

ON THE REMAINDER OF THE FIRST ORDER DEVELOPMENT OF CONVEX FUNCTIONS

DOMINIQUE AZÉ

RÉSUMÉ. On étudie le reste dans le développement conique du premier ordre d'une fonction convexe définie sur un espace normé de dimension infinie. On donne à la fois des majorations et des minoration de ce reste et l'action de la conjugaison sur ces estimations est caractérisée. En particulier, on met en évidence la dualité entre la classe des fonctions convexes coniquement différentiables et celle des fonctions convexes métriquement bien posées.

ABSTRACT. A study of the remainder term in the first order conical development of a convex function is given in the framework of infinite dimensional normed spaces. Lower and upper estimates of this remainder are tackled, and the action of the conjugacy operation on these estimates is characterized. In particular, we shed some light on the duality between the class of conically differentiable convex functions and the class of metrically well-set functions.

1. Introduction. The question of the second order differentiability of the so-called Moreau-Yosida approximation of closed convex functions has aroused a renewed interest recently (see [12], [13], [14], [18], [19]) mainly due to algorithmic considerations. In connection with this problem, C. Lemaréchal and C. Sagastizábal gave in [12] a thorough study of the remainder of the first-order development of convex functions. In the present paper, we generalize and complete the results of [12] in several directions. We get rid of the growth conditions required in the quoted paper, especially the growth condition from below which amounted, more or less, to local strong convexity of the function. Moreover, we work in the setting of normed linear spaces instead of finite dimensional Euclidean spaces. We also improve, in the convex case, the results given in [21]. The paper is organized as follows. In section 2, some technical results on Lipschitz regularization and metrically well-set functions are recalled and established. In section 3 and 4, local upper and lower estimates on the remainder of the first order development of convex functions are tackled and the action of the conjugacy operation on the classes of functions appearing naturally in this setting (see Definitions 2, 3, 4, 5) is characterized.

Reçu le 12 août 1998 et, sous forme définitive, le 12 octobre 1998.

2. Notations and preliminary results. Let us consider two normed vector spaces X and Y whose closed unit balls are denoted by B_X and B_Y . We will denote by $B(x, r)$ the closed ball with center $x \in X$ and radius $r \geq 0$. We will also assume that $X = Y^*$ or $Y = X^*$ in such a way that

$$\|x\| = \sup_{y \in B_Y} \langle x, y \rangle \text{ for all } x \in X,$$

and

$$\|y\| = \sup_{x \in B_X} \langle x, y \rangle \text{ for all } y \in Y,$$

where $\langle \cdot, \cdot \rangle$ denotes the pairing between a space and its dual. In the sequel $\Gamma_0(X)$ (resp. $\Gamma_0(Y)$) will denote the set of convex proper functions which are supremum of a family of affine functions of the type $\langle \cdot, y \rangle + c$ (resp. $\langle x, \cdot \rangle + C$). Given a subset C of X we set

$$d(x, C) = \inf\{\|x - z\| : z \in C\}$$

with the convention $d(x, \emptyset) = +\infty$ and we denote for $C, D \subset X$

$$e(C, D) = \sup_{x \in C} d(x, D) \text{ if } D \neq \emptyset, \quad e(C, \emptyset) = +\infty \text{ if } C \neq \emptyset$$

with $e(\emptyset, D) = 0$. We denote by i_S the indicator function of a set $S \subset X$ defined by

$$i_S(z) = \begin{cases} 0 & \text{if } z \in S \\ +\infty & \text{if } z \in X \setminus S. \end{cases}$$

Recall that given a function $f : X \longrightarrow \mathbb{R} \cup \{+\infty\}$, one defines for all $y \in Y$

$$f^*(y) = \sup\{\langle x, y \rangle - f(x) : x \in X\}.$$

Given a subset C of X its support function σ_C is the function $\sigma_C = (i_C)^*$. For a convex function $f : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ we define for each $\varepsilon \geq 0$

$$\partial_\varepsilon f = \{(x, y) \in X \times Y : f(x) + f^*(y) - \langle x, y \rangle \leq \varepsilon\},$$

and for all $x \in X$,

$$\partial_\varepsilon f(x) = \{y \in Y : (x, y) \in \partial_\varepsilon f\}.$$

If $f \in \Gamma_0(X)$, then for each $x \in \text{dom } f = \{z \in X : f(z) < +\infty\}$ and for each $\varepsilon > 0$ the set $\partial_\varepsilon f(x)$ is nonempty. We shall also use

$$\partial f(x) = \partial_0 f(x) = \bigcap_{\varepsilon > 0} \partial_\varepsilon f(x).$$

Let N be the set of nondecreasing functions $\alpha : \mathbb{R}_+ \longrightarrow \mathbb{R} \cup \{+\infty\}$ such that $\alpha(0) = 0$ and let Φ be the set those $\alpha \in N$ which are firm, that is $\alpha(t) > 0$ for $t > 0$. Let us also denote by Ψ the set of functions $\beta \in N$ such that $\lim_{t \downarrow 0} t^{-1}\beta(t) = 0$. A function α defined on $\mathbb{R}_+$ will be extended to a function $\tilde{\alpha}$ defined on the whole $\mathbb{R}$ by $\tilde{\alpha}(t) = \alpha(|t|)$.

Given such a function $\alpha : \mathbb{R}_+ \longrightarrow \mathbb{R} \cup \{+\infty\}$, we will denote by α^* the function defined for all $s \in \mathbb{R}_+$ by

$$\alpha^*(s) = \sup_{t \geq 0} (st - \alpha(t)).$$

As well-known, one has $\alpha^* = \tilde{\alpha}^*$ on $\mathbb{R}_+$ where $\tilde{\alpha}^*$ stand for the usual Legendre-Fenchel conjugate of the function $\tilde{\alpha}$. A function $\alpha \in N$ is said to be starshaped at 0 if $\alpha(ct) \leq c\alpha(t)$ for all $t \in \mathbb{R}_+$ and $c \in [0, 1]$. As well-known (see Lemma 1 in [1] and Lemma 2.1 in [2]) one has $\alpha^* \in \Psi$ whenever $\alpha \in \Phi$ is starshaped at 0 and $\beta^* \in \Phi$ whenever $\beta \in \Psi$. Given a subset $S \subset X$ and $\alpha \in N$, let us set for all $x \in X, y \in Y$

