FORWARD-BACKWARD APPROXIMATION OF EVOLUTION EQUATIONS IN FINITE AND INFINITE HORIZON

ANDRÉS CONTRERAS AND JUAN PEYPOUQUET

ABSTRACT. This research is concerned with evolution equations and their forward-backward discretizations. 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 variational analysis and optimization. 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 defined using 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 [22, 49], which states that an *unbounded* linear operator A, on a Banach space X, is the infinitessimal 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, it is closed, its domain is dense in X, its spectrum does not intersect $\mathbb{R}_-$, and the resolvents satisfy an appropriate bound. This result was complemented by the Lumer-Phillips Theorem [41, 31], 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 consists in approximating the operator 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, discretizes the time interval $[0, T]$, where $T > 0$ is arbitrary but fixed, constructs approximating trajectories using a sequence of points generated by resolvent iterations, and finally passes to the limit as the partition is refined. Both show the convergence is uniform on $[0, T]$.

2010 *Mathematics Subject Classification.* 34A60, 37L05, 49M25.

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.

Another important landmark was the discovery, two decades later, of sufficient conditions for a nonlinear, 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 $-\dot{u}(t) \in Au(t)$. Yosida's approach was used by Brézis [9] (also Barbu [5] and Pazy [38]), while Hille's path was followed by Crandall and Pazy [16][1], and then simplified and perfected by Rasmussen [42] and Kobayashi [25]. They built a concise and sharp inequality −let us call it (I)− to bound the distance between two sequences of points generated using compositions of resolvents. We shall come back to this point later, since this is the line of research we explore in this paper. Other authors have analyzed the nonautonomous setting [26, 1, 8], 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 [27, 32] and, equivalently, Euler [40] iterations. This issue is addressed in [46, 18], where applications in optimization and game theory are given.

A few years later, Passty [37] 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, as time goes to $+\infty$, based on what is known about those generated by the semigroup it is related to. A similar idea lies behind the notion of *almost-orbit* (see [34]), which helps to prove that every nonexpansive iterative algorithm is robust against summable errors (see [39, 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 generated using products of resolvents of A, with parameters $(\lambda_n)_{n\geq 0}$, more precisely, satisfying $x_n = (I + \lambda_n A)^{-1} x_{n-1}$ for all n, converges strongly (weakly) as $n \to +\infty$ if, and only if, all trajectories generated by the semigroup $(\mathcal{S}_t)_{t\geq 0}$ converge strongly (weakly) as $t \to +\infty$. The process of generating sequences of points using resolvent iterations is also known as the *proximal point algorithm*, as developed by Martinet [33] and further studied by Rockafellar [43] and Brézis-Lions [10], among others. 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 in the next paragraph). 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 [34] and Sugimoto and Koizumi [45] were able to get rid of Passty's superfluous hypotheses by using inequality (I) mentioned above. Inequality (I) also enabled Güler [21] 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 [43] 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 [18], using a variant of inequality (I).

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. Particular cases include: the gradient method, originally introduced by Cauchy in [11]; its variant, the projected gradient method [20, 28]; the proximal point algorithm mentioned above; the proximal-gradient algorithm [37, 30], and its particular instance, ISTA[2] [19, 13], with applications in image and signal processing, data analysis and machine learning. Moreover, some primal dual methods [12, 14, 47]

[1]Although Crandall and Liggett [15] used Yosida's method in their work on Banach spaces.

[2]Iterative Shrinkage Thresholding Algorithm.

can be reduced to these types of iterations. Also, accelerated methods, such as FISTA[3] [35, 6] use a forward-backward engine.

The purpose of this research is to extend, unify and condense the theory on 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 [44] and Kato [24], which uses *double backward* iterations. Double backward iterations require the (costly!) computation of both resolvents. We address this issue, for theoretical curiosity, in a forthcoming paper. 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.

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 main 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* $j: X \to X^*$ 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

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

for all $u, v \in X$ (see [29, 48]). For instance, L^p spaces have this property for $p \geq 2$.

