MATH. SCAND. 47 (1980), 123–134

LIPSCHITZ r-CONTINUITY OF THE APPROXIMATE SUBDIFFERENTIAL OF A CONVEX FUNCTION

J.-B. HIRIART-URRUTY

Abstract.

Earlier results on the continuity of the approximate subdifferential and on the local Lipschitz property for $\varepsilon > 0$ of the ε-subdifferential of a convex function are improved and generalized in various ways.

0. Purpose and scope.

The approximate subdifferential has been proved to be an useful tool in convex analysis, from the theoretical viewpoint as well as for practical purposes. Given a convex function f defined on E, the approximate subdifferential $\partial_. f(.)$ which assigns to $(x, \varepsilon) \in E \times \mathbf{R}_+$ the so-called ε-subdifferential of f at x turns out to have remarkable properties. The properties of the subdifferential $x \rightrightarrows \partial f(x)$ (i.e. for $\varepsilon = 0$) are widely known in convex analysis ([9], [11], [14]). As for the ε-subdifferential $\partial_\varepsilon f$, whose definition is just a "perturbation by ε" of that of ∂f, it enjoys, for $\varepsilon > 0$, some noteworthy properties which are hopeless for the "exact" subdifferential.

One of the main reasons for the nice behavior of the multifunction $\partial_. f(.)$ on $E \times \mathbf{R}_+^*$ is that $\partial_\varepsilon f(x)$ *is not a local notion.* To be more explicit, let us look at the definitions. The subdifferential $\partial f(x_0)$ of a convex function f at x_0 is the set of $x^* \in E^*$ satisfying

$$f(x) \geqq f(x_0) + \langle x^*, x - x_0 \rangle \tag{0.1}$$

for all x. Actually, $\partial f(x_0)$ is a local notion in the sense that it suffices to know the function f in a neighborhood of x_0 to get all of $\partial f(x_0)$. In other words, the set of $x^* \in E^*$ satisfying (0.1) for x in a neighborhood of x_0 is exactly $\partial f(x_0)$. Here, the convexity of f makes that the local character is converted in a global one. The same does not hold for the ε-subdifferential. $\partial_\varepsilon f(x_0)$ is the set of x^* satisfying

$$f(x) \geqq f(x_0) + \langle x^*, x - x_0 \rangle - \varepsilon \tag{0.2}$$

Received November 6, 1979; in revised form February 12, 1980.

for all x. Now, $\partial_\varepsilon f(x_0)$ may be very "sensitive" to the variations of f, even when those variations do not hold in a neighborhood of x_0. As for an example, let $f: x \mapsto |x|$ and $x_0 = 0$. If one modifies this function only on $|x| \geqq \varrho$, the ε-subdifferential at x_0 is altered. The necessity, a priori, of knowing the behavior of f on all of E for the calculation of $\partial_\varepsilon f(x_0)$ is corroborated by the fundamental approximation result due to Brøndsted and Rockafellar [4]. Roughly speaking, this result states that the more precisely you want to know $\partial_\varepsilon f(x_0)$, the farther from x_0 you need to know $\partial f(x)$. However, this disadvantage is weighed against the good effects wrought by the perturbation by ε. Let us recall some results which are typical in that respect. In [3], Brøndsted gave the exact formulation of the subdifferential of the supremum of two lower-semicontinuous convex functions f_1 and f_2 in terms of ε-subdifferentials for f_1 and f_2, without assuming any further continuity property on f_1 or f_2. In an another context, the use of the ε-directional derivative $f'_\varepsilon(x_0; .)$ as a substitute for the usual directional derivative has been proved advantageous in many algorithms of convex optimization; see [2, 10] and references therein. As the continuity property of $\partial_. f(.)$ as a multifunction of both x and ε, the main result goes back to the comprehensive study by Asplund and Rockafellar [1]. In particular, they proved that for a lower-semicontinuous convex function defined on a Banach space, the approximate subdifferential $\partial_. f(.)\colon (x, \varepsilon) \rightrightarrows \partial_\varepsilon f(x)$ was continuous in the Hausdorff sense (i.e. for the Hausdorff-distance h) on $(\operatorname{dom} f)^\circ \times \mathbf{R}_+^*$. More recently, Nurminskii [13] showed the following locally Lipschitzian behavior of the ε-subdifferential: given a convex function $f\colon \mathbf{R}^n \to \mathbf{R}$, $\varepsilon > 0$ and a compact set K, there exists a constant k such that

$$h(\partial_\varepsilon f(x), \partial_\varepsilon f(x')) \leqq \frac{k}{\varepsilon} \|x - x'\| \tag{0.3}$$

whenever x and x' lie in K. This result is rather surprising because, for the subdifferential, it is known that nothing more than the upper-semicontinuity can be claimed whereas the ε-subdifferential enjoys one of the strongest properties which can be required on a multifunction.

