Characterization of Calmness for Banach space mappings

Jan Heerda [1] Bernd Kummer[2]

1 November 2006

Abstract. We characterize calmness of multifunctions explicitly by calmness of level sets to globally Lipschitz functions, by convergence of specific solution methods for the related inclusions as well as by solvability of crucial linear systems. As a main tool, a so-called relative slack function will be applied. In this way, also equivalence between calmness and metric regularity of specific subsystems will be derived.

Key words. calmness, generalized equations, Lipschitz functions, first order methods, crucial linear systems, modified mappings.

Mathematics Subject Classification 2000. 49J52, 49K40, 90C31, 65Y20.

1 Introduction

It is well-known that calmness of multifunctions is a basic property in order to derive optimatity conditions and penalty methods in optimization models and for establishing various duality statements. In this paper, we exploite two recently known facts:
Calmness of a multifunction is nothing but calmness of a canonically assigned Lipschitzian level set map and, on the other hand, calmness is equivalent to the applicability and linear convergence of certain solution methods.
Our basic model is the generalized equation

(1.1) Find $x \in X$ such that $p \in F(x)$, $\quad F : X \rightrightarrows P$,

where $p \in P$ is a canonical parameter, P, X are Banach spaces and F is a closed multifunction, i.e., $F(x) \subset P$ and the graph of F, $\operatorname{gph} F = \{(x,p) \mid p \in F(x)\}$, is a closed set.
System (1.1) describes solutions of equations as well as stationary or critical points of various variational conditions. Several other applications of model (1.1) are known for optimization problems, for describing equilibria and other solutions in games, in so-called MPECs and stochastic and/or multilevel (multiphase) models. We refer e.g. to [6, 1, 39, 31, 2, 8, 22, 13, 23] for the related settings.
We shall consider $S(p) = F^{-1}(p)$ near some particular solution $x^0 \in S(p^0)$ of (1.1) at p^0. In the whole paper, $S = F^{-1} : P \rightrightarrows X$ is a closed multifunction, P, X are Banach spaces and $z^0 = (p^0, x^0)$ is a given point in $\operatorname{gph} S$. By $\operatorname{conv} M$ we denote the convex hull of a set M and $o(t)$ denotes, as usual, a quantity of the type $o(t)/t \to 0$ if $t \downarrow 0$.
We say that some property holds *near* x if it holds for all points in some neighborhood of x. By B we denote the closed unit ball in the related space and

$$x^0 + \varepsilon B := \{x \in X \mid d(x, x^0) \le \varepsilon\}.$$

We often write $d(x, x^0)$ for the (induced) distance in X, for better distinguishing terms in the spaces P and X. In fact, many statements of this paper remain true for a complete metric space X. In particular, one may suppose that $X = M \subset \hat{X}$ where M is a closed subset of a Banach space $\hat{X}$. This situation corresponds to the system

(1.2) Find $x \in M \subset X$ such that $p \in F(x)$, $\quad F : X \rightrightarrows P$

[1] Address: Institut für Mathematik, Humboldt–Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany. E–Mail: janjh@mathematik.hu-berlin.de

[2] Address: Institut für Mathematik, Humboldt–Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany. E–Mail: kummer@mathematik.hu-berlin.de

with the solution map $S_M(p) = F^{-1}(p) \cap M = \{x \in M \mid p \in F(x)\}$. Though (1.2) coincides with (1.1) after setting $F(x) = \emptyset \; \forall x \in X \setminus M$, the explicit consideration of (1.2) may be useful in some situations.
The following definitions generalize typical local properties of the multivalued inverse $S = f^{-1}$ or of level sets $S(p) = \{x \in X \mid f(x) \le p\}$ for functions $f : X \to \mathbb{R}$.

Definition 1. S is said to be *calm* at $z^0 = (p^0, x^0) \in \operatorname{gph} S$ if

$$\exists\, \varepsilon, \delta, L > 0 \text{ such that } S(p) \cap (x^0 + \varepsilon B) \subset S(p^0) + L\|p - p^0\| B \quad \forall p \in p^0 + \delta B. \tag{1.3}$$

S is said to be *Lipschitz lower semicontinuous* (Lipschitz l.s.c.) at z^0 if

$$\exists\, \delta, L > 0 \text{ such that } S(p) \cap (x^0 + L\|p - p^0\| B) \ne \emptyset \quad \forall p \in p^0 + \delta B. \tag{1.4}$$

◇

Notice that (1.3) involves a locally Lipschitzian error estimate, namely

$$\operatorname{dist}(x, S(0)) \le L\|p - p^0\| \qquad \forall x \in S(p) \cap (x^0 + \varepsilon B). \tag{1.5}$$

Remark 1. Using these definitions, other known stability properties can be characterized (we apply the notations of [22]).

(i) S is locally upper Lipschitz at z^0 $\Leftrightarrow$ S is calm at z^0 and x^0 is isolated in $S(p^0)$.

(ii) S is pseudo-Lipschitz (equivalently: S obeys the Aubin property or S^{-1} is metrically regular) at z^0 $\Leftrightarrow$ S is Lipschitz l.s.c. at all points $z \in \operatorname{gph} S$ near z^0 with fixed constants δ and L.

(iii) S is pseudo-Lipschitz at z^0 $\Leftrightarrow$ S is both calm at all $z \in \operatorname{gph} S$ near z^0 with fixed constants ε, δ, L and Lipschitz l.s.c. at z^0.

(iv) S is strongly Lipschitz at z^0 $\Leftrightarrow$ S is pseudo-Lipschitz at z^0 and, for small $\varepsilon > 0$, $S(p) \cap (x^0 + \varepsilon B)$ is single-valued for p near p^0. ◇

The goal of this paper is to characterize calmness, in section 4, by the behavior of methods for solving (1.1) and (1.2), cf. the theorems 4.4, 4.5. We also show that calmness of multifunctions can be transformed into calmness of Lipschitzian level set mappings only, cf. Remark 2. Applying our approach to C^1- inequality systems, we identify the crucial subsystems which have to be metrically regular in order to ensure calmness of the whole system, cf. Theorem 4.6. Before, we discuss, in finite dimension, the meaning of calmness for first-order optimality conditions in section 2 and investigate (more or less known) calmness conditions for inequality systems, section 3.
For basic results concerning the related stability properties we refer to [1, 7, 14, 15, 19, 29, 30] (Aubin property), [5, 26, 35, 38] (strongly Lipschitz), [20, 34, 36] (locally upper Lipschitz) as well as the monographs [2, 8, 22, 31, 39].

2 Comments in view of calmness, KKT points and Abadie CQ

Let us start by recalling the well-known interplay of calmness and the *Abadie constraint qualification* in relation to Karush-Kuhn-Tucker (KKT) points for a usual optimization model

$$\min f_0(x) \quad s.t. \;\; x \in X = \mathbb{R}^n, \;\; f_i(x) \le 0, \quad \text{where } f_0, f_i \in C^1, \; i = 1, ..., m. \tag{2.1}$$

The KKT points $(x, y) \in \mathbb{R}^{n+m}$ are defined by the existence of Lagrange multipliers y with

$$\begin{array}{llll} Df_0(x) + \sum_{i=1}^m y_i \, Df_i(x) = 0, & y_i \ge 0, & \\ y_i \, f_i(x) = 0, & f_i(x) \le 0, & \forall\, i > 0. \end{array} \tag{2.2}$$

For given feasible x, system (2.2) is inconsistent iff there is some $u \in \mathbb{R}^n$ such that

$$(2.3) \qquad Df_0(x)u < 0 \qquad \text{and}$$

$$(2.4) \qquad Df_i(x)u \leq 0 \qquad \forall i : f_i(x) = 0.$$

System (2.3), (2.4) is equivalent to the existence of some $c > 0$ such that, if

$$(2.5) \qquad w(t) \in \mathbb{R}^n \ \text{ and } \ \lim_{t \downarrow 0} \frac{\|w(t)\|}{t} = 0,$$

it holds

$$\lim_{t \downarrow 0} \frac{f_0(x + tu + w(t)) - f_0(x)}{t} \leq -c \ \text{ and } \ \limsup_{t \downarrow 0} \frac{f_i(x + tu + w(t))}{t} \leq 0 \quad \forall\, i > 0.$$

Hence if some w (2.5) even satisfies

$$(2.6) \qquad f_i(x + tu + w(t)) \leq 0 \quad \text{for all } i > 0 \text{ and certain } t = t_k \downarrow 0,$$

then x is never a local minimizer for (2.1). In other words, if x is a local solution to (2.1) satisfying the regularity condition

$$(2.7) \qquad (2.4) \ \text{ implies } \ (2.6) \text{ for some } w \text{ in } (2.5)$$

then some (x, y) fulfills the KKT system (2.2). This motivates the investigation of conditions (constraint qualifications) which ensure (2.7). It is well-known that calmness of

$$S(p) = \{x \in X \mid f_i(x) \leq p_i \ \forall i > 0\}$$

at $(0, x)$ and the Abadie CQ for $S(0)$ at x are conditions of this type.

