

Inclusions in general spaces: Hoelder stability, Solution schemes and Ekeland's principle

Bernd Kummer*

18th June 2008

Abstract. We present two basic lemmas on exact and approximate solutions of inclusions and equations in general spaces. Its applications involve Ekeland's principle, characterize calmness, lower semicontinuity and the Aubin property of solution sets in some Hoelder-type setting and connect these properties with certain iteration schemes of descent type. In this way, the mentioned stability properties can be directly characterized by convergence of more or less abstract solution procedures. New stability conditions will be derived, too. Our basic models are (multi-) functions on a complete metric space with images in a linear normed space.

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

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

1 Introduction

This paper deals with the existence and stability of solutions to a generalized equation:

$$\begin{aligned} &\text{Given a closed multifunction } F : X \rightrightarrows P \text{ and } p \in P \\ &\quad \text{find } x \in X \text{ such that } p \in F(x). \end{aligned} \tag{1.1}$$

X is a complete metric space, P a linear normed space over \mathbb{R} .

The double arrow indicates that $F(x) \subset P$. F is said to be closed if $\text{gph } F := \{(x, p) \mid p \in F(x)\}$ is a closed set. The elements $p \in P$ are canonical parameters of the inclusion. For functions $f : X \rightarrow P$, we identify $f(x)$ and $F(x) = \{f(x)\}$. Then F is closed if f is continuous. In the whole paper, we study the solution sets to (1.1)

$$S(p) := F^{-1}(p) = \{x \in X \mid p \in F(x)\} \tag{1.2}$$

near to some $(\bar{p}, \bar{x}) \in \text{gph } S$ and stability properties of S - mainly calmness and Aubin property with some fixed exponent $q > 0$ in the estimate (Hoelder-type stability).

System (1.1) describes equations and stationary or critical points of various variational conditions. In particular, it reflects level set mappings of extended functionals

$$S(p) = \{x \in X \mid f(x) \leq p\}, \quad f : X \rightarrow \mathbb{R} \cup \{\infty\} \text{ l.s.c.}, \quad p \in \mathbb{R}, \tag{1.3}$$

where $\text{gph } F = \text{epi } f$. Systems

$$S(y) := \{x \in X \mid g(x, y) = 0\} \tag{1.4}$$

are of the form (1.1) after defining a norm for $p := g(\cdot, \bar{y}) - g(\cdot, y)$ (and x in a region of interest) and setting

$$S(p) = \{x \mid g(x, \bar{y}) = p(x)\}, \quad F = S^{-1}. \tag{1.5}$$

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

Many applications of (1.1) are known for optimization problems, for equilibria in games, in so-called MPECs of different type and stochastic and/or multilevel (multiphase) models. We refer to [7, 2, 15, 45, 39, 3, 9, 30] for the related settings. Furthermore, Lipschitz properties (which means $q = 1$ below) of solutions or feasible points are crucial for deriving duality, optimality conditions, local error estimates and penalty methods in optimization. Basic results of this type, as far as they concern the subsequent stabilities with $q = 1$, can be found in [2, 8, 16, 18, 22, 36, 37] (Aubin property), [6, 13, 31, 33, 42, 44] (strongly Lipschitz), [4, 5, 11, 25, 41, 43] (locally upper Lipschitz and calmness) and the monographs [3, 9, 29, 38, 39, 45]. Sufficient conditions for Hoelder stability ($q = \frac{1}{2}$) of stationary points in optimization problems can be found already in [1, 17, 26, 27].

The key of the current paper is Lemma 2.4 (and a modified version Lemma 4.1) on solvability of (1.1). It helps to characterize all subsequent Hoelder-type stabilities of S in Def. 1 below, by the fact that certain iteration schemes find related solutions to (1.1) for all/certain initial data near the reference point. In this way, we point out the connections between stability, approximate solutions and solution procedures directly. Also Ekeland's principle will appear as a consequence of this Lemma. Our approach avoids the known drawbacks of stability criteria by means of (mostly used) generalized derivatives, namely: possibly empty contingent derivatives, the often necessary restriction to Asplund spaces and, last not least, the (not seldom hard) translation of the derivative conditions in terms of the original data.

Three iteration schemes which consist of certain "descent steps" will be used. The first one, S1, involves (in general) global minimization. Hence it is most far from being an applicable "algorithm" in the proper sense. Nevertheless, it helps to describe stability. With the second one, S2, the convergence is linear in the parameter (image) space and the stepsize depends on the Hoelder exponents. The scheme S3 reformulates S2 by using a more familiar stepsize rule.

For $q = 1$, aspects of our approach - based on penalizations, projections and successive approximation - can be already found in [19, 31, 32] for less general spaces and Lipschitz stability. The 1-1 correspondence Lemma 2.2 between calm multifunctions and calm level sets of an assigned Lipschitz functions appeared, perhaps first, in [28]. Newton type methods (again under stronger hypotheses) for showing the Aubin property or calmness have been exploited already in [12] and [20]. Calmness of an (in)finite number of real-valued C^1 -inequalities on a Banach space has been characterized by identifying subsystems which must be metrically regular in [19] (based on the family Ξ_0 of Sect. 4.1) and by a descent method in [32].

Concerning the involved Hoelder exponents q , many questions in view of establishing more verifiable conditions for relevant systems will remain open and shall be studied in forthcoming work. Here, we show how q occurs in the related conditions and which difficulties may appear concerning $q \neq 1$ and the classical case of $q = 1$.

The paper is organized as follows. In Sect. 2, we present and specify the needed definitions as well as the basic Lemma 2.4. In Sect. 3, we interpret it and derive consequences for stability. In Sect. 4, we study the schemes S2 and S3 and certain applications to (Hoelder-) calmness for particular systems.

Notations:

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 $o(t)/t \rightarrow 0$ if $t \downarrow 0$, and $B(\bar{x}, \varepsilon) = \{x \in X \mid d(x, \bar{x}) \leq \varepsilon\}$ denotes the closed ε ball around \bar{x} . For $M \subset X$, we put $B(M, \varepsilon) = \cup_{x \in M} B(x, \varepsilon)$, and $\text{dist}(x, M) = \inf_{\xi \in M} d(x, \xi)$ is the usual point-to set-distance; $\text{dist}(x, \emptyset) = \infty$. Finally, $x \xrightarrow{M} \bar{x}$ denotes convergence $x \rightarrow \bar{x}$ with $x \in M$. Our hypotheses of differentiability, continuity or closeness have to hold near the reference points only.

2 Hoelder type stability and the basic Lemma

2.1 Stability properties

The following definitions describe, for $q = 1$, typical local Lipschitz properties of the multi-function $S = f^{-1}$ or of level sets (1.3) for functions $f : X \rightarrow \mathbb{R}$; mostly called calmness, Aubin property and Lipschitz lower semi-continuity, respectively. 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 notion "Lipschitz lower semi-continuity $[q]$ " we simply write l.s.c. $[q]$.

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

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

$$\exists \varepsilon, \delta, L > 0 : x \in S(p) \cap B(\bar{x}, \varepsilon) \Rightarrow B(x, L\|p - \pi\|^q) \cap S(\pi) \neq \emptyset \quad \forall p, \pi \in B(\bar{p}, \delta). \quad (2.1)$$

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

$$\exists \varepsilon, \delta, L > 0 : x \in S(p) \cap B(\bar{x}, \varepsilon) \Rightarrow B(x, L\|p - \bar{p}\|^q) \cap S(\bar{p}) \neq \emptyset \quad \forall p \in B(\bar{p}, \delta). \quad (2.2)$$

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

$$\exists \delta, L > 0 : B(\bar{x}, L\|\bar{p} - \pi\|^q) \cap S(\pi) \neq \emptyset \quad \forall \pi \in B(\bar{p}, \delta). \quad \diamond \quad (2.3)$$

Compared with (D1), we have $\pi = \bar{p}$ and $p = \bar{p}$ in (D2) and (D3), respectively. The constant L is called *rank* of the related stability.

If $q = 1$, these properties have several applications:

Imposed for feasible sets in optimization models, (standard C^1 - problems in \mathbb{R}^n and several problems with cone-constraints in Banach spaces) these 3 properties are constraint qualifications which ensure the existence of Lagrange multipliers. In addition,

(D1) characterizes the topological behavior of solutions in the inverse function theorem due to Graves and Lyusternik [18, 36].

(D3) required for level sets S (1.3) with $f(\bar{x}) = \bar{p}$, implies that \bar{x} *cannot be* a stationary point of the type

$$f(x) \geq f(\bar{x}) - o(d(x, \bar{x})). \quad (2.4)$$

Using these definitions for $q = 1$, other known stability properties can be defined and characterized (we apply the often used notations of [29]).

Remark 2.1. Let $q = 1$.

- (i) S is locally upper Lipschitz at $\bar{z} \Leftrightarrow S$ is calm at \bar{z} and \bar{x} is isolated in $S(\bar{p})$.
- (ii) 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 .
- (iii) S is strongly Lipschitz at $\bar{z} \Leftrightarrow S$ obeys the Aubin property at \bar{z} and $S(p) \cap B(\bar{x}, \varepsilon)$ is single-valued for some $\varepsilon > 0$ and all p near \bar{p} . \diamond

In the strongest case (iii), the solution mapping S is locally (near \bar{z}) a Lipschitz function.

2.2 Some well-known results for Banach spaces and $q = 1$

For Banach spaces X, P and $f \in C^1(X, P)$, the local upper Lipschitz property holds for (the multivalued inverse) $S = f^{-1}$ at $(f(\bar{x}), \bar{x})$ if $Df(\bar{x})$ is *injective*; the Aubin property is ensured if $Df(\bar{x})$ is *surjective* (Graves-Lyusternik theorem [18, 36]).

To obtain similar statements for F (1.1) recall that the *contingent derivative* [2] $CF(\bar{x}, \bar{p})(u)$ of F at $(\bar{x}, \bar{p}) \in \text{gph } F$ in direction $u \in X$ is the (possibly empty) set of all limits

$$v = \lim t_k^{-1}(p_k - \bar{p}) \quad \text{where } p_k \in F(\bar{x} + t_k u_k), \quad t_k \downarrow 0 \quad \text{and} \quad u_k \rightarrow u. \quad (2.5)$$

Then $CS = CF^{-1}$ satisfies, by definition, $u \in CS(\bar{p}, \bar{x})(v) \Leftrightarrow v \in CF(\bar{x}, \bar{p})(u)$. We refer to [2, 45, 29, 25] for further properties and interrelations to other generalized derivatives. If $F = f$ is a function, the argument $\bar{p} = f(\bar{x})$ can be deleted in the notation of CF .

Definition 2. CF is injective at (\bar{x}, \bar{p}) if $\|v\| \geq c\|u\| \quad \forall v \in CF(\bar{x}, \bar{p})(u)$ holds for some $c > 0$. CF is uniformly surjective near (\bar{x}, \bar{p}) if there is some $c > 0$ such that

$$B_P(0, c) \subset CF(x, p)(B_X(0, 1)) \quad \forall (x, p) \in [B(\bar{x}, c) \times B(\bar{p}, c)] \cap \text{gph } F \quad (2.6)$$

(also called CF is open with uniform rank). \diamond

For F (1.1) with $X = \mathbb{R}^n$ and $P = \mathbb{R}^m$, the local upper Lipschitz property of F^{-1} at (\bar{p}, \bar{x}) coincides with injectivity of CF at (\bar{x}, \bar{p}) [25]; the Aubin property with uniform surjectivity of CF near (\bar{x}, \bar{p}) [2]. Based on the fact that one may weaker require $cB_P \subset \text{conv } CF(x, p)(B_X)$ in (2.6), this also means injectivity of a limiting coderivative [37].

