at $(0,\hat{x})$.

Then there is a subgradient $(\xi,\gamma) \in \partial f(\hat{x},\hat{y})$ *and a normal vector* $\hat{w} \in N_\Xi(\hat{y})$ *such that*

$$0 \in \xi + D^*S(\hat{x},\hat{y})(\gamma) + D^*S(\hat{x},\hat{y})(\hat{w}) + N_\omega(\hat{x}). \tag{6}$$

Proof. Clearly, problem (4) amounts to

$$\begin{aligned} &\text{minimize } \theta(x)\\ &\text{subject to}\\ &\qquad\qquad x \in A, \end{aligned} \tag{7}$$

where $A := \omega \cap s^{-1}(\Xi)$. By virtue of Mordukhovich (1988), Theorem 7.1, one has

$$0 \in \partial\theta(\hat{x}) + N_A(\hat{x}).$$

From Rockafellar and Wets (1998), Theorem 10.49, we have the estimate

$$\partial\theta(\hat{x}) \subset \bigcup \left\{ \xi + D^*s(\hat{x})(\gamma) \middle| (\xi,\gamma) \in \partial f(\hat{x},\hat{y}) \right\},$$

and by Henrion, Jourani and Outrata (2002), Theorem 4.1, we infer that

$$N_A(\hat{x}) \subset N_\omega(\hat{x}) + \bigcup \left\{ D^*s(\hat{x})(w) \middle| w \in N_\Xi(S(\hat{x})) \right\}.$$

Since $D^*s(\hat{x})$ amounts to $D^*S(\hat{x},\hat{y})$ by the definition of the coderivative, the statement has been proved. ∎

Remark 1 *The ImP-hypothesis is in the literature traditionally formulated in the form which we have adopted here. As pointed out by one of the reviewers, however, it could be weakened by assuming merely the existence of a* Hőlder localization *of* S *around a (local) solution. Since* θ *is then not necessarily Lipschitz, this would require imposing additionally the qualification condition*

$$\partial^\infty\theta(\hat{x}) \cap (-N_A(\hat{x})) = \{0\}$$

in Theorem 2, where for the singular subdifferential $\partial^\infty\theta(\hat{x})$ *we have the estimate*

$$\partial^\infty\theta(\hat{x}) \subset D^*s(\hat{x})(0),$$

at our disposal, see Mordukhovich (2006), Theorem 3.38 (iv). Further, one could not directly apply bundle methods in the associated numerical approach discussed in Section 3. Nevertheless, such a weakened ImP-hypothesis still brings a lot of structure into the problem and definitely deserves a careful analysis.

Under the ImP-hypothesis we are entitled to compare the statements of Theorems 1 and 2.

PROPOSITION 1 *Let the ImP-hypothesis hold true around the point* $(\hat{x}, \hat{y})$ *which fulfills condition (3). Then* $(\hat{x}, \hat{y})$ *satisfies condition (6) as well.*

Proof. Let $c \in D^*S(\hat{x}, \hat{y})(a + b) = D^*s(\hat{x})(a + b)$ for some arbitrary vectors $a, b \in \mathbb{R}^m$. Assume, by contradiction, that for any $x_1^{(i)} \to \hat{x}, a^{(i)} \to a, c_1^{(i)} \in \widehat{D}^*s(x_1^{(i)})(a^{(i)})$ and for any $x_2^{(i)} \to \hat{x}, b^{(i)} \to b, c_2^{(i)} \in \widehat{D}^*s(x_2^{(i)})(b^{(i)})$ one has that $c_1^{(i)} + c_2^{(i)}$ does not converge to c. This holds of course even more if we require $x_1^{(i)} = x_2^{(i)}$ $\forall i$. Then, however, by convexity of the regular normal cone one has

$$c_1^{(i)} + c_2^{(i)} \in \widehat{D}^*s(x_1^{(i)})(a^{(i)} + b^{(i)}),$$

which contradicts the relation $c \in D^*s(\hat{x})(a + b)$ and we are done. ■

The reverse inclusion, however, does not hold as shown in the next example.

EXAMPLE 1 *Suppose that* $S(x) = s(x) = -|x|, (\hat{x}, \hat{y}) = (0, 0)$ *and* $\gamma = \hat{\eta} = 1$. *Then*

$$0 \in D^*S(\hat{x})(\gamma) + D^*S(\hat{x})(\hat{\eta}),$$

while

$$0 \notin D^*S(\hat{x})(\gamma + \hat{\eta}).$$

It follows that condition (3) is not less sharp (selective) than (6). Moreover, (3) is definitely more workable, because only one value of $D^*S(\hat{x}, \hat{y})$ has to be computed. Next we compare the calmness qualification conditions related to the multifunctions (2) and (5).

PROPOSITION 2 *Let the ImP-hypothesis be fulfilled around* $(\hat{x}, \hat{y})$. *Then the following two properties are equivalent.*

(i) M *is calm at* $(0, 0, \hat{x}, \hat{y})$;

(ii) $\widetilde{M}$ *is calm at* $(0, \hat{x})$.

