PLISKA

STUDIA MATHEMATICA BULGARICA

ПЛИСКА

БЪЛГАРСКИ МАТЕМАТИЧЕСКИ СТУДИИ

Pliska Stud. Math. Bulgar. **12** (1998), 141-190

PLISKA
STUDIA MATHEMATICA
BULGARICA

WELL-BEHAVIOR, WELL-POSEDNESS AND NONSMOOTH ANALYSIS

Jean-Paul Penot

We survey the relationships between well-posedness and well-behavior. The latter notion means that any critical sequence (x_n) of a lower semicontinuous function f on a Banach space is minimizing. Here "critical" means that the remoteness of the subdifferential $\partial f(x_n)$ of f at x_n (i.e. the distance of 0 to $\partial f(x_n)$) converges to 0. The objective function f is not supposed to be convex or smooth and the subdifferential ∂ is not necessarily the usual Fenchel subdifferential. We are thus led to deal with conditions ensuring that a growth property of the subdifferential (or the derivative) of a function implies a growth property of the function itself. Both qualitative questions and quantitative results are considered.

Keywords: Asymptotical well-behavior, conditioning, critical sequence, error bounds, gage, metrically well-set, minimizing sequence, nice behavior, Palais-Smale condition, Ptak function, quasi-inverse, stationary sequence, well-behavior, well-posed problem.

AMS subject classification: 90C30, 90C33

1 Introduction

In the present survey, we tackle the relationships of the notion of well-behavior of a function with the notion of well-posedness. In fact, we also consider variants of these notions and we introduce quantitative tools which measure nice-behavior or well-behavior. The main question can be roughly expressed as follows: given a differentiable function f on a Banach space, can one deduce growth properties of f from a known growth behavior of its derivative f'? As a simple example, consider a differentiable even function f on $\mathbb{R}$ such that for some $p > 1$ one has $f'(x) \geq px^{p-1}$ for $x > 0$; then one has $f(x) - f(0) \geq |x|^p$ for each $x \in \mathbb{R}$.

In fact, we consider functions which are not necessarily differentiable and we do not limit our study to the convex case as in Auslender [12], Auslender and Crouzeix [14], Auslender, Cominetti and Crouzeix [15], Cominetti [55], Lemaire [109], Angleraud [1], Dolecki-Angleraud [67].

The importance of this question for algorithmic purposes can be illustrated by the following situation (for more information about the use of conditioning for algorithmic questions, in particular about rates of convergence, we refer to [12], [14], [15], [56], [105], [106], [107], [109], [110], [138], [176], [198]). Suppose that in a numerical experience with a minimization algorithm for a differentiable function f on some Euclidean space one has the following outcomes

$$\begin{aligned} \text{for the iteration } k,\ 1 &\leq k < 10 \quad \|\nabla f(x_k)\| \text{ is of order } 10^{-1} \\ \text{for the iteration } k,\ 10 &\leq k < 100 \quad \|\nabla f(x_k)\| \text{ is of order } 10^{-2} \\ \text{for the iteration } k,\ 100 &\leq k < 1000 \quad \|\nabla f(x_k)\| \text{ is of order } 10^{-3} \\ \text{for the iteration } k,\ 1000 &\leq k \qquad\qquad \|\nabla f(x_k)\| \text{ is of order } 10^{-10}. \end{aligned}$$

Is it sensible to stop at iteration $k = 1000$? Such a question occurs when one does not know the value of the infimum of f but one is only able to compute the value of f and of its gradient at each iteration. In such a case, taking the magnitude of the gradient as a stopping rule is tempting. However, as is well known, such a test is not sensible, as the example of the one-variable function f given by $f(x) = 1 - \exp(-x^2)$ shows: for any sequence $(x_n) \to \infty$ one has $(f'(x_n)) \to 0$ but $(f(x_n)) \to 1 = \sup f$. This function is quasiconvex, but not convex. Nonetheless, it is known that even for convex functions there may exist critical (or stationary) sequences which are not minimizing (see [185], [14], [196] and below for appropriate definitions).

It is the purpose of the present paper to survey results linking conditioning, well-posedness and well-behavior. The methods we use are either techniques from convex analysis or tools from nonsmooth analysis. Our main aim consists in extending to non-convex, non differentiable functions results known under convexity or differentiability assumptions.

Since we consider functions with no convexity or differentiability properties, for the sake of versatility we use an unspecified subdifferential, owing to the facts that often a given problem imposes a particular space and that not all subdifferentials have nice properties in an arbitrary space. Also the nature of the functions involved may lead to a specific choice of a subdifferential (for instance, the Fenchel subdifferential in the convex case, a bunch of subdifferentials in the quasiconvex case; see [161] and its bibliography). Thus the choice of a subdifferential may be dictated by the problem at hand. Elementary facts about subdifferentials are recalled in section 3; we endeavour to give a simple approach which minimizes assumptions and the pre-requisites for the reader unfamiliar with the field.

Basic notions about well-posedness and well-behavior are recalled in sections 2 and 7. There we focus our attention on quantitative aspects. Links with the Palais-Smale condition are delineated, apparently for the first time, in section 6. This condition plays a key role in global analysis and nonlinear analysis (see [10], [40], [46], [57], [58], [61]-[63], [82], [95], [96], [99], [130], [134], [136], [137], [148], [188]–[191], [200], [206]...). Sections 4 and 5 deal with critical sequences and critical functions, along with some other classes of functions generalizing the class of convex functions introduced in [89] (see also [52], [88]).

Estimates for the measures of nice-behavior introduced in section 7 are displayed in section 8 under generalized convexity assumptions by using new or known subdifferentials from quasiconvex analysis.

The main part of the paper is section 9 in which three methods providing growth estimates in terms of the behavior of the derivative (or subdifferential) of the function are brought together and formulated in simple terms. This is a first step toward a comparison between different methods yielding akin results.

Section 10 is devoted to the study of growth rates introduced in the convex case in [14], [15]; we generalize the results obtained there and in [12], [55], [56] by relaxing the convexity assumptions and the assumptions on the space. The links with duality theory are not considered here, although a result from [157] dealing with the behavior of the subdifferential of the Legendre-Fenchel conjugate of the function is evoked in section 11; we refer to [1], [24], [29], [30], [55], [165], for such a matter. The links with perturbation theory are not treated either; we refer to [69], [122], [158] and to the lecture by T. Zolezzi in this conference. The relationships with Tikhonov regularization would also deserve developments outside the scope of the present paper; we refer to [4], [47]–[51] and their references. Section 11 and 12 are devoted to applications, namely the study of perturbations of minima and minimizer sets in section 11 and metric regularity and error bounds in section 12. Thus we are back to applications to algorithms.

2 Metrically well-set (or generalized well-posed) functions

Throughout we consider a lower semicontinuous (l.s.c.) function f taking its values in $\mathbb{R}^\bullet = \mathbb{R} \cup \{\infty\}$ defined on a Banach space X whose dual space is denoted by X^*. We set

$$\begin{aligned} \mathrm{dom} f &:= \{x \in X : \ f(x) < +\infty\} \\ m &:= m_f := \inf f(X). \end{aligned}$$

We denote by S (or S_f if there is any risk of confusion) the set of minimizers of f. Recall that a sequence (x_n) is said to be minimizing (for f) if $(f(x_n)) \to m$. For $r \in \mathbb{R}_+$ the sublevel set of f with height r is

$$S(r) := \{x \in X : f(x) \leq r\}.$$

If A is a subset of X and $x \in X$, $\varepsilon \in \mathbb{R}_+$ we set $d(x, A) := \inf_{a \in A} d(x, a)$, $U(A, \varepsilon) := \{x \in X : d(x, A) < \varepsilon\}$, $B(A, \varepsilon) := \{x \in X : d(x, A) \leq \varepsilon\}$. If $A = \{a\}$ we write $U(a, \varepsilon)$ (resp. $B(a, \varepsilon)$) instead of $U(A, \varepsilon)$ (resp. $B(A, \varepsilon)$); in particular, B_X (or B) stands for $B(0, 1)$. The Hausdorff excess of a subset C of X over another subset D is given by

$$e(C, D) := \sup\{d(x, D) : x \in C\}.$$

In [34] the following definition has been introduced as a variant of the classical notion of well-posed minimization problem in the sense of Tikhonov (see [33], [69], [180], [181] for this question) for which any minizing sequence converges to the unique minimizer. Another variant, also dropping the uniqueness assumption (but imposing that S is compact) had been studied before by Furi and Vignoli [76].

Definition 2.1 *A function $f : X \to \mathbb{R} \cup \{\infty\}$ is said to be metrically well-set (in short M-well-set and in symbols $f \in \mathcal{M}(X)$) or metrically generalized well-posed if its set of* *minimizers S is nonempty and if for any minimizing sequence (x_n) one has $(d(x_n, S)) \to 0$.*

This notion is quite natural. The quadratic case is an important illustration.

Example. (see [56] for the finite dimensional case and [206] Proposition 37.33 for the case A is an isomorphism) Suppose $A : X \to X^*$ is a symmetric continuous linear operator with closed range which is positive semi-definite. Then, for each b in the range $R(A)$ of A, the quadratic function f given by $f(x) := \frac{1}{2}\langle A(x), x\rangle - \langle b, x\rangle$ is M-well-set. In fact, for any $u \in A^{-1}(b)$ one has $S = N(A) + u$ where $N(A)$ is the kernel of A and A induces an isomorphism between the quotient space $X/N(A)$ onto $R(A)$, so that, by [44, p. 99], for some $\alpha > 0$ one has for each $x \in X$, with $v := x - u$,

$$f(x) - m = f(u + v) - f(u) = \frac{1}{2}\langle A(v), v\rangle \geq \alpha d(v, N(A))^2 = \alpha d(x, S)^2.$$

3 Subdifferentials

The definition we adopt here for the concept of subdifferential is a very loose one. Making that choice, we are able to consider notions which are not local or which do not necessarily coincide with the usual Fenchel subdifferential in the convex case. Such an enlargement is necessary in order to encompass cases which are useful when dealing with generalized convex functions for which specific concepts exist (see [161] for instance for a general survey). Among them, the Fenchel subdifferential and the Plastria's subdifferential, or lower subdifferential, play a key role. Let us recall that given a point x in the domain $\operatorname{dom} f$ of an extended real-valued function f on a normed vector space X with dual X^* the *lower subdifferential* $\partial^< f(x)$ of Plastria [174] (resp. the *Fenchel subdifferential* or global subdifferential) of f at x is the set of $x^* \in X^*$ such that

$$f(u) - f(x) \geq \langle x^*, u - x\rangle \qquad \forall u \in [f < f(x)] \ \ (\text{resp. } \forall u \in X)$$

where $[f < f(x)] = \{u \in X : f(u) < f(x)\}$. The lower subdifferential has been used for a number of purposes with quasi-convex functions, in particular for cutting planes algorithms. Recall that f is said to be *quasiconvex* if its sublevel sets $[f \leq r] := f^{-1}(]-\infty, r])$ are convex for each $r \in \mathbb{R}$.

There are of course different ways of presenting subdifferentials (see for instance [17], [92], [93], [156] and their references). As noticed by several authors, a unified approach is convenient: in such a way, specific constructions and special properties can be avoided.

Here we define a *subdifferential* ∂ as a mapping which associates to any extended real-valued function f on some normed space X (or to any f in some class of functions $\mathcal{F}(X)$, X belonging to some class $\mathcal{X}$ of normed vector spaces) and to any x in the domain of f a subset $\partial f(x)$ of X^* in such a way that if f attains at x its minimum then $0 \in \partial f(x)$. We do not assume that $\partial(c\|\cdot\|)(x) \subset cB_{X^*}$ for any $c > 0$, $x \in X$, a condition which is satisfied whenever ∂ coincides with the subdifferential of convex analysis on convex functions, a natural requirement; however, this restriction is not satisfied by important subdifferentials of quasiconvex analysis.

As mentioned above, we do not require that ∂ is *local* or *localizable* in the sense that $\partial f(x) = \partial g(x)$ when f and g coincide on a neighborhood of x. We will not impose calculus rules either. However, an additional property of subdifferentials will be used.

Definition 3.1 *A subdifferential ∂ is said to be variational on a class $\mathcal{F}(X)$ of extended real-valued functions on X if for any bounded below lsc function $f \in \mathcal{F}(X)$ and for any $\alpha, \varepsilon, \lambda, \rho > 0$ with $\varepsilon < \lambda\rho$ and for any $x \in X$ such that $f(x) \leq \inf f(X) + \varepsilon$ there exist $w \in B(x, \rho)$ and $w^* \in \partial f(w)$ such that $\|w^*\| \leq \lambda$, $f(w) \leq f(x) + \alpha$.*

This property is satisfied when X is a Banach space, $\mathcal{F}(X)$ is a class of differentiable functions and $\partial f(x) = \{f'(x)\}$ for $x \in X$ and $f \in \mathcal{F}(X)$ (see [71], [72]). In view of the Bronsted-Rockafellar' theorem [41], it is also satisfied if X is a Banach space, $\mathcal{F}(X)$ is a class of l.s.c. convex functions and ∂ is the Fenchel subdifferential. Both cases are encompassed in the case of the class $\mathcal{T}(X)$ of l.s.c. tangentially convex functions. Here f is said to be *tangentially convex* if for each $x \in \operatorname{dom} f$ the *Hadamard lower derivative* (or contingent derivative) of f at x given by

$$f'(x, u) := \lim \inf_{(t,v) \to (0_+, u)} \frac{1}{t}(f(x + tv) - f(x))$$

is a convex function of u; then we take ∂ to be the *Hadamard* (or contingent) *subdifferential* given by

$$\partial f(x) := \{x^* \in X^* : x^*(\cdot) \leq f'(x, \cdot)\}.$$

Observe that f is tangentially convex when, for some differentiable function g and some convex function h, f is of the form (a) $f = g + h$, or of the form (b) $f = g \circ h$ (with g defined on $\mathbb{R}$ and nondecreasing) or of the form (c) $f = h \circ g$ (with g of class C^1 from X to another Banach space Y such that $Y = g'(x)(X) - \mathbb{R}_+(\operatorname{dom} h - g(x))$ for each $x \in \operatorname{dom} f$). In the case of quasiconvex functions, using results in [161], one can find an adapted subdifferential which is variational.

More generally, as shown below, ∂ is variational whenever ∂ is reliable in the sense of [149], [155], [89]. This notion, which is a variant of the concept of trustworthiness due to Ioffe [91] can be defined as follows (see [149], [155], [160]).

Definition 3.2 *A triple $(X, \mathcal{F}(X), \partial)$, where X is a Banach space, $\mathcal{F}(X)$ is a set of functions on X, ∂ is a subdifferential, is said to be reliable (and any of its elements is said to be reliable if the other two are considered as given) if for any lower semicontinuous function $f \in \mathcal{F}(X)$, for any convex Lipschitzian function g on X, for any $x \in \operatorname{dom}(f)$ at which $f + g$ attains its infimum and for any $\varepsilon > 0$ one has*

$$0 \in \partial f(u) + \partial g(v) + \varepsilon B^*,$$

for some $u, v \in B(x, \varepsilon)$ such that $\mid f(u) - f(x) \mid < \varepsilon$.

Thus, this property holds whenever a fuzzy sum rule of weak type (even weaker than the one considered in [156]) is satisfied. In particular, it is satisfied for the Fréchet subdifferential in the wide class of Asplund spaces (whereas this case is not covered under the assumptions of [56]).

Lemma 3.3 *If ∂ is a reliable subdifferential for a class $\mathcal{F}(X)$ of l.s.c. extended real-valued functions on the Banach space X then it is variational.*

Proof. Let $f \in \mathcal{F}(X)$ be bounded below on X, let $\alpha, \varepsilon, \lambda, \rho > 0$ with $\varepsilon < \lambda\rho$ and let $x \in X$ be such that $f(x) \leq \inf f(X) + \varepsilon$. Let $\mu > 0$ be such that $\mu < \lambda$, $\varepsilon < \mu\rho$. The

Ekeland's variational principle (in the form given in [142]) yields some $v \in B(x, \mu^{-1}\varepsilon)$ such that $f(v) + \mu\|v - x\| \leq f(x)$ and

$$\forall u \in X \qquad f(v) \leq f(u) + \mu\|u - v\|.$$

As ∂ is reliable, we can find $w, z \in B(v, \rho - \mu^{-1}\varepsilon)$, $w^* \in \partial f(w)$, $z^* \in \partial(\mu\| \cdot \|)(z)$ such that $\|w^* + z^*\| \leq \lambda - \mu$, $f(w) \leq f(v) + \alpha \leq f(x) + \alpha$. Then $w \in B(x, \rho)$, $\|w^*\| \leq \lambda$ and the result is proved. $\square$

Reliability of a localizable subdifferential entails that the subdifferential is significant in this sense that the domain of subdifferentiability of any l.s.c. function is dense in its domain (see [160, Lemma 2.4]). The same property holds for a localizable variational subdifferential, with a similar proof. It would be interesting to know whether the two assumptions of localizability and significance and the following conditions are enough to ensure a mean value theorem:

$(S1)$ If f is convex then $\partial f(x)$ is the subdifferential in the sense of convex analysis;

$(S2)$ If f attains at x a local minimum then $0 \in \partial f(x)$;

$(S3)$ If h is linear and continuous then $\partial(f + h)(x) = \partial f(x) + h$.

These assumptions are natural requirements and are satisfied by usual subdifferentials including the Hadamard and the Fréchet subdifferentials. With these assumptions, localization and reliability do suffice (see [160] and, for the Fréchet case, see [114]; for another approach, see [17]).

Theorem 3.4 *Let ∂ be a localizable and reliable subdifferential satisfying conditions $(S1) - (S3)$ above and let $f : X \to \mathbb{R}^\bullet = \mathbb{R} \cup \{\infty\}$ be a l.s.c. function finite at $a, b \in X$. Then there exists $c \in [a, b[$ and sequences $(c_n), (c_n^*)$ such that $(c_n) \to c, (f(c_n)) \to f(c)$, $c_n^* \in \partial f(c_n)$ for each n and*

$$\liminf_n \langle c_n^*, b - a\rangle \geq f(b) - f(a). \tag{1}$$

When f is finite and continuous one can find $c \in]a, b[$ and sequences $(c_n), (c_n^*)$ such that $(c_n) \to c$, $c_n^* \in \partial f(c_n) \cup (-\partial(-f)(c_n))$ for each n and

$$\lim_n \langle c_n^*, b - a\rangle = f(b) - f(a). \tag{2}$$

4 Critical and minimizing sequences

The notion of subdifferential enables one to extend to nonsmooth functions the concept of critical sequence and of critical point. This has been done in a number of places, including [46] for the Clarke subdifferential and [89] for the general case.

Definition 4.1 *Given a subdifferential ∂ and a function f we say that a point x is a critical point of f if $0 \in \partial f(x)$. A real number r is said to be a critical value if there exists some critical point x such that $f(x) = r$. We say that a sequence (x_n) is critical or, more precisely, ∂-critical, if there exists a sequence (x_n^*) such that*

$$x_n^* \in \partial f(x_n) \text{ and } (x_n^*) \to 0.$$

Several authors rather use the term "stationary", but we prefer to keep this term for the case (x_n) is critical for f and $-f$.

The following proposition extends a recent result in [40], [148], [52]. It shows that for any minimizing sequence of f one can always find a "nearby" sequence which is both critical and minimizing. The proof adapted to our framework is even simpler than the ones [148], [52], [89].

Proposition 4.2 *Let X be a Banach space, let ∂ be a variational subdifferential and let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a bounded below l.s.c. function. Let (x_n) be a minimizing sequence for f. Then there exists a minimizing sequence (w_n) which is critical and such that* $\lim_{n\to\infty} \|x_n - w_n\| = 0$.

PROOF. Let $\varepsilon_n := f(x_n) - m$. Since the sequence (x_n) is minimizing, we have $(\varepsilon_n) \to 0_+$. When $\varepsilon_n = 0$ we take $w_n := x_n$. When $\varepsilon_n > 0$, the definition of a variational subdifferential, with $\alpha := \varepsilon_n$, $\lambda = \rho := 2\varepsilon_n^{\frac{1}{2}}$ yields some $w_n \in X$ and some $w_n^* \in \partial f(w_n)$ such that $\|x_n - w_n\| \leq 2\varepsilon_n^{\frac{1}{2}}$, $f(w_n) \leq f(x_n) + \varepsilon_n$, $\|w_n^*\| \leq 2\varepsilon_n^{\frac{1}{2}}$ and we get the required sequence. $\square$

Let us note that simple examples show that a minimizing sequence itself is not necessarily a critical sequence, even when the function is differentiable.

As explained in the introduction, the reverse question is more interesting.

The following simple result shows that for a convex function, or more generally, for an invex function, under assumptions which are mild when the space is finite dimensional, the discrepancy between critical and minimizing sequences may only arise with unbounded sequences. Recall that a function f is said to be *invex* (or ∂-invex if there is any risk of confusion) if any critical point of f is a minimizer: $0 \in \partial f(x) \Rightarrow f(x) = \inf f(X)$. See [117], [125], [172], [173], [166] for references and variations. Note that f is invex iff $0 \in \partial f(x) \Rightarrow 0 \in \partial^< f(x)$. Obviously, any pseudoconvex function is invex; in particular any convex function is invex. Here, as [85] (see also [123]-[124] (for the differentiable case), [166] (for the Clarke's subdifferential), [16], [18], [152] (in the general case)), f is said to be *pseudoconvex* if for any $x, y \in X$

$$\exists x^* \in \partial f(x) : \langle x^*, y - x\rangle \geq 0 \Rightarrow \forall z \in [x, y] \quad f(y) \geq f(z).$$

Clearly, for any invex function, local minimizers are global minimizers. Such a property is important and will be used later on. It has been shown in [78], [79] (see also [80]) that this last property is satisfied by integral functionals.

Lemma 4.3 *Let X be a reflexive Banach space and let f be an invex function on X such that the weak limit of a critical sequence is a critical point. Assume one of the following two assumptions*

(a) if a critical sequence (x_n) weakly converges to some x_∞ then $(f(x_n)) \to f(x_\infty)$.

(b) $\partial f(x) \subset \partial^< f(x)$ for each $x \in X$.

Then any bounded critical sequence of f is minimizing.

