From convergence principles to stability and optimality conditions

Diethard Klatte * Alexander Kruger† Bernd Kummer‡

November 29, 2010

Abstract. We show in a rather general setting that Hoelder and Lipschitz stability properties of solutions to variational problems can be characterized by convergence of more or less abstract iteration schemes. Depending on the principle of convergence, new and intrinsic stability conditions can be derived. Our most abstract models are (multi-) functions on complete metric spaces. The relevance of this approach is illustrated by deriving both classical and new results on existence and optimality conditions, stability of feasible and solution sets and convergence behavior of solution procedures.

Key words. Generalized equations, Hoelder stability, iteration schemes, calmness, Aubin property, variational principles.

Mathematics Subject Classification 2000. 49J53, 49K40, 90C31, 65J05.

1 Introduction

In this paper, we shall throughout suppose that

$$X, P \text{ are metric spaces and } X \text{ is complete.} \tag{1.1}$$

We study local stability properties of solution sets to inclusions

$$p \in F(x) \qquad \text{where } F : X \rightrightarrows P \text{ is closed (i.e., has a closed graph)} \tag{1.2}$$

or, in other words, of the inverse mapping S as

$$S(p) := F^{-1}(p) = \{x \in X \mid p \in F(x)\} \tag{1.3}$$

near some $(\bar{p}, \bar{x}) \in \operatorname{gph} S$.

By *local stability* we mean here that given some $(p, x) \in \operatorname{gph} S$ near $(\bar{p}, \bar{x})$ and some π near $\bar{p}$, there exists a solution $\xi \in S(\pi)$ satisfying $d(\xi, x) \le L d(\pi, p)^q$ $(q > 0)$ for some $L > 0$. Additional requirements to p, x and π will specify the type of stability.

A particular and important special case of (1.3) is given by the *level set mapping*

$$S(p) = S_f(p) := \{x \in X \mid f(x) \le p\} \ \text{ where } f : X \to \mathbb{R}_\infty := \mathbb{R} \cup \{\infty\} \ \text{ is l.s.c.} \tag{1.4}$$

There are many further applications of the model (1.2), (1.3) known, in particular, for standard nonlinear programs, in describing equilibria of games, in several types of bi- or multi-level programs, including MPECs, semi-infinite programs and stochastic models. To see how to link the general model with the special ones, we refer e.g. to [1–3, 5, 11, 20, 26, 29].

In many applications, $F = f$ is a function and $S = f^{-1}$ is its multivalued inverse. But the model (1.2), (1.3) *describes not only classical right-hand side perturbations* of inclusions or equations since $S(p)$ may be defined implicitly. Consider, for instance,

*Address: Institut für Operations Research,Universität Zürich, Moussonstrasse 15, CH-8044 Zürich, Switzerland. E–Mail: klatte@ior.uzh.ch

†Address: Centre for Informatics and Applied Optimization, Graduate School of Information Technology and Mathematical Sciences, University of Ballarat, POB 663, Ballarat, Vic, 3350, Australia. E–Mail: a.kruger@ballarat.edu.au

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

Model 1. Given $\Phi : X \times P_1 \rightrightarrows P_2$ put $P = P_1 \times P_2$ and

$$S(p) = \{x \in X \mid p_2 \in \Phi(x, p_1)\}, \quad F(x) = \{p \mid p_2 \in \Phi(x, p_1)\}, \tag{1.5}$$

with equations if $\Phi = f$ is a function. The mapping Φ can describe fixed points or solutions of (some or many) variational problems which depend on x and p_1; e.g., the (stationary or KKT-) solutions to $\min_y \{h(x, y, p_1) \mid w(x, y) \le p_1\}$, solutions to equilibrium problems or to other MPEC- type problems. More generally, Φ may depend on p_2 or other multifunctions, too.

Model 2. Given $h : X \times P_1 \to P_2$ (a linear normed space) and $C \subset P_2$, the mapping

$$S(p) = \{x \in X \mid p_2 + h(x, p_1) \in C\}, \quad p \in P_1 \times P_2 \tag{1.6}$$

describes *set-constraints* (or solution sets) in parametric optimization models. With any analytical description $c \in C \Leftrightarrow g(x) \le 0$, this leads to usual inequality constraints $G(x, p) := g(p_2 + h(x, p_1)) \le 0$, see section 3.3 for polyhedra C. Further, with $\Phi(x, p) = C - h(x, p_1)$, system (1.6) is (1.5), even if C depends on p_1, too.

The main intention of this paper is to show how basic convergence principles can be used to study the connections between local stability, approximate solutions and iterative solution procedures by a unified approach in the general setting of inclusions in complete metric spaces. In this way, we continue and extend the research presented in [14, 22, 23]. Applications to special cases like level set mappings and approximate minimizers are discussed.

In our general approach, we avoid preparations via Ekeland's variational principle [9]. The latter can be done since we do not aim at using the close relations between stability and injectivity of certain generalized derivatives (which do not hold in general spaces). For approaches studying these relations, we refer the reader e.g. to the monographs [1, 3, 8, 11, 20, 26, 29]. However, we also link different view points and approaches, and do this for several relevant special cases of the abstract model.

Primal space approaches to stability, which avoid the use of generalized derivatives, have been already presented in the first part of Ioffe's work [17]. There, Ekeland's principle is applied in several skillful ways. The message of our paper is that primal space stability conditions can be characterized by certain convergence principles and the same few convergence principles characterize both calmness and the Aubin property in a unified way.

The paper is organized as follows. Section 2 is devoted to some convergence principles which are basic for the rest of the paper. A first illustration how to use them is given by deriving (known) convergence properties of cyclic projection and proximal point methods.

In Section 3, we first introduce and discuss some known notions of local stability, in particular, the Aubin property, Lipschitz l.s.c. and calmness and their Hoelder rate equivalents. Then, as a main result of the paper, we present two versions of a theorem on invariance of the Aubin property under Lipschitz perturbations, including concrete estimates between the solutions of two perturbed mappings. The proofs are based on one of the basic convergence principles of Section 2, the results are closely related to [4, 6, 7, 17, 18, 20].

In order to point out specific features of different local stability properties, we then study standard systems of C^1 equations and inequalities. This complements recent studies via different approaches, given e.g. in [8, 10, 15, 16, 19, 22, 23]. We also show how to include set constraints $h(x) \in C$ with a polyhedral set C in these standard schemes. In the last subsection of Section 3, we discuss various view points about the use of generalized derivatives when deriving optimality and stability criteria in nonsmooth settings. In particular, the case of empty subdifferentials is considered.

Section 4 is devoted to connections between stability properties and descent conditions for functionals. This is in particular applied to characterizations of Hoelder calmness of the level set map of a functional, in the standard calmness case this is related to recent results in [10, 17, 23]. Further, it is shown that the main theorem of this section, Theorem 4.1, is equivalent to Ekeland's principle and also leads to a monotonicity criteria for the Aubin (Hoelder-type) property.

In Section 5, stability for general closed multifunctions $F : X \rightrightarrows P$ is studied. If P is even linear normed, the stability characterizations of Section 4 are applied by utilizing the so-called strong closedness of suitable intersection maps. In contrast, if P is a metric space, we need an approach independent on strongly closedness and Ekeland's principle. It turns out that one of our basic (and simple) convergence principles, presented in Lemma 2.4, leads directly to a characterization of (Hoelder-type) calmness and Aubin property in terms of applicability and well-defined convergence behavior of some proper descent method. This new approach and result will be related to results in [17, 21, 22].

Notation

We write $\mathbb{R}_\infty$ for $\mathbb{R} \cup \{\infty\}$ and use the symbol d for both the metric in X and P if the space under consideration is evident. Throughout, we have $x, x', \xi \in X$, $p, p', \pi \in P$. If F is single-valued $F(x) = \{f(x)\}$ we identify F and f. We say that some property holds *near* $\bar{x}$ if it holds for all x in some neighborhood of $\bar{x}$. By $o = o(t)$ we denote a quantity of the type $o(t)/t \to 0$ if $t \downarrow 0$, and $B(\bar{x}, \varepsilon) = \{x \in X \mid d(x, \bar{x}) \le \varepsilon\}$ denotes the closed ε-ball around $\bar{x}$. For real r, r^+ stands as usual for $\max\{r, 0\}$. We write $\dim X < \infty$ in order to say that X is a finite dimensional space, and *locLip* $(\mathbb{R}^n, \mathbb{R}^m)$ denotes the space of locally Lipschitz functions $f : \mathbb{R}^n \to \mathbb{R}^m$. We write $f \in C^{1,1}$ if (Fréchet-) derivatives exist and are locally Lipschitz. Our hypotheses of differentiability, continuity or closedness have to hold near the reference points only.

2 Some principles of convergence

2.1 Convergence of particular sequences

Below, we will apply the following simple statements on convergence.

Lemma 2.1. *Let $g, h : X \to \mathbb{R}_\infty$, g be l.s.c., and let certain x_k, $k = 1, 2, \ldots$, satisfy*

$$g(x_{k+1}) \le g(x_k) \quad \text{and} \quad g(x_{k+1}) \le \ h(x_k) + \varepsilon_k; \quad \varepsilon_k \downarrow 0.$$

Then for any their accumulation point ξ, it holds $g(\xi) \le \liminf_{k\to\infty} h(x_k)$.

Proof. Obviously, if the sequence x_k has an accumulation point ξ then, by monotonicity, the whole sequence $g(x_k)$ is convergent and $g(\xi) \le \lim_{k\to\infty} g(x_{k+1}) \le \liminf_{k\to\infty} [\ h(x_k) + \varepsilon_k\ ] = \liminf_{k\to\infty} h(x_k)$. □

If h is u.s.c. (in our applications it is going to be globally Lipschitz) then the Lemma yields

$$g(\xi) \le h(\xi).$$

The Lemma is one of many possible variations of the well-known Weierstrass theorem for the existence of a minimum where $h(x) \equiv \inf_X g$ is constant and the existence of ξ is ensured by compactness. Evidently, the particular type of the involved functions is essential and depends on the applications we are aiming at. The number of such applications is big, and they may be quite different.

An important setting appears in the context of Ekeland's principle as follows.

Let $\lambda > 0$, $g : X \to \mathbb{R}_\infty$ and let $g(x_0) \in \mathbb{R}$ for some $x_0 \in X$. Define

$$h(u) = \inf_{x\in X} [\ g(x) + \lambda\ d(x, u)\ ] \qquad u \in X. \tag{2.1}$$

Lemma 2.2. *It holds $h \le g$, and either $h(u)$ is finite for all u or $h(u) = -\infty\ \forall u$. In the first case, h is Lipschitz (with rank λ). Furthermore, h is finite if*

$$c_r := \inf_{x\in B(x_0,r)} g(x) > -\infty\ \forall r > 0 \ \text{ and } \ \liminf_{d(x,x_0)\to\infty} g(x)/d(x, x_0) > -\lambda. \tag{2.2}$$

Proof. The inequality $h \le g$ is obvious. We also have $h(u) \le g(x_0) + \lambda\ d(x_0, u) < \infty$. For any $u_1, u_2, x \in X$ it holds

$$h(u_1) \le g(x) + \lambda d(x, u_1) \le g(x) + \lambda(d(x, u_2) + d(u_1, u_2)).$$

Taking the infimum over $x \in X$ we obtain $h(u_1) \le h(u_2) + \lambda d(u_1, u_2)$. Therefore, $h(u_2)$ is finite if so is $h(u_1)$. Since u_1, u_2 are arbitrary, we derive: All $h(u)$ are finite and h is (globally) Lipschitz with rank λ if $h(u)$ is finite for some u. Next assume that $h(u) = -\infty$. Then there are x_n such that $g(x_n) + \lambda d(x_n, u) < -n$.
Case 1: If $d(x_n, x_0) \le r$ for some $r > 0$ then $\inf_{x\in B(x_0,r)} g(x) = -\infty$.
Case 2: If $d(x_n, x_0) \to \infty$, then we have $g(x_n)/d(x_n, u) + \lambda < -n/d(x_n, u) < 0$ and consequently

$$\liminf_{d(x,x_0)\to\infty} \frac{g(x)}{d(x, x_0)} \le \liminf_{n\to\infty} \frac{g(x_n)}{d(x_n, x_0)} = \liminf_{n\to\infty} \frac{g(x_n)}{d(x_n, u)} \le -\lambda.$$

Both these situations are excluded by (2.2). □

If X is a Banach space, the liminf-condition of (2.2) can be replaced by $\liminf_{\|x\|\to\infty} \frac{g(x)}{\|x\|} > -\lambda$.

