

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

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

We study local stability properties of solution sets to inclusions

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

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

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

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

By *local stability* we mean here that given some $(p, x) \in \text{gph } S$ near (\bar{p}, \bar{x}) and some π near \bar{p} , there exists a solution $\xi \in S(\pi)$ satisfying $d(\xi, x) \leq Ld(\pi, p)^q$ ($q > 0$) 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*

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

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

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

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

31 with equations if $\Phi = f$ is a function. The mapping Φ can describe fixed points or solutions of (some
 32 or many) variational problems which depend on x and p_1 ; e.g., the (stationary or KKT-) solutions to
 33 $\min_y \{h(x, y, p_1) \mid w(x, y) \leq p_1\}$, solutions to equilibrium problems or to other MPEC- type problems.
 34 More generally, Φ may depend on p_2 or other multifunctions, too.

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

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

36 describes *set-constraints* (or solution sets) in parametric optimization models. With any analytical
 37 description $c \in C \Leftrightarrow g(x) \leq 0$, this leads to usual inequality constraints $G(x, p) := g(p_2 + h(x, p_1)) \leq 0$,
 38 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
 39 depends on p_1 , too.

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

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

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

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

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

64 In order to point out specific features of different local stability properties, we then study standard
 65 systems of C^1 equations and inequalities. This complements recent studies via different approaches,
 66 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
 67 polyhedral set C in these standard schemes. In the last subsection of Section 3, we discuss various
 68 view points about the use of generalized derivatives when deriving optimality and stability criteria in
 69 nonsmooth settings. In particular, the case of empty subdifferentials is considered.

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

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

82

83 **Notation**

84 We write \mathbb{R}_∞ for $\mathbb{R} \cup \{\infty\}$ and use the symbol d for both the metric in X and P if the space
 85 under consideration is evident. Throughout, we have $x, x', \xi \in X$, $p, p', \pi \in P$. If F is single-valued
 86 $F(x) = \{f(x)\}$ we identify F and f . We say that some property holds *near* \bar{x} if it holds for all
 87 x in some neighborhood of \bar{x} . By $o = o(t)$ we denote a quantity of the type $o(t)/t \rightarrow 0$ if $t \downarrow 0$,
 88 and $B(\bar{x}, \varepsilon) = \{x \in X \mid d(x, \bar{x}) \leq \varepsilon\}$ denotes the closed ε -ball around \bar{x} . For real r , r^+ stands as
 89 usual for $\max\{r, 0\}$. We write $\dim X < \infty$ in order to say that X is a finite dimensional space, and
 90 *locLip* ($\mathbb{R}^n, \mathbb{R}^m$) denotes the space of locally Lipschitz functions $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. We write $f \in C^{1,1}$
 91 if (Fréchet-) derivatives exist and are locally Lipschitz. Our hypotheses of differentiability, continuity
 92 or closedness have to hold near the reference points only.

93 **2 Some principles of convergence**94 **2.1 Convergence of particular sequences**

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

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

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

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

97 *Proof.* Obviously, if the sequence x_k has an accumulation point ξ then, by monotonicity, the whole
 98 sequence $g(x_k)$ is convergent and $g(\xi) \leq \lim_{k \rightarrow \infty} g(x_{k+1}) \leq \liminf_{k \rightarrow \infty} [h(x_k) + \varepsilon_k] =$
 99 $\liminf_{k \rightarrow \infty} h(x_k)$. \square

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

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

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

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

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

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

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

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

Proof. The inequality $h \leq g$ is obvious. We also have $h(u) \leq g(x_0) + \lambda d(x_0, u) < \infty$. For any
 108 $u_1, u_2, x \in X$ it holds

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

Taking the infimum over $x \in X$ we obtain $h(u_1) \leq h(u_2) + \lambda d(u_1, u_2)$. Therefore, $h(u_2)$ is finite if
 109 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 $d(x_n, x_0) \leq r$ for some $r > 0$ then $\inf_{x \in B(x_0, r)} g(x) = -\infty$.

Case 2: If $d(x_n, x_0) \rightarrow \infty$, then we have $g(x_n)/d(x_n, x_0) + \lambda < -n/d(x_n, x_0) < 0$ and consequently

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

108 Both these situations are excluded by (2.2). \square

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

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

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

$$g(x_k) \leq h(x_{k-1}) + 1/k. \quad (2.4)$$

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

$$\lambda d(\xi, x_0) \leq g(x_0) - g(\xi), \quad (2.5)$$

$$g(x) + \lambda d(x, \xi) \geq g(\xi) \quad \forall x \in X. \quad (2.6)$$

110 *Proof.* By Lemma 2.2, h attains only finite values and is globally Lipschitz. Having x_{k-1} for $k > 0$,
 111 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,
 112 the sequence remains constant and the proof is trivial. If $h(x_{k-1}) < g(x_{k-1})$ then there is some x_k
 113 satisfying (2.3) and (2.4) due to definition (2.1). Since $g(x_k) < g(x_{k-1})$ we have $x_k \neq x_0$. Inequality
 114 (2.3) yields for any $n > 0$,