$$h(x) = \alpha(d(x, S)) \text{ and } k(y) = \alpha(\|y\|) + \sigma_S(y).$$

Then (see [4] Lemma 2) one has

$$h^*(y) = \alpha^*(\|y\|) + \sigma_S(y). \quad (1)$$

moreover assuming that S is convex and $\sigma(X, Y)$ -closed one has

$$k^*(x) = \alpha^*(d(x, S)). \quad (2)$$

In the sequel we shall use the following Lipschitz regularization for convex functions (see [8]). Given $f \in \Gamma_0(X)$ and $s > 0$ we set

$$f_{[s]}(x) = (f + s\|\cdot\|)(x) = \inf\{f(z) + s\|z - x\| : z \in X\}.$$

Observe that $f_{[s]}$ is convex and that

$$(f_{[s]})^* = f^* + i_{sB_Y}.$$

Easy adaptation to our setting¹ of Proposition 2.1, 2.2 and 2.3 of [8] given in [3] show that for all $f \in \Gamma_0(X)$, for all $s > d(0, \text{dom } f^*)$, the function $f_{[s]}$ is convex proper Lipschitzian with modulus s and $\sigma(X, Y)$ -l.s.c. Moreover assuming that for each $\varepsilon > 0$, $\partial_\varepsilon f(x) \cap sB_Y \neq \emptyset$, then

$$f_{[s]}(x) = f(x) \quad (3)$$

and

$$\partial_\varepsilon f_{[s]}(x) = \partial_\varepsilon f(x) \cap sB_Y \text{ for all } \varepsilon \geq 0.$$

We will also use the following

Lemma 2.1. *Let $f \in \Gamma_0(X)$ and $s > 0$. Let $x \in X$ be such that*

$$\emptyset \neq \partial f_{[s]}(x) \subset \text{int}(sB_Y),$$

then

$$f_{[s]}(x) = f(x) \text{ and } \partial f_{[s]}(x) = \partial f(x).$$

¹in which for example a convex Lipschitzian function $f \in \Gamma_0(X)$ is not necessarily $\sigma(X, Y)$ -lower semicontinuous.

Proof. Let $y \in \partial f_{[s]}(x)$. One has $x \in \partial(f^* + i_{sB_Y})(y)$. Setting $g = f^*$ and using the fact that $y \in \text{int}(sB_Y)$, it follows that for all $v \in Y$ and for all $t \in [0, 1]$ small enough one has $y + tv \in sB_Y$ and then $g(y + tv) \geq g(y) + t\langle x, v \rangle$. Thus

$$g'(y; v) \geq \langle x, v \rangle \text{ for all } v \in Y$$

which implies $x \in \partial g(y)$ and then $y \in \partial f(x)$ yielding

$$\partial f(x) \cap sB_Y \neq \emptyset \text{ and } \partial f_{[s]}(x) \subset \partial f(x).$$

Using (3) and (4), we get

$$f_{[s]}(x) = f(x) \text{ and } \partial f(x) \cap sB_Y = \partial f_{[s]}(x) \subset \text{int}(sB_Y)$$

thus

$$\partial f(x) \subset \text{int}(sB_Y)$$

which ends the proof of the lemma. $\square$

Given $f \in \Gamma_0(X)$ we set

$$m_f = \inf_{x \in X} f(x) = -f^*(0) \text{ and } S_f = \text{argmin } f = f^{-1}(m_f) = \partial f^*(0).$$

Let us give the main definitions we shall be dealing with in the sequel.

Definition 1. (See [5],[7] and [17]). We say that $f \in \Gamma_0(X)$ is metrically well-set if there exists $\alpha \in \Phi$ such that for all $x \in X$,

$$f(x) \geq -f^*(0) + \alpha(d(x, \partial f^*(0))) = m_f + \alpha(d(x, S_f)). \quad (5)$$

The property required in the previous definition is also called stability in [7]. It amounts to say that $\lim_{n \rightarrow \infty} d(x_n, \text{argmin } f) = 0$ for each minimizing sequence $(x_n)_{n \in \mathbb{N}}$.

Definition 2. Given $g \in \Gamma_0(Y)$, we say that g is firmly subdifferentiable at $y_0 \in Y$ if there exists $x_0 \in X$, $\alpha \in \Phi$ and $\delta > 0$ such that

$$g(y) + i_{B(y_0, \delta)}(y) \geq g(y_0) + \langle x_0, y - y_0 \rangle + \alpha(d(y, \partial g^*(x_0))) \text{ for all } y \in Y \quad (6)$$

that is $g - \langle x_0, \cdot \rangle + i_{B(y_0, \delta)}$ is metrically well-set.

In Definition 2, one has necessarily $x_0 \in \partial g(y_0)$; thus requiring some uniformity in inequality (6) leads to :

Definition 3. Given $g \in \Gamma_0(Y)$, we say that g is uniformly firmly subdifferentiable at $y_0 \in Y$ if there exists $\alpha \in \Phi$ and $\delta > 0$ such that

$$g(y) + i_{B(y_0, \delta)}(y) \geq g(y_0) + \sigma_{\partial g(y_0)}(y - y_0) + \alpha(\|y - y_0\|) \text{ for all } y \in Y.$$

Definition 4. Given $f \in \Gamma_0(X)$, we say that f is firmly differentiable at $x_0 \in \text{dom } f$ if there exists $\beta \in \Psi$ and $y_0 \in \partial f(x_0)$ such that

$$f(x) \leq f(x_0) + \langle x - x_0, y_0 \rangle + \beta(d(x, \partial f^*(y_0))) \text{ for all } x \in X. \quad (7)$$

In (7), one has $\beta(d(x, \partial f^*(y_0))) \leq \beta(\|x - x_0\|)$, so that firm differentiability entails Fréchet differentiability.

