Journal of Convex Analysis
Volume 1 (1994), No.1, 1–29

Variational Sum of Monotone Operators

H. Attouch
Laboratoire d'Analyse Convexe, Université Montpellier II,
F-34095 Montpellier Cedex 5

J.-B. Baillon
Université Lyon I, F-69622 Villeurbanne Cedex.
E-mail baillon@frcpn11.in2p3.fr

M. Théra
LACO, URA 1586, Université de Limoges, F-87060 Limoges Cedex.
E-mail thera@cix.cict.fr

The sum of (nonlinear) maximal monotone operators is reconsidered from the Yosida approximation and graph-convergence point of view. This leads to a new concept, called variational sum, which coincides with the classical (pointwise) sum when the classical sum happens to be maximal monotone. In the case of subdifferentials of convex lower semicontinuous proper functions, the variational sum is equal to the subdifferential of the sum of the functions. A general feature of the variational sum is to involve not only the values of the two operators at the given point but also their values at nearby points.

Keywords : maximal monotone operator, Yosida approximation, convex function, subdifferential operator, set-convergence, graph-convergence, epi-convergence, Mosco-convergence, ρ-Hausdorff distance, graph-distance, epi-distance, inf-convolution, Fenchel duality, Brezis-Crandall-Pazy condition.

1991 Mathematics Subject Classification: 47 H05, 47 H99.

1. The variational sum, an introduction

In this introduction, we present the main lines of the new concept of "variational sum" of maximal monotone operators. The definitions and tools required to make this notion precise are introduced in the next sections. Let H be a real Hilbert space. For an operator $A : H \rightrightarrows H$, we note:

$$D(A) := \{x \in H |\, Ax \neq \emptyset\},$$

the *domain* of A;

$$R(A) := \bigcup_{x \in H} Ax,$$

the range of A;

$$\operatorname{graph} A := \{(u,v) \in H \times H |\, u \in D(A), v \in Au\}$$

the graph of A; A^{-1} the operator defined by

$$x \in A^{-1}y \Longleftrightarrow y \in Ax,$$

ISSN 0944-6532 / \$ 2.50

and $\overline{A}$, the operator defined by

$$\operatorname{graph} \overline{A} := \overline{\operatorname{graph} A}.$$

Given two operators $A, B : H \rightrightarrows H$ possibly nonlinear, multivalued, not everywhere defined, the classical notion of sum is the pointwise sum, that is $A + B : H \rightrightarrows H$ is the operator defined by:

$$D(A+B) = D(A) \cap D(B)$$
$$(A+B)x = Ax + Bx.$$

This is an algebraic notion which, as it has been appearing with more and more evidence, is not always well adapted to problems arising in mathematical analysis. We restrict our attention in this paper to the case of maximal monotone operators A and B. Indeed, as pointed out in the following simple example, although the sum of two maximal monotone operators is a monotone operator, it may fail to be a maximal monotone operator. Define the linear operator A by

$$Au := -u'' \text{ with domain } D(A) := H^2(\mathbb{R}).$$

Fix $f \in L^1(\mathbb{R})$, such that $f \geq 0$, and $f|_\Omega \notin L^2(\Omega)$ [1] for all open set $\Omega \subseteq \mathbb{R}$. Define B by

$$Bu := f.u \text{ and } D(B) := \{u \in L^2(\mathbb{R}) |\; f.u \in L^2(\mathbb{R})\}.$$

A and B are linear self-adjoint operators, and maximal monotone operators in $L^2(\mathbb{R})$. Since their domains have very little in common, ($D(A+B) = D(A) \cap D(B)$ reduces to $\{0\}$), $A + B$ is clearly not a maximal monotone operator.

However, as observed by Brézis, the above "pointwise" sum is quite satisfactory when at least one of the two operators is everywhere defined and Lipschitz continuous. So, a natural idea is, when dealing with two maximal monotone operators, to regularize them, then to take the pointwise sum and to pass to the limit. A good candidate for this approximation procedure is the *Yosida approximation.* We recall that, for any $\lambda > 0$ the *resolvent* of index λ of B is defined by

$$J_\lambda^B := (I + \lambda B)^{-1}.$$

It is a contraction which is everywhere defined. The *Yosida approximate* of index λ of B

$$B_\lambda := \frac{1}{\lambda}(I - J_\lambda^B)$$

[1] As noticed by Lapidus [44], if $\{r_n |\; n \in \mathbb{N}\}$ stands for an enumeration of the rational numbers and if σ is a nonnegative function which belongs to $L^2([0,1])$ and is non integrable near the origin, then, $f(x) := \sum_{n=1}^{\infty} \frac{1}{n^2}\tau(x - r_n)$, where

$$\tau(y) := \begin{cases} \sigma(|y|) & \text{if } |y| \leq 1 \\ 0 & \text{elsewhere} \end{cases}$$

is in $L^1(\mathbb{R})$ and is not in $L^2(\Omega)$ for all open set Ω in $\mathbb{R}$.

is everywhere defined and Lipschitz continuous (with Lipschitz contant $1/\lambda$). So, for any $\lambda \geq 0, \mu \geq 0$, such that $\lambda + \mu \neq 0$, we consider the pointwise sum $A_\lambda + B_\mu$, where A_0 (respectively B_0) denotes A (respectively B). Note that since $\lambda + \mu \neq 0$, at least one of the two operators A_λ, B_μ is Lipschitz continuous and the pointwise sum $A_\lambda + B_\mu$ is maximal monotone according to Brézis [29; Lemma 2.4]. So it is natural to address the general question:

Does the filtered family $\{A_\lambda + B_\mu | (\lambda, \mu) \in \mathbb{R}^2_+, \lambda + \mu \neq 0\}$ graph-converges as $\lambda \to 0$ and $\mu \to 0$?

At this point, a few words are necessary in order to explain what is graph-convergence. It is nothing but the topological set-convergence of the graphs of the operators in the product space $H \times H$ equipped with the product topology. This has been proved to be the right concept for the convergence of sequences of maximal monotone operators. Indeed, it has been shown to be equivalent to the pointwise convergence of the resolvents of the operators (cf. H. Brézis, Ph. Benilan, H. Brézis and A. Pazy, H. Attouch). Note that the Yosida approximation B_λ of an operator B graph-converges to B as λ goes to zero which justifies our convention $B_0 = B$ (the pointwise convergence only provides, for any x in $D(B)$, the element of minimal norm of Bx). So, we are led to the following definition: In the sequel we note $\mathcal{F}$ the filter of all the pointed neighbourhoods of the origin in the set $\mathcal{I} := \{(\lambda, \mu) \in \mathbb{R}^2 | \lambda \geq 0, \mu \geq 0, \lambda + \mu \neq 0\}$ and $\lim\limits_{\mathcal{F}}$ for $\lim\limits_{\substack{\lambda \to 0, \mu \to 0 \\ \lambda + \mu \neq 0 \\ \lambda \geq 0, \mu \geq 0}}$.

The sum (see section 4) of two maximal monotone operators A and B is defined as

$$A \underset{v}{+} B := \liminf_{\mathcal{F}} (A_\lambda + B_\mu),$$

where the limit inferior is taken in the sense of Kuratowski-Painlevé, when we identify the operators with their graphs. The interesting case is when $A \underset{v}{+} B$ is still a maximal monotone operator, in which case the limit inferior is in fact a limit. This is a situation which is considered in sections 4, 5, 6. The graph-convergence $A \underset{v}{+} B = \lim\limits_{\mathcal{F}} (A_\lambda + B_\mu)$ can be equivalently formulated in terms of resolvents: for any $y \in H$, the family $\{u_{\lambda,\mu} | (\lambda, \mu) \in \mathcal{I}\}$ of solutions of

$$u_{\lambda,\mu} + A_\lambda u_{\lambda,\mu} + B_\mu u_{\lambda,\mu} \ni y$$

converges with respect to the filter $\mathcal{F}$ to the solution u of

$$u + (A \underset{v}{+} B)u \ni y.$$

When the pointwise sum $A+B$ is a maximal monotone operator, then, by an argument extending the Brézis-Crandall-Pazy theory, the two above notions (classical and variational sum) coincide. More generally, as proved in section 6, when the graph-closure $\overline{A+B}$ of $A+B$ is maximal monotone, then $\lim\limits_{\mathcal{F}} (A_\lambda + B_\mu)$ exists and we have the following equality:

$$A \underset{v}{+} B = \overline{A + B}.$$

The case of subdifferential operators, which is completely understood, shows that the variational sum is quite an involved concept, which, in general, cannot be reduced to elementary operations like pointwise sum and closure operation. Indeed, as shown in section 7, if $A = \partial f$ and $B = \partial g$ then

$$\partial f \underset{v}{+} \partial g = \partial(f+g).$$

We give a formula for $\partial f \underset{v}{+} \partial g$ which makes clear that the variational sum involves ∂f and ∂g not only at the point x but also at nearby points. Similar type results can be obtained in the case of subdifferentials of closed convex-concave saddle functions. We conclude the paper by giving the example of the Schrödinger equations for which our theory applies. Recent contributions on "generalized" sum of maximal monotone operators have been obtained by several authors in different areas: let us mention M. Lapidus in the study of the Feynman integral via the Trotter formula [44], Ph. Clément and P. Egberts in connection with semilinear elliptic systems [37], F. Kubo in the study of electrical networks via the parallel sum of operarors [43], H. Attouch in the variational averaging of operators and applications to stochastic homogenization of composite media ...

2. Set-convergence, Graph-convergence, Epigraphical convergence

We first recall some basic definitions and results on set-convergences (in the Kuratowski-Painlevé, Mosco and Attouch-Wets sense), just enough in order to make the paper self-contained. The interested reader may find a fairly complete account on these notions in [3], [17], [24], [64] and the preface of the Marseille Meeting on "Convergences en Analyse Multivoque et Unilatérale" [11]. When applying these concepts to operators which are identified with their graphs, we obtain the graph-convergence of sequences of operators. In a parallel way, when real extended-valued functions are identified with their epigraphs, we obtain the epi-convergence of sequences of functions.

In the sequel we will write "$\xrightarrow{s}$", "$\xrightarrow{w}$" and "$\xrightarrow{\tau}$", to denote respectively, the strong (norm) convergence, the weak convergence and the convergence with respect to a topology τ.

