Math. Program., Ser. B 104, 437–464 (2005)

Digital Object Identifier (DOI) 10.1007/s10107-005-0623-2

René Henrion* · Jiří V. Outrata **

Calmness of constraint systems with applications

This paper is dedicated to Prof. R. T. Rockafellar on the occasion of his 70th birthday

Received: April 28, 2004 / Accepted: November 9, 2004
Published online: August 10, 2005 –

Abstract. The paper is devoted to the analysis of the calmness property for constraint set mappings. After some general characterizations, specific results are obtained for various types of constraints, e.g., one single nonsmooth inequality, differentiable constraints modeled by polyhedral sets, finitely and infinitely many differentiable inequalities. The obtained conditions enable the detection of calmness in a number of situations, where the standard criteria (via polyhedrality or the Aubin property) do not work. Their application in the framework of generalized differential calculus is explained and illustrated by examples associated with optimization and stability issues in connection with nonlinear complementarity problems or continuity of the value-at-risk.

Key words. Calmness – constraint sets – nonsmooth calculus – value-at-risk

1. Introduction

There are very many possibilities to define Lipschitz-like properties for a multifunction $Z : Y \rightrightarrows X$ between metric spaces Y and X. For reasons of analogy, it would be appealing, to require at some $\bar{y} \in Y$ the estimate (for some $L, \varepsilon > 0$)

$$d^H(Z(y_1), Z(y_2)) \leq Ld(y_1, y_2) \quad \forall y_1, y_2 \in \mathbb{B}(\bar{y}, \varepsilon).$$

Here, "d^H" refers to the Hausdorff distance between subsets of X, and "$\mathbb{B}(\bar{y}, \varepsilon)$" means a closed ball around $\bar{y}$ with radius ε. Clearly, in the case of a single-valued function Z, d_H reduces to the usual distance in X, and one arrives at the classical local Lipschitz property of functions. More explicitly, the relation above can be formulated as

$$d_{Z(y_1)}(x) \leq Ld(y_1, y_2) \quad \forall x \in Z(y_2)\ \forall y_1, y_2 \in \mathbb{B}(\bar{y}, \varepsilon), \tag{1}$$

where, "d_A" is the distance of a point to a set A. For many applications in variational analysis, nonlinear optimization, nonsmooth calculus etc., this notion is too strong and one

R. Henrion: Weierstrass Institute Berlin, D-10117 Berlin, Germany. e-mail: henrion@wias-berlin.de

J.V. Outrata: Institute of Information Theory and Automation, 18208 Prague, Czech Republic.
e-mail: outrata@utia.cas.cz

Mathematics Subject Classification (2000): 90C30, 49J53

* This research was supported by the DFG Research center MATHEON *Mathematics for key technologies* in Berlin

** Support by grant IAA1030405 of the Grant Agency of the Academy of Sciences of the Czech Republic is acknowledged

rather considers restricted versions of it. The *Aubin property* ([35]), for instance, refers to localized image sets by replacing the expression "$Z(y_2)$" in (1) with "$Z(y_2) \cap \mathbb{B}(\bar{x}, \varepsilon)$", where $\bar{x} \in Z(\bar{y})$ (originally, this concept was introduced under the name pseudo-Lipschitz in [1], and it is closely related to the sub-Lipschitz property in [34]). Another restriction concerns the degree of freedom for the arguments. When fixing $y_1 = \bar{y}$ in (1), Z is said to be *(locally) upper Lipschitz* at $\bar{y}$ ([31]). When combining both mentioned (independent) relaxations of (1), one arrives at the so-called *calmness* property of a multifunction as introduced in [35] (and in [39] under a different name). More explicitly, Z is said to be calm at some $(\bar{y}, \bar{x}) \in \operatorname{Gph} Z$ (graph of Z), if there exist $L, \varepsilon > 0$ such that

$$d_{Z(\bar{y})}(x) \leq Ld(y, \bar{y}) \quad \forall x \in Z(y) \cap \mathbb{B}(\bar{x}, \varepsilon) \; \forall y \in \mathbb{B}(\bar{y}, \varepsilon). \tag{2}$$

Note that, due to the symmetric role of y_1 and y_2, (1) as well as the Aubin property are upper and lower semicontinuity properties at the same time. In contrast, as a consequence of fixing $y_1 = \bar{y}$, calmness and local upper Lipschitzness are just upper semicontinuity properties. The corresponding lower counterparts are obtained when exchanging $\bar{y}$ and y in the respective definitions. A restricted version of calmness, called *calmness on selections* ([10], [19], [22]) substitutes the set $Z(\bar{y})$ by the singleton $\{\bar{x}\}$ in (2). This stronger condition entails that $\mathbb{B}(\bar{x}, \varepsilon) \cap Z(\bar{y}) = \{\bar{x}\}$, i.e., $\{\bar{x}\}$ is isolated in $Z(\bar{y})$.

This paper will focus its attention on the (general) calmness property (2). Of particular importance is the calmness of constraint set mappings as this becomes the key for the existence of local error bounds, exact penalty functions, (nonsmooth) necessary optimality conditions or weak sharp local minimizers. To be more precise, let now Y be a normed space, $\Lambda \subseteq Y$ a closed subset and $g : X \to Y$ a continuous mapping. The multifunction

$$M(y) := \{x \in X \mid g(x) + y \in \Lambda\} \tag{3}$$

may be interpreted as a perturbation of the constraint set $M(0) = g^{-1}(\Lambda)$. Then, at some $\bar{x}$ with $g(\bar{x}) \in \Lambda$, the following statements are equivalent:

1. M is calm at $(0, \bar{x})$.
2. $\exists L, \tilde{\varepsilon} > 0 : d_{g^{-1}(\Lambda)}(x) \leq Ld_\Lambda(g(x)) \quad \forall x \in \mathbb{B}(\bar{x}, \tilde{\varepsilon})$.
3. $\exists L, \tilde{\varepsilon} > 0 : d_{M(0)}(x) \leq L \|y\| \quad \forall y \in Y \, \forall x \in \mathbb{B}(\bar{x}, \tilde{\varepsilon}) \cap M(y)$.

Indeed, one may choose $\tilde{\varepsilon} < \varepsilon$ such that $\|g(x) - g(\bar{x})\| \leq \varepsilon/2$ for all $x \in \mathbb{B}(\bar{x}, \tilde{\varepsilon})$, where ε refers to (2). Now, for arbitrary $x \in \mathbb{B}(\bar{x}, \tilde{\varepsilon})$ and arbitrary $\eta \in (0, \varepsilon/2)$ there is some $\lambda \in \Lambda$ such that

$$\|g(x) - \lambda\| \leq d_\Lambda(g(x)) + \eta \leq \|g(x) - g(\bar{x})\| + \varepsilon/2 \leq \varepsilon.$$

Since $x \in M(\lambda - g(x))$ and $\lambda - g(x) \in \mathbb{B}(0, \varepsilon)$, 1. implies 2. via (2) by taking into account that η was arbitrary:

$$d_{g^{-1}(\Lambda)}(x) = d_{M(0)}(x) \leq L \|\lambda - g(x)\| \leq L \left(d_\Lambda(g(x)) + \eta\right) \quad \forall x \in \mathbb{B}(\bar{x}, \tilde{\varepsilon}).$$

Next, let $y \in Y$ and $x \in \mathbb{B}(\bar{x}, \tilde{\varepsilon}) \cap M(y)$ be arbitrary. Then, $g(x) + y \in \Lambda$, whence $d_\Lambda(g(x)) \leq \|y\|$. Consequently, 2. implies 3. which, in turn, trivially entails 1.

The equivalence between 1. and 3. shows that, for the considered constraint set mappings, the localization of the perturbation parameter y may be omitted when dealing with calmness (in a slightly different context, this was first observed in [4]; see also Prop. 3.4 in [11] for a more general statement). More importantly, the equivalence between 1. and 2. shows that calmness of M amounts to the existence of a local error bound (e.g., [28]) of the constraint function g. It is exactly this equivalence which explains calmness of constraint systems to be the basic condition in the context of penalty functions or constraint qualifications for optimality conditions (see, e.g., [4], [7], [37]). For a recent discussion of these relations, we refer to [20]. A further observation is that the value function φ of some optimization problem having $M(y)$ as a parametric constraint satisfies the inequality

$$\varphi(y) \geq \varphi(0) - c\|y\| \quad (c > 0,\ y \text{ close to } 0),$$

provided that the objective of this problem is locally Lipschitz and that M is calm at solutions. This estimate was the very origin of the calmness concept ([6]). Finally, we note (e.g., [15], Lemma 4.7) that in an optimization problem

$$\min\{f(x) \mid x \in C\}$$

the calmness of the multifunction $y \mapsto \{x \in C \mid f(x) \leq y\}$ at local solutions amounts to these solutions being weak sharp local minimizers (see, e.g.,[5], [36]).

A standard way to ensure calmness of a general multifunction $Z : Y \rightrightarrows X$ consists in the application of some suitable criterion ensuring the (stronger) Aubin property. Alternatively, from [32] we know that, in the finite-dimensional case, Z is calm at each point of its graph whenever this graph is polyhedral (i.e. a union of finitely many convex polyhedral sets). In [14] and [15] the authors derived calmness criteria in the nonpolyhedral case which do not necessarily imply the Aubin property. They consider, however, a specific structure

$$Z(y) = M(y) \cap \Theta, \tag{4}$$

where $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$, $\Theta \subseteq X$ is closed and g in (3) is locally Lipschitz. Additional assumptions like semismoothness or regularity are imposed on g, Λ and Θ. Multifunctions of the type (4) arise frequently in applications. Moreover, as shown in [21], the calmness of a multifunction $\tilde{Z}(y_1, y_2) = Z_1(y_1) \cap Z(y_2)$ can be ensured via the calmness of another map having the form (4). Applying the approach from [14], [15] provides useful information only in case the point of interest $\bar{x}$ belongs to the boundary of Θ. Otherwise, the two main alternative conditions derived there reduce to

$$\ker D^*g(\bar{x}) \cap N_\Lambda(g(\bar{x})) = \{0\}, \tag{5}$$

$$0 \in \operatorname{int} D^*g(\bar{x})(N_\Lambda(g(\bar{x})) \cap \mathbb{B}), \tag{6}$$

where the definitions of the coderivative D^*g and of the limiting normal cone N_Λ can be found in Section 2. Unfortunately, (5) is precisely the standard criterion for the Aubin property of M around $(0, \bar{x})$ which can be derived on the basis of the so-called Mordukhovich criterion ([35]). If g is continuously differentiable and $\Lambda = \mathbb{R}^m_-$, then (5) amounts to the standard Mangasarian-Fromowitz constraint qualification (MFCQ) in dual form

$$0 \notin \operatorname{conv}\{\nabla g_i(\bar{x}) \mid i \in I(\bar{x})\},$$

where $I(\bar{x}) = \{i \in \{1, 2, \ldots, m\} \mid g_i(\bar{x}) = 0\}$. Therefore we will use for condition (5) the name (GMFCQ). Also note that (6) entails not only calmness but even the isolatedness of $\bar{x}$ in $M(0)$, i.e., it is a criterion for the calmness on selections mentioned above (see Remark 3.7 in [15]). Summarizing, the use of the criteria developed in [14], [15] shrinks when applied to interior points of Θ (in particular for $\Theta = X$).

The aim of this paper is to derive new conditions for calmness of (3) which should be weaker than (5) and applicable also in case $\bar{x}$ is not an isolated point of $M(0)$. The paper is organized as follows: Section 3 contains the main results. They are ordered according to the assumptions imposed on the problem data and illustrated by a number of examples. Some of them admit that the spaces X, Y are infinite-dimensional. Section 4 provides applications of the obtained results to generalized differential calculus as well as to stability of the value-at-risk.

2. Notation

The following notation is employed: $\mathbb{B}$ and $\mathbb{S}$ denote the unit ball and the unit sphere, respectively. For a closed cone D with vertex at the origin, D^0 denotes its negative polar cone. $T_\Lambda(x)$ is the contingent (Bouligand) cone to Λ at x and $\bar{\partial} f(x)$ is the Clarke subdifferential of a real-valued function f at x.

For a set $\Pi \subseteq \mathbb{R}^p$ let $a \in \operatorname{cl} \Pi$. The cone

$$\hat{N}_\Pi(a) := \left\{ \xi \in \mathbb{R}^p \mid \limsup_{a' \xrightarrow{\Pi} a} \frac{\langle \xi, a' - a \rangle}{\|a' - a\|} \leq 0 \right\}$$

is called the *Fréchet* normal cone to Π at a.

