HAL open science

Nonlinear Error Bounds via a Change of Function

Dominique Azé, Jean-Noël Corvellec

▶ **To cite this version:**

Dominique Azé, Jean-Noël Corvellec. Nonlinear Error Bounds via a Change of Function. Journal of Optimization Theory and Applications, 2016, pp.9-32. 10.1007/s10957-016-1001-3 . hal-01368648

HAL Id: hal-01368648

https://hal.science/hal-01368648v1

Submitted on 9 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Nonlinear Error Bounds via a Change of Function

Dominique Azé[1] · **Jean-Noël Corvellec**[2]

Abstract This work can be seen as a sequel to our previous paper, which dealt with local nonlinear error bounds for lower semicontinuous functions on complete metric spaces, based on estimates of the strong slope of the function through the distance to a sublevel set. Here, we consider estimates of the strong slope through the values of the function, and we provide characterizations of nonlinear local and global error bounds that are again obtained by a reduction to the linear case. Our main tool is a simple chain rule for the strong slope. We provide several examples showing that recent results in the literature can be recaptured from our general framework.

Keywords Nonlinear error bounds · Asymptotically well-behaved functions · Descent methods in optimization

Mathematics Subject Classification 49J53 · 49J52 · 90C30

1 Introduction

In our previous paper [1], we developed a general approach to nonlinear local error bounds, where the basic assumptions are estimates of the strong slope (of a lower

✉ Jean-Noël Corvellec
corvellec@univ-perp.fr

Dominique Azé
dominique.aze@math.univ-toulouse.fr

1 Université Paul Sabatier, Toulouse, France

2 Université de Perpignan Via Domitia, Perpignan, France

semicontinuous function defined on a complete metric space) through the distance to a sublevel set. The main tool was a so-called change-of-metric principle that allows reducing the nonlinear case to the linear one (systematically studied in [2]), and the results were refinements of those of [3], where the local case was not addressed in a satisfactory manner.

In this paper, we develop the theory of nonlinear error bounds based on assumptions that are estimates of the strong slope through the values of the function outside a sublevel set. This was also treated in [3] (and related therein to the notion of asymptotically well-behaved function) in the *global* case, through an ad hoc argument that is again unsatisfactory, in particular since it does not extend to the local case. Here, we use a simple calculus lemma, a chain rule for the strong slope (Lemma 4.1), that reduces the nonlinear problem to a linear one, both in the local and global case (and which accounts for the title of the paper). We thus start with a presentation of the characterization of error bounds in the linear case, which was already (partly) addressed more than ten years ago in our papers [2,4]. However, not only was the local case not treated in a systematic way in these papers, but we shall here do things in a way that fits the mentioned reduction procedure, and that takes into account more recent developments (both ours and in the literature).

Technically, the results in this paper are obtained in a simpler way than those of [1,3], but it could be said that they do not serve the same purposes. Roughly speaking, the latter are related to the notion of *metric (sub) regularity* and to applications to sensitivity analysis in optimization, while the former fit problems of convergence in descent methods. As a matter of fact, an important motivation for our present study comes from the papers [5–7], which basically deal with that kind of problem, and we conclude this work with an example of convergence of a descent method in optimization, following [5,7].

The paper is organized as follows. In Sect. 2, we recall the notion of strong slope and its relationship with subdifferential notions, while specifying our main notations. In Sect. 3, we establish the characterization of local and global linear error bounds through the variational principle. In Sect. 4, we deduce the nonlinear results through a chain rule for the strong slope. In Sect. 5, we focus on so-called Hölderian error, which has been mainly considered in the literature, some examples of which we describe. In Sect. 6, we make a few remarks in the specific case of convex functions (on Banach spaces). In Sect. 7, we compare our main assumption with that of [1], we give a new result under the former assumption, and we describe an application to the convergence of a descent method in optimization, as mentioned above.

2 Basic Notations and Notions

Throughout the paper, we let X denote a metric space endowed with the metric d, and we let $f : X \to \mathbb{R} \cup \{+\infty\}$ be an extended-valued function that will generally be required to be lower semicontinuous. For $x \in X$ and $\rho > 0$, we denote by $B_\rho(x)$ (resp. $B_\rho[x]$), the open (resp., closed) ball of radius $\rho > 0$ centered at x. For $Y \subset X$, we let

$$d(x, Y) := \inf\{d(x, y) \ : \ y \in Y\},$$

with the usual convention $d(x, \emptyset) = +\infty$. For $-\infty < a < b \leq +\infty$, we let:

$$[f \leq a] := \{x \in X \,:\, f(x) \leq a\}\,, \quad [f < b] := \{x \in X \,:\, f(x) < b\}$$

denote the sublevel sets of f, and we let

$$[a < f < b] := [f < b] \backslash [f \leq a]$$

denote the slice between a and b. When $b = +\infty$, we write

$$\operatorname{dom} f := [f < +\infty]\,, \quad [f > a] := [a < f < +\infty]\,.$$

As usual, for $a \in \mathbb{R}$, we let $a_+ := \max\{a, 0\}$.

Given $\bar{x} \in X$ with $a := f(\bar{x}) \in \mathbb{R}$, we say that the function f has a *nonlinear local error bound* at $\bar{x}$ if there exist $\gamma :]0, +\infty[\to]0, +\infty[$ continuous and $\rho > 0$ such that

$$f(x) \geq a + \gamma(d(x, [f \leq a])) \quad \text{for every } x \in B_\rho(\bar{x}) \backslash [f \leq a]\,.$$

We say that f has a *nonlinear global error bound* with respect to $[f \leq a]$ if the inequality holds for all $x \in [f > a]$ (that is, $B_\rho(\bar{x})$ is replaced by X).

Recall from [8] that for $x \in \operatorname{dom} f$, the *strong slope* of f at x is defined by

$$|\nabla f|(x) := \begin{cases} 0\,, & \text{if } x \text{ is a local minimum point of } f \\ \limsup\limits_{y \to x} \dfrac{f(x) - f(y)}{d(x, y)}\,, & \text{otherwise} \end{cases}\,.$$

(For $x \notin \operatorname{dom} f$, we set $|\nabla f|(x) := +\infty$). The strong slope allows developing general results in variational analysis, while it easily compares with (sub)derivative notions in the case when $(X, \|\cdot\|)$ is a (real) normed vector space—so that these results are both enlightening from the point of view of the abstract theory, and useful from the point of view of applications.

Let us recall a few such comparisons (some of which will be referred to later). Denoting by $(X^*, \|\cdot\|_*)$ the topological dual of $(X, \|\cdot\|)$, for $x \in \operatorname{dom} f$ and $u \in X \backslash \{0\}$, set:

$$f'(x; u) := \liminf_{t \to 0^+,\, v \to u} \frac{f(x + tv) - f(x)}{t}\,,$$

$$\partial^F f(x) := \left\{ x^* \in X^* \,:\, \liminf_{y \to x} \frac{f(y) - f(x) - \langle x^*, y - x \rangle}{\|y - x\|} \geq 0 \right\}\,,$$

which are the *lower contingent derivative* of f at x in the direction u and the *Fréchet subdifferential* of f at x, respectively. From the definitions, it is easy to see that

$$|\nabla f|(x) \geq \sup_{\|u\| = 1} \Big(-f'(x; u) \Big) \tag{1}$$

(with equality if X is finite dimensional), and

$$|\nabla f|(x) \leq d_*(0, \partial^F f(x)) := \inf\{\|x^*\| \,:\, x^* \in \partial^F f(x)\}\,, \tag{2}$$

with equality if f is convex, in which case, if x is not a minimum point of f we also have

$$|\nabla f|(x) = \sup_{y \neq x} \frac{f(x) - f(y)}{\|x - y\|}\,.$$

Roughly speaking, estimate (1) is of *primal* type (in applications, it is related to results involving *tangent* cones), while estimate (2) is of *dual* type (related to results involving *normal* cones). In general, a major estimate is

$$|\nabla f|(x) \geq \liminf_{(y, f(y)) \to (x, f(x))} d_*(0, \partial f(y))\,, \tag{3}$$

where ∂ is a subdifferential operator. From known considerations in "classical" non-smooth analysis, given a specific subdifferential operator ∂, (3) holds for any lower semicontinuous f whenever $(X, \|\cdot\|)$ belongs to some appropriate class of Banach spaces: See, e.g., [2, Section 4], and the beginning of Section 6 in [1] for details and references. For instance, if $\partial = \partial^F$, then (3) holds (for any lower semicontinuous f) if X is an Asplund space. (See, e.g., [9, Example 2.5] for many other examples).

3 Characterization of Linear Error Bounds

In order to obtain sufficient conditions in our error bound results, we use the following straightforward consequence of the variational principle (for which a major reference is [10]). See [1, Proposition 2.1] for a proof, and [1, Remark 2.1] for some comments.