- The Abadie CQ for $S(0)$ at x *requires simply by definition* that (2.7) holds true.
- Calmness of S at $(0, x)$ *implies* that (2.7) holds true since, due to (2.4) and $f_i(x + tu) \leq o(t)\ \forall i$, there are $w(t)$ with $\|w(t)\| \leq L\ o(t)$ and $x + tu + w(t) \in S(0)$ (for small $t > 0$).

Thus calmness is a tool for showing the Abadie CQ. Nevertheless, characterizing any of these conditions in a sharp manner requires – even for $X = \mathbb{R}^n$ – considerable analytical effort provided the involved functions are nonlinear. Concerning the similar role of calmness for optimality conditions under Banach space settings and directional differentiability we refer to [21], sections 4 and 5.
It is worth to mention that the Abadie CQ (hence also calmness) is not necessary for the existence of Lagrange multipliers (2.2) at a solution x (this is again a known fact):

Example 1. The mapping

$$(2.8) \qquad S(p) = \{x \in \mathbb{R} \mid x^2 \leq p_1, \quad -x \leq p_2\}$$

is not calm at $0 \in \mathbb{R}^3$. The cone $K = \{u \in \mathbb{R} \mid Df_i(0)u \leq 0\ \forall i > 0\}$ contains $u = 1$, but the points $tu + w(t)$ are not in $S(0)$ for small $t > 0$. In consequence, the Abadie CQ does not hold for $S(0)$ at the origin. Nevertheless, the KKT-system for the problem $\min x$, *s.t.* $x \in S(0)$ is solvable with $x = 0$ and $y_2 = 1$ while it is unsolvable for the negative objective $f_0(x) = -x$. $\diamond$

3 C^1 constraints in $\mathbb{R}^n$

Previous to study calmness in the context of Banach spaces, the consideration of the finite-dimensional, continuously differentiable case is useful in order to collect possible approaches and to discern possible difficulties. For every constraint system of a usual optimization model in $X = \mathbb{R}^n$, namely

$$(3.1) \qquad S(p_1, p_2) = \{x \in \mathbb{R}^n \mid g(x) \le p_1,\ h(x) = p_2\}, \quad (g, h) \in C^1(\mathbb{R}^n, \mathbb{R}^{m_1+m_2}),$$

the Aubin property can be characterized by elementary and intrinsic means. In the whole section, let $z^0 = (0, x^0) \in \operatorname{gph} S$ and $I(x) = \{i \mid g_i(x) = 0\}$.

Lemma 3.1. *For the multifunction S (3.1), the following statements are equivalent:*

1. *S is Lipschitz l.s.c. at z^0.*
2. *S obeys the Aubin property at z^0.*
3. *The Mangasarian-Fromowitz constraint qualification (MFCQ) holds at z^0, i.e.,*
   $$(3.2) \quad \operatorname{rank} Dh(x^0) = m_2 \text{ and } \exists u \in \ker Dh(x^0) \text{ such that } Dg_i(x^0)u < 0\ \forall i \in I(x^0). \quad \diamond$$

Proof. The proof follows mainly from Robinson's basic paper [32], by taking the equivalence of Aubin property and metric regularity into account. For more details we refer to [25]. □

Analyzing calmness seems to be simpler since it suffices to investigate calmness of the inequality system

$$\widetilde{S}(q) = \{x \in \mathbb{R}^n \mid g_i(x) \le q,\ -q \le h_j(x) \le q,\ \forall\, i = 1, ..., m_1,\ j = 1, ..., m_2\}$$

at $(0, x^0) \in \mathbb{R} \times X$ only, and calmness requires less than the Aubin property. Nevertheless, its equivalent characterization is more complicated, provided the functions involved are not piecewise linear (then calmness holds true). In what follows we assume, for sake of simplicity, that $S(p_1, p_2)$ is written in form of inequalities only, i.e., we suppose

$$(3.3) \qquad S(p) = \{x \in \mathbb{R}^n \mid g_i(x) \le p_i, \quad i = 1, ..., m\}, \qquad g_i \in C^1(\mathbb{R}^n, \mathbb{R}).$$

We already mentioned that calmness implies the Abadie CQ. A non-calm example, satisfying the Abadie CQ, is Example 1 in [18]:

Example 2. $S(p) = \{x \in \mathbb{R} \mid g(x) = x^3 \sin \frac{1}{x} \le p\}, \quad g(0) = 0.$ $\diamond$

For *convex* C^1 inequalities, S is calm at $(0, x^0)$ iff the Abadie CQ holds at all $x \in S(0)$ in some neighborhood of x^0, see [28, 4]. However, checking the latter is nontrivial, too (since - up to now- there is no efficient analytical condition for the Abadie CQ).

3.1 Normal directions

The following calmness condition applies the notion of a limiting normal cone of a closed set $M \subset \mathbb{R}^n$ at x^0:

$$(3.4) \qquad \hat{N}_M(x^0) = \{u \mid u = \lim_{k \to \infty} \lambda_k (x^k - \xi^k),\ \lambda_k \ge 0,\ x^k \to x^0,\ \xi^k \in \operatorname*{argmin}_{\xi \in M} \|x^k - \xi\|\}.$$

With the Euclidean norm, $\xi^k \in M$ is some stationary point of $\max\{\langle u^k, \xi\rangle \mid \xi \in M\}$ where $u^k = \lambda_k(x^k - \xi^k)$, and $\hat{N}_M(x^0)$ is the so-called limiting Fréchet normal cone. Under MFCQ at x^0, the cone $\hat{N}_M(x^0)$ (for $M = S(0)$ in (3.1)) has the representation

$$(3.5) \qquad \hat{N}_M(x^0) = \{u \mid u = \sum_j r_j\, Dh_j(x^0) + \sum_{i:\ g_i(x^0)=0} \lambda_i\, Dg_i(x^0), \quad \lambda_i \ge 0\}$$

and is just the (usual convex) normal cone to the set of all u satisfying (3.2). For other norms, the elements u^k are not necessarily *normals* at ξ^k in the usual sense. Nevertheless one easily shows the auxiliary result

$$\xi^* \in \operatorname*{argmin}_{\xi\in M} \|x-\xi\| \;\Rightarrow\; \xi^* \in \operatorname*{argmin}_{\xi\in M} \|\lambda x + (1-\lambda)\xi^* - \xi\| \;\; \forall \lambda \in (0,1). \tag{3.6}$$

Indeed, otherwise certain $\xi^* \in \operatorname{argmin}_{\xi\in M} \|x-\xi\|$ and $\xi \in M$ satisfy $\|\lambda x + (1-\lambda)\xi^* - \xi\| < \|\lambda x + (1-\lambda)\xi^* - \xi^*\|$ which yields a contradiction:

$$\begin{array}{l} \|x-\xi\| \\ \le \; \|\lambda x + (1-\lambda)\xi^* - \xi\| \; + \; \|(1-\lambda)(x-\xi^*)\| \\ < \; \|\lambda x + (1-\lambda)\xi^* - \xi^*\| \; + \; \|(1-\lambda)(x-\xi^*)\| \\ = \; \lambda\|x-\xi^*\| + (1-\lambda)\|(x-\xi^*)\| \; = \; \|x-\xi^*\|. \end{array}$$

□

Formula (3.6) helps for proving the next lemma with each norm. Next we put

$$M = S(0)$$

and need elements $u = \lim_{k\to\infty} \frac{x^k-\xi^k}{\|x^k-\xi^k\|} \in \hat{N}_M(x^0)$ such that x^k and ξ^k in (3.4) satisfy an additional condition in view of strict inequalities.