For Banach spaces, these conditions are still *sufficient* for the Aubin property, but far from being necessary even if $X = l^2$ and $P = \mathbb{R}$. We refer to [29], example BE2 with the concave function $f(x) = \inf x_k$. A detailed investigation of various criteria for Lipschitz behavior of S was the subject of [31].

Having the Aubin property, the inclusions (or equations) remain Lipschitz stable under various nonlinear perturbations as in (1.4). A first and basic approach was presented in [42] while [10] presents recent investigations and instructive results in this direction.

In the classical case of $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$, all mentioned stability properties ($q = 1$), except for calmness, coincide with $\det Df(\bar{x}) \neq 0$. Calmness makes difficulties since it may disappear after adding small smooth functions: $S = f^{-1}$ for $f \equiv 0$ is calm at 0, not so for $f = \varepsilon x^2$.

2.3 Calm $[q]$ multifunctions and Lipschitz functions

For arbitrary *multifunctions* (1.2), calmness is a monotonicity property for two *Lipschitz functions*,

$$\text{dist}(x, S(\bar{p})) \quad \text{and} \quad \psi_S(x, p) = \text{dist}((p, x), \text{gph } S),$$

defined via $d_{P \times X}((p, x), (p', x')) = \max\{\|p - p'\|, d(x, x')\}$ or some equivalent metric in the product space. For $q = 1$, the next statement is Lemma 3.2 in [28].

Lemma 2.2. S is calm $[q]$ with $0 < q \leq 1$ at $(\bar{p}, \bar{x}) \in \text{gph } S$ if and only if

$$\exists \varepsilon > 0, \alpha > 0 \text{ such that } \alpha \text{dist}(x, S(\bar{p})) \leq \psi_S(x, \bar{p})^q \quad \forall x \in B(\bar{x}, \varepsilon). \quad \diamond \quad (2.7)$$

In other words, calmness $[q]$ at (\bar{p}, \bar{x}) is violated iff

$$0 < \psi_S(x_k, \bar{p})^q = o(\text{dist}(x_k, S(\bar{p}))) \quad \text{holds for some sequence } x_k \rightarrow \bar{x}. \quad (2.8)$$

Proof. Let (2.7) be true. Then, given $x \in S(p) \cap B(\bar{x}, \varepsilon)$, it holds

$$\psi_S(x, \bar{p})^q \leq d((p, x), (\bar{p}, x))^q = \|p - \bar{p}\|^q$$

and $\alpha \text{dist}(x, S(\bar{p})) \leq \psi_S(x, \bar{p})^q \leq \|p - \bar{p}\|^q$, which yields calmness $[q]$ with all $L > \frac{1}{\alpha}$. Conversely, let (2.7) be violated, i.e., (2.8) be true. Given any positive $\delta_k < o(\text{dist}(x_k, S(\bar{p})))$ we find $(p_k, \xi_k) \in \text{gph } S$ such that, for large k ,

$$d((p_k, \xi_k), (\bar{p}, x_k))^q < \psi_S(x_k, \bar{p})^q + \delta_k < b_k := 2 o(\text{dist}(x_k, S(\bar{p}))). \quad (2.9)$$

Particularly, (2.9) implies, due to $0 < q \leq 1$, that $d(\xi_k, x_k)^q < b_k$ and $b_k^{1/q} \leq b_k$. In addition, the triangle inequality $\text{dist}(x_k, S(\bar{p})) \leq d(x_k, \xi_k) + \text{dist}(\xi_k, S(\bar{p}))$ yields

$$\begin{aligned} \text{dist}(\xi_k, S(\bar{p})) &\geq \text{dist}(x_k, S(\bar{p})) - d(\xi_k, x_k) \\ &> \text{dist}(x_k, S(\bar{p})) - b_k^{1/q} &\geq \text{dist}(x_k, S(\bar{p})) - b_k. \end{aligned}$$

Since (2.9) also implies $\|p_k - \bar{p}\|^q < b_k$, we thus obtain for $\xi_k \in S(p_k)$ and $k \rightarrow \infty$,

$$\frac{\|p_k - \bar{p}\|^q}{\text{dist}(\xi_k, S(\bar{p}))} < \frac{b_k}{\text{dist}(x_k, S(\bar{p})) - b_k^{1/q}} \leq \frac{b_k}{\text{dist}(x_k, S(\bar{p})) - b_k} \rightarrow 0$$

Because of $\xi_k \rightarrow \bar{x}$ and $\xi_k \in S(p_k)$, so S is not calm $[q]$ at (\bar{p}, \bar{x}) . \square

Hence calmness $[q]$ for S at (\bar{p}, \bar{x}) coincides with calmness of a *Lipschitzian inequality*:

Corollary 2.3. *For $0 < q \leq 1$, S is calm $[q]$ at $(\bar{p}, \bar{x}) \in \text{gph } S \subset P \times X \Leftrightarrow$ the level set map $\Sigma(r) := \{x \mid \psi_S(x, \bar{p}) \leq r\}$ is calm $[q]$ at $(0, \bar{x}) \in \mathbb{R} \times X$. \diamond*

For this reason, we shall pay particular attention to the mapping S (1.3). The (usually complicated) distance $\psi_S(\cdot, \bar{p})$ may be replaced by any (locally Lipschitz) function ϕ satisfying

$$\alpha \phi(x) \leq \psi_S(x, \bar{p}) \leq \beta \phi(x) \text{ for } x \text{ near } \bar{x} \text{ and certain constants } 0 < \alpha \leq \beta. \quad (2.10)$$

Estimates of ψ_S for (often used) composed systems can be found in [28]. For convex mappings S (i.e. X is a B-space and $\text{gph } S$ is convex), both ψ_S and $d(\cdot, S(\bar{p}))$ are convex functions. For polyhedral S in finite dimension and polyhedral norms, ψ_S and $d(\cdot, S(\bar{p}))$ are piecewise linear. Generally, checking condition (2.7) may be a hard task. But, by the equivalence, one cannot avoid to investigate ψ_S in the original or some equivalent way if calmness or the Aubin property with some exponent q should be characterized. The subsequent statements can be written (even shorter) in terms of ψ_S without using S explicitly. However, though multifunctions and generalized equations play a big role in nonsmooth analysis, we formulate the most statements in this terminology.

2.4 The basic statement

From now on, $q > 0$ denotes any fixed exponent, S, F are related to (1.1) where

$$\begin{aligned} X \text{ is a complete metric space, } P \text{ a linear normed space,} \\ S : P \rightrightarrows X \text{ is closed, } (\bar{p}, \bar{x}) \in \text{gph } S, \quad F = S^{-1}. \end{aligned} \quad (2.11)$$

Clearly, $q > 1$ is out of interest if $F = f$ is a locally Lipschitz function.

Motivation: Given $(p, x) \in \text{gph } S$ and $\pi \in P$ our stability properties claim to verify that some $\xi \in S(\pi) \cap B(x, L\|\pi - p\|^q)$ exists. To show this, we shall use iterations $(p_k, x_k) \in \text{gph } S$ which start at (p, x) and converge to (π, ξ) .

In view of calmness, such iterations may be trivial, e.g., if only $S(p)$ and $S(\pi)$ are non-empty and already (p_2, x_2) is the point we are looking for. This, however, is not the typical situation for basic applications. In contrary, the hypotheses will usually only permit to achieve an approximation of (π, ξ) after a big couple of steps. Furthermore, to apply local informations of F near $(x_k, p_k) \in \text{gph } F$ for constructing $(p_{k+1}, x_{k+1}) \in \text{gph } S$, small steps are usually necessary and, as we will see, possible under the Aubin property.

In what follows, we use the notion *procedure* for indicating the conditions which define the next iterate. We emphasize that these procedures are still far from being *algorithms* which can be practically applied for computation unless specifying them under additional hypotheses

(like in section 4). On the other hand, they indicate which type of next iterates exist and can be used in order to determine a point (π, ξ) in question.

Moving from (p, x) to (p', x') with $\|p' - \pi\| < \|p - \pi\|$ in $\text{gph } S$ can be seen as a descent step for such "procedures". The next lemma asserts

$$S(\pi) \cap B(x, L\|p - \pi\|^q) \neq \emptyset$$

for particular initial points $(p, x) = (p_1, x_1)$ whenever descent steps of the type

$$\|p' - \pi\|^q + \lambda d(x', x) < \|p - \pi\|^q \quad (\lambda > 0). \quad (2.12)$$

are possible for all (p, x) , $p \neq \pi$ in some neighborhood Ω of (\bar{p}, \bar{x}) . Usually, (p_1, x_1) belongs to a small neighborhood $\Omega' \subset \Omega$. To obtain the stability statements we are aiming at, these neighborhoods must be specified.

Below, the constant λ plays the role of L^{-1} , ε and δ are not necessarily small. Concerning the arrangement of the constants in order to satisfy the subsequent crucial condition (2.14) we refer to Remark 2.5. Furthermore, we will put $C = P$ in the most applications, except for (3.8).

Roughly speaking, the lemma is a generalization of a simple fact: *If $f \in C^1(\mathbb{R}^n, \mathbb{R})$, $f(0) = 0$ and $\|Df\| \geq \lambda$ everywhere then $f = \pm\lambda$ has solutions in the unit ball.*

Lemma 2.4. *Let S satisfy (2.11), $C \subset P$ be convex, $\bar{p} \in C$ and $\lambda, q > 0$. In addition, suppose there are $\varepsilon, \delta > 0$ and some $\pi \in B(\bar{p}, \delta) \cap C$ such that*

$$\begin{aligned} & \text{for all } (p, x) \in \text{gph } S \cap [B(\bar{p}, \delta) \times B(\bar{x}, \varepsilon)] \text{ with } p \in C \setminus \{\pi\} \\ & \text{there is some } (p', x') \in \text{gph } S \text{ satisfying (2.12) and } p' \in \text{conv}\{p, \pi\}. \end{aligned} \quad (2.13)$$

Then, if $(p_1, x_1) \in \text{gph } S$, $p_1 \in B(\bar{p}, \delta) \cap C$ and $d(x_1, \bar{x}) + \|p_1 - \pi\|$ is small enough such that

$$d(x_1, \bar{x}) + \lambda^{-1}\|p_1 - \pi\|^q \leq \varepsilon, \quad (2.14)$$

there exists some $\xi \in S(\pi) \cap B(x_1, \lambda^{-1}\|\pi - p_1\|^q)$. \diamond

Proof. We suppose (2.13) and consider any (p_1, x_1) and π satisfying the hypotheses. The Lemma is trivial if $p_1 = \pi$ (put $\xi = x_1$). Let $p_1 \neq \pi$. Now $(p, x) = (p_1, x_1)$ fulfills

$$p \in B(\bar{p}, \delta) \cap C, \quad x \in S(p) \cap B(\bar{x}, \varepsilon), \quad p \neq \pi \quad (2.15)$$

as required in (2.13). Hence (2.12) ensures

$$\mu(p, x) := \inf \{ \|p' - \pi\|^q + \lambda d(x', x) \mid (p', x') \in \text{gph } S, p' \in \text{conv}\{p, \pi\} \} < \|p - \pi\|^q.$$

Next consider

Procedure S1: Beginning with $k = 1$ assign, to (p_k, x_k) , some (p_{k+1}, x_{k+1}) , with

$$\begin{aligned} (i) \quad & \|p_{k+1} - \pi\|^q + \lambda d(x_{k+1}, x_k) < \|p_k - \pi\|^q, \quad (p_{k+1}, x_{k+1}) \in \text{gph } S \\ (ii) \quad & p_{k+1} \in \text{conv}\{p_k, \pi\} \\ (iii) \quad & \|p_{k+1} - \pi\|^q + \lambda d(x_{k+1}, x_k) < \mu(p_k, x_k) + 1/k. \end{aligned} \quad (2.16)$$

Convergence of x_k : From (i), we obtain for step $n \geq 1$, as long as the iterations exist,

$$\begin{aligned} \lambda d(x_{n+1}, x_1) & \leq \lambda \sum_{k=1}^n d(x_{k+1}, x_k) \\ & \leq \|p_1 - \pi\|^q - \|p_2 - \pi\|^q + \dots + \|p_n - \pi\|^q - \|p_{n+1} - \pi\|^q \\ & = \|p_1 - \pi\|^q - \|p_{n+1} - \pi\|^q \leq \|p_1 - \pi\|^q. \end{aligned} \quad (2.17)$$

Thus (2.14) and (2.17) yield

$$d(x_{n+1}, \bar{x}) \leq d(x_{n+1}, x_1) + d(x_1, \bar{x}) \leq \lambda^{-1} \|p_1 - \pi\|^q + d(x_1, \bar{x}) \leq \varepsilon.$$

So, if $p_{k+1} \neq \pi$, (p_{k+1}, x_{k+1}) fulfills (2.15) since $p_{k+1} \in \text{conv}\{p_k, \pi\} \subset \text{conv}\{p_1, \pi\} \subset C$. If $p_n = \pi$ for some n then $\xi = x_n$ satisfies the assertion due to (2.17). Otherwise, since $\sum_k d(x_{k+1}, x_k)$ is bounded, the sequence $\{x_k\}$ converges in the complete space X , $x_k \rightarrow \xi$.

