Strong asymptotic convergence of evolution equations governed by maximal monotone operators with Tikhonov regularization

R. Cominetti [a,1], J. Peypouquet [b,*,2], S. Sorin [c,d,3]

[a] Departamento de Ingeniería Matemática and Centro de Modelamiento Matemático, Universidad de Chile, Blanco Encalada 2120, Santiago, Chile
[b] Universidad Técnica Federico Santa María, Av España 1680, Valparaíso, Chile
[c] Equipe Combinatoire et Optimisation, UFR 929, Université P. et M. Curie, Paris 6, 175 rue du Chevaleret, 75013 Paris, France
[d] Laboratoire d'Econométrie, Ecole Polytechnique, France

ABSTRACT

We consider the Tikhonov-like dynamics $-\dot{u}(t) \in A(u(t)) + \varepsilon(t)u(t)$ where A is a maximal monotone operator on a Hilbert space and the parameter function $\varepsilon(t)$ tends to 0 as $t \to \infty$ with $\int_0^\infty \varepsilon(t)\,dt = \infty$. When A is the subdifferential of a closed proper convex function f, we establish strong convergence of $u(t)$ towards the least-norm minimizer of f. In the general case we prove strong convergence towards the least-norm point in $A^{-1}(0)$ provided that the function $\varepsilon(t)$ has bounded variation, and provide a counterexample when this property fails.

Keywords:
Maximal monotone operators
Tikhonov regularization

1. Introduction

We investigate the asymptotic behavior as $t \to \infty$ of solutions of the differential inclusion

$$-\dot{u}(t) \in A\big(u(t)\big) + \varepsilon(t)u(t); \qquad u(0) = x_0, \tag{D}$$

* Corresponding author.
E-mail addresses: rcominet@dim.uchile.cl (R. Cominetti), juan.peypouquet@usm.cl (J. Peypouquet), sorin@math.jussieu.fr (S. Sorin).

[1] Supported by FONDAP grant in Applied Mathematics, CONICYT-Chile.
[2] Supported by MECESUP grant UCH0009 and FONDAP grant in Applied Mathematics, CONICYT-Chile.
[3] Supported by grant ANR-05-BLAN-0248-01.

where $A:\mathcal{H}\to 2^{\mathcal{H}}$ is a maximal monotone operator on a Hilbert space $\mathcal{H}$, $\varepsilon(t)\geqslant 0$ is measurable, and $x_0\in \mathrm{dom}(A)$. Throughout this paper we assume that (D) admits a (necessarily unique) *strong solution*, namely, an absolutely continuous function $u:[0,\infty)\to\mathcal{H}$ such that (D) holds for almost every $t\geqslant 0$. Sufficient conditions for this existence may be found, among others, in [4,19,20], and [25].

The differential inclusion (D) is a perturbed version of

$$-\dot{u}(t)\in A\big(u(t)\big); \qquad u(0)=x_0. \tag{I}$$

We denote by $S=\{x\in\mathcal{H}:\ 0\in A(x)\}$ the set of rest points of the latter, and we assume that it is nonempty. The monotonicity of A implies that the dynamics (I) are dissipative, so one might expect that they converge to a point in S. This is not always the case as seen by considering a $\frac{\pi}{2}$-rotation in $\mathbb{R}^2$. However, if we perturb these dynamics as in (D) with a fixed $\varepsilon(t)\equiv\varepsilon>0$, the operator $A+\varepsilon I$ is strongly monotone and we have strong convergence to the unique solution of $0\in A(x)+\varepsilon x$. Hence, by introducing a vanishing parameter $\varepsilon(t)\to 0^+$ and under suitable conditions, one may hope to induce weak or even strong convergence of the solutions of (D) towards a point in S.

Several results are available for different classes of maximal monotone operators. In the unperturbed case $\varepsilon(t)\equiv 0$, while convergence does not hold in general, weak convergence was established in the classical paper [14] for the case of demi-positive operators. This class includes the subdifferentials of closed proper convex functions $A=\partial f$, as well as operators of the form $A=I-T$ with T a contraction having fixed points. As shown by the counterexample in [5], even in the case of subdifferential operators one may not expect this convergence to be strong.

Asymptotic results have also been proved for a variety of dynamics coupling a gradient flow with different approximation schemes. In the particular setting of (D) the convergence depends on whether $\varepsilon(t)$ is in $L^1(0,\infty)$ or not. When $\int_0^\infty \varepsilon(t)\,dt<\infty$ the results on asymptotic equivalence described in [32] (see also [2]) imply that the perturbation (D) preserves the qualitative convergence properties of (I). For the case $\int_0^\infty \varepsilon(t)\,dt=\infty$ the most general convergence result available goes back to [33] (based on previous work by [12]) and requires in addition $\varepsilon(t)$ to be non-increasing and convergent to 0 for $t\to\infty$. Under these conditions $u(t)$ converges strongly to x^*, the point of least norm in S. The main contributions in this paper are in the case $\int_0^\infty \varepsilon(t)\,dt=\infty$ with $\varepsilon(t)\to 0$. In Section 2 we consider the subdifferential case $A=\partial f$ and, with no extra assumptions, we prove in Theorem 2 the strong convergence of $u(t)$ towards x^*. For general maximal monotone operators we prove in Theorem 9 of Section 3 that the same result holds if in addition the function $\varepsilon(t)$ has bounded variation. Finally, in Section 4 we provide a counterexample showing that convergence may fail without this bounded variation property.

2. Tikhonov dynamics in convex minimization

Let $f:\mathcal{H}\to\mathbb{R}\cup\{\infty\}$ be closed, proper and convex, and consider the minimization problem

$$\min_{x\in\mathcal{H}} f(x) \tag{P}$$

whose optimal solution set $S=\{x\in\mathcal{H}:\ 0\in\partial f(x)\}$ is assumed to be nonempty. The Tikhonov regularization scheme for (P) is the family of strongly convex problems

$$\min_{x\in\mathcal{H}} f_\varepsilon(x), \tag{P_ε}$$

where $f_\varepsilon(x)=f(x)+\frac{\varepsilon}{2}|x|^2$. It is well known (e.g. [37]) that the unique solution x_ε of (P_ε) converges strongly as $\varepsilon\to 0^+$ to the least-norm element of S, which we denote by x^*.