Proof. (i) ⇒ (ii)

Clearly, the calmness of M at $(0, 0, \hat{x}, \hat{y})$ is equivalent to the calmness of the (localized) multifunction M^ℓ given by

$$M^\ell(p_1, p_2) = \{(x, y) \in \text{Gph}s \,|\, x - p_1 \in \omega, y - p_2 \in \Xi\}$$

at the same point (one just has to appropriately shrink the neighborhoods $\mathcal{U}, \mathcal{V}$ in the definition of calmness). Evidently,

$$M^\ell(p_1, p_2) = \{(x, y) \in \text{Gph}s \,|\, x \in M_1(p_1, p_2)\},$$

where

$$M_1(p_1,p_2) := \{x \,|\, x - p_1 \in \omega, s(x) - p_2 \in \Xi\}.$$

Since s is single-valued and Lipschitz, the calmness of M^ℓ at $(0,0,\hat{x},\hat{y})$ is equivalent to the calmness of M_1 at $(0,0,\hat{x})$. Finally, it is clear that the calmness of M_1 at $(0,0,\hat{x})$ implies the calmness of $\widetilde{M}$ at $(0,\hat{x})$, and we are done.

(ii) $\Rightarrow$ (i)
Taking into account the above mentioned equivalences, assume by contradiction the existence of sequences

$$x^{(i)} \to \hat{x}, (p_1^{(i)}, p_2^{(i)}) \to (0,0) \text{ with } x^{(i)} \in M_1(p_1^{(i)}, p_2^{(i)})$$

such that

$$d_{M_1(0,0)}(x^{(i)}) \geq i(\|p_1^{(i)}\| + \|p_2^{(i)}\|) \;\; \forall\, i.$$

Put $\tilde{x}^{(i)} := x^{(i)} - p_1^{(i)}$ and observe that, due to

$$s(x^{(i)}) - s(\tilde{x}^{(i)}) + s(\tilde{x}^{(i)}) - p_2^{(i)} \in \Xi,$$

one has $s(\tilde{x}^{(i)}) - q^{(i)} \in \Xi$ with $q^{(i)} = s(\tilde{x}^{(i)}) - s(x^{(i)}) + p_2^{(i)}$. By the Lipschitz continuity of s

$$\|q^{(i)}\| \leq \ell\|\tilde{x}^{(i)} - x^{(i)}\| + \|p_2^{(i)}\| = \ell\|p_1^{(i)}\| + \|p_2^{(i)}\| \leq \max\{\ell,1\}(\|p_1^{(i)}\| + \|p_2^{(i)}\|),$$

where ℓ is the Lipschitz constant of s. It follows that

$$d_{\widetilde{M}(0)}(\tilde{x}^{(i)}) \geq d_{M_1(0,0)}(x^{(i)}) - \|p_1^{(i)}\| \geq (i-1)(\|p_1^{(i)}\| + \|p_2^{(i)}\|) \geq \frac{i-1}{\max\{\ell,1\}}\|q^{(i)}\|,$$

whence contradiction with the calmness of $\widetilde{M}$ at $(0,\hat{x})$. The result has been established. ∎

We conclude that the ImP-hypothesis does not enable us to sharpen the optimality condition (3) itself, but instead of the calmness of M we can verify the calmness of a simpler perturbation multifunction $\widetilde{M}$.

Condition (3) is clearly useful only in the case, when we are able to compute the coderivative $D^*S(\hat{x},\hat{y})$ or its tight upper estimate. Then, of course, a certain structure of the equilibrium constraint has to be given. The coderivative $D^*S(\hat{x},\hat{y})$ (or its upper estimate) can then be used also in some available calmness criteria (Henrion and Outrata, 2001; Henrion, Jourani and Outrata, 2002; Ioffe and Outrata, 2008).

Next, we will suppose that

$$S(x) := \{y \in \mathbb{R}^m \,|\, 0 \in F(x,y) + N_\Gamma(y)\}, \tag{8}$$

where $\Gamma \subset \mathbb{R}^m$ is a convex polyhedron and $F[\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m]$ is continuously differentiable. Moreover, at the (local) solution pair $(\hat{x}, \hat{y})$ we will impose the Robinson's *strong regularity condition* (SRC) requiring the existence of neighborhoods $\mathcal{U}$ of $0 \in \mathbb{R}^m$ and $\mathcal{V}$ of $\hat{y}$ such that the mapping

$$p \mapsto \{y \in \mathcal{V} \,|\, p \in F(\hat{x}, \hat{y}) + \nabla_y F(\hat{x}, \hat{y})(y - \hat{y}) + N_\Gamma(y)\}$$

is single-valued and Lipschitz on $\mathcal{U}$. It is well known (Robinson, 1987, 1991) that in this situation the ImP-hypothesis holds true. Further, the respective Lipschitz localization s is directionally differentiable at $\hat{x}$, which will be helpful in the next development. We recall that for an arbitrary direction $d \in \mathbb{R}^n$ the directional derivative $s'(\hat{x}; d)$ amounts to the (unique) solution v of the affine generalized equation

$$0 \in \nabla_x F(\hat{x}, \hat{y})d + \nabla_y F(\hat{x}, \hat{y})v + N_{K(\hat{x})}(v), \tag{9}$$

where $K(\hat{x}) := T_\Gamma(\hat{y}) \cap (F(\hat{x}, \hat{y}))^\perp$ is the *critical cone* to Γ with respect to $\hat{y}$ and $F(\hat{x}, \hat{y})$. Under the imposed assumptions we derive the following optimality conditions.