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

115 and $\lambda \leq (g(x_0) - g(x_n))/d(x_n, x_0)$. Assumption (2.2) ensures $\limsup_{d(x, x_0) \rightarrow \infty} \frac{g(x_0) - g(x)}{d(x, x_0)} < \lambda$.
 116 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
 117 $g(x_0) - g(x_n) \leq g(x_0) - c_r < \infty$. Again by (2.7), so also $\sum_{k=1}^{\infty} d(x_k, x_{k-1})$ is bounded. The latter
 118 obviously implies that $\{x_k\}$ is a Cauchy sequence. Thus the limit $\xi = \lim x_k$ exists in the complete
 119 metric space X . Finally, (2.5) follows from (2.7), while Lemma 2.1 yields $g(\xi) \leq h(\xi)$, which is exactly
 120 (2.6). \square

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

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

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

127 **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 $0 \leq k \leq n$,*

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

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

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

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

130 This proves the first estimate. The claimed convergence follows from the boundedness of the sum
 131 $\sum_{i=0}^k d(x_{i+1}, x_i) \leq L \tau_0$ for all k . Hence we obtain a Cauchy-sequence and $\xi = \lim x_k$ exists. \square

132 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
 133 P and $q > 0$.

134 **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

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

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

137 **Cyclic projections**

138 Given m closed convex subsets $\emptyset \neq C^i \subset \mathbb{R}^n$ we consider the problem of finding some $\xi \in D := \bigcap_i C^i$
139 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
140 C^i . The functions π_{C^i} are Lipschitz continuous with rank 1 (non-expansive). For any $x \in \mathbb{R}^n$, the
141 elementary properties of projections yield

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

$$142 \quad \|\pi_{C^i}(x) - \bar{x}\| = \|x - \bar{x}\| \Leftrightarrow \pi_{C^i}(x) = x \Leftrightarrow x \in C^i. \quad (2.10)$$

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

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

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

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

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

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

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

$$148 \quad \|x_{k+1} - \bar{x}\| \leq \|x_k - \bar{x}\| \leq \dots \leq \|x_0 - \bar{x}\|. \quad (2.12)$$

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

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

149 By (2.9) and (2.10) then ξ remains fixed under all m projections. This ensures $\xi \in D$ for all such
150 accumulation points. Assume there are two of them, ξ^1 and ξ^2 . Since our estimates hold with any
151 $\bar{x} \in D$, they hold for $\bar{x} = \xi^1$, too. From (2.12), then $\xi^2 = \xi^1$ follows. \square

152 **Proximal Points, Moreau-Yosida approximation**

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

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

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

158 **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:

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

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

This entails non-expansivity as above, due to

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

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

161 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 $f : X \rightarrow \mathbb{R}$, 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 $S : P \rightrightarrows X$, $\bar{z} = (\bar{p}, \bar{x}) \in \text{gph } S$.

(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

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

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

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

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 $0 < q \leq 1$. 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 $x_k \rightarrow 0$. Due to $Df(x_k) = 0$ the distances $d_k(\alpha) := \text{dist}(x_k, S(f(x_k) - \alpha))$ cannot satisfy a Lipschitz estimate $d_k(\alpha) \leq L\alpha^q$ 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

$$\text{dist}(x, S(\bar{p})) \leq Ld(p, \bar{p})^q \quad \forall x \in S(p) \cap B(\bar{x}, \varepsilon) \quad \forall p \in P.$$

Proof. If (3.1) holds and $p \in P \setminus B(\bar{p}, \delta)$ then $Ld(p, \bar{p})^q \geq L\delta^q$. Since $\text{dist}(x, S(\bar{p})) \leq d(x, \bar{x})$, it follows that $\text{dist}(x, S(\bar{p})) \leq Ld(p, \bar{p})^q \quad \forall x \in S(p) \cap B(\bar{x}, \varepsilon')$ if $\varepsilon' \leq \min\{\varepsilon, L\delta^q\}$. \square

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

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

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}

$\Leftrightarrow S$ is calm at all $z \in \text{gph } S$ near \bar{z} with fixed constants ε, δ, L and Lipschitz l.s.c. at \bar{z}

$\Leftrightarrow S$ is Lipschitz l.s.c. at all $z \in \text{gph } S$ near \bar{z} with fixed constants δ and L .