The notions of the limiting normal cone, the limiting subdifferential and the coderivative are the cornerstones of the generalized differential calculus of B. Mordukhovich, cf. [24],[25]. The *limiting* normal cone to Π at a, denoted $N_\Pi(a)$ is defined by

$$N_\Pi(a) = \limsup_{a' \xrightarrow{cl\Pi} a} \hat{N}_\Pi(a'),$$

where the "limsup" means the Painlevé-Kuratowski upper (outer) limit. In this finite-dimensional setting one has $\hat{N}_\Pi(a) = (T_\Pi(a))^0$. If $N_\Pi(a) = \hat{N}_\Pi(a)$, we say that Π is *(normally) regular* at a. (In our setting, this regularity concept coincides with the well-known Clarke-regularity). If Π is convex, then $N_\Pi(a) = \hat{N}_\Pi(a)$ at each $a \in \Pi$ and so we will consequently use only the notation $N_\Pi(a)$. Now, let $\varphi : \mathbb{R}^p \to \overline{\mathbb{R}}$ be an arbitrary extended real-valued function and $a \in \operatorname{dom} \varphi$. The set

$$\partial\varphi(a) := \{a^* \in \mathbb{R}^p \mid (a^*, -1) \in N_{\operatorname{epi} \varphi}(a, \varphi(a))\}$$

is called the *limiting subdifferential* of φ at a. Finally, let $\Phi : \mathbb{R}^p \rightrightarrows \mathbb{R}^q$ be an arbitrary multifunction and $(a, b) \in \operatorname{cl} \operatorname{Gph} \Phi$. The multifunction $D^*\Phi(a, b) : \mathbb{R}^q \rightrightarrows \mathbb{R}^p$, defined by

$$D^*\Phi(a, b)(b^*) := \{a^* \in \mathbb{R}^p \mid (a^*, -b^*) \in N_{\operatorname{Gph} \Phi}(a, b)\},$$

is called the *coderivative* of Φ at (a, b).

A function $f : \mathbb{R}^p \to \mathbb{R}$ is called *semismooth* at $\bar{x} \in \mathbb{R}^p$ if it is Lipschitz around $\bar{x}$ and for any sequences $t_n \downarrow 0,\ d_n \to d,\ \xi_n \in \partial f(\bar{x} + t_n d_n)$ the limit $\lim_{n\to\infty} \langle \xi_n, d\rangle$ exists for each $d \in \mathbb{R}^p$. The concept of semismoothness plays an important role both in the numerical methods of nonsmooth analysis ([23]) as well as in the characterization of calmness provided in [14], [15].

3. Characterization of calmness

Throughout the whole paper, we shall be concerned with a multifunction $M : Y \rightrightarrows X$ between normed spaces X, Y, which is defined by

$$M(y) := \{x \in X | g(x) + y \in \Lambda\}, \tag{7}$$

where $g : X \to Y$ and $\Lambda \subseteq Y$ is a closed subset (recall that (7) has been considered before in (3) under the additional assumption of continuity for g).

When inspecting (7), one may wonder if the consideration of canonical perturbations y of g is a serious restriction. The following lemma shows that for Lipschitz data no difference with a general parameterization arises.

Lemma 1. *Let X, U, Y be normed spaces. Consider a multifunction $M^* : U \rightrightarrows X$ defined on the basis of some locally Lipschitzian (with respect to the product topology) function $h : U \times X \to Y$ by means of*

$$M^*(u) := \{x \in X | h(u, x) \in \Lambda\} \quad (\Lambda \subseteq Y).$$

Assume that $h(\bar{u}, \bar{x}) \in \Lambda$ for some $\bar{u} \in U$ and $\bar{x} \in X$. Then, M^* is calm at $(\bar{u}, \bar{x})$ provided that M in (7) is calm at $(0, \bar{x})$ with $g(x) := h(\bar{u}, x) \quad \forall x \in X$.

Proof. The local Lipschitz continuity of h and the calmness of M yield constants $K, L, \varepsilon > 0$ such that

$$\left\|h(u', x) - h(u'', x)\right\| \le K \left\|u' - u''\right\| \quad \forall u', u'' \in \mathbb{B}(\bar{u}, \varepsilon)\, \forall x \in \mathbb{B}(\bar{x}, \varepsilon)$$
$$d_{M(0)}(x) \le L \left\|y\right\| \quad \forall y \in \mathbb{B}(0, \varepsilon)\, \forall x \in \mathbb{B}(\bar{x}, \varepsilon) \cap M(y).$$

Choose ε' such that $0 < \varepsilon' \le \varepsilon$ and $\|h(u, x) - h(\bar{u}, x)\| \le \varepsilon$ for all $(u, x) \in \mathbb{B}(\bar{u}, \varepsilon') \times \mathbb{B}(\bar{x}, \varepsilon')$. Let $x \in M^*(u) \cap \mathbb{B}(\bar{x}, \varepsilon')$ and $u \in \mathbb{B}(\bar{u}, \varepsilon')$ be arbitrary. Then, $x \in M(h(u, x) - g(x)) \cap \mathbb{B}(\bar{x}, \varepsilon')$ by definition of M and M^*. It follows the calmness of M^* at $(\bar{u}, \bar{x})$:

$$d_{M^*(\bar{u})}(x) = d_{M(0)}(x) \le L \left\|h(u, x) - g(x)\right\| \le LK \left\|u - \bar{u}\right\|.$$

□

Statements of the above type can be found in connection with various properties of constraint and variational systems, see e.g., [8]. A general framework for such considerations, based on the notion of *strong approximation* is provided in [33].

The following lemma allows equivalently to reduce the calmness of system (7) to the calmness of a single (nonsmooth) inequality where the distance function is involved.

Lemma 2. *With the multifunction M from (7) we associate a multifunction* $\tilde{M} : \mathbb{R} \rightrightarrows X$ *defined by*

$$\tilde{M}(t) = \{x \in X | d_\Lambda (g(x)) \leq t\}.$$

Then, M is calm at some $(0, \bar{x}) \in \mathrm{Gph}\, M$ *if and only if* $\tilde{M}$ *is calm at* $(0, \bar{x})$.

Proof. Note that $M(0) = \tilde{M}(0)$, hence $(0, \bar{x}) \in \mathrm{Gph}\, M$ if and only if $(0, \bar{x}) \in \mathrm{Gph}\, \tilde{M}$. Assume first that $\tilde{M}$ is calm at $(0, \bar{x})$. By definition, there exist $L, \varepsilon > 0$ such that

$$d_{\tilde{M}(0)}(x) \leq L|t| \quad \forall t \in [-\varepsilon, \varepsilon]\, \forall x \in \tilde{M}(t) \cap \mathbb{B}(\bar{x}, \varepsilon).$$

For any $y \in \mathbb{B}(0, \varepsilon)$ and any $x \in M(y) \cap \mathbb{B}(\bar{x}, \varepsilon)$ one has that $d_\Lambda (g(x)) \leq \|y\| \leq \varepsilon$, hence $x \in \tilde{M}(\|y\|)$ and it follows the calmness of M at $(0, \bar{x})$:

$$d_{M(0)}(x) = d_{\tilde{M}(0)}(x) \leq L\, \|y\| \quad \forall y \in \mathbb{B}(0, \varepsilon)\, \forall x \in M(y) \cap \mathbb{B}(\bar{x}, \varepsilon).$$

Conversely, let M be calm at $(0, \bar{x})$. By definition, there exist $L, \varepsilon > 0$ such that

$$d_{M(0)}(x) \leq L\, \|y\| \quad \forall y \in \mathbb{B}(0, \varepsilon)\, \forall x \in M(y) \cap \mathbb{B}(\bar{x}, \varepsilon).$$

For any $t \in [-\varepsilon/2, \varepsilon/2]$ and any $x \in \tilde{M}(t) \cap \mathbb{B}(\bar{x}, \varepsilon)$ one has that $t \geq 0$ (otherwise $\tilde{M}(t) = \emptyset$) and $d_\Lambda (g(x)) \leq t = |t| \leq \varepsilon/2$. If $t = 0$, then $d_{M(0)}(x) = 0$. Otherwise ($t > 0$), choose $\lambda \in \Lambda$ such that $\|\lambda - g(x)\| \leq 2t$ and put $y := \lambda - g(x)$. Then, $y \in \mathbb{B}(0, \varepsilon)$ and $x \in M(y)$, hence it follows the calmness of $\tilde{M}$ at $(0, \bar{x})$:

$$d_{\tilde{M}(0)}(x) = d_{M(0)}(x) \leq L\, \|y\| \leq 2L|t| \quad \forall t \in [-\varepsilon/2, \varepsilon/2]\, \forall x \in \tilde{M}(t) \cap \mathbb{B}(\bar{x}, \varepsilon/2).$$

□

Obviously, the set $M(0)$ in (7) is the set of global minimizers for the function $d_\Lambda \circ g$. Referring back to the introduction, Lemma 2 then shows that M is calm at $(0, \bar{x})$ if and only if $\bar{x}$ is a weak sharp local minimizer for $d_\Lambda \circ g$.

Corollary 1. *In (7), M fails to be calm at some* $(0, \bar{x}) \in \mathrm{Gph}\, M$ *if and only if there exists a sequence* $x_l \to \bar{x}$ *such that* $d_{M(0)}(x_l) > l d_\Lambda(g(x_l))$*. In particular,* $x_l \notin M(0)$ *or, equivalently,* $g(x_l) \notin \Lambda$.

The next proposition relates the calmness property to the Abadie constraint qualification (ACQ) which is well-known from mathematical programming, (see [3]).

Proposition 1. *In (7), let* $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$ *and g be Lipschitz around* $\bar{x} \in M(0)$ *and directionally differentiable at* $\bar{x}$*. Let* $L_{M(0)}(\bar{x})$ *be the linearized cone to* $M(0)$ *at* $\bar{x}$*, defined by*

$$L_{M(0)}(\bar{x}) = \{h \in \mathbb{R}^n | g'(\bar{x}; h) \in T_\Lambda(g(\bar{x}))\}.$$

If M is calm at $(0, \bar{x})$*, then*

$$T_{M(0)}(\bar{x}) = L_{M(0)}(\bar{x}). \tag{8}$$

Proof. The inclusion $T_{M(0)}(\bar{x}) \subseteq L_{M(0)}(\bar{x})$ holds generally true (without calmness) when g is locally Lipschitz and directionally differentiable. For the reverse inclusion, assume by contradiction the existence of some $h \in \mathbb{R}^n$ such that $g'(\bar{x}; h) \in T_\Lambda(g(\bar{x}))$ but $h \notin T_{M(0)}(\bar{x})$. This amounts to the existence of some $\mu > 0$ with

$$\liminf_{t\downarrow 0} t^{-1} d_{M(0)}(\bar{x} + th) = \mu.$$

On the other hand, there are sequences $k_i \to g'(\bar{x}; h)$ and $t_i \downarrow 0$ such that $g(\bar{x}) + t_i k_i \in \Lambda$ for all i. This means that

$$d_\Lambda(g(\bar{x}) + t_i g'(\bar{x}; h)) \leq t_i \left\| k_i - g'(\bar{x}; h) \right\| \quad \forall i$$

and, consequently,

$$\begin{aligned} t_i^{-1} d_\Lambda(g(\bar{x} + t_i h)) \;&\leq\; t_i^{-1}\{d_\Lambda(g(\bar{x}) + t_i g'(\bar{x}; h)) + \left|g(\bar{x} + t_i h) - g(\bar{x}) - t_i g'(\bar{x}; h)\right|\} \\ &\to_{\; i\to\infty} 0. \end{aligned}$$

For arbitrary $l \in \mathbb{N}$ set $\varepsilon_l := (l+1)^{-1}\mu$. Choose $i_l \in \mathbb{N}$ such that $t_{i_l}^{-1} d_\Lambda(g(\bar{x}+t_{i_l}h)) < \varepsilon_l$ and $t_{i_l}^{-1} d_{M(0)}(\bar{x} + t_{i_l}h) > \mu - \varepsilon_l$. One may assume that i_l is increasing, hence t_{i_l} is a subsequence of t_i. Putting $x_l := \bar{x} + t_{i_l}h$, one gets

$$d_{M(0)}(x_l) > t_{i_l}(\mu - \varepsilon_l) = t_{i_l} l \varepsilon_l > l d_\Lambda(g(x_l)),$$

which contradicts the calmness of M at $(0, \bar{x})$ according to Corollary 1. □

The classical (ACQ) amounts to the identity (8) in the case of standard nonlinear programming, where g is continuously differentiable and $\Lambda = \mathbb{R}_-^{m_1} \times \{0\}_{m_2}$, $m_1 + m_2 = m$. In the following, we will keep the name (ACQ) for (8) under weaker assumptions on g and Λ, specified in Proposition 1.

The following example shows that the converse of Proposition 1 does not apply even in case of a $\mathcal{C}^1$ -function.

Example 1. Put $\Lambda := \mathbb{R}_-$, $\bar{x} = 0$, $g(x) := x^3 \sin x^{-1}$ (with $g(0) = 0$). Then $T_{M(0)}(\bar{x}) = \mathbb{R} = L_{M(0)}(\bar{x})$, i.e., (ACQ) is satisfied but M fails to be calm at (0, 0). To see this in a convenient way, we make a forward reference to Proposition 3 and observe that, in our example, g is continuously differentiable, $g'(0) = 0$ (which is equivalent with the failure of the Aubin property) and 0 is not a local maximizer of g.