Accumulation points of p_k : Again by (i), we observe $\|p_k - \pi\| \rightarrow \beta$ for some β .

Next we need an accumulation point, say $\eta \in C \cap B(\bar{p}, \delta)$, of the sequence $\{p_k\}$. By our assumptions, η exists due to (ii) since all p_k belong to the compact segment $\text{conv}\{p_1, \pi\}$.

Notice that η exists also under other assumptions, discussed in Remark 3.10, below. For this reason, let us only use that some (infinite) subsequence of $\{p_k\}$ converges to η .

Since S is closed, η fulfills $\xi \in S(\eta)$. If $\eta = \pi$, (2.17) implies again the assertion. Let $\eta \neq \pi$. Due to $\eta \in \text{conv}\{p_1, \pi\} \subset C$ and the above estimates, (2.13) holds for $(p, x) := (\eta, \xi)$, too: There exist $(p', x') \in \text{gph } S$ and some $\alpha > 0$, such that

$$\|p' - \pi\|^q + \lambda d(x', \xi) < \|\eta - \pi\|^q - \alpha \quad \text{and} \quad p' \in \text{conv}\{\eta, \pi\}.$$

The shown convergence of some subsequence of (p_k, x_k) yields, for certain large k ,

$$\mu(p_k, x_k) \leq \|p' - \pi\|^q + \lambda d(x', x_k) < \|p_k - \pi\|^q - \alpha$$

and by (iii) also

$$\|p_{k+1} - \pi\|^q + \lambda d(x_{k+1}, x_k) < \mu(p_k, x_k) + 1/k < \beta^q - \alpha/2 + 1/k < \beta^q - \alpha/4.$$

This contradicts $\|p_k - \pi\| \rightarrow \beta$ and implies $\eta = \pi$. So the Lemma is true, indeed. \square

Remark 2.5. If the constants satisfy

$$\lambda^{-1}(2\delta)^q \leq \frac{1}{2}\varepsilon \tag{2.18}$$

then (2.14) holds for all (p_1, x_1) near (\bar{p}, \bar{x}) , namely if $x_1 \in B(\bar{x}, \frac{1}{2}\varepsilon)$ and $p_1 \in B(\bar{p}, \delta)$. \diamond

Lemma 2.6. For the level set map S (1.3) and $\bar{p} = f(\bar{x})$, condition (2.13) is equivalent to

$$\begin{aligned} \forall x \in B(\bar{x}, \varepsilon) \text{ with } \pi < f(x) \leq \bar{p} + \delta \text{ and } f(x) \in C \\ \exists x' : (f_\pi(x') - \pi)^q + \lambda d(x', x) < (f_\pi(x) - \pi)^q \end{aligned} \tag{2.19}$$

where $f_\pi(x') = \max\{\pi, f(x')\}$. \diamond

Proof. By definition, (2.13) requires

$$\begin{aligned} \forall (p, x) \text{ with } f(x) \leq p, \quad x \in B(\bar{x}, \varepsilon) \quad \text{and} \quad p \in [\bar{p} - \delta, \bar{p} + \delta] \cap C, \quad p \neq \pi \\ \exists (p', x') : p' \in \text{conv}\{\pi, p\}, \quad f(x') \leq p' \quad \text{and} \quad |p' - \pi|^q + \lambda d(x', x) < |p - \pi|^q. \end{aligned} \tag{2.20}$$

If $f(x) \leq \pi$ the point $(x', p') = (x, \pi)$ trivially satisfies $f(x') \leq p'$ and $0 = |p' - \pi|^q + \lambda d(x', x) < |p - \pi|^q$. Hence (2.20) coincides with

$$\begin{aligned} \forall (p, x) \text{ with } \pi < f(x) \leq p, \quad x \in B(\bar{x}, \varepsilon) \quad \text{and} \quad \pi < p \leq \bar{p} + \delta, \quad p \in C \\ \exists (p', x') : p' \in [\pi, p], \quad f(x') \leq p' \quad \text{and} \quad (p' - \pi)^q + \lambda d(x', x) < (p - \pi)^q. \end{aligned}$$

Here, it suffices to look at the smallest possible $p' = f_\pi(x')$ and the smallest possible $p = f(x)$ only. This proves the Lemma. \square

3 Consequences and interpretations of Lemma 2.4

Violation of assumption (2.13):

Condition (2.13) fails to hold iff some $(p, x) \in \text{gph } S \cap [B(\bar{p}, \delta) \times B(\bar{x}, \varepsilon)]$, $p \in C$ fulfills

$$\|p' - \pi\|^q + \lambda d(x', x) \geq \|p - \pi\|^q > 0 \quad \forall (p', x') \in \text{gph } S, p' \in \text{conv}\{\pi, p\}. \quad (3.1)$$

In other words, then (p, x) is a global solution of the problem

$$\min_{p', x'} \|p' - \pi\|^q + \lambda d(x', x) \quad \text{s.t.} \quad (p', x') \in \text{gph } S, p' \in \text{conv}\{p, \pi\} \quad (3.2)$$

with optimal value $v > 0$. If (p, x) solves (3.2) with $v = 0$ then $p = \pi$ and $x \in S(\pi) \cap B(\bar{x}, \varepsilon)$ hold trivially. Thus the existence of interesting points $(p, x) \in \text{gph } S$ follows in both cases.

3.1 Calmness and Aubin property $[q]$

Remark 3.1. (Necessity of (2.13)) If S obeys the Aubin property $[q]$ at (\bar{p}, \bar{x}) , condition (2.13) can be satisfied with $C = P$, $\lambda \in (0, L^{-1})$, ε, δ from Def. 1 and *all* $\pi \in B(\bar{p}, \delta)$. Indeed, given (p, x) as in (2.13), put $p' = \pi$ and select $x' \in S(\pi) \cap B(x, L\|p - \pi\|^q)$ which exists by Def. 1. Then (2.12) holds trivially. If S is calm $[q]$ at (\bar{p}, \bar{x}) , condition (2.13) can be satisfied with the same settings; only $\pi = \bar{p}$ is fixed, now. \diamond

Remark 3.2. (Sufficiency of (2.13)) In Lemma 2.4, let $C = P$ and let the constants be taken as in Remark 2.5. Then the assertion becomes

$$S(\pi) \cap B(x_1, L\|p_1 - \pi\|^q) \neq \emptyset \quad \text{if } L = \lambda^{-1}, p_1 \in B(\bar{p}, \delta) \text{ and } x_1 \in S(p_1) \cap B(\bar{x}, \frac{1}{2}\varepsilon). \quad (3.3)$$

If (2.13) is valid with $\pi = \bar{p}$ and any $\lambda, \varepsilon, \delta > 0$, it is also valid for smaller δ satisfying (2.18). Thus (3.3) may be applied and ensures calmness. Similarly, if (2.13) holds for all $\pi \in B(\bar{p}, \delta)$, it holds for smaller δ satisfying (2.18), too. Then (3.3) proves the Aubin property (both with exponent q). \diamond

The latter remarks imply immediately

Proposition 3.3. *Suppose (2.11) and let $C = P$, $\bar{z} = (\bar{p}, \bar{x})$. Then*

(i) *S obeys the Aubin property $[q]$ at $\bar{z} \Leftrightarrow$ there are $\lambda, \varepsilon, \delta > 0$ satisfying condition (2.13) for all $\pi \in B(\bar{p}, \delta)$.*

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

If $\pi = \bar{p}$ we may check whether already $(p', x') = (\bar{p}, \bar{x})$ satisfies (2.12). Having

$$\lambda d(\bar{x}, x) < \|p - \bar{p}\|^q$$

then the calmness estimate is evident and nothing remains to prove. Thus, to verify calmness via the existence of appropriate $\lambda, \varepsilon, \delta$, only points

$$(p, x) \rightarrow (\bar{p}, \bar{x}) \text{ such that } (p, x) \in \text{gph } S, x \neq \bar{x} \text{ and } \lim \|p - \bar{p}\|^q d(x, \bar{x})^{-1} = 0 \quad (3.4)$$

are of interest.

Calm level sets: Applying Prop. 3.3(ii) to level sets, we obtain the following statement for $(\bar{p}, \bar{x}) = (0, \bar{x})$ due to the equivalence between (2.13) and (2.19) and $\pi = 0$.

Proposition 3.4. *S (1.3) is calm $[q]$ at $(0, \bar{x}) \Leftrightarrow$ there are $\lambda, \varepsilon, \delta > 0$ such that*

$$\forall x \in B(\bar{x}, \varepsilon) \text{ with } 0 < f(x) \leq \delta \quad \exists x' : \max\{0, f(x')\}^q - f(x)^q < -\lambda d(x', x). \quad \diamond \quad (3.5)$$

For S (1.3) and $\bar{p} = 0$, (3.4) means $x \rightarrow \bar{x}$, $f(x) > 0$ and $0 = \lim f(x)^q d(x, \bar{x})^{-1}$. If $X = \mathbb{R}^n$, each sequence $x = x_k$ of this type obeys a subsequence such that

$$x_k = \bar{x} + t_k u + o(t_k) w_k, \quad w_k \rightarrow w, \quad t_k \downarrow 0, \quad f(x_k) > 0, \quad \lim f(x_k)^q / t_k = 0. \quad (3.6)$$

This verifies

Corollary 3.5. *If $X = \mathbb{R}^n$ then S (1.3) is calm $[q]$ at $(0, \bar{x}) \Leftrightarrow$ for some $\lambda > 0$ and each sequence of the form (3.6),*

$$\text{some } x'_k \text{ satisfies } \max\{0, f(x'_k)\}^q - f(x_k)^q < -\lambda d(x'_k, x_k) \quad (\text{for large } k). \quad \diamond \quad (3.7)$$

3.2 Ekeland's principle and lower semi-continuity $[q]$

Remark 3.6. To find some $\xi \in S(\pi) \cap B(\bar{x}, \lambda^{-1} \|\pi - \bar{p}\|^q)$ as required for S being l.s.c. $[q]$, put

$$\delta = \|\pi - \bar{p}\|, \quad \varepsilon = \lambda^{-1} \delta^q, \quad C = \text{conv}\{\bar{p}, \pi\}. \quad (3.8)$$

Then, setting $(p_1, x_1) = (\bar{p}, \bar{x})$, (2.14) holds again and Lemma 2.4 yields

$$\begin{aligned} &\text{Either some } \xi \in S(\pi) \cap B(\bar{x}, \varepsilon) \text{ exists or} \\ &\text{some } (p, x) \in \text{gph } S \cap [B(\bar{p}, \delta) \times B(\bar{x}, \varepsilon)], \quad p \in C \setminus \{\pi\} \text{ solves (3.2). } \quad \diamond \end{aligned}$$

In the first case, $(p, x) = (\pi, \xi)$ solves problem (3.2), too. This already proves an existence theorem.

