Approximation and regularization
of arbitrary functions in Hilbert spaces
by the Lasry-Lions method

by

H. ATTOUCH

Université de Montpellier 2, Mathématiques,
place E. Bataillon, 34095 Montpellier Cedex 5, France

and

D. AZE

Université de Perpignan, Mathématiques,
52, Avenue de Villeneuve, 66860 Perpignan Cedex, France

ABSTRACT. — The Lasry-Lions regularization method is extended to arbitrary functions. In particular, to any proper lower semicontinuous function $f: X \to \mathbb{R} \cup \{+\infty\}$ defined on a Hilbert space X and which is quadratically minorized (*i. e.* $f(x) \geq -c(1 + \|x\|^2)$ for some $c \geq 0$), is associated a sequence of differentiable functions with Lipschitz continuous derivatives which approximate f from below. Some variants of the method are considered including the case of non quadratic kernels.

Key words : Hilbert spaces, regularization, approximation, epigraphical sum, Moreau-Yosida approximation.

RÉSUMÉ. — La méthode de régularisation de Lasry-Lions est étendue à des fonctions quelconques. En particulier, à toute fonction $f: X \to \mathbb{R} \cup \{+\infty\}$ semicontinue inférieurement propre sur un espace de Hilbert X et qui est quadratiquement minorée ($f(x) \geq -c(1 + \|x\|^2)$) pour un $c \geq 0$) est associée une suite de fonctions différentiables à dérivées

lipschitziennes qui approche f par valeurs inférieures. Quelques variantes de la méthode sont considérées, notamment le cas de noyaux non quadratiques.

---

1. PRELIMINARIES

Approximation methods play an important role in nonlinear analysis. A number of problems in variational analysis and in optimization theory give rise to nonsmooth functions with possibly infinite values defined on finite or infinite dimensional spaces. By using a regularization procedure based on the infimal convolution or epigraphical sum (*see* [4]), these problems can be attacked with the help of classical analysis tools. Let us mention in this direction the pioneering works of K. Yosida [25], H. Brézis [9], J.-J. Moreau [17]. These authors deal with convex lower semicontinuous functions in Hilbert spaces and with the corresponding subdifferential operators which are maximal monotone. The regularized function is proved to be $C^{1,1}$ (continuously differentiable with Lipschitz continuous gradient). Some direct extensions have been recently obtained in [4] for more general kernels than the square of the norm. A difficult problem is *to extend these results to the non convex case*. A decisive step in this direction has been done recently by J.-M. Lasry and P.-L. Lions in [16]. They were motivated by the study of the Hamilton-Jacobi equations and worked with bounded uniformly continuous functions. In [8] Theorem 2.6, boundedness and uniform continuity assumptions are removed: the approximation/regularization result is obtained assuming that the absolute value of the function is quadratically majorized. Our main results (Theorem 4.1 and Proposition 4.2) state that, given any quadratically minorized function f defined on a Hilbert space X with values in $\mathbb{R} \cup \{ +\infty \}$, the function $(f_\lambda)^\mu$ defined by the formula $(f_\lambda)^\mu (x) = \sup_{y \in X} \inf_{u \in X} (f(u) + (2\lambda)^{-1} \| u - y \|^2 - (2\mu)^{-1} \| y - x \|^2)$ is $C^{1,1}$ whenever $0 < \mu < \lambda$ and approaches f from below as the parameters λ and μ go to 0. Observe that our growth assumption on f allows to treat the case of an indicator function. Clearly, by exchanging the order of the inf-sup operations, one obtains a corresponding approximation from above. The paper is organized with respect to increasing generality: in sections 2, 3, 4 are successively considered the convex, then the convex up to a square case, and finally the general case. A natural question that arises concerns

the flexibility of the method: in section 5 is considered the case of non quadratic kernels. These results open new perspectives in nonsmooth analysis and ask for further developments: one may think defining generalized derivatives by relying on these approximation techniques. Regularization of sets can be considered too by using their indicator functions.

Let $(X, \|.\|)$ be a normed linear space, whose dual is denoted by $(X^*, \|.\|_*)$. To any extended real valued function $f : X \to \bar{\mathbb{R}}$ we can associate

$$\mathrm{epi}\, f = \{ (x, t) \in X \times \mathbb{R};\ t \geqq f(x) \}$$

the epigraph of f, and

$$\mathrm{epi}_s\, f = \{ (x, t) \in X \times \mathbb{R};\ t > f(x) \}$$

the strict epigraph of f.

Probably starting with the work of Fenchel (on convex functions), it has become more and more obvious that most properties of minimization problems can be naturally expressed with the help of *epigraphs*: convexity of f, lower semicontinuity of f are respectively equivalent to convexity, closure property of epi f. Since functions may be seen as sets, it is natural to combine them with the help of set operations. The vectorial sum of sets (also called Minkowski sum) when applied to epigraphs gives rise to the so called epigraphical sum (*see* Attouch and Wets [4]). Given f, $g : X \to \mathbb{R} \cup \{ +\infty \}$ two extended real-valued functions, the epigraphical sum (also called inf-convolution) $f \underset{e}{+} g$ is given by the relation:

$$(f \underset{e}{+} g)(x) = \inf_{u \in X} \{ f(u) + g(x - u) \}.$$

In geometrical terms

$$\mathrm{epi}_s\, (f \underset{e}{+} g) = \mathrm{epi}_s\, f + \mathrm{epi}_s\, g.$$

It plays a basic role in optimization and in the study of variational problems mainly because of the relation

$$(f \underset{e}{+} g)^* = f^* + g^*$$

where for each $y \in X^*$, $f^*(y) = \sup_{x \in X} \{ \langle y, x \rangle - f(x) \}$ is the Legendre-Fenchel transform or conjugate of f. Indeed, historically it has been introduced as the dual operation of the classical sum of (convex) functions. Another important feature of this operation is its *regularization effect*. Given $k : X \to \bar{\mathbb{R}}_+$ a "smoothing kernel" the epigraphical regularization of f is defined by

$$f_k = f \underset{e}{+} k.$$

The case $k_\lambda = \lambda^{-1} \|\,.\,\|$ gives rise to the approximation of lower semiconti-nuous functions by lipschitz continuous functions. This approximation procedure very likely goes back to R. Baire and F. Hausdorff. It has been then considered with increasing generality by E. J. McShane and H. Whitney, for a complete description of the above considered regulariza-tion and extension procedure one may consult [14] and [15]. The case $k_\lambda = 1/2 \lambda \|\,.\,\|^2$ leads to the Moreau-Yosida epigraphical regularization of f (see [2], [9], [17])

$$f_\lambda(x) = \inf_{u \in X} \left\{ f(u) + 1/2\lambda \|x - u\|^2 \right\}.$$

Let us recall in the following proposition (see [2], Theorem 2.64) some of the main properties of this approximation in such a general setting. The particular important case of convex functions will be considered in the next section.

PROPOSITION 1.1. — Let $f: X \to \mathbb{R} \cup \{+\infty\}$ be an extended real-valued function defined on a normed linear space X. Assume that for all $x \in X$, $f(x) \geqq -c/2\,(1 + \|x\|^2)$ where c is some non negative constant. Then provided $0 < \lambda < 1/c$, f_λ is a finitely valued function which is Lipschitz continuous on each bounded subset of X. Moreover $f_\lambda \leqq f$ and for all $x \in X$, $\sup_{\lambda > 0} f_\lambda(x) = \underline{cl}\,f(x)$ where $\underline{cl}\,f$ is the lower semicontinuous regularization of f.

Despite its global definition the Moreau-Yosida regularization operation has a local character as shown by

PROPOSITION 1.2. — Assume that f satisfies the growth assumption of Proposition 1.1.

a) Let $x \in X$ be such that $f(x)$ is finite. Then for each $0 < \lambda < 1/2\,c$ and for each $\rho > \bar{\rho}$,

$$f_\lambda(x) = \inf_{\|x - u\| \leqq \rho} \left\{ f(u) + (2\lambda)^{-1} \|x - u\|^2 \right\}$$

where $\bar{\rho}$ is given by:

$$\bar{\rho} = \bar{\rho}(x, \lambda, c) := \left[\lambda \frac{2f(x) + c\,(2\|x\|^2 + 1)}{1 - 2\lambda c} \right]^{1/2}$$

b) Assume that f, g satisfy the growth assumption of Proposition 1.1 and that $f = g$ in a neighborhood of some point $x \in X$ with $f(x) < +\infty$. Then for each λ small enough, $f_\lambda(x) = g_\lambda(x)$. Moreover assuming that f and g are majorized in a neighborhood of x, there exists a neighborhood of x on which $f_\lambda = g_\lambda$ for each λ small enough.