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

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

198 **Necessary stability conditions**

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

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

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

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

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

210 **Composed mappings**

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

213 **Lemma 3.4.** ([20], Lemma 2.1) *Let $S = S_1 \circ S_2$ be a composed mapping, $S_2 : P \rightrightarrows X_1$, $S_1 : X_1 \rightrightarrows X$.
 214 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})
 215 and S_2 at (\bar{p}, \bar{x}_1) .*

216 Applications:

217 For Banach spaces P, X, X_1 , linear (continuous) operators $F_1 : X \rightarrow X_1$, $F_2 : X_1 \rightarrow P$ and $F =$
 218 $F_2 \circ F_1$ with the assigned inverse multifunctions S_1, S_2, S , the Aubin property simply means that
 219 the images (ranges) satisfy

$$F_2 (\text{Im } F_1) = P \quad (3.3)$$

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

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

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

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

229 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
 230 linearizing involved C^1 functions f_i (w.r. to x_i or both x_i and parameter t_i), the original solution
 231 mapping obeys the Aubin property if and only if this holds for the composed linearizations. Of course,
 232 checking the latter may be still a hard task. For many applications, however, this leads to systems
 233 of linear equations and inequalities with (if the systems reflect optimality conditions) or without (if
 234 they stand for usual constraint sets to variational conditions) complementarity conditions.

235 3.2 Aubin property and small Lipschitzian perturbations

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

238 Next we consider both

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

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

242 **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*
243 *$h_i : B(\bar{x}, \delta_{h_i}) \rightarrow P$ ($i = 1, 2$) be Lipschitz functions with $\alpha := \max\{\alpha(h_i)\} < 1/L_S$. Then there*
244 *is some $\rho > 0$ such that the following holds under the additional assumptions $p_1, p_2 \in B(\bar{p}, \rho)$ and*
245 *$\max\{\beta(h_1), \beta(h_2)\} < \rho$.*

(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*

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

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

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

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

249 **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 $h :$
250 $B(\bar{x}, \delta_h) \rightarrow P$ be a Lipschitz function with $\alpha := \alpha(h) < 1/L_S$, let $(p_0, x_0) \in \text{gph } S \cap [B(\bar{p}, \gamma) \times B(\bar{x}, \gamma)]$
251 and $\pi \in B(\bar{p}, \gamma)$. Then there is a solution ξ to $\pi \in h(x) + F(x)$ such that*

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

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

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

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

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

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

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

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

257 Case $k = 1$. Obviously $\|p_1 - p_0\| = r$, and consequently $\|p_1 - \bar{p}\| \leq \|p_1 - p_0\| + \|p_0 - \bar{p}\| \leq r + \gamma < \delta_S$.
258 The Aubin property ensures the existence of $x_1 \in S(p_1) = \Sigma_\pi(x_0)$ such that $d(x_1, x_0) \leq L_S \|p_1 -$
259 $p_0\| = L_S r < L_S r / (1 - \theta)$. Hence, $x_1 \in B(\bar{x}, \mu)$, and consequently, using the Lipschitzness of h ,
260 $\|p_2 - p_1\| = \|h(x_1) - h(x_0)\| \leq \alpha d(x_1, x_0) \leq \theta r$. It follows that $\|p_2 - p_0\| \leq (\theta + 1)r < r / (1 - \theta)$. So
261 x_1 and p_2 satisfy (3.8).

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

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 $k \leq n + 1$.
Hence, there is again some $x_{n+1} \in S(p_{n+1}) = \Sigma_\pi(x_n)$ with $d(x_{n+1}, x_n) \leq L_S \|p_{n+1} - p_n\|$. Since
 $x_k \in B(\bar{x}, \mu)$ for all $k \leq n$, then, setting $\tau_k = d(x_{k+1}, x_k)$, we have

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

and Lemma 2.4 yields

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

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

265 By Lemma 2.4, we obtain a sequence $x_n \rightarrow \xi$ such that ξ satisfies (3.5), and consequently $\xi \in$
 266 $B(\bar{x}, \mu)$. 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),

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

267 Hence Σ_π is single-valued and Lipschitz with modulus θ on $B(\bar{x}, \mu)$, and $x \in B(\bar{x}, \mu)$ implies $\Sigma_\pi(x) \in$
 268 $B(\Sigma_\pi(\bar{x}), \theta\|x - \bar{x}\|) \subset B(S(\pi - h(\bar{x})), \theta\mu)$. So Σ_π is a self-mapping of $B(\bar{x}, \mu)$ whenever $\|S(\pi -$
 269 $h(\bar{x})) - \bar{x}\| < (1-\theta)\mu$. This is true under (3.7). In consequence, the fixed point $\xi \in B(\bar{x}, \mu)$ of Σ_π is
 270 unique. \square

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

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

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 $\xi (= x_2)$ to $\pi \in h_2(x) + F(x)$ such that

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

271 Assumptions (3.6) of Prop. 3.6 are satisfied for small ρ in Prop. 3.5. This ensures (i) of Prop. 3.5.
 272 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
 273 $p_2 \in h_2(x_2) + F(x_2)$ and $d(x_2, \bar{x}) \leq \frac{L_S}{1-\alpha L_S} \|p_2 - \bar{p} - h_2(x_1)\|$. If $\frac{L_S}{1-\alpha L_S} (\|p_2 - \bar{p}\| + \beta(h_2)) \leq \rho$ so
 274 $x_2 \in B(\bar{x}, \rho)$ follows. After changing the role of h_1 and h_2 this is (ii). Finally, (iii) follows again from
 275 local contractivity of Σ_π since, after decreasing α and ρ if necessary, assumptions (3.7) are satisfied
 276 for $\pi = p_2$ and $h = h_2$. \square

277 **Comments:**

278 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
 279 $h \mapsto S(h) := \{x \mid 0 \in h(x) + F(x)\}$ in view of small Lipschitzian perturbation, measured by $\beta(h_2 - h_1)$,
 280 provided that $\alpha := \max\{\alpha(h_1), \alpha(h_2)\} < \frac{1}{L_S}$.