Proposition 3.1 *Let (X, d) be a complete metric space, $f : X \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous, and U be a (nonempty) open subset of X. Then:*

$$f(x) - \inf_U f \geq \left(\inf_U |\nabla f|\right) d(x, X \backslash U) \quad \textit{for every } x \in U\,. \tag{4}$$

Theorem 3.1 *Let (X, d) be a complete metric space, $f : X \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous, $C \subset X$, $-\infty < a < b \leq +\infty$, and $\sigma, \rho > 0$. Assume that:*

$$|\nabla f|(x) \geq \sigma \quad \textit{for every } x \in B_\rho(C) \cap [a < f < b]\,. \tag{5}$$

The following two properties hold:

(a) *If $[f \leq a] \neq \emptyset$, then*

$$f(x) - a \geq \sigma d(x, [f \leq a]) \quad \textit{for every } x \in C \cap [a < f < b] \textit{ with } d(x, [f \leq a]) \leq \rho\,;$$

(b) *If* $C \cap [f < a+\sigma\rho] \neq \emptyset$, *then* $[f \leq a] \neq \emptyset$, *and*

$$f(x) - a \geq \sigma d(x, [f \leq a]) \quad \textit{for every } x \in C \cap [a < f < a+\sigma\rho] .$$

Proof Let $x \in C \cap [a<f<b]$, set $\widetilde{X} := [f \leq f(x)]$, let $\tilde{f}$ denote the restriction of f to $\widetilde{X}$, and set $U := B_\rho(C) \cap [\tilde{f} > a]$, so that

$$d(x, \widetilde{X} \backslash U) \geq \min\{\rho, d(x, [f \leq a])\} .$$

On the other hand,

$$\tilde{f}(z) = f(z) \text{ and } |\nabla \tilde{f}|(z) = |\nabla f|(z) \quad \text{for every } z \in \widetilde{X} ,$$

so that

$$f(x) - a \geq \sigma d(x, \widetilde{X} \backslash U)$$

according to (4, 5). Thus, if $d(x, [f \leq a]) \leq \rho$, then $d(x, \widetilde{X} \backslash U) \geq d(x, [f \leq a])$, which yields (a). If $f(x) - a < \sigma\rho$, then $d(x, \widetilde{X} \backslash U) < \rho$, which shows that $d(x, \widetilde{X} \backslash U) = d(x, [f \leq a])$ indeed, yielding (b) (in particular, $[f \leq a] \neq \emptyset$). □

Theorem 3.2 (Characterization of linear local error bound) *Let* (X, d) *be a complete metric space,* $f : X \to \mathbb{R} \cup \{+\infty\}$ *be lower semicontinuous,* $-\infty < a < b \leq +\infty$, $\bar{x} \in [f \leq a]$, *and* $\sigma > 0$. *Consider the following statements:*

(a) *There exists* $r > 0$ *such that*

$$|\nabla f|(x) \geq \sigma \quad \textit{for every } x \in B_r(\bar{x}) \cap [a < f < b] ; \tag{6}$$

(b) *There exists* $\rho > 0$ *such that*

$$f(x) - c \geq \sigma d(x, [f \leq c]) \quad \textit{for every } c \in [a, b[\textit{ and } x \in B_\rho[\bar{x}] \cap [c<f<b] . \tag{7}$$

Then, (a) $\Rightarrow$ (b) with $\rho := r/2$, *and* (b) $\Rightarrow$ (a) with $r := \rho$.

Proof (a) $\Rightarrow$ (b): Apply Theorem 3.1(a) with $C := B_{r/2}[\bar{x}]$, $\rho := r/2$, and arbitrary $c \in [a, b[$ in place of a.
(b) $\Rightarrow$ (a): This follows from the definition of the strong slope, as showed in [2, Proposition 2.1]. We give the details for the reader's convenience. Let $x \in B_\rho[\bar{x}] \cap [a<f<b]$ and $\tilde{\sigma} \in]0, \sigma[$. For $n \in \mathbb{N}$ such that $c_n := f(x) - 1/n \geq a$, let $x_n \in [f \leq c_n]$ be such that

$$f(x) - c_n \geq \tilde{\sigma} d(x, x_n) .$$

Then, $0 < d(x, x_n) \to 0$, so that x is not a local minimum point of f, and

$$\frac{f(x) - f(x_n)}{d(x, x_n)} \geq \frac{f(x) - c_n}{d(x, x_n)} \geq \tilde{\sigma} ,$$

so that $|\nabla f|(x) \geq \tilde{\sigma}$, whence $|\nabla f|(x) \geq \sigma$. □

Theorem 3.3 (Characterization of linear global error bound) *Let (X, d) be a complete metric space, $f : X \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous, $-\infty < a < b \leq \infty$ with $[a < f < b] \neq \emptyset$, and $\sigma > 0$. The following are equivalent:*

(a) $|\nabla f|(x) \geq \sigma$ *for every* $x \in [a < f < b]$;
(b) $[f \leq a] \neq \emptyset$, *and*

$$f(x) - c \geq \sigma d(x, [f \leq c]) \quad \textit{for every } c \in [a, b[\textit{ and every } x \in [c < f < b] .$$

Proof (a) $\Rightarrow$ (b): Apply Theorem 3.1(b) with $C := X$, arbitrary $c \in [a, b[$ in place of a, and arbitrary $\rho \in]0, (b - c)/\sigma[$ such that $[f < c + \sigma\rho] \neq \emptyset$.
(b) $\Rightarrow$ (a) is obtained like in the previous theorem. □

Remark 3.1 As far as we know, the first example of (linear) error bound appeared in Rosenbloom [11]. Anticipating the well-known result by Hofmann [12], [11, Lemme 3] provides an error bound for the distance to a convex polyhedral cone, and [11, Lemme 4] provides an error bound for the distance to a convex polyhedron, in a special case. In the general nonconvex case, the first sufficient conditions for local error bounds date back to [13] (more recently, see also [14]). The above three theorems gather and synthesize the approach developed in our earlier papers [2,4] where indeed, the local case was not presented as such (the emphasis was put on (local) *metric regularity*, see the coming remark about this notion). We thus provide here a unified and direct presentation of both cases.

Remark 3.2 Given $f : X \to \mathbb{R} \cup \{+\infty\}$, consider the multifunction

$$\operatorname{epi} f := \{(x, c) \in X \times \mathbb{R} : f(x) \leq c\} ,$$

that is, $\operatorname{epi} f(x) = [f(x), +\infty[$ if $x \in \operatorname{dom} f$ and $\operatorname{epi} f(x) = \emptyset$ otherwise. Under the assumptions (and using the notations) of Theorem 3.3, [6, Theorem 2] states:

The following properties are equivalent and imply that $[f \leq a] \neq \emptyset$:
(i) $|\nabla f|(x) \geq \sigma$ *for every* $x \in [a < f < b]$;
(ii) $(f(x) - c)^+ \geq \sigma d(x, [f \leq c])$ for all $(x, c) \in U := [a < f < b] \times]a, b[$;
(iii) $\operatorname{epi} f$ *is* σ*-metrically regular on* U.

Property (iii) means that for every $(\bar{x}, \bar{c}) \in U$, there exists $\varepsilon > 0$ such that

$$d(c, \operatorname{epi} f(x)) \geq \sigma d(x, (\operatorname{epi} f)^{-1}(c)) \quad \text{for all } (x, c) \in U \cap (B_\varepsilon(\bar{x}) \times]\bar{c} - \varepsilon, \bar{c} + \varepsilon[) .$$

Clearly, for $(x, c) \in X \times \mathbb{R}$ we have

$$(\operatorname{epi} f)^{-1}(c) = [f \leq c] \quad \text{and} \quad d(c, \operatorname{epi} f(x)) = (f(x) - c)^+,$$

so that (iii) reads: For every $(\bar{x}, \bar{c}) \in U$ we have

$$\text{(iii)}' \ (f(x) - c)^+ \geq \sigma d(x, [f \leq c]) \quad \text{for all } (x, c) \in U \cap (B_\varepsilon(\bar{x}) \times]\bar{c} - \varepsilon, \bar{c} + \varepsilon[) ,$$

which is *a priori* weaker than (ii). Moreover, while (i) $\Rightarrow$ (ii) is (a) $\Rightarrow$ (b) in Theorem 3.3 (yielding also $[f \le a] \neq \emptyset$), (iii)$'$ $\Rightarrow$ (i) is similar to (b) $\Rightarrow$ (a) in Theorem 3.2, that is, it depends only on the definition of the strong slope: There is no need to use Zorn's lemma as in [6]. In the sequel, we shall refer several times to [6], which contains many interesting observations on the matter we are dealing with. We already mention that, similarly as what is observed in [6, Remark 3], Theorem 3.3(b) is equivalent to

$$\text{(iv) } \sigma d_{\mathcal{H}}([f \le c_2], [f \le c_1]) \le c_2 - c_1 \quad \text{for all } a \le c_1 < c_2 < b\,,$$

where $d_{\mathcal{H}}$ is the *Hausdorff-Pompeiu metric*:

$$\begin{aligned} d_{\mathcal{H}}([f \le c_2], [f \le c_1]) &:= \max\{e_{\mathcal{H}}([f \le c_2], [f \le c_1]), e_{\mathcal{H}}([f \le c_1], [f \le c_2])\} \\ &= e_{\mathcal{H}}([f \le c_2], [f \le c_1]) := \sup\{d(x, [f \le c_1] : x \in [f \le c_2]\}. \end{aligned}$$