Observe that the assumption that the weak limit of a critical sequence is a critical point is satisfied whenever the graph of ∂f is sequentially closed in the product of the weak topology of X and the strong topology of X^*. In particular, it is satisfied if ∂ is the Fenchel subdifferential. Under our invexity assumption and condition (a), it is also satisfied if ∂ is the Plastria's lower subdifferential, as easily checked.

Proof. Let (x_n) be a bounded critical sequence. Since X is reflexive, any subsequence of (x_n) has a weakly converging subsequence whose limit x_∞ is a critical point by our closedness assumption. As f is invex, x_∞ a minimizer of f. The continuity assumption (a) ensures then that $(f(x_n)) \to f(x_\infty) = \inf f(X)$. When assumption (b) is satisfied we have either $f(x_n) = f(x_\infty)$ or $f(x_n) > f(x_\infty)$ and taking $x_n^* \in \partial f(x_n) \subset \partial^< f(x_n)$ with $(x_n^*) \to 0$ we see that

$$f(x_n) - f(x_\infty) \leq \langle x_n^*, x_n - x_\infty \rangle \to 0,$$

so that, in both cases, (x_n) is minimizing. $\square$

The preceding lemma can be combined with the following result which generalizes [55] Proposition 4.2 to the nonconvex case (see Proposition 5.2 below).

Lemma 4.4 *Suppose $\partial f(x) \subset \partial^< f(x)$ for each $x \in X$ and f is super-coercive in the sense $\liminf_{\|x\| \to \infty} f(x)/\|x\| > 0$. Then any critical sequence of f is bounded.*

Note that both assumptions are satisfied when f is convex and coercive (i.e. $f(x) \to \infty$ as $\|x\| \to \infty$); moreover, in such a case, one has a quantitative relation linking $r := \liminf_{\|x\| \to \infty} f(x)/\|x\|$ to the radius of the greatest open ball centered at 0 on which f^*, the Fenchel conjugate of f, is majorized (see [31], [37], [55], [74], [75], [157], [203]).

Proof. Given a critical sequence (x_n) let $(x_n^*) \to 0$ be such that $x_n^* \in \partial f(x_n)$ for each n. By assumption, there exist $r, s > 0$ such that $f(x) \geq s\|x\|$ whenever $\|x\| \geq r$. If (x_n) is unbounded, taking a subsequence if necessary we may assume $(\|x_n\|) \to \infty$ and that $f(x_n) > f(x_0)$ for each $n > 0$. Then, as $x_n^* \in \partial^< f(x_n)$ we have

$$\begin{aligned} f(x_0) - f(x_n) &\geq \langle x_n^*, x_0 - x_n \rangle, \\ \|x_n\|^{-1} f(x_0) - s &\geq -\|x_n\|^{-1} \|x_n^*\| (\|x_n\| - \|x_0\|), \end{aligned}$$

a contradiction. $\square$

The following criterion ensuring that a critical sequence is minimizing is rather special; it is a variant of [52, Theorem 6.3].

Proposition 4.5 *Let f be l.s.c. and bounded below on X. Suppose $\partial f \subset \partial^< f$. Under the following assumption (A), any critical sequence on which f is bounded is a minimizing sequence:*

(A) *for any $h > m := \inf f(X)$ there exist $k \in [m, h]$ and a nondecreasing function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ such that for each $x \in X$*

$$d(x, [f \leq k]) \leq \varphi((f(x) - k)_+).$$

Proof. Suppose on the contrary that there exists a critical sequence (x_n) such that $\sup_n f(x_n) < \infty$ and which is not minimizing. Taking a subsequence if necessary, we may assume that $(f(x_n))$ converges to some $\ell > m$ and moreover that for some $h \in]m, \ell[$ we have $\ell + 1 > f(x_n) > h$ for each n. Let k and φ be as in assumption (A). Without loss of generality, we may assume $k < h$. For each n we have $\varphi((f(x_n) - k)_+) \geq d(x_n, [f \leq k]) > 0$, so that we can find $u_n \in [f \leq k]$ such that

$$\|u_n - x_n\| \leq 2\varphi((f(x_n) - k)_+).$$

Let $x_n^* \in \partial f(x_n)$ be such that $(x_n^*) \to 0$. Since $\partial f(x_n) \subset \partial^< f(x_n)$ and $f(u_n) \leq k < h < f(x_n)$ for each n, we have

$$\langle x_n^*, u_n - x_n \rangle \leq f(u_n) - f(x_n) < k - h$$

hence

$$\begin{aligned} h - k &< \|x_n^*\| \, \|u_n - x_n\| \\ &\leq 2 \|x_n^*\| \, \varphi((f(x_n) - k)_+) \leq 2 \|x_n^*\| \, \varphi(\ell + 1 - k), \end{aligned}$$

a contradiction with $(\|x_n^*\|) \to 0$. $\square$

5 Well-behaved functions and critical functions

The following definition taken from [89] extends a notion introduced by Auslender [12], Auslender-Crouzeix [14] and Auslender-Cominetti-Crouzeix [15] in the convex case, with X finite dimensional.

Definition 5.1 *A function f on X into $\mathbb{R}^\bullet := (-\infty, +\infty]$ is said to be well-behaved if it is lower semicontinuous, bounded below and if its critical sequences are minimizing. We write $f \in \mathcal{W}$ or $f \in \mathcal{W}(X)$ or $f \in \mathcal{W}^\partial(X)$ if it is necessary to make clear the choice of ∂.*

As mentioned in the introduction, even a convex function may not be well-behaved, as shown by the following counter-example due to Rockafellar.

Example ([185]; see also [14], [196]) Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by $f(0,0) = 0$, $f(r,s) = r^2 s^{-1}$ for $s > 0$, $+\infty$ else. Then, for $x_n = (r_n, r_n^2)$ with $(r_n) \to \infty$ we have $f(x_n) = 1$ and $(f'(x_n)) \to 0$.

When f is bounded and $\partial f \subset \partial^< f$, assumption (A) of the preceding proposition ensures that f is well-behaved. Another case of well-behavior is contained in the following criterion which is an immediate consequence of Lemmas 4.3 and 4.4.

Proposition 5.2 *Suppose ∂ is contained in $\partial^<$. Under the assumptions of Lemma* 4.3, *if f is l.s.c., bounded below and super-coercive, then it is well-behaved.*

The following class of functions has been introduced in [89] as a class related to the class of well-behaved functions.

Definition 5.3 *A function $f : X \to \mathbb{R}^\bullet := (-\infty, +\infty]$ is said to be critical (for the subdifferential ∂) if for any critical sequence (x_n) the sequence $(f(x_n))$ of values converges in $\mathbb{R}$. It is said to be boundedly critical if for any bounded critical sequence (x_n) the sequence $(f(x_n))$ converges in $\mathbb{R}$.*

Obviously, any well-behaved function is critical. Moreover, it is easy to see that a boundedly critical function has at most one critical value. The function $f : \mathbb{R} \to \mathbb{R}$ given by $f(r) = -\exp r$ is critical but not well-behaved.

Example [89] Given a nonempty closed subset S of a Banach space X, let $f := d_S$ be the distance function to S. Then f is a well-behaved (hence critical) function for the

Fréchet subdifferential. Other examples can be given by applying the stability properties below or the criteria given in the next section.

The following result close to [89] Proposition 4.2 presents a criteria for the coincidence of the class of critical functions with the class of well-behaved functions.

Proposition 5.4 *Any well-behaved function is critical. A critical l.s.c. function bounded below on X is well-behaved if the subdifferential ∂ is variational or if the set S of minimizers of f is nonempty.*

PROOF. The first assertion has alredy been observed. Suppose f is critical, l.s.c. and bounded below and ∂ is variational. Let (x_n) be a critical sequence. We know from Proposition 4.2 that there exists a minimizing sequence (w_n) which is also a critical sequence. The sequence (z_n) given by $z_{2n} = x_n$, $z_{2n+1} = w_n$ is critical. Our assumption on f ensures that $(f(z_n))$ converges to $\lim f(w_n) = \inf f(X)$. Thus $(f(x_n)) \to \inf f(X)$.

When the set S is nonempty, any point of S is critical and in the preceding argument we can take for (w_n) the constant sequence with value such a point. $\square$

Although the class of critical functions is defined in a simple way, it does not enjoy good stability properties: the following properties require rather stringent assumptions.

Proposition 5.5 *Let $F : X \to Y$ be a continuously differentiable mapping between two Banach spaces. Suppose there exists $c > 0$ such that for any $v \in Y$ one has $\inf \left\{\|u\| : u \in F'(x)^{-1}(v)\right\} \leq c\|v\|$. Then, for any critical function g on Y such that $\partial(g \circ F)(x) \subset \partial g(F(x)) \circ F'(x)$ for each $x \in F^{-1}(\operatorname{dom} g)$, the function $f := g \circ F$ is critical.*

Proposition 5.6 ([89]) *Let $f = h \circ g$ where $g : X \to \mathbb{R}$ is critical and $h : \mathbb{R} \to \mathbb{R}$ is differentiable. Suppose there exists $c > 0$ such that $h'(r) \geq c$ for each $r \in \mathbb{R}$ and suppose $\partial f(x) \subset h'(g(x))\partial g(x)$ for each $x \in X$. Then f is critical.*

Let us observe that the relation $\partial f(x) \subset h'(g(x))\partial g(x)$ is satisfied by a number of subdifferentials such as the Fréchet and the Hadamard subdifferentials; if h is of class C^1 it is also satisfied by the Clarke subdifferential.

Stability results for sums are presented in [89].

Let us now consider briefly two classes of functions introduced in [89] for which the characterization of well-behavior in terms of sublevel sets can be extended.

Definition 5.7 *A function f from X into $\mathbb{R}^\bullet := \mathbb{R} \cup \{+\infty\}$ is said to be critically convex, in short C-convex, if it satisfies the following property: for any pair of critical sequences $(x_n), (x'_n)$ with $x_n \neq x'_n$ one has*

$$\lim_{n\to\infty} \frac{|f(x_n) - f(x'_n)|}{\|x_n - x'_n\|} = 0. \tag{3}$$

If the preceding property is required for bounded critical sequences only, we say that f is boundedly C-convex, in short BC-convex.

Again, for the Fréchet subdifferential, any distance function is C-convex. Other important examples of C-convex functions are given in the following lemmas. Let us first note an obvious relationship with the class of boundedly critical functions.

Proposition 5.8 *If f is BC-convex and has a bounded critical sequence (z_n) whose values are bounded, then f is boundedly critical.*

The following two results provide examples and justify the terminology.

Lemma 5.9 *If f is convex then it is C-convex. More generally, if $f := h \circ g$ with $h : Y \to \mathbb{R} \cup \{+\infty\}$ convex l.s.c., $g : X \to Y$ Lipschitzian of class C^1 and such that for some $c > 0$ one has for each $x \in X$ and each $v \in Y$ $\partial f(x) \subset g'(x)^T(\partial h(x))$ and* $\inf\{\|u\| : u \in X,\ g'(x)u = v\} \leq c\|v\|$, *then f is C-convex.*

PROOF. Let (x_n), (x'_n) be critical sequences of f with $x_n \neq x'_n$ for each n. For some $y_n^* \in \partial h(y_n)$, with $y_n = g(x_n)$, one has $(x_n^*) = (y_n^* \circ g'(x_n)) \to 0$, hence $(y_n^*) \to 0$, as for each $c' > c$ and each $v \in Y$ one can find $u \in X$ with $g'(x_n)u = v$, $\|u\| \leq c'\|v\|$, so that

$$y_n^*(v) = x_n^*(u) \leq \|x_n^*\| c' \|v\|.$$

Similarly, one can find $y_n'^* \in \partial h(y'_n)$, where $y'_n := g(x'_n)$, such that $(y_n'^*) \to 0$. Now

$$\begin{aligned} f(x'_n) - f(x_n) &= h(y'_n) - h(y_n) \geq \langle y_n^*, y'_n - y_n \rangle \\ &\geq -k\|y_n^*\| \|x'_n - x_n\|. \end{aligned}$$

where k is the Lipschitz rate of g. Thus

$$\|x'_n - x_n\|^{-1} \mid f(x'_n) - f(x_n) \mid \leq k \max(\|y_n^*\|, \|y_n'^*\|) \to 0. \qquad \square$$

Lemma 5.10 [89] *Any quadratic function is C-convex.*

Let us present now some stability results for the class of C-convex functions.

Proposition 5.11 [89] *Let $A : X \to Y$ be a surjective, continuous linear map between two Banach spaces. Then, for any C-convex function h on Y, the function $f := h \circ A$ is C-convex whenever $\partial f(x) \subset A^T(\partial h(x))$*

Proposition 5.12 [89] *Let $f = h \circ g$ where $g : X \to \mathbb{R}$ is C-convex and $h : \mathbb{R} \to \mathbb{R}$ is differentiable and Lipschitzian. Suppose there exists $c > 0$ such that $h'(r) \geq c$ for each $r \in \mathbb{R}$ and suppose $\partial f(x) \subset h'(g(x))\partial g(x)$ for each $x \in X$. Then f is C-convex.*

The following criteria appear in [89] with almost the same assumptions.

Proposition 5.13 *Suppose ∂ is reliable, X is a dual Banach space, f is l.s.c. and satisfies the following assumptions:*
(a) f is constant on the set Z of its critical points;
(b) if (z_n) is a critical sequence weak converging to some z, then $z \in Z$ and $(f(z_n)) \to f(z)$;*
(c) for any critical point z, any sequence (z_n) weak converging to z and any $z_n^* \in \partial f(z_n)$ one has $(z_n^*) \to 0$.*
Then f is BC-convex.

Corollary 5.14 *Suppose X is finite dimensional, f is of class C^1 and is constant on the set of its critical points. Then, (for the usual derivative) f is BC-convex.*

Proposition 5.15 *Suppose ∂ is reliable and suppose f is l.s.c. and satisfies the following assumptions:*
(a) f is BC-convex;
(b) if (z_n) is a critical sequence then $(f(z_n))$ is bounded;
(c) for any sequence (z_n) such that $(\|z_n\|) \to \infty$ and for any $z_n^* \in \partial f(z_n)$ one has $(z_n^*) \to 0$.
Then f is C-convex.

6 Nicely behaved functions and the Palais-Smale condition

One of the following two variants of well-behavior has been introduced by Lemaire [109] (essentially in the convex case and without the terminology we use here in order to avoid confusions).

Definition 6.1 *A function $f : X \to \mathbb{R} \cup \{\infty\}$ is said to be nicely behaved if its set of minimizers S is nonempty and if for any critical sequence (x_n) one has $(d(x_n, S)) \to 0$.*

It is said to be almost nicely behaved if S is nonempty and if this conclusion holds for any critical sequence (x_n) which is minimizing.

It is said to be very nicely behaved if S is nonempty and if for any critical sequence (x_n) one has both $(d(x_n, S)) \to 0$ and $(f(x_n)) \to m$.

We denote by $\mathcal{N}(X)$ (resp. $\mathcal{A}(X)$, resp. $\mathcal{V}(X)$) the set of lsc nicely behaved (resp. almost nicely behaved, resp. very nicely behaved) functions on X.

Obviously, with the preceding notation and the notations of Definitions 2.1 5.1 5 one has

$$\mathcal{V}(X) \subset \mathcal{W}(X) \cap \mathcal{N}(X) \subset \mathcal{N}(X) \subset \mathcal{A}(X) and \mathcal{M}(X) \subset \mathcal{A}(X).$$

The following propositions describe some other relationships between these notions and some other ones we introduced. The first one shows that for Lipschitzian functions well-behavior is a consequence of nice-behavior.

Proposition 6.2 *If f is nicely behaved and if f is uniformly continuous around S, then f is well-behaved. In particular, if S is compact, if f is nicely behaved and is continuous at each point of S, then f is well-behaved.*

Proof. Let $f \in \mathcal{N}(X)$. The uniform continuity of our assumption means that for each $\varepsilon > 0$ there exists $\delta > 0$ such that for each $x \in S$ and for each $w \in X$ satisfying $d(w, x) < \delta$ one has $| f(w) - f(x) | < \varepsilon$. It implies that for each critical sequence (x_n) we have $(f(x_n)) \to \min f$ since $(d(x_n, S)) \to 0$ as f is nicely behaved. □

Proposition 6.3 *Suppose ∂ is variational. If f is l.s.c., almost nicely behaved, then f is M-well-set: $\mathcal{A}(X) \subset \mathcal{M}(X)$, hence $\mathcal{N}(X) \subset \mathcal{M}(X)$.*

Proof. Let $f \in \mathcal{A}(X)$. Given a minimizing sequence (x_n) of f, using Proposition 4.2, we can find a sequence (w_n) which is critical, minimizing and such that $(d(w_n, x_n)) \to 0$. Since $f \in \mathcal{A}(X)$ we have $(d(w_n, S)) \to 0$. Thus $(d(x_n, S)) \to 0$. □

A partial converse is as follows.

Proposition 6.4 *If a critical function f is M-well-set or almost nicely behaved, then it is nicely behaved.*

PROOF. Let f be a critical function and let (x_n) be a critical sequence of f. Since the set S of minimizers of f is nonempty, (x_n) is minimizing by Proposition 5.4 and when f is almost nicely behaved or M-well-set we have $(d(x_n, S)) \to 0$. $\square$

Corollary 6.5 *If a well-behaved function f is M-well-set or almost nicely behaved then it is nicely behaved: $\mathcal{W}(X) \cap \mathcal{M}(X) \subset \mathcal{N}(X)$, $\mathcal{W}(X) \cap \mathcal{A}(X) \subset \mathcal{N}(X)$. Moreover, if ∂ is variational, one has $\mathcal{W}(X) \cap \mathcal{M}(X) = \mathcal{W}(X) \cap \mathcal{N}(X) = \mathcal{W}(X) \cap \mathcal{A}(X)$.*

In particular, for a variational subdifferential, and for a Lipschitzian function f, one has that f is nicely behaved iff f is well-behaved and M-well-set.

Now let us compare the notions of nice-behavior with a Palais-Smale condition. This condition is not exactly the original one due to Palais and Smale, but is a variant considered in a number of works [140], [40], [191], [63], [200]... in the differentiable case.

Definition 6.6 *The function f is said to satisfy the Palais-Smale condition (PS)$_c$ for the value c if any critical sequence (x_n) such that $(f(x_n)) \to c$ has a converging subsequence. We will write $f \in \mathcal{PS}$ or $f \in \mathcal{PS}(X)$ if this property holds for $c = m$, with $m := \inf f(X)$.*

The following observation is immediate.

Proposition 6.7 *If f is l.s.c. and satisfies the Palais-Smale condition (PS)$_m$, then the set S of minimizers of f is compact, nonempty and f is almost nicely behaved. Conversely, if the set S of minimizers of f is compact and if f is almost nicely behaved then f satisfies the Palais-Smale condition (PS)$_m$ for the value $m := \inf f(X)$.*

PROOF. The first assertion follows from the fact that a minimizer is a critical point. Let us prove the second one. Given a critical sequence (x_n) which is minimizing, any subsequence $(x_{k(n)})$ of (x_n) has a further subsequence $(x_{k(h(n))})$ which converges to some x_∞. As f is l.s.c. we have $f(x_\infty) \leq \lim f(x_{k(h(n))}) = m$, hence $x_\infty \in S$ and thus $(d(x_n, S)) \to 0$ by a classical argument about sequences: f is almost nicely behaved.

Conversely, suppose f is almost nicely behaved and S is compact and let (x_n) be a minimizing critical sequence. Taking any sequence (u_n) in S such that $d(u_n, x_n) \leq d(x_n, S) + 1/n$ and using the facts that $(d(x_n, S)) \to 0$ as f is almost nicely behaved, and that (u_n) has a converging subsequence, we see that (x_n) has a converging subsequence. $\square$

The following result has some similarity with Lemma 4.3. Here we say that f satisfies the strong form of the Palais-Smale condition if any critical sequence has a converging subsequence.

Proposition 6.8 *Let f be an invex function on X such that the graph of ∂f is (strongly) closed. Then, if f satisfies the strong form of the Palais-Smale condition it is nicely behaved.*

PROOF. Let (x_n) be a critical sequence. By assumption, it has a converging subsequence whose limit x_∞ is a critical point by the closedness of ∂f. Since f is invex, we have $x_\infty \in S$. Taking subsequences if necessary, we conclude that $(d(x_n, S)) \to 0$. $\square$

Proposition 6.9 *Any bounded below l.s.c. function f satisfying the Palais-Smale condition $(PS)_m$ is almost nicely behaved. If moreover f is well-behaved then it is nicely-behaved.*

PROOF. Let (x_n) be a critical sequence which is minimizing. Using the $(\mathrm{PS})_m$ condition we obtain that it has a subsequence which converges to some x_∞. As f is l.s.c. we get $f(x_\infty) \leq m$ or $x_\infty \in S$. Thus $(d(x_n, S)) \to 0$ as the original sequence can be replaced by an arbitrary subsequence. □

7 A quantitative approach

A quantitative approach can be adopted for the notions introduced above. Such a viewpoint is useful to deal with perturbation questions, as we will show later on. For this purpose, let us recall a piece of terminology and some useful tools. A *modulus* is a nondecreasing function on $\mathbb{R}_+$ which has limit 0 at 0 (see for instance [97, p. 356]). A function $\gamma : \mathbb{R}_+ \to \mathbb{R}_+$ is called a *gage* if it is nondecreasing and *firm* (or admissible or forcing) in the sense : any sequence (r_n) of $\mathbb{R}_+$ such that $\lim_n \gamma(r_n) = 0$ converges to 0. Here we note that for a nondecreasing function γ, firmness is equivalent to the property that γ is positive on the set $\mathbb{P}$ of positive numbers.

As an instance of a way one can quantify the preceding notions, let us present the following definition and lemma.

Definition 7.1 *Given a subdifferential ∂, a function f on X and its associated remoteness $\delta(x) := d(0, \partial f(x))$, the measure of C-convexity of f is the function on $\mathbb{R}_+$ given by*

$$\chi(r) := \chi_f(r) := \sup\left\{ \frac{\mid f(x) - f(y) \mid}{\|x - y\|} : x, y \in X, x \neq y, \delta(x) \leq r, \delta(y) \leq r \right\}.$$

The proof of the following lemma is easy.