Definition 5. Given $f \in \Gamma_0(X)$, we say (see [15]) that f is conically differentiable at $x_0 \in \text{dom } f$ if there exists $\beta \in \Psi$ such that

$$f(x) \leq f(x_0) + \sigma_{\partial f(x_0)}(x - x_0) + \beta(\|x - x_0\|) \text{ for all } x \in X. \quad (8)$$

It is known that $\sigma_{\partial f(x_0)}(\cdot) = f'_-(x_0; \cdot)$ where

$$f'_-(x_0; u) = \liminf_{\substack{t \downarrow 0 \\ v \rightarrow u}} t^{-1} (f(x_0 + tv) - f(x_0))$$

is the l.s.c. regularization of the usual one-sided directional derivative $f'(x_0; \cdot)$, so that conical differentiability means a uniform approximation of Fréchet type for the generalized directional derivative $f'_-(x_0, \cdot)$.

It is also known (see [11, 4, 17]) that the greatest function $\alpha \in \Phi$ satisfying (5) denoted by α_f is given by

$$\alpha_f(t) = \inf\{f(x) - m_f : d(x, S_f) \geq t\} = \inf\{f(x) - m_f : d(x, S_f) = t\},$$

and that its lower semicontinuous regularization $\bar{\alpha}_f$ which is still firm is starshaped at 0. For such a function f , we set

$$\Phi_f = \{\alpha \in \Phi : \alpha \text{ starshaped and } f(x) \geq m_f + \alpha(d(x, S_f)) \text{ for all } x \in X\}.$$

Observe that local estimates are allowed in (7) and (8) whenever $\beta = \bar{\beta} + i_{[0, \varepsilon]}$ with $\bar{\beta} \in \Psi$ and $\varepsilon > 0$. Assuming that $f \in \Gamma_0(X)$ is conically differentiable at $x_0 \in X$, the smallest $\beta \in \Psi$ in (8) is given, for all $t \in \mathbb{R}_+$ by

$$\beta(t) = \sup\{f(x) - f(x_0) - \sigma_{\partial f(x_0)}(x - x_0) : \|x - x_0\| \leq t\},$$

where the supremum is taken over $x \in X$ such that $\sigma_{\partial f(x_0)}(x - x_0)$ is finite, is convex. Assuming that f is conically differentiable at x_0 , we set

$$\Psi_f = \{\beta \in \Psi \cap \Gamma_0(\mathbb{R}_+) : f(x) \leq f(x_0) + \sigma_{\partial f(x_0)}(x - x_0) + \beta(\|x - x_0\|) \text{ for all } x \in X\}.$$

In the finite dimensional case, one has the following simple result.

Proposition 2.2. *Let $f \in \Gamma_0(\mathbb{R}^d)$ be such that $S_f = \text{argmin } f$ is bounded. Then f is metrically well-set.*

Proof. Let us set for $t \in \mathbb{R}_+$

$$\alpha(t) = \inf\{f(x) - m_f : d(x, S_f) = t\}.$$

As f is lower semicontinuous and $\{x \in \mathbb{R}^d : d(x, S_f) = t\}$ is compact, the infimum is attained in the definition of α in such a way that $\alpha(t) > 0$ for all $t > 0$. Let us prove that α is nondecreasing. Let $t_1, t_2 \in \mathbb{R}_+$ be such that $t_1 \leq t_2$ and $\alpha(t_2) < +\infty$. Given $\varepsilon > 0$, there exists $x_2 \in \{x \in \mathbb{R}^d : d(x, S_f) = t_2\}$ such that $f(x_2) - m_f \leq \alpha(t_2) + \varepsilon$. Let $x \in S_f$ be the projection of x_2 on S_f and let $x_1 = (t_1/t_2)x_2 + (1 - t_1/t_2)x$. One has $d(x_1, S_f) = t_1$ so that

$$\alpha(t_1) \leq f(x_1) - m_f \leq \frac{t_1}{t_2}f(x_2) - m_f \leq \frac{t_1}{t_2}(\alpha(t_2) + \varepsilon).$$

Letting ε decrease to 0 yields $\alpha(t_1) \leq t_1/t_2\alpha(t_2)$. $\square$

Using these notations it can be proved

Theorem 2.3. ([4, Proposition 2.4]). *Let $f, h \in \Gamma_0(X)$. Then*

- a) *f metrically well-set implies that f^* is conically differentiable at 0.*
- b) *h conically differentiable at 0 implies that h^* is metrically well-set.*

Moreover one has

$$(\Phi_f)^* \subset \Psi_{f^*} \text{ and } (\Psi_h)^* \subset \Phi_{h^*}.$$

Example 1.

- a) The function $f = i_{\mathbb{R}_-}$ is metrically well-set with $\alpha_f = i_{\{0\}}$ and its conjugate $f^* = i_{\mathbb{R}_+}$ is conically differentiable at 0 with $\beta = 0$.
- b) The function f defined on $\mathbb{R}^2$ by $f = \sigma_{\mathcal{P}}$ where

$$\mathcal{P} = \left\{ (x, y) \in \mathbb{R}^2 : y \leq -\frac{x^2}{4} \right\}$$

is not metrically well-set. Indeed it is equal to

$$f(x, y) = \begin{cases} +\infty & \text{if } y \leq 0 \text{ and } x \neq 0 \\ 0 & \text{if } (x, y) = (0, 0) \\ x^2/y & \text{if } y > 0 \end{cases}$$

so that $S_f = \{0\} \times \mathbb{R}_+$. Thus for $\varepsilon > 0$ and $y > 0$ one has

$$d((\varepsilon, y), S_f) = \varepsilon \text{ and } \lim_{y \rightarrow +\infty} f(\varepsilon, y) = 0.$$

The following result will be also useful. We give its proof for the convenience of the reader.

