Journal of Convex Analysis
Volume 3 (1996), No. 2, 207–219

Subdifferential Calculus Without Qualification Assumptions

Jean-Paul Penot
Université de Pau et des Pays de l'Adour, Lab. de Math. Appliquées, URA – CNRS 1204, Avenue de l'Université, 64000 Pau, France.
e-mail: jean-paul.penot@univ-pau.fr

Received October 10, 1995
Revised manuscript received January 9, 1996

Some classical formulas of convex subdifferential calculus in a general Banach space are presented without any qualification condition, but in a limiting form. Classical formulas are recovered when a qualification condition such as the Attouch-Brézis condition is assumed.

Several decades after the blossoming of convex analysis, the subject of subdifferential calculus rule is still very active (see [1]–[4], [15], [16], [23], [30]–[32] for instance). There are several reasons for that; among them is the attention devoted to calculus rules. Simple formulas such as

$$\partial\left(f_1+f_2\right)(x)=\partial f_1(x)+\partial f_2(x) \tag{0.1}$$

$$\partial(g \circ A)(x)=\partial g(A(x)) \circ A, \tag{0.2}$$

in which A is linear continuous, f_1, f_2 and g are closed proper convex functions finite at x and $A(x)$ respectively, are not valid without additional assumptions, a well known and disappointing fact, especially for beginners in convex analysis. A general formula for $\partial\left(f_1+f_2\right)(x)$ which does not require any qualification condition has been proposed in [15] (see also [16]). But it involves the approximate subdifferential $\partial_{\varepsilon} f$ of convex analysis. Thus this condition is specific to the convex situation. It is our purpose here to present a formula which is valid without any additional condition and which uses only the ordinary subdifferential of convex analysis. Such an aim has already been pursued in [2], [3], [14] and [29]; moreover, in [2], [3] a connection with a notion of variational sum of operators is pointed out. However the conditions we impose on the sequences involved in our formula differ from the conditions of Attouch-Baillon-Théra and Thibault: they are more precise and their interpretation is easier and more amenable to the general nonconvex, nonsmooth case. Since our approach aims at a link with the general situation of nonsmooth analysis, an interest for such conditions seems to be justified. It may have consequences on the evolution of nonsmooth analysis. Moreover, our conditions bear on the coupling functional and it is well known in nonlinear functional analysis and in the study of partial differential equations that delicate limiting procedures are to be found for such terms (see [7], [10], [19] for instance). Thus the analysis of convex functionals which are obtained in terms of integrals or integro-differential terms as in [5], [8], [13], [25], [26] may benefit from a closer analysis of calculus rules.

ISSN 0944-6532 / $ 2.50 © Heldermann Verlag

1. Preliminaries

The following result will be our basic tool. It consists in a slight supplement to the famous Brønsted-Rockafellar's theorem [9] and to an extension due to J. Borwein [6], [23]. It has its own interest. It involves the ε-approximate subdifferential of f given by

$$\partial_\varepsilon f(u) := \{u^* \in X^* : \quad \forall x \in X \quad f(x) - f(u) \geq \langle u^*, x - u\rangle - \varepsilon\}.$$

This set is a nonempty weak* closed convex set whose support function $h(\cdot, \partial_\varepsilon f(x_0)) := \sup\{\langle \cdot, x_0^*\rangle : x_0^* \in \partial_\varepsilon f(x_0)\}$ satisfies

$$h(v, \partial_\varepsilon f(x_0)) = \inf_{t>0} \frac{1}{t}\left(f(x_0 + tv) - f(x_0) + \varepsilon\right) \tag{1.1}$$

for each $v \in X$. Moreover one has

$$\partial f(x_0) = \bigcap_{\varepsilon>0} \partial_\varepsilon f(x_0). \tag{1.2}$$

Proposition 1.1. *Given a closed proper convex function f on a Banach space X, $\alpha > 0, \varepsilon > 0,$ $u_0 \in \operatorname{dom} f,$ $u_0^* \in \partial_\varepsilon f(u_0)$ one can find $(x_\varepsilon, x_\varepsilon^*) \in \partial f$ and $\gamma \in [-1, 1]$ such that*

$$\begin{aligned}
\|x_\varepsilon - u_0\| + \alpha \mid \langle u_0^*, x_\varepsilon - u_0\rangle \mid \ &\leq \sqrt{\varepsilon},\\
\left\|x_\varepsilon^* - (1 + \alpha\gamma\sqrt{\varepsilon})u_0^*\right\| \ &\leq \sqrt{\varepsilon},\\
|\langle x_\varepsilon^* - u_0^*, x_\varepsilon - u_0\rangle| \ &\leq \varepsilon,\\
|\langle x_\varepsilon^*, x_\varepsilon - u_0\rangle| \ &\leq \varepsilon + \alpha^{-1}\sqrt{\varepsilon},\\
|f(x_\varepsilon) - f(u_0)| \ &\leq \varepsilon + \alpha^{-1}\sqrt{\varepsilon}.
\end{aligned}$$

Further on we will take $\alpha = 1$, but it may be useful to dispose of such a parameter: the choice $\alpha = \varepsilon^{-\frac{1}{2}}$ is also an interesting one.