Lemma 3.2. *The mapping S (3.3) is not calm at $z^0 = (0, x^0)$ $\Leftrightarrow$*

(3.7) $\exists u \in \hat{N}_M(x^0)$ *such that* $u = \lim_{k\to\infty} \frac{x^k-\xi^k}{\|x^k-\xi^k\|}$
holds for certain $x^k \neq \xi^k$ *satisfying (3.4) as well as*
$g_i(x^k) < 0$ *if both* $i \in I(x^0)$ *and* $Dg_i(x^0)u > 0$. ◇

Supplement: The inequality in requirement (3.7) can be sharpened,

$$g_i(\xi^k) < g_i(x^k) < -\frac{\|x^k-\xi^k\|}{2}\, Dg_i(x^0)u \;\; \text{if } i \in I(x^0) \text{ and } Dg_i(x^0)u > 0. \tag{3.8}$$

Proof. Obviously, all components g_i with $g_i(x^0) < 0$ can be deleted since $g_i(x) < 0$ remains true for all x near x^0. Hence let $g(x^0) = 0$ to simplify the proof.
($\Leftarrow$) Given u as in (3.7), let $t_k = \|x^k - \xi^k\|$. Since $x^k, \xi^k \to x^0$ and $\xi^k \in M$, the C^1 functions satisfy $g_i(x^k) \le o(t_k)$ if $Dg_i(x^0)u \le 0$. Setting now $p_i^k = g_i(x^k)^+ := \max\{0,\ g_i(x^k)\}$, S cannot be calm since $x^k \in S(p^k)$ and $\min_{\xi\in M} \|x^k - \xi\| = \|x^k - \xi^k\| = t_k >> \|p^k\|$.
($\Rightarrow$) Let S be not calm. Then there are $(p^k, x^k) \in \operatorname{gph} S$ such that $(p^k, x^k) \to (0, x^0)$ and certain $\xi^k \in \operatorname{argmin}_{\xi\in M} \|x^k - \xi\|$ fulfill

$$\frac{\|p^k\|}{\|x^k-\xi^k\|} \to 0 \;\text{ and }\; \xi^k \to x^0. \tag{3.9}$$

Again let $t_k = \|x^k - \xi^k\|$. For $u^k = \frac{x^k-\xi^k}{t_k}$, some cluster point $u \in \hat{N}_M(x^0)$ exists. We may assume that $u^k \to u$ (otherwise pass to some subsequence). Next apply

$$p_i^k \ge g_i(x^k) = g_i(\xi^k) + t_k Dg_i(x^0)u^k + o_i(t_k) \text{ and } g_i(\xi^k) \le 0.$$

If $Dg_i(x^0)u > 0$, this yields $g_i(\xi^k) < -\frac{3}{4} t_k Dg_i(x^0)u$. So the the points $y^k = \xi^k + \frac{1}{2} t_k u^k$ $(= \frac{x^k+\xi^k}{2})$ satisfy

$$g_i(\xi^k) < g_i(y^k) < -\frac{1}{4} t_k Dg_i(x^0)u.$$

After replacing x^k by y^k, which gives new $t_k := \frac{1}{2}t_k$ and again $\xi^k \in \operatorname{argmin}_{\xi\in M} \|y^k - \xi\|$ due to (3.6), this tells us that (3.7) is satisfied even with the requirement (3.8). □

Example 3. Consider the complementarity map $S(p) = \{x \in \mathbb{R}^2 \mid x_1 \le p_1,\ x_2 \le p_2,\ \ x_1x_2 \le p_3\}$ where $S(0)$ consists of the negative half-lines and the origin $x^0 = 0$. We apply the Euclidean norm and put $x^k = (-1/k, -1/k)$, $\xi^k = (-1/k, 0)$. Then $u = (0, -1)$ satisfies $u \in \hat{N}_M(x^0)$ and $Dg_i(x^0)u \le 0\ \forall i$. So (3.7) holds true; calmness at the origin is violated.
Let $x^0 = (-1, 0)$, then $I(x^0) = \{2, 3\}$. For the related sequences x^k, ξ^k, we obtain $u = (0, 1) \in \hat{N}_M(x^0)$ if $x_2^k > 0\ \forall k$. Since $Dg_2(x^0)u > 0$ and $g_2(x^k) > 0$ now (3.7) is violated.
If $x_2^k < 0\ \forall k$ one obtains $u = (0, -1) \in \hat{N}_M(x^0)$, $g_3(x^k) > 0$ and $Dg_3(x^0)u = 1 > 0$. Hence (3.7) is again violated, S is calm at $(-1, 0)$. ◇

3.2 Reduction of inequalities

Lemma 3.2 is still far from a condition which can be checked for complicated constraint systems. However, it allows a reduction of inequalities until the set $I^+(u) := \{i \in I(x^0) \mid Dg_i(x^0)u > 0\}$ is empty.
To see this, let again $g(x^0) = 0$, assume that (3.7) holds with $I^+(u) \ne \emptyset$ and define a *reduced subsystem* by deleting in (3.1) all constraints assigned to $I^+(u)$; let its solution set mapping be denoted by S_{red}.
Again, S_{red} is not calm at $(0, x^0)$.
Indeed, otherwise calmness of S_{red} together with non-calmness of S imply that for some sequence $\{(p^k, x^k, \xi^k)\} \subset \operatorname{gph} S \times S(0)$ satisfying $(p^k, x^k) \to (0, x^0)$, $\xi^k \in \operatorname{argmin}_{\xi \in\ S(0)} \|x^k - \xi\|$ and property (3.9), there are certain points $\xi^k_{red} \in S_{red}(0)$ with $d(x^k, \xi^k_{red}) \le L\|p^k\| = o(t_k)$, where $t_k = d(x^k, \xi^k)$ and $o(.)$ and $o_i(.)$ are as in the proof of Lemma 3.2.
On the other hand, one has $g_i(\xi^k_r) > 0$ for at least one $i \in I^+(u)$ since S is not calm. Because of $g_i(\xi^k_r) = g_i(x^k) + o_i(t_k)$ then (3.8) leads to a contradiction:

$$0 < g_i(\xi^k_r) = g_i(x^k) + o_i(t_k) < -\tfrac{1}{2} t_k Dg_i(x^0)u + o_i(t_k) < 0.$$

Therefore, S_{red} is not calm.
Repeating this reduction with $S = S_{red}$ as long as possible, one obtains a non-calm subsystem of the original (3.1) one such that $I^+(u) = \emptyset$ (with some new u and with I^+ for this subsystem). In consequence, calmness holds true, if (3.7) with $I^+(u) = \emptyset$ can be excluded for all subsystems. Obviously, $I^+(u) = \emptyset$ means that u belongs to the cone

$$K = \{u \in \mathbb{R}^n \mid\ Dg_i(x^0)u \le 0\ \forall i \in I(x^0)\},$$

similarly if I^+ is considered for subsystems. So we have proved the following

Corollary 3.3. *The mapping S (3.3) is calm at $z^0 = (0, x^0)$ if for all sets $J \subset I(x^0)$, $(u \in \hat{N}_{M(J)}(x^0)$ and $Dg_i(x^0)u \le 0\ \forall i \in J)$ implies $u = 0$, provided that $M(J) = \{x \in \mathbb{R}^n \mid\ g_i(x) \le 0\, \forall i \in J\}$.* ◇ □

For instance, if MFCQ holds true for the initial system at z^0, then also for each subsystem. With $I^+(u) = \emptyset$, now (3.5) shows that (3.7) cannot hold since this would imply

$$\langle u, u\rangle = \sum_{i:\ g_i(x^0)=0} \lambda_i\ Dg_i(x^0)u \le 0.$$

Example 4. The condition of the corollary is not necessary; take the calm system

$$S(p) = \{x \in \mathbb{R} \mid x^2 \le p_1,\ x \le p_2,\ -x \le p_3\} \tag{3.10}$$

with the non-calm subsystem $x^2 \le p_1$ and $J = \{1\}$, $x^0 = 0$. ◇

One may also criticize that, without supposing MFCQ or calmness, there is no (simple) rule for determining the cone $\hat{N}_{M(J)}(x^0)$ by studying the given functions and their derivatives only. This is a drawback of many stability conditions.

There are several other sufficient calmness conditions which fit to our problem class (3.1), see e.g. [18, 17, 16]. The idea of imposing conditions for particular subsystems can be found also in Theorem 3 of [18].

Theorem 3.4. *[18] The mapping S (3.3) is calm at $(0, x^0) \in \operatorname{gph} S$ if, at x^0,*
(i) the Abadie CQ holds true and
(ii) some $u \in \mathbb{R}^n$ satisfies $Dg_i(x^0)u < 0 \ \forall i \in J$ whenever J fulfills $g_i(\xi^k) = 0 \ \forall i \in J$ for some sequence $\xi^k \to x^0, \ \xi^k \in \operatorname{bd} S(0) \setminus \{x^0\}$ (MFCQ with respect to J). $\diamond$

Proof. The original proof needs two previous theorems as well as a chain rule for directional derivatives of composed functions in [39]. So let us add a short proof, here.
Assume that S is not calm. Then one finds x^k and minimizer ξ^k in (3.4) such that the Euclidean norm fulfills $\|x^k - \xi^k\| >> \phi(x^k) := \max_i g_i(x^k) > 0, \ \xi^k, x^k \to x^0$. Here, $\xi^k \in \operatorname{bd} M$ is obvious. Passing to some subsequence, one may assume that $I(\xi^k) = J$ is constant for all k and either $\xi^k = x^0 \forall k$ or $\xi^k \neq x^0 \forall k$ holds true. Furthermore, convergence of $u^k = \frac{x^k - \xi^k}{\|x^k - \xi^k\|} \to u$ may be supposed, and

$$g_i(x^k) = Dg_i(\theta_{k,i})(x^k - \xi^k) \text{ holds for some } \theta_{k,i} \in \operatorname{conv}\{\xi^k, x^k\} \quad (\forall i \in J). \tag{3.11}$$