Theorem 2.4. ([4, Theorem 2.2]). *Assume that $f \in \Gamma_0(Y)$ with $Y = X^*$ is such that $S_f = \operatorname{argmin} f$ is bounded. Assume also that for some $\alpha \in N$ one has*

$$f(y) \geq m_f + \alpha(d(y, S_f))$$

for all $y \in Y$. Then, for all $\varepsilon > 0$, one has

$$\partial f^*(x) \subset \partial f^*(0) + \varepsilon B_Y \text{ for all } x \in X \text{ such that } \|x\| < \varepsilon^{-1}\alpha(\varepsilon).$$

Proof. Let $\|x\| < \varepsilon^{-1}\alpha(\varepsilon)$, one has

$$\inf_{X^*}(f - \langle x, \cdot \rangle) = \min_{N_\varepsilon(S_f)}(f - \langle x, \cdot \rangle),$$

where $N_\varepsilon(S_f) = \{y \in X^* : d(y, S_f) \leq \varepsilon\}$. Indeed, it is clear that

$$\inf_{X^*}(f - \langle x, \cdot \rangle) \leq \min_{N_\varepsilon(S_f)}(f - \langle x, \cdot \rangle).$$

On the other hand, given $y \in \operatorname{argmin}(f - \langle x, \cdot \rangle + i_{N_\varepsilon(S_f)})$, one has $d(y, S_f) < \varepsilon$. If not as

$$f(y) - \langle x, y \rangle \leq m_f - \langle x, z \rangle \text{ for each } z \in S_f$$

and $d(y, S_f) \geq \varepsilon$, we should get

$$\alpha(\varepsilon) \leq f(y) - m_f \leq \langle x, y - z \rangle \leq \|x\|\|y - z\|$$

for all $z \in S_f$ yielding

$$\alpha(\varepsilon) \leq \|x\|\varepsilon,$$

a contradiction. Thus we get $y \in \operatorname{argmin}(f - \langle x, \cdot \rangle)$. From what precedes we obtain,

$$\operatorname{argmin}(f - \langle x, \cdot \rangle) \cap \{y \in X^* : d(y, S_f) < \varepsilon\} \neq \emptyset,$$

$$\operatorname{argmin}(f - \langle x, \cdot \rangle) \cap \{y \in X^* : d(y, S_f) = \varepsilon\} \neq \emptyset.$$

It follows that

$$\operatorname{argmin}(f - \langle x, \cdot \rangle) \subset \{y \in X^* : d(y, S_f) < \varepsilon\},$$

thus

$$\partial f^*(x) \subset \{y \in X^* : d(y, S_f) < \varepsilon\} = \{y \in X^* : d(y, \partial f^*(0)) < \varepsilon\}. \quad \square$$

Corollary 2.5. *Let $\beta \in \Phi \cap \Gamma_0(\mathbb{R})$. Then for all $\delta > 0$ one has*

$$\partial \beta^*(t) \subset [-\delta, \delta]$$

for all $t \in \mathbb{R}_+$ such that $t < \delta^{-1}\beta(\delta)$.

Proof. As $\operatorname{argmin} \beta = \{0\}$ is bounded one gets $\beta(t) \geq \beta(d(t, \operatorname{argmin} \beta))$ for all $t \in \mathbb{R}_+$, thus we can apply theorem 2.4 with $f = \beta$ and $\alpha = \beta$ yielding the announced result since $\partial \beta^*(0) = \operatorname{argmin} \beta = \{0\}$. $\square$

3. Upper estimates on the remainder. In this section, we prove that a function $\varphi \in \Gamma_0(X)$ is conically differentiable at x_0 with $\partial \varphi(x_0)$ bounded if and only there exists $y_0 \in \partial \varphi(x_0)$ such that φ^* is firmly subdifferentiable at y_0 .

Proposition 3.1. *Let $\varphi \in \Gamma_0(X)$ with $Y = X^*$ be conically differentiable at $x_0 \in \text{dom } \varphi$. Assume that $\partial\varphi(x_0)$ is bounded. Then, for all $\beta \in \Psi_\varphi$ and for all $\varepsilon > 0$,*

$$\partial\varphi(x) \subset \partial\varphi(x_0) + \varepsilon B_{X^*} \text{ for all } x \in X \text{ such that } \|x - x_0\| < \varepsilon^{-1}\beta^*(\varepsilon). \quad (9)$$

Moreover, assuming that $\beta = \bar{\beta} + i_{[0,\delta]}$ with $\bar{\beta}$ firm, then for all $\varepsilon \in (0, \delta^{-1}\bar{\beta}(\delta))$, one has

$$\partial\varphi(x) \subset \partial\varphi(x_0) + \varepsilon B_{X^*} \text{ for all } x \in X \text{ such that } \|x - x_0\| < \varepsilon^{-1}\bar{\beta}^*(\varepsilon). \quad (10)$$

Proof. Let us define $f \in \Gamma_0(X^*)$ by $f = \varphi^* - \langle x_0, \cdot \rangle$. As $f^* = \varphi(x_0 + \cdot)$ is conically differentiable at 0, we deduce from Theorem 2.3 that f is metrically well-set with $\alpha = \beta^*$. Observe that $S_f = \partial\varphi(x_0)$ is bounded. Thus (9) follows from Theorem 2.4. In the case where $\beta = \bar{\beta} + i_{[0,\delta]}$, one has $\beta^* = (\bar{\beta}^*)_{[\delta]}$. Now $\bar{\beta}^*(\varepsilon) = (\bar{\beta}^*)_{[\delta]}(\varepsilon)$ provided that $\partial\bar{\beta}^*(\varepsilon) \cap [-\delta, \delta] \neq \emptyset$. Applying Corollary 2.5 we get $\partial\bar{\beta}^*(\varepsilon) \subset [-\delta, \delta]$ whenever $0 < \varepsilon < \delta^{-1}\bar{\beta}(\delta)$ yielding (10). $\square$

Corollary 3.2. *Let $\varphi \in \Gamma_0(X)$ with $Y = X^*$ and let $x_0 \in \text{dom } \varphi$ be such that for some $c > 0$ and $\delta > 0$*

$$\varphi(x) \leq \varphi(x_0) + \sigma_{\partial\varphi(x_0)}(x - x_0) + \frac{c}{2}\|x - x_0\|^2 + i_{[0,\delta]}(\|x - x_0\|).$$

with $\partial\varphi(x_0)$ bounded. Then, for all $x \in U(x_0, \delta/4)$, one has

$$e(\partial\varphi(x), \partial\varphi(x_0)) \leq 2c\|x - x_0\|.$$

Proof. Let us set $\bar{\beta}(t) = ct^2/2$ and let us choose

$$\varepsilon \in \left(2c\|x - x_0\|, \frac{c}{2}\delta\right).$$

We have $\varepsilon \in (0, \delta^{-1}\bar{\beta}(\delta))$ and $\|x - x_0\| < \varepsilon^{-1}\bar{\beta}^*(\varepsilon)$. Thus, from Proposition 3.1, we obtain

$$e(\partial\varphi(x), \partial\varphi(x_0)) \leq \varepsilon.$$

Hence, the conclusion of the corollary follows by letting ε decrease to $2c\|x - x_0\|$. $\square$

In fact, Proposition 3.1 and next Proposition 3.3 show that a function $\varphi \in \Gamma_0(X)$ is conically differentiable at x_0 with $\partial\varphi(x_0)$ bounded if and only if the multimapping $x \rightrightarrows \partial\varphi(x)$ is Hausdorff upper semicontinuous at x_0 with $\partial\varphi(x_0)$ bounded.