THEOREM 3 *Consider the MPEC (1) with S given by (8). Assume that $(\hat{x}, \hat{y})$ is its (local) solution, and*

(i) f is continuously differentiable,
(ii) Γ is a convex polyhedron,
(iii) $\Xi = \mathbb{R}^m$ (i.e. no state constraints), and
(iv) SRC holds at $(\hat{x}, \hat{y})$.

Then there are multipliers $\hat{u}, \hat{b}$ with

$$\begin{pmatrix} \hat{u} \\ \hat{b} \end{pmatrix} \in N_{\mathrm{Gph} N_{K(\hat{x})}}(0, 0)$$

such that

$$\begin{aligned} &0 \in \nabla_x f(\hat{x}, \hat{y}) - (\nabla_x F(\hat{x}, \hat{y}))^\top \hat{b} + N_\omega(\hat{x}) \\ &0 = \nabla_y f(\hat{x}, \hat{y}) + \hat{u} - (\nabla_y F(\hat{x}, \hat{y}))^\top \hat{b}. \end{aligned} \tag{10}$$

Proof. Denote by Λ the constraint set in the considered MPEC, i.e.,

$$\Lambda = \mathrm{Gph} S \cap (\omega \times \mathbb{R}^m).$$

By virtue of the directional differentiability of s it is easy to show that

$$T_\Lambda(\hat{x}, \hat{y}) = \{(d, v) \in T_\omega(\hat{x}) \times \mathbb{R}^m \,|\, v = s'(\hat{x}; d)\}.$$

It follows from the (local) optimality of $(\hat{x}, \hat{y})$ that (0,0) is a solution of the "linearized" MPEC

$$\begin{aligned} &\text{minimize } \langle \nabla_x f(\hat{x}, \hat{y}), d\rangle + \langle \nabla_y f(\hat{x}, \hat{y}), v\rangle \\ &\text{subject to} \\ &\qquad 0 \in \nabla_x F(\hat{x}, \hat{y})d + \nabla_y F(\hat{x}, \hat{y})v + N_{K(\hat{x})}(v) \\ &\qquad d \in T_\omega(\hat{x}) \end{aligned} \tag{11}$$

in variables (d, v). Due to assumptions (ii) and (iv), the generalized equation in (11) satisfies SRC at (0,0), see Outrata, Kočvara and Zowe (1998), Theorem 5.3. We can thus invoke Outrata (2000), Proposition 3.2, according to which the classical MPEC constraint qualification

$$\left.\begin{array}{r}\begin{pmatrix} 0 & -(\nabla_x F(\hat{x},\hat{y}))^\top \\ Id & -(\nabla_y F(\hat{x},\hat{y}))^\top \end{pmatrix}\begin{pmatrix} u \\ b \end{pmatrix} \in -N_{T_\omega(\hat{x})}(0) \times \{0\} \\ (u,b) \in N_{\mathrm{Gph} N_{K(\hat{x})}}(0,0)\end{array}\right\} \Rightarrow \begin{cases} u = 0 \\ b = 0 \end{cases}$$

holds. This enables us now to apply Outrata (2000), Theorem 3.1, to the MPEC (11), which yields the existence of multipliers $(\hat{u}, \hat{b}) \in N_{\mathrm{Gph} N_{K(\hat{x})}}(0,0)$ such that relations (10) hold with $N_\omega(\hat{x})$ replaced by $N_{T_\omega(\hat{x})}(0)$. It remains to observe that by virtue of Rockafellar and Wets (1998), Theorem 6.27 (a),

$$N_{T_\omega(\hat{x})}(0) \subset N_\omega(\hat{x}),$$

and so the statement has been established. ∎

From the above proof it is clear that the statement holds even for nonpolyhedral sets Γ under the assumptions that, in addition to (i), (iii) and (iv),

- $K(\hat{x})$ is a polyhedral cone;
- s is directionally differentiable at $\hat{x}$ with

  $$s'(\hat{x}; d) = (G \circ \nabla_x F(\hat{x},\hat{y}))(d),$$

  where G is a single-valued map from $\mathbb{R}^m$ to $\mathbb{R}^m$ defined by

  $$G(h) := \{v \in \mathbb{R}^m | 0 \in h + Z(\hat{x})v + N_{K(\hat{x})}(v)\}$$

  and $Z(\hat{x})$ is an $m \times m$ matrix.

The above assumptions enable us to apply Outrata (2000), Proposition 3.2 and Theorem 3.1, exactly as in the fully polyhedral case. Indeed, from the single-valuedness of G and the polyhedrality of $K(\hat{x})$ it follows that the generalized equation

$$0 \in \nabla_x F(\hat{x},\hat{y})d + Z(\hat{x})v + N_{K(\hat{x})}(v)$$

satisfies SRC at $(0,0)$, because a single-valued polyhedral mapping is Lipschitz.