Our study is along the lines of the Lipschitz properties of the approximate subdifferential. It is divided into three parts. In the first part, we list some basic calculus rules on the ε-subdifferential. Besides their own interest since they cover the case where $\varepsilon = 0$, they will be of use in the next sections. Part II is devoted to a smoothing operation, namely to the infimal convolution of a convex function with a norm. From the viewpoint of the study of the ε-subdifferential, performing this operation has regularizing effects. In particular, it will be shown that, to a certain extent, an adequate Lipschitz function can be substituted for the given function. Part III contains the main results. We begin by proving the Lipschitz behavior of $\partial_{.} f(.)$ as a multifunction of x and $\varepsilon > 0$ for a Lipschitz convex function f. The proof which hinges on Lipschitz properties of the ε-directional derivatives and on Hörmander's relation is quite simple. The smoothing operating studied in Part II then allows us to consider the general case and to handle unbounded ε-subdifferentials. For unbounded subsets, the Hausdorff distance has no meaning and thus must be adapted. In the way of "quantifying" convergence for closed convex sets, Salinetti and Wets [16] were led to the notion of r-distance. For $r > 0$, the r-distance between two subsets C and D of a metric space E is given by

$$h_r(C, D) = h(C \cap rB, D \cap rB) \tag{0.4}$$

where B denotes the unit closed ball in E. The r-convergence (i.e. the convergence in the sense of h_r, for $r \geqq r_0$) quantifies, to a certain extent, convergence in the sense of Kuratowski [16, Theorem 4]. Actually, we shall prove what we called "Lipschitz r-continuity of the approximate subdifferential", that is to say a Lipschitz behavior of $h_r\ (\partial_{.} f(.), \partial_{.} f(.))$ as a function of both x and $\varepsilon > 0$. So, the given result improves the continuity result quoted above and generalizes in various ways Nurminskii's result. Finally, consequences to first-order approximations of convex functions are indicated.

We assume that the reader is familiar with basic definitions and properties from convex analysis [9, 11, 14].

I. Basic calculus rules on the ε-subdifferential.

Let E be a locally convex Hausdorff topological vector space with dual E^*. The canonical bilinear form on $E \times E^*$ is denoted by $\langle .,. \rangle$. Throughout, we shall use the following notations:

$\mathrm{Conv}_0\ (E)$ for the set of *proper convex* functions (a function f is said to be proper if f is not identically equal to $+\infty$ and if $f(x) > -\infty$ for all x),

$\Gamma_0(E)$ for the set of functions in $\mathrm{Conv}_0\ (E)$ which are *lower-semicontinuous* (l.s.c.).

Given a proper function f, the *ε-subdifferential* of f at $x_0 \in \mathrm{dom}\, f$ (dom f is the set where f is finite) is defined for each $\varepsilon \geqq 0$ as the set of vectors $x^* \in E^*$ satisfying

$$f(x) \geqq f(x_0) + \langle x^*, x - x_0 \rangle - \varepsilon \tag{1.1}$$

for all $x \in E$.

The set of such vectors, denoted by $\partial_\varepsilon f(x_0)$, is a $\sigma(E^*, E)$-closed convex set in E^* which reduces to the subdifferential for $\varepsilon = 0$. Moreover, if $f \in \Gamma_0(E)$, $\partial_\varepsilon f(x_0)$ is nonempty whenever ($x_0 \in \mathrm{dom}\, f$ and) $\varepsilon > 0$. There are two fundamental ways of characterizing $\partial_\varepsilon f(x)$: through the conjugate function f^* and with its support function. Let us recall the first one which every computation concerning the ε-subdifferentials hinges on.

PROPOSITION 1.1. *$x^* \in E^*$ belongs to $\partial_\varepsilon f(x_0)$ if and only if we have:*

$$f(x_0)+f^*(x^*)-\langle x_0, x^*\rangle \leqq \varepsilon . \tag{1.2}$$

If $f \in \Gamma_0(E)$, f and $f^*(\in \Gamma_0(E^*))$ play a symmetric role; then (1.2) is also equivalent to: $x_0 \in \partial_\varepsilon f^*(x^*)$.