Proposition 2.3. *Let $g : X \to \mathbb{R}_\infty$ be l.s.c., $\lambda > 0$, $g(x_0) < \infty$ and suppose (2.2). Then there exist x_k, $k = 1, 2, \ldots$, such that*

$$g(x_k) + \lambda\, d(x_k, x_{k-1}) \;\le\; g(x_{k-1}) \quad (\,\le g(x_0)\,), \tag{2.3}$$

$$g(x_k) \;\le\; h(x_{k-1}) + 1/k. \tag{2.4}$$

For any such sequence, the limit $\xi := \lim x_k$ exists and fulfills

$$\lambda\, d(\xi, x_0) \;\le\; g(x_0) - g(\xi), \tag{2.5}$$

$$g(x) + \lambda d(x,\xi) \;\ge\; g(\xi) \quad \forall x \in X. \tag{2.6}$$

Proof. By Lemma 2.2, h attains only finite values and is globally Lipschitz. Having x_{k-1} for $k > 0$, an appropriate x_k can be found as follows. If $h(x_{k-1}) = g(x_{k-1})$ then take $x_k = x_{k-1}$. In this case, the sequence remains constant and the proof is trivial. If $h(x_{k-1}) < g(x_{k-1})$ then there is some x_k satisfying (2.3) and (2.4) due to definition (2.1). Since $g(x_k) < g(x_{k-1})$ we have $x_k \ne x_0$. Inequality (2.3) yields for any $n > 0$,

$$\lambda d(x_n, x_0) \le \lambda \sum_{k=1}^{n} d(x_k, x_{k-1}) \le \sum_{k=1}^{n} [g(x_{k-1}) - g(x_k)] = g(x_0) - g(x_n) \tag{2.7}$$

and $\lambda \le (g(x_0) - g(x_n)/d(x_n, x_0)$. Assumption (2.2) ensures $\limsup_{d(x,x_0)\to\infty} \frac{g(x_0)-g(x)}{d(x,x_0)} < \lambda$. This tells us that $d(x_n, x_0)$ remains bounded, say $x_n \in B(x_0, r)$. Since $c_r > -\infty$ we conclude that $g(x_0) - g(x_n) \le g(x_0) - c_r < \infty$. Again by (2.7), so also $\sum_{k=1}^{\infty} d(x_k, x_{k-1})$ is bounded. The latter obviously implies that $\{x_k\}$ is a Cauchy sequence. Thus the limit $\xi = \lim x_k$ exists in the complete metric space X. Finally, (2.5) follows from (2.7), while Lemma 2.1 yields $g(\xi) \le h(\xi)$, which is exactly (2.6). ☐

Notice that (2.2) holds true if $\inf_X g$ is finite. Then the existence of ξ is just Ekeland's principle, cf. proposition 4.4. If $\dim X < \infty$, the property $c_r > -\infty$ follows from compactness and lower semi-continuity of g.

The conclusion of Proposition 2.3 is obviously stable with respect to small Lipschitz perturbations of g.

The next lemma provides another simple convergence tool which will be used in the sequel.

Lemma 2.4. *Let $\theta \in (0, 1)$, and $L = (1 - \theta)^{-1}$. Let certain $x_k \in X$, $\tau_k \in \mathbb{R}_+$ satisfy, for $0 \le k \le n$,*

$$d(x_{k+1}, x_k) \le \tau_k \quad \text{and} \quad \tau_{k+1} \le \theta \tau_k. \tag{2.8}$$

Then $x_k \in B(x_0, L\,\tau_0)$ for all $k \le n + 1$. If (2.8) holds for all $k \ge 0$ then the limit $\xi := \lim x_k$ exists and satisfies $\xi \in B(x_0, L\,\tau_0)$.

Proof. It holds for $0 \le k \le n$,

$$\tau_{k+1} \le \theta^{k+1}\,\tau_0, \quad d(x_{k+1}, x_k) \le \theta^k \tau_0, \quad \text{and}$$
$$d(x_{k+1}, x_0) \le \textstyle\sum_{i=0}^{k} d(x_{i+1}, x_i) \le \sum_{i=0}^{k} \theta^i\, \tau_0 \le L\,\tau_0.$$

This proves the first estimate. The claimed convergence follows from the boundedness of the sum $\sum_{i=0}^{k} d(x_{i+1}, x_i) \le L\,\tau_0$ for all k. Hence we obtain a Cauchy-sequence and $\xi = \lim x_k$ exists. ☐

In section 5.2.2 we shall put $\tau_k = d(p_k, \pi)^q$ where p_k, assigned to x_k, and π are specified elements of P and $q > 0$.

2.2 Applications: Convergence via compactness and projections

In this subsection, the function h in Lemma 2.1 is defined by the next iteration point $x' := T(x)$ of some procedure as

$$h(x) := d(\ T(x),\ \bar{x})$$

where $\bar{x}$ is a solution we are interested in. The error constants ε_k are zero. We show how to use Lemma 2.1 in deriving two well-known convergence results.

Cyclic projections

Given m closed convex subsets $\emptyset \neq C^i \subset \mathbb{R}^n$ we consider the problem of finding some $\xi \in D := \cap_i C^i$ where we assume that $D \neq \emptyset$. Let $\bar{x} \in D$ and $\pi_{C^i}(x)$ denote the Euclidean projection of $x \in \mathbb{R}^n$ onto C^i. The functions π_{C^i} are Lipschitz continuous with rank 1 (non-expansive). For any $x \in \mathbb{R}^n$, the elementary properties of projections yield

$$\|\pi_{C^i}(x) - \bar{x}\| \leq \|x - \bar{x}\| \qquad \text{and} \tag{2.9}$$

$$\|\pi_{C^i}(x) - \bar{x}\| = \|x - \bar{x}\| \ \Leftrightarrow\ \pi_{C^i}(x) = x \ \Leftrightarrow\ x \in C^i. \tag{2.10}$$

Let $x^{(m)}$ be the result after a cyclic projection of x, i.e., after applying the m projections as

$$x' := \pi_{C^1}(x), \quad x'' := \pi_{C^2}(x'), \quad \ldots, \quad x^{(m)} := \pi_{C^m}(x^{(m-1)}).$$

$$\text{Put} \qquad T(x) := x^{(m)}, \qquad g(x) := d(x,\ \bar{x}), \qquad h(x) := d(T(x),\ \bar{x}), \qquad x_{k+1} := T(x_k)$$

for any initial point x_0. The latter defines the procedure of *cyclic projections* (also known as *Feijer method*). We verify the known result

Proposition 2.5. *The sequence $\{x_k\}$ converges to some $\xi \in D$.*

Proof. Obviously, g, T (as a composition of projections), and h are continuous. Because of (2.9), it holds

$$h(x_k) = g(x_{k+1}) \leq g(x_k), \tag{2.11}$$

$$\|x_{k+1} - \bar{x}\| \leq \|x_k - \bar{x}\| \leq \ldots \leq \|x_0 - \bar{x}\|. \tag{2.12}$$

Thus the bounded sequence has an accumulation point ξ. Due to (2.11), it follows from Lemma 2.1 that

$$d(\xi, \bar{x}) \leq d(T(\xi), \bar{x}) \leq d(\xi, \bar{x}), \quad \text{hence} \quad d(\xi, \bar{x}) = d(T(\xi), \bar{x}).$$

By (2.9) and (2.10) then ξ remains fixed under all m projections. This ensures $\xi \in D$ for all such accumulation points. Assume there are two of them, ξ^1 and ξ^2. Since our estimates hold with any $\bar{x} \in D$, they hold for $\bar{x} = \xi^1$, too. From (2.12), then $\xi^2 = \xi^1$ follows. □

Proximal Points, Moreau-Yosida approximation

For minimizing a convex function $f : \mathbb{R}^n \to \mathbb{R}$ which has a minimizer, one may consider the so-called Moreau-Yosida approximation $F_y(x) = f(x) + \frac{1}{2}\|x - y\|^2$. Its minimizer $x = x(y)$ is unique since F_y is strongly convex, and is characterized by

$$0 \in \partial F_y(x) \ = \ x - y + \partial f(x). \tag{2.13}$$

Hence, the solutions $\hat{x} \in \operatorname{argmin} f$ are just the fixed points of the function $y \mapsto x(y)$. The proximal point method generates a sequence by setting $x_{k+1} = T(x_k) := \operatorname{argmin} F_{x_k}$ where x_0 is arbitrary.

Proposition 2.6. *If $\operatorname{argmin} f \neq \emptyset$ then the sequence $\{x_k\}$ converges to a minimizer of f.*

Proof. Every x_{k+1} is the unique solution to (2.13) for $y = x_k$. Since ∂f is monotone, it holds for related solutions x and x' corresponding to y and y' respectively:

$$y - x \in \partial f(x), \qquad y' - x' \in \partial f(x'),$$

$$0 \leq \langle y' - x' - (y - x),\ x' - x \rangle = \langle y' - y,\ x' - x \rangle - \|x' - x\|^2 \leq \|y' - y\|\ \|x' - x\| - \|x' - x\|^2.$$

This entails non-expansivity as above, due to

$$\|x' - x\|^2 \leq \|y' - y\|\ \|x' - x\| \ \text{ and } \ \|x' - x\| \leq \|y' - y\|.$$

Discussing here the equation, it should be evident, that convergence follows in the same manner as for the cyclic projections. □

If $\mathbb{R}^n$ is replaced by a Hilbert space, one obtains still weak convergence of $\{x_k\}$ by the same proof.

3 Hoelder type stability

3.1 Stability properties

The following definitions describe, for $q = 1$, typical local Lipschitz properties of the multifunction $S = F^{-1}$ or of level sets for functions $f : X \to \mathbb{R}$, called *Aubin property*, *calmness*, and *Lipschitz lower semi-continuity.* In what follows we will speak about the analogue properties *with exponent* $q > 0$ and add $[q]$ in order to indicate this fact. To avoid the misleading term "Lipschitz lower semi-continuity $[q]$" we write "lower semi-continuity (l.s.c.) $[q]$".

Definition 1. Let $S : P \rightrightarrows X, \quad \bar z = (\bar p, \bar x) \in \operatorname{gph} S$.

(D1) S obeys the *Aubin property* $[q]$ at $\bar z$ if

$$\exists\, \varepsilon, \delta, L > 0 :\; x \in S(p) \cap B(\bar x, \varepsilon) \Rightarrow B(x, Ld(p, \pi)^q) \cap S(\pi) \neq \emptyset \quad \forall p, \pi \in B(\bar p, \delta).$$

(D2) S is *calm* $[q]$ at $\bar z$ if

$$\exists\, \varepsilon, \delta, L > 0 :\; x \in S(p) \cap B(\bar x, \varepsilon) \Rightarrow B(x, Ld(p, \bar p)^q) \cap S(\bar p) \neq \emptyset \quad \forall p \in B(\bar p, \delta). \tag{3.1}$$

(D3) S is *lower semi-continuous* $[q]$ *(l.s.c.* $[q]$ *)* at $\bar z$ if

$$\exists\, \delta, L > 0 : \qquad B(\bar x, Ld(\bar p, \pi)^q) \cap S(\pi) \neq \emptyset \qquad \forall \pi \in B(\bar p, \delta).$$

Conditions (D2) and (D3) correspond to fixing in (D1) $\pi = \bar p$ and $p = \bar p$, respectively. The constant L is called a *rank* of the related stability.

Obviously, these requirements correspond to statements of implicit function type for $F = S^{-1}$ near $(\bar p, \bar x)$ along with an appropriate estimate. If F stands for a sufficiently smooth function f, its derivative plays a crucial role. Next we mention possible problems for $f \notin C^1$.

Example 1. Let $0 < q \le 1$. The locally Lipschitz function

$$f(x) = \begin{cases} x + x^2 \sin(1/x) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0 \end{cases}$$

is differentiable, but Df is discontinuous at 0. Since $Df(0) \neq 0$, $S = S_f$ is both *calm and Lipschitz l.s.c.* at the origin $(0, 0)$ with the given $[q]$. At the same time, f has (positive and negative) local minimizers $x_k \to 0$. Due to $Df(x_k) = 0$ the distances $d_k(\alpha) := \operatorname{dist}(x_k,\ S(f(x_k) - \alpha))$ cannot satisfy a Lipschitz estimate $d_k(\alpha) \le L\alpha^q$ as $\alpha \downarrow 0$. Hence S is not Lipschitz l.s.c. $[q]$ at $(f(x_k), x_k)$ and, in consequence, the Aubin property $[q]$ at the origin is violated, too.

Remark 3.1. Calmness (D2) allows $S(p) = \emptyset$ and can be written without δ and the requirement $p \in B(\bar p, \delta)$ in (3.1). It stands for error estimates near $\bar x$: There are positive ε and L such that

$$\operatorname{dist}(x, S(\bar p)) \le Ld(p, \bar p)^q \qquad \forall x \in S(p) \cap B(\bar x, \varepsilon)\ \forall p \in P.$$

Proof. If (3.1) holds and $p \in P \setminus B(\bar p, \delta)$ then $Ld(p, \bar p)^q \ge L\delta^q$. Since $\operatorname{dist}(x, S(\bar p)) \le d(x, \bar x)$, it follows that $\operatorname{dist}(x, S(\bar p)) \le Ld(p, \bar p)^q \;\; \forall x \in S(p) \cap B(\bar x, \varepsilon')$ if $\varepsilon' \le \min\{\varepsilon, L\delta^q\}$. □

In consequence, (D2) for S_f is equivalent to the *error bound* property:

$$\exists\, \varepsilon, L > 0 :\; x \in B(\bar x, \varepsilon) \;\Rightarrow \operatorname{dist}(x, S_f(\bar p)) \le L((f(x) - f(\bar x))^+)^q.$$

Using our definitions for $q = 1$, other known stability properties can be defined and characterized. We recall some relations which are needed below, for details we refer to [20].

Remark 3.2. Let $q = 1$.

(i) S is called *locally upper Lipschitz* at $\bar z$ if S is calm at $\bar z$ and $\bar x$ is isolated in $S(\bar p)$.

(ii) S is called *strongly stable* at $\bar z$ if S obeys the Aubin property at $\bar z$ and $S(p) \cap B(\bar x, \varepsilon)$ is single-valued for all $p \in B(\bar p, \delta)$.