Proof. We follow the general line of the proof of [6]. Here we endow X with the equivalent norm $\|\cdot\|_\alpha$ given by $\|x\|_\alpha := \|x\| + \alpha \mid \langle u_0^*, x\rangle \mid$ for $x \in X$.
Since $u_0^* \in \partial_\varepsilon f(u_0),$ u_0 is an ε-approximate minimizer of the function $g := f - \langle u_0^*, .\rangle$. The Ekeland's variational principle (see for instance [12]) applied to g yields $x_\varepsilon \in X$ such that

$$g(x_\varepsilon) + \sqrt{\varepsilon}\|x_\varepsilon - u_0\|_\alpha \leq g(u_0),$$

$$0 \in \partial\left(g + \sqrt{\varepsilon}\,\|\cdot - x_\varepsilon\|_\alpha\right)(x_\varepsilon).$$

Since $g(u_0) \leq g(x_\varepsilon) + \varepsilon$, the first relation implies that

$$\|x_\varepsilon - u_0\| + \alpha \mid \langle u_0^*, x_\varepsilon - u_0\rangle \mid \leq \sqrt{\varepsilon}.$$

Using familiar subdifferential calculus rules, and denoting by B^* the closed unit ball of X^*, the second relation yields

$$u_0^* \in \partial\left(f + \sqrt{\varepsilon}\,\|\cdot - x_\varepsilon\| + \alpha\sqrt{\varepsilon} \mid \langle u_0^*, \cdot - x_\varepsilon\rangle \mid\right)(x_\varepsilon) \subset \partial f(x_\varepsilon) + \sqrt{\varepsilon}B^* + \alpha\sqrt{\varepsilon}\,[-1, 1]\,u_0^*.$$

In other terms, there exists $x_\varepsilon^* \in \partial f(x_\varepsilon)$, $\gamma \in [-1,1]$ and $v^* \in B^*$ such that

$$x_\varepsilon^* = (1 + \alpha\gamma\sqrt{\varepsilon})u_0^* + \sqrt{\varepsilon}v^*.$$

It follows that

$$|\langle x_\varepsilon^* - u_0^*, x_\varepsilon - u_0\rangle| \leq | \alpha\gamma\sqrt{\varepsilon}\langle u_0^*, x_\varepsilon - u_0\rangle | + | \sqrt{\varepsilon}\langle v^*, x_\varepsilon - u_0\rangle | \leq \varepsilon$$

and

$$|\langle x_\varepsilon^*, x_\varepsilon - u_0\rangle| \leq \varepsilon + | \langle u_0^*, x_\varepsilon - u_0\rangle | \leq \varepsilon + \alpha^{-1}\sqrt{\varepsilon}.$$

Since $x_\varepsilon^* \in \partial f(x_\varepsilon)$ one has

$$f(u_0) - f(x_\varepsilon) \geq \langle x_\varepsilon^*, u_0 - x_\varepsilon\rangle \geq -\varepsilon - \alpha^{-1}\sqrt{\varepsilon}.$$

Since $u_0^* \in \partial_\varepsilon f(u_0)$ one has

$$f(x_\varepsilon) - f(u_0) \geq \langle u_0^*, x_\varepsilon - u_0\rangle - \varepsilon \geq -\alpha^{-1}\sqrt{\varepsilon} - \varepsilon$$

and the last inequality of the statement holds. □

The last assertion of the following corollary could not be obtained without the additional information we got. Thus, it shows the importance of this information. It suggests a new definition of stabilized or limiting normals in the nonconvex case; such a proposal will not be considered here.

Corollary 1.2. *Given x_0 in the domain* $\operatorname{dom} f$ *of a closed proper convex function f on X one can find a sequence (x_n, x_n^*) in the graph of ∂f such that $(x_n) \to x_0$, $(f(x_n)) \to f(x_0)$, $(\langle x_n^*, x_n - x_0\rangle) \to 0$. Moreover, each possible weak* cluster point of a sequence or a net (x_n^*) satisfying these properties belongs to $\partial f(x_0)$.*

Proof. For each $n \in \mathbb{N} \setminus \{0\}$ we can find $y_n^* \in \partial_{1/n^2} f(x_0)$. Setting $u_0 = x_0$, $u_0^* = y_n^*$, $\alpha = 1$, $\varepsilon = 1/n^2$ in the preceding proposition, we get $(x_n, x_n^*) \in \partial f$ such that

$$\begin{aligned} \|x_n - x_0\| &\leq 1/n, \\ |f(x_n) - f(x_0)| &\leq 2/n, \\ |\langle x_n^*, x_n - x_0\rangle| &\leq 2/n. \end{aligned}$$

If x_0^* is a weak* cluster point of (x_n^*), for each $x \in X$ one has

$$\begin{aligned} f(x) &\geq \limsup_n (f(x_n) + \langle x_n^*, x - x_0\rangle + \langle x_n^*, x_0 - x_n\rangle) \\ &\geq f(x_0) + \langle x_0^*, x - x_0\rangle. \end{aligned}$$

so that $x_0^* \in \partial f(x_0)$. □

We will make use of the following recent results whose proofs are elementary.

Proposition 1.3. ([15], [16]) *Let A be a linear continuous operator between two Banach spaces X, Y, and let $g : Y \to \mathbb{R}^\bullet = \mathbb{R} \cup \{\infty\}$ be a closed convex function finite at some*

$y_0 \in Y$. Then, denoting by A^T the transpose of A and by cl^*C the weak-closure of a subset C of X^* one has, for any $x_0 \in A^{-1}(y_0)$,*

$$\partial(g \circ A)(x_0) = \bigcap_{\varepsilon>0} cl^* A^T(\partial_\varepsilon g(y_0)). \tag{1.3}$$

When X is reflexive, the weak closure can be replaced by the norm closure.*

This result is a consequence of relations ((1.1), (1.2)) and of the relation

$$\partial_\varepsilon f(x_0) = cl^*(A^T(\partial_\varepsilon g(y_0))), \tag{1.4}$$

where $f := g \circ A$; this relation amounts to the obvious equalities

$$\begin{aligned} h(v, cl^*(A^T(\partial_\varepsilon g(y_0)))) &= h(v, A^T(\partial_\varepsilon g(y_0))) = h(Av, \partial_\varepsilon g(y_0)), \\ \inf_{t>0} \frac{1}{t}\,(f(x_0+tv) - f(x_0) + \varepsilon) &= \inf_{t>0} \frac{1}{t}\,(g(y_0+tAv) - g(y_0) + \varepsilon). \end{aligned}$$