Proof. − *a*) For each $u \in X$, one has

$$f(u) + (2\lambda)^{-1} \| x - u \|^2 \geq -\frac{c}{2} (\| u \|^2 + 1) + (2\lambda)^{-1} \| x - u \|^2$$

$$\geq ((2\lambda)^{-1} - c) \| x - u \|^2 - c \| x \|^2 - \frac{c}{2}.$$

It ensues that given $\eta > 0$,

$$\inf_{\| x - u \| > r} \{ f(u) + (2\lambda)^{-1} \| x - u \|^2 \} \geq f(x) + \eta > f_\lambda(x),$$

where $r = \left[\lambda \dfrac{2f(x) + 2\eta + c(2\|x\|^2 + 1)}{1 - 2\lambda c} \right]^{1/2}$, thus

$$f_\lambda(x) = \inf_{\| x - u \| \leq r} \{ f(u) + (2\lambda)^{-1} \| x - u \|^2 \}.$$

Let $\rho > \bar{\rho}$ and let $\eta > 0$ be such that $\rho > r > \bar{\rho}$, it is clear that

$$f_\lambda(x) = \inf_{\| x - u \| \leq \rho} \{ f(u) + (2\lambda)^{-1} \| x - u \|^2 \}.$$

b) Assume that $f = g$ on $B(x, \delta)$, the closed ball with center x and radius $\delta > 0$. For each λ small enough one has $\bar{\rho}(x, \lambda, c) < \delta$. Let $\bar{\rho}(x, \lambda, c) < \rho < \delta$. Observing that f and g coincide on $B(x, \rho)$, it is clear from part *a*) that $f_\lambda(x) = g_\lambda(x)$.

Assume now that for some $\delta > 0$, $M \geq 0$, $f(x) = g(x) \leq M$ on $B(x, 2\delta)$. For λ small enough one has,

$$\left[\lambda \frac{2M + c(2\|x\|^2 + 1)}{1 - 2\lambda c} \right]^{1/2} < \delta.$$

Observe that for each $z \in B(x, \delta)$, $\bar{\rho}(z, \lambda, c) < \delta$ and that $f = g$ on $B(z, \delta)$. It ensues from part *a*) that $f_\lambda(z) = g_\lambda(z)$, which ends the proof of the proposition. ∎

As we already stressed, these epigraphical notions are well fitted to minimization problems, approximation of lower semicontinuous functions, convex duality... Since we have in mind to regularize and approximate arbitrary functions it is natural to consider their symmetric *hypographical* version. Given $f, g : X \to \mathbb{R} \cup \{ -\infty \}$, the hypographical sum of f and g is defined by

$$(f \overset{h}{+} g)(x) = \sup_{u \in X} \{ f(u) + g(x - u) \}.$$

The Moreau-Yosida hypographical approximate of index $\mu > 0$ of f is defined by

$$f^\mu(x) = \sup_{u \in X} \{ f(u) - (2\mu)^{-1} \| x - u \|^2 \}.$$

Noticing that

$$f^\mu(x) = -(-f)_\mu(x)$$

we can convert the proposition 1.1 into

PROPOSITION 1.3. — *Let $f: X \to \mathbb{R} \cup \{-\infty\}$ be an extended real-valued function defined on a normed linear space* X. *Assume that for all $x \in X$,*
$f(x) \leqq \dfrac{d}{2}(1 + \|x\|^2)$ *where d is some non negative constant. Then provided*
$0 < \mu < 1/d$, f^μ *is a finitely valued function which is Lipschitz continuous on bounded subsets of* X. *Moreover $f \leqq f^\mu$ and for all $x \in X$,* $\displaystyle\inf_{\mu > 0} f^\mu(x) = \overline{cl} f(x)$
where $\overline{cl} f$ is the upper semicontinuous regularization of f.

Clearly the hypographical regularization has also a local character: for each $0 < \mu < 1/2d$ and for each point x where $f(x) > -\infty$,

$$f^\mu(x) = \sup_{\|x-u\| \leqq \sigma} \left\{ f(u) - \frac{1}{2\mu} \|x - u\|^2 \right\}$$

for each $\sigma > \bar\sigma$ where

$$\bar\sigma = \bar\sigma(x, \mu, d) := \left[\mu \frac{d(2\|x\|^2 + 1) - 2g(x)}{1 - 2\mu d} \right]^{1/2}.$$

2. THE CONVEX CASE

Let us now assume that X is a Hilbert space whose norm $\|.\|$ is associated to a scalar product denoted by $\langle ., . \rangle$. Let us denote by $\Gamma_0(X)$ the convex cone of the convex lower semicontinuous proper ($\not\equiv +\infty$) functions from X into $\mathbb{R} \cup \{+\infty\}$. It is a classical result that the Legendre-Fenchel transform $f \mapsto f^*$ is a one to one correspondence from $\Gamma_0(X)$ onto itself. For any function f belonging to $\Gamma_0(X)$ its subdifferential is the multivalued operator $\partial f: X \to X^*$ whose graph is defined by

$$\partial f = \{ (x, y) \in X \times X^*; f(x) + f^*(y) - \langle x, y \rangle = 0 \}.$$

From this relation, we can observe that

$$\partial f^* = (\partial f)^{-1}.$$

For any $\lambda > 0$, and any $x \in X$ we denote by $J_\lambda^f x$ or briefly $J_\lambda x$ (when there is no ambiguity) the unique point of X where the function

$f(.) + \dfrac{1}{2\lambda} \| x - . \|^2$ achieves its minimum. Hence

$$f_\lambda(x) = f(J_\lambda x) + \dfrac{1}{2\lambda} \| x - J_\lambda x \|^2.$$

Observe that $J_\lambda x = (I + \lambda \partial f)^{-1} x$, *i.e.* J_λ is the resolvent of index $\lambda > 0$ of the maximal monotone operator $A = \partial f$. The following classical result due to J.-J. Moreau [18] summarizes the properties of the Moreau-Yosida approximation in the convex setting (*see* also [2], [9], [19], [26]). It is the starting point of our study.

THEOREM 2.1. − *Let* X *be a Hilbert space and let* $f: X \to \mathbb{R} \cup \{ +\infty \}$ *be a convex, lower semicontinuous proper function. Then for any* $\lambda > 0$ *the Moreau-Yosida approximate* f_λ *of* f *satisfies the following properties:*

a) f_λ *is a convex* $C^{1,1}$ *function (continuously differentiable with a Lipschitz continuous gradient). More precisely, for every* $x \in X$, $Df_\lambda(x) = A_\lambda(x)$ *where* A_λ *is the Yosida approximation of the maximal monotone operator* $A = \partial f$, $A_\lambda(x) := \lambda^{-1}(x - J_\lambda x)$. *Moreover, the operator* A_λ *is* λ^{-1}-*Lipschitz continuous.*

b) *As* λ *decreases to zero,* f_λ *increases to* f, *while* $Df_\lambda \to \partial f$ *in the graph sense* (i. e. *in the Kuratowski-Painlevé set convergence sense, see* [2] *for more details). Moreover for every* $x \in \text{dom } \partial f$, $A_\lambda x \to \partial f^0(x)$ *as* $\lambda \to 0$, *where* $\partial f^0(x)$ *is the element of minimal norm of the closed convex set* $\partial f(x)$.

Note that the set convergence of Df_λ to ∂f has been obtained in [2], Prop. 3.56. The above result has been extended in [4] to the case of more general kernels than $\| . \|^2$.

3. THE CONVEX UP TO A SQUARE CASE

Our purpose is to extend the class of functions which can be regularized into $C^{1,1}$ functions. One step in this direction has been done for the class of weakly convex or paraconvex functions (*see* [3], [6], [7], [8], [23] and [19] for the Yosida approximation of weakly monotone operators). Let us recall and complete these results.

DEFINITION 3.1. − A function $f: X \to \mathbb{R} \cup \{ +\infty \}$ is said to be weakly convex, or convex up to a square, or paraconvex if there exists some constant $c \geq 0$ such that $f(.) + \dfrac{c}{2} \| . \|^2$ is convex, that is

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) + t(1-t)\dfrac{c}{2} \| x - y \|^2$$

for all x, y belonging to X and all t in [0, 1].