A set-valued operator $A : X \to 2^X$ 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 [23, Lemma 1.1] that A is accretive if, and only if, it is *monotone*, which means that

$$\langle j(x-y), u-v\rangle \geq 0,$$ [4]

[3] Fast Iterative Shrinkage Thresholding Algorithm.

[4] In Hilbert spaces, this terminology is preferred, and the inequality reads $(x-y, u-v) \geq 0$, where $(\cdot,\cdot)$ is the inner product.

whenever $u \in Ax$ and $v \in Ay$. Next, an operator $B : X \to X$ 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 $E_\lambda : X \to X$, 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* $T_\lambda : X \to X$, 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 2.1. Actually, the minimal hypotheses 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

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

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

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

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

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

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

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

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 obtained 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 $t \to \infty$, as the sequences satisfying (4), when $k \to \infty$, 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 2.2. *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*

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

where $\|\!|Au|\!\| = \inf\limits_{v \in Au} \|v\|$, $\sigma_k = \sum\limits_{i=1}^{k} \lambda_i$, $\tau_k = \sum\limits_{i=1}^{k} \lambda_i^2$ *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 [21], where Güler attributes it to Kobayashi [25] (see also [40]). However, the main arguments were given by Rasmussen [42], who simplified the proof of Crandall and Liggett [15], ultimately based on that of Hille [22]. Similar estimations are given in [26, 1] (still for $B = 0$, but for a time-dependent A) and in [46, 18] for $A = 0$.

3. Approximation in finite horizon

Theorem 2.2 provides existence and regularity results for the evolution equation

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

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

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

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

Proof. We may assume that $u_0 \in D(A)$. Extension to $\overline{D(A)}$ will then be possible in view of the Lipschitz (thus uniform) continuity. Given $t, s > 0$ and $n, m \in \mathbb{N}$, define $u_m(t)$ and $u_n(s)$ as above. By Theorem 2.2, we have

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

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 $u : [0, \infty) \to X$. Passing to the limit in (11), as $m, n \to \infty$, we obtain

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

for all $t, s > 0$. $\square$

Remark 3.2. Given $S > 0$ and $m \geq 1$, define $v_m : [0, S] \to X$ by

$$v_m(t) = \left[T_{\frac{S}{m}}\right]^{\mu(t)} u_0, \quad \text{where} \quad \mu(t) = \left\lfloor m\frac{t}{S} \right\rfloor \quad \text{and} \quad t \in [0, S]. \tag{12}$$

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

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

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

Theorem 3.3. *The function u, given by Proposition 3.1, satisfies* (9)*.*

Proof. We shall verify that u is an integral solution of (9) in the sense of Bénilan (see [7]), which means that, whenever $y \in (A + B)x$ and $S \geq t > s \geq 0$, we have

$$\|u(t) - x\|^2 - \|u(s) - x\|^2 \leq 2 \int_s^t \langle j(x - u(\tau)), y\rangle d\tau. \tag{13}$$

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,

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

Now, let us choose $x_0 = u_0$, $\lambda_n \equiv \frac{S}{m}$, where m is fixed but arbitrary. In view of Remark 3.2, there is a constant $K > 0$ such that $2\theta^{-1}\|x - x_n\| \leq K$ for $n = 1, \dots, m$. Summing for $n = \mu(s), \cdots, \mu(t)$, we obtain

$$\begin{array}{rcl}
\|v_m(t) - x\|^2 - \|u(s) - x\|^2 & \leq & 2 \displaystyle\sum_{n=\mu(s)}^{\mu(t)} \frac{S}{m}\left[\langle j(x - x_n), y\rangle + K\|x_n - x_{n-1}\|\right] \\
& \leq & 2 \displaystyle\sum_{n=\mu(s)}^{\mu(t)} \frac{S}{m}\langle j(x - x_n), y\rangle + \sum_{n=\mu(s)}^{\mu(t)} \frac{6S^2K\|\!|(A + B)u_0\|\!|}{m\sqrt{m}} \\
& = & 2 \displaystyle\sum_{n=\mu(s)}^{\mu(t)} \frac{S}{m}\langle j(x - x_n), y\rangle + (\mu(t) - \mu(s))\frac{6S^2K\|\!|(A + B)u_0\|\!|}{m\sqrt{m}} \\
& \leq & 2 \displaystyle\sum_{n=\mu(s)}^{\mu(t)} \frac{S}{m}\langle j(x - x_n), y\rangle + \frac{6S^2K\|\!|(A + B)u_0\|\!|}{\sqrt{m}}.
\end{array}$$