EXAMPLE 1. Let $f \in \Gamma_0(E)$ be positively homogeneous (i.e. $f(\lambda x)=\lambda f(x)$ for all $x \in E$ and $\lambda>0$). For such an f, f^* is the indicator of $\partial f(o)$. Hence, $\partial_\varepsilon f(o) = \partial f(o)$ for all $\varepsilon \geqq 0$.

Consider now two proper functions f and g; one defines a function h in the following way:

$$h(x) = \inf_{\substack{x_1+x_2 \in E \\ x_1+x_2=x}} \{f(x_1)+g(x_2)\} .$$

h is said to be the *infimal convolution* of f and g, and we shall use the notation $h = f \nabla g$. The infimal convolution is said to be exact at $x_0 = x_0^1 + x_0^2$ if one has

$$f(x_0^1)+g(x_0^2) = \min_{\substack{u, v \in E \\ u+v=x_0}} \{f(u)+g(v)\} .$$

General properties of the infimal convolution, particularly those related to convex analysis, may be found in [11, § 3], [9, § 6.5] or [14].

PROPOSITION 1.2. *Let f, g be two proper functions, let $x_0 = x_0^1 + x_0^2$ be a point where the infimal convolution is finite and exact. Then*

$$\partial_\varepsilon(f \nabla g)(x_0) = \bigcup_{\substack{\varepsilon_1 \geqq 0, \varepsilon_2 \geqq 0 \\ \varepsilon_1+\varepsilon_2=\varepsilon}} \{\partial_{\varepsilon_1} f(x_0^1) \cap \partial_{\varepsilon_2} g(x_0^2)\} . \tag{1.3}$$

PROOF. We have that $(f \nabla g)(x_0) = f(x_0^1)+g(x_0^2)$. Moreover, $(f \nabla g)^*$ equals f^*+g^* [9, Theorem 6.5.4]. Therefore, accordingly to Proposition 1.1, $x^* \in \partial_\varepsilon(f \nabla g)(x_0)$ if and only if

$$f(x_0^1)+f^*(x^*)-\langle x_0^1, x^*\rangle + g(x_0^2)+g^*(x^*)-\langle x_0^2, x^*\rangle \leqq \varepsilon .$$

Hence, equality (1.3) is easily deduced.

For the sake of completeness, let us note the dual version of the proposition above.

PROPOSITION 1.3. *Let $f, g \in \Gamma_0(E)$. Assume there exists a point at which f and g are finite, one of them being continuous at this point. Then*

$$\partial_\varepsilon(f+g)(x_0) = \bigcup_{\substack{\varepsilon_1 \geqq 0, \varepsilon_2 \geqq 0 \\ \varepsilon_1+\varepsilon_2=\varepsilon}} \{\partial_{\varepsilon_1} f(x_0) + \partial_{\varepsilon_2} g(x_0)\} \tag{1.4}$$

for all $x_0 \in \operatorname{dom} f \cap \operatorname{dom} g$.

PROOF. Under the assumptions made on f and g, $(f+g)^*$ is $f^* \nabla g^*$ and the infimal convolution is exact at all x^* of E^* [9, Theorem 6.5.8]. Then (1.4) is derived through the characterization (1.2).

REMARK. If $E = \mathbf{R}^n$, assumptions different from those made in Propositions 1.2 and 1.3 yield the same results. Conditions are expressed in terms of relative interiors of domains of functions involved in the problem; these conditions themselves can be deleted if the functions are polyhedral. For all these materials, see [14].

II. Smoothing a convex function.

In this Section, we suppose that E is a *Banach space*. The norm (function) on E will be denoted by $\|.\|$ and the associate dual norm on E^* by $\|.\|_*$. Let $f \in \operatorname{Conv}_0(E)$ and let r *be positive*. We are interested in the function $f_r = f \nabla r\|.\|$. Due to the properties of the infimal convolution, f_r appears as the result of a sort of "regularization". Actually, f_r does have some peculiar properties that we describe now. First, observe that $f_r(x) < +\infty$ for all $x \in E$. Hence, since f_r is a convex function (as the infimal convolution of two convex functions), there are only two possibilities: either f_r identically equals $-\infty$ or f_r is finite everywhere. In the latter case, we note the following "regularizing effect".

PROPOSITION 2.1. *If there exists $\bar{x} \in E$ at which $f_r(\bar{x}) > -\infty$, then f_r is Lipschitz on E with Lipschitz constant r.*

PROOF. [7, Corollary 1].