Proposition 1.4. ([15], [16]) *Given two closed proper convex functions f_1, f_2 on the Banach space X and $x_0 \in \operatorname{dom} f_1 \cap \operatorname{dom} f_2$ one has*

$$\partial\,(f_1 + f_2)\,(x_0) = \bigcap_{\varepsilon>0} cl^*\,(\partial_\varepsilon f_1\,(x_0) + \partial_\varepsilon f_2\,(x_0))\,.$$

2. The main results

The main results we have in view are the following ones. Let us start with composition, the fundamental operation from the point of view of category theory. We first consider a simple case.

Proposition 2.1. *Let A be a linear continuous operator between two Banach spaces X, Y, let $b \in Y$ and let $g : Y \to \mathbb{R}^\bullet := \mathbb{R} \cup \{\infty\}$ be a closed proper convex function. Let $f := g \circ (A+b)$, $x_0 \in \operatorname{dom} f$. Suppose that for each element ε of a subset E of $(0,\infty)$ containing 0 in its closure the set $A^T(\partial_\varepsilon g(y_0))$ is weak*-closed. Then for each $x_0^* \in \partial f(x_0)$ there exist sequences $(y_n)_{n\in\mathbb{N}}$, $(y_n^*)_{n\in\mathbb{N}}$ such that $y_n^* \in \partial g(y_n)$ for each $n \in \mathbb{N}$, $(y_n)_{n\in\mathbb{N}} \to y_0 := A(x_0) + b$, $(g(y_n))_{n\in\mathbb{N}} \to g(y_0)$, $(A^T(y_n^*))_{n\in\mathbb{N}} \to x_0^*$, and $(\langle y_n^*, y_n - y_0\rangle)_{n\in\mathbb{N}} \to 0$.*

Proof. Let (ε_n) be a sequence of E converging to 0. Proposition 1.3 enables one to find some $u_n^* \in \partial_{\varepsilon_n} g(y_0)$ such that $x_0^* = A^T(u_n^*)$. Then Proposition 1.1 provides some $y_n \in B(y_0, \varepsilon_n)$, $y_n^* \in \partial g(y_n)$, $\gamma_n \in [-1,1]$ such that $\|y_n^* - (1+\gamma_n\sqrt{\varepsilon_n})u_n^*\| \le \sqrt{\varepsilon_n}$, $|\langle y_n^*, y_n - y_0\rangle| \le \varepsilon_n + \sqrt{\varepsilon_n}$, $|g(y_n) - g(y_0)| \le \varepsilon_n + \sqrt{\varepsilon_n}$. As $(1+\gamma_n\sqrt{\varepsilon_n})A^T(u_n^*) = (1+\gamma_n\sqrt{\varepsilon_n})x_0^* \to x_0^*$, we get the result. □

When X is reflexive, the condition we obtained characterizes the elements of $\partial f(x_0)$, as shown at the end of the proof of the following statement which deals with the general case.

Theorem 2.2. *Let A, b, f, g be as above. Then, for any $x_0 \in \operatorname{dom} f$ the subdifferential of f at x_0 is the set of weak*-limit points x_0^* of nets $(x_n^*)_{n\in N}$ such that there exists a net $(y_n)_{n\in N}$ converging to $y_0 := A(x_0) + b$ and for each $n \in N$, some $y_n^* \in \partial g(y_n)$ with $x_n^* = A^T(y_n^*)$, $(g(y_n))_{n\in N} \to g(y_0)$ and $(\langle y_n^*, y_n - y_0\rangle)_{n\in N} \to 0$.*

If y_0^* is any weak-limit point of such a net $(y_n^*)_{n\in N}$ then $x_0^* = A^T(y_0^*)$ and $y_0^* \in \partial g(y_0)$. If X is reflexive one can take sequences instead of nets and impose convergence of $(x_n^*)_{n\in N}$ to x_0^* in norm.*

Proof. Without loss of generality we may suppose $b = 0$. Let P be a family of semi-norms on X^* whose unit balls form a base of neighborhoods of 0 for the weak* topology. Given $x_0^* \in \partial f(x_0)$, using Proposition 1.3, for any $k \in K := \mathbb{N}\setminus\{0\}$ and for any $p \in P$ one can find $z_{k,p}^* \in \partial_{1/k^2} g(y_0)$ such that $p(A^T(z_{k,p}^*) - x_0^*) < 1/k$. Then, by Proposition 1.1 and Corollary 1.2, one can pick $(y_{k,p}, y_{k,p}^*) \in \partial g$ and $\gamma_{k,p} \in [-1,1]$ such that $\|y_{k,p} - y_0\| < 1/k$, $| g(y_{k,p}) - g(y_0) | < 2/k$, $\langle y_{k,p}^*, y_{k,p} - y_0\rangle \le 2/k$ and $\|z_{k,p}^* - (1 + k^{-1}\gamma_{k,p})y_{k,p}^*\| < 1/k$. Setting $N := K \times P$ with the product order and $x_n^* := A^T y_{k,p}^*$ for $n := (k,p)$, and observing that

$$(1 + k^{-1}\gamma_{k,p})^{-1} A^T(z_{k,p}^*) \to x_0^*$$

for the weak* topology, so that

$$A^T(y_{k,p}^*) \to x_0^*$$

by the continuity of A^T, we get the announced nets. The assertion about limit points of the net $(y_{k,p}^*)$ follows from Corollary 1.2 and the weak* continuity of A^T. Moreover, we observe that when X is reflexive we can replace the weak*-closure by the norm closure and take $P := \left\{k^{-1}\|\cdot\| : k \in K\right\}$, so that we get a sequence rather than a net and strong convergence instead of weak* convergence.

