JOURNAL OF APPROXIMATION THEORY 64, 245–270 (1991)

Towards Minimal Assumptions for the Infimal Convolution Regularization

MIREILLE BOUGEARD

I.U.T. Ville d'Avray, 1 chemin Desvallières, 92410 Ville d'Avray, France

JEAN-PAUL PENOT

Faculté des Sciences, UA CNRS 1204, Avenue de l'Université, 64000 Pau, France

AND

ALAIN POMMELLET

8 Avenue de Celle, 92360 Meudon La Foret, France

Communicated by Charles K. Chui

Received May 13, 1986; revised April 5, 1989

We encompass the Moreau–Yosida regularization process by infimal convolution into a general framework. This sheds light on the assumptions required for obtaining the usual properties. In particular the class of lower-T^2 mappings is shown to be a suitable class for performing the usual proximal regularization in open subsets of Hilbert spaces. The role of growth conditions is pointed out.

The present work can be seen as an introduction to [21] and a supplement to it. Here we focus our attention on a generalization of the Moreau–Yosida regularization process of a real-valued function f in a metric space (X, d) given by

$$f_\varepsilon(x) = \inf_{w \in X} \left[f(w) + \tfrac{1}{2} \varepsilon^{-1} d(w, x)^2 \right] \tag{1}$$

for $x \in X$, $\varepsilon > 0$. It is obtained by replacing the quadratic term $\frac{1}{2} d(w, x)^2$ above by $K(w, x)$ where $K: X^2 \to \mathbb{R}_+$ is a continuous mapping null on the diagonal called a (regularization) kernel. Under some conditions, the regularity properties of K can be transferred to the approximations f_ε. This idea occurred to several authors as far ago as R. Baire (see [4, 6, 9, 14, 22]

0021-9045/91 $3.00

for instance). Here we try to give a systematic treatment of this idea, going a step further than in [21] for what concerns differentiability of the approximations. Nevertheless we do not try to use the most general framework which would be Banach manifolds, as we are not convinced that (for the time being) the potential applications would justify the amount of work required for dealing with the geometrical problems. Still we hope that our study will make clearer what conditions are required for regularization, in particular in the case of an open subset X of a reflexive Banach space when some kind of local convexity can be invoked on f. Here, as in [21] we stress the favorable class of lower-C^2 mappings (or its extension to infinite dimensional Hilbert spaces [20, 24]). Some results were obtained in [21] when f is allowed to take the value $+\infty$. Here we reject this extension, realizing that it leads to non-trivial problems. For instance, when f is the indicator function of some subset A of X (i.e., f is zero on A and $+\infty$ elsewhere) then f_ε is nothing but $\frac{1}{2}\varepsilon^{-1}d_A^2$ where $d_A = d(\cdot, A)$ is the distance to A; thus one is led to problems such as the existence of proximal points and the like (see [12, 19] and their references). For a study in the important case of a subset A defined by equalities and inequalities as in mathematical programming see [3].

After a short comparison of the merits of the approximation process by infimal convolution with those of the approximation by mollifiers in Section 1, we reveal the utmost importance of growth conditions and describe the elementary properties of the infimal convolution approximation (Section 2). Section 3 is devoted to differentiability properties of the approximations; it contains a study of the limit behavior of the derivatives $(f'_\varepsilon)_{\varepsilon>0}$ as $\varepsilon \to 0_+$ which seems to be new, at least in the nonconvex case. We conclude with an extension of the classical use of regularization in epiconvergence (see [1, 2, 8, 13]).

Throughout, the open ball with center x and radius r in a metric space is denoted by $B(x, r)$ and the set of positive real numbers is denoted by $\mathbb{P}$, while $\mathbb{R}_+ = \mathbb{P} \cup \{0\}$, $\mathbb{R}^{\bullet} = \mathbb{R} \cup \{+\infty\}$, $\bar{\mathbb{R}} = \mathbb{R}^{\bullet} \cup \{-\infty\}$. For a subset A of a metric space (E, d) and $x \in E$ we set $d(x, A) = \inf\{d(x, a) : a \in A\}$.

1. Regularization via Convolution versus Regularization via Infimal Convolution

The most usual way of regularizing a locally integrable function f on some open subset X of an euclidean space E of dimension d consists of taking a mollifier M on E (i.e., a C^∞ function with compact support such that $\int_E M(x)\,dx = 1$) and in setting

$$R_\varepsilon f(x) = \varepsilon^{-d} \int_X M(\varepsilon^{-1}(x - v)) f(v)\,dv. \tag{2}$$

This regularization process is no more valid in infinite dimensional spaces (unless some more sophisticated tools such as Wiener measures are used). On the other hand $R_\varepsilon f$ is easily seen to be of class C^∞ and can be defined even when f takes its values in a Banach space.

These properties do not carry over to the regularization process (1) by infimal convolution. On the other hand it can be used when E is an infinite dimensional Hilbert space (and in even more general situations, as shown below), provided f satisfies a mild growth condition. When X and f are convex f_ε is convex over E (for $R_\varepsilon f$ this is true only on a subset X_α of X strongly contained in X in a sense made precise below). Moreover one has the following properties.

1.1. PROPOSITION. *Let $f: X \to \bar{\mathbb{R}}$ and for $\varepsilon \in \mathbb{P}$ let f_ε be defined by (1). Then*

$$\inf f_\varepsilon = \inf f.$$

Moreover any minimizer for f is a minimizer for f_ε, and if f is lower semi-continuous (l.s.c.) any minimizer for f_ε is a minimizer for f.

These assertions carry over to the more general process considered in the next section. Moreover one can show that critical points, when properly defined, are preserved [21].

Proof. The first assertion is a consequence of the equality

$$\inf_x \inf_w \left(f(w) + \tfrac{1}{2}\varepsilon^{-1} d(w, x)^2\right) = \inf_w \inf_x \left(f(w) + \tfrac{1}{2}\varepsilon^{-1} d(w, x)^2\right) = \inf_w f(w).$$

If $x \in X$ is such that $f(x) \leqslant f(w)$ for each $w \in X$ then obviously $f_\varepsilon(x) = f(x) \leqslant \inf_{w \in X} f_\varepsilon(w)$ and x is a minimizer of f_ε. Finally let x be a minimizer of f_ε and let (w_n) be a sequence such that

$$f(w_n) + \varepsilon^{-1} d(w_n, x)^2 \leqslant f_\varepsilon(x) + \frac{1}{n}.$$

Then we have

$$\varepsilon^{-1} d(w_n, x)^2 \leqslant f_\varepsilon(x) + \frac{1}{n} - \inf f = \frac{1}{n}$$

so that $(w_n) \to x$. Therefore, if f is l.s.c. at x we get

$$f(x) \leqslant \liminf_n f(w_n) \leqslant \liminf \left(f_\varepsilon(x) + \frac{1}{n}\right) = \inf f. \quad \blacksquare$$

In the following proposition for a subset A of X and $\alpha \in \mathbb{P}$ we set

$$A_\alpha = \{x \in E : d(x, A) < \alpha\}$$

while for a mapping h: $W \to \bar{\mathbb{R}}$ and $r \in \mathbb{R}$ we denote the strict r-level set of h by

$$S(h, r) = \{w \in X : h(w) < r\}.$$

1.2. PROPOSITION. *For each f: $X \to \mathbb{R}^{\bullet}$ and each $\varepsilon \in \mathbb{P}$ the strict level sets of f and f_ε: $E \to \bar{\mathbb{R}}$ are related via the formula*

$$S(f_\varepsilon, r) = \bigcup_{t > 0} S(f, r - t)_{\sqrt{2\varepsilon t}}.$$

The proof of this assertion is easy. Let us note that it might prove to be useful for giving a proof of Theorem 4.1 or Corollary 4.2 below in the spirit of [27] or for duality results in the spirit of [26].

As a further motivation for considering more general regularizing terms than the quadratic term $\frac{1}{2}d(w, x)^2$ in (1) let us note that the regularization given by

$$f_\lambda(x) = \inf_{w \in X} (f(w) + \lambda^{-1} d(w, x)), \tag{3}$$

for $x \in X$, $\lambda \in \mathbb{P}$, has been used in [6, 14] for extending lipschitzian functions and approaching lower semicontinuous (l.s.c.) functions by lipschitzian ones; here we use it to rephrase a famous result.

1.3. PROPOSITION (Ekeland's Variational Principle [10]). *Let (X, d) be a complete metric space and let f: $X \to \mathbb{R}^{\bullet}$ be a l.s.c. function bounded from below. Let $m = \inf f$. Then for any positive numbers α, λ and any α-approximate minimizer x_0 of f (i.e., $x_0 \in f^{-1}(]-\infty, m + \alpha])$) there exists $\bar{x} \in B(x_0, \alpha\lambda)$ with $f_\lambda(\bar{x}) = f(\bar{x})$.*

2. THE PROMINENT ROLE OF GROWTH CONDITIONS

Whereas the regularization process by mollifiers applies to any continuous function on a finite dimensional space, the use of the Moreau–Yosida approximation scheme is limited to functions satisfying a growth condition. This fact already noted in [1, 4, 21] becomes still more important when one deals with a function f defined on an open subset X of a Hilbert space E. The extension of f by $+\infty$ on $E \backslash X$ is l.s.c. only if

$f(x) \to +\infty$ as $x \to \bar{x}$, $x \in X$ for each $\bar{x} \in \mathrm{cl}(X) \setminus X$. A way of circumventing this difficulty consists in replacing the usual quadratic term $\frac{1}{2}|w-x|^2$ in (1) be a more general term $K(w, x)$ so that

$$f_\varepsilon(x) = \inf_{w \in X} [f(w) + \varepsilon^{-1} K(w, x)]$$

for $x \in X$, $\varepsilon \in \mathbb{P} =]0, +\infty[$. Thus, even if $f(w)$ does not converges to $+\infty$ as w converges to a boundary point, one may ensure that the infimum is attained by requiring that $K(w, x) \to +\infty$ as w converges to a boundary point of X, along with some compactness assumption or growth condition.

