

EXTENSIONS OF SUBDIFFERENTIAL CALCULUS RULES IN BANACH SPACES

A. JOURANI AND L. THIBAUT

ABSTRACT. This paper is devoted to extending formulas for the geometric approximate subdifferential and the Clarke subdifferential of extended-real-valued functions on Banach spaces. The results are strong enough to include completely the finite dimensional setting.

0. Introduction. This paper continues a research program introduced in [14] whose results will be frequently used throughout the text. It concerns the basic rules of subdifferential calculus for nonsmooth functions on arbitrary Banach spaces. Ioffe [12, 13] and Rockafellar [26] have respectively proved for the Ioffe's geometric approximate subdifferential and the Clarke subdifferential that

$$\partial(f + g)(x_0) \subset \partial f(x_0) + \partial g(x_0)$$

whenever one of the functions is directionally Lipschitzian (see also Kruger [18] for the particular case where the Banach space admits an equivalent norm which is Fréchet differentiable away from zero). Note that their results do not include the finite dimensional case (see Ioffe [11], Mordukhovich [23], Rockafellar [27], Ward and Borwein [29]) since for finite dimensional spaces neither function f nor g need be directionally Lipschitzian.

More generally combining Ioffe's sum rule (see Theorem 5.6 in [13]) and chain rule (see Corollary 7.8.1 in [13]) it is not difficult to see that the Ioffe geometric approximate subdifferential satisfies (under a constraint qualification)

$$(0) \quad \partial(f + h \circ g)(x_0) \subset \partial f(x_0) + \bigcup_{y^* \in \partial h(g(x_0))} \partial(y^* \circ g)(x_0)$$

whenever h is directionally Lipschitzian at $g(x_0)$ and whenever g is Lipschitz near x_0 and admits a strict prederivative having norm compact values. This is the best sufficient condition known so far ensuring the important subdifferential rule (0). This condition is not completely satisfactory since it does not include what can be said in the finite dimensional case, for it requires h to be directionally Lipschitzian. Furthermore the existence of a strict prederivative (with norm compact values) is not always easy to check. The aim of the present paper is to weaken significantly the above conditions. We prove formula (0)

Received by the editors June 3, 1994; revised June 29, 1994.

AMS subject classification: 49J52, 46B20, 58C20.

Key words and phrases: subdifferential, epi-Lipschitz-like sets, compactly epi-Lipschitzian sets.

© Canadian Mathematical Society 1996.

with the geometric approximate subdifferential (and a similar one with the Clarke subdifferential) for general classes of functions including both the finite dimensional situation and directionally Lipschitzian functions on Banach spaces. More precisely we only assume in our proof of (0) that g is compactly Lipschitzian at x_0 (a notion introduced in Thibault [28]) and that h is compactly epi-Lipschitz at x_0 in the sense of Borwein and Strojwas [3]. Although this result extends the above one of Ioffe and those known so far in the finite dimensional setting, it also provides a unified way to prove all these theorems.

The paper is organized as follows. Section 1 develops some tangency formulae in terms of Ioffe geometric approximate subdifferentials and Clarke subdifferentials. These formulae allow us to extend formula (0) to the compactly-epi-Lipschitz and epi-Lipschitz-like functions in Section 2.

Before concluding this introduction we have to say that Borwein and Strojwas informed us that they also obtained some different extended calculus rules for the Clarke subdifferential.

1. Formulae for G -normal cones and Clarke normal cones. In all the paper X and Y are Banach spaces, \mathbb{B}_X the closed unit ball centered at the origin in X and X^* the topological dual of X . If not specified, the norm in a product of two Banach spaces is defined as $\|(u_1, u_2)\| = \|u_1\| + \|u_2\|$. We denote by $d(x, S) = \inf_{u \in S} \|x - u\|$ the usual distance function to the set S .

For a function $f: X \rightarrow \mathbb{R} \cup \{+\infty\}$ and $x \in X$ we denote by $\partial_C f(x)$ and $\partial_G f(x)$ the Clarke and the Ioffe geometric approximate subdifferential of f at x and we refer to Clarke [6] and Ioffe [13] for the definitions.

We recall our chain rule proved in Jourani and Thibault [16]. A similar result for Lipschitz mappings with compact prederivatives (hence which are strongly compactly Lipschitzian (see [28] or [16])) had been established before by Ioffe [13]. We state our result in terms of G -subdifferentials.

THEOREM 1.1. *Let $f: Y \rightarrow \mathbb{R}$ be Lipschitz around y_0 and $g: X \rightarrow Y$ be strongly compactly Lipschitzian at $x_0 \in g^{-1}(y_0)$. Then $f \circ g$ is Lipschitz around x_0 and*

$$\partial_G(f \circ g)(x_0) \subset \bigcup_{y^* \in \partial_G f(y_0)} \partial_G(y^* \circ g)(x_0). \quad \blacksquare$$

Borwein and Strojwas [3] have defined a subset S of Y to be compactly epi-Lipschitz at $y_0 \in S$ if there exist a real number $r > 0$ and a $\|\cdot\|$ -compact subset H of Y such that for every $t \in]0, r]$

$$(1.1) \quad (y_0 + r\mathbb{B}_Y) \cap S + t\mathbb{B}_Y \subset S - tH.$$

It is not difficult to see that any epi-Lipschitzian set (in the sense of Rockafellar [25]) or any subset of a finite dimensional space Y is compactly epi-Lipschitz. Several other examples are given in [3].

The following Theorem, whose proof is given in Jourani and Thibault [14], will be crucial in our approach of formulae for normal cones in Theorems 1.10 and 1.11. We state it in terms of G -approximate subdifferentials.

THEOREM 1.2 ([14]). *Let C and D be two closed subsets of X and Y respectively with D compactly epi-Lipschitz at $y_0 \in D$. Let $g: X \rightarrow Y$ be strongly compactly Lipschitzian at $x_0 \in C \cap g^{-1}(y_0)$. Suppose that*

$$(R) \quad y^* \in \partial_G d(y_0, D), \quad 0 \in \partial_G (y^* \circ g + d(\cdot, C))(x_0) \Rightarrow y^* = 0.$$

Then there exist $r > 0$ and $\gamma \geq 0$ such that

$$d(x, C \cap g^{-1}(D)) \leq \gamma d(g(x), D)$$

for all $x \in (x_0 + r\mathbb{B}_X) \cap C$.