2.1. Set-convergence

Definition 2.1. Let (X,τ) be a first countable topological space. Given a sequence $\{C_n \subseteq X | n \in \mathbb{N}\}$ of subsets of X, the τ-lower limit of the sequence $\{C_n | n \in \mathbb{N}\}$, denoted by $\tau - \liminf_{n\to\infty} C_n$ is the closed subset of X defined by

$$\tau - \liminf_{n\to\infty} C_n := \{x \in X | \exists (x_n)_{n\in\mathbb{N}}, \forall n \in \mathbb{N}\ x_n \in C_n \text{ and } x_n \xrightarrow{\tau} x\}. \tag{2.1}$$

The τ-upper limit of the sequence $\{C_n | n \in \mathbb{N}\}$, denoted by $\tau - \limsup_{n\to\infty} C_n$ is the closed subset of X defined by

$$\begin{gathered}\tau - \limsup_{n\to\infty} C_n := \\ \{x \in X | \exists (n_k)_{k\in\mathbb{N}}, \exists \{x_k\}_{k\in\mathbb{N}}, \forall k \in \mathbb{N}\ x_k \in C_{n_k} \text{ and } x_k \xrightarrow{\tau} x\}.\end{gathered} \tag{2.2}$$

The sequence $\{C_n | n \in \mathbb{N}\}$ is declared *Kuratowski-Painlevé* convergent for the topology τ, or briefly τ-convergent, if the following equality holds :

$$\tau - \limsup_{n\to\infty} C_n = \tau - \liminf_{n\to\infty} C_n.$$

Its limit, denoted by C = τ-lim C_n, is the closed subset of X equal to this common value

$$C = \tau - \liminf_{n\to\infty} C_n = \tau - \limsup_{n\to\infty} C_n = \tau - \lim_{n\to\infty} C_n.$$

When X is a normed linear space, the sequential weak upper limit of a sequence $\{C_n | \, n \in \mathbb{N}\}$ of subsets of X is defined by

$$\mathrm{seq} - \mathrm{w} - \limsup_{n\to\infty} C_n := \ \{x \in X | \ \exists (n_k)_{k\in\mathbb{N}}, \exists (x_k)_{k\in\mathbb{N}}, \forall k \in \mathbb{N}, x_k \in C_{n_k} \text{ and } x_k \stackrel{w}{\to} x\}.$$

Definition 2.2. A sequence $\{C_n | \, n \in \mathbb{N}\}$ of subsets of a normed linear space X is said to Mosco converge to a set C, and we write $C = \mathrm{M-} \lim_{n\to\infty} C_n$, if :

$$seq - w - \limsup_{n\to\infty} C_n \subseteq C \subseteq s - \liminf_{n\to\infty} C_n.$$

Equivalently, $C = \mathrm{M}\text{-} \lim_{n\to\infty} C_n$, if and only if both of the following conditions hold:

(i) for each $x \in C$, there exists a sequence $\{x_n | \, n \in \mathbb{N}\}$ norm converging to x such that $x_n \in C_n$ for each $n \in \mathbb{N}$;
(ii) for each subsequence $\{n_k | \, k \in \mathbb{N}\}$ and $\{x_k | \, k \in \mathbb{N}\}$ such that $x_k \in C_{n_k}$, the weak convergence of $\{x_k | \, k \in \mathbb{N}\}$ to $x \in C$ forces x to belong to C.

It is an immediate consequence of these definitions that the Mosco convergence implies the Kuratowski-Painlevé convergence and that the two notions coincide whenever X is finite dimensional. It turns out that the Mosco-convergence is a basic concept when considering sequences of convex sets in reflexive Banach spaces. For instance, the mapping which assigns to every f its Fenchel conjugate f^* is an homeomorphism from the set of convex lower semicontinuous functions on X to the set of convex lower semicontinuous functions on X^* ([47], [11], [25]).
Let us now turn our attention to the concept of *ρ-Hausdorff distance* which has been recently introduced by Attouch-Wets [14] and further developed by Beer [24], Azé-Penot [21], Beer-Lucchetti [26]. Let X be a normed space with norm $||.||$ and unit ball B. For any $C, D \subseteq X$, the *excess* of C on D is defined as

$$e(C,D) := \ \sup_{x\in C} d(x,D),$$

where $d(x,C) := \ \inf_{x\in C} ||x-y||$ is the distance from x to C. We adopt the convention that if $C = \emptyset$, then $d(x,C) = \infty$ and $e(C,D) = 0$.

For any $\rho > 0$, the ρ-Hausdorff distance between C and D is given by :

$$\mathrm{haus}_\rho(C,D) := \ \max\{e(C_\rho, D), e(D_\rho, C)\}$$

where C_ρ (resp.D_ρ) is defined by $C_\rho = C \cap \rho B$ (resp. $D_\rho = D \cap \rho B$).

Definition 2.3. A sequence $\{C_n \subseteq X | \, n \in \mathbb{N}\}$ of subsets of X is said to converge with respect to the ρ-Hausdorff distances to a set C, if for each $\rho > 0$,

$$\lim_{n\to\infty} \mathrm{haus}_\rho(C_n, C) = 0.$$

It can be checked, when restricting our attention to convex sets, that $\lim_{n\to\infty} \text{haus}_\rho(C_n, C) = 0$ forces $C = \text{M}-\lim_{n\to\infty} C_n$ (see [14]).

2.2. Graph-convergence of operators

Definition 2.4. Let (X,τ) and (Y,β) be two first countable topological spaces. A sequence $\{A_n | n \in \mathbb{N}\}$ of operators $A_n : X \rightrightarrows Y$, is said to be $\tau\times\beta-$graph convergent to an operator $A : X \rightrightarrows Y$ if the sequence of sets $\{\text{graph } A_n | n \in \mathbb{N}\}, \tau\times\beta$-converges in $X \times Y$, in the Kuratowski-Painlevé sense, to the set graph A.

We then write

$$\tau\times\beta - \text{graph} - \lim_{n\to\infty} A_n = A.$$

In other words, $\{A_n | n \in \mathbb{N}\}$ $\tau\times\beta$- graph-converges to A if and only if the following inclusions hold :

$$\tau\times\beta - \limsup_{n\to\infty} \text{graph } A_n \subseteq \text{graph } A \subseteq \tau\times\beta - \liminf_{n\to\infty} \text{graph } A_n,$$

with

$$\tau\times\beta - \liminf_{n\to\infty} \text{graph } A_n = \\ \{(x,y)\in X\times Y | \,\exists\, (x_n,y_n),\, \forall n\in\mathbb{N}, y_n \in A_n(x_n), x_n \xrightarrow{\tau} x, y_n \xrightarrow{\beta} y\}, \quad (2.3)$$

$$\tau\times\beta - \limsup_{n\to\infty} \text{graph } A_n = \\ \{(x,y)\in X\times Y | \,\exists\, (n_k),\, \exists (x_k,y_k), \forall k\in\mathbb{N}, y_k \in A_{n_k}(x_k), x_k \xrightarrow{\tau} x, y_k \xrightarrow{\beta} y\}. \quad (2.4)$$

When X and Y are normed spaces and $\tau = s_X$ and $\beta = s_Y$ are respectively the strong topologies on X and Y, which is often the case of interest, we shall adopt a simplified terminology and notation: we shall write briefly graph-$\lim_{n\to\infty} A_n = A$ or even simpler $A = \text{G}-\lim_{n\to\infty} A_n$ in order to denote the $s_X \times s_Y-$graph-convergence of the sequence $\{A_n | n \in \mathbb{N}\}$ to A. We shall say that the sequence of operators $\{A_n | n \in \mathbb{N}\}$ graph-converges or briefly G-converges to A.
We may equally define, for any $\rho \geq 0$, the ρ-Hausdorff graph-distance between two operators A_1 and $A_2 : X \rightrightarrows Y$:

$$\text{haus}_\rho(A_1, A_2) := \text{haus}_\rho(\text{graph } A_1, \text{graph } A_2)$$

where X and Y are two normed spaces and $X \times Y$ is equipped with the box norm defined by $|||(x,y)||| = \max\{||x||,||y||\}$.

Definition 2.5. A sequence of operators $\{A_n : X \rightrightarrows Y | n \in \mathbb{N}\}$ converges with respect to the ρ-Hausdorff distances to an operator $A : X \rightrightarrows Y$, and we write graph-dist-$\lim_{n\to\infty} A_n = A$, if for each $\rho > 0$,

$$\lim_{n\to\infty} \text{haus}_\rho(A_n, A) = 0.$$

Note that this notion is independant of the choice of the product norm on $X \times Y$.

2.3. Epigraphical convergence of functions

Definition 2.6. Let $\{f, f_n |\ X \to \overline{\mathbb{R}}, n \in \mathbb{N}\}$ be a sequence of extended real-valued functions. If the sequence $\{\text{epi } f_n | n \in \mathbb{N}\}$ Kuratowski-Painlevé converges to epi$f$ in $X \times \mathbb{R}$ for the product topology, then we say that the sequence $\{f_n | n \in \mathbb{N}\}$ epiconverges to f and we write epi-$\lim\limits_{n\to\infty} f_n = f$.

Definition 2.7. Let X be a normed space and $\{f, f_n | n \in \mathbb{N}\}$ be a sequence of functions from X into $\mathbb{R}\cup\{+\infty\}$. We say that f is the Mosco-epi-limit of the sequence $\{f_n | n \in \mathbb{N}\}$, and we write $f = \mathrm{M} - \text{epi} \lim\limits_{n\to\infty} f_n$, if the sequence $\{\text{epi } f_n | n \in \mathbb{N}\}$ Mosco converges to epi f.

This is equivalent to say that, for any $x \in X$, the two following statements hold:

(i) for any $\{x_n\} \overset{w}{\to} x$ then $f(x) \le \liminf\limits_{n\to\infty} f_n(x_n)$

and

(ii) there exists $\{x_n\} \overset{s}{\to} x$ such that $\limsup\limits_{n\to\infty} f_n(x_n) \le f(x)$.

Definition 2.8. For any $\rho \ge 0$, the ρ-Hausdorff epi-distance between two extended real-valued functions $f, g : X \to \overline{\mathbb{R}}$ is defined by :

$$\text{haus}_\rho(f, g) := \text{haus}_\rho(\text{epi}f, \text{epi}g)$$

where the unit ball of $X \times \mathbb{R}$ is the set $B := B_X \times \mathbb{R} = \{(x, \alpha) : ||x|| \le 1, |\alpha| \le 1\}$.

Note that we may equally consider any equivalent norm on the product space $X \times \mathbb{R}$. It turns out that, for practical purpose and computation, the box norm is more convenient. Let us now state some basic facts about the topology induced by the pseudo-distances $\{\text{haus}_\rho | \rho > 0\}$ on the space of extended real-valued functions.

Definition 2.9. Let $\overline{\mathbb{R}}^X$ be the space of extended real-valued functions defined on the normed linear space X. The topology on $\overline{\mathbb{R}}^X$ generated by the pseudo-distances $\{\text{haus}_\rho | \rho > 0\}$ is called the epi-distance topology or the bounded Hausdorff topology or the Attouch-Wets topology. In other words, for a sequence $\{f_n | n \in \mathbb{N}\}$:

$$f = \text{epi} - \text{dist} \lim_{n\to\infty} f_n \Leftrightarrow \lim_{n\to\infty} \text{haus}_\rho(f_n, f) = 0 \quad \text{for all } \rho > 0.$$

Proposition 2.10. [14, 24] *When $X = \mathbb{R}^m$ is a finite dimensional space, the epi-distance topology coincide with the topology of the epi-convergence. When X is a Banach space and $\{f, f_n : X \to \mathbb{R} \cup \{+\infty\}, n \in \mathbb{N}\}$ is a sequence of lower semicontinuous convex proper functions, the following implication holds:*

$$\lim_{n\to\infty} \text{haus}_\rho(f, f_n) = 0 \ \text{ for all } \rho \text{ sufficiently large, implies that } f = \mathrm{M} - \text{epi} \lim_{n\to\infty} f_n.$$

