Preservation or not of the maximally monotone property by graph-convergence

Samir ADLY[*], Hedy ATTOUCH[†] and R. Tyrrell ROCKAFELLAR[‡]

This paper is dedicated to Professor Roger J-B Wets on the occasion of his 85th birthday.

ABSTRACT. In a general real Hilbert space $\mathcal{H}$, given a sequence $(A_n)_{n\in\mathbb{N}}$ of maximally monotone operators $A_n : \mathcal{H} \rightrightarrows \mathcal{H}$, which graphically converges to an operator A whose domain is nonempty, we analyze if the limit operator A is still maximally monotone. This question is justified by the fact that, as we show on an example in infinite dimension, the graph limit in the sense of Painlevé-Kuratowski of a sequence of maximally monotone operators may not be maximally monotone. Indeed, the answer depends on the type of graph convergence which is considered. In the case of the Painlevé-Kuratowski convergence, we give a positive answer under a local compactness assumption on the graphs of the operators A_n. Under this assumption, the sequence $(A_n)_{n\in\mathbb{N}}$ turns out to be convergent for the bounded Hausdorff topology. Inspired by this result, we show that, more generally, when the sequence $(A_n)_{n\in\mathbb{N}}$ of maximally monotone operators converges for the bounded Hausdorff topology to an operator whose domain is nonempty, then the limit is still maximally monotone. The answer to these questions plays a crucial role in the analysis of the sensitivity of monotone variational inclusions, and makes it possible to understand these questions in a unified way thanks to the concept of protodifferentiability. It also leads to revisit several notions which are based on the convergence of sequences of maximally monotone operators, in particular the notion of variational sum of maximally monotone operators.

1 Introduction

Throughout the paper $\mathcal{H}$ is a real Hilbert space, endowed with the scalar product $\langle \cdot, \cdot \rangle$ and the associated norm $\| \cdot \|$. The aim of this paper is to answer the following question, that was raised for the first time in [2, Remark 6] as an open question. Given a sequence $(A_n)_{n\in\mathbb{N}}$ of maximally monotone operators $A_n : \mathcal{H} \rightrightarrows \mathcal{H}$ which graph converges to an operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$ whose domain is nonempty [1], can we conclude that the limit operator A is still maximally monotone? In the case of a negative answer, would it possible to give a sufficient condition preserving the maximal monotonicity property under the graph-convergence limit? After identifying an operator with its graph, an equivalent formulation of the first question would be: is the class of maximally monotone operators a closed subset of the hyperspace of closed subsets of $\mathcal{H} \times \mathcal{H}$, equipped with the Painlevé-Kuratowski (PK) set convergence? It is well-known that in finite dimensional settings, the maximal monotonicity is preserved under graphical convergence in the sense of Painlevé-Kuratowski set convergence. This may fail in infinite dimension as shown by the counterexample presented in Section 5. In relation to this phenomenon, we give a sufficient condition ensuring the maximality of the limit operator, which reflects a local compactness property of the graphs of the operators A_n. The difficulty linked to the use of convergence in the sense of Painlevé-Kuratowski

[*]Laboratoire XLIM, Université de Limoges, 123, avenue Albert Thomas, 87060 Limoges, France. E-mail: samir.adly@unilim.fr

[†]IMAG, Université Montpellier, CNRS, Place Eugène Bataillon, 34095 Montpellier CEDEX 5, France. E-mail: hedy.attouch@umontpellier.fr

[‡]University of Washington, Department of Mathematics, Box 354350, Seattle, WA 98195-4350. E-mail: rtr@uw.edu

[1]this allows to avoid convergence to the empty set

is that it is not attached to a metrizable topology in infinite dimension. This therefore calls into question its usefulness for such an approximation theory, and leads to consider other notions of convergence of sets implying the Painlevé-Kuratowski convergence. One of them is the "bounded Hausdorff" convergence, which is equivalent to the Painlevé-Kuratowski convergence in finite dimension, and is attached to a metrizable topology even in infinite dimension. We will show that the bounded Hausdorff convergence is the right vehicle to deal with the quantitative stability of variational systems involving maximally monotone operators. We will therefore give a positive answer to the initial question simply by replacing the Painlevé-Kuratowski convergence by the bounded Hausdorff convergence.
The above question naturally arises in several situations involving the graph limit of sequences of maximally monotone operators. We will pay attention to two of these particular situations:

(i) In the analysis of the sensitivity of monotone variational inclusions, the concept of protodifferentiability makes it possible to understand these questions in a unified way.

(ii) The notion of variational sum of maximally monotone operators.

2 Preliminary results

For a set-valued map $A : \mathcal{H} \rightrightarrows \mathcal{H}$, the *domain* of A is given by $\mathrm{Dom}(A) := \{x \in \mathcal{H} \mid A(x) \neq \emptyset\}$, and its *graph* is defined by $\mathrm{Gph}\,(A) := \{(x, y) \in \mathcal{H} \times \mathcal{H} \mid y \in A(x)\}$. The product space $\mathcal{H} \times \mathcal{H}$, where are located the graphs of the operators, will be equipped with the usual classical norm $\|(x, y)\|_{\mathcal{H}\times\mathcal{H}} = \sqrt{\|x\|^2 + \|y\|^2}$.
We denote by $A^{-1} : \mathcal{H} \rightrightarrows \mathcal{H}$ the inverse of A, which is the set-valued map defined by: for all $y \in \mathcal{H}$

$$A^{-1}(y) := \{x \in \mathcal{H} \mid y \in A(x)\}.$$

The range of A is defined by

$$\mathrm{Rge}(A) = \bigcup_{x\in\mathcal{H}} A(x).$$

Given a sequence (x_n) in $\mathcal{H}$ that converges to some $x \in \mathcal{H}$, the strong (resp. weak) convergence is denoted by $s\text{-}\lim_{n\to+\infty} x_n = x$ or $x_n \to x$ (resp. $w\text{-}\lim_{n\to+\infty} x_n = x$ or $x_n \rightharpoonup x$) .

Let's now recall some basic facts from the theory of maximally monotone operators, and from the set convergence theory that will be useful for our developments.

2.1 Basic facts concerning maximally monotone operators

The set-valued mapping $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is called monotone if it has the property

$$\langle x^* - y^*, x - y\rangle \geq 0 \quad \text{whenever } (x, x^*) \in \mathrm{Gph}\,(A),\ (y, y^*) \in \mathrm{Gph}\,(A).$$

The set-valued map A is said to be maximally monotone if and only if it is monotone and its graph is maximal in the class of monotone operators for the relation of inclusion, *i.e.* $\mathrm{Gph}\,(A)$ is not properly contained in the graph of any other monotone operator.
For a given set-valued map $A : \mathcal{H} \rightrightarrows \mathcal{H}$, the resolvent of A of index $\lambda > 0$ is given by

$$J_{\lambda A} = (\mathrm{Id} + \lambda A)^{-1}, \tag{2.1}$$

where Id stands for the identity operator on $\mathcal{H}$. It is well known that if $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximally monotone, then, for any $\lambda > 0$, its resolvent $J_{\lambda A} : \mathcal{H} \to \mathcal{H}$ is a single-valued and nonexpansive mapping, *i.e.*

$$\|J_{\lambda A}(x) - J_{\lambda A}(y)\| \leq \|x - y\|,\ \forall x,\ y \in \mathcal{H}.$$

Moreover, $J_{\lambda A} : \mathcal{H} \to \mathcal{H}$ is firmly nonexpansive, *i.e.*

$$\|J_{\lambda A}(x) - J_{\lambda A}(y)\|^2 + \|(\mathrm{Id} - J_{\lambda A})(x) - (\mathrm{Id} - J_{\lambda A})(y)\|^2 \le \|x-y\|^2, \ \ \forall x,\ y \in \mathcal{H}.$$

The resolvents are tied by the resolvent equation: for any $\lambda > 0, \mu > 0$, for any $x \in \mathcal{H}$

$$J_{\lambda A}(x) = J_{\mu A}\left(\frac{\mu}{\lambda}x + \frac{\lambda-\mu}{\lambda}J_{\lambda A}(x)\right).$$

The Yosida approximation of index $\lambda > 0$ associated with the operator A is given by

$$A_\lambda = \frac{1}{\lambda}(\mathrm{Id} - J_{\lambda A}).$$

The Yosida approximation plays an important tool in the theory of monotone operators. It can be considered as a single-valued Lipschitz continuous regularization of a given set-valued maximally monotone operator. It is well-known that if $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximally monotone, then its Yosida approximation A_λ is single-valued, everywhere defined and Lipschitz continuous with modulus $\frac{1}{\lambda}$. Moreover, for any $x \in \mathcal{H}$ and $\lambda > 0$,

$$\big(J_{\lambda A}(x), A_\lambda(x)\big) \in \mathrm{Gph}\,(A).$$

The above formula reflects the fact that $\mathrm{Gph}\,(A)$ is a Lipschitzian manifold in the product space $\mathcal{H} \times \mathcal{H}$, see [24] for more details.

Given a maximally monotone operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$, for any $x \in \mathrm{Dom}(A)$, $A^0(x)$ is the unique element of minimal norm of the closed convex set $A(x)$, *i.e.*

$$A^0(x) = \mathrm{Proj}_{A(x)}(0),\ x \in \mathrm{Dom}(A),$$

where $\mathrm{Proj}_{A(x)}$ denotes the projection operator onto the closed convex set $A(x)$, $x \in \mathrm{Dom}(A)$. The operator A^0 is called the minimal section of the operator A.

For any $x \in \mathrm{Dom}(A)$, we have

$$\lim_{\lambda \to 0} A_\lambda(x) = A^0(x) \in A(x).$$

We recall the following fundamental theorem, known in the literature as Minty's Theorem.

Theorem 2.1 (Minty, 1962) *Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a **monotone** operator. The following statements are equivalent:*

(i) *A is a maximally monotone operator;*

(ii) *The operator $(\mathrm{Id} + A)$ is surjective, i.e.* $\mathrm{Rge}(\mathrm{Id} + A) = \mathcal{H}$.

We note that A is maximally monotone if and only if λA is also maximally monotone, for every $\lambda > 0$. Thus, property (ii) is equivalent to $\mathrm{Rge}(\mathrm{Id} + \lambda A) = \mathcal{H}$, for every $\lambda > 0$.

2.2 Basic facts concerning set convergence

Let $\mathcal{H}$ be a real Hilbert space, and let $(C_n)_{n\in\mathbb{N}}$ be a sequence of subsets of $\mathcal{H}$. The *outer* and the *inner limits* of $(C_n)_{n\in\mathbb{N}}$ when $n \to +\infty$ are defined respectively by

$$\begin{aligned} \mathrm{Limsup}\ C_n &:= \{x \in \mathcal{H} \mid \exists N \in \mathcal{N}^\sharp,\ \forall n \in N,\ \exists x_n \in C_n\ :\ s\text{-}\lim_{n\in N} x_n = x\}, \\ \mathrm{Liminf}\ C_n &:= \{x \in \mathcal{H} \mid \exists N \in \mathcal{N},\ \forall n \in N,\ \exists x_n \in C_n\ :\ s\text{-}\lim_{n\in N} x_n = x\}, \end{aligned}$$

where $\mathcal{N}^\sharp := \{N \subset \mathbb{N} \mid \ N \text{ is infinite}\}$ and $\mathcal{N} := \{N \subset \mathbb{N} \ \mid \ \mathbb{N} \setminus N \text{ is finite}\}$. We can define in the same manner w-Limsup and w-Liminf by replacing the strong convergences in the definitions above by weak ones. The following inclusion holds true in general:

$$\text{Liminf } C_n \subset \text{Limsup } C_n.$$

The *Painlevé-Kuratowski convergence* is defined by this inclusion being an equality.
The following inclusions hold true

$$\text{Liminf } C_n \subset \text{Limsup } C_n \subset \text{w-Limsup } C_n \ \text{ and } \ \text{Liminf } C_n \subset \text{w-Liminf } C_n \subset \text{w-Limsup } C_n.$$

The *Mosco convergence* [2] of the sequence (C_n) to C, denoted by $C_n \xrightarrow{\text{M}} C$, holds if

$$\text{w-Limsup } C_n \subset C \subset \text{Liminf } C_n.$$

When considering graphs of operators from $\mathcal{H}$ to $\mathcal{H}$, the above definitions should be applied to subsets of the product space $\mathcal{H} \times \mathcal{H}$.