In this setting, the dynamics (D) with $A=\partial f$ correspond to the coupling of the Tikhonov regularization scheme with a steepest descent dynamics, namely

$$-\dot{u}(t)\in\partial f_{\varepsilon(t)}\big(u(t)\big)=\partial f\big(u(t)\big)+\varepsilon(t)u(t); \qquad u(0)=x_0. \tag{T}$$

Since (T) is a perturbed steepest descent method for $f(\cdot)$, we expect $u(t)$ to converge towards a point $x_\infty \in S$. The following slight variant of Gronwall's inequality will be used in the analysis.

Lemma 1. *Let $\theta : [0,\infty) \to \mathbb{R}$ be absolutely continuous with $\dot\theta(t) + \varepsilon(t)\theta(t) \leqslant \varepsilon(t)h(t)$ for almost all $t \geqslant 0$, where $h(t)$ is bounded and $\varepsilon(t) \geqslant 0$ with $\varepsilon \in L^1_{\mathrm{loc}}(\mathbb{R}_+)$. Then the function $\theta(t)$ is bounded and if $\int_0^\infty \varepsilon(\tau)\,d\tau = \infty$ we have $\limsup_{t\to\infty}\theta(t) \leqslant \limsup_{t\to\infty} h(t)$.*

Proof. Let $\kappa_s = \sup\{h(t)\colon\ t \geqslant s\}$ so that $\dot\theta(t) + \varepsilon(t)[\theta(t) - \kappa_s] \leqslant 0$ for $t \geqslant s$. Multiplying by $\exp(\int_0^t \varepsilon(\tau)\,d\tau)$ and integrating over $[s,t]$ we get

$$\bigl[\theta(t) - \kappa_s\bigr] \leqslant \bigl[\theta(s) - \kappa_s\bigr]\exp\Biggl(-\int_s^t \varepsilon(\tau)\,d\tau\Biggr). \tag{1}$$

It follows that $\theta(t)$ is bounded and, if $\int_0^\infty \varepsilon(\tau)\,d\tau = \infty$, then letting $t \to \infty$ in (1) we get $\limsup_{t\to\infty}\theta(t) \leqslant \kappa_s$, so that $s \to \infty$ yields $\limsup_{t\to\infty}\theta(t) \leqslant \limsup_{t\to\infty} h(t)$. $\square$

In this section we improve the known results, showing that the asymptotic convergence of Tikhonov dynamics holds as soon as $\varepsilon(t) \to 0^+$ when $t \to \infty$, without any extra assumption (not even monotonicity of $\varepsilon(t)$).

Theorem 2. *Let $u : [0,\infty) \to \mathcal{H}$ be the strong solution of (T) with $\varepsilon(t) \to 0^+$ as $t \to \infty$.*

(i) *If $\int_0^\infty \varepsilon(t)\,dt = \infty$ then $u(t) \to x^*$.*
(ii) *If $\int_0^\infty \varepsilon(t)\,dt < \infty$ then $u(t) \rightharpoonup x_\infty$ for some $x_\infty \in S$.*

Proof. (i) Let $\theta(t) = \frac{1}{2}|u(t) - x^*|^2$ so that $\dot\theta(t) = \langle \dot u(t), u(t) - x^*\rangle$. Using (T) and the strong convexity of $f_\varepsilon(\cdot)$ we get

$$f_{\varepsilon(t)}\bigl(u(t)\bigr) + \bigl\langle -\dot u(t), x^* - u(t)\bigr\rangle + \frac{1}{2}\varepsilon(t)\bigl|u(t) - x^*\bigr|^2 \leqslant f_{\varepsilon(t)}(x^*)$$

which may be rewritten as

$$\dot\theta(t) + \varepsilon(t)\theta(t) \leqslant f_{\varepsilon(t)}(x^*) - f_{\varepsilon(t)}\bigl(u(t)\bigr).$$

Since $f_\varepsilon(x_\varepsilon) \leqslant f_\varepsilon(u(t))$ and $f(x^*) \leqslant f(x_\varepsilon)$ we deduce

$$\dot\theta(t) + \varepsilon(t)\theta(t) \leqslant \frac{1}{2}\varepsilon(t)\bigl[|x^*|^2 - |x_{\varepsilon(t)}|^2\bigr]$$

and since $x_\varepsilon \to x^*$ as $\varepsilon \to 0^+$ (see for instance [37]), we may use Lemma 1 with $h(t) = \frac{1}{2}[|x^*|^2 - |x_{\varepsilon(t)}|^2]$ to conclude $\limsup_{t\to\infty}\theta(t) \leqslant 0$, hence $u(t) \to x^*$.

(ii) The proof is based on a result by [10]. Let $\bar x \in S$ and set $\theta(t) = \frac{1}{2}|u(t) - \bar x|^2$. Proceeding as in part (i) we get

$$\dot\theta(t) + \varepsilon(t)\theta(t) \leqslant f(\bar x) - f\bigl(u(t)\bigr) + \frac{1}{2}\varepsilon(t)\bigl[|\bar x|^2 - \bigl|u(t)\bigr|^2\bigr] \tag{2}$$

from which it follows that $\dot\theta(t) \leqslant \frac{1}{2}|\bar x|^2\varepsilon(t)$. Thus $\theta(t) - \frac{1}{2}|\bar x|^2\int_0^t \varepsilon(\tau)\,d\tau$ is decreasing and hence convergent so that $\theta(t)$ has a limit for $t \to \infty$. Invoking Opial's Lemma [30] the proof will follow if we show that every weak accumulation point of $u(t)$ belongs to S, for which it suffices to establish that $f(u(t)) \to \alpha := \inf_{x\in\mathcal{H}} f(x)$. To prove the latter we note that (T) may be written as $-\dot{u}(t) \in \partial f(u(t)) + v(t)$ with $v(t) = \varepsilon(t)u(t) \in L^1(0,\infty;\mathcal{H})$, so that [10, Lemma 3.3] implies that $f(u(t))$ is absolutely continuous with

$$\frac{d}{dt}\big[f\big(u(t)\big)\big] = -\big\langle \dot{u}(t) + \varepsilon(t)u(t), \dot{u}(t)\big\rangle \quad \text{a.e. } t \geqslant 0.$$