Such a situation appears, e.g., in Ralph and Dempe (1995), Corollary 4, statement (2), or in Luo, Pang and Ralph (1996), Theorem 4.2.25, where

$$\Gamma = \{y \in \mathbb{R}^m | q^i(y) \le 0, i = 1, \ldots, l\}$$

with $q^i[\mathbb{R}^m \to \mathbb{R}], i = 1, \ldots, l$, convex and twice continuously differentiable. Moreover, one requires that Mangasarian Fromowitz and constant rank constraint qualifications are fulfilled at the reference point. In such a case,

$$Z(\hat{x}) = \nabla_y F(\hat{x},\hat{y}) + \sum_{i=1}^{l} \lambda^i \nabla^2 q^i(\hat{y}),$$

where $\lambda \in \mathbb{R}^l_+$ is *any* Lagrange multiplier associated with $(\hat{x}, \hat{y})$. The standard optimality conditions as in Outrata (2000), Theorem 3.1, may be not too sharp because we are able to compute only an upper estimate of $N_{\mathrm{Gph} N_\Gamma}$ (according to Mordukhovich and Outrata, 2007, Theorem 4.3). On the other hand, $N_{\mathrm{Gph} N_{K(\hat{x})}}$ can be computed exactly, see Dontchev and Rockafellar (1996), Henrion and Roemisch (2007) and Henrion, Outrata and Surowiec (2009). We believe, however, that the conditions of Theorem 3 can be useful also for a polyhedral Γ given by linearly dependent inequalities, provided we are able to express $K(\hat{x})$ in a suitable way. This is shown in the next example.

EXAMPLE 2 *Consider the MPEC*

$$\begin{aligned} &\textit{minimize } -2x_1 - \frac{x_2}{2} - y_2 \\ &\textit{subject to} \\ &\qquad 0 \in a + x + y + N_\Gamma(y) \\ &\qquad x_1 \leq 0, \end{aligned} \tag{12}$$

where $a = (1, -1)^\top$ *and* $\Gamma = \{y \in \mathbb{R}^2 | y_2 \leq 0, y_2 \leq y_1, y_2 \leq -y_1\}$.

One can verify that the point, $(\hat{x}, \hat{y}) = (0, 0, 0, 0)$ *is a solution of (12). Moreover, the generalized equation in (12) is strongly regular at* $(0, 0, 0, 0)$. *Note that at* $(0, 0)$ *the active inequalities defining* Γ *are linearly dependent.*

At $(0, 0, 0, 0)$ *we can compute the critical cone*

$$K(\hat{x}) = \{(v_1, v_2) \in \mathbb{R}^2 | -v_1 + v_2 \leq 0, v_1 \leq 0\},$$

and the normal cone

$$\begin{aligned} N_{GphN_{K(\hat{x})}}(0,0) = &\left\{(u,b) \in \mathbb{R}^2 \times \mathbb{R}^2 \,|\, b = 0\right\} \bigcup \\ &\left\{(u,b) \in \mathbb{R}^2 \times \mathbb{R}^2 \,|\, u_1 = -u_2, b_1 = b_2\right\} \bigcup \\ &\left\{(u,b) \in \mathbb{R}^2 \times \mathbb{R}^2 \,|\, u_1 \geq -u_2, b_1 = b_2, b_2 \leq 0\right\}, \end{aligned}$$

where we followed the approach from Dontchev and Rockafellar (1996), proof of Theorem 2.

The conditions (10) reduce to

$$\begin{aligned} 0 &= -2 - b_1 + \xi, \\ 0 &= -\frac{1}{2} - b_2, \\ 0 &= u_1 - b_1, \\ 0 &= -1 + u_2 - b_2, \end{aligned}$$

with $(u, b) \in N_{GphN_{K(\hat{x})}}(0, 0)$ *and* $\xi \geq 0$. *We observe that they are satisfied only for the multipliers* $(u_1, u_2, b_1, b_2, \xi) = (-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{3}{2})$.

3. Exact penalization under ImP-hypothesis

To simplify the formulations, let us replace throughout this whole section the ImP-hypothesis by the stronger assumption that S is single-valued and locally Lipschitz over ω.

The notion of calmness is closely linked to exact penalization in nonlinear programming, see Clarke (1983), Proposition 6.4.3, Burke (1991) and Henrion and Outrata (2001). Since f is locally Lipschitz, the exact penalization property of $d_\Xi \circ S$ can be proved under the calmness of the multifunction $\widetilde{M}$. For the readers' convenience we give here a short proof of this well-known statement.