Since $\frac{g_i(x^k)}{\|\xi^k - x^k\|} \leq \frac{\phi(x^k)}{\|\xi^k - x^k\|} \to 0$, also

$$\limsup_{k \to \infty} \frac{g_i(x^k)}{\|\xi^k - x^k\|} \leq 0 \text{ and } Dg_i(x^0)u = \lim Dg_i(\theta_{k,i})u^k \leq 0 \quad (\forall i \in J) \tag{3.12}$$

follow. Now, we may apply the existence of Lagrange multipliers for the minimizers ξ^k. Assume first $\xi^k = x^0 \forall k$. Then, it holds $J = I(x^0)$ and - because of (i) - solvability of

$$P(u^k): \qquad u^k = \sum_{i \in J} \lambda_i \ Dg_i(x^0), \quad \lambda_i \geq 0, \qquad \text{where } \lambda = \lambda(k)$$

is ensured, which yields solvability of the linear system $P(u)$. Thus $1 = \sum_{i \in J} \lambda_i \ Dg_i(x^0)u$ holds with certain $\lambda_i \geq 0$, in contradiction to $Dg_i(x^0)u \leq 0 \forall i \in J$ from (3.12).
Let $\xi^k \neq x^0 \forall k$. By (ii), MFCQ holds w.r. to the subsystem $(g_i \leq 0, \ i \in J)$ at x^0, so it also holds at ξ^k near x^0. Hence there are $\lambda_i \geq 0$ (depending on k) such that

$$x^k - \xi^k = \sum_{i \in J} \lambda_i \ Dg_i(\xi^k). \tag{3.13}$$

Moreover, there is some C, not depending on k, such that

$$\|\lambda\| \leq C \|x^k - \xi^k\| \tag{3.14}$$

is valid for large k (multiply in (3.13) with a MFCQ- direction v for $\xi = x^0$). Using (3.11) ... (3.14) we obtain again a contradiction, namely

$$\begin{array}{ll} \|x^k - \xi^k\|^2 & \\ = \sum_{i \in J} \lambda_i \ (\ Dg_i(\theta_{k,i})(x^k - \xi^k) & + \ [\ Dg_i(\xi^k) - Dg_i(\theta_{k,i}) \ ](x^k - \xi^k) \ ) \\ = \sum_{i \in J} \lambda_i \ g_i(x^k) & + \ \sum_{i \in J} \lambda_i \ [Dg_i(\xi^k) - Dg_i(\theta_{k,i})](x^k - \xi^k) \\ \leq \ o(\|x^k - \xi^k\|^2) & + \ o(\|x^k - \xi^k\|^2). \end{array}$$

Hence S is calm. $\square$

Example 5. Again, this sufficient condition is not necessary, take the linear and calm mapping

$$S(p) = \{(x_1, x_2) |\; x_2 \le p_1,\; -x_2 \le p_2\} : \text{ check (ii) for } J = \{1, 2\},\; \xi^k = (\frac{1}{k}, 0) \to (0, 0). \quad \diamond$$

The reason for the gap between necessity and sufficiency in Corollary 3.3 and Theorem 3.4 consists in an inappropriate definition of the sets J, cf. Theorem 4.6.

3.3 Crucial limits

To obtain *explicit necessary or sufficient calmness conditions* from Lemma 3.2, the local structure of $M = S(0)$ plays a decisive role, even if the vectors $u^k = \frac{x^k - \xi^k}{\|x^k - \xi^k\|}$ can be written by Lagrange multipliers

$$u^k = \sum_{i:\; g_i(\xi^k)=0} \lambda_i^k \, Dg_i(\xi^k); \qquad \lambda_i^k \ge 0. \tag{3.15}$$

The latter is guaranteed if S satisfies the Abadi CQ at all $\xi^k \in M$ near x^0. In this case, violation of calmness means (equivalently) by Lemma 3.2, that a limit of the form

$$0 < \langle u, u \rangle = \lim_{k\to\infty} \langle u^k, u \rangle = \lim_{k\to\infty} \sum_{i:\; g_i(\xi^k)=0,\; \lambda_i^k \ge 0} \lambda_i^k \, Dg_i(\xi^k) u$$

is positive though $(u^k, \xi^k) \to (u, x^0)$ and $Dg_i(x^0)u \le 0$ hold for the involved constraints. Evidently, this may happen only if certain λ_i^k diverge and some gradient is not constant (hence not under MFCQ at z^0 or for linear systems). After selecting an appropriate subsequence, the limits of u^k (3.15) can be written (more abstractly) as

$$\limsup_{\xi \to x^0,\; \lambda \in \Lambda} \sum_i \lambda_i \, Dg_i(\xi), \qquad \Lambda \text{ a polyhedral cone} \tag{3.16}$$

where ξ satisfies the "face condition" $g(\xi) = 0$. Without this face condition, the upper Hausdorff (or Kuratovski-Painlevé) limit (3.16) is also crucial for characterizing the *strong Lipschitz property* – cf. Remark 1 – of stationary points (the x-components of KKT -tuples) in parametric C^2 and $C^{1.1}$ optimization [23]. There, one also finds a formula for the limits (3.16) if they represent linear constraints with at most one quadratic condition (e.g. a complementarity condition). On the other hand, given any ν, the limits in (3.16) do not depend on the first ν derivatives at x^0 only (convex polynomial examples are given).
Since the same limits (3.15) (or lim sup (3.16)) are important for quite different stability problems, it remains a callenge for the future to describe them in some more involved way.

3.4 Intersections

A simple way of dealing with system (3.1), even with arbitrary locally Lipschitz functions g and h, consists in a splitting approach. Split all constraints into two families such that

$$S(p) = U(y) \cap V(z) \;\text{ and }\; p = (y, z). \tag{3.17}$$

For instance, one could put $y = p_1, z = p_2$, whereafter U and V represent the inequalities and equations in (3.1), respectively. Alternatively, one may assume (e.g.) that U collects all linear constraints and V the remaining ones. Obviously, S is calm at $(0, x^0)$ only if the both mappings

$$U_0(y) = U(y) \cap V(0) \quad \text{ and } \quad V_0(z) = U(0) \cap V(z)$$

are calm at $(0, x^0)$. By Theorem 3.6 in [21], also some reverse statement holds true for a big class of calm multifunctions U and V in metric spaces. In particular, it holds

Lemma 3.5. *The mapping S (3.1, 3.17) is calm at $(0, x^0)$ if so are U, V and V_0.* $\diamond$

Example 6. The KKT system for problem (2.1) can be witten as intersection of

$$\begin{array}{l} U(y) = \{(\xi,\eta) \mid -\eta \le y \ \} \\ V(z) = \{(\xi,\eta) \mid Df_0(\xi) + \sum_{i=1}^m \eta_i Df_i(\xi) = z^1, \ f_i(\xi) \le z_i^2, \ \eta_i f_i(\xi) = z_i^3 \ \} \ \text{ at } \ (y,z) = 0. \end{array}$$

The lemma says that $U \cap V$ is calm if V and V_0 are calm (U is trivally calm here). $\diamond$

The calmness-hypothesis for V_0 is essential.

Example 7. Let $U(y) = \{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 - 1 \le y\}$, $V(z) = \{x \in \mathbb{R}^2 \mid 1 - x_2 \le z\}$. Both, U and V are calm at $(0, \ (0,1))$. The mapping $S(y,z) = U(y) \cap V(z)$ is not calm at $((0,0) \ (0,1))$. Indeed, it holds $S(0,0) = \{(0,1)\}$, and already the both mappings $U_0(y) = U(y) \cap V(0)$ as well as $V_0(z) = V(z) \cap U(0)$ are not calm at $(0 \ (0,1))$ since $(\sqrt{y}, \ 1) \in U_0(y)$ and $(\sqrt{2z - z^2}, 1 - z) \in V_0(z)$ violate the calmness conditions for small positive y, z. $\diamond$

Once more however, the condition of the lemma is not necessary since, for calm $U \cap V$, it may happen that one of the mappings U, V is not calm. We refer to (3.10), where the quadratic constraint alone (or together with only one linear constraint) forms a non-calm subsystem.

Summarizing, we may state that a sharp characterization of calmness for finite- dimensional nonlinear (C^k -) systems is not possible up to now (at least by our knowledge) in terms of the original functions and their derivatives (until some fixed order). An important exception occurs for piecewise linear systems (or polyhedral multifunctions) since such systems can be reformulated as (a finite union of) linear systems, cf. [33, 36]. Furthermore, though weaker than the Aubin property or MFCQ, calmness may turn out to be a quite strong sufficient condition for ensuring the existence of Lagrange multipliers to an optimization problem. This reduces the meaning of calmness for this purpose.
Nevertheless, calmness does not only describe a useful error estimate for inclusions. We are now going to show that calmness implies linear convergence for certain solution methods and vice versa. Surprisingly, the latter holds under rather general hypotheses.

4 Calmness of general mappings and of Lipschitzian level sets

4.1 Basic transformations

Though we are speaking now about closed *multifunctions* $S : P \rightrightarrows X$ which act between Banach spaces, calmness is a *monotonicity property* with respect to two canonically assigned *Lipschitz functions*: the distance of x to $S(p^0)$ and the graph-distance

$$\psi_S(x,p) = \operatorname{dist}((p,x), \ \operatorname{gph} S),$$

defined via the norm $\|(p,x)\| = \max\{\|p\|, \|x\|\}$ or some equivalent norm in $P \times X$.