A function $g : X \to \mathbb{R} \cup \{ -\infty \}$ is said to be weakly concave, or concave up to a square or paraconcave if $-g$ is weakly convex. This is equivalent to the existence of some constant $c \geq 0$ such that $g(.) - \frac{c}{2} \| . \|^2$ is concave, that is

$$g(tx + (1-t)y) \geq tg(x) + (1-t)g(y) - t(1-t)\frac{c}{2} \| x - y \|^2$$

for all x, y belonging to X and all t in $[0, 1]$.

Let us denote by $\Gamma_c(X)$ the set of functions f such that $f(.) + \frac{c}{2} \| . \|^2$ belongs to $\Gamma_0(X)$.

PROPOSITION 3.2. — *Let* $f : X \to \mathbb{R} \cup \{ +\infty \}$, $g : X \to \mathbb{R} \cup \{ -\infty \}$. *Assume that f and g satisfy the growth assumptions of Propositions* 1.1 *and* 1.3. *Then for any* $0 < \lambda < \frac{1}{c}$ *and* $0 < \mu < \frac{1}{d}$, $-f_\lambda \in \Gamma_{1/\lambda}(X)$ *and* $g^\mu \in \Gamma_{1/\mu}(X)$.

Proof. — It immediately follows from the definition of f_λ that

$$f_\lambda(x) = \frac{1}{2\lambda} \| x \|^2 - \sup_{u \in X} \left(\frac{1}{\lambda} \langle x, u \rangle - \left(f(u) + \frac{1}{2\lambda} \| u \|^2 \right) \right)$$

which can be reexpressed with the help of the Legendre-Fenchel transform as

$$f_\lambda(x) = \frac{1}{2\lambda} \| x \|^2 - \left(f(.) + \frac{1}{2\lambda} \| . \|^2 \right)^* \left(\frac{x}{\lambda} \right). \tag{9}$$

This means that f_λ is $\frac{1}{\lambda}$-weakly concave. It is also proper thanks to Proposition 1.1. The conclusion concerning g^μ is obtained in a similar way. ∎

When f is equal to Ψ_C, the indicator function of a set $C \subset X$, it follows from the preceding proposition that the function $d^2(., C)$ is weakly concave (a result due to E. Asplund).

Remark 3.3. — If $A : X \to X$ is a linear continuous symmetric operator, then the function $f(x) = \frac{1}{2} \langle A(x), x \rangle$ is weakly convex. Indeed, denoting by c the norm of A, we obtain $f(x) + \frac{c}{2} \| x \|^2 \geq 0$ and then the quadratic function $f(.) + \frac{c}{2} \| . \|^2$ is convex.

Given a weakly convex (resp. weakly concave) function f (resp. g), and given $x \in \text{dom} f$ (resp. $x \in \text{dom} g$), we denote by $\partial f(x)$ (resp. $\partial g(x)$) the set of lower (resp. upper) subgradients of f (resp. g) in the sense of R. T. Rockafellar (*see* [22]). If

$$f(.) + \frac{c}{2}\|.\|^2 = \varphi(.) \in \Gamma_0(X) \left[\text{resp.} f(.) - \frac{c}{2}\|.\|^2 = \psi, \ -\psi \in \Gamma_0(X) \right]$$

then $\partial f^2(x) = -cx + \partial \varphi(x)$ [resp. $\partial g(x) = cx^2 + \partial \psi(x)$], where $\partial \varphi(x)$ [resp. $\partial \psi(x)$] is the classical subdifferential (resp. upperdifferential) of convex analysis. Observe that the notation ∂f and ∂g are not ambiguous since a function which is both weakly convex and weakly concave is easily shown to be Fréchet differentiable. The following result, is a slight sharpening of [3] Proposition 3.3 (*see* also [6]).

THEOREM 3.4. — *Let $f \in \Gamma_c(X)$ be a weakly convex function. Then for any $0 < \lambda < c^{-1}$, f_λ belongs to the class $C^{1,1}$. Moreover, introducing $\varphi(.) = f(.) + \dfrac{c}{2}\|.\|^2$ (which is a convex function) the following relations hold:*

a) *for all $x \in X, f_\lambda(x) = \dfrac{\|x\|^2}{2\lambda} - (\varphi^*)_{(1/\lambda)-c}\left(\dfrac{x}{\lambda}\right)$, and f_λ is λ^{-1}-weakly concave.*

b) *For all $x \in X, f_\lambda(x) = -\dfrac{c}{1-\lambda c} \dfrac{\|x\|^2}{2} + \varphi_{\lambda/(1-\lambda c)}\left(\dfrac{x}{1-\lambda c}\right)$, and f_λ is $\dfrac{c}{1-c\lambda}$ - weakly convex.*

c) *If $\lambda + \mu < \dfrac{1}{c}, \ \forall x \in X, f_{\lambda+\mu}(x) = (f_\lambda)_\mu(x)$.*

d) *Let us denote by J_λ and A_λ the resolvent and Yosida approximates of $\partial \varphi$. For any $0 < \lambda < c^{-1}$, the function $f(.) + (2\lambda)^{-1}\|x - .\|^2$ attains its minimum at a unique point*

$$\tilde{J}_\lambda(x) := J_{\lambda/(1-\lambda c)}\left(\dfrac{x}{1-\lambda c}\right),$$

and, the operator $\tilde{J}_\lambda(.)$ is $\dfrac{1}{1-\lambda c}$-Lipschitz continuous. Moreover,

$$Df_\lambda(x) = \tilde{A}_\lambda(x) := \dfrac{x - \tilde{J}_\lambda(x)}{\lambda},$$

and the operator $\tilde{A}_\lambda(.)$ is $\max(\lambda^{-1}, (1-\lambda c)^{-1}c)$-Lipschitz continuous and satisfies $\tilde{A}_\lambda(x) \in \partial f(\tilde{J}_\lambda(x))$.

e) *As* λ *decreases to zero the following convergences hold:* f_λ *increases to* f *and* $\mathrm{D}f_\lambda \to \partial f$ *in the graph sense.*

Proof:

a)

$$
\begin{aligned}
f_\lambda(x) &= \inf_{u \in X} \left(\varphi(u) - \frac{c}{2} \|u\|^2 + \frac{1}{2\lambda} \|x\|^2 - \frac{1}{\lambda} \langle x, u \rangle + \frac{1}{2\lambda} \|u\|^2 \right) \\
&= \frac{\|x\|^2}{2\lambda} - \left(\varphi(.) + \left(\frac{1}{\lambda} - c \right) \frac{\|.\|^2}{2} \right)^* \left(\frac{x}{\lambda} \right) \\
&= \frac{\|x\|^2}{2\lambda} - \left(\varphi^* \underset{e}{+} \frac{1}{2(1/\lambda) - c} |.|^2 \right) \left(\frac{x}{\lambda} \right) \\
&= \frac{\|x\|^2}{2\lambda} - (\varphi^*)_{(1/\lambda) - c} \left(\frac{x}{\lambda} \right)
\end{aligned}
$$

which proves that f_λ is $C^{1, 1}$ since it is the case for the square of the norm and for the Moreau-Yosida approximate of a convex lower semicontinuous function.

b) Elementary computations yield

$$
\begin{aligned}
\varphi_{\lambda/(1-\lambda c)} \left(\frac{x}{1-\lambda c} \right) - \frac{c\|x\|^2}{2(1-\lambda c)} \quad &= \inf_{u \in X} \left(\varphi(u) + \frac{1-\lambda c}{2\lambda} \left\| \frac{x}{1-\lambda c} - u \right\|^2 \right) \\
&\quad - \frac{c\|x\|^2}{2(1-\lambda c)} \\
&= \inf_{u \in X} \left(\varphi(u) - \frac{c}{2} \|u\|^2 + \frac{1}{2\lambda} \|x-u\|^2 \right) \\
&= f_\lambda(x).
\end{aligned}
$$

c) From *b*), we can write

$$
f_\lambda(x) = - \frac{d}{2} \|x\|^2 + \psi(x)
$$

where

$$
d = \frac{c}{1-\lambda c} \qquad and \qquad \psi(x) = \varphi_{\lambda/(1-\lambda c)} \left(\frac{x}{1-\lambda c} \right).
$$