Conversely, if x_0^*, $(y_n^*)_{n\in N}$, $(y_n)_{n\in N}$ are as in the statement, for each $x \in X$ we have

$$\begin{aligned}\langle x_0^*, x - x_0\rangle &= \lim_n \langle y_n^*, A(x) - A(x_0)\rangle = \lim_n \langle y_n^*, A(x) - y_n\rangle \\ &\le \limsup_n (g(A(x)) - g(y_n)) \le g(A(x)) - g(A(x_0))\end{aligned}$$

by the lower semicontinuity of g, and thus x_0^* belongs to $\partial f(x_0)$. □

Now let us turn to the sum.

Theorem 2.3. *Let X be a Banach space, let $f = f_1 + f_2$ with f_i closed proper convex on X and let $x_0 \in \operatorname{dom} f$. Then $\partial f(x_0)$ is the set of weak*- limits of nets of the form $(x_{1,n}^* + x_{2,n}^*)_{n\in N}$ with $x_{i,n}^* \in \partial f_i(x_{i,n})$, $(x_{i,n})_{n\in N} \to x_0$, $(f_i(x_{i,n}))_{n\in N} \to f_i(x_0)$, $\left(\langle x_{i,n}^*, x_{i,n} - x_0\rangle\right)_{n\in N} \to 0$ for $i = 1,2$. If X is reflexive the preceding nets can be replaced by sequences and the convergence of $x_{1,n}^* + x_{2,n}^*$ can be taken in the strong topology.*

Proof. Let us show how this result can be deduced from the composition result. Given $f = f_1 + f_2$, we set $Y := X \times X$ and we define A and g by $A(x) := (x,x)$, $g(y) :=$

$f_1(x_1) + f_2(x_2)$ for $y = (x_1, x_2)$; then $f = g \circ A$. Given $x_0^* \in \partial f(x_0)$, let $(y_n^*)_{n \in N}$, $(y_n)_{n\in N}$ be as in Theorem 2.1, with $y_n^* := (x_{1,n}^*, x_{2,n}^*)$, $y_n := (x_{1,n}, x_{2,n})$, $y_0 = (x_0, x_0)$. Since $x_{i,n}^* \in \partial f_i(x_{i,n})$ as easily seen, we have

$$f_1(x_{1,n}) - f_1(x_0) \leq \langle x_{1,n}^*, x_{1,n} - x_0 \rangle \leq \langle y_n^*, y_n - y_0 \rangle - (f_2(x_{2,n}) - f_2(x_0)).$$

Since $(g(y_n)) \to g(y_0)$, the lower semicontinuity of f_1 and f_2 entails that $(f_i(x_{i,n}) - f_i(x_0)) \to 0$ and we get $(\langle x_{1,n}^*, x_{1,n} - x_0 \rangle) \to 0$ from the convergence $(\langle y_n^*, y_n - y_0 \rangle) \to 0$; similarly $(\langle x_{2,n}^*, x_{2,n} - x_0 \rangle) \to 0$. Moreover

$$A^T(x_{1,n}^*, x_{2,n}^*) = x_{1,n}^* + x_{2,n}^* \to x_0^*$$

for the weak* convergence.

Conversely, let us suppose

$$x_0^* = \lim_n \left(x_{1,n}^* + x_{2,n}^* \right)$$

with

$$x_{i,n}^* \in \partial f_i\left(x_{i,n}\right), \ \left(x_{i,n}\right) \to x_0, \ \left(\alpha_{i,n}\right) := \left(f_i\left(x_{i,n}\right) - f_i\left(x_0\right)\right) \to 0,$$

$$\left(\gamma_{i,n}\right) := \left(\langle x_{i,n}^*, x_{i,n} - x_0 \rangle\right) \to 0 \quad \text{for } i = 1, 2.$$

Then for each $x \in \operatorname{dom} f$ we have

$$\begin{aligned} & f\left(x\right) - f\left(x_0\right) - \langle x_0^*, x - x_0 \rangle \\ = & \sum_{i=1}^{2} (f_i\left(x\right) - f_i\left(x_{i,n}\right) + \alpha_{i,n} - \gamma_{i,n} - \langle x_{i,n}^*, x - x_{i,n} \rangle) + \langle x_{1,n}^* + x_{2,n}^* - x_0^*, x - x_0 \rangle. \end{aligned}$$

Passing to the limit and using the relations

$$f_i\left(x\right) - f_i\left(x_{i,n}\right) - \langle x_{i,n}^*, x - x_{i,n} \rangle \geq 0$$

we get

$$f\left(x\right) - f\left(x_0\right) - \langle x_0^*, x - x_0 \rangle \geq 0,$$

and $x_0^* \in \partial f(x_0)$. □

Remark 2.4. The preceding statement shows that the familiar conditions

$$\left(x_{i,n}\right) \to x_0, \ \left(f_i\left(x_{i,n}\right)\right) \to f_i\left(x_0\right)$$

which are of common use in nonsmooth analysis have to be supplemented with the third condition

$$\left(\langle x_{i,n}^*, x_{i,n} - x_0 \rangle\right) \to 0$$

which represents a coupling condition on the sequence $\left(x_{i,n}, x_{i,n}^*\right)$. Obviously, the three conditions imply the condition

$$\left(f_i\left(x_{i,n}\right) - \langle x_{i,n}^*, x_{i,n} - x_0 \rangle\right) \to f_i\left(x_0\right)$$

imposed in [29]. The latter condition seems to be more difficult to interpret. Moreover, our conditions being more stringent, given an element x_0^* of $\partial f(x_0)$ we get more information about its limiting decomposition. This additional information may be precious, for instance for writing optimality conditions in terms of subdifferentials.