3. Graph-convergence of maximal monotone operators and Yosida approximation

From now on, unless specified, we work in a Hilbert space setting. Let us denote by H a real Hilbert space and $||\cdot||$ the norm generated by the scalar product $\langle\cdot,\cdot\rangle$. Given an operator $A : H \rightrightarrows H$, possibly multivalued and not everywhere defined, it is identified with its graph. Let us first review some basic facts about such operators. $A : H \rightrightarrows H$ is *monotone,* if for each $x_1 \in D(A), x_2 \in D(A)$ and $y_1 \in Ax_1, y_2 \in Ax_2$ we have

$$\langle y_2 - y_1, x_2 - x_1\rangle \geq 0.$$

It is said to be *maximal monotone* if its graph is maximal for the inclusion among all monotone operators. From Minty's Theorem, this maximality property is equivalent to say that the range of the operator $I + A$ is the whole space, that is $R(I + A) = H$. The resolvent of index $\lambda > 0$ of the maximal monotone operator A is the operator

$$J_\lambda^A := \ (I + \lambda A)^{-1}.$$

It is a contraction which is everywhere defined. The resolvents are tied by the so-called resolvent equation:

$$\text{for any } \lambda > 0,\ \mu > 0, \quad J_\lambda^A x = J_\mu^A\Big[\frac{\mu}{\lambda}x + (1 - \frac{\mu}{\lambda})J_\lambda^A x\Big]. \tag{3.5}$$

When $A = \partial f$ (the subdifferential of a lower semicontinuous convex proper function $f : H \to \mathbb{R}\cup\{+\infty\}$) we will write J_λ^f instead of $J_\lambda^{\partial f}$. The Yosida approximate of index $\lambda > 0$ of A, $A_\lambda := \frac{1}{\lambda}(I - J_\lambda)$ is a Lipschitz continuous monotone operator (with Lipschitz constant $1/\lambda$) which is everywhere defined. The following relation will be quite useful in the sequel:

$$\text{for any } x \in H, \text{ for any } \lambda > 0, \quad A_\lambda x \in A(J_\lambda^A x). \tag{3.6}$$

(see [29] for further details). For any x belonging to $D(A)$, the filtered family $\{A_\lambda x |\, \lambda \to 0\}$ norm-converges to A^0x, the unique element of minimal norm of the nonempty closed convex subset Ax of H.
Let us now review a few facts about the graph-convergence and the graph-distance convergence of maximal monotone operators (see [3]):

Proposition 3.1.
(i) *The class of maximal monotone operators on a finite dimensional space H is closed with respect to the Kuratowski-Painlevé set-convergence, the space $H \times H$ being equipped with the product topology; in other words, if a sequence $\{A_n |\, n \in \mathbb{N}\}$ of maximal monotone operators is such that*

$$s \times s - \lim_{n\to\infty} \text{graph}\, A_n = A,$$

then A is the graph of a maximal monotone operator and graph-$\lim_{n\to\infty} A_n = A$.
(ii) *For any sequence $\{A, A_n |\, n \in \mathbb{N}\}$ of maximal monotone operators on a general Hilbert space H the following equivalence holds:*

$$\text{graph} - \lim_{n\to\infty} A_n = A \Leftrightarrow \text{graph}\, A \subseteq s \times s - \liminf_{n\to\infty} \text{graph}\, A_n.$$

In other words, the sequence $\{A_n | n \in \mathbb{N}\}$ *graph-converges to* A *if and only if, for any* $(x,y) \in \operatorname{graph} A$, *there exists a sequence* $\{(x_n, y_n) | n \in \mathbb{N}\}$ *such that for all* $n \in \mathbb{N}$ $y_n \in A_n(x_n)$ *and* $\{x_n | n \in \mathbb{N}\}, \{y_n | n \in \mathbb{N}\}$ *norm-converge respectively to* x *and* y.

(iii) *For any sequence* $\{A, A_n | n \in \mathbb{N}\}$ *of maximal monotone operators the following implications hold:*

$$\begin{cases} \text{graph} - \lim\limits_{n\to\infty} A_n = A \\ \Rightarrow s \times w - \limsup\limits_{n\to\infty} \operatorname{graph} A_n \subseteq \operatorname{graph} A \\ \Rightarrow w \times s - \limsup\limits_{n\to\infty} \operatorname{graph} A_n \subseteq \operatorname{graph} A. \end{cases}$$

In other words, for any strictly increasing sequence $n_1 < n_2 < n_3 < \dots$ *and any sequence* $\{(x_k, y_k) | k \in \mathbb{N}\}$ *such that for all* $k \in \mathbb{N}$ $y_k \in A_{n_k}(x_k)$, *the strong convergence of* $\{x_k | k \in \mathbb{N}\}$ *to* x *and the weak convergence of* $\{y_k | k \in \mathbb{N}\}$ *to* y *force* (x,y) *to belong to* $\operatorname{graph} A$. *The same property holds by exchanging strong and weak convergences.*

This last property is related to the fact that a maximal monotone operator is closed in $s - H \times w - H$ and in $w - H \times s - H$.
Let us now describe the graph convergence of sequences of maximal monotone operators with the help of their resolvents [3]:

Proposition 3.2. *For any sequence* $\{A, A_n | n \in \mathbb{N}\}$ *of maximal monotone operators the following equivalences hold:*

(i) graph$-\lim\limits_{n\to\infty} A_n = A$;

(ii) *for any* $\lambda > 0$, *for any* $x \in H$, *the sequence* $J_\lambda^{A_n} x$ *norm-converges to* $J_\lambda^A x$ *as* $n \to \infty$;

(iii) *for some* $\lambda_0 > 0$, *for any* $x \in H$, *the sequence* $J_{\lambda_0}^{A_n} x$ *norm-converges to* $J_{\lambda_0}^A x$ *as* $n \to \infty$.

So, the graph-convergence of maximal monotone operators is equivalent to the pointwise convergence of their resolvents, which is clearly equivalent to the pointwise convergence of their Yosida approximates[2]. Let us now examine the graph-distance convergence of such operators and show that it is equivalent to the uniform convergence on bounded subsets of H of the resolvents. To that end, following Attouch & Wets [14] let us introduce a second type of distance between maximal monotone operators expressed with the help of the Yosida approximates. For all $\lambda > 0$, $\rho \geq 0$,

$$d_{\lambda,\rho}(A,B) := \sup_{||x||\leq\rho} ||J_\lambda^A x - J_\lambda^B x|| = \lambda. \sup_{||x||\leq\rho} ||A_\lambda x - B_\lambda x||,$$

[2] Note that the class of maximal monotone operators on a general Hilbert space H is closed with respect to the topology of the pointwise convergence of the resolvents, see [3]. Compare with Proposition 3.2 (i), where a similar property holds with respect to the Kuratowski-Painlevé convergence, but with dim H $< +\infty$!

see also B. Lemaire [45], P. Tossings [68], H. Attouch & A. Moudafi & H. Riahi [10], A. Moudafi [49].

Proposition 3.3. [10; Prop 1.1] *Let A and B be two maximal monotone operators. Then for every $\lambda > 0$ and $\rho \geq 0$,*

$$\mathrm{haus}_\rho(A, B) \leq \max(1, \frac{1}{\lambda}).d_{\lambda,\rho'}(A, B)$$

where

$$\rho' := (1 + \lambda)\rho.$$

By taking $\lambda = 1$, we have $\mathrm{haus}_\rho(A, B) \leq d_{1,2\rho}(A, B)$.

Proposition 3.4. [10; Prop 1.2] *Let A and B be two maximal monotone operators. Then, for every $\lambda > 0$, $\rho \geq 0$*

$$d_{\lambda,\rho}(A, B) \leq (2 + \lambda)\mathrm{haus}_{\rho''}(A, B)$$

where

$$\rho'' = \max\left\{\rho + ||J_\lambda^A 0||,\ \frac{1}{\lambda}(\rho + ||J_\lambda^A 0||)\right\}.$$

Taking $\lambda = 1$, $d_{1,\rho}(A, B) \leq 3\,\mathrm{haus}_{\rho+||J_1^A(0)||}(A, B)$.

Proposition 3.5. [10; Prop 1.3] *Let B, B' be two maximal monotone operators. Then for any $\lambda > 0, \rho \geq 0$*

$$d_{\lambda,\rho}(B, B') \leq (1 + |1 - \lambda|)d_{1,\rho'}(B, B')$$

where,

$$\rho' = \frac{1}{\inf(1, \lambda)}(\rho + |1 - \lambda|.||J_1^B 0||).$$

The following result shows that the Yosida approximations A_λ graph-distance converge to A (and hence graph-converge) as λ goes to zero. This property and the key role played by this approximation in the definition of the variational sum, justifies our interest for the graph-distance convergence of sequences of maximal monotone operators.

Proposition 3.6. *Let A be a maximal monotone operator. Then, for every $\rho > 0$, the following approximation result holds:*

$$\lim_{\lambda \to 0} \mathrm{haus}_\rho(A, A_\lambda) = 0.$$

In other words, A_λ converges to A , as λ goes to zero, for the ρ-graph distance. More precisely, the following estimation holds:

$$d_{1,\rho}(A, A_\lambda) \leq \frac{2\lambda}{\lambda + 1}(2\rho + ||J_1 0||).$$

4. The variational sum of maximal monotone operators. Definition.

We denote by $\mathcal{I} := \{(\lambda, \mu) \in \mathbb{R}^2 | \, \lambda \geq 0, \mu \geq 0, \lambda + \mu \neq 0\}$. For any $(\lambda, \mu) \in \mathcal{I}$, we consider the pointwise sum $C_{\lambda,\mu} := A_\lambda + B_\mu$, where A_0 by convention (respectively B_0) denotes A (respectively B). Note that, since $\lambda + \mu \neq 0$, at least one of the two operators A_λ, B_μ is Lipschitz continuous and therefore $C_{\lambda,\mu}$ is maximal monotone (c.f [29]). In the sequel, we denote by $\mathcal{F}$ the filter of all the pointed neighbourhoods of the origin in $\mathcal{I}$, and by $\text{graph} - \liminf_{\mathcal{F}} C_{\lambda,\mu}$ (respectively $\text{graph} - \limsup_{\mathcal{F}} C_{\lambda,\mu}$) the *lower limit* (respectively *upper limit*) of the family $\{\text{graph}\, C_{\lambda,\mu} | \, (\lambda, \mu) \in \mathcal{I}\}$ along the filter $\mathcal{F}$ with respect to the box topology on $H \times H$. Recall that

- (x, y) belongs to $\text{graph} - \liminf_{\mathcal{F}} C_{\lambda,\mu}$, whenever for every neighbourhood Q of (x, y), there is $F \in \mathcal{F}$ such that for $(\lambda, \mu) \in F$, $\text{graph}\, C_{\lambda,\mu} \cap Q \neq \emptyset$

while,

- (x, y) belongs to $\text{graph} - \limsup_{\mathcal{F}} C_{\lambda,\mu}$, if for every neighbourhood Q of (x, y) and every $F \in \mathcal{F}$, there is $(\lambda, \mu) \in F$ such that $\text{graph}\, C_{\lambda,\mu} \cap Q \neq \emptyset$.

The filtered family $\{C_{\lambda,\mu} | \, (\lambda, \mu) \in \mathcal{I}\}$ is declared *graph-convergent* to C and we note $\text{graph-}\lim_{\mathcal{F}} C_{\lambda,\mu} = C$, if

$$\text{graph} - \limsup_{\mathcal{F}} C_{\lambda,\mu} = \text{graph} - \liminf_{\mathcal{F}} C_{\lambda,\mu} = C.$$

Definition 4.1. The variational sum of two maximal monotone operators A and B is defined as

$$A \underset{v}{+} B := \text{graph} - \liminf_{\mathcal{F}} (A_\lambda + B_\mu).$$

Since the box topology on $H \times H$ is first countable and the filter $\mathcal{F}$ is countably based, it amounts to say that:

$(x, y) \in \text{graph}\, (A \underset{v}{+} B)$, if and only if, for every sequences $\{\lambda_n | \, n \in \mathbb{N}\}$, $\{\mu_n | \, n \in \mathbb{N}\}$ with $\lambda_n \geq 0$, $\mu_n \geq 0, \lambda_n + \mu_n \neq 0$, $\lim_{n\to\infty} \lambda_n = \lim_{n\to\infty} \mu_n = 0$, there exist sequences $\{x_n | \, n \in \mathbb{N}\}$ and $\{y_n | \, n \in \mathbb{N}\}$ such that $\lim_{n\to+\infty} x_n = x$, $\lim_{n\to+\infty} y_n = y$ and $y_n \in (A_{\lambda_n} + B_{\mu_n})(x_n)$.