Definition 2.1 (Graph-convergence) *Let $A_n : \mathcal{H} \rightrightarrows \mathcal{H}$, $n = 1, 2...$ be a sequence of operators. The sequence $(A_n)_{n\in\mathbb{N}}$ is said to be graph-convergent to the operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$, if* $\text{Gph }(A_n)$ *converges to* $\text{Gph }(A)$ *in the sense of Painlevé-Kuratowski (PK). We then denote $A_n \xrightarrow{\text{G}} A$.*

For any subset $C \subset \mathcal{H}$, the distance from a point $x \in \mathcal{H}$ to C is defined by

$$d(x, C) = \inf_{y\in C} \|x - y\|.$$

By convention, we set $d(x, \emptyset) = \infty$.
Let C and D be two subsets of $\mathcal{H}$. The excess function of C on D is defined by

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

with the convention $e(\emptyset, D) = 0$.
For any nonegative real number $\rho \geq 0$, the closed ball centered at the origin and with radius ρ is denoted by $\rho\mathbb{B}$. For any subset $C \subset \mathcal{H}$, we denote by C_ρ the intersection of C with $\rho\mathbb{B}$ i.e. $C_\rho := C \cap \rho\mathbb{B}$.
Following Attouch-Wets [8, 9, 10], Attouch-Lucchetti-Wets [6], Azé-Penot [12], Beer [14] we have the following quantitative notion.

Definition 2.2 (ρ-Hausdorff distance) *For any $\rho > 0$, the ρ-Hausdorff distance between two subsets C and D of $\mathcal{H}$ is defined by*

$$\text{haus}_\rho(C, D) = \max\Big(e(C_\rho, D), e(D_\rho, C)\Big).$$

A sequence $(C_n)_{n\in\mathbb{N}}$ of subsets of $\mathcal{H}$ is said to converge to a set $C \subset \mathcal{H}$ with respect to the ρ-Hausdorff distance, if for any $\rho > 0$, $\lim_{n\to+\infty} \text{haus}_\rho(C_n, C) = 0$, i.e. for any $\varepsilon > 0$ and any $\rho > 0$, there exists $N > 0$

$$C_n \cap \rho\mathbb{B} \subset C + \varepsilon\mathbb{B} \ \textit{ and } \ C \cap \rho\mathbb{B} \subset C_n + \varepsilon\mathbb{B}, \textit{ for all } n \geq N.$$

[2] In general, the Mosco convergence is used for sequences of closed convex sets, since it necessarily implies that the limit set is weakly closed

We recall that the ρ-Hausdorff distance is not a metric.
The convergence for the ρ-Hausdorff distance is associated with a topology that is metrizable. Following [6], [14], it is induced by the metric which is defined with the help of the semi-distances

$$d_\rho(C,D) := \sup_{\|x\|\le\rho} |d(x,C) - d(x,D)|$$

by the following formula

$$d(C,D) := \sum_{n=1}^{+\infty} 2^{-n} \frac{d_{\rho_n}(C,D)}{1+d_{\rho_n}(C,D)} \tag{2.2}$$

where (ρ_n) is a sequence of increasing positive numbers that tends to $+\infty$ (this sequence can be taken arbitrarily, the topology and the uniform structure remain the same). The link between the ρ-Hausdorff distance and the above metric follows from the following result

Proposition 2.1 *For any* $\rho > 0$ *and* $\rho_0 > \max\{d(0,C), d(0,D)\}$

$$\mathrm{haus}_\rho(C,D) \le d_\rho(C,D) \le \mathrm{haus}_{2\rho+\rho_0}(C,D).$$

Proof. The first inequality is immediate. For the second inequality, we refer to [15, Lemma 3.1]. ■

As an important result, we have the following completeness property.

Theorem 2.2 (Theorem 2.1 [6], Theorem 3.1.3 [14]) *The hyperspace of closed sets is complete for the metric defined in* (2.2).

We therefore use the two terminologies interchangeably: convergence for the ρ-Hausdorff distance, or convergence for the ρ-Hausdorff topology. This contrasts sharply with the Painlevé-Kuratowski convergence which is associated with a topology only when the underlying space is locally compact (see e.g. [3, section 2.8], [14], [22]).

Remark 2.1 (i) If the Hilbert space $\mathcal{H}$ is of finite dimension, then the Painlevé-Kuratowski and the ρ-Hausdorff set-convergence coincide.

(ii) Note also that if the Hilbert space $\mathcal{H}$ is of infinite dimension, and when considering sequences of closed convex sets, then the ρ-Hausdorff convergence implies the Mosco convergence, *i.e.*

$$\Big(\text{for all } \ \rho > 0, \ \lim_{n\to+\infty} \mathrm{haus}_\rho(C_n, C) = 0\Big) \ \text{ implies } C_n \xrightarrow{\mathrm{M}} C \text{ as } n \to +\infty.$$

The notion of ρ-Hausdorff distance can be specialized to functions via their identification with their epigraphs, and to operators via their identification with their graphs.
For an extended real-valued function $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ its epigraph, denoted by epi (f), is defined by

$$\mathrm{epi}\,(f) = \{(x,\alpha) \in \mathcal{H} \times \mathbb{R} \ : \ f(x) \le \alpha\}.$$

Let $f, \ g \ : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ be two extended real valued functions. For any $\rho \ge 0$, the ρ-Hausdorff epi-distance between f and g is defined by

$$\mathrm{haus}_\rho(f,g) = \mathrm{haus}_\rho(\mathrm{epi}\,(f), \mathrm{epi}\,(g)).$$

We denote by $\Gamma_0(\mathcal{H})$ the set of all convex, proper and closed extended real valued functions.
Let $A, \ B : \mathcal{H} \rightrightarrows \mathcal{H}$ be set-valued maps and $\rho \ge 0$. The ρ-graph distance between A and B is defined by

$$\mathrm{haus}_\rho(A,B) := \mathrm{haus}_\rho\Big(\mathrm{Gph}\,(A), \mathrm{Gph}\,(B)\Big).$$

Definition 2.3 (Epigraphical-convergence) *Let $(f_n)_{n\in\mathbb{N}}$ and f be a sequence of extended real valued functions on $\mathcal{H}$.*

(i) *We say that $(f_n)_{n\in\mathbb{N}}$ epiconverges to f if and only if* $\mathrm{epi}\,(f_n)$ *converges to* $\mathrm{epi}\,(f)$ *in the sense of Painlevé-Kuratowski in the product space $\mathcal{H}\times\mathbb{R}$. We write in this case, $f = \mathrm{epi}\text{ - }\lim f_n$ or $f_n \xrightarrow{\text{epi}} f$ as $n \to +\infty$.*

(ii) *We say that $(f_n)_{n\in\mathbb{N}}$ Mosco-epiconverges to f if and only if* $\mathrm{epi}\,(f_n)$ *converges to* $\mathrm{epi}\,(f)$ *in the sense of Mosco in the product space $\mathcal{H}\times\mathbb{R}$. We write in this case, $f = M\text{-epi - }\lim f_n$ or $f_n \xrightarrow{\text{M-epi}} f$ as $n \to +\infty$.*

(iii) *We say that $(f_n)_{n\in\mathbb{N}}$ epiconverges to f in the sense of the bounded-Hausdorff topology if and only if $\lim\limits_{n\to+\infty} \mathrm{haus}_\rho(f_n, f) = 0$ for all $\rho > 0$. We write in this case, $f = \mathrm{epi}\text{ -dist- }\lim f_n$ or $f_n \xrightarrow{\text{epi-dist}} f$ as $n \to +\infty$.*

Remark 2.2 Note that a sequence of closed proper and convex functions $(f_n)_n$ Mosco-epiconverges to some f if and only if for all $x \in \mathcal{H}$,

- for any sequence $x_n \rightharpoonup x$, we have $\liminf\limits_{n\to+\infty} f_n(x_n) \geq f(x)$;
- there exists $x_n \to x$ such that $\limsup\limits_{n\to+\infty} f_n(x_n) \leq f(x)$.

2.3 Basic facts concerning the convergence of sequences of maximally monotone operators

The following proposition gives a characterization of the graph-convergence in the sense of Painlevé-Kuratowski (PK) of a sequence of maximally monotone operators. Note that in the following statement we assume that the limit operator is maximally monotone (we refer to [3]).

Proposition 2.2 *Let $(A_n)_n$ be a sequence of maximally monotone operators $A_n : \mathcal{H} \rightrightarrows \mathcal{H}$, and let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximally monotone operator. Then the following statements are equivalent:*

(i) $A_n \xrightarrow{\text{G}} A$*, i.e. (A_n) graph-converges to A in the sense of Painlevé-Kuratowski;*

(ii) $\mathrm{Gph}\,(A) \subset \mathrm{Liminf}\ \mathrm{Gph}\,(A_n)$.

Equivalently, for every $(x, y) \in \mathrm{Gph}\,(A)$, there exists a sequence $(x_n, y_n) \in \mathrm{Gph}\,(A_n)$ such that: $x_n \to x$ and $y_n \to y$ as $n \to +\infty$.

The above proposition results from the fact that, under the assumption A maximally monotone operator, the inclusion $\mathrm{Limsup}\ \mathrm{Gph}\,(A_n) \subset \mathrm{Gph}\,(A)$ is automatically satisfied. Moreover, we have the following property, see [3, Proposition 3.59] which is an extension to the parametrized case of the strong-$\times$-weak closure property of the graph of a maximally monotone operator

Proposition 2.3 *Let $(A_n)_n$ be a sequence of maximally monotone operators $A_n : \mathcal{H} \rightrightarrows \mathcal{H}$ that graph-converges to a maximally monotone operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$. Then the following property is satisfied: Whenever a sequence $(x_n, y_n) \in \mathrm{Gph}\,(A_n)$ satisfies $x_n \to x$ strongly in $\mathcal{H}$ and $y_n \to y$ weakly in $\mathcal{H}$ as $n \to +\infty$, then $y \in A(x)$.*

Remarkably, when working with sequences of maximally monotone operators, the various notions of graph convergence can be equivalently formulated with the help of the resolvents. Some of them are summarized in the following proposition (see e.g. [3]).

Proposition 2.4 *Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of maximally monotone operators, and let A be a maximally monotone operator. Then the following statements are equivalent:*

(i) $A_n \xrightarrow{\mathrm{G}} A$ *as* $n \to +\infty$*;*

(ii) $J_{\lambda A_n}(x) \to J_{\lambda A}(x)$ *strongly in* $\mathcal{H}$*, for every* $x \in \mathcal{H}$*, for every* $\lambda > 0$*;*

(iii) $J_{\lambda_0 A_n}(x) \to J_{\lambda_0 A}(x)$ *strongly in* $\mathcal{H}$*, for every* $x \in \mathcal{H}$*, for some* $\lambda_0 > 0$.