We obtain (13) by letting $m \to \infty$. $\square$

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

Corollary 3.4. *The differential inclusion* (9) *has a unique solution.*

Uniqueness follows from monotonicity. Another consequence of the results above is:

Corollary 3.5. *Let* (x_k) *be a sequence generated by* (3) *and let* $u : [0, S] \to X$ *be a solution of* (9). *Then*

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

4. Approximation in infinite horizon

In this section, we show that the forward-backward sequence generated by (3), have the same asymptotic behavior, as the number of iterations goes to infinity, as the solutions of the evolution equation (9), when time does. The key argument is the idea of asymptotic equality introduced by Passty [37], closely related to the notion of almost-orbit, introduced by Miyadera and Kobayasi [34]. Further commentaries on this topic can be found in [1, 2, 3].

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

$$\mathcal{S}_t x = u(t), \tag{14}$$

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

$$U_{\mathcal{S}}(t, s) = \mathcal{S}(t - s). \tag{15}$$

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

$$\mathcal{T}_n x = T_{\lambda_n} \circ \cdots \circ T_{\lambda_1} x. \tag{16}$$

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

$$U_{\mathcal{T}}(t, s) = \prod_{i=\nu(s)+1}^{\nu(t)} T_{\lambda_i}, \tag{17}$$

where the product denotes composition of functions and the empty composition is the identity.

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

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

Example 4.1. The families $\big(U_{\mathcal{S}}\big)$ and $\big(U_{\mathcal{T}}\big)$, defined in (15) and (17), 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 $\phi : [0, \infty) \to X$ is an *almost-orbit* of the nonexpansive evolution system U if

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

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

Proposition 4.2. *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 $t \to \infty$, 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, 18].

The following result establishes a relationship between the trajectories generated by $U_{\mathcal{S}}$ and $U_{\mathcal{T}}$:

Theorem 4.3. *Let $(\lambda_n) \in \ell^2 \setminus \ell^1$, and fix $x \in X$. For each $t > 0$, define $\phi_{\mathcal{S}}(t) = \mathcal{S}_t x$ and $\phi_{\mathcal{T}}(t) = \mathcal{T}_{\nu(t)} x$[5]. Then, $\phi_{\mathcal{S}}$ is an almost-orbit of $U_{\mathcal{T}}$, and $\phi_{\mathcal{T}}$ is an almost-orbit of $U_{\mathcal{S}}$.*

Proof. We first prove that $\phi_{\mathcal{S}}$ is an almost-orbit of $U_{\mathcal{T}}$. By Theorem 2.2 and Corollary 3.5, we have

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

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 $t \to \infty$. Passing to the limit as $m \to \infty$, we obtain

$$\| \mathcal{S}_h \mathcal{S}_t x - U_{\mathcal{T}}(t+h,t) \mathcal{S}_t x \| \leq \| (A+B)x \| \sqrt{4\rho^2(t) + \tau_{\nu(t)+1}^{\infty}},$$

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

$$\lim_{t \to \infty} \sup_{h \geq 0} \| \phi_{\mathcal{S}}(t+h) - U_{\mathcal{T}}(t+h,t) \phi_{\mathcal{S}}(t) \| = 0.$$

To prove that $\phi_{\mathcal{T}}$ is an almost-orbit of $U_{\mathcal{S}}$, we proceed in a similar fashion, to obtain