Equivalently, in terms of resolvents this means that, for any $y \in R(I + (A \underset{v}{+} B))$ (in particular, by Minty's Theorem whenever $A \underset{v}{+} B$ is maximal monotone for every $y \in H$), the family $\{u_{\lambda,\mu} | \, (\lambda, \mu) \in \mathcal{I}\}$ of solutions of

$$u_{\lambda,\mu} + A_\lambda u_{\lambda,\mu} + B_\mu u_{\lambda,\mu} \ni y$$

converges for $(\lambda, \mu) \in \mathcal{I}$ to the solution u of

$$u + (A \underset{v}{+} B)u \ni y.$$

Proposition 4.2. *Let A, B be two maximal monotone operators. Then,*

(1) *If $D(A) \cap D(B) \neq \emptyset$ then $D(A \underset{v}{+} B) \neq \emptyset$;*

(2) *$A \underset{v}{+} B$ is monotone;*

(3) *If $A \underset{v}{+} B$ is maximal monotone, then the family $\{A_\lambda + B_\mu \,|\, (\lambda, \mu) \in \mathcal{I}\}$ graph-converges with respect to the filter $\mathcal{F}$;*

(4) *$A \underset{v}{+} B = B \underset{v}{+} A$ (commutativity).*

Proof. (1), (2) and (4) are a direct consequence of the definition. For (1), notice that $A_\lambda x + B_\mu x \to A^0 x + B^0 x$. Let us prove (3). If $A \underset{v}{+} B$ is maximal monotone, then by virtue of Proposition 3.1 part (ii), we have that $A \underset{v}{+} B \subseteq \text{graph}-\liminf_{\mathcal{F}}(A_\lambda + B_\mu)$ implies that $A \underset{v}{+} B = \text{graph} - \lim_{\mathcal{F}}(A_\lambda + B_\mu)$. □

5. The case $A + B$ maximal monotone: $A + B = A \underset{v}{+} B$

In this section, some classical results on the maximal monotonicity of the pointwise sum of two maximal monotone operators are revisited from the approximation and graph-convergence point of view. Given two maximal monotone operators A and B, the pointwise sum $A+B$ is still monotone. From Minty's Theorem, the maximal monotonicity of $A+B$ is equivalent to the solvability for any y in H of the equation:

$$u + Au + Bu \ni y. \tag{5.7}$$

The Brézis-Crandall-Pazy approach, which will be our guideline in this paper (even in the case where $A+B$ is not maximal monotone) consists in solving the approximate equation

$$u_\lambda + Au_\lambda + B_\lambda u_\lambda \ni y. \tag{5.8}$$

The equation (5.8) has always a solution because of the maximal monotonicity of the operator $A+B_\lambda$, which follows from the fact that B_λ is everywhere defined and continuous [29]. So doing, the problem of the maximality of the sum has been converted into a convergence problem: Under what conditions can one pass to the limit in the equation (5.8) as λ goes to zero?

Theorem 5.1. [31; Theorem 2.1] *The equation*

$$u + Au + Bu \ni y$$

has a solution if and only if the family $\{ B_\lambda u_\lambda \,|\, \lambda \to 0 \}$ remains bounded. When this condition is satisfied, the family $\{ u_\lambda \,|\, \lambda > 0 \}$ norm-converges to u as λ goes to 0 and the family $\{ B_\lambda u_\lambda \,|\, \lambda > 0 \}$ norm-converges to the element of minimal norm of the convex set $Bu \cap (y - u - Au)$ as λ goes to 0.

We are led to introduce the following definition:

Definition 5.2. A pair (A, B) of maximal monotone operators A and B satisfies the Brézis-Crandall-Pazy condition, if for any y in H the family $\{B_\lambda u_\lambda |\, \lambda > 0\}$, which is defined by (5.8), remains bounded in H.

We can reformulate the Brézis-Crandall-Pazy theorem as follows:

Proposition 5.3. *For any pair (A, B) of maximal monotone operators A and B the following statements are equivalent:*

(i) *The pair (A, B) satisfies the Brézis-Crandall-Pazy condition;*
(ii) *$A + B$ is a maximal monotone operator;*
(iii) $\mathrm{graph}-\lim\limits_{\mathcal{F}}(A + B_\mu) = A + B$;
(iv) $\mathrm{graph}-\lim\limits_{\mathcal{F}}(A + B_\mu) = \mathrm{graph}-\lim\limits_{\mathcal{F}}(A_\lambda + B) = \mathrm{graph}-\lim\limits_{\mathcal{F}}(A_\lambda + B_\mu) = A + B$;

and then the following equality holds: $A + B = A \underset{v}{+} B$.

This proposition says more than the initial result of Brézis-Crandall-Pazy and we need Theorem 6.1 below to conclude that $\mathrm{graph}-\lim\limits_{\substack{\lambda\to 0,\mu\to 0\\ \lambda>0,\mu>0}} A_\lambda + B_\mu = A + B$.

In view of Proposition 3.6, which states that A_λ converges to A , as λ goes to zero, for the graph-distance, it is a natural question to know if the convergence (iii) holds too for the graph-distance. To answer this question, we need to introduce a reinforced Brézis-Crandall-Pazy condition (see [10; Theorem 3.1] for further details): Denote by $m_{A,B} : H \to \mathbb{R}$ the finite valued function

$$m_{A,B}(y) := \limsup_{\lambda\to 0,\lambda>0} ||B_\lambda u_\lambda||.$$

Definition 5.4. A pair (A, B) of maximal monotone operators A and B satisfies the *uniform Brézis-Crandall-Pazy condition* if for any $\rho > 0$

$$M_{A,B}(\rho) := \sup_{||y||\le\rho} m_{A,B}(y) < +\infty.$$

In other words, the bound on the family $\{\, B_\lambda u_\lambda |\, \lambda > 0 \,\}$ has to be independant of y when y remains in a bounded subset of H.

Theorem 5.5. [10] *Suppose that a pair (A, B) of maximal monotone operators A and B satisfies the uniform Brézis-Crandall-Pazy condition. Then*

$$A + B = \mathrm{graph} - \mathrm{dist}-\lim_{\lambda\to 0}(A + B_\lambda).$$

Moreover, the following estimation holds:

$$d_{1,\rho}(A + B, A + B_\lambda) \le \frac{\lambda}{2} M_{A,B}(\rho).$$

Indeed, in most practical situations where the sum is maximal monotone, the uniform Brézis-Crandall-Pazy condition is satisfied. Let us illustrate this fact in the case of the *acute angle* condition:

$$\text{for all } v \in D(A) \text{ and all } \lambda > 0 \quad \langle Av, B_\lambda v\rangle \geq 0,$$

Proposition 5.6. *Suppose that A and B are two maximal monotone operators such that $D(A) \cap D(B) \neq \emptyset$ and which satisfy the acute angle condition. Then they verify the uniform Brézis-Crandall-Pazy condition with*

$$M_{A,B}(\rho) = \Big[2(\rho + ||v_o||) + ||\xi|| + ||\eta||\Big]^{1/2},$$

where $v_o \in D(A) \cap D(B), \xi \in Av_o$ and $\eta \in Bv_o$ have been taken arbitrary.

Proof. Let us multiply the equation (5.8) by $B_\lambda u_\lambda$ and use the acute angle condition. We obtain

$$||B_\lambda u_\lambda||^2 \leq ||y|| + ||u_\lambda||.$$

On the other hand, taking an arbitrary point v_o in D(A) ∩ D(B) and using the accretivity of A+ B_λ, we infer

$$||u_\lambda - v_o|| \leq ||y - (v_o + \xi + B_\lambda v_0)||.$$

Hence, $||B_\lambda u_\lambda||^2 \leq 2(||y|| + ||v_o||) + ||\xi|| + ||\eta||$. □

6. The case $\overline{A+B}$ maximal monotone: $A \underset{v}{+} B = \overline{A+B}$

Theorem 6.1. *Let A and B be two maximal monotone operators such that $\overline{A+B}$ is maximal monotone. Given any $y \in H$, let us denote, for each $(\lambda, \mu) \in \mathcal{I}$, by $u_{\lambda,\mu}$ the solution of*

$$u_{\lambda,\mu} + A_\lambda u_{\lambda,\mu} + B_\mu u_{\lambda,\mu} \ni y. \tag{6.9}$$

Then, the filtered family $\{u_{\lambda,\mu} | (\lambda,\mu) \in \mathcal{I}\}$ converges to the solution u of $u + \overline{A+B}(u) \ni y$.

In other words,

$$\text{graph}-\lim_{\mathcal{F}} (A_\lambda + B_\mu) = \overline{A+B}.$$

Proof. Take arbitrary $v \in D(A) \cap D(B)$, $\xi \in Av$ and $\eta \in Bv$. By virtue of (3.6) we have

$$B_\mu u_{\lambda,\mu} \in B(J_\mu^B u_{\lambda,\mu}),$$

and from the monotonicity of B we obtain

$$\langle \eta - B_\mu u_{\lambda,\mu}, v - J_\mu^B u_{\lambda,\mu}\rangle \geq 0. \tag{6.10}$$

Similarly we also have

$$\langle \xi - A_\lambda u_{\lambda,\mu}, v - J_\lambda^A u_{\lambda,\mu}\rangle \geq 0. \tag{6.11}$$

Using the relations $J_\mu^B u_{\lambda,\mu} = u_{\lambda,\mu} - \mu \cdot B_\mu u_{\lambda,\mu}$ and $J_\lambda^A u_{\lambda,\mu} = u_{\lambda,\mu} - \lambda \cdot A_\lambda u_{\lambda,\mu}$ we rewrite (6.10) and (6.11), as

$$\langle \eta - B_\mu u_{\lambda,\mu}, v - u_{\lambda,\mu} + \mu B_\mu u_{\lambda,\mu}\rangle \geq 0 \tag{6.12}$$

and

$$\langle \xi - A_\lambda u_{\lambda,\mu}, v - u_{\lambda,\mu} + \lambda A_\lambda u_{\lambda,\mu} \rangle \ \geq \ 0. \tag{6.13}$$

Then adding (6.12) and (6.13) we derive

$$\begin{aligned} &\langle \eta + \xi - (A_\lambda u_{\lambda,\mu} + B_\mu u_{\lambda,\mu}), v - u_{\lambda,\mu} \rangle \\ &+ \langle \eta - B_\mu u_{\lambda,\mu}, \mu B_\mu u_{\lambda,\mu} \rangle + \langle \xi - A_\lambda u_{\lambda,\mu}, \lambda A_\lambda u_{\lambda,\mu} \rangle \geq 0. \end{aligned} \tag{6.14}$$