Lemma 7.2 *The function f is C-convex iff χ_f is a modulus.*

Some of the preceding statements can be given a quantitative form by using the preceding measure of C-convexity. For instance, if f is a convex function, one has $\chi_f(r) \leq r$ for each $r \geq 0$. Indeed, for any $x, y \in \operatorname{dom} \partial f$ with $x \neq y$, $\delta(x) \leq r$, $\delta(y) \leq r$, for each $x^* \in \partial f(x)$, one has $f(y) - f(x) \geq -\|x^*\| \|y - x\|$ hence $f(x) - f(y) \leq r \|y - x\|$ and a similar inequality with x and y interchanged. If f is quadratic on a Hilbert space, $f(x) = \frac{1}{2}(Ax \mid x) - (b \mid x)$, one also has $\chi_f(r) \leq r$. In fact, whenever $x, y \in X$ satisfy $\|\nabla f(x)\| \leq r$, $\|\nabla f(y)\| \leq r$, by the mean value theorem, there exists some $t \in [0, 1]$ such that, for $z := (1 - T)x + ty$, one has

$$\begin{aligned} |f(x) - f(y)| &= |(\nabla f(z) \mid x - y)| = |(Az - b \mid x - y)| \\ &\leq ((1 - t) \|Ax - b\| + t \|Ay - b\|) \|x - y\| \leq r \|x - y\| . \end{aligned}$$

We will use the following two natural quasi-inverses of an element φ of the set N of nondecreasing functions on $\mathbb{R}_+$ which take the value 0 at 0 :

$$\varphi^e(s) = \inf\{r \in \mathbb{R}_+ : s \leq \varphi(r)\},$$

$$\varphi^h(s) = \sup\{r \in \mathbb{R}_+ : s \geq \varphi(r)\}.$$

When φ is strictly increasing, these two quasi-inverses coincide, but in general they are different and any ψ such that $\varphi^e \leq \psi \leq \varphi^h$ is a quasi-inverse of φ in the sense of [170]: $\varphi(r) < s \Rightarrow r \leq \psi(s),\ s < \varphi(r) \Rightarrow \psi(s) \leq r$.

These definitions are convenient to deal with a number of topics such as rearrangements (see for instance [65], [133], [179], [193]). We will use them for the study of conditioning and of well-behavior. Let us give a simple direct proof of the following result from [143] Proposition 2.8 which is a key fact for our purposes.

Lemma 7.3 *(a) If μ is an element of the set M of modulus, then $\gamma := \mu^e$ is an element of the set G of gages. The same is true for any other quasi-inverse.*

(b) If γ is an element of the set G of gages, then γ^h, as any quasi-inverse of γ, is an element of the set M of modulus.

Proof. (a) Given $\mu \in M$ and $s > 0$ there exists $\delta > 0$ such that $\mu(r) < s$ for any $r \in [0, \delta]$. Thus

$$\mu^e(s) = \inf\{r \in \mathbb{R}_+ : s \leq \mu(r)\} \geq \delta > 0.$$

(b) Given $\gamma \in G$ and given $\varepsilon > 0$, we have $\delta := \gamma(\varepsilon) > 0$, so that, for $s \in \mathbb{R}_+,\ s < \delta$ we get

$$\gamma^h(s) = \inf\{r \in \mathbb{R}_+ : s < \gamma(r)\} \leq \varepsilon.$$

Thus $\mu := \gamma^h$ is continuous at 0 and $\mu(0) = 0 : \mu \in M$. $\square$

The preceding lemma can be combined with the following result.

Lemma 7.4 *For any two functions φ, ψ from X to $\mathbb{R}_+ \cup \{\infty\}$, the following assertions are equivalent:*

(a) $\varphi(x) \to 0 \Rightarrow \psi(x) \to 0$;

(b) there exists a modulus μ such that $\psi \leq \mu \circ \varphi$

(c) there exists a gage γ such that $\gamma \circ \psi \leq \varphi$.

Proof. The implication (b)⇒(a) is obvious. Assuming (a), let

$$\mu(s) := \sup\{\psi(x) : \varphi(x) \leq s\}.$$

Then clearly μ is a modulus and $\psi \leq \mu \circ \varphi$: (b) holds.

Using the fact that for any φ one has $\varphi^e \circ \varphi \leq I$, the identity mapping, we see that (b)⇒(c), taking $\gamma := \mu^e$, so that $\gamma \circ \psi \leq \mu^e \circ \mu \circ \varphi \leq \varphi$, and using the fact that γ is a gage.

Now let us show that (c)⇒(a). Given $\varepsilon > 0$, for any $r < \delta := \gamma(\varepsilon)$, $r \geq 0$ we have $\psi(r) < \varepsilon$ whenever $\varphi(r) < \delta$ since when $\psi(r) \geq \varepsilon$ we have $\varphi(r) \geq \gamma(\psi(r)) \geq \gamma(\varepsilon) = \delta$. $\square$

From the preceding two observations, one can deduce interesting consequences for our purposes. A piece of terminology will ease our statements.

Given an extended real-valued function $f : X \to \mathbb{R}^\bullet$ such that $m := \inf f(X) \in \mathbb{R}$ and $S := \operatorname{Arg\,min} f$ is nonempty, a function $\mu : \mathbb{R}_+ \to \mathbb{R}_+^\bullet := \mathbb{R}_+ \cup \{\infty\}$ has been called a *conditioner* for f in [7], [157] if

$$\forall x \in X \quad d(x, S) \leq \mu(f(x) - m).$$

There exists a smallest conditioner and a smallest nondecreasing conditioner called the *canonical conditioner of* f. The latter is given by

$$\mu_f(t) := e(S(m+t), S) := \sup \{d(x, S) : x \in S(m+r)\}$$

where $S(r) := \{x \in X : f(x) \leq r\}$. We are interested in situations in which μ_f is a modulus. In general, it is not easy to determine a conditioner; it is usually easier to find a function $\gamma : \mathbb{R}_+ \to \mathbb{R}_+^\bullet$ such that

$$\forall x \in X \qquad \gamma(d(x, S)) \leq f(x) - m.$$

Such a function is called a *conditioning function* ([7]) or a *growth function*. The largest nondecreasing function γ satisfying this inequality (called the *canonical growth* (*or conditioning*) *function* of f) is the function γ_f given by

$$\gamma_f(r) = \inf \{f(x) - m : x \in X,\ d(x, S) \geq r\}.$$

When γ_f is a gage, we call it the *canonical conditioning gage* of f.

The following result which relates the canonical growth function γ_f to the canonical conditioner μ_f of f could be improved by substituting strict inequalities to inequalities in the definitions, but we refrain to do so.

Proposition 7.5 ([157, Lemma 2.2]) *The canonical growth function γ_f and the canonical conditioner μ_f of f are quasi-inverses.*

As a consequence, one disposes of the following characterization of metrically well-set functions ([157]; see also [67]).

Proposition 7.6 *For any function $f : X \to \mathbb{R}^\bullet := \mathbb{R} \cup \{\infty\}$ with finite infimum m and nonempty set of minimizers S the following conditions are equivalent:*

(a) f is metrically well-set;

(b) $\lim_{\varepsilon \to 0} e(S(m+\varepsilon), S) = 0$;

(c) f has a conditioner which is a modulus;

(e) the canonical conditioner μ_f of f is a modulus;

(f) the canonical growth function γ_f of f is a gage;

(g) there exists a growth function which is a gage.

Proof. The equivalence (a)⇔(b), (b)⇔(e), (c)⇔(e), (f)⇔(g) are obvious or simple reformulations. For the other equivalences, in Lemma 7.4 one can take $\varphi(x) = f(x) - m$, $\psi(x) = d(x, S)$. □

The following result partially proved in [157] (see also Lemma 9.8 below) completes [7, Prop. 5.2], [202, Prop. 2]: here S is not supposed to be a singleton and the convexity assumption is slightly relaxed. It will justify the next definition.

Proposition 7.7 *Suppose X is a normed vector space (n.v.s.), S is nonempty, m is finite, and f is convex (or starshaped at any $x \in S$) and metrically well-set. Then the* *function*

$$\rho_f(r) := \inf\{f(x) - m : x \in X,\ d(x,S) = r\}$$

coincides with the canonical conditioning function γ_f and is starshaped. Moreover the canonical conditioner μ_f of f is such that $-\mu_f$ is starshaped.

Part of this result remains under a relaxed convexity assumption.

Proposition 7.8 *Suppose X is a n.v.s., S is nonempty, m is finite, and f is quasiconvex. Then the function ρ_f is nondecreasing, hence coincides with the canonical conditioning function γ_f.*

Another striking property of the function $\rho := \rho_f$ is the following one, inspired by [40] Propositions 2, 3 and Lemma 1. Note that here S could be an arbitrary subset of X, not just the set of minimizers of f; in particular, S could be the set of critical points of f. Recall that a function $\rho : \mathbb{R} \to \mathbb{R}$ is said to be *strictly quasiconcave* on an interval $]a,b[$ if for any $r < s < t$ in $]a,b[$ one has $\rho(s) > \min(\rho(r),\rho(t))$.

Proposition 7.9 *Suppose ∂ is a reliable subdifferential. Suppose f is nicely behaved, or more generally, suppose f is l.s.c., S is nonempty and that for some $a,b \in [0,\infty]$, $a < b$ and any $r < t$ in $]a,b[$ one has*

$$\inf\{\|x^*\| : x \in X,\ r \le d(x,S) \le t,\ x^* \in \partial f(x)\} > 0.$$

Then the function $\rho := \rho_f$ defined above is strictly quasiconcave on $]a,b[$. In particular, there exists $c \in [a,b]$ such that ρ is increasing on $]a,c[$ and decreasing on $]c,b[$.

Proof. Suppose on the contrary that there exist positive numbers $r < s < t < \infty$ in $[a,b]$ such that $\rho(s) \le \min(\rho(r),\rho(t))$. Then one can find a sequence (x_n) in the set

$$F(s) := \{x \in X : d(x,S) = s\}$$

such that $f(x_n) - m \le \rho(s) + 1/n^2$. The Ekeland's principle in the form given in [142] Theorem A with $\gamma = 1/n$ yields some

$$y_n \in C := B(S,t)\backslash U(S,r) := \{x \in X : r \le d(x,S) \le t\}$$

such that

$$\begin{aligned} f(y_n) &\le f(z) + \frac{1}{n}\|z - y_n\| \quad \forall z \in C, \\ f(y_n) &\le f(x_n) - \frac{1}{n}\|x_n - y_n\|. \end{aligned}$$

We cannot have $y_n \in F(r)$ for n large as otherwise we would have $\|x_n - y_n\| \ge s - r$ hence

$$\rho(r) \le f(y_n) - m \le f(x_n) - \frac{1}{n}\|x_n - y_n\| - m \le \rho(s) + \frac{1}{n^2} - \frac{1}{n}(s-r) < \rho(r)$$

for n large enough, a contradiction. Similarly, $y_n \notin F(t)$ for n large. Therefore $y_n \in C\backslash(F(r) \cup F(t)) \subset \mathrm{int}C$ and y_n is a local minimizer of $f + \frac{1}{n}\|\cdot - y_n\|$. Since ∂ is reliable, there exists some critical sequence (z_n) with z_n so close to y_n that $z_n \in C\backslash(F(r) \cup F(t))$, a contradiction with our assumption.

In order to prove the last assertion, let us set

$$R := \{s \in]a,b[: \forall t > s \ \rho(t) < \rho(s)\}.$$

We observe that for any $r \in R$ and any $s > r$ we have $s \in R$ as otherwise we would find some $t > s$ with $\rho(t) \geq \rho(s)$, a contradiction with $\rho(s) < \rho(r)$ and the strict concavity of ρ. Let $c := \inf R$ and let us observe that what precedes shows that ρ is decreasing on $]c,b[$. Let us prove that ρ is increasing on $]a,c[$. If this is not the case, we can find some $r,s \in]a,c[$ such that $r < s$ and $\rho(r) \geq \rho(s)$. As $s \notin R$ we can find some $t > s$ such that $\rho(t) \geq \rho(s)$; this is another contradiction with the strict concavity of ρ. □

As shown in the previous propositions dealing with the convex case, one has sometimes a property stronger than well-setness.

Definition 7.10 *A function f on a metric space X is said to be very well-conditioned if m is finite, if the set of minimizers S is nonempty and if there exists a starshaped gage γ such that $f(\cdot) - m \geq \gamma(d(\cdot, S))$. The (nondecreasing) function $\widehat{\gamma}$ given by $\widehat{\gamma}(0) = 0$, $\widehat{\gamma}(t) := \gamma(t)/t$ is called a reduced growth function of f.*

The case in which a linear growth function can be found (hence a constant reduced growth function exists) is extremely important; see [56], [105], [107], [106], [138] and their references for characterizations and applications.

The following simple criteria is taken from [54] (in the case $p(t) = qt$ for some $q \in]0,1[$) and from Ptak [177] in the general case. Here we say that a nondecreasing function $p : \mathbb{R}_+ \to \mathbb{R}_+$ is a *Ptak function* if the series $s(t) := \sum_{n \geq 1} p^{(n)}(t)$ is convergent, where $p^{(0)}(t) = t$, $p^{(1)}(t) = p(t)$, $p^{(n)}(t) = p^{(n-1)}(p(t))$ and we say that s is the associated sum.

Lemma 7.11 *Let f be a bounded below real-valued function on a complete metric space X which satisfies the following assumptions for some $c > 0$ and some Ptak function p with associated sum s:*

(a) if $(x_n) \to x$ and $(f(x_n)) \to m := \inf f(X)$ then $f(x) = m$;

(b) there exist $c > 0$ such that for each $v \in X$ there exists $w \in X$ satisfying $d(w,v) \leq c(f(v) - m)$, $f(w) - m \leq p(f(v) - m)$.

Then the set S of minimizers of f is nonempty and $d(x,S) \leq cs(f(x) - m)$ for each $x \in X$. In particular, for $p(t) = qt$ with $0 < q < 1$ we have $f(x) - m \geq c^{-1}(1-q)d(x,S)$.

Proof. Without loss of generality we may suppose $m = 0$. Given $x \in X$ we define inductively a sequence by setting $x_0 := x$ and by associating to x_n (supposed to be already obtained) some $x_{n+1} \in X$ such that $d(x_{n+1}, x_n) \leq cf(x_n)$, $f(x_{n+1}) \leq p(f(x_n))$. Then we have $f(x_n) \leq p^{(n)}(f(x))$ for $n > 0$ and (x_n) is a Cauchy sequence, hence has a limit x_∞. By assumption (a) we have $x_\infty \in S$ and

$$d(x, x_\infty) \leq \sum_{n=0}^{\infty} d(x_n, x_{n+1}) \leq \sum_{n=0}^{\infty} cp^{(n)}(f(x)) = cs(f(x)). \qquad \square$$

Let us now deal with nice-behavior and well-behavior and introduce the following notions.

Definition 7.12 *The measure of niceness of f is the function $\nu := \nu_f : \mathbb{R}_+ \to \mathbb{R}_+^\bullet$ given by*

$$\nu(r) := \sup\{d(x,S) : x \in X,\ \delta(x) := d(0, \partial f(x)) \le r\}.$$

The niceness index of f is the function $\eta := \eta_f$ given by

$$\eta(s) := \inf\{\delta(x) := d(0, \partial f(x)) : x \in X,\ d(x,S) \ge s\}$$

Definition 7.13 *The measure of well-behavior of f is the function $\omega := \omega_f : \mathbb{R}_+ \to \mathbb{R}_+^\bullet$ given by*

$$\omega(r) := \sup\{f(x) - m : x \in X,\ \delta(x) := d(0, \partial f(x)) \le r\}.$$

The well behavior index of f is the function $\beta := \beta_f$ given by

$$\beta(s) := \inf\{\delta(x) := d(0, \partial f(x)) : x \in X,\ f(x) \ge m + s\}$$

Using again Lemma 7.4, but with $\varphi(x) := \delta(x) := d(0, \partial f(x))$, $\psi(x) := d(x,S)$ (resp. $\psi(x) := f(x) - m$), we obtain characterizations of nice-behavior and well-behavior.

Proposition 7.14 *For any function $f : X \to \mathbb{R}^\bullet$ with finite infimum m and nonempty set of minimizers S, the following conditions are equivalent :*
(a) f has a nice-behavior;
(b) $\delta(x) := d(0, \partial f(x)) \to 0 \Rightarrow d(x,S) \to 0$;
(c) there exists a modulus μ such that $d(\cdot, S) \le \mu(\delta(\cdot))$;
(e) the measure of niceness ν_f of f is a modulus ;
(f) the niceness index η_f of f is a gage ;
(g) there exists a gage γ such that $\delta(\cdot) := d(0, \partial f(\cdot)) \ge \gamma(d(\cdot, S))$.

Proposition 7.15 *For any function $f : X \to \mathbb{R}^\bullet$ with finite infimum m and nonempty set of minimizers S, the following conditions are equivalent:*
(a) f is well behaved;
(b) $\delta(x) := d(0, \partial f(x)) \to 0 \Rightarrow f(x) \to m$;
(c) there exists a modulus μ such that $f(x) - m \le \mu(\delta(\cdot))$;
(e) the measure of well-behavior ω_f of f is a modulus;
(f) the well-behavior index β_f of f is a gage ;
(g) there exists a gage γ such that $\delta(\cdot) := d(0, \partial f(\cdot)) \ge \gamma(f(\cdot) - m)$.

The quantities introduced above are not unrelated (see [164]). In fact, if λ_f is given by $\lambda_f(s) := \sup\{f(x) - m : d(x,S) \le s\}$, and if λ_f^e is its lower quasi-inverse, the following relations can be proved:

$$\begin{array}{rclcl} \nu_f & \le & \mu_f \circ \omega_f & \qquad & \beta_f \circ \gamma_f \le \eta_f \\ \omega_f & \le & \lambda_f \circ \nu_f & \qquad & \eta_f \circ \lambda_f^e \le \beta_f. \end{array}$$

Note that λ_f is a modulus if f is uniformly continuous.

We refer to [59] for another index of strong quasiconvexity which may have some relationships with the preceding indexes. Let us also note the following fact.

Lemma 7.16 *If $\widehat{\gamma}$ is a reduced growth function of f and if χ is the measure of C-convexity of f then one has $\widehat{\gamma} \circ d_S \le \chi \circ \delta$. If ν is the measure of niceness of f and if $\widehat{\gamma}$ is l.s.c. at $s = \nu(r)$, then one has $(\widehat{\gamma} \circ \nu)(r) \le \chi(r)$*

Proof. Since any element u of S is a critical point, for each $x \in X$ satisfying $\delta(x) = r$ and each $u \in S$ one has $f(x) - f(u) \leq \chi(r) \|x - u\|$ hence $f(x) - m \leq \chi(r) d(x, S)$ and $\widehat{\gamma}(d_S(x)) \leq \chi(r)$. When $\widehat{\gamma}$ is l.s.c. at $s = \nu(r)$, taking the supremum over $x \in \delta^{-1}([0, r])$ we get the second assertion. $\square$

8 Nice behavior in the quasiconvex and the convex cases

In the present section we do not make the blanket assumption that f is convex, but we examine some consequences of convexity or quasiconvexity assumptions and we use the classical subdifferentials of quasiconvex (and convex) analysis. Besides the Plastria's lower subdifferential $\partial^{<}$ which definition has been introduced in section 3 and the Gutiérrez *infradifferential* $\partial^{\leq}$ [84], we will use two other related subdifferentials which seem to be new but whose importances are not as great. Recall that *infradifferential* $\partial^{\leq} f(x)$ of f at x is the set of $x^* \in X^*$ such that

$$f(u) - f(x) \geq \langle x^*, u - x \rangle \quad \forall u \in S(f(x)) := [f \leq f(x)].$$

Definition 8.1 *The upper subdifferential of f at $x \in \operatorname{dom} f$ is the set $\partial^{>} f(x)$ of $y \in X^*$ such that*

$$\forall\, u \in [f > f(x)] \quad f(u) - f(x) \geq \langle y, u - x \rangle.$$

The supradifferential of f at $x \in \operatorname{dom} f$ is the set $\partial^{\geq} f(x)$ of $y \in X^*$ such that

$$\forall\, u \in [f \geq f(x)] \quad f(u) - f(x) \geq \langle y, u - x \rangle.$$

Thus, the Fenchel subdifferential of f is given by

$$\partial f(x) = \partial^{<} f(x) \cap \partial^{\geq} f(x) = \partial^{>} f(x) \cap \partial^{\leq} f(x).$$

Conditions ensuring that the lower subdifferential $\partial^{<} f(x)$ is nonempty are presented in [174], [127] and [161]. It would be interesting to devise criteria for the upper subdifferential or the supradifferential. The following examples show the upper and the lower subdifferentials of a quasiconvex function maybe large.

Examples. Let $f(x) = \min(x^2 - 1, 0)$ for $x \in X := \mathbb{R}$. Then $\partial^{>} f(1) = X^*$, $\partial^{\geq} f(1) = \{0\}$ while $\partial^{\leq} f(1)$ and the Fenchel subdifferential of f at 1 are empty and $\partial^{<} f(1) = [2, \infty[$. For f given by $f(x) = -(1 - x^2)^{1/2}$ for $x \in [-1, 1]$, $f(x) = 0$ otherwise, one has again $\partial^{>} f(1) = X^*$, $\partial^{\geq} f(1) = \{0\}$ but $\partial^{<} f(1) = \partial^{\leq} f(1) = \emptyset$. Note that these two functions are quasiconvex.

The following lemma and proposition complete [109] Lemme 3.1 and Proposition 3.1 and are our starting points in going outside the class of convex functions while using subdifferentials which may differ from the Fenchel subdifferential. Note that $\partial^{<} f(x) = X^*$ when $x \in S$, so that the first relation in the following statement cannot be valid for $x \in S$.

Lemma 8.2 *Suppose S is nonempty. Then, for the lower subdifferential $\partial^{<} f$ one has*

$$\begin{aligned} (4) \qquad \inf_{u \in S} \inf_{y \in \partial^{<} f(x)} \langle y, x - u \rangle \quad &\geq \quad f(x) - m, \quad \textit{for each } x \in (\operatorname{dom} \partial^{<} f\,) \backslash S \\ d(x, S) d(0, \partial^{<} f(x)) \quad &\geq \quad f(x) - m \quad \textit{for each } x \in \operatorname{dom} \partial^{<} f. \end{aligned}$$

Thus, for any subdifferential ∂ satisfying $\partial f \subset \partial^{<} f$, any bounded critical sequence is minimizing; in particular f is boundedly critical.