This property naturally leads to the introduction of the topology of the convergence of the resolvents on the class of maximally monotone operators acting on $\mathcal{H}$, *i.e.* the weakest topology making continuous all the applications $A \mapsto J_{\lambda A}(x) \in \mathcal{H},\ \lambda > 0,\ x \in \mathcal{H}.$ As an important result, when $\mathcal{H}$ is separable, this topology is metrizable. An example of such distance is given by

$$d(A,B) = \sum_k \frac{1}{2^k} \inf\Big\{1,\ \|J_{\lambda_0 A}(x_k) - J_{\lambda_0 B}(x_k)\|\Big\}, \tag{2.3}$$

where (x_k) is in a countable dense subset of $\mathcal{H}$. The following result due to Attouch (see Theorem 3.62 [3]) is in this sense.

Theorem 2.3 (Theorem 3.62 [3]) *Suppose that $\mathcal{H}$ is a separable Hilbert space. Then the class of the maximally monotone operators acting on $\mathcal{H}$ equipped with the topology of the convergence of the resolvents is a metrizable, separable, complete space. For any sequence $(A_n)_{n\in\mathbb{N}}$ of maximally monotone operators, and A maximally monotone operator, the following properties are equivalent:*

(i) $A_n \xrightarrow{\mathrm{G}} A$ *as* $n \to +\infty$*;*

(ii) $d(A_n, A) \to 0$ *as* $n \to +\infty$*;*

(iii) $J_{\lambda_0 A_n}(x) \to J_{\lambda_0 A}(x)$ *strongly in* $\mathcal{H}$*, for every* $x \in \mathcal{H}$*, for some* $\lambda_0 > 0$ *as* $n \to +\infty$.

We have parallel results for the bounded Hausdorff convergence, which are summarized in the proposition below.

Proposition 2.5 *Suppose that $\mathcal{H}$ is a general Hilbert space. Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of maximally monotone operators, and let A be a maximally monotone operator. Then the following statements are equivalent:*

(i) $\forall \rho > 0\ \mathrm{haus}_\rho(A_n, A) \to 0$ *as* $n \to +\infty$*;*

(ii) $\forall \lambda > 0, \forall r > 0\ \ \sup_{\|x\|\le r} \|J_{\lambda A_n}(x) - J_{\lambda A}(x)\| \to 0$ *as* $n \to +\infty$*;*

(iii) $\exists \lambda_0 > 0\ \forall r > 0\ \ \sup_{\|x\|\le r} \|J_{\lambda_0 A_n}(x) - J_{\lambda_0 A}(x)\| \to 0$ *as* $n \to +\infty$.

Proposition 2.5 is a direct consequence of the more general Proposition 2.6 below. In fact, we will need the following more precise result which shows that on the set of maximally monotone operators, the uniform structure attached to the bounded Hausdorff distances is equivalent to the uniform structure attached to the family of pseudo-distances

$$d_{\lambda,r}(A,B) := \sup_{\|x\|\le r} \|J_{\lambda A}(x) - J_{\lambda B}(x)\|.$$

Precisely, according to Attouch-Moudafi-Riahi [7], we have

Proposition 2.6 [7, Proposition 1.1 and 1.2.] Let A and B be two maximally monotone operators. Then, for any $\lambda > 0$, ρ and $r > 0$

$$\mathrm{haus}_\rho(A, B) \leq \max\left(1, \frac{1}{\lambda}\right) d_{\lambda,(1+\lambda)\rho}(A, B) \tag{2.4}$$

$$d_{\lambda,r}(A, B) \leq (2 + \lambda)\,\mathrm{haus}_\rho(A, B) \tag{2.5}$$

where in (2.5) $\rho = \max\Big(r + \|J_{\lambda A}(0)\|, \frac{1}{\lambda}(r + \|J_{\lambda A}(0)\|)\Big)$.

Let us notice that, according to the inequalities (2.4) and (2.5), to generate the above uniform structure, we don't need to consider all the pseudo-distances $(d_{\lambda,r})$. It is sufficient to consider the pseudo-distances $(d_{\lambda_0,r})$ for a given $\lambda_0 > 0$.

Proposition 2.7 [7, Proposition 1.4] *Let A be a maximally monotone operator, and let A_λ be its Yosida approximation of index $\lambda > 0$. Then as $\lambda \to 0$*

$$\forall \rho > 0 \;\; \mathrm{haus}_\rho(A_\lambda, A) \to 0.$$

For further results concerning the theory of maximally monotone operators we refer to [13], [16], [30]. For the variational convergences of sequences of maximally monotone operators see [3], [7], [9]. For properties concerning the set convergence theory and its link with the variational analysis see [28] in finite dimensional spaces and [6], [14] in general normed spaces.

3 Main results

In this section, we state and prove our main results. We start by giving a sufficient compactness assumption $(\mathcal{A}_c)$ ensuring the maximally monotone property of the limit operator under the graph-convergence in the sense of Painlevé-Kuratowski (see Theorem 3.1). We show in Theorem 3.2 that the answer to our initial question is positive without any additional assumption when the graph convergence is taken in the sense of the bounded Hausdorff topology. Finally, we show in Theorem 3.3, that under the same compactness assumption $(\mathcal{A}_c)$ both graph-convergences in the sense of Painlevé-Kuratowski and bounded Hausdorff topology coincide.

Theorem 3.1 *Let (A_n) be a sequence of maximally monotone operators $A_n : \mathcal{H} \rightrightarrows \mathcal{H}$ that converges graphically in the Painlevé-Kuratowski sense to an operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$ whose domain is nonempty. Suppose that the following compactness assumption $(\mathcal{A}_c)$ is satisfied: every sequence (x_n) of $\mathcal{H}$ such that*

$$\left(\sup_n \|x_n\| < +\infty \quad \textit{and} \quad \sup_n \|A_n^0(x_n)\| < +\infty\right)$$

is contained in a compact subset of $\mathcal{H}$.

Then, the limit operator A is maximally monotone.

Proof. The first five claims are elementary. Only the last claim 6 requires a detailed proof.

Claim 1: The limit operator A is monotone.

Claim 2: $J_A := (\mathrm{Id} + A)^{-1}$ is single valued.

Claim 3: $\mathrm{Id} + A_n \xrightarrow{\mathrm{G}} \mathrm{Id} + A$ in the graph sense.

Claim 4: $J_{A_n} \xrightarrow{\mathrm{G}} J_A$ in the graph sense.

Claim 5: $J_{A_n}(x) \to J_A(x)$ for all $x \in \mathrm{Rge}(\mathrm{Id} + A)$ (pointwise convergence).

Claim 6: $\mathrm{Rge}(\mathrm{Id} + A) = \mathcal{H}$.
Given $y \in \mathcal{H}$, we have to solve

$$x + Ax \ni y.$$

According to Minty's Theorem (see Theorem 2.1) and the fact that A_n is maximally monotone, for every $n \in \mathbb{N}$, there exists $x_n \in \mathcal{H}$ such that

$$x_n + A_n(x_n) \ni y,$$

that is $x_n = J_{A_n}(y)$. Let's first verify that the sequence (x_n) remains bounded. Indeed, taking $x_0 \in \mathrm{Rge}(\mathrm{Id} + A)$ (recall that A has been assumed to have a nonempty domain), and using that the resolvents are nonexpansive operators, we have

$$\|x_n - J_{A_n}(x_0)\| = \|J_{A_n}(y) - J_{A_n}(x_0)\| \leq \|y - x_0\|.$$

Therefore, by the triangle inequality

$$\|x_n\| \leq \|y - x_0\| + \|J_{A_n}(x_0)\|. \tag{3.1}$$

According to claim 5, the sequence $(J_{A_n}(x_0))_n$ is convergent in $\mathcal{H}$, and hence bounded. Using (3.1), we obtain that the sequence (x_n) is bounded.

From $x_n + A_n(x_n) \ni y$, we deduce that $\|A_n^0(x_n)\| \leq \|y\| + \|x_n\|$, and hence $\sup_n \|A_n^0(x_n)\| < +\infty$. By the compactness assumption $(\mathcal{A}_c)$, we deduce that (x_n) remains in a compact subset of $\mathcal{H}$. Therefore we can extract a subsequence (x_{n_k}) such that

$$x_{n_k} \to \bar{x} \text{ strongly in } \mathcal{H}.$$

We have

$$y - x_{n_k} \in A_{n_k}(x_{n_k}),$$

with $y - x_{n_k} \to y - \bar{x}$ and $x_{n_k} \to \bar{x}$ strongly in $\mathcal{H}$. According to the graph convergence of the sequence (A_n) to A, we deduce that

$$y - \bar{x} \in A(\bar{x}),$$

which expresses that $y \in \mathrm{Rge}(\mathrm{Id} + A)$. We have obtained that A is monotone and satisfies $\mathrm{Rge}(\mathrm{Id} + A) = \mathcal{H}$. Therefore, according to Minty's theorem, A is maximally monotone. ∎

Remark 3.1 During the reviewing process, we received the following simple proof of Theorem 3.1, from one of the two anonymous referees. Since the domain of A is nonempty, we select some $y \in A(x)$. By the graphical convergence of (A_n) to A, there exists a sequence $(x_n, y_n) \to (x, y)$ with $y_n \in A_n(x_n)$. Hence, $x_n = J_{A_n}(x_n + y_n)$. Let $z \in \mathcal{H}$ be arbitrary and set $v_n = J_{A_n}(z)$. By the nonexpansiveness of J_{A_n}, we have

$$\|x_n - v_n\| \leq \|x_n + y_n - z\|.$$

From this last inequality, we deduce that the sequence $(\|v_n\|)_n$ is bounded. Since $z - v_n \in A_n(v_n)$, we get the boundedness of $A_n^0(v_n)$ as well. By the compactness assumption $(\mathcal{A}_c)$, we deduce that (v_n) remains in a compact subset of $\mathcal{H}$. Therefore we can extract a subsequence, still denoted (v_n) such that $v_n \to v$. Passing to the limit in $z - v_n \in A_n(v_n)$ gives $z - v \in A(v)$. The conclusion follows from Minty's Theorem.

Remark 3.2 The compactness assumption $(\mathcal{A}_c)$ is verified in the following situations:

(a) $\mathcal{H}$ is a finite dimensional Hilbert space. This is a clear consequence of the fact that, in this case, bounded sets are relatively compact.

(b) $A_n = \partial\Phi_n$ where, for each $n \in \mathbb{N}$, $\Phi_n \in \Gamma_0(\mathcal{H})$, the set of extended real-valued functions which are convex lsc. and proper, and the sequence (Φ_n) satisfies the inf-compactness property

$$\left(\sup_n \|x_n\| < +\infty \text{ and } \sup_n \Phi_n(x_n) < +\infty\right) \Longrightarrow (x_n) \text{ is contained in a compact subset of } \mathcal{H}, \tag{3.2}$$

together with: there exists a sequence (ω_n) such that $\sup_n \|\omega_n\| < +\infty$ and $\sup_n \Phi_n(\omega_n) < +\infty$. This results immediately from the subdifferential inequality

$$\Phi_n(\omega_n) \geq \Phi_n(x_n) + \left\langle (\partial\Phi_n)^0 (x_n), \omega_n - x_n \right\rangle,$$

which gives

$$\Phi_n(x_n) \leq C + \| (\partial\Phi_n)^0 (x_n)\| \; \|\omega_n - x_n\|.$$

From the boundedness of the sequences (x_n), (ω_n) and $((\partial\Phi_n)^0 (x_n))$, we infer $\sup_n \Phi_n(x_n) < +\infty$, which gives the relative compactness of the sequence (x_n).