(iii) S obeys the Aubin property (equivalently: $F = S^{-1}$ is *metrically regular*, S is *pseudo-Lipschitz*) at $\bar z$
⇔ S is calm at all $z \in \operatorname{gph} S$ near $\bar z$ with fixed constants ε, δ, L and Lipschitz l.s.c. at $\bar z$
⇔ S is Lipschitz l.s.c. at all $z \in \operatorname{gph} S$ near $\bar z$ with fixed constants δ and L.

In the strongest case (ii), the mapping S is locally (near $\bar{z}$) a Lipschitz function, and one also says that S is strongly Lipschitz.

(D1) characterizes, for $q = 1$, locally the behavior of A^{-1} for linear, continuous and surjective operators A between Banach spaces as well as the topological behavior of solutions in the inverse function theorem due to Graves and Lyusternik [13, 24].

Necessary stability conditions

Remark 3.3. (D1) implies with $0 < \lambda < L^{-1}$,

$$\begin{array}{c} \text{For all } \ (p,x) \in \operatorname{gph} S \ \cap \ [B(\bar{p},\delta) \times B(\bar{x},\varepsilon)] \ \text{ and } \ \pi \in B(\bar{p},\delta) \setminus \{p\} \\ \text{there is some } (p',x') \in \operatorname{gph} S \ \text{ with } d(p',\pi)^q + \lambda\, d(x',x) < d(p,\pi)^q \end{array} \tag{3.2}$$

since we can choose $(p', x') \in \operatorname{gph} S$ with $p' = \pi$. (D2) and (D3) imply the same for $\pi = \bar{p}$ and $p = \bar{p}$, respectively.

Our paper shows that (3.2) is also sufficient for the related stability if some extra supposition is imposed which is always satisfied if $F = f$ is a continuous function or $\dim P < \infty$.

Our main argument is constructive and quite simple: For initial points (p_0, x_0) near $(\bar{p}, \bar{x})$, we construct a sequence where (p_{k+1}, x_{k+1}) is just some *particular* point (p', x') which exists for $(p, x) = (p_k, x_k)$ by condition (3.2), and we show that the limit exists and fulfills the stability requirements. This direct approach, which needs only some simple statements about convergence of appropriate sequences, has been already used to derive stability characterizations for $q = 1$ in [21, 22] and, for normed spaces P, in [23].

Composed mappings

It is important for many applications that the Aubin property of composed mappings is persistent and can be simplified by differentiation.

Lemma 3.4. *([20], Lemma 2.1) Let $S = S_1 \circ S_2$ be a composed mapping, $S_2 : P \rightrightarrows X_1$, $S_1 : X_1 \rightrightarrows X$. Let $\bar{x} \in S_1(\bar{x}_1)$, $\bar{x}_1 \in S_2(\bar{p})$. Then the Aubin property holds for S at $(\bar{p}, \bar{x})$ if it holds for S_1 at $(\bar{x}_1, \bar{x})$ and S_2 at $(\bar{p}, \bar{x}_1)$.*

Applications:

For Banach spaces P, X, X_1, linear (continuous) operators $F_1 : X \to X_1$, $F_2 : X_1 \to P$ and $F = F_2 \circ F_1$ with the assigned inverse multifunctions S_1, S_2, S, the Aubin property simply means that the images (ranges) satisfy

$$F_2\ (\mathrm{Im}\ F_1) = P \tag{3.3}$$

since $\mathrm{Im}\ F = F_2\ (\mathrm{Im}\ F_1) \subset \mathrm{Im}\ F_2 \subset P$ and we need just $\mathrm{Im}\ F = P$ (by Banach's inverse mapping theorem) for the Aubin property of $S = F^{-1}$. Hence (3.3) is the crucial condition: F_2 has to be surjective and the image of the inner map F_1 must be "sufficiently large" in X_1. Clearly, surjectivity of both operators is sufficient.

If F_1, F_2 are C^1 functions, one may pass to the linearizations $F_{1,\,\mathrm{lin}}$ of F_1 at $\bar{x}$ and $F_{2,\,\mathrm{lin}}$ of F_2 at $\bar{x}_1 = F_1(\bar{x})$ and obtains: $S_{\mathrm{lin}} = [(F_{2,\,\mathrm{lin}} \circ F_{1,\,\mathrm{lin}}]^{-1}$ obeys the Aubin property if and only if (3.3) holds for the linearizations (at the related points), i.e.,

$$DF_2(F_1(\bar{x})) \circ DF_1(\bar{x}) \ \text{ maps } X \text{ onto } \ P. \tag{3.4}$$

In the next section, we see that (3.4) is equivalent to the Aubin property of the original mapping S and that this equivalence can be extended to linearized generalized equations.

Hence, as long as any composed generalized equation $p_i \in f_i(x_i, t_i) + F_i(x_i)$ can be simplified by linearizing involved C^1 functions f_i (w.r. to x_i or both x_i and parameter t_i), the original solution mapping obeys the Aubin property if and only if this holds for the composed linearizations. Of course, checking the latter may be still a hard task. For many applications, however, this leads to systems of linear equations and inequalities with (if the systems reflect optimality conditions) or without (if they stand for usual constraint sets to variational conditions) complementarity conditions.

3.2 Aubin property and small Lipschitzian perturbations

Let P be a normed space, $\delta_h > 0$ and $h : B(\bar x, \delta_h) \subset X \to P$ be a Lipschitz function. Let $\alpha(h)$ be the smallest Lipschitz rank of h on $B(\bar x, \delta_h)$, $\beta(h) = \sup_{x \in B(\bar x, \delta_h)} \|h(x)\|$ and $\|h\|_{C^{0,1}} = \alpha(h) + \beta(h)$.

Next we consider both

$$F : X \rightrightarrows P \ (1.2) \quad \text{and} \quad F_h := h + F : B(\bar x, \delta_h) \subset X \rightrightarrows P \text{ near } (\bar x, \bar p) \in \operatorname{gph} F$$

and show, in particular, invariance of the Aubin property for the inverse mappings S, S_h near the reference point provided that $\|h\|_{C^{0,1}}$ is small enough. Additionally, we estimate solutions $x_i \in S_{h_i}$ for two different functions h_i.

Proposition 3.5. *Let S obey the Aubin property with rank L_S and constants ε_S, δ_S at $(\bar p, \bar x)$. Let $h_i : B(\bar x, \delta_{h_i}) \to P$ $(i = 1, 2)$ be Lipschitz functions with $\alpha := \max\{\alpha(h_i)\} < 1/L_S$. Then there is some $\rho > 0$ such that the following holds under the additional assumptions $p_1, p_2 \in B(\bar p, \rho)$ and $\max\{\beta(h_1), \beta(h_2)\} < \rho$.*

(i) If $x_1 \in B(\bar x, \rho)$, $p_1 \in h_1(x_1) + F(x_1)$ then there is some x_2 with $p_2 \in h_2(x_2) + F(x_2)$ such that

$$d(x_2, x_1) \le \frac{L_S}{1 - \alpha L_S} \ \|(p_2 - p_1) + (h_1(x_1) - h_2(x_1))\|.$$

(ii) If $\frac{L_S}{1-\alpha L_S} \ (\|p_i - \bar p\| + \beta(h_i)) \le \rho$ then $x_i \in B(\bar x, \rho)$ satisfying $p_i \in h_i(x_i) + F(x_i)$ exist.

(iii) If S is strongly stable, x_i under (ii) are unique for possibly smaller positive α and ρ.

We prove first a modified version under the same assumptions on S.

Proposition 3.6. *Let S obey the Aubin property with rank L_S and constants ε_S, δ_S at $(\bar p, \bar x)$. Let $h : B(\bar x, \delta_h) \to P$ be a Lipschitz function with $\alpha := \alpha(h) < 1/L_S$, let $(p_0, x_0) \in \operatorname{gph} S \cap [B(\bar p, \gamma) \times B(\bar x, \gamma)]$ and $\pi \in B(\bar p, \gamma)$. Then there is a solution ξ to $\pi \in h(x) + F(x)$ such that*

$$d(\xi, x_0) \le \frac{L_S}{1 - \alpha L_S} \ \|\pi - p_0 - h(x_0)\| \tag{3.5}$$

provided that both the norm $r := \|\pi - p_0 - h(x_0)\|$ and γ are sufficiently small, namely if

$$\frac{r}{1 - \theta} + \gamma < \delta_S \quad \text{and} \quad \gamma + \frac{L_S \ r}{1 - \theta} \ < \mu \text{ where } \theta = \alpha L_S \quad \text{and} \quad \mu = \min\{\varepsilon_S, \delta_h\}. \tag{3.6}$$

Moreover, ξ belongs to $B(\bar x, \mu)$. If, additionally, S is strongly Lipschitz then ξ is unique for possibly smaller α, γ and r, namely if

$$\|\pi - \bar p - h(\bar x)\| + \alpha\mu < \delta_S \quad \text{and} \quad \|\pi - \bar p - h(\bar x)\| < (1 - \theta)\mu \ L_S^{-1}. \tag{3.7}$$

Proof. It holds $\pi \in h(x) + F(x) \iff x \in \Sigma_\pi(x) := S(\pi - h(x))$. Thus, we are looking for a fixed point of Σ_π. For this purpose, we will construct successively a sequence $x_k \in X$ starting with the given x_0 and the corresponding sequence $p_k := \pi - h(x_{k-1}) \in P$ $(k > 0)$ and satisfying for $k > 0$ the conditions

$$x_k \in \Sigma_\pi(x_{k-1}), \quad d(x_k, x_{k-1}) \le L_S\|p_k - p_{k-1}\|, \quad d(x_k, x_0) \le \frac{L_S \ r}{1 - \theta}, \quad \|p_{k+1} - p_0\| \le \frac{r}{1 - \theta}. \tag{3.8}$$

First notice that if x_k and p_{k+1} satisfy the last two inequalities in (3.8), then, by (3.6), $x_k \in B(\bar x, \mu)$ and $p_{k+1} \in B(\bar x, \delta_S)$.

Case $k = 1$. Obviously $\|p_1 - p_0\| = r$, and consequently $\|p_1 - \bar p\| \le \|p_1 - p_0\| + \|p_0 - \bar p\| \le r + \gamma < \delta_S$. The Aubin property ensures the existence of $x_1 \in S(p_1) = \Sigma_\pi(x_0)$ such that $d(x_1, x_0) \le L_S\|p_1 - p_0\| = L_S r < L_S r/(1 - \theta)$. Hence, $x_1 \in B(\bar x, \mu)$, and consequently, using the Lipschitzness of h, $\|p_2 - p_1\| = \|h(x_1) - h(x_0)\| \le \alpha d(x_1, x_0) \le \theta r$. It follows that $\|p_2 - p_0\| \le (\theta + 1)r < r/(1 - \theta)$. So x_1 and p_2 satisfy (3.8).

Now assume that $n > 0$ and the points satisfying (3.8) have been constructed for all $k \le n$.

Case $k = n + 1$. By the last inequality in (3.8) and case $k = 1$ above, $p_k \in B(\bar x, \delta_S)$ for all $k \le n+1$. Hence, there is again some $x_{n+1} \in S(p_{n+1}) = \Sigma_\pi(x_n)$ with $d(x_{n+1}, x_n) \le L_S\|p_{n+1} - p_n\|$. Since $x_k \in B(\bar x, \mu)$ for all $k \le n$, then, setting $\tau_k = d(x_{k+1}, x_k)$, we have

$$\tau_k \le L_S\|p_{k+1} - p_k\| = L_S\|h(x_k) - h(x_{k-1})\| \le \theta\tau_{k-1},$$

and Lemma 2.4 yields

$$d(x_{n+1}, x_0) \le \frac{\tau_0}{1-\theta} = \frac{d(x_1, x_0)}{1-\theta} \le \frac{L_S\ r}{1-\theta}.$$

It follows that $x_{n+1} \in B(\bar{x}, \mu)$ and $\|p_{n+2} - p_0\| \le \|p_{n+2} - p_1\| + \|p_1 - p_0\| \le \theta r/(1-\theta) + r = r/(1-\theta)$. So x_{n+1} and p_{n+2} satisfy (3.8).

By Lemma 2.4, we obtain a sequence $x_n \to \xi$ such that ξ satisfies (3.5), and consequently $\xi \in B(\bar{x}, \mu)$. Since Σ is closed and $x_{k+1} \in \Sigma_\pi(x_k)$ we conclude that $\xi \in \Sigma_\pi(\xi)$, i.e., $\pi \in h(\xi) + F(\xi)$.