Remark 2.5. When receiving the first version of this paper, L. Thibault has kindly informed us that a related condition appears in [2] and [3] in the Hilbertian case, where it is obtained with quite different methods. Again, our three conditions are more precise when X is reflexive, so that one uses sequences : the conditions $\left(x_{1,n}^* + x_{2,n}^*\right) \to x_0^*$, $\left(\langle x_{i,n}^*, x_{i,n} - x_0\rangle\right) \to 0$ for $i = 1, 2$ imply the conditions

$$\left(\langle x_{i,n}^*, x_{1,n} - x_{2,n}\rangle\right) \to 0 \text{ for } i = 1, 2.$$

However, the variational interpretation obtained in [3] is no more present here and there is no connection with sums of monotone operators as in [3].

Remark 2.6. The preceding two theorems are equivalent statements. In order to show that the sum theorem implies the composition theorem let f, g, A, b be given as in Theorem 2.2, and let us set $f_1(x, y) = g(y)$, $f_2(x, y) = \iota_G(x, y)$, where $G = \{(x, Ax + b) : x \in X\}$ and ι_G is the indicator function of G, so that $f = f_1 + f_2$. The derivation is then straightforward in view of the obvious relationship between ∂f_1 and ∂g.

Remark 2.7. Let us observe that for any nets $(y_n)_{n\in N}, (y_n^*)_{n\in N}$ as in Theorem 2.2 and for each $\varepsilon > 0$ one has $y_n^* \in \partial_\varepsilon g\ (y_0)$ for $n \in N$, $n \geq n(\varepsilon)$ where $n(\varepsilon)$ is chosen in such a way that $\mid \langle y_n^*, y_n - y_0\rangle \mid + \mid g(y_n) - g(y_0) \mid \leq \varepsilon$ for $n \geq n(\varepsilon)$ since for each $y \in Y$ one has

$$g(y) - g(y_0) \geq \langle y_n^*, y - y_0\rangle + \langle y_n^*, y_0 - y_n\rangle + g(y_n) - g(y_0).$$

□

Let us derive a sufficient condition for a composition formula.

Corollary 2.8. *Suppose with the notations of* Proposition 2.1 *that for some $\eta > 0$ and each $r > 0$ there exists some $s > 0$ such that*

$$y^* \in \partial_\eta g(y_0), \quad \|A^T(y^*)\| \leq r \Rightarrow \|y^*\| \leq s.$$

Then

$$\partial(g \circ A)(x_0) = A^T(\partial g(y_0)).$$

Proof. The Krein-Smulian Theorem shows that for $\varepsilon < \eta$ the convex set $C := A^T(\partial_\varepsilon g(y_0))$ is weak*-closed since for each $r > 0$

$$C \cap rB_{X^*} = A^T(\partial_\varepsilon g(y_0) \cap sB_{Y^*}) \cap rB_{X^*}$$

and the image by A^T of the weak* compact set $\partial_\varepsilon g(y_0) \cap sB_{Y^*}$ is weak* compact. Then if we take sequences $(u_n^*)_{n\in N}, (y_n)_{n\in N}, (y_n^*)_{n\in N}$ as in Proposition 2.1, the assumption shows that $(u_n^*)_{n\in N}$ is bounded, hence has a converging subnet whose limit y_0^* is also the limit of $(y_n^*)_{n\in N}$, and belongs to $\partial g(y_0)$ by Corollary 1.2 and obviously $x_0^* = A^T(y_0^*)$. □

In order to derive sufficient conditions for the validity of formula (0.1) let us introduce the following notion which differs slightly from analogous concepts of nonsmooth analysis (see [20], [21], [27], [17], [28] for instance).

Definition 2.9. Given a closed proper convex function f on X, $x_0 \in \operatorname{dom} f$, the asymptotic subdifferential of f at x_0 is the set $\partial_\infty f(x_0)$ of $u^* \in X^*$ such that there exist nets $(t_n)_{n\in N} \to 0_+, ((x_n, x_n^*))_{n\in N}$ in ∂f with $(x_n)_{n\in N} \to x_0$, $(f(x_n))_{n\in N} \to f(x_0)$, $(\langle x_n^*, x_n - x_0\rangle)_{n\in N} \to 0$ and $(t_n x_n^*)_{n\in N} \to u^*$ in X^* weak*.
When $\partial f(x_0)$ is nonempty, taking $x_n = x_0$ for each n we see that this set is larger than the asymptotic cone $(\partial f(x_0))_\infty$ of $\partial f(x_0)$ and is larger than the corresponding sets in [17], [20], [21], [27], [28] given by

$$\partial^\infty f(x_0) = \left\{u^* \in X^* : (u^*, 0) \in N\left(E_f, (x_0, r_0)\right)\right\} \tag{2.1}$$

where $r_0 = f(x_0)$, E_f is the epigraph of f and $N(C, a)$ denotes the normal cone to C at a :

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

Here we recall that $\partial f(x_0) \times \{-1\} = N\left(E_f, (x_0, r_0)\right) \cap (X^* \times \{-1\})$. In fact, given $y_0^* \in \partial f(x_0)$, for any $u^* \in \partial^\infty f(x_0)$ and any $(t_n) \to 0_+$ we have $x_n^* := t_n^{-1} u^* + y_0^* \in \partial f(x_0)$ as $N\left(E_f, (x_0, r_0)\right)$ is a convex cone, and $(t_n x_n^*) \to u^*$. Let us note that the coupling condition $(\langle x_n^*, x_n - x_0\rangle) \to 0$ is useful in order to shrink $\partial_\infty f(x_0)$ as most as possible, so that the qualification condition presented below will not be too exacting.
Let us first give a more precise identification of $\partial_\infty f(x_0)$.