Lemma 4.1. *S is calm at $(p^0, x^0) \in \operatorname{gph} S$ if and only if*

$$(4.1) \qquad \exists \varepsilon > 0, \ \alpha > 0 \textit{ such that } \ \alpha \ \operatorname{dist}(x, S(p^0)) \le \ \psi_S(x,p^0) \quad \forall x \in x^0 + \varepsilon B. \quad \diamond$$

In other words, calmness at (p^0, x^0) is violated iff

$$(4.2) \qquad 0 < \psi_S(x^k, p^0) = o(\ \operatorname{dist}(x^k, S(p^0)) \ ) \text{ holds for some sequence } x^k \to x^0.$$

Proof. A proof is possible as for Lemma 3.2 in [21]; we verify Lemma 4.1 for completeness. Let (4.1) hold true. Then, given $x \in S(p) \cap (x^0 + \varepsilon B)$, it holds $\psi_S(x,p^0) \le d((p,x),(p^0,x)) = \|p - p^0\|$ and, in consequence, $\alpha \ \operatorname{dist}(x, S(p^0)) \le \psi_S(x,p^0) \le \|p - p^0\|$ which yields calmness

with rank $L = \frac{1}{\alpha}$.
Conversely, let (4.1) be violated, i.e., (4.2) be true. Given any positive $\delta_k < o(\operatorname{dist}(x^k, S(p^0)))$, we find $(p^k, \xi^k) \in \operatorname{gph} S$ such that

$$d((p^k, \xi^k),\ (p^0, x^k)) < \psi_S(x^k, p^0) + \delta_k < \ b_k := 2\ o(\operatorname{dist}(x^k, S(p^0))\ ).$$

In addition, the triangle inequality $\operatorname{dist}(x^k, S(p^0)) \le d(x^k, \xi^k) + \operatorname{dist}(\xi^k, S(p^0))$ yields

$$\operatorname{dist}(\xi^k, S(p^0)) \ge \operatorname{dist}(x^k, S(p^0)) - d(\xi^k, x^k) > \operatorname{dist}(x^k, S(p^0)) - b_k.$$

Using also the evident inequality $\|p^k - p^0\| < b_k$, we thus obtain for $\xi^k \in S(p^k)$,

$$\frac{\|p^k - p^0\|}{\operatorname{dist}(\xi^k, S(0))} < \frac{b_k}{\operatorname{dist}(x^k, S(p^0)) - b_k} \to 0 \quad \text{as } k \to \infty.$$

Hence, since $\xi^k \to x^0$, S cannot be calm at (p^0, x^0). □

Estimates of ψ_S, for composed systems, can be found in [21]. Condition (4.1) requires that $\psi_S(., p^0)$ increases in a Lipschitzian manner if x (near x^0) leaves $S(p^0)$. Clearly, this property depends on the local structure of the boundaries of $\operatorname{gph} S$ and $S(p^0)$. For convex multifunctions (i.e. $\operatorname{gph} S$ is convex), ψ_S and $d(., S(p^0))$ are even convex.
Combined with Remark 1(iii), condition (4.1) characterizes the Aubin property, too. Concerning similar characterizations of other stability properties we refer to [24]. The distance ψ_S can be applied also for both characterizing optimality and computing solutions in optimization models via penalization [27, 21] and [22, Chapt. 2]; for the particular context of exact penalization techniques, see also [9, 6, 3]. The approximate minimization of ψ_S, defined via the norm $\|(p, x)\| = \|p\| + \lambda\|x\|$ ($\lambda > 0$ fixed), plays a key role in [25].

Evidently, setting $G = \psi_S$ we obtain a (globally) Lipschitz function $G : X \times P \to \mathbb{R}$, assigned to S, such that

$$(p, x) \in \operatorname{gph} S \ \Leftrightarrow\ G(x, p) \le 0. \tag{4.3}$$

For every such description of $\operatorname{gph} S$, it follows

Lemma 4.2. *S is calm at (p^0, x^0) if*

$$\exists \varepsilon > 0,\ \alpha > 0 \ \textit{such that}\ \ \alpha\ \operatorname{dist}(x, S(p^0)) \le\ G(x, p^0) \qquad \forall x \in x^0 + \varepsilon B. \qquad \diamond \tag{4.4}$$

Proof. Given any $\delta > 0$ choose $(p', x') \in \operatorname{gph} S$ with $d((p', x'), (p^0, x)) < \psi_S(x, p^0) + \delta$. Then $G(x', p') \le 0$ yields with some Lipschitz constant L, $G(x, p^0) \le L\ (\psi_S(x, p^0) + \delta)$ and $\frac{\alpha}{L}\ \operatorname{dist}(x, S(p^0)) \le \frac{G(x,p^0)}{L} \le \psi_S(x, p^0) + \delta$. Hence one obtains, via $\delta \downarrow 0$, that (4.1) holds with some new $\alpha := \frac{\alpha}{L}$. □

On the other hand, the conditions (4.3) and (4.4) are only sufficient for calmness (put, e.g., $G = \psi_S^2$). Nevertheless, the lemmata obviously ensure

Corollary 4.3. *A multifunction S is calm at (p^0, x^0) if and only if there is some Lipschitz function $G : X \times P \to \mathbb{R}$ satisfying (4.3) and (4.4).* ◇ □

Finally, with any locally Lipschitz function $\phi : X \to \mathbb{R}$ such that

$$c_1\phi(x) \le \psi_S(x, p^0) \le c_2\phi(x) \ \text{ for } x \text{ near } x^0 \text{ and certain constants } 0 < c_1 \le c_2 \tag{4.5}$$

and with the mapping

$$\Sigma(q) = \{x \in X \mid \phi(x) \le q\}, \tag{4.6}$$

condition (4.1) of Lemma 4.1 is equivalent to

$$(4.7) \qquad \exists \varepsilon > 0,\ \alpha > 0 \text{ such that } \alpha \ \operatorname{dist}(x,\ \Sigma(0)) \le\ q \quad \forall x \in x^0 + \varepsilon B \text{ with } \phi(x) = q > 0.$$

This verifies

Remark 2. Calmness for any closed multifunction S at (p^0, x^0) can be reduced to the particular case of calmness of a Lipschitzian inequality only, namely to calmness of Σ (4.6) at $(0, x^0)$ where $\phi = \psi_S(., p^0)$ or ϕ is another Lipschitz function satisfying (4.5).

4.2 Level sets and the algorithmic approach

According to Remark 2, we study calmness of Σ (4.6) for any (locally) Lipschitz function $\phi : X \to \mathbb{R}$ on a Banach space X. In particular, we pay attention to the case of

$$(4.8) \qquad \phi(x) = \max_{i \in I} g_i(x) \quad \text{where } g_i \in C^1(X, \mathbb{R}) \ \text{ and } I = \{1, 2, ..., m\}$$

which is of interest for many applications (for a compact topological space I we refer to Remark 4). The next statement follows from Theorem 3 in [25] and shows the big difference between the (general) Lipschitzian and piecewise differentiable case (4.8). We add a self-contained, constructive proof which presents the related constants directly. First of all, we define some *relative slack of* g_i in comparison with ϕ (4.8) as in [25].

$$(4.9) \qquad s_i(x) = \frac{\phi(x) - g_i(x)}{\phi(x)} \qquad \text{if } \phi(x) > 0.$$

Theorem 4.4. *Let $\phi : X \to \mathbb{R}$ be (locally) Lipschitz and $\phi(x^0) = 0$.*

(i) Then Σ (4.6) is calm at $(0, x^0)$ if and only if there are $\lambda, \varepsilon \in (0, 1)$ such that, for all $x \in\ x^0 + \varepsilon B$ with $\phi(x) > 0$, there exist $u \in B$ and $t > 0$ satisfying

$$(4.10) \qquad \frac{\phi(x + tu) - \phi(x)}{t} \le -\lambda \quad \text{and} \quad \lambda\phi(x) \le t \le \frac{1}{\lambda}\phi(x).$$

(ii) For the maximum function ϕ (4.8), one may delete t and replace condition (4.10) by

$$(4.11) \quad Dg_i(x)u \le \frac{s_i(x)}{\lambda} - \lambda \quad \text{or alternatively by} \quad Dg_i(x^0)u \le \frac{s_i(x)}{\lambda} - \lambda \quad \forall i \in I. \qquad \diamond$$

Notice that nothing is required if $\phi \le 0$ on $x^0 + \varepsilon B$.

Proof. Let $L_\phi \ge 1$ be a Lipschitz constant for ϕ (near x^0).