The latter may be bounded from above by $\delta(t) = \frac{1}{4}\varepsilon(t)^2|u(t)|^2 \in L^1(0,\infty;\mathbb{R})$, so that $\frac{d}{dt}[f(u(t)) - \int_0^t \delta(\tau)\,d\tau] \leqslant 0$ implying that $f(u(t)) - \int_0^t \delta(\tau)\,d\tau$ is decreasing and hence convergent. It follows that $f(u(t))$ converges as well. Now, using (2) we get $0 \leqslant f(u(t)) - f(\bar{x}) \leqslant -\dot{\theta}(t) + \frac{1}{2}|\bar{x}|^2\varepsilon(t)$ so that

$$\int_0^T \big[f\big(u(t)\big) - \alpha\big]\,dt \leqslant \theta(0) - \theta(T) + \frac{1}{2}|\bar{x}|^2 \int_0^T \varepsilon(t)\,dt \leqslant \theta(0) + \frac{1}{2}|\bar{x}|^2 \int_0^\infty \varepsilon(t)\,dt < \infty$$

which allows to conclude that the limit of $f(u(t))$ is indeed α as claimed. □

Remark. As mentioned in the introduction, when $\varepsilon(t)$ is non-increasing, part (i) was proved in [33]. This result went unnoticed and several special cases of it were rediscovered in [3,7,15] as examples of couplings of the steepest descent method with general approximation schemes. Particular cases of (ii) were described in [15,17], though we note that this may be deduced from the general results in [20] or, alternatively, from the results on asymptotic equivalence presented in [32].

Theorem 2 still holds, with essentially the same proof, when the regularizing kernel $\frac{1}{2}|x|^2$ is replaced by any strongly convex term. Moreover, part (i) admits the following straightforward generalization.

Proposition 3. *Let $f_\varepsilon(\cdot)$ be strongly convex with parameter $\beta(\varepsilon) > 0$, namely, for each $x \in \mathcal{H}$ and $y \in \partial f_\varepsilon(x)$*

$$f_\varepsilon(x) + \langle y, z - x\rangle + \frac{1}{2}\beta(\varepsilon)|z - x|^2 \leqslant f_\varepsilon(z), \quad \forall z \in \mathcal{H}.$$

Assume that the minimum x_ε of $f_\varepsilon(\cdot)$ has a strong limit x^* as $\varepsilon \to 0^+$. Suppose further that there is $y_\epsilon \in \partial f_\varepsilon(x^*)$ with $|y_\epsilon| \leqslant M\beta(\epsilon)$ for some $M \geqslant 0$. If $\int_0^\infty \beta(\varepsilon(t))\,dt = \infty$ then any solution of $-\dot{u}(t) \in \partial f_{\varepsilon(t)}(u(t))$ satisfies $u(t) \to x^*$ for $t \to \infty$.

Proof. Proceeding as in the previous proof we get

$$\begin{aligned}\dot{\theta}(t) + \beta\big(\varepsilon(t)\big)\theta(t) &\leqslant f_{\varepsilon(t)}(x^*) - f_{\varepsilon(t)}(x_{\varepsilon(t)}) \\ &\leqslant \langle y_{\varepsilon(t)}, x^* - x_{\varepsilon(t)}\rangle \\ &\leqslant M\beta\big(\varepsilon(t)\big)|x^* - x_{\varepsilon(t)}|\end{aligned}$$

so the conclusion follows again from Lemma 1 since $h(t) := M|x^* - x_{\varepsilon(t)}| \to 0$. □

3. Tikhonov dynamics for maximal monotone maps

Let us consider now the case of a maximal monotone operator $A : \mathcal{H} \to 2^{\mathcal{H}}$, and let $S = A^{-1}(0)$ denote the solution set of the inclusion $0 \in A(x)$. We suppose that S is nonempty and, as before, we denote x^* its least-norm element (recall that S is closed and convex). In contrast with the subdifferential case, the strong solution of (I) need not converge when $t \to \infty$ towards a point in S, unless some further restriction is imposed on the operator A. On the other hand, for any fixed $\varepsilon > 0$, the perturbed operator $A_\varepsilon = A + \varepsilon I$ is strongly monotone and the solution of the differential inclusion

$$-\dot{u}(t) \in A\big(u(t)\big) + \varepsilon u(t)$$

converges strongly to $x_\varepsilon = A_\varepsilon^{-1}(0)$.

Before analyzing the conditions for convergence in the non-autonomous case $\varepsilon(t)$ as in (D), we recall the following asymptotic property for the trajectory $\varepsilon \mapsto x_\varepsilon$. This corresponds to Lemma 1 in [13] and can be traced back to [29]. See also [16] for a recent extension with the identity operator replaced by a c-uniformly maximal monotone operator V. For the reader's convenience we include a short proof.