Proposition 3 *Let $\hat{x}$ be a local solution of (4), let ℓ_1 be a Lipschitz constant of f around $(\hat{x}, S(\hat{x}))$ and ℓ_2 be a Lipschitz constant of S around $\hat{x}$. Further, assume that the map $\widetilde{M}$, defined in (5) is calm at $(0, \hat{x})$ with modulus L. Then, for any $R \geq \ell_1(\ell_2 + 1)L$ the vector $\hat{x}$ solves the penalized problem*

$$\begin{aligned} &\text{minimize } \theta(x) + R d_\Xi(S(\hat{x})) \\ &\text{subject to} \\ &\qquad x \in \omega. \end{aligned} \tag{13}$$

Proof. Recall that $A = \omega \cap S^{-1}(\Xi)$. Applying Clarke (1983), Proposition 2.4.3, to (7), for any $\hat{\ell} \geq \ell_1(\ell_2 + 1)$, the function

$$g(x) = \theta(x) + \hat{\ell} d_A(x)$$

attains a minimum at $\hat{x}$.

From the definition of calmness of $\widetilde{M}$ at $(0, \hat{x})$ there is a neighborhood $\mathcal{U}$ of $\hat{x}$ and a modulus $L \geq 0$ such that for all $x \in \mathcal{U}$

$$d_A(x) = d_{\widetilde{M}(0)}(x) \leq L d_{\widetilde{M}^{-1}(x)}(0).$$

Clearly,

$$d_{\widetilde{M}^{-1}(x)}(0) = \begin{cases} d_{S(x)-\Xi}(0) = d_\Xi(S(x)) & \text{if } x \in \omega, \\ +\infty & \text{otherwise.} \end{cases}$$

This finishes the proof. ∎

Proposition 3 provides a lower bound for the penalty parameter R in relation with the Lipschitz and calmness moduli. This lower bound can also be related to the multipliers in the optimality conditions of Theorem 2.

To this end, let us rewrite the penalized problem (13) in the form

$$\begin{aligned} &\text{minimize } \theta(x) + R\|v\| \\ &\text{subject to} \\ &\qquad x \in \omega \\ &\qquad S(x) - v \in \Xi \end{aligned} \tag{14}$$

in variables x and v. Then, for the optimal solution $(\hat{x}, 0)$ of (14) we have from Mordukhovich (1988), Theorem 7.1, that

$$0 \in \begin{pmatrix} \partial\theta(\hat{x}) \\ R\mathbb{B} \end{pmatrix} + N_{\widetilde{A}}(\hat{x}, 0),$$

where $\widetilde{A} = (\omega \times \mathbb{R}^m) \cap D$ with $D = \{(x, v) | S(x) - v \in \Xi\}$.

Since

$$N_D(\hat{x}, 0) = \begin{pmatrix} D^*S(\hat{x}) \\ -Id \end{pmatrix} N_\Xi(S(\hat{x}))$$

by virtue of Rockafellar and Wets (1998), Corollary 10.50, the standard constraint qualification

$$N_{\omega\cap\mathbb{R}^m}(\hat{x}, 0) \cap (-N_D(\hat{x}, 0)) = \{0\}$$

for the intersection of sets holds true. Consequently,

$$N_{\widetilde{A}}(\hat{x}, 0) \subset \begin{pmatrix} N_\omega(\hat{x}) \\ 0 \end{pmatrix} + \begin{pmatrix} D^*S(\hat{x}) \\ -Id \end{pmatrix} N_\Xi(S(\hat{x})).$$

Hence, there is a multiplier $w \in N_\Xi(S(\hat{x}))$ such that

$$\begin{aligned} &0 \in \partial\theta(\hat{x}) + N_\omega(\hat{x}) + D^*S(\hat{x})(w), \\ &0 \in R\mathbb{B} - w. \end{aligned} \tag{15}$$

The first line of (15) clearly amounts to the condition (2.5), whereas the second line of (15) provides us with the desired estimate

$$R \geq \|w\|.$$

For the numerical solution of (13) we can propose a variant of a classical implicit programming technique. It has been developed in connection with the Stackelberg problem in Outrata (1990) with the use of a standard bundle method in nonsmooth optimization. In Outrata, Kočvara and Zowe (1998), this method has been applied in combination with the classical Bundle-Trust region (BT) algorithm from Schramm and Zowe (1992).

The main idea of BT is to construct an approximation of an objective ψ based on the bundle information $(\psi(x_i), g_i)$ for $i \in J_k$, where g_i is an arbitrary element from the Clarke subdifferential $\bar{\partial}\psi(x_i)$ and $J_k \subset \{0, 1, \ldots, k\}$ is a set which determines the part of bundle information used in the current iteration.

The key part in the successful application of the BT code to the penalized problem (13) is the computation of a Clarke subgradient of the nonsmooth function

$$\widetilde{\theta}(x) := \theta(x) + Rd_\Xi(S(x)) = (f + Rd_\Xi) \circ \begin{pmatrix} Id \\ S \end{pmatrix} (x) \tag{16}$$

at the current iteration point $\bar{x}$.