Let us begin by proving, with the help of Theorem 1.2, the following result. Similar less general results have been proved in Hilbert spaces (in a different way) by Clarke and Raïssi [7], with $g = \text{Id}$. See also Ioffe [13] for the case where D is epi-Lipschitzian and g is the identity mapping. Recall (see Ioffe [13]) that $\hat{N}_G(C; x_0) := \mathbb{R}_+ \partial_G d(x_0, C)$ is the nucleus of the G -normal to C at x_0 .

THEOREM 1.3. *Under the assumptions of Theorem 2.4 there exists some real number $K > 0$ such that*

$$\partial_G d(x_0, C \cap g^{-1}(D)) \subset \bigcup_{y^* \in \partial_G d(g(x_0), D)} K \partial_G (y^* \circ g + d(\cdot, C))(x_0)$$

and hence

$$\hat{N}_G(C \cap g^{-1}(D); x_0) \subset \hat{N}_G(C; x_0) + \bigcup_{y^* \in \hat{N}_G(D; g(x_0))} \hat{\partial}(y^* \circ g)(x_0).$$

PROOF. By Theorem 1.2, the function

$$x \mapsto \gamma d(g(x), D) - d(x, C \cap g^{-1}(D))$$

attains a local minimum at x_0 relative to C and hence (see Proposition 2.4.3 in [6]) x_0 is a local minimum, over some ball $x_0 + s\mathbb{B}_X$, of the function

$$x \mapsto (1 + \gamma k_g) d(x, C) + \gamma d(g(x), D) - d(x, C \cap g^{-1}(D))$$

where k_g is a Lipschitz constant for g at x_0 . Setting $K := \max(\gamma, 1 + \gamma k_g)$ we get for all $x \in x_0 + s\mathbb{B}_X$

$$(1) \quad d(x, C \cap g^{-1}(D)) \leq K \left(d(x, C) + d(g(x), D) \right).$$

Set $S := C \cap g^{-1}(D)$, $H(x) := (g(x), d(x, C)) \in Y \times \mathbb{R}$ and $h(y, r) := r + d(y, D)$. If we put $\Delta := d(\cdot, S)$ and $\Delta_M(x) = \Delta(x)$ if $x \in M$ and $\Delta_M(x) = +\infty$ otherwise, and if we denote by $\mathcal{F}(X)$ the collection of all finite dimensional subspaces of X , we have by Proposition 2.4 in Ioffe [13]

$$\partial_G d(x_0, S) = \bigcap_{L \in \mathcal{F}(X)} \limsup_{\substack{x \rightarrow x_0 \\ \varepsilon \downarrow 0}} \partial_\varepsilon^- \Delta_{x+L}(x)$$

(where ∂_ε^- denotes the Dini ε -subdifferential in [13] and $x \xrightarrow{S} x_0$ means $x \rightarrow x_0$ and $x \in S$) and hence it follows from (1) that

$$(2) \quad \partial_G d(x_0, S) \subset K \partial_G (h \circ H)(x_0).$$

But by Theorem 1.1, since H is strongly compactly Lipschitzian at x_0 , we may write

$$\partial_G (h \circ H)(x_0) \subset \bigcup_{y^* \in \partial_G d(g(x_0), D)} \partial_G (y^* \circ g + d(\cdot, C))(x_0).$$

This and relation (2) complete the proof. ■

REMARK. The arguments above also show that the inequality (1) (in the proof of Theorem 1.3) is equivalent to the (metric regularity) inequality in the statement of Theorem 1.2. This inequality (1) has been used by Dolecki [8], when g is the identity mapping, as a starting point in the calculus of the Clarke tangent cone to the intersection of two sets. ■

Taking g as the identity mapping in the above theorem, we obtain the following corollary.

COROLLARY 1.4. *Let C and D be two closed subsets of X with D compactly epi-Lipschitz at $x_0 \in C \cap D$. Suppose that*

$$\partial_G d(x_0, C) \cap (-\partial_G d(x_0, D)) = \{0\}.$$

Then there exists $K > 0$ such that

$$\partial_G d(x_0, C \cap D) \subset K(\partial_G d(x_0, C) + \partial_G d(x_0, D)).$$

Although the following corollary could be proved with the help of Corollary 1.4, we prefer to provide a proof using Theorem 1.2 and the Clarke subdifferential calculus rules for locally Lipschitz functions directly.

COROLLARY 1.5. *Let C and D be two closed subsets of X and Y respectively and let $g: X \rightarrow Y$ be strictly differentiable at $x_0 \in C \cap g^{-1}(D)$. Assume that D is compactly epi-Lipschitz at $y_0 := g(x_0)$ and that the condition (corresponding to (R) in Theorem 1.2)*

$$\Lambda^{-1}(\partial_G d(C, x_0)) \cap (-\partial_G d(D, y_0)) = \{0\},$$

holds with $\Lambda := g'(x_0)^$ (the adjoint linear map of $g'(x_0)$). Then for some $K \geq 0$ one has*

$$\partial_c d(x_0, C \cap g^{-1}(D)) \subset K \partial_c (d_C + d_D \circ g)(x_0) \subset K[\partial_c d(x_0, C) + \Lambda(\partial_c d(y_0, D))],$$

where $d_C := d(\cdot, C)$.