The problem, however, is to give conditions ensuring that there exists a point where f_r is finite. For that purpose, let $\operatorname{rg}(\partial_\varepsilon f)$ denote (for $\varepsilon \geqq 0$) the range of $\partial_\varepsilon f$, that is to say

$$\operatorname{rg}(\partial_\varepsilon f) = \bigcup_{x \in \operatorname{dom} f} \partial_\varepsilon f(x).$$

PROPOSITION 2.2. *The following statements are equivalent:*
(i) *there exists* $\bar{x} \in E$ *at which* $f_r(\bar{x}) > -\infty$,
(ii) *for all* $\varepsilon > 0$, *there exists* $x^* \in \operatorname{rg}(\partial_\varepsilon f)$ *such that* $\|x^*\|_* \leqq r$,
(iii) *there exist* $\varepsilon \geqq 0$ *and* $x^* \in \operatorname{rg}(\partial_\varepsilon f)$ *such that* $\|x^*\|_* \leqq r$.

PROOF. (i) $\Rightarrow$ (ii). According to Proposition 2.1, f_r is (convex and) Lipschitz on E with Lipschitz constant r. The conjugate of f_r is $f^* + \delta(.\,|\,rB^*)$ where B^* denotes the closed unit ball in E^*. Hence, the (nonempty) subdifferential $\partial f_r(\bar{x})$ can be characterized as follows:

$$(2.1) \qquad \partial f_r(\bar{x}) = \{x^* \in rB^*, \ f^*(x^*) + f_r(\bar{x}) - \langle \bar{x}, x^* \rangle = 0\}.$$

Now, for $\varepsilon > 0$, let us consider x_0 satisfying

$$(2.2) \qquad f(x_0) + r\|x_0 - \bar{x}\| \leqq f_r(\bar{x}) + \varepsilon.$$

From (2.1) and (2.2), one easily checks that $\partial f_r(\bar{x}) \subset \partial_\varepsilon f(x_0)$.
. (ii) $\Rightarrow$ (iii). Obvious.
. (iii) $\Rightarrow$ (i). Let $\varepsilon \geqq 0$, $\bar{x} \in E$ and $x^* \in \partial_\varepsilon f(\bar{x})$ with $\|x^*\|_* \leqq r$.

Following the definition of the ε-subdifferential, we have that

$$\begin{aligned} f(x) &\geqq f(\bar{x}) + \langle x^*, x - \bar{x} \rangle - \varepsilon \\ &\geqq f(\bar{x}) - r\|x - \bar{x}\| - \varepsilon \text{ for all } x \in E. \end{aligned}$$

Hence, $f_r(\bar{x}) \geqq f(\bar{x}) - \varepsilon$ and (i) is proved.

Besides the interesting equivalence (i) $\Leftrightarrow$ (ii), we note that the condition (iii) written for $\varepsilon = 0$, i.e.

(H_0) "there exists $x^* \in \operatorname{rg}(\partial f)$ such that $\|x^*\|_* \leqq r$",

is a *sufficient* condition for (i) to hold. However, this condition is not necessary, as shown by the next example.

EXAMPLE 2. Let $f \in \Gamma_0(\mathbf{R})$ be defined by $f(x) = 1/x$ if $x > 0$ and $+\infty$ if $x \leqq 0$. By taking $r = 0$, we get a function f_0 identically null on $\mathbf{R}$. However, $0 \notin \operatorname{rg}(\partial f)$.

For the "smoothing operation" $f \rightsquigarrow f \nabla r\|.\|$ we are concerned with, we define $C_r(f)$ as the *coincidence set of* f *and* f_r, that is to say

$$C_r(f) = \{x \in E, \ f_r(x) = f(x)\}.$$

Actually, $C_r(f)$ can be expressed in a more usable way, in connection with the assumption (H_0).

PROPOSITION 2.3. *The coincidence set of f and f_r is the "inverse image" of rB^* by the multifunction ∂f, i.e.*

$$C_r(f) = \{x \in E,\ \partial f(x) \cap rB^* \neq \emptyset\} . \tag{2.3}$$

Moreover,

$$\partial_\varepsilon f_r(x_0) = \partial_\varepsilon f(x_0) \cap rB^* \tag{2.4}$$

for all $x_0 \in C_r(f)$ and all $\varepsilon \geqq 0$.

PROOF. If f_r identically equals $-\infty$, rg $(\partial f) \cap rB^*$ is empty (Proposition 2.2). Hence, (2.3) trivially holds.