3.1. Special Cases

In this section, we collect criteria for calmness in certain special cases. For a function $g : \mathbb{R}^n \to \mathbb{R}$ denote by

$$\begin{aligned} g^\downarrow(\bar{x}; h) &:= \liminf_{t\downarrow 0, h'\to h} t^{-1}(g(\bar{x} + th') - g(\bar{x})) \\ g^\uparrow(\bar{x}; h) &:= \limsup_{t\downarrow 0, h'\to h} t^{-1}(g(\bar{x} + th') - g(\bar{x})) \end{aligned}$$

the lower and the upper Hadamard derivative at $\bar{x}$ in direction h. We start with the simple situation of an inequality defined by a real function.

Proposition 2. *In (7), let $X = \mathbb{R}$, $Y = \mathbb{R}$, $\Lambda = \mathbb{R}_-$ and g be lower semicontinuous at some $\bar{x}$ with $g(\bar{x}) = 0$. Then, M is calm at $(0, \bar{x})$ if the following two conditions hold true:*

$$0 \in \left[g^{\downarrow}(\bar{x}; 1), g^{\uparrow}(\bar{x}; 1)\right] \Longrightarrow \exists\, \varepsilon > 0\, \exists\, \eta > 0\, \forall x \in [\bar{x}, \bar{x} + \varepsilon] : \quad g(x) \le 0 \text{ or } g(x) \ge \eta(x - \bar{x}). \tag{9}$$

$$0 \in \left[g^{\downarrow}(\bar{x}; -1), g^{\uparrow}(\bar{x}; -1)\right] \Longrightarrow \exists\, \varepsilon > 0\, \exists\, \eta > 0\, \forall x \in [\bar{x} - \varepsilon, \bar{x}] : \quad g(x) \le 0 \text{ or } g(x) \ge \eta(\bar{x} - x). \tag{10}$$

If, moreover, g is semismooth at $\bar{x}$ (see sect. 2), then the pair of conditions

$$g'(\bar{x}; 1) = 0 \Longrightarrow \exists\, \varepsilon > 0\, \forall x \in [\bar{x}, \bar{x} + \varepsilon] : g(x) \le> 0 \tag{11}$$

$$g'(\bar{x}; -1) = 0 \Longrightarrow \exists\, \varepsilon > 0\, \forall x \in [\bar{x} - \varepsilon, \bar{x}] : g(x) \le> 0 \tag{12}$$

is equivalent with M being calm at $(0, \bar{x})$.

Proof. Assuming violation of calmness, Corollary 1 provides a sequence $x_l \to \bar{x}$ such that

$$0 < g(x_l) < l^{-1} d_{M(0)}(x_l) \le l^{-1} |x_l - \bar{x}| \ \ \forall\, l \in \mathbb{N}. \tag{13}$$

Without loss of generality, we may assume that, upon passing to a subsequence, $x_l > \bar{x}$ or $x_l < \bar{x}$ for all l. Assume first that $x_l > \bar{x}$ for all l. Then, (13) amounts to $g^{\downarrow}(\bar{x}; 1) \le 0$. On the other hand, since $g(x_l) > 0$, we also have that $g^{\uparrow}(\bar{x}; 1) \ge 0$. However, the inequalities $g(x_l) > 0$ and $g(x_l) < l^{-1}(x_l - \bar{x})$ contradict directly condition (9). Similarly, in case of $x_l < \bar{x}$ for all l, condition (10) is violated. In this way the first part of the statement has been established. Now assume that g is semismooth. According to the previous result, all we have to show now is that violation of one of the conditions (11) or (12) leads to a violation of calmness. Without loss of generality, let (11) be violated (the proof running analogously in the second case). Then, $g'(\bar{x}; 1) = 0$ and there is some sequence $x_l \downarrow \bar{x}$ such that $g(x_l) > 0$. If calmness held true, then $d_{M(0)}(x_l) \le L g(x_l)$ for some $L > 0$ and for l large enough. Choose $z_l \in M(0)$ such that $|z_l - x_l| = d_{M(0)}(x_l)$. In particular, $z_l \ge \bar{x}$, $z_l \ne x_l$, $g(z_l) \le 0$ and, by the mean value theorem for Clarke's subdifferential,

$$L^{-1} |z_l - x_l| \le g(x_l) \le g(x_l) - g(z_l) \le |\xi_l|\, |z_l - x_l|, \tag{14}$$

where $\xi_l \in \bar{\partial} g(u_l)$ and u_l belongs to the line segment joining x_l and z_l. Since $|z_l - x_l| \le |x_l - \bar{x}| \to 0$, we get $u_l \downarrow \bar{x}$. Now, the semismoothness of g at $\bar{x}$ entails that $\xi_l \to g'(\bar{x}; 1) = 0$. Since $z_l \ne x_l$, (14) provides the contradiction $L^{-1} \le 0$. Consequently, calmness is violated. □

The importance of the "or"- part in conditions (9),(10) can be illustrated by the function

$$g(x) = \begin{cases} -x & \text{if } x = n^{-1} \text{ for some } n \in \mathbb{N} \\ x & \text{otherwise,} \end{cases}$$

where the corresponding M is calm at (0, 0), but one also has that

$$0 \in \left[g^{\downarrow}(\bar{x}; 1), g^{\uparrow}(\bar{x}; 1)\right]$$

and g fails to be nonpositive on an interval $[\bar{x}, \bar{x} + \varepsilon]$.

Remark 1. The first result of Proposition 2 requires that $g(\bar{x}) = 0$. Indeed, the example

$$g(x) = \begin{cases} x - 1 & \text{if } x \leq 0 \\ x^2 & \text{if } x > 0 \end{cases}$$

shows that calmness of M may be violated for a lower semicontinuous function g which satisfies conditions (9),(10). The reason is that $g(\bar{x}) = -1$. However, as soon as g is continuous, calmness of M holds automatically true at any $\bar{x}$ with $g(\bar{x}) < 0$ due to $\bar{x}$ being an interior point of $M(0)$ then. Consequently, for investigating calmness of M when g is continuous (as in the second result of Proposition 2), one may assume $g(\bar{x}) = 0$ without loss of generality.

A trivial consequence of the definition is that calmness of M holds true whenever $\bar{x}$ is a local maximizer of g. If g is differentiable, this situation even covers the gap between calmness and the Aubin property in Banach spaces:

Proposition 3. *In (7), let X be a Banach space, $\Lambda = \mathbb{R}_-$ and $g : X \to \mathbb{R}$ be continuously differentiable in a neighborhood of $\bar{x} \in X$ such that $g(\bar{x}) = 0$. Then, M is calm at $(0, \bar{x})$ if and only if either this multifunction has the Aubin property around $(0, \bar{x})$ or $\bar{x}$ is a local maximizer of g.*

Proof. The Aubin property being equivalent with $\nabla g(\bar{x}) \neq 0$ here, all we have to show is that calmness is violated in the case when $\nabla g(\bar{x}) = 0$ and there exists a sequence $x_l \to \bar{x}$ with $g(x_l) > 0$. If calmness held true, then, as in the last lines of the proof of Proposition 2, there would exist a sequence z_l such that the following modification of (14) is valid with u_l belonging to the line segment $[x_l, z_l]$:

$$L^{-1} \|z_l - x_l\| \leq g(x_l) \leq g(x_l) - g(z_l) \leq \|\nabla g(u_l)\| \, \|z_l - x_l\| .$$

As in the proof of Proposition 2, $u_l \to \bar{x}$, whence $\nabla g(u_l) \to 0$. Again, the contradiction $L^{-1} \leq 0$ results. □

Remark 2. The differentiability of g is essential in the statement of Proposition 3, as one can see from the example $X = \mathbb{R}$, $g(x) = \max\{-x^2, x\}$, and $\bar{x} = 0$. Here, M is calm although neither it has the Aubin property nor $\bar{x}$ is a local maximizer of g. However, since g is semismooth, one may apply the second result of Proposition 2 in order to detect calmness.

3.2. Calmness of a single nonsmooth inequality

According to the previous section, there are simple criteria for calmness in the special case of a single inequality. In those criteria either the respective constraint function g is defined on $\mathbb{R}$ and then may be rather general or it is defined on a general Banach space and then has to be continuously differentiable. In many applications, of course, one will be faced with several differentiable inequalities or with a nondifferentiable inequality defined on more general spaces than $\mathbb{R}$. As far as calmness is concerned, Lemma 2 indicates, that the former task could be reduced to the latter one via the distance function. The following theorem provides a sufficient condition for calmness of a single nonsmooth inequality. This result will be exploited in later sections for the situation of several smooth constraints (not necessarily inequalities). In the following, for notational convenience, the expression bd $M(0) \setminus \{\bar{x}\}$ is supposed to mean (bd $M(0)) \setminus \{\bar{x}\}$, where "bd" refers to the topological boundary.

Theorem 1. *In (7), let $X = \mathbb{R}^n$, $Y = \mathbb{R}$, $\Lambda = \mathbb{R}_-$ and g be lower semicontinuous. M is calm at $(0, \bar{x})$, where $g(\bar{x}) = 0$, if the following conditions are satisfied:*

1. $g^{\downarrow}(\bar{x}; h) > 0 \quad \forall h \in \hat{N}_{M(0)}(\bar{x}) \setminus \{0\}$;
2. $\liminf\limits_{\substack{(z,h)\to(\bar{x},0) \\ (z,h)\in [\text{bd}\, M(0)\setminus\{\bar{x}\}]\times\left[\hat{N}_{M(0)}(z)\setminus\{0\}\right]}} \frac{g(z+h)}{\|h\|} > 0.$

Proof. Using the fact that $\hat{N}_{M(0)}(\bar{x})$ is a closed cone, it is easy to see that the two conditions of our Theorem imply the relation

$$\liminf_{\substack{(z,h)\to(\bar{x},0) \\ (z,h)\in \text{bd}\, M(0)\times\left[\hat{N}_{M(0)}(z)\setminus\{0\}\right]}} \frac{g(z+h)}{\|h\|} > 0. \tag{15}$$

By Corollary 1, violation of calmness entails the existence of some sequence $x_l \to \bar{x}$ such that $x_l \notin M(0)$ and $d_{M(0)}(x_l) > l g(x_l)$ for all $l \in \mathbb{N}$. Denote by z_l the Euclidean projection of x_l onto $M(0)$ and set $h_l := x_l - z_l$. Then, $z_l \in$ bd $M(0)$ and $h_l \in \hat{N}_{M(0)}(z_l) \setminus \{0\}$. From $x_l \to \bar{x} \in M(0)$ and $d_{M(0)}(x_l) = \|h_l\|$, it follows that $h_l \to 0$. Since

$$\frac{g(z_l + h_l)}{\|h_l\|} = \frac{g(x_l)}{d_{M(0)}(x_l)} < \frac{1}{l},$$

we get a contradiction with (15):

$$\liminf_{l\to\infty} \frac{g(z_l + h_l)}{\|h_l\|} \leq 0.$$

□

Remark 3. Conditions similar to those of the last Theorem can be found in many references on local error bounds or weak sharp minima, respectively (see, e.g., [2], [26], [36], [38]). The formulation in Theorem 1 appears to be quite favorable for the subsequent analysis. The reason to keep the conditions of Theorem 1 separate, rather than combining them to a single one as done in the proof, is to illustrate the addition to (ACQ) (related to condition 1.) which is necessary in order to obtain the stronger calmness property (compare Proposition 1 and Example 1). The two conditions of Theorem 1 will figure in adapted forms in several of the results below.

3.3. Calmness of differentiable constraints modeled by a finite union of polyhedra

In the following, we consider (7) for a continuously differentiable mapping g between finite-dimensional spaces and for a set Λ which is union of p convex polyhedra Λ_j. This framework allows the modeling of certain equilibrium constraints and incorporates conventional feasible sets of nonlinear optimization. It is easy to see (cf. [9]) that only finitely many cones can occur as $N_\Lambda(u)$, where $u \in \Lambda$. This allows to introduce the following finite family of cones for some fixed $\bar{x} \in \mathbb{R}^n$:

$$\mathcal{N} := \{N | \exists\, x_i \overset{\operatorname{bd} M(0)\setminus\{\bar{x}\}}{\longrightarrow} \bar{x}\ \exists j \in \{1, 2, \dots, p\} : \\ g(x_i) \in \Lambda_j \text{ and } N = N_{\Lambda_j}(g(x_i)) \text{ for all } i \in \mathbb{N}\}.$$

In the following, ∇g shall refer to the Jacobian of g.