PROOF. As in the proof of Corollary 1.4 we get for all $x \in x_0 + r\mathbb{B}_X$

$$d(x, C \cap g^{-1}(D)) \leq K[d(x, C) + h \circ g(x)]$$

where $h(y) = d(y, D)$. So (see [6]), for $S := C \cap g^{-1}(D)$ the Clarke generalized directional derivative $d_S^0(x_0; \cdot)$ of the distance function to S satisfies

$$\begin{aligned} d_S^0(x_0; h) &= \limsup_{\substack{S \\ x \rightarrow x_0 \\ t \downarrow 0}} t^{-1} d(x + th, S) \\ &\leq K(d_C + h \circ g)^0(x_0; h) \\ &\leq K[d_C^0(x_0; h) + d_S^0(y_0; g'(x_0)h)] \end{aligned}$$

for all $h \in X$. So the conclusion of the corollary is a direct consequence of these last inequalities and the subdifferential calculus rules for continuous convex functions. ■

REMARK. More generally under the assumptions of Theorem 1.2 one can prove that

$$\partial_c d(x_0, C \cap g^{-1}(D)) \subset K \left[\partial_c d(x_0, C) + \text{cl}_{w^*} c \circ \bigcup_{y^* \in \partial_c d(y_0, D)} \partial_c (y^* \circ g)(x_0) \right].$$

In Theorem 1.3, we have obtained a general formula for the nucleus \hat{N}_G . We proceed now to establish some calculus rules much sharper and broader than those available for the G -normal cone as well as for the Clarke normal cone. First we need to prove Lemmas 1.7 and 1.8 which will be some keys to the proofs of Theorems 1.10 and 1.11. Let us recall before this lemma of Jourani and Thibault [14].

LEMMA 1.6 ([14]). *Let S be a subset of X which is compactly epi-Lipschitz at $x_0 \in S$ and let H be the compact subset of X as given in relation (1.1). There exist a real number $\beta > 0$ and a neighbourhood V of x_0 such that for each $\varepsilon > 0$ there are h_1, \dots, h_m in H satisfying*

$$\|x^*\| \leq \varepsilon + \beta \text{Max}_{1 \leq i \leq m} \langle x^*, h_i \rangle$$

for all $x \in V$ and $x^* \in \partial_G d(x, S)$.

From this lemma we can deduce the following inequality for the G -normal cone. We refer the reader to Loewen [20] for many other important results for the limiting Fréchet normal cones in reflexive spaces. Recall that the G -normal cone $N_G(X; x)$ to S at $x \in S$ is, by definition, the weak-star closure of $\mathbb{R}_+ \partial_G(x, S)$.

LEMMA 1.7. *Let S be a subset of X which is compactly epi-Lipschitz at $x_0 \in S$ and let H be the compact subset of X as given in relation (1.1). Then there exists a real number $\beta > 0$ and a neighbourhood V of x_0 such that for each $\varepsilon \in [0, \beta^{-1}]$ there are h_1, \dots, h_m in H satisfying*

$$(1 - \beta\varepsilon)\|x^*\| \leq \beta \text{Max}_{1 \leq i \leq m} \langle x^*, h_i \rangle$$

for all $x \in V \cap S$ and $x^* \in N_G(S; x)$.

PROOF. Choose β and V as given by Lemma 1.6. It is obvious that

$$(1) \quad \|x^*\| \leq \beta \text{Max}_{h \in H} \langle x^*, h \rangle$$

for all $x \in V \cap S$ and $x^* \in \partial_G d(x, S)$. By positive homogeneity, (1) also holds for all $x \in V \cap S$ and $x^* \in \mathbb{R}_+ \partial_G d(x, S)$. Fix $\varepsilon \in]0, 1]$ and choose elements h_1, \dots, h_m in H such that $H \subset \{h_1, \dots, h_m\} + \varepsilon \mathbb{B}_X$. We deduce from (1) that, for any $x^* \in \mathbb{R}_+ \partial_G d(x, S)$ with $x \in V \cap S$, one has

$$\|x^*\| \leq \beta \varepsilon \|x^*\| + \beta \text{Max}_{1 \leq i \leq m} \langle x^*, h_i \rangle$$

and hence

$$(2) \quad (1 - \beta \varepsilon) \|x^*\| \leq \beta \text{Max}_{1 \leq i \leq m} \langle x^*, h_i \rangle.$$

Then by the weak-star lower semicontinuity of the function $(1 - \beta \varepsilon) \| \cdot \|$ and the weak-star continuity of $x^* \mapsto \text{Max}_{1 \leq i \leq m} \langle x^*, h_i \rangle$, it follows that relation (2) still holds for all $x^* \in N_G(S; x) = \text{cl}_{w^*} [\mathbb{R}_+ \partial_G d(x, S)]$ and $x \in V \cap S$. So the proof of the lemma is complete. ■

Concerning the Clarke normal cone $N_c(S; \cdot)$: we can prove the following similar result for epi-Lipschitz-like sets. Recall first that Borwein [2] has defined a subset S of X to be epi-Lipschitz-like at $x_0 \in S$ if there exist a real number $r > 0$, a vector $v \in Y$ and a nonempty closed convex set Ω containing zero such that Ω^0 is weak-star locally compact and

$$(x_0 + r\mathbb{B}_X) \cap S +]0, r][v + \Omega) \subset S,$$

where Ω^0 is the polar of Ω , that is, $\Omega^0 = \{x^* \in X^* : \langle x^*, x \rangle \leq 1, \forall x \in \Omega\}$. He has also proved that any epi-Lipschitz-like set is compactly epi-Lipschitz. Note that epi-Lipschitzian sets and subsets of any finite dimensional space X are always epi-Lipschitz-like.

LEMMA 1.8. *Let S be a subset of X which is epi-Lipschitz-like at $x_0 \in S$. Then there exist a neighbourhood V of x_0 in X and elements $h_1, h_2, \dots, h_m \in X$ such that*

$$\|x^*\| \leq \sum_{i=1}^m |\langle x^*, h_i \rangle|$$

for all $x^* \in N_c(S; x)$ and $x \in S \cap V$.

PROOF. Let Ω be a closed convex subset with $0 \in \Omega$ and Ω^0 weak-star locally compact, $r \in]0, \infty[$ and $v \in X$ such that

$$(1) \quad (x_0 + r\mathbb{B}_X) \cap S +]0, r][v + \Omega) \subset S.$$

By Lemma 2.1 in Borwein [1] there exist a compact subset H in X and $s \in]0, \infty[$ such that $2s\mathbb{B}_X \subset \Omega + sH$. Choose $h_2, \dots, h_m \in H$ such that $H \subset \mathbb{B}_X + \{h_2, \dots, h_m\}$. Then

$$2s\mathbb{B}_X \subset \Omega + s\mathbb{B}_X + \text{co}\{sh_2, \dots, sh_m\}$$

and hence by the Rådstrom law (see [24])