Theorem 3.7. *Suppose (2.11). Given any $\pi \in P$ and $\lambda, q > 0$, choose δ, ε, C as in (3.8). Then some $(p, x) \in \text{gph } S \cap [B(\bar{p}, \delta) \times B(\bar{x}, \varepsilon)], \quad p \in C$ solves (3.2), i.e.,*

$$(p, x) \in \underset{p', x'}{\text{argmin}} \{ \|p' - \pi\|^q + \lambda d(x', x) \mid (p', x') \in \text{gph } S, \quad p' \in \text{conv}\{p, \pi\} \}. \quad \diamond \quad (3.9)$$

The application to level sets (1.3) leads us to

Proposition 3.8. *Let $f : X \rightarrow \mathbb{R} \cup \{\infty\}$ be l.s.c., $\pi < f(\bar{x}) < \infty$, $\lambda, q > 0$ and $f_\pi(\cdot) = \max\{\pi, f(\cdot)\}$. Then some $z \in B(\bar{x}, \lambda^{-1}(f(\bar{x}) - \pi)^q)$ fulfills*

$$f(z) \leq f(\bar{x}) \quad \text{and} \quad (f_\pi(x') - \pi)^q + \lambda d(x', z) \geq (f_\pi(z) - \pi)^q \quad \forall x' \in X. \quad \diamond \quad (3.10)$$

Proof. Put $\bar{p} = f(\bar{x})$, choose δ, ε, C as in (3.8) and study condition (2.19) of Lemma 2.6. If (2.19) is violated then (3.10) holds for some $z \in B(\bar{x}, \varepsilon)$ with $\pi < f(z) \leq \bar{p}$ and nothing remains to prove. Otherwise (2.19) and (2.13) are satisfied. So we can apply Lemma 2.4 like in Remark 3.6, with $(p_1, x_1) = (\bar{p}, \bar{x})$ since (2.14) is valid. In consequence, some $\xi \in S(\pi) \cap B(\bar{x}, \varepsilon)$ exists, and we may put $z = \xi$ because of $(f_\pi(\xi) - \pi)^q = 0$. \square

Let us again specify:

Proposition 3.9. *(Ekeland's principle for $g = f^q$) If $\inf_X f = 0$ holds in Prop. 3.8, then some $z \in X$ fulfills*

$$f(z) \leq f(\bar{x}), \quad d(z, \bar{x}) \leq \lambda^{-1} f(\bar{x})^q \quad \text{and} \quad f(x')^q + \lambda d(x', z) \geq f(z)^q \quad \forall x' \in X. \quad \diamond \quad (3.11)$$

The usual constants in Ekeland's principle [14] ($q = 1$) for $\inf f = 0$ are

$$f(\bar{x}) = \varepsilon_E, \quad \alpha_E > 0, \quad \lambda_E = \varepsilon_E / \alpha_E, \quad d(z, \bar{x}) \leq \alpha_E. \quad (3.12)$$

Here, we obtain the same by setting $\varepsilon_E = f(\bar{x}) = \delta$ and $\alpha_E = f(\bar{x}) / \lambda$.

(Generalized) derivatives :For $0 < q \leq 1$, the function r^q is concave on \mathbb{R}^+ . So (3.11) also yields

$$q (f(x') - f(z)) f(z)^{q-1} \geq f(x')^q - f(z)^q \geq -\lambda d(x', z) \quad \forall x' \in X. \quad (3.13)$$

If X is a Banach space and $f \in C^1$ then (3.13) implies via $x' = z + tu$, $t \rightarrow 0$,

$$q \|Df(z)\| \leq \lambda f(z)^{1-q} \leq \lambda f(\bar{x})^{1-q}. \quad (3.14)$$

For small $f(\bar{x})$ and $q < 1$, the estimate (3.14) is better than $\|Df(z)\| \leq \lambda$ for $q = 1$, whereas the bound $\lambda^{-1} f(\bar{x})^q$ of $d(z, \bar{x})$ is larger than $\lambda^{-1} f(\bar{x})$.

For arbitrary l.s.c. f on a Banach space, (3.13) can be applied in order to estimate $Cf(x)$ and $D_-f(x; u) = \inf Cf(x)(u)$, the contingent and the lower Dini derivative in direction u . By the definitions only, then (3.13) implies (as for $f \in C^1$) that

$$\inf_{\|u\|=1} D_-f(z; u) \geq -\lambda f(z)^{1-q} \geq -\lambda f(\bar{x})^{1-q}.$$

If f is locally Lipschitz, $Cf(x)(u)$ is a non-empty compact interval, and one may put $u_k \equiv u$ in Def. (2.5).

3.3 The role of C , $\text{conv}\{p, \pi\}$ and q

The convex set C restricts the variation of p to regions of interest, e.g. a subspace (closed or not) of P or a line-segment only. If C is closed, one can pass to the (again closed) mappings $F_C(x) = F(x) \cap C$ and $S = F_C^{-1}$ in order to avoid this restriction.

The condition $p' \in \text{conv}\{p, \pi\}$ of (2.13) requires to study solutions for the homotopy

$$p_\alpha = \alpha \pi + (1 - \alpha)p, \quad 0 \leq \alpha \leq 1.$$

Obviously, $\|p_\alpha - p\| = \alpha \|\pi - p\|$, $\|p_\alpha - \pi\| = (1 - \alpha)\|\pi - p\|$. So the bounds $\|p' - p\|^q$ and $b := \|p - \pi\|^q - \|p' - \pi\|^q$ of $d(x', x)$ in Def. 1 and (2.12) can be easily compared

$$\|p' - p\|^q \geq b \quad \text{if } q \leq 1, \quad \|p' - p\|^q \leq b \quad \text{if } q \geq 1. \quad (3.15)$$

Remark 3.10. If $\dim P < \infty$, the extra requirement

$$\|p_1 - \pi\| + \|\pi - \bar{p}\| \leq \delta \quad (3.16)$$

under (2.14) allows to delete all condition which involve C or $\text{conv}\{p, \pi\}$ in Lemma 2.4. \diamond

Proof. Indeed, the proof of Lemma 2.4 shows that, for (non-convex) closed C of finite dimension, the conditions $p' \in \text{conv}\{p, \pi\}$ of (2.13) and (2.16)(ii) can be replaced by $p' \in C$ and $p_{k+1} \in C$, respectively. To ensure that $p_{k+1} \in B(\bar{p}, \delta)$ remains true for the constructed points, (3.16) suffices since $\|p_{k+1} - \bar{p}\| \leq \|p_{k+1} - \pi\| + \|\pi - \bar{p}\| \leq \|p_1 - \pi\| + \|\pi - \bar{p}\| \leq \delta$. \square

In general, the convergence of sequences satisfying $\|p_{k+1} - \pi\| < \|p_k - \pi\|$ is connected with the drop property in the parameter space P [34].

3.4 Local and global aspects

Let X be a Banach space. Then, conditions for the discussed stabilities ($q = 1$) are usually given via generalized derivatives or subgradients, cf. section 2.2. However, for calmness, they imply only sufficient conditions, in general. The key consists in the fact that, in (2.13), $d((p', x'), (p, x))$ may be too large in order to apply informations on derivatives at (p, x) . So it becomes important to know whether or not

$$\text{condition (2.13) implies that } (p, x) \text{ is not a local minimizer of problem (3.2).} \quad (3.17)$$

Having (3.17), local optimality conditions could be used to check (2.13) in equivalent way. Obviously, (3.17) holds if $\text{gph } S$ is convex and $q = 1$ (connect (p', x') and (p, x) by a line). It also holds if condition (2.13) is required for all π near \bar{p} (as needed for the Aubin property).

Lemma 3.11. *In Lemma 2.4, let $C = P$, $q \geq 1$ and $\lambda, \varepsilon, \delta$ be constants satisfying (2.13) for all $\pi \in B(\bar{p}, \delta)$. Then, with possibly smaller constants $\lambda', \varepsilon', \delta'$, statement (3.17) is valid. \diamond*

Proof. By Prop. 3.3, S obeys the Aubin property $[q]$ at (\bar{p}, \bar{x}) with rank $L = \lambda^{-1}$ and certain $\varepsilon', \delta' > 0$ in Def. 1. Thus, given $p \in B(\bar{p}, \delta') \setminus \{\pi\}$ and $x \in S(p) \cap B(\bar{x}, \varepsilon')$ we can first choose $p' \in \text{relint conv}\{p, \pi\}$ arbitrarily close to p and obtain next the existence of $x' \in S(p') \cap B(x, L\|p' - p\|^q)$. With $\lambda' = \lambda/2$, this yields $\lambda'd(x', x) < \|p' - p\|^q$. Due to $q \geq 1$ and (3.15) we may continue, $\|p' - p\|^q \leq \|\pi - p\|^q - \|p' - \pi\|^q$. Thus inequality (2.14) holds as required for (p', x') arbitrarily close to (p, x) . \square

Generally, (3.17) can be violated.

Example 1. The locally Lipschitz function (differentiable everywhere, but not C^1)

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

has local minima and maxima arbitrarily close to 0 though $S = f^{-1}$ is calm at the origin. Hence (3.17) fails to hold in spite of the fact that the hypotheses of Lemma 2.4 are satisfied; put $p' = 0$ for $(\bar{p}, \bar{x}) = (0, 0)$, $\pi = 0$ and small constants. \diamond

Without (3.17), derivative-conditions (which exclude that (p, x) solves (3.2) locally) are only sufficient for the desired stability. This also applies to the calmness conditions for level sets (1.3) in terms of slopes and subdifferentials in [24], Thm. 2.1. Having, e.g., directional derivatives f' of f at x , it suffices to know - by Prop. 3.4 - that, for $\bar{p} = f(\bar{x}) = 0$,

$$\inf_{u \in B} f'(x; u) \leq -\lambda \quad \text{for all } x \text{ near } \bar{x} \text{ with } 0 < f(x) \leq \delta \text{ (for some } \delta > 0\text{)}. \quad (3.18)$$

Similarly, contingent derivatives $Cf(x)$ of a locally Lipschitz function f on \mathbb{R}^n can be used to obtain the sufficient calmness condition

$$\inf_{u \in B} \max_{v \in Cf(x)(u)} v \leq -\lambda \quad \text{for all } (x, f(x)) \text{ near } (\bar{x}, 0), f(x) > 0. \quad (3.19)$$

For convex f on \mathbb{R}^n , these conditions coincide and are equivalent to

$$\min_{x^* \in \partial\phi(x)} \|x^*\| \geq \lambda \quad \text{for all } (x, f(x)) \text{ near } (\bar{x}, 0), f(x) > 0. \quad (3.20)$$

by the basic relation between f' and the convex subdifferential. Thus, these are sharp criteria for calmness of S (1.3) in the convex case, but not in the Lipschitzian one. To obtain necessity, one needs additional hypotheses which ensure that $f(x')^q - f(x)^q$ can be sufficiently sharp estimated by the used generalized derivative of f at x with $f(x) > 0$, provided that $x', x \in B(\bar{x}, \varepsilon)$ and ε is small enough, cf. Thm. 4.7.

