



Forward-backward approximation of nonlinear semigroups in finite and infinite horizon

Andrés Contreras, Juan Peypouquet

► To cite this version:

Andrés Contreras, Juan Peypouquet. Forward-backward approximation of nonlinear semigroups in finite and infinite horizon. Communications on Pure and Applied Analysis, 2021, 20 (5), pp.1893-1906. 10.3934/cpaa.2021051 . hal-03528849

HAL Id: hal-03528849

<https://hal.science/hal-03528849>

Submitted on 17 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

FORWARD-BACKWARD APPROXIMATION OF NONLINEAR SEMIGROUPS IN FINITE AND INFINITE HORIZON

ANDRÉS CONTRERAS

University of Paris-Saclay, CentraleSuplec, OPIS, Inria Saclay 91190, Gif-sur-Yvette, France.
Departamento de Ingeniería Matemática & Centro de Modelamiento Matemático (CNRS UMI2807).
FCFM, Universidad de Chile, Beauchef 851, Santiago, Chile.

JUAN PEYPOUQUET*

University of Groningen, Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence.
Bernoulliborg, Nijenborgh 9, 9747 AG Groningen, The Netherlands.
Departamento de Ingeniería Matemática & Centro de Modelamiento Matemático (CNRS UMI2807).
FCFM, Universidad de Chile, Beauchef 851, Santiago, Chile.

(Communicated by the associate editor name)

ABSTRACT. This work is concerned with evolution equations and their forward-backward discretizations, and aims at building bridges between differential equations and variational analysis. Our first contribution is an estimation for the distance between iterates of sequences generated by forward-backward schemes, useful in the convergence and robustness analysis of iterative algorithms of widespread use in numerical optimization and variational inequalities. Our second contribution is the approximation, on a bounded time frame, of the solutions of evolution equations governed by accretive (monotone) operators with an additive structure, by trajectories constructed by interpolating forward-backward sequences. This provides a short, simple and self-contained proof of existence and regularity for such solutions; unifies and extends a number of classical results; and offers a guide for the development of numerical methods. Finally, our third contribution is a mathematical methodology that allows us to deduce the behavior, as the number of iterations tends to $+\infty$, of sequences generated by forward-backward algorithms, based solely on the knowledge of the behavior, as time goes to $+\infty$, of the solutions of differential inclusions, and viceversa.

1. Introduction. Semigroup theory is a relevant tool in the study of ordinary and partial differential equations, as well as differential inclusions, which appear, for instance, in contact mechanics, optimization, variational analysis and game theory. Among its applications, it helps analyze the evolution of flows in mechanical systems, and establish convergence and convergence rates for numerical optimization algorithms. One of its cornerstones was the Hille-Yosida Theorem [10, 29], which states that an *unbounded* linear operator A , on a Banach space X , is the

2020 *Mathematics Subject Classification.* 34A60, 37L05, 65K15.

Key words and phrases. Nonlinear semigroups, differential inclusions, accretive operators, monotone operators, discrete approximations, forward-backward iterations, asymptotic equivalence.

Supported by FONDECYT Grant 1181179, CMM-Conicyt PIA AFB170001, ECOS-CONICYT Grant C18E04 and CONICYT-PFCHA/DOCTORADO NACIONAL/2016 21160994.

* Corresponding author: Juan Peypouquet.

infinitesimal generator of a strongly continuous semigroup $(\mathcal{S}_t)_{t \geq 0}$ of nonexpansive linear operators on X , satisfying $-\dot{u}(t) = Au(t)$ if, and only if, A is closed and densely defined, its spectrum does not intersect \mathbb{R}_- , and the resolvents satisfy an appropriate bound. This result was complemented by the Lumer-Phillips Theorem [15], which provides an alternative—and perhaps more practical—characterization in terms of semidefiniteness. It is important to mention that Hille and Yosida used different strategies to construct the semigroup (that is, to show the necessity). Yosida’s approach consisted in approximating A by a family $(A_\lambda)_{\lambda > 0}$ of *bounded* ones, establishing the existence of solution to the regularized differential equation $-\dot{u}_\lambda(t) = A_\lambda u_\lambda(t)$ by classical arguments, and then passing to the limit while showing that the regularized solutions u_λ converge to a true solution of the original problem. Hille, in turn, discretized the time interval $[0, T]$, where $T > 0$ is arbitrary but fixed, constructed approximating trajectories using a sequence of points generated by resolvent iterations, and finally passed to the limit as the partition is refined. Both showed the convergence is uniform on $[0, T]$.

Two decades later, sufficient conditions for a nonlinear—and possibly multi-valued—operator A to generate a strongly continuous semigroup $(\mathcal{S}_t)_{t \geq 0}$ of nonexpansive nonlinear operators that solves the differential inclusion $-\dot{u}(t) \in Au(t)$ were discovered. Yosida’s approach was used by Brézis [6], while Hille’s path was followed by Crandall and Pazy [7], and then simplified and perfected by Rasmussen [22] and Kobayashi [13]. The Rasmussen-Kobayashi method provided a concise and sharp inequality—a special case of Theorem 2.1 below—to bound the distance between two sequences of points generated using compositions of resolvents. Other authors have analyzed the nonautonomous setting [14, 1], where there is a function $t \mapsto A(t)$ that generates an *evolution system* that, of course, is not a semigroup, in general. In some relevant special cases, resolvents may be replaced by Krasnosel’skiĭ-Mann iterations. This issue is addressed in [27, 8], where applications in optimization and game theory are given.

A few years later, Passty [19] introduced the notion of an *asymptotic semigroup*, which is, roughly speaking, a possibly nonautonomous evolution system that asymptotically behaves like a semigroup. This concept allows us to deduce several convergence properties of the trajectories generated by an asymptotic semigroup, based on what is known about those generated by the semigroup itself. A similar idea lies behind the notion of *almost-orbit* (see [17]), which helps to prove that every nonexpansive iterative algorithm is robust against summable errors (see [20, Lemma 5.3]). The interested reader is referred to [1, 2, 3] for further details and applications. Passty proved, under some restrictive assumptions, that every sequence $(x_n)_{n \in \mathbb{N}}$ generated using compositions of resolvents of A converges strongly (weakly) as $n \rightarrow +\infty$ if, and only if, all trajectories generated by the semigroup $(\mathcal{S}_t)_{t \geq 0}$ converge strongly (weakly) as $t \rightarrow +\infty$.

The process of generating sequences of points using resolvent iterations is also known as the *proximal point algorithm*, as developed by Martinet [16]. It is one of the fundamental building blocks of first order methods used to solve nonsmooth optimization problems and variational inequalities in practice (see the note on forward-backward iterations below).