Proposition 3.3. *Let $\varphi \in \Gamma_0(X)$ with $Y = X^*$, let $x_0 \in X$ be such that $\partial\varphi(x_0)$ is bounded and*

$$e(\partial\varphi(x), \partial\varphi(x_0)) \leq \gamma(\|x - x_0\|)$$

for some $\gamma \in N$ such that $\lim_{t \downarrow 0} \gamma(t) = 0$. Then

$$\varphi(x) \leq \varphi(x_0) + \sigma_{\partial\varphi(x_0)}(x - x_0) + \beta(\|x - x_0\|)$$

for all x close enough to x_0 where $\beta \in \Psi$ is defined by

$$\beta(t) = t\gamma(t).$$

Proof. Let $s > 0$ and $\eta > 0$ be such that $\partial\varphi(x) \subset sB_Y$ for all $x \in B(x_0, \eta)$. It follows that, for all $x \in B(x_0, \eta)$, one has $\varphi_{[s]}(x) = \varphi(x)$ and $\partial\varphi_{[s]}(x) = \partial\varphi(x)$ leading to

$$\varphi_{[s]}(x) - \varphi_{[s]}(x_0) = \varphi(x) - \varphi(x_0) = \langle x - x_0, y_\theta \rangle$$

for some $\theta \in (0, 1)$ and $y_\theta \in \partial\varphi_{[s]}(x + \theta(x - x_0))$. Thus, we get

$$\varphi(x) - \varphi(x_0) \leq \sigma_{\partial\varphi(x_0)}(x - x_0) + \min_{y \in \partial\varphi(x_0)} \|y_\theta - y\| \|x - x_0\|$$

yielding

$$\varphi(x) - \varphi(x_0) \leq \sigma_{\partial\varphi(x_0)}(x - x_0) + \|x - x_0\| \gamma(\|x - x_0\|)$$

for all $x \in B(x_0, \eta)$ which ends the proof of the proposition. $\square$

Theorem 3.4. Let $\varphi \in \Gamma_0(X)$ with $Y = X^*$, let $x_0 \in X$, let $y_0 \in \partial\varphi(x_0)$ and let $\beta \in \Psi$ be such that

$$\varphi(x) \leq \varphi(x_0) + \sigma_{\partial\varphi(x_0)}(x - x_0) + \beta(\|x - x_0\|) \text{ for all } x \in X. \quad (11)$$

Then, for all $y \in Y$

$$\varphi^*(y) \geq \varphi^*(y_0) + \langle x_0, y - y_0 \rangle + \beta^*(d(y, \partial\varphi(x_0))). \quad (12)$$

Moreover, if $\beta(t) = \bar{\beta}(t) + i_{[0, \varepsilon]}(t)$ with $\bar{\beta} \in \Psi \cap$ and $\varepsilon > 0$ then

$$\varphi^*(y) \geq \varphi^*(y_0) + \langle x_0, y - y_0 \rangle + \bar{\beta}^*(d(y, \partial\varphi(x_0))) \quad (13)$$

for all $y \in Y$ with $\|y - y_0\| < \varepsilon^{-1}\bar{\beta}(\varepsilon)$.

Proof. Let us define $g \in \Gamma_0(X)$ by $g(s) = \varphi(x_0 + s)$ and $f = g^*$. One has $f(y) = \varphi^*(y) - \langle x_0, y \rangle$, $m_f = -\varphi(x_0)$ and $S_f := \operatorname{argmin} f = \partial\varphi(x_0)$. Assumption (11) amounts to say that

$$g(s) \leq g(0) + \sigma_{\partial g(0)}(s) + \beta(\|s\|) \text{ for all } s \in X.$$

Applying Theorem 2.3, we get

$$f(y) \geq m_f + \beta^*(d(y, S_f)) \text{ for all } y \in Y$$

which is exactly (12). Assume now that $\beta(t) = \bar{\beta}(t) + i_{[0, \varepsilon]}(t)$. We have $\beta^*(t) = (\bar{\beta}^*)_{[\varepsilon]}$. Applying Corollary 2.5, we obtain that

$$\partial\bar{\beta}^*(t) \subset [-\varepsilon, \varepsilon] \text{ for all } t \in \mathbb{R}_+ \text{ with } t < \varepsilon^{-1}\bar{\beta}(\varepsilon),$$

hence,

$$\beta^*(t) = \bar{\beta}^*(t) \text{ assuming that } t < \varepsilon^{-1}\bar{\beta}(\varepsilon)$$

yielding (13). $\square$

The preceding theorem admits the following converse.

Theorem 3.5. *Let $\psi \in \Gamma_0(Y)$ with $Y = X^*$, let $y_0 \in Y$, let $x_0 \in \partial\psi(y_0)$ and let $\alpha \in \Phi$ be such that*

$$\psi(y) \geq \psi(y_0) + \langle x_0, y - y_0 \rangle + \alpha(d(y, \partial\psi^*(x_0))) \text{ for all } y \in Y. \quad (14)$$

Then

$$\psi^*(x) \leq \psi^*(x_0) + \sigma_{\partial\psi^*(x_0)}(x - x_0) + \alpha^*(\|x - x_0\|) \text{ for all } x \in X. \quad (15)$$

Moreover, assuming that $\partial\psi^*(x_0) \subset B(y_0, r)$ and that, for some $\delta > r$,

$$\psi(y) + i_{B(y_0, \delta)} \geq \psi(y_0) + \langle x_0, y - y_0 \rangle + \alpha(d(y, \partial\psi^*(x_0))) \text{ for all } y \in Y \quad (16)$$

then,

$$\psi^*(x) \leq \psi^*(x_0) + \sigma_{\partial\psi^*(x_0)}(x - x_0) + \alpha^*(\|x - x_0\|) \quad (17)$$

for all $x \in X$ such that $\|x - x_0\| < (\delta - r)^{-1}\alpha(\delta - r)$.