3.5 Uniform Lipschitzian lower semi-continuity and Aubin property

We call S l.s.c. $[q]$ near $(\bar{p}, \bar{x}) \in \text{gph } S$ with uniform rank L if, for all $(p, x) \in \text{gph } S$ near (\bar{p}, \bar{x}) , there exists some $\delta(p, x) > 0$ such that

$$S(p') \cap B(x, L\|p' - p\|^q) \neq \emptyset \quad \forall p' \in B(p, \delta(p, x)). \quad (3.21)$$

Compared with remark 2.1 (ii), now the radii of the balls $B(p, \delta)$ may depend on (p, x) .

Corollary 3.12. *If $q \geq 1$ then S (2.11) obeys the Aubin property $[q]$ at $(\bar{p}, \bar{x}) \Leftrightarrow S$ is l.s.c. $[q]$ near (\bar{p}, \bar{x}) with uniform rank L . \diamond*

Proof. Direction (\Rightarrow) is trivial, we consider (\Leftarrow) . Suppose (3.21) for all $(p, x) \in \text{gph } S \cap [B(\bar{p}, \delta_0) \times B(\bar{x}, \varepsilon_0)]$. Let $0 < \lambda < L^{-1}$. Given $\pi \in B(\bar{p}, \delta_0)$ and $p \neq \pi$ select any $p' \in \text{relint conv}\{\pi, p\} \cap B(p, \delta(p, x))$ and next x' in the intersection of (3.21). Because of (3.15) and $q \geq 1$ now $\lambda d(x', x) < \|p' - p\|^q \leq \|\pi - p\|^q - \|p' - \pi\|^q$ yields the existence of x' as required in (2.13) for $\varepsilon = \varepsilon_0$, $\delta = \delta_0$ and all $\pi \in B(\bar{p}, \delta)$. So Prop. 3.3 ensures the assertion. \square

For $q = 1$ and under additional hypotheses (namely that $x \mapsto \text{dist}(\pi, F(x))$ is l.s.c., and projections onto $F(x)$ exist if $F(x) \neq \emptyset$), this statement can be also found in [29] (Thm. 3.4) or [31] (Thm. 1). There, the additional hypotheses were needed in order to apply Ekeland's principle for deriving the related stability. Here, conversely, his principle just follows from the sufficiently general stability statements.

4 Proper descent steps

To modify condition (2.12) let $0 < \lambda < 1$, $\pi \in P$ and require:

$$\begin{aligned} & \text{For all } (p, x) \in \text{gph } S \cap [B(\bar{p}, \delta) \times B(\bar{x}, \varepsilon)] \\ & \text{there is some } (p', x') \in \text{gph } S \text{ satisfying} \\ (i) \quad & \lambda d(x', x) \leq \|p - \pi\|^q \text{ and } (ii) \quad \|p' - \pi\| \leq (1 - \lambda) \|p - \pi\|. \end{aligned} \quad (4.1)$$

Condition (2.12) is now weakened by deleting the term $\|p' - \pi\|^q$, but (ii) is added. The set C does not appear.

Again we specify the condition for level sets S (1.3) and $\pi = 0$. If $(\bar{p}, \bar{x}) = (0, \bar{x})$ and $f(\bar{x}) = 0$ (as used for calmness [q]), (4.1) claims

$$\begin{aligned} \forall x \in B(\bar{x}, \varepsilon) \text{ with } 0 < f(x) \leq \delta \text{ there is some } x' \text{ satisfying} \\ \lambda d(x', x) \leq f(x)^q \text{ and } f(x') \leq (1 - \lambda)f(x). \end{aligned} \quad (4.2)$$

If $f(\bar{x}) = \delta > 0$ and $(\bar{p}, \bar{x}) = (\delta, \bar{x})$ (as for l.s.c. [q]), (4.1) claims the same.

Using (4.1), let us replace procedure S1 by a simpler one without involving global infima.

Procedure S2: Find any $(p_{k+1}, x_{k+1}) \in \text{gph } S$ such that

$$(i) \quad \lambda d(x_{k+1}, x_k) \leq \|p_k - \pi\|^q \text{ and } (ii) \quad \|p_{k+1} - \pi\| \leq (1 - \lambda) \|p_k - \pi\|. \quad (4.3)$$

Our basic lemma now attains the following form.

Lemma 4.1. *Suppose $\lambda \in (0, 1)$ and (4.3) for any sequence of (p_k, x_k) , $k \geq 1$ (not necessarily in $\text{gph } S$). Then it holds, with $\theta = 1 - \lambda$ and $L = [\lambda(1 - \theta^q)]^{-1}$,*

$$d(x_{k+1}, x_1) \leq \sum_{i=1}^k d(x_{i+1}, x_i) \leq L \|p_1 - \pi\|^q, \quad (4.4)$$

after which $\xi = \lim x_k \in B(x_1, L \|p_1 - \pi\|^q)$ in the complete space X exists. Moreover, if $\varepsilon, \delta > 0$, $\pi, p_1 \in B(\bar{p}, \delta)$ and $d(x_1, \bar{x}) + \|p_1 - \pi\| + \|\pi - \bar{p}\|$ is small enough such that

$$d(x_1, \bar{x}) + L \|p_1 - \pi\|^q \leq \varepsilon \text{ and } \|p_1 - \pi\| + \|\pi - \bar{p}\| \leq \delta, \quad (4.5)$$

then all (p_k, x_k) satisfy (4.5). \diamond

Proof. If $p_k = \pi$ we have trivially $(p_{k+1}, x_{k+1}) = (p_k, x_k)$. Otherwise, (4.3)(ii) implies

$$\|p_{k+1} - \pi\|^q \leq \theta^q \|p_k - \pi\|^q.$$

Along with (i) this yields both $\lambda d(x_{k+1}, x_k) \leq \|p_k - \pi\|^q \leq (\theta^q)^{k-1} \|p_1 - \pi\|^q$ and

$$d(x_{k+1}, x_1) \leq \sum_{i=1}^k d(x_{i+1}, x_i) \leq \lambda^{-1} \sum_{i=1}^k (\theta^q)^{i-1} \|p_1 - \pi\|^q \leq L \|p_1 - \pi\|^q. \quad (4.6)$$

Thus convergence of (p_k, x_k) is ensured by (4.3) only. Finally, let (p_1, x_1) satisfy the remaining assumptions. The choice of the constants (4.5) then yields

$$\begin{aligned} d(x_{k+1}, \bar{x}) &\leq d(x_1, \bar{x}) + d(x_{k+1}, x_1) \leq d(x_1, \bar{x}) + L \|p_1 - \pi\|^q \leq \varepsilon, \\ \|p_k - \bar{p}\| &\leq \|p_k - \pi\| + \|\pi - \bar{p}\| \leq \|p_1 - \pi\| + \|\pi - \bar{p}\| \leq \delta. \end{aligned} \quad (4.7)$$

Hence the lemma is valid. \square

Remark 4.2. If the constants satisfy

$$L (2\delta)^q \leq \varepsilon/2 \quad \text{and} \quad \|\pi - \bar{p}\| \leq \delta/3. \quad (4.8)$$

then (4.5) holds for all (p_1, x_1) near (\bar{p}, \bar{x}) , namely if $x_1 \in B(\bar{x}, \varepsilon/2)$ and $p_1 \in B(\bar{p}, \delta/3)$. \diamond

Remark 4.3. Even if (p_k, x_k) , (p_{k+1}, x_{k+1}) are not in $\text{gph } S$, Lemma 4.1 ensures convergence

$$(p_k, x_k) \rightarrow (\pi, \xi), \quad \xi \in B(x_1, L \|p_1 - \pi\|^q).$$

To obtain $\xi \in S(\pi)$, it obviously suffices that points $(p'_{k+1}, x'_{k+1}) \in \text{gph } S$ exist with $d(x'_{k+1}, x_{k+1}) + \|p'_{k+1} - p_{k+1}\| \rightarrow 0$. This allows approximations with respect to $\text{gph } S$. \diamond

Theorem 4.4. *For the mapping S (2.11), suppose that $\lambda \in (0, 1)$, $\varepsilon, \delta > 0$ and some $\pi \in B(\bar{p}, \delta)$ satisfy (4.1). Then, if $(p_1, x_1) \in \text{gph } S$ and the remaining hypotheses of Lemma 4.1 are fulfilled, i.e., $p_1 \in B(\bar{p}, \delta)$ and*

$$d(x_1, \bar{x}) + L \|p_1 - \pi\|^q \leq \varepsilon \quad \text{and} \quad \|p_1 - \pi\| + \|\pi - \bar{p}\| \leq \delta, \quad (4.9)$$

procedure S2 defines an infinite sequence satisfying

$$\lim x_k = \xi \in S(\pi) \quad \text{and} \quad d(\xi, x_1) \leq L \|p_1 - \pi\|^q. \quad \diamond \quad (4.10)$$

Proof. By Lemma 4.1, we may apply the hypothesis (4.1) to (p_1, x_1) and all generated points (p_k, x_k) . Thus the points (p_{k+1}, x_{k+1}) in question exist in $\text{gph } S$, indeed. Evidently, (4.3)(ii) implies $p_k \rightarrow \pi$. Applying now $(p_k, x_k) \in \text{gph } S$ and closeness of S , the limit $\xi = \lim x_k$ belongs to $S(\pi)$. So nothing remains to prove. \square

The theorem allows us to replace the procedures and conditions in order to derive criteria for calmness and the Aubin property with geometrically decreasing $\|p' - \pi\|$ and the stepsize estimate (4.1)(i).

Corollary 4.5. *Suppose (2.11). Then*

(i) *S obeys the Aubin property $[q]$ at $(\bar{p}, \bar{x}) \Leftrightarrow$ there are $\lambda \in (0, 1)$ and $\varepsilon, \delta > 0$ satisfying (4.1) for all $\pi \in B(\bar{p}, \delta)$.*

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

Proof. Repeat the proof of Prop. 3.3. Necessity (\Rightarrow) follows again via Remark 3.1 while Thm. 4.4 and Remark 4.2 ensure the sufficiency. \square

The equivalence in terms of $S2$ can be written in a more convenient algorithmic manner.

Procedure S3: Given $(p_1, x_1) \in \text{gph } S$ put $\lambda_1 = 1$ and determine $(p_{k+1}, x_{k+1}) \in \text{gph } S$ satisfying

$$(i) \quad \lambda_k d(x_{k+1}, x_k) \leq \|p_k - \pi\|^q, \quad (ii) \quad \|p_{k+1} - \pi\| \leq (1 - \lambda_k) \|p_k - \pi\|. \quad (4.11)$$

If (p_{k+1}, x_{k+1}) exists put $\lambda_{k+1} := \lambda_k$. Otherwise $(p_{k+1}, x_{k+1}) := (p_k, x_k)$, $\lambda_{k+1} = \frac{1}{2}\lambda_k$. $k := k + 1$, repeat. \diamond

Having $\lambda_k \geq \lambda > 0$ and initial points near (\bar{p}, \bar{x}) , then both the convergence and the estimate (4.10) are ensured by Thm. 4.4. More precisely, involving also Prop. 3.3 and Cor. 4.5, we may thus summarize

Theorem 4.6. *Suppose (2.11) and $C = P$. Then*

- (i) *S obeys the Aubin property $[q]$ at (\bar{p}, \bar{x})*
 \iff *There exist $\lambda, \varepsilon, \delta > 0$ satisfying (2.13) for all $\pi \in B(\bar{p}, \delta)$.*
 \iff *There exist $\lambda \in (0, 1)$ and $\varepsilon, \delta > 0$ satisfying (4.1) for all $\pi \in B(\bar{p}, \delta)$.*
 \iff *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*

$$d(x_1, \bar{x}) + \|p_1 - \bar{p}\| + \|\pi - \bar{p}\| < \alpha \quad \text{and} \quad x_1 \in S(p_1). \quad (4.12)$$

\iff *There are $\alpha > 0$ and $\lambda \in (0, 1)$ such that $\lim \lambda_k \geq \lambda$ holds for all initial points of S3 satisfying (4.12).*

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

For $q = 1$ and less general spaces, the equivalence between the stability properties and the related behavior of S3 is known from [31, 32].