Remark 3.3 In [15, Definition 2], Chao and Cheng introduce the notion of *subslope* which is denoted and defined by

$$^{\dagger}|\nabla f|(x) := \sup_{y \in D(x)} \frac{f(x) - f(y)^{+}}{d(x, y)}$$

for $x \in \operatorname{dom} f$, and $^{\dagger}|\nabla f|(x) := +\infty$ for $x \notin \operatorname{dom} f$, where

$$D(x) := \{y \in X \,:\, d(y, [f \le 0]) \le d(x, [f \le 0])\}\,.$$

In [15, Theorem 2], the authors establish the following result (we use our notations):

Let (X, d) be a complete metric space, let $f : X \to \mathbb{R} \cup \{+\infty\}$ be proper and lower semicontinuous, and let $\sigma > 0$. The following are equivalent:
(i) $^{\dagger}|\nabla f|(x) \ge \sigma$ for every $x \in [f > 0]$.
(ii) $[f \le 0] \neq \emptyset$, and $f(x) \ge \sigma d(x, [f \le 0])$ *for every* $x \in [f > 0]$.

Indeed, (i) $\Rightarrow$ (ii) is contained in [14, Theorem 3] (see also [2, Remark 2.2 (b)]), while (ii) $\Rightarrow$ (i) readily follows from the definition of the subslope. Since it does not involve all values $c \ge 0$, this statement is formally simpler than that of Theorem 3.3. However, since the subslope is a *global* notion (as is the *global slope* introduced in [16]), we prefer to stick with results based on the strong slope, sharing Ioffe's opinion at the end of [17, Section 5], that the interest of the sufficient condition in Theorem 3.3 lies in the fact that it "gives a global estimate based on purely infinitesimal information."

We conclude this section with a "bilateral" result, where for $a \in \mathbb{R}$, $[f = a] := \{x \in X \,:\, f(x) = a\}$.

Proposition 3.2 *Let (X, d) be a complete metric space with connected balls, and $f : X \to \mathbb{R}$ be continuous with*

$$|\nabla f|(x) = |\nabla(-f)|(x) \quad \textit{for every } x \in X\,.$$

(a) *Let* $\bar{x} \in X$ *with* $a := f(\bar{x}) \in \mathbb{R}$, *and* $\sigma, \rho > 0$. *Assume that*

$$|\nabla f|(x) \geq \sigma \quad \textit{for every } x \in B_{2\rho}(\bar{x}) \backslash [f = a] \,.$$

Then,

$$|f(x) - a| \geq \sigma d(x, [f = a]) \quad \textit{for every } x \in B_{\rho}(\bar{x}) \,.$$

(b) *Let* $-\infty < a < b \leq +\infty$ *with* $[a < f < b] \neq \emptyset$ *and* $\sigma > 0$. *The following are equivalent:*

(i) $|\nabla f|(x) \geq \sigma$ *for every* $x \in [a < f < b]$;
(ii) $[f = c] \neq \emptyset$ *for every* $c \in [a, b[$, *and*

$$|f(x) - c| \geq \sigma d(x, [f = c]) \quad \textit{for all } (c, x) \in [a, b[\times [a < f < b] \,.$$

Proof Since X has connected balls and f is continuous, for $c \in \mathbb{R}$ such that $[f \leq c] \neq \emptyset$ and for $x \in [f > c]$ we have $d(x, [f \leq c]) = d(x, [f = c])$. Thus, the conclusions are obtained applying Theorems 3.2 and 3.3 to both f and $-f$. □

Remark 3.4 Observe that (ii) above is equivalent to

$$\text{(v) } \sigma d_{\mathcal{H}}([f = c_2], [f = c_1]) \leq c_2 - c_1 \quad \text{for every } a \leq c_1 < c_2 < b$$

(recall the end of Remark 3.2), which was observed in [6, Proposition 6], from which we borrowed the assumption that X has connected balls.

4 Characterization of Nonlinear Error Bounds

In this section, we establish a general (though elementary) chain rule for the strong slope that is used to reduce our main results on nonlinear error bounds to the linear ones in the previous section.

Lemma 4.1 *Let* (X, d) *be a metric space,* $f : X \to \mathbb{R} \cup \{+\infty\}$ *be lower semicontinuous,* $-\infty < a < b \leq +\infty$, *and* $\varphi :]a, b[\to \mathbb{R}$ *be of class* C^1 *with* $\varphi' > 0$. *Then,*

$$|\nabla(\varphi \circ f)|(x) = \varphi'(f(x))|\nabla f|(x) \quad \textit{for every } x \in [a < f < b] \,.$$

Proof Of course, $\varphi \circ f : [a < f < b] \to \mathbb{R}$. Let $x \in [a < f < b]$. We first show that $|\nabla(\varphi \circ f)|(x) \leq \varphi'(f(x))|\nabla f|(x)$. We may assume that $|\nabla f|(x) < +\infty$ and that $|\nabla(\varphi \circ f)|(x) > 0$. Thus, there exists a sequence (x_n) converging to x such that

$$|\nabla(\varphi \circ f)|(x) = \lim_{n \to \infty} \frac{\varphi(f(x)) - \varphi(f(x_n))}{d(x, x_n)} > 0 \,,$$

so that $f(x_n) < f(x)$ (because φ is increasing) and $f(x_n) \to f(x)$ (because $|\nabla f|(x) < \infty$). Let $z_n \in]f(x_n), f(x)[$ be such that

$$\varphi(f(x)) - \varphi(f(x_n)) = \varphi'(z_n)(f(x) - f(x_n)),$$

then

$$\begin{aligned}|\nabla(\varphi \circ f)|(x) &= \lim_{n\to\infty} \varphi'(z_n)\,\frac{f(x) - f(x_n)}{d(x, x_n)} \\ &= \varphi'(f(x)) \lim_{n\to\infty} \frac{f(x) - f(x_n)}{d(x, x_n)} \le \varphi'(f(x))|\nabla f|(x).\end{aligned}$$

We then show that $|\nabla(\varphi\circ f)|(x) \ge \varphi'(f(x))|\nabla f|(x)$. We may assume that $|\nabla f|(x) > 0$ and that $|\nabla(\varphi \circ f)|(x) < +\infty$. Thus, there exists a sequence (x_n) converging to x such that

$$|\nabla f|(x) = \lim_{n\to\infty} \frac{f(x) - f(x_n)}{d(x, x_n)} > 0,$$

so that (for n large enough) $0 < f(x_n) < f(x)$ (because f is lower semicontinuous at x). From which $\varphi(f(x_n)) < \varphi(f(x))$ (because φ is increasing) and $\varphi(f(x_n)) \to \varphi(f(x))$ (because $|\nabla(\varphi \circ f)|(x) < \infty$), so that $f(x_n) \to f(x)$ (because φ is increasing, again). Let $z_n \in]f(x_n), f(x)[$ be such that

$$\varphi(f(x)) - \varphi(f(x_n)) = \varphi'(z_n)(f(x) - f(x_n)),$$

then

$$|\nabla(\varphi \circ f)|(x) \ge \lim_{n\to\infty} \varphi'(z_n)\,\frac{f(x) - f(x_n)}{d(x, x_n)} = \varphi'(f(x))|\nabla f|(x).$$

□

Notation 4.1 Given $b \in [0, +\infty]$, we denote by Φ_b the set of continuous $\alpha :]0, b[\to]0, +\infty[$ such that for some (hence for all) $t \in]0, b[$ we have:

$$\varphi_\alpha(t) := \int_0^t \frac{ds}{\alpha(s)} < +\infty.$$

Thus, if $\alpha \in \Phi_b$, then $\varphi_\alpha :]0, b[\to]0, +\infty[$ is of class $\mathcal{C}^1$, is increasing, extends continuously to $[0, b[$ with $\varphi_\alpha(0) = 0$, and we may set:

$$\varphi_\alpha(b) := \sup_{0<t<b} \varphi_\alpha(t).$$

Keeping this notation in mind, the following are the two main results of this paper.