Strong stability: By assumption, the mapping $p \mapsto S(p) \cap B(\bar{x}, \varepsilon_S)$ is single-valued and Lipschitz with modulus L_S on $B(\bar{p}, \delta_S)$. Without loss of generality we suppose that $S(p) = S(p) \cap B(\bar{x}, \varepsilon_S)$ if $p \in B(\bar{p}, \delta_S)$. For $x \in B(\bar{x}, \mu)$ we have $h(x) \in B(h(\bar{x}), \alpha\|x - \bar{x}\|)$, and $p := \pi - h(x)$ fulfills by (3.7),

$$\|p - \bar{p}\| = \|\pi - h(x) - \bar{p}\| \le \|\pi - h(\bar{x}) - \bar{p}\| + \alpha\|x - \bar{x}\| \le \|\pi - \bar{p} - h(\bar{x})\| + \alpha\mu < \delta_S.$$

Hence Σ_π is single-valued and Lipschitz with modulus θ on $B(\bar{x}, \mu)$, and $x \in B(\bar{x}, \mu)$ implies $\Sigma_\pi(x) \in B(\Sigma_\pi(\bar{x}), \theta\|x - \bar{x}\|) \subset B(S(\pi - h(\bar{x})),\ \theta\mu)$. So Σ_π is a self-mapping of $B(\bar{x}, \mu)$ whenever $\|S(\pi - h(\bar{x})) - \bar{x}\| < (1-\theta)\mu$. This is true under (3.7). In consequence, the fixed point $\xi \in B(\bar{x}, \mu)$ of Σ_π is unique. □

Proof. (of Prop. 3.5). Consider $F_i = h_i + F$ with $S_i = F_i^{-1}$ and select any $(p_1, x_1) \in \operatorname{gph} S_1$. Then we have, setting $p_0 = p_1 - h_1(x_0)$,

$$(p_1, x_1) \in \operatorname{gph} S_1 \Leftrightarrow p_1 \in h_1(x_1) + F(x_1) \Leftrightarrow p_0 \in F(x_1) \Leftrightarrow (p_0, x_1) \in \operatorname{gph} S.$$

Thus, if $d(p_0, \bar{p}) < \gamma$, Prop. 3.6 can be applied; now with $x_0 := x_1$, $\pi := p_2$ and $h := h_2$. This yields, under the remaining assumptions: there is a solution ξ $(= x_2)$ to $\pi \in h_2(x) + F(x)$ such that

$$d(\xi, x_0) \le \frac{L_S}{1 - \alpha L_S}\ \|\pi - p_0 - h_2(x_0)\| = \frac{L_S}{1 - \alpha L_S}\ \|(\pi - p_1) + (h_1(x_0) - h_2(x_0))\|.$$

Assumptions (3.6) of Prop. 3.6 are satisfied for small ρ in Prop. 3.5. This ensures (i) of Prop. 3.5. Solvability (ii) follows by applying (i) to $(p_1, x_1) = (\bar{p}, \bar{x}) \in S$ and $h_1 \equiv 0$. Hence some x_2 fulfills $p_2 \in h_2(x_2) + F(x_2)$ and $d(x_2, \bar{x}) \le \frac{L_S}{1-\alpha L_S}\ \|p_2 - \bar{p} - h_2(x_1)\|$. If $\frac{L_S}{1-\alpha L_S}(\|p_2 - \bar{p}\| + \beta(h_2)) \le \rho$ so $x_2 \in B(\bar{x}, \rho)$ follows. After changing the role of h_1 and h_2 this is (ii). Finally, (iii) follows again from local contractivity of Σ_π since, after decreasing α and ρ if necessary, assumptions (3.7) are satisfied for $\pi = p_2$ and $h = h_2$. □

Comments:
With $h_2 = h_1$, Prop. 3.5 yields the Aubin property of S_{h_1}; with $p_1 = p_2$, this is the Aubin property of $h \mapsto S(h) := \{x \mid 0 \in h(x) + F(x)\}$ in view of small Lipschitzian perturbation, measured by $\beta(h_2 - h_1)$, provided that $\alpha := \max\{\alpha(h_1), \alpha(h_2)\} < \frac{1}{L_S}$.

The first proof of the fact that the strong Lipschitz property of S is invariant w.r. to adding small C^1 functions h was given in [27], while [4, 6, 7, 18] present investigations around the invariance of the Aubin property for Lipschitz functions. Some estimates in terms of $\beta(h)$ - less sharp than above, but derived in a more general setting - are included in [20].

The invariance principle is important for Banach spaces X, P.

(a) Let $f \in C^1(X, P)$ and $f_{\text{lin}\ \bar{x}}(x) = f(\bar{x}) + Df(\bar{x})(x - \bar{x})$ be its linearization at $\bar{x}$. It follows that one of the inclusions

$$p \in f(x) + F(x) \quad \text{and} \quad p \in f_{\text{lin}\ \bar{x}}(x) + F(x)$$

obeys the Aubin (or strong Lipschitz) property if so does the other.
Indeed, setting $h = f - f_{\text{lin}\ \bar{x}}$ on $B(\bar{x}, \delta_h)$, the Lipschitz rank $\alpha(h)$ vanishes as $\delta_h \downarrow 0$ [apply the mean-value theorem to $h(x') - h(x)$], while $\beta(h) = o(\delta_h)$ is obvious.

(b) If f is only *strictly differentiable* at $\bar{x}$ (see e.g. [29] for the definition), the arguments of (a) still hold by definition since α and β have the same properties. They also hold for $f \in C^1$ and $f_{\text{lin}\ x_0}$ if $\|x_0 - \bar{x}\|$ is sufficiently small. Solving the linearized generalized equation and replacing, in the next step, x_0 by the solution x_1, one obtains methods of Newton type.

(c) In the same manner, one can study variations of the type $f(x, t) \in F(x)$ where $h = f(\cdot, t) - f(\cdot, \bar{t})$ and $t, \bar{t} \in T$, provided (e.g.) that T is a Banach space and $f \in C^1$. Replacing also $f(\cdot, \bar{t})$ by its linearization at $\bar{x}$, is possible due to (a).

Unfortunately, these propositions fail to hold for calmness (replacing the Aubin property), cf. Example 2 below.

3.3 Particular C^1 systems for $q = 1$

Let X, P be Banach spaces (on $\mathbb{R}$) and $f \in C^1(X, P)$. We suppose $q = 1$.

Theorem 3.7. *Let* $S(p) = \{x \in X \mid g_2(x) = p_2,\ g_1(x) - p_1 \in K\}$, *where* $p = (p_1, p_2) \in P = P_1 \times P_2$, P_1 *and* P_2 *are Banach spaces,* $K \subset P_1$ *is a closed convex cone,* $\operatorname{int} K \neq \emptyset$, $\bar{x} \in S(0)$, $g_i \in C^1(X, P_i)$ $(i = 1, 2)$. *Then, if*

$$Dg_2(\bar{x})X = P_2 \quad \text{and} \quad \exists u \in \ker Dg_2(\bar{x}) : \ g_1(\bar{x}) + Dg_1(\bar{x})u \in \operatorname{int} K \tag{3.9}$$

the Aubin property of S *at* $(0, \bar{x}) \in \operatorname{gph} S$ *is ensured.*

The proof can be based on the Robinson-Ursescu open mapping theorem and observation (a) at the end of section 3.2, cf. [4]. For non-differentiable (multi-) functions g_i and necessity of the suppositions we refer to the intersection theorem 2.22 in [20].

Remark 3.8. Under the assumptions of Theorem 3.7, conditions (3.9) are also necessary for the Aubin property.

Proof. By section 3.2, we may consider the linearized system only. The Aubin property (even the weaker lower Lipschitz property) then yields, using solvability only:

For all p_2, there is some u such that $Dg_2(\bar{x})u = p_2$. Thus $Dg_2(\bar{x})X = P_2$.

For $p_1 \in \operatorname{int} K$ and $p_2 = 0$, there is some u such that $Dg_2(\bar{x})u = 0$ and $k := g_1(\bar{x}) + Dg_1(\bar{x})u - p_1 \in K$. Since K is a convex cone, it follows $g_1(\bar{x}) + Dg_1(\bar{x})u = p_1 + k \in \operatorname{int} K$. □

Lemma 3.9. *If* $S = f^{-1}$ *is locally upper Lipschitz at* $(f(\bar{x}), \bar{x})$ *then* $Df(\bar{x})$ *is injective. If* $\dim X < \infty$, *the reverse is also true.*

Proof. Suppose that $Df(\bar{x})u = 0$ and $u \neq 0$. Then $x(t) := \bar{x} + tu$ fulfills $\|f(x(t)) - f(\bar{x})\| = o(t) \ll d(x(t), \bar{x}))$, i.e., S is not locally upper Lipschitz. Let $\dim X < \infty$. If $Df(\bar{x})$ is not locally upper Lipschitz, there are $x_k \to \bar{x}$ with $\|f(x_k) - f(\bar{x})\| \ll d(x_k, \bar{x})$. Setting now $u_k = (x_k - \bar{x})/\|x_k - \bar{x}\|$, one obtains $Df(\bar{x})u = 0$ for each accumulation point u of $\{u_k\}$. Since $\|u\| = 1$, $Df(\bar{x})$ is not injective. □

In the classical case of $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ and $S = f^{-1}$, all mentioned stability properties ($q = 1$), except for calmness, coincide with $\det Df(\bar{x}) \neq 0$. Calmness is excepted since it may disappear after adding small smooth functions; compare S_f for $f \equiv 0$ and $f(x) = \varepsilon x^2$ or

Example 2. Let $q > 0$. The function $f(x) = \begin{cases} e^{-1/x^2} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$ (known from discussing Taylor's theorem) fulfills $f^{(n)}(0) = 0\ \forall n$, and S_f is not calm [q] at the origin. On the other hand, the level set map S_g for $g \equiv 0$ with the same derivatives is calm [q] everywhere. Hence, even for C^∞ functions and $q = 1$, the derivatives of f at the reference point $\bar{x}$ do not say enough for determining calmness of f^{-1} and the level sets S_f (if $Df(\bar{x}) = 0$). The unpleasant effect comes from a gap of dimensions.

Proposition 3.10. *Let* $f \in C^1(X, \mathbb{R})$, $f(\bar{x}) = 0$, $d = \dim\,[Df(\bar{x})X]$, *and for* $\varepsilon > 0$, $d_\varepsilon = \dim\,[f(B(\bar{x}, \varepsilon)) \cap \mathbb{R}^+]$. *Then,* $S = S_f$ *is calm at* $(0, \bar{x}) \Leftrightarrow \exists\, \varepsilon_0 > 0$ *such that* $d = d_\varepsilon\ \forall \varepsilon \in (0, \varepsilon_0)$.

The condition also means equivalently: $f(B(\bar{x}, \varepsilon)) \cap \mathbb{R}^+ \subset \operatorname{Im} Df(\bar{x})\ \forall \varepsilon \in (0, \varepsilon_0)$ and [$Df(\bar{x}) \neq 0$ or $f(x) \leq 0$ for all x near $\bar{x}$], respectively.

Proof. Notice that d_ε is constant for small $\varepsilon > 0$ and that $f \in C^1$ yields $d \leq d_\varepsilon$.

($\Rightarrow$) Assume, in contrary, $d \neq d_\varepsilon$. Then it holds $d = 0 < d_\varepsilon = 1$ and there are $x_k \to \bar{x}$ such that $f(x_k) > 0$. Using calmness, there are $\xi_k \in S(0)$ such that $\|x_k - \xi_k\| \leq L f(x_k)$ (for large k). Thus also $\xi_k \to \bar{x}$ and $f(\xi_k) \leq 0$ hold true. It follows $(f(x_k) - f(\xi_k))\ \|x_k - \xi_k\|^{-1} \geq L^{-1}$. Additionally, $f(x_k) - f(\xi_k) = Df(\theta_k)(x_k - \xi_k)$ holds with some $\theta_k \in \operatorname{conv}\{x_k, \xi_k\}$. Setting $u_k = (x_k - \xi_k)/\|x_k - \xi_k\|$ and taking $\theta_k \to \bar{x}$ into account, this ensures

$$Df(\theta_k)u_k \geq L^{-1}, \quad \|u_k\| = 1 \quad \text{and} \quad Df(\theta_k) \to Df(\bar{x}) \text{ in } X^*.$$

Recalling $d = 0$ and $Df(\bar{x}) = 0$, also $\|Df(\theta_k)\|_* \to 0$ and $Df(\theta_k)u_k \to 0$ are true. This contradiction to $Df(\theta_k)u_k \geq L^{-1}$ proves the first part.

($\Leftarrow$) If $d = d_\varepsilon = 0$ then $f \leq 0$ holds near $\bar{x}$ and calmness is trivial. If $d = d_\varepsilon = 1$, we obtain $Df(\bar{x}) \neq 0$ which ensures even the Aubin property. □

All introduced stability properties can be exactly characterized for finite dimensional systems of equations and inequalities with RHS perturbations. The knowledge of these characterizations is the key for understanding all generalizations.

Let $S(p)$ be given, with $P = \mathbb{R}^{m_1+m_2}$, $g \in C^1(\mathbb{R}^n, P)$, as

$$S(p) = \{x \in \mathbb{R}^n \mid g_1(x) \le p_1,\ g_2(x) = p_2\} \text{ where } p = (p_1, p_2) \in P. \tag{3.10}$$

These sets have the form as in Theorem 3.7 if K is the closed negative orthant of $\mathbb{R}^{m_1}$. Without loss of generality let $g(\bar{x}) = 0 \in P$ (delete non-active inequality constraints). Then, from classical results in stability analysis, the necessary and sufficient condition (3.9) for the Aubin property coincides with the Mangasarian-Fromowitz constraint qualification, while the linear independence constraint qualification requires stronger $Dg(\bar{x})\mathbb{R}^n = P$.

LICQ for set constraints

Let $C \subset \mathbb{R}^m$ be a convex, polyhedral cone and

$$S(p) = \{x \in \mathbb{R}^n \mid h(x) - p \in C\}, \quad h \in C^1(\mathbb{R}^n, \mathbb{R}^m),\ p \in \mathbb{R}^m. \tag{3.11}$$

For discussing stability we may assume that $\bar{p} = 0$ and $h(\bar{x}) = 0$. If $h(\bar{x}) \ne 0$ or C is a polyhedron, one can replace C by its (contingent-) tangent cone at $h(\bar{x})$. Similarly, additional constraints like $x \in D$ (a polyhedron) can be handled by introducing the function $\hat{h} = h \times id$ where $id(x) = x$.