Suppose now that $f_r(x) > -\infty$ for all x. Let $\bar{x}$ be such that $\partial f(\bar{x})$ meets rB^*. We proved earlier that $f_r(\bar{x}) = f(\bar{x})$ (let $\varepsilon = 0$ in the third part of the proof of Proposition 2.2).

Conversely, if $f_r(\bar{x}) = f(\bar{x})$, the infimal convolution $f \nabla r\|.\|$ is exact at $\bar{x} = \bar{x} + 0$ and, according to Proposition 1.2, we have that

$$\partial f_r(\bar{x}) = \partial f(\bar{x}) \cap rB^* .$$

Hence the result is proved since $\partial f_r(\bar{x})$ is nonempty.

As for the result (2.4), see Example 1 and Proposition 1.2.

As for an example, we calculate the ε-subdifferential of the distance function. Let S be a nonempty convex subset of E and $x_0 \in S$. The distance function d_S is nothing more than $\delta(.\,|S) \nabla \|.\|$; hence, for all $\varepsilon \geqq 0$, we have that

$$\partial_\varepsilon d_S(x_0) = N_\varepsilon(S;\, x_0) \cap B^*$$

where $N_\varepsilon(S;\, x_0)$ designates the set of $x^* \in E^*$ satisfying $\langle x^*, x - x_0 \rangle \leqq \varepsilon$ for all $x \in S$.

REMARK 1. f and its l.s.c. hull f^{**} both yield the same smoothed version f_r. Hence, one can suppose, without lack of generality, that $f \in \Gamma_0(E)$. In such a case, $C_r(f)$ is *closed* and can be viewed as the image of rB^* by ∂f^*. Moreover, if $rB^* \subset (\operatorname{dom} f^*)^\circ$ (that expresses an inf-compactness condition on f), $C_r(f)$ is *$\sigma(E, E^*)$-compact and connected.*

REMARK 2. Proposition 2.3 also points out that, under suitable assumptions, the knowledge of $\partial_\varepsilon f(x)$ for $x \in S$ and $\varepsilon \in [0, \bar{\varepsilon}]$ amounts to the knowledge of the ε-subdifferential of a more workable function, namely of a Lipschitz function. To be more precise, let S be a nonempty subset of dom f and let $\bar{\varepsilon} > 0$. We assume that

(H$_1$) $\sup\{\|x^*\|_*, \ x^* \in \partial_{\bar\varepsilon} f(x), \ x \in S\} = r < +\infty$.

Then f_r is Lipschitz on E with Lipschitz constant r and

$$f_r(x) = f(x) \qquad \forall x \in S ,$$

$$\partial_\varepsilon f_r(x) = \partial_\varepsilon f(x) \quad \forall x \in S, \ \forall \varepsilon \in [0, \bar\varepsilon] .$$

Actually, (H$_1$) is assumed in order to get a *finite* smoothed function f_r. Along the same lines, let us indicate what happens in the general case. As before, let $S \subset \operatorname{dom} f$ and $\bar\varepsilon > 0$. We set

$$f_S = f \nabla \delta^*(.\,|B)$$

where B designates $\bigcup_{x\in S} \partial_{\bar\varepsilon} f(x)$. Then one easily checks that

$$f_S(x) = f(x) \qquad \forall x \in S$$

$$\partial_\varepsilon f_S(x) = \partial_\varepsilon f(x) \quad \forall x \in S, \ \forall \varepsilon \in [0, \bar\varepsilon] .$$

As an illustration of the results of this Section, let us look at the following simple example on R.

EXAMPLE 3. Let $f \in \Gamma_0(\mathrm{R})$ be defined by $f(x) = -\sqrt{x}$ if $x \geqq 0$ and $+\infty$ if $x < 0$. f_r is finite if and only if $r > 0$. For each $r > 0$, the coincidence interval of f and f_r is $I_r = [1/4r^2, +\infty[$.

According to Remark 2 above, the calculation of $\partial_\varepsilon f(x)$ for $x \in S$ and $\varepsilon \in [0, \bar\varepsilon]$ amounts to the computation of $\partial_\varepsilon f_r(x)$ for a certain Lipschitz function f_r, provided that $S \subset \mathrm{R}_+^*$.