Theorem 4.1 (Characterization of nonlinear local error bound) *Let (X, d) be a complete metric space, $f : X \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous, $-\infty < a < b \le +\infty$, $\bar{x} \in [f \le a]$, and $\alpha \in \Phi_{b-a}$. Consider the following statements:*

(a) *There exists* $r > 0$ *such that*

$$|\nabla f|(x) \geq \alpha(f(x) - a) \quad \textit{for every } x \in B_r(\bar{x}) \cap [a < f < b]\,. \tag{8}$$

(b) *There exists* $\rho > 0$ *such that*

$$\varphi_\alpha(f(x) - a) \geq \varphi_\alpha(c - a) + d(x, [f \leq c])$$

for every $c \in [a, b[$ *and every* $x \in B_\rho[\bar{x}] \cap [c < f < b]$.
Then, (a) $\Rightarrow$ (b) with $\rho := r/2$, *and* (b) $\Rightarrow$ (a) with $r := \rho$.

Proof According to Lemma 4.1, (8) reads

$$|\nabla(\varphi_\alpha \circ f)|(x) = \varphi'_\alpha(f(x) - a)|\nabla f|(x) = \frac{|\nabla f|(x)}{\alpha(f(x) - a)} \geq 1$$

for every $x \in B_r(\bar{x}) \cap [a < f < b]$. Since $\varphi_\alpha : [0, b-a[\to [0, \varphi_\alpha(b-a)[$ is one-to-one increasing, the result readily follows from Theorem 3.2 applied to φ_α. □

In a similar way, we deduce from Theorem 3.3 the following characterization of a nonlinear global error bound.

Theorem 4.2 (Characterization of nonlinear global error bound) *Let* (X, d) *be a complete metric space,* $f : X \to \mathbb{R} \cup \{+\infty\}$ *be lower semicontinuous,* $-\infty < a < b \leq \infty$ *with* $[a < f < b] \neq \emptyset$, *and* $\alpha \in \Phi_{b-a}$. *The following are equivalent:*

(a) $|\nabla f|(x) \geq \alpha(f(x) - a)$ *for every* $x \in [a < f < b]$.
(b) $[f \leq a] \neq \emptyset$, *and*

$$\varphi_\alpha(f(x) - a) \geq \varphi_\alpha(c - a) + d(x, [f \leq c]) \quad \textit{for all } c \in [a, b[\textit{ and all } x \in [c < f < b]\,.$$

Remark 4.1 Recalling the last part of Remark 3.2, note that the assertions of Theorem 4.2 are equivalent to:

(c) $d_{\mathcal{H}}([f \leq c_2], [f \leq c_1]) \leq \varphi_\alpha(c_2 - a) - \varphi_\alpha(c_1 - a)$ for all $a \leq c_1 < c_2 < b$.

The main part in the above two theorems is of course the sufficient condition (a) $\Rightarrow$ (b), with $c := a$ in (b). In the case of Theorem 4.2, it reads (i) $\Rightarrow$ (ii) with:

(i) $|\nabla f|(x) \geq \alpha(f(x) - a)$ for every $x \in [a < f < b]$;
(ii) $[f \leq a] \neq \emptyset$, and $f(x) - a \geq \varphi_\alpha^{-1}(d(x, [f \leq a]))$ for every $x \in [a < f < b]$.

This was already proved in [3, Theorem 6.1], through a different argument, essentially based on property (c) above (an argument previously used in [4, Theorem 4.2] in a particular case). Here, we provide a more satisfactory (truly variational) approach, by reducing the nonlinear case to the linear one through Lemma 4.1, which naturally yields a *local* result as well (see [3, Remark 6.1 (b)]). We note that in the quoted results of [3,4], the function α is not assumed continuous, but it is assumed nondecreasing, in which case the function φ_α is *concave*. A typical such situation is

$$\varphi_\alpha(t) := \frac{1}{\sigma}\, t^\gamma \quad \text{with } \sigma > 0 \text{ and } 0 < \gamma < 1\,, \tag{9}$$

which we shall emphasize in the following section. As pointed out in [3, Section 6], the existence of a (continuous) nondecreasing $\alpha :]0, b-a[\to]0, +\infty[$ such that (i) above holds is equivalent to the fact that f is so-called *asymptotically well behaved* in $[a<f<b]$, that is: For every sequence $(x_k) \subset [a<f<b]$,

$$|\nabla f|(x_k) \to 0 \quad \Rightarrow \quad f(x_k) \to a\,.$$

See [3, Remark 6.1], and Example 5.3 below, for bibliographical comments on this notion.

Remark 4.2 A main step in the above proof of Theorem 4.1 is to write assumption (8) as

$$\varphi'(f(x)-a)|\nabla f|(x) \geq 1 \quad \text{for every } x \in B_r(\bar{x}) \cap [a<f<b] \tag{10}$$

(with $\varphi = \varphi_\alpha$). According to the terminology used in [6] (also used in [7]), inequality (10) could be termed a *generalized Kurdyka–Łojasiewicz inequality*. Indeed, such inequality was considered by Kurdyka in [18] in the case when f is a C^1 function defined on a Hilbert space X. The original inequality was established by Łojasiewicz in the 1960s, in the case when f is (sub)analytic on $\mathbb{R}^n$ and φ is of the type (9), an inequality that was used to establish the convergence of the trajectories of dynamical systems, see, e.g., the short survey [19].

5 Error Bounds of Order $\gamma > 0$ and Examples from the Literature

In this section, we specialize the sufficient parts of Theorems 4.1 and 4.2 to the case when $\alpha(s) := \tau s^\theta$ with $\tau > 0$ and $\theta < 1$. This is the case that has been essentially considered in the literature, which is natural due to its simplicity and its applicability. We thus give several examples of recent results that can be recovered from our general approach.

Corollary 5.1 *Let (X, d) be a complete metric space, $f : X \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous, $-\infty < a < b \leq +\infty$, and $\sigma, \gamma > 0$.*

(a) *Let $\bar{x} \in [f \leq a]$, and $\rho > 0$. Assume that*

$$|\nabla f|(x) \geq \frac{\sigma}{\gamma}(f(x)-a)^{1-\gamma} \quad \textit{for every } x \in B_{2\rho}(\bar{x}) \cap [a<f<b]\,.$$

Then,

$$(f(x)-a)^\gamma \geq \sigma d(x, [f \leq a]) \quad \textit{for every } x \in B_\rho[\bar{x}] \cap [a<f<b]\,.$$

(b) *Assume that $[a<f<b] \neq \emptyset$, and that*

$$|\nabla f|(x) \geq \frac{\sigma}{\gamma}(f(x)-a)^{1-\gamma} \quad \textit{for every } x \in [a<f<b]\,.$$

Then, $[f \le a] \neq \emptyset$ *and*

$$(f(x) - a)^\gamma \ge \sigma d(x, [f \le a]) \quad \textit{for every } x \in [a < f < b] .$$

Example 5.1 Let $(X, \langle \cdot , \cdot \rangle)$ be a real Hilbert space, let $A : X \to X$ be a bounded self-adjoint operator, and define $f : X \to \mathbb{R} \cup \{+\infty\}$ by

$$f(x) := \langle Ax, x \rangle \quad \text{if } \|x\| = 1 ,$$

$f(x) := +\infty$ otherwise. It is immediate that for $x \in \operatorname{dom} f$:

$$\partial^F f(x) = 2Ax + \mathbb{R}x$$

(of course, we identify X with its dual), from which it is readily seen that

$$d(0, \partial^F f(x)) = 2\|Ax - \langle Ax, x \rangle x\| = 2\|Ax - f(x)x\| .$$

Let $\lambda_0 = \inf_X f$ be the lower bound of the spectrum of A. When λ_0 is an eigenvalue of A, $\operatorname{argmin} f$ is the intersection of the associated eigenspace X_0 with the unit sphere. Assume that λ_0 is isolated in the spectrum of A, let $X_1 := X_0^\perp$, and let λ_1 be the lower bound of the spectrum of the restriction of A to X_1, so that $\lambda_0 < \lambda_1$. We have (see, e.g., [20, (10.11)]):

$$\|Ax - \lambda_0 x\|^2 \ge (\lambda_1 - \lambda_0)\langle Ax - \lambda_0 x, x \rangle = (\lambda_1 - \lambda_0)(f(x) - \lambda_0) \quad \text{for } x \in \operatorname{dom} f ,$$

so that for $x \in [f < \lambda_1]$ we obtain

$$\begin{aligned} d(0, \partial^F f(x))^2 = 4\|Ax - f(x)x\|^2 &= 4\|Ax - \lambda_0 x - (f(x)x - \lambda_0 x)\|^2 \\ &= 4(\|Ax - \lambda_0 x\|^2 - (f(x) - \lambda_0)^2) \\ &\ge 4(\lambda_1 - f(x))(f(x) - \lambda_0) \end{aligned}$$

(see [20, Lemma 10.2]). Thus, given $\lambda \in]\lambda_0, \lambda_1[$, we have

$$d(0, \partial^F f(x))^2 \ge 4(\lambda_1 - \lambda)(f(x) - \lambda_0) \quad \text{for every } x \in [f \le \lambda] ,$$

or, according to (3):

$$|\nabla f|(x) \ge 2(\lambda_1 - \lambda)^{\frac{1}{2}} (f(x) - \lambda_0)^{\frac{1}{2}} > 0 \quad \text{for every } x \in [\lambda_0 < f \le \lambda] .$$

In all the results quoted from the literature, in the examples below, we freely use our notations in order to provide a more straightforward comparison with our results. We note that all these results were obtained using the variational principle.