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

285 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

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

286 obeys the Aubin (or strong Lipschitz) property if so does the other.

287 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
 288 mean-value theorem to $h(x') - h(x)$], while $\beta(h) = o(\delta_h)$ is obvious.

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

293 (c) In the same manner, one can study variations of the type $f(x, t) \in F(x)$ where $h = f(\cdot, t) - f(\cdot, \bar{t})$
 294 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
 295 linearization at \bar{x} , is possible due to (a).

296 Unfortunately, these propositions fail to hold for calmness (replacing the Aubin property), cf. Exam-
 297 ple 2 below.

298 3.3 Particular C^1 systems for $q = 1$

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

300 **Theorem 3.7.** *Let $S(p) = \{x \in X \mid g_2(x) = p_2, g_1(x) - p_1 \in K\}$, where $p = (p_1, p_2) \in P =$
301 $P_1 \times P_2$, P_1 and P_2 are Banach spaces, $K \subset P_1$ is a closed convex cone, $\text{int } K \neq \emptyset$, $\bar{x} \in S(0)$,
302 $g_i \in C^1(X, P_i)$ ($i = 1, 2$). Then, if*

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

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

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

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

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

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

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

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

316 *Proof.* Suppose that $Df(\bar{x})u = 0$ and $u \neq 0$. Then $x(t) := \bar{x} + tu$ fulfills $\|f(x(t)) - f(\bar{x})\| = o(t) \ll$
317 $d(x(t), \bar{x})$, i.e., S is not locally upper Lipschitz. Let $\dim X < \infty$. If $Df(\bar{x})$ is not locally upper
318 Lipschitz, there are $x_k \rightarrow \bar{x}$ 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
319 obtains $Df(\bar{x})u = 0$ for each accumulation point u of $\{u_k\}$. Since $\|u\| = 1$, $Df(\bar{x})$ is not injective. \square

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

323 *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

324 theorem) fulfills $f^{(n)}(0) = 0 \forall n$, and S_f is not calm $[q]$ at the origin. On the other hand, the level
325 set map S_g for $g \equiv 0$ with the same derivatives is calm $[q]$ everywhere. Hence, even for C^∞ functions
326 and $q = 1$, the derivatives of f at the reference point \bar{x} do not say enough for determining calmness
327 of f^{-1} and the level sets S_f (if $Df(\bar{x}) = 0$). The unpleasant effect comes from a gap of dimensions.

328 **Proposition 3.10.** *Let $f \in C^1(X, \mathbb{R})$, $f(\bar{x}) = 0$, $d = \dim [Df(\bar{x})X]$, and for $\varepsilon > 0$, $d_\varepsilon = \dim [f(B(\bar{x}, \varepsilon)) \cap$
329 $\mathbb{R}^+]$. Then, $S = S_f$ is calm at $(0, \bar{x}) \Leftrightarrow \exists \varepsilon_0 > 0$ such that $d = d_\varepsilon \forall \varepsilon \in (0, \varepsilon_0)$.*

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

332 *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 $x_k \rightarrow \bar{x}$ such that
 $f(x_k) > 0$. Using calmness, there are $\xi_k \in S(0)$ such that $\|x_k - \xi_k\| \leq Lf(x_k)$ (for large k). Thus
also $\xi_k \rightarrow \bar{x}$ 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 $\theta_k \in \text{conv}\{x_k, \xi_k\}$. Setting $u_k = (x_k - \xi_k)/\|x_k - \xi_k\|$
and taking $\theta_k \rightarrow \bar{x}$ into account, this ensures

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

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

335 (\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 $Df(\bar{x}) \neq 0$
336 which ensures even the Aubin property. \square

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

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

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

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

346 LICQ for set constraints

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

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

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

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

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

354 Setting

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

355 we thus obtain

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

356 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 $\text{Im } A \subset \mathbb{R}^\mu$ and that, instead of the formal LICQ with respect to \mathbb{R}^μ , one only needs (for all analytical consequences) that $Dg(\bar{x})$ maps onto the parameter space in question. Hence LICQ for (3.13) becomes

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

357 This is exactly the point for applying - as usually - the inverse and implicit function theorems with
 358 the parameters $b = Ap$ of interest. Setting $F = F_2 \circ F_1$; $F_2 = A$ and $F_1 = h$, $(LICQ)_A$ is condition
 359 (3.4) for $F^{-1}(b) = \{x \mid Ah(x) = b\}$ and the parameter space $\text{Im } A$.