Theorem 2. *Consider (1) with $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$, $g \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^m)$ and $\Lambda = \bigcup_{j=1}^p \Lambda_j \subseteq \mathbb{R}^m$, where each Λ_j is a convex polyhedron. Then, M is calm at some $(0, \bar{x}) \in \operatorname{Gph} M$ under the following two assumptions:*

1. *$T_{M(0)}(\bar{x}) = \{h \in \mathbb{R}^n \mid \nabla g(\bar{x})h \in T_\Lambda(g(\bar{x}))\}$ (i. e. (ACQ) holds at $\bar{x}$);*
2. *$N \cap \ker (\nabla g(\bar{x}))^T = \{0\} \quad \forall N \in \mathcal{N}$.*

Proof. By Lemma 2, it is sufficient to show the calmness of the multifunction

$$\tilde{M}(t) = \{x \in X \mid d_\Lambda(g(x)) \le t\}$$

at $(0, \bar{x})$. This will be done on the basis of Theorem 1 applied to the function $b := d_\Lambda \circ g$. Put

$$\mathbb{I}(x) := \{j \in \{1, \dots, p\} \mid g(x) \in \Lambda_j\}.$$

Since d_{Λ_j} is convex continuous, the composition $b_j := d_{\Lambda_j} \circ g$ is directionally differentiable, and for all $j \in \mathbb{I}(x)$ and $h \in \mathbb{R}^n$ one has

$${b_j}'(x; h) = d'_{\Lambda_j}(g(x); \nabla g(x)h) = d_{T_{\Lambda_j}(g(x))}(\nabla g(x)h),$$

(cf. [35], Example 8.53). Clearly, $b = \min\{b_j | j \in \{1, 2, \dots, p\}\}$. By a continuity argument one even has the identity

$$b(x + u) = \min_{j \in \mathbb{I}(x)} b_j(x + u) \tag{16}$$

for all $x \in M(0)$ and all u sufficiently close to x. Consequently, for all $x \in M(0)$ and all h,

$$\begin{aligned} b'(x; h) &= \lim_{\lambda \downarrow 0} \lambda^{-1} (b(x + \lambda h) - b(x)) = \lim_{\lambda \downarrow 0} \lambda^{-1} \left(\min_{j \in \mathbb{I}(x)} b_j(x + \lambda h) \right) \\ &= \min_{j \in \mathbb{I}(x)} \lim_{\lambda \downarrow 0} \lambda^{-1} \left(b_j(x + \lambda h) - b_j(x) \right) = \min_{j \in \mathbb{I}(x)} {b_j}'(x; h) \\ &= \min_{j \in \mathbb{I}(x)} d_{T_{\Lambda_j}(g(x))}(\nabla g(x)h) = d_{\cup\{T_{\Lambda_j}(g(x)) | \, j \in \mathbb{I}(x)\}}(\nabla g(x)h) \\ &= d_{T_\Lambda(g(x))}(\nabla g(x)h). \end{aligned}$$

Here, we used that $b(x) = b_j(x) = 0$ for all $j \in \mathbb{I}(x)$. Along with our assumption 1., the obtained relation yields that $b^{\downarrow}(\bar{x}; h) = b'(\bar{x}; h) > 0$ for all $h \in \hat{N}_{M(0)}(\bar{x})\backslash\{0\}$, which is the first condition of Theorem 1. To verify the second one, consider an arbitrary sequence

$$(z_l, h_l) \to (\bar{x}, 0), z_l \in \text{bd } M(0)\backslash\{\bar{x}\}, h_l \in \hat{N}_{M(0)}(z_l)\backslash\{0\}.$$

Clearly, $g(z_l) \in \Lambda$, and, by the finiteness argument, one may pass to a subsequence (which will not be relabeled) such that $\mathbb{I}(z_l)$ amounts to a fixed index set $\mathbb{I}^*$ and, for each $j \in \mathbb{I}^*$, the normal cones $N_{\Lambda_j}(g(z_l))$ reduce to some fixed closed convex cones N_j for all $l \in \mathbb{N}$. By definition, all these cones N_j belong to $\mathcal{N}$. Setting $\tilde{h}_l := \|h_l\|^{-1}h_l$, one may pass to another subsequence (again not relabeled) such that $\tilde{h}_l \to \tilde{h}$ with $\|\tilde{h}\| = 1$. Since $h_l \in \hat{N}_{M(0)}(z_l)$ and $M(0) = \cup_{j=1}^{p} g^{-1}(\Lambda_j)$, it follows that $h_l \in \cap_{j\in\mathbb{I}^*}\hat{N}_{g^{-1}(\Lambda_j)}(z_l)$. Here, we have used the existence of some open neighbourhood U of z_l such that

$$M(0) \cap U = \left(\cup_{j\in\mathbb{I}^*} g^{-1}(\Lambda_j)\right) \cap U.$$

On the other hand, our assuption 2. ensures that $N_j \cap \ker(\nabla g(z_l))^T = \{0\}$ for l sufficiently large. This constraint qualification allows to apply Theorem 6.14 in [35] and to derive that $\hat{N}_{g^{-1}(\Lambda_j)}(z_l) = (\nabla g(z_l))^T N_j$. We show now that

$$\tilde{h} \in (\nabla g(\bar{x}))^T N_j \cap \mathbb{S} \quad \forall j \in \mathbb{I}^*. \tag{17}$$

Indeed, for an arbitrary fixed $j \in \mathbb{I}^*$, one has that $\tilde{h}_l = (\nabla g(z_l))^T k_l$ with $k_l \in N_j$ and it suffices to verify that the sequence $\{k_l\}$ is bounded. Taking into account that $\left\|(\nabla g(z_l))^T k_l\right\| = 1$, this follows, however, immediately from our assumption 2. Therefore, relation (17) holds true.

Now, since each Λ_j is convex, one has for all $j \in \mathbb{I}^*$ that $\Lambda_j - g(z_l) \subset T_{\Lambda_j}(g(z_l))$. Consequently,

$$\begin{aligned} b_j(z_l + h_l) &= d_{\Lambda_j}(g(z_l + h_l) \geq d_{T_{\Lambda_j}(g(z_l))}(g(z_l + h_l) - g(z_l)) \\ &= d'_{\Lambda_j}(g(z_l); (g(z_l + h_l) - g(z_l))) = \max_{\xi\in N_j\cap\mathbb{B}} \langle \xi, g(z_l + h_l) - g(z_l)\rangle, \end{aligned}$$

where the last two equalities follow from Example 8.53 in [35]. Since g is continuously differentiable, it is strictly differentiable at $\bar{x}$ and one has

$$\|h_l\|^{-1}(g_i(z_l + h_l) - g_i(z_l)) \to \left\langle \nabla g_i(\bar{x}), \tilde{h}\right\rangle,$$

so that

$$\left\langle \xi, \|h_l\|^{-1}(g(z_l + h_l) - g(z_l))\right\rangle \to \left\langle (\nabla g(\bar{x}))^T \xi, \tilde{h}\right\rangle.$$

From (17), we know that $\tilde{h} = (\nabla g(\bar{x}))^T \tilde{k}$ for some $\tilde{k} \in N_j\backslash\{0\}$. Recalling, that a function $\max_{\xi\in K}\langle\xi, \Psi(\cdot)\rangle$ with K convex compact and Ψ continuous is continuous, we may summarize that, for all $j \in \mathbb{I}^*$,

$$\begin{aligned}\liminf_{l\to\infty} \|h_l\|^{-1} b_j(z_l + h_l) &\geq \liminf_{l\to\infty} \max_{\xi\in N_j\cap\mathbb{B}} \left\langle \xi, \|h_l\|^{-1} \left(g(z_l + h_l) - g(z_l)\right)\right\rangle \\ &= \max_{\xi\in N_j\cap\mathbb{B}} \left\langle (\nabla g(\bar{x}))^T \xi, \tilde{h}\right\rangle \\ &\geq \left\langle (\nabla g(\bar{x}))^T \left(\|\tilde{k}\|^{-1}\tilde{k}\right), (\nabla g(\bar{x}))^T \tilde{k}\right\rangle \\ &= \|\tilde{k}\|^{-1} \|(\nabla g(\bar{x}))^T \tilde{k}\,\|^2 > 0\end{aligned}$$

in view of our assumption 2. Referring to (16), it follows that

$$\begin{aligned}\liminf_{l\to\infty} \|h_l\|^{-1} b(z_l + h_l) &= \liminf_{l\to\infty} \min_{j\in\mathbb{I}^*} \|h_l\|^{-1} b_j(z_l + h_l) \\ &= \min_{j\in\mathbb{I}^*} \liminf_{l\to\infty} \|h_l\|^{-1} b_j(z_l + h_l) > 0.\end{aligned}$$

This establishes condition 2. of Theorem 1 and completes the proof. □

Remark 4. From the proof of Theorem 2 it is clear that one may replace condition 1. by the weaker condition

$$\hat{N}_{M(0)}(\bar{x}) \cap \{h \in \mathbb{R}^n | \nabla g(\bar{x})h \in T_\Lambda(g(\bar{x}))\} = \{0\}.$$

This is particularly efficient in situations where $\hat{N}_{M(0)}(\bar{x}) = \{0\}$ as in Example 3 below. With this condition, however, there is no real gain in the statement of Theorem 2 because calmness implies (ACQ) (see Prop. 1).

Three examples shall illustrate the application of Theorem 2.

Example 2. Consider the nonlinear complementarity problem (NCP) governed by the generalized equation (GE)

$$0 \in f(x) + N_{\mathbb{R}_+}(x) \tag{18}$$

with

$$f(x) = \begin{cases} -x^2 & \text{for } x < 0 \\ 0 & \text{for } x \in [0, 1] \\ (x-1)^2 & \text{for } x > 1 \end{cases}.$$

Clearly, this problem can be rewritten as $g(x) \in \Lambda$ with

$$g(x) = (x, -f(x))^T \text{ and } \Lambda = \operatorname{Gph} N_{\mathbb{R}_+} = (\mathbb{R}_+ \times \{0\}) \cup (\{0\} \times \mathbb{R}_-).$$

Note that Λ is the union of two convex polyhedra (half lines). It is easily seen that $M(0) = [0, 1]$ holds true for the multifunction M in (7). We examine calmness of M at $(0, 0) \in \operatorname{Gph} M$. Condition 2. of Theorem 2 is automatically fulfilled because there is no sequence $x_i \to 0$ with $x_i \in \operatorname{bd} M(0)\backslash\{0\}$. Condition 1. of Theorem 2 is also satisfied due to

$$T_{M(0)}(0) = \mathbb{R}_+ = \{h \in \mathbb{R} | (h, 0) \in \Lambda\} = \{h \in \mathbb{R} | \nabla g(0)h \in T_\Lambda(g(0))\}.$$

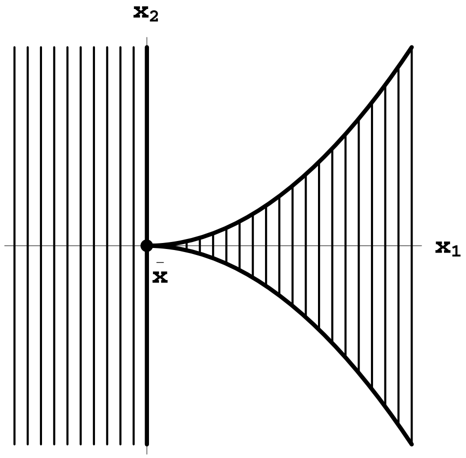

Fig. 1. Illustration of the set $M(0)$ in Example 3

Consequently, M is calm at $(0,0)$. Observe, however, that M does not possess the Aubin property at $(0,0)$. Indeed, one has $M(0,\varepsilon) = \{1+\sqrt{\varepsilon}\}$ for $\varepsilon > 0$ which implies that $M(0,\varepsilon) \cap \mathbb{B}(0,1) = \emptyset$ in contradiction with the Aubin property. Therefore, calmness cannot be detected here as a consequence of the Aubin property.

Example 3. Let

$$g(x_1,x_2) = (-x_1^2 + x_2, -x_1^2 - x_2, x_1)^T,$$

$\bar{x} = 0$ and $\Lambda = \Lambda_1 \cup \Lambda_2$ with $\Lambda_1 = \mathbb{R}^2 \times \mathbb{R}_-$ and $\Lambda_2 = \mathbb{R}^2_- \times \mathbb{R}_+$. The set $M(0)$ is illustrated in Figure 1.

It is easily calculated that $(1,1,0) \in N_\Lambda(g(\bar{x})) \cap \ker\,(\nabla g(\bar{x}))^T$. Hence, the calmness of the multifunction M in (7) at $(0,0)$ cannot be ensured by (GMFCQ) (cf. (5)). On the other hand, the condition of Remark 4 is trivially fulfilled due to $\hat{N}_{M(0)}(\bar{x}) = \{0\}$. This entails the weakened condition 1. of Theorem 2. As for condition 2. of that theorem, note that the family $\mathcal{N}$ consists of the three cones

$$N_1 = \mathbb{R}_+ \times \{0\} \times \{0\},\, N_2 = \{0\} \times \mathbb{R}_+ \times \{0\},\, N_3 = \{0\} \times \{0\} \times \mathbb{R}_+.$$

Since $N_i \cap \ker\,(\nabla g(\bar{x}))^T = \{0\}$ for $i = 1,2,3$, condition 2. holds true as well and calmness follows.