Given a (regularization) *kernel* K on a topological space X, i.e., a continuous function $K: X \times X \to \mathbb{R}_+$ such that $K(x, x) = 0$ for each $x \in X$, one defines the coefficient of *K-minorization* (or K-decrease) of $f: X \to \mathbb{R}^{\bullet}$ as the infimum $d_K(f)$ (or $d(f)$ if no confusion can arise) of the set of $c \in \mathbb{R}_+$ such that $f + cK(\cdot, x)$ is bounded below for each $x \in X$. If K is *coherent* in the sense that for any x, y in X and each $p > 1$ there exists $r \in \mathbb{R}$ with

$$K(w, x) \leqslant pK(w, y) + r \qquad \text{for each} \quad w \in X,$$

then

$$d_K(f) = \inf\{c \in \mathbb{R}_+ : \exists x_0 \in X\, \exists b \in \mathbb{R} : f \geqslant b - cK(\cdot, x_0)\}.$$

When X is a subset of a topological vector space (t.v.s.) E a general way of obtaining a kernel consists in setting

$$K(w, x) = k(w - x),$$

where $k: E \to \mathbb{R}_+$ is continuous such that $k(0) = 0$. When X is a subset of a metric space one can use an arbitrary continuous mapping $h: \mathbb{R}_+ \to \mathbb{R}_+$ with $h(0) = 0$ for setting

$$K(w, x) = h(d(w, x)).$$

In particular, for $\alpha \in \mathbb{P}$ the kernel associated in this way to $h_\alpha: r \to (1/\alpha) r^\alpha$ is denoted by K_α. When $\alpha = 2$ (the usual case) and $d_K(f) < +\infty$ f is said to be *quadratically minorized.*

When X is an open subset of a metric space (E, d), it may be advantageous to take into account the geometry of X by modifying the distance function on X. For instance one can set

$$d_X(x, y) = d(x, y) + |d(x, X^c)^{-1} - d(y, X^c)^{-1}|,$$

where $X^c = E \setminus X$ is supposed to be nonempty; when E is complete X is complete for d_X and d_X induces the usual topology on X. Moreover functions f on X which are not coercive on X (where f is said to be coercive if $f(x_n) \to +\infty$ as (x_n) converges to some boundary point of X or $(d(x_n, x_0)) \to +\infty$) can be taken into account. We may even allow $f(x)$ to converge to $-\infty$ as x converges to some boundary point of X as in the example $E = \mathbb{R}$, $X = \mathbb{P}$, $f = ln$.

Another way of defining a new distance on X which can be used in a kernel consists in taking the geodesic distance on X associated to a suitable Finsler (or Riemannian structure) [16, 17, 21],

$$d(x, y) = \inf \left\{ \int_0^1 g(c(t)) \, |\dot{c}(t)| \, dt : c \in C^1([0, 1], X), c(0) = x, c(1) = y \right\},$$

where $g: X \to \mathbb{P}$ is a continuous function such as $d(x, X^c)^{-1}$ for instance.

Another example of interest is the case of a kernel on a n.v.s. E given by $K(w, x) = \frac{1}{2} \langle A(w - x), w - x \rangle$ where $A: E \to E'$ is a positive linear operator from E into its topological dual space E'. For such a kernel the firmness condition introduced below is not satisfied unless A is definite positive (for some $\alpha \in \mathbb{P}$ one has $\langle Ax, x \rangle \geqslant \alpha |x|^2$ for each $x \in E$). However, when A is strictly positive ($\langle Ax, x \rangle > 0$ for $x \in E \backslash \{0\}$) the norm $| \ |_A$ given by $|x|_A = (\langle Ax, x \rangle)^{1/2}$ may be used instead of the norm of E, along with some alterations of what follows.

Finally, let us note that when E is some L_p-space, $p \geqslant 1$, a kernel of the form $K(w, x) = (1/p) \, |w - x|^p$ seems to fit more to the structure of the space than the usual quadratic kernel. A similar remark is valid for Orlicz spaces.

In this respect let us note (see also [13]) the following fact which is a direct consequence of [23, Theorem 3.A] to which we refer for the notions used below. Let $(S, \mathscr{S}, \sigma)$ be a σ-finite measured space and let $f: S \times E \to \mathbb{R}^{\bullet}$ be a normal integrand, where E is some separable Banach space. Let $k: E \to \mathbb{R}_+$ be a convex continuous function with $k(0) = 0$ and let K be the associated kernel on E given by $K(w, x) = k(w - x)$. Let X be a decomposable linear space of measurable mappings from S into E such that for each $x \in X$ $\int_S k(x(s)) \, d\sigma < +\infty$ (for instance $X = L_p$, $k(e) = (1/p) \, |e|^p$). Then we get a kernel K^S on X setting

$$K^S(w, x) = \int_S k(w(s) - x(s)) \, ds \qquad \text{for} \quad w, x \in X.$$

If we denote by f^S the integral functional defined on X by

$$f^S(x) = \int_S f(s, x(s)) \, ds$$

and by f_ε^S the similar integral associated with the ε-approximate integrand $f_\varepsilon(s, \cdot)$ we have

$$(f^S)_\varepsilon (x) = f_\varepsilon^S(x)$$

for each $x \in X$ such that the ε-approximate $(f^S)_\varepsilon$ of f^S (with respect to K^S) is finite at x, since

$$\inf_{w \in X} \int_S [f(s, w(s)) + \varepsilon^{-1} k(w(s) - x(s))] \, d\sigma$$

$$= \int_S \inf_{v \in E} [f(s, v) + \varepsilon^{-1} k(v - x(s))] \, d\sigma.$$

Therefore the knowledge of the regularization of the integrand yields the regularized integral functional.

2.1. PROPOSITION. *Let $f: X \to \mathbb{R}^\bullet$ be such that $d_K(f) < +\infty$. Then for each $\varepsilon \in]0, d_K(f)^{-1}[$ f_ε does not assume the value $-\infty$. If moreover f is proper (i.e., is finite somewhere) then f_ε is everywhere finite. Furthermore for $0 < \varepsilon < \delta < d_K(f)^{-1}$ one has*

$$f_\delta \leqslant f_\varepsilon \leqslant f.$$

Proof. Given $\varepsilon \in]0, d_K(f)^{-1}[$ and $x \in X$ we choose $c \in]d_K(f), \varepsilon^{-1}[$ and $b \in \mathbb{R}$ such that $f \geqslant b - cK(\cdot, x)$. Then we have $f_\varepsilon(x) \geqslant b$. If f takes a finite value at $z \in X$ then for each $x \in X$ we have

$$f_\varepsilon(x) \leqslant f(z) + \varepsilon^{-1} K(z, x) < +\infty.$$

The last inequalities are obvious as K is nonnegative and $f_\varepsilon(x) \leqslant f(x) + \varepsilon^{-1} K(x, x) = f(x)$. ∎

In the sequel K is said to be (locally) *firm* if for each $x \in X$ and each sequence (w_n) in X with $\lim_n K(w_n, x) = 0$ one has $(w_n) \to x$; K is said to be *locally strictly firm* if for each $x \in X$ and any sequences (w_n), (x_n) in X with $(x_n) \to x$, $(K(w_n, x_n)) \to 0$ one has $(w_n) \to x$. When X is an open subset of a n.v.s. E and $K(w, x) = k(w - x)$ for $k: E \to \mathbb{R}_+$, K is locally strictly firm iff K is firm and this is the case if k is firm in this sense that a sequence (e_n) of E has limit 0 iff $(k(e_n)) \to 0$; the converse is true when $X = E$.

2.2. PROPOSITION. *Suppose K is a firm kernel. Let $f: X \to \mathbb{R}^\bullet$ be such that $d(f) < +\infty$. Then $(f_\varepsilon)_{\varepsilon > 0}$ converges pointwise as $\varepsilon \to 0_+$ to the lower semi-continuous hull $\bar{f}(x) = \liminf_{v \to x} f(v)$.*

Proof. We may suppose f is proper since the result is trivial when $f \equiv +\infty$.

For each $\varepsilon \in \mathbb{P}$ and each net $(x_i)_{i \in I}$ with limit x in X we have $f_\varepsilon(x) \leqslant f(x_i) + \varepsilon^{-1} K(x_i, x)$ hence $f_\varepsilon(x) \leqslant \liminf_{i \in I} f(x_i)$. Therefore $\lim_{\varepsilon \to 0} f_\varepsilon(x) =$

$\sup_{\varepsilon>0} f_\varepsilon(x) \leqslant \bar{f}(x)$. Let $s < \bar{f}(x)$: there exists a neighborhood V of x such that $f(v) > s$ for each $v \in V$. As K is firm we can find $t > 0$ such that $K(w, x) \geqslant t$ for each $w \in X \backslash V$. Let $\delta \in]0, d(f)^{-1}[$ and let $\bar{\varepsilon} = (\delta^{-1} + t^{-1} |s - f_\delta(x)|)^{-1}$. Then for $\varepsilon \in]0, \bar{\varepsilon}]$ and any $w \in X \backslash V$ we have

$$f(w) + \varepsilon^{-1} K(w, x) \geqslant f_\delta(x) + (\varepsilon^{-1} - \delta^{-1})\, K(w, x)$$
$$\geqslant f_\delta(x) + (\varepsilon^{-1} - \delta^{-1})\, t \geqslant s,$$

hence $f_\varepsilon(x) \geqslant s$ as K takes nonnegative values on $V \times \{x\}$. ∎

Obviously for each $\varepsilon \in \mathbb{P}$, f_ε is upper semicontinuous (u.s.c.) as an infimum of continuous functions; in particular when X is an open convex subset of a t.v.s. and when f and K are convex we get that f_ε is continuous. More general assumptions will be given later on guaranteeing the continuity of f_ε.

Let us now suppose X is an open subset of a metric space (E, d). Then, under some conditions on K, a lipschitzian property of the mappings $K(w, \cdot)$, $w \in X$ can be transferred to the approximates f_ε of f. The result we present below is an easy variant of [21, Proposition 3.5]. It uses the family $\mathscr{B}(X)$ of bounded subsets of X which are strongly contained in X, where B is said to be *strongly contained in* X if there exists $r \in \mathbb{P}$ such that $B_r := \{x \in E : \exists y \in B,\ d(x, y) < r\}$ is contained in X.

2.3. PROPOSITION. *Let K be a coherent kernel satisfying the following conditions for some $x_0 \in X$, p, q, r in $\mathbb{P}$:*

(a) *for each $(w, x) \in X^2$ $K(w, x_0) \leqslant pK(w, x) + qK(x, x_0) + r$;*

(b) *if $K(\cdot, x_0)$ is bounded on a subset B of X then $B \in \mathscr{B}(X)$;*

(c) *for each $B \in \mathscr{B}(X)$ there exists $l \in \mathbb{R}_+$ such that*

$$|K(w, x) - K(w, y)| \leqslant l d(x, y) \qquad \textit{for each} \quad (w, x, y) \in B^3.$$

Then for any proper $f: X \to \mathbb{R}^\bullet = \mathbb{R} \cup \{+\infty\}$ with $d_K(f) < +\infty$ and any $\varepsilon \in]0, d_K(f)^{-1}[$ with $\varepsilon q d_K(f) < 1$, f_ε is lipschitzian on any member of $\mathscr{B}(X)$.