(i) Necessity of (4.10): Calmness with rank $L > 0$ allows to put $u = \frac{\xi - x}{\|\xi - x\|}$ and $t = \|\xi - x\|$ where $\xi \in \Sigma(0)$ and $t \le L \cdot \phi(x)$. Since $\phi(x + tu) \le 0$ this yields for $\phi(x) > 0$:

$$\frac{\phi(x + tu) - \phi(x)}{t} \le \frac{-\phi(x)}{t} \le -L^{-1} \qquad \text{and} \ \ t \le L\ \phi(x).$$

On the other hand, the Lipschitz estimate for points near x^0 yields

$$\frac{\phi(x)}{t} \le \frac{|\phi(x + tu) - \phi(x)|}{t} \le L_\phi \|u\| = L_\phi \quad \text{and} \quad \frac{1}{L_\phi}\phi(x) \le t.$$

Therefore, (4.10) holds true if $0 < \lambda \le \frac{1}{L_\phi}$ and $\frac{1}{\lambda} \ge L$, i.e., if $0 < \lambda \le \min\{L_\phi^{-1},\ L^{-1}\}$.
Sufficiency of (4.10): Put $\theta = 1 - \lambda^2$. Taking a sufficently small $\delta \in (0,\ \frac{1}{2}\varepsilon)$ we have

$$(4.12) \qquad \frac{\lambda^{-1}}{1 - \theta}\ \phi(x) \le \frac{\lambda^{-1}}{1 - \theta} L_\phi\ d(x, x^0) < \tfrac{1}{2}\varepsilon \quad \forall x \in x^0 + \delta B.$$

Now let any $x \in x^0 + \delta B$ with $\phi(x) > 0$ be arbitrarily given. Selecting, for $x^1 = x$, related u^1 and t_1 from (4.10), we obtain for $x^2 = x^1 + t_1 u^1$:

$$\phi(x^2) \le \phi(x^1) - \lambda\, t_1 \quad \text{and} \quad \lambda^2 \phi(x^1) \le \lambda\, t_1 \le \phi(x^1),$$

hence also

$$\phi(x^2) \le (1-\lambda^2)\phi(x^1) = \theta\phi(x^1) \quad \text{and} \quad \|x^2 - x^1\| \le t_1 \le \lambda^{-1}\phi(x^1). \tag{4.13}$$

Because of (4.12) we have $x^2 \in x^0 + \varepsilon B$. This allows us to construct a sequence $x^{k+1} = x^k + t_k u^k$ which, beginning with $k = 1$, satisfies

$$\phi(x^{k+1}) \le \theta^k \phi(x^1) \quad \text{and} \quad t_k \le \lambda^{-1}\phi(x^k) \le \theta^{k-1}\lambda^{-1}\phi(x^1) \tag{4.14}$$

whenever $x^k \in x^0 + \varepsilon B$ and $\phi(x^k) > 0$ (if $\phi(x^k) \le 0$ put $x^{k+1} = x^k,\ t_k = 0$). Indeed, due to

$$\|x^{k+1} - x\| \le \|x^{k+1} - x^k\| + \ldots + \|x^2 - x^1\| \le \Big(\sum_{j=0}^{\infty} \theta^j\Big)\lambda^{-1}\phi(x) = \frac{\lambda^{-1}}{1-\theta}\,\phi(x) < \tfrac{1}{2}\varepsilon$$

and $\|x - x^0\| < \delta$, the hypothesis (4.10) can be applied to all x^k as long as $\phi(x^k) > 0$. Thus, we generate a Cauchy sequence in $x^0 + \varepsilon B$. The existing limit $\xi = \lim x^k$ fulfills, by (4.14), $\phi(\xi) = 0$ as well as the calmness condition $d(\xi, x) \le L\ \phi(x)$ with $L = \frac{\lambda^{-1}}{1-\theta} = \lambda^{-3}$.

(ii) We show that the assertion follows from the first part (i) and the uniform convergence

$$\limsup_{i\in I,\ x\to x^0,\ t\downarrow 0,\ \|u\|\le 1} \left| \frac{g_i(x+tu) - g_i(x)}{t} - Dg_i(x^0)u \right| = 0. \tag{4.15}$$

The sequence $x^{k+1} = x^k + t_k u^k$ can be constructed by setting $t_k = \lambda\phi(x^k)$ now.
We verify first that (4.10) implies (4.11). Indeed, the first condition of (4.10) becomes

$$\begin{array}{rlll} & \frac{\phi(x+tu)-\phi(x)}{t} & \le -\lambda & \\ \Leftrightarrow & g_i(x+tu) - \phi(x) & \le -\lambda t & \forall i \\ \Leftrightarrow & \frac{g_i(x+tu)-g_i(x)}{t} & \le -\lambda + \frac{\phi(x)-g_i(x)}{t} & \forall i. \end{array} \tag{4.16}$$

Applying $\lambda\phi(x) \le t \le \frac{1}{\lambda}\phi(x)$, which ensures $t \downarrow 0$ as $x \to x^0$, this also yields

$$\frac{g_i(x+tu) - g_i(x)}{t} \le \frac{\phi(x) - g_i(x)}{t} - \lambda \le \frac{\phi(x) - g_i(x)}{\lambda\ \phi(x)} - \lambda = \frac{s_i(x)}{\lambda} - \lambda \quad \forall i. \tag{4.17}$$

With $\lambda' = \frac{1}{2}\lambda$ and x near x^0, we thus obtain from (4.15),

$$Dg_i(x^0)u \le \frac{s_i(x)}{\lambda'} - \lambda' \ \text{ and } \ Dg_i(x)u \le \frac{s_i(x)}{\lambda'} - \lambda' \quad (\forall i). \tag{4.18}$$

Hence (4.10) implies (4.11) (with new λ) for the max-function.
Conversely, having (4.18) for all x near x^0 with $\phi(x) > 0$ and $u = u(x) \in B$, we may conclude that, for $\lambda = \frac{1}{2}\lambda'$ and $t = \lambda\phi(x)$,

$$\frac{g_i(x+tu) - g_i(x)}{t} \le \frac{s_i(x)}{\lambda} - \lambda = \phi(x)\frac{s_i(x)}{t} - \lambda = \frac{\phi(x) - g_i(x)}{t} - \lambda \qquad (\forall i).$$

By (4.16) the latter yields (4.10). Hence also (4.11) implies (4.10) (with new λ). □

Remark 3. (*Applying generalized derivatives*) While the first condition of (4.10) is a usual descent condition, the second one looks strange and does not appear in the context of known generalized derivatives or co-derivatives for (multi-) functions. Both estimates of t are essential: the upper one for obtaining a convergent sequence $\{x^k\}$ as well as a Lipschitz estimate of $d(\xi, x)$, the lower one for $\phi(x^k) \to 0$. So it is not surprising that all sufficient calmness conditions, based on known concepts of generalized (co-) derivatives for arbitrary Lipschitz functions or multifunctions, are not necessary – even for finite-dimensional systems. ◇

4.3 Solution method and calmness for systems of C^1 inequalities

Again, let X be a Banach space in this subsection.
It is important that the proof of Theorem 4.4 involves a procedure which finds some element $\xi \in \Sigma(0)$ such that $d(\xi, x) \le L\phi(x)$ (if $\phi(x) > 0$). This procedure can be rewritten as a locally convergent algorithm for solving $\xi \in \Sigma(0)$ whenever u and t in (4.10) can be determined.
As a typical situation, we continue in considering the case of Σ (4.6) with the max-function (4.8), i.e., we study

$$S(p) = \{x \in X \mid g_i(x) \le p_i \ \forall i \in I\}, \quad g \in C^1(X, \mathbb{R}^m), \quad I = \{1, 2, ..., m\} \tag{4.19}$$

and know that calmness of S at $(0, x^0) \in \operatorname{gph} S$ is equivalent, by (4.5), to calmness of

$$\Sigma(q) = \{x \in X \mid \phi(x) \le q \ \forall i \in I\}, \qquad \phi(x) = \max_i g_i(x)$$

at $(0, x^0) \in \mathbb{R} \times X$.
Next, the calm situation will be completely characterized by an algorithm called ALG3 in [25] which uses the relative slack s_i (4.9) and the quantities

$$b_i(x, \lambda) = \frac{s_i(x)}{\lambda} - \lambda \qquad \text{for } \phi(x) > 0, \ \lambda > 0. \tag{4.20}$$

Obviously, $b_i(x, \ .)$ is decreasing in λ.

ALG3: Given $x^k \in X$ and $\lambda_k > 0$, put $x^{k+1} = x^k$ and $\lambda_{k+1} = \lambda_k$ in the trivial case of $\phi(x^k) \le 0$. Otherwise solve the (convex) system

$$Dg_i(x^k)u \ \le \ b_i(x^k, \lambda_k) \quad \forall i \in I, \qquad \|u\| \le 1. \tag{4.21}$$

Having a solution u, put $\quad x^{k+1} = x^k + \lambda_k \phi(x^k) u, \quad \lambda_{k+1} = \lambda_k,$
otherwise put $\quad x^{k+1} = x^k, \quad \lambda_{k+1} = \frac{1}{2}\lambda_k.$

For $X = \mathbb{R}^n$, sum-norm $\|.\|_1$ and $\phi(x^k) > 0$, it suffices to solve the linear program

$$\min \sum u_i^+ + u_i^- \ \ s.t. \ \ Dg_i(x^k)(u^+ - u^-)^T \le b_i(x^k, \lambda_k) \quad \forall i \in I, \ u^+ \ge 0, u^- \ge 0 \tag{4.22}$$

and to check whether $u = u^+ - u^-$ satisfies $\|u\|_1 \le 1$ (in case of solvability).