III. Lipschitz r-continuity of the approximate subdifferential.

Let us denote by $\mathscr{C}_{\sigma,b}(E^*)$ the collection of all nonempty $\sigma(E^*, E)$-closed bounded convex subsets of E^*. We recall that the Hausdorff-topology on $\mathscr{C}_{\sigma,b}(E^*)$ is the topology in which, for each $C \in \mathscr{C}_{\sigma,b}(E^*)$, the sets of the form

$$\{D \in \mathscr{C}_{\sigma,b}(E^*) \mid D \subset C + \alpha B^* \text{ and } C \subset D + \alpha B^*\}$$

constitute a fundamental system of neighborhoods of C as α ranges over R_+^*. Given $f \in \Gamma_0(E)$, $\partial_\varepsilon f(x)$ is $\sigma(E^*, E)$-bounded whenever $x \in (\operatorname{dom} f)^\circ$. So, the multifunction $(x, \varepsilon) \rightrightarrows \partial_\varepsilon f(x)$, restricted to $(\operatorname{dom} f)^\circ \times \mathrm{R}_+$ ranges over $\mathscr{C}_{\sigma,b}(E^*)$. The following result asserting the *continuity* of the approximate subdifferential is due to Asplund and Rockafellar [1, pp. 456–458].

THEOREM 3.1. *Let $f \in \Gamma_0(E)$. Then the multifunction $(x, \varepsilon) \rightrightarrows \partial_\varepsilon f(x)$ is continuous from $(\operatorname{dom} f)^\circ \times \mathrm{R}_+^*$ (endowed with the strong topology) into $\mathscr{C}_{\sigma,b}(E^*)$ in the Hausdorff-topology.*

In the theorem below, we shall "quantify" this continuity, and what actually will be proved is a *local Lipschitz property* of the approximate subdifferential on $(\operatorname{dom} f)^\circ \times \mathbf{R}_+^*$.

The Hausdorff-topology on $\mathscr{C}_{\sigma,b}(E^*)$ can be defined by a metric (the so-called Hausdorff-distance h) but, for our purposes, it will be more convenient to use the Hausdorff-topology such as defined in a dual way. In fact, the cornerstone relation turns out to be the *Hörmander's equality* [8]:

$$\forall C, D \in \mathscr{C}_{\sigma,b}(E^*): \ h(C,D) = \sup_{\|d\| \leqq 1} |\delta^*(d \mid C) - \delta^*(d \mid D)| \, .$$

In view of results of Section II, we begin by convex *Lipschitz* functions.

THEOREM 3.2. *Let f: $E \to \mathbf{R}$ be a convex Lipschitz function. Then there exists k such that*

$$h(\partial_\varepsilon f(x), \partial_{\varepsilon'} f(x')) \leqq \frac{k}{\min(\varepsilon, \varepsilon')} (\|x - x'\| + |\varepsilon - \varepsilon'|) \tag{3.1}$$

for all x, x' in E and all $\varepsilon, \varepsilon'$ in $\mathbf{R}_+^*$.

PROOF. Let d be any direction in B, unit ball of E. Given $x \in E$ and $\varepsilon \in \mathbf{R}_+$, define φ on $\mathbf{R}_+^*$ by

$$\lambda \mapsto \varphi(\lambda) = \frac{f(x+\lambda d) - f(x) + \varepsilon}{\lambda} \, .$$

Strictly speaking, φ depends on d, x and ε. The infimum of $\varphi(\lambda)$ over $\mathbf{R}_+^*$ is just $f'_\varepsilon(x; d)$, support function of $\partial_\varepsilon f(x)$ in the d direction [14, pp. 219–220].

Actually, due to the Lipschitz property of f (with Lipschitz constant $r > 0$), one easily checks that, for any positive α and ε, all the λ_α satisfying.

$$\inf_{\lambda > 0} \varphi(\lambda) \leqq \varphi(\lambda_\alpha) \leqq \inf_{\lambda > 0} \varphi(\lambda) + \alpha$$

belong to the interval $[\varepsilon/(2r+\alpha), +\infty[$, whatever $x \in E$ and $d \in B$. Let now x, $x' \in E$ and $\varepsilon, \varepsilon' \in \mathbf{R}_+^*$ be arbitrarily chosen. According to what has been claimed just above and to the Lipschitz property of f, we have

$$f'_\varepsilon(x; d) - f'_{\varepsilon'}(x'; d) \leqq \frac{2r+\alpha}{\varepsilon'} (2r\|x - x'\| + |\varepsilon - \varepsilon'|)$$

$$f'_{\varepsilon'}(x'; d) - f'_\varepsilon(x; d) \leqq \frac{2r+\alpha}{\varepsilon} (2r\|x - x'\| + |\varepsilon' - \varepsilon|)$$

for all $\alpha > 0$ and all $d \in B$.