For the sake of simplicity we will further assume that f is continuously differentiable and that Ξ is convex. Then we distinguish the following three cases:

i) $S(\bar{x}) \notin \Xi$.
Upon choosing d to be the Euclidean distance, d is a C^1 function around $\bar{y} = S(\bar{x})$ and

$$\begin{aligned}\bar{\partial}\widetilde{\theta}(\bar{x}) &= \bar{\partial}(f + Rd_\Xi) \circ \begin{pmatrix} Id \\ S \end{pmatrix} (\bar{x}) \\ &= \nabla_x f(\bar{x}, \bar{y}) + \left\{ H^\top \left(\nabla_y f(\bar{x}, \bar{y}) + R\, \frac{\bar{y} - P_\Xi(\bar{y})}{d_\Xi(\bar{y})} \right) \middle| \, H \in \bar{\partial} S(\bar{x}) \right\} \\ &\supset \nabla_x f(\bar{x}, \bar{y}) + D^* S(\bar{x}) \left(\nabla_y f(\bar{x}, \bar{y}) + R\, \frac{\bar{y} - P_\Xi(\bar{y})}{d_\Xi(\bar{y})} \right),\end{aligned}$$

where the second equality follows from Clarke (1983), Theorem 2.6.6, and Rockafellar and Wets (1998), Example 8.53, and the inclusion follows from Mordukhovich (1994), relation (2.23).

ii) $S(\bar{x}) \in \operatorname{int} \Xi$.
Clearly, locally around $\bar{x}$ the functions $\widetilde{\theta}$ and θ coincide and so

$$\begin{aligned}\bar{\partial}\widetilde{\theta}(\bar{x}) = \bar{\partial}\theta(\bar{x}) &= \nabla_x f(\bar{x}, \bar{y}) + \left\{ H^\top \nabla_y f(\bar{x}, \bar{y}) \,\middle|\, H \in \bar{\partial} S(\bar{x}) \right\} \\ &\supset \nabla_x f(\bar{x}, \bar{y}) + D^* S(\bar{x})(\nabla_y f(\bar{x}, \bar{y})).\end{aligned}$$

iii) $S(\bar{x}) \in \operatorname{bd} \Xi$.
In this case, the distance function is not differentiable and thus the composition in (16) involves two nonsmooth functions. This leaves us only with the inclusions

$$\bar{\partial}\widetilde{\theta}(\bar{x}) \subset \nabla_x f(\bar{x}, \bar{y}) + \bigcup \left\{ D^* S(\bar{x})(\nabla_y f(\bar{x}, \bar{y}) + R\bar{\xi}) \,\middle|\, \bar{\xi} \in N_\Xi(\bar{y}) \cap \mathbb{B} \right\} \tag{17}$$

and

$$\begin{aligned}\bar{\partial}\widetilde{\theta}(\bar{x}) &\supset \hat{\partial}\widetilde{\theta}(\bar{x}) \\ &\supset \nabla_x f(\bar{x}, \bar{y}) + \bigcup \left\{ \widehat{D}^* S(\bar{x})(\nabla_y f(\bar{x}, \bar{y}) + R\bar{\xi}) \,\middle|\, \bar{\xi} \in N_\Xi(\bar{y}) \cap \mathbb{B} \right\},\end{aligned} \tag{18}$$

which follow from Rockafellar and Wets (1998), Corollary 10.9 and Theorem 10.49.

The above formulas require an arbitrary value of $D^* S(\bar{x})(\cdot)$ or $\widehat{D}^* S(\bar{x})(\cdot)$ evaluated at the appropriate argument. These quantities can be computed via the co-called limiting or regular adjoint generalized equation introduced in Kočvara and Outrata (2004). If S is given by (8), these adjoint generalized equations involve the respective coderivatives of the normal-cone mapping $N_\Gamma(\cdot)$ that have been analyzed for a number of frequently occurring sets Γ, see, e.g., Henrion and Roemisch (2007), Outrata and Sun (2008) and Henrion, Outrata and Surowiec (2009). Notice that, since the limiting normal cone to Gph N_Γ contains typically a linear subspace, the respective limiting adjoint generalized equation amounts to a linear equation, and so it is substantially easier to solve than its regular counterpart.

For this reason, in the case iii) we mostly put $\bar{\xi} = 0$ and compute the subgradient just like in the case ii). Unfortunately, this means that in some cases we might provide BT with a false subgradient, which could destroy the convergence. Despite this possibility, we did not observe such phenomenon in our numerical examples.

To test the performance of the BT method for MPECs with state constraints we modify the oligopolistic market example from Murphy, Sherali and Soyster (1982), see also Outrata, Kočvara and Zowe (1998), Section 12.1.

EXAMPLE 3 *Consider an example of five firms supplying a quantity* $z_i \in \mathbb{R}_+, i = 1, \ldots, 5,$ *of some homogeneous product on the market with the* inverse demand function

$$p(T) = 5000^{\frac{1}{\gamma}} T^{-\frac{1}{\gamma}},$$

where γ *is a positive parameter termed* demand elasticity *and* $T = \sum_{i=1}^{5} z_i$ *denotes the total supply.*

Let all the production cost functions *be in the form*

$$c_i(z_i) = b_i z_i + \frac{\beta_i}{1+\beta_i} K_i^{-\frac{1}{\beta_i}} (z_i)^{\frac{1+\beta_i}{\beta_i}},$$

where b_i, K_i *and* $\beta_i,\ i = 1, \ldots, 5,$ *are positive parameters given in Table 1.*