As $\mu < \dfrac{1}{d}$, we obtain, using *b*),

$$
(f_\lambda)_\mu(x) = - \frac{d\|x\|^2}{2(1-\mu d)} + \psi_{\mu/(1-\mu d)} \left(\frac{x}{1-\mu d} \right).
$$

Observing that $\dfrac{d}{1-\mu d} = \dfrac{c}{1-(\lambda+\mu)c}$ and that

$$\psi_{\mu/(1-\mu d)}\left(\frac{x}{1-\mu d}\right) = \inf_{u \in X}\left(\varphi_{\lambda/(1-\lambda c)}\left(\frac{u}{1-\lambda c}\right) + \frac{1-\mu d}{2\mu}\left\|\frac{x}{1-\mu d} - u\right\|^2\right)$$

$$= \varphi_{\lambda/(1-\lambda c) + (\mu/(1-\mu d))(1-\lambda c)}\left(\frac{x}{(1-\lambda c)(1-\mu d)}\right)$$

$$= \varphi_{(\lambda+\mu)/(1-(\lambda+\mu)c)}\left(\frac{x}{1-(\lambda+\mu)c}\right),$$

we obtain the announced result.

d) The first part of *d*) is an immediate consequence of *b*) and of the fact that the operator $J_\gamma(.)$ is a contraction (*see* [9] for example). From *b*), we obtain that

$$Df_\lambda(x) = -\frac{cx}{1-\lambda c} + \frac{1}{1-\lambda c}A_{\lambda/(1-\lambda c)}\left(\frac{x}{1-\lambda c}\right)$$

$$= \frac{x - \tilde{J}_\lambda(x)}{\lambda}.$$

The Lipschitz constant of Df_λ is obtained from [16], p. 265, observing that

$$f_\lambda(.) + \frac{d\|.\|^2}{2} \qquad \text{and} \qquad \frac{d\|.\|^2}{2} - f_\lambda(.)$$

are convex with $d = \max\left(\dfrac{1}{\lambda}, \dfrac{c}{1-\lambda c}\right)$. The fact that $\tilde{A}_\lambda(x) \in \partial f(\tilde{J}_\lambda(x))$ follows from a straightforward computation. Indeed

$$A_{\lambda/(1-\lambda c)}\left(\frac{x}{1-\lambda c}\right) \in \partial \varphi\left(J_{\lambda/(1-\lambda c)}\left(\frac{x}{1-\lambda c}\right)\right)$$
$$cx + (1-\lambda c)\tilde{A}_\lambda(x) \in \partial \varphi(\tilde{J}_\lambda(x))$$
$$\tilde{A}_\lambda(x) \in -c\tilde{J}_\lambda(x) + \partial \varphi(\tilde{J}_\lambda(x)) = \partial f(\tilde{J}_\lambda(x)).$$

e) The first part of *e*) is well known and can be recovered from *b*) since $-\dfrac{c}{1-\lambda c}\dfrac{\|x\|^2}{2}$ goes to $-c\dfrac{\|x\|^2}{2}$ and $\varphi_{\lambda/(1-\lambda c)}\left(\dfrac{x}{1-\lambda c}\right)$ goes to $\varphi(x)$ as λ decreases to 0. For the second part, it suffices to observe that

$$Df_\lambda(x) = \frac{1}{1-\lambda c}\left(-cx + A_{\lambda/(1-\lambda c)}\left(\frac{x}{1-\lambda c}\right)\right)$$

and that the operator $A_{\lambda/(1-\lambda c)}\left(\dfrac{\cdot}{1-\lambda c}\right)$ graph converges to $\partial\varphi$ as λ goes to 0 thanks to Theorem 2.1. ∎

Remark 3.5. — If $\lambda < \dfrac{1}{2c}$, then $\max(\lambda^{-1}, (1-\lambda c)^{-1}) = \lambda^{-1}$. Thus it

follows from Proposition 3.4, *d*) that $\tilde{A}_\lambda$ is $\dfrac{1}{\lambda}$-Lipschitzian (*see* also [19],

p. 376-377). Moreover one can easily prove that

$$\lim_{\lambda \to 0} \tilde{J}_\lambda(x) = \mathrm{Proj}_{\overline{\mathrm{dom}\,f}}(x),$$

and

$$\lim_{\lambda \to 0} \tilde{A}_\lambda(x) = \mathrm{Proj}_{\partial f(x)}(0).$$

PROPOSITION 3.6. — *The functions f and f_λ have same critical points and critical values.*

Proof. — Observe that

$$Df_\lambda(x) = 0 \Leftrightarrow \frac{x}{\lambda} \in \frac{1}{\lambda} \partial\left((\varphi^*)_{(1/\lambda)-c}\right)\left(\frac{x}{\lambda}\right)$$

$$\Leftrightarrow \frac{x}{\lambda} \in \partial\left((\varphi^*)_{(1/\lambda)-c}\right)^*(x)$$

$$\Leftrightarrow \frac{x}{\lambda} \in \partial\left(\varphi(.) + \left(\frac{1}{\lambda} - c\right)\frac{\|x\|^2}{2}\right)(x)$$

$$\Leftrightarrow 0 \in \partial f(x),$$

moreover, $0 \in \partial f(x)$ is equivalent to

$$\varphi(x) + \left(\frac{1}{\lambda} - c\right)\frac{\|x\|^2}{2} + \left(\varphi(.) + \left(\frac{1}{\lambda} - c\right)\frac{\|.\|^2}{2}\right)^*\left(\frac{x}{\lambda}\right) = \frac{\|x\|^2}{\lambda}$$

$$\varphi(x) - c\frac{\|x\|^2}{2} = \frac{\|x\|^2}{2\lambda} - \left(\varphi(.) + \left(\frac{1}{\lambda} - c\right)\frac{\|.\|^2}{2}\right)^*\left(\frac{x}{\lambda}\right)$$

$$f(x) = f_\lambda(x). \quad \blacksquare$$

4. THE GENERAL CASE

In [16], J.-M. Lasry and P.-L. Lions introduced a method providing $C^{1,1}$ regularization of bounded uniformly continuous functions defined on a Hilbert space. In this section, we extend the class of regularizable functions, allowing infinite values for the functions without requiring any regularity assumptions.

Let $f: X \to \mathbb{R} \cup \{ +\infty \}$ and $g: X \to \mathbb{R} \cup \{ -\infty \}$ be real extended valued functions and $\lambda, \mu > 0$. Following [16], we introduce

$$(f_\lambda)^\mu (x) = \sup_{y \in X} \ \inf_{u \in X} \left(f(u) + \frac{1}{2\lambda} \| u - y \|^2 - \frac{1}{2\mu} \| y - x \|^2 \right),$$

$$(g^\lambda)_\mu (x) = \inf_{y \in X} \ \sup_{u \in X} \left(g(u) - \frac{1}{2\lambda} \| u - y \|^2 + \frac{1}{2\mu} \| y - x \|^2 \right).$$

THEOREM 4.1. − Regularization: *Assume there exists $c, d \geqq 0$ such that, for every $x \in X$*

$$f(x) \geqq -\frac{c}{2}(\| x \|^2 + 1), \qquad g(x) \leqq \frac{d}{2}(\| x \|^2 + 1).$$

Then, for all $0 < \mu < \lambda < \dfrac{1}{c}$ (resp. $0 < \mu < \lambda < \dfrac{1}{d}$), $(f_\lambda)^\mu$ (resp. $(g^\lambda)_\mu$) is a $C^{1,1}$ function whose gradient is $\max\left(\dfrac{1}{\mu}, \dfrac{1}{\lambda - \mu}\right)$-Lipschitz continuous. One has $(f_\lambda)^\mu \in \Gamma_{1/\mu}(X)$ and $-(f_\lambda)^\mu \in \Gamma_{1/(\lambda - \mu)}(X)$. Moreover, $(f_\lambda)^\mu \leqq f$ and $(g^\lambda)_\mu \geqq g$.