Lemma 4. *If $S \neq \emptyset$ then $x_\varepsilon \to x^*$ as $\varepsilon \to 0^+$.*

Proof. Monotonicity of A gives $\langle -\varepsilon x_\varepsilon, x_\varepsilon - x^* \rangle \geqslant 0$ so that $|x_\varepsilon| \leqslant |x^*|$ and x_ε remains bounded as $\varepsilon \to 0^+$. Thus $\varepsilon x_\varepsilon \to 0$ and since $\mathrm{gph}(A)$ is weak–strong sequentially closed, it follows that every weak cluster point $x_\infty = w - \lim x_{\varepsilon_k}$ with $\varepsilon_k \to 0$ belongs to S. The inequality $|x_{\varepsilon_k}| \leqslant |x^*|$ then gives $|x_\infty| \leqslant |x^*|$ by weak lower-semicontinuity of the norm, and then $x_\infty = x^*$ so that $x_\varepsilon \rightharpoonup x^*$. Since we also have $|x_\varepsilon| \to |x^*|$, the convergence is strong. □

Let us go back to the Tikhonov dynamics (D) with $\varepsilon(t) \to 0^+$ as $t \to \infty$. The case when $\int_0^\infty \varepsilon(t)\,dt < \infty$ may be completely analyzed by combining [32, Proposition 7.9] and [32, Proposition 8.5]: the trajectories of (D) converge (either weakly or strongly) to a point in S if and only if the corresponding property holds for the unperturbed dynamics (I). Let us then address the question whether $\int_0^\infty \varepsilon(t)\,dt = \infty$ is enough to ensure the convergence of the trajectories. We shall see that the answer is negative in general, but under some additional assumptions one can establish strong convergence to x^*. For instance, adapting the arguments in [3], we can easily prove the following:

Proposition 5. *Suppose $\varepsilon(t)$ is decreasing to 0 and let $u(t)$ be the strong solution of (D). Assume $\int_0^\infty \varepsilon(t)\,dt = \infty$ and also that either the path $\varepsilon \mapsto x_\varepsilon$ has finite length or the parameter function satisfies $\dot{\varepsilon}(t)/\varepsilon(t)^2 \to 0$ as $t \to \infty$. Then $u(t) \to x^*$ strongly.*

Proof. The proof consists in showing that $\theta(t) = \frac{1}{2}|u(t) - x_{\varepsilon(t)}|^2$ tends to 0. We recall that $x_\varepsilon = (A + \varepsilon I)^{-1}(0)$ is absolutely continuous on $(0, \infty)$ (see e.g. [3, p. 530]). Differentiating we get

$$\dot{\theta}(t) = \left\langle \dot{u}(t) - \dot{\varepsilon}(t)\frac{d}{d\varepsilon}x_{\varepsilon(t)}, u(t) - x_{\varepsilon(t)} \right\rangle$$

for almost all $t \geqslant 0$, and then using the strong monotonicity of $A + \varepsilon I$ we deduce

$$\dot{\theta}(t) \leqslant -2\varepsilon(t)\theta(t) - \dot{\varepsilon}(t)\left|\frac{d}{d\varepsilon}x_{\varepsilon(t)}\right|\sqrt{2\theta(t)}$$

which is the same inequality obtained in [3] so that the arguments in that paper yield $\theta(t) \to 0$ as required. □

This extension, included here for completeness, was suggested in [28] and it appeared in the recent thesis [22]. Now, the case $\dot{\varepsilon}(t)/\varepsilon(t)^2 \to 0$ was already studied in [24] and, as a matter of fact, it may be obtained as a particular case of a more general statement [33, Theorem 1.4] which can be itself traced back to [12, Theorem 10.12] for a special class of operators (see also [34,35]). These more general results do not require finite length of $\varepsilon \mapsto x_\varepsilon$ nor $\dot{\varepsilon}(t)/\varepsilon(t)^2 \to 0$, but only $\varepsilon(t)$ to be decreasing. We shall prove that even this monotonicity condition can be relaxed. We begin by characterizing the strong convergence of the solutions of (D).

Proposition 6. *The strong solution $u(t)$ of (D) is bounded and if $\int_0^\infty \varepsilon(\tau)\,d\tau = \infty$ then the following properties are equivalent*:

(a) *all weak cluster points of $u(t)$ for $t \to \infty$ belong to S*,
(b) $\liminf_{t\to\infty} |u(t)| \geqslant |x^*|$,
(c) $u(t) \to x^*$ *strongly.*

Proof. Let $\theta(t) = \frac{1}{2}|u(t) - x^*|^2$. Differentiating and using the monotonicity of A we get

$$\begin{aligned}\dot\theta(t) &= \langle \dot u(t), u(t) - x^*\rangle \\ &= \langle \dot u(t) + \varepsilon(t)u(t), u(t) - x^*\rangle + \varepsilon(t)\langle u(t), x^* - u(t)\rangle \\ &\leqslant \varepsilon(t)\langle u(t), x^* - u(t)\rangle \\ &= \frac{\varepsilon(t)}{2}\left[|x^*|^2 - |u(t)|^2 - |x^* - u(t)|^2\right]\end{aligned}$$

so that setting $h(t) = \frac{1}{2}[|x^*|^2 - |u(t)|^2]$ we obtain

$$\dot\theta(t) + \varepsilon(t)\theta(t) \leqslant \varepsilon(t)h(t).$$

Applying Lemma 1 we deduce that $\theta(t)$ is bounded and therefore so is $u(t)$. On the other hand, (a) $\Rightarrow$ (b) follows from the weak lower-semicontinuity of the norm, while (c) $\Rightarrow$ (a) is straightforward (both implications hold no matter what the value of $\int_0^\infty \varepsilon(\tau)\,d\tau$ is). Finally, (b) $\Rightarrow$ (c) follows from Lemma 1 provided that $\int_0^\infty \varepsilon(\tau)\,d\tau = \infty$ since then $\limsup_{t\to\infty}\theta(t) \leqslant \limsup_{t\to\infty} h(t) \leqslant 0$ so that $\theta(t) \to 0$. $\square$

Remark. The implication (b) $\Rightarrow$ (c) may fail if $\int_0^\infty \varepsilon(\tau)\,d\tau < \infty$. To see this, take $A = \partial f$ given by Baillon's counterexample for strong convergence in [5]: the solutions of (D) converge weakly but not strongly to some element of S, thus they satisfy (a) and (b), but not (c). To see the latter we invoke the equivalence result in [32] to deduce that the systems with or without $\varepsilon(t)$ have the same asymptotic behavior.

The next lemmas provide tools to check that condition (a) in Proposition 6 holds. From now on we exploit the fact that the function $\varepsilon(t)$ has bounded variation.