Theorem 4.5. *The mapping S (4.19) is calm at $(0, x^0)$ if and only if there is some $\alpha > 0$ such that, for $\|x^1 - x^0\|$ small enough and $\lambda_1 = 1$, it follows $\lambda_k \ge \alpha \ \forall k$ for ALG3. In this case, the sequence $\{x^k\}_{k \ge 1}$ converges to some $\xi \in S(0)$, and satisfies, for $\phi(x^k) > 0$,*

$$\phi(x^{k+1}) \le (1 - \beta^2)\phi(x^k) \ \text{ whenever } \ 0 < \beta < \frac{\alpha^2}{1 + \sup_i \|Dg_i(x^0)\|}. \qquad \diamond \tag{4.23}$$

Proof. The first statement follows immediately from Theorem 4.4. For a proof of the estimate, we refer to Theorem 4 in [25]. $\square$

Notice that (4.23) yields

$$\|x^{k+1} - x^k\| \le \lambda_k \phi(x^k) \le \phi(x^k) \le (1 - \beta^2)^{k-1} \phi(x^1).$$

In consequence, calmness holds with rank

$$L = \beta^{-2}, \ \textit{since} \ \ \|\xi - x^1\| \ \le \ \phi(x^1) \sum_{k \ge 1} (1 - \beta^2)^{k-1} = \frac{1}{\beta^2} \, \phi(x^1).$$

Remark 4. (*Infinitely many constraints.*) As in semi-infinite programs (but without supposing $\dim X < \infty$ here), one can consider S (4.19) with a compact topological space I, $\|p\| = \sup_i |p_i|$, and a continuous map $(i,x) \mapsto g_i(x)$ which is uniformly (in view of $i \in I$) locally Lipschitz w.r. to x near x^0. Further, write $g \in C^1$ if all derivatives $Dg_i(x)$ w.r. to x exist and are continuous on $I \times X$. Then, the Theorems 4.4 and 4.5 remain true, due to (4.15), with the same proof. $\diamond$

Remark 5. (*Intersection with closed sets.*) Suppose the mapping S (4.19) or the level set map Σ of the Lipschitz function ϕ in the Theorems 4.4 and 4.5 are restricted to some additional fixed condition $x \in M$ where $M \subset X$ is a closed set,

$$\begin{aligned} S(p) = S_M(p) &= \{x \in M \mid \quad g_i(x) \le p_i \ \forall i \in I\}, \\ \Sigma(q) = \Sigma_M(q) &= \{x \in M \mid \quad \phi(x) \le q \ \}. \end{aligned} \tag{4.24}$$

Then, the statements are again true with the same proof, provided the points x, x^k are taken in M, the C^1- property holds on an open set containing M, and the extra conditions $x+tu \in M$, $x + \lambda\phi(x)u \in M$ and $x^k + \lambda_k \phi(x^k)u \in M$ are added in (4.10), (4.11) and (4.21), respectively. $\diamond$

4.4 Assigned linear inequality systems

We continue in considering the mapping S (4.19) in order to clarify that certain inequality systems of the kind $Dg_j(x^0)u < 0 \ \forall j \in J$ are crucial for calmness, and to indicate the sets J which play the essential role.

Theorem 4.6. *Let* $\phi(x^0) = \max_i g_i(x^0) = 0$. *Then, the mapping S (4.19) is calm at* $(0, x^0)$ *if and only if each system*

$$Dg_i(x^0)u < 0 \quad \forall i \in J \tag{4.25}$$

is solvable, whenever J fulfills $J = \{i \mid \lim_{k\to\infty} s_i(x^k) = 0\}$ *for certain* $x^k \to x^0, \phi(x^k) > 0$. $\diamond$

Comments:
(i) The set J collects the active ($g_i = \phi$) and "almost active" functions g_i for the given sequence of $x^k \notin S(0)$. It holds $J \subset I(x^0) = \{i \mid g_i(x^0) = 0\}$, and $J = \emptyset$ is possible (e.g. if $g(x) \equiv 0$). For $J = \emptyset$, system (4.25) is solvable by definition.
(ii) Well-known duality statements yield: (4.25) is unsolvable $\Leftrightarrow 0 \in \operatorname{conv}\{Dg_i(x^0) \mid i \in J\} \Leftrightarrow u = 0$ minimizes $\max_{i\in J} Dg_i(x^0)u$. For (affine-) linear g_i, so calmness follows from the simple fact that the same holds at x^k, too. Because of $g_i(x^k) > 0 \ \forall i \in J$ this would imply $\max_{i\in J} g_i(x^k + u) > 0$, a contradiction for $u = x^0 - x^k$.
(iii) Solvability of (4.25) means that the mapping $S^J(p) = \{x \in X \mid g_i(x) \le p_i \ \forall i \in J\}$ obeys the Aubin property at $(0, x^0)$.
(iv) With the (larger) sets $J = \{i \mid g_i(x^k) > 0\}$ and some nonlinear g_i, the given condition is no longer necessary, cf. (3.10).

Proof. We consider sequences $x = x^k \to x^0$ with $\phi(x) > 0$ and $\lambda = \lambda_k \downarrow 0$ such that, for

$$b_i = \frac{s_i(x)}{\lambda} - \lambda \qquad (\text{ where } b = b(k) \text{ depends on } k \to \infty \text{ }),$$

the limits $l_i = \lim_{k\to\infty} b_i \in [0, \infty]$ exist. We call such a sequence (x^k, λ_k) *critical.*
By definition, $l_i = 0$ yields, due to $s_i = \lambda^2 + \lambda b_i$, that $s_i = o_i(\lambda)$. Conversely, $s_i = o_i(\lambda) = \alpha_i\lambda$ for $\alpha_i \to 0$, implies $s_i = \lambda^2 + \lambda b_i$ with $b_i = \alpha_i - \lambda \to 0$. Thus,

$$l_i = 0 \quad \text{simply means} \quad s_i(x) = o_i(\lambda).$$

Define as above,

$$J = \{i \mid \lim_{k\to\infty} s_i(x^k) = 0\} \quad \text{and} \quad \mu(x) = \max_{i\in J} s_i(x).$$

Next we shall modify λ, for a critical sequence, in such a way that

$$l_i = 0 \text{ and } b_i < 0 \qquad \forall i \in J. \tag{4.26}$$

If $\mu(x) = 0$ or $s_i(x) = 0$, we have nothing to do since $s_i(x) = o_i(\lambda)$ and $b_i = -\lambda_k < 0$ follow immediately. Otherwise, both $s_i = o_i(\lambda)$ and $b_i < 0$ can be satisfied for $i \in J$ by increasing the elements $\lambda_k \downarrow 0$ if necessary. So it suffices to put

$$\lambda_k = 2\sqrt{\mu(x^k)}$$

which ensures

$$s_i(x^k)/\lambda_k \le \tfrac{1}{2}\sqrt{s_i(x^k)} \to 0 \quad \text{and} \quad b_i\sqrt{\mu(x^k)} = \tfrac{1}{2}s_i(x) - 2\mu(x^k) < 0.$$

Hence, given any sequence $x^k \to x^0$ with $\phi(x^k) > 0$, there are λ_k (depending on x^k) such that (x^k, λ_k) is a critical sequence satisfying (4.26). We call such a sequence *critical**.
Under calmness, we know by Theorem 4.4 that the system

$$Dg_i(x^0)u \le b_i \ \forall i \in I, \quad \|u\| \le 1 \qquad \text{where } b = b(k) \tag{4.27}$$

is solvable for all k (sufficiently large), even if $b(k)$ is defined by a critical* sequence (x^k, λ_k). Conversely, if calmness is violated, Theorem 4.4 ensures, for certain $x^k \to x^0$, $\phi(x^k) > 0$, $\lambda_k \downarrow 0$, that (4.27) is inconsistent for all k. By passing to some subsequence, (x^k, λ_k) is critical. By increasing λ_k if necessary up to $2\sqrt{\mu(x^k)}$ (this makes b_i smaller) (4.27) remains inconsistent and also (4.26) holds true.
Therefore, calmness at $(0, x^0)$ is equivalent to solvability of (4.27) for all critical* sequences. For such sequences, we may omit all inequalities of (4.27) which are assigned to $i \notin J$ since, due to $b_i \to l_i > 0$, these inequalities already hold for small $\|u\|$ (and large k), namely if $\|u\| \le \min_{i\in I\setminus J} l_i \ \|1 + Dg_i(x^0)\|^{-1}$. In consequence, (4.27) may be replaced by

$$Dg_i(x^0)u \le b_i \ \forall i \in J, \quad \|u\| \le 1; \quad \text{where } 0 > b_i \to 0 \ as\ k \to \infty. \tag{4.28}$$