4.1 Calm C^1 systems

We derive two criteria for calmness (using Lemma 2.4 and Thm. 4.4) and show that calmness holds true iff a zero of the related system can be found by a simple (proper) algorithm. Let X be a Banach space,

$$S(p) = \{x \in X \mid g_i(x) \leq p_i, \ i = 1, \dots, m\}, \quad p \in \mathbb{R}^m \quad \text{and} \quad g_i \in C^1(X, \mathbb{R}). \quad (4.13)$$

To investigate calmness of S at $(0, \bar{x}) \in \mathbb{R}^m \times X$, we set

$$f(x) = \max_i \{0, g_i(x)\}. \quad (4.14)$$

Obviously, the level sets to f are calm at $(0, \bar{x}) \in \mathbb{R} \times X$ iff so is S at $(0, \bar{x}) \in \mathbb{R}^m \times X$. Put

$$I(x) = \{i \mid g_i(x) = f(x)\}, \quad F^+ = \{x \mid f(x) > 0\} \quad (4.15)$$

and let Ξ denote the family of all $J \subset \{1, \dots, m\}$ such that $J \equiv I(x)$ holds for some sequence of $x = x_k \xrightarrow{F^+} \bar{x}$. If $\Xi = \{\emptyset\}$, condition (4.16) holds trivially.

Theorem 4.7. *S is calm at $(0, \bar{x})$ if and only if*

$$\text{For all } J \in \Xi, \text{ there is some } u_J \in X \text{ satisfying } Dg_i(\bar{x})u_J < 0 \ \forall i \in J. \quad \diamond \quad (4.16)$$

Proof. By Prop. 3.4, verifying (proper) calmness means to find $\varepsilon, \lambda > 0$ such that, for each $x \in B(\bar{x}, \varepsilon) \cap F^+$, there is some x' such that

$$f(x') + \lambda d(x', x) < f(x). \quad (4.17)$$

Necessity of (4.16): Inequality (4.17) yields $d(x', x) \rightarrow 0$ as $x \rightarrow \bar{x}$ and

$$\frac{g_i(x') - g_i(x)}{d(x', x)} \leq \frac{f(x') - g_i(x)}{d(x', x)} < -\lambda \quad \forall i \in I(x). \quad (4.18)$$

Recalling $g_i \in C^1$ and setting $u_{x,x'} = \frac{x' - x}{d(x', x)}$ we have, if $\varepsilon < \varepsilon_0(\lambda)$ is small enough,

$$\left| \frac{g_i(x') - g_i(x)}{d(x', x)} - Dg_i(y)u_{x,x'} \right| < \frac{1}{2}\lambda \quad \forall i \in I(x) \ \forall y \in B(\bar{x}, \varepsilon). \quad (4.19)$$

Decreasing ε once more if necessary, also $I(x) = J \in \Xi$ has to hold. So we can assign, to each $J \in \Xi$, some descent direction u_J for all g_i ($i \in J$), by setting $u_J = u_{x,x'}$ with arbitrarily fixed

$x = \hat{x} \in B(\bar{x}, \varepsilon) \cap F^+$ satisfying $I(\hat{x}) = J$ and with any related x' , assigned to \hat{x} . This still yields, due to (4.18) and (4.19),

$$Dg_i(y)u_J < -\frac{1}{2}\lambda \quad \forall i \in J \quad \forall y \in B(\bar{x}, \varepsilon). \quad (4.20)$$

Thus (4.16) is necessary for calmness.

Sufficiency of (4.16): For $x \in F^+$ near \bar{x} , $I(x)$ coincides with some $J \in \Xi$ and (4.16) implies

$$\lambda := -\frac{1}{2} \max_{J \in \Xi} \max_{i \in J} Dg_i(\bar{x})u_J > 0.$$

Thus, for verifying (4.17), it suffices to choose $x' = x + t_x u_J$ with small $t_x > 0$ since the inequalities $g_j(x) < f(x)$ if $j \notin J$ remain true for x' if t_x is small enough. \square

Recalling that only sequences (3.4) are essential, Thm. 4.7 still holds if, given any $c > 0$, only sequences $x \rightarrow \bar{x}$ with $0 < f(x) < c d(x, \bar{x})$ are taken into account for defining the index sets J of Ξ .

Remark 4.8. There is a formally similar statement, namely Thm. 3 in [21] for $X = \mathbb{R}^n$: *S is calm at $(0, \bar{x})$ if, at \bar{x} , the Abadie Constraint Qualification (i.e., system $S(0)$ and its linearization have the same contingent cone; we refer to [35] for this condition) and (4.16) hold true.* In [21] however, the family Ξ consists of all J such that $J \equiv \{i \mid g_i(\xi_k) = 0\}$ holds for certain $\xi_k \rightarrow \bar{x}$, $\xi_k \in \text{bd } S(0) \setminus \{\bar{x}\}$. In addition, as noted in [32], then condition (4.16) is very strong and even not necessary for linear systems, consider

Example 2. $S(p) = \{x \in \mathbb{R}^2 \mid x_1 \leq p_1, -x_1 \leq p_2\}$. \diamond

Equivalent conditions: Clearly, condition (4.16) also means

$$\text{If } J \in \Xi \text{ then } S_J(p) = \{x \mid g_J(x) \leq p_J\} \text{ obeys the Aubin property at } (0_J, \bar{x}). \quad (4.21)$$

For the piecewise C^1 maximum function f and $X = \mathbb{R}^n$, this can be written in terms of several generalized derivatives of f at $x \in F^+$ near \bar{x} .

With Clarke's subdifferential and the directional derivative at $x \in F^+$ [7], it holds $\partial^c f(x) = \text{conv}\{Dg_i(x) \mid i \in I(x)\}$ and $f'(x; u) = \max \{\langle x^*, u \rangle \mid x^* \in \partial^c f(x)\}$. Thus (4.16) is also equivalent to the existence of some $\gamma > 0$ such that, for $x \in F^+$ near \bar{x} , it holds (3.18), i.e.,

$$\inf_{\|u\|=1} f'(x; u) \leq -\gamma \quad \text{or in dual formulation} \quad \inf_{x^* \in \partial^c f(x)} \|x^*\| \geq \gamma. \quad (4.22)$$

The procedures S2 and S3: To characterize calmness via S2 and S3, the relative slack

$$\sigma_i(x) = f(x)^{-1} (f(x) - g_i(x)) \quad \text{if } x \in F^+ \quad (4.23)$$

can be used in order to rewrite S3 in Thm. 4.6, as in [32], in form of a proper algorithm which claims to solve, in each step, linear inequalities with arguments in a ball:

Beginning with $\lambda_1 = 1$ and any x_1 , stop if $x_k \notin F^+$.

Otherwise find $u \in B$ such that $Dg_i(x_k)u \leq \lambda_k^{-1} \sigma_i(x_k) - \lambda_k \quad \forall i$. (4.24)

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

Theorem 4.9. *S (4.13) is calm at $(0, \bar{x})$ iff there are positive α and β such that, for all initial points $x_1 \in B(\bar{x}, \alpha)$, it follows $\lambda_k \geq \beta$ for procedure (4.24).* \diamond

Proof. Outline of the proof: The detailed proof in [32] makes use of the fact that, for the maximum f (4.14) under consideration, condition (4.1) can be written, with new $\lambda > 0$ and $x' = x + tu$, $u \in B$, $t > 0$ as

$$t^{-1}(g_i(x + tu) - g_i(x)) \leq (\lambda')^{-1} \sigma_i(x) - \lambda' \quad \forall i, \quad \lambda' f(x) \leq t \leq (\lambda')^{-1} f(x); \quad (4.25)$$

and conversely. Specifying $t = \lambda' f(x)$ this leads to the (equivalent) calmness characterization

$$\exists \lambda > 0 : \forall x \in F^+ \text{ near } \bar{x}, \text{ some } u_x \in B \text{ fulfills } Dg_i(x)u_x \leq \lambda^{-1} \sigma_i(x) - \lambda \quad \forall i. \quad (4.26) \square$$

Remark 4.10. Based on (4.26) and method (4.24), another calmness condition can be imposed: Given $x_k \rightarrow \bar{x}$ in F^+ , put $I_0\{x_k\} = \{i \mid \lim_{k \rightarrow \infty} \sigma_i(x_k) = 0\}$, and let Ξ_0 denote the family of all $J \subset \{1, \dots, m\}$ such that $J \equiv I_0\{x_k\}$ holds for such a sequence. With $\Xi_0 \supset \Xi$ in place of Ξ , then Theorem 4.7 holds again, cf. [19]. \diamond

Finally, by the same arguments as under Thm. 4.13 below, it follows from $\lambda_k \rightarrow 0$ that the sequence, generated by algorithm (4.24) with any initial point x_1 having a bounded level set $S(x_1)$, converges to some \hat{x} with $0 \in \partial^c f(\hat{x})$.

4.2 Hoelder calm C^2 systems, $q = \frac{1}{2}$.

We consider $S_h(p) = \{x \in \mathbb{R}^n \mid h_i(x) \leq p_i, i = 1, \dots, m\}$ at $(0, \bar{x})$, suppose $h(\bar{x}) = 0$, $h \in C^2$ and use the notations (4.15) with

$$g_i = \sqrt{\max\{0, h_i\}}, \quad f = \max_i g_i \quad \text{and} \quad H = \max_i h_i.$$

Now, calmness of S_h for $q = \frac{1}{2}$ and (proper) calmness of the level sets to f coincide. Since $h \in C^2$, the contingent derivative $C(\partial^c H)(\bar{x}, x^*)$ can be determined [40], [29]. It depends linearly on $Dh(\bar{x})$ and $D^2h(\bar{x})$ only. If $0 \in \partial^c H(\bar{x})$, now *injectivity* of $C(\partial^c H)$ at $(\bar{x}, 0)$ plays a role. Setting

$$|DH(x)| = \text{dist}(0, \text{conv}\{Dh_i(x) \mid i \in I(x)\}) = \min \{ \|x^*\| \mid x^* \in \partial^c H(x) \},$$

this injectivity requires nothing but the existence of some $K > 0$ such that

$$|DH(x)| \geq K \|x - \bar{x}\| \quad \text{for } x \text{ near } \bar{x}. \quad (4.27)$$

In particular, (4.27) holds if all Hessians $D^2h_i(\bar{x})$ are positive definite. For $m = 1$, (4.27) requires just regularity of $D^2H(\bar{x})$.

Theorem 4.11. *Using the above notations, S_h is calm at $(0, \bar{x})$ with $q = \frac{1}{2}$ if $0 \notin \partial^c H(\bar{x})$ or if (otherwise) the contingent derivative $C(\partial^c H)$ is injective at $(\bar{x}, 0)$. \diamond*

Put $h \equiv 0$ in order to see that the condition is not a necessary one.