Passty’s innovative idea is remarkable, since it makes it possible to use calculus techniques, such as derivation and integration, to analyze the behavior of iterative algorithms. A few years later, Miyadera and Kobayashi [17] and Sugimoto and Koizumi [26] were able to get rid of Passty’s superfluous hypotheses. This approach

also enabled Güler [9] to show, based on an example of Baillon [4], that there is a proper, lower-semicontinuous, convex function for which the proximal point algorithm produces sequences that converge weakly but not strongly, settling an open question in optimization theory posed by Rockafellar [23] fifteen years earlier. As a matter of fact, this function may be chosen differentiable and with Lipschitz-continuous gradient, as proved by the authors in [8].

Forward-backward iterations combine the principles of proximal, Krasnosel'skiĭ-Mann and Euler iterations. They are fundamental in the numerical analysis of structured optimization problems and variational inequalities, since they represent the core of first order methods that can be applied to minimize functions with smooth and nonsmooth features.

The purpose of this research is to extend, unify and condense the theory concerning the generation of strongly continuous semigroups of nonlinear and nonexpansive mappings by multi-valued operators with an additive structure. On the one hand, we analyze the approximation of solutions for the differential inclusion $-\dot{u}(t) \in (A+B)u(t)$ by trajectories constructed by interpolation of sequences generated using forward-backward iterations, on a compact time interval. This approach is different from the one by Trotter [24] and Kato [12], which uses *double-backward* iterations. Double-backward iterations require the (costly!) computation of both resolvents. On the other hand, we establish asymptotic equivalence results that link the behavior, as the number of iterations tends to $+\infty$, of sequences generated by forward-backward iterations, with the behavior of the solutions of the differential inclusion $-\dot{u}(t) \in (A+B)u(t)$, as time t tends to $+\infty$. We obtain new strong convergence results for forward-backward sequences as straightforward corollaries. We have aimed at presenting these findings in a simple and pedagogic manner, accessible to researchers in functional analysis, differential equations and optimization alike.

Although the Hilbert space setting is suitable for many applications, our results may be stated and proved in a class of Banach spaces with no additional effort. The extension to general Banach spaces is an open question.

The paper is organized as follows: In Section 2, we give the notation and definitions, along with a description of the fundamental technical tool required to prove our main results. The approximation in a finite time horizon is discussed in Section 3. Section 4 is devoted to the approximation in an infinite time horizon and contains new convergence results for forward-backward sequences. The technical proofs are given in Section 5.

2. Forward-backward iterations defined by accretive and cocoercive operators. Let X be a Banach space with topological dual X^* . Their norms and the duality product are denoted by $\|\cdot\|$, $\|\cdot\|_*$ and $\langle \cdot, \cdot \rangle$, respectively. The *duality mapping* $j : X \rightarrow X^*$ is defined by

$$j(x) = \{x^* \in X^* : \langle x^*, x \rangle = \|x\|^2 = \|x^*\|_*^2\}.$$

In what follows, we assume that X^* is 2-uniformly convex, which implies that X is reflexive, the duality mapping is single valued, and there is a constant $\kappa > 0$ such that

$$\|u + v\|^2 \leq \|u\|^2 + 2\langle j(u), v \rangle + \kappa\|v\|^2, \quad (1)$$

for all $u, v \in X$ (see [28]). For instance, Hilbert spaces have this property, as well as L^p spaces, for $p \geq 2$.

A set-valued operator $A : X \rightarrow 2^X$ is *accretive* if, whenever $u \in Ax$ and $v \in Ay$, we have

$$\|x - y + \lambda(u - v)\| \geq \|x - y\|$$

for all $\lambda > 0$. If, moreover, $I + \lambda A$ is surjective for all $\lambda > 0$, we say A is *m-accretive*. In this case, its *resolvent*, defined as $J_\lambda = (I + \lambda A)^{-1}$, is single-valued, everywhere defined and nonexpansive. It follows from [11, Lemma 1.1] that A is accretive if, and only if, it is *monotone*, which means that

$$\langle j(x - y), u - v \rangle \geq 0,^1$$

whenever $u \in Ax$ and $v \in Ay$. Next, an operator $B : X \rightarrow X$ is *cocoercive* with parameter $\theta > 0$ if

$$\langle j(x - y), Bx - By \rangle \geq \theta \|Bx - By\|^2,$$

for all $x, y \in X$. Clearly, if B is cocoercive with parameter θ , it is Lipschitz-continuous with constant $\frac{1}{\theta}$. Moreover, the operator $E_\lambda : X \rightarrow X$, defined by

$$E_\lambda = I - \lambda B, \tag{2}$$

is nonexpansive for all $\lambda \in [0, \frac{2\theta}{\kappa}]$. Finally, if A is *m-accretive* and B is cocoercive, then $A + B$ is *m-accretive*, and the *forward-backward splitting operator* $T_\lambda : X \rightarrow X$, defined by

$$T_\lambda = J_\lambda \circ E_\lambda,$$

is single-valued, everywhere defined and nonexpansive. These are the standing assumptions on X , A and B for the rest of the paper.

Remark 1. Even in \mathbb{R} , the behavior of E_λ may become erratic if the bound $\lambda\kappa \leq 2\theta$ does not hold. To see this, set $Bx = f'(x)$, where $f(x) = x^2$. Here, $\kappa = 1$ and $\theta = 1/2$. If we take $\lambda = 1 + \varepsilon$, with $\varepsilon > 0$, then $x_n = (-1)^n(1 + 2\varepsilon)^n x_0$. If $x_0 \neq 0$, $|x_n|$ goes to infinity exponentially as $n \rightarrow \infty$. On the other hand, for every initial condition the solution to $-\dot{x}(t) = Bx(t) = 2x(t)$ remains bounded and converges exponentially to zero as $t \rightarrow \infty$.

Remark 2. Actually, the minimal hypothesis on λ , B and X , required for our proofs to hold, is that E_λ be nonexpansive for all $\lambda \in [0, \Lambda]$ for some $\Lambda > 0$. Some definitions and proofs must be slightly adjusted if the duality mapping j is not single-valued. If $B = 0$, no assumptions need be made on X or λ .