Using the quadratic kernel $K = K_2$ yields several important properties (see, for instance, [15, Propositions 3.6 and 3.7] and [21]). Let us note in particular the following two useful results.

2.4. LEMMA. *Suppose X is a subset of a Hilbert space E and $f: X \to \mathbb{R}$ is quadratically minorized. Then, for $\lambda \in]0, d(f)^{-1}[$, $f_\lambda - (1/2\lambda) |\ |^2$ is concave and u.s.c. on E.*

Proof. This follows from the fact that $f_\lambda - (1/2\lambda)\,|\;|^2$ is the infimum of the family of continuous affine functions $(g_v)_{v\in X}$ given by

$$g_v(x) = -\frac{1}{\lambda}(x \mid v) + \frac{1}{2\lambda}|v|^2 + f(v). \quad \blacksquare$$

2.5. LEMMA. *Suppose X is a convex subset of a Hilbert space E and $f: X \to \mathbb{R}^{\cdot}$ is quadratically minorized and such that $f + (1/2\lambda)\,|\;|^2$ is convex. Then, for each $\mu \in\,]0, \lambda[$, $f_\mu + (1/2(\lambda-\mu))\,|\;|^2$ is convex on E.*

Proof. It is well known that if $g: X \times E \to \mathbb{R}^{\cdot}$ is convex then $m: E \to \bar{\mathbb{R}}$ given by $m(x) = \inf\{g(w, x) : w \in X\}$ is convex. Thus, as

$$\frac{1}{\lambda-\mu}|x|^2 + \frac{1}{\mu}|w-x|^2 - \frac{1}{\lambda}|w|^2$$
$$= \frac{1}{\mu(\lambda-\mu)}\left|\left(1-\frac{\lambda}{\mu}\right)(x-w) + \frac{\lambda}{\mu}x\right|^2$$

the result follows from

$$f_\mu(x) + \frac{1}{2(\lambda-\mu)}|x|^2$$
$$= \inf_{w\in X}\left[\left(f(w) + \frac{1}{2\lambda}|w|^2\right) + \frac{1}{2\mu(\lambda-\mu)}\left|\left(1-\frac{\lambda}{\mu}\right)(x-w) + \frac{\lambda}{\mu}x\right|^2\right]. \quad \blacksquare$$

The proof shows that for $K = K_2$, $d_K(f_\mu) \geqslant (d_K(f)^{-1} - \mu)^{-1}$ for $\mu \in\,]0, d_K(f)^{-1}[$; this type of result can be extended to a general kernel K satisfying a "metric-like" condition [21, Proposition 3.2].

Now we would like to give a short account of a nice recent work of J.-M. Lasry and P.-L. Lions [15]. Rather than insisting on the uniform continuity of the functions involved, we intend to put in full light the role of a growth condition. Recall that given a mapping

$$F: X \times Y \to \bar{\mathbb{R}}$$

on the product of two metric spaces (X, d_X), (Y, d_Y) its lower Moreau–Yosida approximate (with parameters λ, μ) has been defined by H. Attouch and R. J.-B. Wets [2] as

$$F^{\downarrow}(\lambda, \mu, x, y) = \sup_{v\in Y}\inf_{u\in X}\left[F(u, v) + \frac{1}{2\lambda}d_X^2(u, x) - \frac{1}{2\mu}d_Y^2(v, y)\right].$$

Given $f: X \to \mathbb{R}$ on a Hilbert space X, J.-M. Lasry and P.-L. Lions introduce the (λ, μ)-approximate of f by

$$f^{\downarrow}_{\lambda,\mu}(x) := F^{\downarrow}(\lambda, \mu, x, 0), \qquad \text{where} \quad F(x, y) = f(x-y),$$

so that, setting $w=u-v$, $z=x-v$ one has

$$f^{\downarrow}_{\lambda,\mu}(x)=\sup_{v\in Y}\inf_{w\in X}\left[f(w)+\frac{1}{2\lambda}|v+w-x|^2-\frac{1}{2\mu}|v|^2\right]$$

$$=\sup_{z\in Y}\inf_{w\in X}\left[f(w)+\frac{1}{2\lambda}|w-z|^2-\frac{1}{2\mu}|z-x|^2\right].$$

Therefore $f^{\downarrow}_{\lambda,\mu}=-(-f_\lambda)_\mu$.

2.6. THEOREM (compare with [15]). *Let X be a nonempty subset of a Hilbert space E and let $f\colon X\to\mathbb{R}$ be such that for some b, c in $\mathbb{R}_+$ $|f(x)|\leqslant\frac{1}{2}c|x|^2+b$ for each $x\in X$. Then for $\lambda\in\,]0,c^{-1}[$, $\mu\in\,]0,\lambda[$, $f^{\downarrow}_{\lambda,\mu}$ is a mapping of class C^1 with lipschitzian derivative of lipschitzian rate* $\max(\mu^{-1},(\lambda-\mu)^{-1})$.

Proof. Let $b, c\in\mathbb{R}_+$ be such that $|f|\leqslant\frac{1}{2}c|\cdot|^2+b$. Then, for $\lambda\in\,]0,c^{-1}[$, $-f_\lambda\geqslant -f\geqslant -\frac{1}{2}c|\ |^2-b$ so that, using Lemma 2.4, we get that $f^{\downarrow}_{\lambda,\mu}+(1/2\mu)|\cdot|^2=-((-f_\lambda)_\mu-(1/2\mu)|\cdot|^2)$ is convex on E. Therefore, for each $x\in E$, the directional subderivative of $f^{\downarrow}_{\lambda,\mu}$ at x given by

$$f^{\downarrow}_{\lambda,\mu}(x,y)=\liminf_{(t,z)\to(0_+,y)}\frac{1}{t}(f^{\downarrow}_{\lambda,\mu}(x+tz)-f^{\downarrow}_{\lambda,\mu}(x))\qquad\text{for}\quad y\in E$$

is a l.s.c. sublinear mapping in y. On the other hand, as $-f_\lambda\geqslant-\frac{1}{2}c|\cdot|^2-b$ and $-f_\lambda+(1/2\lambda)|\cdot|^2$ is convex on E by Lemma 2.4 we can conclude from Lemma 2.5 that for $\mu\in\,]0,\lambda[$, $f^{\downarrow}_{\lambda,\mu}-(1/2(\lambda-\mu))|\cdot|^2=-[(-f_\lambda)_\mu+(1/2(\lambda-\mu))|\cdot|^2]$ is concave on E. Therefore $f^{\downarrow}_{\lambda,\mu}$ has a continuous linear directional derivative at each point of E.

It remains to apply the following result to $h=f^{\downarrow}_{\lambda,\mu}$.

2.7. LEMMA. *Let $h\colon E\to\mathbb{R}$ be a continuous mapping such that for some $v\in\mathbb{P}=\,]0,+\infty[$, $h+\frac{1}{2}v|\cdot|^2$ and $-h+\frac{1}{2}v|\cdot|^2$ are convex. Then h is of class C^1 and its derivative is lipschitzian with rate v.*

Proof. What precedes shows that h is directionally differentiable; thus it suffices to show that ∇h has Lipschitz rate v (this will ensure Fréchet differentiability). Now, by a classical result of Alexandroff, for each finite dimensional subspace F of E the restriction h_F of h to F is twice differentiable on the complement of a null set N of F. Moreover, by a well-known property of symmetric bilinear functionals, for each $z\in F\setminus N$

$$\|h''_F(z)\|=\max(\sup_{|v|\leqslant 1}h''_F(z)\cdot v\cdot v,\ \sup_{|v|\leqslant 1}-h''_F(z)\cdot v\cdot v)\leqslant v.$$

It follows that h'_F has Lipschitz rate v. Therefore, for each v, x, y in E, taking any finite dimensional subspace F containing v, x, y we get

$$|h'(x)\,v - h'(y)\,v| = |h'_F(x)\,v - h'_F(y)\,v| \leqslant v\;|v| \cdot |x - y|.$$

As v is arbitrary we get that h' has Lipschitz rate v. ∎

The fact that the preceding result applies to uniformly continuous functions follows from the following simple observation.

2.8. LEMMA. *Let $f\colon X \to \mathbb{R}$ be a uniformly continuous function on a convex subset X of a n.v.s. E. Then there exists $b, c \in \mathbb{R}_+$ such that $|f(x)| \leqslant c\,|x| + b$.*

Proof. Let us define $m\colon \mathbb{R}_+ \to \bar{\mathbb{R}}_+$ by

$$m(r) = \sup\{|f(x) - f(y)| : (x, y) \in X^2, |x - y| \leqslant r\},$$

so that $\lim_{r \to 0_+} m(r) = 0$. Subdividing any segment $[x, y]$ of X into k segments we observe that for any $k \in \mathbb{N}$, $m(kr) \leqslant km(r)$ and $m(r+s) \leqslant m(r) + m(s)$ for any $r, s \in \mathbb{R}_+$. Thus m is finite valued and

$$m(r) \leqslant [r]\, m(1) + m(1) \leqslant m(1)(r+1)$$

for $[r] = \max\{k \in \mathbb{N} : k \leqslant r\}$, so that, for any $(x, x_0) \in X^2$

$$\begin{aligned} |f(x)| &< |f(x_0)| + m(1)(|x - x_0| + 1) \\ &\leqslant m(1)\,|x| + |f(x_0)| + m(1)(|x_0| + 1). \end{aligned}$$ ∎

3. EXACTNESS AND DIFFERENTIABILITY

Let us call the ε-approximate f_ε of f *exact at* x (resp. *strictly exact* at x) if the infimum

$$f_\varepsilon(x) = \inf_{w \in X} [f(w) + \varepsilon^{-1} K(w, x)]$$

is attained (resp. attained at a unique point). This property is intimately linked with differentiability properties of f_ε when X is an open subset of a n.v.s. E and K is differentiable.

Given a kernel K on an open subset X of a n.v.s. E let us define the *index of K-nonconvexity* of $f\colon X \to \mathbb{R}^\bullet$ at $\bar{x}$ as the infimum $c_K(f, \bar{x})$ (or $c(f, \bar{x})$ if no confusion can arise) of the set of $c \in \mathbb{R}_+$ such that there exists $r \in \mathbb{P}$ for which the mapping $w \mapsto f(w) + cK(w, x)$ is convex and proper (i.e., $\not\equiv +\infty$) on $B(\bar{x}, r)$ for each $x \in B(\bar{x}, r)$. Some properties of this index are described in [21] when $K = K_2$ (but there the properness condition was not required).