If we combine (6.9) and (6.14) we obtain:

$$\begin{aligned} &\langle \eta + \xi + (u_{\lambda,\mu} - y), v - u_{\lambda,\mu} \rangle \\ &+ \langle \eta - B_\mu u_{\lambda,\mu}, \mu B_\mu u_{\lambda,\mu} \rangle + \langle \xi - A_\lambda u_{\lambda,\mu}, \lambda A_\lambda u_{\lambda,\mu} \rangle \geq 0, \end{aligned} \tag{6.15}$$

that is

$$\begin{aligned} &||u_{\lambda,\mu}||^2 + \lambda ||A_\lambda u_{\lambda,\mu}||^2 + \mu ||B_\mu u_{\lambda,\mu}||^2 \\ &\leq \langle u_{\lambda,\mu}, y - \eta - \xi + v \rangle + \langle \eta, \mu B_\mu u_{\lambda,\mu}, \rangle + \langle \xi, \lambda A_\lambda u_{\lambda,\mu} \rangle + \langle v, \xi + \eta - y \rangle. \end{aligned} \tag{6.16}$$

From (6.16) we obtain the existence of a constant C such that

$$||u_{\lambda,\mu}|| \ \leq \ C \tag{6.17}$$

$$||\mu^{\frac{1}{2}} B_\mu u_{\lambda,\mu}|| \ \leq \ C \tag{6.18}$$

and

$$||\lambda^{\frac{1}{2}} A_\lambda u_{\lambda,\mu}|| \ \leq \ C. \tag{6.19}$$

Let $\hat{u}$ be a weak limit point in H of the family $\{u_{\lambda,\mu} | \, (\lambda, \mu) \in \mathcal{I}\}$. Clearly by (6.18) and (6.19), $\lim_{\mathcal{F}} \mu B_\mu u_{\lambda,\mu} = \lim_{\mathcal{F}} \lambda A_\lambda u_{\lambda,\mu} = 0$. Hence, on passing to the limit in (6.16), and using the lower semicontinuity of $|| \cdot ||^2$ for the weak topology we obtain

$$||\hat{u}||^2 \leq \langle \hat{u}, y - \xi - \eta + v \rangle + \langle v, \xi + \eta - y \rangle,$$

that is

$$\langle v - \hat{u}, \xi + \eta - y + \hat{u} \rangle \geq 0.$$

This being true for any $\xi + \eta \in (A + B)v$, it follows that for any $\theta \in (\overline{A + B})(v)$

$$\langle v - \hat{u}, \theta - y + \hat{u} \rangle \geq 0.$$

From the maximal monotonicity of $\overline{A + B}$, we infer $y - \hat{u} \in (\overline{A + B})(\hat{u})$, that is

$$\hat{u} + \overline{A + B}(\hat{u}) \ni y.$$

Hence, $\hat{u} = J_\lambda^{\overline{A+B}}(y)$ and the whole family $\{u_{\lambda,\mu} | \, (\lambda, \mu) \in \mathcal{I}\}$ weakly converges to this element. Returning to (6.16) we derive

$$\limsup_{\mathcal{F}} ||u_{\lambda,\mu}||^2 \leq \langle \hat{u}, y - \xi - \eta + v \rangle + \langle v, \xi + \eta - y \rangle. \tag{6.20}$$

This being true for any $v \in D(A) \cap D(B)$, by letting v tend to $\hat{u}$, we infer

$$\limsup_{\mathcal{F}} ||u_{\lambda,\mu}||^2 \leq ||\hat{u}||^2. \tag{6.21}$$

(6.21) yields $\lim_{\mathcal{F}} ||u_{\lambda,\mu}|| = ||\hat{u}||$ and therefore, the family $\{u_{\lambda,\mu} | (\lambda, \mu) \in \mathcal{I}\}$ norm converges to $\hat{u} = J_\lambda^{\overline{A+B}}(y)$.

To conclude the proof, we need also to consider the two cases $A_\lambda + B$ and $A + B_\mu$ for λ and μ strictly positive. Indeed, in theses cases the proof works in the same way (c.f Brézis [29], Clément & Egberts [37]). □

7. The case of subdifferentials of convex functions

Let us first recall that for any convex lower semicontinuous function $f : H \to \mathbb{R} \cup \{\infty\}$,

$$\operatorname{dom} f := \{x \in H |\, f(x) < +\infty\}$$

is called the *domain* of f and that f is declared *proper* if $\operatorname{dom} f \neq \emptyset$. It is well-known that for any convex lower semicontinuous proper function $f : H \to \mathbb{R} \cup \{\infty\}$ its subdifferential operator ∂f is a maximal monotone operator (see [60] for further details). Graph-convergence of sequences of subdifferential operators has been first investigated in the case of the set-convergence of the corresponding graphs, the set-convergence being taken in the Kuratowski-Painlevé sense (equality between the topological limsup and liminf), which corresponds to the pointwise convergence of the resolvent operators. Denoting by "s" the topology of the norm in H and by "s×s " the product topology on $H \times H$ and recalling that the Mosco epiconvergence is the set-convergence of the epigraphs both for the strong and the weak topologies of $H \times H$ (see [47], [4]) this result can be stated as follows (indeed it is formulated below in a slightly more general setting, the space being only assumed to be reflexive):

Theorem 7.1. [2, 4] *Let X be a reflexive Banach space. For any sequence $\{f, f_n$: $X \to \mathbb{R} \cup \{+\infty\} | n \in \mathbb{N}\}$ of lower semicontinuous convex proper functions, the two following properties are equivalent :*

(i) $f = \mathrm{M} - \mathrm{epi} - \lim_{n\to\infty} f_n$;

(ii) $\operatorname{graph} \partial f = s \times s - \lim_{n\to\infty} \operatorname{graph} \partial f_n$ *and the "normalization condition (NC)".*

The normalization condition comes from the fact that f is determined by ∂f up to an additive constant and is described below :

$$\begin{cases} \exists (x_0, x_0^*) \in \operatorname{graph} \partial f,\ \exists (x_{0n}, x_{0n}^*) \in \operatorname{graph} \partial f_n \quad \text{for every } n \in \mathbb{N} \\ \text{such that} \\ x_{0n} \xrightarrow{s} x_0, x_{0n}^* \xrightarrow{s} x_0^* \ \text{ and }\ f_n(x_{0n}) \to f(x_0). \end{cases}$$

The recent monograph of Aubin and Frankowska [17] gives a convenient access to this result. The extension of Theorem 7.1 to arbitrary Banach spaces has been recently obtained

by Attouch & Beer [8] and requires the introduction of the notion of slice convergence for sequences of convex functions, a notion which has been introduced by G. Beer precisely to extend the Mosco convergence to general Banach spaces.
We recall that, given X a normed linear space and f a proper extended real-valued function from X into $\mathbb{R} \cup \{+\infty\}$, the *Moreau-Yosida approximation* of index λ of f is the function f_λ defined by:

$$f_\lambda(x) = \inf_{u \in X} \{f(u) + \frac{1}{2\lambda}||x-u||^2\}. \tag{7.22}$$

When X is an Hilbert space and f is a convex lower semicontinuous proper function, then f_λ is a continuously differentiable function which derivative is precisely the Yosida approximation A_λ of the maximal monotone operator $A = \partial f$ (see [29], [13–15]):

$$(\partial f)_\lambda = \partial f_\lambda. \tag{7.23}$$

We can now state the following result:

Theorem 7.2. *Let $A = \partial f$ and $B = \partial g$ be the subdifferential operators of two convex lower semicontinuous proper functions f and $g : H \to \mathbb{R} \cup \{+\infty\}$. If $\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$, then the variational sum of ∂f and ∂g is a maximal monotone operator and is equal to $\partial(f+g)$:*

$$\partial f \underset{v}{+} \partial g = \partial(f+g).$$

Equivalently,

$$\text{graph} - \lim_{\mathcal{F}} [(\partial f)_\lambda + (\partial g)_\mu] = \text{graph} - \lim_{\mathcal{F}} \partial(f_\lambda + g_\mu) = \partial(f+g).$$

Proof. We first notice that, because of the assumption $\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$, the function $f+g$ is a convex lower semicontinuous proper function, and its subdifferential operator $\partial(f+g)$ is therefore by Rockafellar's Theorem a maximal monotone operator. By definition of the variational sum, we need to prove that $\lim_{\mathcal{F}} [(\partial f)_\lambda + (\partial g)_\mu]$ exists. Let us examine the operator $(\partial f)_\lambda + (\partial g)_\mu$. Since the mappings f_λ and g_μ $(\lambda > 0, \mu > 0)$ are everywhere continuous since $\mathcal{C}^{1,1}$, we know by the Moreau-Rockafellar Theorem that $\partial f_\lambda + \partial g_\mu = \partial(f_\lambda + g_\mu)$ $((\lambda, \mu) \in \mathcal{I})$. Since by (7.23), $(\partial f)_\lambda = \partial f_\lambda$ (respectively $(\partial g)_\mu = \partial g_\mu$) we derive,

$$(\partial f)_\lambda + (\partial g)_\mu = \partial(f_\lambda + g_\mu).$$

Since the filtered family $\{f_\lambda | \, \lambda \to 0\}$, (respectively $\{g_\mu | \, \mu \to 0\}$) monotonically increases to f, (respectively to g) as λ, (respectively μ) decreases to zero, by [11; Theorem 3.20] the family $\{f_\lambda + g_\mu | \, \lambda \to 0, \mu \to 0\}$ Mosco-converges to $f+g$ as λ and μ decrease to zero. We conclude thanks to Theorem 7.1 that $\lim_{\mathcal{F}} \partial(f_\lambda + g_\mu) = \partial(f+g)$, which yields the desired result. $\square$

In the following result we give a formulation of $\partial(f+g)$ in terms of ∂f and ∂g only assuming, as in Theorem 7.2, that $\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$. In particular, the pointwise sum $\partial f + \partial g$ is not assumed to be maximal monotone.

In the sequel, it is convenient to use the notation $\lim_{n\to\infty}\{\partial f(u_n)+\partial g(v_n)\}$ for $\lim_{n\to\infty}(u_n^*+v_n^*)$ with $u_n^*\in\partial f(u_n), v_n^*\in\partial g(v_n)$

Theorem 7.3. *Let f and $g : H\to \mathbb{R}\cup\{+\infty\}$ be two convex lower semicontinuous proper functions such that $\operatorname{dom} f\cap\operatorname{dom} g\neq\emptyset$. Then, for any $x\in\operatorname{dom}\partial(f+g)$, the following equality holds:*

$$\partial(f+g)(x)=$$

$$=\left\{\lim_{n\to\infty}\{\partial f(u_n)+\partial g(v_n)\}\mid x=\lim_{n\to\infty}u_n=\lim_{n\to\infty}v_n \text{ and } \liminf_{n\to\infty}\langle\partial g(v_n),u_n-v_n\rangle\geq 0\right\}.$$

Proof. Set $A:=\partial(f+g)$ and

$$B:=\left\{\lim_{n\to\infty}\{\partial f(u_n)+\partial g(v_n)\}\mid\ x=\lim_{n\to\infty}u_n=\lim_{n\to\infty}v_n \text{ and } \liminf_{n\to\infty}\langle\partial g(v_n),u_n-v_n\rangle\geq 0\right\}$$

a) Let us first prove the inclusion $B\supseteq A$. We know from Theorem 7.2 that

$$\text{graph}-\lim_{\mathcal{F}}\partial(f_\lambda+g_\mu)=\partial(f+g).$$

As a result, for any $x\in D(\partial(f+g))$ and $y\in\partial(f+g)(x)$, there exists some net $\{x_{\lambda,\mu}\mid(\lambda,\mu)\in\mathcal{I}\}$ norm converging to x such that $y=\lim_{\mathcal{F}}\partial(f_\lambda+g_\mu)(x_{\lambda,\mu})$. Since f_λ, (respectively g_μ) is everywhere defined and continuous, by the Moreau-Rockafellar Theorem on the additivity of the sum of subdifferentials,

$$y=\lim_{\mathcal{F}}[\partial f_\lambda(x_{\lambda,\mu})+\partial g_\mu(x_{\lambda,\mu})],$$

which by relation (3.6) gives

$$y=\lim_{\mathcal{F}}[\partial f(J_\lambda^f x_{\lambda,\mu})+\partial g(J_\mu^g x_{\lambda,\mu})]. \tag{7.24}$$