We are interested in the study of sequences satisfying

$$x_k = T_{\lambda_k}(x_{k-1}) = J_{\lambda_k}(E_{\lambda_k}(x_{k-1})) \tag{3}$$

for $k \geq 1$, where (λ_k) is a sequence of positive numbers, called *step sizes*, and $x_0 \in X$ is the *initial point*. We mentioned earlier that these sequences are fundamental in the numerical analysis of optimization problems, variational inequalities and fixed-point problems. However, our purpose here is to analyze them as discrete approximations of an evolution equation governed by the sum $A + B$. To this end, it is useful to rewrite (3) as

$$-\frac{x_k - x_{k-1}}{\lambda_k} \in Ax_k + Bx_{k-1}, \quad k \geq 1, \tag{4}$$

or, more generally, as

$$-\frac{x_k - x_{k-1}}{\lambda_k} + \varepsilon_k \in Ax_k + Bx_{k-1}, \quad k \geq 1, \tag{5}$$

¹In Hilbert spaces, this terminology is preferred, and the inequality reads $\langle x - y, u - v \rangle \geq 0$, where $\langle \cdot, \cdot \rangle$ is the inner product.

where ε_k accounts for possible perturbations or computational errors. In the notation of formula (3), this is

$$x_k = J_{\lambda_k}(E_{\lambda_k}(x_{k-1}) + \lambda_k \varepsilon_k). \quad (6)$$

Back to the exact version (4), the left-hand side can be interpreted as a discretization of the velocity for a trajectory $t \mapsto u(t)$, so (4) can be related to the differential inclusion

$$-\dot{u}(t) \in Au(t) + Bu(t), \quad (7)$$

for $t > 0$. In the following sections, we shall establish the nature of this relationship. On the one hand, we shall prove that the iterations described in (4) can be used, in at least two different ways, to construct a sequence of curves that approximate the solutions of (7) uniformly on each compact time interval. The existence of such solutions is recovered as a byproduct. On the other hand, we shall show that, given A and B , the trajectories satisfying (7) will have the same convergence properties, when $t \rightarrow \infty$, as the sequences satisfying (4), when $k \rightarrow \infty$, provided the step sizes are sufficiently small. The key mathematical tool is the following inequality, whose proof is technical, and will be given in Section 5.

Theorem 2.1. *Let (x_k) , (\hat{x}_l) be two sequences generated by (5), with step sizes (λ_k) and $(\hat{\lambda}_l)$, as well as error sequences (ε_k) and $(\hat{\varepsilon}_l)$. Assume $\lambda_k, \hat{\lambda}_l \leq \frac{\theta}{\kappa}$ for all $k, l \in \mathbb{N}$. Then, for $u \in D(A)$ fixed, and each $k, l \in \mathbb{N}$, we have*

$$\|x_k - \hat{x}_l\| \leq \|x_0 - u\| + \|\hat{x}_0 - u\| + \|(A + B)u\| \sqrt{(\sigma_k - \hat{\sigma}_l)^2 + \tau_k + \hat{\tau}_l} + e_k + \hat{e}_l, \quad (8)$$

where $\|Au\| = \inf_{v \in Au} \|v\|$, $\sigma_k = \sum_{i=1}^k \lambda_i$, $\tau_k = \sum_{i=1}^k \lambda_i^2$ and $e_k = \sum_{i=1}^k \lambda_i \|\varepsilon_i\|$ (similarly for $\hat{\sigma}_l$, $\hat{\tau}_l$ and \hat{e}_l).

We first became aware of an inequality of this sort (for $B \equiv 0$ and slightly less sharp) in [9], where Güler attributes it to Kobayashi [13] (see also [21]). However, the main arguments were given by Rasmussen [22], who simplified the proof of Crandall and Liggett [?], ultimately based on that of Hille [10]. Similar estimations are given in [14, 1] (still for $B = 0$, but for a time-dependent A) and in [27, 8] for $A = 0$.

3. Approximation in finite horizon. Theorem 2.1 provides existence and regularity results for the evolution equation

$$\begin{cases} -\dot{u}(t) & \in (A + B)u(t), \quad \text{for almost every } t > 0, \\ u(0) & = u_0 \in D(A), \end{cases} \quad (9)$$

by means of an approximation scheme. For each $t \geq 0$ and $m \geq 1$, set

$$u_m(t) = \left[T_{\frac{t}{m}} \right]^m u_0. \quad (10)$$

In other words, $u_m(t)$ is the m -th term of the forward-backward sequence generated by (3) from u_0 using the constant step size $\lambda_k \equiv t/m$. We shall prove that (u_m) converges uniformly on compact intervals to a Lipschitz-continuous function satisfying (9). We begin by establishing the convergence.

Proposition 1. *The sequence (u_m) converges pointwise on $[0, \infty)$, and uniformly on $[0, S]$ for each $S > 0$, to a function $u : [0, \infty) \rightarrow X$, which is globally Lipschitz-continuous with constant $\|(A + B)u_0\|$.*

Proof. Take $u_0 \in D(A)$. Given $t, s > 0$ and $n, m \in \mathbb{N}$, define $u_m(t)$ and $u_n(s)$ as above. By Theorem 2.1, we have

$$\|u_m(t) - u_n(s)\| \leq \|(A + B)u_0\| \sqrt{(t - s)^2 + \frac{t^2}{m} + \frac{s^2}{n}}. \quad (11)$$

For $s = t$, this gives

$$\|u_m(t) - u_n(t)\| \leq t \|(A + B)u_0\| \sqrt{\frac{1}{m} + \frac{1}{n}}.$$

It follows that (u_m) converges pointwise on $[0, \infty)$, and uniformly on $[0, S]$ for each $S > 0$, to a function $u : [0, \infty) \rightarrow X$. Passing to the limit in (11), as $m, n \rightarrow \infty$, we obtain

$$\|u(t) - u(s)\| \leq \|(A + B)u_0\| |t - s|$$

for all $t, s > 0$. □

Theorem 3.1. *The function u , given by Proposition 1, satisfies (9).*

Proof. We shall verify that u is an integral solution of (9) in the sense of B enilan (see [5]), which means that, whenever $y \in (A + B)x$ and $S \geq t > s \geq 0$, we have

$$\|u(t) - x\|^2 - \|u(s) - x\|^2 \leq 2 \int_s^t \langle j(x - u(\tau)), y \rangle d\tau. \quad (12)$$

If (x_n) is any sequence generated by (4) with steps sizes (λ_n) , then

$$-(x_n - x_{n-1}) - \lambda_n Bx_{n-1} + \lambda_n Bx_n \in \lambda_n Ax_n + \lambda_n Bx_n$$

for each $n \geq 1$. In view of the monotonicity of $A + B$, we have

$$\langle j(x - x_n), \lambda_n y + x_n - x_{n-1} + \lambda_n Bx_{n-1} - \lambda_n Bx_n \rangle \geq 0,$$

whenever $y \in Ax + Bx$. Whence,

$$\begin{aligned} 2\lambda_n \langle j(x - x_n), y \rangle &\geq 2\langle j(x - x_n), x_{n-1} - x_n \rangle + 2\lambda_n \langle j(x - x_n), Bx_n - Bx_{n-1} \rangle \\ &= 2\|x_n - x\|^2 + 2\langle j(x - x_n), x_{n-1} - x \rangle \\ &\quad + 2\lambda_n \langle j(x - x_n), Bx_n - Bx_{n-1} \rangle \\ &\geq \|x_n - x\|^2 - \|x_{n-1} - x\|^2 + 2\lambda_n \langle j(x - x_n), Bx_n - Bx_{n-1} \rangle \\ &\geq \|x_n - x\|^2 - \|x_{n-1} - x\|^2 \\ &\quad - 2\theta^{-1}\lambda_n \|x - x_n\| \|x_n - x_{n-1}\|, \end{aligned} \quad (13)$$

where θ is the cocoercivity parameter of B .