(c) Let us give a concrete example where $A_n = \partial\Phi_n$ and (3.2) is satisfied by the sequence (Φ_n). Let Ω be a bounded open set of $\mathbb{R}^N$, and let $\mathcal{H}$ be equal to the Lebesgue space $L^2(\Omega)$ endowed with its classical Hilbert structure. Various situations in calculus of variations involve dealing with a sequence of convex integral functionals (Φ_n) of the following form, see [3] for examples (the elements of $\mathcal{H}$ are functions which are denoted generically by $\omega \mapsto u(\omega)$)

$$\Phi_n(u) = \begin{cases} \displaystyle\int_\Omega a_n(\omega)\|\nabla u(\omega)\|^2 d\omega \;\text{ if }\; u \in H_0^1(\Omega) \\ +\infty \;\text{ if }\; u \in L^2(\Omega)\setminus H_0^1(\Omega). \end{cases}$$

Suppose that the coefficients (a_n) (diffusion or elasticity coefficients for example) belong to $L^\infty(\Omega)$ and verify $a_n(\omega) \geq c > 0$ with c independent of n and ω. According to the Poincaré inequality, we have that the functions Φ_n satisfy the coercivity property:

$$\Phi_n(u) \geq r\|u\|^2_{H_0^1(\Omega)} \tag{3.3}$$

for some $r > 0$, independent of n. It follows that, for each $n \in \mathbb{N}$, the function Φ_n is lower semicontinuous on $\mathcal{H} = L^2(\Omega)$. Indeed whenever $u_k \to u$ in $L^2(\Omega)$ and $(\Phi_n(u_k))_k$ is bounded, it follows from (3.3) that the sequence $(u_k)_k$ is bounded in $H_0^1(\Omega)$, and therefore converges weakly to u in $H_0^1(\Omega)$. Since Φ_n is convex continuous on $H_0^1(\Omega)$, it is lower semicontinuous for the weak topology of $H_0^1(\Omega)$. Then, returning to prove (3.2), using a similar argument, we have that whenever a sequence a sequence $(u_n)_n$ satisfies $\sup_n \Phi_n(u_n) < +\infty$, then it is bounded in $H_0^1(\Omega)$. The conclusion follows from the Rellich-Kondrakov compact embedding of the Sobolev space $H_0^1(\Omega)$ into $L^2(\Omega)$, which tells us that the sequence $(u_n)_n$ is contained in a compact subset of $L^2(\Omega)$. For related general results see [6, Theorem 3.2].

Let's give a positive answer to the initial question, without additional conditions, when the graph convergence is taken for the bounded Hausdorff topology.

Theorem 3.2 *Suppose that $\mathcal{H}$ is a general real Hilbert space. Let (A_n) be a sequence of maximally monotone operators $A_n : \mathcal{H} \rightrightarrows \mathcal{H}$ that graph-converges in the sense of the bounded Hausdorff topology to an operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$, whose domain is nonempty. Then, the limit operator A is maximally monotone.*

Proof. By definition of the convergence for the bounded Hausdorff topology, we have

$$\lim_n \text{haus}_\rho(A_n, A) = 0$$

for all $\rho > 0$. By the "generalized triangle inequality", see [9, Proposition 1.2], we have for all $n, m \in \mathbb{N}$

$$\text{haus}_\rho(A_n, A_m) \leq \text{haus}_{3\rho}(A_n, A) + \text{haus}_{3\rho}(A, A_m).$$

Hence for all $\rho > 0$

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

The above property expresses that (A_n) is a Cauchy sequence with respect to the metric of the bounded Hausdorff distances. We now rely on Proposition 2.6 which shows the uniform equivalence (*i.e.* same Cauchy sequences) between the metric of the bounded Hausdorff distances and the metric of the uniform convergence of the resolvents on the bounded sets. Precisely to apply Proposition 2.6 we need a uniform bound $\sup_n \|J_{\lambda A_n}(0)\| < +\infty$. Indeed this last property follows from the graph convergence of (A_n) to A, and from the fact that the domain of A is nonempty, as explained in Remark 3.1.

Let us fix some $\lambda > 0$. As a consequence, for each $x \in \mathcal{H}$, $(J_{\lambda A_n}(x))_n$ is a Cauchy sequence in $\mathcal{H}$, which is a complete metric space. Therefore, it converges strongly. This gives the existence of an operator $T : \mathcal{H} \to \mathcal{H}$ such that, for each $x \in \mathcal{H}$, as $n \to +\infty$

$$J_{\lambda A_n}(x) \to T(x), \text{ and} \tag{3.4}$$

$$(A_n)_\lambda := \frac{1}{\lambda}\left(x - J_{\lambda A_n}(x)\right) \to U(x) = \frac{1}{\lambda}\left(x - T(x)\right). \tag{3.5}$$

Let us define the operator $B : \mathcal{H} \rightrightarrows \mathcal{H}$ by

$$\text{Gph}\,(B) := \{(T(x), U(x)) \mid x \in \mathcal{H}\}, \tag{3.6}$$

i.e. $U(x) \in B(T(x))$ for every $x \in \mathcal{H}$. Note that $\text{Dom}(B) = \text{Rge}(T)$. Let us show that B is a monotone operator. According to the relation

$$\frac{1}{\lambda}\left(x - J_{\lambda A_n}(x)\right) \in A_n(J_{\lambda A_n}(x)),$$

and the monotonicity of A_n, we have for any $x, y \in \mathcal{H}$

$$\left\langle \frac{1}{\lambda}\left(x - J_{\lambda A_n}(x)\right) - \frac{1}{\lambda}\left(y - J_{\lambda A_n}(y)\right), J_{\lambda A_n}(x)) - J_{\lambda A_n}(y) \right\rangle \geq 0.$$

According to (3.4) and (3.5), and by passing to the limit in the above equation we obtain, for any $x, y \in \mathcal{H}$

$$\langle U(x) - U(y), T(x) - T(y)\rangle \geq 0. \tag{3.7}$$

Using the definition of B in (3.6) and (3.7), it follows that B is a monotone operator. Moreover, according to the trivial relation

$$J_{\lambda A_n}(x) + \lambda\frac{1}{\lambda}\left(x - J_{\lambda A_n}(x)\right) = x, \tag{3.8}$$

we obtain, by passing to the limit in (3.8), as $n \to +\infty$

$$T(x) + \lambda U(x) = x. \tag{3.9}$$

Since $U(x) \in B(T(x))$, we deduce that for all $x \in \mathcal{H}$

$$T(x) + \lambda B(T(x)) \ni x.$$

So, B is a monotone operator, which satisfies $\mathrm{Rge}(\mathrm{Id} + \lambda B) = \mathcal{H}$. According to Minty's theorem, this implies that B is maximally monotone. Using (3.9), we have that T is equal to resolvent of B of index λ, *i.e.* $T = J_{\lambda B}$.
Finally, we have obtained that for any $x \in \mathcal{H}$,

$$J_{\lambda A_n}(x) \to T(x) = J_{\lambda B}(x),$$

which expresses that the sequence (A_n) graph converges to the maximally monotone operator B. Since the bounded Hausdorff convergence implies the graph convergence, and by uniqueness of the limit, we get $A = B$, which gives that A is maximally monotone. ■

Remark 3.3 When $\mathcal{H}$ is separable, we can give a different proof. Let us return to the fact that we have obtained a Cauchy sequence for the metric d of the pointwise convergence of the resolvents defined in (2.3). We note that by applying, the dominated convergence theorem, we have

$$\lim_{n,m\to+\infty} d(A_n, A_m) = \lim_{n,m\to+\infty} \sum_k \frac{1}{2^k} \inf(1, \|J_{A_n}(x_k) - J_{A_m}(x_k)\|) = 0.$$

Since the metric d is complete (see Theorem 2.3), we have the existence of a maximally monotone operator $\tilde{A}$ such that

$$A_n \to \tilde{A} \quad \text{in the sense of the pointwise convergence of the resolvents.}$$

According to the equivalence between the pointwise convergence of the resolvents and the graph convergence, see Theorem 2.3), we obtain

$$A_n \xrightarrow{\mathrm{G}} \tilde{A} \quad \text{in the sense of the Kuratowski-Painlevé graph convergence.}$$

Since the bounded Hausdorff convergence implies the Kuratowski-Painlevé convergence we have $A = \tilde{A}$, and hence A is a maximally monotone operator.

We show in the following theorem that, under the compactness assumption $(\mathcal{A}_c)$, the graph-convergence in the sense of Painlevé-Kuratowski and the convergence for the bounded Hausdorff topology coincide.

Theorem 3.3 *Let (A_n) be a sequence of maximally monotone operators $A_n : \mathcal{H} \rightrightarrows \mathcal{H}$ that satisfies the compactness assumption $(\mathcal{A}_c)$. Then we have the following equivalence:* (i) $\Longleftrightarrow$ (ii)

(i) *(A_n) graph-converges to A, with a nonempty domain, in the Painlevé-Kuratowski sense.*

(ii) *(A_n) converges to A, with a nonempty domain, for the bounded Hausdorff topology.*

In this case, the limit A is maximally monotone.

Proof. As a general result, convergence for the bounded Hausdorff topology implies convergence in the Painlevé-Kuratowski sense. So the implication (ii) $\Rightarrow$ (i) is automatically satisfied, and we just need to show (i) $\Rightarrow$ (ii). So suppose (i). We have previously shown in Theorem 3.1 that A is maximally monotone, and that for all $x \in \mathcal{H}$, for any $\lambda > 0$

$$J_{\lambda A_n}(x) \to J_{\lambda A}(x), \tag{3.10}$$

where the convergence holds for the strong topology of $\mathcal{H}$. According to Proposition 2.5 we must prove that the resolvents converge uniformly on the bounded subsets of $\mathcal{H}$. Let us argue by contradiction, and suppose that this is not true. Then according to Proposition 2.5 (ii), there exists $\lambda_0 > 0$ and $R > 0$ such that the following property is not satisfied

$$\sup_{x\in B(0,R)} \|J_{\lambda_0 A_n}(x) - J_{\lambda_0 A}(x)\| \to 0 \text{ as } n \to +\infty.$$

This gives the existence of $\epsilon > 0$, and of a subsequence (n_k) such that, for all $k \in \mathbb{N}$

$$\sup_{x \in B(0,R)} \|J_{\lambda_0 A_{n_k}}(x) - J_{\lambda_0 A}(x)\| \geq \epsilon.$$

In turn, this implies the existence of a sequence (x_k) such that $x_k \in B(0,R)$ for all $k \in \mathbb{N}$, and

$$\|J_{\lambda_0 A_{n_k}}(x_k) - J_{\lambda_0 A}(x_k)\| \geq \frac{\epsilon}{2}. \tag{3.11}$$

According to (3.10), for k given we have that as $n \to +\infty$

$$J_{\lambda_0 A_n}(x_k) \to J_{\lambda_0 A}(x_k). \tag{3.12}$$

Therefore there exists a sequence of positive integers m_k which tends to infinity such that for all $k \in \mathbb{N}$

$$\|J_{\lambda_0 A_{m_k}}(x_k) - J_{\lambda_0 A}(x_k)\| \leq \frac{\epsilon}{4}. \tag{3.13}$$

According to (3.11) and (3.13), and by using the triangle inequality, we deduce that for all $k \in \mathbb{N}$

$$\|J_{\lambda_0 A_{n_k}}(x_k) - J_{\lambda_0 A_{m_k}}(x_k)\| \geq \frac{\epsilon}{4}. \tag{3.14}$$