Formally, the stability theory of (3.11) generalizes the related theory for usual systems (3.10) where C is some orthant. Thus constraints (3.11) are not less general than the "traditional ones". On the other hand,

$$C = \{y \in \mathbb{R}^m \mid Ay \le 0\} \text{ holds with some (not unique) } (\mu, m) \text{ matrix } A. \tag{3.12}$$

Setting

$$G(b) = \{x \in \mathbb{R}^n \mid g(x) := Ah(x) \le b\} \quad \text{and} \quad b = Ap \in \mathbb{R}^\mu, \tag{3.13}$$

we thus obtain

$$x \in S(p) \Leftrightarrow A(h(x) - p) \le 0 \ \Leftrightarrow\ x \in G(b) \text{ with } g = A \circ h \text{ and } b = Ap. \tag{3.14}$$

So S is a particular case of the "traditional mapping" $G = G(b)$.

To see possible differences, note that $\mu > n$ is possible. Then the μ active gradients $Dg_i(\bar{x}) = A_i Dh(\bar{x}) \in \mathbb{R}^n$ are linearly dependent. Hence LICQ (requiring linear independence of the active gradients) is necessarily violated. This was the main justification for studying set constraints in [28] without using "classical" results. However, it was nowhere mentioned that all parameters b of interest belong to the image Im $A \subset \mathbb{R}^\mu$ and that, instead of the formal LICQ with respect to $\mathbb{R}^\mu$, one only needs (for all analytical consequences) that $Dg(\bar{x})$ maps onto the parameter space in question. Hence LICQ for (3.13) becomes

$$(LICQ)_A \qquad \text{Im } A = \text{Im } (ADh(\bar{x})) \text{ or equivalently } \ker(Dh(\bar{x})^T A^T) = \ker A^T.$$

This is exactly the point for applying - as usually - the inverse and implicit function theorems with the parameters $b = Ap$ of interest. Setting $F = F_2 \circ F_1$; $F_2 = A$ and $F_1 = h$, $(LICQ)_A$ is condition (3.4) for $F^{-1}(b) = \{x \mid Ah(x) = b\}$ and the parameter space Im A.

Let C_{ver} be the set of vertexes in C. Then $C_{ver} = ker\ A$, and $(LICQ)_A$ follows immediately (multiply with A) from the *non-degeneracy condition* in [28],

$$(LICQ)_h \qquad \mathbb{R}^m \subset C_{ver} + \text{Im } Dh(\bar{x}). \tag{3.15}$$

Conversely, having $(LICQ)_A$ and any $y \in \mathbb{R}^m$, there is some $u \in \mathbb{R}^n$ such that $Ay = ADh(\bar{x})u$. With $v = Dh(\bar{x})u$, this yields $y - v \in ker\ A = C_{ver}$, $v \in \text{Im } Dh(\bar{x})$ and via $y = (y - v) + v$ also (3.15). Thus $(LICQ)_A \Leftrightarrow (LICQ)_h$. Consequently, $(LICQ)_A$ is invariant with respect to the choice of A and μ in (3.12).

Calmness for C^1 systems

In contrast to the well-known characterization of the Aubin property by MFCQ (which is often hidden in equivalent, but less intrinsic co-derivative conditions), sharp conditions for calmness of S (3.10), have been established only recently. Concerning calmness of S and G (3.14) at $(0, \bar{x})$ one easily shows that both conditions coincide, also without restricting b to Im A, since $G(b) = \emptyset$ is permitted for $b \neq 0$. Writing S as inequality system is now important since it allows a simple description in the propositions 3.11 and 3.12 below. To formulate them we delete the equations in (3.10) (write two inequalities instead). Thus we assume

$$S(p) = \{x \in \mathbb{R}^n \mid g(x) \leq p\}, \quad p \in \mathbb{R}^m, \; g \in C^1(\mathbb{R}^n, \mathbb{R}^m).$$

The next statements from [14,23] and [22], respectively, are still true if x belongs to a Banach space X. Put

$$\phi(x) = \max_i \; g_i(x) \;\text{ and }\; I(x) = \{i \mid g_i(x) = \phi(x)\}.$$

Let $\phi(\bar{x}) = 0$ and Σ be (the possibly empty) family of all index sets $J \subset \{1, ..., m\}$ such that some sequence $x_k \to \bar{x}$ satisfies $\phi(x_k) > 0$ and $I(x_k) \equiv J$. Obviously, $J \subset I(\bar{x})$.

Proposition 3.11 ([14,23]). *Under these assumptions, S is calm at $(0, \bar{x}) \Leftrightarrow$ for all $J \in \Sigma$ there is some $u(J) \in X$ such that $Dg_j(\bar{x})\, u(J) < 0 \;\forall j \in J$.*

In other words, calmness of S means that MFCQ (or the Aubin property) has to hold for all subsystems given by $J \in \Sigma$. An alternative condition can be based on an algorithm for solving $g(x) \leq 0$ which uses the (computable) *relative slack*

$$s_i(x) = (\phi(x) - g_i(x))/\phi(x) \;\text{ if } \phi(x) > 0.$$

ALG0: Let $x_0 \in X$, $\lambda_0 = 1$. For $k \geq 0$, put $x_{k+1} = x_k$ and $\lambda_{k+1} = \lambda_k$ if $\phi(x_k) \leq 0$. Otherwise find some $u \in X$ such that

$$Dg_i(x_k)u \leq \frac{s_i(x_k)}{\lambda_k} - \lambda_k \;\forall i \;\text{ and }\; \|u\| = 1.$$

If a solution exists, put $x_{k+1} = x_k + \lambda_k \phi(x_k)u, \lambda_{k+1} = \lambda_k$, else $x_{k+1} = x_k, \lambda_{k+1} = \frac{1}{2}\lambda_k$.

Proposition 3.12 ([22]). *S is calm at $(0, \bar{x}) \Leftrightarrow$ there are $\varepsilon, \alpha > 0$ such that, for all sequences of ALG0 with $x_0 \in B(\bar{x}, \varepsilon)$, it follows $\lambda_k \geq \alpha \;\forall k$. Then the sequence x_k converges to some $\xi \in S(0)$, and it holds: $\phi(x_{k+1}) \leq (1 - \beta^2)\phi(x_k)$ whenever $0 < \beta < \alpha$ and $x_{k+1} \neq x_k$.*

3.4 Stability and optimality conditions in terms of generalized derivatives

3.4.1 Stability

Let X and P be Banach spaces.

To obtain stability for multifunctions or nonsmooth functions, *generalized derivatives* are widely used in the literature, and there is meanwhile a big collection of such derivatives D^{gen}, see, e.g., [1,3,12,20, 25,26,29]. However, all these generalizations describe a specific behavior of f or F near a reference point $(\bar{x}, \bar{p}) \in gphF$, and it depends on our goals (deriving optimality conditions, some stability, Newton-type solution methods ...) whether the application of a particular derivative D^{gen} makes sense at all. In addition, the tools of computing them are far behind the C^1-calculus. As the main reason, already chain rules for arbitrary Lipschitz functions in finite (appropriate) dimension usually hold - if at all - only in the form of inclusions

$$\text{if } \; h(x) = f(g(x)) \; \text{ then } \; D^{gen}h(x) \subset D^{gen}f(g(x)) \circ D^{gen}g(x) \tag{3.16}$$

with a big gap between both sides. The gap can already occur if $g \in C^1$ and $D^{gen}g = Dg$ (namely if Dg maps into proper subspaces). Similar effects appear for sums, products and for total and partial derivatives as well. Hence even if some injectivity/surjectivity or another property of $D^{gen}h(x)$ is crucial for our goal, the replacement of $D^{gen}h(x)$ by the (often simpler) right-hand side can be questionable.

The exact chain rule (equality in (3.16)) holds for $f \in C^1$ and most of the generalized derivatives D^{gen}. For stability of solutions to optimization problems, this implies that the involved functions have to be C^2. But this is usually violated when one of them is a marginal (or solution) function of a second (lower level) optimization problem, i.e., for multilevel problems [5] where solutions are, in the best case, unique and locally Lipschitz, and the assigned optimal values are only $C^{1,1}$.

3.4.2 Optimality

Insufficient chain rules may have consequences for optimality conditions to $x \in \operatorname{argmin}_X f$ if we try to write them via sums of non-empty subdifferentials as in the convex case. To explain the situation, we suppose

$$X \text{ is a closed subset of } Z, \quad \dim Z < \infty, \quad \bar{x} \in X \quad \text{and} \quad f \in locLip\ (Z, \mathbb{R}). \tag{3.17}$$

With the usual indicator function $i_X : Z \to \{0, \infty\}$ and $h = f + i_X$ then $\operatorname{argmin}_X f = \operatorname{argmin}_Z h$ holds globally and locally. Next consider the obvious local optimality condition

$$h(x) \geq h(\bar{x}) - o(d(x, \bar{x})) \tag{3.18}$$

for some o-type function $o(\cdot)$. It can be used to define a convex subset $\partial^F h(\bar{x}) \subset (Z)^*$, called the *Fréchet subdifferential*, by writing $x^* \in \partial^F h(\bar{x})$ if $h - x^*$ fulfills (3.18). Then we have

$$x^* \in \partial^F h(\bar{x}) \Leftrightarrow 0 \in \partial^F (h - x^*)(\bar{x}) \Leftrightarrow \ h - x^* \text{ fulfills (3.18)}. \tag{3.19}$$

Furthermore (due to finite dimension), the convex *Fréchet normal cone* $N^F_X(\bar{x}) := \partial^F i_X(\bar{x})$ is polar to the generally non-convex *contingent cone*

$$T_X^{cont}\ (\bar{x}) = \{u \mid \exists t_k \downarrow 0, u_k \to u : \bar{x} + t_k u_k \in X\}; \qquad N^F_X(\bar{x}) = [T_X^{cont}(\bar{x})]^*.$$

Passing from f to $h = f + i_X$ implies for the contingent derivative

$$Ch(\bar{x})(u) := \{v \in \mathbb{R}_\infty \mid v = \lim t_k^{-1}(h(\bar{x} + t_k u_k) - h(\bar{x})) \ \text{ where } \ t_k \downarrow 0 \ \text{ and } \ u_k \to u\},$$

that $\infty \in Ch(\bar{x})(u)$ iff $u \in Z \setminus \operatorname{int} T_X^{cont}\ (\bar{x})$ while $\min Ch(\bar{x})(u) < \infty\ \forall u \in T_X^{cont}(\bar{x})$.
In any case, under the assumptions (3.17) the equivalences (3.19) ensure a simple and sharp characterization of $\partial^F h$ and of the optimality condition in terms of the contingent derivative

$$0 \in \partial^F h(\bar{x}) \ \Leftrightarrow \ \min Ch(\bar{x})(u) \geq 0\ \forall u \in Z \ \Leftrightarrow \ h \text{ fulfills (3.18)}. \tag{3.20}$$

Moreover, again by the definitions only, we have a (relatively) simple condition for $\min Ch(\bar{x})(u)$ to be finite: $\infty > r \in Ch(\bar{x})(u) \ \Leftrightarrow$

$$\exists t_k \downarrow 0, u_k \to u : \bar{x} + t_k u_k \in X \ \text{ and } \ r = \lim t_k^{-1}[f(\bar{x} + t_k u_k) - f(\bar{x})]. \tag{3.21}$$

Generally, this says much more than the obvious consequence

$$r \in Cf(\bar{x})(u) + Ci_X(\bar{x})(u), \tag{3.22}$$

where different sequences (t_k, u_k), (t'_k, u'_k) are hidden in the limits assigned to Cf and Ci_X.
If the particular choice of these sequences plays no role, e.g., if directional derivatives $f'(\bar{x}, u)$ exist or if X is polyhedral, then both (3.21) and (3.22) coincide with

$$u \in T_X^{cont}\ (\bar{x}) \ \text{ and } \ f'(\bar{x}, u) = r,$$

and $C(f + i_X)$ in optimality condition (3.20) satisfies additionally the exact chain rule

$$C(f + i_X)(\bar{x})(u) = Cf(\bar{x})(u) + Ci_X(\bar{x})(u). \tag{3.23}$$

Empty and non-empty subdifferentials

The problems begin if we want to have non-empty subdifferentials or want to use the exact chain rule in terms of ∂^F (like above or in convex optimization) as

$$\partial^F (f + i_X)(\bar{x}) = \partial f^F(\bar{x}) + \partial^F i_X(\bar{x})$$

or in inclusion $\subset$ form. The latter (nowhere needed above) may fail while (3.23) holds true.

Example 3. Put $f = \min\{x, 0\}$ and $X = \mathbb{R}^+$ where $0 \in \partial^F (f + i_X)(0)$ and $\partial^F f(0) = \emptyset$.

Thus, in contrast to (3.20), condition

$$0 \in \partial^F f(\bar{x}) + \partial^F i_X(\bar{x}) \tag{3.24}$$

does not necessarily hold for $\bar{x} \in \operatorname{argmin}_X f$.

Remark 3.13. Inclusion (3.24) yields that $u = 0$ solves the *convex* problem $\min\{c(u) \mid u \in C\}$ where $C = \operatorname{conv} T_X^F(\bar{x})$ and $c(u) = \sup\{\langle x^*, u\rangle \mid x^* \in \partial^F f(\bar{x})\}$.