Now, let us choose $x_0 = u_0$, $\lambda_n \equiv \frac{S}{m}$, where m is fixed but arbitrary. Let us bound the factors $\|x - x_n\|$ and $\|x_n - x_{n-1}\|$ using Theorem 2.1. On the one hand, we use $k = n$, $\ell = 0$ and $\hat{x}_0 = u = x$ to obtain

$$\|x_n - x\| \leq \|x_0 - x\| + \|(A + B)x\| \sqrt{\sigma_n^2 + \tau_n} \leq \|x_0 - x\| + \sqrt{2S} \|(A + B)x\|.$$

On the other, we write $k = n$, $\ell = n - 1$ and $x_0 = \hat{x}_0 = u = u_0$ to deduce that

$$\|x_n - x_{n-1}\| \leq \|(A + B)u_0\| \sqrt{\lambda_n^2 + \tau_n + \tau_{n-1}} \leq \frac{\sqrt{2S}}{\sqrt{m}} \|(A + B)u_0\|.$$

Using these bounds, (13) gives

$$\|x_n - x\|^2 - \|x_{n-1} - x\|^2 \leq 2\lambda_n \langle j(x - x_n), y \rangle + \frac{K}{m\sqrt{m}},$$

where K depends on θ , S , x and u_0 . In order to construct a Riemann sum for the integral in (12), we first take $\nu_m : [0, S] \rightarrow \{0, 1, \dots, m\}$ such that $\lim_{m \rightarrow \infty} \frac{\nu_m(t)}{m} \rightarrow \frac{t}{S}$. Then, we sum for $n = \nu_m(s), \dots, \nu_m(t)$, to obtain

$$\begin{aligned} \|v_m(t) - x\|^2 - \|u(s) - x\|^2 &\leq 2 \sum_{n=\nu_m(s)}^{\nu_m(t)} \frac{S}{m} \langle j(x - x_n), y \rangle + \sum_{n=\nu_m(s)}^{\nu_m(t)} \frac{K}{m\sqrt{m}} \\ &\leq 2 \sum_{n=\nu_m(s)}^{\nu_m(t)} \frac{S}{m} \langle j(x - x_n), y \rangle + \frac{K}{\sqrt{m}}. \end{aligned}$$

We obtain (12) by letting $m \rightarrow \infty$. \square

Existence of solution for (9) can be recovered as a consequence of the preceding arguments.

Corollary 1. *The differential inclusion (9) has a unique solution.*

Proof. We have constructed a sequence (u_m) that converges pointwise on $[0, \infty)$, and uniformly on $[0, S]$ for each $S > 0$, to a Lipschitz-continuous function $u : [0, \infty) \rightarrow X$, which satisfies (9) by Theorem 3.1. For the uniqueness, suppose that $u, v : [0, \infty)$ satisfy (9). For almost every t , we have

$$\frac{d}{dt} \|u(t) - v(t)\|^2 = \langle u(t) - v(t), \dot{u}(t) - \dot{v}(t) \rangle \leq 0, \quad (14)$$

since $-\dot{u}(t) \in (A + B)u(t)$, $-\dot{v}(t) \in (A + B)v(t)$, and $A + B$ is monotone. As a consequence, $\|u(t) - v(t)\| \leq \|u(0) - v(0)\| = 0$ for all $t > 0$. \square

Another consequence of the results above is:

Corollary 2. *Let (x_k) be a sequence generated by (3) and let $u : [0, S] \rightarrow X$ be a solution of (9). Then*

- (i) *The function $t \mapsto \|(A + B)u(t)\|$ is nonincreasing.*
- (ii) $\|x_k - u(t)\| \leq \|x_0 - u_0\| + \min \{ \|(A + B)x_0\|, \|(A + B)u_0\| \} \sqrt{(\sigma_k - t)^2 + \tau_k}.$

Proof. For (i), fix $t \geq t_0 \geq 0$ and $h > 0$, and follow the proof of Proposition 1, but starting from the initial condition $u(t_0)$, to deduce that

$$\|u(t+h) - u(t)\| \leq h \|(A + B)u(t_0)\|.$$

This implies $\|\dot{u}(t)\| \leq \|(A + B)u(t_0)\|$. However, since $-\dot{u}(t) \in (A + B)u(t)$, we have $\|(A + B)u(t)\| \leq \|\dot{u}(t)\|$ (actually, [6, Théorème 3.1] shows they are equal). The proof of (ii) is also similar to that of Proposition 1, but passing to the limit in only one of the sequences. More precisely, use Theorem 2.1 to obtain

$$\|x_k - u_n(t)\| \leq \|x_0 - u\| + \|u - u_0\| + \|(A + B)u\| \sqrt{(\sigma_k - t)^2 + \tau_k + \frac{t^2}{n}},$$

then let $n \rightarrow \infty$, and finally take either $u = u_0$ or $u = x_0$ to achieve the minimum. \square

Remark 3. The function u in Proposition 1 and Theorem 3.1 can be approximated by a sequence of piecewise constant functions. Given $S > 0$ and $m \geq 1$, define $v_m : [0, S] \rightarrow X$ by

$$v_m(t) = \left[T_{\frac{S}{m}} \right]^{\mu(t)} u_0, \quad \text{where} \quad \mu(t) = \left\lfloor m \frac{t}{S} \right\rfloor \quad \text{and} \quad t \in [0, S]. \quad (15)$$

This is a piecewise constant interpolation of the forward-backward sequence generated with $\frac{S}{m}$ as step sizes, and initial point u_0 for $k = 1, \dots, m$. In order to estimate the distance between v_m and u_m (defined in (10)), we use (8) to obtain

$$\|u_m(t) - v_m(t)\| \leq \|(A + B)u_0\| \sqrt{\frac{S^2}{m^2} + \frac{t^2}{m} + \frac{tS}{m}} \leq \frac{\sqrt{3}S}{\sqrt{m}} \|(A + B)u_0\|.$$

Whence, as $m \rightarrow \infty$, v_m converges uniformly on $[0, S]$, for each $S > 0$, to the same function u .

4. Approximation in infinite horizon. In this section, we show that, as the number of iterations goes to infinity, the forward-backward sequences generated by (3) have the same asymptotic behavior as the solutions of the evolution equation (9) do when $t \rightarrow \infty$. The key argument is the idea of asymptotic equality introduced by Passty [19], closely related to the notion of almost-orbit, introduced by Miyadera and Kobayasi [17]. Further commentaries on this topic can be found in [1, 2, 3].