Since the sequence (x_k) is contained in a bounded set of $\mathcal{H}$, we can extract a sequence (still denoted (x_k)) such that $x_k \rightharpoonup \bar{x}$ in $\mathcal{H}$ as $k \to +\infty$. By the same argument as in Theorem 3.1, we have that the sequences $(J_{\lambda_0 A_{n_k}}(x_k))_k$ and $(J_{\lambda_0 A_{m_k}}(x_k))_k$ are bounded. Moreover,

$$\lambda_0 A_{n_k}(J_{\lambda_0 A_{n_k}}(x_k)) \ni x_k - J_{\lambda_0 A_{n_k}}(x_k)$$

and hence

$$\sup_k \|A^0_{n_k}(J_{\lambda_0 A_{n_k}}(x_k))\| < +\infty.$$

According to the compactness assumption $(\mathcal{A}_c)$, we deduce that the sequence $(J_{\lambda_0 A_{n_k}}(x_k))_k$ is relatively compact in $\mathcal{H}$. Similarly, we obtain that the sequence $(J_{\lambda_0 A_{m_k}}(x_k))_k$ is relatively compact in $\mathcal{H}$. Let us extract subsequences (we still keep the same notation for subsequences) such that

$$J_{\lambda_0 A_{n_k}}(x_k) \to z_1 \quad \text{strongly in } \mathcal{H} \quad \text{and} \quad J_{\lambda_0 A_{m_k}}(x_k) \to z_2 \quad \text{strongly in } \mathcal{H}.$$

On the one hand, by passing to the limit in (3.14), thanks to the above strong convergence property

$$\|z_1 - z_2\| \geq \frac{\epsilon}{4}. \tag{3.15}$$

On the other hand using

$$x_k \rightharpoonup \bar{x} \text{ weakly and } J_{\lambda_0 A_{m_k}}(x_k) \to z_2 \text{ strongly},$$

and the convergence of the sequence of maximally monotone operators $J_{\lambda_0 A_{m_k}}$ to $J_{\lambda_0 A}$, we obtain by applying Proposition 2.3 that

$$z_2 = J_{\lambda_0 A}(\bar{x}).$$

Similarly from

$$x_k \rightharpoonup \bar{x} \text{ weakly and } J_{\lambda_0 A_{n_k}}(x_k) \to z_1 \text{ strongly},$$

and the convergence of the sequence of maximally monotone operators $J_{\lambda_0 A_{n_k}}$ to $J_{\lambda_0 A}$ we get

$$z_1 = J_{\lambda_0 A}(\bar{x}).$$

Therefore $z_2 = z_1 = J_{\lambda_0 A}(\bar{x})$, a clear contradiction with (3.15). The proof of Theorem 3.3 is thereby completed. ■

The following result completes Theorem 3.3 by showing that the limit operator A has resolvents which are compact operators. Note that this is not an immediate result since the compactness assumption $(\mathcal{A}_c)$ implies the operators (A_n) and not their potential limit.

Proposition 3.1 *Let (A_n) be a sequence of maximally monotone operators $A_n : \mathcal{H} \rightrightarrows \mathcal{H}$ that graph-converges and that satisfies the compactness assumption $(\mathcal{A}_c)$. Then (A_n) converges for the bounded Hausdorff topology to a maximally monotone operator A which has compact resolvents. This means that for any $r > 0$ and $\lambda > 0$*

$$J_{\lambda A}(\mathbb{B}(0,r)) \text{ is compact in } \mathcal{H}.$$

Proof. According to Theorem 3.3 the sequence (A_n) converges for the bounded Hausdorff topology to a maximally monotone operator A. Let us give a bounded sequence (x_n) such that $x_n \in \mathbb{B}(0,r)$ for all $n \in \mathbb{N}$ and $\lambda_0 > 0$. By the triangle inequality and the definition of the Yosida approximate, we have

$$\begin{aligned}
\|A_{n,\lambda_0}(x_n)\| &\leq \|A_{\lambda_0}(x_n)\| + \frac{1}{\lambda_0}\|J_{\lambda_0 A_n}(x_n) - J_{\lambda_0 A}(x_n)\| \\
&\leq \|A_{\lambda_0}(x_n)\| + \frac{1}{\lambda_0}\sup_{\|x\|\leq r}\|J_{\lambda_0 A_n}(x) - J_{\lambda_0 A}(x)\| \\
&\leq \|A_{\lambda_0}(x_n)\| + \frac{1}{\lambda_0} d_{\lambda_0,r}(A_n, A). \qquad (3.16)
\end{aligned}$$

According to Proposition 2.5 and the fact that the sequence (A_n) converges for the bounded Hausdorff topology to the maximally monotone operator A, we have

$$d_{\lambda_0,r}(A_n, A) \to 0 \text{ as } n \to +\infty. \qquad (3.17)$$

Moreover, according to the Lipschitz continuity property of A_{λ_0} and (x_n) bounded we have

$$\sup_n \|A_{\lambda_0}(x_n)\| < +\infty. \qquad (3.18)$$

Combining (3.16) with (3.17) with (3.18) we deduce that

$$\sup_n \|A_{n,\lambda_0}(x_n)\| < +\infty.$$

The same argument as in Theorem 3.3 gives that the sequence $(J_{\lambda_0 A_n}(x_n))_n$ is relatively compact in $\mathcal{H}$. Let us extract subsequences such that

$$x_{n_k} \rightharpoonup \xi \text{ weakly in } \mathcal{H}$$

and

$$J_{\lambda_0 A_{n_k}}(x_{n_k}) \longrightarrow \eta \text{ strongly in } \mathcal{H}.$$

From (3.17) we deduce that

$$J_{\lambda_0 A}(x_{n_k}) \to \eta \text{ strongly in } \mathcal{H}.$$

By the demi-closedness property of the resolvent operator $J_{\lambda_0 A}$, we get $\eta = J_{\lambda_0 A}(\xi)$. Therefore

$$J_{\lambda_0 A}(x_{n_k}) \longrightarrow J_{\lambda_0 A}(\xi) \text{ strongly in } \mathcal{H},$$

with $\xi \in \mathbb{B}(0,r)$. This proves the claim. ■

4 Closedness of the family of maximally monotone operators with respect to the pointwise convergence of the resolvents

The following result plays a central role in the proof of Theorem 3.2. It's interesting to figure it out. It will also help us to give a counterexample showing that the answer to our initial question may be negative.

Theorem 4.1 *Suppose that $\mathcal{H}$ is a general Hilbert space. Let (A_n) be a sequence of maximally monotone operators $A_n : \mathcal{H} \rightrightarrows \mathcal{H}$ that converges in the following sense: for some $\lambda > 0$ and for all $x \in \mathcal{H}$*

$$s\text{-}\lim J_{\lambda A_n}(x) \quad \textit{exists}.$$

Then, the sequence (A_n) graph converges to a maximally monotone operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$, and we have, for all $x \in \mathcal{H}$

$$J_{\lambda A}(x) = s\text{-}\lim J_{\lambda A_n}(x).$$

Proof. Since rescaling by a positive parameter $\lambda > 0$ does not affect the maximal monotonicity of an operator, without loss of generality, we assume that $\lambda = 1$. The proof is based on an argument similar to the one used in the second part of the proof of Theorem 3.2. It is based on the introduction of the operator $T : \mathcal{H} \to \mathcal{H}$ such that, for each $x \in \mathcal{H}$, as $n \to +\infty$

$$J_{A_n}(x) \to T(x) \ \text{ in } \mathcal{H} \tag{4.1}$$

$$(x - J_{A_n}(x)) \to U(x) = (x - T(x)) \ \text{ in } \mathcal{H}. \tag{4.2}$$

Then define the operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$ by

$$\text{Gph}\,(A) := \{(T(x), U(x)) \mid x \in \mathcal{H}\}\,, \tag{4.3}$$

i.e. $U(x) \in A(T(x))$ for every $x \in \mathcal{H}$. Note that $\text{Dom}(A) = \text{Rge}(T)$. Then, following the proof of Theorem 3.2, we can show that A is a monotone operator, which satisfies $\text{Rge}(\text{Id} + A) = \mathcal{H}$. According to Minty's theorem, this implies that A is maximally monotone. Morever, according to (4.2) we have $T(x) + U(x) = x$, which gives

$$T(x) + A(T(x)) \ni x.$$

Therefore, T is equal to resolvent of A of index 1, *i.e.* $T = J_A$. We have proved that, for all $x \in \mathcal{H}$

$$J_{A_n}(x) \to J_A(x)$$

which, according to Theorem 2.3, gives the graph convergence of the sequence (A_n) to A. ∎

5 A counterexample

We are going to exhibit a sequence (A_n) of maximally monotone operators $A_n : \mathcal{H} \rightrightarrows \mathcal{H}$ such that A_n graph converges to A, and A is monotone but not maximally monotone.

Theorem 5.1 *In an infinite dimensional Hilbert space, the graph limit in the sense of Painlevé-Kuratowski of a sequence (A_n) of maximally monotone operators $A_n : \mathcal{H} \rightrightarrows \mathcal{H}$ might be no more maximally monotone.*

Proof. Take $\mathcal{H} = L^2(0,1)$ the space of square integrable functions with respect to the Lebesgue measure on $[0,1]$. Classically, this is a Hilbert space, when equipped with the scalar product $\langle u, v\rangle = \int_0^1 u(t)v(t)dt$ and the associated norm.

Consider the sequence of functions (a_n) which oscillates more and more rapidly between two positive values, take 1 and 3. For example

$$a_n(t) = \begin{cases} 1 \ \text{ if } \ t \in [0, \frac{1}{n}] \cup [\frac{2}{n}, \frac{3}{n}] \cup ... \\ 3 \ \text{ if } \ t \in [\frac{1}{n}, \frac{2}{n}] \cup [\frac{3}{n}, \frac{4}{n}].... \end{cases}$$

This is a model situation for a sequence that converges weakly and not strongly in $\mathcal{H} = L^2(0,1)$. Indeed, the sequence (a_n) converges weakly to its mean value, which is the constant function equal to 2. But it does not converge strongly since $\|a_n - 2\|_{\mathcal{H}} = 1$.
For each $n \geq 1$ define the operator A_n by

$$A_n(u) = a_n u$$

that is, $(A_n(u))(t) = a_n(t)u(t)$. Clearly, $A_n : L^2(0,1) \to L^2(0,1)$ is a linear continuous and monotone operator. Indeed

$$\langle A_n(u), u\rangle = \int_0^1 a_n(t)u^2(t)dt \geq \int_0^1 u^2(t)dt,$$

and

$$\|A_n(u)\|_{\mathcal{H}} \leq 3\|u\|_{\mathcal{H}}.$$

So, for each $n \geq 1$, the operator A_n is maximally monotone. We note that the operator A_n is a subdifferential of a convex and continuous quadratic function. Let us compute its resolvent. Given $f \in \mathcal{H}$, $u_n = J_{A_n} f$ is the solution of

$$u_n + a_n u_n = f,$$

which gives

$$J_{A_n} f = \frac{1}{1+a_n} f.$$

Let us show that the sequence $(J_{A_n} f)_n$ converges strongly in $\mathcal{H} = L^2(0,1)$ if and only if $f = 0$. Note that the sequence $(\frac{1}{1+a_n})$ converges weakly $\sigma(L^\infty, L^1)$ to its mean value which is the constant function equal to $\frac{3}{8}$. Therefore, the sequence $(J_{A_n} f)_n$ converges weakly in $\mathcal{H}$ to $\frac{3}{8} f$. Let us compute

$$\|\frac{1}{1+a_n} f - \frac{3}{8} f\|_{\mathcal{H}} = \frac{1}{8}\|f\|_{\mathcal{H}}.$$

Therefore, the resolvents $(J_{A_n} f)_n$ converge strongly only for $f = 0$. As a consequence, the sequence (A_n) does not converge to a maximally monotone operator.