Proposition 2.10. *For any closed proper convex function f on X and any $x_0 \in D_f := \operatorname{dom} f$ one has $\partial_\infty f(x_0) \subset \partial^\infty f(x_0) = N(D_f, x_0)$. Equality holds if $\partial f(x_0)$ is nonempty.*

Proof. The second equality is a direct consequence of the definitions. Since we already observed the inclusion $\partial_\infty f(x_0) \supset \partial^\infty f(x_0)$ when $\partial f(x_0)$ is nonempty, and since for u^*, t_n, x_n, x_n^* as in Definition 2.9 we have $((x_n, r_n)) \to (x_0, r_0)$ for $r_0 = f(x_0)$, $r_n = f(x_n)$, $(t_n(x_n^*, -1)) \to (u^*, 0)$ weak* in $(X \times \mathbb{R})^*$,

$$\langle t_n(x_n^*, -1), (x_n, r_n) - (x_0, r_0)\rangle = t_n\langle x_n^*, x_n - x_0\rangle - t_n(r_n - r_0) \to 0,$$

the first inclusion and the proposition are consequences of the following result of independent interest. □

Lemma 2.11. *Let C be a convex subset of a normed vector space Z, let $(a_n)_{n\in N}$ be a net in C with limit a, and let $a_n^* \in N(C, a_n)$ be such that $(\langle a_n^*, a_n - a\rangle) \to 0$ and (a_n^*) weak* converges to some a^*. Then $a^* \in N(C, a)$.*

Let us observe that if we deal with a sequence rather than a net the condition $(\langle a_n^*, a_n - a\rangle) \to 0$ is automatically satisfied.

Proof. The result is a special case of the last assertion of Corollary 1.2 for the indicator function of C. □

We are ready to present a qualification condition which is new in the context of convex analysis. We need to introduce some terminology.

Definition 2.12. A convex subset C of a n.v.s. Z is said to be normally compact at $z \in C$ if for any net $(z_n)_{n\in N}$ of C with limit z, any net $(z_n^*)_{n\in N}$ such that $z_n^* \in N(C, z_n)$, $\|z_n^*\| = 1$ for each $n \in N$ has a non null cluster weak* point z^*. A convex function $f : X \to \mathbb{R}^\bullet$ on a n.v.s. X is said to be subdifferentially compact at $x_0 \in \operatorname{dom} f$ if, for any net $((x_n, x_n^*))_{n\in N}$ in ∂f with $(x_n) \to x_0$, $(\|x_n^*\|) \to \infty$, the net $(\|x_n^*\|^{-1}x_n^*)$ has a non null weak* cluster point.

Clearly, in a finite dimensional n.v.s., any convex set is normally compact and any function is subdifferentially compact. Moreover one has the following criteria. Note that assertion (a) below is valid for the more general class of epi-Lipschitzian sets, but here we do not wish to leave the framework of convex analysis.

Lemma 2.13.

a) *If the interior of a convex set C is nonempty, then C is normally compact at each $z \in C$;*

b) *If the epigraph E_f of a convex function f is normally compact at $z_0 := (x_0, f(x_0))$, then f is subdifferentially compact at x_0;*

c) *If the convex function f is continuous at some point of its domain, then f is subdifferentially compact at each point of its domain.*

Proof. a) Let $a \in C$ and $r > 0$ be such that the ball $B(a, r)$ is contained in C. Then, for any net $(z_n)_{n\in N}$ of C with limit $z \in C$, any net $(z_n^*)_{n\in N}$ such that $z_n^* \in N(C, z_n)$, $\|z_n^*\| = 1$ for each $n \in N$, one has for each u in the unit ball of Z

$$\langle z_n^*, a + ru - z_n\rangle \le 0$$

hence

$$\langle z_n^*, a - z_n\rangle \le -r\|z_n^*\| = -r,$$

so that each cluster weak* point z^* of $(z_n^*)_{n\in N}$ satisfies

$$\langle z^*, a - z\rangle \le -r$$

and is nonzero.

b) Let $((x_n, x_n^*))_{n\in N}$ be a net in ∂f with $(x_n) \to x_0$, $(\|x_n^*\|) \to \infty$, for $z_n := (x_n, f(x_n))$. For n large enough $(u_n^*, r_n) := (\|x_n^*\|^{-1}x_n^*, -\|x_n^*\|^{-1}) \in N(E_f, z_n)$ and is a unit vector; then, if (u^*, r) is a non null weak* cluster point of $((u_n^*, r_n))$ one has $r = 0$, $u^* \neq 0$ and and the net $(\|x_n^*\|^{-1}x_n^*)$ has a non null weak* cluster point u^*.

Assertion c) follows from the preceding two assertions. □

Corollary 2.14. *Let X be a Banach space and let f_1, f_2 be two closed proper convex functions on X. Let $x_0 \in \operatorname{dom} f_1 \cap \operatorname{dom} f_2$ be such that*

$$\partial_\infty f_1(x_0) \cap (-\partial_\infty f_2(x_0)) = \{0\} \tag{2.2}$$

and such that f_1 is subdifferentially compact at x_0. Then

$$\partial(f_1 + f_2)(x_0) = \partial f_1(x_0) + \partial f_2(x_0).$$

Proof. Given $x_0^* \in \partial (f_1 + f_2)(x_0)$, Theorem 2.1 ensures that for $i = 1, 2$ there exists nets $(x_{i,n})_{n\in N} \to x_0$, $\left(x_{i,n}^*\right)_{n\in N}$ such that

$$x_{i,n}^* \in \partial f_i(x_{i,n}),\ (f_i(x_{i,n})) \to f_i(x_0),\ \left(\langle x_{i,n}^*, x_{i,n} - x_0\rangle\right) \to 0 \text{ with } \left(x_{1,n}^* + x_{2,n}^*\right) \xrightarrow{w^*} x_0^*.$$