Example 5.2 In [21, Corollary 2 *i), ii)*], Ngai and Théra establish the following, which is their main result on nonlinear error bounds:

Let X be an Asplund space, and let $f : X \to \mathbb{R}\cup\{+\infty\}$ be lower semicontinuous. Then:
i) If there are $\sigma, \gamma > 0$ such that $\gamma\|x^*\| f(x)^{\gamma-1} \geq \sigma$ for all $x \notin [f \leq 0]$, $x^* \in \partial^F f(x)$, then

$$d(x, [f \leq 0]) \leq \frac{1}{\sigma} f(x)_+^\gamma \quad \text{for all } x \in X\,.$$

ii) Let $\bar{x}$ be in the boundary of $[f \leq 0]$. Let us suppose that there exist real$\sigma, \gamma, \rho > 0$ such that $\gamma\|x^*\| f(x)^{\gamma-1} \geq \sigma$, for all $x \in B_{2\rho}(\bar{x})\backslash[f \leq 0]$, $x^* \in \partial^F f(x)$, then

$$d(x, [f \leq 0]) \leq \frac{1}{\sigma} f(x)_+^\gamma \quad \text{for all } x \in B_\rho(\bar{x})\,.$$

Indeed, as recalled earlier, (3) holds in the above setting, so that for $x \notin [f \leq 0]$ (resp., $x \in B_{2\rho}(\bar{x})\backslash[f \leq 0]$) we have:

$$|\nabla f|(x) \geq \liminf_{(y, f(y)) \to (x, f(x))} d_*(0, \partial^F f(x)) \geq \frac{\sigma}{\gamma} f(x)^{1-\gamma}\,,$$

yielding the assumptions of Corollary 5.1. Note that the assumption that $\bar{x}$ belong to the *boundary* of $[f \leq 0]$, though not essential in an abstract result, is the interesting case in practice since it yields that $B_r(\bar{x}) \cap [f > 0] \neq \emptyset$ for any $r > 0$.

Example 5.3 In [22, Corollary 6.5.4], Facchinei and Pang establish the following result:

Let $S \subset \mathcal{D} \subset \mathbb{R}^n$ with S closed convex and $\mathcal{D}$open, let $f : \mathcal{D} \to \mathbb{R}$ be of class C^1 with $a := \inf_S f \in \mathbb{R}$, and let $\gamma \in]0, 1]$ and $\delta > 0$. Assume that

$$d(-\nabla f(x), \mathcal{N}(x; S)) \geq \delta(f(x) - \mu)^{1-\gamma} \quad \text{for every } x \in [f > a]\,. \tag{11}$$

The following statements hold:
(a) For every sequence $(x_k) \subset S$,

$$\lim_{k\to\infty} d(-\nabla f(x_k), \mathcal{N}(x_k; S)) = 0 \quad \Rightarrow \quad \lim_{k\to\infty} f(x_k) = a\,.$$

(b) $[f \leq a] \neq \emptyset$.
(c) $(f(x) - a)^\gamma \geq \gamma^\delta d(x, [f \leq a])$ for every $x \in S$.

Indeed, let $g := f + \iota_S$, where $\iota(x) := 0$ if $x \in S$ and $\iota(x) := +\infty$ if $x \notin S$. From elementary calculus, we have

$$\partial^F g(x) = \nabla f(x) + \mathcal{N}(x; S) \quad \text{for every } x \in S\,,$$

where $\mathcal{N}(x; S) := \{z \in \mathbb{R}^n : \langle z, y - x\rangle \leq 0 \ \forall\, y \in S\}$ is the normal cone to S at x. Thus, $d(0, \partial^F g(x)) = d(-\nabla f(x), \mathcal{N}(x; S))$, so that

$$|\nabla g|(x) \geq \delta(f(x)-a)^{1-\gamma} \quad \text{for every } x \in S\,,$$

according to (3) and (11), and (b), (c) follow from Corollary 5.1(b).

Observe that property (a) is immediate from (11). Facchinei and Pang mention it in order to relate their result with the notion of *asymptotical good behavior* (of f in $[f>0]$) introduced by Auslender and Crouzeix in [23] for the study of non-coercive convex minimization problems on $\mathbb{R}^n$. Such property might be called a *concrete* asymptotical good behavior property, to be checked in applications (using (sub)differential notions fitting the problem). Note that, due to (3), such concrete property implies the abstract one (using the strong slope) mentioned at the end of Remark 4.1, so that the abstract theory may indeed be applied to the problem under study.

Example 5.4 In [17, Theorem 7], Ioffe establishes the following result:

Let X be a complete metric space, let $f : X \to \mathbb{R}\cup\{+\infty\}$ be lower semicontinuous, let U be an open subset of X, let β be a 1-Lipschitz function on X which is positive on U, and set

$$U_\beta := \bigcup_{x\in U} B_{\beta(x)}(x)\,.$$

Assume that for some $k \geq 1$ and $\sigma > 0$:

$$|\nabla^{(k)} f|(x) > \sigma \quad \text{for every } x \in [f>0] \cap U_\beta\,.$$

Then, $[f \leq 0] \neq \emptyset$ and

$$f(x)^{\frac{1}{k}} \geq \sigma d(x,[f\leq 0]) \quad \text{for every} x \in U \text{ with } 0 < f(x) < (\sigma\beta(x))^k\,.$$

Here, $|\nabla^{(k)} f| := |\nabla[f]^{\frac{1}{k}}|$, where $[f]^{\frac{1}{k}}(x) := |f(x)|^{\frac{1}{k}}\,\mathrm{sign}(f(x))$. Indeed, in order to obtain that $[f \leq 0] \neq \emptyset$, it must first be assumed that there *is* some $x \in U$ with $0 < f(x) < (\sigma\beta(x))^k$. Then, set $\rho := \beta(x)$. According to the definition of $|\nabla^{(k)} f|$, we have

$$|\nabla[f]^{\frac{1}{k}}|(y) \geq \sigma \quad \text{for every } y \in B_\rho(x) \cap [0<[f]^{\frac{1}{k}}<\sigma\rho]\,,$$

and the conclusion follows from Theorem 3.1(b), applied to $[f]^{\frac{1}{k}}$ with $C := \{x\}$. Note that the inequality need not be strict in the main assumption, while we may consider any constant $k > 0$, i.e., we need not restrict ourselves to the "concave case" $k \geq 1$ (an assumption used in Ioffe's proof for the computation of $|\nabla^{(k)} f|$).

We conclude this section with a nonlinear version of Proposition 3.2(a).

Proposition 5.1 *Let (X,d) be a complete metric space with connected balls, and $f : X \to \mathbb{R}$ be continuous with*

$$|\nabla f|(x) = |\nabla(-f)|(x) \quad \textit{for every } x \in X\,.$$

Let $\bar{x} \in X$ with $a := f(\bar{x}) \in \mathbb{R}$, and $\sigma, \rho > 0$. Assume that

$$|\nabla f|(x) \geq \frac{\sigma}{\gamma}|f(x) - a|^{1-\gamma} \quad \textit{for every } x \in B_{2\rho}(\bar{x})\backslash[f = a]\,.$$

Then,

$$|f(x) - a|^{\gamma} \geq \sigma d(x, [f = a]) \quad \textit{for every } x \in B_{\rho}(\bar{x})\,.$$

Remark 5.1 Let f be analytic in a neighborhood of $0 \in \mathbb{R}^n$, with $f(0) = 0$. In [19], Łojasiewicz mentions that the following assertions are equivalent:

(i) There exist $\theta \in]0, 1[$ and $r > 0$ such that

$$\|\nabla f(x)\| \geq |f(x)|^{\theta} \quad \text{for every } x \in B_r(0)\,;$$

(ii) There exist $\sigma, \gamma, \rho > 0$ such that

$$|f(x)| \geq \sigma d(x, [f{=}0])^{\gamma} \quad \text{for every } x \in B_{\rho}[0]$$

Of course, implication (ii) $\Rightarrow$ (i) relies on the fact that f is analytic, and is definitely not true in general, in the continuous case.

