Proceedings of the Royal Society of Edinburgh, **131A**, 945–966, *2001*

On the Yosida approximation of operators

Jean-Paul Penot and **Robert Ratsimahalo**
Laboratoire de Mathématiques Appliquées UPRES-A 5033 CNRS, Université de Pau et des Pays de l'Adour, Av. de l'Université, 64000 Pau, France (jean-paul.penot@univ-pau.fr)

(MS received 23 November 1999; accepted 18 May 2000)

A generalized Yosida approximation of monotone (and non-monotone) operators in Banach space is introduced. It uses a general potential that is not necessarily the square of the norm. It is therefore advisable to use it in cases where some other more convenient potentials are available, such as in L_p-spaces. As an illustration, the case of Nemyckii operators is considered.

1. Introduction

It is well known that monotone operators on Hilbert spaces can be regularized into single-valued Lipschitzian monotone operators via a process known as the Yosida approximation (see [5, 6, 10, 14, 15, 40]). In [5, 16–21, 31, 41, 48, 52, 55], an extension of this process is given in Banach spaces satisfying some assumptions.

The purpose of the present paper is not just a search for minimal assumptions for this process. Having in view that in many concrete situations the usual duality mapping is difficult to handle, we introduce a regularization process using a general duality mapping associated with a potential. It appears that most known properties are still valid when one uses a potential that is not necessarily the square of the norm. In an L_p-space (or in a Sobolev space $W^{m,p}(\Omega)$), the function $p^{-1}\|.\|^p$ is more appropriate than the square of the norm and the use of such a potential enables one to avoid technicalities when writing down in concrete terms the relations involved by the regularization process.

Another incentive to use a general potential arises from the regularization of functions. For such a purpose, it may be convenient to use a general potential; it is even compulsory if the function has a severe decrease at infinity (see [12]).

We point out in § 3 that the general regularization process we define is a special case of what is called a parallel sum. Such an operation, which has been introduced in connection with the modelling of electrical networks, plays an important role in nonlinear analysis and convexity theory (see [3, 7, 33–37, 39]). Sufficient conditions ensuring that the regularization has the whole space for its domain are presented. They rely on surjectivity results involving maximal monotone operators in reflexive and non-reflexive spaces that are of independent interest. We devote § 4 to the study of basic properties of this regularization, such as single valuedness, continuity properties and preservation of monotonicity.

Section 5, which deals with the links between the Moreau–Yosida regularization of functions and the regularization of their subdifferentials, illustrates the abovementioned fact that these processes are not limited to the convex (respectively, monotone) case. Here we use a general subdifferential satisfying appropriate properties.

In § 6 we tackle the case in which the space is non-reflexive but is a dual Banach space. Introducing a density criterion for an appropriate topology, we get a surjectivity result and a condition in order that the regularization is defined on the whole space. We compare our results with previous ones obtained by Gossez [28], in which the topology is different.

We close the paper by a study of the regularization of Nemyckii operators on L_p-spaces. In such a case, the regularization process is simply derived from a pointwise regularization. Such a result cannot be obtained with the classical regularization when $p \neq 2$.

2. Preliminaries

Since we will make strong use of duality mappings with weights, let us recall briefly some basic facts about such mappings (see also [4, 20, 23, 31, 54]). Let h be a continuous and non-decreasing function from the set $\mathbb{R}_+$ of non-negative real numbers into $\mathbb{R}_+$ such that $h(0) = 0$ and $\lim_{t\to+\infty} h(t) = +\infty$ (or, equivalently, $\lim_{t\to+\infty} t^{-1}H(t) = +\infty$, where $H(t) = \int_0^t h(s)\,\mathrm{d}s$). Note that H is a convex function. The multimapping J_h from a Banach space X into its topological dual X^*, given by

$$J_h(x) = \{x^* \in X^* : \langle x^*, x\rangle = \|x\| \cdot \|x^*\|,\ \|x^*\| = h(\|x\|)\},$$

is called the *duality mapping of weight* (or gauge) h.

In the sequel, we denote by ∂ the subdifferential in the sense of convex analysis: for an extended real-valued function f on X finite at x,

$$\partial f(x) = \{x^* \in X^* : f(y) \geqslant f(x) + \langle x^*, y - x\rangle\ \forall y \in X\};$$

although this definition is valid for an arbitrary function, its use is generally limited to the case f is convex. It is known that when f is a lower semicontinuous (LSC) proper convex function, ∂f is a maximal monotone operator from X into X^*. When X is the dual of some Banach space X_*, we set

$$\partial_* f(x) = \{x_* \in X_* : f(y) \geqslant f(x) + \langle x_*, y - x\rangle\ \forall y \in X\}.$$

If, moreover, f is LSC for the weak-star topology, proper and convex, $\partial_* f$ is a maximal monotone operator from X into X_* as f is the conjugate of an LSC proper convex function f_* on X_* and $\partial_* f = (\partial f_*)^{-1}$. For the concept of maximal monotone operatosr, we refer to [10, 14, 19, 24, 55] for instance.

The duality mapping J_h has close similarities with the classical duality mapping J given by $J(\cot) = \partial(\frac{1}{2}\|\cdot\|^2)$ (i.e. the duality mapping associated with the identity function on $\mathbb{R}_+$). In particular, one has the following characterization.

Lemma 2.1 (cf. [1, 18, 19, 54]). *The multimapping J_h satisfies $J_h = \partial j_h$, where $j_h = H \circ \|\cdot\|$.*

Let us note that $J_h(\cdot) = \|\cdot\|^{-1}h(\|\cdot\|)J(\cdot)$ on $X \setminus \{0\}$, as it is easily seen from the very definition, or from the preceding characterization by taking directional derivatives. As $J_h(0) = \{0\}$, the multimapping J_h is single valued on X if and only if the norm of X is Gâteaux differentiable on $X \setminus \{0\}$.

Although J_h is similar to J, it may have some specific advantages. For example, when $X = L_p(\Omega)$, with $1 < p < +\infty$, $p \neq 2$, where $(\Omega, \mathcal{A}, \mu)$ is a measured space, it can be convenient to take $h(t) = t^{p-1}$ instead of $h(t) = t$. In this case, the duality mapping has the simple form

$$J_h(x)(\omega) = |x(\omega)|^{p-2} x(\omega),$$

which does not involve any integration on Ω.

LEMMA 2.2. *The multimapping J_h is onto if and only if X is reflexive.*

Proof. Suppose X is reflexive. Then, for any $x^* \in X^*$, there exists some $x \in X$, $\|x\| = 1$, such that $\langle x^*, x\rangle = \|x^*\|$. Since for $s := \|x^*\|$ there exists some $t \geqslant 0$ such that $h(t) = s$, taking $y = tx$, we get $x^* \in J_h(y)$. Conversely, by James' theorem (see, for instance, [11, 25]), it suffices to show that, for any $x^* \in X^*$ such that $\|x^*\| = 1$, there exists some $x \in X$ satisfying $\langle x^*, x\rangle = \|x^*\|$, $\|x\| = 1$. As J_h is onto, we can find some $u \in X$ such that $\langle x^*, u\rangle = \|u\|$ and $h(\|u\|) = 1$. We cannot have $\|u\| = 0$, hence $x := \|u\|^{-1} u$ is well defined and fulfils the requirements. □

In the sequel, we identify a multimapping with its graph; such an identification is convenient, but when using operations on images such as sums, we define $F + G$ as $(F + G)(x) = F(x) + G(x)$ and $(rF)(x) = rF(x)$ when F and G are two multimappings with values in some vector space and $r \in \mathbb{R}$. The domain of a multimapping M is the set $\operatorname{dom} M := \{x : M(x) \neq \emptyset\}$.

The following result from [19, proposition 1] will be useful. We present a short proof for the reader's convenience.

LEMMA 2.3. *The multimapping J_h is upper semicontinuous from X endowed with the strong topology into X^* equipped with the weak-star topology (and even for X^* endowed with the bounded weak-star topology $\sigma_b(X^*, X)$).*

Proof. The result is valid for the subdifferential of any continuous convex function f and not just for $J_h = \partial j_h$. Let V be an open subset of X^* for the bounded weak-star topology such that $\partial f(x) \subset V$. Suppose there exists a sequence $(x_n) \subset X$ strongly convergent to x such that $\partial f(x_n) \cap (X \setminus V) \neq \emptyset$. Let $x_n^* \in \partial f(x_n) \cap (X \setminus V)$. As f is continuous, there exists a weak-star closed bounded subset B of X^* such that $\partial f(x_n) \subset B$ for each n. Since B is (weak-star) compact, there exists a cluster point $x^* \in B$ of x_n^*; as $B \setminus V$ is closed in B endowed with the weak-star topology, we have $x^* \in B \setminus V$. Then, by the continuity of the coupling function on $B \times X$, for each $u \in X$, we get

$$\langle x^*, x - u\rangle \geqslant \liminf_n \langle x_n^*, x_n - u\rangle \geqslant \liminf_n f(x_n) - f(u) \geqslant f(x) - f(u)$$

and $x^* \in \partial f(x)$; a contradiction. □

Using an argument similar to the one given in [55] for instance, we see that the following properties hold.