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

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

362 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
 363 $v = Dh(\bar{x})u$, 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
 364 $(LICQ)_A \Leftrightarrow (LICQ)_h$. Consequently, $(LICQ)_A$ is invariant with respect to the choice of A and μ in
 365 (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 $\text{Im } A$, 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

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

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

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

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

368 **Proposition 3.11** ([14, 23]). *Under these assumptions, S is calm at $(0, \bar{x}) \Leftrightarrow$ for all $J \in \Sigma$ there
369 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*

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

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

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

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

371 **Proposition 3.12** ([22]). *S is calm at $(0, \bar{x}) \Leftrightarrow$ there are $\varepsilon, \alpha > 0$ such that, for all sequences of
372 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 $\xi \in S(0)$,
373 and it holds: $\phi(x_{k+1}) \leq (1 - \beta^2)\phi(x_k)$ whenever $0 < \beta < \alpha$ and $x_{k+1} \neq x_k$.*

374 3.4 Stability and optimality conditions in terms of generalized derivatives

375 3.4.1 Stability

376 Let X and P be Banach spaces.

377 To obtain stability for multifunctions or nonsmooth functions, *generalized derivatives* are widely used
378 in the literature, and there is meanwhile a big collection of such derivatives D^{gen} , see, e.g., [1, 3, 12, 20,
379 25, 26, 29]. However, all these generalizations describe a specific behavior of f or F near a reference
380 point $(\bar{x}, \bar{p}) \in gphF$, and it depends on our goals (deriving optimality conditions, some stability,
381 Newton-type solution methods ...) whether the application of a particular derivative D^{gen} makes
382 sense at all. In addition, the tools of computing them are far behind the C^1 -calculus. As the main
383 reason, already chain rules for arbitrary Lipschitz functions in finite (appropriate) dimension usually
384 hold - if at all - only in the form of inclusions

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

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

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

395 3.4.2 Optimality

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

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

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

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

401 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
 402 *Fréchet subdifferential*, by writing $x^* \in \partial^F h(\bar{x})$ if $h - x^*$ fulfills (3.18). Then we have

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

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

$$T_X^{\operatorname{cont}}(\bar{x}) = \{u \mid \exists t_k \downarrow 0, u_k \rightarrow u : \bar{x} + t_k u_k \in X\}; \quad N_X^F(\bar{x}) = [T_X^{\operatorname{cont}}(\bar{x})]^*.$$

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

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

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

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

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

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

408 Generally, this says much more than the obvious consequence

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

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

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

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

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

410 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

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

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

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

413 Thus, in contrast to (3.20), condition

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

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

415 *Remark 3.13.* Inclusion (3.24) yields that $u = 0$ solves the *convex* problem $\min\{c(u) \mid u \in C\}$ where
416 $C = \operatorname{conv} T_X^F(\bar{x})$ and $c(u) = \sup\{\langle x^*, u \rangle \mid x^* \in \partial^F f(\bar{x})\}$.

417 *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
418 $\partial c(0) = \partial^F f(\bar{x})$ (Minkowski-duality), so $0 \in \partial c(0) + C^*$ and optimality of $u = 0$ follow. Having
419 $\partial^F f(\bar{x}) \neq \emptyset$ the reverse direction holds similarly. \square

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:

$$\begin{aligned} x^* \in \partial_{\lim}^F f(x) & \quad \text{if } \exists (x_k^*, x_k) \rightarrow (x^*, x) \text{ such that } x_k^* \in \partial^F f(x_k), \\ x^* \in N_{\lim}^F X(x) & \quad \text{if } \exists (x_k^*, x_k) \rightarrow (x^*, x) \text{ such that } x_k^* \in N_X^F(x_k), \quad x_k \in X. \end{aligned}$$

420 Then also

$$0 \in \partial_{\lim}^F f(\bar{x}) + N_{\lim}^F X(\bar{x}) \quad (3.25)$$

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

422 *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.
423 Then (3.25) requires at $\bar{x} = 0$: $-Df(0) \in X = N_{\lim}^F X(0)$. In other words, (3.25) requires that *one*
424 partial derivative must vanish. With Clarke's [3] normal cone $N_X^c(x)$, one even obtains $N_X^c(0) = \mathbb{R}^2$.
425 So the corresponding necessary optimality condition is satisfied at the origin for any $f \in C^1$.

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

429 **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
430 that*

$$\sum_{\nu} u_\nu = 0 \quad \text{and} \quad \sum_{\nu} \min C f(\bar{x})(u_\nu) = -1. \quad (3.26)$$

Proof. Let $q(u) := \min C f(\bar{x})(u)$.