6 Remarks on the Convex Case

In this section, $(X, \|\cdot\|)$ is a Banach space with topological dual $(X^*, \|\cdot\|_*)$, $f : X \to \mathbb{R} \cup \{+\infty\}$ is convex and lower semicontinuous, and for $x \in \operatorname{dom} f$,

$$\partial f(x) := \{x^* \in X^* \,:\, f(z) - f(x) \geq \langle x^*, z - x\rangle\}$$

denotes the Fenchel subdifferential of f at x.

Proposition 6.1 *For $x \in \operatorname{dom} f$, we have*

$$|\nabla f|(x) = d_*(0, \partial f(x))\,,$$

and if x is not a minimum point of f we have

$$|\nabla f|(x) = \sup_{f(z) < f(x)} \frac{f(x) - f(z)}{\|x - z\|}\,.$$

Thus, if $a \in \mathbb{R}$ is such that $[f \leq a] \neq \emptyset$, we have:

$$|\nabla f|(x) \geq \frac{f(x) - a}{d(x, [f \leq a])} \quad \textit{for every } x \in [f > a]\,.$$

Proof See, e.g., [2, Proposition 3.1] and [3, Proposition 5.2]. □

Let $-\infty < a < b \leq +\infty$ with $[f \leq a] \neq \emptyset$, and assume that there exists an increasing, continuous $\psi :]0, b-a[\to]0, +\infty[$ such that

$$\psi(f(x) - a) \geq d(x, [f \leq a]) \quad \text{for every } x \in [a < f < b],$$

so that

$$|\nabla f|(x) \geq \frac{f(x) - a}{\psi(f(x) - a)} \quad \text{for every } x \in [a < f < b],$$

according to Proposition 6.1. Thus, if

$$\varphi(t) := \int_0^t \frac{\psi(s)}{s}\, ds < +\infty \quad \text{for } t \in]0, b-a[,$$

then f satisfies the inequality:

$$\varphi'(f(x) - a)|\nabla f|(x) \geq 1 \quad \text{for every } x \in [a < f < b].$$

This was observed in [6, Theorem 30], in the case when X is a Hilbert space (the Hilbertian structure being used in the proof), and $a := \min f$.

Theorem 6.1 *Let X be a Banach space, $f : X \to \mathbb{R} \cup \{+\infty\}$ be convex and lower semicontinuous, $-\infty < a < b \leq +\infty$, and $0 < \gamma < 1$. Consider the following statements:*

(a) *There exists $r > 0$ such that*

$$|\nabla f|(x) \geq r(f(x) - a)^{1-\gamma} \quad \textit{for every } x \in [a < f < b];$$

(b) *$[f \leq a] \neq \emptyset$, and there exists $\rho > 0$ such that*

$$(f(x) - a)^{\gamma} \geq \rho\, d(x, [f \leq a]) \quad \textit{for every } x \in [a < f < b].$$

Then, (a) $\Rightarrow$ (b) with $\rho := \gamma r$, *and* (b) $\Rightarrow$ (a) with $r := \rho$.

Proof (a) $\Rightarrow$ (b) is given by Corollary 5.1, (b) $\Rightarrow$ (a) follows from the above considerations with $\psi(s) := s^{\gamma}/\rho$. □

Note that in the previous result, and with respect to Corollary 5.1, we restricted γ to be less than 1. The reason is that in the convex case, "the slope is nondecreasing with altitude," so that an inequality of the type

$$|\nabla f|(x) \geq \alpha(f(x) - a) \quad \text{for every } x \in [a < f < b]$$

should involve a nondecreasing function α. Let us be more specific about this.

Proposition 6.2 *Let X be a Banach space, $f : X \to \mathbb{R} \cup \{+\infty\}$ be convex and lower semicontinuous, and $a \in \mathbb{R}$ with $[f \le a] \neq \emptyset$. Then,*

$$\inf_{[f>a]} |\nabla f| \ge \inf_{[f=a]} |\nabla f| .$$

Proof See [2, Proposition 3.2]. (Note that $[f = a] \neq \emptyset$.) □

Proposition 6.3 *Let X be a Banach space, $f : X \to \mathbb{R} \cup \{+\infty\}$ be convex and lower semicontinuous, and $-\infty < a < b \le +\infty$. Assume that there exists a continuous $\alpha :]0, b-a[\to]0, +\infty[$ with*

$$\int_0^t \frac{ds}{\alpha(s)} < +\infty \quad \textit{for } t \in]0, b-a[,$$

such that

$$|\nabla f|(x) \ge \alpha(f(x) - a) \quad \textit{for every } x \in [a<f<b] .$$

Then, there exists a nondecreasing, continuous $\tilde{\alpha} :]0, +\infty[\to]0, +\infty[$ with

$$\int_0^t \frac{ds}{\tilde{\alpha}(s)} < +\infty \quad \textit{for } t > 0 ,$$

such that

$$|\nabla f|(x) \ge \tilde{\alpha}(f(x) - a) \quad \textit{for every } x \in [f>a] .$$

Proof Of course, we assume $[a<f<b] \neq \emptyset$. For $s > 0$, set

$$\hat{\alpha}(s) := \inf\{|\nabla f|(x) \ : \ f(x) - a = s\} .$$

Then, $\hat{\alpha} \not\equiv +\infty$ (Proposition 3.1 readily implies that $\mathrm{dom}|\nabla f|$ is dense in $\mathrm{dom} f$), and $\hat{\alpha}$ is nondecreasing, according to Proposition 6.2. By definition of $\hat{\alpha}$, we also have

$$|\nabla f|(x) \ge \hat{\alpha}(f(x) - a) \ge \alpha(f(x) - a) > 0 \quad \text{for every } x \in [a<f<b] ,$$

whence

$$\int_0^t \frac{ds}{\hat{\alpha}(s)} \le \int_0^t \frac{ds}{\alpha(s)} < +\infty \quad \text{for } t \in]0, b-a[.$$

Standard inf-convolution arguments then produce a continuous, nondecreasing $\tilde{\alpha} :]0, +\infty[\to]0, +\infty[$ with $\tilde{\alpha} \le \hat{\alpha}$ and

$$\int_0^t \frac{ds}{\tilde{\alpha}(s)} < +\infty \quad \text{for } t \in]0, b-a[,$$

and the result follows. □

Remark 6.1 This result was obtained in [6, Theorem 29] (where it is formulated in terms of the generalized Kurdyka–Łojasiewicz inequality), with X a Hilbert space (and $a := \inf f$), through more involved arguments. Note that the "standard inf-convolution arguments" mentioned in the above proof can be found, e.g., in [6, Lemma 45].

7 A Further Comparison and an Application to Descent Methods

Let $-\infty < a < b \leq +\infty$, and let $\alpha \in \Phi_{b-a}$ (recall Notation 4.1) be *nondecreasing*, so that $\varphi :]0, b-a[\to]0, +\infty[$ defined by

$$\varphi(t) := \int_0^t \frac{ds}{\alpha(s)}$$

is of class C^1, increasing, and concave, so that $\beta :]0, \varphi(b-a)[\to]0, +\infty[$ defined by

$$\beta(s) := (\varphi^{-1})'(s)$$

is continuous and increasing.

Proposition 7.1 *Let (X, d) be a complete metric space, $f : X \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous, and a b, α, φ, β be as above.*

(a) *Assume that for some $\bar{x} \in [f \leq a]$ and $\rho > 0$ we have*

$$|\nabla f|(x) \geq \alpha(f(x) - a) \quad \textit{for every } x \in B_{2\rho}(\bar{x}) \cap [a < f < b]\,.$$

Then,

$$|\nabla f|(x) \geq \beta(d(x, [f \leq a])) \quad \textit{for every } x \in B_\rho[\bar{x}] \cap [a < f < b]\,. \tag{12}$$

(b) *Assume that*

$$|\nabla f|(x) \geq \alpha(f(x) - a) \quad \textit{for every } x \in [a < f < b]\,.$$

Then, $[f \leq a] \neq \emptyset$ and

$$|\nabla f|(x) \geq \beta(d(x, [f \leq a])) \quad \textit{for every } x \in [a < f < b]\,. \tag{13}$$

Proof From Theorem 4.1, we have

$$f(x) - a \geq \varphi^{-1}(d(x, [f \leq a])) \quad \text{for every } x \in B_\rho[\bar{x}] \cap [a < f < b]\,.$$

Thus, since φ' is nonincreasing, for every $x \in B_\rho[\bar{x}] \cap [a<f<b]$ we have

$$\begin{aligned}\frac{|\nabla f|(x)}{\beta(d(x,[f\le a]))} &= \varphi'(\varphi^{-1}(d(x,[f\le a])))|\nabla f|(x)\\ &\ge \varphi'(f(x)-a)|\nabla f|(x) = \frac{|\nabla f|(x)}{\alpha(f(x)-a)} \ge 1\,.\end{aligned}$$

Similarly, (b) follows from Theorem 4.2. □

Remark 7.1 Properties (12) and (13) are the main assumptions in [1,3], where nonlinear error bound results are derived from the linear case through a "change of metric." Though quite natural, that technique is more involved than the one used in this paper, so that the implications above are only fitting. Proposition 7.1(b) was not observed in [3], nor was it observed there that the global result [3, Theorem 4.3] indeed holds without assuming that the function β be *nondecreasing*, a result we now establish, after which we provide some comparison between the two types of results.