Proof. Indeed, (3.24) says that some $x^* \in \partial^F f(\bar{x}) \cap -N_X^F(\bar{x})$ exists. Because of $C^* = N_X^F(\bar{x})$ and $\partial c(0) = \partial^F f(\bar{x})$ (Minkowski-duality), so $0 \in \partial c(0) + C^*$ and optimality of $u = 0$ follow. Having $\partial^F f(\bar{x}) \neq \emptyset$ the reverse direction holds similarly. □

Since $\partial^F f(\bar{x}) = \emptyset$ is possible and $\partial^F(f+g)(\bar{x}) \subset \partial^F f(\bar{x}) + \partial^F g(\bar{x})$ can be violated, *limiting subdifferentials and limiting normal cones* (via i_X) are often applied:

$$\begin{array}{lll} x^* \in \partial^F_{\lim} f(x) & \text{if } \exists\ (x_k^*, x_k) \to (x^*, x) \text{ such that } x_k^* \in \partial^F f(x_k), & \\ x^* \in N^F_{\lim\ X}(x) & \text{if } \exists\ (x_k^*, x_k) \to (x^*, x) \text{ such that } x_k^* \in N_X^F(x_k), & x_k \in X. \end{array}$$

Then also

$$0 \in \partial^F_{\lim} f(\bar{x}) + N^F_{\lim\ X}(\bar{x}) \tag{3.25}$$

is a frequently used optimality condition. We study it for $f \in C^1$ and polyhedral X.

Example 4. Let $f \in C^1$ and $X = \{x \in \mathbb{R}^2 \mid x_1 x_2 = 0\}$ which is crucial for complementarity problems. Then (3.25) requires at $\bar{x} = 0$: $-Df(0) \in X = N^F_{\lim\ X}(0)$. In other words, (3.25) requires that *one* partial derivative must vanish. With Clarke's [3] normal cone $N_X^c(x)$, one even obtains $N_X^c(0) = \mathbb{R}^2$. So the corresponding necessary optimality condition is satisfied at the origin for any $f \in C^1$.

Notice that $\partial^F f(\bar{x}) = \emptyset$ provides additional information, namely: $\bar{x}$ cannot satisfy the necessary optimality condition for $\min_Z f$ even if we change f by adding any linear function.

Proposition 3.14. *Let $Z = \mathbb{R}^n$. It holds $\partial^F f(\bar{x}) = \emptyset \Leftrightarrow$ there are $n+2$ directions $u_\nu \in \mathbb{R}^n$ such that*

$$\sum_\nu u_\nu = 0 \quad \textit{and} \quad \sum_\nu \min Cf(\bar{x})(u_\nu) = -1. \tag{3.26}$$

Proof. Let $q(u) := \min Cf(\bar{x})(u)$.
($\Leftarrow$) Condition (3.26) implies $0 \notin \partial^F f(\bar{x})$ since $q(u_\nu) < 0$ holds for some ν. Take $x^* \in Z^*$. Considering $\hat{f} := f - x^*$ and using that $C\hat{f}(\bar{x})(u) = Cf(\bar{x})(u) - \langle x^*, u\rangle$, (3.26) also holds for $\hat{f}$. Thus, it holds $0 \notin \partial^F \hat{f}(\bar{x})$ and, equivalently, $x^* \notin \partial^F f(\bar{x})$.
($\Rightarrow$) Let $\partial^F f(\bar{x}) = \emptyset$. This means by (3.19) and (3.20): $\forall x^*\ \exists u$ such that $q(u) - \langle x^*, u\rangle < 0$. Thus the set $H = \{\ x^* \mid \langle x^*, u\rangle \leq q(u)\ \forall u\}$ is empty. Let $Q = \operatorname{epi} q \subset \mathbb{R}^{n+1}$, $Q^c = \operatorname{conv} Q$. Then $0 \in Q^c$. If $0 \notin \operatorname{int} Q^c$, we obtain a contradiction by separation as follows: Some $(x^*, \tau^*) \neq 0$ fulfills $\langle x^*, u\rangle + \tau^* t \leq 0\ \forall (u,t) : t \geq q(u)$. Since $q(u) < \infty\ \forall u$, then $\tau^* \geq 0$ is impossible. Hence $\tau^* < 0$ and, without loss of generality, $\tau^* = -1$. But this yields with $t = q(u)$ that $x^* \in H$, a contradiction. Hence $0_{n+1} \in \operatorname{int} Q^c$. Now $(0_n, -\varepsilon) \in Q^c$ holds for some $\varepsilon > 0$ (the subscript shows the dimension). Using Caratheodory's theorem there are $n+2$ elements $(u_\nu, t_\nu) \in Q \subset \mathbb{R}^{n+1}$ and $\lambda_\nu \geq 0$ such that $\sum \lambda_\nu = 1$ and $\sum \lambda_\nu (u_\nu, t_\nu) = (0, -\varepsilon)$. Setting $u'_\nu = \lambda_\nu u_\nu$, this yields $q(u'_\nu) = \lambda_\nu q(u_\nu) \leq \lambda_\nu t_\nu$ as well as

$$\sum_\nu u'_\nu = 0 \ \text{ and } \ s := \sum_\nu q(u'_\nu) \leq -\varepsilon.$$

Multiplying all u'_ν with $1/|s|$ yields the assertion. □

Since (3.18) implies that $S_h = S_f$ is not Lipschitz l.s.c. at $(f(\bar{x}), \bar{x})$, it follows

$$\begin{array}{rll} \bar{x} \in \operatorname{argmin}_X f & \Rightarrow & 0 \in \partial^F (f + i_X)(\bar{x}) \\ & \Rightarrow & S_f \text{ is not Lipschitz l.s.c. at } (f(\bar{x}), \bar{x}) \\ & \Rightarrow & S_f \text{ violates the Aubin property at } (f(\bar{x}), \bar{x}). \end{array} \tag{3.27}$$

Thus optimality also yields that some stability of the mapping (1.4) is violated at a solution. Any analytical condition for this fact is a necessary optimality condition.

The normal cone

Calmness, which does not appear in (3.27), comes into the play when $N_X^F(\bar{x})$ or $T_X^{cont}(\bar{x})$ must be written in terms of describing functions. For C^1 systems

$$S(p) = \{x \in \mathbb{R}^n \mid g_1(x) \le p_1,\ g_2(x) = p_2\}, \quad g \in C^1(\mathbb{R}^n, \mathbb{R}^{m_1+m_2}) \ \text{ and } \ \bar{x} \in X := S(0),$$

it is well-known that calmness of S at $(0, \bar{x})$ yields for the tangents

$$u \in T_X^{cont}(\bar{x}) \ \Leftrightarrow \ Dg_2(\bar{x})u = 0 \ \text{ and } \ Dg_{1,i}(\bar{x})u \le 0 \ \text{ if } g_{1,i}(\bar{x}) = 0. \tag{3.28}$$

Then the form of $N_X^F(\bar{x}) = T_X^{cont}(\bar{x})^*$ follows from LP-duality. The known Abadie constraint qualification (weaker than calmness) requires $\Leftarrow$ in (3.28). But direction ($\Rightarrow$) is trivial by the mean-value theorem. So Abadie's condition simply requires (3.28) which says equivalently that

$$\begin{array}{c} T_X^{cont}(\bar{x}) \text{ does not change if we replace } g \text{ by} \\ \text{the linearization } \ g_{lin\ \bar{x}} \ \text{ at } \bar{x}. \end{array} \tag{3.29}$$

Hence, calmness remains the weakest proper condition for ensuring (3.28) and (3.29).

4 Approximate minimizers and stable level sets

Above (in section 2.2), the existence of an accumulation point was a consequence of boundedness and finite dimensions, and of $g(T(\xi)) = g(\xi)$ being equivalent to $T(\xi) = \xi$. Now we are going to ensure convergence by using some proper descent condition for functionals.

4.1 Existence and estimates for solutions

The next theorem connects stability with some monotonicity.

Theorem 4.1. *Let $q > 0$, $f : X \to \mathbb{R}_\infty$ be l.s.c., $\bar{x}, x_0 \in X$ and $c < f(x_0) < \infty$. Put $g_c(x) = (f(x) - c)^+$ and suppose that there are positive λ and ε such that*

$$\begin{array}{c} \textit{for all } \ x \in B(\bar{x}, \varepsilon) \ \textit{ with } \ c < f(x) \le f(x_0) \\ \exists x' \ \textit{satisfying} \ \ g_c(x')^q - g_c(x)^q < -\lambda d(x', x). \end{array} \tag{4.1}$$

Additionally, let $d(x_0, \bar{x})$ and $f(x_0) - c$ be small enough, such that

$$d(x_0, \bar{x}) + \lambda^{-1}(f(x_0) - c)^q \ \le \ \varepsilon. \tag{4.2}$$

Then, if $y = x_0$ or, more generally, $y \in X$, $d(y, \bar{x}) \le d(x_0, \bar{x})$ and $c < f(y) \le f(x_0)$, there is some ξ_y satisfying

$$f(\xi_y) \le c \ \ \textit{and} \ \ d(\xi_y, y) \le \ \lambda^{-1} \ [\ f(y) - c\ ]^q.$$

Proof. We consider first $y = x_0$ and apply proposition 2.3 to the function $g = (g_c)^q$. This ensures, for the related sequence and the limit $\xi = \lim x_k$, inequalities (2.5) and (2.6). The first inequality implies $g_c(\xi) \ \le \ g_c(x_0)$ and consequently $f(\xi) \le f(x_0)$. We also obtain from (2.5),

$$\lambda \ d(\xi, x_0) \le \ g_c(x_0)^q = [f(x_0) - c]^q.$$

Using (4.2), we have

$$d(\xi, \bar{x}) \le d(\xi, x_0) + d(x_0, \bar{x}) \le \lambda^{-1} \ (f(x_0) - c)^q + d(x_0, \bar{x}) \le \varepsilon.$$

In consequence, if $f(\xi) > c$ then (4.1) can be applied to ξ but this contradicts (2.6). Hence $f(\xi) \le c$ and the proof is finished for $y = x_0$. The general assertion follows simply from the fact, that the considered points y satisfy all hypotheses imposed on x_0, □

Notice that (theoretically) ξ can be found by the sequence of proposition 2.3 with $g = (g_c)^q$.

4.2 Remarks, corollaries and interpretations

We call (4.1) the *uniform descent condition.*

Remark 4.2. Condition (4.2) is obviously satisfied, if $[f(x_0) - c]^q \le \frac{1}{2}\lambda\varepsilon$ and $x_0 \in B(\bar{x}, \frac{1}{2}\varepsilon)$. In some situations, we have $x_0 = \bar{x}$. Then, again trivially, $[f(x_0) - c\,]^q \le \lambda\varepsilon$ is sufficient.

Consequences

1. Calmness [q]: Let $f(\bar{x}) = c < \infty$. Then (4.1) implies that $S = S_f$ is calm $[q]$ at $(f(\bar{x}), \bar{x})$ with rank $L = \lambda^{-1}$. Conversely, (4.1) is satisfied if S is calm $[q]$ at $(f(\bar{x}), \bar{x})$ by Remark 3.3. Hence, with $c = f(\bar{x})$, (4.1) is a necessary and sufficient calmness [q] - condition. This yields

Corollary 4.3. *Let $q > 0$, $f : X \to \mathbb{R}_\infty$ be l.s.c. and $f(\bar{x}) = 0$. The level set map $S = S_f$ is calm [q] at $(0, \bar{x})$ if and only if, with $g(x) := f(x)^+$, the following condition holds:*

$$\begin{array}{c} \exists \lambda, \delta > 0 \text{ such that } \forall x \in B(\bar{x}, \delta) \text{ with } g(x) > 0 \\ \exists x' \text{ satisfying } g(x')^q - g(x)^q < -\lambda d(x', x). \end{array} \tag{4.3}$$

If $g(x)^q > \lambda d(x, \bar{x})$, the condition is obviously satisfied for $x' = \bar{x}$. Thus, in (4.3), one may additionally require that x fulfills $g(x)^q \le \lambda d(x, \bar{x})$ or $g(x)^q \le \lambda\delta$. In consequence, for $q = 1$, condition (4.3) can be written as

$$\liminf_{x \to \bar{x},\ g(x) > 0} s_1(x) > 0 \quad \text{with } s_1(x) = \sup_{x' \ne x} \frac{g(x) - g(x')}{d(x', x)}, \tag{4.4}$$

where the convention $\inf \emptyset = \infty$ is in use, but equivalently also by the conditions

$$\liminf_{x \to \bar{x},\ g(x) \downarrow 0} s_1(x) > 0, \tag{4.5}$$

$$\liminf_{x \to \bar{x},\ g(x)/d(x, \bar{x}) \downarrow 0} s_1(x) > 0.$$

Condition (4.4) (slightly modified) already appeared in the Basic Lemma of [17] as a sufficient calmness condition, the same for condition (4.5) in [10] where the left-hand side is called *middle uniform strict slope.*

2. Aubin-property [q] at $(f(\bar{x}), \bar{x})$: Suppose that $c < f(\bar{x}) < f(x_0)$ fulfill the estimate (4.2) and that (4.1) holds for all $c' \in (c, f(x_0))$ (with the related function $g_{c'} \le g_c$ and the same ε and λ). Then the Aubin-property [q] follows from Theorem 4.1, and the required condition (4.1) is necessary by Remark 3.3.

3. Ekeland's principle: Let $\bar{x} = x_0$, $c = \inf_X f$, $q = 1$ and, for any $\lambda > 0$,

$$\varepsilon = \lambda^{-1} \ (f(x_0) - \inf_X f). \tag{4.6}$$

Then (4.2) is satisfied.