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

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

$$\| \phi_{\mathcal{T}}(t+h) - \mathcal{S}_h \phi_{\mathcal{T}}(t) \| \leq \| (A+B)x \| \sqrt{4\rho^2(t) + \tau_{\nu(t)+1}^{\infty}},$$

and conclude. □

Theorem 4.3 implies [37, Lemmas 4 & 6], [45, Proposition 2.3], [34, Proposition 7.4], [40, Propositions 8.6 i) & 8.7] and [18, Theorem 3.1]. Combining Theorem 4.3 with Proposition 4.2, and using [39, Lemma 5.3], we obtain

Theorem 4.4. *The following statements are equivalent:*

i) *For every $z \in \overline{D(A)}$, $\mathcal{S}_t z$ converges strongly (weakly), as $t \to +\infty$.*
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 $n \to +\infty$.*
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 $n \to +\infty$.*

Theorem 4.4 implies [37, Theorems 1 & 2], [45, Theorem], [34, Theorem 7.5], as well as [18, Theorem 3.2].

[5] This is a piecewise constant interpolation of the sequence $\mathcal{T}_n x$.

New convergence results for forward backward sequences on Banach spaces. Theorem 4.4 can automatically give new convergence results for forward-backward sequences by translating the information available on the behavior of the semigroup. Theorem 4.5 below is provided as a *methodological* example, to show how this indirect analysis can be carried out. Therefore, we have priviledged 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, supose 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 4.5. *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 $n\to+\infty$, to a point in Σ.

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

i) The hypotheses of [36, Theorem 1] are easily verified.
ii) It suffices to combine [36, Proposition 1] and [36, Theorem 1].
iii) We use [36, Theorem 4].

We conclude by applying Theorem 4.4. □

5. Proof of the fundamental inequality

This last section is devoted to the proof of Theorem 2.2. 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

$$\alpha=\frac{\lambda(\Theta-\mu)}{\Theta(\lambda+\mu)-\lambda\mu},\ \beta=\frac{\mu(\Theta-\lambda)}{\Theta(\lambda+\mu)-\lambda\mu},\ \gamma=\frac{\lambda\mu}{\Theta(\lambda+\mu)-\lambda\mu}. \tag{18}$$

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

$$\|T_\lambda^\varepsilon(x)-T_\mu^\eta(y)\|\leq\alpha\|T_\lambda^\varepsilon(x)-y\|+\beta\|x-T_\mu^\eta(y)\|+\gamma\|x-y\|+\gamma\Theta\|\varepsilon-\eta\|. \tag{19}$$

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

$$\begin{aligned}\Theta(\lambda+\mu)\|T_\lambda^\varepsilon(x)-T_\mu^\eta(y)\|^2 &= \Theta(\lambda+\mu)\langle T_\lambda^\varepsilon(x)-T_\mu^\eta(y),\Delta\rangle\\ &= \Theta\lambda\langle T_\lambda^\varepsilon(x)-E_\mu^\eta(y),\Delta\rangle+\Theta\mu\langle E_\lambda^\varepsilon(x)-T_\mu^\eta(y),\Delta\rangle\\ &\quad +\Theta\lambda\mu\left\langle\frac{E_\mu^\eta(y)-T_\mu^\eta(y)}{\mu}-\frac{E_\lambda^\varepsilon(x)-T_\lambda^\varepsilon(x)}{\lambda},\Delta\right\rangle\\ &\leq \Theta\lambda\langle T_\lambda^\varepsilon(x)-E_\mu^\eta(y),\Delta\rangle+\Theta\mu\langle E_\lambda^\varepsilon(x)-T_\mu^\eta(y),\Delta\rangle,\end{aligned} \tag{20}$$

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 $z, \varepsilon \in X$. We can rewrite (20) as

$$\Theta(\lambda+\mu)\|T_\lambda^\varepsilon(x)-T_\mu^\eta(y)\|^2 \leq \Theta\lambda\langle T_\lambda^\varepsilon(x)-y,\Delta\rangle+\Theta\mu\langle x-T_\mu^\eta(y),\Delta\rangle-\lambda\Theta\mu\langle Bx-\varepsilon-By+\eta,\Delta\rangle. \tag{21}$$

Notice also that

$$-\lambda\mu\|T_\lambda^\varepsilon(x) - T_\mu^\eta(y)\|^2 = -\lambda\mu\langle T_\lambda^\varepsilon(x) - y, \Delta\rangle - \lambda\mu\langle x - T_\mu^\eta(y), \Delta\rangle + \lambda\mu\langle x - y, \Delta\rangle. \tag{22}$$

Combining (21) and (22), we obtain

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

Since E_Θ is nonexpansive and $\|\Delta\| = \|T_\lambda^\varepsilon(x) - T_\mu^\eta(y)\|$, we finally get (19). □

We are now in a position to conclude.