Example 4. Consider the parameter-dependent NCP governed by the GE $0 \in f(x_1,x_2) + N_{\mathbb{R}_+}(x_2)$ with $f(x_1,x_2) = x_1^2 - x_2$ together with the parameter constraint $x_1 \le 0$. Again, this can be written as $g(x) \in \Lambda$, where

$$g(x) = (x_1, x_2, -f(x_1,x_2))^T \text{ and } \Lambda = \mathbb{R}_- \times \operatorname{Gph} N_{\mathbb{R}_+}.$$

Now, Λ is the union of two convex polyhedra. For the multifunction M in (7) one computes

$$M(0) = (\mathbb{R}_- \times \{0\}) \cup \{(x_1,x_2) \in \mathbb{R}_- \times \mathbb{R} | x_1^2 = x_2\}.$$

Calmness of M shall be examined at $(0, 0) \in \text{Gph}\, M$. First note that

$$(0, -1, 1)^T \in N_\Lambda\,(g(0, 0)) \cap \ker\,(\nabla g(0, 0))^T \neq \{0\},$$

which means that, again, (GMFCQ) is violated and, thus, cannot be applied in order to detect calmness. On the other hand, condition 1. of Theorem 2 is fulfilled because

$$\begin{aligned} T_{M(0)}(0) = \mathbb{R}_- \times \{0\} &= \{h \in \mathbb{R}^2 | \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \in \Lambda\} \\ &= \{h \in \mathbb{R}^2 | \nabla g(0) h \in T_\Lambda(g(0))\}. \end{aligned}$$

Further note that the family $\mathcal{N}$ in Theorem 2 consists of the two cones

$$N_1 = \{0\} \times \{0\} \times \mathbb{R}, \quad N_2 = \{0\} \times \mathbb{R} \times \{0\}.$$

Since

$$N_i \cap \ker\,(\nabla g(0))^T = \{0\} \quad (i = 1, 2),$$

condition 2. of Theorem 2 is also satisfied and calmness of M at the origin has been established.

As an application of Theorem 2 consider the special case

$$g(x) = Ax + c, \tag{19}$$

for some (m, n)- matrix A and some $c \in \mathbb{R}^m$. From Robinson's well-known theorem in [32] it follows that the multifunction M in (7) with g defined in (19) is calm at $(0, \bar{x})$ for each $\bar{x} \in M(0)$. Next we show, how this result can alternatively be derived from Theorem 2. We start with a preparatory statement.

Proposition 4. *Consider the setting of Theorem 2 with* $p = 1$ *(i.e.,* Λ *itself is a convex polyhedron). Then* M *in (7) with* g *defined in (19) is calm at* $(0, \bar{x})$ *for each* $\bar{x} \in M(0)$.

Proof. It is well-known that condition 1. of Theorem 2 is satisfied for our data (see [3]). Concerning condition 2. of Theorem 2 we get back to the sequences $\{z_l\}$, $\{\tilde{h}_l\}$ specified in the proof of that theorem. Due to the form of g, one has $\hat{N}_{M(0)}(z_l) = A^T N$ with some fixed closed convex cone N whenever l is sufficiently large. This implies that $\tilde{h} \in A^T N$ as well. Simultaneously, $T_{M(0)}(z_l) = (A^T N)^0 = \{k \in \mathbb{R}^n \mid Ak \in N^0\}$ and we denote this fixed convex cone by T. Following the proof of Theorem 2, it remains to show that

$$\max_{\xi \in N \cap \mathbb{B}} \langle A^T \xi, \tilde{h} \rangle > 0. \tag{20}$$

Assume by contradiction that

$$\langle \xi, A\tilde{h} \rangle \leq 0 \,\forall \xi \in N \cap \mathbb{B}.$$

This implies, however, that $A\tilde{h} \in N^0$, i.e., $\tilde{h} \in T$. On the other hand, the intersection of negative polar cones cannot contain a nonzero element. Thus, inequality (20) holds true and we conclude that condition 2. of Theorem 2 is satisfied. □

Consider now the multifunction M with g given by (19) and

$$\Lambda = \bigcup_{j=1}^{p} \Lambda_j,$$

where the Λ_j are convex polyhedra. With $M_j : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ defined by

$$M_j(y) := \{x \in \mathbb{R}^n | Ax + c + y \in \Lambda_j\} \quad (j = 1, \dots, p),$$

it is easy to see that

$$\text{Gph}\, M = \bigcup_{j=1}^{p} \text{Gph}\, M_j.$$

Now we may use an idea from [35] (Example 9.57): Let $(\bar{x}, 0) \in \text{Gph}\, M$ so that $(\bar{x}, 0) \in \text{Gph}\, M_j$ for $j \in \mathbb{I}\,(\bar{x})$. By virtue of Proposition 4, there exist $l_j, \varepsilon_j \geq 0$, such that

$$d_{M_j(0)}(x) \leq l_j \|y\| \quad \forall y \in \mathbb{B}(0, \varepsilon_j)\, \forall x \in \mathbb{B}(\bar{x}, \varepsilon_j) \cap M_j(y).$$

Consequently, with

$$l := \max_{j \in \mathbb{I}(\bar{x})} l_j, \quad \varepsilon := \min_{j \in \mathbb{I}(\bar{x})} \varepsilon_j,$$

one has

$$d_{M_j(0)}(x) \leq l \|y\| \quad \forall y \in \mathbb{B}(0, \varepsilon)\, \forall x \in \mathbb{B}(\bar{x}, \varepsilon) \cap M_j(y)\, \forall j \in \mathbb{I}\,(\bar{x})\,.$$

This amounts, however, to the calmness of M at $(\bar{x}, 0)$.

3.4. Calmness of finitely many differentiable inequalities

As a further application of Theorem 2 we characterize calmness of a finite system of smooth inequalities, i.e., $\Lambda = \mathbb{R}^m_-$. Let

$$I(x) := \{i \in \{1, \dots, m\} \mid g_i(x) = 0\}$$

be the set of active indices at x. The standard results on characterization of calmness of M mentioned in the introduction amount to the following conditions:

$$\text{(MFCQ)} \;\; 0 \notin \text{conv}\, \{\nabla g_i(\bar{x}) | i \in I(\bar{x})\} \tag{21}$$

$$\text{(see (6))} \;\; 0 \in \text{int}\, \text{conv}\, \{\nabla g_i(\bar{x}) | i \in I(\bar{x})\} \;\; . \tag{22}$$

Simple examples show that in the remaining case $0 \in \text{bd}\, \text{conv}\, \{\nabla g_i(\bar{x}) | i \in I(\bar{x})\}$ calmness can be violated or satisfied (take $g_1(x) = x$ and $g_2(x) = 0$ or $g_2(x) = x^2$). The application of Theorem 2, however, will provide a condition which allows to detect calmness of M also in this case. Let $\mathcal{J}$ be the family of critical index sets $I \subseteq I(\bar{x})$, defined by

$$\mathcal{J} := \{I | \exists\, x_i \stackrel{\text{bd}\, M(0) \setminus \{\bar{x}\}}{\longrightarrow} \bar{x} : \; I = I(x_i)\, \forall\, i \in \mathbb{N}\}.$$

Theorem 3. *Consider (7) with* $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$, $g \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^m)$ *and* $\Lambda = \mathbb{R}^m_-$. *Then, M is calm at some* $(0, \bar{x}) \in \mathrm{Gph}\, M$ *under the following two assumptions:*

1. $T_{M(0)}(\bar{x}) = \{h \in \mathbb{R}^n | \nabla g_i(\bar{x})h \leq 0 \quad \forall i \in I(\bar{x})\}$;
2. $0 \notin \mathrm{conv}\ \{\nabla g_i(\bar{x}) | i \in I\}\ \forall I \in \mathcal{J}$.

Proof. Condition 1. above is just the classical (ACQ), i. e., the specification of condition 1. in Theorem 2 to the setting considered here. Since for an arbitrary point $x \in M(0)$

$$\hat{N}_{\mathbb{R}^m_-}(g(x)) = \{k \in \mathbb{R}^m_+ | k_i = 0 \text{ for } i \notin I(x)\},$$

condition 2. of Theorem 2 reduces to the condition that, for all $I \in \mathcal{J}$ one has the implication

$$(\nabla g(\bar{x}))^T k = 0, k \in \mathbb{R}^m_+, k_i = 0 \text{ if } i \notin I \Longrightarrow k = 0.$$

This, however, is equivalent to $0 \notin \mathrm{conv}\ \{\nabla g_i(\bar{x}),\ i \in I\}\ \forall I \in \mathcal{J}$. □

Corollary 2. *In the setting of Theorem 3, let $\bar{x}$ be an isolated point of $M(0)$, i.e., $M(0) \cap \mathcal{U} = \{\bar{x}\}$ for some neighborhood $\mathcal{U}$ of $\bar{x}$. Then, calmness of M at $(0, \bar{x})$ is equivalent with (ACQ) at $\bar{x}$.*

Proof. The isolatedness assumption immediately implies that $\mathcal{J} = \emptyset$, whence condition 2. of Theorem 3 is automatically satisfied. Therefore, condition 1. implies calmness. The reverse implication follows from Proposition 1. □

Corollary 2 clarifies, why Example 1 could work as an even smooth counter-example to Proposition 1: The point $\bar{x} = 0$ in that example failed to be isolated in $M(0)$.

In the setting of Theorem 3, one gets the following criterion for *calmness on selections* (see Introduction):

Corollary 3. *M is calm on selections at $(0, \bar{x}) \Longleftrightarrow \bar{x}$ is an isolated point of $M(0)$ and (ACQ) holds at $\bar{x}$.*

Proof. The definition of calmness on selections (see Introduction) implies its equivalence with usual calmness complemented by local isolatedness of $\bar{x}$ in $M(0)$. This, by virtue of Corollary 2, is equivalent with (ACQ) complemented by local isolatedness of $\bar{x}$. □

The use of Theorem 3 as a condition for calmness is emphasized by the following observation:

Proposition 5. *Each of the conditions (21) and (22) implies the two conditions of Theorem 3. Both implications are strict.*

Proof. It is well known that (MFCQ) implies (ACQ) (see, e. g., [35, Theorem 6.31]). The upper semi-continuity of the index set mapping $x \mapsto I(x)$ entails that $I \subseteq I(x)$ for all $I \in \mathcal{I}$. Thus condition 2. of Theorem 3 follows from (21) too. As to (22), it implies that $\bar{x}$ is an isolated point of $M(0)$ (see discussion of (6) in the introduction). Hence, condition 2. of Theorem 3 is trivially satisfied. Moreover, isolatedness means that $T_{M(0)}(\bar{x}) = \{0\}$. Finally, under (22), 0 cannot be separated from the set conv $\{\nabla g_i(\bar{x}) | i \in I(\bar{x})\}$. Therefore, the right-hand side of condition 1. in Theorem 3 reduces to $\{0\}$, so this condition is met as well. The subsequent examples will show that the implications in this proposition are strict. □

The first two of the following examples illustrate the application of Theorem 3. In both of them, conditions (21) and (22) are violated. In the third example the respective M is not calm. We always put $\bar{x} = 0$.