Table 1. Parameter specification for the production costs

	Firm 1	Firm 2	Firm 3	Firm 4	Firm 5
b_i	2	8	6	4	2
K_i	5	5	5	5	5
β_i	1.2	1.1	1.0	0.9	0.8

The aim of each firm is then to minimize the loss function

$$f_i(z) = c_i(z_i) - z_i p(T).$$

Each production cost function is convex and twice continuously differentiable on some open set containing the feasible set of strategies of a corresponding player. The inverse demand curve is twice continuously differentiable on int $\mathbb{R}_+$, *strictly decreasing, and convex. Observe that the so-called* industry revenue curve

$$Tp(T) = 5000^{\frac{1}{\gamma}} T^{\frac{\gamma-1}{\gamma}}$$

is concave on int $\mathbb{R}_+$ *for* $\gamma \geq 1$.

We suppose that the first firm is the so-called market leader*, i.e., has a temporal advantage and is able to decide about this production before the others, so-called* market followers*. To be consistent with the notation from the first part of the paper, put*

$$\begin{aligned} x :=& z_1 \\ y :=& (z_2, \ldots, z_5). \end{aligned}$$

The leader thus aims to solve the following MPEC

$$\begin{aligned} &\text{minimize } f_1(x, y) \\ &\text{subject to} \\ &\qquad 0 \in F(x, y) + N_{\mathbb{R}^4_+}(y) \\ &\qquad x \in \mathbb{R}_+, \end{aligned} \tag{19}$$

where

$$F(x, y) := \begin{pmatrix} \nabla_{y_1} f_2(x, y) \\ \nabla_{y_2} f_3(x, y) \\ \nabla_{y_3} f_4(x, y) \\ \nabla_{y_4} f_5(x, y) \end{pmatrix}.$$

We assume that the leader is producing some positive production quantity. By Outrata, Kočvara and Zowe (1998), Lemma 12.2, the partial Jacobian $\nabla_y F(x, y)$ *is positive definite at each feasible pair* (x, y)*. This implies that* S *is single-valued and the ImP-hypothesis holds true at each feasible pair* (x, y).

The first section of Table 2 shows the productions and profits of all firms for $\gamma = 1.0$*. Now, suppose there is a state constraint in the form* $y_i \leq N, i = 1, \ldots, 4$*, imposed on the followers. Applying the BT code to the respective penalized problem, the second and third section of Table 2 show the productions and profits for* $N = 45$ *and* $N = 40$*, respectively. In the last section we suggest, additionally to* $N = 40$*, to impose also an upper production bound* $M = 150$ *on the leader. In this case, the penalty term fails to vanish at the optimal point despite the calmness of the solution map simply due to the fact that this MPEC is infeasible. This situation can be easily detected: for any choice of penalty parameter* R*, the penalized distance at the optimal point remains the same.*

We point out that the penalty parameter R*, for which the penalty term becomes exact, is not a priori known in most examples. Table 3 then illustrates the dependence of the results on the choice of* R *for* $N = 40$.

Interestingly, we were able to obtain the same results also with the Euclidean norm replaced by the sum norm. Hence the differentiability of the distance function in out-of-set points does not seem to be of a great importance.

Table 2. Productions and profits

	Firm 1	Firm 2	Firm 3	Firm 4	Firm 5
No state constraints					
Production	99.5329	44.3804	45.8893	44.2806	40.2357
Profit	958.6348	284.6830	350.5039	393.2798	410.5312
$N = 45, R = 15$					
Production	108.1305	43.2615	45.0000	43.5921	39.7215
Profit	952.5061	266.1045	331.9174	375.8856	395.1587
Penalized distance	0				
$N = 40, R = 100$					
Production	157.6612	36.9306	40.0000	39.7606	36.8917
Profit	691.8176	178.6969	242.5824	291.1216	319.4797
Penalized distance	$1.5857 \cdot 10^{-6}$				
$N = 40, M = 150$					
Production	150.000	37.8935	40.7579	40.3377	37.3149
Profit	673.1138	190.2799	254.6497	302.7105	329.9166
Penalized distance	0.8297				

Table 3. Productions and profits: $N = 40$

	Firm 1	Firm 2	Firm 3	Firm 4	Firm 5
$R = 10$					
Production	110.1426	43.0002	44.7926	43.4319	39.6022
Profit	883.1784	261.9189	327.7146	371.9429	391.6679
Penalized distance	6.6143				
$R = 50$					
Production	139.4624	39.2288	41.8101	41.1409	37.9056
Profit	723.5776	207.2848	277.2178	319.4936	344.9748
Penalized distance	2.1397				
$R = 100$					
Production	157.6612	36.9306	40.0000	39.7606	36.8917
Profit	691.8176	178.6969	242.5824	291.1216	319.4797
Penalized distance	$1.5857 \cdot 10^{-6}$				

Conclusion

We have studied the influence of the ImP-hypothesis on a class of MPECs. It has been proved that this hypothesis does not lead to sharper optimality conditions, but enables us to simplify the respective constraint qualification. Then we have considered equilibria governed by strongly regular variational inequalities and derived a new variant of Mordukhovich stationarity conditions. It is well-suited to the case when the constraint set of the variational inequality is given by linearly dependent inequalities. Finally, we have shown the ability of a numerical method from Outrata (1990) and Outrata, Kočvara and Zowe (1998) to solve also some MPECs with state constraints.