Proof. − Relying on Proposition 3.2, $-(f_\lambda)$ is $\dfrac{1}{\lambda}$-weakly convex and finitely valued, thus for $\mu < \lambda$, we obtain from Theorem 3.4 that the function $(f_\lambda)^\mu = -(-(f_\lambda))_\mu$ is $\dfrac{1}{\mu}$-weakly convex, is $\dfrac{1}{\lambda - \mu}$-concave and is $C^{1,1}$. Using again Theorem 3.4, we derive that the gradient $D(f_\lambda)^\mu$ is Lipschitz continuous with constant $\max\left(\dfrac{1}{\mu}, \dfrac{1}{\lambda - \mu}\right)$. As $(g^\lambda)_\mu = -((-g)_\lambda)^\mu$, the proof of the first part is complete. Choosing $u = x$ in the definition of $(f_\lambda)^\mu$ and $(g^\lambda)_\mu$ and taking into account that $\mu < \lambda$, we obtain $(f_\lambda)^\mu \leqq f$ and $(g^\mu)_\lambda \geqq g$. ■

THEOREM 4.2. − Approximation: *Assume that f, g verify the growth conditions of theorem 4.1. Then,*

$$\lim_{\lambda \to 0, \, \mu \to 0, \, \lambda > \mu} (f_\lambda)^\mu (x) = \underline{cl} \, f(x),$$
$$\lim_{\lambda \to 0, \, \mu \to 0, \, \lambda > \mu} (g^\lambda)_\mu (x) = \overline{cl} \, g(x).$$

where $\underline{cl}$ and $\overline{cl}$ denote respectively the l.s.c. and the u.s.c. regularization operation.

Proof. − We prove the first equality. The other can be obtained in a similar way. Observe that $f_\lambda = (\underline{cl} \, f)_\lambda$. Since the inequality $(f_\lambda)^\mu \leqq f$ is valid for any function f, we derive from the above inequalities that $(f_\lambda)^\mu \leqq \underline{cl} \, f$

for all $0 < \mu < \lambda < \dfrac{1}{c}$. Hence we obtain $\displaystyle\limsup_{\lambda \to 0, \, \mu \to 0, \, \mu < \lambda} (f_\lambda)^\mu (x) \leqq \underline{\mathrm{cl}}\, f(x)$ for each $x \in X$. On the other hand, $(f_\lambda)^\mu \geqq f_\lambda$, hence

$$\liminf_{\lambda \to 0, \, \mu \to 0, \, \mu < \lambda} (f_\lambda)^\mu (x) \geqq \liminf_{\lambda \to 0} f_\lambda (x) = \underline{\mathrm{cl}}\, f(x)$$

for each $x \in X$, hence the result. ■

Remark 4.3. — *a*) The preceding approximation result can be reinforced in the following way. Let $x \in X$ and let $(x_{\lambda, \, \mu})$ converging to x as λ and μ go to 0 with $0 < \mu < \lambda < \dfrac{1}{c}$. We observe that

$$\liminf_{\lambda \to 0, \, \mu \to 0, \, \mu < \lambda} (f_\lambda)^\mu (x_{\lambda, \, \mu}) \geqq \liminf_{\lambda \to 0, \, \mu \to 0, \, \mu < \lambda} f_\lambda (x_{\lambda, \, \mu}).$$

Since (f_λ) increases as λ decreases to 0, for any $\lambda_0 > 0$,

$$\liminf_{\lambda \to 0, \, \mu \to 0, \, \mu < \lambda} (f_\lambda)^\mu (x_{\lambda, \, \mu}) \geqq \liminf_{\lambda \to 0, \, \mu \to 0, \, \mu < \lambda} f_{\lambda_0} (x_{\lambda, \, \mu}).$$

As f_{λ_0} is continuous (and hence lower semicontinuous), it ensues

$$\liminf_{\lambda \to 0, \, \mu \to 0, \, \mu < \lambda} (f_\lambda)^\mu (x_{\lambda, \, \mu}) \geqq f_{\lambda_0} (x).$$

This being true for any λ_0, taking the supremum with respect to λ_0 yields

$$\liminf_{\lambda \to 0, \, \mu \to 0, \, \mu < \lambda} (f_\lambda)^\mu (x_{\lambda, \, \mu}) \geqq \underline{\mathrm{cl}}\, f(x).$$

It follows that $(f_\lambda)^\mu$ converges to $\underline{\mathrm{cl}} f$ in the sense of epi convergence (*see* [12]). This fact is important since it provides a path for a connection with non smooth analysis via extensions of the Attouch's Theorem relating epi convergence of functions and convergence of the graphs of their derivatives.

b) From Theorem 3.4*c*), we deduce that $(f_\lambda)^{\mu + \nu} = ((f_\lambda)^\mu)^\nu$ and that $(g^\lambda)_{\mu + \nu} = ((g^\lambda)_\mu)_\nu$.

c) From Theorem 4.1, we derive that, if $\mu \leqq \dfrac{\lambda}{2}$, $D(f_\lambda)^\mu$ is $\dfrac{1}{\mu}$ — Lipschitz continuous.

d) If $f = \psi_s$ the indicator function of a subset S of X, we obtain $0 \leqq \dfrac{1}{2\lambda} d^2 (x, \mathrm{S}) \leqq (f_\lambda)^\mu (x) \leqq f(x)$. It follows that $x \in \mathrm{S}$ if and only if $(f_\lambda)^\mu (x) = 0$.

e) Observe that, thanks to Theorem 3.4, the supremum is attained in the sup convolution defining $(f_\lambda)^\mu (x)$ at a unique point $\mathrm{J}_{\lambda, \, \mu} (x)$ characterized by

$$\frac{\mathrm{J}_{\lambda, \, \mu} (x) - x}{\mu} \in \partial f_\lambda (\mathrm{J}_{\lambda, \, \mu} (x)).$$

Moreover $D(f_\lambda)^\mu(x) = \dfrac{J_{\lambda,\mu}(x) - x}{\mu}$.

Example 4.4. − *a*) If f or $-g$ belongs to $\Gamma_0(X)$, then $(f_\lambda)^\mu = f_{\lambda-\mu}$ and $(g^\lambda)_\mu = g^{\lambda-\mu}$. Indeed,

$$-f_\lambda(x) = \varphi(x) - \frac{\|x\|^2}{2\lambda}$$

where

$$\varphi(x) = \left(f(.) + \frac{\|.\|^2}{2\lambda} \right)^* \left(\frac{x}{\lambda} \right),$$

thus, using Theorem 3.4 *a*)

$$
\begin{aligned}
(-f_\lambda)_\mu(x) &= \frac{\|x\|^2}{2\mu} - (\varphi^*)_{(1/\mu)-(1/\lambda)}\left(\frac{x}{\mu} \right) \\
&= \frac{\|x\|^2}{2\mu} - \inf_{u \in X}\left(f(\lambda u) + \frac{\lambda}{2}\|u\|^2 + \frac{\mu\lambda}{2(\lambda-\mu)}\left\| \frac{x}{\mu} - u \right\|^2 \right) \\
&= -f_{\lambda-\mu}(x).
\end{aligned}
$$

b) Let us consider the function $f \colon \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = -|x| \quad \text{if } |x| \leq 1, \qquad f(x) = |x| - 2 \quad \text{if } |x| \geq 1.$$

An easy computation leads to

$$
f_\lambda(x) =
\begin{cases}
-|x| - \dfrac{\lambda}{2} & \text{if } |x| \leq 1 - \lambda \\[2mm]
-1 + \dfrac{(|x|-1)^2}{2\lambda} & \text{if } 1 - \lambda \leq |x| \leq 1 + \lambda, \\[2mm]
|x| - 2 - \dfrac{\lambda}{2} & \text{if } |x| \geq 1 + \lambda
\end{cases}
$$

and

$$
(f_\lambda)^\mu(x) =
\begin{cases}
-\dfrac{\lambda}{2} - \dfrac{|x|^2}{2\mu} & \text{if } |x| \leq \mu \\[2mm]
-|x| - \dfrac{\lambda-\mu}{2} & \text{if } u \leq |x| \leq 1 - (\lambda-\mu) \\[2mm]
-1 + \dfrac{(|x|-1)^2}{2(\lambda-\mu)} & \text{if } 1 - (\lambda-\mu) \leq |x| \leq 1 + (\lambda-\mu). \\[2mm]
|x| - 2 - \dfrac{\lambda-\mu}{2} & \text{if } |x| \geq 1 + (\lambda-\mu)
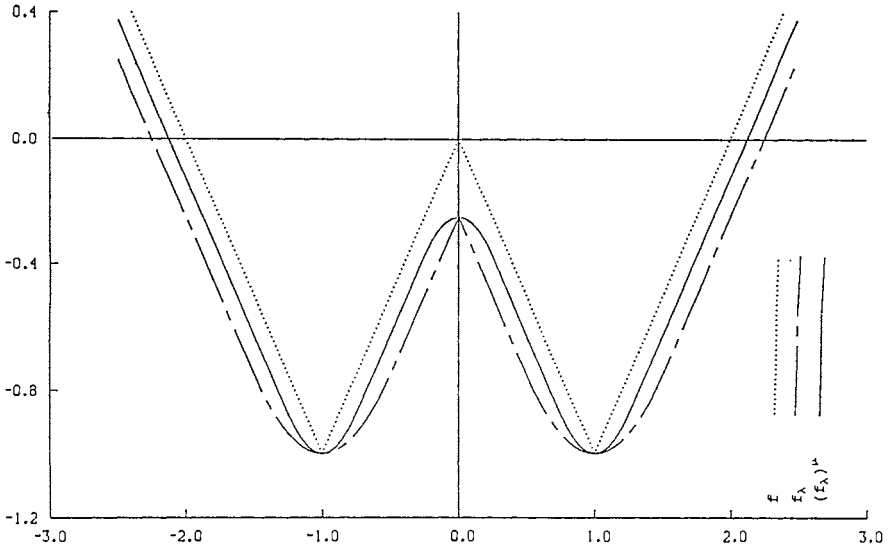\end{cases}
$$