Proposition 5.2. *Theorem 2.2 is true.*

Proof. 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

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

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

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

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 (19) with $x = x_{k-1}$, $y = \hat{x}_{l-1}$, $\lambda = \lambda_k$ and $\mu = \hat{\lambda}_l$:

$$\|x_k - \hat{x}_l\| \leq \alpha_{k,l}\|x_k - \hat{x}_{l-1}\| + \beta_{k,l}\|x_{k-1} - \hat{x}_l\| + \gamma_{k,l}\|x_{k-1} - \hat{x}_{l-1}\| + \gamma_{k,l}\Theta\|\varepsilon_k - \hat{\varepsilon}_l\|. \tag{23}$$

Using the induction hypothesis in (23) and the fact that $\alpha_{k,l}+\beta_{k,l}+\gamma_{k,l}=1$, we deduce that

$$\begin{aligned}
\|x_k-\hat{x}_l\| &\leq \|x_0-u\|+\|\hat{x}_0-u\|+|\!|\!|(A+B)u|\!|\!|\,(\alpha_{k,l}c_{k,l-1}+\beta_{k,l}c_{k-1,l}+\gamma_{k,l}c_{k-1,l-1})\\
&\quad+\alpha_{k,l}(e_k+\hat{e}_{l-1})+\beta_{k,l}(e_{k-1}+\hat{e}_l)+\gamma_{k,l}(e_{k-1}+\hat{e}_{l-1})+\gamma_{k,l}\Theta(\|\varepsilon_k\|+\|\hat{\varepsilon}_l\|)\\
&= \|x_0-u\|+\|\hat{x}_0-u\|+|\!|\!|(A+B)u|\!|\!|\,(\alpha_{k,l}c_{k,l-1}+\beta_{k,l}c_{k-1,l}+\gamma_{k,l}c_{k-1,l-1})\\
&\quad+e_{k-1}+\hat{e}_{l-1}+(\alpha_{k,l}\lambda_k+\gamma_{k,l}\Theta)\|\varepsilon_k\|+(\beta_{k,l}\hat{\lambda}_l+\gamma_{k,l}\Theta)\|\hat{\varepsilon}_l\|\\
&= \|x_0-u\|+\|\hat{x}_0-u\|+|\!|\!|(A+B)u|\!|\!|\,(\alpha_{k,l}c_{k,l-1}+\beta_{k,l}c_{k-1,l}+\gamma_{k,l}c_{k-1,l-1})+e_k+\hat{e}_l,
\end{aligned} \tag{24}$$

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

$$\begin{aligned}
\alpha_{k,l}c_{k,l-1}+\beta_{k,l}c_{k-1,l}+\gamma_{k,l}c_{k-1,l-1} &\leq \sqrt{\alpha_{k,l}+\beta_{k,l}+\gamma_{k,l}}\sqrt{\alpha_{k,l}c_{k,l-1}^2+\beta_{k,l}c_{k-1,l}^2+\gamma_{k,l}c_{k-1,l-1}^2}\\
&= \sqrt{\alpha_{k,l}c_{k,l-1}^2+\beta_{k,l}c_{k-1,l}^2+\gamma_{k,l}c_{k-1,l-1}^2},
\end{aligned} \tag{25}$$

and

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

Therefore,

$$\alpha_{k,l}c_{k,l-1}^2+\beta_{k,l}c_{k-1,l}^2+\gamma_{k,l}c_{k-1,l-1}^2=c_{k,l}^2-2\gamma_{k,l}\lambda_k\hat{\lambda}_l\leq c_{k,l}^2. \tag{26}$$

Combining (24), (25) and (26), we obtain (8). □

References

[1] Álvarez, F.; Peypouquet, J. Asymptotic equivalence and Kobayashi-type estimates for nonautonomous monotone operators in Banach spaces. Discrete and Continuous Dynamical Systems 25 (2009), no. 4, 1109-1128.