Proof. Assumption (14) says that the function $f = \psi - \langle x_0, \cdot \rangle$ is metrically well-set. Thus we derive from Theorem 2.3 that function $f^* = \psi^*(x_0 + \cdot)$ is conically differentiable at 0 with $\beta = \alpha^*$ which is exactly (15). Now let us introduce the function $f = \psi - \langle x_0, \cdot \rangle + i_{B(y_0, \delta)}$. As $\partial\psi^*(x_0) \subset B(y_0, r)$, it follows that $\partial\psi^*(x_0) \subset B(y_0, \delta)$ thus $\partial\psi^*(x_0) \subset \operatorname{argmin} f$ so that f is metrically well-set. Observing that $f^* = (\psi^*)_{[\delta]}(x_0 + \cdot)$, we obtain from Theorem 2.3 that

$$(\psi^*)_{[\delta]}(x) \leq (\psi^*)_{[\delta]}(x_0) + \sigma_{\partial(\psi^*)_{[\delta]}(x_0)}(x - x_0) + \alpha^*(\|x - x_0\|) \text{ for all } x \in X.$$

Relying on the assumption $\partial\psi^*(x_0) \subset B(y_0, r)$ and using (4), we get

$$\partial(\psi^*)_{[\delta]}(x_0) = \partial\psi^*(x_0) \cap B(0, \delta) = \partial\psi^*(x_0).$$

Moreover, from Theorem 2.4, we obtain

$$\partial(\psi^*)_{[\delta]}(x_0 + s) \subset \partial\psi^*(x_0) + (\delta - r)B_Y \subset \operatorname{int}(B(y_0, \delta))$$

for all $s \in X$ such that $\|s\| < (\delta - r)^{-1}\alpha(\delta - r)$. Thus, from Lemma 2.1, we get

$$(\psi^*)_{[\delta]}(x) = \psi^*(x) \text{ for all } x \in X \text{ with } \|x - x_0\| < (\delta - r)^{-1}\alpha(\delta - r)$$

which ends the proof of the theorem. $\square$

4. Lower estimates on the remainder. In this section, we prove that a function $\psi \in \Gamma_0(Y)$ is uniformly firmly subdifferentiable at y_0 if and only if its conjugate ψ^* is firmly differentiable at all $x_0 \in \partial\psi(y_0)$.

Lemma 4.1. *Let $\psi \in \Gamma_0(Y)$ with $Y = X^*$ and let $\alpha \in \Phi$ be such that, for some $\delta > 0$*

$$\psi(h) + i_{B(0, \delta)}(h) \geq \psi(0) + \sigma_{\partial\psi(0)}(h) + \alpha(\|h\|) \text{ for all } h \in Y \quad (18)$$

and let $\varepsilon \in]0, \delta[$. Then

$$\partial\psi^*(x) \subset B(0, \varepsilon) \text{ for all } x \in X \text{ with } d(x, \partial\psi(0)) < \varepsilon^{-1}\alpha(\varepsilon). \quad (19)$$

Proof. Let $x_0 \in \partial\psi(0)$ and let $f = \psi - \langle x_0, \cdot \rangle + i_{B(0,\delta)}$. From assumption (18), one gets

$$f(h) \geq m_f + \alpha(\|h\|) \text{ for all } h \in Y$$

which entails $\operatorname{argmin} f = \{0\}$ and then

$$\partial f^*(0) = \partial(\psi + i_{B(0,\delta)})^*(x_0) = \{0\}.$$

From Theorem 2.4, we derive that

$$\partial f^*(s) \subset \partial f^*(0) + \varepsilon B_Y \text{ for all } s \in X \text{ with } \|s\| < \varepsilon^{-1}\alpha(\varepsilon),$$

yielding

$$\partial(\psi^*)_{[\delta]}(x_0 + s) \subset \varepsilon B_Y \text{ for all } x_0 \in \partial\psi(0) \text{ and for all } s \in X \text{ with } \|s\| < \varepsilon^{-1}\alpha(\varepsilon).$$

As $0 < \varepsilon < \delta$, it follows that for all $x \in X$ with $d(x, \partial\psi(0)) < \varepsilon^{-1}\alpha(\varepsilon)$ one has

$$\partial(\psi^*)_{[\delta]}(x) \subset \operatorname{int}(B(0, \delta)),$$

thus, from Lemma 2.1 we get $\partial(\psi^*)_{[\delta]}(x) = \partial\psi^*(x)$ and then $\partial\psi^*(x) \subset B(0, \varepsilon)$ which ends the proof of the lemma. $\square$

Theorem 4.2. *Let $\psi \in \Gamma_0(Y)$ with $Y = X^*$ and let $\alpha \in \Phi$ be such that, for some $\delta > 0$*

$$\psi(y) + i_{B(y_0, \delta)}(y) \geq \psi(y_0) + \sigma_{\delta\psi(y_0)}(y - y_0) + \alpha(\|y - y_0\|) \text{ for all } y \in Y.$$

Then for all $x_0 \in \partial\psi(y_0)$ the function ψ^* is firmly differentiable at x_0 and given $\varepsilon \in (0, \delta)$, one has

$$\psi^*(x) \leq \psi^*(x_0) + \langle x - x_0, y_0 \rangle + \alpha^*(d(x, \partial\psi(y_0))) \quad (20)$$

for all $x \in X$ with $d(x, \partial\psi(y_0)) < \varepsilon^{-1}\alpha(\varepsilon)$.