In the sequel we suppose X is an open subset of a reflexive Banach space E and K is *locally convex* in the following sense: for each $\bar{x} \in X$ there exists $\alpha \in \mathbb{P}$ such that $B(\bar{x}, \alpha) \subset X$ and for each $x \in B(\bar{x}, \alpha)$, $K(\cdot, x)$ is convex on $B(\bar{x}, \alpha)$. When in the preceding condition $K(\cdot, x)$ is strictly convex on $B(\bar{x}, \alpha)$, K is said to be *locally strictly convex*.

Let us introduce some conditions on the kernel K. The first one is rather mild; in particular it is a weakening of the metric-like condition of [21]:

(m) for each $\bar{x} \in X$ there exist p, q, r, s in $\mathbb{P} =]0, +\infty[$ such that

$$K(w, \bar{x}) \leqslant pK(w, x) + qK(x, \bar{x}) + r \qquad \text{for each} \quad (w, x) \in X \times B(\bar{x}, s).$$

When $K(w, x) = h(|w - x|)$ where $h: \mathbb{R}_+ \to \mathbb{R}_+$ is a convex continuous function with $h(0) = 0$ this condition is satisfied with $p = q = \frac{1}{2}c$, $r = d$, s arbitrary whenever h satisfies the following classical condition:

(Δ_2) there exists $c \in \mathbb{P}$, $d \in \mathbb{R}_+$ such that $h(2t) \leqslant ch(t) + d$ for each $t \in \mathbb{R}_+$.

In particular this condition is satisfied for $h(t) = (1/\alpha)\, t^\alpha$, $\alpha \geqslant 1$.

Our second condition is a strengthening of the firmness condition, so that K will be said to be *strongly firm* if it satisfies it. It reads as follows (here $\bar{B}(x, \alpha)$ denotes the closed ball with center x and radius α):

(f) for each $\bar{x} \in X$, each $\alpha \in \mathbb{P}$ with $\bar{B}(\bar{x}, \alpha) \subset X$, each $z \in B(\bar{x}, \alpha)$ there exist β, γ, δ in $\mathbb{P}$ such that

$$K(w, x) \geqslant \beta + \gamma,\ \beta \geqslant K(z, x) \qquad \text{for any} \quad w \in X \backslash B(\bar{x}, \alpha),\ x \in B(\bar{x}, \delta).$$

When $K(w, x) = k(w - x)$ for some $k: E \to \mathbb{R}_+$ with $k(0) = 0$ this condition is satisfied whenever k enjoys the property:

(f_0) for each ρ, σ in $\mathbb{P}$, with $\rho > \sigma$, $\inf\{k(u) : |u| \geqslant \rho\} > \sup\{k(v) \leqslant \sigma\}$.

In particular (f) and (f_0) are satisfied when $k(v) = h(|v|)$ where $h: \mathbb{R}_+ \to \mathbb{R}_+$ is continuous, strictly increasing with $h(0) = 0$. This is the case for $h(t) = (1/\alpha)\, t^\alpha$ with $\alpha \geqslant 1$.

On the other hand, when the following variant (m') of condition (m) holds condition (f) can be simplified into

(f') for each $\bar{x} \in X$, each $\alpha \in \mathbb{P}$ with $\bar{B}(\bar{x}, \alpha) \subset X$, each $z \in B(\bar{x}, \alpha)$

$$\inf\{K(w, \bar{x}) : w \in X \backslash B(\bar{x}, \alpha)\} > K(z, \bar{x}),$$

where condition (m') is

(m') for any $\bar{x} \in X$, $r \in \mathbb{P}$, $p \in \mathbb{R}$ with $p > 1$ there exist $q \in \mathbb{R}_+$, $s \in \mathbb{P}$ such that for each $(w, x) \in X \times B(\bar{x}, s)$

$$K(w, \bar{x}) \leqslant pK(w, x) + qK(x, \bar{x}) + r.$$

LEMMA. *When condition (m') holds true conditions (f) and (f') are equivalent.*

Proof. It suffices to show that (f) holds true when (f') and (m') are satisfied. Let $\theta \in \mathbb{P}$ be such that $K(w, \bar{x}) \geqslant K(z, \bar{x}) + \theta$ for each $w \in X \setminus B(\bar{x}, \alpha)$ and let $\gamma = r = \frac{1}{8}\theta$. Let us choose $p \in]1, 2[$ such that $p^{-1}(K(z, \bar{x}) + \theta) > K(z, \bar{x}) + \frac{1}{2}\theta$ and let $\delta \in]0, s[$ be such that $qK(x, \bar{x}) < r$, $|K(z, \bar{x}) - K(z, x)| < r$ for $x \in B(\bar{x}, \delta)$. Then for $w \in X \setminus B(\bar{x}, \alpha)$, $x \in B(\bar{x}, \delta)$

$$\beta := K(z, \bar{x}) + r \geqslant K(z, x),$$

$$K(w, x) \geqslant p^{-1}K(w, \bar{x}) - p^{-1}qK(x, \bar{x}) - p^{-1}r$$

$$\geqslant K(z, \bar{x}) + \tfrac{1}{2}\theta - 2r = \beta + \gamma. \quad \blacksquare$$

3.1. PROPOSITION. *Let K be a locally convex (resp. locally strictly convex) kernel on X satisfying conditions (f) and (m) or conditions (f') and (m') above. Let $f: X \to \mathbb{R}^{\bullet} = \mathbb{R} \cup \{+\infty\}$ be* l.s.c. *with $d_K(f) < +\infty$ and $c_K(f, x) < +\infty$ for each $x \in X$. Then there exists an open subset $\hat{X}$ of $X \times \mathbb{P}$ such that $\hat{X} \cup X \times \{0\}$ is a neighborhood of $X \times \{0\}$ in $X \times \mathbb{R}_+$ and such that for each $(x, \varepsilon) \in \hat{X}$ the ε-approximate of f is exact (resp. strictly exact) at x.*

Proof. Let $d > d_K(f)$. For each $\bar{x} \in X$ we can find $b = b(\bar{x}) \in \mathbb{R}$ such that $f \geqslant b - dK(\cdot, \bar{x})$ and $\alpha = \alpha(\bar{x}) \in \mathbb{P}$ such that for some $c = c(\bar{x}) \in \mathbb{R}_+$, $\bar{B}(\bar{x}, \alpha)$ is contained in X and for each $x \in B(\bar{x}, \alpha)$, $f + cK(\cdot, x)$ and $K(\cdot, x)$ are convex on $B(\bar{x}, \alpha)$, f being proper on $B(\bar{x}, \alpha)$. Let $z \in B(\bar{x}, \alpha)$ be such that $f(z)$ is finite. Using condition (f) we can find β, γ, δ in $\mathbb{P}$ such that

(f) $K(w, x) \geqslant \beta + \gamma$, $\beta \geqslant K(z, x)$ for any $w \in X \setminus B(\bar{x}, \alpha)$, $x \in B(\bar{x}, \delta)$.

Let p, q, r, s be as in condition (m), the dependence on $\bar{x}$ of these numbers being omitted for the moment for the sake of simplicity:

(m) $K(w, \bar{x}) \leqslant pK(w, x) + qK(x, \bar{x}) + r$ for any $w \in X$, $x \in B(\bar{x}, s)$.

We may take $\delta \in]0, s[$ so small that $qK(x, \bar{x}) < r$ for each $x \in B(\bar{x}, \delta)$. Let

$$\bar{\varepsilon} = \bar{\varepsilon}(\bar{x}) = \min(c^{-1}, d^{-1}p^{-1}, \gamma\, |f(z) + 2dr + dp(\beta + \gamma) - b|^{-1}).$$

Then for $\varepsilon \in]0, \bar{\varepsilon}[$, $w \in X \setminus B(\bar{x}, \alpha)$, $x \in B(\bar{x}, \delta)$ we have

$$(\varepsilon^{-1} - dp)\, K(w, x) - \varepsilon^{-1}K(z, x) \geqslant \varepsilon^{-1}\gamma - dp(\beta + \gamma)$$

so that

$$
\begin{aligned}
f(w)+\varepsilon^{-1}K(w, x) &\geq b - dK(w, \bar{x}) + \varepsilon^{-1}K(w, x) \\
&\geq b + (\varepsilon^{-1} - dp)\, K(w, x) - dqK(x, \bar{x}) - dr \\
&> f(z) + \varepsilon^{-1}K(z, x).
\end{aligned}
$$

This shows that for each $\varepsilon \in]0, \bar{\varepsilon}[$ and each $x \in B(\bar{x}, \delta)$

$$F_\varepsilon(\cdot, x) := f + \varepsilon^{-1}K(\cdot, x)$$

cannot attain its minimum on X but on $B(\bar{x}, \alpha)$. As the closure $\bar{B}(\bar{x}, \alpha)$ of $B(x, \alpha)$ is weakly compact and as for $x \in B(\bar{x}, \alpha)$

$$F_\varepsilon(\cdot, x) = F_{c^{-1}}(\cdot, x) + (\varepsilon^{-1} - c)\, K(\cdot, x)$$

is convex and weakly l.s.c. on $\bar{B}(\bar{x}, \alpha)$, this function does attain its minimum on $B(\bar{x}, \alpha)$, hence on X. When K is locally strictly convex this minimizer is unique.

Let

$$\hat{X} = \bigcup_{\bar{x} \in X} B(\bar{x}, \delta(\bar{x})) \times]0, \bar{\varepsilon}(\bar{x})[,$$

where now the dependence of α, δ, $\bar{\varepsilon}$ on $\bar{x}$ is taken into account. Then $\hat{X}$ is open, $\hat{X} \cup (X \times \{0\})$ is a neighborhood of $X \times \{0\}$ in $X \times \mathbb{R}_+$, and for each $(x, \varepsilon) \in \hat{X}$ we can find $\bar{x} \in X$ with $x \in B(\bar{x}, \delta(\bar{x}))$, $\varepsilon \in]0, \bar{\varepsilon}(\bar{x})[$ so that f_ε is exact at x. ∎

When f is supposed to be finite everywhere the proof of the preceding result becomes simpler and its conclusion can be made more complete. More generally, when the domain of f is dense in X, in the preceding result one can replace assumption (f) by the condition that K is (locally) strictly firm.