If (4.1) is violated then there is some $x \in B(x_0, \varepsilon)$ with $c < f(x) \le f(x_0)$ such that, due to $g_c(x') - g_c(x) = f(x') - f(x)$,

$$f(x') - f(x) \ \ge -\lambda d(x', x) \quad \forall x' \in X. \tag{4.7}$$

If (4.1) holds true then $\xi \in B(x_0, \varepsilon)$ minimizes f, and $x = \xi$ fulfills (4.7), too. Thus we obtain, in both cases,

Proposition 4.4. *Ekeland's principle [9]: Let $f : X \to \mathbb{R}_\infty$ be l.s.c. and $\inf_X f$ as well as $f(x_0)$ be finite. Then, for any $\lambda > 0$ and ε given by (4.6), there is some $x \in B(x_0, \varepsilon)$ which fulfills $f(x) \le f(x_0)$ and (4.7).*

Thus Ekeland's principle, often used for showing stability, is equivalent to Theorem 4.1.

4.3 Discussion of the calmness condition.

Let $q = 1$ in this subsection. We already know that the calmness condition (4.3) of Corollary 4.3, with $g(x) = f(x)^+$, and the assigned limit conditions can be modified in several ways: the strict inequality of (4.3) can be replaced by the non-strict one,

$$g(x') - g(x) \le -\lambda d(x', x) \ \text{ and } \ x' \ne x.$$

(as in [17] and [23]) or one considers only (the crucial) points $x \to \bar{x}$ such that $g(x)/d(x,\bar{x}) \downarrow 0$ in the limit conditions. Accordingly, there are several equivalent conditions of the type (4.4).

Notice however, that, for a fixed x, the inequality defining x' in (4.3) is NOT a local condition: it does not require that x' can be chosen arbitrarily close to x. In other words, the obvious inequality

$$s_0(x) := \limsup_{x' \to x,\ x' \neq x} \frac{g(x) - g(x')}{d(x', x)} \ \leq \ s_1(x) = \sup_{x' \neq x} \frac{g(x) - g(x')}{d(x', x)}$$

can be strict. Replacing, in (4.5) or (4.4) $s_1(x)$ by the (possibly smaller) upper limit $s_0(x)$ (the *slope* of g at x – in [17]) one arrives at a sufficient calmness condition (used, e.g., in [19, Theorem 2.1 (e)]), which can be far from necessary. Indeed, consider the points $x_k \downarrow 0$ of example 1 where $s_0(x_k)$ vanishes while $\liminf_{x \to \bar{x},\ \ g(x) > 0} \ s_1(x) = 1$. To obtain necessity, an extra condition of the type

$$s_1(x) - s_0(x) \to 0 \ \text{ as } \ x \to \bar{x},\ g(x) > 0$$

must be imposed. It is satisfied, for instance, if g is convex.

For locally Lipschitz f, the calmness criterion Coroll. 2 of [22] (applied to $g = f^+$) requires, with different λ,

$$\exists \delta, \lambda > 0 : \forall x \in B(\bar{x}, \delta)\ \exists x' \text{ with } g(x') - g(x) \ \leq -\lambda\ d(x', x) \ \text{ and } \ d(x', x) \geq \lambda g(x). \tag{4.8}$$

Hence it has the same form as (4.3) while $d(x', x) \geq \lambda g(x)$ is a consequence of the Lipschitz property. For Banach spaces X, condition (4.8) was used in [22], Theorem 4.

5 Closed multifunctions

Following [20, 21], where this notion has been introduced for Banach space mappings, we call a closed multifunction $F : X \rightrightarrows P$ between metric spaces *strongly closed* if, for each $\pi \in P$, the distance function $f(x) = \operatorname{dist}(\pi, F(x))$ obeys the properties

(P1) If $f(x)$ is finite then the distance is attained at some $p(x) \in F(x)$, and

(P2) f is l.s.c.

These properties are satisfied, for instance, if $\operatorname{gph} F$ is closed and $\dim P < \infty$ or F is single-valued and continuous. In [20], Lemma 2.13, the reader can find other examples, namely: $F(x) = \phi(x) + \Phi(x)$ where ϕ is continuous and Φ is locally compact or $F(x) = \phi(x) + K$ where ϕ is continuous and K is a closed convex subset of a Hilbert space.

In [21], the application of Ekeland's principle to strongly closed mappings was demonstrated, and Theorem 1 therein is our Thm. 5.1 restricted to $q = 1$ and Banach spaces X, P with modified constants. In a similar manner, Ekeland points for strongly closed mappings have been applied in order to characterize the Aubin property in [20], Lemma 2.18.

5.1 P is a linear normed space

We study the closed mappings F (1.2) and $S = F^{-1}$ (1.3) first in the case of a linear normed space P of parameters. Our goal consists in applying Theorem 4.1 and the assigned sequence x_k for stability characterizations. The next theorem is a modified version of the basic Lemma 2.4 in [23].

Theorem 5.1. *Let $q > 0$, $(\bar{p}, \bar{x}) \in P \times X$, $(p_0, x_0) \in \operatorname{gph} S$, $\pi \in P$ and $C = \operatorname{conv}\{p_0, \pi\}$. Suppose there are positive $\varepsilon, \delta, \lambda$ such that*

$$\begin{array}{c} \textit{for all} \quad (p, x) \in \operatorname{gph} S \ \cap \ [\ B(\bar{p}, \delta) \times B(\bar{x}, \varepsilon)\ ] \ \ \textit{with} \ \ p \in C \setminus \{\pi\} \\ \exists (p', x') \in \operatorname{gph} S \ \ \textit{with} \ \ \|p' - \pi\|^q + \lambda\ d(x', x) < \|p - \pi\|^q \quad \textit{and} \ \ p' \in C. \end{array} \tag{5.1}$$

Additionally, let $p_0, \pi \in B(\bar{p}, \ \frac{1}{3}\delta)$ and $d(x_0, \bar{x})$ and $\|p_0 - \pi\|$ be small enough such that

$$d(x_0, \bar{x}) + \lambda^{-1} \|p_0 - \pi\|^q \ \leq \ \varepsilon. \tag{5.2}$$

Then there exists some $\xi \in S(\pi) \cap B(x_0, \ \lambda^{-1}\|p_0 - \pi\|^q)$.

Proof. We put $F_C(x) := F(x) \cap C$, $f(x) = \text{dist}(\,\pi,\ F_C(x)\ )$ and show that Theorem 4.1 can be applied to f. Since C is compact (we shall not explicitly use that $C = \text{conv}\{p_0, \pi\}$, but we need $\pi, p_0 \in C$) and F is closed, it follows that F_C is strongly closed. Because of $(p_0, x_0) \in \text{gph}\, S$ it holds $f(x_0) \le d(\pi, p_0) < \infty$. Let $f(x_0) > 0$ (otherwise we may put $\xi = x_0$) and consider any $x \in B(\bar{x}, \varepsilon)$ with $0 < f(x) \le f(x_0)$. Let $p(x) \in F_C(x)$ realize the distance $f(x)$. Then we have

$$0 < f(x) = \|p(x) - \pi\|, \quad p(x) \in C, \quad (p(x), x) \in \text{gph}\, S.$$

Since $(p_0, x_0) \in \text{gph}\, S$, $p_0, \pi \in B(\bar{p},\ \frac{1}{3}\delta)$ and $p_0, \pi \in C$, it holds

$$\|p(x) - \pi\| \le f(x_0) = \|p(x_0) - \pi\| \le \|p_0 - \pi\| \le \frac{2}{3}\delta,$$

which yields $p(x) \in B(\bar{p}, \delta)$. Hence (5.1) may be applied to $(p(x), x)$ and guarantees the existence of some $(p', x') \in \text{gph}\, S$ with $p' \in C$ such that

$$\|p' - \pi\|^q + \lambda\ d(x', x) < \|p(x) - \pi\|^q.$$

Since $f(x') \le \|\pi - p'\|$ and $f(x) = \|p(x) - \pi\|$ we also obtain

$$f(x')^q - f(x)^q < -\lambda d(x', x).$$

Summarizing, so all hypotheses of Theorem 4.1 are satisfied with $c = 0$ and $g_c = f$. The related point ξ, assigned to $y = x_0$, now satisfies

$$f(\xi) \le 0 \ \text{ and } \ d(\xi, x_0) \le \lambda^{-1}[f(x_0) - c]^q = \lambda^{-1} f(x_0)^q \le \lambda^{-1}\ \|p_0 - \pi\|^q.$$

This yields both $\xi \in S(\pi)$ and the required estimate. □

Remark 5.2. If δ is sufficiently small (compared with ε) such that $\lambda^{-1}\ (\ 2\delta/3\ )^q \le \frac{1}{2}\varepsilon$ then inequality (5.2) holds true whenever $p_0, \pi \in B(\bar{p}, \delta/3)$ and $x_0 \in B(\bar{x}, \frac{1}{2}\varepsilon)$.

Comments:

Let $(\bar{p}, \bar{x}) \in \text{gph}\, S$ in Theorem 5.1. By Remark 3.3, condition (5.1) necessarily holds for π near $\bar{p}$ under the Aubin property [q] of S at $(\bar{p}, \bar{x})$. The same is true for calmness [q] when $\pi = \bar{p}$ is fixed. Conversely, if (5.1) holds for all $(p_0, x_0) \in \text{gph}\, S$ near $(\bar{p}, \bar{x})$ and π near $\bar{p}$, the existence of $\xi \in S(\pi) \cap B(x_0,\ \lambda^{-1}\|p_0 - \pi\|^q)$ implies the Aubin-property [q] at $(\bar{p}, \bar{x})$. If (5.1) holds for all $(p_0, x_0) \in \text{gph}\, S$ near $(\bar{p}, \bar{x})$ *and fixed* $\pi = \bar{p}$, then S is calm [q] at $(\bar{p}, \bar{x})$. Hence, depending on the choice of π, condition (5.1) is necessary and sufficient for calmness [q] and the Aubin-property [q] at $(\bar{p}, \bar{x})$.

Now let $(\bar{p}, \bar{x}) \notin \text{gph}\, S$ and assume that we are interested in solutions to $\bar{p} \in F(x)$. Setting again $\pi = \bar{p}$, Theorem 5.1 says: if $(p_0, x_0) \in \text{gph}\, S$ (e.g., a starting point for some algorithm) is sufficiently close to $(\bar{p}, \bar{x})$ and (5.1) is valid, then a solution ξ to $\bar{p} \in F(x)$ exists in $B(x_0,\ \lambda^{-1}\|p_0 - \bar{p}\|^q)$. Clearly, to satisfy the hypotheses, the distance $d((\bar{p}, \bar{x}), \text{gph}\, S)$ has to be small enough.

5.2 P is a metric space

Concerning C in the proof of Theorem 5.1, we only used that

$$\pi, p_0 \in C \ \text{ and } \ x \mapsto F(x) \cap C \ \text{ is strongly closed.}$$

This tells us that the theorem remains true when P is a general metric space and C is any set of this type. Notice however that, with the simplest setting $C = \{p_0, \pi\}$, the descent condition (5.1) implies $p' = \pi$, and the whole statement becomes trivial. This makes reasonable definitions of C for *metric spaces* difficult unless F itself is strongly closed and we can put $C = P$.

Our setting $C = \text{conv}\{p_0, \pi\}$ for normed P requires the investigation of S on 1-dimensional segments of the parameter space P only and seems, thus, sufficiently reasonable. But, without supposing strong closedness, we need for metric spaces P, an approach, independent on strong closedness and on Ekeland's principle. This will be demonstrated now.

5.2.1 Stability in terms of approximate projections

In this subsection, we suppose that $q = 1$.
The following *approximate projection method* of [22] (onto $\operatorname{gph} S$) characterizes "stability" by linear order of convergence. Define, in $P \times X$, a distance depending on $\lambda > 0$ as

$$d_\lambda((p', x'), (p, x)) = d(p', p) + \lambda d(x', x)$$

$$\text{and} \qquad H_\lambda(p, x) = \operatorname{dist}_\lambda((p, x), \operatorname{gph} S) = \inf_{(p',x') \in \operatorname{gph} S} d_\lambda((p', x'), (p, x)).$$

We assume that $\pi \in P$, $\gamma \geq 0$ and $\lambda > 0$ are fixed.
Procedure S1: Let $(p_0, x_0) \in \operatorname{gph} S$. Given (p_k, x_k), $k \geq 0$ choose any approximate minimizer $(p_{k+1}, x_{k+1}) \in \operatorname{gph} S$ of the distance in the definition of $H_\lambda(\pi, x_k)$ such that

$$d_\lambda((p_{k+1}, x_{k+1}), (\pi, x_k)) \leq H_\lambda(\pi, x_k) + \gamma\lambda\, d(p_k, \pi).$$

Notice that, for any $\gamma > 0$, some next iteration points exist. The case $\gamma = 0$ can be of interest if $\operatorname{gph} S$ is locally compact, particularly, if $\dim X < \infty$.

Theorem 5.3. *[22] Let $\gamma > 0$.*

(i) The Aubin property of S holds at $(\bar p, \bar x)$ $\Leftrightarrow$ there exist $\lambda > 0$ and $\alpha > 0$ such that, for all initial points $(p_0, x_0) \in \operatorname{gph} S \cap (B(\bar p, \alpha) \times B(\bar x, \alpha))$ and $\pi \in B(\bar p, \alpha)$, Procedure S1 generates a sequence (p_k, x_k) satisfying

$$d_\lambda((p_{k+1}, x_{k+1}), (\pi, x_k)) \leq \ \theta\, d(p_k, \pi) \qquad \textit{with some fixed } \theta < 1. \tag{5.3}$$

(ii) The same statement, with fixed $\pi \equiv \bar p$, holds in view of calmness of S at $(\bar p, \bar x)$.