1. $g_1(x) = -x^2$, $g_2(x) = x$: Then,

$$M(0) = T_{M(0)}(\bar{x}) = \{h \in \mathbb{R} | \nabla g_i(\bar{x})h \leq 0 \quad \forall i \in I(\bar{x}) = \{1, 2\}\} = \mathbb{R}_-.$$

   Since bd $M(0) = \{\bar{x}\}$, it results that $\mathcal{J}$ is an empty family of index sets and, hence, condition 2. of Theorem 3 is trivially fulfilled. Therefore, M is calm at $(0, 0)$.
2. $g_1(x_1, x_2) = x_2 - x_1^2$, $g_2(x_1, x_2) = -x_2 - x_1^2$, $g_3(x_1, x_2) = -x_1$: Then,

$$M(0) = \{(x_1, x_2) | \, |x_2| \leq x_1^2, \ x_1 \geq 0\} \text{ and}$$
$$T_{M(0)}(\bar{x}) = \{h \in \mathbb{R}^2 | \nabla g_i(\bar{x})h \leq 0 \quad \forall i \in I(\bar{x}) = \{1, 2, 3\}\} = \mathbb{R}_+ \times \{0\}.$$

   Moreover, we have that $\mathcal{J} = \{\{1\}, \{2\}\}$ (the third inequality never becomes active at $M(0) \setminus \{\bar{x}\}$). Since $\nabla g_1(\bar{x}) = (0, 1) \neq 0$ and $\nabla g_2(\bar{x}) = (0, -1) \neq 0$, condition 2. of Theorem 3 is fulfilled. Thus, M is calm at $(0, 0)$.
3. $g_1(x) = x^2$, $g_2(x) = x$: One easily verifies that M is not calm at $(0, 0)$. Then, (ACQ) is violated because $M(0) = T_{M(0)}(\bar{x}) = \{0\}$ and

$$\{h \in \mathbb{R} | \nabla g_i(\bar{x})h \leq 0 \quad \forall i \in I(\bar{x}) = \{1, 2\}\} = \mathbb{R}_-.$$

3.5. Calmness of infinitely many differentiable inequalities

The idea developed in Theorem 3 can be also applied to the case of another multifunction M, where y is an infinite-dimensional parameter. Let $T \subseteq \mathbb{R}^m$ be compact and denote by $\mathcal{C}(T)$ the Banach space of continuous functions on T equipped with the maximum norm. Let $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ be continuously differentiable such that $\nabla_x g$ is locally Lipschitzian (which is satisfied, for instance, if g is of class $\mathcal{C}^2$ or even $\mathcal{C}^{1,1}$). Consider the multifunction $M : \mathcal{C}(T) \rightrightarrows \mathbb{R}^n$ defined by

$$M(y) := \{x \in \mathbb{R}^n | g(x, z) \leq -y(z) \quad \forall z \in T\}. \tag{23}$$

Evidently, one may equivalently write (23) as

$$M(y) := \{x \in \mathbb{R}^n | \tilde{g}(x) + y \in \Lambda\}, \tag{24}$$

where $\tilde{g}(x) := g(x, \cdot)$ and Λ refers to the cone of nonpositive, continuous functions on T. For any $x \in \mathbb{R}^n$, the set of active indices will be denoted by

$$I(x) := \{z \in T | g(x, z) = G(x)\}, \quad \text{where} \quad G(x) = \max\{g(x, z) | z \in T\}. \tag{25}$$

It is well known that G is locally Lipschitzian and Clarke-regular. In particular, G is directionally differentiable and one has

$$G'(x; h) = \max\{\langle \nabla_x g(x, z), h \rangle \, | z \in I(x)\} \tag{26}$$

(note that writing “max” is justified here due to the compactness of $I(x)$). Assume that $\bar{x} \in \mathbb{R}^n$ satisfies $G(\bar{x}) = 0$, hence $(0, \bar{x}) \in \operatorname{Gph} M$. Finally, we introduce the following family of critical index sets:

$$\mathcal{J} := \{S \subseteq T \,|\, \exists x_i \overset{\operatorname{bd} M(0)\setminus\{\bar{x}\}}{\longrightarrow} \bar{x} : d_H(S, I(x_i)) \to 0\}.$$

Here, d_H refers to the Hausdorff distance between compact sets.

We shall need the following auxiliary result:

Lemma 3. *Let $K \subset \mathbb{R}^n$ be a closed convex set such that $0 \notin K \subseteq L\mathbb{B}$ for some $L > 0$. Then,*

$$\max_{k \in K} \langle k, h \rangle \geq L^{-1} \|\xi\|^2 \|h\| \quad \forall h \in \mathbb{R}_+ K,$$

where ξ is the norm-minimal element in K.

Proof. Since ξ is a norm-minimal element in K, one has $\|\xi\|^2 \leq \langle \xi, h \rangle$ for all $h \in K$. Consequently,

$$\max_{k \in K} \langle k, h \rangle \geq \langle \xi, h \rangle \geq L^{-1} \|\xi\|^2 \|h\| \quad \forall h \in K.$$

Since both sides of the last inequality are positively homogeneous in h, the same inequality holds true for all $h \in \mathbb{R}_+ K$. □

Theorem 4. *Consider (7) with $X := \mathbb{R}^n$, $Y := \mathcal{C}(T)$ and M given by (24) (where $\tilde{g}$ plays the role of g in (7)). Let $(0, \bar{x}) \in \mathcal{C}(T) \times \mathbb{R}^n$ such that $G(\bar{x}) = 0$, i.e., $g(\bar{x}, z) \leq 0$ for all $z \in T$, and there exists some $\bar{z} \in T$ with $g(\bar{x}, \bar{z}) = 0$. Assume that*

1. *$T_{M(0)}(\bar{x}) = \{h \in \mathbb{R}^n \,|\, \langle \nabla_x g(\bar{x}, z), h \rangle \leq 0 \quad \forall z \in I(\bar{x})\}$.*
2. *There is some $\rho > 0$ such that $d_{\operatorname{conv} \{\nabla_x g(\bar{x}, z) | z \in S\}}(0) \geq \rho \quad$ for all $S \in \mathcal{J}$.*

Then, M is calm at $(0, \bar{x})$.

Proof. According to Lemma 2, calmness of M at $(0, \bar{x}) \in \mathcal{C}(T) \times \mathbb{R}^n$ is equivalent with the calmness of

$$\tilde{M}(t) := \{x \in \mathbb{R}^n | d_\Lambda \tilde{g}(x) \leq t\} = \{x \in \mathbb{R}^n \,|\, \max\{G(x), 0\} \leq t\}$$

at $(0, \bar{x}) \in \mathbb{R} \times \mathbb{R}^n$. The definition of calmness immediately yields that, another time, calmness of $\tilde{M}$ at $(0, \bar{x})$ is equivalent with the calmness at $(0, \bar{x})$ of

$$M^*(t) := \{x \in \mathbb{R}^n | G(x) \leq t\}.$$

Hence, we are going to verify this last property on the basis of Theorem 1 (with the function g there replaced by our function G here). By our assumption 1. we have that $\hat{N}_{M(0)}(\bar{x}) = \left(L_{M(0)}(\bar{x})\right)^0$. Then, (26) provides condition 1. of Theorem 1:

$$G^{\downarrow}(\bar{x}; h) = G'(\bar{x}; h) = \max\{\langle \nabla_x g(\bar{x}, z), h \rangle \,| z \in I(\bar{x})\} > 0 \quad \forall h \in \hat{N}_{M(0)}(\bar{x}) \diagdown \{0\}.$$

In order to check condition 2. of Theorem 1, consider arbitrary sequences $x_l \to \bar{x}$ and $h_l \to 0$ such that $x_l \in \operatorname{bd} M(0) \diagdown \{\bar{x}\}$ and $h_l \in \hat{N}_{M(0)}(\bar{x}) \diagdown \{0\}$. Denote by $c > 0$ a

Lipschitz modulus of $\nabla_x g$ on the compact set $\mathbb{B}(\bar{x}, 1) \times T$. We verify the following relation:

$$\exists l_0 \, \forall l \geq l_0 \, \exists S \in \mathcal{J} : I(x_l) \subseteq S + \mathbb{B}(0, (4c)^{-1}\rho), \tag{27}$$

where $\rho > 0$ refers to our condition 2. If the relation did not hold true, then there would be subsequences $\{x_l\}$, $\{z_l\}$ which we do not relabel, such that $z_l \in I(x_l)$ and $d_S(z_l) > (4c)^{-1}\rho$ for all l and all $S \in \mathcal{J}$. Since the space of compact subsets of $\mathbb{R}^m$ endowed with the Hausdorff metric is itself compact, there is some compact $\tilde{S} \subseteq T$ along with another subsequence $\{x_l\}$, which again we do not relabel, such that $d_H(\tilde{S}, I(x_l)) \to 0$. By definition, $\tilde{S} \in \mathcal{J}$. Finally, after passing yet to another subsequence, we have that $z_l \to \bar{z}$ for some $\bar{z} \in T$. Consequently, $\bar{z} \in \tilde{S}$, which contradicts $d_{\tilde{S}}(z_l) > (4c)^{-1}\rho$ for all l. This proves (27).

In addition to (27), we may assume that $\|x_l - \bar{x}\| < (4c)^{-1}\rho$ for all $l \geq l_0$. Now, we fix an arbitrary $l \geq l_0$ and an arbitrary $z \in I(x_l)$. By $S \in \mathcal{J}$, we denote the set whose existence is guaranteed in (27) and by $z^* \in S$ the Euclidean projection of z onto S. Then, due to (27), we get

$$\left\| \nabla_x g(x_l, z) - \nabla_x g(\bar{x}, z^*) \right\| \leq c(\|x_l - \bar{x}\| + \left\| z - z^* \right\|) \leq \rho/2.$$

Our assumption 2., along with a separation argument, ensures the existence of some x^* with $\|x^*\| = 1$ and

$$\left\langle x^*, v \right\rangle \geq \rho \geq \left\langle x^*, u \right\rangle \quad \forall v \in \operatorname{conv}\{\nabla_x g(\bar{x}, z) | z \in S\} \, \forall u \in \mathbb{B}(0, \rho).$$

Then, since $z \in I(x_l)$ was arbitrary, one derives

$$\begin{aligned} \left\langle x^*, \nabla_x g(x_l, z) \right\rangle &\geq \left\langle x^*, \nabla_x g(\bar{x}, z^*) \right\rangle - \left\| \nabla_x g(x_l, z) - \nabla_x g(\bar{x}, z^*) \right\| \\ &\geq \rho - \rho/2 = \rho/2 \geq \left\langle x^*, u \right\rangle \quad \forall z \in I(x_l) \, \forall u \in \mathbb{B}(0, \rho/2). \end{aligned}$$

It follows that $\operatorname{conv}\{\nabla_x g(x_l, z) | z \in I(x_l)\} \cap \operatorname{int} \mathbb{B}(0, \rho/2) = \emptyset$. Since $l \geq l_0$ was arbitrary, we have that

$$d_{\operatorname{conv}\{\nabla_x g(x_l, z) | z \in I(x_l)\}}(0) \geq \rho/2 \quad \forall l \geq l_0. \tag{28}$$

In particular, $0 \notin \operatorname{conv}\{\nabla_x g(x_l, z) | z \in I(x_l)\} = \partial G(x_l)$. This constraint qualification along with the regularity of G ensures that $\hat{N}_{M(0)}(x_l) = \mathbb{R}_+ \partial G(x_l)$ (cf. Prop. 10.3. in [35]). Accordingly, $\|h_l\|^{-1} h_l \in \mathbb{R}_+ \partial G(x_l)$. The continuity of the gradients $\nabla_x g$ implies the existence of some $L > 0$ such that $K_l \subseteq L\mathbb{B}$ for l large enough. Now, Lemma 3 and (28) ensure that

$$\max_{k \in \partial G(x_l)} \langle k, \|h_l\|^{-1} h_l \rangle \geq L^{-1} \left(d_{\partial G(x_l)}(0) \right)^2 \geq L^{-1} \rho^2 / 4 \quad \forall l \geq l_0.$$

We assume also l_0 large enough to meet the condition $\max\{\|x_l - \bar{x}\|, \|h_l\|\} \leq 1/2$ whenever $l \geq l_0$. Now, fix an arbitrary $l \geq l_0$ and put

$$\alpha(h, z) := g(x_l + h, z) - g(x_l, z) - \langle \nabla_x g(x_l, z), h \rangle.$$

Clearly, α is continuous and, by the mean value theorem and by $\nabla_x g$ having Lipschitz modulus $c > 0$ on $\mathbb{B}(\bar{x}, 1) \times T$, one gets that

$$|\alpha(h, z)| \leq |\langle \nabla_x g(x_l + \Theta_{h,z} h, z) - \nabla_x g(x_l, z), h \rangle| \leq c\Theta_{h,z} \|h\|^2 \\ \forall (h, z) \in \mathbb{B}(0, 1/2) \times T,$$

where $\Theta_{h,z} \in [0, 1]$. This implies

$$\|h\|^{-1} |\alpha(h, z)| \leq c \|h\| \quad \forall (h, z) \in (\mathbb{B}(0, 1/2) \setminus \{0\}) \times T.$$

We note that $x_l \in \operatorname{bd} M(0)$ entails $G(x_l) = 0$ by continuity of G and, hence, $g(x_l, z) = 0$ for all $z \in I(x_l)$. Then, the following estimation holds true for all $l \geq l_0$:

$$\begin{aligned} \frac{G(x_l + h_l)}{\|h_l\|} &\geq \max_{z \in I(x_l)} \frac{g(x_l + h_l, z) - g(x_l, z)}{\|h_l\|} \\ &= \max_{z \in I(x_l)} \left\{ \langle \nabla_x g(x_l, z), \|h_l\|^{-1} h_l \rangle + \|h_l\|^{-1} \alpha(h_l, z) \right\} \\ &\geq \max_{z \in I(x_l)} \left\{ \langle \nabla_x g(x_l, z), \|h_l\|^{-1} h_l \rangle \right\} - \max_{z \in I(x_l)} \left\{ \|h_l\|^{-1} |\alpha(h_l, z)| \right\} \\ &\geq L^{-1} \rho^2 / 4 - c\|h_l\|. \end{aligned}$$

Choosing l_0 large enough to satisfy $\|h_l\| \leq (8cL)^{-1} \rho^2$ for all $l \geq l_0$, it follows that

$$\frac{G(x_l + h_l)}{\|h_l\|} \geq L^{-1} \rho^2 / 8 > 0 \quad \forall l \geq l_0.$$

This last relation eventually entails condition 2. of Theorem 1. □

4. Applications

4.1. Nonsmooth Calculus

This section is devoted to two applications of the preceding theory in nonsmooth calculus. The first one concerns the computation of the limiting normal cone to the set $M(0) = \{x \in \mathbb{R}^n \mid g(x) \in \Lambda\}$, where g maps $\mathbb{R}^n$ into $\mathbb{R}^m$ and $\Lambda \subset \mathbb{R}^m$ has a special structure.

