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

Roberto Cominetti* Juan Peypouquet† Sylvain Sorin‡

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 and the parameter function $\varepsilon(t)$ tends to 0 for $t \to \infty$ 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.

1 Introduction

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

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

where $A : \mathcal{H} \to 2^{\mathcal{H}}$ is a maximal monotone operator on a Hilbert space $\mathcal{H}$, $\varepsilon(t) \geq 0$ 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 $t \geq 0$. Sufficient conditions for this existence may be found, among others, in [4, Attouch and Damlamian], [14, Crandall and Pazy], [15, Furuya *et al.*], and [19, Kenmochi].

The differential inclusion (D) is a perturbed version of

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

*Departamento de Ingeniería Matemática and Centro de Modelamiento Matemático, Universidad de Chile ⟨*rcominet@dim.uchile.cl*⟩. Supported by FONDAP grant in Applied Mathematics, CONICYT-Chile.

†Departamento de Ingeniería Matemática, Universidad de Chile ⟨*jpeypou@dim.uchile.cl*⟩. Supported by MECESUP-Chile.

‡Equipe Combinatoire et Optimisation, UFR 929, Université P. et M. Curie – Paris 6, 175 rue du Chevaleret, 75013 Paris and Laboratoire d'Econométrie – Ecole Polytechnique ⟨*sorin@math.jussieu.fr*⟩. Supported by grant ANR-05-BLAN-0248-01.

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 [10, Bruck] 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, Baillon], 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 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 [25, Peypouquet] (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 [26, Reich] (based on previous work by [8, Browder]) 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 §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 §3 that the same result holds if in addition the function $\varepsilon(t)$ has bounded variation. Finally in §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

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

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

$$(P_\varepsilon) \qquad \min_{x \in \mathcal{H}} f_\varepsilon(x)$$

where $f_\varepsilon(x) = f(x) + \frac{\varepsilon}{2}|x|^2$. It is well known (*e.g.* [28, Tikhonov and Arsenin]) 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

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

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 $\dot{\theta}(t)+\varepsilon(t)\theta(t) \leq \varepsilon(t)h(t)$ for almost all $t \geq 0$, where $h(t)$ is bounded and $\varepsilon(t) \geq 0$ with $\varepsilon \in L^1_{\mathrm{loc}}(\mathbb{R}_+)$. Then $\theta(t)$ is bounded and if $\int_0^\infty \varepsilon(\tau)d\tau = \infty$ we have $\limsup_{t\to\infty} \theta(t) \leq \limsup_{t\to\infty} h(t)$.*

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

$$[\theta(t) - \kappa_s] \leq [\theta(s) - \kappa_s] \exp(-\textstyle\int_s^t \varepsilon(\tau)d\tau). \tag{1}$$

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 $\limsup_{t\to\infty} \theta(t) \leq \kappa_s$, so that $s \to \infty$ yields $\limsup_{t\to\infty} \theta(t) \leq \limsup_{t\to\infty} h(t)$. □

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

$$f_{\varepsilon(t)}(u(t)) + \langle -\dot{u}(t), x^* - u(t) \rangle + \tfrac{1}{2}\varepsilon(t)|u(t) - x^*|^2 \leq f_{\varepsilon(t)}(x^*)$$

which may be rewritten as

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

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

$$\dot{\theta}(t) + \varepsilon(t)\theta(t) \leq \tfrac{1}{2}\varepsilon(t) \left[|x^*|^2 - |x_{\varepsilon(t)}|^2\right]$$

and since $x_\varepsilon \to x^*$ as $\varepsilon \to 0$ (see for instance [28]), we may use Lemma 1 with $h(t) = \frac{1}{2}[|x^*|^2 - |x_{\varepsilon(t)}|^2]$ to conclude $\limsup_{t\to\infty} \theta(t) \le 0$, hence $u(t) \to x^*$.

(ii) The proof is based on a result by [7, Brézis]. Let $\bar{x} \in S$ and set $\theta(t) = \frac{1}{2}|u(t) - \bar{x}|^2$. Proceeding as in part *(i)* we get

$$\dot{\theta}(t) + \varepsilon(t)\theta(t) \le f(\bar{x}) - f(u(t)) + \tfrac{1}{2}\varepsilon(t)\left[|\bar{x}|^2 - |u(t)|^2\right] \tag{2}$$

from which it follows that $\dot{\theta}(t) \le \frac{1}{2}|\bar{x}|^2\varepsilon(t)$. 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 [23] 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 [7, Lemma 3.3] implies that $f(u(t))$ is absolutely continuous with

$$\tfrac{d}{dt}\left[f(u(t))\right] = -\langle \dot{u}(t) + \varepsilon(t)u(t), \dot{u}(t)\rangle \qquad \text{a.e. } t \ge 0.$$

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 $\frac{d}{dt}\left[f(u(t)) - \int_0^t \delta(\tau)\, d\tau\right] \le 0$ 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 $0 \le f(u(t)) - f(\bar{x}) \le -\dot{\theta}(t) + \frac{1}{2}|\bar{x}|^2\varepsilon(t)$ so that

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

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 [26, Reich]. This result went unnoticed and several special cases of it were rediscovered in [3, 6, 11] as examples of couplings of the steepest descent method with general approximation schemes. Particular cases of (ii) were described in [11, 13], though we note that this may be deduced from the general results in [15, Furuya, Miyashiba and Kenmochi] or, alternatively, from the results on asymptotic equivalence presented in [25, Peypouquet].