Lambda = 1/2 and mu = 1/4

FIG. 1

In figure 1, are drawn the graphs of f (the doted curve), f_λ (the dashed curve) and $(f_\lambda)^\mu$ for the values $\lambda = \frac{1}{2}$ and $\mu = \frac{1}{4}$. Observe that the first regularization will smooth the lower corners while the second regularization will smooth the upper corners without introducing lower corners.

THEOREM 4.5. — Critical points: *Assume that* $f(.) \geqq -\frac{c}{2}(1 + \|.\|^2)$ *for some* $c > 0$. *Then for all* $0 < \mu < \lambda < \frac{1}{c}$

a) $D((f_\lambda)^\mu)(x) = 0$ *if and only if* $0 \in \partial f_\lambda(x)$ *(the upperdifferential of* f_λ) *and in this case* $f_\lambda(x) = (f_\lambda)^\mu(x)$.

b) *Assume that* f *is lower semicontinuous. Then*

$$\inf_x f = \inf_x (f_\lambda)^\mu,$$
$$\operatorname{argmin} f = \operatorname{argmin} (f_\lambda)^\mu.$$

c) *Assume that* $\bar{x} \in X$ *is a local minimum of* f *such that* f *is majorized near* $\bar{x}$. *Then if* $0 < \mu < \lambda$ *are small enough,* $\bar{x}$ *is a local minimum of* $(f_\lambda)^\mu$ *and* $(f_\lambda)^\mu(\bar{x}) = f(\bar{x})$.

Proof. $-$ *a*) Follows from Proposition 3.6 observing that $(f_\lambda)^\mu = -(-f_\lambda)_\mu$ and that $(-f_\lambda) \in \Gamma_{1/\lambda}$.

b) As $(f_\lambda)^\mu \leq f$, one derives $\inf_x (f_\lambda)^\mu \leq \inf f$. Moreover $(f_\lambda)^\mu \geq f_\lambda$, hence $\inf_x (f_\lambda)^\mu \geq \inf_x f_\lambda = \inf f$. Thus we obtain $\inf_x (f_\lambda)^\mu = \inf_x f$. As $(f_\lambda)^\mu \leq f$, it ensues that

$$\operatorname{argmin} f \subset \operatorname{armin} (f_\lambda)^\mu.$$

Conversely let $\bar{x} \in \operatorname{argmin} (f_\lambda)^\mu$, we derive that

$$\inf_x (f_\lambda)^\mu = (f_\lambda)^\mu (\bar{x}) \geq f_\lambda (\bar{x}) \geq \inf_x f_\lambda = \inf f$$

which combined with $\inf_x f = \inf_x (f_\lambda)^\mu$ yields $\bar{x} \in \operatorname{argmin} f_\lambda$. Thus from [3] Proposition 3.1 we obtain that $\bar{x} \in \operatorname{argmin} f$. ∎

c) There exists a ball B with center $\bar{x}$ and positive radius such that $f + \Psi_B$ attains its minimum over X on $\bar{x}$ and such that f is majorized on B. From Proposition 1.2., *b*), we derive that f_λ and $(f + \Psi_B)_\lambda$ coincide near $\bar{x}$ for all λ small enough. Observing that f_λ and $(f + \Psi_B)_\lambda$ are uniformly minorized by $f(\bar{x})$ near $\bar{x}$, we can apply again Proposition 1.2, *b*) yielding that $(f_\lambda)^\mu$ and $((f + \Psi_B)_\lambda)^\mu$ coincide near $\bar{x}$ for all $0 < \mu < \lambda$ small enough. From part *b*) of this theorem the point $\bar{x}$ minimizes $((f + \Psi_B)_\lambda)^\mu$ over X thus $\bar{x}$ is a local minimizer of $(f_\lambda)^\mu$. Moreover

$$(f_\lambda)^\mu (\bar{x}) = ((f + \Psi_B)_\lambda)^\mu (\bar{x}) = f(\bar{x}). \quad \blacksquare$$

5. EXTENSION TO NON QUADRATIC KERNELS

Let us begin by some definitions. We denote by P the set of even convex functions $\pi : \mathbb{R} \to \bar{\mathbb{R}}_+$ with $\pi(0) = 0$. Let us set

$$B = \{ \beta \in P : \beta(t) > 0 \quad \text{for } t \neq 0 \}$$

and

$$A = \left\{ \alpha \in P : \lim_{t \to 0} \frac{\alpha(t)}{t} = 0 \right\}.$$

It is known ([1], Lemma 1) that $\beta \in B$ if and only if $\beta^* \in A$. A function $f : X \to \mathbb{R} \cup \{ +\infty \}$ is said to be β-convex if there exists $\beta \in P$ such that, for each $x, z \in X$, $t \in [0, 1]$,

$$f(x_t) \leq t f(x) + (1 - t) f(z) - t(1 - t) \beta(\| x - z \|),$$

where $x_t = tx + (1-t)z$. A function $g : X \to \mathbb{R} \cup \{ +\infty \}$ is said to be α-smooth if there exists $\alpha \in P$ such that, for each x, $z \in X$ and for each $t \in [0, 1]$,

$$tg(x) + (1-t)g(z) \leqq g(x_t) + t(1-t)\alpha(\| x - z \|).$$

We say that a β-convex function f (resp. a α-smooth function g) is β-uniformly convex (resp. α-uniformly smooth) if $\beta \in B$ (resp. $\alpha \in A$). For further details about theses classes of functions, the reader may consult ([5], [24], [27]).

The following lemma whose proof is a direct consequence of the definitions, is taken from [5] (Proposition 1.8 and Corollary 1.10).

LEMMA 5.1. − Let $f \in \Gamma_0(X)$, then f is β-convex (resp. β-uniformly convex) if and only if f^* is β^*-smooth (resp. β^*-uniformly smooth). ■

Given a function $f : X \to \mathbb{R} \cup \{ +\infty \}$ and x, $u \in X$, we set when the limit exists

$$f'(x; u) = \lim_{t \downarrow 0} \frac{f(x + tu) - f(x)}{t}.$$

For a locally lipschitz function f, we define

$$f^0(x; u) = \limsup_{t \downarrow 0, \, z \to x} \frac{f(z + tu) - f(z)}{t}.$$

The function $f^0(x; u)$ is the Clarke directional derivative (see [10]). The subdifferential of f at x is the closed bounded convex subset $\partial f(x)$ whose support function is $f^0(x; \cdot)$.

LEMMA 5.2. − Let $f : X \to \mathbb{R}$ be a locally lipschitz function such that $g = -f$ is α-uniformly smooth for some $\alpha \in A$. Then,

$$f'(x; u) = f^0(x; u) \quad \text{for all} \quad u \in X.$$

Proof. − We first observe that,

$$\limsup_{t \downarrow 0} \frac{f(x + tu) - f(x)}{t} \leqq f^0(x; u).$$

Moreover, setting $y = z + u$, we obtain, for each $t \in [0, 1]$,

$$f(z + tu) \leqq tf(y) + (1-t)f(z) + t(1-t)\alpha(\| u \|),$$

and then,

$$\frac{f(z + tu) - f(z)}{t} \leqq f(z + u) - f(z) + (1-t)\alpha(\| u \|),$$

hence

$$f^0(x; u) \leqq f(x + u) - f(x) + \alpha(\| u \|).$$

Replacing u by tu with $t > 0$, it ensues

$$f^0(x; u) \leqq \frac{f(x+tu)-f(x)}{t} + \frac{\alpha(t\|u\|)}{t},$$

hence

$$f^0(x; u) \leqq \liminf_{t \downarrow 0} \frac{f(x+tu)-f(x)}{t}. \quad \blacksquare$$

As a consequence, we get,

LEMMA 5.3. − *Let* $f : X \to \mathbb{R}$ *be a locally lipschitz function such that f and $-f$ are uniformly smooth. Then f is Fréchet differentiable and* $Df(x) = f^0(x; .)$.