Proof. Let us set $\tilde{\psi}(h) = \psi(y_0 + h)$. One gets

$$\tilde{\psi}(h) + i_{B(0, \delta)}(h) \geq \tilde{\psi}(0) + \sigma_{\partial\psi(y_0)}(h) + \alpha(\|h\|) \text{ for all } h \in Y, \quad (21)$$

and $x_0 \in \partial\tilde{\psi}(0)$. Taking the conjugate on both sides of (21) and using (2), we derive that

$$\tilde{\psi}_{[\delta]}^*(x) \leq -\tilde{\psi}(0) + \alpha^*(d(x, \partial\psi(y_0))) = \psi^*(x_0) - \langle x_0, y_0 \rangle + \alpha^*(d(x, \partial\psi(y_0))) \quad (22)$$

for all $x \in X$. Applying Lemma 4.1 to the function $\tilde{\psi}$, we get

$$\partial\tilde{\psi}(x_0 + s) \subset B(0, \varepsilon)$$

for all $s \in X$ with $\|s\| < \varepsilon^{-1}\alpha(\varepsilon)$. It follows that

$$\tilde{\psi}_{[\delta]}^*(x_0 + s) = \tilde{\psi}^*(x_0 + s) = \psi^*(x_0 + s) - \langle x_0 + s, y_0 \rangle,$$

$s \in X$ with $\|s\| < \varepsilon^{-1}\alpha(\varepsilon)$. Returning to (22), we get (20) and the proof of the theorem is complete. $\square$

The preceding theorem admits the following converse.

Theorem 4.3. *Let $\varphi \in \Gamma_0(X)$ with $Y = X^*$ be firmly differentiable at $x_0 \in X$ and let $\bar{\beta} \in \Psi$ be firm, $\delta > 0$, $x_0 \in X$, $y_0 \in \partial\varphi(x_0)$ be such that*

$$\varphi(x) \leq \varphi(x_0) + \langle x - x_0, y_0 \rangle + \bar{\beta}(d(x, \partial\varphi^*(y_0))), \quad (23)$$

for all $x \in X$ such that $d(x, \partial\varphi^*(y_0)) \leq \delta$. Then φ^* is uniformly firmly subdifferentiable at y_0 . More precisely, for any $\varepsilon < \delta^{-1}\bar{\beta}(\delta)$ one has

$$\varphi^*(y) + i_{B(y_0, \varepsilon)}(y) \geq \varphi^*(y_0) + \sigma_{\partial\varphi^*(y_0)}(y - y_0) + \bar{\beta}^*(\|y - y_0\|) \text{ for all } y \in Y.$$

Proof. Observe that (23) is equivalent to

$$\varphi(x) \leq \varphi(x_0) + \langle x - x_0, y_0 \rangle + \beta(d(x, \partial\varphi^*(y_0))) \text{ for all } x \in X$$

where $\beta = \bar{\beta} + i_{[0, \delta]}$. Taking the conjugate on both sides, we get using (1)

$$\varphi^*(y) + i_{B(y_0, \varepsilon)}(y) \geq \varphi^*(y_0) + \sigma_{\partial\varphi^*(y_0)}(y - y_0) + \beta^*(\|y - y_0\|) \text{ for all } y \in Y.$$

Now we derive from Corollary 2.5 that

$$\beta^*(\|y - y_0\|) = \bar{\beta}_{[\delta]}^*(\|y - y_0\|) = \bar{\beta}^*(\|y - y_0\|)$$

whenever $\|y - y_0\| < \delta^{-1}\bar{\beta}(\delta)$, hence the result. $\square$

Résumé substantiel en français. Dans leurs articles [13] et [14] C. Lemaréchal et Q. Sagastizábal ont mis en évidence l'intérêt d'estimations précises du reste dans le développement du premier ordre d'une fonction convexe, quand on s'intéresse par exemple à la différentiabilité du second ordre de la régularisée de Moreau-Yosida d'une telle fonction. Dans ce travail, nous introduisons diverses notions de bonne position et de différentiabilité faible sur la classe des fonctions convexes semi-continues inférieurement et propres, et nous montrons que ces notions se correspondent par dualité. On généralise ainsi les résultats de [13] dans plusieurs directions : on ne suppose pas l'espace de dimension finie, et l'on montre que l'on peut supprimer dans [13] une condition de croissance qui restreignait la validité des résultats à la classe des fonctions fortement convexes. Dans la section 2, on donne quelques résultats techniques sur la régularisée Lipschitzienne des fonctions convexes, et dans les sections 3 et 4, on donne les résultats principaux.

Nous utilisons notations et définitions suivantes. Dans la suite, X et Y désignent des espaces normés avec $Y = X^*$ ou $X = Y^*$ et $\Gamma_0(X)$ désigne l'ensemble des fonctions convexes à valeurs dans $\mathbb{R} \cup \{+\infty\}$ qui sont $\sigma(X, Y)$ -semi-continues inférieurement et propres (i.e. $\not\equiv +\infty$). On note N l'ensemble des fonctions croissantes $\alpha : \mathbb{R}_+ \rightarrow \mathbb{R} \cup +\infty$ telles que $\alpha(0) = 0$; Φ , l'ensemble des fonctions $\alpha \in N$ qui sont fermes (i.e. $\alpha(t) > 0$ pour $t > 0$), et Ψ l'ensemble des fonctions $\beta \in N$ telles que $\lim_{t \downarrow 0} t^{-1}\beta(t) = 0$. On utilise les notations classiques en analyse convexe pour la conjuguée et le sous-différentiel d'une fonction convexe.

Définition 1. On dit que $f \in \Gamma_0(X)$ est métriquement bien posée, s'il existe $\alpha \in \Phi$ telle que, pour tout $x \in X$,

$$f(x) \geq -f^*(0) + \alpha(d(x, \partial f^*(0))) = m_f + \alpha(d(x, S_f))$$

où $m_f = \inf_x f$ et $S_f = f^{-1}(m_f)$.

Définition 2. On dit que $g \in \Gamma_0(Y)$ est fermement sous-différentiable en $y_0 \in Y$ s'il existe $x_0 \in X$, $\alpha \in \Phi$ et $\delta > 0$ tels que

$$g(y) + i_{B(y_0, \delta)}(y) \geq g(y_0) + \langle x_0, y - y_0 \rangle + \alpha(d(y, \partial g^*(x_0))) \text{ pour tout } y \in Y.$$

où i_S désigne la fonction indicatrice de l'ensemble S , c'est-à-dire que $g - \langle x_0, \cdot \rangle + i_{B(y_0, \delta)}$ est métriquement bien posée.