Set $u_{\lambda,\mu}:=J_\lambda^f x_{\lambda,\mu}$ and $v_{\lambda,\mu}:=J_\mu^g x_{\lambda,\mu}$. Let us first verify that $x=\lim_{\mathcal{F}}u_{\lambda,\mu}=\lim_{\mathcal{F}}v_{\lambda,\mu}$. To that end, we notice that, from the contraction property of J_μ^g,

$$||J_\mu^g x_{\lambda,\mu}-J_\mu^g x||\leq||x_{\lambda,\mu}-x||$$

and that the limit of $J_\mu^g x$, as μ goes to zero, is equal to the projection of x on the closure of the domain of g, that is x. Hence, $\lim_{\mathcal{F}}u_{\lambda,\mu}=x$ and similarly, $\lim_{\mathcal{F}}v_{\lambda,\mu}=x$. Let us now consider,

$$\begin{aligned}
\langle\partial g(v_{\lambda,\mu}),u_{\lambda,\mu}-v_{\lambda,\mu}\rangle &= \langle\partial g_\mu(x_{\lambda,\mu}),J_\lambda^f x_{\lambda,\mu}-J_\mu^g x_{\lambda,\mu}\rangle\\
&= \langle\partial g_\mu(x_{\lambda,\mu}),(J_\lambda^f x_{\lambda,\mu}-x_{\lambda,\mu})+(x_{\lambda,\mu}-J_\mu^g x_{\lambda,\mu})\rangle\\
&= \langle\partial g_\mu(x_{\lambda,\mu}),\mu\partial g_\mu(x_{\lambda,\mu})-\lambda\partial f_\lambda(x_{\lambda,\mu})\rangle\\
&\geq -\lambda\langle\partial g_\mu(x_{\lambda,\mu}),\partial f_\lambda(x_{\lambda,\mu})\rangle\\
&\geq -\lambda\langle\partial g_\mu(x_{\lambda,\mu})+\partial f_\lambda(x_{\lambda,\mu}),\partial f_\lambda(x_{\lambda,\mu})\rangle+\lambda|\partial f_\lambda(x_{\lambda,\mu})|^2\\
&\geq -\lambda|\partial g_\mu(x_{\lambda,\mu})+\partial f_\lambda(x_{\lambda,\mu})|.|\partial f_\lambda(x_{\lambda,\mu})|+\lambda|\partial f_\lambda(x_{\lambda,\mu})|^2.
\end{aligned}$$

Now use the elementary minorization:

$$t^2 - \alpha t \geq -\frac{\alpha^2}{4}$$

to obtain (take $t = |\partial f_\lambda(x_{\lambda,\mu})|$)

$$\langle \partial g(v_{\lambda,\mu}), u_{\lambda,\mu} - v_{\lambda,\mu}\rangle \geq -\frac{\lambda}{4}|\partial g_\mu(x_{\lambda,\mu}) + \partial f_\lambda(x_{\lambda,\mu})|^2.$$

Since by assumption, $\partial g_\mu(x_{\lambda,\mu}) + \partial f_\lambda(x_{\lambda,\mu})$ norm-converges to $y \in \partial(f+g)(x)$, and hence remains bounded, we eventually infer

$$\liminf_{\mathcal{F}} \langle \partial g(v_{\lambda,\mu}), u_{\lambda,\mu} - v_{\lambda,\mu}\rangle \geq 0.$$

b) Let us now prove the inclusion $A \supseteq B$. Take $u_\nu \in D(\partial f), v_\nu \in D(\partial g)$ such that

$$Bx = \lim_{\nu\to\infty} [\partial f(u_\nu) + \partial g(v_\nu)],$$

with $x = \lim_{\nu\to\infty} u_\nu = \lim_{\nu\to\infty} v_\nu$ and $\liminf_{\nu\to\infty} \langle \partial g(v_\nu), u_\nu - v_\nu\rangle \geq 0$.

The subdifferential inequality yields that for any $\xi \in H$

$$\begin{aligned} f(\xi) &\geq f(u_\nu) + \langle \partial f(u_\nu), \xi - u_\nu\rangle \\ g(\xi) &\geq g(v_\nu) + \langle \partial g(v_\nu), \xi - v_\nu\rangle. \end{aligned}$$

By adding these two inequalities we obtain

$$\begin{aligned} f(\xi) + g(\xi) &\geq f(u_\nu) + g(v_\nu) + \langle \partial f(u_\nu), \xi - u_\nu\rangle + \langle \partial g(v_\nu), \xi - v_\nu\rangle \\ &\geq f(u_\nu) + g(v_\nu) + \langle \partial f(u_\nu) + \partial g(v_\nu), \xi - u_\nu\rangle + \langle \partial g(v_\nu), u_\nu - v_\nu\rangle. \end{aligned}$$

Then passing to the limit and using the definition of Bx we obtain

$$\begin{aligned} f(\xi) + g(\xi) &\geq \liminf_{\nu\to\infty} [f(u_\nu) + g(v_\nu)] + \liminf_{\nu\to\infty} \langle \partial f(u_\nu) + \partial g(v_\nu), \xi - u_\nu\rangle + \\ &\qquad \liminf_{\nu\to\infty} \langle \partial g(v_\nu), u_\nu - v_\nu\rangle \\ &\geq f(x) + g(x) + \langle Bx, \xi - x\rangle. \end{aligned}$$

The above inequality being true for any $\xi \in H$, we infer $Ax = \partial(f+g)(x) \supseteq Bx$ which completes the proof of the theorem. $\square$

Remark 7.4. Let us notice that the sign condition " $\liminf \langle \partial g(v_\nu), u_\nu - v_\nu\rangle \geq 0$ " which at a first glance is not symmetric is indeed quite symmetric.

Let us write

$$\langle \partial f(u_\nu), v_\nu - u_\nu\rangle = \langle \partial f(u_\nu) + \partial g(v_\nu), v_\nu - u_\nu\rangle + \langle \partial g(v_\nu), u_\nu - v_\nu\rangle.$$

Hence,

$$\liminf_{\nu\to\infty}\langle \partial f(u_\nu), v_\nu - u_\nu\rangle = \lim_{\nu\to\infty}\langle \partial f(u_\nu)+\partial g(v_\nu), v_\nu - u_\nu\rangle + \liminf_{\nu\to\infty}\langle \partial g(v_\nu), u_\nu - v_\nu\rangle$$

is nonnegative (the first term of the right hand side is equal to zero and the second term is nonnegative).
Let us now state the following result from Attouch, Ndoutoume and Théra that relates the convergence for the epi-distance of a sequence of convex functions to the convergence for the graph-distance of the associated sequence of subdifferentials.

Theorem 7.5. [12; Theorem 2.3] *Let X be a Banach space. For any sequence $\{f, f_n|$ $X \to \mathbb{R}\cup\{+\infty\}\}$ of lower semicontinuous convex proper functions, the following implication* (i) $\Rightarrow$ (ii) *holds :*

(i) $f = \mathrm{epi}-\mathrm{dist}-\lim \mathrm{f_n}$;
(ii) $\partial f = \text{graph-dist}\lim \partial f_n$ *and the " normalization condition (NC) ".*

If furthermore X is super-reflexive, then the converse implication holds, that is, (i) *and* (ii) *are equivalent.*

Theorem 7.4 is a topological result. Indeed, in the Hilbert space setting one can give a metric (quantitative) version of it and prove an Hölder continuous type estimate which extends the result of Schultz :

Theorem 7.6. [15; Theorem 5.2] *Let f and g be two convex lower semicontinuous proper functions from H into $\mathbb{R}\cup\{+\infty\}$. To any $\rho > \max[d(0, \mathrm{epi}\, f), d(0, \mathrm{epi}\, g)]$ there correspond some constants γ and κ (that depend on ρ) such that*

$$\mathrm{haus}_\rho(\partial f, \partial g) \le \kappa[\mathrm{haus}_\gamma(f,g)]^{1/2}.$$

In turn, Theorem 7.5 has been extended to the case of general Banach spaces by Riahi [57] and by Aze & Penot [20] , the main ingredient of the proof being the Ekeland's $\epsilon-$variational principle.
Theorem 7.4 and the estimation below of independant interest show the convergence property of the Yosida approximation (Proposition 3.6) in a different light :

Proposition 7.7. [15; Lemma 3.6] *Let X be a normed linear space and f a proper extended real-valued function from X into $\mathbb{R}\cup\{+\infty\}$, minorized by $-\alpha(1+||\cdot||^2)$ for some $\alpha \ge 0$. Then, for any $0 < \lambda < \frac{1}{4\alpha}$ and $\rho \ge 0$,*

$$\mathrm{haus}_\rho(f_\lambda, f) \le C\lambda^{1/2}$$

where the "constant" C is given by

$$C = \left\{\frac{2(2\alpha\rho^2+\rho+\alpha)}{1-4\alpha\lambda}\right\}^{1/2}.$$

8. The role of the approximation

It is a natural question to ask whether the Yosida approximation plays a particular role in this theory. In other words, what happens when we replace A_λ (respectively B_μ) in the definition of the variational sum by A_ν (respectively B_θ) such that:

1) $A = \text{graph} - \lim\limits_{\nu \to +\infty} A_\nu$, (respectively, $B = \text{graph} - \lim\limits_{\theta \to +\infty} B_\theta$);

2) A_ν and B_θ are everywhere defined and Lipschitzian.

Figure 1

For example, if we take for A, A_ν (respectively B, B_θ) the subdifferentials of some convex proper lower semicontinuous extended real-valued functions, then the previous question may be formulated as follows:

Let $\{\Phi_n | n \in \mathbb{N}\}, \{\Psi_n | n \in \mathbb{N}\}$ be two sequences of convex functions everywhere defined on an Hilbert space H and Mosco-converging respectively to the functions Φ and Ψ. Does the sequence $\{\Phi_n + \Psi_n |\ n \in \mathbb{N}\}$ Mosco-converge to $\Phi + \Psi$?