Proof. When $x \in \operatorname{dom} \partial^{<} f$, $x \notin S$, for any $u \in S$ and any $y \in \partial^{<} f(x)$ one has $f(u) < f(x)$, hence

$$f(x) - f(u) \leq \langle y, x - u \rangle,$$

so that the first inequality follows by taking the infima on $u \in S$ and $y \in \partial^{<} f(x)$. The second inequality ensues from the relation

$$\langle y, x - u \rangle \leq \|y\|.\|x - u\|$$

for $x \notin S$ and is obvious for $x \in S$. $\square$

The following consequence has some similarity with [67, Proposition 3] and [56, Theorem 5.2] which deal with the convex case.

Proposition 8.3 *If ∂ is contained in the lower subdifferential $\partial^{<}$, then any very well-conditioned function f is nicely behaved. Moreover, for any reduced growth function $\widehat{\gamma}$ of f, the niceness index η_f of f satisfies $\eta_f \geq \widehat{\gamma}$.*

Proof. Let $\widehat{\gamma}$ be a firm function such that for each $x \in X$

$$d(x, S)\widehat{\gamma}(d(x, S)) \leq f(x) - m.$$

For $x \in X \backslash S$, the second inequality of the preceding lemma yields

$$\widehat{\gamma}(d(x, S)) \leq d(0, \partial^{<} f(x)) \leq d(0, \partial f(x))$$

and the result follows from the definitions. $\square$

In the following statements we tackle a converse of the preceding result. Here we say that a subset D of a convex subset C of X is *radially dense* in C if, for any $x_0 \in C,\ x_1 \in D$, the set $D \cap [x_0, x_1]$ is dense in the segment $[x_0, x_1]$. This condition is satisfied if C has a nonempty interior int$C$ and $D :=$int$C$. If $f = g + i_C$ where i_C is the indicator function of C and g is a convex function on X continuous at each point of C, this condition is satisfied with $D = \operatorname{dom} \partial f$. Thus our statement contains [34] Prop. 4.6 and its generalization in [109] Prop. 4.2 in which $D = C = \operatorname{dom} f = \operatorname{dom} \partial f$. Recall that f is said to be *quasiconvex* if $f((1-t)x_0 + tx_1) \leq \max(f(x_0), f(x_1))$ whenever $t \in]0, 1[$, $x_0, x_1 \in X$; it is said to be *semi-strictly quasiconvex* if this inequality is strict when $f(x_0) \neq f(x_1)$ (see [19], [201], for example). Note that any convex function is semi-strictly quasiconvex and if $f = h \circ g$ with g semi-strictly quasiconvex and h increasing, then f is semi-strictly quasiconvex.

Proposition 8.4 *Let f be a l.s.c. quasiconvex function whose set of minimizers S is nonempty. Let ∂ be a subdifferential such that $\partial f \subset \partial^{\geq} f$. Suppose that there exists some subset $D \subset \operatorname{dom} \partial f$ containing S and radially dense in $C := \operatorname{dom} f$ and a nondecreasing function $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ such that the following relation holds for each $x \in D \backslash S$*

$$\inf_{u \in S} \sup_{y \in \partial f(x)} \langle y, x - u \rangle \geq \psi(d_S(x)). \tag{5}$$

Then, for $\widetilde{\psi}(t) := \sup_{0<c<1} c^{-1}(1-c)\psi(ct)$ *one has for each* $x \in X \backslash S$

$$f(x) - m \geq \widetilde{\psi}(d_S(x)). \tag{6}$$

If f *is semi-strictly quasiconvex one may substitute* $\partial^{>} f$ *to* $\partial^{\geq} f$ *in what precedes.*

Proof. The inequality (6) is obvious for $x \in X \backslash C$. Let $x \in C \backslash S$ and let (ε_n) be a sequence of positive numbers with limit 0. Let $r := d(x, S)$ and let (u_n) be a sequence of points of S such that $\|u_n - x\| = q_n r$, with $q_n \leq 1 + \varepsilon_n$. Given $c \in]0,1[$, for n large enough we can find $c_n \in [c + \varepsilon_n, c + 2\varepsilon_n] \subset [0,1]$ such that $w_n := (1-c_n)u_n + c_n x \in D$. Then, we have $c_n \geq c + \varepsilon_n \geq {q_n}^{-1}(c + q_n - 1)$. As $\|x - w_n\| = (1-c_n)\|x - u_n\| = (1-c_n)q_n r$ it follows that

$$\begin{aligned} d(w_n, S) &\geq d(x, S) - \|x - w_n\| \\ &\geq r - (1-c_n)q_n r \geq rc > 0. \end{aligned}$$

Our assumption yields some $y_n \in \partial f(w_n) \subset \partial^{\geq} f(w_n)$ (resp. $\partial^{>} f(w_n)$ when f is semi-strictly quasiconvex) such that

$$\langle y_n, w_n - u_n \rangle \geq \psi(d_S(w_n)) - \varepsilon_n \geq \psi(rc) - \varepsilon_n,$$

as ψ is nondecreasing. Moreover, by quasiconvexity (semi-strict quasiconvexity), we have $f(x) \geq f(w_n)$ (resp. $f(x) > f(w_n)$). Thus we get

$$\begin{aligned} f(x) - m &\geq f(x) - f(w_n) \geq \langle y_n, x - w_n \rangle \\ &\geq c_n^{-1}(1-c_n)\langle y_n, w_n - u_n \rangle \\ &\geq c_n^{-1}(1-c_n)(\psi(cd_S(x)) - \varepsilon_n). \end{aligned}$$

Passing to the limits we get

$$f(x) - m \geq c^{-1}(1-c)(\psi(cd_S(x)),$$

hence the announced inequality by taking the supremum on c. □

In the convex case, one can drop the assumption about the domain of ∂f provided one makes stronger assumptions (compare with [119, Theorem 5.12], [208, Theorems 6 and 7] and [34, Proposition 4.6]).

Proposition 8.5 *Suppose* X *is a dual Banach space and* f *is convex weak* l.s.c. with a nonempty set of minimizers* S. *Suppose that* ∂ *is the Fenchel subdifferential and that for some nondecreasing* $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ *one has for each* $x \in (\mathrm{dom}\partial f) \backslash S$

$$\inf_{u \in S} \inf_{y \in \partial f(x)} \langle y, x - u \rangle \geq \psi(d_S(x)). \tag{7}$$

Then, for $\widetilde{\psi}(t) := \sup_{0<c<1} c^{-1}(1-c)\psi(ct)$ *one has for each* $x \in X \backslash S$

$$f(x) - m \geq \widetilde{\psi}(d_S(x)).$$

Proof. Since S is a nonempty weak* closed convex subset of X, given $x \in X \backslash S$ there exists some $u \in S$ such that $\|x - u\| = d(x, S)$. Given $\overline{c} \in]0,1[$, let $c > \overline{c}$ with $c < 1$ and let $w := (1-c)u + cx$, $\overline{w} := (1-\overline{c})u + \overline{c}x$. The Bronsted-Rockafellar's theorem yields some $w_n \in B(w, \varepsilon_n) \cap D$ where D is the domain of ∂f and (ε_n) is a sequence with limit

0 in $]0,(c-\overline{c})\|x-u\|[$. One easily sees that $d(w,S)=\|w-u\|>\|\overline{w}-u\|=d(\overline{w},S)$ and that for n large enough one has

$$d_S(w_n)\geq d_S(w)-\|w-w_n\|\geq d_S(\overline{w}).$$

Moreover, using precisions of the Bronsted-Rockafellar's theorem displayed in [36], [171], [153] Prop. 1.1, we can find some $y_n\in\partial f(w_n)$ such that $|\langle y_n,w-w_n\rangle|<\varepsilon_n$. Then we have

$$\begin{array}{rcl} f(x)-m & \geq & f(x)-f(w_n)\geq\langle y_n,x-w_n\rangle \\ & \geq & \langle y_n,x-w\rangle-\varepsilon_n \\ & \geq & c^{-1}(1-c)\langle y_n,w-u\rangle-\varepsilon_n \\ & \geq & c^{-1}(1-c)\langle y_n,w_n-u\rangle-c^{-1}\varepsilon_n \\ & \geq & c^{-1}(1-c)\psi(d_S(w_n))-c^{-1}\varepsilon_n \\ & \geq & c^{-1}(1-c)\psi(d_S(\overline{w}))-c^{-1}\varepsilon_n \\ & \geq & c^{-1}(1-c)\psi(\overline{c}d_S(x))-c^{-1}\varepsilon_n. \end{array}$$

Taking the limit on n and the supremum on $c>\overline{c}$ we get

$$f(x)-m\geq\overline{c}^{-1}(1-\overline{c})\psi(\overline{c}d_S(x))$$

Since $\overline{c}$ is arbitrary in $]0,1[$, we get the result. □

Thus, for convex functions we have a wide circle of characterizations (see [56] Theorem 5.2). In the following statement which displays them, we denote by

$$h_S(y):=\sup_{x\in S}\langle x,y\rangle$$

the support function of the subset S of X (a positively homogeneous function extensively used by L. Hörmander) and we denote by N (resp. G, resp. H) the set of nondecreasing functions $\psi:\mathbb{R}_+\to\mathbb{R}_+$ (resp. the set of gages, resp. the set of functions ψ such that $t\mapsto\psi(t)/t$ is a gage).

Proposition 8.6 *Let f be a l.s.c. quasiconvex function on X whose set of minimizers S is nonempty. Let ∂ be the Fenchel subdifferential. If there exists a subset D of* dom∂f *radially dense in* dom$f$ *and containing S then the assertions (a)–(g) below are equivalent. If f is convex and X is reflexive then the assertions (a)–(c) are equivalent. Moreover, in the implications (a)⇒(b)...⇒(g) one can keep the same function ψ while in the implications (c)⇒(a), (g)⇒(a) one has to change ψ into $\widetilde{\psi}$ given by $\widetilde{\psi}(t):=\sup_{0<c<1}c^{-1}(1-c)\psi(ct)$.*

(a) there exists some ψ in N (resp. G, resp. H) such that for any $x\in X$ one has $f(x)-m\geq\psi(d_S(x))$;

(b) there exists some ψ in N (resp. G, resp. H) such that for any $x\in D$, $y\in\partial f(x)$, $u\in S$ one has $\langle y,x-u\rangle\geq\psi(d_S(x))$;

(c) there exists some ψ in N (resp. G, resp. H) such that for any $x\in D$, $y\in\partial f(x)$, one has $\langle y,x\rangle-h_S(y)\geq\psi(d_S(x))$;

(d) there exists some ψ in N (resp. G, resp. H) such that for any $x\in D$, there exists $y\in\partial f(x)$ such that $\langle y,x\rangle-h_S(y)\geq\psi(d_S(x))$;

(e) there exists some ψ in N (resp. G, resp. H) such that for any $x\in D$, one has $sup\{\langle y,x\rangle-h_S(y):y\in\partial f(x)\}\geq\psi(d_S(x))$;

(f) there exists some ψ in N (resp. G, resp. H) such that for any $x \in D$ one has $\sup_{y \in \partial f(x)} \inf_{u \in S} \langle y, x-u \rangle \geq \psi(d_S(x))$;

(g) there exists some ψ in N (resp. G, resp. H) such that for any $x \in D$, one has $inf_{u \in S} sup_{y \in \partial f(x)} \langle y, x-u \rangle \geq \psi(d_S(x))$.

Proof. The implication (a)⇒(b) is a consequence of the inequality

$$\langle y, x-u \rangle \geq f(x) - f(u).$$

The implication (b)⇒(c) follows by taking the infimum on $u \in S$ on both sides; the implications (c)⇒(d)⇒(e)⇒(f)⇒(g) are trivial.

The implication (c)⇒(a) (resp. (g)⇒(a)) has been proved in Proposition 8.5 (resp. in Proposition 8.4). □

9 Conditioning in the general case

In this section we gather three different techniques to get growth estimates which complete the criteria obtained in Lemma 7.11 and in the preceding section. The first one is inspired by [207] and [56] and relies on the concept of variational subdifferential (hence implicitely on the Ekeland's principle). The second one is related to what is known as deformation techniques. The third one is a simple but rough approach to the results of [67] using sublevel sets, Dini derivatives and the Zygmund lemma.

It will be convenient to formulate our first estimate by using the following operation (see [159]). Given a function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$, its *rounding* $\overset{\cap}{\varphi}$ is defined for $t \in \mathbb{R}_+$ by

$$\overset{\cap}{\varphi}(t) := \sup_{0<c<1} (1-c)\varphi(ct).$$

Such an operation (which would deserve more attention) is useful in the study of the geometry of Banach spaces (see for instance [73, Lemma2] which asserts that if ψ is a nonegative starshaped function then its biconjugate ψ^{**} satisfies $\psi^{**}(t) \geq \widetilde{\psi}(t) = t\overset{\cap}{\varphi}(t)$, where $\varphi(0) := 0$, $\varphi(t) := t^{-1}\psi(t)$ for $t > 0$). Observe that $\overset{\cap}{\varphi} \geq \frac{1}{2}\varphi(\frac{1}{2}\cdot)$, so that $\overset{\cap}{\varphi}$ is a gage when φ is a gage and if φ is starshaped one has $\overset{\cap}{\varphi} \leq \frac{1}{4}\varphi$. Moreover, when φ is nondecreasing, one has the following relations

$$\sup_{0<c<1} c^{-1}(1-c)\int_0^{ct} \varphi(s)ds \leq t\overset{\cap}{\varphi}(t) \leq \int_0^t \varphi(s)ds,$$

which can be used to compare the following estimates.

Theorem 9.1 ([207], [56]) *Suppose ∂ is variational on a class $\mathcal{F}(X)$ of l.s.c. extended real valued functions on X. Let $f \in \mathcal{F}(X)$ be l.s.c., bounded below and satisfying the following estimate for some $r \in \mathbb{R}_+$ and some nondecreasing function φ on $\mathbb{R}_+$*

$$\varphi(d_S(x)) \leq \|x^*\| \quad \text{for any } x \in B(S,r)\backslash S,\ x^* \in \partial f(x). \tag{8}$$

Then $\overset{\cap}{\varphi}$ is a reduced growth function for f on $B(S, r/2)$:

$$f(x) - m \geq \overset{\cap}{\varphi}(d_S(x))d_S(x) \quad \text{for any } x \in B(S, r/2).$$

Thus, any nicely behaved function in $\mathcal{F}(X)$ is metrically well-set: $\mathcal{N}(X) \cap \mathcal{F}(X) \subset \mathcal{M}(X)$. Moreover, the canonical reduced growth function $\widehat{\gamma_f}$ of f and the niceness index η_f of f satisfy $\widehat{\gamma_f} \geq \overset{\cap}{\eta}_f$.

Proof. Suppose, on the contrary, that there exists some $z \in B(S, r/2)$ such that

$$f(z) - m < \overset{\cap}{\varphi} (d_S(z)) d_S(z).$$

Then we have $d_S(z) > 0$ and we can find $c \in]0, 1[$ such that

$$f(z) - m < (1 - c)\varphi(cd_S(z))d_S(z).$$

Taking $\rho < (1-c)d_S(z),\ \lambda < \varphi(cd_S(z))$ such that $\lambda\rho > f(z)-m$, we can find $w \in B(z, \rho)$, $w^* \in \partial f(w)$ such that $\|w^*\| < \lambda$. It follows that

$$d_S(w) \geq d_S(z) - \|w - z\| \geq cd_S(z) > 0$$

and $d_S(w) \leq (2 - c)d_S(z) < r$, so that

$$\lambda \geq \|w^*\| \geq \varphi(d_S(w)) \geq \varphi(cd_S(z)),$$

a contradiction with our choice of λ. The last assertion is a consequence of the definitions and of the inequality $\gamma_f(s) \geq \overset{\cap}{\varphi} (s)s$ for $s \in \mathbb{R}_+$ and $\varphi = \eta$. $\square$

The preceding proof is close to proofs in [207] and [56]. However the growth function in our conclusion is different from the ones in these references and we do not make the assumption that φ is bounded. In fact, it is possible to enlarge the domain in which one disposes of a growth estimate for f (see [164]).

It is interesting to compare the preceding result with Propositions 8.4 and 8.5. We first note that, setting $\psi(t) = t\varphi(t)$, the conclusions of Theorem 9.1, Propositions 8.4 and 8.5 are the same since $\widetilde{\psi}(t) = t \overset{\cap}{\varphi} (t)$. Moreover, relation (5) (resp. (7)) made in Proposition 8.4 (resp. 8.5) implies

$$\varphi(d_S(x)) \leq \sup \{\|x^*\| :\ x \in X,\ x^* \in \partial f(x)\}$$

(resp.

$$\varphi(d_S(x)) \leq \inf \{\|x^*\| :\ x \in X,\ x^* \in \partial f(x)\}$$

which coincides with (8) when $r = \infty$). Thus, as the Fenchel subdifferential is variational in the class of convex functions, we can conclude that, in the convex case, Theorem 9.1 is a result better than Proposition 8.5 but that it cannot replace Proposition 8.4 inasmuch $\partial^{\geq}$ is not known to be variational.

In order to illustrate the preceding estimate, let us note that when $\varphi(t) := \kappa$ we have $\widetilde{\psi}(t) = t \overset{\cap}{\varphi} (t) = \kappa t$ and we get linear conditioning:

$$f(x) - m \geq \kappa d(x, S).$$

Conversely, when ∂ is contained in the lower subdifferential, by Proposition 8.3 above such a conditioning implies relation (8) with $\varphi(t) = \kappa t$. When $\varphi(t) = \kappa t$, we have $\widetilde{\psi}(t) = t \overset{\cap}{\varphi} (t) = \kappa t^2/4$ and the correspondence is not as accurate. The two methods we consider next will give in that case the better estimate

$$f(x) - m \geq \frac{1}{2}\kappa d(x, S)^2.$$

Now let us give a short account of a quite different technique known as the deformation method. The presentation we adopt differs from the ones in [40], [42], [58], [95], [96], [99] [178], but is clearly related to these contributions. However, we do not consider here the crucial passage from local notions to global ones, and as we limit our aims to a study of growth properties, we only retain simple ideas. A more complex approach would be required for other objectives such as the Ljusternik-Schnirelmann theory or the Morse theory. The framework is a general metric space (X, d), a l.s.c. function f on X, and a "critical" subset C of X associated to f; in this theory, C is not necessarily the minimizer set S of f, but we are primarily interested in the case $C = S$. For the sake of simplicity, we do not consider the case X is some open subset of some larger space, although such a localization is important (see [95], [96]).

In the following definitions, we endeavour to capture the essence of the deformation method.

Definition 9.2 *A stream for f on X is a mapping $h : \mathbb{R}_+ \times X \to X$ satisfying the following conditions:*

(a) for each $x \in X$ the functions $h(\cdot, x)$ and $f(h(\cdot, x))$ are continuous on $\mathbb{R}_+$;

(b) for each $x \in X$ one has $h(0, x) = x$.

If h is jointly continuous it is called a *deformation* of f or an homotopy for f. Homotopies are often defined on $[0, 1] \times X$ only, but any mapping defined on $[0, 1] \times X$ and satisfying there conditions (a) and (b) can be extended to $\mathbb{R}_+ \times X$ into a stream by setting

$$h(t, x) := h(t - n, h_1^{(n)}(x)) \qquad \text{for } n \leq t < n + 1,\ x \in X,$$

where $h_1 := h(1, \cdot)$. If h satisfies the semigroup property

$$h(t, h(s, x)) = h(s + t, x) \qquad \forall s, t \in \mathbb{R}_+\ \ \forall x \in X, \tag{9}$$

h is called a *flow* (or rather a *semiflow*). Usually, flows arise from solutions of differential equations or differential inclusions. For instance, if X is a Hilbert space, identifying X and its dual, one may wish to consider solutions of the differential inclusion

$$-\dot{u}(t) \in \partial f(u(t)), \quad u(0) = x. \tag{10}$$

Let us note that the reverse process of associating a functional to a vector field is a fruitful method in the study of dynamical systems known as the Liapunov method. Here we do not impose the stringent property (9) nor joint continuity. Instead we suppose given a nondecreasing function $\zeta : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\zeta(t) = 0$ implies $t = 0$ (i.e. ζ is a gage) and we say that h is a ζ-stream if for some countable subset D of $\mathbb{R}_+$ the following two conditions are satisfied for each $(t, x) \in (\mathbb{R}_+ \backslash D) \times X$:

(a) $v(t, x) := \liminf_{s \searrow 0} \frac{1}{s} d(h(t + s, x), h(t, x)) \geq \zeta(d(h(t, x), C))$;

(b) $D_+(-f \circ h(\cdot, x))(t) \geq v(t, x)\zeta(d(h(t, x), C))$,

where $D_+g(t) := \liminf_{s \searrow 0} s^{-1}(g(t + s) - g(t))$ is the lower right Dini derivative of the function g. We observe that the preceding conditions are intermediate between the conditions of [58], [95], [99] which are purely metric conditions and differential conditions such as the ones obtained by taking gradient or pseudo-gradient vector fields when X is a Finsler manifold (see [61], [130], [136], [184], [188], [197], [200], [206] for instance). A connection with the basic deformation lemmas of these works is the following elementary result which relies on the Zygmund's lemma ([192, Chapter 1], for instance).

Lemma 9.3 *Suppose f is bounded below and h is a ζ-stream for f on X. Then for any $0 < q < r$ and any $x \in X$ such that $d(x, C) = r$ there exists some $\tau := \tau(x) > 0$ such that*

(a) $d(h(\tau, x), C) = q, \quad d(h(t, x), C) > q$ for $t < \tau$;

(b) $f(x) - f(h(\tau, x) \geq \zeta(q) d(h(\tau, x), x) \geq \zeta(q)(r - q)$.

From this result it is easy to derive the following growth estimate.