$$s\mathbb{B}_X \subset \Omega + \text{co}\{sh_2, \dots, sh_m\}.$$

Let V be an open neighbourhood of x_0 with $V \subset x_0 + r\mathbb{B}_X$. Fix $b \in \mathbb{B}_X$ and choose $h \in \text{co}\{-h_2, \dots, -h_m\}$ such that $sb + sh \in \Omega$. Fix $x \in S \cap V$ and $x^* \in N_c(S; x)$. We get by (1) that $v + sb + sh \in T_c(S; x)$ (the Clarke tangent cone to S at x , see Clarke [6]) and hence $\langle x^*, v + sb + sh \rangle \leq 0$. Set $h_1 := -s^{-1}v$ and write $h = -\sum_{i=2}^m \lambda_i h_i$ with $\sum_{i=2}^m \lambda_i = 1$ and $\lambda_i \geq 0$. Then we obtain

$$\begin{aligned} \langle x^*, b \rangle &\leq \langle x^*, h_1 \rangle + \sum_{i=2}^m \lambda_i \langle x^*, h_i \rangle \\ &\leq |\langle x^*, h_1 \rangle| + \sum_{i=2}^m \lambda_i |\langle x^*, h_i \rangle| \\ &\leq \sum_{i=1}^m |\langle x^*, h_i \rangle| \end{aligned}$$

and hence

$$\|x^*\| \leq \sum_{i=1}^m |\langle x^*, h_i \rangle|. \quad \blacksquare$$

Before proving the theorems of extended calculus for normal cones let us mention the following important consequences of the above lemmas.

PROPOSITION 1.9. *Let S be a subset of X which is compactly epi-Lipschitz (resp. epi-Lipschitz-like) at $x_0 \in S$. Then for any x^* in $N_G(S; x_0)$ (resp. in $N_c(S; x)$) there exist nets $(\lambda_i)_{i \in I}$ in $]0, \infty[$, $(x_i^*)_{i \in I}$ in $\partial_G d(x_0, S)$ (resp. in $\partial_c d(x_0, S)$) such that the net $(\lambda_i x_i^*)_{i \in I}$ is bounded and w^* -converges to x^* .*

PROOF. As $N_G(S; x_0) = \text{cl}_{w^*}(]0, \infty[\partial_G d(x_0, S))$ by definition, there are two nets $(\gamma_j)_{j \in J}$ in $]0, \infty[$ and $(u_j^*)_{j \in J}$ in $\partial_G d(x_0, S)$ such that $\gamma_j u_j^* \xrightarrow{w} x^*$. Let h_1, \dots, h_m be given by Lemma 1.7 with $\varepsilon = \frac{1}{2}\beta^{-1}$. Then for all $j \in J$

$$(1) \quad \|\gamma_j u_j^*\| \leq 2\beta \text{Max}_{1 \leq k \leq m} \langle \gamma_j u_j^*, h_k \rangle.$$

As the net $\text{Max}_{1 \leq k \leq m} \langle \gamma_j u_j^*, h_k \rangle_{j \in J}$ converges in \mathbb{R} , there exists a subnet $(\text{Max}_{1 \leq k \leq m} \langle \gamma_{\alpha(i)} u_{\alpha(i)}^*, h_k \rangle)_{i \in I}$ which is bounded. By (1) the subnet $(\gamma_{\alpha(i)} u_{\alpha(i)}^*)_{i \in I}$ is also bounded and hence it is enough to set $\lambda_i := \gamma_{\alpha(i)}$ and $x_i^* := u_{\alpha(i)}^*$. This completes the proof of the assertion relative to the geometric approximate normal cone.

Concerning the Clarke normal cone it is enough to recall (see Clarke [6]) that $N_c(S; x_0) = \text{cl}_{w^*}(]0, \infty[\partial_c d(x_0, S))$ and to repeat the above argument with Lemma 1.8 instead of Lemma 2.9. \blacksquare

A new general rule for calculating or estimating G -normal cones can now be given.

THEOREM 1.10. *Let C and D be two closed subsets of X and Y respectively with D compactly epi-Lipschitz at $y_0 \in D$. Let $g: X \rightarrow Y$ be strongly compactly Lipschitzian at $x_0 \in C \cap g(y_0)$. Suppose that*

$$(R') \quad y^* \in N_G(D, y_0), \quad 0 \in \partial_G(y^* \circ g)(x_0) + N_G(C, x_0) \Rightarrow y^* = 0.$$

Then one has

$$N_G(C \cap g^{-1}(D), x_0) \subset N_G(C, x_0) + \bigcup_{y^* \in N_G(D, g(x_0))} \partial_G(y^* \circ g)(x_0).$$

PROOF. The proof uses some ideas of Theorem 5.4 in Ioffe [13]. Let $x^* \in N_G(C \cap g^{-1}(D))$. By definition there are nets $(\lambda_i)_{i \in I}$ in $]0, \infty[$, $(x_i)_{i \in I}$ in $\partial_G d(x_0, C \cap g^{-1}(D))$ such that $\lambda_i x_i^* \xrightarrow{w^*} x^*$. By Theorem 1.3 there are $v_i^* \in K \partial_G d(x_0, C)$, $y_i^* \in K \partial_G d(g(x_0), D)$ and $(u_i^*) \in \partial_G(y_i^* \circ g)(x_0)$ such that

$$(1) \quad x_i^* = u_i^* + v_i^*.$$

Assume for a moment that the net $(\lambda_i y_i^*)_{i \in I}$ has a bounded subnet $(\lambda_{\alpha(j)} y_{\alpha(j)}^*)_{j \in J}$. Put $\gamma_j := \lambda_{\alpha(j)}$, $k_j^* := x_{\alpha(j)}^*$, $\ell_j^* := u_{\alpha(j)}^*$, $p_j^* := v_{\alpha(j)}^*$, $q_j^* := y_{\alpha(j)}^*$. Then the net $(\gamma_j \ell_j^*)_{j \in J}$ is also bounded since $\gamma_j \ell_j^* \in \partial_G(\gamma_j q_j^* \circ g)(x_0)$. Extracting a subnet if necessary, we may suppose that the net $(\gamma_j \ell_j^*, \gamma_j q_j^*)_{j \in J}$ weakly-star converges to some point (u^*, y^*) . Then $y^* \in N_G(D, y_0)$ and by Lemma 2.5 in [9] one has $u^* \in \partial_G(y^* \circ g)(x_0)$. Since $\gamma_j k_j^* \xrightarrow{w^*} x^*$ we also see by (1) that the net $(\gamma_j p_j^*)_{j \in J}$ weakly-star converges to some point $v^* \in N_G(C; x_0)$ and that $x^* = u^* + v^*$.