In order to simplify the notation, given $x \in D(A)$ and $t \geq 0$, we write

$$\mathcal{S}_t x = u(t), \quad (16)$$

where u satisfies (9) with $u_0 = x$. Also, for $0 \leq s \leq t$, we write

$$U_{\mathcal{S}}(t, s) = \mathcal{S}(t - s). \quad (17)$$

In a similar fashion, if $n \in \mathbb{N}$ and $x \in H$, we denote

$$\mathcal{T}_n x = T_{\lambda_n} \circ \dots \circ T_{\lambda_1} x. \quad (18)$$

In other words, $\mathcal{T}_n x$ is the n -th term of the forward-backward sequence starting from $x \in D(A)$. Assume $(\lambda_n) \notin \ell^1$, and write $\nu(t) = \max\{n \in \mathbb{N} : \sigma_n \leq t\}$. For $0 \leq s \leq t$, we set

$$U_{\mathcal{T}}(t, s) = \prod_{i=\nu(s)+1}^{\nu(t)} T_{\lambda_i}, \quad (19)$$

where the product denotes composition of functions and the empty composition is the identity. With this notation, the function $t \mapsto \mathcal{T}_{\nu(t)} x$ defines a piecewise constant interpolation of the sequence $\mathcal{T}_n x$.

A *nonexpansive evolution system* on X is a family $(U(t, s))_{0 \leq s \leq t}$ such that

- (i) $U(t, t)z = z$ for all $z \in X$ and $t \geq 0$.
- (ii) $U(t, s)U(s, r)z = U(t, r)z$ for all $z \in X$ and all $t \geq s \geq r \geq 0$.
- (iii) $\|U(t, s)x - U(t, s)y\| \leq \|x - y\|$ for all $x, y \in X$ and $t \geq s \geq 0$.

Example 1. The families $(U_{\mathcal{S}})$ and $(U_{\mathcal{T}})$, defined in (17) and (19), respectively, are nonexpansive evolution systems. Actually, the same is true if \mathcal{S} is replaced by any other semigroup of nonexpansive functions on X , and if each T_{λ_i} is replaced by any other nonexpansive function on X .

A function $\phi : [0, \infty) \rightarrow X$ is an *almost-orbit* of the nonexpansive evolution system U if

$$\lim_{t \rightarrow \infty} \sup_{h \geq 0} \|\phi(t + h) - U(t + h, t)\phi(t)\| = 0.$$

The following result (see [2, Theorem 3.3]) reveals the usefulness of the concept of almost-orbit.

Proposition 2. *Let U be a nonexpansive evolution system and let ϕ be an almost-orbit of U . If, for each $x \in X$ and $s \geq 0$, $U(t, s)x$ converges weakly (resp. strongly) as $t \rightarrow \infty$, then so does $\phi(t)$. The same holds if the word “converges” is replaced by “almost-converges” or “converges in average”.*

Several examples and applications, along with additional commentaries can be found in [2, 8].

The following result establishes a relationship between the trajectories generated by U_S and U_T :

Theorem 4.1. *Let $(\lambda_n) \in \ell^2 \setminus \ell^1$, and fix $x \in X$. For each $t > 0$, define $\phi_S(t) = S_t x$ and $\phi_T(t) = T_{\nu(t)} x$. Then, ϕ_S is an almost-orbit of U_T , and ϕ_T is an almost-orbit of U_S .*

Proof. We first prove that ϕ_S is an almost-orbit of U_T . By Theorem 2.1 and Corollary 2, we have

$$\begin{aligned} \left\| \left[\prod_{k=1}^m T_{\frac{h}{m}} \right] S_t x - \left[\prod_{i=\nu(t)+1}^{\nu(t+h)} T_{\lambda_i} \right] S_t x \right\| &\leq \|(A+B)S_t x\| \sqrt{\left(\sigma_{\nu(t)+1}^{\nu(t+h)} - h \right)^2 + \tau_{\nu(t)+1}^{\nu(t+h)} + \frac{h^2}{m}} \\ &\leq \|(A+B)x\| \sqrt{4\rho^2(t) + \tau_{\nu(t)+1}^\infty + \frac{h^2}{m}}, \end{aligned}$$

where $\sigma_k^n = \sigma_n - \sigma_k$, $\tau_k^n = \tau_n - \tau_k$ and $\rho(t) := \sup\{\lambda_n : n \geq \nu(t) - 1\}$, which vanishes as $t \rightarrow \infty$. Passing to the limit as $m \rightarrow \infty$, we obtain

$$\|S_h S_t x - U_T(t+h, t) S_t x\| \leq \|(A+B)x\| \sqrt{4\rho^2(t) + \tau_{\nu(t)+1}^\infty},$$

which tends to 0 as $t \rightarrow \infty$, uniformly in $h \geq 0$. It follows that

$$\lim_{t \rightarrow \infty} \sup_{h \geq 0} \|\phi_S(t+h) - U_T(t+h, t) \phi_S(t)\| = 0.$$

To prove that ϕ_T is an almost-orbit of U_S , we proceed in a similar fashion, to obtain

$$\left\| \left[\prod_{i=\nu(t)+1}^{\nu(t+h)} T_{\lambda_i} \right] T_{\nu(t)} x - \left[\prod_{k=1}^m T_{\frac{h}{m}} \right] T_{\nu(t)} x \right\| \leq \|(A+B)x\| \sqrt{4\rho^2(t) + \tau_{\nu(t)+1}^\infty + \frac{h^2}{m}}.$$

Then, we pass to the limit as $m \rightarrow \infty$ to deduce that

$$\|\phi_T(t+h) - S_h \phi_T(t)\| \leq \|(A+B)x\| \sqrt{4\rho^2(t) + \tau_{\nu(t)+1}^\infty},$$

and conclude. \square

Theorem 4.1 implies [19, Lemmas 4 & 6], [26, Proposition 2.3], [17, Proposition 7.4], [21, Propositions 8.6 i) & 8.7] and [8, Theorem 3.1]. Combining Theorem 4.1 with Proposition 2, and using [20, Lemma 5.3], we obtain

Theorem 4.2. *The following statements are equivalent:*

- i) *For every $z \in \overline{D(A)}$, $S_t z$ converges strongly (weakly), as $t \rightarrow +\infty$.*
- ii) *For every initial point $x_0 \in X$, every sequence of step sizes $(\lambda_n)_{n \geq 1} \in \ell^2 \setminus \ell^1$, and every sequence of errors $(\varepsilon_k)_{k \geq 1}$ such that $\sum_{k \geq 1} \|\varepsilon_k\| < +\infty$, the sequence (x_n) , generated by (5), converges strongly (weakly), as $n \rightarrow +\infty$.*
- iii) *There exists a sequence of step sizes $(\lambda_n)_{n \geq 1} \in \ell^2 \setminus \ell^1$ such that, for every initial point $x_0 \in X$, the sequence (x_n) , generated by (4), converges strongly (weakly), as $n \rightarrow +\infty$.*

Theorem 4.2 implies [19, Theorems 1 & 2], [26, Theorem], [17, Theorem 7.5], as well as [8, Theorem 3.2].