Theorem 9.4 *Suppose that for some l.s.c. gage ζ and for each $c \in]0, 1[$ one can find a $c\zeta$-stream for f. Let $m := \inf f(X)$ be finite and let γ be given by $\gamma(t) := \int_0^t \zeta(s) ds$ be finite. Then the following growth estimate holds for each $x \in X$:*

$$f(x) - m \geq \gamma(d(x, C)).$$

In order to relate this result to more classical ones, let us suppose we are given a function $\sigma : X \to \mathbb{R}_+$ called the *index of criticality* of f (or the *slope* of f) such that for each $c \in]0, 1[$ one can find a stream h satisfying the following two conditions for any $(t, x) \in (\mathbb{R}_+ \backslash D) \times X$ where D is a countable subset of $\mathbb{R}_+$:

(a) $v(t, x) := \liminf_{s \searrow 0} \frac{1}{s} d(h(t + s, x), h(t, x)) \geq c\sigma(h(t, x))$

(b) $D_+(-f \circ h(\cdot, x))(t) \geq cv(t, x)\sigma(h(t, x))$,

If for instance X is a Banach space and ∂ is a subdifferential one may wish to take $\sigma(x) = \delta(x) := d(0, \partial f(x))$ as before; such a choice is known to be valid for the Clarke subdifferential $\partial^{\uparrow}$ for the slope used in [58], [61]–[62] and [95]. When X is a Hilbert space and f is a l.s.c. convex function, the differential inclusion

$$u'(t) \in -\partial f(u(t)), \;\; u(0) = x$$

is known to have a unique solution [39, Théorème 3.2]; moreover, the map $u(\cdot)$ has a right derivative u'_+ on $]0, \infty[$ which satisfies

$$\begin{aligned} -u'_+(t) &\in \partial f(u(t)), \; \|u'_+(t)\| = d\left(0, \partial f(u(t))\right), \\ -(f \circ u)'_+(t) &= \|u'_+(t)\|^2 = \|u'_+(t)\| d(0, \partial f(u(t)), \end{aligned} \tag{11}$$

and since

$$v(t, x) := \liminf_{s \searrow 0} \frac{1}{s} \|u(t + s) - u(t)\| = \|u'(t)\|$$

by continuity of the norm, the preceding conditions are fulfilled if one takes $\sigma(x) = d(0, \partial f(x))$. Taking for ∂ a subdifferential containing the Hadamard (or contingent) subdifferential and contained in the Clarke subdifferential, the preceding case has been extended to the case $f = g + h$, where g is of class C^1 (resp. of class C^1 with a Lipschitzian derivative) and h is convex l.s.c. in [190] (resp. [39, Proposition 3.12]), and, more generally, to the case $f = h \circ g$, with $g : X \to Y$ of class C^1, h convex l.s.c., with a natural qualification condition in [83]. Let us note that in each of the preceding cases the function f is tangentially convex,

$$f'(x, v) = \sup_{y \in \partial f(x)} \langle y, v \rangle,$$

where $f'(x, \cdot)$ is the lower derivative of f at x, so that relation (11) ensures that (by weak compactness of closed balls and Moreau's minimax theorem)

$$\inf_{\|v\| \leq \sigma(u(t))} f'(u(t), v) \quad = \quad \inf_{\|v\| \leq \sigma(u(t))} \sup_{y \in \partial f(u(t))} \langle y, v \rangle$$

$$\begin{aligned}
&= \sup_{y\in\partial f(u(t))} \inf_{\|v\|\le\sigma(u(t))} \langle y, v\rangle \\
&= \sup_{y\in\partial f(u(t))} -\|y\|\sigma(u(t)) = -\|u'_+(t)\|d(0,\partial f(u(t)) \\
&= (f\circ u)'_+(t) \ge f'(u(t), u'(t)),
\end{aligned}$$

so that $u'(t)$ does represent the steepest descent direction in $u(t)$.

Then the expected connection with what precedes can be stated as follows.

Corollary 9.5 *Suppose that for some l.s.c. gage ζ the slope σ satisfies $\sigma(x) \ge \zeta(d(x,C))$ for each $x \in X$. Then, for $m := \inf f(X)$ and γ given by $\gamma(t) := \int_0^t \zeta(s)ds$, the following growth estimate holds for each $x \in X$:*

$$f(x) - m \ge \gamma(d(x,C)).$$

Let us close this section by presenting a criteria for f to be M-well-set related to a result in [67] (which uses the quite different technique of nondiscrete induction of Ptak [177]) and in [164]. We need to introduce a notion of growth rate for f and to use the following simple lemma, the proof of which we give for the reader's convenience.

Lemma 9.6 *For any nonempty subsets A, B, C of a metric space one has*

$$e(C,A) \le e(C,B) + e(B,A).$$

Proof. As $e(B,A) := \sup_{b\in B} d(b,A)$, and as for any $c \in C,\ b \in B$ one has

$$d(c,A) \le d(c,b) + d(b,A),$$

switching $d(c,b)$ to the left hand side and taking the supremum over $b \in B$ we get

$$d(c,A) - d(c,B) \le e(B,A).$$

The result follows by a similar operation involving the supremum over $c \in C$. □

Let us introduce the *growth rate* of f as the function $g : [m,\infty[\to \mathbb{R}_+ \cup \{\infty\}$ given by

$$g(r) := \lim_{s\searrow r} \vartheta(r,s) = \sup_{s>r} \vartheta(r,s)$$

where, for $s > r \ge m$,

$$\vartheta(r,s) := \inf\left\{\frac{f(w) - r}{d(w, S(r))} : w \in S(s)\backslash S(r)\right\}.$$

We observe that for each $r \ge m$ the function $\vartheta(r,\cdot)$ is nonincreasing and that, when f is convex it is a constant function on $]r,\infty[$, as shown in the next lemma. We observe that the growth rate of f satisfies $g(r) \ge \vartheta(r) := \inf_{s>r} \vartheta(r,s)$ for each $r \ge m$, with equality when f is convex, in view of Lemma 9.8 below.

The assumptions of our criteria are rather strong, but the proof is simple. Given $m' > m$, these assumptions imply in particular that any local minimizer of f with value in $[m,m']$ is a global minimizer (see [204], [205] in this connection). In the following statement, given some $m' > m$ we say (as in [66, p. 159]) that a function $h : [m,m'] \to \mathbb{R}$ has a primitive H if the function h is (right) differentiable for each $r \in [m,m']\backslash D$, where D is at most countable and $H'(r) = h(r)$ for $r \in [m,m']\backslash D$; then we write

$H(r) = \int_m^{m+r} h(t)dt$, a notation consistent with the case h is regulated and the fact that H is determined up to a constant, so that we can impose $H(m) = 0$.

Proposition 9.7 *Suppose the function canonical growth function $\mu := \mu_f : r \mapsto e(S(m+r), S)$ is continuous on $[0, m' - m]$ for some $m' > m$ and there exists a function $h : [m, m'] \to \mathbb{R}$ which has a primitive and is such that $1/g \leq h$. Then, for each $r \in [0, m'-m]$ one has*

$$\mu_f(r) := e(S(m+r), S) \leq \int_m^{m+r} h(t)dt.$$

PROOF. Since the definition of $\vartheta(r, s)$ for $s > r$ ensures that

$$\vartheta(r,s)e(S(s), S(r)) \leq s - r,$$

the preceding lemma yields

$$\begin{aligned} D^{-}\mu(m+r) \quad &:= \liminf_{s \searrow r} \frac{\mu(m+s) - \mu(m+r)}{s-r} \leq \liminf_{s \searrow r} \frac{e(S(m+s), S(m+r))}{s-r} \\ &\leq \liminf_{s \searrow r} \frac{1}{\vartheta(r,s)} = \frac{1}{g(r)} \leq h(r). \end{aligned}$$

Using $\mu(r) = \mu(r) - \mu(m) \leq \int_m^{m+r} h(t)dt$ [192, Chapter 1] we get the announced estimate. □

The following lemma shows that the assumptions we made are sensible in the convex case.

Lemma 9.8 *Suppose f is convex and $r > m$ or $r = m$ and $S(r)$ $(= S)$ is nonempty. Then, for $r < s < t$, $\vartheta(r,s) = \vartheta(r,t)$, so that $g(\cdot) = c(\cdot)$ where*

$$c(r) := \inf \left\{ \frac{f(w) - r}{d(w, S(r))} : w \in X \backslash S(r) \right\}.$$

Moreover g is nondecreasing, hence on each interval on which g is positive the function $1/g$ is regulated and has a primitive.

PROOF. It suffices to prove that $\vartheta(r,s) \leq \vartheta(r,t)$. Let $z \in S(t) \backslash S(r)$ and let $u_n \in S(r)$ be such that $\|u_n - z\| \leq d(z, S(r)) + \varepsilon_n$, where $(\varepsilon_n) \to 0_+$. For $q = (t-r)^{-1}(s-r)$ and

$$w_n := (1-q)u_n + qz$$

we have $f(w_n) \leq (1-q)f(u_n) + qf(z) \leq s$. Moreover we have

$$d(w_n, S(r)) \geq \|w_n - u_n\| - \varepsilon_n;$$

otherwise we could find $v_n \in S(r)$ such that

$$\|w_n - v_n\| < \|w_n - u_n\| - \varepsilon_n$$

and we would get

$$\begin{aligned} \|z - v_n\| \quad &< \quad \|z - w_n\| + \|w_n - u_n\| - \varepsilon_n \\ \|z - v_n\| \quad &< \quad (1-q)\|z - u_n\| + q\|z - u_n\| - \varepsilon_n \leq d(z, S(r)), \end{aligned}$$

a contradiction. It follows that

$$\vartheta(r,s) \leq \frac{f(w_n) - r}{d(w_n, S(r))} \leq \frac{(1-q)r + qf(z) - r}{\|w_n - u_n\| - \varepsilon_n} \leq \frac{q(f(z) - r)}{q\|z - u_n\| - \varepsilon_n}$$

Taking limits on n and the infimum on z, we get $\vartheta(r,s) \leq \vartheta(r,t)$.

A proof of the fact that c is nondecreasing is given in [14], [15] for X finite dimensional, in [55] for X reflexive and in [149, Lemma 4.1], [164] in the general case. □

10 Auslender-Crouzeix's rates of well-behavior

The following behavior rates have been introduced (but the second one) by Auslender and Crouzeix in [14] in the convex case, with X finite dimensional. In order to extend them to the general case we have to use the duality multimapping $J : X \rightrightarrows X^*$ given by

$$J(x) := \left\{x^* \in X^* : \langle x^*, x\rangle = \|x\|^2,\ \|x^*\| = \|x\|\right\}.$$

When X is a Hilbert space and X is identified with its dual, then J is simply the identity mapping. In the nonconvex case, these rates which are functions defined below on the interval $[m, \infty[$, with $m := \inf f(X)$, depend on the choice of the subdifferential . However, this dependence does not appear in the notation we use, as long as there is no risk of confusion:

$$\begin{aligned}
a(r) &:= \inf\left\{\|x^*\| : x \in f^{-1}(r),\ x^* \in \partial f(x)\right\},\\
b(r) &:= \inf\left\{\|x^*\| : x \in f^{-1}([r,\infty[),\ x^* \in \partial f(x)\right\},\\
c(r) &:= \inf\left\{\frac{f(w)-r}{d(w,S(r))} : w \in X\backslash S(r)\right\},\\
k(r) &:= \inf\left\{f'(x, \frac{v}{\|v\|}) : x \in f^{-1}(r),\ v \in J^{-1}(\partial f(x)),\ v \neq 0\right\}.
\end{aligned}$$

Here $f'(x,\cdot)$ is the *Hadamard lower derivative* (or contingent derivative) of f at x defined in sections 3 and 9. The original definition of k involved instead the radial derivative of f. Under the assumption that $\mathbb{R}_+\mathrm{dom}f$ is a closed vector subspace, both derivatives coincide (see [45]). We note that $b(r) = \beta(r-m)$ for each $r \geq m$ and that when ∂f is contained in $\partial^< f$ it follows from Lemma 8.2 that for $r > m$

$$a(r) \geq \frac{r-m}{\mu_f(r-m)}.$$

Let us present some comparison between these quantities. Their proofs are often similar to the ones in [14, Lemma 2.1] or [55, Lemma 5.2 and Theorem 5.3] but they require some supplementary assumptions (as we do not assume convexity) and the following characterization of best approximations .

Lemma 10.1 *For a nonempty convex subset C of X and $w \in X$, $x \in C$ the following assertions are equivalent:*

(a) x is a metric projection of w in C, that is $\|x-w\| = d(w,C)$;

(b) $J(w-x) \cap N(C,x) \neq \emptyset$ where $N(C,x)$ is the normal cone to C at x given by

$$N(C,x) := \{x^* \in X^* : \forall u \in C\ \langle x^*, u-x\rangle \leq 0\}.$$

If f is a convex function and if $C := f^{-1}(\mathbb{R}_-)$, $w \in \mathrm{dom}f\backslash C$, $x \in f^{-1}(0)$ then the preceding assertions are satisfied whenever there exists some $r > 0$, $x^* \in \partial f(x)$ such that $r^{-1}x^* \in J(w-x)$. If f takes some finite negative value, this last condition is equivalent to (a) and (b).

Proof. The first part is proved in [167, Theorem 3.1]. A proof of the second part in the reflexive case is given in [55, Lemma 5.1] using duality arguments. Let us provide a simple direct proof of it. Since $\mathbb{R}_+\partial f(x)$ is obviously included in $N(C,x)$, it suffices to show that any element u^* of $J(w-x)\cap N(C,x)$ belongs to $]0,\infty[\partial f(x)$ when f takes some finite negative value. As $u^* \in N(C,x)$ iff $(u^*,0) \in N(E\cap(X\times\mathbb{R}_-),(x,0))$, where E is the epigraph of f and as the indicator function $i_{X\times\mathbb{R}_-}$ of $X\times\mathbb{R}_-$ is continuous at some point of E, by the Moreau-Rockafellar theorem one has

$$N(E\cap(X\times\mathbb{R})_-,(x,0)) = \partial(i_E + i_{X\times\mathbb{R}_-})(x,0) = \partial i_E(x,0) + \partial i_{X\times\mathbb{R}_-}(x,0).$$

Thus, there exists $s \geq 0$ such that $(u^*,-s) \in \partial i_E(x,0) = N(E,(x,0))$. If $s=0$, we have $u^* \in N(\mathrm{dom} f, x)$. Then, as $w \in \mathrm{dom} f$,

$$\|w-x\|^2 = \langle u^*, w-x\rangle \leq 0,$$

a contradiction with $w \in X\backslash C$. Thus, $s>0$ and for $r := s^{-1}$ we get $x^* := ru^* \in \partial f(x)$. $\square$

Remark. The preceding proof shows that if X is reflexive and if f is convex, finite and takes negative values, then, for $x \in f^{-1}(0)$ and $C := f^{-1}(\mathbb{R}_-)$ one has $N(C,x)\backslash\{0\} = \mathbb{R}_+\partial f(x)\backslash\{0\}$ (given $u^* \in N(C,x)\backslash\{0\}$ one can find $w \in X\backslash C$ such that $u^* \in J(w-x)$).

Let us start with the simplest comparison which is as follows. Here we use the *Hadamard (or contingent) subdifferential* of f at $x \in \mathrm{dom}\, f$ which is the set of $x^* \in X^*$ which are bounded above by the contingent derivative $f'(x,\cdot)$ of f at x.

Proposition 10.2 *Suppose $\partial f(x)$ is contained in the contingent subdifferential of f at x for each $x \in f^{-1}(r)$, where $r \geq m$. Then*

$$a(r) \leq k(r).$$

Proof. Let $x \in f^{-1}(r)$, $v \in J^{-1}(x^*)$, for some $x^* \in \partial f(x)$, $v \neq 0$ and let $u := \|v\|^{-1}v$. The definitions and the assumption yield

$$f'(x,u) \geq \langle x^*, u\rangle = \|x^*\| \geq a(r),$$

so that the result follows by taking the infimum on v and x. $\square$

Proposition 10.3 *Suppose f is quasiconvex and for some $r \geq m$ and each $x \in f^{-1}(r)$ one has $\partial f(x) \subset N(S(r),x)$ where ∂ is an arbitrary subdifferential. Then*

$$c(r) \leq k(r).$$

As observed in [161] the inclusion $\partial f(x) \subset N(S(r),x)$ is satisfied by all known subdifferentials of quasiconvex analysis when there is no local minimizer of f in $f^{-1}(r)$. It also holds without any assumption for the infradifferential $\partial^{\leq}$ of Gutiérrez [84], hence for the Fenchel subdifferential. Thus, if $r > m$ and if f is semi-strictly quasiconvex, that inclusion holds for the Plastria's subdifferential [161, Prop. 18].

Proof. Suppose f is quasiconvex, or, more generally, that $S(r)$ is convex. Let $x \in f^{-1}(r)$, $v \in J^{-1}(\partial f(x))$, $v \neq 0$ and let $u := \|v\|^{-1}v$. By definition of the Hadamard lower derivative, there exist sequences $(t_n) \searrow 0$, $(u_n) \to u$ such that

$$f'(x,u) := \lim_n t_n^{-1}\left(f(x+t_nu_n) - f(x)\right).$$

The inclusion $\partial f(x) \subset N(S(r), x)$ and Lemma 10.1 ensure that for each n the point x is a best approximation of $w_n := x + t_n u$ in $S(r)$. Thus

$$t_n = \|w_n - x\| = d(w_n, S(r)).$$

Let $w'_n := x + t_n u_n$, $t'_n := d(w'_n, S(r))$, so that, $d(\cdot, S(r))$ being Lipschitzian with rate one,

$$| t'_n - t_n | = | d(w'_n, S(r)) - d(w_n, S(r)) | \leq \|w'_n - w_n\| = t_n \|u_n - u\|$$

and $\left(t'_n t_n^{-1}\right) \to 1$. Thus $t'_n > 0$ for n large enough and

$$f'(x, u) = \lim_n \frac{f(w'_n) - f(x)}{t'_n} = \lim_n \frac{f(w'_n) - r}{d(w'_n, S(r))} \geq c(r).$$

Taking the infimum on x and v we get the result. $\square$

Proposition 10.4 *Suppose that for some $r \geq m$ and each $x \in X \backslash S(r)$ one has $\partial f(x) \subset \partial^{<} f(x)$, the lower subdifferential of Plastria. Then for each $r' > r$ one has*

$$c(r) \leq b(r_+) := \lim_{s \to r_+} b(s) \leq b(r') \leq a(r').$$

If for each $x \in f^{-1}([r, \infty[)$ one has $\partial f(x) \subset \partial^{\leq} f(x)$, the infradifferential of Gutiérrez, then one has

$$c(r) \leq b(r) \leq a(r).$$

Proof. Let $w \in f^{-1}([r, \infty[)$, $w^* \in \partial f(w)$ and let (x_n) be a sequence of $S(r)$ such that $(\|w - x_n\|) \to d(w, S)$. Then, as $f(x_n) \leq r < r' \leq f(w)$ and as $\partial f(w) \subset \partial^{<} f(w)$, we have

$$\begin{aligned} c(r) &\leq \lim_n \frac{f(w) - r}{\|w - x_n\|} \leq \limsup_n \frac{f(w) - f(x_n)}{\|w - x_n\|} \\ &\leq \limsup_n \frac{\langle w^*, w - x_n \rangle}{\|w - x_n\|} \leq \|w^*\|. \end{aligned}$$

The first assertion follows by taking the infimum over w and w^*. The proof of the second assertion is similar: when $\partial f(w) \subset \partial^{\leq} f(w)$ for each $w \in f^{-1}([r, \infty[)$ the preceding inequalities are valid in so far as $f(x_n) \leq f(w)$. $\square$

Let us consider a reverse inequality. The idea of considering truncations of the function f appears in [108] for more qualitative aims.

Proposition 10.5 *Suppose ∂ is local and variational in some class $\mathcal{F}(X)$ of l.s.c. functions on X. Then, given f and r such that $S(r) := [f \leq r]$ is nonempty (in particular for $r > m$) and $f_r := \max(f, r) \in \mathcal{F}(X)$, one has*

$$c(r) \geq b(r_+) := \inf_{r' > r} a(r') \geq b(r).$$

Note that the condition about the truncations of f is satisfied if $\mathcal{F}(X)$ is the class of convex functions or the class of quasiconvex functions on X.

Proof. Since $X \backslash S(r)$ is open and f_r coincides with f on $X \backslash S(r)$ we have $\partial f_r(x) = \partial f(x)$ for each $x \in X \backslash S(r)$. Then, by Theorem 9.1, the estimate $\|x^*\| \geq b(r_+)$ for each $x \in X \backslash S(r)$ and each $x^* \in \partial f(x)$ yields for each $x \in X$

$$f_r(x) - r \geq b(r_+) d(x, S(r)),$$

as $S(r)$ is the minimizer set of f_r. The result follows. $\square$

In order to get another inequality, we use the *upper subdifferential* $\partial^> f(x)$ of f at x, introduced above as the set of $x^* \in X^*$ such that

$$f(u) - f(x) \geq \langle x^*, u - x \rangle \qquad \forall u \in [f > f(x)].$$

It seems that this subdifferential has not been used yet and is not of much interest. However, it is larger than the Fenchel subdifferential ∂^F, so that condition (13) below is slightly more general than the one in which $\partial^F f(x)$ stands in place of $\partial^> f(x)$. In that case, relation (13) reduces to condition (12). Observe that if $r = \max f(X)$, condition (13) is trivially satisfied whatever f is (since in that case one has $\partial^> f(x) = X^*$). That would not be the case with the Fenchel subdifferential, as the following example shows: $X = \mathbb{R}$, $x = 1$, $f(u) = -\sqrt{1-u^2}$ for $u \in [-1, 1]$, $f(u) = 1$ for $\mid u \mid > 1$.

Hereafter we say that f is *radially usc on its domain* if for any $w, x \in \mathrm{dom} f$ the function $t \mapsto f((1-t)x + tw)$ is usc on $[0, 1]$. Let us observe that this property is satisfied if f is convex as for any $z \in [w, x]$ one has

$$f((1-t)z + tw) \leq (1-t)f(z) + tf(w),$$

hence $\mathrm{limsup}_{t \to 0_+} f((1-t)z + tw) \leq f(z)$ and similarly $\mathrm{limsup}_{t \to 0_+} f((1-t)z + tx) \leq f(z)$; see [161].