To get the inclusion it thus remains to prove that the net $(\lambda_i y_i^*)_{i \in I}$ admits a bounded subnet. Suppose the contrary. Then for each $n \in \mathbb{N}$ this last supposition ensures that there exists, for each $i \in I$, some $\alpha(n, i) \in I$ such that

$$(2) \quad \alpha(n, i) \geq i \text{ and } \|\lambda_{\alpha(n, i)} y_{\alpha(n, i)}^*\| \geq n.$$

If we consider the preorder on $\mathbb{N} \times I$ defined by

$$(n, i) \geq (n', i') \Leftrightarrow n \geq n' \text{ and } i \geq i',$$

then for each $i_0 \in I$ there exists $(n_0, i_0) \in \mathbb{N} \times I$ (take any $n_0 \in \mathbb{N}$) such that $\alpha(n, i) \geq i_0$ for all (n, i) in $\mathbb{N} \times I$ satisfying $(n, i) \geq (n_0, i_0)$. Then $(\lambda_{\alpha(n, i)} \|y_{\alpha(n, i)}^*\|)_{(n, i) \in \mathbb{N} \times I}$ is a subnet (see [17]) and by (2)

$$\lambda_{\alpha(n, i)} \|y_{\alpha(n, i)}^*\| \xrightarrow{(n, i) \in \mathbb{N} \times I} +\infty,$$

that is there exists some subnet $(\lambda_{\alpha(j)} \|y_{\alpha(j)}^*\|)_{j \in J}$ converging to $+\infty$. We may suppose $\|y_{\alpha(j)}^*\| \neq 0$ for all $j \in J$. Put

$$a_j^* := \|y_{\alpha(j)}^*\|^{-1} x_{\alpha(j)}^*, \quad b_j^* := \|y_{\alpha(j)}^*\|^{-1} u_{\alpha(j)}^*, \\ c_j^* := \|y_{\alpha(j)}^*\|^{-1} v_{\alpha(j)}^* \text{ and } z_j^* := \|y_{\alpha(j)}^*\|^{-1} y_{\alpha(j)}^*.$$

Then $b_j^* \in \partial_G(z_j^* \circ g)(x_0)$, $c_j^* \in N_G(C, x_0)$, $z_j^* \in N_G(D, g(x_0))$ and

$$(3) \quad a_j^* = b_j^* + c_j^*.$$

But $a_j^* \xrightarrow[k \in J]{w^*} 0$ since $a_j^* = (\lambda_{\alpha(j)} \|y_{\alpha(j)}^*\|)^{-1} \cdot (\lambda_{\alpha(j)} x_{\alpha(j)}^*)$, and $(\lambda_{\alpha(j)} x_{\alpha(j)}^*)$ weakly-star converges and $\lambda_{\alpha(j)} \|y_{\alpha(j)}^*\| \xrightarrow[j \in J]{} +\infty$. Moreover since $\|z_j^*\| = 1$ and $\|b_j^*\| \leq k_g$ (where k_g is a Lipschitz constant of g at x_0), we may assume (extracting another subnet if necessary) that $z_j^* \xrightarrow[k \in J]{w^*} z^* \in N_G(D; g(x_0))$ and $b_j^* \xrightarrow[k \in J]{w^*} b^*$. By Lemma 2.5 in [9] we have $b^* \in \partial_G(z^* \circ g)(x_0)$ and by (3)

$$(4) \quad 0 = b^* + c^* \text{ for some } c^* \in N_G(C, x_0).$$

Let us prove that $z^* \neq 0$. Choose, by Lemma 1.7, elements $h_1, \dots, h_m \in X$ such that

$$(5) \quad \|\zeta^*\| \leq \text{Max}_{1 \leq k \leq m} \langle \zeta^*, h_k \rangle$$

for all $\zeta^* \in N_G(D, g(x_0))$. As $\|z_j^*\| = 1$, it follows from (5) that for all $j \in J$

$$1 \leq \text{Max}_{1 \leq k \leq m} \langle z_j^*, h_k \rangle$$

and hence

$$1 \leq \text{Max}_{1 \leq k \leq m} \langle z^*, h_k \rangle$$

which implies that $z^* \neq 0$. Thus relation (4) is in contradiction with assumption (R'). This completes the proof of the theorem. ■

REMARK. 1) The above proof also shows that the theorem still holds whenever for nets of points belonging to $N_G(D, \cdot)$ weak-star and norm convergences to zero coincide. 2) The method of Ioffe [13] also shows the second member of the inclusion of the theorem is w^* -closed. ■

A similar result also holds for the Clarke normal cone. See also Aubin [1], Rockafellar [27] and Ward and Borwein [29] for a similar result under conditions like (R'') but in the finite dimensional setting.

THEOREM 1.11. *Let C and D be two closed subsets of X and Y respectively and let $g: X \rightarrow Y$ be strictly differentiable at $x_0 \in C \cap g^{-1}(D)$. Assume that D is epi-Lipschitz-like at $y_0 := g(x_0)$ and assume also that (for $\Lambda := g'(x_0)^*$) the following regularity condition holds*

$$(R'') \quad \Lambda^{-1}(N_c(C; x_0)) \cap (-N_c(D; y_0)) = \{0\}.$$

Then

$$N_c(C \cap g^{-1}(D); x_0) \subset N_c(C; x_0) + \Lambda(N_c(D; y_0)).$$

PROOF. It is enough to use the result of Corollary 1.5 and to repeat, with the appropriate modifications, the arguments of Theorem 1.10. ■

2. Subdifferential calculus. The result of the preceding section can be employed to obtain some general new subdifferential calculus rules. For use in subsequent arguments we need first to recall some notions.