Acknowledgment

The authors are indebted to S. Dempe, D. Ralph and S. Scholtes for helpful discussions about parameterized variational inequalities with constraint sets given by linearly dependent inequality systems. Additionally, the authors would like to thank to anonymous referees for careful reading and important suggestions and comments which significantly improved the text.

References

Burke, J.V. (1991) Calmness and exact penalization. *SIAM J. Control Optim.* **29**, 493–497.

Clarke, F.H. (1983) *Optimization and Nonsmooth Analysis.* Wiley, New York.

Dempe, S. (2002) *Foundations of Bi-level Programming.* Kluwer Academic Publishers, Dordrecht, The Netherlands.

Dontchev, A.D. and Rockafellar, R.T. (1996) Characterization of strong regularity for variational inequalities over polyhedral convex sets. *SIAM J. Optim.* **7**, 1087–1105.

Henrion, R. and Outrata, J.V. (2001) A subdifferential criterion for calmness of multifunctions. *J. Math. Anal. Appl.* **258**, 110–130.

Henrion, R. and Roemisch, W. (2007) On M-stationary points for a stochastic equilibrium problem under equilibrium constraints in electricity spot market modeling. *Applications of Mathematics* **52**, 6, 473–494.

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

Henrion, R., Outrata, J.V. and Surowiec, T. (2009) On the co-derivative of normal cone mappings to inequality systems. *Nonlinear Analysis: Theory, Methods and Applications* **71**, 1213–1226.

Ioffe, A.D. and Outrata, J.V. (2008) On metric and calmness qualification conditions in subdifferential calculus. *Set-Valued Analysis* **16**, 199–227.

Kočvara, M. and Outrata, J.V. (2004) Optimization problems with equilibrium constraints and their numerical solution. *Math. Prog. B* **101**, 119–150.

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

Mordukhovich, B.S. (1988) *Approximation Methods in Problems of Optimization and Control.* Nauka, Moscow (in Russian).

Mordukhovich, B.S. (1994) Generalized differential calculus for nonsmooth and set-valued mappings. *J. Math. Anal. Appl.* **183**, 250–288.

Mordukhovich, B.S. (2006) *Variational Analysis and Generalized Differentiation: I (Basic Theory) and II (Applications).* Springer-Verlag, Berlin, Heidelberg.

Mordukhovich, B.S. and Outrata, J.V. (2001) Second order subdifferentials and their applications. *SIAM J. Optim.* **12**, 139–169.

MORDUKHOVICH, B.S. and OUTRATA, J.V. (2007) Coderivative analysis of quasivariational inequalities with applications to stability and optimization. *SIAM J. Optim.* **18**, 389–412.

MURPHY, F.H., SHERALI, H.D. and SOYSTER, A.L. (1982) A mathematical programming approach for determining oligopolistic market equilibrium. *Math. Prog.* **24**, 92–106.

OUTRATA, J.V. (1990) On the numerical solution of a class of Stackelberg problems. *Zeit. f. Operationsresearch* **34**, 255–277.

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

OUTRATA, J.V. (2000) Generalized mathematical program with equilibrium constraints. *SIAM J. Control Optim.* **38**, 1623–1638.

OUTRATA, J.V. and SUN, D. (2008) On the coderivative of the projection operator onto the second-order cone. *Set-Valued Analysis* **16**, 999–1014.

OUTRATA, J.V., KOČVARA, M. and ZOWE, J. (1998) *Nonsmooth Approach to Optimization Problems with Equilibrium Contraints.* Kluwer, Dordrecht.

RALPH, D. and DEMPE, S. (1995) Directional derivatives of the solution of a parametric nonlinear program. *Math. Prog.* **70**, 159–172.

ROBINSON, S.M. (1980) Strongly regular generalized equations. *Mathematics of Operations Research* **5**, 43–62.

ROBINSON, S.M. (1987) Local structure of feasible sets in nonlinear programming, Part III: Stability and Sensitivity. *Mathematical Programming Study* **30**, 45–66.

ROBINSON, S.M. (1991) An implicit-function theorem for a class of nonsmooth functions. *Mathematics of Operations Research* **16**, 292— 309.

ROCKAFELLAR, R.T. and WETS, R. (1998) *Variational Analysis.* Springer Verlag, Berlin.

SCHRAMM, H. and ZOWE, J. (1992) A version of bundle idea for minimizing a nonsmooth function: conceptual idea, convergence analysis, numerical results. *SIAM J. Optim.* **2**, 121–152.

YE, J.J. (2000) Constraint qualifications and necesary optimality conditions for optimization problems with variational inequality constraints. *SIAM J. Optim.* **10**, 943–962.

YE, J.J. and YE, X.Y. (1997) Necessary optimality conditions for optimization problems with variational inequality constraints. *Math. Oper. Res.* **22**, 977–997.