Theorem 5. *Let g be continuously differentiable and $\Lambda = \cup_{j=1}^{p} \Lambda_j$, where each $\Lambda_j \subseteq \mathbb{R}^m$ is a convex polyhedron. Suppose that $g(\bar{x}) \in \Lambda$ and both assumptions of Theorem 2 are fulfilled. Then one has*

$$N_{M(0)}(\bar{x}) \subseteq (\nabla g(\bar{x}))^T N_\Lambda(g(\bar{x})). \tag{29}$$

If Λ happens to be regular at $g(\bar{x})$, then $M(0)$ is regular at $\bar{x}$ and inclusion (29) becomes an equality.

Proof. The first assertion follows immediately from the calmness of the respective map M at $(0, \bar{x})$ by virtue of [15, Theorem 4.1]. To prove the second assertion, note that

$$N_{M(0)}(\bar{x}) \supseteq \hat{N}_{M(0)}(\bar{x}) \supseteq (\nabla g(\bar{x}))^T \hat{N}_\Lambda(g(\bar{x})) \tag{30}$$

without any assumptions. Since $\hat{N}_\Lambda(g(\bar{x})) = N_\Lambda(g(\bar{x}))$ by the regularity of Λ at $g(\bar{x})$, it suffices to combine (29) and (30) to get

$$\hat{N}_{M(0)}(\bar{x}) = N_{M(0)}(\bar{x}) = (\nabla g(\bar{x}))^T N_\Lambda(g(\bar{x})),$$

and we are done. □

The preceding result can be utilized, e.g., in deriving optimality conditions for the program

$$\min\{\varphi(x)|g(x) \in \Lambda\}, \tag{31}$$

where $\varphi : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz and g, Λ satisfy the assumptions of Theorem 2. Let $\hat{x}$ be a local solution of (31) and assume that $T_{M(0)}(\hat{x})$ is not convex. Then, one usually employs the optimality conditions from [24]

$$0 \in \partial\varphi(\hat{x}) + N_{M(0)}(\hat{x}).$$

On the basis of Theorem 5 we arrive in this way at the desired relation

$$0 \in \partial\varphi(\hat{x}) + (\nabla g(\hat{x}))^T N_\Lambda(g(\hat{x})) \tag{32}$$

even in the case when (GMFCQ) does not hold at $\bar{x}$.

This situation can be illustrated by means of the constraint system analyzed in Example 3.

Example 5. Consider the mathematical program (31) with

$$\varphi(x_1, x_2) = 2\,|\,x_1 - x_2\,| - (x_1 + x_2) \tag{33}$$

and g, Λ being given in Example 3. On the basis of Figure 1 and the objective (33) one easily deduces that $\bar{x} = 0$ is a local minimizer in this program. From Example 3 we know that the respective map M is calm at $(0, \bar{x})$. Therefore, by virtue of Theorem 5, it follows that

$$N_{M(0)}(\bar{x}) \subseteq \begin{bmatrix} 0 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} N_\Lambda(0). \tag{34}$$

One readily computes that

$$N_\Lambda(0) = \left(\mathbb{R}^2_+ \times \{0\}\right) \cup (\{0\} \times \{0\} \times \mathbb{R}_+)\,.$$

Furthermore,

$$\partial\varphi(\bar{x}) = \begin{bmatrix} 2 \\ -2 \end{bmatrix} \mathbb{B} - \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

and we observe that the vector $(-2, 0)^T \in \partial\varphi(\bar{x})$ and the vector $(2, 0)^T$ belong to the cone on the right-hand side of (34). This implies that the optimality conditions (32) are fulfilled.

Calmness plays also a crucial role in the computation of coderivatives of composite multifunctions. This concerns the general situation considered in [25, Theorem 5.1], but here we restrict ourselves only to the multifunction

$$S(u) := \{x \in \Theta \mid h(u, x) \in \Lambda\}, \tag{35}$$

where $h : \mathbb{R}^p \times \mathbb{R}^n \to \mathbb{R}^m$ is locally Lipschitz and the sets $\Theta \subseteq \mathbb{R}^n$, $\Lambda \subseteq \mathbb{R}^m$ are closed. We start with a modification of [25, Theorem 6.10] and introduce for this purpose the multifunction $P : \mathbb{R}^m \rightrightarrows \mathbb{R}^p \times \mathbb{R}^n$ defined by

$$P(y) := \{(u, x) \in \mathbb{R}^p \times \Theta \mid h(u, x) + y \in \Lambda\}. \tag{36}$$

Clearly, $x \in S(u)$ iff $(u, x) \in P(0)$, i.e., $\operatorname{Gph} S = P(0)$.

Theorem 6. *Let* $(\bar{u}, \bar{x}) \in \operatorname{Gph} S$ *and assume that* P *is calm at* $(0, \bar{u}, \bar{x})$. *Then one has for all* $x^* \in \mathbb{R}^n$ *the inclusion*

$$D^*S(\bar{u}, \bar{x})(x^*) \subseteq \left\{u^* \in \mathbb{R}^p \mid \begin{bmatrix} u^* \\ -x^* \end{bmatrix} \in D^*h(\bar{u}, \bar{x}) \circ N_\Lambda(h(\bar{u}, \bar{x})) + \begin{bmatrix} 0 \\ N_\Theta(\bar{x}) \end{bmatrix}\right\}. \tag{37}$$

Proof. According to the definition,

$$D^*S(\bar{u}, \bar{x})(x^*) = \left\{u^* \in \mathbb{R}^p \mid \begin{bmatrix} u^* \\ -x^* \end{bmatrix} \in N_{P(0)}(\bar{u}, \bar{x})\right\}.$$

Due to the required calmness of P we can invoke [15, Theorem 4.1] which yields the inclusion

$$N_{P(0)}(\bar{u}, \bar{x}) \subseteq D^*h(\bar{u}, \bar{x}) \circ N_\Lambda(h(\bar{u}, \bar{x})) + \begin{bmatrix} 0 \\ N_\Theta(\bar{x}) \end{bmatrix}$$

and completes the proof. □

Formula (37) is useful, e.g., for testing the Aubin property of S around $(\bar{u}, \bar{x})$ via the Mordukhovich criterion $D^*S(\bar{u}, \bar{x})(0) = \{0\}$. If we connect this criterion with the qualification conditions from [25, Theorem 6.10], ensuring the validity of inclusion (37), we arrive at the condition

$$\left.\begin{array}{c} \begin{bmatrix} u^* \\ 0 \end{bmatrix} \in D^*h(\bar{u}, \bar{x})(v) + \begin{bmatrix} 0 \\ N_\Theta(\bar{x}) \end{bmatrix} \\ v \in N_\Lambda(h(\bar{u}, \bar{x})) \end{array}\right\} \Rightarrow \begin{cases} u^* = 0 \\ v = 0. \end{cases} \tag{38}$$

If we, however, ensure the validity of (37) via the calmness of P at $(0, \bar{u}, \bar{x})$, then S possesses the Aubin property around $(\bar{u}, \bar{x})$ provided

$$\left.\begin{array}{c} \begin{bmatrix} u^* \\ 0 \end{bmatrix} \in D^*h(\bar{u}, \bar{x})(v) + \begin{bmatrix} 0 \\ N_\Theta(\bar{x}) \end{bmatrix} \\ v \in N_\Lambda(h(\bar{u}, \bar{x})) \end{array}\right\} \Rightarrow u^* = 0. \tag{39}$$

The importance of the difference between (38) and (39) is strikingly illustrated by the following NCP.

Example 6. Let $S : \mathbb{R} \to \mathbb{R}^2$ be the map which assigns to the parameter u the set of solutions to the complementarity problem, governed by the GE

$$0 \in \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ u+2 \end{bmatrix} + N_{\mathbb{R}^2_+}(x).$$

We want to examine the Aubin property of S at $(\bar{u}, \bar{x}) = (0, 1, 0)$. This problem can be converted to the form (35) in the same way as it was done in Example 2; thereby $\Theta = \mathbb{R}^2$ and the corresponding map h is affine. We easily realize that condition (38) is not fulfilled (each vector $(v_1, v_2) \in \mathbb{R} \times \{0\}$ belongs to $N_\Lambda(h(\bar{u}, \bar{x})) \cap \ker (\nabla h(\bar{u}, \bar{x}))^T$). On the other hand, since h is affine, the corresponding map P is calm and condition (39) is fulfilled. This implies that S has the Aubin property around $(\bar{u}, \bar{x})$, which could not be detected by the standard technique.

The theory, developed in Section 2, does not allow one to establish the calmness of P in the above general setting in a new way. If, however, $\Theta = \mathbb{R}^n$, Λ is as in Theorem 2 and h happens to be continuously differentiable, then one can try to apply Theorem 6 whenever the qualification conditions of [25, Theorem 6.10] are not fulfilled.

4.2. Continuity of the Value-at-Risk

A prominent risk measure used in mathematics of finance or in stochastic optimization is the *value at risk*. For a given random variable X and a given probability level $p \in (0, 1]$, this value at risk is defined as

$$\mathrm{VaR}_p(X) := \inf\{r \in \mathbb{R} | P(X \leq r) \geq p\} = \inf\{r \in \mathbb{R} | F_X(r) \geq p\}.$$

Here, P denotes some probability measure and F_X is the distribution function of X. It is well known, and sometimes stated as a shortcoming of this risk measure, that, in general, $\mathrm{VaR_p}$ does not depend continuously on X. The fact that a suitable growth condition has to be imposed on the distribution function in order to obtain a Lipschitz-like property for the value-at-risk, has already been observed in [29] (Proposition 8). The following theorem uses Proposition 2 in order to derive the analogous result under the assumption that X has a density f_X, i.e., $F_X(x) = \int_{-\infty}^{x} f_X(t)dt$. The deviation between two random variables X and Y shall be measured by

$$\Delta(X, Y) := \sup_{t \in \mathbb{R}} |F_X(t) - F_Y(t)|$$

which is the Kolmogorov distance between the distributions induced by X and Y, respectively. For convenience of notation, we put $\bar{x} := \mathrm{VaR}_p(X)$. Furthermore, denoting by λ the Lebesgue measure in $\mathbb{R}$, we introduce the quantities

$$\varphi^{\uparrow}(\varepsilon, \alpha) := \lambda\{x \in [\bar{x}, \bar{x} + \varepsilon] \,|\, f_X(x) \geq \alpha\}$$
$$\varphi^{\downarrow}(\varepsilon, \alpha) := \lambda\{x \in [\bar{x} - \varepsilon, \bar{x}] \,|\, f_X(x) \geq \alpha\}.$$

Theorem 7. *Let X be a fixed random variable. Assume that $p \in (0, 1)$ and that*

$$\liminf_{\alpha, \varepsilon \downarrow 0} \varepsilon^{-1} \varphi^{\uparrow}(\varepsilon, \alpha) > 0 \text{ and } \liminf_{\alpha, \varepsilon \downarrow 0} \varepsilon^{-1} \varphi^{\downarrow}(\varepsilon, \alpha) > 0. \quad (40)$$

Then, there exist constants $L, \delta > 0$, such that

$$\left|\mathrm{VaR}_p(X) - \mathrm{VaR}_p(Y)\right| \leq L\Delta(X,Y) \text{ for all } Y \text{ with } \Delta(X,Y) < \delta.$$

Proof. As a distribution function with density, F_X is continuous and satisfies $\lim_{x \to -\infty} F_X(x) = 0$. From here, it follows immediately that, under our assumption $p \in (0,1)$, one has that $F_X(\bar{x}) = p$. The second condition in (40) provides the existence of $\alpha, \gamma, \delta > 0$ such that

$$\varphi^{\downarrow}(\varepsilon, \alpha) \geq \gamma\varepsilon \quad \forall \varepsilon \in (0, \delta).$$

Consequently,

$$F_X(\bar{x}) - F_X(\bar{x} - \varepsilon) = \int_{\bar{x}-\varepsilon}^{\bar{x}} f_X(t)dt \geq \alpha\varphi^{\downarrow}(\varepsilon, \alpha) \geq \alpha\gamma\varepsilon \quad \forall \varepsilon \in (0, \delta), \tag{41}$$

With $g(x) := p - F_X(x)$, this yields in the notation of Proposition 2 that $g^{\downarrow}(\bar{x}; -1) > 0$. Consequently, $0 \notin \left[g^{\downarrow}(\bar{x}; -1), g^{\uparrow}(\bar{x}; -1)\right]$ and the implication (10) holds trivially true. On the other hand, because F_X is nondecreasing as a distribution function, one has that

$$g(x) = p - F_X(x) \leq p - F_X(\bar{x}) = 0 \quad \forall x \geq \bar{x}.$$

Thus, (the conclusion of) the implication (9) holds true. Summarizing, Proposition 2 may be applied to derive calmness of the mapping

$$t \mapsto \{x | g(x) \leq -t\}$$

at $(0, \bar{x})$ which amounts to the calmness of the mapping

$$t \mapsto \{x | F_X(x) \geq t\}$$

at $(p, \bar{x})$. By definition, there are constants $L, \delta_1 > 0$ such that

$$d_{[\bar{x},\infty)}(r) \leq L|t - p| \quad \forall r \in [\bar{x} - \delta_1, \bar{x} + \delta_1] : F_X(r) \geq t \quad \forall t \in [p - \delta_1, p + \delta_1].$$