Let $f: X \rightarrow \bar{\mathbb{R}} := \mathbb{R} \cup \{+\infty, -\infty\}$ be lower semicontinuous near x_0 with $|f(x_0)| < \infty$. The function f is called (see [2]) *compactly epi-Lipschitz* or *epi-Lipschitz-like* at x_0 if $\text{epi} f$ is compactly epi-Lipschitz or epi-Lipschitz-like at $(x_0, f(x_0))$. We refer to [2] for many examples of such functions. Note that it is not difficult to see that f is compactly epi-Lipschitz at x_0 iff there exist two real numbers β and $r > 0$ and a $\|\cdot\|$ -compact subset K of X such that

$$\inf_{k \in K} t^{-1}[f(x + tb + tk) - f(x)] \leq \beta$$

for all $t \in]0, r]$, $b \in rB_X$ and $x \in x_0 + rB_X$ satisfying $|f(x) - f(x_0)| \leq r$.

Borwein [2] has proved

- (a) if f is directionally Lipschitzian at x_0 then f is epi-Lipschitz-like at x_0 ,
- (b) if f is epi-Lipschitz-like at x_0 then it is compactly epi-Lipschitz at x_0 ,
- (c) if X is finite dimensional then each function is epi-Lipschitz-like at each point of its domain.

We are now ready to prove our results concerning the G -approximate and Clarke subdifferentials of the sum and the composition of functions.

The first theorem is an extension of Theorem 5.6 of Ioffe [13] and Theorem 4.1 of Mordukhovich [23]. Recall (Ioffe [13]) that $x^* \in \partial_G f(x)$ iff $(x^*, -1) \in N_G(\text{epi} f; x, f(x))$.

THEOREM 2.1. *Let $g: X \rightarrow Y$ be strongly compactly Lipschitzian at x_0 and let $f: X \rightarrow \mathbb{R} \cup \{+\infty\}$ and $h: Y \rightarrow \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous near x_0 and $y_0 := g(x_0)$ respectively (with $f(x_0) < \infty$ and $h(y_0) < \infty$). Assume that h is compactly epi-Lipschitz at y_0 and*

$$(H_G) \quad y^* \in \partial_G^\infty h(y_0) \text{ and } 0 \in \partial_G(y^* \circ g)(x_0) + \partial_G^\infty f(x_0) \Rightarrow y^* = 0.$$

Then

$$\partial_G(f + h \circ g)(x_0) \subset \partial_G f(x_0) + \bigcup_{y^* \in \partial_G h(y_0)} \partial_G(y^* \circ g)(x_0).$$

PROOF. Our results in Section 1 allow us to follow several parts of the proof of Theorem 5.6 in Ioffe [13]. Put

$$\begin{aligned} C &:= \{(x, r, s) \in X \times \mathbb{R} \times \mathbb{R} : f(x) \leq r\} \\ D &:= \{(y, r, s) \in Y \times \mathbb{R} \times \mathbb{R} : h(y) \leq s\} \\ B &:= \{(x, r, s) \in X \times \mathbb{R} \times \mathbb{R} : f(x) + h \circ g(x) \leq r + s\}. \end{aligned}$$

By Lemma 5.5 in [11] we have $N_G(B; z_0) \subset N_G(C \cap \hat{g}^{-1}(D); z_0)$ where $z_0 := (x_0, f(x_0), h(y_0))$ and \hat{g} is the mapping from $X \times \mathbb{R} \times \mathbb{R}$ into itself defined by $\hat{g}(x, r, s) =$

$(g(x), r, s)$. Let us show that the regularity condition in Theorem 1.10 holds. Observe first that

$$(1) \quad N_G(C; x_0, f(x_0), s) = \{(x^*, r^*, 0) \in X^* \times \mathbb{R} \times \mathbb{R} : (x^*, r^*) \in N_G(\text{epi } f; x_0, f(x_0))\}$$

and

$$(2) \quad N_G(D; y_0, r, h(y_0)) = \{(y^*, 0, s^*) \in Y^* \times \mathbb{R} \times \mathbb{R} : (y^*, s^*) \in N_G(\text{epi } h; y_0, h(y_0))\}.$$

Then for all $(y^*, 0, s^*) \in N_G(D; \hat{g}(z_0))$ and $(x^*, r^*, 0) \in N_G(C; z_0)$ satisfying

$$0 \in (x^*, r^*, 0) + \partial_G((y^*, 0, s^*) \circ \hat{g})(z_0)$$

we have

$$0 \in (x^*, r^*, 0) + \partial_G(y^* \circ g)(x_0) \times \{0\} \times \{s^*\}$$

and hence $r^* = s^* = 0$ and

$$(3) \quad 0 \in x^* + \partial_G(y^* \circ g)(x_0).$$

Therefore we have $x^* \in \partial_G^\infty f(x_0)$ and $y^* \in \partial_G^\infty h(y_0)$ and hence (3) and the assumption (H_G) imply that $y^* = 0$. The regularity condition in Theorem 1.10 is thus satisfied and hence this theorem ensures that

$$(4) \quad N_G(B; z_0) \subset N_G(C; z_0) + \bigcup_{z^* \in N_G(D; \hat{g}(z_0))} \partial_G(z^* \circ \hat{g})(z_0).$$

Now let $x^* \in \partial_G(f+h \circ g)(x_0)$. It is not difficult to verify that $(x^*, -1, -1) \in N_G(B; z_0)$ and hence, by (1), (2) and (4), there exist $(u^*, r^*) \in N_G(\text{epi } f; x_0, f(x_0))$ and $(y^*, s^*) \in N_G(\text{epi } h; y_0, h(y_0))$ such that

$$(x^*, -1, -1) \in (y^*, r^*, 0) + \partial_G(y^* \circ g)(x_0) \times \{0\} \times \{s^*\}.$$

So $r^* = s^* = -1$, $u^* \in \partial_G f(x_0)$, $y^* \in \partial_G h(y_0)$ and

$$x^* \in u^* + \partial_G(y^* \circ g)(x_0).$$

This completes the proof of the theorem. ■

COROLLARY 2.2. *Let $f, h: X \rightarrow \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous near x_0 and finite at x_0 . Assume that h is compactly epi-Lipschitz at x_0 and*

$$\partial_G^\infty f(x_0) \cap (-\partial_G^\infty h(x_0)) = \{0\}.$$

Then

$$\partial_G(f+h)(x_0) \subset \partial_G f(x_0) + \partial_G h(x_0). \quad \blacksquare$$

Although this result extends Theorem 4.1 of Mordukhovich [23] where $X = Y = \mathbb{R}^n$ and Theorem 5.6 of Ioffe [13] where X and Y are general Banach spaces but h is directionally Lipschitz, it also provides a unified version for both theorems. Applying Theorem 1.3 instead of Theorem 1.10, the reader can easily see that the results of Theorem 2.1 and Corollary 2.2 are also valid for the G -nuclei of the G -subdifferentials.