Lemma 7. *Suppose $\varepsilon(t) \to 0^+$ for $t \to \infty$ and $\dot u(t) \to 0$ when $t \to \infty$, $t \in D$, where D is a dense subset of $[0,\infty)$. Then all weak cluster points of $u(t)$ for $t \to \infty$ are in S.*

Proof. Let $\bar x$ be a weak cluster point of $u(t)$ and choose $t_k \to \infty$ with $u(t_k) \rightharpoonup \bar x$. Since $u(\cdot)$ is continuous we may find $\tilde t_k \in D$ close enough to t_k so that $|u(\tilde t_k) - u(t_k)| \leqslant \frac{1}{k}$ and therefore $u(\tilde t_k) \rightharpoonup \bar x$. Then $\dot u(\tilde t_k) \to 0$ and since $\varepsilon(t) \to 0$ and $u(t)$ is bounded it follows that $v_k := -\dot u(\tilde t_k) - \varepsilon(\tilde t_k)u(\tilde t_k) \to 0$ with $v_k \in A(u(\tilde t_k))$, from which we conclude $0 \in A(\bar x)$ as required. $\square$

Lemma 8. *If $\int_0^\infty \varepsilon(t)\,dt = \infty$ and $\int_0^\infty |\dot\varepsilon(t)|\,dt < \infty$ then there exists $D \subset [0,\infty)$ with full measure such that $\dot u(t) \to 0$ when $t \to \infty$, $t \in D$.*

Proof. Let $\theta(t) = \frac{1}{2}|u(t+\delta) - u(t)|^2$ with $\delta > 0$ so that

$$\begin{aligned}\dot\theta(t) &= \langle \dot u(t+\delta) - \dot u(t), u(t+\delta) - u(t)\rangle \\ &\leqslant \varepsilon(t+\delta)\langle u(t+\delta), u(t) - u(t+\delta)\rangle + \varepsilon(t)\langle u(t), u(t+\delta) - u(t)\rangle\end{aligned}$$

$$= -\big[\varepsilon(t+\delta)+\varepsilon(t)\big]\theta(t) + \frac{1}{2}\big[\varepsilon(t)-\varepsilon(t+\delta)\big]\big[\big|u(t+\delta)\big|^2 - \big|u(t)\big|^2\big].$$

Multiplying this inequality by $\exp(E_t^\delta)$ where $E_t^\delta = \int_0^t [\varepsilon(\tau+\delta)+\varepsilon(\tau)]\,d\tau$, we may integrate over $[s,t]$ in order to obtain

$$\exp\big(E_t^\delta\big)\theta(t) \leqslant \exp\big(E_s^\delta\big)\theta(s) + \frac{1}{2}\int_s^t \exp\big(E_\tau^\delta\big)\big[\varepsilon(\tau)-\varepsilon(\tau+\delta)\big]\big[\big|u(\tau+\delta)\big|^2 - \big|u(\tau)\big|^2\big]\,d\tau.$$

Now $u(\cdot)$ is differentiable on a set $D \subseteq [0,\infty)$ of full measure, so that multiplying the previous inequality by $2/\delta^2$ and letting $\delta \to 0^+$ it follows that for all $s,t \in D$ with $s \leqslant t$ we have

$$\begin{aligned}\exp\big(E_t^0\big)\big|\dot u(t)\big|^2 &\leqslant \exp\big(E_s^0\big)\big|\dot u(s)\big|^2 - 2\int_s^t \exp\big(E_\tau^0\big)\dot\varepsilon(\tau)\big\langle \dot u(\tau), u(\tau)\big\rangle\,d\tau \\ &\leqslant \exp\big(E_s^0\big)\big|\dot u(s)\big|^2 + \int_s^t \exp\big(E_\tau^0\big)\big|\dot\varepsilon(\tau)\big|\big[\big|\dot u(\tau)\big|^2 + \big|u(\tau)\big|^2\big]\,d\tau.\end{aligned}$$

Denoting $\phi(t) = \exp(E_t^0)|\dot u(t)|^2$ and $R = \sup_{\tau\geqslant 0}|u(\tau)|$ we get

$$\phi(t) \leqslant \phi(s) + R^2\int_s^t \exp\big(E_\tau^0\big)\big|\dot\varepsilon(\tau)\big|\,d\tau + \int_s^t \big|\dot\varepsilon(\tau)\big|\phi(\tau)\,d\tau$$

and since the quantity $\kappa(s,t) = \phi(s) + R^2\int_s^t \exp(E_\tau^0)|\dot\varepsilon(\tau)|\,d\tau$ is non-decreasing in t, we may use Gronwall's inequality to deduce

$$\phi(z) \leqslant \kappa(s,t)\exp\left(\int_s^z \big|\dot\varepsilon(\tau)\big|\,d\tau\right), \quad \forall z \in [s,t].$$

In particular, for $z = t$ this gives

$$\begin{aligned}\big|\dot u(t)\big|^2 &\leqslant \left[\phi(s)\exp\big(-E_t^0\big) + R^2\int_s^t \exp\big(E_\tau^0 - E_t^0\big)\big|\dot\varepsilon(\tau)\big|\,d\tau\right]\exp\left(\int_s^t \big|\dot\varepsilon(\tau)\big|\,d\tau\right) \\ &\leqslant \left[\phi(s)\exp\big(-E_t^0\big) + R^2\int_s^t \big|\dot\varepsilon(\tau)\big|\,d\tau\right]\exp\left(\int_s^t \big|\dot\varepsilon(\tau)\big|\,d\tau\right)\end{aligned}$$

and letting $t \to \infty$ with $t \in D$ we obtain

$$\limsup_{t\to\infty,\, t\in D}\big|\dot u(t)\big|^2 \leqslant R^2\exp\left(\int_s^\infty \big|\dot\varepsilon(\tau)\big|\,d\tau\right)\int_s^\infty \big|\dot\varepsilon(\tau)\big|\,d\tau.$$

Since the right-hand side expression tends to 0 for $s \to \infty$, we conclude that $\dot u(t) \to 0$ for $t \to \infty$, $t \in D$. □

Combining Proposition 6 with Lemmas 7 and 8 we obtain the announced extension of [33, Theorem 1.4].