Hence the result (3.1) is achieved with $k = \max(2r, 4r^2)$ by a straightforward application of Hörmander's relation.

THEOREM 3.3. *Let* $f \in \mathrm{Conv}_0\,(E)$, *let* S *be a subset of* dom f. *We suppose that, for a certain* $r_0 > 0$, $\partial f(x) \cap r_0 B^*$ *is nonempty for all* $x \in S$. *Then, for all* $r \geqq r_0$, *there exists* k_r *such that*

$$(3.2) \qquad h(\partial_\varepsilon f(x) \cap rB^*,\ \partial_{\varepsilon'} f(x') \cap rB^*) \leqq \frac{k_r}{\min(\varepsilon, \varepsilon')} (\|x - x'\| + |\varepsilon - \varepsilon'|)$$

for all x, x' *in* S *and* $\varepsilon, \varepsilon'$ *in* $\mathbf{R}_+^*$.

PROOF. In view of Proposition 2.3, for all $r \geqq r_0$, the smoothed function f_r coincides with f on S and $\partial_\varepsilon f_r(x) = \partial_\varepsilon f(x) \cap rB^*$ for all $\varepsilon \geqq 0$ and all $x \in S$. It remains to apply Theorem 3.2 to the function f_r.

The formula (3.2) can be looked at under different angles: Lipschitz r-continuity of $\partial_\varepsilon f(x)$ as a multifunction of ε when x is fixed and $\varepsilon \geqq \underline{\varepsilon} > 0$, Lipschitz r-continuity of $\partial_\varepsilon f(x)$ as multifunction of x when $\varepsilon > 0$ is fixed, and so on ... As a particular corollary, we shall derive Nurminskii's result on the (local) Lipschitz continuity of $\partial_\varepsilon f(x)$ as a multifunction of x.

COROLLARY 3.4. *Let* $f \in \mathrm{Conv}_0\,(E)$ *and assume that there exists a nonempty open set on which* f *is bounded above. Then for all* $\bar{\varepsilon} > 0$ *and for all compact* $K \subset (\mathrm{dom}\, f)^\circ$, *there exists* k *such that*

$$(3.3) \qquad h(\partial_\varepsilon f(x), \partial_\varepsilon f(x')) \leqq \frac{k}{\varepsilon} \|x - x'\|$$

for all x, x' *in* K *and all* $\varepsilon \in \,]0, \bar{\varepsilon}]$.

PROOF. Since f is bounded above on a nonempty open set, $(\mathrm{dom}\, f)^\circ$ is nonempty and f is locally Lipschitz on it. Therefore, there exist $\eta > 0$ and $\varrho > 0$ such that

$$K_\eta = \{u \mid \exists x \in K \text{ with } \|u - x\| < \eta\} \subset (\mathrm{dom}\, f)^\circ$$

and

$$\|f(x) - f(x')\| \leqq \varrho \|x - x'\| \quad \text{for all } x, x' \text{ in } K_\eta .$$

Consequently, there exists r such that

$$\bigcup_{\substack{\varepsilon \in]0, \bar{\varepsilon}] \\ x \in K}} \partial_\varepsilon f(x) \subset rB^* .$$

Hence we get the result (3.3) by a straightforward application of Theorem 3.3.

A consequence which is worth mentioning concerns *first-order approximations of convex functions.*

COROLLARY 3.5. *Let $f \in \Gamma_0(E)$ be bounded above on a nonempty open set. Consider a convex compact subset $K \subset (\operatorname{dom} f)^\circ$ and $\bar{\varepsilon} > 0$. Then, for all x, x' in K and all $\varepsilon \in]0, \bar{\varepsilon}]$, there exists $x_\varepsilon^* \in \partial_\varepsilon f(x)$ satisfying*

$$f(x) - f(x') = \langle x_\varepsilon^*, x' - x \rangle + \frac{O(1)}{\varepsilon} \|x - x'\|^2 . \tag{3.4}$$

PROOF. Let x and x' be arbitrarily chosen in K. According to the mean value theorem for convex functions (see, for example, Theorem 7 in [6]), there exist $x'' \in]x, x'[$ and $u^* \in \partial f(x'')$ such that

$$f(x) - f(x') = \langle u^*, x - x' \rangle . \tag{3.5}$$

Let now x_ε^* in $\partial_\varepsilon f(x)$ satisfying:

$$\|u^* - x_\varepsilon^*\|_* = \min_{x^* \in \partial_\varepsilon f(x)} \|u^* - x^*\|_*$$

(such a point exists since the norm function $\|.\|_*$ is $\sigma(E^*, E)$-inf-compact). According to Corollary 3.4, we have that

$$\|u^* - x_\varepsilon^*\|_* \leqq \frac{k}{\varepsilon} \|x'' - x\| .$$

Hence the announced result is proved.