Proof. If $0 \notin \partial^c H(\bar{x})$ even the (proper) Aubin property is satisfied for S_h . Hence let $0 \in \partial^c H(\bar{x})$. We investigate (proper) calmness for the level sets to f at $(0, \bar{x})$ via the calmness criterion (4.17). By Corollary 3.5, it suffices to consider only $x \rightarrow \bar{x}$ such that $0 < f(x) < \|x - \bar{x}\|$. Because all g_i , $i \in I(x)$ are C^1 near $x \in F^+$ with $Dg_i = \frac{1}{2}Dh_i/f$, we may first notice that calmness holds true if there is some $\lambda > 0$ such that

$$\text{some } u_x \in \text{bd } B \text{ fulfills } \frac{Dh_i(x)}{2f(x)} u_x \leq -2\lambda \quad \forall i \in I(x) \quad \text{if } x \in F^+ \text{ near } \bar{x}. \quad (4.28)$$

Indeed, (4.28) yields for sufficiently small $t = t_x > 0$ (since \sqrt{r} is concave),

$$g_i(x + tu_x) \leq g_i(x) + t \frac{Dh_i(x)}{2f(x)} u_x + o_{i,x}(t) < f(x) - \lambda t \quad \forall i \in I(x)$$

and implies (4.17). On the other hand, (4.28) is equivalent to

$$|DH(x)| \geq 4\lambda f(x) \quad \text{if } x \in F^+ \text{ near } \bar{x}.$$

Because of $\|x - \bar{x}\| > f(x)$, so already

$$|DH(x)| \geq 4\lambda \|x - \bar{x}\| \quad \text{if } x \in F^+ \text{ near } \bar{x} \quad (4.29)$$

is sufficient for calmness. The latter follows with $0 < \lambda < K/4$ from (4.27). \square

4.3 Procedure S3 for Hoelder calm C^2 and $C^{1,1}$ level sets

We study the case of $f \in C^2(\mathbb{R}^n, \mathbb{R})$ in order to demonstrate possible concrete steps of procedure S3 to find some ξ such that $f(\xi) \leq 0$ and $d(\xi, x_1) \leq L \max\{0, f(x_1)\}^q$. Clearly, we have to start near a zero \bar{x} of f and, as long as $f(x_k) > 0$, to find some x_{k+1} satisfying

$$(i) \lambda_k d(x_{k+1}, x_k) \leq f(x_k)^q, \quad (ii) f(x_{k+1}) \leq (1 - \lambda_k) f(x_k) \quad (4.30)$$

or to decrease $\lambda_k := \frac{1}{2}\lambda_k$. If $Df(\bar{x}) \neq 0$ even the Aubin property holds trivially. Hence let $Df(\bar{x}) = 0$. Additionally, suppose the Hessian $D^2f(\bar{x})$ to be regular.

Lemma 4.12. *Under these assumptions, calmness with $q = \frac{1}{2}$ is satisfied. The steps of S3 can be realized (for small λ_k) by using any $u \in B$ with $Df(x_k)u \leq -\rho \|Df(x_k)\|$ for fixed $\rho \in (0, 1)$ and setting*

$$x_{k+1} = x_k + tu, \quad t = \lambda_k^q f(x_k)^q. \quad \diamond \quad (4.31)$$

Proof. Calmness $[q]$ already follows from Thm. 4.11. By our hypotheses, there are $\delta, c, C > 0$ such that, for x_k near \bar{x} , $t > 0$ and $\|u\| \leq 1$,

$$f(x_k) = f(x_k) - f(\bar{x}) \leq C \|x_k - \bar{x}\|^2, \quad (4.32)$$

$$f(x_k + tu) - f(x_k) = t Df(x_k)u + o_{u,k}(t) \quad \text{where } |o_{u,k}(t)| \leq c t^2. \quad (4.33)$$

$$\|Df(x_k)\| \geq \delta \|x_k - \bar{x}\|. \quad (4.34)$$

By (4.31), condition (4.30)(i) holds true, and (4.30)(ii) becomes

$$t Df(x_k)u + o_{u,k}(t) \leq -\lambda_k f(x_k),$$

which is ensured if $\lambda_k^q f(x_k)^q Df(x_k)u + c\lambda_k^{2q} f(x_k)^{2q} \leq -\lambda_k f(x_k)$, i.e.,

$$Df(x_k)u \leq -(\lambda_k^{1-q} f(x_k)^{1-q} + c\lambda_k^q f(x_k)^q) = -\sqrt{\lambda_k} (1+c) \sqrt{f(x_k)}. \quad (4.35)$$

Taking (4.32) into account, (4.35) holds under the stronger condition

$$Df(x_k)u \leq -\sqrt{\lambda_k} (1+c) \sqrt{C} \|x_k - \bar{x}\|. \quad (4.36)$$

By (4.34), our specified $u \in B$ satisfies $Df(x_k)u \leq -\rho \|Df(x_k)\| \leq -\rho \delta \|x_k - \bar{x}\|$. Thus, (4.36) holds if $\sqrt{\lambda_k} \leq \rho \delta [(1+c)\sqrt{C}]^{-1}$. In consequence, the steps of S3 can be realized in the given manner and λ_k will not vanish. Hence Thm. 4.6 implies calmness $[q]$, too. \square

Explicitly, our settings allow to put, for the essential steps of S3,

$$x_{k+1} = x_k - \lambda_k^q f(x_k)^q Df(x_k) \|Df(x_k)\|^{-1}, \quad q = \frac{1}{2}. \quad (4.37)$$

Switching from $q = 1$ to $q = \frac{1}{2}$ (i.e. changing the stepsize rule) is possible according to the given estimates. For instance, choose any $\gamma > 0$ and apply S3 and (4.37) with $q = 1$ as long as $\|Df(x_k)\| \geq \gamma$ and with $q = \frac{1}{2}$ otherwise.

Applying Newton steps ? By the supposed regularity of $D^2f(\bar{x})$, the point \bar{x} is a regular zero of the gradient $g = Df$, and Newton steps for solving $g = 0$ are locally well-defined. Though Newton's method finds very fast an element in $S(0)$, the computed zero may be too far from the initial point in order to verify calmness $[q]$ of S by procedure S3.

Example 3. Let $f = x_1 x_2$. Given $q \in (0, 1]$ put $x = (s, s^m)$ where $s > 0$ is small and m fulfills $m+1 > 1/q$. Each Newton step, applied to the linear function $g = Df$, leads us to $x_{k+1} = \bar{x} = 0$. Since, for fixed $\beta > 0$ and small $s > 0$, we have $\beta d(\bar{x}, x) \approx \beta s > s^{q(m+1)} = f(x)^q$, the estimate (4.30)(i) cannot hold for all sufficiently small $\lambda_k > 0$ and x near \bar{x} . In other words, $\inf \lambda_k \geq \beta > 0$ is not true for all initial points of the form $x_1 = (s, s^m)$ in (4.12). \diamond

The case of $f \in C^{1,1}(\mathbb{R}^n, \mathbb{R})$:

We applied $f \in C^2$ only for obtaining (4.32), (4.33) and (4.34). These properties are still ensured if Df is locally Lipschitz, i.e., for so-called $C^{1,1}$ functions. Then (4.32) and (4.33) remain valid without additional assumptions. To obtain (4.34) (for some $\delta > 0$) it suffices to suppose that the contingent derivative $CDf(\bar{x})$ of Df at \bar{x} is injective, which replaces regularity of the Hessian $D^2f(\bar{x})$.

4.4 Arbitrary initial points

Next let us start $S3$ at any point $(p_1, x_1) \in \text{gph } S$. If $\lambda_k \geq \lambda > 0$ does not vanish, we have

$$(p_k, x_k) \rightarrow (\pi, \xi) \in \text{gph } S, \quad \xi \in B(x_1, L\|p_1 - \pi\|^q)$$

by Lemma 4.1. Otherwise, it holds $\lambda_k \downarrow 0$ and with $F = S^{-1}$,

$$0 < (1 - \lambda_k) \|p_k - \pi\| \leq \inf_p \{ \|p - \pi\| \mid p \in F(B(x_k, \lambda_k^{-1}\|p_k - \pi\|^q)) \} \quad (4.38)$$

where p_k realizes the infimum up to error $\lambda_k\|p_k - \pi\|$. We discuss this situation for our standard level sets under additional assumptions. The subsequent constant L depends on λ and q as in Lemma 4.1, $L = [\lambda(1 - \theta^q)]^{-1}$ where $\theta = 1 - \lambda$.

Theorem 4.13. *Let S (1.3) be given on a reflexive B -space X , $0 < q \leq 1$, $\pi = 0$ and $S(f(x_1))$ be bounded. Then, procedure $S3$, with start at x_1 and $p_k = f(x_k)$, determines some $\xi \in S(0) \cap B(x_1, L|f(x_1)|^q)$ if $\lambda_k \geq \lambda > 0$. Otherwise, the sequence $\{x_k\}$ has a weak accumulation point \hat{x} which fulfills $0 \leq f(\hat{x}) \leq f(x_1)$ and is stationary in the following sense: There are points z_k such that $d(z_k, x_k) \rightarrow 0$ and*

$$\liminf_{k \rightarrow \infty} \inf_{\|u\|=1} D_-f(z_k; u) \geq 0 \quad (\text{i.e., } Df(\hat{x}) = 0 \text{ if } f \in C^1(\mathbb{R}^n, \mathbb{R})). \quad \diamond \quad (4.39)$$

Proof. By Lemma 4.1, we have to assume that $\lambda_k \rightarrow 0$. Condition (4.38) yields

$$0 < (1 - \lambda_k) f(x_k) \leq \inf \{ f(x) \mid x \in B(x_k, \lambda_k^{-1}f(x_k)^q) \}. \quad (4.40)$$

Applying Ekeland's principle to f and $x_k \in X_k := B(x_k, \lambda_k^{-1}f(x_k)^q)$ with

$$\varepsilon_E := \varepsilon_k = \lambda_k f(x_k) \quad \text{and} \quad \alpha_E := \alpha_k = r_k \lambda_k^{-1} f(x_k)^q, \quad 0 < r_k < 1,$$

formula (3.12) ensures the existence of $z_k \in B(x_k, \alpha_k) \subset \text{int } X_k$ such that $(1 - \lambda_k) f(x_k) \leq f(z_k) \leq f(x_k)$ and, since $\varepsilon_E/\alpha_E = \rho_k := r_k^{-1} \lambda_k^2 f(x_k)^{1-q}$,

$$f(x') + \rho_k d(x', z_k) \geq f(z_k) \quad \forall x' \in X_k. \quad (4.41)$$

Setting $r_k = \lambda_k^{3/2}$ we obtain that both $\rho_k = \sqrt{\lambda_k} f(x_k)^{1-q}$ and $\alpha_k = \sqrt{\lambda_k} f(x_k)^q$ are vanishing (recall $0 < q \leq 1$). This implies $d(z_k, x_k) \rightarrow 0$. Finally, condition (4.39) follows from $\rho_k \rightarrow 0$ and (4.41). Since $S(f(x_1))$ is bounded and X is a reflexive, now x_k, z_k are bounded and there exists a common weak accumulation point \hat{x} of x_k and z_k in X . \square

The assumption of boundedness for $S(f(x_1))$ is natural for many applications, but cannot be deleted.

Example 4. For $f = e^x$, $q = 1$ and $r_k = \lambda_k^{3/2}$, (4.41) yields $Df(z_k) \leq \sqrt{\lambda_k}$ and $x_k, z_k \leq \ln(\sqrt{\lambda_k}) \rightarrow -\infty$. Hence convergence is not guaranteed if $S(f(x_1))$ is unbounded. \diamond

Let us add two consequences of the theorem.