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 $y \in \partial f_\varepsilon(x)$*

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

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 $|y_\epsilon| \leq M\beta(\epsilon)$ for some $M \geq 0$. 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

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

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. 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

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

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 [9, Bruck] and can be traced back to [22, Minty]. See also [12, Combettes and Hirstoaga] 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 $\langle -\varepsilon x_\varepsilon, x_\varepsilon - x^*\rangle \geq 0$ so that $|x_\varepsilon| \leq |x^*|$ 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 $|x_{\varepsilon_k}| \leq |x^*|$ then gives $|x_\infty| \leq |x^*|$ 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 [25, Proposition 7.9] and [25, Proposition 8.5]: the trajectories of (D) converge (weak or strong) 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, page 530]). Differentiating we get

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

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

$$\dot\theta(t) \leq -2\varepsilon(t)\theta(t) - \dot\varepsilon(t)|\tfrac{d}{d\varepsilon}x_{\varepsilon(t)}|\sqrt{2\theta(t)}$$

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

This extension, included here for completeness, was suggested in [21, Lemaire] and it appeared in the recent thesis [17, Hirstoaga]. Now, the case $\dot\varepsilon(t)/\varepsilon(t)^2 \to 0$ was already studied in [18, Proposition 10] and, as a matter of fact, it may be obtained as a particular case of a more general statement [26, Theorem 1.4] which can be itself traced back to [8, Theorem 10.12] (for a special class of operators A). 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) $\liminf_{t\to\infty} |u(t)| \geq |x^*|$,
(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

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

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

$$\dot{\theta}(t) + \varepsilon(t)\theta(t) \leq \varepsilon(t)h(t).$$

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

Remark. The implication *(b)⇒(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 [25, Peypouquet] 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 $|u(\tilde{t}_k) - u(t_k)| \le \frac{1}{k}$ 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. □

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

$$\begin{aligned}\dot{\theta}(t) &= \langle \dot{u}(t+\delta) - \dot{u}(t), u(t+\delta) - u(t)\rangle \\ &\le \varepsilon(t+\delta)\langle u(t+\delta), u(t) - u(t+\delta)\rangle + \varepsilon(t)\langle u(t), u(t+\delta) - u(t)\rangle \\ &= -[\varepsilon(t+\delta) + \varepsilon(t)]\theta(t) + \tfrac{1}{2}[\varepsilon(t) - \varepsilon(t+\delta)][|u(t+\delta)|^2 - |u(t)|^2].\end{aligned}$$

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

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

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 $s \le t$ we have

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

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

$$\phi(t) \le \phi(s) + R^2\textstyle\int_s^t \exp(E_\tau^0)|\dot{\varepsilon}(\tau)|d\tau + \int_s^t |\dot{\varepsilon}(\tau)|\phi(\tau)d\tau$$

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

$$\phi(z) \le \kappa(s,t)\exp(\textstyle\int_s^z |\dot{\varepsilon}(\tau)|d\tau) \quad \forall\, z \in [s,t].$$

In particular, for $z = t$ this gives

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

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

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

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 [26, Theorem 1.4].

Theorem 9 *Let $u(t)$ be the strong solution of (D) and assume that $\varepsilon(t) \to 0$ for $t \to \infty$ 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 nonconvergent Tikhonov-like trajectory

In this subsection we give a counter-example 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 Figure 1 below). 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 Figure 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 $p = 1+i$, equation (D) may be rewritten as

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

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

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

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

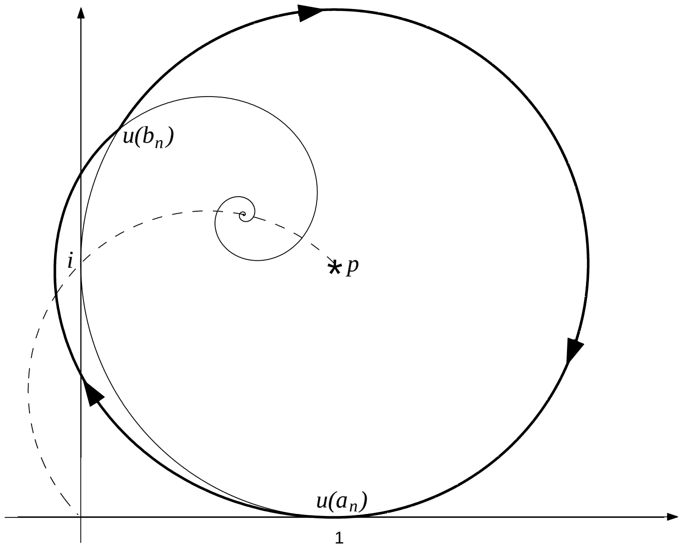

Figure 1: The trajectory $u(t)$ on the interval $[a_n, a_{n+1}]$, starting from 1 and back.