(\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
 $\hat{f} := f - \langle x^*, \cdot \rangle$ and using that $C \hat{f}(\bar{x})(u) = C f(\bar{x})(u) - \langle x^*, u \rangle$, (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

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

431 Multiplying all u'_ν with $1/|s|$ yields the assertion. \square

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

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

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

435 **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

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

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

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

437 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

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

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

441 **4 Approximate minimizers and stable level sets**

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

445 **4.1 Existence and estimates for solutions**

446 The next theorem connects stability with some monotonicity.

447 **Theorem 4.1.** *Let $q > 0$, $f : X \rightarrow \mathbb{R}_\infty$ be l.s.c., $\bar{x}, x_0 \in X$ and $c < f(x_0) < \infty$. Put $g_c(x) =$
448 $(f(x) - c)^+$ and suppose that there are positive λ and ε such that*

$$\text{for all } x \in B(\bar{x}, \varepsilon) \text{ with } c < f(x) \leq f(x_0) \\ \exists x' \text{ satisfying } g_c(x')^q - g_c(x)^q < -\lambda d(x', x). \quad (4.1)$$

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

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

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

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

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 $g_c(\xi) \leq g_c(x_0)$ and consequently $f(\xi) \leq f(x_0)$. We also obtain from (2.5),

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

Using (4.2), we have

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

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

453 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 $[f(x_0) - c]^q \leq \frac{1}{2}\lambda\varepsilon$ and $x_0 \in B(\bar{x}, \frac{1}{2}\varepsilon)$. In some situations, we have $x_0 = \bar{x}$. Then, again trivially, $[f(x_0) - c]^q \leq \lambda\varepsilon$ 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 $q > 0$, $f : X \rightarrow \mathbb{R}_\infty$ 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:*

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

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 $g(x)^q \leq \lambda d(x, \bar{x})$ or $g(x)^q \leq \lambda \delta$. In consequence, for $q = 1$, condition (4.3) can be written as

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

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

$$\begin{aligned} \liminf_{x \rightarrow \bar{x}, g(x) \downarrow 0} s_1(x) > 0, \\ \liminf_{x \rightarrow \bar{x}, g(x)/d(x, \bar{x}) \downarrow 0} s_1(x) > 0. \end{aligned} \quad (4.5)$$

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 $g_{c'} \leq g_c$ 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 $\lambda > 0$,

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

Then (4.2) is satisfied.

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

$$f(x') - f(x) \geq -\lambda d(x', x) \quad \forall x' \in X. \quad (4.7)$$

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 $f : X \rightarrow \mathbb{R}_\infty$ 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 $f(x) \leq f(x_0)$ 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,

$$g(x') - g(x) \leq -\lambda d(x', x) \quad \text{and } x' \neq x.$$

485 (as in [17] and [23]) or one considers only (the crucial) points $x \rightarrow \bar{x}$ such that $g(x)/d(x, \bar{x}) \downarrow 0$ in the
 486 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

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

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 $\liminf_{x \rightarrow \bar{x}, g(x) > 0} s_1(x) = 1$. To obtain necessity, an extra condition of the type

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

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

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

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

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

492 5 Closed multifunctions

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

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

497 (P2) f is l.s.c.

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

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

506 5.1 P is a linear normed space

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

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

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

512 *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*

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

513 *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 $f(x_0) \leq d(\pi, p_0) < \infty$. Let $f(x_0) > 0$ (otherwise we may put $\xi = x_0$) and consider any $x \in B(\bar{x}, \varepsilon)$ with $0 < f(x) \leq f(x_0)$. 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

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

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 $f(x') \leq \|\pi - p'\|$ 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

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

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

515 *Remark 5.2.* If δ is sufficiently small (compared with ε) such that $\lambda^{-1}(2\delta/3)^q \leq \frac{1}{2}\varepsilon$ then inequality
516 (5.2) holds true whenever $p_0, \pi \in B(\bar{p}, \delta/3)$ and $x_0 \in B(\bar{x}, \frac{1}{2}\varepsilon)$.

517 *Comments:*

518 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
519 \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
520 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
521 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
522 $(p_0, x_0) \in \text{gph } S$ near (\bar{p}, \bar{x}) and fixed $\pi = \bar{p}$, then S is calm [q] at (\bar{p}, \bar{x}) . Hence, depending on the
523 choice of π , condition (5.1) is necessary and sufficient for calmness [q] and the Aubin-property [q] at
524 (\bar{p}, \bar{x}) .

525 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
526 $\pi = \bar{p}$, Theorem 5.1 says: if $(p_0, x_0) \in \text{gph } S$ (e.g., a starting point for some algorithm) is sufficiently
527 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,
528 to satisfy the hypotheses, the distance $d((\bar{p}, \bar{x}), \text{gph } S)$ has to be small enough.

529 5.2 P is a metric space

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

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

530 This tells us that the theorem remains true when P is a general metric space and C is any set of this
531 type. Notice however that, with the simplest setting $C = \{p_0, \pi\}$, the descent condition (5.1) implies
532 $p' = \pi$, and the whole statement becomes trivial. This makes reasonable definitions of C for *metric*
533 *spaces* difficult unless F itself is strongly closed and we can put $C = P$.