Proposition 10.6 *Suppose X is a dual Banach space, f is quasiconvex, l.s.c. for the weak* topology and radially usc on its domain.*
(a) Suppose that for some $r \geq m$ and for each $x \in f^{-1}(r)$

$$N(S(r), x) \subset \mathbb{R}_+(\partial^> f(x) \cap \partial f(x)). \tag{12}$$

Then one has

$$a(r) \leq c(r).$$

(b) If $f'(x, u) \leq \sup \{\langle x^*, u \rangle : x^* \in \partial^> f(x)\}$ and

$$N(S(r), x) \subset \mathbb{R}_+ \partial f(x) \tag{13}$$

for each $x \in f^{-1}(r)$ and each $u \in X$ then one has

$$k(r) \leq c(r).$$

Proof. Let $w \in X \backslash S(r)$. As $S(r)$ is weak* closed there exists a best approximation x to w in $C := S(r)$. Moreover, we have $f(x) = r$ since otherwise we would have $f(x') < r$ for some $x' \in [x, w]$, $x' \neq x$ by the radial upper semicontinuity of f and this would lead to $x' \in S(r)$, $\|x' - w\| < \|x - w\|$, a contradiction. In view of Lemma 10.1, we can find $u^* \in J(w - x) \cap N(C, x)$. As $\|u^*\| = \|w - x\| > 0$, relation (12) yields some $s > 0$ such that $x^* := s^{-1}u^* \in \partial^> f(x) \cap \partial f(x)$. As $f(w) > r = f(x)$, we can write

$$f(w) - f(x) \geq \langle x^*, w - x \rangle = \langle s^{-1}u^*, w - x \rangle = s^{-1}\|u^*\|.\|w - x\|$$

hence, as $x^* \in \partial f(x)$,

$$\frac{f(w) - r}{d(w, C)} \geq s^{-1}\|u^*\| = \|x^*\| \geq a(r).$$

Taking the infimum on $w \in X \backslash S(r)$ we get the announced inequality.

If the assumption of assertion (b) holds, setting $u := \|w - x\|^{-1}(w - x) = \|v\|^{-1}v$ with $v := s^{-1}(w - x)$ and observing that by relation (13) we can find $s > 0$ and $x^* \in \partial f(x)$

such that $u^* = sx^*$ so that $J(v) = s^{-1}J(w-x)$ contains x^*, the definitions of $\partial^> f(x)$ and k yield

$$\frac{f(w)-r}{d(w,C)} \geq \frac{1}{\|w-x\|} \sup_{z^* \in \partial^> f(x)} \langle z^*, w-x\rangle \geq f'(x,u) \geq k(r)$$

and the second assertion follows by taking the infimum on $w \in X \backslash S(r)$. □

Theorem 10.7 *Let a, b, c be the rates associated with an arbitrary function f. Among the following assertions:*

(a) $a(r) > 0$ for each $r > m$;

(b) $b(r) > 0$ for each $r > m$;

(c) $c(r) > 0$ for each $r > m$;

(d) f is well-behaved;

(f) any critical sequence on which f is bounded is minimizing

one always has (d)⇔(b)⇒(a), (d)⇒(f)⇒(a); if $\partial f \subset \partial^< f$ one has (c)⇒(d). If moreover the assumptions of Proposition 10.6 (a) hold, all these assertions are equivalent.

Proof. (b)⇒(d). If f is not well behaved, one can find $r > m$ and sequences (x_n) in $f^{-1}([r, \infty[)$, (x_n^*) in X^* such that $(x_n^*) \to 0$ and $x_n^* \in \partial f(x_n)$ for each n. Then one has $b(r) = 0$. (d)⇒(b). If $b(r) = 0$ for some $r > m$ there exist sequences (x_n) in $f^{-1}([r, \infty[)$, (x_n^*) in X^* such that $(x_n^*) \to 0$ and $x_n^* \in \partial f(x_n)$ for each n; thus f is not well-behaved. The implication (f)⇒(a) is similar. The implications (d)⇒(f) and (b)⇒(a) are obvious, as $a(r) \geq b(r)$.

(c)⇒(d) when $\partial f \subset \partial^< f$. If f is not well-behaved one can find $r > m$ and sequences (x_n), $(x_n^*) \to 0$ such that $f(x_n) > r$, $x_n^* \in \partial f(x_n)$ for each n. Then, as $\partial f(x_n) \subset \partial^< f(x_n)$, for each $u \in C := S(r)$ we have

$$r - f(x_n) \geq f(u) - f(x_n) \geq \langle x_n^*, u - x_n\rangle \geq -\|x_n^*\|.\|u - x_n\|,$$

hence, taking the infimum on $u \in C$ we get

$$f(x_n) - r \leq \|x_n^*\| d(x_n, C) \leq \|x_n^*\| c(r)^{-1}(f(x_n) - r),$$

a contradiction for n so large that $\|x_n^*\| < c(r)$.

(a)⇒(c) follows from the inequality $a(r) \leq c(r)$ when the assumptions of Proposition 10.6 (a) hold. □

Gathering different hypothesis, we can get more precise results, as in [14], but with relaxed convexity assumptions.

Proposition 10.8 *Suppose X is a dual Banach space, f is quasiconvex, l.s.c. for the weak* topology and radially usc on its domain. Suppose that for some $r \geq m$ and for any $w \in f^{-1}([r, \infty[)$, $x \in f^{-1}(r)$ one has $\partial f(w) \subset \partial^\leq f(w)$ and inclusion* (12). *Then one has*

$$a(r) = b(r) = c(r).$$

If, instead of (12)*, one has*

$$\sup\{\langle x^*, u\rangle : x^* \in \partial f(x)\} \leq f'(x,u) \leq \sup\left\{\langle x^*, u\rangle : x^* \in \partial^> f(x)\right\} \tag{14}$$

then

$$a(r) = b(r) = c(r) = k(r).$$

Proof. Once one observes that the first part of relation (14) implies that $\partial f(x)$ is contained in the Hadamard subdifferential of f at x, everything is a consequence of the previous estimates. □

Proposition 10.9 *Suppose f is convex, l.s.c. and bounded below and ∂ is the Fenchel subdifferential. Then for each $r > m := \inf f(X)$ one has $b(r) = c(r) \leq a(r) \leq k(r)$. If X is a dual Banach space and f is l.s.c. for the weak* topology then one has*

$$a(r) = b(r) = c(r) = k(r).$$

Proof. The first assertion is a consequence of Propositions 10.2, 10.4 and 10.5. Under the assumptions of the second assertion, the hypothesis of Proposition 10.6 (b) are satisfied but relation (13) as for each $x \in \mathrm{dom} f$ one has

$$\sup \{\langle x^*, u\rangle : x^* \in \partial f(x)\} = f'(x, u).$$

However, the proof of Proposition 10.6 and the proof of Lemma 10.1 show that when f is convex and $r > m$ one can replace condition (13) by the weaker one

$$N(S(r), x) \cap J^{-1}(\mathrm{dom} f \backslash S(r) - x) \subset \mathbb{R}_+ \partial f(x),$$

so that the conclusion $k(r) \leq c(r)$ holds. □

The preceding equalities are given in [15] in the finite dimensional case. It would be interesting to know whether they all hold in an arbitrary Banach space.

For the question of dual characterizations of well-behavior, we refer the reader to [1], [13], [55] and the paper [24] we received during the revision of the present article.

11 Application to perturbation of minimizer sets

In this short section we evoke the use of the notion of metrically well-set function for studying perturbations questions, following [7], [8], [30], [34] and [157]. Let us construct a topology appropriate to this aim. We observe that even for a very strong topology such as the topology of uniform convergence one cannot preserve the quality of conditioning under small perturbations and one cannot get lower semicontinuity of the minimizer set. Take for instance the one variable function f given by $f(x) := \max(\mid x \mid -1, 0)$ and the nearby functions $f_{\varepsilon,p}$ given by $f_{\varepsilon,p}(x) = \varepsilon(\mid x \mid^p -1)$ for $x \in [-1, 1]$, $f_{\varepsilon,p}(x) = f(x)$ for $\mid x \mid > 1$. Still the topology of uniform convergence is too strong a topology and it is not adapted to the case the functions take the value $+\infty$. Thus, we use a topology related to (but weaker than) the Pompeiu-Hausdorff convergence of epigraphs. Given a family $\mathcal{A}$ of closed subsets of a normed space X (which could be an arbitrary metric space for parts of what follows) and two subsets C, D of X we set

$$e_A(C, D) \quad : \quad = e(C \cap A, D) := \sup_{x \in C \cap A} d(x, D) \qquad A \in \mathcal{A}.$$
$$d_A(C, D) \quad : \quad = \max(e_A(C, D), e_A(D, C)) \qquad A \in \mathcal{A}$$

In the usual cases one takes for $\mathcal{A}$ either $\{X\}$ or $\{B(S, r) : r > 0\}$ for some $S \subset X$ or the family of closed bounded sets or the family of closed balls centered at 0. In the latter case, we set, for $r \in \mathbb{R}_+$

$$e_r\,(C, D) := e_{B(0,r)}(C, D) := \sup \{d\,(x, D) : x \in C \cap B(0, r)\}$$

$$d_r(C,D) := \max\left(e_r(C,D),\ e_r(D,C)\right),$$

with the usual convention $\inf \emptyset = \infty$, $\sup \emptyset = 0$ in $\mathbb{R}_+$ and we call them the bounded (or truncated) Hausdorff excesses (resp. the bounded Hausdorff distances) of C over D. These quantities have been introduced in [98], [135, p. 169] to study the variation of vector subspaces; they are also mentioned in the work of Mosco (Ref. [132]) and Moreau. These local (or truncated) excesses (or hemimetrics) and metrics have been used in 1986 by the author in [147] in order to replace the Mosco convergence by a stronger one preserving usual operations and extensively used for a quantitative study of conditioning in [6], [7], [8], [157] and for dealing with analytical and geometrical operations in [6], [7], [8], [9], [25], [26], [27], [32], [147], [146], [186], [187]). A comparison with other convergences is made in many other references. It has been shown that the Legendre-Young-Fenchel transform is continuous for the topology these local metrics induce ([29], [35], [144]) and that they satisfy a collective triangle inequality [145, Corollary 1.3] which justifies the name "polymetrics".

Identifying a function $f : X \to \overline{\mathbb{R}}$ with its epigraph

$$E(f) := \{(x,r) \in X \times \mathbb{R} : r \geq f(x)\},$$

the preceding hemimetrics and polymetrics induce hemimetrics and polymetrics on the space $\overline{\mathbb{R}}^X$ of extended real-valued functions on X :

$$e_A(f,g) := e(E(f), E(g)) = e(E(f) \cap A, E(g)),$$

$$d_A(f,g) := \max\left(e_A\left(E(f), E(g)\right), e_A\left(E(g), E(f)\right)\right),$$

with a corresponding simplification of the notation when $A = B(0,r) \times \mathbb{R}$.

Then one can define a topology called the *topology of $\mathcal{A}$-convergence* (and when $\mathcal{A}$ is the family of balls the topology of bounded hemiconvergence, or bounded-Hausdorff convergence) on $\overline{\mathbb{R}}^X$ by taking as a base of open sets the open "balls"

$$V_A(f,\varepsilon) := \left\{g \in \overline{\mathbb{R}}^X : d_A(f,g) < \varepsilon\right\}.$$

When $\mathcal{A}$ is the family of balls, this topology is metrizable and, if X is complete, the space $\mathcal{C}(X)$ of closed proper functions is complete in the associated metric ([9]). Moreover, when X is a Hilbert space, the topology of bounded hemiconvergence just described coincides with the topology of Attouch-Wets previously defined through infimal convolution ([5]).

In what follows, we endow the set $\overline{\mathbb{R}}^X$ of extended real valued functions on X the topology associated with the family $\mathcal{A}$ of cylinders $A := A_X \times \mathbb{R}$ where A_X belongs to a family $\mathcal{A}_X$ of subsets of X, for instance the family of balls $B_X(0,r)$ with $r > 0$. This choice will enable us to convey the main ideas (in particular the notion of reef) and to avoid some technical details in presenting results about the perturbation of infima and minimizers. Let us observe however that this topology induced on $\overline{\mathbb{R}}^X$ by the family of polymetrics $(d_A)_{A \in \mathcal{A}}$ is stronger than the topology of bounded hemiconvergence induced by the family $(d_B)_{B \in \mathcal{B}}$, where $\mathcal{B}$ is the family of balls centered at 0 in $X \times \mathbb{R}$. In fact, given subsets C, D of $X \times \mathbb{R}$, one has $e_A(C,D) \geq e_B(C,D)$ when $A \supset B$, in particular when $A = B_X(0,r) \times \mathbb{R}$ and $B = B_X(0,r) \times [-r,r]$. Thus the following theorem differs from previous results. Note that a connection can be established thanks to the following lemma.

Lemma 11.1 *Using the preceding notation, given a function $f \in \mathbb{R}^X$ which is bounded below on bounded sets, and given $\varepsilon > 0$, $A \in \mathcal{A}$, there exists $B \in \mathcal{B}$ such that $e_A(f,g) < \varepsilon$ whenever $g \in \overline{\mathbb{R}}^X$ satisfies $e_B(f,g) < \varepsilon$. Therefore, if $\mathcal{F}$ is a subset of $\overline{\mathbb{R}}^X$ such that for each $r > 0$ there exists some $c(r) > 0$ satisfying* $\inf\{f(x) : x \in B_X(0,r)\} \geq -c(r)$ *for each $f \in \mathcal{F}$, the topologies induced on $\mathcal{F}$ by the families $(d_A)_{A\in\mathcal{A}}$ and $(d_B)_{B\in\mathcal{B}}$ coincide.*

Proof. Let $A := B_X(0,r)\times \mathbb{R} \in\mathcal{A}$ and let $c := c(r) \geq r$ be such that $f(x) \geq -c$ for each $x \in B_X(0,r)$. Given $(x,s) \in E(f)\cap A$ we observe that $-c \leq f(x) \leq s$ so that $s' := \min(s,c) \in [-c,c]$. As $(x,s') \in B := B_X(0,r) \times [-c,c]$, given $g \in \overline{\mathbb{R}}^X$ such that $e_B(f,g) < \varepsilon$, we can find $(y,t') \in E(g)$ such that $d((x,s'),(y,t')) < \varepsilon$. Then, for $t := t' + s' - s$, we have $(y,t) \in E(g)$ and $d((x,s),E(g)) \leq d((x,s),(y,t)) = d((x,s'),(y,t')) < \varepsilon$. Taking the supremum on $(x,s) \in E(f)\cap A$ we get $e_A(f,g) < \varepsilon$. The proof of the second assertion is similar, interchanging the roles of f and g. □

The following central result for the study of perturbed minimization problems has two faces: the value function $f \mapsto m_f$ appears to satisfy a Lipschitz type property, but the set of minimizers S_f varies in a way similar to a Stepanov behavior, i.e. one can estimate $e(S_g,S_f)$ but not $e(S_g,S_h)$ for g,h close to f. Moreover $e(S_g,S_f)$ is sensitive to the conditioning of f (thus $g \mapsto e(S_g,S_f)$ is of Hölder type, usually). Here we say that f is *inf-connected* if its sublevel sets are connected. This class of functions has been extensively studied by Avriel, Zang and their co-authors ([20], [204]...). Quasiconvex functions on a normed space are obviously inf-connected. Note that here we do not suppose the set S_f is bounded as in [157] [149] and we simplify some arguments thanks to the use of the family $\mathcal{A}$ of cylinders described above. The case of the genuine bounded Hausdorff topology is similar.

Theorem 11.2 *Suppose the set S_f of minimizers of f is nonempty and $A = A_X \times \mathbb{R}$, where $A_X \cap S_f$ is nonempty. Then*

$$m_g \leq m_f + e_A(f,g).$$

In particular, for $r > d(0,S_f)$ one has $m_g \leq m_f + e_A(f,g)$, where $A = B_X(0,r)\times \mathbb{R}$.

Suppose f is metrically well-set, with growth gage γ. Given $\varepsilon > 0$ let $\delta > 0$ be such that $\delta \leq \frac{1}{2}\min(\varepsilon,\gamma(\frac{1}{2}\varepsilon))$, or, more generally, $\delta \leq \min(\varepsilon,\frac{1}{2}\gamma(\varepsilon-\delta))$. If A_X contains $B(S_f,\varepsilon)$ and if g is some inf-connected function on X satisfying $d_A(f,g) < \delta$, then one has $e(S_g,S_f) \leq \varepsilon$. Moreover, if h is another inf-connected function satisfying $d_A(f,h) < \delta$, then one has

$$|m_g - m_h| \leq d_A(g,h).$$

Proof. The first assertion is immediate: given $x \in S_f$, $\delta > e_A(f,g)$ we can find $(y,t) \in E(g)$ such that $d(x,y) < \delta$,

$$g(y) - f(x) \leq t - f(x) \leq \|(y,t) - (x,f(x))\| \leq \delta.$$

Taking the infimum on y and δ we get the announced inequality.

In order to prove the second assertion, let us consider $\varepsilon > 0$, $\delta > 0$, $A \in \mathcal{A}$ such that $\delta \leq \min(\varepsilon,\frac{1}{2}\gamma(\varepsilon-\delta))$, $B(S_f,\varepsilon) \subset A_X$ and an inf-connected function g such that $d_A(f,g) < \delta$. For any $y \in X$ such that $d(y,S_f) = \varepsilon$ there exists $x \in U(y,\delta)$ such that $f(x) < g(y) + \delta$. Then we have $d(x,S_f) > \varepsilon - \delta$, hence $f(x) \geq m_f + \gamma(\varepsilon-\delta)$. It follows that we cannot have $g(y) \leq m_f + \gamma(\varepsilon - \delta) - \delta$ and a fortiori we cannot have $g(y) \leq m_f + \delta$. Thus, the sublevel set $[g \leq m_f + \delta]$ is contained in the union of $U(S_f, \varepsilon)$ and $X \backslash B(S_f, \varepsilon)$. Since these sets are open and since the connected set $[g \leq m_f + \delta]$ meets the first one by what precedes, it is contained in it. As S_g is contained in $[g \leq m_f + \delta]$ by the begining of the proof, we get $e(S_g, S_f) \leq \varepsilon$. Note that we can take $\delta \leq \sup_{0<c<1} \min(\varepsilon(1-c), \frac{1}{2}\gamma(c\varepsilon))$, in particular $\delta \leq \frac{1}{2}\min(\varepsilon, \gamma(\frac{1}{2}\varepsilon))$.

The last assertion follows from the fact that for any $\beta > 0$ with $d_A(f,g) < \delta - \beta$ and any

$$y \in [g < m_g + \beta] \subset [g < m_f + e_A(f,g) + \beta] \subset [g < m_f + \delta]$$

one has $y \in U(S_f, \varepsilon) \subset A$ so that we can find $z \in B(y, d_A(g,h) + \beta)$ with $h(z) < g(y) + d_A(g,h) + \beta$. As β is arbitrarily small, we get $m_h \leq m_g + d_A(g,h)$. Moreover one can interchange the roles of g and h. $\square$

Remark. In a similar way one can prove that if φ is a conditioner for f, if g is an inf-connected function on X such that $d_A(f,g) < \delta$ for some set $A := A_X \times \mathbb{R}$ with A_X containing $B(S_f, \delta + \varphi(\delta))$, then one has

$$e\left(S_g, S_f\right) \leq d_A\left(f,g\right) + \varphi\left(2d_A\left(f,g\right)\right).$$

This result can be applied to the study of projection operators ; see [7], [157], and for a more direct approach, [163], [168] and their references.

One can also deduce from the preceding result a quantitative form of an upper hemicontinuity result of [29], [30], [157], for the subdifferential ∂f of a convex function f satisfying a smoothness condition or whose conjugate f^* satisfies a growth condition. Here we use the Legendre-Fenchel conjugate of the function f, given by the usual formula $f^*(y) := \sup_{x \in X}(\langle y, x \rangle - f(x))$ and we observe that for any $x_0 \in X$ the set $\partial f(x_0)$ is the minimizer set of $f^* - \langle x_0, \cdot \rangle$.

Theorem 11.3 ([157]) *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper convex l.s.c. function subdifferentiable at some $x_0 \in X$ such that $\partial f(x_0)$ is nonempty and bounded and such that for some gage γ one has for each $y \in X^*$*

$$f^*(y) - \langle x_0, y \rangle \geq \inf_{X^*}\left(f^* - \langle x_0, \cdot \rangle\right) + \gamma\left(d\left(y, \partial f\left(x_0\right)\right)\right). \tag{15}$$

Then, given $r > \sup\{\|y\| : y \in \partial f(x_0)\}$, there exists some $\varepsilon > 0$ such that for $x \in B(x_0, \varepsilon)$ one has

$$e\left(\partial f\left(x\right), \partial f\left(x_0\right)\right) \leq r\left\|x - x_0\right\| + \gamma^h\left(2r\left\|x - x_0\right\|\right).$$

Before presenting a variant of this result, let us observe that condition (15) is related to a smoothness property of f. Given a modulus μ one defines f to be μ-smooth at x_0 if there exists some $y_0 \in X^*$ such that

$$\forall x \in X \quad f\left(x\right) \leq f\left(x_0\right) + \langle x - x_0, y_0 \rangle + \mu\left(\left\|x - x_0\right\|\right).$$

Then one can show (see [28, Theorem 2.1, Prop. 3.1] and [21], [22], [165]) that

$$\forall y \in Y \quad f^*\left(y\right) \geq f^*\left(y_0\right) + \langle x_0, y - y_0 \rangle + \mu^*\left(\left\|y - y_0\right\|\right),$$

and if $\widehat{\mu}$ given by $\widehat{\mu}(t) = t^{-1}\mu(t)$ is a modulus, one sees easily that f is Fréchet differentiable at x_0 with $f'(x_0) = y_0$. Then one can apply the preceding corollary with $\gamma = \mu^*$.

Using the results of [28] and [202] we get a quantitative form of a well known upper semicontinuity result.

Variants along this line have been given in [29], [30]. Let us present here a simple approach which avoids the weak* compactness argument of [30, Theorem 3]. It relies on the following lemma of independent interest.