LEMMA 2.4. *For any $(x, x^*) \in J_h$, $(y, y^*) \in J_h$, one has*

$$\langle x^* - y^*, x - y \rangle \geqslant (\|x\| - \|y\|)(h(\|x\|) - h(\|y\|)).$$

Proof. It suffices to use the definition of J_h and to apply the inequality $\langle u^*, u \rangle \leqslant \|u^*\| \|u\|$ for $(u, u^*) = (x, x^*), (y, y^*)$. □

COROLLARY 2.5. *If the Banach space X is strictly convex and h is increasing, then the multimapping J_h is strictly monotone. For any $x, y \in X$, and any $x^* \in J_h(x)$, $y^* \in J_h(y)$, one has*

$$\langle x^* - y^*, x - y \rangle \leqslant 0 \quad \Rightarrow \quad x = y.$$

Proof. Suppose the Banach space X is strictly convex. Let $x, y \in X$, $(x^*, y^*) \in J_h(x) \times J_h(y)$ be such that

$$\langle x^* - y^*, x - y \rangle \leqslant 0.$$

Then, taking $z^* \in J_h(z)$ with $z = (\frac{1}{2})(x + y)$, one has

$$0 = \langle x^* - z^*, x - z \rangle + \langle y^* - z^*, y - z \rangle.$$

Using lemma 2.4, it follows that

$$0 \geqslant (\|x\| - \|z\|)(h(\|x\|) - h(\|z\|)) + (\|y\| - \|z\|)(h(\|y\|) - h(\|z\|)).$$

Since the weight h is increasing, one gets $\|x\| = \|z\| = \|y\|$, and since the Banach space X is strictly convex, one concludes that $x = y$. □

3. Definition of the regularization process

We devote the present section to the formal definition of the regularization process we have in view. It stems from the notion of the parallel sum of two multimappings. Recall that if M and N are two multimappings from an additive semi-group X (or a group or a vector space) into a set Y, their parallel sum $M \parallel N$ is defined as

$$M \parallel N = (M^{-1} + N^{-1})^{-1}.$$

This notion has been introduced for the needs of the study of electrical networks [3, 33–37]; it is also connected with the notion of harmonic means of real numbers.

DEFINITION 3.1. The generalized Moreau–Yosida regularization of the multimapping M from X into X^* (associated with the weight h) is given for $\lambda > 0$ by

$$M_\lambda := (M^{-1} + \lambda J_h^{-1})^{-1}.$$

Let us note that such an operator is always well defined whatever M is.

DEFINITION 3.2. The multimapping M is said to be *regularizable* (respectively, eventually regularizable) if, for any $\lambda > 0$ (respectively, if, for $\lambda > 0$ small enough), the multimapping $M^{-1} + \lambda J_h^{-1}$ is onto, i.e.

$$(M^{-1} + \lambda J_h^{-1})(X^*) = X,$$

or, equivalently, the domain of M_λ is the whole space X.

EXAMPLE 3.3. The multimapping J_h is regularizable since, for any $\lambda > 0$, $w \in X$, one has

$$(1+\lambda)J_h^{-1}(X^*) = X \quad \text{and} \quad (J_h)_\lambda(w) = J_h((1+\lambda)^{-1}w).$$

This is also the case of J_k, where k is another weight such that $k(t) > 0$ for $t > 0$. As h, k are non-decreasing continuous functions from $\mathbb{R}_+$ into $\mathbb{R}_+$, $h(0) = 0$, $k(t) > 0$ for $t > 0$, for any $w \in X \setminus \{0\}$, one can find a positive real number $s \leqslant \|w\|\lambda^{-1}$ such that

$$h(s) = k(-\lambda s + \|w\|).$$

Then, by setting $q := \|w\|s^{-1} - \lambda$, $z := (q+\lambda)^{-1}w$, $y := qz$, we get

$$k(\|y\|) = k\left(\frac{q}{q+\lambda}\|w\|\right) = h\left(\frac{\|w\|}{q+\lambda}\right) = h(\|z\|).$$

Thus one has

$$J_k(y) = \frac{k(\|y\|)}{\|y\|}J(y) = \frac{k(\|y\|)}{\|z\|}J(z) = \frac{h(\|z\|)}{\|z\|}J(z) = J_h(z),$$

and, for any $x^* \in J_k(y) \cap J_h(z)$, we get $w = y + \lambda z \in (J_k^{-1} + \lambda J_h^{-1})(x^*)$. Let us also note that

$$(J_k)_\lambda(w) = \bigcup_{\substack{y,z \\ y+\lambda z = w}} J_k(y) \cap J_h(z).$$

We will give below sufficient conditions ensuring that the condition of the preceding definition is satisfied. We first connect this definition with the notion of a resolvent.

DEFINITION 3.4. The resolvent or proximal multimapping P_λ^M from X into X associated with the multimapping M (and the weight h) is given, for $\lambda > 0$ and $w \in X$, by

$$P_\lambda^M(w) = \left\{x \in X : 0 \in J_h\left(\frac{x-w}{\lambda}\right) + M(x)\right\}. \tag{3.1}$$

The terminology we adopt arises from the case $M = \partial i_C$, where C is a non-empty closed convex subset of X and i_C is the indicator function on C given by $i_C(x) = 0$ if $x \in C$, ∞ otherwise. The multimapping $P_\lambda^M(w)$ is then the metric projection of w onto C. In fact, whenever f is an LSC proper convex function on X and $M = \partial f$, $P_\lambda^M(w)$ is the set of minimizers of

$$x \mapsto f(x) + \lambda j_h\left(\frac{x-w}{\lambda}\right).$$

Indeed, since the function $j_{w,\lambda}(x) := \lambda j_h((x-w)/\lambda)$ is finite, convex and continuous, the usual subdifferential calculus rules ensure that

$$0 \in \partial(f + j_{w,\lambda})(x) = \partial f(x) + \partial j_{w,\lambda}(x) = \partial f(x) + J_h\left(\frac{x-w}{\lambda}\right)$$

is a necessary and sufficient condition for x to be a minimizer of $f + j_{w,\lambda}$ (see, for example, [48, 50], and for recent monographs, [8, 30]). The multimapping $P_\lambda^{\partial f}$ is then the Moreau–Yosida proximal operator of $M = \partial f$ [32].

Since $J_h(-v) = -J_h(v)$, the definition of the resolvent P_λ^M can be reformulated as follows:

$$P_\lambda^M(w) = \left\{ x \in X : J_h\left(\frac{w-x}{\lambda}\right) \cap M(x) \neq \emptyset \right\}.$$

This reformulation yields the connection between M_λ and the resolvent P_λ^M we announced.

Proposition 3.5. *For any $\lambda > 0$, one has $\operatorname{dom} M_\lambda = \operatorname{dom} P_\lambda^M$, and, for any $w \in X$, one has*

$$M_\lambda(w) = \bigcup_{x \in P_\lambda^M(w)} J_h\left(\frac{w-x}{\lambda}\right) \cap M(x).$$

Proof. Given $w \in \operatorname{dom} P_\lambda^M$ and $x \in P_\lambda^M(w)$, there exists

$$x^* \in J_h\left(\frac{w-x}{\lambda}\right) \cap M(x).$$

For any such x^* one has $x \in M^{-1}(x^*)$ and $\lambda^{-1}(w-x) \in J_h^{-1}(x^*)$. Hence, one has

$$w \in (M^{-1} + \lambda J_h^{-1})(x^*) \quad \text{or} \quad x^* \in (M^{-1} + \lambda J_h^{-1})^{-1}(w).$$

Thus $x^* \in M_\lambda(w)$. Conversely, if $x^* \in M_\lambda(w)$, then there exists $x \in M^{-1}(x^*)$, $v \in J_h^{-1}(x^*)$ with $w = x + \lambda v$. Thus $x^* \in J_h(\lambda^{-1}(w-x)) \cap M(x)$ and $x \in P_\lambda^M(w)$. □

Under a differentiability assumption, one recovers a familiar formula.

Corollary 3.6. *Suppose J_h is single valued. Then, for any $\lambda > 0$, $w \in X$, one has*

$$M_\lambda(w) = J_h(\lambda^{-1} w - \lambda^{-1} P_\lambda^M(w)).$$

Proof. It suffices to prove that if $w \in \operatorname{dom} P_\lambda^M$, $x \in P_\lambda^M(w)$, $x^* \in J_h(\lambda^{-1}w - \lambda^{-1}x)$, then $x^* \in M_\lambda(w)$. Let $v^* \in J_h(\lambda^{-1}(w-x)) \cap M(x)$, which is non-empty by the preceding characterization of $P_\lambda^M(w)$. Then, as $J_h(\lambda^{-1}(w-x))$ is a singleton, we must have $x^* = v^*$, so that

$$x^* \in J_h(\lambda^{-1}(w-x)) \cap M(x) \subset M_\lambda(w).$$

□

The regularization process we study is not limited to monotone operators, although its usual application concerns this class of operators. As an illustration of this fact, let us consider the following example.

Example 3.7. Let $X = \mathbb{R}$ and let M be the multimapping from X into X^* given by

$$M(x) := \{y : y^3 - y^2 = x\},$$

so that the domain of M is the whole space, but its values may contain one, two or three elements and M is not monotone. Taking h as given by $h(t) = t^{p-1}$ with $p > \frac{4}{3}$ so that $J_h^{-1}(y) = |y|^{q-2}y$ with $p^{-1} + q^{-1} = 1$, we get that M is regularizable and, for $\lambda > 0$,

$$M_\lambda(x) = \{y : y^3 - y^2 + \lambda |y|^{q-2} y = x\}.$$

Moreover, for $\lambda \geqslant 1$, the operator M_λ is monotone and single valued.