The answer is generally negative as the following example shows:

Take $H = \mathbb{R}^2$ and define C_n and D_n as in the Figure 1.

$$C_n := \ \{(x, y) \in \mathbb{R}^2 | \, x \leq 0,\ y \leq -t_n x\} \text{ and } D_n := \ \{(x, y) \in \mathbb{R}^2 | \, x \geq 0,\ y \leq t_n x\}.$$

Let $\{t_n | n \in \mathbb{N}\}$ be a sequence of real numbers tending to $+\infty$. Set

$$\Phi_n = \frac{1}{2\lambda_n} \mathrm{dist}^2(., \mathrm{C}_n) = (\mathrm{I}_{\mathrm{C}_n})_{\lambda_n} \text{ and } \Psi_n = \frac{1}{2\lambda_n} \mathrm{dist}^2(., \mathrm{D}_n) = (\mathrm{I}_{\mathrm{D}_n})_{\lambda_n},$$

where I_{C_n} (respectively, I_{D_n}) stands for the indicator function of the sets C_n (respectively, D_n). Since, Φ_n (respectively, Ψ_n) is the Moreau-Yosida approximate of I_{C_n} (respectively, I_{D_n}), it is $\mathrm{C}^{1,1}$, hence continuous on $\mathbb{R}^2$. Set $k := \lim\limits_{n \to +\infty} \lambda_n t_n^2$. An easy calculation shows that: $\{\Phi_n | n \in \mathbb{N}\}$ (respectively, $\{\Psi_n | n \in \mathbb{N}\}$) epiconverges to the indicator $\Phi = I_C$

(respectively, $\Psi = I_D$) of the set C:= $\{(x,y) \in \mathbb{R}^2 | \, x \leq 0\}$ (respectively, D: = $\{(x,y) \in \mathbb{R}^2 | \, x \geq 0\}$), and

1) if k $= +\infty$, then $\Phi_n + \Psi_n$ epiconverges to $\mathrm{I}_{C\cap D}$;

2) if k $= 0$, then $\Phi_n + \Psi_n$ epiconverges to $\mathrm{I}_{\{x=0, y\leq 0\}}$;

3) if k$\in (0, +\infty)$, then $\Phi_n + \Psi_n$ epiconverges to a mapping which depends on k.

The limit function is equal to $+\infty$ outside $\{0\} \times \mathbb{R}$ and its restriction to $\{0\} \times \mathbb{R}$ has the profile (which depends on k), as described in Figure 2.

In terms of operators, set $A_n := \partial \Phi_n$ and $B_n := \partial \Psi_n$. By virtue of the Attouch Theorem on the continuity (for convex functions) of the operation $f \rightrightarrows \partial f$, A_n (respectively, B_n) graph-converges to $\partial \Phi := A$ (respectively, $\partial \Psi := B$). Hence, $A_n + B_n$ which is a maximal operator since it is equal to $\partial(\Phi_n + \Psi_n)$ necessarily graph-converges to a maximal monotone operator C_k. According to Rockafellar's Theorem ([60]), C_k is therefore the subdifferential of a convex function F_k, which depends on the parameter k, linked to the chosen approximation. We have $C_k = \partial(\Phi + \Psi) = A + B = A \underset{v}{+} B$ only whenever $k = +\infty$.

As a result, the Yosida approximation plays a central role in the theory developed in this presentation.

Figure 2

9. An example: The Schrödinger equations

The motivation for the introduction of various generalizations of the concept of sum for operators comes in particular from the Schrödinger equations and problems arising in quantum theory ([30], [61], [62], [66]). Let us briefly recall the mathematical setting of these equations. Let $\Omega \subset \mathbb{R}^N$ be an open subset of $\mathbb{R}^N$ (possibly unbounded). Let $V : \Omega \to \mathbb{R}^+$ be a locally integrable function with some singularities (for example, $V(x) = 1/||x||^p$). The differential operator attached to the Schrödinger equations can be formally written as

$$Au = -\Delta u + Vu.$$

This operator, clearly appears as a "sum" of two operators

$$Bu := -\Delta u \text{ and } Cu := Vu.$$

A natural objective is to define adequately the domain of A in order to obtain an operator A which is maximal monotone, self-adjoint. Equivalently, this amounts to define in which sense one has to add the two operators B and C. Take $H = L^2(\Omega)$, $B = \partial\Phi$, and $C = \partial\Psi$ the subdifferential operators of the convex lower semicontinuous proper functions

$$\Phi(u) := \begin{cases} \frac{1}{2}\int_\Omega |Du|^2 dx & \text{if } u \in H_0^1(\Omega) \\ +\infty & \text{on } L^2(\Omega)\setminus H_0^1(\Omega) \end{cases}$$

and

$$\Psi(u) := \begin{cases} \frac{1}{2}\int_\Omega Vu^2 dx & \text{if } Vu^2 \in L^1(\Omega) \\ +\infty & \text{elsewhere.} \end{cases}$$

$$\begin{cases} Bu & := \ -\Delta u \text{ with} \\ D(B) & := \ H^2(\Omega)\cap H_0^1(\Omega) \end{cases}$$

and

$$\begin{cases} Cu & := \ Vu \text{ with} \\ D(C) & := \ \{u \in L^2(\Omega) |\, Vu \in L^2(\Omega)\}. \end{cases}$$

A naive approach which consists to take for $A = B + C$ the pointwise sum of B and C is not well adapted to the situation. If u belongs to $D(B)\cap D(C)$, then $u \in H^2(\Omega)$ and $Vu \in L^2(\Omega)$, these two conditions are usually incompatible. For example, if $N = 2, u \in H^2(\Omega)$ implies u continuous and $Vu \in L^2(\Omega)$ with u continuous and $V(x) = \dfrac{1}{||x||^{1/2}}$ requires that $u(0) = 0$. This is a too strong requirement with respect to the physical conditions. On the opposite, the variational approach developped in this paper consists in taking

$$\hat{A} := B \underset{v}{+} C.$$

Since $B = \partial\Phi$ and $C = \partial\Psi$ are subdifferential of convex lower semicontinuous proper functions, by Theorem 7.2

$$\hat{A} = \partial(\Phi + \Psi).$$

Let us compute $\hat{A}$. Indeed, if we follow the approximation method which led us to the above concept

$$\hat{A} = \text{graph}-\lim_{\lambda\to 0}(\partial\Phi + \partial\Psi_\lambda) = \text{graph}-\lim_{\lambda\to 0}(B + C_\lambda).$$

An elementary computation gives

$$C_\lambda u = \frac{V}{1+\lambda V}u.$$

Let us introduce $T_\lambda : \mathbb{R}^+ \to \mathbb{R}^+$ the *truncation mapping* defined by $T_\lambda(r) := \dfrac{r}{1+\lambda r}$ and set $K_\lambda(r) := \inf(r, 1/\lambda)$.

Figure 3

Then, $C_\lambda u = T_\lambda(V)u$ is a continuous linear operator on $L^2(\Omega)$, and in order to pass from C to C_λ, we have "smoothly" truncated V, so that $T_\lambda(V)$ belongs to $L^\infty(\Omega)$; indeed, $0 \leq T_\lambda(V) \leq 1/\lambda$. This method is very closely related to the direct truncation method of V by $K_\lambda(V) = \inf(V, 1/\lambda)$ followed by Brézis [30]. Note that T_λ and K_λ satisfy $0 \leq T_\lambda \leq K_\lambda$ and they are tangent at zero and $+\infty$. Using the resolvent formulation of the graph-convergence, we have to compute the limit u of u_λ, where u_λ is the solution of

$$u_\lambda + Bu_\lambda + C_\lambda u_\lambda = f.$$

Note that we know that u_λ converges to u which satisfies

$$u + (B \underset{v}{+} C)(u) = f.$$

So we have just to identify the limit equation satisfied by u. In the following, in order to simplify the expository, we suppose that Ω is bounded. We have

$$\begin{cases} u_\lambda - \Delta u_\lambda + T_\lambda(V)u_\lambda = f & \text{on } \Omega \\ u_\lambda = 0 & \text{on } \partial\Omega. \end{cases} \tag{9.25}$$

Multiplying (9.25) by u_λ and integrating by part classicaly yield the following estimations:

$$\|u_\lambda\|_{H_0^1(\Omega)} \leq C$$

and

$$\int_\Omega T_\lambda(V)u_\lambda^2 \leq C.$$

On relabeling (we still denote by u_λ the extracted subnet) u_λ tends weakly to u in $H_0^1(\Omega)$ and almost everywhere (by the Rellich-Kondrakov compact embedding of H_0^1 into L^2). Using Fatou's lemma, and noticing that $T_\lambda(V) \to V$ a.e., we derive

$$\int_\Omega Vu^2 \leq \liminf_{\lambda \to 0} \int_\Omega T_\lambda(V)u_\lambda^2 < +\infty.$$

From,

$$\int_\Omega Vu \leq \int_\Omega V(1+u^2)dx < +\infty,$$

we infer $Vu \in L^1(\Omega)$ if $V \in L^1(\Omega)$ and $Vu \in L^1_{\rm loc}(\Omega)$ if $V \in L^1_{\rm loc}(\Omega)$. Hence, $T_\lambda(V)u_\lambda \to Vu$ a.e. . The difficulty is to prove that the convergence holds in the distribution sense. To that end we use the argument from Strauss [66] and prove, in order to apply the Vitali Theorem, that the net $\{T_\lambda(V)u_\lambda |\, \lambda \to 0\}$ is equi-integrable. Indeed,

$$\begin{aligned} T_\lambda(V)|u_\lambda| &= T_\lambda(V)(\tfrac{1}{\sqrt{R}}|u_\lambda|\sqrt{R}) \\ &\leq \tfrac{1}{2}T_\lambda(V)(\tfrac{1}{R}u_\lambda^2 + R). \end{aligned}$$

Hence, for each $\Delta \subseteq \Omega$ we have

$$\left\{ \begin{aligned} \int_\Delta T_\lambda(V)|u_\lambda| &\leq \tfrac{1}{2R}\textstyle\int_\Delta T_\lambda(V)u_\lambda^2 + \tfrac{R}{2}\int_\Delta T_\lambda(V) \\ &\leq \frac{C}{2R} + \frac{R}{2}\int_\Delta V. \end{aligned} \right. \tag{9.26}$$

The equi-integrability of the net $\{T_\lambda(V)u_\lambda |\, \lambda \to 0\}$ follows easily from (9.26). By passing to the limit on (9.25) we obtain:

$$\left\{ \begin{array}{ll} u - \Delta u + Vu = f & \text{on } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{array} \right. \tag{9.27}$$

with $u \in H^1_0(\Omega)$, $\displaystyle\int_\Omega Vu^2 < +\infty$, $Vu \in L^1_{\rm loc}(\Omega)$ and $-\Delta u + Vu \in L^2(\Omega)$. Let us summarize the above results in the following theorem:

Theorem 9.1. *Given $V \in L^1_{\rm loc}(\Omega), V \geq 0$ a.e., let us define*

$$\left\{ \begin{array}{l} \hat{A}u = -\Delta u + Vu \\ D(\hat{A}) = \{u \in H^1_0(\Omega) |\, Vu^2 \in L^1(\Omega),\ -\Delta u + Vu \in L^2(\Omega)\}. \end{array} \right.$$

Then $\hat{A}$ is a self-adjoint linear maximal monotone operator in $L^2(\Omega)$, and $\hat{A} = \partial\Phi \underset{v}{+} \partial\Psi$ $= \partial(\Phi + \Psi)$, where $\Phi(u) = \dfrac{1}{2}\displaystyle\int_\Omega |Du|^2$ and $\Psi(u) = \dfrac{1}{2}\displaystyle\int_\Omega Vu^2$.

As a result, the variational sum approach provides the same operator than the one obtained by Simon [61, 62], Kato [41] and Brézis [30].

Note, in this direction, the following result obtained by Brézis in the case where $\Omega = \mathbb{R}^N$: $\hat{A} = A^\sharp$ with

$$\left\{ \begin{array}{l} A^\sharp u = -\Delta u + Vu \\ D(A^\sharp) = \{u \in L^2(\Omega) |\, Vu \in L^1_{\rm loc}(\Omega), -\Delta u + Vu \in L^2(\Omega)\}. \end{array} \right.$$

References

[1] H. Attouch: Convergence de fonctions convexes, des sous-différentiels et semi-groupes associés, C.R. Acad. Sci. Paris 284, 539-542, 1977.