REMARK. As a consequence of (3.4), note the following "adjustment" property: let $\{x_n\}$ and $\{x_n'\}$ be sequences converging in K to x_0; then there exists a sequence $\{x_n^*\}$ in $\partial_\varepsilon f(K)$ satisfying

$$\lim_{n \to \infty} \frac{\|f(x_n) - f(x_n') - \langle x_n^*, x_n - x_n' \rangle\|}{\|x_n - x_n'\|} = 0 . \tag{3.6}$$

This kind of adjustment is usually hopeless with sequences $\{x_n^*\} \subset \partial f(K)$. In the same vein, expansions like (3.4) are useful for subgradient methods. As for a very simple example, let f: $\mathsf{R} \to \mathsf{R}$ be a convex function; assume that $\{x_n\}$ is a bounded sequence and that $\{\varepsilon_n\}$ is a given positive sequence converging to 0. Then there exists $x_n^* \in \partial_{\varepsilon_n} f(x_n)$ such that

$$\frac{f(x_n + \lambda_n) - f(x_n)}{\lambda_n} = x_n^* + O(1) \frac{\lambda_n}{\varepsilon_n} .$$

So, by using only the values of the function, the quotient above provides an element of $\partial_{\varepsilon_n} f(x_n)$ with an accuracy which can be controlled.

REFERENCES

1. E. Asplund and R. T. Rockafellar, *Gradients of convex functions*, Trans. Amer. Math. Soc. 139 (1969), 443–467.
2. M. L. Balinski and P. Wolfe, editors, *Nondifferentiable Optimization, Math. Programming Study* 3, North-Holland Publ. co., Amsterdam - New York, 1975.
3. A. Brøndsted, *On the subdifferential of the supremum of two convex functions*, Maht. Scand. 31 (1972), 225–230.
4. A. Brøndsted and R. T. Rockafellar, *On the subdifferentiability of convex functions*, Proc. Amer. Math. Soc. 16 (1965), 605–611.
5. J.-B. Hiriart-Urruty, *Etude de quelques propriétés de la fonctionnelle moyenne et de l'inf-convolution continue en analyse convexe stochastique*, C.R. Acad. Sci. Paris, Série A 280 (1975), 129–132.
6. J.-B. Hiriart-Urruty, *A note on the mean value theorem for convex functions*, to appear in Boll. Un. Mat. Ital.
7. J.-B. Hiriart-Urruty, *Extension of Lipschitz functions*, J. Math. Anal. Appl., to appear.
8. L. Hörmander, *Sur la fonction d'appui des ensembles convexes dans un espace localement convexe*, Ark. Mat. 3 (1954), 181–186.
9. P.-J. Laurent, *Approximation et Optimisation*, Hermann, Paris, 1972.
10. C. Lemarechal and R. Mifflin, editors, *Nonsmooth Optimization*, I.I.A.S.A. Proceedings 3, Pergamon Press, 1978.
11. J.-J. Moreau, *Fonctionnelles convexes, Séminaire sur les équations aux dérivées partielles* II, Collège de France, 1966–1967.
12. E. A. Nurminskii, *Continuity of ε-subgradient mappings*, Kibernetika (Kiev), 5 (1977), 148–149 (= Cybernetics 5 (1978), 790–791).
13. E. A. Nurminskii, *On ε-differential mappings and their applications in nondifferentiable optimization*, Working paper 78-58, I.I.A.S.A., December 1978.
14. R. T. Rockafellar, *Convex Analysis* (Princeton Mathematical Series 28), Princeton University Press, Princeton, 1970.
15. R. T. Rockafellar, *Integrals which are convex functionals*, II, Pacific J. Math. 39 (1971), 439–469.
16. G. Salinetti and R. J.-B. Wets, *On the convergence of sequence of convex sets in finite dimensions*, SIAM rev. 21 (1979), 18–33.

UNIVERSITÉ DE CLERMONT II
COMPLEXE SCIENTIFIQUE DE CÉZEAUX
DÉPARTMENT DE MATHÉMATIQUES APPLIQUÉES
BOÎTE POSTALE n° 45
63170 AUBIERE
FRANCE