Using finally that (4.28) is solvable for all sufficiently large k iff system (4.25) is consistent, we obtain the claimed result. □

The situation for $S = S_M$.
For $S = S_M$, calmness at $(0, x^0)$ means similarly the existence of solutions u to (4.27) for all critical*- sequences with $x^k \in M$ and $x^k + \lambda_k\phi(x^k)u \in M$. One obtains now a sufficient calmness condition after replacing (4.25) by

$$Dg_i(x^0)u < 0 \quad \forall i \in J, \quad u \in T^C_M(x^0) \tag{4.29}$$

where

$$T^C_M(x^0) = \{u \mid \lim_{k\to\infty} \frac{\operatorname{dist}(x^k + t_k u, M)}{t_k} = 0 \quad \forall\ t_k \downarrow 0, x^k \to x^0, x^k \in M\}$$

is Clarke's tangent cone of M at x^0. The condition is only sufficient since we consider *particular* $t_k = \lambda_k\phi(x^k)$. It is known [6], [39] that the possibilities for an analytical description of this cone depend on the description of M.
The main problem for direct applications of Theorem 4.6 consists in finding the crucial sets J. Less directly, it can be also used to see that certain functions are not important for calmness.

Corollary 4.7. *(Complementarity)*
Suppose that system S (4.19) contains (among others) three conditions of the type

$$g_1(x) = u(x) \le p_1, \quad g_2(x) = v(x) \le p_2, \quad g_3(x) = u(x)v(x) \le p_3$$

which require complementarity $u \le 0,\ v \le 0,\ uv = 0$ *for* $p = 0$*. Let* x *be restricted to the set* $M = \{x \mid \max\{u(x), v(x)\} \ge 0\}$*. Then calmness of* $S = S_M$ *at* $(0, x^0)$ *does not depend on the condition* $g_3 \le p_3$ *if strict complementarity is violated, i.e., if* $u(x^0) = v(x^0) = 0$*.*

Proof. Let $x^k \to x^0$, $\phi(x^k) > 0$. Because of $u(x^0)v(x^0) = 0$ it holds $u(x^k)v(x^k) << \max\{u(x^k), v(x^k)\} \le \phi(x^k)$ whenever $g_3(x^k) > 0$ and $x^k \in M$. This implies $l_3 = \lim s_3(x^k) \ge 1$, hence $3 \notin J$. $\square$

If the usual complementarity condition $x_1 \le 0,\ x_2 \le 0,\ x_1x_2 = 0$ is involved, the setting $M = \{x \mid x_1x_2 = 0\}$ is well-known standard. Of course, here and in the corollary, also more than 1 complementarity pair can be considered.

Acknowledgement
The authors are indebted to D. Klatte for various helpful discussions and constructive comments concerning the subject of this paper.

References

[1] J.-P. Aubin and I. Ekeland. *Applied Nonlinear Analysis.* Wiley, New York, 1984.

[2] J.F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems.* Springer, New York, 2000.

[3] J. V. Burke. *Calmness and exact penalization*, *SIAM J. Control Optim.*, 29: 493–497, 1991.

[4] J. V. Burke and S. Deng. *Weak sharp minima revisited, Part III: Error bounds for differentiable convex inclusions*, Manuscript 2006, to appear in *Mathematical Programming.*

[5] F.H. Clarke. On the inverse function theorem. *Pacific Journal of Mathematics*, 64: 97–102, 1976.

[6] F.H. Clarke. *Optimization and Nonsmooth Analysis.* Wiley, New York, 1983.

[7] R. Cominetti. *Metric regularity, tangent sets and second-order optimality conditions. Applied Mathematics and Optimization*, 21: 265–287, 1990.

[8] S. Dempe. *Foundations of Bilevel Programming.* Kluwer Academic Publ., Dordrecht-Boston-London, 2002.

[9] S. Dolecki and S. Rolewicz. *Exact penalties for local minima*, *SIAM J. Control Optim.*, 17: 596–606, 1979.

[10] A. Dontchev. *Local convergence of the Newton method for generalized equations. Comptes Rendus de l'Acédemie des Sciences de Paris*, 332: 327–331, 1996.

[11] A. Dontchev and R.T. Rockafellar. *Regularity and conditioning of solution mappings in variational analysis. Set-Valued Analysis*, 12: 79–109, 2004.

[12] I. Ekeland. *On the variational principle. Journal of Mathematical Analysis and Applications*, 47: 324–353, 1974.

[13] F. Facchinei and J.-S. Pang. *Finite-Dimensional Variational Inequalities and Complementary Problems, Vol I and Vol II.* Springer, New York, 2003.

[14] P. Fusek. *Isolated zeros of Lipschitzian metrically regular R^n functions. Optimization*, 49: 425–446, 2001.

[15] L.M. Graves. *Some mapping theorems. Duke Mathematical Journal*, 17: 11–114, 1950.

[16] R. Henrion and J. Outrata. *A subdifferential condition for calmness of multifunctions. Journal of Mathematical Analysis and Applications*, 258: 110–130, 2001.

[17] R. Henrion, A. Jourani and J. Outrata. *On the calmness of a class of multifunctions. SIAM Journal Optim.*, 13 (2): 603–618, 2002.

[18] R. Henrion and J. Outrata. *Calmness of constraint systems with applications. Mathematical Programming Ser. B*, 104: 437–464, 2005.

[19] A.D. Ioffe. *Metric regularity and subdifferential calculus. Russ. Mathematical Surveys*, 55: 501–558, 2000.

[20] A. King and R.T. Rockafellar. *Sensitivity analysis for nonsmooth generalized equations. Mathematical Programming*, 55: 341–364, 1992.

[21] D. Klatte and B. Kummer. *Constrained minima and Lipschitzian penalties in metric spaces. SIAM Journal Optim.*, 13 (2): 619–633, 2002.

[22] D. Klatte and B. Kummer. *Nonsmooth Equations in Optimization - Regularity, Calculus, Methods and Applications.* Nonconvex Optimization and Its Applications. Kluwer Academic Publ., Dordrecht-Boston-London, 2002.

[23] D. Klatte and B. Kummer. *Strong Lipschitz Stability of Stationary Solutions for Nonlinear Programs and Variational Inequalities. SIAM J. Optimization*, 16: 96–119, 2005.

[24] D. Klatte and B. Kummer. *Stability of inclusions: Characterizations via suitable Lipschitz functions and algorithms.* to appear in Optimization.

[25] D. Klatte and B. Kummer. *Optimization methods and stability of inclusions in Banach spaces.* to appear in Mathematical Programming, Series B.

[26] B. Kummer. Lipschitzian inverse functions, directional derivatives and application in $C^{1,1}$ optimization. *Journal of Optimization Theory and Applications*, 70:559–580, 1991.

[27] B. Kummer. *Inverse functions of pseudo regular mappings and regularity conditions.* Mathematical Programming, Series B, 88: 313–339, 2000.

[28] W. Li. *Abadie's constraint qualification, metric regularity, and error bounds for differentiable convex inequalities. SIAM Journal on Optimization*, 7: 966-978, 1997.

[29] L. Lyusternik. *Conditional extrema of functions. Mathematical Sbornik*, 41: 390–401, 1934.

[30] B.S. Mordukhovich. Complete characterization of openness, metric regularity and Lipschitzian properties of multifunctions. *Transactions of the American Mathematical Society.* 340: 1–35, 1993.

[31] J. Outrata and M. Kočvara and J. Zowe. *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints.* Kluwer Academic Publ., Dordrecht-Boston-London, 1998.

[32] S.M. Robinson. *Stability theorems for systems of inequalities. Part II: Differentiable nonlinear systems*, *SIAM Journal on Numerical Analysis*, 13: 497–513, 1976.

[33] S.M. Robinson, *Stability theorems for systems of inequalities. Part I: Linear systems*, *SIAM Journal on Numerical Analysis*, 12: 754–769, 1976.

[34] S.M. Robinson. *Generalized equations and their solutions, Part I: Basic theory. Mathematical Programming Study*, 10: 128–141, 1979.

[35] S.M. Robinson. *Strongly regular generalized equations. Mathematics of Operations Research*, 5:43–62, 1980.

[36] S.M. Robinson. *Some continuity properties of polyhedral multifunctions. Mathematical Programming Study.* 14: 206–214, 1981.

[37] S.M. Robinson. *Generalized equations and their solutions. Part II: Applications to nonlinear programming. Mathematical Programming Study*, 19:200–221, 1982.

[38] S.M. Robinson. *Variational conditions with smooth constraints: structure and analysis. Mathematical Programming*, 97: 245–265, 2003.

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