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

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

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

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

We claim that if $\varepsilon_n \le \frac{1}{2}$ 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

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

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

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

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 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, equation (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 $\eta \in \mathcal{C}^\infty(\mathbb{R}_+; \mathbb{R}_{++})$ 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 [25] show that equation (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, a natural question related to the results in [16, 24, 27] 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.

Let $\varepsilon(t)$ be the function defined in §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 [20] as well as [16] and [1]) states that

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

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

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

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 $V(t,s)x = \prod_{k=\nu(s)+1}^{\nu(t)} \left[I + \lambda_k(A + \theta_k I)\right]^{-1} x$, where $\nu(t) = \max\{k \in \mathbb{N} \mid \Lambda_k \le t\}$. Applying inequality (5) repeatedly for $B_n = A + \theta_n I$ in the appropriate subintervals one gets

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

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 $4^{k-1} < n \le 4^k$ set $\lambda_n = 4^{-k}m$. We then have $\sum_{n\ge 0} \lambda_n = \infty$, while $\sum_{n\ge 1} \sqrt{\lambda_{J_n}} \le m \sum_{n\ge 0} 2^{-n} < \infty$.

References

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

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

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

[4] Attouch H. and Damlamian A. [1978], *Strong solutions for parabolic variational inequalities*, Journal of Nonlinear Analysis 2, 329–353.

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

[6] Baillon J.-B. and Cominetti R. [2001], *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, 263–273.

[7] Brézis H. [1973], Opérateurs maximaux monotones et semigroupes de contractions dans les espaces de Hilbert, *North-Holland Mathematical Studies* 5, North-Holland, Amsterdam.

[8] Browder F.E. [1976], Nonlinear operators and nonlinear equations of evolution in Banach spaces, *Proc. Symp. Pure Math 18 (part 2), Amer. Math. Soc.*, Providence RI.

[9] Bruck R.E. [1974], *A strongly convergent iterative solution of* $0 \in U(x)$ *for a maximal monotone operator* U *in Hilbert space*, J. Math. Anal. Appl. 48, 114–126.

[10] Bruck R.E. [1975], *Asymptotic convergence of nonlinear contraction semigroups in Hilbert space*, J. Funct. Anal. 18, 15–26.

[11] Cabot A. [2004], *The steepest descent dynamical system with control. Applications to constrained minimization*, ESAIM Control, Optimisation and Calculus of Variations 10, 243–258.

[12] Combettes P.L. and Hirstoaga S.A. [2006], *Approximating curves for nonexpansive and monotone operators*, Journal of Convex Analysis 13, 633–646.

[13] Cominetti R. and Alemany O. [1999], *Steepest descent evolution equations: asymptotic behavior of solutions and rate of convergence*, Transactions of the AMS 351, 4847–4860.

[14] Crandall M.G. and Pazy A. [1972], *Nonlinear evolution equations in Banach spaces*, Israel J. of Math. 11, 57–94.

[15] Furuya H., Miyashiba K. and Kenmochi N. [1986], *Asymptotic behavior of solutions to a class of nonlinear evolution equations*, J. Diff. Equations 62, 73–94.

[16] Güler O. [1991], *On the convergence of the proximal point algorithm for convex minimization*, SIAM J. Control & Opt. 29, 403-419.

[17] Hirstoaga S. [2006], Approximation et résolution de problèmes d'équilibre, de point fixe et d'inclusion monotone, Ph.D. Thesis, UPMC Paris 6.

[18] Israel M.M. and Reich S. [1981], *Asymptotic behavior of solutions of a nonlinear evolution equation*, J. Math. Anal. and Appl. 83, 43–53.

[19] Kenmochi N. [1981], *Solvability of nonlinear evolution equations with time-dependent constraints and applications*, Bull. Fac. Ed. Chiba Univ. 30, 1–87.

[20] Kobayashi Y. [1975], *Difference approximation of Cauchy problems for quasi-dissipative operators and generation of nonlinear semigroups*, J. Math Soc. Japan 27, 640-665.

[21] Lemaire B. [2001], *Staircase parametrization in dynamical selection*, Set-Valued Analysis 9, 111–121.

[22] Minty G.J. [1963], *On a monotonicity method for the solution of nonlinear equations in Banach spaces*, Proc. Nat. Acad. Sci. USA 50, 1038–1041.

[23] Opial Z. [1967], *Weak convergence of the sequence of successive approximations for nonexpansive mappings*, Bulletin of the American Mathematical Society 73, 591–597.

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

[25] Peypouquet J. [2007], Analyse asymptotique de systèmes d'évolution et applications en optimisation, Ph.D. Thesis, UPMC Paris 6 and U. de Chile.

[26] Reich, S. [1976], *Nonlinear evolution equations and nonlinear ergodic theorems*, Nonlinear Analysis: Theory, Methods & Applications 1, 319-330.

[27] Rockafellar R.T. [1976], *Monotone operators and the proximal point algorithm*, SIAM J. Control & Opt. 14, 877-898.

[28] Tikhonov A. and Arsenin V. [1974], Méthodes de résolution de problèmes mal posés, MIR.