EXAMPLE 3.8. For X and h as in the preceding example, let M be given for $u \in \mathbb{R}$ by

$$M(u) = \begin{cases} 1 & \text{if } u \leqslant 0, \\ [-1,1] & \text{if } u = 0, \\ -1 & \text{if } u \geqslant 0. \end{cases}$$

The multimapping M is not monotone but is regularizable. Its regularization M_λ is given for $w \in \mathbb{R}$, $p \in]1, +\infty[$ by

$$M_\lambda(w) = \begin{cases} 1 & \text{if } w \leqslant -\lambda, \\ \{|w/\lambda|^{p-2}(w/\lambda)\} \cup \{-1\} \cup \{1\} & \text{if } w \in [-\lambda, \lambda], \\ -1 & \text{if } w \geqslant \lambda. \end{cases}$$

For monotone operators, one has the following necessary condition for an operator to be regularizable. It is similar to the classical one (see, for instance, [10, theorem 1.2 p. 39], [17, lemma 1.1] and [49, corollary to proposition 1, p. 78]). For completeness, we present a simple proof, as here h may be different from the identity mapping.

PROPOSITION 3.9. *Suppose X is reflexive, strictly convex, with X^* strictly convex and h is increasing. If M is monotone and regularizable, then M is maximal monotone.*

Proof. Let $(x, x^*) \in X \times X^*$ be such that

$$\langle u^* - x^*, u - x \rangle \geqslant 0 \quad \forall (u, u^*) \in M. \tag{3.2}$$

By lemma 2.2, we can find $x' \in X$ such that $x^* \in J_h(x')$. By hypothesis, taking $\lambda > 0, w := x + \lambda x' \in X = \operatorname{dom} P_\lambda^M$ in relation (3.1), there exist $(z, z^*) \in M$, $z' \in J_h^{-1}(z^*)$ such that

$$\lambda z' + z = w = \lambda x' + x.$$

Replacing (u, u^*) by (z, z^*) in equation (3.2), we obtain

$$\langle z^* - x^*, z' - x' \rangle = \lambda^{-1} \langle z^* - x^*, x - z \rangle \leqslant 0. \tag{3.3}$$

Since X is strictly convex, corollary 2.5 implies $z' = x'$, and since J_h is single valued as X is reflexive and X^* is strictly convex, $x^* = z^*$, so that we have $(x, x^*) \in M$ and M is maximal monotone. ∎

We will deduce a sufficient condition from the following fundamental surjectivity result. Here we say that an operator T is *bounding* if it maps bounded subsets into bounded subsets; following [19, theorem 1], we say that T is *coercive* if there exists a function $c : \mathbb{R}_+ \to \mathbb{R}$ satisfying $\lim_{r \to +\infty} c(r) = \infty$ such that, for each $(x, x^*) \in T$, one has

$$\langle x^*, x \rangle \geqslant c(\|x\|)\|x\|.$$

THEOREM 3.10 (cf. theorem 2$'$ of [19]). *Let X be a reflexive Banach space. Let S and T be two maximal monotone operators from X into X^* such that*

(i) $0 \in \operatorname{dom} S$;

(ii) T *is upper semicontinuous from* X *into* X^**, endowed with the bounded weak-star topology, bounding and coercive with* $\operatorname{dom} T = X$.

Then $S + T$ *is surjective.*

A surjectivity result using J_h can readily be derived from the preceding statement. Note that here we do not make any assumption of strict convexity as in [5,10,24], [49, proposition 1] and [55, theorem 32F].

THEOREM 3.11. *Suppose* X *is reflexive and the multimapping* M *of* X *into* X^* *is maximal monotone. Then, for each* $\lambda > 0$*, the operator* $M + \lambda J_h$ *is surjective.*

Proof. Without loss of generality, we suppose $\lambda = 1$. Given $u \in \operatorname{dom} M$ (such a point exists, as $M \neq \emptyset$, M being maximal monotone), let us set

$$S(x) := M(u + x), T(x) := J_h(u + x).$$

Then $0 \in \operatorname{dom} S$, the multimappings S, T are maximal monotone, and T is upper semicontinuous and bounding with $\operatorname{dom} T = X$. The coercivity of T is a consequence of the following estimate in which $x \in X$, $x^* \in T(x)$,

$$\begin{aligned} \langle x^*, x \rangle &\geqslant \langle x^*, u + x \rangle - \|x^*\| \|u\| \\ &\geqslant h(\|u + x\|)\|u + x\| - h(\|u + x\|)\|u\| \\ &\geqslant c(\|x\|)\|x\|, \end{aligned}$$

where $c(0) = 0$ and, for $r > 0$,

$$c(r) := \inf\{h(\|u + x\|)\|x\|^{-1}(\|u + x\| - \|u\|) : x \in X,\ \|x\| = r\}$$

and $c(r) \to \infty$ as $r \to \infty$. Then the surjectivity of $M + \lambda J_h$ is a consequence of the preceding theorem. The result can also be deduced from [21, theorem 7.2] or [49, theorem 1]. □

An immediate consequence of this result can be phrased as follows.

PROPOSITION 3.12. *Suppose* X *is reflexive. If the multimapping* M *of* X *into* X^* *is maximal monotone, then* M *is regularizable: for any* $\lambda > 0$, $\operatorname{dom} P_\lambda^M$ *is the whole space* X.

Proof. It suffices to observe that when M is maximal monotone, for any $\lambda > 0$, $w \in X$, the multimapping $M_{w,\lambda}$ given by $M_{w,\lambda}(u) := M(\lambda u + w)$ is maximal monotone. By theorem 3.11, one has $R(M_{w,\lambda} + J_h) = X^*$. Thus one has

$$P_\lambda^M(w) = \{x \in X : x = \lambda u + w,\ 0 \in J_h(u) + M_{w,\lambda}(u)\} \neq \emptyset.$$

□

Let us recall (see, for example, [11,25]) that any reflexive Banach space X can be renormed in such a way that X is strictly convex with a strictly convex dual norm satisfying the Kadec–Klee property (with such a norm, the space X is also said to be a C-normed space; see [5]). It is often considered that such norms are convenient for the regularization of maximal monotone operators; propositions 4.4 and 4.5 below justify this opinion. However, such renormings may be unrelated to the nature of the problem, and keeping the original norm for the regularization may be advisable.

4. Some properties of regularized multimappings

In some situations, we present in this section basic properties of the regularized multimapping M_λ of M. We will first prove that the semi-group property holds without any assumption. In order to do so, let us recall a few elementary facts. Given a non-decreasing function $h : \mathbb{R}_+ \to \mathbb{R}_+$, a non-decreasing function $k : \mathbb{R}_+ \to \mathbb{R}_+$ is a *quasi-inverse* of h if it satisfies

$$h^\flat(r) := \inf\{r \in \mathbb{R} : h(r) \leqslant s\} \leqslant k(s) \leqslant h^\sharp(r) := \sup\{r \in \mathbb{R} : h(r) \leqslant s\}$$

(see [47]). Given a convex function f from X into $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$, its conjugate f^* from X^* into $\bar{\mathbb{R}}$ is defined by

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

We need the following lemma of independent interest.

Lemma 4.1. *Let k be a quasi-inverse of h. Then J_h^{-1} is the co-restriction to X of the multimapping J_k^* from X^* into X^{**}:*

$$J_h^{-1}(x^*) = J_k^*(x^*) \cap X.$$

In particular, $J_h^{-1}(x^*)$ is closed convex.

Proof. Since j_h is proper convex continuous, we have (see [10, 20] and [50, corollary 12A] for instance)

$$J_h^{-1} = (\partial j_h)^{-1} = \partial_* j_h^*,$$

where j_h^* is the conjugate of j_h. Now, by [8, proposition IV 7], $j_h^*(x^*) = H^*(\|x^*\|)$, where $\|x^*\|$ is the dual norm of x^*; using [47, proposition 4.7] for instance, one has

$$H^*(t) = \int_0^t k(s)\,\mathrm{d}s,$$

where k is a quasi-inverse of h. Thus one recognizes the definition of J_k^*. □

Note that when $h(t) = t^{p-1}$, then $k(t) = t^{q-1}$, where $p^{-1} + q^{-1} = 1$.

Now the announced semi-group property for M_λ can be given.

Proposition 4.2. *For any operator M and for any λ, μ in the set $\mathbb{P}$ of (strictly) positive numbers, one has $M_{\lambda+\mu} = (M_\lambda)_\mu$.*

Proof. By definitions of $(M_\lambda)_\mu$ and $M_{\lambda+\mu}$, one has

$$\begin{aligned}((M_\lambda)_\mu)^{-1}(x^*) &= (M_\lambda)^{-1}(x^*) + \mu J_h^{-1}(x^*)\\ &= M^{-1}(x^*) + \lambda J_h^{-1}(x^*) + \mu J_h^{-1}(x^*),\\ (M_{\lambda+\mu})^{-1}(x^*) &= M^{-1}(x^*) + (\lambda+\mu) J_h^{-1}(x^*).\end{aligned}$$

As $J_h^{-1}(x^*)$ is closed convex by lemma 4.1, we get

$$(\lambda + \mu) J_h^{-1}(x^*) = \lambda J_h^{-1}(x^*) + \mu J_h^{-1}(x^*),$$
$$((M_\lambda)_\mu)^{-1}(x^*) = (M_{\lambda+\mu})^{-1}(x^*),$$

hence $M_{\lambda+\mu} = (M_\lambda)_\mu$. □

Let us note that, as expected, monotonicity is preserved by regularization.