The second theorem is the following new result which concerns the Clarke subdifferential. It generalizes in several ways the main calculus rules proved by Clarke [6], Ioffe [11], Rockafellar [25, 26], Ward and Borwein [29].

THEOREM 2.3. *Let $g: X \rightarrow Y$ be strictly differentiable at x_0 and let $f: X \rightarrow \mathbb{R} \cup \{+\infty\}$ and $h: Y \rightarrow \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous near x_0 and $y_0 = g(x_0)$ respectively (with $f(x_0) < \infty$ and $h(y_0) < \infty$). Assume that h is epi-Lipschitz-like at y_0 and that, for $\Lambda = g'(x_0)^*$,*

$$(H_c) \quad (-\partial_c^\infty f(x_0)) \cap \Lambda^{-1}(\partial_c^\infty h(x_0)) = \{0\}.$$

Then

$$\partial_c(f + h \circ g)(x_0) \subset \partial_c f(x_0) + \Lambda(\partial h(x_0)).$$

Before giving the proof of this theorem, let us consider the following lemma which has been stated without proof by Ioffe [11] in the finite dimensional setting. Because of its importance we sketch a proof.

LEMMA 2.4. *Let B, C, D and \hat{g} be defined as in the proof of Theorem 2.1. Assume that f and h are lower semicontinuous at x_0 and $y_0 := g(x_0)$ respectively and that g is continuous at x_0 . Then*

$$T_c(C \cap \hat{g}^{-1}(D); z_0) \subset T_c(B; z_0)$$

where $z_0 := (x_0, f(x_0), h(y_0))$.

PROOF. Note that the Clarke tangent cone is not isotone and hence we may not say that the lemma follows from the inclusion $C \cap \hat{g}^{-1}(D) \subset B$. Set $S := C \cap \hat{g}^{-1}(D)$ and take $(v, r, s) \in T_c(S; z_0)$. Let $(x_n, \alpha_n, \beta_n) \xrightarrow{B} z_0$ and $t_n \downarrow 0$. The lower semicontinuity of $f + h \circ g$ ensures for any $\varepsilon > 0$ that for n large enough

$$f(x_0) + h \circ g(x_0) + \varepsilon \geq \alpha_n + \beta_n \geq f(x_n) + h \circ g(x_n) \geq f(x_0) + h \circ g(x_0) - \varepsilon$$

and hence

$$(1) \quad (f + h \circ g)(x_n) \rightarrow f(x_0) + h \circ g(x_0).$$

We claim that $f(x_n) \rightarrow f(x_0)$ and $h \circ g(x_n) \rightarrow h \circ g(x_0)$. Indeed for any $\varepsilon > 0$, there exists, by the lower semicontinuity of f and $h \circ g$ at x_0 and by relation (1), some $N \in \mathbb{N}$ such that for every integer number $n \geq N$

$$f(x_0) - \varepsilon \leq f(x_n), \quad g(x_0) - \varepsilon/2 \leq g(x_n) \text{ and } f(x_n) + g(x_n) \leq f(x_0) + g(x_0) + \varepsilon/2.$$

The last two inequalities ensure, for every integer $n \geq N$, that $f(x_n) \leq f(x_0) + \varepsilon$; hence $f(x_0) - \varepsilon \leq f(x_n) \leq f(x_0) + \varepsilon$. So $f(x_n) \rightarrow f(x_0)$ and by (1) $h \circ g(x_n) \rightarrow h \circ g(x_0)$.

Set $\alpha'_n := f(x_n)$ and $\beta'_n := \alpha_n + \beta_n - \alpha'_n$. Then $\beta'_n \geq h \circ g(x_n)$, $\alpha'_n + \beta'_n = \alpha_n + \beta_n$, $\alpha'_n \rightarrow f(x_0)$, $\beta'_n \rightarrow g(x_0)$ and $(x_n, \alpha'_n, \beta'_n) \in S$. As $(v, r, s) \in T_c(S; z_0)$, it follows that there exists $(v_n, r_n, s_n) \rightarrow (v, r, s)$ such that for all $n \in \mathbb{N}$

$$(x_n, \alpha'_n, \beta'_n) + t_n(v_n, r_n, s_n) \in S \subset B.$$

As $\alpha'_n + \beta'_n = \alpha_n + \beta_n$ we obtain by definition of B that

$$(x_n, \alpha_n, \beta_n) + t_n(v_n, r_n, s_n) \in B$$

for all $n \in \mathbb{N}$ and hence $(v, r, s) \in T_c(B; z_0)$, by the sequential characterization of the Clarke tangent cone (Clarke [6]). ■

We can now follow the main idea of the proof of Corollary 4.4 in Ioffe [11].

PROOF OF THEOREM 2.3. Let B , C and D be as in the proof of Theorem 2.1. It is not difficult to see that the assumption (H_c) ensures the assumption of Theorem 1.11. Applying the theorem and Lemma 2.4, we obtain (for $S = C \cap \hat{g}^{-1}(D)$)

$$(1) \quad N_c(B; z_0) \subset N_c(S; z_0) \subset N_c(C; x_0) + \hat{\Lambda}(N_c(D; \hat{g}(z_0)))$$

where $\hat{\Lambda}$ is the adjoint mapping of the derivative of \hat{g} at z_0 . Moreover it is easily seen that

$$T_c(B; z_0) = \{(v, r, z) \in X \times \mathbb{R} \times \mathbb{R} : (v, r + s) \in T_c(\text{epi}(f + h \circ g); x_0, (f + h \circ g)(x_0))\},$$

$$T_c(C; z_0) = \{(v, r, s) \in X \times \mathbb{R} \times \mathbb{R} : (v, r) \in T_c(\text{epi}f; x_0, f(x_0))\}$$

and

$$T_c(D; z_0) = \{(v, r, s) \in X \times \mathbb{R} \times \mathbb{R} : (v, s) \in T_c(\text{epi}h \circ g; x_0, h \circ g(x_0))\}.$$