Proof. − Let $x, u \in X$. Assume that f (resp. $-f$) is α_1-uniformly smooth (resp. α_2-uniformly smooth). From the proof of Lemma 5.2, it ensues, $f^0(x; u) \leqq f(x+u) - f(x) + \alpha_2(\|u\|)$. Thus we obtain

$$(-f)^0(x; u) \leqq -f(x+u) + f(x) + \alpha_1(\|u\|).$$

From Lemma 5.2, we get

$$(-f)^0(x; u) = (-f)'(x; u) = -f'(x; u) = -f^0(x; u).$$

It ensues that $f^0(x; .)$ is linear continuous and that,

$$-\alpha_2(\|u\|) \leqq f(x+u) - f(x) - f^0(x; u) \leqq \alpha_1(\|u\|),$$

which ends the proof of the Lemma 5.3. $\quad \blacksquare$

In the sequel, we shall work with smoothing kernels,

$$\varphi, \psi : \mathbb{R}_+ \to \mathbb{R}_+ \qquad \text{with} \quad \varphi(0) = \psi(0) = 0.$$

We make some assumptions (*see* [2], [7], [8]), ensuring that, given f, $g : X \to \mathbb{R} \cup \{+\infty\}$, the functions

$$f_\varphi = f + (\varphi \circ \|.\|) \qquad \text{and} \qquad g^\psi = g + (-\psi \circ \|.\|)$$

are finite and locally lipschitz. From now one, we shall assume that the functions $\tilde{\varphi} = \varphi \circ \|.\|$ and $\tilde{\psi} = \psi \circ \|.\|$ satisfy

a) $\tilde{\varphi}(x_n) \to 0$ implies $x_n \to 0$.

b) For each $x \in X$ and each $k > 1$, there exists $m \in \mathbb{R}_+$ such that, for each $v \in X$,

$$\tilde{\varphi}(x+v) \leqq k \tilde{\varphi}(v) + m.$$

c) There exist, $p, q, r \in \mathbb{R}_+$ such that, for each $u, v \in X$,

$$\tilde{\varphi}(u+v) \leqq p \tilde{\varphi}(u) + q \tilde{\varphi}(v) + r.$$

d) $\tilde{\varphi}$ is lipschitz on bounded subsets of X.

e) $B \subset X$ is bounded if $\tilde{\varphi}$ is bounded on B.

Given $f : X \to \mathbb{R} \cup \{ + \infty \}$, we also assume that there exist $c \in \mathbb{R}_+$ and $x_0 \in X$ such that $f + c\, \varphi\, (x_0 - .)$ is bounded from below. When assumptions $a), \ldots, e)$ hold, Proposition 3.5 of [7] ensures that, for each $\varepsilon < \min \left(\dfrac{1}{c}, \dfrac{1}{cp} \right)$, the function $f_\varphi = f \underset{e}{+} \dfrac{1}{\varepsilon} (\varphi \circ \| \, . \, \|)$ is finite and locally lipschitz.

We are going now to explore the uniform convexity and uniform smoothness of the above mentioned regularized functions.

PROPOSITION 5.4. $-$ *Let* $f : X \to \mathbb{R} \cup \{ + \infty \}$ *and let* $\varphi : X \to \mathbb{R}$ *such that* φ *is* α-*smooth for some* $\alpha \in P$. *Then* $f \underset{e}{+} \varphi$ *is* α-*smooth provided it is everywhere finite.*

Proof. $-$ Let x, $z \in X$ and $t \in [0, 1]$. By definition of the epigraphical sum, for each $u \in X$

$$t \,(f \underset{e}{+} \varphi)\, (x) \leqq tf(u) + t\, \varphi\, (x - u),$$

$$(1 - t)\,(f \underset{e}{+} \varphi)\, (z) \leqq (1 - t)\, f(u) + (1 - t)\, \varphi\, (z - u).$$

Thus,

$$t \,(f \underset{e}{+} \varphi)\, (x) + (1 - t)\,(f \underset{e}{+} \varphi)\, (z) \leqq f(u) + \varphi\,(x_t - u) + t\,(1 - t)\, \alpha\,(\| x - z \|),$$

where $x_t = tx + (1 - t)\, z$. The result follows by taking the infimum over $u \in X$ in the right hand-side of the above inequality. ∎

PROPOSITION 5.5. $-$ *Let* $g : X \to \mathbb{R} \cup \{ + \infty \}$ *be* α-*smooth and let* $\psi : X \to \mathbb{R}$ *be* β-*convex. Assume that* $\delta = \alpha \overset{h}{+} (-\beta)$ *belongs to* P *and that* $g \overset{h}{+} (-\psi)$ *is everywhere finite. Then,* $g \overset{h}{+} (-\psi)$ *is* δ-*smooth.*

Proof. $-$ Let x, $z \in X$, $t \in [0, 1]$. Given $\varepsilon > 0$, there exists v, $w \in X$ such that

$$t \,(g \overset{h}{+} (-\psi))\, (x) \leqq t\,(g\,(v) - \psi\,(x - v)) + \varepsilon$$

and

$$(1 - t)\,(g \overset{h}{+} (-\psi))\, (z) \leqq (1 - t)\,(g\,(w) - \psi\,(z - w)) + \varepsilon.$$

It ensues,

$$
\begin{aligned}
t \,(g \overset{h}{+} (-\psi))\, (x) + (1 - t)\,(g \overset{h}{+} (-\psi))\, (z) & \\
\leqq g\,(tv + (1 - t)\, w) - \psi\,(x_t - (tv + (1 - t)\, w)) & \\
+ t\,(1 - t)\,(\alpha\,(\| v - w \|) - \beta\,(\| (x - z) - (v - w) \|)) + 2\,\varepsilon. &
\end{aligned}
$$

Observing that $\|(x-z)-(v-x)\| \geqq |\ \|x-z\| - \|v-w\|\ |$, that $\beta(.)$ is non decreasing on $\mathbb{R}_+$ and that $\alpha \overset{h}{+} (-\beta)$ is even, we obtain

$$\alpha(\|v-w\|) - \beta(\|(x-z)-(v-w)\|) \leqq (\alpha \overset{h}{+} (-\beta))(\|x-z\|).$$

Thus, letting ε go to 0,

$$t(g \overset{h}{+} (-\psi))(x) + (1-t)(g \overset{h}{+} (-\psi))(z)$$

$$\leqq (g \overset{h}{+} (-\psi))(x_t) + t(1-t)(\alpha \overset{h}{+} (-\beta))(\|x-z\|),$$

hence the result. ∎

Remark 5.6. − The preceding result is sharp when considering $g(.) = \dfrac{\|.\|^2}{2\lambda}$ and $\psi(.) = \dfrac{\|.\|^2}{2\mu}$ where $\mu < \lambda$. The functions g and ψ are respectively α-smooth and β-convex with $\alpha(t) = \dfrac{t^2}{2\lambda}$ and $\beta(t) = \dfrac{t^2}{2\mu}$. An elementary computation leads to $g \overset{h}{+} (-\psi) = \dfrac{\|.\|^2}{2(\lambda-\mu)}$ and this function is exactly $(\alpha \overset{h}{+} (-\beta))$-smooth.