Proposition 4.3. *If M is a monotone multimapping, then M_λ is monotone.*

Proof. For $i = 1, 2$, let $x_i^* \in M_\lambda(w_i)$. By definition, for $i = 1, 2$, there exist $x_i^* \in J_h(\lambda^{-1}(w_i - x_i)) \cap M(x_i)$; the monotonicity of M also implies that

$$\langle x_1^* - x_2^*, w_1 - w_2 \rangle \geqslant \langle x_1^* - x_2^*, x_1 - x_2 \rangle \geqslant 0.$$

Thus M_λ is monotone. □

Proposition 4.4. *Suppose M is strictly monotone or M is monotone and the duality mapping J_h is strictly monotone. Then the resolvent P_λ^M is single valued. If, moreover, J_h or M is single valued, the regularization M_λ of M is single valued.*

Proof. Let $x^*, y^* \in M_\lambda(w)$. Then

$$x^* \in J_h(\lambda^{-1}(w - x)) \cap M(x) \quad \text{and} \quad y^* \in J_h(\lambda^{-1}(w - y)) \cap M(y),$$

with $x, y \in P_\lambda^M(w)$. Using the monotonicity of M and J_h, one has

$$0 \leqslant \left\langle x^* - y^*, \frac{w - x}{\lambda} - \frac{w - y}{\lambda} \right\rangle = \lambda^{-1} \langle x^* - y^*, y - x \rangle \leqslant 0.$$

Since M or J_h is strictly monotone, $x = y$. When J_h or M is single valued, $x^* = y^*$. □

We recall for the reader's convenience that a *gage* is a non-decreasing function γ such that $\gamma(0) = 0$ and $\gamma(t) > 0$ for all $t > 0$. We also recall that any quasi-inverse of a gage is a *modulus*, i.e. a non-decreasing function μ such that $\lim_{t \to 0_+} \mu(t) = 0$. A multimapping M from X into X^* is said to be *uniformly monotone* if there exists a gage γ (called *a gage of monotonicity*) γ, such that, for any $(x_i, x_i^*) \in M$, $i = 1, 2$, one has

$$\langle x_1^* - x_2^*, x_1 - x_2 \rangle \geqslant \|x_1 - x_2\| \gamma(\|x_1 - x_2\|).$$

The multimapping M is said to be *strongly monotone* if $\gamma(t) = ct$ for some $c > 0$. It is known that when the space is uniformly convex, then, for any $p > 1$, the function $j_h : x \mapsto p^{-1}\|x\|^p$ is uniformly convex and J_h is uniformly monotone (see [2], [9, lemma 5.1, theorem 4.4], [43, 46, 53]). Note that when M is uniformly monotone, its inverse is single valued. Moreover, one has the following property.

Proposition 4.5. *Suppose the multimapping J_h^{-1} is uniformly monotone with gage of monotonicity γ, and M is monotone. Then M_λ is single valued and uniformly continuous with modulus of uniform continuity $\gamma^\sharp(\lambda^{-1}\cdot)$, where $\gamma^\sharp$ is the greatest quasi-inverse of γ. In particular, if J_h^{-1} is strongly monotone, M_λ is Lipschitzian with rate $(\lambda c)^{-1}$.*

Proof. For $i = 1, 2$, let $x_i^* \in J_h(\lambda^{-1}(w_i - x_i)) \cap M(x_i)$, with $x_i \in P_\lambda^M(w_i)$. As J_h^{-1} is uniformly monotone and $\lambda^{-1}(w_i - x_i) \in J_h^{-1}(x_i^*)$, one has

$$\|x_1^* - x_2^*\|\gamma(\|x_1^* - x_2^*\|) \leqslant \lambda^{-1}\langle x_1^* - x_2^*, (w_1 - x_1) - (w_2 - x_2)\rangle.$$

Since M is monotone, one gets

$$\|x_1^* - x_2^*\|\gamma(\|x_1^* - x_2^*\|) \leqslant \lambda^{-1}\langle x_1^* - x_2^*, w_1 - w_2\rangle \leqslant \lambda^{-1}\|x_1^* - x_2^*\|\|w_1 - w_2\|.$$

Hence

$$\gamma(\|x_1^* - x_2^*\|) \leqslant \lambda^{-1}\|w_1 - w_2\|.$$

Taking $w_1 = w_2$, we get that M_λ is single valued and, by definition of $\gamma^\sharp$,

$$\|x_1^* - x_2^*\| \leqslant \gamma^\sharp(\lambda^{-1}\|w_1 - w_2\|).$$

□

Let us prove boundedness and convergence properties. The first result below is quite simple and natural.

Proposition 4.6. *If X is reflexive, for any multimapping M, one has*

$$M \subset \liminf_{\lambda \to 0_+} M_\lambda.$$

If M is maximal monotone, then M_λ converges in graph to M.

Proof. Given $(x, x^*) \in M$, we will find some $(u_\lambda) \to x$ such that $(u_\lambda, x^*) \to (x, x^*)$ and $(u_\lambda, x^*) \in M_\lambda$, which will prove the first assertion. The second one follows by a well-known argument, which uses the fact that a limit superior of monotone multimappings is monotone (see [5]). As X is reflexive, we can find some $z \in X$ such that $x^* \in J_h(z)$. Then we take $u_\lambda := x + \lambda z$, so that

$$x^* \in M(x) \cap J_h(\lambda^{-1}(u_\lambda - x)).$$

Thus $x \in P_\lambda^M(u_\lambda)$ and $x^* \in M_\lambda(u_\lambda)$. □

Our next result provides precise estimates; it is similar to [17, lemma 1.3], but here we do not make any assumption about the norms of X and X^*, and we may have $h(t) \neq t$.

Proposition 4.7. *For any monotone multimapping M and for any $\lambda > 0$, the mappings P_λ^M and M_λ are bounding, i.e. map bounded sets into bounded sets. Moreover, for each $x \in \operatorname{dom} M_\lambda \cap \operatorname{dom} M$ and any $x_\lambda \in P_\lambda^M(x)$, $x_\lambda^* \in M_\lambda(x)$, one has*

$$h(\|\lambda^{-1}(x - x_\lambda)\|) \leqslant |M(x)| := \inf\{\|x^*\| : x^* \in M(x)\},$$
$$\|x_\lambda^*\| \leqslant |M(x)|,$$

and $x_\lambda \to x$ as $\lambda \to 0_+$. If M is maximal monotone, then any weak-star cluster point of (x_λ^*) as $\lambda \to 0_+$ belongs to the set $M^0(x)$ of points of $M(x)$ with least norm. If, furthermore, the norm of X^* is strictly convex, then $(x_\lambda^*) \to M^0(x)$ for the weak-star topology as $\lambda \to 0_+$.

Proof. Given a monotone multimapping M and $\lambda > 0$, let $x \in \operatorname{dom} M_\lambda$, $x_\lambda^* \in J_h(\lambda^{-1}(x - x_\lambda)) \cap M(x_\lambda)$, where $x_\lambda \in P_\lambda^M(x)$. Then, for $r_x := r_{x,\lambda} := \|x_\lambda - x\|$ and for any (u, u^*) in the graph of M, we have, by the definitions and the monotonicity of M,

$$\begin{aligned} r_x h(\lambda^{-1} r_x) &= \langle x_\lambda^*, x - x_\lambda \rangle \\ &= \langle x_\lambda^*, x - u \rangle + \langle x_\lambda^*, u - x_\lambda \rangle \\ &\leqslant h(\lambda^{-1} r_x)\|x - u\| + \langle u^*, u - x_\lambda \rangle \\ &\leqslant h(\lambda^{-1} r_x)\|x - u\| + \|u^*\|(\|x - u\| + r_x). \end{aligned}$$

It follows that when x remains in a bounded subset B, the number r_x is bounded: P_λ^M is bounding. Since J_h is bounding, M_λ is bounding too.

If we suppose $x \in \operatorname{dom} M_\lambda \cap \operatorname{dom} M$, we can take $u = x$ above, so that

$$r_{x,\lambda} h(\lambda^{-1} r_{x,\lambda}) \leqslant \|u^*\| r_{x,\lambda}$$

for any $u^* \in M(x)$. Taking the infimum, it follows that

$$h(\lambda^{-1} r_{x,\lambda}) \leqslant | M(x) | := \inf\{\|u^*\| : u^* \in M(x)\}$$

and $\lambda^{-1} r_{x,\lambda}$ is bounded. Thus $r_{x,\lambda} \to 0$ as $\lambda \to 0$.