Theorem 9. *Let $u(t)$ be the strong solution of (D) and assume that $\varepsilon(t) \to 0$ as $t \to \infty$ with $\int_0^\infty \varepsilon(t)\,dt = \infty$ and $\int_0^\infty |\dot{\varepsilon}(t)|\,dt < \infty$. Then $u(t) \to x^*$ strongly.*

4. Counterexamples

4.1. A non-convergent Tikhonov-like trajectory

In this subsection we give a counterexample showing that Theorem 9 may fail if $\varepsilon(t)$ is not of bounded variation. The idea is as follows. Consider $A(x) = (1 - x_2, x_1 - 1)$ the $\frac{\pi}{2}$-rotation around the unique rest point $p = (1, 1)$. The Tikhonov trajectory is $x_\varepsilon = \frac{1}{1+\varepsilon^2}(1 - \varepsilon, 1 + \varepsilon)$ and describes a half-circle with center at $(\frac{1}{2}, \frac{1}{2})$ and radius $\frac{1}{\sqrt{2}}$ (see dotted line in Fig. 1). For the dynamics, let us start from a point x_0 on the other half of this circle and let d be its distance to p. Fix $\varepsilon > 0$ and follow the trajectory of $-\dot{u}(t) = Au(t) + \varepsilon u(t)$ which spirals towards x_ε. On a first phase $u(t)$ increases its distance to p and afterwards it comes closer again (see Fig. 1). Stop exactly when the distance is again d and shift to $\varepsilon = 0$ in such a way that the trajectory now turns around p until it comes back to the initial point x_0, from where we restart a new cycle with a smaller ε. To make this idea more precise and to simplify the computations we use complex numbers, identifying $\mathbb{R}^2$ with $\mathbb{C}$.

The operator: Since A is the $\frac{\pi}{2}$ clockwise rotation in the plane around the point $p = 1 + i$, Eq. (D) may be rewritten as

$$\dot{u}(t) = -i\big(u(t) - p\big) - \varepsilon(t)u(t). \tag{3}$$

The parameter function: Let ε_n be a sequence of positive real numbers with $\varepsilon_n \to 0$ and $\sum \varepsilon_n = \infty$. Take $a_0 = 0$ and let $b_n = a_n + \tau_n$, $a_{n+1} = b_n + \sigma_n$ with $\tau_n > 0$, $\sigma_n > 0$ to be fixed later on, and consider the step function

$$\varepsilon(t) = \begin{cases} \varepsilon_n & \text{if } a_n \leqslant t < b_n, \\ 0 & \text{if } b_n \leqslant t < a_{n+1}. \end{cases}$$

Clearly $\varepsilon(t) \to 0^+$ and we get $\int_0^\infty \varepsilon(t)\,dt = \infty$ provided τ_n is bounded away from zero.

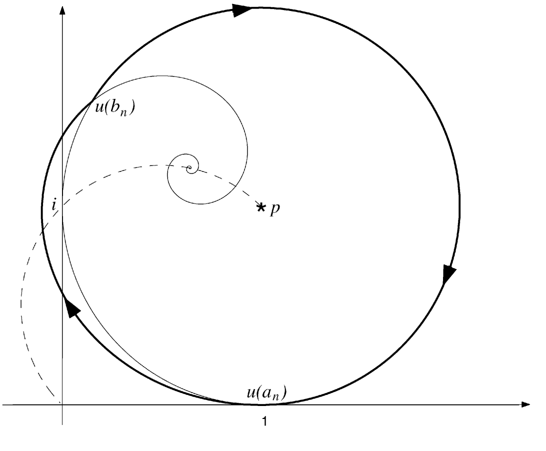

Fig. 1. The trajectory $u(t)$ on the interval $[a_n, a_{n+1}]$, starting from 1 and back.

The dynamics: Let $u(a_n) = 1 \in \mathbb{C}$. On the interval $[a_n, b_n)$ the solution of (3) is

$$u(t) = \frac{1}{\varepsilon_n + i}\big[i - 1 + (1+\varepsilon_n)e^{-(\varepsilon_n+i)(t-a_n)}\big]. \tag{4}$$

Let $t = b_n$ be the first time after a_n with $|u(t) - p| = 1$, so that $\tau_n = b_n - a_n$ may be characterized as the first positive zero of the function

$$\psi_n(s) = (1+\varepsilon_n)e^{-2\varepsilon_n s} + 2\varepsilon_n e^{-\varepsilon_n s}\big[\sin(s) - \cos(s)\big] + \varepsilon_n - 1.$$

We claim that if $\varepsilon_n \leqslant \frac{1}{2}$ then $\tau_n \in [\frac{1}{4}, \frac{3}{2}\pi]$. For the lower bound, since $\psi_n(0) = 0$ it suffices to show that $\psi_n'(s) > 0$ for all $s \in (0, \frac{1}{4})$. Now, $\psi_n'(s) = 2\varepsilon_n e^{-\varepsilon_n s}\phi_n(s)$ with $\phi_n(s) = (1+\varepsilon_n)\cos(s) + (1-\varepsilon_n)\sin(s) - (1+\varepsilon_n)e^{-\varepsilon_n s}$, and since $\phi_n(0) = 0$ it suffices to check $\phi_n'(s) > 0$ for $s \in (0, \frac{1}{4})$, which follows from

$$\phi_n'(s) = (1-\varepsilon_n)\cos(s) - (1+\varepsilon_n)\sin(s) + \varepsilon_n(1+\varepsilon_n)e^{-\varepsilon_n s} > \frac{1}{2}\big[\cos(s) - 3\sin(s)\big] > 0.$$

For the upper bound we just prove that $\psi_n(\frac{3}{2}\pi) < 0$. To this end we set $\rho = e^{-\frac{3}{2}\pi\varepsilon_n}$ so that $\rho \in (0,1)$ and therefore

$$\psi_n\left(\frac{3}{2}\pi\right) = (\rho - 1)\big[1 + \rho + \varepsilon_n(\rho - 1)\big] = (\rho - 1)\big[2\rho + (1-\varepsilon_n)(1-\rho)\big] < 0.$$

On the interval $[b_n, a_{n+1})$ the solution is $u(t) = p + (u(b_n) - p)e^{-i(t-b_n)}$, and we may pick σ_n such that $u(a_{n+1}) = 1$ in order for the solution to cycle indefinitely. More precisely, let σ_n be the first positive solution of $e^{is} = i(u(b_n) - p)$. Such a positive solution exists because $|u(b_n) - p| = 1$. On the interval $[b_n, a_{n+1})$, the trajectory $u(t)$ travels from $u(b_n)$ to 1 along the circle $|z - p| = 1$. Now, Eq. (4) implies that the real part of $u(b_n)$ is strictly less than 1. Therefore, the trajectory covers at least the arc joining (clockwise) the points $1 + 2i$ and 1 on the circle $|z - p| = 1$ as t goes from b_n to a_{n+1}, so it cannot converge as $t \to \infty$.