Theorem 7.1 *Let (X,d) be a complete metric space, $f : X \to \mathbb{R}\cup\{+\infty\}$ be lower semicontinuous, $-\infty < a < b \le \infty$, and $\beta : [0,+\infty[\to [0,+\infty[$ be continuous, with $\beta(s) > 0$ for $s > 0$, and such that*

$$\int_0^{+\infty} \beta(s)\,ds = +\infty\,. \tag{14}$$

Assume that $[f\le a] \ne \emptyset$, and that

$$|\nabla f|(x) \ge \beta(d(x,[f\le a])) \quad \textit{for every } x \in [a<f<b]\,.$$

Then,

$$f(x)-a \ge \int_0^{d(x,[f\le a])} \beta(s)\,ds \quad \textit{for every } x \in [a<f<b]\,.$$

Proof Assume first that $\beta(0) = 0$. We may assume that $[a<f<b] \ne \emptyset$, so that for $c \in]a,b[$ sufficiently close to b and for $\delta > 0$ small enough, we have $[f\le c]\backslash C_\delta \ne \emptyset$, where

$$C_\delta := \{x \in X \,:\, d(x,[f\le a]) \le \delta\}\,.$$

For such δ, define $\beta_\delta : [0,\infty[\to]0,\infty[$ by

$$\beta_\delta(s) := \begin{cases}\beta(\delta) \text{ if } s \le \delta\\ \beta(s) \text{ if } s \ge \delta\end{cases},$$

and consider the metric $\tilde{d} = \tilde{d}([f\le a],\beta_\delta)$ given by the so-called change-of-metric principle: See [24, Theorem 4.1], or [25, Theorem 2.2]. Then, $\tilde{d}$ is topologically

equivalent to d, $(X, \tilde{d})$ is complete (due to (14)), and for every $x \in U := [f \le c] \setminus C_\delta$:

$$\tilde{d}(x, [f \le a]) = \int_0^{d(x,[f \le a])} \beta_\delta(s)\, ds \ge \int_\delta^{d(x,[f \le a])} \beta(s)\, ds\,,$$

and

$$|\widetilde{\nabla} \tilde{f}|(x) = |\widetilde{\nabla} f|(x) = \frac{|\nabla f|(x)}{\beta_\delta(d(x, [f \le a]))} = \frac{|\nabla f|(x)}{\beta(d(x, [f \le a]))} \ge 1\,,$$

where $|\widetilde{\nabla}\cdot|$ denotes the strong slope with respect to the metric $\tilde{d}$, and $\tilde{f}$ denotes the restriction of f to $[f \le c]$. Thus, applying Proposition 3.1 in $([f \le c], \tilde{d})$ to the function $\tilde{f}$, we obtain that for every $x \in U$:

$$f(x) - a \ge f(x) - \inf_U f \ge \tilde{d}(x, X \setminus U) \ge \tilde{d}(x, [f \le a]) \ge \int_\delta^{d(x,[f \le a])} \beta(s)\, ds\,,$$

and the conclusion follows by letting $\delta \to 0$ first, then $c \to b$.

In the case when $\beta(0) > 0$, the proof is just simpler, since we may directly use the metric $\tilde{d} := \tilde{d}([f \le a], \beta)$. □

Remark 7.2 This result extends [3, Theorem 4.3], where β is assumed nondecreasing. Comparing Theorem 7.1 with Theorem 4.2, observe that in the former, $[f \le a] \ne \emptyset$ is an assumption (not a conclusion). Indeed, this set is involved in the definition of the modified metric. Consequently, Theorem 7.1 does not provide the characterization obtained in Theorem 4.2. Let us mention that property (12) is naturally linked with the notion of *metric (sub)regularity*, see, e.g., [1], see also [26] for a quite recent survey on this notion.

We conclude the paper by illustrating how the results of Sect. 4 can be applied to the convergence of descent methods. This example is borrowed from the papers [5,7], as we explain in more detail in Remark 7.3 below.

Example 7.1 Let $(X, \|\cdot\|)$ be a Banach space, let U be an open subset of X, and let $f : U \to \mathbb{R}$ be differentiable, with a uniformly continuous differential $Df : U \to X^*$. Set

$$m(s) := \sup\{\|Df(x) - Df(z)\|_* \,:\, \|x - z\| \le s\}\,,$$

so that $m :]0, \infty[\to]0, \infty[$ is nondecreasing, finite for small s, with $m(s) \to 0$ as $s \to 0$. It follows from the mean value inequality that

$$|f(x) - f(z) - Df(z)(x - z)| \le M(\|x - z\|) \quad \text{whenever } [x, z] \subset U\,, \tag{15}$$

where $M(t) := \int_0^t m(s)\, ds \le t m(t)$. Let us consider the following algorithm.

Parameters: $\tau, \theta \in]0, 1[$.

Step 1. *Initialization*. Choose $x_0 \in U$ and $\delta > 0$ such that $m(\delta)$ is finite. We assume that

$$\{x \in X \,:\, d(x, [f \leq f(x_0)]) \leq \delta\} \subset U \,.$$

Step 2. *Stopping test*. Given iterate x_k, stop if $Df(x_k) = 0$. Otherwise, let $\|d_k\| = 1$ be such that

$$Df(x_k)(d_k) \leq -\tau \|Df(x_k)\|_* . \tag{16}$$

Step 3. *Linesearch*. If $f(x_k + \delta d_k) \leq f(x_k) + \theta\delta\, Df(x_k)(d_k)$, set $t_k := \delta$, $x_{k+1} := x_k + t_k d_k$ and go to **Step 2**. Otherwise, by a classical backtracking argument, find $t_k \in]0, \delta/2]$ such that

$$f(x_k + t_k d_k) \leq f(x_k) + \theta t_k Df(x_k)(d_k) \,, \tag{17}$$

$$f(x_k + 2t_k d_k) > f(x_k) + 2\theta t_k Df(x_k)(d_k) \,, \tag{18}$$

set $x_{k+1} := x_k + t_k d_k$ and go to **Step 2**.

From (16) to (17), since $t_k = \|x_{k+1} - x_k\|$ we have

$$f(x_{k+1}) \leq f(x_k) - \tau\theta \, \|x_{k+1} - x_k\| \, \|Df(x_k)\|_* \,, \tag{19}$$

while from (15) to (18), if $t_{k+1} < \delta$ we have

$$\begin{aligned} 0 &< f(x_k + 2t_k d_k) - f(x_k + t_k d_k) - \theta t_k Df(x_k)(d_k) \\ &\leq M(t_k) + t_k(Df(x_k + t_k d_k) - Df(x_k))(d_k) + (1-\theta) t_k Df(x_k)(d_k) \\ &\leq 2t_k m(t_k) - t_k \tau (1-\theta) \|Df(x_k)\|_* \,, \end{aligned}$$

that is,

$$2m(t_k) \geq \tau(1-\theta)\|Df(x_k)\|_* \,. \tag{20}$$

(a) Assume that the sequence of iterates is infinite [so that the sequence $(f(x_k))$ is decreasing according to (19], and has a subsequence converging to some $\bar{x} \in [f < f(x_0)]$ (which is the case, e.g., if $X = \mathbb{R}^n$ and $[f < f(x_0)]$ is bounded), and set $a := f(\bar{x})$. Assume further that for some $r > 0$ and $b > a$ we have:

$$\|Df(x)\|_* \geq \alpha(f(x) - a) \quad \text{for every } x \in B_r(\bar{x}) \cap [a < f < b] \,, \tag{21}$$

for some nondecreasing $\alpha \in \Phi_{b-a}$ (see Notation 4.1). If $x_k \in B_r(\bar{x}) \cap [a<f<b]$, then

$$\tau\theta \, \|x_{k+1} - x_k\| \leq \frac{f(x_k) - f(x_{k+1})}{\alpha(f(x_k) - a)} \leq \varphi_\alpha(f(x_k) - a) - \varphi_\alpha(f(x_{k+1}) - a)$$

due to (19), (21), and the fact that α is nondecreasing. Consequently, given $0 < \rho \leq \min\{r, b/\tau\theta\}$, if $k_1 < k_2$ are such that $\varphi_\alpha(f(x_{k_1}) - a) < \tau\theta\rho$ and $x_k \in B_\rho(\bar{x})$ for $k = k_1, \ldots, k_2 - 1$, then

$$\tau\theta \sum_{k=k_1}^{k_2-1} \|x_{k+1} - x_k\| \leq \varphi_\alpha(f(x_{k_1}) - a) < \tau\theta\rho$$