Moreover, as $x_\lambda^* \in J_h(\lambda^{-1}(x - x_\lambda))$, the definition of J_h yields

$$\|x_\lambda^*\| = h(\|\lambda^{-1}(x - x_\lambda)\|) = h(\lambda^{-1} r_{x,\lambda}) \leqslant |M(x)|.$$

As in [5], we can prove that any weak-star cluster point x^* of the bounded net (x_λ^*) when $\lambda \to 0_+$ belongs to $M(x)$ when M is maximal monotone. The lower semicontinuity of the dual norm ensures that $x^* \in M^0(x)$. The last assertion follows from weak-star compactness of balls and the fact that $M^0(x)$ is a singleton when the dual norm is strictly convex. □

For other convergence properties, in particular, for convergence with respect to different topologies, we refer to [6, 22, 41] and their references.

5. Links with the regularization of functions

In this section, we relate our regularization process to the regularization of functions via infimal convolution. Let us recall that the infimal regularization (or extended Moreau–Yosida regularization) of a function f from X into $\mathbb{R} \cup \{+\infty\}$, relative to the weight h, is given for $\lambda > 0$ and $w \in X$ by

$$f_\lambda(w) := \inf_{x \in X} (f(x) + k_\lambda(w - x)),$$

where k_λ is the function $v \mapsto \lambda j_h(\lambda^{-1} v)$ (see, for instance, [12]). One recognize here the infimal convolution of f and $g := k_\lambda$ given by

$$(f \,\Box\, g)(w) = \inf\{f(x) + g(y) : x, y \in X,\ x + y = w\}.$$

The λ-regularization will be said to be exact at w if the infimal convolution is *exact* at w, i.e. if the above infimum is attained; it is said to be exact if it is exact at w for each $w \in X$.

The following lemma includes a sufficient condition for exactness. The growth condition it uses is satisfied for proper convex functions; in fact, much more general growth conditions can be used provided $\lambda > 0$ is taken small enough (see [12], for example). In particular, if $h(t) = t^{p-1}$, with $p > 1$, it suffices to suppose that, for some real numbers b, c, one has the following estimate for $\|x\|$ large enough:

$$f(x) \geqslant c - b\|x\|^p.$$

LEMMA 5.1. *Suppose X is the dual of a Banach space X_*, and the norm of X is a dual norm. Suppose that the function f from X into $\mathbb{R} \cup \{+\infty\}$ is LSC for the weak-star topology $\sigma(X, X_*)$ and bounded below by a weak-star continuous affine function. Then, for each $w \in X$, the function*

$$f_{\lambda,w} : x \mapsto f(x) + \lambda j_h\left(\frac{w-x}{\lambda}\right)$$

is LSC for the weak-star topology $\sigma(X, X_*)$ and coercive, and thus attains its infimum on X. Furthermore, if H is increasing and the norm of X^* satisfies the Kadec–Klee condition (K) then, for any bounded net $(w_i)_{i\in I}$ with (strong) limit w such that $(f_\lambda(w_i)) \to f_\lambda(w)$ and for any x_i in the set $S(w_i)$ of minimizers of f_{λ,w_i}, the net $(x_i)_{i\in I}$ has a (strongly) converging subnet with limit in $S(w)$.

Recall that the Kadec–Klee condition is as follows.

(K) Any net $(u_i)_{i\in I}$ in X that weak-star converges to some $u \in X$ such that $\|u\| = \lim_i \|u_i\|$ is strongly convergent.

Proof. Clearly, j_h is LSC for the weak-star topology $\sigma(X, X_*)$, as is the chosen dual norm. The lower semicontinuity of $f_{\lambda,w}$ ensues; let us prove its coercivity. By assumption, there exists $(a, b) \in X_* \times \mathbb{R}$ such that

$$f(x) \geqslant \langle a, x\rangle - b.$$

Then, for $x \in X \setminus \{0\}$, one has

$$\begin{aligned} f_{\lambda,w}(x) &\geqslant \langle a, x\rangle - b + \lambda H(\lambda^{-1}(\|x\| - \|w\|)) \\ &\geqslant \|x\|(-\|a\| + \lambda\|x\|^{-1}H(\lambda^{-1}(\|x\| - \|w\|))) - b. \end{aligned}$$

Since $\lim_{t\to+\infty} t^{-1}H(t) = +\infty$, we get $f_{\lambda,w}(x) \to +\infty$ as $\|x\| \to +\infty$.

In order to prove the last assertion, let $(w_i)_{i\in I}$ be bounded and convergent to w. Then, by the preceding estimates, for any $x_i \in S(w_i)$, the net $(x_i)_{i\in I}$ is bounded, hence has a converging subnet $(x_j)_{j\in J}$ for the weak-star topology. By the weak-star lower semicontinuity of f and j_h, the limit x of $(x_j)_{j\in J}$ satisfies $f_{\lambda,w}(x) \leqslant f_{\lambda,w}(u)$ for each $u \in X : x \in S(w)$. Taking a further subnet if necessary, we may suppose $\|w_j - x_j\|$ converges to some $r \geqslant \|w - x\|$. If we had $r > \|w - x\|$, since H is increasing and f is weak-star LSC, we would get

$$f_\lambda(w) = f_{\lambda,w}(x) < \liminf_j \left[f(x_j) + \lambda H\left(\frac{\|w_j - x_j\|}{\lambda}\right)\right] = \lim_j f_{\lambda,w_j}(x_j) = f_\lambda(w);$$

a contradiction. Thus $(\|w_j - x_j\|)_{j\in J}$ converges to $\|w - x\|$, and condition (K) implies that $(x_j)_{j\in J} \to x$. □

Now let us consider the case $M = \partial^? f$, where $\partial^?$ is an abstract subdifferential. By this we mean that for any Banach space X of some class $\mathcal{X}$, for any function f of some class $\mathcal{F}(X)$ and any $x \in \operatorname{dom} f$, $\partial^? f(x)$ is a subset of X^* (see [26,27,42,44] and their references for recent developments of this notion). We will assume some of the following properties.

(S) If f is an arbitrary function finite at x, if g is a continuous convex function and if $f + g$ attains its infimum at x, then $0 \in \partial^? f(x) + \partial g(x)$.

(H) If f is an arbitrary function finite at x and if g is convex and Gâteaux differentiable at x such that $f \leqslant g$, $f(x) = g(x)$, then $\partial^? f(x) \subset \{g'(x)\}$.

(P) Given $f : X \times X \to \mathbb{R} \cup \{\infty\}$, $p(\cdot) := \inf_{x \in X} f(\cdot, x)$,

$$S(w) = \{x \in X : f(w, x) = p(w)\},$$

and given $w \in \operatorname{dom} p$, $w^* \in \partial^? p(w)$, there exists $x \in S(w)$ such that $w^* \in \partial^? f(\cdot, x)(w)$ whenever the following condition is satisfied. Given a bounded net $(w_i, x_i)_{i \in I}$ in the graph of S such that $(w_i)_{i \in I} \to w$, $(p(w_i))_{i \in I} \to p(w)$, there exists a subnet of $(x_i)_{i \in I}$ that converges to some element of $S(w)$.

This last condition is close to a condition about performance functions considered in [27], where it is shown that a number of classical subdifferentials satisfy it (at least in the class of Asplund spaces for what concerns the case of the limiting Fréchet subdifferential). Condition (H) is satisfied by all bornological subdifferentials, but not by all subdifferentials. Condition (S) is satisfied by the limiting Fréchet subdifferential in any Asplund space and by the Clarke subdifferential and the Ioffe subdifferential in any space.

Proposition 5.2. *Let $\partial^?$ be an arbitrary subdifferential satisfying (S). Then if, for some $\lambda > 0$, the λ-regularization of f is exact, one has*

$$\operatorname{dom} \partial^? f_\lambda \subset \operatorname{dom}(\partial^? f)_\lambda.$$

If the norm of X is Gâteaux differentiable and if condition (H) is satisfied, then

$$\partial^? f_\lambda \subset (\partial^? f)_\lambda.$$

The same conclusion holds if instead of (H) one assumes conditions (K) and (P).

Proof. Let $w \in \operatorname{dom} \partial^? f_\lambda$. By assumption, there exists a minimizer x_λ of

$$f_{\lambda,w} : x \mapsto f(x) + k_\lambda(x - w).$$

Condition (S) then ensures that, for each such element x_λ,

$$0 \in \partial^? f(x_\lambda) + \partial k_\lambda(x_\lambda - w) = \partial^? f(x_\lambda) + J_h(\lambda^{-1}(x_\lambda - w)),$$

so that $x_\lambda \in \operatorname{dom} P_\lambda^{\partial^? f} = \operatorname{dom}(\partial^? f)_\lambda$.

When j_h is Gâteaux differentiable and condition (H) is satisfied, for any $w^* \in \partial^? f_\lambda(w)$, we have $w^* = J_h(\lambda^{-1}(w - x_\lambda))$, as, for each $v \in X$,

$$f_\lambda(v) \leqslant g_\lambda(v) := f_{\lambda,v}(x_\lambda) := f(x_\lambda) + \lambda j_h(\lambda^{-1}(v - x_\lambda))$$

and $f_\lambda(w) = f_{\lambda,w}(x_\lambda) = g_\lambda(w)$. Thus $w^* \in (\partial f)_\lambda(w)$ by corollary 3.6.