[2] H. Attouch: Convergence de fonctionnelles convexes, Journées d'Analyse Non Linéaire, Proceedings, Besançon, France, édité par Ph. Bénilan et J. Robert, Lecture Notes in Mathematics 665, 1977.

[3] H. Attouch: Familles d'opérateurs maximaux monotones et mesurabilité , Ann. Mat. Pura Appl.,(4) 120, 35-111, 1979.

[4] H. Attouch: On the maximality of the sum of two maximal monotone operators, Nonlinear Analysis, Theory, Methods & Applications, Vol. 5, n^0 2, pp 143-147, 1981.

[5] H. Attouch: Variational convergence for functions and operators, Applicable Maths. Series, Pitman, London, 1984.

[6] H. Attouch, D. Azé, R. Wets: Convergence of convex-concave saddle functions, Annales de l'Institut H. Poincaré, Analyse non linéaire, 5, 532-572, 1988.

[7] H. Attouch, J.-B. Baillon, M. Théra: Somme variationelle d'opérateurs maximaux monotones, C. R. Acad. Sci. Paris, 317, n^0 5, 485-490, 1993.

[8] H. Attouch, G. Beer: On the convergence of subdifferentials, Arkiv der Mathematik , Vol 60, 389-400, 1993.

[9] H. Attouch, R. Lucchetti, R. Wets: The topology of the ρ-Hausdorff distance, Annali Mat. Pura. Appl., IV, Vol.CLX, 303-320, 1991.

[10] H. Attouch, A. Moudafi, H. Riahi: Quantitative stability analysis for maximal monotone operators and semigroups of contractions, Nonlinear Anal. Vol. 21, n^0 9, 497-723, 1993.

[11] H. Attouch, M. Théra: Convergences en analyse multivoque et unilatérale, MATAPLI, Bulletin de liaison n^0 36, 23-40, 1993.

[12] H. Attouch, J.-L. Ndoutoume, M. Théra: Epigraphical convergence of functions and convergence of their derivatives in Banach spaces, Séminaire d'Analyse Convexe de Montpellier, Vol. 20, exposé 9, 1990.

[13] H. Attouch, R. Wets: Isometries for the Legendre-Fenchel transform, Transactions of the American Mathematical Society, 296, 33-60, 1986.

[14] H. Attouch, R. Wets: Quantitative stability of variational systems : I. The epigraphical distance, Transactions of the American Mathematical Society, 328, n^0 2, 695-729 1991.

[15] H. Attouch, R. Wets: Epigraphical analysis, Analyse non linéaire, eds. H. Attouch, J.P. Aubin, F. Clarke, I. Ekeland, 73-100, Gauthier-Villars, Paris, 1989.

[16] J-.P. Aubin, I. Ekeland: Applied Nonlinear Analysis, Wiley, New-York 1984.

[17] J.-P. Aubin, H. Frankowska: Set-valued Analysis, Birkhäuser, 1990.

[18] J.-P. Aubin, R. Wets: Stable approximations of set-valued maps, Ann. Inst. Henri Poincaré, Vol 5, n^0 6, 519-535, 1988.

[19] D. Azé: Rapidité de convergence pour une somme de fonctions convexes, Publication Avamac, Université de Perpignan, Vol. I, 1986.

[20] D. Azé, J.-P. Penot: Recent Quantitative Results about the Convergence of sets and Functions, Functional analysis and approximation, P.-L. Papini ed., Pitagora Editrice, Bologna, 90-110, 1988.

[21] D. Azé, J.-P. Penot: Operations on convergent families of sets and functions, Optimization, 21, 4, 521-534, 1990.

[22] J.-B. Baillon, Ph. Clément: Examples of unbounded imaginary powers of operators, J. Func. Anal., 100, 419-434, 1991.

[23] H. T. Banks, S. Reich, I. G. Rosen: An approximation theory for the identification of nonlinear distributed parameter systems. W.P. Lefschetz Center, Brown Univ. Providence, 1988.

[24] G. Beer: Conjugate convex functions and the epidistance topology, Proc. Amer. Math. Soc., 108, 117-128, 1990.

[25] G. Beer: Topologies on Closed and Closed Convex Sets, Kluwer, 1993.

[26] G. Beer, R. Lucchetti: Convex Optimization and the Epi-distance Topology, Trans. Amer. Math. Soc. 327, 795-813,1991.

[27] G. Beer, M. Théra: Attouch-Wets convergence and a differential operator for convex functions, to appear in Proc. Amer. Math. Soc. 1993.

[28] Ph. Benilan: Une remarque sur la convergence des semi-groupes non linéaires, C. R. Acad. Sci. 272, 1182-1184, 1971.

[29] H. Brézis: Opérateurs maximaux monotones et semi groupes de contractions dans les espaces de Hilbert, North Holland, 1973.

[30] H. Brézis: Cours Université Paris 6, 1978.

[31] H. Brézis, M.G. Crandall, A. Pazy: Perturbations of nonlinear maximal monotone sets in Banach spaces, Comm. pure and Appl. Math., 23, 123-144, 1970.

[32] H. Brézis, A. Pazy: Convergence and approximation of nonlinear semi-groups in Banach spaces, J. Funct. Anal., n^0 9, 63-74, 1972.

[33] F. Browder: Nonlinear operators and nonlinear equations of evolution. Proc. Symp. Pure Math. 18, Part 2, Amer. Math. Soc., 1976.

[34] Z. Chbani: Caractérisation de la convergence d'intégrales définies à partir d'opérateurs elliptiques par la convergence des coefficients, Séminaire d'Analyse Convexe de Montpellier, Vol 21, exposé 12, 1991.

[35] P. R. Chernoff: Product formulas, nonlinear semigroups, and addition of unbounded operators, Mem. Amer. Math. Soc., 140, 1-121, 1974.

[36] V. Chiado Piat, G. Dal Maso, Defranceschi: G-convergence of monotone operators, Annales de l'Institut H. Poincaré, Analyse non linéaire, Vol 7, n^0 3, 123-160, 1990.

[37] Ph. Clément, P. Egberts: On the sum of two maximal monotone operators, Differential and Integral Equations, Vol. 3 . n^0 6, 1127-1138, 1990.

[38] M. Crandall: Semi-groups of nonlinear transformations in Banach spaces, Contributions to nonlinear functional analysis, E. Zarantonello ed., Acad. Press, 1971.

[39] G. Da Prato, P. Grisvard: Sommes d'opérateurs linéaires et équations différentielles opé rationnelles, J. Math. Pures Appl., 54, 305-387, 1975.

[40] A. Haraux: Nonlinear evolution equations- Global behavior of solutions, Lecture notes in Mathematics 841, Springer-Verlag, 1981.

[41] T. Kato: Perturbation theory for linear operators, Springer, Berlin, 1966.

[42] D. Kravvaritis, N. Papageorgiou: Multivalued perturbations of subdifferential type evolution

equations in Hilbert spaces, J. of Dif. Equations, 76, 2, 238-255, 1988.

[43] F. Kubo: Conditional expectations and operations derived from network connections, J. of Math. Anal. and Appl. 80, 477-489, 1981.

[44] M. Lapidus: Formules de Trotter et calcul opérationnel de Feynman, Thèse d' Etat, Université Paris VI, Juin 1986.

[45] B. Lemaire: Coupling optimization methods and variational convergence, trends in mathematical optimization. Series of Num. Math., K.H. Hoffmann, J.B. Hiriart-Urruty, C. Lemarechal, J. Zowe eds, Birkhäuser, Basel, Vol 84, 163-179, 1988.

[46] M. Matzeu: Su un tipo di continuità dell'operatore subdifferenziale, Boll. U.M.I. 11,107-124, 1975.

[47] U. Mosco: Convergence of convex sets and of solutions of variational inequalities, Advances in Math. 3, 510-585, 1969.

[48] K. Mouallif: Convergence variationnelle et méthodes perturbées pour les problèmes d'optimisation et de point selle, Thèse d'État, Université de Liège, 1989.

[49] A. Moudafi: Perturbation de la méthode de l'inverse partiel. Publication Avamac 87- 05, Université de Perpignan, Vol. II, 1987.

[50] J.-L. Ndoutoume: Epi-convergence en calcul différentiel gé néralisé. Résultats de convergence et approximation, Thè se, Université de Perpignan, 1987.

[51] J.-L. Ndoutoume, M. Théra: Generalized second-order derivatives of convex functions in reflexive Banach spaces, to appear in J. of Austr. Math. Soc.

[52] N. Papageorgiou, A. Kandikalis: Convergence in approximation and nonsmooth analysis, J. Appr. Theory, 49, 41-54, 1987.

[53] G.B. Passty: The parallel sum of nonlinear monotone operators, Nonlinear Analysis, Theory, Methods & applications, Vol. 10, n^0. 3, 215-227, 1986.

[54] J.-P. Penot: Topologies and convergences on the space of convex functions, Preprint. Univ. Pau.

[55] J.-P. Penot: The cosmic Hausdorff topology, the bounded-Hausdorff topology and continuity of polarity, Proc. Amer. Math. Soc. (to appear).

[56] S. Reich: Convergence and approximation of nonlinear semigroups, J. Math. Anal. Appl., 76, 77-83, 1980.

[57] H. Riahi: Quelques résultats de stabilité en analyse épigraphique : approche topologique et quantitative. Thèse, Université de Montpellier, 1989.

[58] R.T. Rockafellar: Conjugate duality and optimization, SIAM Coll. Regional Conference Series in Applied Mathematics $n^0$16, Philadelphia, 1976.

[59] R.T. Rockafellar: Maximal monotone relations and the second derivative of non-smooth functions, Annales de l'Institut H. Poincaré, Analyse non linéaire, 2, 510-585, 1985.

[60] R.T. Rockafellar: On the Maximal Monotonicity of the Subdifferential Mapping, Pacific Journal of Mathematics, 33, 209-216, 1970.

[61] B. Simon: Quantum Mechanics for Hamiltonians defined as Quadratic Forms, Princeton, N. J. 1973.

[62] B. Simon: Schrödinger semigroups, Bull. Amer. Math., Soc., 7, 447-526, 1982.

[63] P. Shunmugaraj, D.V. Pai: On stability of approximate solutions of minimization problems, W.P. Indian Institute of Technology, Bombay, to appear in Numerical Functional Analysis and Optimization.

[64] Y. Sonntag: Convergence au sens de Mosco; théorie et applications à l'approximation des solutions d'inéquations. Thèse d'Etat. Université de Provence, 1982.

[65] M. Soueycatt: Analyse epi/hypo-graphique des problèmes de points selles, Thèse Université de Montpellier, 1991.

[66] A. Strauss: The nonlinear Schrödinger equation, in Contempory Developments in Continuum Mechanics and PDE edited by G. de la Penha and L. Mediros, 452-465, North Holland, Amsterdam, 1978.

[67] R. M. Torrejon: Perturbation of nonlinear operators and their parallel sum, Nonlinear Analysis, Theory, Methods & Applications, Vol.10, n^0 5, 483-490, 1986.

[68] P. Tossings: Convergence variationnelle et opérateurs maximaux monotones, W.P. n^0 90014, Université de Liège, 1990.

[69] H. P. Trotter: On the product of semigroups of operators, Proc. Amer. Math. Soc., 8, 545-551, 1959.

[70] M. Tsukada: Convergence of best approximations in smooth Banach spaces, J. Approx. Theory, 40, 301-309, 1984.

[71] T. Zolezzi: On stability analysis in mathematical programming, North-Holland, Mathematical Programming Study, 21, 227-242, 1984.