New convergence results for forward-backward sequences on Banach spaces. Theorem 4.2 can automatically give new convergence results for forward-backward sequences by translating the information available on the behavior of the semigroup. Theorem 4.3 below is provided as a *methodological* example, to show how this indirect analysis can be carried out. Therefore, we have privileged statement simplicity over generality.

Recall from Section 2 that X is a Banach space with 2-uniformly convex dual, A is m -accretive and B is cocoercive. Let $(\varepsilon_k)_{k \geq 1}$ be a sequence representing computational errors and let $(x_k)_{k \geq 0}$ satisfy (5). We assume that $\sum_{k \geq 1} \|\varepsilon_k\| < +\infty$. Finally, set $\mathcal{A} = A + B$ and $\Sigma = \mathcal{A}^{-1}0$, and assume $\Sigma \neq \emptyset$. To simplify the statements and arguments, suppose X is uniformly convex. We know that Σ is closed and convex, and the projection P_Σ is well defined, single-valued and continuous.

Theorem 4.3. *Let $(\lambda_n)_{n \geq 1} \in \ell^2 \setminus \ell^1$. Assume one of the following conditions holds:*

- i) *There is $\alpha > 0$ such that for every $x \notin \Sigma$ and every $y \in \mathcal{A}x$, $\langle j(x - P_\Sigma x), y \rangle \geq \alpha \|x - P_\Sigma x\|^2$;*
- ii) *J_1 is compact and, for every $x \notin \Sigma$ and every $y \in \mathcal{A}x$, $\langle j(x - P_\Sigma x), y \rangle > 0$; or*
- iii) *The interior of Σ is not empty.*

Then, x_n converges strongly, as $n \rightarrow +\infty$, to a point in Σ .

Proof. In all three cases, we first prove that for each $z \in D(A)$, $\mathcal{S}_t z$ converges strongly, as $t \rightarrow +\infty$, to a point in Σ .

- i) The hypotheses of [18, Theorem 1] are easily verified.
- ii) It suffices to combine [18, Proposition 1] and [18, Theorem 1].
- iii) We use [18, Theorem 4].

We conclude by applying Theorem 4.2. □

5. Proof of the fundamental inequality. This last section is devoted to the proof of Theorem 2.1. In order to simplify the notation, given $\nu > 0$ and $z, d \in X$, write

$$E_\lambda^\varepsilon(z) = E_\lambda(z) + \lambda\varepsilon, \quad \text{and} \quad T_\lambda^\varepsilon(z) = J_\lambda(E_\lambda^\varepsilon(z)),$$

so that (6) reads

$$x_k = T_{\lambda_k}^{\varepsilon_k}(x_{k-1}).$$

Next, given $\Theta > 0$ and $\lambda, \mu \in (0, \Theta]$, set

$$\alpha = \frac{\lambda(\Theta - \mu)}{\Theta(\lambda + \mu) - \lambda\mu}, \quad \beta = \frac{\mu(\Theta - \lambda)}{\Theta(\lambda + \mu) - \lambda\mu}, \quad \gamma = \frac{\lambda\mu}{\Theta(\lambda + \mu) - \lambda\mu}. \quad (20)$$

Lemma 5.1. *Write $\Theta = \frac{\theta}{\kappa}$. For $\lambda, \mu \in (0, \Theta]$ and $x, y, \varepsilon, \eta \in X$, we have*

$$\|T_\lambda^\varepsilon(x) - T_\mu^\eta(y)\| \leq \alpha \|T_\lambda^\varepsilon(x) - y\| + \beta \|x - T_\mu^\eta(y)\| + \gamma \|x - y\| + \gamma \Theta \|\varepsilon - \eta\|. \quad (21)$$

Proof. Set $\Delta = j(T_\lambda^\varepsilon(x) - T_\mu^\eta(y))$. We have

$$\begin{aligned}
\Theta(\lambda + \mu)\|T_\lambda^\varepsilon(x) - T_\mu^\eta(y)\|^2 &= \Theta(\lambda + \mu)\langle T_\lambda^\varepsilon(x) - T_\mu^\eta(y), \Delta \rangle \\
&= \Theta\lambda\langle T_\lambda^\varepsilon(x) - E_\mu^\eta(y), \Delta \rangle + \Theta\mu\langle E_\lambda^\varepsilon(x) - T_\mu^\eta(y), \Delta \rangle \\
&\quad - \Theta\lambda\mu\left\langle \frac{E_\lambda^\varepsilon(x) - T_\lambda^\varepsilon(x)}{\lambda} - \frac{E_\mu^\eta(y) - T_\mu^\eta(y)}{\mu}, \Delta \right\rangle \\
&\leq \Theta\lambda\langle T_\lambda^\varepsilon(x) - E_\mu^\eta(y), \Delta \rangle \\
&\quad + \Theta\mu\langle E_\lambda^\varepsilon(x) - T_\mu^\eta(y), \Delta \rangle,
\end{aligned} \tag{22}$$

since A is accretive and

$$\frac{E_\nu^\varepsilon(z) - T_\nu^\varepsilon(z)}{\nu} \in A(T_\nu^\varepsilon(z))$$

for all $\nu > 0$ and $z, \varepsilon \in X$. Using the definition of E_λ^ε and E_μ^η , we can rewrite (22) as

$$\begin{aligned}
\Theta(\lambda + \mu)\|T_\lambda^\varepsilon(x) - T_\mu^\eta(y)\|^2 &\leq \Theta\lambda\langle T_\lambda^\varepsilon(x) - y, \Delta \rangle + \Theta\mu\langle x - T_\mu^\eta(y), \Delta \rangle \\
&\quad - \lambda\Theta\mu\langle Bx - \varepsilon - By + \eta, \Delta \rangle.
\end{aligned} \tag{23}$$

Notice also that

$$-\lambda\mu\|T_\lambda^\varepsilon(x) - T_\mu^\eta(y)\|^2 = -\lambda\mu\langle T_\lambda^\varepsilon(x) - y, \Delta \rangle - \lambda\mu\langle x - T_\mu^\eta(y), \Delta \rangle + \lambda\mu\langle x - y, \Delta \rangle. \tag{24}$$

Combining (23) and (24), we obtain

$$\begin{aligned}
[\Theta(\lambda + \mu) - \lambda\mu]\|T_\lambda^\varepsilon(x) - T_\mu^\eta(y)\|^2 &\leq \lambda(\Theta - \mu)\langle T_\lambda^\varepsilon(x) - y, \Delta \rangle \\
&\quad + \mu(\Theta - \lambda)\langle x - T_\mu^\eta(y), \Delta \rangle \\
&\quad + \lambda\mu\langle E_\Theta(x) - E_\Theta(y), \Delta \rangle + \lambda\mu\Theta\langle \varepsilon - \eta, \Delta \rangle.
\end{aligned}$$