Remark. The lack of continuity of the function $\varepsilon(t)$ is not the problem, nor is it the fact that $\varepsilon(t)$ vanishes in some intervals. In fact, one can find $\eta \in C^\infty(\mathbb{R}_+; \mathbb{R}_{++})$ such that $\eta \notin L^1(0,\infty)$ while $\varepsilon - \eta \in L^1(0,\infty)$. Obviously this η will not be of bounded variation. The arguments in [32] show that Eq. (4) with $\eta(t)$ instead of the previous $\varepsilon(t)$ has the same asymptotic behavior and therefore it will not converge.

4.2. A non-convergent discrete trajectory

Given the close connection between evolution equations and the proximal point method [18,19,26, 27,31,32,35], a natural question is whether one may find sequences $\{\lambda_n\}$ and $\{\theta_n\}$ with $\sum \lambda_n\theta_n = \infty$ and such that the discrete trajectory generated by the (perturbed) proximal point algorithm

$$\frac{x_{n-1} - x_n}{\lambda_n} \in Ax_n + \theta_n x_n$$

does not converge. This is strongly related to [34]. Observe that in the unperturbed case ($\theta_n \equiv 0$) the sequence x_n converges weakly in average [6]. For $A = \partial f$ the sequence converges weakly [11], but the counterexample in [21] (based on that of [5]) shows that this convergence need not be strong; answering a question posed earlier in [36]. More examples of this kind have appeared recently in [8,9], based on results of [23].

Let $\varepsilon(t)$ be the function defined in Section 4.1. One can select a non-increasing sequence $\{\lambda_n\}$ in such a way that the function ε is constant on each interval of the form $[\Lambda_n, \Lambda_{n+1})$, where $\Lambda_n = \sum_{k=1}^{n} \lambda_k \to \infty$. Define $\theta_n = \varepsilon(\Lambda_n)$ and observe that

$$\sum_{n=1}^{\infty} \lambda_n \theta_n = \int_0^{\infty} \varepsilon(t)\, dt = \infty.$$

With these conditions, a corollary of Kobayashi's inequality (see [26] as well as [21], [1] or [32]) states that

$$\left|u(t) - x_n\right| \leqslant \left|u(s) - x_k\right| + |Bx_k| \sqrt{\left[(\Lambda_n - \Lambda_k) - (t-s)\right]^2 + \sum_{j=k+1}^{n} \lambda_j^2}, \tag{5}$$

where B is any maximal monotone operator, $x_n = \prod_{j=1}^{n}(I + \lambda_j B)^{-1}x$ is a corresponding proximal sequence, and u satisfies $-\dot{u}(t) \in Bu(t)$.

Consider now the indices J_n such that the discontinuities of the function $\varepsilon(t)$ lie precisely on the set $\{\Lambda_{J_n}\}$. We have

$$\sum_{k=J_n+1}^{J_{n+1}} \lambda_k^2 \leqslant \lambda_{J_n+1}(\Lambda_{J_{n+1}} - \Lambda_{J_n}) \leqslant 2M\lambda_{J_n},$$

where M is an upper bound for the τ_n's and the σ_n's.

Let $U(t,s)x = u(t)$, where $-\dot{u}(t) = Au(t) + \varepsilon(t)u(t)$ and $u(s) = x$. Define also $V(t,s)x = \prod_{k=\nu(s)+1}^{\nu(t)}[I + \lambda_k(A + \theta_k I)]^{-1}x$, where $\nu(t) = \max\{k \in \mathbb{N} \mid \Lambda_k \leqslant t\}$. Applying inequality (5) repeatedly for $B_n = A + \theta_n I$ in the appropriate subintervals one gets

$$\left|U(t,s)x - V(t,s)x\right| \leqslant K \sum_{n=\nu(s)+1}^{\nu(t)} \sqrt{\lambda_{J_n}}$$

for some constant K, which depends on a bound for the sequence $\{Ax_n + \varepsilon(\Lambda_n)x_n\}$. If $\sum_{k=1}^{\infty} \sqrt{\lambda_{J_k}}$ is finite, this implies that the trajectories $t \mapsto U(t,s)x$ converge if and only if the same holds for $t \mapsto V(t,s)x$. Therefore the proximal point algorithm cannot always converge.

Sequences satisfying $\sum_{k=1}^{\infty} \sqrt{\lambda_{J_k}} < \infty$ and not being in ℓ^1 are difficult to characterize. However we can provide a very simple example. First, let m be a positive lower bound for the τ_n's and the σ_n's. Define $\{\lambda_n\}$ as follows: for $4^{k-1} < n \leqslant 4^k$ set $\lambda_n = 4^{-k}m$. We then have $\sum_{n\geqslant 0} \lambda_n = \infty$, while $\sum_{n\geqslant 1} \sqrt{\lambda_{J_n}} \leqslant m \sum_{n\geqslant 0} 2^{-n} < \infty$.

References

[1] F. Álvarez, J. Peypouquet, Asymptotic equivalence and Kobayashi-type estimates for nonautonomous monotone operators in Banach spaces, CMM Technical Report CMM_B_07_08_191, 2007.

[2] F. Álvarez, J. Peypouquet, Asymptotic almost-equivalence of abstract evolution systems, CMM Technical Report CMM_B_07_08_190, 2007.