The main result of this section is,

THEOREM 5.7. − *Let* $f: X \to \mathbb{R} \cup \{+\infty\}$, $\varphi: \mathbb{R}_+ \to \mathbb{R}_+$, $\psi: \mathbb{R}_+ \to \mathbb{R}_+$ *be such that* $f \underset{e}{+} (\varphi \circ \|.\|)$ *and* $(f \underset{e}{+} (\varphi \circ \|.\|)) \overset{h}{+} (-\psi \circ \|.\|)$ *are everywhere finite and locally lipschitz. Assume that* $\varphi \circ \|.\|$ *is* α-*uniformly smooth that* $\psi \circ \|.\|$ *is* β-*strongly convex and uniformly smooth, and that* $\alpha \overset{h}{+} (-\beta) = \delta$ *with* $\delta \in A$. *Then*

$$(f \underset{e}{+} (\varphi \circ \|.\|)) \overset{h}{+} (-\psi \circ \|.\|)$$

is Fréchet differentiable on X. *Moreover, this function is* C^1 *if* X *is finite dimensional.*

Proof. − From Proposition 5.4, $f \underset{e}{+} (\varphi \circ \|.\|)$ is α-smooth. From Proposition 5.5, $(f \underset{e}{+} (\varphi \circ \|.\|)) \overset{h}{+} (-\psi \circ \|.\|)$ is δ-smooth and then δ-

uniformly smooth since $\delta \in A$. Applying again Proposition 5.4,

$$-((f \underset{e}{+} (\varphi \circ \|.\|)) \overset{h}{+} (-\psi \circ \|.\|)) = (-(f \underset{e}{+} (\varphi \circ \|.\|))) \underset{e}{+} (\psi \circ \|.\|)$$

is uniformly smooth. It follows from lemma 5.3 that

$$(f \underset{e}{+} (\varphi \circ \|.\|)) \overset{h}{+} (-\psi \circ \|.\|)$$

is Fréchet differentiable. When X is finite dimensional, we obtain from Lemma 5.3 that the Fréchet derivative of the function $(f \underset{e}{+} (\varphi \circ \|.\|)) \overset{h}{+} (-\psi \circ \|.\|)$ is equal to its Clarke derivative, thus it is continuous (see [10]). ■

Remark 5.8. — The condition $\alpha \overset{h}{+} (-\beta) = \delta \in A$ is equivalent to the condition $\alpha^* - \beta^* \in B$, observing that $(\alpha^* - \beta^*)^* = \alpha \overset{h}{+} (-\beta)$ thanks to the Hiriart-Urruty's formula (see [12]).

Example 5.9. — Let us give an example of non quadratic kernels which satisfy the assumptions of Theorem 5.7. From ([24], Theorem 3), we know that, given $\xi : \mathbb{R}_+ \to \mathbb{R}_+$ with $\xi(0) = 0$, and $\xi(cs) \geqq c \xi(s)$ for each $s \in \mathbb{R}_+$ and each $c \geqq 1$, the function $\psi(u) = \int_0^{\|u\|} \xi(s)\,ds$ is uniformly convex with modulus $\beta(t) = \int_0^t \xi\left(\dfrac{s}{2}\right) ds$. Let us introduce $\xi : \mathbb{R}_+ \to \mathbb{R}_+$ defined by $\xi(t) = e^t - 1$. The function $\psi_0(u) = e^{\|u\|} - \|u\| - 1$ is β_0-convex with $\beta_0(t) = 2e^{t/2} - t - 2$. Let us set $\psi = \psi_0 \underset{e}{+} \dfrac{\|.\|^2}{2}$, this function is uniformly smooth and β-uniformly convex with $\beta = \beta_0 \underset{e}{+} \dfrac{(.)^2}{2}$. Then the function $\varphi = \psi^*$ is α-smooth with $\alpha(t) = \beta^*(t) = 2(t+1)\ln(t+1) - 2t$. By using a formula on a conjugate of a difference of convex functions due to J.-B. Hiriart-Urruty [12], we obtain that

$$\alpha \overset{h}{+} (-\beta) = (\alpha^* - \beta^*)^* = (\beta - \alpha)^*.$$

An easy computation shows that $\beta - \alpha \in B$. Hence $(\beta - \alpha)^*$ belongs to A and Theorem 5.7 applies.

REFERENCES

[1] E. ASPLUND, Fréchet differentiability of convex functions, *Acta Math.*, Vol. **121**, 1968, pp. 31-47.

[2] H. ATTOUCH, *Variational Convergences for Functions and Operators*, Applicable Mathematics Series, Pitman London, 1984.

[3] H. ATTOUCH, D. AZÉ and R. J. B. WETS, On continuity properties of the partial Legendre-Fenchel transform: convergence of sequences of augmented Lagrangians functions, Moreau-Yosida approximates and subdifferential operators, *Fermat days 85: Mathematics for Optimization*, J.-B. HIRIART-URRUTY Ed., North-Holland Amsterdam, 1986, pp. 1-42.

[4] H. ATTOUCH and R. J. B. WETS, Epigraphical analysis, *Analyse non linéaire*, H. ATTOUCH, J.-P. AUBIN, F. H. CLARKE and I. EKELAND Eds., Gauthier-Villars, Paris, 1989, pp. 73-100.

[5] D. AZÉ and J.-P. PENOT, Uniformly convex and uniformly smooth convex functions (submitted).

[6] M. BOUGEARD, Contribution à la théorie de Morse en dimension finie, *Thèse de 3e cycle*, Université de Paris-IX, 1978.

[7] J.-P. PENOT and M. BOUGEARD, Approximation and decomposition properties of some classes of locally d.c. functions, *Math. Progr.*, Vol. **41**, 1989, pp. 195-227.

[8] M. BOUGEARD, J.-P. PENOT and A. POMMELET, Towards minimal assumptions for the infimal convolution regularization, *J. of Approx. Theory*, Vol. **64**, 1991, pp. 245-270.

[9] H. BRÉZIS, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland, Amsterdam, 1973.

[10] F. H. CLARKE, *Optimization and Nonsmooth Analysis*, J.Wiley, New York, 1983.

[11] I. EKELAND and J.-M. LASRY, Problèmes variationnels non convexes en dualité, *C. R. Acad. Sci. Paris*, **291**, Series I, 1980, pp. 493-497.

[12] J.-B. HIRIART-URRUTY, A general formula on the conjugate of the difference of functions, *Can. Math. Bull.*, Vol. **29**, 1986, pp. 482-485.

[13] J.-B. HIRIART-URRUTY and Ph. PLAZANET, Moreau's decomposition Theorem revisited, *Analyse non linéaire*, H. ATTOUCH, J.-P. AUBIN, F. H. CLARKE, I. EKELAND Eds., Gauthier-Villars, Paris, 1989.

[14] J.-B. HIRIART-URRUTY, Extension of Lipschitz functions, *J. Math. Anal. Appl.*, Vol. **72**, 1980, pp. 539-554.

[15] J.-B. HIRIART-URRUTY, Lipschitz *r*-continuity of the approximate subdifferential of a convex function, *Math. Scand.*, Vol. **47**, 1980, pp. 123-134.

[16] J.-M. LASRY and P.-L. LIONS, A remark on regularization in Hilbert spaces, *Israel Journal of Mathematics*, Vol. **55**, 1986, pp. 257-266.

[17] J.-J. MOREAU, Fonctionnelles convexes, *Lecture Notes*, Collège de France, Paris, 1967.

[18] J.-J. MOREAU, Proximité et dualité dans un espace Hilbertien, *Bull. Soc. Math. Fr.*, Vol. **93**, 1965, pp. 273-299.

[19] A. PAZY, Semi-groups of nonlinear contractions in Hilbert spaces, in *Problems in Nonlinear Analysis*, Edizioni Cremonese, Roma, 1971, pp. 343-430.

[20] A. POMMELET, Analyse convexe et théorie de Morse, *Thèse de 3e cycle*, Université de Paris-IX, 1982.

[21] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, 1966.

[22] R. T. ROCKAFELLAR, Generalized directional derivatives and subgradients of nonconvex functions, *Canadian J. Math.*, Vol. **32**, 1980, pp. 157-180.

[23] S. ROLEWICZ, On paraconvex multifunctions, *Proceedings of III Symposium uber Operation Research*, Mannheim, Sept. 1978, pp. 539-546.

[24] A. A. VLADIMIROV, Yu. E. NESTEROV and Yu. N. CHEKANOV, On uniformly convex functionals, *Vest. Mosk. Univ.*, Vol. **3**, 1978, pp. 12-23 (Russian).

[25] J. C. WELLS, Differentiable functions on Banach spaces with lipschitz derivative, *J. Differential Geometry*, Vol. **8**, 1973, pp. 135-152.

[26] K. YOSIDA, *Functional analysis*, third ed., Springer, Berlin, Heidelberg, New York, 1971.

[27] C. ZALINESCU, On uniformly convex functions, *Jour. Math. Anal. Appl.*, Vol. **95**, 1983, pp. 344-374.