Let us show that the sequence (A_n) graph converges to the operator $A : \mathcal{H} \to \mathcal{H}$ whose domain is reduced to the singleton $\{0\}$, and such that $A(0) = 0$. Clearly, since $A_n(0) = 0$, we have

$$\text{Gph } A \subset \text{Liminf Gph } A_n.$$

We must show that

$$\text{Limsup Gph } A_n \subset \text{Gph } A.$$

So, let us give a sequence (u_n) such that

$$u_n \to u \ \text{ strongly in } \mathcal{H} \quad \text{ and } \quad a_n u_n \to f \ \text{ strongly in } \mathcal{H}.$$

By the triangle inequality we have

$$\|a_n u - f\| \leq \|a_n u - a_n u_n\| + \|a_n u_n - f\| \leq 3\|u - u_n\| + \|a_n u_n - f\|,$$

which gives that the sequence $(a_n u)$ converges strongly to f. Since the sequence (a_n) converges weakly to its mean value which is the constant function equal to 2, we have that $(a_n u)$ converges weakly to $2u$. Hence $f = 2u$, and $(a_n u)$ converges strongly to $2u$. Then note that

$$\|a_n u - 2u\| = \|u\|.$$

This forces u to be equal to 0, which in turn gives $f = 0$. This gives the claim. ■

Remark 5.1 Based on the above counterexample, we are led to modify our initial question as follows: Given a sequence (A_n) of maximally monotone operators $A_n : \mathcal{H} \rightrightarrows \mathcal{H}$ that graph converges to an operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$ in the following sense

$$\Big(s \times w \text{ - Limsup Gph } A_n\Big) \cup \Big(w \times s \text{ - Limsup Gph } A_n\Big) \subset \text{Gph } A \subset s \times s \text{ - Liminf Gph } A_n$$

is the limit operator A maximally monotone?
This is a subject for further investigation.

Remark 5.2 When finalizing this paper, we received a note by G. Wachsmuth [29] where an other counter-example is given. This counter-example is based on the construction of a sequence of maximally monotone nonlinear operators $A_n : \ell^2 \to \ell^2$ such that its graphical limit (in the sense of Painlevé-Kuratowski) is the operator $A : \ell^2 \to \ell^2$ such that its graph $\text{Gph}\,(A) = \{(0,0)\}$, which is clearly not maximally monotone (see [29, Theorem 2] for more details).

6 The proto-differentiability of a maximally monotone operator

Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a set-valued map. For all $t > 0$, $x \in \mathcal{H}$ and $x^* \in A(x)$, we define the following difference quotient

$$\Delta_t A(x|x^*)(\omega) := \frac{1}{t}\left(A(x + t\omega) - x^*\right). \tag{6.1}$$

The concept of proto-differentiability was introduced by Rockafellar [25] and is associated with the graph convergence properties of the net of operators $(\Delta_t A(x|x^*))_{t>0}$.

We will successively consider the case of the Painlevé-Kuratowski (PK), then of the bounded-Hausdorff (BH) convergences. In infinite dimensional spaces, they give rise to different concepts. In what follows, we will use the notation D_p (respectively $\widetilde{D}_p$) for the proto-differentiability associated with the PK (respectively the BH) convergence.

6.1 Proto-differentiability with respect to the Painlevé-Kuratowski graph convergence

Definition 6.1 (PK-Proto-differentiability) *Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a set-valued map. We say that A is PK-proto-differentiable at $x \in \mathcal{H}$ relative to $x^* \in A(x)$ if $(\Delta_t A(x|x^*))_{t>0}$ PK-graph converges. In that case, we denote the graph limit by*

$$D_p A(x|x^*) := \text{PK- Graph-lim}\, \Delta_t A(x|x^*).$$

The set-valued map $D_p A(x|x^*) : \mathcal{H} \rightrightarrows \mathcal{H}$ is called the PK-proto-derivative of A at x relative to x^*.

It is easy to check that if $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximally monotone, then, for each $t \in [0, T[$, the mapping $\omega \mapsto \Delta_t A(x|x^*)(\omega) := \frac{1}{t}\left(A(x + t\omega) - x^*\right)$ is also maximally monotone (see [2, Lemma 1]). This is a direct consequence of the preservation of the maximality with respect to translation and homothety.
Question: Suppose that $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximally monotone and proto-differentiable at $x \in \mathcal{H}$ relative to

$x^* \in A(x)$. Under which condition the monotone operator $D_pA(x|x^*) : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximally monotone? Let us translate the compactness condition $(\mathcal{A}_c)$ for the sequence of maximally monotone operators: $\Delta_{t_n}A(x|x^*)(\cdot)$ with $t_n \to 0^+$. Let's call this new compactness condition $(\tilde{\mathcal{A}}_c)$: every sequence $(\omega_n) \subset \mathcal{H}$ satisfying the following two conditions (i)-(ii) is contained in a compact subset of $\mathcal{H}$:

$$\begin{cases} \text{(i)} \ \ \sup\limits_n \|\omega_n\| < +\infty \\ \text{(ii)} \ \ \exists M > 0 \text{ such that } \forall t_n \searrow 0^+, \ \|\mathrm{Proj}_{A(x+t_n\omega_n)}(x^*) - x^*\| \leq M t_n. \end{cases}$$

Proposition 6.1 *Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximally monotone operator. Assume that A is proto-differentiable at $x \in \mathcal{H}$ relative to $x^* \in A(x)$. If the condition $(\tilde{\mathcal{A}}_c)$ is satisfied, then $D_pA(x|x^*)$ is a maximally monotone operator.*

Proof. Let us first notice that if C is a nonempty closed convex subset of $\mathcal{H}$, then for every $x, \ x^* \in \mathcal{H}$ and for every $t > 0$, we have

$$\mathrm{Proj}_{\frac{1}{t}(C-x^*)}(0) = \frac{1}{t}(\mathrm{Proj}_C(x^*) - x^*) = \frac{1}{t}\mathrm{Proj}_{C-x^*}(0). \tag{6.2}$$

In fact,

$$\begin{aligned} y = \mathrm{Proj}_{\frac{1}{t}(C-x^*)}(0) &\iff \langle 0 - y, \frac{1}{t}(a - x^*) - y\rangle, \ \forall a \in C \\ &\iff \langle 0 - y, \frac{1}{t}(a - x^*) - y\rangle, \ \forall a \in C \\ &\iff \langle x^* - (x^* + ty), a - (x^* + ty)\rangle, \ \forall a \in C \\ &\iff x^* + ty = \mathrm{Proj}_C(x^*). \end{aligned}$$

We set $C = A(x + t_n\omega_n)$. Using (6.2), we have

$$[\Delta_{t_n}A(x|x^*)]^0(\omega_n) = \mathrm{Proj}_{\frac{1}{t_n}(A(x+t_n\omega_n)-x^*)}(0) = \frac{1}{t_n}(\mathrm{Proj}_{A(x+t_n\omega_n)}(x^*) - x^*).$$

It is easy to see that condition $(\mathcal{A}_c)$ is satisfied for $\Delta_{t_n}A(x|x^*)(\omega_n)$.
Since we automatically have $0 \in D_pA(x|x^*)(0)$, we do not have to assume that the domain of the operator $D_pA(x|x^*)$ is nonempty. The conclusion of Proposition 6.1 follows from Theorem 3.1. ■

Remark 6.1 The counterexample in [29] shows that the protoderivative of a maximally monotone operator may not be maximally monotone. This means that, the assertion of Proposition 6.1 may fail if we drop the compactness assumption.

6.2 Proto-differentiability with respect to the bounded Hausdorff topology

In the Definition 6.1 of proto-differentiability, let us replace the Graph-convergence in the sense of Painlevé-Kuratowski by the Graph-convergence in the sense of the bounded Hausdorff topology. The properties of the new proto-differentiability notion can be studied in the same way. By specializing our results to the case of $A = \partial\Phi$ with $\Phi \in \Gamma_0(\mathcal{H})$ we can define the twice epidifferentiability of Φ with respect to the bounded Hausdorff topology.

Definition 6.2 (BH-Proto-differentiability) *Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a set-valued map. We say that A is BH-proto-differentiable at $x \in \mathcal{H}$ relative to $x^* \in A(x)$ if $(\Delta_t A(x|x^*))_{t>0}$ graph converges in the sense of the bounded Hausdorff topology. In that case, we denote the graph limit by*

$$\widetilde{D}_p A(x|x^*) := \text{BH- Graph-lim}\, \Delta_t A(x|x^*).$$

The set-valued map $\widetilde{D}_p A(x|x^*) : \mathcal{H} \rightrightarrows \mathcal{H}$ is called the BH-proto-derivative of A at x relative to x^*.

The notion of proto-differentiability is defined in terms of graphs of first-order difference quotient set-valued maps $(\Delta_t A(x|x^*))_{t>0}$. It is easy to prove that this concept is preserved by the inverse operation. The following proposition is in this sense.

Proposition 6.2 *A set-valued map $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is BH-proto-differentiable at $x \in \mathcal{H}$ relative to $x^* \in A(x)$ if and only if its inverse A^{-1} is BH-proto-differentiable at x^* relative to $x \in A^{-1}(x^*)$ and we have*

$$\left[\widetilde{D}_p A(x|x^*)\right]^{-1} = \widetilde{D}_p A^{-1}(x^*|x).$$

Proof. We notice first that for every $x^* \in A(x)$, we have

$$\text{Gph}\,(\Delta_t A(x|x^*)) = \frac{1}{t}(\text{Gph}\,(A) - (x, x^*)) \text{ and } \text{Gph}\,(\Delta_t A^{-1}(x^*|x)) = \frac{1}{t}(\text{Gph}\,(A^{-1}) - (x^*, x)).$$

Let $L : \mathcal{H} \times \mathcal{H} \to \mathcal{H} \times \mathcal{H}$ be the automorphism defined by $(x, x^*) \mapsto L(x, x^*) = (x^*, x)$. Let us define the set-valued map $LA : \mathcal{H} \rightrightarrows \mathcal{H}$ by:

$$\text{Gph}\,(LA) := L\,\text{Gph}\,(A) = \text{Gph}\,(A^{-1}).$$

We have,

$$\begin{aligned} L[\text{Gph}\,(\Delta_t A(x|x^*))] &= \text{Gph}\,(\Delta_t (LA)(x^*|x) \\ &= \frac{1}{t}(\text{Gph}\,(LA) - (x^*|x)) \\ &= \frac{1}{t}(L\text{Gph}\,(A) - (x^*|x)) \\ &= \frac{1}{t}(\text{Gph}\,(A^{-1}) - (x^*|x)) \\ &= \text{Gph}\,(\Delta_t A^{-1}(x^*|x)). \end{aligned}$$

On the other hand, we have

$$L[\text{Gph}\,(\Delta_t A(x|x^*))] = \text{Gph}\,([\Delta_t A(x|x^*)]^{-1}).$$

Hence,

$$\text{Gph}\,(\Delta_t A^{-1}(x^*|x)) = \text{Gph}\,([\Delta_t A(x|x^*)]^{-1}).$$

The conclusion follows from the fact that for every sequence of set-valued operators $A_n,\ A : \mathcal{H} \rightrightarrows \mathcal{H}$ and every $\rho > 0$, we have

$$\lim_{n\to+\infty} \text{haus}_\rho \Big(\text{Gph}\,(A_n), \text{Gph}\,(A)\Big) = 0 \Longleftrightarrow \lim_{n\to+\infty} \text{haus}_\rho \Big(\text{Gph}\,(A_n^{-1}), \text{Gph}\,(A^{-1})\Big) = 0,$$

which completes the proof of Proposition 6.2. ■

Remark 6.2 (i) Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a set-valued map and $x \in \mathcal{H}$ with $x^* \in x + A(x)$. It is easy to prove that $(\mathrm{Id} + A)$ is BH-proto-differentiable at x relative to x^* if and only if A is BH-proto-differentiable at x relative to $x^* - x$. In that case we have

$$\widetilde{D}_p(\mathrm{Id} + A)(x|x^*) = \mathrm{Id} + \widetilde{D}_p A(x|x^* - x). \tag{6.3}$$

(ii) Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximally monotone operator. By Proposition 6.2 and (i), we have A is BH-proto-differentiable at x relative to x^* if and only if $J_A := (\mathrm{Id} + A)^{-1}$ is BH-proto-differentiable at $x + x^*$ relative to x. In this case, it holds

$$\widetilde{D}_p J_A(x + x^*|x) = J_{\widetilde{D}_p A(x|x^* - x)}. \tag{6.4}$$

In fact, by Proposition 6.2 and (6.3) we have

$$\begin{aligned}
\widetilde{D}_p J_A(x + x^*|x) &= \widetilde{D}_p(\mathrm{Id} + A)^{-1}(x + x^*|x) \\
&= [\widetilde{D}_p(\mathrm{Id} + A)(x|x + x^*)]^{-1} \\
&= [\mathrm{Id} + \widetilde{D}_p(A)(x|x^* - x)]^{-1} \\
&= J_{\widetilde{D}_p A(x|x^* - x)}.
\end{aligned}$$

(iii) Proposition 6.2 and the items (i)-(ii) above are known if we consider the Painlevé-Kuratowski convergence (see for example [17, 18, 27]).