Since E_Θ is nonexpansive and $\|\Delta\| = \|T_\lambda^\varepsilon(x) - T_\mu^\eta(y)\|$, we finally get (21). \square

Proof of Theorem 2.1. To simplify notation set

$$c_{k,l} = \sqrt{(\sigma_k - \hat{\sigma}_l)^2 + \tau_k + \hat{\tau}_l}.$$

In view of the characterization (6) of the sequence (x_k) , for each $k \geq 1$, we have

$$y_k := \frac{E_{\lambda_k}(x_{k-1}) + \lambda_k \varepsilon_k - x_k}{\lambda_k} \in Ax_k.$$

Given any $v \in Au$, the accretivity of A implies

$$\begin{aligned}
\|x_k - u\| &\leq \|x_k + \lambda_k y_k - u - \lambda v\| \\
&= \|E_{\lambda_k}(x_{k-1}) - E_{\lambda_k}(u) - \lambda_k(v + Bu) + \lambda_k \varepsilon_k\| \\
&\leq \|E_{\lambda_k}(x_{k-1}) - E_{\lambda_k}(u)\| + \lambda_k\|(v + Bu)\| + \lambda_k\|\varepsilon_k\|.
\end{aligned}$$

Since E_{λ_k} is nonexpansive and $v \in Au$ is arbitrary, we deduce that

$$\|x_k - u\| \leq \|x_{k-1} - u\| + \lambda_k\|(A + B)u\| + \lambda_k\|\varepsilon_k\|.$$

Iterating this inequality, we obtain

$$\|x_k - u\| \leq \|x_0 - u\| + \sigma_k\|(A + B)u\| + e_k,$$

and, noticing that $\sigma_k \leq c_{k,0}$, we conclude that

$$\|x_k - \hat{x}_0\| \leq \|x_0 - u\| + \|\hat{x}_0 - u\| + c_{k,0}\|(A + B)u\| + e_k,$$

thus inequality (8) holds for the pair $(k, 0)$. For $(0, l)$, with $l \geq 0$, the argument is analogous.

The proof will continue using induction on the pair (k, l) . Let us assume inequality (8) holds for the pairs $(k-1, l-1)$, $(k, l-1)$ and $(k-1, l)$, and show that it also holds for the pair (k, l) . To this end, we use the inequality (21) with $x = x_{k-1}$, $y = \hat{x}_{l-1}$, $\lambda = \lambda_k$ and $\mu = \hat{\lambda}_l$:

$$\begin{aligned} \|x_k - \hat{x}_l\| &\leq \alpha_{k,l} \|x_k - \hat{x}_{l-1}\| + \beta_{k,l} \|x_{k-1} - \hat{x}_l\| + \gamma_{k,l} \|x_{k-1} - \hat{x}_{l-1}\| \\ &\quad + \gamma_{k,l} \Theta \|\varepsilon_k - \hat{\varepsilon}_l\|. \end{aligned} \quad (25)$$

Using the induction hypothesis in (25) and the fact that $\alpha_{k,l} + \beta_{k,l} + \gamma_{k,l} = 1$, we deduce that

$$\begin{aligned} \|x_k - \hat{x}_l\| &\leq \|x_0 - u\| + \|\hat{x}_0 - u\| \\ &\quad + \|(A+B)u\| (\alpha_{k,l} c_{k,l-1} + \beta_{k,l} c_{k-1,l} + \gamma_{k,l} c_{k-1,l-1}) \\ &\quad + \alpha_{k,l} (e_k + \hat{e}_{l-1}) + \beta_{k,l} (e_{k-1} + \hat{e}_l) + \gamma_{k,l} (e_{k-1} + \hat{e}_{l-1}) \\ &\quad + \gamma_{k,l} \Theta (\|\varepsilon_k\| + \|\hat{\varepsilon}_l\|) \\ &= \|x_0 - u\| + \|\hat{x}_0 - u\| \\ &\quad + \|(A+B)u\| (\alpha_{k,l} c_{k,l-1} + \beta_{k,l} c_{k-1,l} + \gamma_{k,l} c_{k-1,l-1}) \\ &\quad + e_{k-1} + \hat{e}_{l-1} + (\alpha_{k,l} \lambda_k + \gamma_{k,l} \Theta) \|\varepsilon_k\| + (\beta_{k,l} \hat{\lambda}_l + \gamma_{k,l} \Theta) \|\hat{\varepsilon}_l\| \\ &= \|x_0 - u\| + \|\hat{x}_0 - u\| \\ &\quad + \|(A+B)u\| (\alpha_{k,l} c_{k,l-1} + \beta_{k,l} c_{k-1,l} + \gamma_{k,l} c_{k-1,l-1}) \\ &\quad + e_k + \hat{e}_l, \end{aligned} \quad (26)$$

since $\alpha_{k,l} \lambda_k + \gamma_{k,l} \Theta = \lambda_k$ and $\beta_{k,l} \hat{\lambda}_l + \gamma_{k,l} \Theta = \hat{\lambda}_l$. On the other hand, we have

$$\alpha_{k,l} c_{k,l-1} + \beta_{k,l} c_{k-1,l} + \gamma_{k,l} c_{k-1,l-1} \leq \sqrt{\alpha_{k,l} c_{k,l-1}^2 + \beta_{k,l} c_{k-1,l}^2 + \gamma_{k,l} c_{k-1,l-1}^2}, \quad (27)$$

in view of the Cauchy-Schwarz inequality. A simple computation yields

$$\begin{aligned} c_{k,l-1}^2 &= c_{k,l}^2 + 2\hat{\lambda}_l(\sigma_k - \hat{\sigma}_l) \\ c_{k-1,l}^2 &= c_{k,l}^2 + 2\lambda_k(\sigma_k - \hat{\sigma}_l) \\ c_{k-1,l-1}^2 &= c_{k,l}^2 + 2(\hat{\lambda}_l - \lambda_k)(\sigma_k - \hat{\sigma}_l) - 2\lambda_k \hat{\lambda}_l. \end{aligned}$$

Therefore,

$$\alpha_{k,l} c_{k,l-1}^2 + \beta_{k,l} c_{k-1,l}^2 + \gamma_{k,l} c_{k-1,l-1}^2 = c_{k,l}^2 - 2\gamma_{k,l} \lambda_k \hat{\lambda}_l \leq c_{k,l}^2. \quad (28)$$

Combining (26), (27) and (28), we obtain (8). \square

Concluding remarks. We have provided an approximating scheme for nonlinear semigroups generated by a sum of a m -accretive operator and a cocoercive one. In a finite horizon, we built a family of approximating curves generated by means of forward-backward iterations, and we proved the uniform convergence to the solutions of the differential inclusion, quantifying precisely the distance between the generated curves and their limit. In an infinite horizon, we showed that the semigroup and the forward-backward algorithm have the same behavior as time goes to infinity. As a consequence, we derived new convergence results for the forward-backward iterations in Banach spaces.