[2] Álvarez F.; Peypouquet, J. Asymptotic almost-equivalence of Lipschitz evolution systems in Banach spaces. Nonlinear Analysis: Theory, Methods & Applications 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 Analysis: Theory, Methods & Applications 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] Barbu V, Nonlinear semigroups and differential equations in Banach spaces. Noordhoff, Leyden, 1976.

[6] Beck, A.; Teboulle, M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imaging Sci. 2 (2009), no. 1, 183-202.

[7] Bénilan, P. Équations d'Évolution san un Espace de Banach Quelconque et Applications, Thése, Orsay (1972).

[8] Bianchi, P.; Hachem, W. Dynamical behavior of a stochastic forward-backward algorithm using random monotone operators. J. Optim. Theory Appl. 171 (2016), no. 1, 90120.

[9] Brézis, H. Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert, North Holland Publishing Company, Amsterdam, 1973.

[10] Brézis, H.; Lions, P.L. Produits infinis de résolvantes, Israel J. Math., 29 (1978), 329-345.

[11] Cauchy, A.L. Méthode générale pour la résolution des systemes déquations simultanées. Comp. Rend. Sci. Paris, 25 (1847), 536-538.

[12] Chambolle, A.; Pock, T. A first-order primal-dual algorithm for convex problems with applications to imaging. J. Math. Imaging Vision 40 (2011), no. 1, 120145.

[13] Combettes, P.L.; Wajs, V.R. Signal recovery by proximal forward-backward splitting. Multiscale Model. Simul. 4 (2005), no. 4, 11681200.

[14] Condat, L. A primal-dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms. J. Optim. Theory Appl. 158 (2013), 460-479.

[15] Crandall, M.G.; Liggett, T.M. Generation of semigroups of nonlinear transformations on general Banach spaces, Am. J. Math, 93 (1971), 265-298.

[16] Crandall, M.G., Pazy, A. Semi-groups of nonlinear contractions and dissipative sets, J Funct. Anal., 3 (1969), 376-418.

[17] Cartan, H. Differential calculus. London: Kershaw Publishing Co, (1971).

[18] Contreras, A.; Peypouquet, J. Asymptotic equivalence of evolution equation governed by cocoercive operators and their forward discretizations. Under review.

[19] Daubechies, I.; Defrise, M.; De Mol, C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. Comm. Pure Appl. Math. 57 (2004), no. 11, 14131457.

[20] Goldstein, A.A. Convex programming in Hilbert space. Bull. Amer. Math. Soc. 70 (1964), 709710.

[21] Güler, O. On the convergence of the proximal point algorithm for convex optimization, SIAM J. Control Opt., 29 (1991), 403-419.

[22] E. Hille, On the generation of semi-groups and the theory of conjugate functions. Kungl. Fysiografiska Sllskapets i Lund Frhandlingar [Proc. Roy. Physiog. Soc. Lund] 21, (1952). no. 14, 13 pp.

[23] Kato, T. Nonlinear semigroups and evolution equations, J. Math Soc. Japan, 19 (1973), 508-520.

[24] Kato, T. On the Trotter-Lie product formula, Proc. Japan Acad. 50 (1974), 694698.

[25] Kobayashi, Y. Difference aproximation of Cauchy problems for quasi-dissipative operators and generation of nonlinear semigroups, J. Math Soc. Japan, 27 (1975), 640-665.

[26] Kobayasi, K; Kobayashi, Y.; Oharu, S. Nonlinear evolution operators in Banach spaces, Osaka J. Math, 21 (1984), 281–310.