Lemma 11.4 *Suppose the set S_f of minimizers of f is nonempty, let γ be a growth function for f and let $\beta := \widehat{\gamma}$ be the slope of γ given by $\beta(0) = 0$, $\beta(t) = t^{-1}\gamma(t)$ for $t > 0$. Then, for any function g such that $g - f$ is Lipschitzian with rate c and for any $\varepsilon > 0$ such that $\beta(\varepsilon) > c$ one has $e(S_g, S_f) \leq \varepsilon$. In other terms*

$$e(S_g, S_f) \leq \beta^h(c).$$

Proof. The result is obvious when S_g is empty. Otherwise, given $y \in S_g$, for each $x \in S_f$ we have

$$\begin{aligned} f(y) - f(x) &\leq g(y) - g(x) + cd(y, x) \\ &\leq cd(y, x). \end{aligned}$$

Taking the infimum over $x \in S_f$ it follows that

$$\gamma(d(y, S_f)) \leq f(y) - m_f \leq cd(y, S_f),$$

hence

$$\beta(d(y, S_f)) \leq c. \tag{16}$$

As β is nondecreasing, for any $\varepsilon > 0$ such that $c < \beta(\varepsilon)$ one has $d(y, S_f) \leq \varepsilon$, hence $e(S_g, S_f) \leq \varepsilon$. Then the definition of β^h yields $e(S_g, S_f) \leq \beta^h(c)$. When β is l.s.c., we deduce from (16) that $\beta\left(e\left(S_g, S_f\right)\right) \leq c$. $\square$

Proposition 11.5 *Let f be an arbitrary function on the normed space X, let ∂ be the Fenchel subdifferential and let $w^* \in \partial f(X)$ be such that $f - w^*$ is very-well conditioned with reduced growth gage β. Then, for each $x^* \in X^*$ one has*

$$e\left((\partial f)^{-1}(x^*), (\partial f)^{-1}(w^*)\right) \leq \beta^h(\|x^* - w^*\|).$$

If X is the dual space of a normed space X_*, if w^* corresponds to some element $w_* \in X_*$ in the canonical injection $X_* \to X^*$ and if f is the conjugate function of a l.s.c. proper convex function f_*, then for each $x_* \in X_*$ one has

$$e\left(\partial f_*(x_*), \partial f_*(w_*)\right) \leq \beta^h(\|x_* - w_*\|).$$

Proof. Setting $j := f - w^*$, $k := f - x^*$, the result follows from the preceding lemma in which f and g are replaced by j and k respectively, using the equivalence

$$w \in (\partial f)^{-1}(w^*) \Leftrightarrow w^* \in \partial f(w) \Leftrightarrow w \in S_j$$

and the similar one with w, w^* replaced by x, x^* respectively. The last assertion, which is close to [30, Theorem 3], is a consequence of the equivalence

$$x \in (\partial f)^{-1}(x_*) \Leftrightarrow x_* \in \partial f(x) \Leftrightarrow x \in \partial f_*(x_*). \qquad \square$$

12 Error bounds, penalization and metric regularity

Among the outcomes of the study of conditioning and error bounds are their uses for exact penalization. For such applications, we refer the reader to the thorough survey by Pang [138] and to its abundant bibliography. Let recall briefly the principle of this procedure, which is quite simple. Consider the constrained minimization problem

$$(\mathcal{P}) \qquad \text{minimize } j(x) \text{ subject to } x \in F,$$

where the feasible set F is a nonempty subset of a metric space X and the objective function j is Lipschitzian with rate ℓ on X. Let $f := j + i_F$, where i_F is the indicator function of F given by $i_F(x) = 0$ for $x \in F$, ∞ for $x \in X \backslash F$. The value of $(\mathcal{P})$ is clearly the infimum m of f on X, but the new objective function is difficult to handle. It is advisable to substitute to it the more tractable function

$$f_r := j + rh$$

when one disposes of an auxiliary function $h : X \to \mathbb{R}$ which satisfies $h \mid F = 0$,

$$h(x) \geq cd(x, F) \quad \forall x \in X.$$

This is possible, thanks to the following result which is a slight variant of the well-known [53, Proposition 2.4.3], [138, Theorem 3], as here we do not assume that problem $(\mathcal{P})$ has an optimal solution.

Lemma 12.1 *Suppose the auxiliary function h satisfies the preceding assumptions and j is Lipschitzian with rate ℓ. Then for any $r \geq p := \ell c^{-1}$ one has*

$$\inf_{x \in X} f_r(x) = \inf_{x \in F} j(x).$$

If moreover F is closed and $r > p$ one has

$$\operatorname{Arg}\min_X f_r = \operatorname{Arg}\min_F j.$$

PROOF. As for any $u \in F$, $x \in X$ we have

$$j(u) \leq j(x) + \ell d(x, u),$$

taking the infimum on $u \in F$ and then on $x \in X$ we get

$$m := \inf_{u \in F} j(u) \leq \inf_{x \in X} (j(x) + \ell d(x, F)) \leq \inf_{x \in X} (j(x) + ph(x)) \leq \inf_{x \in X} (j(x) + rh(x)).$$

Since $h(x) = 0$ for $x \in F$, these inequalities are equalities. The same reason yields $\operatorname{Arg}\min_F j \subset \operatorname{Arg}\min_X f_r$. Now if $x \in \operatorname{Arg}\min_X f_r$ then

$$\begin{aligned} rh(x) &= f_r(x) - j(x) = m - j(x) \\ &\leq m - (m - \ell d(x, F)) = pcd(x, F) \leq ph(x). \end{aligned}$$

Since $r > p$ we must have $h(x) = 0$, hence $d(x, F) = 0$ and $x \in F$ as F is closed. □

The preceding assumptions on h are satisfied when F is the set of minimizers of a linearly conditioned function h. The growth criteria obtained in the preceding sections may be useful in such a connection.

An important case is when the feasible set F is defined by equalities and inequalities or, more generally, as an implicit constraint of the form

$$F = g^{-1}(C),$$

where $g : X \to Z$ is a mapping in some other metric space Z, C is a subset of Z and when g is *metrically regular* in the sense that for some $c > 0$ one has

$$d(x, F) \le cd(g(x), C) \qquad \forall x \in X.$$

Then one can take $h(x) = cd(g(x), C)$, a function which is usually easier to compute than $d(\cdot, F)$. A famous result due to Hoffman [87] asserts that such a situation occurs when X and Z are finite dimensional normed spaces, g is affine and C is a polyhedral cone. Such a result has been extended by Ioffe [90] to the case X and Z are Banach spaces, Z is a product $V \times W$, $C = \{0\} \times D$ where D is a polyhedral cone in the finite dimensional space W; see also [68], [100]–[103], [104]. In each of these cases, one has a penalization result which is standard in mathematical programming.

The preceding inequality pertains to the notion called metric regularity; however, this notion is usually adopted in a local sense. Recall that a multimapping $M : X \rightrightarrows Y$ between two normed spaces is said to be *metrically regular* around some (x_0, y_0) if there exist $c > 0$ and neighborhoods V, W of x_0 and y_0 respectively such that for each $y \in W$ and each $x \in V$ satisfying $d(x, M^{-1}(y)) < \delta$ one has $d(x, M^{-1}(y)) \le cd(y, M(x))$. It has been shown that such a property is equivalent to an openness property (see [23], [38], [143]). This property has been widely used, in the convex case as in the nonconvex case (see for instance [23], and [150] for recent contributions with numerous references). In particular, it is a useful tool for computing tangent cones and normal cones (see [141], [162] and their references).

When one considers a convex multimapping, the usual notion of metric regularity can be simplified, thanks to the following nice result of Li and Singer [113, Theorem 4]. In the general (local) case, see [86] and [186, Lemma 9.39].

Proposition 12.2 *Suppose $M : X \rightrightarrows Y$ is a multimapping with convex graph. Then the following assertions, in which c and δ are positive numbers and W is a subset of Y, are equivalent:*

(a) for each $y \in W$ and each $x \in X$ satisfying $d(x, M^{-1}(y)) < \delta$ one has

$$d(x, M^{-1}(y)) \le cd(y, M(x));$$

(b) for each $y \in W$ and each $x \in X$ one has $d(x, M^{-1}(y)) \le cd(y, M(x))$;

The equivalence is an easy consequence of the following lemma in which one takes $h(x) := d(y, M(x))$, $S := M^{-1}(y)$.

Lemma 12.3 *Let S be an arbitrary nonempty subset of X, let $\delta > 0$ and let h be a continuous convex function on X, null on S and such that $d(x, S) \le h(x)$ for each $x \in X$ satisfying $d(x, S) \le \delta$. Then $d(x, S) \le h(x)$ for each $x \in X$.*

Proof. Given $x \in X$, let (u_n) be a sequence of S such that $r_n := d(x, u_n) \to r := d(x, S)$. Let $t \in]0, 1]$ be such that $td(x, S) < \delta$. We may assume that $td(x, u_n) < \delta$ for each n. Let $z_n := (1-t)u_n + tx$. As $d(z_n, S) \le \|z_n - u_n\| < \delta$ and as $h(u_n) = 0$, we have

$$d(z_n, S) \le h(z_n) \le th(x). \tag{17}$$

Now we claim that $\limsup_n d(z_n, S) \ge td(x, S)$. Otherwise, there exist $\varepsilon > 0$ and an infinite set of integers K such that $d(z_k, S) < td(x, S) - \varepsilon$ for each $k \in K$. Then we can find some $w_k \in S$ such that $d(z_k, w_k) < tr - \varepsilon$ and we get for $k \in K$ large enough

$$d(x, w_k) \leq d(x, z_k) + tr - \varepsilon \leq (1-t)r_k + tr - \varepsilon < r,$$

a contradiction. Taking limsup in relation (17) we get $d(x, S) \leq h(x)$. $\square$

We refer to the recent papers of Auslender [13] and Lewis and Pang [111] for more results in the direction of error bounds for convex systems and their relationships with conditioning and recession conditions.

Acknowledgements. We are grateful to A. Daniilidis, D. Motreanu, S. Shiraishi and C. Zălinescu for their careful readings and criticisms and to D. Azé for providing us with reference [135].

REFERENCES

[1] P. Angleraud. Caractérisation duale du bon comportement de fonctions convexes. *C. R. Acad. Sci. Paris Sér. I Math.* **314** (1992), 583-586.

[2] E. Asplund, R. T. Rockafellar. Gradients of convex functions. *Trans. Amer. Math. Soc.* **139** (1966), 443-467.

[3] H. Attouch. Variational Convergence for Functions and Operators. Pitman, Boston, 1984.

[4] H. Attouch. Viscosity solutions of minimization problems. Epi-convergence and scaling. *Séminaire d'Anal. Convexe, Montpellier* **22**, *8* (1992), 1-48.

[5] H. Attouch, R. J.-B. Wets. Isometries for the Legendre-Fenchel transform. *Trans. Amer. Math. Soc.* **296** (1986), 33-60.

[6] H. Attouch, R. J.-B. Wets. Quantitative stability of variational systems: I. The epigraphical distance. *Trans. Amer. Math. Soc.* **328**, *2* (1992), 695-729.

[7] H. Attouch, R. J.-B. Wets. Quantitative stability of variational systems: II. A framework for nonlinear conditioning. *SIAM J. Optim.* **3** (1992), 359-381.

[8] H. Attouch, R. J.-B. Wets. Quantitative stability of variational systems: III. ε-approximate solutions. *Math. Programming* **61** (1993), 197-214.

[9] H. Attouch, R. Lucchetti, R. J.-B. Wets. The topology of the ρ-Hausdorff distance. *Ann. Mat. Pura Appl. (4)* **160** (1991), 303-320.

[10] J.-P. Aubin, I. Ekeland. Applied Nonlinear Analysis. Wiley-Interscience, New York, 1984.

[11] A. Auslender. Asymptotic properties of the Fenchel's dual functional and their applications to decomposition problems. *J. Optim. Theory Appl.* **73** (1992), 427-450.

[12] A. Auslender. Convergence of critical sequences for variational inequalities with maximal monotone operators. *Appl. Math. Optim.* **28** (1993), 161-172.

[13] A. Auslender. How to deal with the unbounded in optimization: theory and algorithms. *Math. Programming* **79** (1997), 3-18.

[14] A. Auslender, J.-P. Crouzeix. Well behaved asymptotical convex functions. In: Anal. Non-linéaire (eds. J.-P. Aubin et al), Gauthier-Villars, Paris, 1989, 101-122.

[15] A. Auslender, R. Cominetti, J.-P. Crouzeix. Convex functions with unbounded level sets and applications to duality theory. *SIAM J. Optim.* **3** (1993), 669-687.

[16] D. Aussel. Subdifferential properties of quasiconvex and pseudoconvex functions: a unified approach. Preprint, Univ. B. Pascal, Clermont-Ferrand, April 1995.

[17] D. Aussel, J.-N. Corvellec, M. Lassonde. Mean value property and subdifferential criteria for lower semicontinuous functions. *Trans. Amer. Math. Soc.* **347** (1995), 4147-4161.

[18] D. Aussel. Théorème de la valeur moyenne et convexité généralisée en analyse non régulière. Thesis, Univ. B. Pascal, Clermont-Ferrand, Nov. 1994.

[19] M. Avriel, W. E. Diewert, S. Schaible, I. Zang. Generalized Concavity. Plenum, New York, 1988.

[20] M. Avriel, I. Zang. Generalized arcwise-connected functions and characterizations of local-global minimum properties. *J. Optim. Theory Appl.* **32** (1980), 407-425.

[21] D. Azé. On the remainder of the first order development of convex functions. Preprint, Univ. of Perpignan, 1997.

[22] D. Azé. Eléments d'Analyse Convexe et Variationnelle. Ellipses, Paris, 1997.

[23] D. Azé, C. C. Chou, J.-P. Penot. Subtraction theorems and approximate openness for multifunctions: topological and infinitesimal viewpoints. *J. Math. Anal. Appl.* **221** (1998), 33-58.

[24] D. Azé, L. Michel. Computable dual characterizations of asymptotically well behaved convex functions. Preprint, Univ. of Perpignan, January 1998.

[25] D. Azé, J.-P. Penot. Recent quantitative results about the convergence of convex sets and functions. In: Functional analysis and approximation. Bagni di Lucca, Italy, 16-20 May 1988, Pitagora, Bologne, 1989, 90-110.

[26] D. Azé, J.-P. Penot. Operations on convergent families of sets and functions. *Optimization* **21** (1990), 521-534.

[27] D. Azé, J.-P. Penot. Qualitative results about the convergence of convex sets and convex functions. In: Optimization and Nonsmooth Analysis, (eds. A. D. Ioffe et al.). Pitman Research Notes, vol. **244**, Longman, Harlow, 1992, 1-25.

[28] D. Azé, J.-P. Penot. Uniformly convex and uniformly smooth convex functions. *Ann. Fac. Sci. Toulouse* **4**, *4* (1995), 705-730.

[29] D. Azé, A. Rahmouni. Lipschitz behavior of the Legendre-Fenchel transform. *Set-Valued Anal.* **2**, *1–2* (1994), 35-48.

[30] D. Azé, A. Rahmouni. Intrinsic bounds for Kuhn-Tucker points of perturbed convex programs. In: Recent developments in optimization, Seventh French-German Conference on Optimization, (eds. R. Durier, C. Michelot), Lecture notes in Economics and Math. systems, vol. **429**, Springer Verlag, Berlin, 1995, 17-35.

[31] D. Azé, A. Rahmouni. On primal dual stability in convex optimization. *J. Convex Anal.* **3** (1996), 309-329.

[32] D. Azé, M. Volle. A stability result in quasi-convex programming. *J. Optim. Theory Appl.* **67**, *1* (1990), 175-184.

[33] E. Bednarczuk, J.-P. Penot. On the position of the notions of well-posed minimization problems. *Boll. Un. Mat. Ital. B. (7)* **6-B**, (1992), 665-683.

[34] E. Bednarczuk, J.-P. Penot. Metrically well-set minimization problems. *Appl. Math. Optim.* **26** (1992), 273-285.

[35] G. Beer. Conjugate convex functions and the epi-distance topology. *Proc. Amer. Math. Soc.* **108** (1990), 117-126.

[36] J. M. Borwein. A note on ε-subgradients and maximal monotonicity. *Pacific J. Math.* **103** (1982), 307-314.

[37] J. M. Borwein, J. D. Vanderwerff. Convergence of Lipschitz regularization of convex functions. *J. Funct. Anal.* **128** (1995), 139-162.

[38] J. M. Borwein, D. M. Zhuang. Verifiable necessary and sufficient conditions for openness and regularity of set-valued and singled-valued maps. *J. Math. Anal. Appl.* **134** (1988), 441-459.

[39] H. Brézis. Opérateurs Maximaux Monotones. Math. Studies, North Holland, Amsterdam, 1973.

[40] H. Brézis, L. Nirenberg. Remarks on finding critical points. *Comm. Pure Appl. Math.* **44** (1991), 939-963.

[41] A. Brøndsted, R. T. Rockafellar. On the subdifferentiability of convex functions. *Proc. Amer. Math. Soc.* **16** (1965), 605-611.

[42] F. E. Browder. Nonlinear eigenvalue problems and group invariance. In: Functional analysis and related fields (ed. F. E. Browder), Springer Verlag, New York, 1970, 1-58.

[43] I. Campa, M. Degiovanni. Subdifferential calculus and nonsmooth critical point theory. Preprint, Univ. di Brescia, 1997.

[44] H. Cartan. Cours de calcul différentiel. Hermann, Paris, 1967, 1977.

[45] O. Chadly, Z. Chbani, H. Riahi. Directional derivatives and applications to subdifferential calculus. Preprint, Univ. of Marrakech, 1997.

[46] K. C. Chang. Variational methods for nondifferentiable functionals and their applications to partial differential equations. *J. Math. Anal. Appl.* **80**, *1* (1981), 102-129.

[47] G. Chavent. Strategies for the regularization of nonlinear least squares problems. In: Inverse problems in diffusion processes. Proceedings of the GAMM-SIAM symposium, (eds. H. W. Engl, W. Rundell), SIAM, Philadelphia, 1995, 217-232.

[48] G. Chavent. Quasi-convex sets and size $\times$ curvature condition, application to nonlinear inversion. *Appl. Math. Optim.* **24** (1991), 129-169.

[49] G. Chavent. New size $\times$ curvature conditions, for strict quasi-convexity of sets. *SIAM J. Control Optim.* **29** (1991), 1348-1372.

[50] G. Chavent, K. Kunish. Convergence of Tikhonov regularization for constrained ill-posed inverse problems. *Inverse Problems* **10** (1994), 63-76.

[51] G. Chavent, K. Kunish. On weakly nonlinear inverse problems. *Siam J. Appl. Math.* **56**, *2* (1996), 542-576.

[52] C. C. Chou, K. F. Ng, J. S. Pang. Minimizing and stationary sequences of optimization problems. *Siam J. Control Opt.*, to appear.

[53] F. H. Clarke. Optimization and Nonsmooth Analysis. Wiley-Interscience, New York, 1983.

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

[55] R. Cominetti. Some remarks on convex duality in normed spaces with and without compactness. *Control and Cybernetics* **23**, *1/2* (1994), 123-138.

[56] O. Cornejo, A. Jourani, C. Zălinescu. Conditioning and upper-Lipschitz inverse subdifferentials in nonsmooth optimization problems. *J. Optim. Theory Appl.* **95**, *1* (1997), 127-148.

[57] J.-N. Corvellec. Morse theory for continuous functionals. *J. Math. Anal. Appl.* **196** (1995), 1050-172.

[58] J.-N. Corvellec, M. Degiovanni, M. Marzocchi. Deformation properties of continuous functionals and critical point theory. *Topological Methods in Nonlinear Anal.* **1** (1993), 151-171.

[59] J.-P. Crouzeix, J. Ferland, C. Zălinescu. α-convex sets and strong quasiconvexity. *Math. Oper. Res.* **22**, *4* (1997), 998-1022.

[60] J. Danes. On local and global moduli of convexity. *Commentationes Math. Universitatis Carolinea.* **18**, *2* (1977), 393-400.

[61] M. Degiovanni. Nonsmooth critical point theory and applications. Second World Congress of Nonlinear Analysts, Athens, 1996.

[62] M. Degiovanni, A. Marino, M. Tosques. Evolution equations with lack of convexity. *Nonlinear Anal. Theory Methods Appl.* **9** (1985), 1401-1443.

[63] M. Degiovanni, M. Marzocchi. A critical point theory for nonsmooth functionals. *Ann. Mat. Pura Appl. (4)* **167** (1994), 73-100.

[64] R. Deville, G. Godefroy, V. Zizler. A smooth variational principle with applications to Hamilton-Jacobi equations in infinite dimensions. *J. Funct. Anal.* **111** (1993), 197-212.

[65] J. I. Diaz. On a nonlocal elliptic problem arising in the magnetic confinement of a plasma in a stellarator. Proc. 2nd World Congress of Nonlinear Analysts. *Nonlinear Anal. Theory Methods Appl.* **30**, *7* (1997), 3963-3974.