Let us now suppose condition (P) and all the assumptions of the preceding lemma are satisfied by X and f. Then, by its conclusion and condition (P), for some element $x_\lambda \in S(w)$,

$$w^* \in \partial^?(f(x_\lambda) + \lambda j_h(\lambda^{-1}(\cdot - x_\lambda)))(w) = J_h(\lambda^{-1}(w - x_\lambda)),$$

as j_h is convex continuous. Again, corollary 3.6 yields $w^* \in (\partial f)_\lambda(w)$ when j_h is differentiable. $\square$

Example 5.3. Given $X = \mathbb{R}$, let M be the multimapping given for $u \in \mathbb{R}$ by

$$M(u) = \begin{cases} 1 - u & \text{if } u > 0, \\ [-1, 1] & \text{if } u = 0, \\ -1 - u & \text{if } u < 0. \end{cases}$$

Then M is the Fréchet (and Hadamard) subdifferential of the one-variable function f given by $f(x) = |x| - \frac{1}{2}x^2$. Although M is not monotone, the multimapping M is eventually regularizable (for the weight $h(t) = t$). For $\lambda \in]0, 1[$, M_λ is everywhere defined, single valued, Lipschitzian and, for $w \in \mathbb{R}$, one has

$$M_\lambda(w) = \begin{cases} \dfrac{1 - w}{1 - \lambda} & \text{if } w > \lambda, \\ \dfrac{w}{\lambda} & \text{if } w \in [-\lambda, \lambda], \\ \dfrac{-w - 1}{1 - \lambda} & \text{if } w < -\lambda. \end{cases}$$

Moreover, M_λ is the subdifferential of the function f_λ. Note that f is paraconvex and M is paramonotone (or ω-monotone in the terminology of [40]).

Now let us turn to the convex case. For the usual Fenchel subdifferential, conditions (S), (H), (P) are satisfied on the class of convex functions and the preceding proposition applies. However, better results can be obtained by more specific arguments. They involve the following result, which is proved in [38, proposition 6.6.4] by duality arguments; it can also be proved by direct elementary computations (see [45]).

Proposition 5.4. *Let f, g be two convex functions from a Banach space X into $\bar{\mathbb{R}}$. If the infimal convolution of f and g is exact at $w \in \operatorname{dom} f + \operatorname{dom} g$, then for $w = x + y$, with*

$$(f \,\square\, g)(w) = f(x) + g(y),$$

one has

$$\partial(f \,\square\, g)(w) = \partial f(x) \cap \partial g(y).$$

If X is a dual Banach space, a similar relation holds, with ∂ replaced with ∂_*.

Now let us relate ∂f_λ to $(\partial f)_\lambda$.

Proposition 5.5. *Let f be a proper convex function from a Banach space X into $\bar{\mathbb{R}}$. Then $(\partial f)_\lambda \subset \partial f_\lambda$. If X is the dual space of a Banach space X_*, one has $(\partial f)_\lambda \cap X \times X_* \subset \partial_* f_\lambda$.*

Proof. Given $w^* \in (\partial f)_\lambda(w)$, there exists $x \in (\partial f)^{-1}(w^*)$, $v \in J_h^{-1}(w^*)$ with $w = x + \lambda v$. Thus $w^* \in \partial f(x)$, $-w^* \in J_h(-v)$ and

$$0 \in \partial f(x) + J_h\left(\frac{x-w}{\lambda}\right),$$

so that, j_h being even, x is a minimizer of

$$f_{\lambda,w} : x \mapsto f(x) + \lambda j_h\left(\frac{w-x}{\lambda}\right)$$

and by proposition 5.4 we get that

$$w^* \in \partial f(x) \cap \partial j_h\left(\frac{w-x}{\lambda}\right) = \partial f_\lambda(w).$$

The case where w^* belongs to a predual space of X is similar. □

When f is a proper convex LSC function on X, one deduces from the following proposition that $\operatorname{dom} \partial f_\lambda = X$; although the result is probably known, we have not been able to find an appropriate reference, so that we provide a proof.

PROPOSITION 5.6. *Let f be a LSC proper convex function from a Banach space X into $\bar{\mathbb{R}}$. Then f_λ is a continuous convex function on X.*

Proof. The convexity of f_λ is obvious. Since f is bounded below by a continuous affine function and k_λ is coercive, f_λ does not take the value $-\infty$. As f takes at least one finite value and k_λ is finite, f_λ is always finite. Then the result follows from [51, corollary 3]. It can also be proved directly as follows. Let us prove that f_λ is locally bounded above around any point w of X. Let $z \in \operatorname{dom} f$, let $m \in \mathbb{R}$ and let V be a neighbourhood of w such that $\lambda j_h(\lambda^{-1}(v-z)) \leqslant m$ for each $v \in V$. Thus

$$f_\lambda(v) \leqslant f(z) + \lambda j_h(\lambda^{-1}(v-z)) \leqslant f(z) + m.$$

□

In certain cases, a much stronger result is available.

THEOREM 5.7. *Suppose that f is a proper convex function from X into $\mathbb{R} \cup \{+\infty\}$ and that its λ-regularization is exact for some $\lambda > 0$. Then $\partial f_\lambda = (\partial f)_\lambda$. In particular, this conclusion holds when X is the dual of a Banach space X_*, the norm of X is a dual norm and f is proper convex and LSC for the weak-star topology $\sigma(X, X_*)$.*

Proof. We have seen that the λ-regularization of f is exact under the assumptions of the last assertion. Using proposition 5.4, for each minimizer x of the function $f_{\lambda,w}$, we have that

$$\partial f_\lambda(w) = \partial f(x) \cap J_h\left(\frac{w-x}{\lambda}\right).$$

Moreover, by proposition 5.6, $\partial f_\lambda(w)$ is non-empty. Thus, for $M := \partial f$, the characterization of P_λ^M we gave ensures that $x \in P_\lambda^M(w)$. Then proposition 3.5 yields that any $w^* \in \partial f(x) \cap J_h(\lambda^{-1}(w-x))$ belongs to $M_\lambda(w)$. Therefore, $\partial f_\lambda(w) \subset (\partial f)_\lambda(w)$ and, by proposition 5.5, equality holds. □

For h the identity mapping, the following consequence is known under additional assumptions on the norm we do not make here (see [5] for instance).

COROLLARY 5.8. *For any LSC proper convex function f on a reflexive Banach space and for any $\lambda > 0$, one has $\partial f_\lambda = (\partial f)_\lambda$.*

6. The case of non-reflexive Banach spaces

The preceding section represents an incentive to consider the case of operators in non-reflexive spaces. Let X be the dual of Banach space X_* and let us denote by $\mathcal{T}$ the topology on $X \times X^*$ that is the weakest among those for which the mappings

$$\begin{aligned} &(x, x^*) \mapsto \langle w_*, x\rangle \quad (w_* \in X_*),\\ &(x, x^*) \mapsto \langle x^*, w\rangle \quad (w \in X),\\ &(x, x^*) \mapsto \langle x^*, x\rangle \end{aligned}$$

are continuous. In particular, this topology is weaker than the product convergence of the strong convergence on X with the weak-star convergence of bounded nets on X^*. It is also weaker than the topology considered in [28], which is defined as the product of the strong topology on X with the weakest topology on X^* for which the functions $x^* \mapsto \|x^*\|$ and $x^* \mapsto \langle x^*, w\rangle$ for $w \in X$ are continuous. In fact, if a net $((x_i, x_i^*))_{i\in I}$ converges to (x, x^*) for the Gossez's topology, there exist $k \in I$ and $r > 0$ such that $\|x_i^*\| \leqslant r$ for $i \in I$, $i < k$, so that

$$\begin{aligned} |\langle x_i^*, x_i\rangle - \langle x^*, x\rangle| &\leqslant |\langle x_i^*, x_i - x\rangle| + |\langle x_i^* - x^*, x\rangle|\\ &\leqslant \|x_i^*\|.\|x_i - x\| + |\langle x_i^* - x^*, x\rangle| \to 0 \end{aligned}$$

and $((x_i, x_i^*))_{i\in I}$ converges to (x, x^*) for our topology. Nonetheless, the closure of a monotone operator for this topology is still monotone, and the graph of a maximal monotone operator is closed in this topology, so that it seems to be adapted to the study of nonlinear monotone operators. The following definition is closely related to a notion due to Gossez [28]; however, it is limited to the case of a dual space, it is taken in a reverse way and the topology is weaker.

DEFINITION 6.1. An operator M from X into X^* is said to be of type D if M is contained in the closure of $M_* := M \cap X \times X_*$ in the topology $\mathcal{T}$ introduced above.

Obviously, when X is reflexive, any operator M from X into X^* is of type D. When M is maximal monotone and of type D, M coincides with the closure of M_*, since this closure is monotone. Section 5 above and the following examples motivate our study; the second one presents some analogy with [28, theorem 3.1], but does not seem to be a consequence of it.

LEMMA 6.2. *If A is an operator from X_* into X, then $M = A^{-1}$, considered as an operator from X into X^*, is of type D. If f is a proper convex function on X, which is LSC for the weak-star topology $\sigma(X, X_*)$, then $M := \partial f$ is of type D.*

Proof. The first assertion is obvious. The second one follows, as our assumption ensures that there exists an LSC proper convex function f_* on X_* such that f is the conjugate of f_*. Then, by [50, lemma 12A, theorem 3.1], $\partial f = (\partial f_*)^{-1}$. $\square$

We omit the case of maximal monotone operators arising from the subdifferentials of saddle functions, but we observe that not all maximal monotone operators are of type D (use proposition 1 in [29]).