Définition 3. On dit que $g \in \Gamma_0(Y)$ est uniformément fermement sous-différentiable en $y_0 \in Y$ s'il existe $y_0 \in \Phi$ et $\delta > 0$ tels que

$$g(y) + i_{B(y_0, \delta)}(y) \geq g(y_0) + \sigma_{\partial g(y_0)}(y - y_0) + \alpha \|y - y_0\| \text{ pour tout } y \in Y.$$

où $\sigma_{\partial g(y_0)} = i_{\partial g(y_0)}^*$ est la fonction d'appui de $\partial g(y_0)$.

Définition 4. On dit que $g \in \Gamma_0(Y)$ est fermement différentiable en $x_0 \in \text{dom} f$, s'il existe $\beta \in \Psi$ et $y_0 \in \partial f(x_0)$ tels que

$$f(x) \leq f(x_0) + \langle x - x_0, y_0 \rangle + \beta(d(x, \partial f^*(y_0))) \text{ pour tout } x \in X.$$

Définition 5. On dit que $f \in \Gamma_0(X)$ est coniquement différentiable s'il existe $\beta \in \Psi$ tels que

$$f(x) \leq f(x_0) + \sigma_{\partial f(x_0)}(x - x_0) + \beta(\|x - x_0\|) \text{ pour tout } x \in X.$$

Il est à noter que l'on peut avoir des estimations locales dans les définitions précédentes, car on utilise des modules α et β qui peuvent prendre la valeur $+\infty$.

Les résultats principaux sont alors qu'une fonction $\phi \in \Gamma_0(X)$ est coniquement différentiable en x_0 avec $\partial\phi(x_0)$ borné si et seulement si la multiapplication $\partial\phi$ est semi-continue supérieurement au sens de Hausdorff en x_0 , c'est-à-dire que

$$\lim_{x \rightarrow x_0} \sup_{y \in \partial\phi(x)} d(y, \partial\phi(x_0)) = 0$$

avec $\partial\phi(x_0)$ borné, ce qui équivaut aussi à dire qu'il existe $y_0 \in \partial\phi(x_0)$ tel que ϕ^* est fermement sous-différentiable en y_0 . Enfin, il est montré qu'une fonction $\psi \in \Gamma_0(Y)$ est uniformément fermement sous-différentiable en y_0 si et seulement si sa conjuguée ψ^* est fermement différentiable en tout $x_0 \in \partial\psi(y_0)$.

REFERENCES

1. E. Asplund, *Fréchet differentiability of convex functions*, Acta Math. **121** (1968), 31–47.
2. D. Azé and J.-P. Penot, *Uniformly convex and uniformly smooth convex functions*, Ann. Fac. Sci. Toulouse Math. (6) **4** (1995), 705–730.
3. D. Azé and A. Rahmouni, *Lipschitz behaviour of the Legendre-Fenchel transform. Set convergence in nonlinear analysis and optimization*, Set-Valued Anal. **2** (1994), 35–48.
4. D. Azé and A. Rahmouni, *Intrinsic bounds for Kuhn-Tucker points of perturbed convex programs*, Recent developments in optimization (Dijon, 1994), Lecture Notes in Econom. and Math. Systems, vol. 429, Springer-Verlag, Berlin, 1995, pp. 17–35.
5. E. Bednarczyk and J.-P. Penot, *Metrically well-set minimization problems*, Appl. Math. Optim. **26** (1992), 273–285.
6. J.-P. Crouzeix, *A relationship between the second derivative of a convex function and of its conjugate*, Math. Programming **13** (1977), 364–365.
7. A.L. Donchev and T. Zollezzi, *Well-posed optimization problems*, Lecture Notes in Math., vol. 1543, Springer-Verlag, Berlin, 1991.
8. J.-B. Hiriart-Urruty, *Lipschitz r -continuity of the approximate subdifferential of a convex function*, Math. Scand. **47** (1980), 123–134.
9. J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex analysis and minimization algorithms*, Springer-Verlag, Berlin, 1993.
10. B. Lemaire, *Bons comportements et conditionnement linéaire*, Preprint, University of Montpellier, 1991.
11. B. Lemaire, *Bonne position, conditionnement, et bon comportement en optimisation*, Preprint, University of Montpellier, 1991.
12. C. Lemaréchal and C. Sagastizábal, *More than first order developments of convex functions: primal-dual relations*, J. Convex Analysis **3** (1996), 255–268.
13. C. Lemaréchal and C. Sagastizábal, *Practical aspect of the Moreau-Yosida approximation: theoretical preliminaries*, SIAM J. Optim. **7** (1997), 367–385.
14. Y. Lucet, *Formule explicite du Hessien de la régularisée de Moreau-Yosida d'une fonction convexe f en fonction de l'épi-différentielle seconde de f* , Preprint, University of Toulouse, 1995.
15. F. Mignot, *Contrôle dans les inéquations variationnelles elliptiques*, J. Functional Analysis **22** (1976), 130–185.
16. J.-P. Pennot, *Subhessians, superhessians and conjugation*, Nonlinear Anal. **23** (1994), 689–702.
17. J.-P. Pennot, *Conditioning convex and nonconvex problems*, J. Optim. Theory Appl. **90** (1996), 535–554.
18. R. Poliquin and R.T. Rockafellar, *Generalized Hessian properties of regularized nonsmooth functions*, SIAM J. Optim. **6** (1996), 1121–1137.
19. L. Qi, *Second-order analysis of the Moreau-Yosida regularization*, Nonlinear analysis and convex analysis (Niigata, 1998), World Scientific, River Edge, NJ, 1999, pp. 16–25.
20. A. Rahmouni, *Étude de l'effet d'une perturbation variationnelle sur le comportement primal-dual en optimisation convexe*, Thèse de doctorat, Université de Perpignan.
21. R. Zhang and J. Treiman, *Upper-Lipschitz multifunctions and inverse subdifferentials*, Nonlinear Anal. **24** (1995), 273–286.

D. AZÉ

UNIVERSITÉ DE PERPIGNAN, MATHÉMATIQUES,
52, AV. DE VILLENEUVE,
66860 PERPIGNAN CEDEX
FRANCE