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

538 **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 $\text{gph } S$) characterizes “stability” by linear order of convergence. Define, in $P \times X$, a distance depending on $\lambda > 0$ as

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

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

We assume that $\pi \in P$, $\gamma \geq 0$ and $\lambda > 0$ are fixed.

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

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

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

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

542 (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
543 points $(p_0, x_0) \in \text{gph } S \cap (B(\bar{p}, \alpha) \times B(\bar{x}, \alpha))$ and $\pi \in B(\bar{p}, \alpha)$, Procedure S1 generates a sequence
544 (p_k, x_k) satisfying

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

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

546 (iii) These statements remain true if we additionally require that P is a linear normed space and
547 $p_{k+1} \in \text{conv}\{p_k, \pi\}$.

548 **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
549 $p_k \rightarrow \pi$, $x_k \rightarrow \xi \in S(\pi)$ and $d(\xi, x_0) \leq \lambda^{-1} d(p_0, \pi)$. Statement (iii) shows a connection to Theorem 5.1.

550 **5.2.2 Calmness [q] via proper descent steps**

551 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:

$$\text{For all } (p, x) \in \text{gph } S \cap [B(\bar{p}, \delta) \times B(\bar{x}, \varepsilon)], \text{ some } (p', x') \in \text{gph } S \text{ satisfies} \quad (5.4)$$

$$(i) \lambda d(x', x) \leq d(p, \pi)^q \quad \text{and} \quad (ii) d(p', \pi) \leq (1 - \lambda) d(p, \pi).$$

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

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

552 Next assume $q > 0$, $(p_0, x_0) \in \text{gph } S$ and consider

553 **Procedure S2:** Beginning with $k = 0$, find any $(p_{k+1}, x_{k+1}) \in \text{gph } S$ such that

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

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

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

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

557 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

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

558 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

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

559 Hence the lemma is valid. □

560 **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
561 (5.4). Then, if $(p_0, x_0) \in \text{gph } S$ and π satisfy (5.7), Procedure S2 is applicable and defines a sequence
562 $\{x_k\}$ converging to some $\xi \in S(\pi)$ satisfying (5.6).

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

565 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
566 there is no solution with the current λ . Being applicable now means $\lambda_k \geq \lambda > 0$ for all initial points
567 $(p_0, x_0) \in \text{gph } S$ and π satisfying (5.7). Similarly, one could use varying q , beginning with $q_1 = 1$. The
568 estimates then hold with exponent \bar{q} if also $q_k \geq \bar{q} > 0$.

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

571 **Corollary 5.6.** Suppose (1.3) and $(\bar{p}, \bar{x}) \in \text{gph } S$. Then

572 (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)
573 is satisfied for all $\pi \in B(\bar{p}, \delta)$.

574 (ii) With fixed $\pi = \bar{p}$, the same holds in view of calmness $[q]$.

575 *Proof.* Necessity (\Rightarrow) follows easily from the stability definitions while Prop. 5.5 ensures the suffi-
576 ciency. □

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

579 **Theorem 5.7.** Suppose (1.3) and $(\bar{p}, \bar{x}) \in \text{gph } S$. Then

580 (i) S obeys the Aubin property $[q]$ at (\bar{p}, \bar{x})

581 \Leftrightarrow There exist $\lambda \in (0, 1)$ and $\varepsilon, \delta > 0$ such that (5.4) is satisfied for all $\pi \in B(\bar{p}, \delta)$.

582 \Leftrightarrow There are $\alpha > 0$ and $\lambda \in (0, 1)$ such that iterates (p_{k+1}, x_{k+1}) for procedure S2 exist in each
583 step, whenever the initial points satisfy $d(x_0, \bar{x}) + d(p_0, \bar{p}) + d(\pi, \bar{p}) < \alpha$ and $x_0 \in S(p_0)$.

584 (ii) With fixed $\pi \equiv \bar{p}$, the same holds in view of calmness $[q]$. □