3.2. THEOREM. *Let $f: X \to \mathbb{R}$ be l.s.c. such that $d_K(f) < +\infty$, $c_K(f, x) < +\infty$ for each $x \in X$, where K is a strictly firm and locally strictly convex kernel on X satisfying condition (m). Then there exists an open subset $\hat{X}$ of $X \times \mathbb{P}$ containing the trace on $X \times \mathbb{P}$ of a neighborhood of $X \times \{0\}$ in $X \times \mathbb{R}$ such that for each $(x, \varepsilon) \in X$ there exists a unique $J_\varepsilon x \in X$ verifying*

$$f_\varepsilon(x) = f(J_\varepsilon x) + \varepsilon^{-1}K(J_\varepsilon x, x).$$

Moreover for each $x \in X$, $(J_\varepsilon x)_{\varepsilon > 0}$ converges to x as $\varepsilon \to 0_+$.

Proof. Let us first observe that, by a well-known argument about l.s.c. convex functions on Banach spaces, f is continuous on X.

Let $\bar{x} \in X$, let $d > d_K(f)$, and let $b \in \mathbb{R}$, $c \in \mathbb{R}_+$, $\alpha \in \mathbb{P}$ be such that $f \geqslant b - dK(\cdot, x)$, $f + cK(\cdot, x)$ and $K(\cdot, x)$ are convex on $\bar{B}(\bar{x}, \alpha) \subset X$ for each $x \in B(\bar{x}, \alpha)$. Let p, q, r, s be as in condition (m); we may suppose $\alpha \leqslant s$. Let $\eta \in]0, \alpha]$. As K is locally strictly firm we can find $\sigma \in \mathbb{P}$ and $\delta \in]0, \eta]$ such that $K(w, x) \geqslant \sigma$ for any $w \in X \setminus B(\bar{x}, \eta)$ and any $x \in B(\bar{x}, \delta)$. We take δ so small that $qK(x, \bar{x}) < r$, $f(x) < f(\bar{x}) + r$ for each $x \in B(\bar{x}, \delta)$. Let

$$\bar{\varepsilon} := \bar{\varepsilon}(\bar{x}) := \min(c^{-1}, d^{-1}p^{-1}, \sigma\, |f(\bar{x}) + r + 2dr + dp\sigma - b|^{-1}).$$

Then for $\varepsilon \in]0, \bar{\varepsilon}]$, $w \in X \setminus B(\bar{x}, \eta)$, $x \in B(\bar{x}, \delta)$ we have

$$\begin{aligned} f(w) + \varepsilon^{-1} K(w, x) &\geqslant b - dK(w, \bar{x}) + \varepsilon^{-1} K(w, x) \\ &\geqslant b + (\varepsilon^{-1} - dp)\, K(w, x) - dqK(x, \bar{x}) - dr \\ &\geqslant b + (\varepsilon^{-1} - dp)\, \sigma - 2dr \\ &\geqslant f(\bar{x}) + r > f(x) + \varepsilon^{-1} K(x, x). \end{aligned}$$

Therefore the minimum of $F_\varepsilon(\cdot, x) := f + \varepsilon^{-1} K(\cdot, x)$ on X is attained on $B(\bar{x}, \eta)$ and not elsewhere. As $\eta \leqslant \alpha$ the minimizer $J_\varepsilon x$ is unique and as $\eta \in]0, \alpha]$ is arbitrary we get that $(J_\varepsilon x) \to x$ as $\varepsilon \to 0_+$. Finally we take $\hat{X}$ as in the preceding proof with $\bar{\varepsilon}(\bar{x})$ as above, $\delta(\bar{x})$ being the $\delta \in]0, \eta]$ corresponding to $\eta = \alpha = \alpha(\bar{x})$.

Let us now consider the question of continuity for J_ε; we give two results in this direction.

Let us recall that a mapping $J: D \to E$ with $D \subset E$ is said to be *mildly continuous* if it is continuous when D is endowed with the strong topology and E is endowed with the weak topology. We shall require on K the following equicontinuity condition on the members of the family $\mathscr{B}(X)$ of bounded subsets which are strongly contained in X:

(e) for each $B \in \mathscr{B}(X)$ the family $\{K(w, \cdot) : w \in B\}$ is equicontinuous on B.

In other terms, for each sequence (w_n) in B and each sequence (x_n) in B with limit x one has $\lim_n (K(w_n, x_n) - K(w_n, x)) = 0$. This condition is satisfied if the Lipschitz condition (c) of Proposition 2.3 holds true.

Ordinary continuity of J_ε will be obtained either under a strong convexity assumption or under the following condition on K:

(h) if (w_n) has weak limit w in X and $(K(w_n, x))$ converges to $K(w, x)$ for some $x \in X$ then (w_n) converges to w.

When $K(w, x) = k(|w - x|)$ where $k: \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous strictly increasing convex function satisfying $k(0)=0$, and when the norm of E satisfies condition (H) below then condition (h) is satisfied:

(H) if (e_n) has weak limit e and if $(|e_n|)$ has limit $|e|$ then $\lim |e_n - e| = 0$.

3.3. PROPOSITION. (a) *Suppose the assumptions of the preceding theorem are in force and condition* (e) *holds true. Then for some choice of* $\hat{X}$ *the mappings* J_ε *are mildly continuous from* $X_\varepsilon = \{x \in X : (x, \varepsilon) \in \hat{X}\}$ *into* E.

(b) *If moreover condition* (h) *holds true then* J_ε *is continuous.*

Proof. (a) Let us keep the notations of the preceding proof; for each $\bar{x} \in X$ we shrink $\alpha(\bar{x})$ if necessary so that $B(\bar{x}, \alpha(\bar{x}))$ is strongly contained in X. Let (x_n) be a sequence with limit x in X_ε. Without loss of generality we may suppose that x and the whole sequence (x_n) are contained in some ball $B(\bar{x}, \delta(\bar{x}))$. As $J_\varepsilon x_n \in B(\bar{x}, \alpha(\bar{x}))$ for each n, a subsequence $(J_\varepsilon x_n)_{n \in N}$ (with N an infinite subset of $\mathbb{N}$) has a weak limit $w \in \bar{B}(\bar{x}, \alpha(\bar{x}))$. Then, setting $w_n = J_\varepsilon x_n$ and using assumption (e) and the fact that $F_\varepsilon(\cdot, x)$ is weakly l.s.c. on $\bar{B}(\bar{x}, \alpha(\bar{x}))$ as any continuous convex function we get

$$\begin{aligned} f(J_\varepsilon x) + \varepsilon^{-1} K(J_\varepsilon x, x) &= \lim_n \left(f(J_\varepsilon x) + \varepsilon^{-1} K(J_\varepsilon x, x_n) \right) \\ &\geqslant \limsup_n \left(f(J_\varepsilon x_n) + \varepsilon^{-1} K(J_\varepsilon x_n, x_n) \right) \\ &\geqslant \liminf_n \left(f(w_n) + \varepsilon^{-1} K(w_n, x) \right) \\ &\quad + \lim_n \varepsilon^{-1} \left(K(w_n, x_n) - K(w_n, x) \right) \\ &\geqslant f(w) + \varepsilon^{-1} K(w, x), \end{aligned}$$

so that, by uniqueness, $w = J_\varepsilon x$. As N can be chosen to be a subset of any given infinite subset M of $\mathbb{N}$, the whole sequence $(J_\varepsilon x_n)_{n \in \mathbb{N}}$ converges weakly to $J_\varepsilon x$.

(b) Let (x_n) be as above and let $\lambda \in]\varepsilon, \varepsilon(\bar{x})[$ so that $f + \lambda^{-1} K(\cdot, x)$ is convex on $B(\bar{x}, \alpha(\bar{x}))$. Observing that the preceding inequalities yield

$$\begin{aligned} f(J_\varepsilon x) + \varepsilon^{-1} K(J_\varepsilon x, x) &= \lim \left(f(w_n) + \varepsilon^{-1} K(w_n, x) \right) \\ &\geqslant \liminf \left(f(w_n) + \lambda^{-1} K(w_n, x) \right) \\ &\quad + (\varepsilon^{-1} - \lambda^{-1}) \limsup_n K(w_n, x) \\ &\geqslant f(J_\varepsilon x) + \lambda^{-1} K(J_\varepsilon x, x) \\ &\quad + (\varepsilon^{-1} - \lambda^{-1}) K(J_\varepsilon x, x) \end{aligned}$$

we get $\limsup K(w_n, x) = \liminf K(w_n, x) = K(J_\varepsilon x, x)$. Using condition (h) we obtain that $(J_\varepsilon x_n)_n$ converges to $J_\varepsilon x$. ∎

3.4. *Remark.* The hypothesis of Theorem 3.2 and condition (b) guarantee that f_ε is continuous on X_ε.

3.5. PROPOSITION. *Suppose the assumptions of Theorem 3.2 are in force, suppose f_ε is continuous and K satisfies the following strong convexity assumption:*

(c) *for each $\bar{x} \in X$ there exists $\rho \in \mathbb{P}$ and $c \in \mathbb{P}$ such that for $(x, y, z) \in B(\bar{x}, \rho)^2$*

$$K(\tfrac{1}{2}y + \tfrac{1}{2}z, x) \leqslant \tfrac{1}{2}K(y, x) + \tfrac{1}{2}K(z, x) - c\,|y - z|^2.$$

Then J_ε is continuous on X_ε, where $X_\varepsilon = \{x \in X : (x, \varepsilon) \in \hat{X}\}$.