Next we exploit that $F_X(\bar{x} - \delta_1) < F_X(\bar{x})$ (otherwise the fact that F_X is nondecreasing implies the contradiction $F_X(r) = F_X(\bar{x})$ for all $r \in [\bar{x} - \delta_1, \bar{x}]$ with (41)). Therefore, taking into account once more that F_X is nondecreasing and observing that $d_{[\bar{x},\infty)}(r) = 0$ for $r \geq \bar{x}$, the above relation can be extended to

$$d_{[\bar{x},\infty)}(r) \leq L|t - p| \quad \forall r \in \mathbb{R} : F_X(r) \geq t \quad \forall t \in [p - \delta_2, p + \delta_2], \tag{42}$$

where $\delta_2 := \min\{\delta_1, (F_X(\bar{x}) - F_X(\bar{x} - \delta_1))/2\} > 0$. Now, consider an arbitrary random variable Y and an arbitrary $r \in \mathbb{R}$ with $F_Y(r) \geq p$. By definition, $F_X(r) \geq p - \Delta(X,Y)$. If Y is such that $\Delta(X,Y) \leq \delta_2$, then we may put $t := p - \Delta(X,Y)$ in (42) and get that $d_{[\bar{x},\infty)}(r) \leq L\Delta(X,Y)$. Consequently,

$$\bar{x} \leq r + d_{[\bar{x},\infty)}(r) \leq r + L\Delta(X,Y) \quad \forall r : F_Y(r) \geq p \quad \forall Y : \Delta(X,Y) \leq \delta_2.$$

Passing to the infimum over all r with $F_Y(r) \geq p$, yields

$$\mathrm{VaR}_p(X) \leq \mathrm{VaR}_p(Y) + L\Delta(X, Y) \quad \forall Y : \Delta(X, Y) \leq \delta_2. \tag{43}$$

Repeating the analogous argumentation, but now based on the first condition in (40), one deduces calmness of the mapping

$$t \mapsto \{x | F_X(x) \leq t\}$$

at $(p, \bar{x})$ and arrives at a relation similar to (42):

$$d_{(-\infty,\bar{x}]}(r) \leq L|t - p| \quad \forall r \in \mathbb{R} : F_X(r) \leq t \quad \forall t \in [p - \delta_2, p + \delta_2],$$

Now, we put $\bar{y} := \mathrm{VaR}_p(Y)$. Since, in contrast to the given variable X, we do not assume that the perturbed variable Y has a density, we cannot except F_Y to be continuous. Therefore, $F_Y(\bar{y}) > p$ is possible (note that F_Y is always upper semicontinuous as a distribution function, whence the relation $F_Y(\bar{y}) < p$ is excluded by the definition of VaR_p). However, a simple argument shows that, due to continuity of F_X, one always has that $\Delta(X, Y) \geq (F_Y(\bar{y}) - p)/2$. Hence,

$$F_X(\bar{y}) \leq F_Y(\bar{y}) + \Delta(X, Y) \leq p + 3\Delta(X, Y)$$

Now, if Y is such that $\Delta(X, Y) \leq \delta_2/3$, then we may put $t := p + 3\Delta(X, Y)$ and $r := \bar{y}$ in (44) to derive that $d_{(-\infty,\bar{x}]}(\bar{y}) \geq 3L\Delta(X, Y)$. Therefore

$$\mathrm{VaR}_p(X) = \bar{x} \geq \bar{y} - d_{(-\infty,\bar{x}]}(\bar{y}) \geq \mathrm{VaR}_p(Y) - 3L\Delta(X, Y)$$

for all Y with $\Delta(X, Y) \leq \delta_2/3$. This Combines with (43) to the assertion of the theorem. □

Remark 5. Using Theorem 1 in [13], the conclusion of the last theorem could be obtained without condition (40) but under the assumption that the density f_X is log-concave, i.e., log f_X is concave (this holds true, for instance, for the normal, Gamma, Dirichlet, uniform, lognormal and many other distributions, see [30]).

Remark 6. Instead of (40) one might consider the simpler condition

$$\exists \varepsilon > 0 : f_X(x) \geq \varepsilon \quad \text{for almost all } x \in [\bar{x} - \varepsilon, \bar{x} + \varepsilon],$$

which obviously implies that

$$\liminf_{\alpha,\varepsilon \downarrow 0} \varepsilon^{-1} \varphi^{\uparrow}(\varepsilon, \alpha) = \liminf_{\alpha,\varepsilon \downarrow 0} \varepsilon^{-1} \varphi^{\downarrow}(\varepsilon, \alpha) = 1,$$

and, hence is stronger than (40). Indeed, this condition was shown in [12, Theorem 6] to imply the Aubin property of the mapping

$$t \mapsto \{x | F_X(x) \geq t\}$$

around $(p, \bar{x})$. From here, one might expect now a stronger Lipschitz result as compared to Theorem 7, e.g.:

$$\left|\mathrm{VaR}_p(Y_1) - \mathrm{VaR}_p(Y_2)\right| \leq L\Delta(Y_1, Y_2) \quad \forall Y_1, Y_2 : \Delta(X, Y_1), \Delta(X, Y_2) < \delta.$$

This, however, does not hold true as is confirmed by an example in [16] (Example 1), which is easily translated to the “value-at-risk”- setting considered here.

The following example demonstrates the use of condition (40) in Theorem 7 as compared to the condition in the last remark:

Example 7. Consider a random variable X with its distribution having density

$$f_X(x) := Ke^{-x^2} \max\{\sin x^{-2}, 0\},$$

where we put $f_X(0) := 0$, $p := 0.5$ and K is a normalizing constant such that $\int f_X(x)dx = 1$. Due to symmetry of f, it follows that $\bar{x} := \mathrm{VaR}_p(X) = 0$. Some calculation shows that

$$\liminf_{\alpha,\varepsilon\downarrow 0} \varepsilon^{-1}\varphi^{\uparrow}(\varepsilon, \alpha) = \liminf_{\alpha,\varepsilon\downarrow 0} \varepsilon^{-1}\varphi^{\downarrow}(\varepsilon, \alpha) = 0.5,$$

so that (40) is satisfied and the result of Theorem 7 may be derived, but the condition of Remark 6 is violated.

Acknowledgements. The authors are indebted to J. Jarušek (Czech Academy of Sciences, Prague) and A. Kruger (University of Ballarat, Australia) for helpful advice and numerous discussions on the subject of this paper. They would also like to express their gratitude to two anonymous referees for many useful suggestions.

References

1. Aubin, J.-P.: Lipschitz behavior of solutions to convex minimization problems. Math. Oper. Res. **9**, 87–111 (1984)
2. Azé, D., Corvellec, J.-N.: Characterizations of error bounds for lower semicontinuous functions on metric spaces. ESAIM Control Optim. Calc. Var. **10**, 409–425 (2004)
3. Bazaraa, M.S., Shetty C.M.: Nonlinear Programming. Theory and Algorithms. Wiley, Chichester, 1979
4. Burke, J. V.: Calmness and exact penalization. SIAM J. Control Optim. **29**, 493–497 (1991)
5. Burke, J. V., Deng, S.: Weak sharp minima revisited. I. Basic theory. Control Cybernet. **31**, 439–469 (2002)
6. Clarke, F.H.: A new approach to Lagrange multipliers. Math. Oper. Res. **2**, 165–174 (1976)
7. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley, New York, 1983
8. Dontchev, A.L., Hager, W.W.: Implicit functions, Lipschitz maps and stability in optimization. Math. Oper. Res. **19**, 753–768 (1994)
9. Dontchev, A.L., Rockafellar, R.T.: Characterization of strong regularity for variational inequalities over polyhedral convex sets. SIAM J. Optim. **7**, 1087–1105 (1996)
10. Dontchev, A.L., Rockafellar, R.T.: Ample parameterization of variational inclusions, SIAM J. Optim. **12**, 170–187 (2001)
11. Dontchev, A.L., Rockafellar, R.T.: Regularity and conditioning of solution mappings in variational analysis, Set-Valued Anal. **12**, 79–109 (2004)
12. Henrion, R., Römisch, W.: Metric regularity and quantitative stability in stochastic programs with probabilistic constraints. Math. Program. **84**, 55–88 (1999)
13. Henrion, R.: Qualitative stability in convex programs with probabilistic constraints. In: V.H. Nguyen et al. (eds.), Optimization, Lecture Notes in Economics and Mathematical Systems, **481**, Springer, Berlin, 2000, pp. 164–180
14. Henrion, R., Outrata, J.V.: A subdifferential criterion for calmness of multifunctions. J. Math. Anal. Appl. **258**, 110–130 (2001)
15. Henrion, R., Jourani, A., Outrata, J.V.: On the calmness of a class of multifunctions. SIAM J. Optim. **13**, 603–618 (2002)
16. Henrion, R.: Perturbation Analysis of chance-constrained programs under variation of all constraint data. In: K. Marti et al. (eds.), Dynamic Stochastic Optimization, Lecture Notes in Economics and Mathematical Systems, **532**, Springer, Berlin, 2004, pp. 257–274
17. Ioffe, A.D.: Necessary and sufficient conditions for a local minimum, Part 1: A reduction theorem and first order conditions. SIAM J. Control Optim. **17**, 245–250 (1979)
18. Ioffe, A.D.: Metric regularity and subdifferential calculus. Russian Math. Surveys **55**, 501–558 (2000)

19. King, A.J., Rockafellar, R.T.: Sensitivity analysis for nonsmooth generalized equations, Math. Programming **55**, 193–212 (1992)
20. Klatte, D., Kummer, B.: Constrained minima and Lipschitzian penalties in metric spaces, SIAM J. Optim. **13**, 619–633 (2002)
21. Klatte, D., Kummer, B.: Nonsmooth Equations in Optimization-Regularity, Calculus, Methods and Applications. Kluwer, Dordrecht, 2002
22. Levy, A.B.: Implicit multifunction theorems for the sensitivity analysis of variational conditions, Math. Programming **74**, 333–350 (1996)
23. Mifflin, R.: Semismooth and semiconvex functions in constrained optimization. SIAM J. Control Optim. **15**, 959–972 (1979)
24. Mordukhovich, B.S.: Approximation Methods in Problems of Optimization and Control. Nauka, Moscow, 1988 (in Russian)
25. Mordukhovich, B.S.: Generalized differential calculus for nonsmooth and set-valued mappings. J. Math. Anal. Appl. **183**, 250–288 (1994)
26. Ng, K.F., Zheng, X.Y.: Error bounds for lower semicontinuous functions in normed spaces. SIAM J. Optim. **12**, 1–17 (2001)
27. Outrata, J.V.: A generalized mathematical program with equilibrium constraints. SIAM J. Control Optim. **38**, 1623–1638 (2000)
28. Pang, J.-S.: Error bounds in mathematical programming. Math. Program. **79**, 299–332 (1997)
29. Pflug, G.Ch.: Stochastic Optimization and Statistical Inference. In: A. Ruszczynski and A. Shapiro (eds.) Stochastic Programming. Handbooks in Operations Research and Management Science, Vol. 10, Elsevier, Amsterdam 2003, chapter 7
30. Prékopa, A.: Stochastic Programming. Kluwer, Dordrecht, 1995
31. Robinson, S.M.: Generalized equations and their solutions, Part I: Basic theory, Math. Programming Study **10**, 128–141 (1979)
32. Robinson, S.M.: Some continuity properties of polyhedral multifunctions. Math. Programming Study **14**, 206–214 (1981)
33. Robinson, S.M.: An implicit-function theorem for a class of nonsmooth functions. Math. Oper. Res. **16**, 292–309 (1991)
34. Rockafellar, R.T.: Lipschitzian properties of multifunctions, Nonlinear Anal. **9**, 867–885 (1985)
35. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Springer, New York, 1998
36. Studniarski, M., Ward, D.E.: Weak sharp minima: characterizations and sufficient conditions. SIAM J. Control and Optim. **38**, 219–236 (1999)
37. Thibault, L.: Propriétés des Sous-Différentiels de Fonctions Localement Lipschitziennes Définies sur un espace de Banach Séparable. PhD Thesis, Department of Mathematics, University of Montpellier, 1975
38. Wu, Z., Ye, J.J.: Sufficient conditions for error bounds. SIAM J. Optim. **12**, 421–435 (2001)
39. Ye, J.J., Ye, X.Y.: Necessary optimality conditions for optimization problems with variational inequality constraints. Math. Oper. Res. **22**, 977–997 (1997)