(iii) These statements remain true if we additionally require that P is a linear normed space and $p_{k+1} \in \operatorname{conv}\{p_k, \pi\}$.

Note. Explicitly, (5.3) means $d(p_{k+1}, \pi) \leq \theta\, d(p_k, \pi) - \lambda\, d(x_{k+1}, x_k)$, which implies again convergence $p_k \to \pi$, $x_k \to \xi \in S(\pi)$ and $d(\xi, x_0) \leq \lambda^{-1} d(p_0, \pi)$. Statement (iii) shows a connection to Theorem 5.1.

5.2.2 Calmness [q] via proper descent steps

We study again S (1.3). Let $q, \varepsilon, \delta > 0$, $\lambda \in (0, 1)$, $\pi \in P$, $(\bar p, \bar x) \in P \times X$ and require:

$$\begin{array}{c} \text{For all } (p, x) \in \operatorname{gph} S \ \cap \ [\ B(\bar p, \delta) \times B(\bar x, \varepsilon)\ ], \text{ some } (p', x') \in \operatorname{gph} S \text{ satisfies} \\ (i) \ \ \lambda\, d(x', x) \ \leq \ d(p, \pi)^q \ \text{ and } \ (ii) \ \ d(p', \pi) \ \leq \ (1 - \lambda)\, d(p, \pi). \end{array} \tag{5.4}$$

In consequence, for $q = 1$, multiplying (i) by $\lambda/2$ and adding it with (ii) we obtain

$$d(p', \pi) \ + (\lambda^2/2)\ d(x', x) \leq (1 - \lambda/2)\ d(p, \pi).$$

Thus $d(p', \pi) \ + \beta_1\ d(x', x) \leq \beta_2\ d(p, \pi)$ holds with constants $\beta_1, \beta_2 \in (0, 1)$. This (formally weaker) condition in place of (i) and (ii) has been used to verify calmness and the Aubin property in [17]. There, the proof needs Ekeland's principle whereas the relations between (5.4) and stability are direct and almost trivial (while (5.4) is still necessary, see below). For comparing with Corollary 4.3 and level sets S (1.4), put $\pi = 0, (\bar p, \bar x) = (0, \bar x)$ and $f(\bar x) = 0$. Then condition (5.4) claims

$$\forall x \in B(\bar x, \varepsilon) \text{ with } 0 < f(x) \leq \delta\ \exists x' \text{ with } \lambda d(x', x) \leq f(x)^q \ \text{ and } \ f(x') \leq (1 - \lambda) f(x).$$

Next assume $q > 0$, $(p_0, x_0) \in \operatorname{gph} S$ and consider
Procedure S2: Beginning with $k = 0$, find any $(p_{k+1}, x_{k+1}) \in \operatorname{gph} S$ such that

$$(i)\ \lambda d(x_{k+1}, x_k) \leq d(p_k, \pi)^q \ \text{ and } \ (ii)\ d(p_{k+1}, \pi) \leq (1 - \lambda) d(p_k, \pi). \tag{5.5}$$

If such points can be found for all k then $p_k \to \pi$ holds trivially, and we call S2 applicable.

Lemma 5.4. *Suppose $\lambda \in (0,1)$, $\theta = (1-\lambda)^q$, and (5.5) holds true for some sequence (p_k, x_k), $k \geq 0$ (not necessarily in $\operatorname{gph} S$). Then the limit $\xi = \lim x_k$ exists and satisfies*

$$d(\xi, x_0) \leq L d(p_0, \pi)^q \quad \textit{with} \quad L = [\ \lambda\ (1-\theta)\ ]^{-1}. \tag{5.6}$$

Moreover, if $\varepsilon, \delta > 0$ and $d(x_0, \bar{x})$, $d(\pi, \bar{p})$, and $d(p_0, \pi)$ are small enough such that

$$d(x_0, \bar{x})\ +\ L\ d(p_0, \pi)^q \leq\ \varepsilon \quad \textit{and} \quad d(p_0, \pi) +\ d(\pi, \bar{p}) \leq\ \delta, \tag{5.7}$$

then $x_k \in B(\bar{x}, \varepsilon)$ and $p_k \in B(\bar{p}, \delta)$ hold for all $k \geq 0$.

Proof. With p_k, assigned to x_k, we may put $\tau_k = \lambda^{-1} d(p_k, \pi)^q$ and apply Lemma 2.4. This yields $d(x_k, x_0) \leq (1-\theta)^{-1}\tau_0 = L\ d(p_0, \pi)^q$ and the existence of the limit $\xi = \lim x_k$ satisfying (5.6). If (p_0, x_0) satisfies (5.7), then for any $k \geq 0$ we have

$$\begin{array}{llll} d(x_k,\ \bar{x}) & \leq\ d(x_0, \bar{x})\ +\ d(x_k, x_0) & \leq\ d(x_0, \bar{x})\ + L\ d(p_0, \pi)^q & \leq\ \varepsilon, \\ d(p_k,\ \bar{p}) & \leq\ d(p_k, \pi)\ +\ d(\pi, \bar{p}) & \leq\ d(p_0, \pi)\ + d(\pi, \bar{p}) & \leq\ \delta. \end{array}$$

Hence the lemma is valid. □

Proposition 5.5. *For S defined by (1.3), suppose that $\lambda \in (0,1)$, $\varepsilon, \delta > 0$ and $\pi \in B(\bar{p},\ \delta)$ satisfy (5.4). Then, if $(p_0, x_0) \in \operatorname{gph} S$ and π satisfy (5.7), Procedure S2 is applicable and defines a sequence $\{x_k\}$ converging to some $\xi \in S(\pi)$ satisfying (5.6).*

Proof. By Lemma 5.4, hypothesis (5.4) is applicable to (p_0, x_0) and all generated points (p_k, x_k). Thus all (p_k, x_k) can be chosen in $\operatorname{gph} S$ which ensures $(p_k, x_k) \to (\pi, \xi) \in \operatorname{gph} S$. □

As is all step-size algorithms, one can start with fixed $\lambda_1 = 1$ and put $\lambda_{k+1} := \lambda_k/2$, $x_{k+1} = x_k$ if there is no solution with the current λ. Being applicable now means $\lambda_k \geq \bar{\lambda} > 0$ for all initial points $(p_0, x_0) \in \operatorname{gph} S$ and π satisfying (5.7). Similarly, one could use varying q, beginning with $q_1 = 1$. The estimates then hold with exponent $\bar{q}$ if also $q_k \geq \bar{q} > 0$.

Again, criteria for calmness and the Aubin property with exponent q can be derived in a unified manner.

Corollary 5.6. *Suppose (1.3) and $(\bar{p}, \bar{x}) \in \operatorname{gph} S$. Then*
(i) S obeys the Aubin property $[q]$ at $(\bar{p}, \bar{x})$ $\Leftrightarrow$ there are $\lambda \in (0,1)$ and $\varepsilon, \delta > 0$ such that (5.4) is satisfied for all $\pi \in B(\bar{p},\ \delta)$.
(ii) With fixed $\pi = \bar{p}$, the same holds in view of calmness $[q]$.

Proof. Necessity ($\Rightarrow$) follows easily from the stability definitions while Prop. 5.5 ensures the sufficiency. □

For *q=1 and strongly closed mappings acting between Banach spaces,* this statement is Theorem 3 in [21]. By Prop. 5.5 and Corollary 5.6, we may thus summarize

Theorem 5.7. *Suppose (1.3) and $(\bar{p}, \bar{x}) \in \operatorname{gph} S$. Then*
(i) S obeys the Aubin property $[q]$ at $(\bar{p}, \bar{x})$
$\Longleftrightarrow$ There exist $\lambda \in (0,1)$ and $\varepsilon, \delta > 0$ such that (5.4) is satisfied for all $\pi \in B(\bar{p},\ \delta)$.
$\Longleftrightarrow$ There are $\alpha > 0$ and $\lambda \in (0,1)$ such that iterates (p_{k+1}, x_{k+1}) for procedure S2 exist in each step, whenever the initial points satisfy $d(x_0, \bar{x}) + d(p_0, \bar{p}) + d(\pi, \bar{p}) < \alpha$ and $x_0 \in S(p_0)$.
(ii) With fixed $\pi \equiv \bar{p}$, the same holds in view of calmness $[q]$. □

For $q = 1$ and less general spaces, the equivalence between the stability properties and the related behavior of S2 is known from [21, 22].

As a consequence of the theorem, conditions (5.4) and (5.1), for $C = P$ and $(\bar{p}, \bar{x}) \in \operatorname{gph} S$, are equivalent whenever S (1.3) is strongly closed.

References

[1] Aubin, J.-P., and Ekeland, I. *Applied Nonlinear Analysis.* Pure and Applied Mathematics. John Wiley & Sons Inc., New York, 1984.

[2] Bonnans, J. F., and Shapiro, A. *Perturbation Analysis of Optimization Problems.* Springer Series in Operations Research. Springer-Verlag, New York, 2000.

[3] Clarke, F. H. *Optimization and Nonsmooth Analysis.* Canadian Mathematical Society Series of Monographs and Advanced Texts. John Wiley & Sons Inc., New York, 1983.

[4] Cominetti, R. Metric regularity, tangent sets, and second-order optimality conditions. *Appl. Math. Optim. 21*, 3 (1990), 265–287.

[5] Dempe, S. *Foundations of Bilevel Programming*, vol. 61 of *Nonconvex Optimization and its Applications.* Kluwer Academic Publishers, Dordrecht, 2002.

[6] Dmitruk, A. V., and Kruger, A. Y. Metric regularity and systems of generalized equations. *J. Math. Anal. Appl. 342*, 2 (2008), 864–873.

[7] Dontchev, A. L., and Hager, W. W. An inverse mapping theorem for set-valued maps. *Proc. Amer. Math. Soc. 121*, 2 (1994), 481–489.

[8] Dontchev, A. L., and Rockafellar, R. T. *Implicit Functions and Solution Mappings. A View from Variational Analysis.* Springer Monographs in Mathematics. Springer, Dordrecht, 2009.

[9] Ekeland, I. On the variational principle. *J. Math. Anal. Appl. 47* (1974), 324–353.

[10] Fabian, M., Henrion, R., Kruger, A. Y., and Outrata, J. V. Error bounds: necessary and sufficient conditions. *Set-Valued Var. Anal. 18*, 4 (2010), 121–149.

[11] Facchinei, F., and Pang, J.-S. *Finite-Dimensional Variational Inequalities and Complementarity Problems. Vol. I and II.* Springer Series in Operations Research. Springer-Verlag, New York, 2003.

[12] Frankowska, H. High order inverse function theorems. *Ann. Inst. H. Poincaré Anal. Non Linéaire 6*, suppl. (1989), 283–303.

[13] Graves, L. M. Some mapping theorems. *Duke Math. J. 17* (1950), 111–114.

[14] Heerda, J., and Kummer, B. Characterization of calmness for Banach space mappings. Preprint 06-26, Institut of Mathematics, Humboldt-University Berlin, 2006. Available at http://www.mathematik.hu-berlin.de/publ/pre/2006/p-list-06.html.

[15] Henrion, R., and Outrata, J. V. A subdifferential condition for calmness of multifunctions. *J. Math. Anal. Appl. 258*, 1 (2001), 110–130.

[16] Henrion, R., and Outrata, J. V. Calmness of constraint systems with applications. *Math. Program., Ser. B 104*, 2-3 (2005), 437–464.

[17] Ioffe, A. D. Metric regularity and subdifferential calculus. *Russian Math. Surveys 55* (2000), 501–558.

[18] Ioffe, A. D. On regularity concepts in variational analysis. *J. Fixed Point Theory Appl. Online first* (2010). DOI: 10.1007/s11784-010-0021-0.

[19] Ioffe, A. D., and Outrata, J. V. On metric and calmness qualification conditions in subdifferential calculus. *Set-Valued Anal. 16*, 2–3 (2008), 199–227.

[20] Klatte, D., and Kummer, B. *Nonsmooth Equations in Optimization. Regularity, Calculus, Methods and Applications*, vol. 60 of *Nonconvex Optimization and its Applications.* Kluwer Academic Publishers, Dordrecht, 2002.

[21] Klatte, D., and Kummer, B. Stability of inclusions: characterizations via suitable Lipschitz functions and algorithms. *Optimization 55*, 5-6 (2006), 627–660.

[22] Klatte, D., and Kummer, B. Optimization methods and stability of inclusions in Banach spaces. *Math. Program. 117*, 1-2, Ser. B (2009), 305–330.

[23] Kummer, B. Inclusions in general spaces: Hoelder stability, solution schemes and Ekeland's principle. *J. Math. Anal. Appl. 358*, 2 (2009), 327–344.

[24] Lyusternik, L. A. On conditional extrema of functionals. *Math. Sbornik 41* (1934), 390–401. In Russian.

[25] Minchenko, L. I. Multivalued analysis and differential properties of multivalued mappings and marginal functions. *J. Math. Sci. 116*, 3 (2003), 3266–3302.

[26] Mordukhovich, B. S. *Variational Analysis and Generalized Differentiation. I: Basic Theory*, vol. 330 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2006.

[27] Robinson, S. M. Strongly regular generalized equations. *Math. Oper. Res. 5*, 1 (1980), 43–62.

[28] Robinson, S. M. Variational conditions with smooth constraints: structure and analysis. *Math. Program., Ser. B 97*, 1-2 (2003), 245–265.

[29] Rockafellar, R. T., and Wets, R. J.-B. *Variational Analysis*, vol. 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998.