[27] Krasnosel'skiĭ, M. A. Two remarks on the method of successive approximations. (Russian) Uspehi Mat. Nauk (N.S.) 10, (1955). no. 1(63), 123127

[28] Levitin, E.S.; Poljak, B.T. Minimization methods in the presence of constraints. (Russian) Z. Vycisl. Mat i Mat. Fiz. 6 (1966), 787823.

[29] Lindenstrauss, J.; Tzafriri, L. Classical Banach spaces. II. Function spaces, Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and related areas], 97, Berlin-New York: Springer-Velarg, pp. x+243, 1079.

[30] Lions, P.L.; Mercier, B. Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. 16 (1979), no. 6, 964979.

[31] Lumer, G.; Phillips, R.S. Dissipative operators in a Banach space, Pacific J. Math. 11 (1961), 679698.

[32] Mann, W.R. Mean value methods in iteration. Proc. Amer. Math. Soc. 4 (1953), 506510.

[33] Martinet, B. Régularisation d'inéquations variationnelles par approximations successives. Rev. fr. autom. inform. rech. opér., 4 (1970), 154-158.

[34] Miyadera, I.; Kobayasi, K. On the asymptotic behavior of almost-orbits of nonlinear contractions in Banach spaces, Nonlinear Anal., 6 (1982), 349-365.

[35] Nesterov, Y. A method of solving a convex programming problem with convergence rate $\mathcal{O}(1/k^2)$. Dokl. Akad. Nauk SSSR, 27 (1983), 372-376.

[36] 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.

[37] Passty, G.B. Preservation of the asymptotic behavior of a nonlinear contraction semigroup by backward differencing, Houston J. Math., 7 (1981), 103-110.

[38] Pazy A, Semigroups of nonlinear contractions and their asymptotic behavior. Nonlinear Analysis and Mechanics: Heriot-Watt Symposium, vol III, 1979, Pitman, 36-134.

[39] Peypouquet, J. Convex optimization in normed spaces: theory, methods and examples. Springer, 2015.

[40] Peypouquet, J.; Sorin, S. Evolution equations for maximal monotone operators: asymptotic analysis in continuous and discrete time. J. Convex Anal., 17 (2010), 1113-1163.

[41] Phillips, R.S. Dissipative operators and hyperbolic systems of partial differential equations, Trans. Amer. Math. Soc. 90 (1959), 193254.

[42] Rasmussen, S. Nonlinear semigroups, evolution equations and product integral representations. Various Publication Series, Vol 20, Aarhus Universitet (1971/72).

[43] Rockafellar, R.T. Monotone operators and the proximal point algorithm, SIAM J. Control Optim., 14 (1976), 877-898.

[44] Trotter, H. F. On the product of semi-groups of operators, Proc. Amer. Math. Soc. 10 (1959), 545551

[45] Sugimoto, T.; Koizumi, M. On the asymptotic behaviour of a nonlinear contraction semigroup and the resolvente iteration. Proc. Japan Acad. Ser. A. Math. Sci., 59 (1983), no. 6, 238-240.

[46] Vigeral, G. Evolution equations in discrete and continuous time for nonexpansive opreators in Banach spaces, ESAIM, Control Optim. Calc. Var., 16 (2010), 809-832.
[47] Vu, B.C. A splitting algorithm for dual monotone inclusions involving cocoercive operators. Adv. Comput. Math., 38 (2013), no. 3, 667-681.
[48] Xu, H.K. Inequalities in Banach spaces with applications, Nonlinear Anal., 16 (1991), 1127-1138.
[49] Yosida, K. On the differentiability and the representation of one-parameter semi-group of linear operators. J. Math. Soc. Japan, 1 (1948), 15-21.

Departamento de Ingeniería Matemática & Centro de Modelamiento Matemático (CNRS UMI2807), FCFM, Universidad de Chile, Beauchef 851, Santiago, Chile
E-mail address: acontreras@dim.uchile.cl

Departamento de Ingeniería Matemática & Centro de Modelamiento Matemático (CNRS UMI2807), FCFM, Universidad de Chile, Beauchef 851, Santiago, Chile
E-mail address: jpeypou@dim.uchile.cl