Let $r_n = \left\|x_{1,n}^*\right\|$. If (r_n) has a bounded subnet we may assume that $\left(x_{1,n}^*\right)$ w^*-converges and then $\left(x_{2,n}^*\right)$ w^*-converges too. In view of Corollary 1.2, the limit (x_1^*, x_2^*) of $\left(x_{1,n}^*, x_{2,n}^*\right)$ belongs to $\partial f_1(x_0) \times \partial f_2(x_0)$, and $x_1^* + x_2^* = x_0^*$. If $(r_n) \to \infty$, setting $u_{1,n}^* = r_n^{-1} x_{1,n}^*$ we use our assumption to get that $\left(u_{1,n}^*\right)$ has a converging subnet whose limit u_1^* is non null and belongs to $\partial_\infty f_1(x_0)$. Since

$$\lim r_n^{-1} x_{2,n}^* = \lim r_n^{-1} x_0^* - \lim r_n^{-1} x_{1,n}^* = -u_1^*$$

belongs to $\partial_\infty f_2(x_0)$ we reach a contradiction with our assumption (2.2).
The opposite inclusion being obvious, we get the announced equality. □

Remark 2.15. The qualification condition (2.2) seems to have been used for the first time (and in a nonconvex framework, but in finite dimensions) in [20], [21].The familiar qualification condition: f_1 is finite and continuous at some point a of $\operatorname{dom} f_2$ implies assumption (2.2). In fact, if $B(a, r) \subset \operatorname{dom} f_1$ and if $u^* \in \partial_\infty f_1(x_0) \cap (-\partial_\infty f_2(x_0))$, in view of Proposition 2.10 we have $\langle u^*, a + ru - x_0\rangle \leq 0$ for each u in the unit ball B_X, and $\langle -u^*, a - x_0\rangle \leq 0$, hence $-r\langle u^*, u\rangle \leq 0$ for each $u \in B_X$ and $u^* = 0$. □

Let us conclude by observing that the classical qualification condition of [27]

$$\mathbb{R}_+(\operatorname{dom} f_1 - \operatorname{dom} f_2) = X \tag{2.3}$$

entails condition (2.2) since it can be equivalently written

$$\mathbb{R}_+(\operatorname{dom} f_1 - x_0) - \mathbb{R}_+(\operatorname{dom} f_2 - x_0) = X$$

so that for any $x \in X$ and any $u^* \in N(\operatorname{dom} f_1, x_0) \cap (-N(\operatorname{dom} f_2, x_0))$ one has $\langle u^*, x\rangle \leq 0$, hence $u^* = 0$.
It is shown in [29] that condition (2.3) is a qualification condition in order to get a sum rule. In fact it can be shown that the more general condition

$$X_0 := \mathbb{R}_+(\operatorname{dom} f_1 - x_0) - \mathbb{R}_+(\operatorname{dom} f_2 - x_0)) \text{ is a closed vector subspace} \tag{2.4}$$

of Attouch and Brézis ([1], see also [18] Th 3.6.1 for a similar condition in an algebraic framework) can be deduced from Theorem 2.3. Similarly, the following qualification condition

$$Y_0 := \mathbb{R}_+(\operatorname{dom} g - y_0) - A(X) \text{ is a closed vector subspace} \tag{2.5}$$

can be derived from Theorem 2.2.

Corollary 2.16. *Given two closed proper convex functions f_1, f_2 on the Banach space X such that condition* (2.4) *is satisfied, then for any $x_0 \in \operatorname{dom} f_1 \cap \operatorname{dom} f_2$ one has*

$$\partial (f_1 + f_2)(x_0) = \partial f_1 (x_0) + \partial f_2 (x_0).$$

Given f, g, X, Y, A, b, x_0, y_0 as in Theorem 2.2 *such that condition* (2.5) *is satisfied one has*

$$\partial f(x_0) = A^T(\partial g(y_0)).$$

Proof. For the sake of brievity, we prove only the second assertion, both being related, as seen in a previous remark. Here one has to observe that if C_1 and C_2 are two convex cones of X such that $X_0 := C_1 - C_2$ is a closed vector subspace of X, then $Y_0 := C_1 \times C_2 - D$ is a closed vector subspace of $Y := X \times X$, D being the diagonal of Y.
Let us first suppose $Y_0 = Y$. Let us show that the assumption of Corollary 2.8 is satisfied. Given $r > 0$ let us show that the set $D := \partial_\varepsilon g(y_0) \cap (A^T)^{-1}(rB_{X^*})$ is bounded. This follows from the Banach-Steinhaus Theorem and the fact that for each $y \in Y$ one can find $t > 0$, $z \in \operatorname{dom} g$, $x \in X$ such that $y = t(z - y_0) - A(x)$ hence for each $y^* \in D$

$$\begin{aligned} \langle y^*, y \rangle &= t\langle y^*, z - y_0 \rangle - \langle A^T(y^*), x \rangle \\ &\leq t(g(z) - g(y_0) + \varepsilon) + r\|x\|, \end{aligned}$$

a number independent of y^*.
Now, as is well-known (see [1]), the general case can be deduced from the special case by setting $g_0(y) := g(y+y_0)$ for $y \in Y_0$ and by observing that any continuous linear extension $y^* \in Y^*$ of some $y_0^* \in \partial g_0(0)$ belongs to $\partial g(y_0)$ as Y_0 contains $A(X)$ and $\operatorname{dom} g - y_0$, (so that

$$g'(y_0, v) = \infty$$

for any $v \in Y \setminus Y_0$) and is such that $A^T(y^*) = x_0^*$ if $A^T(y_0^*) = x_0^*$. □

Acknowledgment. The author is grateful to L. Barbet for a careful reading, to A.D. Ioffe, M. Théra and an anonymous referee for some references and to L. Thibault for helpful discussions. He also acknowledges useful comments and criticisms by another anonymous referee.