Proof. For each $\bar{x} \in X$ we choose the associated $\alpha \in \mathbb{P}$ of the proof of Theorem 3.2 so that $\alpha < \rho$. Suppose J_ε is not continuous at some $x \in B(\bar{x}, \alpha)$: there exists $\sigma \in \mathbb{P}$ and a sequence (x_n) of $B(\bar{x}, \alpha)$ with limit x such that $|J_\varepsilon x_n - J_\varepsilon x| \geqslant \sigma$ for each $n \in \mathbb{N}$. Then, with the notations of the proof of Theorem 3.2 we observe that for $\lambda, \varepsilon \in \mathbb{P}$ with $\varepsilon < \lambda < \bar{\varepsilon}$

$$F_\varepsilon(\cdot, x) = F_\lambda(\cdot, x) + (\varepsilon^{-1} - \lambda^{-1})\, K(\cdot, x)$$

is strongly convex on $B(\bar{x}, \alpha)$ for each $x \in B(\bar{x}, \alpha)$. In particular

$$F_\varepsilon(\tfrac{1}{2}J_\varepsilon x_n + \tfrac{1}{2}J_\varepsilon x, x_n) \leqslant \tfrac{1}{2}F_\varepsilon(J_\varepsilon x_n, x_n) + \tfrac{1}{2}F_\varepsilon(J_\varepsilon x, x_n) - c\,|J_\varepsilon x_n - J_\varepsilon x|^2.$$

Taking the limits as $n \to +\infty$ we get, since $K(J_\varepsilon x, \cdot)$ is continuous,

$$\begin{aligned} f_\varepsilon(x) = \lim_n f_\varepsilon(x_n) &\leqslant \liminf F_\varepsilon(\tfrac{1}{2}J_\varepsilon x_n + \tfrac{1}{2}J_\varepsilon x, x_n) \\ &\leqslant \limsup_n \tfrac{1}{2}f_\varepsilon(x_n) + \limsup_n \tfrac{1}{2}F_\varepsilon(J_\varepsilon x, x_n) - c\sigma^2 \\ &\leqslant \tfrac{1}{2}f_\varepsilon(x) + \tfrac{1}{2}F_\varepsilon(J_\varepsilon x, x) - c\sigma^2 = f_\varepsilon(x) - c\sigma^2, \end{aligned}$$

a contradiction. ∎

Some differentiability assumption must be made on K in order that f_ε be differentiable. The following assumption (d) is in particular satisfied when $K(w, x) = k(w - x)$ with k strictly differentiable.

3.6. PROPOSITION. *Suppose with the assumptions of Theorem 3.2 that J_ε is continuous and that K satisfies the following differentiability assumption:*

(d) *for each* (w, x) *in* $X \times X$ *there exists a continuous linear functional* $D_2K(w, x)$ *on* E *such that*

$$\lim_{\substack{v \to 0, v \neq 0 \\ u \to w}} \frac{1}{|v|} (K(u, x+v) - K(u, x) - D_2K(w, x)\, v) = 0.$$

Then f_ε *is Fréchet differentiable on* X_ε *with*

$$f'_\varepsilon(x) = \varepsilon^{-1} D_2 K(J_\varepsilon x, x).$$

Proof. For each $(x, y) \in X_\varepsilon^2$ we have

$$f_\varepsilon(y) \leqslant f(J_\varepsilon x) + \varepsilon^{-1} K(J_\varepsilon x, y)$$

hence

$$\begin{aligned} f_\varepsilon(y) - f_\varepsilon(x) &\leqslant f(J_\varepsilon x) + \varepsilon^{-1} K(J_\varepsilon x, y) \\ &\quad - (f(J_\varepsilon x) + \varepsilon^{-1} K(J_\varepsilon x, x)) \\ &\leqslant \varepsilon^{-1}(K(J_\varepsilon x, y) - K(J_\varepsilon x, x)) \\ &\leqslant \varepsilon^{-1} D_2 K(J_\varepsilon x, x)(y-x) + \varepsilon^{-1} R(J_\varepsilon x, x, y) \end{aligned}$$

with $R(w, x, y) = K(w, y) - K(w, x) - D_2K(w, x)(y-x)$ so that

$$\lim_{\substack{y \to x \\ \neq}} |y-x|^{-1} R(J_\varepsilon x, x, y) = 0.$$

Interchanging the role of x and y we get

$$f_\varepsilon(x) - f_\varepsilon(y) \leqslant \varepsilon^{-1}(K(J_\varepsilon y, x) - K(J_\varepsilon y, y))$$

so that

$$f_\varepsilon(y) - f_\varepsilon(x) - \varepsilon^{-1} D_2K(J_\varepsilon x, x)(y-x) \geqslant \varepsilon^{-1} R(J_\varepsilon y, x, y)$$

with $\lim_{y \to x, \neq} |y-x|^{-1} R(J_\varepsilon y, x, y) = 0$ by our assumption on K and the fact that J_ε is continuous at x. ∎

3.7. *Remark.* When the assumptions of Theorem 3.2 are satisfied, J_ε is continuous, $K(w, x) = k(w-x)$ where $k: E \to \mathbb{R}_+$ is convex and Gâteaux differentiable, the preceding estimates show that f_ε is Gâteaux differentiable on X_ε and in fact is Hadamard differentiable on X_ε with

$$f'_\varepsilon(x) = -\varepsilon^{-1} k'(J_\varepsilon x - x).$$

Let us now tackle the important question of the behavior of the family

(f'_ε) as $\varepsilon \to 0_+$. In order to do so we have to recall that the (lower) subdifferential of f at x (where $f(x)$ is finite) is given as in [19] by

$$\partial f(x) = \{x' \in E' : \forall v \in E\, f^!(x, v) \leqslant \langle x', v \rangle\},$$

where E' is the topological dual of E and $f^!(x, \cdot)$ is given by

$$f^!(x, v) = \liminf_{(t,u) \to (0_+, v)} t^{-1}(f(x + tu) - f(x)).$$

When $f = g + h$ with g convex and h of class C^1 this subdifferential coincides with Clarke's famous strict subdifferential [7].

3.8. THEOREM. *Suppose the assumptions of Theorem 3.2 are in force, with $K(w, x) = k(w - x)$ where $k: E \to \mathbb{R}_+$ is a Gâteaux differentiable convex function, $k(0) = 0$, and J_ε is continuous. Then for each $x \in X$*

(a) *any weak* cluster point of $(f'_\varepsilon(x))_{\varepsilon > 0}$ as $\varepsilon \to 0_+$ belongs to $\partial f(x)$;*

(b) *if moreover $\partial f(x)$ is nonempty and if $k(z) = h(|z|)$ for $z \in E$ where $h: \mathbb{R}_+ \to \mathbb{R}_+$ is convex, strictly increasing, differentiable with $h'(0) = 0$ and the norm of E is Gâteaux-differentiable off 0 then $(f'_\varepsilon(x))_{\varepsilon > 0}$ converges weakly to the element of $\partial f(x)$ with least norm.*

Proof. (a) Let $(x, \varepsilon) \in \hat{X}$ so that $x \in B(\bar{x}, \delta(\bar{x}))$ for some $\bar{x} \in X$ and some $\delta(\bar{x}) \in \mathbb{P}$ as in the proof of Theorem 3.2. Taking $\lambda \in]\varepsilon, \bar{\varepsilon}[$ as before and observing that $F_\lambda(\cdot, x) = f + \lambda^{-1} K(\cdot, x)$ is convex, we observe that $J_\varepsilon x$ is characterized by

$$0 \in (\varepsilon^{-1} - \lambda^{-1})\, D_1 K(J_\varepsilon x, x) + \partial(f + \lambda^{-1} K(\cdot, x))(J_\varepsilon x)$$

or

$$-\varepsilon^{-1} D_1 K(J_\varepsilon x, x) \in \partial f(J_\varepsilon x).$$

Here this can be written

$$-\varepsilon^{-1} k'(J_\varepsilon x - x) \in \partial f(J_\varepsilon x).$$

Now $f'_\varepsilon(x) = -\varepsilon^{-1} k'(J_\varepsilon x - x)$. As the graph of

$$w \mapsto \partial f(w) = \partial F_\lambda(\cdot, \bar{x})(w) - \lambda^{-1} k'(w - \bar{x})$$

is closed in the product topology of the strong topology on X and the $\sigma(E', E)$-topology on E' since this is the case for $\partial F_\lambda(\cdot, \bar{x})$ and k' is continuous, we get that any cluster point x' of $(-\varepsilon^{-1} k'(J_\varepsilon x - x))_{\varepsilon > 0}$ belongs to $\partial f(x)$.

(b) Now let us suppose $k = h \circ N$ where $N: E \to \mathbb{R}_+$ is the norm of E and $h: \mathbb{R}_+ \to \mathbb{R}_+$ is convex and differentiable. Let $\lambda \in]0, \bar{\varepsilon}[$ and let

$\varepsilon \in]0, \lambda[$. As $g: w \mapsto f(w) + \lambda^{-1} k(w - x)$ is convex on $B(\bar{x}, \alpha(\bar{x}))$ for each $x \in B(\bar{x}, \delta(\bar{x}))$, using the monotonicity of ∂g on $B(\bar{x}, \alpha(\bar{x}))$ we can write, for any $x' \in \partial f(x)$, with $x_\varepsilon = J_\varepsilon x$, $x'_\varepsilon = -\varepsilon^{-1} k'(x_\varepsilon - x)$,

$$\langle x' - x'_\varepsilon + \lambda^{-1} k'(0) - \lambda^{-1} k'(x_\varepsilon - x), x - x_\varepsilon \rangle \geqslant 0$$

or, as $k'(0) = 0$

$$\langle x', x - x_\varepsilon \rangle \geqslant (1 - \lambda^{-1}\varepsilon) < x'_\varepsilon, x - x_\varepsilon \rangle.$$

Let us first suppose $x_\varepsilon \neq x$ for ε small enough so that $\langle N'(x_\varepsilon - x), x_\varepsilon - x \rangle = |x_\varepsilon - x|$, $|N'(x_\varepsilon - x)| = 1$.
Then

$$x'_\varepsilon = -\varepsilon^{-1} h'(|x_\varepsilon - x|)\, N'(x_\varepsilon - x)$$

so that

$$\langle x'_\varepsilon, x - x_\varepsilon \rangle = \varepsilon^{-1}\, |x_\varepsilon - x|\; h'(|x_\varepsilon - x|) \leqslant (1 - \lambda^{-1}\varepsilon)^{-1} \langle x', x - x_\varepsilon \rangle.$$

It follows that

$$\limsup_{\varepsilon \to 0_+} |x'_\varepsilon| = \limsup\, \varepsilon^{-1} h'(|x_\varepsilon - x|) \leqslant |x'|.$$

Therefore $(x'_\varepsilon)_{\varepsilon > 0}$ has weak* cluster points as $\varepsilon \to 0_+$. As the norm is weakly* l.s.c. on E', each of these cluster points $\bar{x}'$ satisfies $|\bar{x}'| \leqslant \liminf_{\varepsilon \to 0_+} |x'|$ for each $x' \in \partial f(x)$.

As N is Gâteaux differentiable on $E \backslash \{0\}$, the dual norm is strictly convex, hence the closed convex set $\partial f(x)$ has at most one point with smallest norm. This uniqueness of cluster points ensures that $(x'_\varepsilon)_{\varepsilon > 0}$ converges weakly.