In the sequel we endow X with the dual norm of the norm of X_*. For the sake of simplicity, we omit mention of the weights for the duality mappings of X_*, X, X^*, which are denoted by J_*, J, J^*, respectively, $J_* = \partial j_*$, $J = \partial j$, $J^* = \partial j^*$, where $j_*(x_*) := k(\|x_*\|)$, $j(x) = h(\|x\|)$, $j^*(x^*) = k(\|x^*\|)$, where k is a quasi-inverse of h as in lemma 4.1. For F in the family of finite-dimensional subspaces of X_*, we denote by $i_F : F \to X_*$ the canonical injection and we denote by $i_F^{\mathrm{T}} : X \to F^*$ its transpose. For an operator M from X into X^*, we set $M_F^{-1} := i_F^{\mathrm{T}} \circ M^{-1} \circ i_F$, X_* being identified with its image in X^*. The following lemma is a consequence of [19, proposition 7].

Lemma 6.3. *For any monotone mapping M from X into X^* such that $M(X) \cap X_* \neq \emptyset$, and for any $y_0 \in M(X) \cap X_*$, there exists $r > 0$ such that, for any F in the family $\mathcal{F}$ of finite-dimensional subspaces of X_* containing y_0, one can find $(w_*^F, w^F) \in F \times F^*$ satisfying $w^F \in (i_F^{\mathrm{T}} \circ J_* \circ i_F)(w_*^F)$, $\|w_*^F\| \leqslant r$,*

$$\langle w_*^F - x_*, w^F + x\rangle \leqslant 0 \quad \forall (x_*, x) \in M_F^{-1}. \tag{6.1}$$

Proof. Let $z_0 \in M^{-1}(y_0)$ and let us set

$$G_F := M_F^{-1} - (y_0, 0), \qquad T_F := i_F^{\mathrm{T}} \circ J_* \circ i_F - (y_0, 0),$$

so that $(0, z_F) \in G_F$, where $z_F := i_F^{\mathrm{T}}(z_0)$. An estimate similar to the one in the proof of theorem 3.11 shows that there exists a function $c : \mathbb{R}_+ \to \mathbb{R}$ such that $c(t) \to +\infty$ as $t \to +\infty$ and

$$\langle w, u\rangle \geqslant c(\|u\|)\|u\| \quad \text{for any } u \in F, \quad w \in T_F(u).$$

Then [19, proposition 7] ensures that there exists $r > 0$, independent of F and $(u^F, w^F) \in T_F$, such that $\|u^F\| \leqslant r$,

$$\langle w^F - x_*, u^F + x\rangle \leqslant 0 \quad \forall (x_*, x) \in G_F.$$

Setting $w_*^F := u^F + y_0$, we get the result. □

This lemma leads us to the following surjectivity result, which has some similarity with [28, theorem 4.1], although in that statement an approximate duality mapping appears and not the exact duality mapping.

Proposition 6.4. *Let M be a maximal monotone operator from X into X^* of type D. Then, for each $\lambda > 0$, the operator $M^{-1} + \lambda J^{-1}$ is surjective.*

Proof. Again, it suffices to suppose $\lambda = 1$ and to prove that 0 belongs to the range of $M^{-1} + \lambda J^{-1}$. As M is of type D, one has $M(X) \cap X_* \neq \emptyset$ and one can pick $y_0 \in M(X) \cap X_*$. For each F in the family $\mathcal{F}$ defined as above, let us set $v_*^F := i_F(w_*^F)$ and let us pick $v^F \in J_*(v_*^F)$ such that $w^F = i_F^{\mathrm{T}}(v^F)$. Since (v_*^F) and (v^F) are bounded nets in X_* and X, respectively, we can find $(v_*^{**}, v_*^*) \in X_*^{**} \times X_*^*$, which is a cluster point of $(v_*^{F_i}, v^{F_i})_{i \in I}$ for the product of the weak-star topologies along a subnet $(v_*^{F_i}, v^{F_i})_{i \in I}$ of (v^F, v^F), for which $(\langle v_*^{F_i}, v^{F_i}\rangle)_{i \in I}$ converges to some real number s. Writing $(w^*, w) := (v_*^{**}, v_*^*)$ considered as an element of $X^* \times X = X_*^{**} \times X_*^*$ and taking limits in (6.1), we get

$$s \leqslant \langle w^*, -x\rangle + \langle x_*, w\rangle + \langle x_*, x\rangle \quad \forall (x_*, x) \in M_*^{-1} = M^{-1} \cap X_* \times X. \tag{6.2}$$

On the other hand, the lower semicontinuity of j and j^* and the relations $v^{F_i} \in \partial j_*(v_*^{F_i})$ or $\langle v_*^{F_i}, v^{F_i}\rangle = j_*(v_*^{F_i}) + j(v^{F_i})$ yield, by the Young–Fenchel inequality,

$$\begin{aligned}\langle w^*, w\rangle &\leqslant j^*(w^*) + j(w)\\ &\leqslant \liminf_i j^*(v_*^{F_i}) + \liminf_i j(v^{F_i})\\ &\leqslant \liminf_i (j^*(v_*^{F_i}) + j(v^{F_i}))\\ &= \lim_i \langle v_*^{F_i}, v^{F_i}\rangle = s.\end{aligned} \tag{6.3}$$

Gathering inequalities (6.2) and (6.3), we get

$$\langle w^* - x_*, w + x\rangle \leqslant 0 \quad \forall (x_*, x) \in M_*^{-1} = M^{-1} \cap X_* \times X.$$

Since the coupling function is continuous for the topology $\mathcal{T}$ and since M_*^{-1} is dense in M^{-1} for this topology, we obtain

$$\langle w^* - x^*, w + x\rangle \leqslant 0 \quad \forall (x^*, x) \in M^{-1},$$

hence $w^* \in M(-w)$ by maximal monotonicity of M. As the right-hand side of inequality (6.2) is continuous for $\mathcal{T}$, we also get

$$s \leqslant \langle w^*, -x\rangle + \langle x^*, w\rangle + \langle x^*, x\rangle \quad \forall (x, x^*) \in M.$$

In particular, for $(x, x^*) = (-w, w^*)$, this inequality shows that $s \leqslant \langle w^*, w\rangle$. Then, by inequality (6.3), we get

$$s \leqslant \langle w^*, w\rangle \leqslant j^*(w^*) + j(w) \leqslant s$$

and $w^* \in \partial j(w)$. Thus $0 = -w + w \in M^{-1}(w^*) + J^{-1}(w^*)$. □

The following theorem ensues.

THEOREM 6.5. *Let X be a dual space and let M be a maximal monotone operator from X into X^* of type D. Then M is regularizable.*

7. The case of Nemyckii operators

Let us consider in this section the case of an important class of operators, namely Nemyckii operators (see [24,55] for instance). Given a measured space T, $p > 1$, a Banach space E with dual space E^* and a multimapping F from E into E^*, the Nemyckii operator associated to F is the operator N_F from $X := L_p(T, E)$ into $X^* = L_q(T, E^*)$ given by

$$N_F(x) := \{y \in L_q(T, E^*) : y(t) \in F(x(t)) \text{ a.e. } t \in T\}.$$

We suppose the domain of N_F is non-empty, which can be ensured by measurability and growth conditions. The use of the duality mapping associated with the weight h given by $h(t) = t^{p-1}$, rather than the classical duality mapping, means we can obtain the following result, which states that the regularization can be made pointwise.

PROPOSITION 7.1. *Suppose E and $X := L_p(T, E)$ are provided with the duality mappings associated with the weight $h(t) = t^{p-1}$. Then, for any $\lambda > 0$, the λ-regularization $(N_F)_\lambda$ of the Nemyckii operator N_F is a pointwise regularization in the sense that $(N_F)_\lambda$ is the Nemyckii operator associated with the λ-regularization F_λ of F,*

$$(N_F)_\lambda = N_{F_\lambda}.$$

Proof. Let $M := N_F$. Then

$$M^{-1}(y) = \{x \in L_p(T, E) : x(t) \in F^{-1}(y(t)) \text{ a.e. } t \in T\} = N_{F^{-1}}(y).$$

Therefore, as $J_h^{-1}(y) = \|y\|_E^{q-2} y$, with $q > 0$ such that $p^{-1} + q^{-1} = 1$, one has

$$\begin{aligned} y \in M_\lambda(x) \quad &\Rightarrow \quad x \in M^{-1}(y) + \lambda J_h^{-1}(y) \\ &\Rightarrow \quad x(t) - \lambda \|y(t)\|_E^{q-2} y(t) \in F^{-1}(y(t)) \quad \text{a.e. } t \in T \\ &\Rightarrow \quad y(t) \in F_\lambda(x(t)) \quad \text{a.e. } t \in T \end{aligned}$$

and this last relation corresponds to the definition of N_{F_λ}. □

Let us conclude by observing that the regularization we have dealt with is by no means the most general one. One may wish to take a regularization through non-isotropic kernels, whereas the kernel $j_h = H \circ \|\cdot\|$ we have chosen is isotropic. Such a circumstance occurs in the following case. Suppose $M = \partial f$, where $f = g \circ A$, A being a continuous linear mapping from X onto another Banach space Y and g an LSC proper convex function on Y. Then it may be convenient to replace j_h by $\hat{j}_h = H \circ \|A(\cdot)\|$. In such a case, $M_{[\lambda]} := \partial f_{[\lambda]}$, with $f_{[\lambda]} := f \,\Box\, \lambda \hat{j}_h(\lambda^{-1}\cdot)$ or $f_{[\lambda]} := (g \,\Box\, \lambda j_h(\lambda^{-1}\cdot)) \circ A$ can be considered as good approximations of M. Let us note that both coincide when A is surjective, but none of them coincide with the one we dealt with.