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

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

References

- 589 [1] AUBIN, J.-P., AND EKELAND, I. *Applied Nonlinear Analysis*. Pure and Applied Mathematics.
590 John Wiley & Sons Inc., New York, 1984.
- 592 [2] BONNANS, J. F., AND SHAPIRO, A. *Perturbation Analysis of Optimization Problems*. Springer
593 Series in Operations Research. Springer-Verlag, New York, 2000.
- 594 [3] CLARKE, F. H. *Optimization and Nonsmooth Analysis*. Canadian Mathematical Society Series
595 of Monographs and Advanced Texts. John Wiley & Sons Inc., New York, 1983.
- 596 [4] COMINETTI, R. Metric regularity, tangent sets, and second-order optimality conditions. *Appl.*
597 *Math. Optim.* 21, 3 (1990), 265–287.
- 598 [5] DEMPE, S. *Foundations of Bilevel Programming*, vol. 61 of *Nonconvex Optimization and its*
599 *Applications*. Kluwer Academic Publishers, Dordrecht, 2002.
- 600 [6] DMITRUK, A. V., AND KRUGER, A. Y. Metric regularity and systems of generalized equations.
601 *J. Math. Anal. Appl.* 342, 2 (2008), 864–873.
- 602 [7] DONTCHEV, A. L., AND HAGER, W. W. An inverse mapping theorem for set-valued maps.
603 *Proc. Amer. Math. Soc.* 121, 2 (1994), 481–489.
- 604 [8] DONTCHEV, A. L., AND ROCKAFELLAR, R. T. *Implicit Functions and Solution Mappings. A*
605 *View from Variational Analysis*. Springer Monographs in Mathematics. Springer, Dordrecht,
606 2009.
- 607 [9] EKELAND, I. On the variational principle. *J. Math. Anal. Appl.* 47 (1974), 324–353.
- 608 [10] FABIAN, M., HENRION, R., KRUGER, A. Y., AND OUTRATA, J. V. Error bounds: necessary
609 and sufficient conditions. *Set-Valued Var. Anal.* 18, 4 (2010), 121–149.
- 610 [11] FACCHINEI, F., AND PANG, J.-S. *Finite-Dimensional Variational Inequalities and Complementarity*
611 *Problems. Vol. I and II*. Springer Series in Operations Research. Springer-Verlag, New
612 York, 2003.
- 613 [12] FRANKOWSKA, H. High order inverse function theorems. *Ann. Inst. H. Poincaré Anal. Non*
614 *Linéaire* 6, suppl. (1989), 283–303.
- 615 [13] GRAVES, L. M. Some mapping theorems. *Duke Math. J.* 17 (1950), 111–114.
- 616 [14] HEERDA, J., AND KUMMER, B. Characterization of calmness for Banach space mappings.
617 Preprint 06-26, Institut of Mathematics, Humboldt-University Berlin, 2006. Available at
618 <http://www.mathematik.hu-berlin.de/publ/pre/2006/p-list-06.html>.
- 619 [15] HENRION, R., AND OUTRATA, J. V. A subdifferential condition for calmness of multifunctions.
620 *J. Math. Anal. Appl.* 258, 1 (2001), 110–130.
- 621 [16] HENRION, R., AND OUTRATA, J. V. Calmness of constraint systems with applications. *Math.*
622 *Program., Ser. B* 104, 2-3 (2005), 437–464.
- 623 [17] IOFFE, A. D. Metric regularity and subdifferential calculus. *Russian Math. Surveys* 55 (2000),
624 501–558.
- 625 [18] IOFFE, A. D. On regularity concepts in variational analysis. *J. Fixed Point Theory Appl. Online*
626 *first* (2010). DOI: 10.1007/s11784-010-0021-0.
- 627 [19] IOFFE, A. D., AND OUTRATA, J. V. On metric and calmness qualification conditions in sub-
628 differential calculus. *Set-Valued Anal.* 16, 2–3 (2008), 199–227.
- 629 [20] KLATTE, D., AND KUMMER, B. *Nonsmooth Equations in Optimization. Regularity, Calculus,*
630 *Methods and Applications*, vol. 60 of *Nonconvex Optimization and its Applications*. Kluwer Aca-
631 demic Publishers, Dordrecht, 2002.
- 632 [21] KLATTE, D., AND KUMMER, B. Stability of inclusions: characterizations via suitable Lipschitz
633 functions and algorithms. *Optimization* 55, 5-6 (2006), 627–660.
- 634 [22] KLATTE, D., AND KUMMER, B. Optimization methods and stability of inclusions in Banach
635 spaces. *Math. Program.* 117, 1-2, Ser. B (2009), 305–330.

- 636 [23] KUMMER, B. Inclusions in general spaces: Hoelder stability, solution schemes and Ekeland's
637 principle. *J. Math. Anal. Appl.* 358, 2 (2009), 327–344.
- 638 [24] LYUSTERNIK, L. A. On conditional extrema of functionals. *Math. Sbornik* 41 (1934), 390–401.
639 In Russian.
- 640 [25] MINCHENKO, L. I. Multivalued analysis and differential properties of multivalued mappings and
641 marginal functions. *J. Math. Sci.* 116, 3 (2003), 3266–3302.
- 642 [26] MORDUKHOVICH, B. S. *Variational Analysis and Generalized Differentiation. I: Basic Theory*,
643 vol. 330 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Math-*
644 *ematical Sciences]*. Springer-Verlag, Berlin, 2006.
- 645 [27] ROBINSON, S. M. Strongly regular generalized equations. *Math. Oper. Res.* 5, 1 (1980), 43–62.
- 646 [28] ROBINSON, S. M. Variational conditions with smooth constraints: structure and analysis. *Math.*
647 *Program., Ser. B* 97, 1-2 (2003), 245–265.
- 648 [29] ROCKAFELLAR, R. T., AND WETS, R. J.-B. *Variational Analysis*, vol. 317 of *Grundlehren der*
649 *Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-
650 Verlag, Berlin, 1998.