Therefore

$$N_c(B; z_0) = \{(x^*, t, t) \in X^* \times \mathbb{R} \times \mathbb{R} : (x^*, t) \in N_c(\text{epi}(f + h \circ g); x_0, (f + h \circ g)(x_0))\},$$

$$(2) \quad N_c(C; z_0) = \{(x^*, t, 0) \in X^* \times \mathbb{R} \times \mathbb{R} : (x^*, t) \in N_c(\text{epi}f; x_0, f(x_0))\}$$

and

$$(3) \quad N_c(D; z_0) = \{(x^*, 0, t) \in X^* \times \mathbb{R} \times \mathbb{R} : (x^*, t) \in N_c(\text{epi}h \circ g; x_0, h \circ g(x_0))\}.$$

Now take any $x^* \in \partial_c(f + h \circ g)(x_0)$. Then $(x^*, -1, -1) \in N_c(B; z_0)$ and, from (1), (2) and (3), there are $(u^*, -1) \in N_c(\text{epi}f; x_0, f(x_0))$ and $(v^*, -1) \in N_c(\text{epi}h \circ g; x_0, h \circ g(x_0))$ such that $x^* = u^* + \Lambda(v^*)$. This completes the proof. ■

Obviously applications to necessary optimality conditions for general nonsmooth infinite optimization problems can be derived. This is left to the reader.

ACKNOWLEDGMENT. We owe thanks to a referee for several useful suggestions that allowed to improve the presentation of the paper.

REFERENCES

1. J. P. Aubin, *Lipschitz behaviour of solutions to convex minimization problems*, Math. Oper. Res. **9**(1984), 87–111.
2. J. M. Borwein, *Epi-Lipschitz-like sets in Banach space: Theorems and Examples*, Nonlinear Anal. **11**(1987), 1207–1217.
3. J. M. Borwein and H. M. Strojwas, *Tangential approximations*, Nonlinear Anal. **9**(1985), 1347–1366.
4. F. H. Clarke, *Necessary conditions for nonsmooth problems in optimal control and the calculus of variations*, Thesis, University of Washington, Seattle (1973).
5. ———, *A new approach to Lagrange multipliers*, Math. Oper. Res. **1**(1976), 165–174.
6. ———, *Optimization and Nonsmooth Analysis*, John Wiley, New York (1983).
7. F. H. Clarke and N. Raissi, *Formules d'intersection en analyse non lisse*, Ann. Sci. Math. Québec, **14**(1990), 121–129.
8. S. Dolecki, *Tangency and differentiation: some applications of convergence*, Theory, Ann. Mat. Pura Appl. **130**(1982), 235–255.
9. B. El Abdouni and L. Thibault, *Lagrange multipliers for Pareto nonsmooth programming problems in Banach spaces*, Optimization **26**(1992), 277–285.
10. A. D. Ioffe, *Nonsmooth analysis: differential calculus of non-differentiable mappings*, Trans. Amer. Math. Soc. **266**(1981), 1–56.
11. ———, *Approximate subdifferentials and applications. I: The finite dimensional theory*, Trans. Amer. Math. Soc., **281**(1984), 389–416.
12. ———, *Approximate subdifferentials and applications II: Functions on locally convex spaces*, Matematika, **33**(1986), 111–128.
13. ———, *Approximate subdifferentials and applications III: The metric theory*, Mathematika, **36**(1989), 1–38.
14. A. Jourani and L. Thibault, *Metric regularity for strongly compactly Lipschitzian mappings*, Nonlinear Anal., to appear.
15. ———, *The use of metric graphical regularity in approximate subdifferential calculus rules in finite dimensions*, Optimization **21**(1990), 509–519.
16. ———, *Approximate subdifferential of composite functions*, Bull. Austral. Math. Soc., **47**(1993), 1–13.
17. J. L. Kelley, *General topology*, Springer-Verlag, New York, (1975).
18. A. Ya Kruger, *Properties of generalized differentials*, Sib. Math. J. **26**(1985), 1822–1832.
19. A. Ya Kruger and B. S. Mordukhovich, *Extreme points and the Euler equation in nondifferentiable optimization problems*, Dokl. Akad. Nauk. BSSR, **24**(1980), 684–687.
20. P. D. Loewen, *Limits of Fréchet normals in nonsmooth analysis*, In: Optimization and Nonlinear Analysis (eds. A. Ioffe, M. Marcus and S. Reich), Pitman Research Notes in Mathematics Series, Great Britain (1992).
21. B. S. Mordukhovich, *Maximum principle in the problem of time optimal control with nonsmooth constraints*, J. Appl. Math. Mech., **40**(1976), 960–969.
22. ———, *Nonsmooth analysis with nonconvex generalized differentials and adjoint mappings*, Dokl. Akad. Nauk. BSSR, **28**(1984), 976–979.
23. ———, *Approximation Methods in Problems of Optimization and Control*, Nauka, Moscow (1988).
24. H. Rådström, *An imbedding theorem for spaces of convex sets*, Proc. Amer. Math. Soc., **3**(1952), 165–169.
25. R. T. Rockafellar, *Generalized directional derivatives and subgradients of nonconvex functions*, Canad. J. Math. **32**(1980), 257–280.
26. ———, *Directionally Lipschitzian functions and subdifferential calculus*, Proc. London Math. Soc., **39**(1979), 331–355.
27. ———, *Extensions of subgradient calculus with applications to optimization*, Nonlinear Anal. **9**(1985), 665–698.

28. L. Thibault, *Subdifferentials of compactly Lipschitzian vector valued functions*, Ann. Mat. Pura Appl. **125**(1980), 157–192.
29. D. E. Ward and J. M. Borwein, *Nonsmooth calculus in finite dimensions*, SIAM J. Control Optim. **25**(1987), 1312–1340.

Université de Bourgogne
Laboratoire d'Analyse Appliquée et Optimisation
B.P. 138
21004 Dijon cedex
France

Université Montpellier II
Département des Sciences Mathématiques
34095 Montpellier cedex 5
France