The following proposition shows that without additional assumptions, the BH-proto-derivative of a maximally monotone operator is still maximally monotone.

Proposition 6.3 *Suppose that $\mathcal{H}$ is a general Hilbert space. Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximally monotone operator. Assume that A is BH-proto-differentiable at $x \in \mathcal{H}$ relative to $x^* \in A(x)$. Then its BH-proto-derivative $\widetilde{D}_p A(x|x^*)$ is a maximally monotone operator.*

Proof. The sequence $(\Delta_t A(x|x^*))_{t>0}$ graph converges in the sense of the bounded Hausdorff topology to $\widetilde{D}_p A(x|x^*)$ means that for every $t_n \searrow 0^+$ and for every $\rho > 0$, we have

$$\lim_{n\to+\infty} \mathrm{haus}_\rho\Big(\Delta_{t_n} A(x|x^*), \widetilde{D}_p A(x|x^*)\Big) = 0.$$

Since for every $t_n > 0$, $\Delta_{t_n} A(x|x^*)$ is a maximally monotone operator, by Theorem 3.2, the limit $\widetilde{D}_p A(x|x^*)$ is also maximally monotone. ■

Remark 6.3 Under the compactness condition $(\tilde{\mathcal{A}}_c)$, by Theorem 3.3 the following two conditions are equivalent

(i) $(\Delta_t A(x|x^*))_{t>0}$ graph-converges to $D_p A(x|x^*)$ in the Painlevé-Kuratowski sense.

(ii) $(\Delta_t A(x|x^*))_{t>0}$ graph-converges to $\widetilde{D}_p A(x|x^*)$ in the bounded Hausdorff topology.

In this case, the unique limit $D_p A(x|x^*) = \widetilde{D}_p A(x|x^*)$ is maximally monotone.

Remark 6.4 In order to take into account the perturbation of all the data in a given variational problem it is possible to extend the notion of proto-differentiability to a parametrized sequence of set-valued maps as e.g. in [1, 2]. Let $A : [0,T[\times\mathcal{H} \rightrightarrows \mathcal{H}$ be a parameterized set-valued map. For all $t > 0$, $x \in \mathcal{H}$ and $x^* \in A(0,x)$, we define the following difference quotient

$$\Delta_t A(x|x^*)(\omega) := \frac{1}{t}\left(A(t, x + t\omega) - x^*\right). \tag{6.5}$$

The concept of proto-differentiability is associated to the graph convergence properties of the net of operators $(\Delta_t A(x|x^*))_{t>0}$.

We say that A is *proto-differentiable* at $x \in \mathcal{H}$ relative to $x^* \in A(0,x)$ if $(\Delta_t A(x|x^*))_{t>0}$ graph converges.
For the PK and the BH convergences, the *proto-derivatives* of A at x relative to x^* will be denoted respectively by $D_pA(x|x^*)$ and $\widetilde{D}_pA(x|x^*)$.

It is easy to check that if $A : [0,T[\times\mathcal{H} \rightrightarrows \mathcal{H}$ is such that $A(t,\cdot)$ is maximally monotone for all $t \in [0,T[$, then, for each $t \in [0,T[$, the mapping $\omega \mapsto \Delta_t A(x|x^*)(\omega) := \frac{1}{t}\left(A(t, x + t\omega) - x^*\right)$ is also maximally monotone. We adapt the compactness condition $(\tilde{\mathcal{A}}_c)$ to a parameterized set-valued map: every sequence $(\omega_n) \subset \mathcal{H}$ satisfying the following two conditions (i)-(ii) is contained in a compact subset of $\mathcal{H}$:

$$\begin{cases} \text{(i)} \;\; \sup\limits_n \|\omega_n\| < +\infty \\ \text{(ii)} \;\; \exists M > 0 \text{ such that } \forall t_n \searrow 0^+, \; \|\mathrm{Proj}_{A(t_n, x+t_n\omega_n)}(x^*) - x^*\| \le Mt_n. \end{cases}$$

We note that Proposition 6.1 and Proposition 6.3 are still valid for the case of a parametrized maximally monotone operator $A : [0,T[\times\mathcal{H} \rightrightarrows \mathcal{H}$.

Remark 6.5 To define a second-order epi-derivative [17, 26, 27] associated with a closed convex and proper function in a Hilbert space $\Phi \in \Gamma_0(\mathcal{H})$, the Mosco epi-convergence plays a central role. The link between the PK-proto-differentiability of the subdifferential $\partial\Phi$ of the function Φ and its twice epi-differentiability is tied to Attouch's theorem (see [3] Theorem 3.66). It is possible to explore the notion of twice epi-differentiability of a function $\Phi \in \Gamma_0(\mathcal{H})$ in the sense of the bounded Hausdorff convergence and its link with the BH-proto-differentiability of its subdifferential $\partial\Phi$. In this case, instead of Attouch's Theorem, one can use [11, Theorem 2.3].

7 Application to the sensitivity analysis of monotone inclusions

Let us consider the following perturbed variational inclusion with respect to the parameter $t \in [0,T[$, with $T > 0$

$$\text{find } x(t) \in \mathcal{H} \text{ such that } \xi(t) \in x(t) + A(t, x(t)), \tag{7.1}$$

where

(i) $A : [0,T[\times\mathcal{H} \rightrightarrows \mathcal{H}$ is a maximally monotone set-valued operator, i.e. for all $t \in [0,T[$, $A(t,\cdot)$ is a maximally monotone operator.

(ii) $\xi : [0,T[\to \mathcal{H},\; t \mapsto \xi(t) \in \mathcal{H}$ is a given right hand-term.

It is easy to see that (7.1) is equivalent to

$$x(t) = (\mathrm{Id} + A(t,\cdot))^{-1}(\xi(t)) =: J_{A(t,\cdot)}(\xi(t)). \tag{7.2}$$

Our aim is to derive sufficient conditions on the data A, and ξ ensuring the right-differentiability at $t = 0$ of the solution $x : [0,T[\to \mathcal{H}$ of (7.1), and to provide an explicit formula for its right-derivative $x'(0)$. We show that the right-derivative of x at 0 is solution of a variational inclusion involving the proto-derivative of the operators A. Precisely, we prove that $x'(0)$ is the solution of the following variational inclusion

$$\xi'(0) \in x'(0) + D_pA\big(x(0)|x^*(0)\big)\big(x'(0)\big), \tag{7.3}$$

with $x^*(0) := \xi(0) - x(0) \in A(0, x(0))$, and $D_pA\big(x(0)|x^*(0)\big)$ is the proto-derivative of A at $x(0)$ relative to $x^*(0)$.

Theorem 7.1 *Let $A : [0,T[\times\mathcal{H} \rightrightarrows \mathcal{H}$ be a maximally monotone operator, and let $\xi : [0,T[\to \mathcal{H}$ be a function. We consider a solution $x : [0,T[\to \mathcal{H}$ to problem* (7.1)*. If the following assertions are satisfied:*

(i) *ξ is right-differentiable at $t = 0$;*

(ii) *A is PK-proto-differentiable at $x(0)$ relative to $x^*(0) := \xi(0) - x(0) \in A(0, x(0))$ with the proto-derivative operator having a nonempty domain;*

(iii) *The compactness assumption $(\tilde{\mathcal{A}}_c)$ is satisfied.*

Then $x : [0,T[\to \mathcal{H}$ is right-differentiable at $t = 0$ with

$$x'(0) = J_{D_pA(x(0)|x^*(0))}(\xi'(0)),$$

which means that $x'(0)$ is the unique solution of the following variational inclusion

$$\xi'(0) \in x'(0) + D_pA(x(0)|x^*(0))(x'(0)).$$

Proof. By Proposition 6.1, the compactness condition $(\tilde{\mathcal{A}}_c)$ ensures the maximality of the proto-derivative $D_pA(x(0)|x^*(0))(x'(0))$. The rest of the proof is similar to the one given in [2]. ∎

We now give now an equivalent result of Theorem 7.1 by replacing the PK-convergence by the BH-convergence.

Theorem 7.2 *Assume that $\mathcal{H}$ is general Hilbert space. Let $A : [0,T[\times\mathcal{H} \rightrightarrows \mathcal{H}$ be a maximally monotone operator, and let $\xi : [0,T[\to \mathcal{H}$ be a given function. We consider a solution $x : [0,T[\to \mathcal{H}$ to problem* (7.1)*. If the following assertions are satisfied:*

(i) *ξ is right-differentiable at $t = 0$;*

(ii) *A is BH-proto-differentiable at $x(0)$ relative to $x^*(0) := \xi(0) - x(0) \in A(0, x(0))$ with the proto-derivative operator having a nonempty domain;*

then $x : [0,T[\to \mathcal{H}$ is right-differentiable at $t = 0$ with

$$x'(0) = J_{\widetilde{D}_pA(x(0)|x^*(0))}(\xi'(0)),$$

which means that $x'(0)$ is the unique solution of the following variational inclusion

$$\xi'(0) \in x'(0) + \widetilde{D}_pA(x(0)|x^*(0))(x'(0)).$$

Proof. By Remark 6.2 (ii) and (6.4), we deduce from (ii) that the operator $\mathrm{J}_A := J_{A(t,\cdot)}(\cdot)$ is BH-proto-differentiable at $\xi(0)$ relative to $x(0)$ and its BH-proto-derivative is given by

$$\widetilde{D}_p\mathrm{J}_A(\xi(0)|x(0)) = \mathrm{J}_{\widetilde{D}_pA(x(0)|x^*(0))}.$$

In particular, we have
$$\widetilde{D}_p \mathrm{J}_A(\xi(0)|x(0))(\xi'(0)) = \mathrm{J}_{\widetilde{D}_p A(x(0)|x^*(0))}(\xi'(0)).$$
The function $x : [0, T[\to \mathcal{H}$ is a solution to problem (7.1) if and only if

$$x(t) = J_{A(t,\cdot)}(\xi(t)) = \mathrm{J}_A(t, \xi(t)).$$

Since the convergence for the bounded Hausdorff topology implies the convergence in the Painlevé-Kuratowski sense, we get by (ii) that A is PK-proto-differentiable at $x(0)$ relative to $x^*(0) := \xi(0) - x(0) \in A(0, x(0))$. Using the same argument as in the proof of [2, Theorem 1], we deduce the right-differentiability of $x(\cdot)$ at $t = 0$ and that

$$x'(0) = J_{D_p A(x(0)|x^*(0))}(\xi'(0)) = J_{\widetilde{D}_p A(x(0)|x^*(0))}(\xi'(0)),$$

which completes the proof. ∎

Remark 7.1 It is possible to extend the analysis to the more general variational inclusion

$$\text{find } x(t) \in \mathcal{H} \text{ such that } \xi(t) \in f(t, x(t)) + A(t, x(t)),$$

where $f : [0, T[\times \mathcal{H} \to \mathcal{H}$ is a single-valued map assumed to be uniformly Lipschitz continuous and uniformly strongly monotone. Using the notion of semi-differentiability for f, we can obtain the same result as in as in [2, Theorem 1] by replacing the PK-proto-differentiability of A by the BH-proto-differentiability.