Now if $x_\varepsilon = x$ for ε in a subset Q of $\mathbb{P}$ with 0 in its closure, we have $x'_\varepsilon = -\varepsilon^{-1} k'(0) = 0$ for each $\varepsilon \in Q$, hence $0 \in \partial f(x_\varepsilon) = \partial f(x)$ while the limit of (x'_ε) as $\varepsilon \to 0$, $\varepsilon \in \mathbb{R} \backslash Q$ is 0 by what precedes, so that $(x'_\varepsilon) \to 0$ as $\varepsilon \to 0$ in $\mathbb{P}$. ∎

3.9. *Remark.* When the assumptions of Theorem 3.8(b) are satisfied and when the norm of E' satisfies the condition

(H') if (x'_n) converges weakly to x' and if $(|x'_n|)$ converges to $|x'|$ then (x'_n) converges strongly to x'

then the preceding proof shows that $(f'_\varepsilon(x))$ converges strongly to the element of least norm in $\partial f(x)$.

Let us conclude this section by giving a positive partial answer to a question raised by J.-M. Lasry and P.-L. Lions [15]. For simplicity we suppose

E is a Hilbert space and $K=K_2$ is the usual quadratic kernel. Given $f\colon X\to\mathbb{R}$ and $x\in X$ with $\partial f(x)\neq\varnothing$, we denote by $\partial_0 f(x)$ the element of least norm in $\partial f(x)$.

3.10. THEOREM. *Let $f\colon X\to\mathbb{R}$ be l.s.c., quadratically minorized ($d_F(f)<+\infty$) and such that $c_K(f,x)<+\infty$ for each $x\in X$. Suppose f satisfies the following Palais–Smale condition:*

(C_0) *each sequence (x_n) such that $\partial f(x_n)\neq\varnothing$ and $(\partial_0 f(x_n))\to 0$ has a cluster point.*

Then for each $\varepsilon\in\mathbb{P}$, f_ε satisfies the usual Palais–Smale condition on X_ε:

(C) *each sequence (x_n) such that $(f'_\varepsilon(x_n))_{n\geqslant 0}\to 0$ has a cluster point.*

Let us observe that for $X=E$, $\sup_{x\in X} c_K(f,x)\leqslant c$, we have $X_\varepsilon=E$ for $\varepsilon\in\,]0,c^{-1}[$.

Proof. We have seen that for $x_n\in X_\varepsilon$

$$\nabla f_\varepsilon(x_n)=\varepsilon^{-1}(x_n-J_\varepsilon x_n),$$

where $J_\varepsilon x_n$ is characterized by

$$\varepsilon^{-1}(x_n-J_\varepsilon x_n)\in\partial f(J_\varepsilon x_n).$$

As $(\nabla f_\varepsilon(x_n))_{n\geqslant 0}\to 0$ we get $(\partial_0 f(J_\varepsilon x_n))_{n\geqslant 0}\to 0$ and by our assumption $(J_\varepsilon x_n)_{n\geqslant 0}$ has a converging subsequence $(J_\varepsilon x_k)_{k\in K}$. Since $(|x_k-J_\varepsilon x_k|)_{k\in K}$ converges to 0, $(x_k)_{k\in K}$ has the same limit. ∎

Some higher differentiability results will be found in [21]. Let us here just note an observation showing that even in a simple case some extra assumptions are needed.

It is easy to see that if f is a polyhedral convex function on some open interval X of $\mathbb{R}$ then for each $x\in X$ there exists $\varepsilon>0$ and a neighborhood U of x on which f_ε is of class C^∞. This is no more true in higher dimensions, as shown by the following example.

3.11. EXAMPLE. Let $X=\mathbb{R}^2$, $f(z)=\max(x,y,0)$ for $z=(x,y)\in\mathbb{R}^2$. For any $\varepsilon>0$ let U_ε be the open ball with center $(0,0)$ and radius $\varepsilon(\sqrt{2}/2)$. Then for $z\in U_\varepsilon\cap\mathbb{R}^2_+$ we have $J_\varepsilon z=(0,0)$ as $\varepsilon^{-1}z\in\partial f(0)=\mathrm{co}(0,e_1,e_2)$, where $e_1=(1,0)$, $e_2=(0,1)$, so that $f_\varepsilon(z)=(2\varepsilon)^{-1}(x^2+y^2)$. For $z=(x,y)\in U_\varepsilon\cap(\mathbb{R}_+\times(-\mathbb{P}_+))$, we have $J_\varepsilon z=(0,y)$ since $\varepsilon^{-1}(z-J_\varepsilon z)\in\partial f(J_\varepsilon z)=\mathrm{co}(0,e_1)$, so that $f_\varepsilon(z)=(2\varepsilon)^{-1}x^2$. Similarly, for $z=(x,y)\in U_\varepsilon\cap((-\mathbb{R}_+)\times\mathbb{P}_+)$ we get $f_\varepsilon(z)=(2\varepsilon)^{-1}y^2$. Finally for $z=(x,y)$ with $x\leqslant 0$,

$y \leqslant 0$ we have $J_\varepsilon z = z$, $f_\varepsilon(z) = 0$. Therefore f_ε is of class $C^{1,1}$ on U_ε, with $f_\varepsilon(z) = (2\varepsilon)^{-1}((x^+)^2 + (y^+)^2)$,

$$\nabla f_\varepsilon(x, y) = \varepsilon^{-1}(x^+, y^+),$$

but there is no neighborhood of (0, 0) on which f_ε is of class C^2. ∎

The preceding example enhances the role of the transversality conditions given in [21, Proposition 5.3 and its corollaries] in order that f_ε be of class C^2 around a point. In particular, we observe that in the preceding example, condition (b) of Corollaries 5.7 and 5.8 of [21] is not satisfied although the other conditions are met with $A = \{(0, 0)\}$.

4. Epiconvergence and Approximation by Infimal Convolution

It is well known that the Moreau–Yosida's approximation scheme enables one to reduce the epiconvergence of a family of functions to ordinary pointwise convergence of the families of approximate functions (see [1, 13], for instance). Here we show that this fact remains true when the approximation is given by a general firm kernel. Our proof is a simple direct consequence of the definitions.

Let X be a topological space and let $(f^p)_{p \in P}$ be a family of extended real-valued functions on X indexed by a parameter p belonging to a subset P of a topological space $P^\bullet$. Given a particular point $\bar{\omega}$ of the closure cl P of P in $P^\bullet$ we denote by $\mathscr{Q}$ the trace on P of the family $\mathscr{Q}^\bullet$ of neighborhoods of $\bar{\omega}$ in $P^\bullet$: $\mathscr{Q} = \{Q = Q^\bullet \cap P : Q^\bullet \in \mathscr{Q}^\bullet\}$. Given $h\colon P \to \bar{\mathbb{R}}$ we write $\liminf_p h(p)$ for $\sup_{Q \in \mathscr{Q}} \inf_{p \in Q} h(p)$, omitting the inclusion $p \in P$ and the convergence $p \to \bar{\omega}$. Convergence with respect to a filter $\mathscr{F}$ in P can be set into this familiar framework by adding a "point at infinity" $\bar{\omega}$ to P and putting on $P^\bullet = P \cup \{\bar{\omega}\}$ a topology τ inducing on P the discrete topology and such that $\mathscr{F}^\bullet = \{F^\bullet = F \cup \{\bar{\omega}\} : F \in \mathscr{F}\}$ is the family of neighborhoods of $\bar{\omega}$ in $P^\bullet$. Let us recall that the epilimit inferior and the epi-limit superior of the family $(f^p)_{p \in P}$ are given by

$$\operatorname*{e\,li}_p f^p(x) = \sup_{V \in \mathscr{N}(x)} \liminf_p \inf_{v \in V} f^p(v)$$

$$\operatorname*{e\,ls}_p f^p(x) = \sup_{V \in \mathscr{N}(x)} \limsup_p \inf_{v \in V} f^p(v),$$

where $\mathscr{N}(x)$ is the filter of neighborhoods of x in X. Setting

$$f^Q(x) = \inf_{p \in Q} f^p(x)$$

for a subset Q of P and $x \in X$ we observe that

$$\operatorname*{e\,li}_{p} f^p(x) = \liminf_{\substack{(p,v)\to(\bar\omega,x)\\ p\in P}} f^p(v) = \sup_{Q\in\mathcal{Q}} \overline{f^Q}(x),$$

where $\bar{g}$ is the lower-semicontinuous hull of $g\colon X \to \overline{\mathbb{R}}$ given by

$$\bar{g}(x) = \liminf_{v\to x} g(x) = \sup_{V\in\mathcal{N}(x)} \inf_{v\in V} g(v).$$

The following result relates the preceding epi-limits of the family (f^p) to ordinary pointwise limits of the approximate functions (f^p_ε).

4.1. THEOREM. *For any parametrized family* $(f^p)_{p\in P}$ *of extended real-valued functions on* X *one has for each* $x \in X$

(a) $\operatorname*{e\,li}_{p} f^p(x) \geqslant \sup_{\varepsilon>0} \liminf_{p} f^p_\varepsilon(x),$

(b) $\operatorname*{e\,ls}_{p} f^p(x) \geqslant \sup_{\varepsilon>0} \limsup_{p} f^p_\varepsilon(x).$

If $(f^p)_{p\in P}$ *is a* K*-equiminorized family and if the kernel* K *is firm equality holds in* (a) *and* (b).

Here the family $(f^p)_{p\in P}$ is said to be *K-equiminorized* if for each $x \in X$ there exists b and c in $\mathbb{R}_+$ such that

$$f^p(w) \geqslant b - cK(w, x) \qquad \text{for each} \quad w\in X, \text{each } p \in P.$$

In fact we could assume that this inequality holds true for each $w \in X$ and each p belonging to some member Q of $\mathcal{Q}$.