Cocoercive operators are m -accretive and Lipschitz-continuous, but the converse is not true in general. Skew-symmetric matrices are a canonical counterexample. Forward-backward iterations are not suitable to approximate nonlinear semigroups

generated by the sum of two m -accretive operators, one of which is Lipschitz-continuous but not cocoercive, since the *forward* step may stray away from the continuous trajectory in view of the orthogonality that characterizes skew-symmetry². We conjecture that Tseng's algorithm [25] provides a sufficiently close approximation of the continuous trajectory that may yield asymptotic equivalence results. Nevertheless, this is out of the scope of this paper and a subject for future research.

REFERENCES

- [1] Álvarez, F.; Peypouquet, J. Asymptotic equivalence and Kobayashi-type estimates for nonautonomous monotone operators in Banach spaces. *Discrete Contin. Dyn. Syst.* 25 (2009), no. 4, 1109-1128.
- [2] Álvarez F.; Peypouquet, J. Asymptotic almost-equivalence of Lipschitz evolution systems in Banach spaces. *Nonlinear Anal.* 73 (2010), no. 9, 3018-3033
- [3] Álvarez F.; Peypouquet, J. A unified approach to the asymptotic almost-equivalence of evolution systems without Lipschitz conditions. *Nonlinear Anal.* 74 (2011), no. 11, 3440-3444
- [4] Baillon J.B. Un exemple concernant le comportement asymptotique de la solution du problème $du/dt + \partial\varphi(u) \ni 0$. *J. Funct. Anal.* 28 (1978), 369-376.
- [5] Bénéilan, P. Équations d'Évolution dans un Espace de Banach Quelconque et Applications, Thèse, Orsay (1972).
- [6] Brézis, H. Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert, North Holland Publishing Company, Amsterdam, 1973.
- [7] Crandall, M.G., Pazy, A. Semi-groups of nonlinear contractions and dissipative sets, *J. Funct. Anal.*, 3 (1969), 376-418.
- [8] Contreras, A.; Peypouquet, J. Asymptotic equivalence of evolution equation governed by cocoercive operators and their forward discretizations. *J. Optim. Theory Appl.* 182 (2019), no. 1, 30-48.
- [9] Güler, O. On the convergence of the proximal point algorithm for convex optimization, *SIAM J. Control Opt.*, 29 (1991), 403-419.
- [10] E. Hille, On the generation of semi-groups and the theory of conjugate functions. *Kungl. Fysiografiska Sällskapets i Lund Föreläsningar* [Proc. Roy. Physiol. Soc. Lund] 21, (1952). no. 14, 13 pp.
- [11] Kato, T. Nonlinear semigroups and evolution equations, *J. Math. Soc. Japan*, 19 (1973), 508-520.
- [12] Kato, T. On the Trotter-Lie product formula, *Proc. Japan Acad.* 50 (1974), 694-698.
- [13] Kobayashi, Y. Difference approximation of Cauchy problems for quasi-dissipative operators and generation of nonlinear semigroups, *J. Math. Soc. Japan*, 27 (1975), 640-665.
- [14] Kobayashi, K.; Kobayashi, Y.; Oharu, S. Nonlinear evolution operators in Banach spaces, *Osaka J. Math.*, 21 (1984), 281-310.
- [15] Lumer, G.; Phillips, R.S. Dissipative operators in a Banach space, *Pacific J. Math.* 11 (1961), 679-698.
- [16] Martinet, B. Régularisation d'inéquations variationnelles par approximations successives. *Rev. fr. autom. inform. rech. opér.*, 4 (1970), 154-158.
- [17] Miyadera, I.; Kobayashi, K. On the asymptotic behavior of almost-orbits of nonlinear contractions in Banach spaces, *Nonlinear Anal.*, 6 (1982), 349-365.
- [18] Nevanlinna, O.; Reich, S. Strong convergence of contraction semigroups and of iterative methods for accretive operators in Banach spaces, *Israel J. Math.* 32 (1979), 44-58.
- [19] Passty, G.B. Preservation of the asymptotic behavior of a nonlinear contraction semigroup by backward differencing, *Houston J. Math.*, 7 (1981), 103-110.
- [20] Peypouquet, J. *Convex optimization in normed spaces: theory, methods and examples*. Springer, 2015.
- [21] Peypouquet, J.; Sorin, S. Evolution equations for maximal monotone operators: asymptotic analysis in continuous and discrete time. *J. Convex Anal.*, 17 (2010), 1113-1163.

²For an example in \mathbb{R}^2 , take $B(x, y) = (-y, x)$. Each solution to $-(\dot{x}(t), \dot{y}(t)) = B(x(t), y(t))$ draws a circle around the origin. In turn, sequences generated by $(x_{n+1}, y_{n+1}) = (x_n, y_n) - \lambda B(x_n, y_n) = (x_n - \lambda y_n, y_n + \lambda x_n)$ satisfy $\lim_{n \rightarrow \infty} \|(x_n, y_n)\| = \infty$, no matter how small $\lambda > 0$ is, unless $(x_0, y_0) = (0, 0)$.

- [22] Rasmussen, S. Nonlinear semigroups, evolution equations and product integral representations. Various Publication Series, Vol 20, Aarhus Universitet (1971/72).
- [23] Rockafellar, R.T. Monotone operators and the proximal point algorithm, SIAM J. Control Optim., 14 (1976), 877-898.
- [24] Trotter, H. F. On the product of semi-groups of operators, Proc. Amer. Math. Soc. 10 (1959), 545–551
- [25] Tseng, P. A modified forward-backward splitting method for maximal monotone mappings, SIAM J. Control Optim., 38 (2000), 431-446.
- [26] Sugimoto, T.; Koizumi, M. On the asymptotic behaviour of a nonlinear contraction semigroup and the resolvent iteration. Proc. Japan Acad. Ser. A. Math. Sci., 59 (1983), no. 6, 238-240.
- [27] Vigerel, G. Evolution equations in discrete and continuous time for nonexpansive operators in Banach spaces, ESAIM, Control Optim. Calc. Var., 16 (2010), 809-832.
- [28] Xu, H.K. Inequalities in Banach spaces with applications, Nonlinear Anal., 16 (1991), 1127-1138.
- [29] Yosida, K. On the differentiability and the representation of one-parameter semi-group of linear operators. J. Math. Soc. Japan, 1 (1948), 15-21.

Received xxxx 20xx; revised xxxx 20xx.

E-mail address: andres-antonio.contreras-tavarez@inria.fr, acontreras@dim.uchile.cl

E-mail address: j.g.peypouquet@rug.nl, jpeypou@dim.uchile.cl