8 Application to the variational sum of maximally monotone operator

Let $A, B : \mathcal{H} \rightrightarrows \mathcal{H}$ be two maximally monotone operators such that $\mathrm{Dom}(A) \cap \mathrm{Dom}(B) \neq \emptyset$. In general, their pointwise sum $A + B$, with $\mathrm{Dom}(A + B) = \mathrm{Dom}(A) \cap \mathrm{Dom}(B)$, is not a maximally monotone operator. An intensive literature has been devoted to establishing sufficient conditions on the two operators to ensure that their pointwise sum is also a maximally monotone operator. In a general way, these conditions express that the intersection of the domains of the two operators is sufficiently large, and they are called qualification condition. They give rise to the resolution of the monotone inclusion: given $y \in \mathcal{H}$, find $x \in \mathcal{H}$ solution of
$$A(x) + B(x) \ni y, \tag{8.1}$$
where the sum is taken in the pointwise sense. Yet, it has been discovered that taking the pointwise sum is often a too rigid approach, and that different notions of sum can be defined, with interesting variational properties. They give rise to generalized solutions of (8.1), with often rich physical interpretation such as viscosity solutions or entropy solutions. Historically, the sum defined via the Trotter-Lie-Kato formula naturally emerged to solve Schrödinger's equation with a singular potential. More recently, and related to solving optimization problems and monotone inclusions, the concept of variational sum naturally emerged. The starting point is the following remark:

On the one hand, for any $\mu > 0$, the operator $A + B_\mu$ is maximally monotone, where B_μ is the Yosida approximation of index $\mu > 0$ of B. On the other hand, B_μ graph converges to B as $\mu \to 0$. So it is natural to consider the graph convergence properties of the filtered sequence $(A + B_\mu)$, as $\mu \to 0$.

To obtain a commutative sum, one must be able to commute the role of A and B, and consider instead $A_\lambda + B$, which leads to the following definition.

We denote by $\mathcal{I} = \big\{(\lambda, \mu) \in \mathbb{R}^2 \;|\; \lambda \geq 0, \mu \geq 0,\ \lambda + \mu \neq 0\big\}$, and by $\mathcal{F}$ the filter of all the pointed neighborhoods of the origin in $\mathcal{I}$. We adopt the convention that $A_0 = A, B_0 = B$, so that when $(\lambda, \mu) \in \mathcal{I}$, at least one of the two operators A_λ or B_μ is continuous, which makes the sum $A_\lambda + B_\mu$ maximally monotone.

Definition 8.1 (Attouch-Baillon-Théra [4]) *Let $A, B : \mathcal{H} \rightrightarrows \mathcal{H}$ be two maximally monotone operators such that* $\mathrm{Dom}(A) \cap \mathrm{Dom}(B) \neq \emptyset$*. The variational sum $A \underset{v}{+} B$ is the monotone operator defined by*

$$\mathrm{Gph}\,(A \underset{v}{+} B) = \mathrm{Liminf}_{\mathcal{F}}\ \mathrm{Gph}\,(A_\lambda + B_\mu).$$

Of course, the central question is to identify the operator $A \underset{v}{+} B$, and to know if it is maximally monotone. When the operator $A \underset{v}{+} B$ is maximally monotone, then the filtered sequence $(A_\lambda + B_\mu)$ graph converges to $A \underset{v}{+} B$, see Proposition 2.2. Such positive answer has been obtained in [4] in the following situations:

a) The pointwise sum $A + B$ is maximally monotone. In this case $A \underset{v}{+} B = A + B$;

b) The closure $\overline{A+B}$ is maximally monotone. In this case $A \underset{v}{+} B = \overline{A+B}$;

c) The operators A and B are subdifferentials of closed convex proper functions, let $A = \partial f$, $B = \partial g$ with $\mathrm{dom} f \cap \mathrm{dom} g \neq \emptyset$. Then $A \underset{v}{+} B = \partial(f+g)$.

Many questions remain to be solved. In view of our developments, it would be interesting to answer the following question:

Suppose that the filtered sequence $(A_\lambda + B_\mu)$ graph converges. So is the limit operator maximally monotone? According to Theorem 3.2, if the convergence holds for the bounded Hausdorff topology, then $A \underset{v}{+} B$ is a maximally monotone operator.

Transposing Theorem 3.1 to our situation gives the following result

Theorem 8.1 *Suppose that the filtered sequence $(A_\lambda + B_\mu)$ graph converges, let $A \underset{v}{+} B$ be its limit whose domain is assumed to be nonempty. Suppose that the following compactness assumption $(\mathcal{A}_c)$ is satisfied: every filtered sequence $(x_{\lambda,\mu})$ of $\mathcal{H}$ such that*

$$\left(\sup_{\lambda,\mu} \|x_{\lambda,\mu}\| < +\infty \quad \textit{and} \quad \sup_{\lambda,\mu} \|A_\lambda x_{\lambda,\mu} + B_\mu x_{\lambda,\mu}\| < +\infty \right)$$

is contained in a compact subset of $\mathcal{H}$.

Then, $A \underset{v}{+} B$ is also a maximally monotone operator.

Due to the specific structure of the approximating sequence of operators $(A_\lambda + B_\mu)$, finding a counterexample as in section 5 would require taking a more sophisticated situation.

Acknowledgements. The research of the first and the third authors benefited from the support of the FMJH "Program Gaspard Monge for optimization and operations research and their interactions with data science", and from the support from EDF, Thales and Orange.

References

[1] S. Adly and L. Bourdin. *Sensitivity analysis of variational inequalities via twice epi-differentiability and proto-differentiability of the proximity operator*. SIAM J. Optim. Vol. 28, No. 2, pp. 1699-1725 (2018).

[2] S. Adly, R.T. Rockafellar, *Sensitivity analysis of monotone inclusions via the proto- differentiability of the resolvent operator*, Math. Program. 189, no. 1-2, Ser. B, 37-54 (2021).

[3] H. Attouch, *Variational convergence for functions and operators*, Pitman Advanced Publishing Program, Applicable Mathematics Series, 1984.

[4] H. Attouch, J.-B. Baillon, M. Théra, *Variational sum of maximal monotone operators*, J. Convex Analysis, 1 (1994), No. 1, pp. 1–29.

[5] H. Attouch, G. Buttazzo, G. Michaille, *Variational analysis in Sobolev and BV spaces. Applications to PDEs and optimization.* Second Edition, MOS/SIAM Series on Optimization, MO 17, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, (2014).

[6] H. Attouch, R. Lucchetti, R. Wets, *The topology of the ρ-Hausdorff distance*, Annali di Matematica Pura ed Applicata, 160 (1991), No. 2, pp. 303–320.

[7] H. Attouch, A. Moudafi, H. Riahi, *Quantitative stability analysis for maximal monotone operators and semi-groups of contractions*, Nonlinear Analysis, TMA, 21 (1993), No. 9, pp. 697–723.

[8] H. Attouch, R. Wets, *Quantitative stability of variational systems: I, The epigraphical distance*, Transactions of the American Mathematical Society, 328 (1991), No. 2, pp. 695–729.

[9] H. Attouch, R. Wets, *Quantitative stability of variational systems: II, A framework for nonlinear conditioning*, SIAM J. Optim., 3 (1993), pp. 359–381.

[10] H. Attouch, R. Wets, *Quantitative stability of variational systems: III, epsilon-approximate solutions*, Math. Program., 61 (1993), No. 2, pp. 197–214.

[11] H. Attouch, H., J.L. Ndoutoume, M. Thra. *Epigraphical convergence of functions and convergence of their derivatives in Banach spaces*. Sém. Anal. Convexe 20 (1990), Exp. No. 9, 45.

[12] D. Aze, J.-P. Penot, *Operations on convergent families of sets and functions*, Optimization, 21 (4) (1990), pp. 521-534.

[13] H. Bauschke, P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert spaces*, CMS Books in Mathematics, Springer, (2011).

[14] G. Beer, *Topologies on closed and closed convex sets*, Mathematics and its Applications, Vol. 268, Kluwer, 1993.

[15] G. Beer, *Conjugate convex functions and the epi-distance topology*. Proc. Amer. Math. Soc. 108 (1990), no. 1, 117-126.

[16] H. Brézis, *Opérateurs maximaux monotones dans les espaces de Hilbert et équations d'évolution*, Lecture Notes 5, North Holland, 1972.

[17] C. N. Do. *Generalized second-order derivatives of convex functions in reflexive Banach spaces*. Trans. Amer. Math. Soc., 334(1):281–301, 1992.

[18] A. B. Levy and R. T. Rockafellar. *Sensitivity analysis of solutions to generalized equations*. Trans. Amer. Math. Soc., 345(2):661–671, 1994.

[19] A. B. Levy and R. T. Rockafellar. Sensitivity of solutions in nonlinear programming problems with nonunique multipliers. In *Recent advances in nonsmooth optimization*, pages 215–223. World Sci. Publ., River Edge, NJ, 1995.

[20] A. B. Levy and R. T. Rockafellar. Variational conditions and the proto-differentiation of partial subgradient mappings. *Nonlinear Anal.*, 26(12):1951–1964, 1996.

[21] A. B. Levy, R. Poliquin, and L. Thibault. *Partial extensions of Attouch's theorem with applications to proto-derivatives of subgradient mappings.* Trans. Amer. Math. Soc., 347(4):1269–1294, 1995.

[22] S. Mrowka, *On the convergence of nets of sets*, Fund. Math., 45 (1958), pp. 237–246.

[23] J. Peypouquet, S. Sorin, *Evolution equations for maximal monotone operators: asymptotic analysis in continuous and discrete time*, J. Convex Anal, 17 (3-4) (2010), pp. 1113–1163.

[24] R. T. Rockafellar. *Maximal monotone relations and the second derivatives of nonsmooth functions*. Ann. Inst. H. Poincaré Anal. Non Linéaire, 2(3):167–184, 1985.

[25] R. T. Rockafellar. *Proto-differentiability of set-valued mappings and its applications in optimization*. Ann. Inst. H. Poincaré Anal. Non Linéaire, 6:449–482, 1989.

[26] R. T. Rockafellar. *Generalized second derivatives of convex functions and saddle functions*. Trans. Amer. Math. Soc., 322(1):51–77, 1990.

[27] R. T. Rockafellar. *Second-order convex analysis*. J. Nonlinear Convex Anal., 1(1):1-16, 2000.

[28] R.T. Rockafellar and R. Wets, *Variational analysis*, Springer, Berlin, 1998.

[29] G. Wachsmuth, *Maximal monotone operators with non-maximal graphical limit*. Examples and Counterexamples, Volume 2, November 2022, 100073.

[30] E. Zeidler, *Nonlinear Functional Analysis and its Applications*, II/B, Springer-Verlag, 1990.