Proof. (a) As $\bar{g} \geqslant \sup_{\varepsilon>0} g_\varepsilon$ for any $g \in \overline{\mathbb{R}}^X$, in particular for $g = f^Q$, $Q \in \mathcal{Q}$ we have

$$\begin{aligned}
\operatorname*{e\,li}_{p} f^p(x) &= \sup_{Q\in\mathcal{Q}} \overline{f^Q}(x) \geqslant \sup_{Q\in\mathcal{Q}}\ \sup_{\varepsilon>0} f^Q_\varepsilon(x) \\
&= \sup_{\varepsilon>0}\ \sup_{Q\in\mathcal{Q}}\ \inf_{w\in X}\ (\inf_{p\in Q} f^p(w) + \varepsilon^{-1}K(w,x)) \\
&= \sup_{\varepsilon>0}\ \sup_{Q\in\mathcal{Q}}\ \inf_{p\in Q}\ \inf_{w\in X}\ (f^p(w) + \varepsilon^{-1}K(w,x)) \\
&= \sup_{\varepsilon>0}\ \liminf_{p} f^p_\varepsilon(x).
\end{aligned}$$

When $(f^p)_{p\in P}$ is K-equiminorized, for each $x \in X$ we can find b and c such that

$$f^p(w) \geqslant b - cK(w,x) \qquad \text{for each} \quad (w,p) \in X \times P$$

so that for each $Q \in \mathcal{Q}$ (or each $Q \in \mathcal{Q}$ contained in some $Q_0 \in \mathcal{Q}$)

$$f^Q(w) \geqslant b - cK(w, x) \qquad \text{for each} \quad w \in X.$$

Then, if K is firm, Proposition 2.2 asserts that

$$\overline{f^Q}(x) = \sup_{\varepsilon > 0} f^Q_\varepsilon(x)$$

so that equality holds everywhere above.

(b) In order to prove the announced inequality it suffices to show that for any $\varepsilon > 0$ and any $r \in \mathbb{R}$ such that $r > \text{e ls}_p f^p(x)$ one has

$$r > \limsup_p f^p_\varepsilon(x)$$

since we may suppose $\text{e ls}_p f^p(x) < +\infty$. Let $\alpha > 0$ be such that $r - \alpha > \text{e ls}_p f^p(x)$ and let $U \in \mathcal{N}(x)$ be such that $K(u, x) \leqslant \varepsilon\alpha$ for each $u \in U$. Then

$$\begin{aligned}
\limsup_p f^p_\varepsilon(x) &\leqslant \limsup_p \inf_{u \in U} (f^p(u) + \varepsilon^{-1} K(u, x)) \\
&\leqslant \limsup_p \inf_{u \in U} f^p(u) + \alpha \\
&\leqslant \sup_{V \in \mathcal{N}(x)} \limsup_p \inf_{v \in V} f^p(v) + \alpha \\
&= \operatorname*{e\,ls}_p f^p(x) + \alpha < r
\end{aligned}$$

Now let us prove the opposite inequality when K is firm and (f^p) is K-equiminorized. We may suppose $\text{e ls}\, f^p(x) \neq -\infty$. Let $r \in \mathbb{R}$ be such that $r < \text{e ls}_p f^p(x)$. By definition of the epi-limit superior we can find $U \in \mathcal{N}(x)$ such that

$$r < \limsup_p \inf_{u \in U} f^p(u).$$

Let b and c in $\mathbb{R}_+$ be such that $f^p \geqslant b - cK(\cdot, x)$ for each $p \in P$. As K is firm we can find $\varepsilon_U \in]0, c^{-1}[$ such that $(\varepsilon^{-1} - c) K(w, x) \geqslant r - b$ for each $w \in X \backslash U$, $\varepsilon \in]0, \varepsilon_U[$. As

$$r < \inf_{Q \in \mathcal{Q}} \sup_{p \in Q} \inf_{u \in U} f^p(u)$$

for each $Q \in \mathcal{Q}$ we can find $q \in Q$ such that

$$r < \inf_{u \in U} f^q(u).$$

As for $w \in X \setminus U$, $\varepsilon \in]0, \varepsilon_U[$ we have

$$f^q(w) + \varepsilon^{-1}K(w, x) \geqslant b - cK(w, x) + \varepsilon^{-1}K(w, x) \geqslant r$$

we get

$$f_\varepsilon^q(x) \geqslant \min(r, \inf_{u \in U} f^q(u) + \varepsilon^{-1}K(u, x)) \geqslant r.$$

Therefore, for $\varepsilon \in]0, \varepsilon_U[$

$$\limsup_p f_\varepsilon^p(x) \geqslant r. \quad \blacksquare$$

Let $f^0 \in \bar{\mathbb{R}}^X$ and let $(f^p)_{p \in P} \subset \bar{\mathbb{R}}^X$; we define a family (f^q) of extended real-valued functions on X parametrized by $P \cup \{\bar{\omega}\}$ by setting $f^{\bar{\omega}} = f^0$ and following [25] declare that this extended family is *epi-l.s.c.* (resp. *epi-u.s.c.*) at ($\bar{\omega}$ and) x if $\mathrm{eli}_p f^p(x) \geqslant f^0(x)$ (resp. $\mathrm{els}_p f^p(x) \leqslant f^0(x)$).

Then the main assertion of Theorem 4.1 can be rephrased as follows.

4.2. Corollary. *Suppose K is firm and that the family $(f^p)_{p \in P}$ is K-equiminorized. For any $f^0 \in \bar{\mathbb{R}}^X$ the extended family $(f^q)_{q \in P \cup \{\omega\}}$ is epi-l.s.c. (resp. epi-u.s.c.) at x iff*

$$\sup_{\varepsilon > 0} \liminf_p f_\varepsilon^p(x) \geqslant f^0(x)$$

(*resp.* $\sup_{\varepsilon > 0} \limsup_p f_\varepsilon^p(x) \leqslant f^0(x)$).
In particular $(f^p)_{p \in P}$ epi-converges at x iff

$$\sup_{\varepsilon > 0} \liminf_p f_\varepsilon^p(x) = \sup_{\varepsilon > 0} \limsup_p f_\varepsilon^p(x).$$

Therefore the family $(f^p)_{p \in P}$ epi-converges at x whenever for each $\varepsilon > 0$, $(f_\varepsilon^p(x))_{p \in P}$ converges.

References

1. H. Attouch, "Variational Convergence for Functions and Operators," Applicable Math. Series, Pitman, London, 1984.
2. H. Attouch and R. J.-B. Wets, A convergence theory for saddle functions, *Trans. Amer. Math. Soc.* **280** No. 1 (1983), 1–41.
3. D. P. Bertsekas, Convexification procedures and decomposition methods for nonconvex optimization problems, *J. Optim. Theory Appl.* **29** (1979), 169–197.
4. M. L. Bougeard, "Contribution à la Théorie de Morse en dimension finie," Thèse de 3ème cycle, Université Paris-Dauphine, Paris, 1978.
5. M. L. Bougeard, Morse theory for some lower-C^2 functions in finite dimensions, *Math. Programming* **41**, No. 2 (1988), 141–169.

6. C. CASTAING, A propos de l'existence des sections séparément mesurables et séparement semi-continues d'une multiapplication séparément semi-continue inférieurement, *in* "Travaux du Séminaire d'Analyse Convexe, Montpellier, 1976," Exposé 6, pp. 6.1–6.6.
7. F. H. CLARKE, "Optimization and Nonsmooth Analysis," Wiley, New York, 1983.
8. E. DE GIORGI, Convergence problems for functionals and operators, *in* "Proceedings, Int. Meeting on Recent Methods in Nonlinear Analysis" (E. De Giorgi, E. Magenes, and U. Mosco, Eds.), pp. 131–188, Roma, 1978, Pitagora, Bologna, 1979.
9. S. DOLECKI, On inf-convolutions, generalized convexity and seminormality, preprint, Université de Limoges, 1985.
10. I. EKELAND, On the variational principle, *J. Math. Anal. Appl.* **47** (1974), 324–353.
11. I. EKELAND AND J. M. LASRY, On the number of closed trajectories for a Hamiltonian flow on a convex energy surface, *Ann. of Math* **112** (1980), 283–319.
12. S. FITZPATRICK, Metric projections and the differentiability of distance functions, *Bull. Austral. Math. Soc.* **22**, No. 2 (1980), 291–312.
13. A. FOUGERES AND A. TRUFFERT, "Régularisation s.c.i. et Γ-convergence: Approximations inf-convolutives associées à un référentiel (version augmentée)," Publications AVAMAC, Université de Perpignan 84-08/15, 1984.
14. J.-B. HIRIART-URRUTY, Extension of Lipschitz functions, *J. Math. Anal. Appl.* **77**, No. 2 (1980), 539–554.
15. J.-M. LASRY AND P. L. LIONS, Remark on regularization in Hilbert spaces, *Israël J. Math.* **55**, No. 3 (1986), 257–266.
16. K. NOMIZU AND H. OZEKI, The existence of complete Riemannian metrics, *Proc. Sympos. Pure Math. Soc.* **12** (1961), 889–891.
17. R. S. PALAIS, Critical point theory and the minimax principle, *in* "Global Analysis" (S. Chern and S. Smale, Eds.), pp. 185–212, Proc. Symp. Pure Math. Vol. 15, Amer. Math. Soc., Providence, RI, 1970.
18. J.-P. PENOT, Calcul sous-différentiel et optimisation, *J. Funct. Anal.* **27** (1978), 248–276.
19. J.-P. PENOT, A characterization of tangential regularity, *Nonlinear Anal.* **5**, No. 6 (1981), 625–643.
20. J.-P. PENOT, On favorable classes of mappings in nonliner analysis and optimization, preprint, Univ. Pau, 1985.
21. J.-P. PENOT AND M.-L. BOUGEARD, Approximation and decomposition properties of some classes of locally d.c. functions, *Math. Programming* **41**, No. 2 (1988), 195–228.
22. A. POMMELLET, "Analyse convexe et théorie de Morse," Thèse de 3ème cycle, Université Paris IX, 1982.
23. R.-T. ROCKAFELLAR, Integral functionals, normal integrands and measurable selections, *in* "Nonlinear Operators and the Calculus of Variations" (L. Waelbroeck, Ed.), pp. 157–207, Lecture Notes in Math. Vol. 543, Springer-Verlag, Berlin/New York, 1976.
24. R.-T. ROCKAFELLAR, Favorable classes of Lipschitz continuous functions in subgradient optimization, *in* "Progress in Nondifferentiable Optimization" (E. Nurminski, Ed.), pp. 125–143, I.I.A.S.A., Laxemburg, Australia, 1982.
25. R.-T. ROCKAFELLAR AND R. J.-B. WETS, Variational systems, an introduction, pp. 1–54, *in* "Multifunctions and Integrands" (G. Salinetti, Ed.), Vol. 1091, Lecture Notes in Mathematics, Springer-Verlag, Berlin/New York, 1984.
26. M. VOLLE, Conjugaisons par tranches, *Ann. Mat. Pura Appl.* **139** (1985), 279–312.
27. M. VOLLE, Quelques résultats relatifs à l'approche par les tranches de l'épi-convergence, submitted for publication.
28. R. WETS, Convergence of convex functions, variational inequalities and convex optimization problems, *in* "Variational Inequalities and Complementary Problems" (P. Cottle *et al.*, Eds.), Wiley, Chichester, 1980.