(note that $\varphi_\alpha(f(x_k) - a) < \varphi_\alpha(f(x_{k_1}) - a)$ for $k > k_1$), which shows that $x_{k_2} \in B_{2\rho}(\bar{x})$, whence

$$\bar{x} = \lim_{k\to\infty} x_k \,. \tag{22}$$

Moreover, it follows from (20) that

$$Df(\bar{x}) = \lim_{k\to\infty} Df(x_k) = 0 \,.$$

Also, Theorem 4.1 and Proposition 7.1 yield

$$d(x_k, [f = a]) \leq \varphi_\alpha(f(x_k) - a) \,,$$
$$\beta(d(x_k, [f = a])) \leq \|Df(x_k)\|_* \leq \frac{2m(t_{k+1})}{\tau(1-\theta)}$$

for all k large enough, where $\beta := (\varphi_\alpha^{-1})'$ (recall from the proof of Proposition 3.2 that $d(x_k, [f \leq a]) = d(x_k, [f = a])$).
(b) Assume now that $U = X$, that f is bounded below, and that, letting $a := \inf f < b := f(x_0)$, there exists a nondecreasing $\alpha \in \Phi_{b-a}$ such that

$$\|Df(x)\|_* \geq \alpha(f(x) - a) \quad \text{for every } x \in [a < f < b] \,. \tag{23}$$

Theorem 4.2 and Proposition 7.1 yield that $\operatorname{argmin} f = [f = a] = [f \leq a]$ is nonempty and

$$\varphi_\alpha(f(x) - a) \geq d(x, [f = a]) \quad \text{for every } x \in [a < f < b] \,,$$

$$\|Df(x)\|_* \geq \beta(d(x, [f = a])) \quad \text{for every } x \in [a < f < b] \,,$$

where $\beta := (\varphi_\alpha^{-1})'$. Of course, $f(x_k) = a$ if $Df(x_k) = 0$. On the other hand, if the sequence of iterates (x_k) is infinite, for any $K \in \mathbb{N}$ we have (similarly as before):

$$\tau\theta \sum_{k=1}^{K} \|x_{k+1} - x_k\| \leq \varphi_\alpha(f(x_1) - a) < +\infty$$

due to (19) and (23), which shows that (x_k) converges to some $\bar{x} \in [f < b]$, and that $t_k = \|x_{k+1} - x_k\| \to 0$, so that $Df(\bar{x}) = 0$ due to (20), which shows that $f(\bar{x}) = a$.

Remark 7.3 The above example synthesizes some results developed in [5, Sects. 3, 4.1] and in [7, Section. 3] (where $X = \mathbb{R}^n$). In particular, the very nice argument leading to (22) is given in the proof of [5, Theorem 3.2], which applies to Theorem 3.4 therein, which is itself expressed with a condition of the type of (21). Still, in [5] as well as in [7], much emphasis is put on the Kurdyka–Łojasiewicz property, which, as already said, is but another formulation of (21) (in the local case), or of (23) (in the global case):

$$\varphi'(f(x) - a)\|Df(x)\|_* \geq 1 \quad \text{for every } x \in [a < f < b]\,.$$

We note that this condition was used for the study of convergence of descent methods, back in [27, p. 49], with $\varphi(t) := \sigma\sqrt{t}$ for some $\sigma > 0$.

We finally observe that in [7], it is assumed that the function f is of class $C^{1,1}$. Assuming only that Df is uniformly continuous seems as good and natural. Moreover, with respect to [5,7], we added the error bound estimates provided by our abstract results.

8 Conclusions

With respect to earlier work of ours, we refined the abstract approach to nonlinear error bounds for lower semicontinuous functions on complete metric spaces, in the case when the main assumptions are lower estimates of the strong slope with respect to the values of the function. Such estimates are linked to the notion of asymptotically well-behaved functions, and to the Kurdyka–Łojasiewicz inequality, so that our approach can be used in various settings, as we illustrated through many examples from the literature. We hope this can help develop systematic approaches that may simplify and clarify various theoretical and technical aspects of nonsmooth analysis and optimization.

References

1. Azé, D., Corvellec, J.-N.: Nonlinear local error bounds via a change of metric. J. Fixed Point Theory Appl. **16**, 351–372 (2014)
2. Azé, D., Corvellec, J.-N.: Characterizations of error bounds for lower semicontinuous functions on metric spaces. ESAIM Control Optim. Calc. Var. **10**, 409–425 (2004)
3. Corvellec, J.-N., Motreanu, V.V.: Nonlinear error bounds for lower semicontinuous functions on metric spaces. Math. Program., Ser. A **114**, 291–319 (2008)
4. Azé, D., Corvellec, J.-N., Lucchetti, R.E.: Variational pairs and applications to stability in nonsmooth analysis. Nonlinear Anal. **49**(2002), 643–670 (2002)
5. Absil, P.-A., Mahony, R., Andrews, B.: Convergence of the iterates of descent methods for analytic cost functions. SIAM J. Optim. **16**, 531–547 (2005)

6. Bolte, J., Daniilidis, A., Ley, O., Mazet, L.: Characterizations of Łojasiewicz inequalities: subgradient flows, talweg, convexity. Trans. Am. Math. Soc. **362**, 3319–3363 (2010)
7. Noll, D., Rondepierre, A.: Convergence of linesearch and trust-region methods using the Kurdyka-Łojasiewicz inequality. In: Computational and Analytical Mathematics, Springer Proc. Math. Stat., vol. 50, pp. 593–611. Springer-Verlag, New York (2013)
8. De Giorgi, E., Marino, A., Tosques, M.: Problemi di evoluzione in spazi metrici e curve di massima pendenza. Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur **68**, 180–187 (1980)
9. Azé, D.: A unified theory for metric regularity of multifunctions. J. Convex Anal. **13**, 225–252 (2006)
10. Ekeland, I.: On the variational principle. J. Math. Anal. Appl. **47**, 324–353 (1974)
11. Rosenbloom, P.C.: Quelques classes de problèmes extrémaux. Bull. Soc. Math. France **79**, 1–58 (1951)
12. Hoffman, A.J.: On approximate solutions of systems of linear inequalities. J. Res. Nat. Bur. Stand. **49**, 263–265 (1952)
13. Ioffe, A.D.: Regular points of Lipschitz functions. Trans. Am. Math. Soc. **251**, 61–69 (1979)
14. Wu, Z., Ye, J.: First-order and second-order conditions for error bounds. SIAM J. Optim. **14**, 621–645 (2003)
15. Chao, M., Cheng, C.: Linear and nonlinear error bounds for lower semicontinuous functions. Optim. Lett. **8**, 1301–1312 (2014)
16. Ngai, H.V., Théra, M.: Error bounds in metric spaces and applications to the perturbation stability of metric regularity. SIAM J. Optim. **19**, 1–20 (2008)
17. Ioffe, A.D.: Nonlinear regularity models. Math. Program., Ser. B **139**, 223–242 (2013)
18. Kurdyka, K.: On gradients of functions definable in o-minimal structures. Ann. Inst. Fourier (Grenoble) **48**, 769–783 (1998)
19. Łojasiewicz, S.: On semi-analytic and subanalytic geometry. In: Panoramas of Mathematics (Warsaw 1992/94). Banach Center Publ, vol. 34, pp. 89–104. Polish Acad. Sci, Warsaw (1995)
20. Lyubich, YuI, Maistrowskii, G.D.: The general theory of relaxation processes for convex functionals. Russ. Math. Surv. **25**, 57–117 (1970)
21. Ngai, H.V., Théra, M.: Error bounds for systems of lower semicontinuous functions in Asplund spaces. Math. Program., Ser. B **116**, 397–427 (2009)
22. Facchinei, F., Pang, J.-S.: Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer Series in Operations Research, vol. I. Springer, New York (2003)
23. Auslender, A.: Well behaved asymptotical convex functions. Ann. Inst. H. Poincaré, Anal. Non Linéaire **6**, 101–121 (1989)
24. Corvellec, J.-N.: Quantitative deformation theorems and critical point theory. Pac. J. Math. **187**, 263–279 (1999)
25. Corvellec, J.-N., Hantoute, A.: Homotopical stability of isolated critical points of continuous functionals. Set-Valued Anal. **10**, 143–164 (2002)
26. Ioffe A.D.: Metric regularity. Theory and applications—a survey. arXiv:1505.07920v1
27. Pchenitchny, B., Daniline, Y.: Numerical Methods in Extremal Problems. Mir, Moscow (1978)