References

[1] H. Attouch, H. Brézis: Duality for the sum of convex functions in general Banach spaces, in Aspects of Mathematics and its applications, J.A. Barroso ed., Elsevier (1986) 125–133.

[2] H. Attouch, J.-B. Baillon and M. Théra: Somme variationnelle d'opérateurs maximaux monotones, C.R. Acad. Sci. Paris I 317 (1993) 485–490.

[3] H. Attouch, J.-B. Baillon and M. Théra: Variational sum of maximal monotone operators, J. Convex Anal. 1 (1) (1994) 1–29.

[4] D. Azé: Duality for the sum of convex functions in general normed spaces, Arch. Math. 6 (1994) 554–561.

[5] M.-F. Bidault-Véron: Nonlinear elliptic equations of order 2m and subdifferentials, J. Optim. Th. Appl. 40 (3) (1983) 405–432.

[6] J. Borwein: A note on ε-subgradients and maximal monotonicity, Pacific J. Math. 103 (1982) 307–314.

[7] H. Brézis: Equations et inéquations non linéaires dans les espaces vectoriels en dualité, Ann. Inst. Fourier 18 (1968) 115–175.

[8] H. Brézis: Intégrales convexes dans les espaces de Sobolev, Israel J. Math. 13 (1–2) (1972) 9–23.

[9] A. Brønsted, R.T. Rockafellar: On the subdifferentiability of convex functions, Proc. Amer. Math. Soc. 16 (1965) 605–611.

[10] F.E. Browder: Nonlinear operators and nonlinear equations of evolution in Banach spaces, Proc. Symp. Pure Math. 18, Amer. Math. Soc., Providence, 1976.

[11] F.H. Clarke: Optimization and nonsmooth analysis, SIAM, Philadelphia, 1990.

[12] I. Ekeland: Nonconvex variational principles, Bull. Amer. Math. Soc. 1 (1979) 443–474.

[13] I. Ekeland, R. Témam: Analyse convexe et problèmes variationnels, Dunod-Gauthier-Villars, Paris 1974 (English translation, North Holland , Amsterdam, 1976).

[14] A. Fougères: Mathematical comparison between the variational representations of a mechanical equilibrium governed by localisable complementary energies, Journées Fermat, Toulouse, May 1985.

[15] J.-B. Hiriart-Urruty, R. Phelps: Subdifferential calculus, using ε-subdifferentials, J. Funct. Anal. 118 (1993) 154–166.

[16] J.-B. Hiriart-Urruty: M. Moussaoui, A. Seeger and M. Volle, Subdifferential calculus without qualification conditions, using approximate subdifferentials : a survey, Nonlinear Analysis, Theory, Methods, Appl. 24 (12) (1995) 1727–1754.

[17] A.D. Ioffe: Approximate subdifferentials and applications I. The finite dimensional theory, Trans. Amer. Math. Soc. 281 (1984) 389–416.

[18] S.S. Kutateladze: Convex operators, Uspekhi Mat. Nauk 34 (1) (1979) 167–196, English transl. Russian Math. Surveys 34 (1) (1979) 181–214.

[19] J.-L. Lions: Quelques méthodes de résolution des problèmes aux limites non linéaires, Dunod- Gauthier-Villars, Paris 1969.

[20] B.S. Mordukhovich: Nonsmooth analysis with nonconvex generalized differentials and adjoint mappings, Dokl. Akad. Nauk BSSR 28 (1994) 976–979.

[21] B.S. Mordukhovich: Approximation methods in problems of optimization and control, Nauka, Moscow, 1988.

[22] F. Murat: Compacité par compensation, in Analyse non convexe, Bull. Soc. Math. France, Mémoire 60 (1979) 125–127.

[23] R. Phelps: Convex Functions, Monotone Operators and Differentiability, Lecture Notes in Math 1364, Springer Verlag, New York (1989).

[24] S.M. Robinson: Regularity and stability for convex multivalued functions, Math. Oper. Res. 1 (1976) 130–143.

[25] R.T. Rockafellar: Integrals which are convex functionals , Pacific J. Math.24 (1966) 525–539.

[26] R.T. Rockafellar: Integrals which are convex functionals II, Pacific J. Math.39 (1971) 439–469.

[27] R.T. Rockafellar: Conjugate duality and optimization, SIAM CBMS 16, Philadelphia (1974).

[28] R.T. Rockafellar, Directionally Lipschitzian functions and subdifferential calculus, Proc. London Math. Soc. 39 (1979) 331–355.

[29] L. Thibault: A generalized sequential formula for subdifferentials of sums of convex functions defined on Banach spaces, in Recent Developments in Optimization, Proc. Seventh French-German Conference on Optimization, Dijon, July 1994, R. Duriez and C. Michelot (eds) Lecture Notes in Economics and Math. Systems 429, Springer Verlag, Berlin (1995) 340–345.

[30] M. Volle: Some applications of the Attouch-Brézis condition to closedness criterions, optimization and duality, Sémin. Anal. Conv. Univ. Montpellier (1992) Exposé 16.

[31] M. Volle: A general formula for the subdifferential of an upper envelope of convex functionals, preprint, Univ. Avignon, 1993.

[32] C. Zalinescu: On some open problems in convex analysis, Arch. Math. 59 (1992) 566–571.

HIER :

Leere Seite
220