[3] H. Attouch, R. Cominetti, A dynamical approach to convex minimization coupling approximation with the steepest descent method, J. Differential Equations 128 (1996) 519–540.

[4] H. Attouch, A. Damlamian, Strong solutions for parabolic variational inequalities, Nonlinear Anal. 2 (1978) 329–353.

[5] J.B. Baillon, An example concernant le comportement asymptotique de la solution du problème $\frac{du}{dt} + \partial\varphi(u) \ni 0$, J. Funct. Anal. 28 (1978) 369–376.

[6] J.B. Baillon, H. Brézis, Une remarque sur le comportement asymptotique des semi-groupes non linéaires, Houston J. Math. 2 (1976) 5–7.

[7] J.B. Baillon, R. Cominetti, A convergence result for non-autonomous subgradient evolution equations and its application to the steepest descent exponential penalty trajectory in linear programming, J. Funct. Anal. 187 (2001) 263–273.
[8] H.H. Bauschke, J.V. Burke, F.R. Deutsch, H.S. Hundal, J.D. Vanderwerff, A new proximal point iteration that converges weakly but not in norm, Proc. Amer. Math. Soc. 133 (6) (2005) 1829–1835.
[9] H.H. Bauschke, E. Matoušková, S. Reich, Projection and proximal point methods: Convergence results and counterexamples, Nonlinear Anal. 56 (2004) 715–738.
[10] H. Brézis, Opérateurs maximaux monotones et semigroupes de contractions dans les espaces de Hilbert, North-Holland Math. Stud., vol. 5, North-Holland, Amsterdam, 1973.
[11] H. Brézis, P.L. Lions, Produits infinis de résolvantes, Israel J. Math. 29 (1978) 329–345.
[12] F.E. Browder, Nonlinear Operators and Nonlinear Equations of Evolution in Banach Spaces, Proc. Sympos. Pure Math., vol. 18 (part 2), Amer. Math. Soc., Providence, RI, 1976.
[13] R.E. Bruck, A strongly convergent iterative solution of $0 \in U(x)$ for a maximal monotone operator U in Hilbert space, J. Math. Anal. Appl. 48 (1974) 114–126.
[14] R.E. Bruck, Asymptotic convergence of nonlinear contraction semigroups in Hilbert space, J. Funct. Anal. 18 (1975) 15–26.
[15] A. Cabot, The steepest descent dynamical system with control. Applications to constrained minimization, ESAIM Control Optim. Calc. Var. 10 (2004) 243–258.
[16] P.L. Combettes, S.A. Hirstoaga, Approximating curves for nonexpansive and monotone operators, J. Convex Anal. 13 (2006) 633–646.
[17] R. Cominetti, O. Alemany, Steepest descent evolution equations: Asymptotic behavior of solutions and rate of convergence, Trans. Amer. Math. Soc. 351 (1999) 4847–4860.
[18] M.G. Crandall, T.M. Liggett, Generation of semigroups of nonlinear transformations on general Banach spaces, Amer. J. Math. 93 (1971) 265–298.
[19] M.G. Crandall, A. Pazy, Nonlinear evolution equations in Banach spaces, Israel J. Math. 11 (1972) 57–94.
[20] H. Furuya, K. Miyashiba, N. Kenmochi, Asymptotic behavior of solutions to a class of nonlinear evolution equations, J. Differential Equations 62 (1986) 73–94.
[21] O. Güler, On the convergence of the proximal point algorithm for convex minimization, SIAM J. Control Optim. 29 (1991) 403–419.
[22] S. Hirstoaga, Approximation et résolution de problèmes d'équilibre, de point fixe et d'inclusion monotone, PhD thesis, UPMC Paris 6, 2006.
[23] H.S. Hundal, An alternating projection that does not converge in norm, Nonlinear Anal. 57 (1) (2004) 35–61.
[24] M.M. Israel Jr., S. Reich, Asymptotic behavior of solutions of a nonlinear evolution equation, J. Math. Anal. Appl. 83 (1981) 43–53.
[25] N. Kenmochi, Solvability of nonlinear evolution equations with time-dependent constraints and applications, Bull. Fac. Ed. Chiba Univ. 30 (1981) 1–87.
[26] Y. Kobayashi, Difference approximation of Cauchy problems for quasi-dissipative operators and generation of nonlinear semigroups, J. Math. Soc. Japan 27 (1975) 640–665.
[27] K. Kobayasi, Y. Kobayashi, S. Oharu, Nonlinear evolution operators in Banach spaces, Osaka J. Math. 21 (1984) 281–310.
[28] B. Lemaire, Staircase parametrization in dynamical selection, Set-Valued Anal. 9 (2001) 111–121.
[29] G.J. Minty, On a monotonicity method for the solution of nonlinear equations in Banach spaces, Proc. Natl. Acad. Sci. USA 50 (1963) 1038–1041.
[30] Z. Opial, Weak convergence of the sequence of successive approximations for nonexpansive mappings, Bull. Amer. Math. Soc. 73 (1967) 591–597.
[31] G. Passty, Preservation of the asymptotic behavior of a nonlinear contraction semigroup by backward differencing, Houston J. Math. 7 (1981) 103–110.
[32] J. Peypouquet, Analyse asymptotique de systèmes d'évolution et applications en optimisation, PhD thesis, UPMC Paris 6 and U. de Chile, 2007.
[33] S. Reich, Nonlinear evolution equations and nonlinear ergodic theorems, Nonlinear Anal. 1 (1976) 319–330.
[34] S. Reich, An iterative procedure for constructing zeros of accretive sets in Banach spaces, Nonlinear Anal. 2 (1978) 85–92.
[35] S. Reich, Strong convergence theorems for resolvents of accretive operators in Banach spaces, J. Math. Anal. Appl. 75 (1980) 287–292.
[36] R.T. Rockafellar, Monotone operators and the proximal point algorithm, SIAM J. Control Optim. 14 (1976) 877–898.
[37] A. Tikhonov, V. Arsenine, Méthodes de résolution de problèmes mal posés, Mir, Moscow, 1974.