Computing feasible points:

Clearly, if a stationary point \hat{x} as in Thm. 4.13 cannot exist (e.g., if f is convex on \mathbb{R}^n and $\inf f < 0$) then λ_k cannot vanish and procedure S3 determines necessarily some point $\xi = \lim x_k \in S(0) \cap B(x_1, L|f(x_1)|^q)$. For more details, if f is a maximum of C^1 functions, we refer to section 4.1 where we considered concrete steps of S3 under (4.24) by applying the relative slack ($q = 1$).

Computing stationary points in optimization problems:

Let

$$f = \max\{rg_0, g_1, \dots, g_m\} \quad (4.42)$$

for any $r > 0$ and functions $g_i \in C^1(\mathbb{R}^n, \mathbb{R})$ of an optimization problem

$$\min \{g_0(x) \mid x \in \mathbb{R}^n, g_i(x) \leq 0; i = 1, \dots, m\} \quad (4.43)$$

with optimal value $v > 0$. Evidently, under solvability, $v > 0$ can be arranged by replacing g_0 with $g_0 + C$ where C is a big constant. Then, because of $S(0) = \emptyset$, the first case cannot happen in Thm. 4.13. In consequence, the obtained point \hat{x} now satisfies by (4.39)

$$\max_{i \in I^0} Dg_i(\hat{x})u \geq 0 \quad \forall u \in \mathbb{R}^n \quad \text{where} \\ i \in I^0 \quad \text{if } (i > 0 \text{ and } g_i(\hat{x}) = f(\hat{x}) \text{ or } (i = 0 \text{ and } rg_0(\hat{x}) = f(\hat{x})).$$

Now apply standard arguments of optimization: Firstly (by the Farkas Lemma) the origin is a nontrivial and non-negative linear combination of the active gradients

$$0 = \sum_{i \in I^0} \gamma_i Dg_i(\hat{x}) \quad \text{where } \gamma_i \geq 0 \quad \text{and} \quad \sum_{i \in I^0} \gamma_i > 0.$$

If $0 \notin I^0$ or $\gamma_0 = 0$, so \hat{x} is stationary for the constraint function $f_{-0} = \max_{i>0} g_i$. Provided this was excluded by a regularity condition (like some extended MFCQ, imposed also for points outside the feasible set, i.e., $0 \notin \partial f_{-0}^c(x)$ for $f_{-0}(x) \geq 0$ in terms of Clarke's subdifferential or alternatively by convexity along with a Slater point as above), it follows via $\gamma_0 > 0$ that \hat{x} satisfies the Lagrange condition and is a stationary point of the penalty problem

$$\min_{x \in \mathbb{R}^n} g_0(x) + \frac{1}{r} \sum_{i>0} \max\{0, g_i(x)\}. \quad (4.44)$$

For its well-known relations to Karush-Kuhn-Tucker points under calmness of the constraints we refer to [7]. We only mention that, under calmness of the constraints at a local solution \hat{x} of (4.43), \hat{x} is also a local solution of (4.44) whenever $r > 0$ is small enough. This yields

Corollary 4.14. *If both the level set $S(f(x_1))$ to (4.42) is bounded and an extended MFCQ holds on $S(f(x_1))$ then procedure S3 (with any $0 < q \leq 1$) determines a stationary point \hat{x} to problem (4.43) whenever r is sufficiently small.*

Since we study a maximum of C^1 functions, the concrete simple steps of S3 under (4.24) (which require to solve linear inequalities only) can be again applied.

Using the more general assumptions of Thm. 4.13, the point \hat{x} can be similarly interpreted as a (weak) limit of approximations (due to the involved Ekeland term $\rho d(x', z)$) of such penalty points. If all Dg_i are even weakly continuous, \hat{x} has the same properties as just mentioned for $x \in \mathbb{R}^n$.

Even for non-convex problems in \mathbb{R}^n , the point \hat{x} is not necessarily a stationary point in the usual sense (2.4); consider the origin for $f(x) = \min\{x, x^2\}$. In terms of subdifferentials and for $X = \mathbb{R}^n$, one only obtains the weak optimality condition that the origin belongs to the so-called limiting Fréchet subdifferential of f at \hat{x} .

Acknowledgement

The author is indebted to Diethard Klatte and Jan Heerda for various helpful discussions and constructive comments concerning the subject of this paper.

References

- [1] W. Alt. *Lipschitzian perturbations of infinite optimization problems*. in A.V. Fiacco, editor: *Mathematical Programming with Data Perturbations*. M. Decker, New York, 7-21, 1983.
- [2] J.-P. Aubin and I. Ekeland. *Applied Nonlinear Analysis*. Wiley, New York, 1984.
- [3] J.F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer, New York, 2000.
- [4] J. V. Burke. *Calmness and exact penalization*. *SIAM J. Control Optim.*, 29: 493–497, 1991.
- [5] J. V. Burke and S. Deng. *Weak sharp minima revisited, Part III: Error bounds for differentiable convex inclusions*. Manuscript 2006, to appear in *Mathematical Programming*.
- [6] F.H. Clarke. *On the inverse function theorem*. *Pacific Journal of Mathematics*, 64: 97–102, 1976.
- [7] F.H. Clarke. *Optimization and Nonsmooth Analysis*. Wiley, New York, 1983.
- [8] R. Cominetti. *Metric regularity, tangent sets and second-order optimality conditions*. *Applied Mathematics and Optimization*, 21: 265–287, 1990.
- [9] S. Dempe. *Foundations of Bilevel Programming*. Kluwer Academic Publ., Dordrecht-Boston-London, 2002.
- [10] A. V. Dmitruk and A. Kruger. Metric regularity and systems of generalized equations. *J. Math. Anal. Appl.* 342 (2008), no. 2, 864–873. DOI: 10.1016/j.jmaa.2007.12.057 Preprint : Optimization Online
- [11] S. Dolecki and S. Rolewicz. *Exact penalties for local minima*. *SIAM J. Control Optim.*, 17: 596–606, 1979.
- [12] A. Dontchev, Local convergence of the Newton method for generalized equations. *Comptes Rendus de l'Académie des Sciences de Paris*, **332**, 327–331, 1996.
- [13] A. Dontchev and R.T. Rockafellar. *Characterizations of strong regularity for variational inequalities over polyhedral convex sets*. *SIAM Journal on Optimization*, 6:1087–1105, 1996.
- [14] I. Ekeland. *On the variational principle*. *Journal of Mathematical Analysis and Applications*, 47: 324–353, 1974.
- [15] F. Facchinei and J.-S. Pang. *Finite-Dimensional Variational Inequalities and Complementary Problems, Vol I and Vol II*. Springer, New York, 2003.
- [16] P. Fusek. *Isolated zeros of Lipschitzian metrically regular R^n functions*. *Optimization*, 49: 425–446, 2001.
- [17] H. Gfrerer. *Hoelder continuity of solutions of perturbed optimization problems under Mangasarian-Fromovitz Constraint Qualification*. In *Parametric Optimization and Related Topics*, eds. J. Guddat, H.Th. Jongen, B. Kummer and F. Nožička, Akademie-Verlag, Berlin, 113–124, 1987.
- [18] L.M. Graves. *Some mapping theorems*. *Duke Mathematical Journal*, 17: 11–114, 1950.
- [19] J. Heerda and B. Kummer. *Characterization of Calmness for Banach space mappings*. Preprint 06-26 of the Institut of Mathematics, Humboldt-University Berlin, 2006 (<http://www.mathematik.hu-berlin.de/publ/pre/2006/p-list-06.html>)

- [20] R. Henrion and J. Outrata. *A subdifferential condition for calmness of multifunctions*. *Journal of Mathematical Analysis and Applications*, 258: 110–130, 2001.
- [21] R. Henrion and J. Outrata. *Calmness of constraint systems with applications*. *Mathematical Programming B*, 104: 437–464, 2005.
- [22] A.D. Ioffe. *Metric regularity and subdifferential calculus*. *Russ. Mathematical Surveys*, 55: 501–558, 2000.
- [23] A.D. Ioffe. *On regularity concepts in variational analysis Preprint: Department of Mathematics Technion, Haifa, Israel 2006*.
- [24] A.D. Ioffe and J. Outrata. *On metric and calmness qualification conditions in subdifferential calculus. Preprint: Department of Mathematics Technion, Haifa, Israel 2006*.
- [25] A. King and R.T. Rockafellar. *Sensitivity analysis for nonsmooth generalized equations*. *Mathematical Programming*, 55: 341–364, 1992.
- [26] D. Klatté. *On the stability of local and global optimal solutions in parametric problems of nonlinear programming*. Seminarbericht Nr. 80 der Sektion Mathematik, Humboldt-Universität Berlin., 1–21, 1985.
- [27] D. Klatté. *On quantitative stability for non-isolated minima*. *Control and Cybernetics*, 23, 183–200, 1994.
- [28] D. Klatté and B. Kummer. *Constrained minima and Lipschitzian penalties in metric spaces*. *SIAM Journal Optim.*, 13 (2): 619–633, 2002.
- [29] D. Klatté and B. Kummer. *Nonsmooth Equations in Optimization - Regularity, Calculus, Methods and Applications*. Nonconvex Optimization and Its Applications. Kluwer Academic Publ., Dordrecht-Boston-London, 2002.
- [30] D. Klatté and B. Kummer. *Strong Lipschitz Stability of Stationary Solutions for Nonlinear Programs and Variational Inequalities*. *SIAM J. Optimization*, 16: 96–119, 2005.
- [31] D. Klatté and B. Kummer. *Stability of inclusions: Characterizations via suitable Lipschitz functions and algorithms*. *Optimization*, 55: Issue 5,6: 627–660, Oct. 2006.
- [32] D. Klatté and B. Kummer. *Optimization methods and stability of inclusions in Banach spaces*. to appear in *Mathematical Programming, Series B*.
- [33] B. Kummer. *An implicit function theorem for $C^{0,1}$ -equations and parametric $C^{1,1}$ -optimization*. *Journal of Mathematical Analysis and Applications*, 158: 35–46, 1991.
- [34] D. N. Kutzarova and S. Rolewicz. *On drop property for convex sets*. *Archiv der Mathematik*, Vol 56, Nr. 5 : 501–511, 1991.
- [35] W. Li. *Abadie's constraint qualification, metric regularity, and error bounds for differentiable convex inequalities*. *SIAM Journal on Optimization*, 7: 966–978, 1997.
- [36] L. Lyusternik. *Conditional extrema of functions*. *Mathematical Sbornik*, 41: 390–401, 1934.
- [37] B.S. Mordukhovich. *Complete characterization of openness, metric regularity and Lipschitzian properties of multifunctions*. *Transactions of the American Mathematical Society*, 340: 1–35, 1993.
- [38] B.S. Mordukhovich. *Variational Analysis and generalized differentiation. Vol. I: Basic Theory, Vol II: Applications Springer, Berlin, 2005*.
- [39] J. Outrata and M. Kočvara and J. Zowe. *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*. Kluwer Academic Publ., Dordrecht-Boston-London, 1998.
- [40] R.A. Poliquin and R.T. Rockafellar. *Proto-derivative formulas for basic subgradient mappings in mathematical programming*. *Set-Valued Analysis*, 2: 275–290, 1994.
- [41] S.M. Robinson. *Generalized equations and their solutions, Part I: Basic theory*. *Mathematical Programming Study*, 10: 128–141, 1979.
- [42] S.M. Robinson. *Strongly regular generalized equations*. *Mathematics of Operations Research*, 5: 43–62, 1980.
- [43] S.M. Robinson. *Some continuity properties of polyhedral multifunctions*. *Mathematical Programming Study*, 14: 206–214, 1981.
- [44] S.M. Robinson. *Variational conditions with smooth constraints: structure and analysis*. *Mathematical Programming*, 97: 245–265, 2003.
- [45] R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*. Springer, Berlin, 1998.