[66] DIEUDONNÉ. Foundations of Modern Analysis. Academic Press, New York, 1960.
[67] S. DOLECKI, P. ANGLERAUD. When a well behaving convex function is well-conditioned. *Southeast Asian Bull. Math.* **20** (1996), 59-63.
[68] A. L. DONTCHEV, R. T. ROCKAFELLAR. Characterizations of Lipschitzian stability in nonlinear programming. In: Mathematical Programming with Data Perturbations, (ed. A. V. Fiacco.), Lecture Notes in Pure and Appl. Math., vol. **195**, M. Dekker, New York, 1998, 65-82.
[69] A. L. DONTCHEV, T. ZOLEZZI. Well-posed optimization problems. Lecture Notes in Maths., vol. **1543**, Springer Verlag, Berlin, 1993.
[70] I. EKELAND. Sur les problèmes variationnels. *C. R. Acad. Sci. Paris Sér. I Math.* **275** (1972), 1057-1059; **276** (1973), 1347-1348.
[71] I. EKELAND. On the variational principle. J. MATH. ANAL. APPL. **47** (1974), 324-353.
[72] I. EKELAND. Nonconvex minimization problems. *Bull. Amer. Math. Soc.* **1** (1979), 443-474.
[73] T. FIGIEL. On the moduli of convexity and smoothness. *Studia Math.* **56**, *2* (1978), 121-155.
[74] A. FOUGÈRES. Propriétés géométriques et minoration des intégrandes convexes normales coercives. *C. R. Acad. Sci. Paris Sér. I Math.* **A-B 284** (1977), A 873-876.
[75] A. FOUGÈRES. Analyse variationnelle I: convexité et structure semi-normée intrinsèque. Preprint, Univ. Perpignan, October 1980.
[76] M. FURI, A. VIGNOLI. About well-posed minimum problems for functionals in metric spaces. *J. Optim. Theory Appl.* **5** (1970), 225-229.
[77] M. FURI, A. VIGNOLI. A characterization of well-posed minimum problems in a complete metric space. *J. Optim. Theory Appl.* **5** (1970), 452-461.
[78] E. GINER. Local minimizers of integral functionals are global minimizers. *Proc. Amer. Math. Soc.* **123**, *3* (1995), 755-757.
[79] E. GINER. Minima sous contrainte de fonctionnelles intégrales. *C. R. Acad. Sci. Paris Sér. I Math.* **321**, *4* (1995), 429-431.
[80] E. GINER. On Pareto minima of vector-valued integral functionals. Preprint, Univ. Toulouse, 1997.
[81] M. GOBERNA, A. LOPEZ. Optimal value function in semi-infinite programming. *J. Optim. Theory Appl.* **59** (1988), 261-278.
[82] D. GOELEVEN. Noncoercive variational problems and related results. Pitman Research Notes in Maths. vol. **357**, Longman, Harlow, 1997.
[83] S. GUILLAUME. Méthode de plus grande pente en analyse convexe composite. *C. R. Acad. Sci. Paris Sér. I Math.* **322** (1996), 9-14.
[84] J. M. GUTIÉRREZ. Infragradientes y direcciones de decrecimiento. *Rev. real Acad. C. Ex., Fis. y Nat. Madrid* **78** (1984), 523-532.
[85] N. HADJISAVVAS, S. SCHAIBLE. Generalized monotone maps: concepts and characterizations (multivalued case), to appear.
[86] R. HENRION. The approximate subdifferential and parametric optimization. Habilitation thesis, Humbold Univ., Berlin, 1997.
[87] A. J. HOFFMAN. On approximate solutions of systems of linear inequalities. *J. Research Nat. Bureau of Standards* **49** (1952), 263-265.
[88] L. R. HUANG, X. B. LI. Minimizing sequences in nonsmooth optimization. Preprint, The Chinese Univ. of Hong Kong, 1995.
[89] L. R. HUANG, K. F. NG, J.-P. PENOT. On minimizing and critical sequences in nonsmooth optimization. Preprint, February 1997.
[90] A. D. IOFFE. Regular points of Lipschitz functions. *Trans. Amer. Math. Soc.* **251** (1979), 61-69.

[91] A. D. Ioffe. On subdifferentiability spaces. *Ann. New York Acad. Sci.* **410** (1983), 107-119.

[92] A. D. Ioffe. Calculus of Dini subdifferentials of functions and contingent coderivatives of set-valued maps. *Nonlinear Anal. Theory, Methods Appl.* **8** (1984), 517-539.

[93] A. D. Ioffe. Codirectional compactness, metric regularity and subdifferential calculus. Preprint, Technion, Haifa 1996.

[94] A. D. Ioffe, J.-P. Penot. Subdifferentials of performance functions and calculus of coderivatives of set-valued mappings. *Serdica Math. J.* **22** (1996), 359-384.

[95] A. Ioffe, E. Schwartzman. Metric critical point theory 1. Morse regularity and homotopic stability of a minimum. *J. Math. Pures Appl.* **75** (1996), 125-153.

[96] A. Ioff, E. Schwartzman. Metric critical point theory 2. Deformation techniques. *Oper. Theory: Adv. Appl.* **98** (1997), 131-144.

[97] A. D. Ioffe, V. M. Tihomirov. Theory of extremal problems. Stud. Math. Appl. vol. **6**, North Holland, Amsterdam, 1979.

[98] T. Kato. Perturbation Theory for Linear Operators. Springer Verlag, Berlin, 1966.

[99] G. Katriel. Mountain pass theorems and global homeomorphism theorems. *Ann. Inst. H. Poincaré, Anal. Non Linéaire* **11** (1994), 189-209.

[100] D. Klatte. Lipschitz stability and Hoffmann's error bound for convex inequality systems. In: Parametric optimization and related topics IV. Proceedings Intern. Conference on parametric optimization and related topics IV Enschede (NL), June 6-9, 1995 (eds. J. Guddat and al.), Peter Lang, 1996, 201-212.

[101] D. Klatte. Hoffmann's error bound for systems of convex inequalities. In: Mathematical programming with data perturbations, (ed. A. V. Fiacco), Lecture Notes in Pure and Appl. Math., vol. **195**, M. Dekker, New York, 1998, 185-199.

[102] D. Klatte, Wu Li. Asymptotic constraint qualification and global error bounds for convex inequalities. *Math. Programming*, to appear.

[103] D. Klatte. Error bounds for solutions of linear equations and inequalities. *Z. Oper. Res.* **41** (1995), 191-214.

[104] B. Kummer. Lipschitzian and pseudo-Lipschitzian inverse functions and applications to nonlinear optimization. In: Mathematical programming with data perturbations, (eds. A. V. Fiacco), Lecture Notes in Pure and Appl. Math., vol. **195**, M. Dekker, New York, 1998, 201-222.

[105] O. Lefebvre, C. Michelot. About the finite convergence of the proximal point algorithm. In: Trends in Mathematical Optimization, (eds. K. H. Hoffmann et al.), Intern. Series in Numer. Math., vol. **84**, Birkhäuser, Basel, (1988), 153-179.

[106] B. Lemaire. The proximal algorithm. In: New Methods in Optimization and their Industrial Uses, (ed. J.-P. Penot), Intern. Series in Numer. Math., vol. **84**, Birkhäuser, Basel, 1989, 73-88.

[107] B. Lemaire. About the convergence of the proximal method. In: Advances in Optimization. Proceedings, Lambrecht, FRG, 1991, (eds. W. Oettli and D. Pallaschke). Lecture Notes in Econom. and Math. Systems vol. **382**, Springer Verlag, Berlin, 1992, 39-51.

[108] B. Lemaire. Bons comportements et conditionnement linéaire, hand-out. Journées d'analyse non-linéaire et optimisation Nov. 1991, Avignon.

[109] B. Lemaire. Bonne position, conditionnement, et bon comportement asymptotique. *Séminaire d'Analyse Convexe, Univ. of Montpellier* **22** (1992), 5.1-5.12.

[110] C. Lemaire-Misonne. Conditionnement de problèmes. Application aux statistiques. Preprint, Univ. of Montpellier.

[111] A. S. Lewis, J. S. Pang. Error bounds for convex inequality systems. In: Generalized convexity and generalized monotonicity (eds. J.-P. Crouzeix, J.-E. Martinez-Legaz, M. Volle), Kluwer, Dordrecht, 1998, 75-110.

[112] Wu Li. Error bounds for piecewise convex quadratic programs and applications. *SIAM J. Control Optim.* **33**, *5* (1995), 1510-1529.
[113] W. Li, I. Singer. Global error bounds for convex multifunctions and applications. Preprint, Old Dominion University, July 1996.
[114] Ph. Loewen. A mean value theorem for Fréchet subgradients. *Nonlinear Anal. Theory Methods Appl.* **23** (1994), 1365-1381.
[115] D. T. Luc. On generalised convex nonsmooth functions. *Bull. Austral. Math. Soc.* **49** (1994), 139-149.
[116] D. T. Luc. Characterizations of quasiconvex functions. *Bull. Austral. Math. Soc.* **48** (1993), 393-405.
[117] D. T. Luc, C. Malivert. Invex optimization problems. *Bull. Austral. Math. Soc.* **46**, *1* (1992), 47-66.
[118] D. T. Luc, S. Swaminathan. A characterization of convex functions. *Nonlinear Anal. Theory, Methods Appl.* **30** (1993), 697-701.
[119] R. Lucchetti, F. Patrone. A characterization of Tychonov well-posedness for minimum problems with applications to variational inequalities governed by linear operators. *Numer. Funct. Anal. Optim.* **3** (1981), 461-476.
[120] R. Lucchetti, F. Patrone. Some properties of well-posed variational inequalities. *Numer. Funct. Anal. Optim.* **5** (1982-1983), 341-361.
[121] R. Lucchetti, F. Patrone. Hadamard and Tychonov well-posedness for a class of convex functions. *J. Math. Anal. Appl.* **88** (1982), 204-215.
[122] R. Lucchetti, T. Zolezzi. On well-posedness and stability analysis in optimization. In: Mathematical programming with data perturbations, (ed. A. V. Fiacco.), Lecture Notes in Pure and Appl. Math., vol. **195**, M. Dekker, New York, 1998, 223-251.
[123] O. L. Mangasarian. Pseudoconvex functions. *SIAM J. Control* **3** (1965), 281-290.
[124] O. L. Mangasarian. Nonlinear Programming. Mc Graw-Hill, New-York, 1969.
[125] D. H. Martin. The essence of invexity. *J. Optim. Theory Appl.* **47** (1985), 65-76.
[126] J.-E. Martinez-Legaz. Level sets and the minimal time function of linear control processes. *Numer. Funct. Anal. Optim.* **9**, *1-2* (1987), 105-129.
[127] J.-E. Martinez-Legaz. On lower subdifferentiable functions. Trends in Mathematical Optimization, (eds. K. H. Hoffmann et al.) Int. Series Numer. Math., vol. **84**, Birkhauser, Basel, 1988, 197-232.
[128] J.-E. Martinez-Legaz, P. H. Sach. A new subdifferential in quasiconvex analysis. Preprint 95/9, Hanoi Institute of Math, 1995.
[129] B. Martos. Nonlinear Programming, Theory and Methods. North Holland, Amsterdam, 1975.
[130] J. Mawhin, M. Willem. Critical point theory and Hamiltonian systems. Appl. Math. Sci., vol. **74**, Springer Verlag, New York 1989.
[131] P. Mazzoleni. Generalized concavity for economic applications. Proc. Workshop Pisa 1992, Univ. Verona.
[132] U. Mosco. Convergence of convex sets and solutions of variational inequalities. *Adv. in Math.* **3** (1969), 510-585.
[133] J. Mossino. Inégalités isopérimétriques et applications en physique. Hermann, Paris, 1984.
[134] D. Motreanu. Existence of critical points in a general setting. *Set-Valued Anal.* **3** (1995), 295-305.
[135] J. Necas. Les méthodes directes en théorie des équations elliptiques. Academia, Prague, Masson, Paris, 1967.
[136] R. S. Palais. Critical point theory and minimax principle. In: Global Analysis, Proc. Symp. Pure Applied Maths., vol. **15**, 1970, 185-212.
[137] R. S. Palais, S. Smale. A generalized Morse theory. *Bull. Amer. Math. Soc.* **70** (1964), 165-171.

[138] J. S. Pang. Error bounds in mathematical programming. *Math. Prog.* **79** (1997), 299-332.
[139] B. N. Pchenitchny, Y. Daniline. Méthodes numériques dans les problèmes d'extrémum. Mir, French transl., Moscow, 1975.
[140] J.-P. Penot. Méthodes de descente. Lecture Notes Univ. of Pau, 1974.
[141] J.-P. Penot. On regularity conditions in mathematical programming. *Math. Progr. Study* **19** (1982), 167-199.
[142] J.-P. Penot. The drop theorem, the petal theorem and Ekeland's variational principle. *J. Nonlinear Anal., Th., Methods, Appl.* **10**, *9* (1986), 813-822.
[143] J.-P. Penot. Metric regularity, openness and Lipschitzian behavior of multifunctions. *J. Nonlinear Anal., Th., Methods, Appl.* **13**, *6* (1989), 629-643.
[144] J.-P. Penot. The cosmic Hausdorff topology, the bounded Hausdorff topology and continuity of polarity. *Proc. Amer. Math. Soc.* **113**, *1* (1991), 275-285.
[145] J.-P. Penot. Topologies and convergences on the space of convex functions. *Nonlinear Anal. Theory Methods Appl.* **18**, *10* (1992), 905-916.
[146] J.-P. Penot. On the convergence of subdifferentials of convex functions. *Nonlinear Anal. Theory Methods Appl.* **21**, *2* (1993), 87-101.
[147] J.-P. Penot. Preservation of persistence and stability under intersections and operations, Part I: Persistence. *J. Optim. Theory Appl.* **79**, *3* (1993), 525-551.
[148] J.-P. Penot. Palais-Smale condition and coercivity. Preprint, Univ. of Pau, 1994.
[149] J.-P. Penot. Miscellaneous incidences of convergence theories in optimization and nonlinear analysis I: behavior of solutions. *Set-Valued Anal.* **2** (1994), 259-274.
[150] J.-P. Penot. Inverse functions theorems for mappings and multimappings. *South East Asian Bull. Math.* **19**, *2* (1995), 1-16.
[151] J.-P. Penot. Miscellaneous incidences of convergence theories in optimization and nonlinear analysis II: applications in nonsmooth analysis. In: Recent Advances in nonsmooth optimization (eds. D.-Z. Du, L. Qi, R. S. Womersley), World Scientific, Singapore, 1995, 289-321.
[152] J.-P. Penot. Generalized convexity in the light of nonsmooth analysis. In: Recent developments in optimization, Seventh French-German Conference on Optimization, (eds. R. Durier, C. Michelot), Lecture Notes in Econom. and Math. Systems, vol. **429**, Springer Verlag, Berlin, 1995, 269-290.
[153] J.-P. Penot. Subdifferential calculus without qualification assumptions. *J. Convex Anal.* **3**, *2* (1996), 207-220.
[154] J.-P. Penot. Proximal mappings. Preprint, Univ. of Pau, 1995; (to appear in J. Approx. Theory).
[155] J.-P. Penot. Compactness properties, openness criteria and coderivatives. Preprint, Univ. of Pau, Sept. 1995.
[156] J.-P. Penot. Subdifferential calculus and subdifferential compactness. In: Proceedings of the Second Catalan Days on Applied Mathematics (eds. M. Sofonea, J.-N. Corvellec), Presses universitaires de Perpignan, Perpignan, 1995, 209-226.
[157] J.-P. Penot. Conditioning convex and nonconvex problems. *J. Optim. Theory Appl.* **90**, *3* (1995), 539-558.
[158] J.-P. Penot. Genericity of well-posedness, perturbations and smooth variational principles. Lecture in the conference on "Well-posed optimization problems", Marseille, Sept. 1995 and preprint, Univ. of Pau, 1995.
[159] J.-P. Penot. Starshaped functions and conjugacy. Preprint Univ. of Pau, 1996.
[160] J.-P. Penot. A mean value theorem with small subdifferentials. *J. Optim. Theory Appl.* **94**, *1* (1997), 209-221.
[161] J.-P. Penot. Are generalized derivatives useful for generalized convexity? In: Generalized convexity and generalized monotonicity (eds. J.-P. Crouzeix, J.-E. Martinez-Legaz, M. Volle), Kluwer, Dordrecht, 1998, 3-59.

[162] J.-P. PENOT. Metric estimates for the calculus of multimappings. *Set-Valued Anal.* **5**, *4* (1997), 291-308.

[163] J.-P. PENOT. Continuity of projection operators. Preprint, Univ. Pau, 1997.

[164] J.-P. PENOT. Conditioning and nice-behavior of generalized convex functions, in preparation.

[165] J.-P. PENOT. Growth conditions and duality, in preparation.

[166] J.-P. PENOT, P. H. QUANG. On generalized convexity offunctions and generalized monotonicity of set-valued maps. *J. Optim. Theory Appl.* **92**, *2* (1997), 343-356.

[167] J.-P. PENOT, R. RATSIMAHALO. Characterizations of metric projections in Banach spaces. *Abstract and Applied Anal.*, to appear.

[168] J.-P. PENOT, R. RATSIMAHALO. Uniform continuity of projection mappings in Banach spaces. Preprint, Univ. of Pau, 1997.

[169] J.-P. PENOT, P. H. SACH. Generalized monotonicity of subdifferentials and generalized convexity. *J. Optim. Theory Appl.* **64**, *1* (1997), 251-262.

[170] J.-P. PENOT, M. VOLLE. Inversion of real-valued functions and applications. *Z. Oper. Res.* **34** (1990), 117-141.

[171] R. PHELPS. Convex functions, monotone operators and differentiability. Lecture Notes in Math., vol. **1364**, Springer Verlag, New York, 1989.

[172] R. PINI. Invexity and generalized convexity. *Optimization* **22** (1991), 513-525.

[173] R. PINI, C. SINGH. A survey of recent advances in generalized convexity with applications to duality theory and optimality conditions (1985-1995).*Optimization* **39**, *4* (1997), 311-360.

[174] F. PLASTRIA. Lower subdifferentiable functions and their minimization by cutting plane. *J. Optim. Theory Appl.* **46**, *1* (1985), 37-54.

[175] R. A. POLIQUIN. Subgradient monotonicity and convex functions. *Nonlinear Anal. Theory Methods Appl.* **14** (1990), 305-317.

[176] B. T. POLYAK. Sharp minimum. International worshop on augmented Lagrangians, Vienna, 1979.

[177] V. PTAK. Nondiscrete mathematical induction and iterative existence proofs. *Linear Algebra Appl.* **13** (1976), 223-236.

[178] P. H. RABINOWITZ. Minimax methods in critical point theory with applications to differential equations. CBMS Regional Conference series Math., vol. **65**, Amer. Math. Soc., Providence, 1986.

[179] J. M. RAKOTOSON. Some properties of the relative rearrangement. *J. Math. Anal. Appl.* **135** (1988), 475-495.

[180] J. P. REVALSKI. Generic properties concerning well-posed optimization problems. *C. R. Acad. Bulgare Sci.* **38**, *9* (1985), 1431-1434.

[181] J. P. REVALSKI. Generic well-posedness in some classes of optimization problems. *Acta Univ. Carolin. – Math. Phys.* **8**, *2* (1987), 117-125.

[182] J. P. REVALSKI. Well-posedness of optimization problems – a survey. In: Functional analysis and approximation, (ed. P. L. Papini), Bagni di Lucca (Italy) 16-20 May 1988, Pitagora, Bologne, 1989, 238-255.

[183] J. P. REVALSKI. Hadamard and strong well-posedness for convex programs. *SIAM J. Optim.* **7**, *2* (1997), 519-526.

[184] N. K. RIBARSKA, T .V. TSACHEV, M. I. KRASTANOV. Speculating about mountains. *Serdica Math. J.* **22**, *3* (1996), 341-358.

[185] R. T. ROCKAFELLAR. Convex Analysis. Princeton Univ. Press, Princeton, 1970.

[186] R. T. ROCKAFELLAR, R. J.-B. WETS. Variational systems, an introduction. In: Multifunctions and Integrands, Proc. Conf. Catania, 1983, (ed. G. Salinetti). Lecture Notes in Math., vol. **1091**, Springer-Verlag, Berlin, 1984, 112-129.

[187] R. T. ROCKAFELLAR, R. J.-B. WETS. Variational Analysis. Springer-Verlag, Berlin, 1997.

[188] M. Struwe. Variational methods. Applications to Nonlinear Partial Differential Equations and Hamiltonian Systems. Erg. Math., vol. **34**, Springer, Berlin, 1990, 1996.

[189] P. Shunmugaraj. Well-set and well-posed minimization problems. *Set-Valued Analysis* **3**, *3* (1995), 281-294,

[190] A. Szulkin. Ljusternik-Schnirelmann theory on C^1-manifolds. *Ann. Inst. H. Poincaré, Anal. Non Linéaire* **5** (1988), 119-139.

[191] A. Szulkin. Minimax principles for lower semicontinuous functions and applications to nonlinear boundary value problems. *Ann. Inst. H. Poincaré, Anal. Non Linéaire* **3**, *2* (1986), 77-109.

[192] J. Szarski. Differential Inequalities. Polish Scientific Publishers, Warszawa, 1967.

[193] G. Talenti. Elliptic equations and rearrangements. *Ann. Scuola Norm. Sup. Pisa, Cl. Sci IV* **4** (1976), 697-718.

[194] A. V. Tikhonov. Solution to incorrectly formulated problems and the regularization method. *Dokl. Akad. Nauk SSSR* **151** (1963), 501-504.

[195] A. V. Tikhonov, V. Arsénine. Méthodes de résolution de problèmes mal posés. MIR, Moscow, 1974; French translation, 1976.

[196] M. J. Todd. On convergence properties of algorithms for unconstrained minimization, *IMA J. Numer. Anal.* **9** (1989), 435-441.

[197] M. Tosques. Quasi-autonomous parabolic evolution equations associated with a class of non linear operators. *Ricerche Mat.* **38** (1989), 63-92.

[198] P. Tseng, D. Bersekas. Relaxation methods for problems with strictly convex costs and linear constraints. *Math. Programming* **38** (1987), 303-321.

[199] D. Ward. Sufficient conditions for weak sharp minima of order two and directional derivatives of the value function. In: Mathematical programming with data perturbations, (ed. A. V. Fiacco), Lect. Notes in Pure and Appl. Math., vol. **195**, M. Dekker, New York, 1998, 419-436.

[200] M. Willem. Lectures on critical point theory. *Trab. Dep. Mat.* **199**, 1983.

[201] X. M. Yang. Semistrictly convex functions. *Opsearch* **31** (1994), 15-27.

[202] C. Zalinescu. On uniformly convex functions. *J. Math. Anal. Appl.* **95** (1983), 344-374.

[203] C. Zalinescu. Mathematical Programming in Infinite Dimensional Normed Spaces. Editura Academiei, Bucharest, and French translation, to appear.

[204] I. Zang, M. Avriel. On functions whose local minimum are global. *J. Optim. Theory Appl.* **16** (1975), 183-190.

[205] I. Zang, E. V. Choo, M. Avriel. A note on functions whose local minimum are global. *J. Optim. Theory Appl.* **18** (1976), 555-559.

[206] E. Zeidler. Nonlinear Functional Analysis and its Applications, II B. Springer, Berlin, 1990.

[207] R. Zhang, J. Treiman. Upper-Lipschitz multifunctions and inverse subdifferentials. *Nonlinear Anal. Theory Methods Appl.* **24**, *2* (1995), 273-286.

[208] T. Zolezzi. On equiwellset minimum problems. *Appl. Math. Optim.* **4** (1978), 209-223.

Laboratoire de Mathématiques Appliquées
CNRS UPRES A 5033
Faculté des Sciences
Av. de l'Université
64000 Pau
France