Acknowledgments

We are grateful to Professor Ya. Al'ber for providing us with some references and to an anonymous referee for constructive criticisms.

References

1 J.-C. Aggeri and C. Lescarret. Sur une application de la théorie de la sous-différentiabilité à des fonctions convexes duales associées à un couple d'ensembles mutuellement polaires (Montpellier, 1965). (Mimeographed manuscript.)

2 Ya. I. Al'ber and A. I. Notik. On some estimates for projection operator in Banach spaces. *Commun. Appl. Nonlinear Analysis* **2** (1995), 47–56.

3 W. N. Anderson Jr and R. J. Duffin. Series and parallel addition of matrices. *J. Math. Analysis Appl.* **26** (1969), 576–594.

4 E. Asplund. Positivity of duality mappings. *Bull. Am. Math. Soc.* **73** (1967), 200–203.

5 H. Attouch. *Variational convergence for functions and operators* (Boston, MA: Pitman, 1984).

6 H. Attouch, A. Moudafi and H. Riahi. Quantitative stability analysis for maximal monotone operators and semi-groups of contractions. *Séminaire Analyse Convexe Montpellier* **21** (1991), 9.1–9.38.

7 H. Attouch, Z. Chbani and A. Moudafi. Une notion d'opérateur de récession pour les maximaux monotones. *Séminaire Analyse Convexe Montpellier* **22** (1992), 12.1–12.37.

8 D. Azé. *Eléments d'analyse convexe et variationnelle* (Paris: Ellipse, 1997).

9 D. Azé and J.-P. Penot. Uniformly convex and uniformly smooth convex functions. *Annls Fac. Sci. Toulouse* **4**, (1995), 705–730.

10 V. Barbu. *Nonlinear semigroups and differential equations in Banach spaces* (Groningen: Noordhoff, 1976).

11 B. Beauzamy. *Introduction to Banach spaces and their geometry*. Mathematic Studies, vol. 68 (Amsterdam: North-Holland, 1982).

12 M. Bougeard, J.-P. Penot and A. Pommellet. Towards minimal assumptions for the infimal convolution regularization. *J. Approx. Theory* **64** (1991), 245–270.

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

14 H. Brézis. *Opérateurs maximaux monotones* (Amsterdam: North Holland, 1973).

15 H. Brézis. *Analyse fonctionnelle* (Paris: Masson, 1985; Paris: Dunod, 1999).

16 H. Brézis and M. Sibony. Méthodes d'approximation et d'itération pour les opérateurs monotones. *Arch. Ration. Mech. Analysis* **28** (1968), 59–82.

17 H. Brézis, M. G. Crandall and A. Pazy. Perturbations of nonlinear maximal monotone sets. *Commun. Pure Appl. Math.* **23** (1970), 123–144.

18 F. E. Browder. Nonlinear maximal monotone operators in Banach spaces. *Math. Annln* **175** (1968), 89–113.

19 F. E. Browder. Nonlinear variational inequalities and maximal monotone mappings in Banach spaces. *Math. Annln* **183** (1969), 213–231.

20 F. E. Browder. *Existence theorems for nonlinear partial differential equations*. Proceedings of Symposia in Pure Mathematics, vol. 16, pp. 1–60 (Providence, RI: American Mathematical Society, 1970).

21 F. E. Browder. *Nonlinear operators and nonlinear equations of evolution in Banach spaces*. Proceedings of Symposia in Pure Mathematics, vol. 18 (Providence, RI: American Mathematical Society, 1976).

22 Z. Chabni and J. Lahrache. Convergence des sous-différentiels et topologies intermédiaires. *Séminaire Analyse Convexe Montpellier* **22** (1992), 4.1–4.18.

23 R. Dautray and J. L. Lions. *Analyse mathématique et calcul numérique pour les sciences et techniques*, vol. 8 (Paris: Masson, 1984).

24 K. Deimling. *Nonlinear functional analysis* (Springer, 1985).

25 J. Diestel. *Geometry of Banach spaces. Selected topics*. Lecture Notes in Mathematics, vol. 485 (Springer, 1975).

26 A. D. Ioffe. Codirectional compactness, metric regularity and subdifferential calculus. In *Constructive, experimental, and nonlinear analysis* (ed. M. Theria). *Can. Math. Soc. Conf. Proc. 27*, pp. 123–163 (Providence, RI: American Mathematical Society, 2000).

27 A. D. Ioffe and J.-P. Penot. Subdifferentials of performance functions and calculus of coderivatives of set-valued mappings. *Serdica Math. J.* **22** (1996), 359–384.

28 J.-P. Gossez. Opérateurs monotones non-linéaires dans les espaces de Banach non réflexifs. *J. Math. Analysis Appl.* **34** (1971), 371–395.

29 J.-P. Gossez. On the range of a coercive maximal monotone operator in a nonreflexive Banach space. *Proc. Am. Math. Soc.* **35** (1972), 88–92.

30 J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex analysis and minimization algorithms* (Springer, 1993).

31 J.-L. Lions. *Quelques méthodes de résolution des problèmes aux limites* (Paris: Dunod-Gauthier-Villars, 1970).

32 J.-J. Moreau. Proximité et dualité dans un espace hilbertien. *Bull. Soc. Math. France* **93** (1965), 273–299.

33 F. Kubo. Conditional expectations and operations derived from network connections. *J. Math. Analysis Appl.* **80** (1981), 477–489.

34 F. Kubo and T. Ando. Means of positive linear operators. *Math. Annln* **246** (1980), 205–224.

35 J. C. Maxwell. *A treatise on electricity and magnetism*, 3rd edn (New York: Dover, 1953).

36 M.-L. Mazure. L'addition parallèle d'opérateurs interprétée comme inf-convolution de formes quadratiques convexes. *Modélisation Math. Analysis Numérique* **20** (1986), 497–515.

37 T. D. Morley. Parallel summation, Maxwell principle and the infimum of projections. *J. Math. Analysis Appl.* **70** (1979), 33–40.

38 P.-J. Laurent. *Approximation et optimisation* (Paris: Hermann, 1972).

39 G. B. Passty. The parallel sum of nonlinear monotone operators. *Nonlinear Analysis* **10** (1986), 215–227.

40 A. Pazy. *Semigroups of nonlinear contractions in Hilbert spaces*, pp. 343–430 (CIME, Varenna (1970), Cremonese (1971)).

41 J.-P. Penot. On the approximation of dissipative operators (1994). (Manuscript.)

42 J.-P. Penot. Are generalized derivatives useful for generalized convex functions? In *Generalized convexity, generalized monotonicity: recent results* (ed. J.-P. Crouzeix, J.-E. Martínez-Legaz and M. Volle), pp. 3–59 (Dordrecht: Kluwer Academic, 1997).

43 J.-P. Penot. *Continuity properties of projection operators* (University of Pau, 1997). (Preprint.)

44 J.-P. Penot. *The compatibility with order of some subdifferentials* (University of Pau, 1997). (Preprint.)

45 J.-P. Penot. Proximal mappings. *J. Approx. Theory* **94** (1998), 203–221.

46 J.-P. Penot and R. Ratsimahalo. Uniform continuity of projection mappings in Banach spaces. *Commun. Appl. Nonlinear Analysis* **6** (1999), 1–17.

47 J.-P. Penot and M. Volle. Inversion of real-valued functions and applications. *ZOR Meth. Models Operat. Res.* **34** (1990), 117–141.

48 R. R. Phelps. *Convex functions, monotone operators and differentiability*. Lecture Notes in Mathematics, vol. 1364 (Springer, 1988).

49 R. T. Rockafellar. On the maximality of sum of nonlinear operators. *Trans. Am. Math. Soc.* **149** (1970), 75–88.

50 R. T. Rockafellar. *Conjugate duality and optimization*. CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 16 (Philadelphia, PA: SIAM, 1974).

51 M. Volle. Sur quelques formules de dualité convexe et non convexe. *Set-Valued Analysis* **2** (1994), 369–379.

52 J. J. Vrabie. *Compactness methods for nonlinear evolutions*. Pitman Monographs and Surveys in Pure and Applied Mathematics, vol. 32 (New York: Longman, 1987).

53 Z.-B. Xu and G. F. Roach. Characteristic inequalities of uniformly convex and uniformly smooth Banach spaces. *J. Math. Analysis Appl.* **157** (1991), 189–210.

54 C. Zalinescu. On uniformly convex functions. *J. Math. Analysis Appl.* **95** (1983), 344–374.

55 E. Zeidler. *Nonlinear functional analysis and its applications. Nonlinear monotone operators*, vol. II B (Springer, 1990).

(*Issued 17 August 2001*)