



HAL
open science

Weak Versus Strong Convergence of a Regularized Newton Dynamic for Maximal Monotone Operators

Hedy Attouch, Jean-Bernard Baillon

► **To cite this version:**

Hedy Attouch, Jean-Bernard Baillon. Weak Versus Strong Convergence of a Regularized Newton Dynamic for Maximal Monotone Operators. Vietnam Journal of Mathematics, 2018, 46 (1), pp.177 - 195. <10.1007/s10013-017-0267-6>. <hal-01821835>

HAL Id: hal-01821835

<https://hal.science/hal-01821835v1>

Submitted on 22 Jun 2018

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

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



HAL Authorization

WEAK VERSUS STRONG CONVERGENCE OF A REGULARIZED NEWTON DYNAMIC FOR MAXIMAL MONOTONE OPERATORS

HEDY ATTOUCH AND JEAN-BERNARD BAILLON

Abstract. In a Hilbert space \mathcal{H} , given $A : \mathcal{H} \rightrightarrows \mathcal{H}$ a general maximal monotone operator whose solution set is assumed to be non-empty, and $\lambda(\cdot)$ a time-dependent positive regularization parameter, we analyze, when $t \rightarrow +\infty$, the weak versus strong convergence properties of the trajectories of the Regularized Newton dynamic

$$(RN) \quad \begin{cases} v(t) \in A(x(t)), \\ \lambda(t)\dot{x}(t) + \dot{v}(t) + v(t) = 0. \end{cases}$$

The term $\lambda(t)\dot{x}(t)$ acts as a Levenberg-Marquardt regularization of the continuous Newton dynamic associated with A , which makes (RN) a well-posed system. The coefficient $\lambda(t)$ is allowed to tend to zero as $t \rightarrow +\infty$, which makes (RN) asymptotically close to the Newton continuous dynamic. As a striking property, when $\lambda(t)$ does not converge too rapidly to zero as $t \rightarrow +\infty$ (with $\lambda(t) = e^{-t}$ as the critical size), Attouch and Svaiter showed that each trajectory generated by (RN) converges weakly to a zero of A . By adapting Baillon's counterexample, we show a situation where A is the gradient of a smooth convex function, and there is a trajectory of the corresponding system (RN) that does not converge strongly. On the positive side, under certain particular assumptions about the operator A , or on the regularization parameter $\lambda(\cdot)$, we show the strong convergence when $t \rightarrow +\infty$ of the (RN) trajectories .

Key-words. maximal monotone operators, Newton-like continuous dynamic, Levenberg-Marquardt regularization, Baillon's counterexample, weak versus strong asymptotic convergence.

AMS subject classification. 34G25, 47J25, 47J30, 47J35, 49M15, 49M37, 65K15, 90C25, 90C53.

1. INTRODUCTION

Throughout the paper, \mathcal{H} is a real Hilbert space, and $A : \mathcal{H} \rightrightarrows \mathcal{H}$ stands for a general maximal monotone operator whose solution set $S = \{x \in \mathcal{H} : 0 \in Ax\}$ is supposed to be nonempty. We will analyze the weak versus strong asymptotic convergence properties of a Newton-like continuous dynamics attached to solving the monotone inclusion

$$(1) \quad \text{find } x \in \mathcal{H} \text{ such that } 0 \in Ax.$$

These questions, which have been intensively studied for semigroups of contractions generated by maximal monotone operators, see [8, 10, 13, 16, 29] and references therein, are seemingly new for the dynamics associated with Newton's method. Our analysis deals with the Regularized Newton dynamic (called (RN) for short) introduced by Attouch and Svaiter in [6]

$$(RN) \quad \begin{cases} v(t) \in A(x(t)), \\ \lambda(t)\dot{x}(t) + \dot{v}(t) + v(t) = 0. \end{cases}$$

Recall that, when A is a smooth operator, the classical Newton method writes $A'(x)(x_+ - x) + Ax = 0$, where A' is the derivative of A . Its continuous version gives $A'(x(t))\dot{x}(t) + Ax(t) = 0$. Equivalently

$$(N) \quad \begin{cases} v(t) = A(x(t)), \\ \dot{v}(t) + v(t) = 0. \end{cases}$$

Newton's method, however, requires a lot of regularity both in its discrete and continuous version. The introduction of the additional term $\lambda(t)\dot{x}(t)$ in (RN) acts as a Levenberg-Marquard regularization (see [22, 24]) of the ill-posed continuous Newton dynamic (N) associated to A . The parameter $\lambda(t)$ is taken positive and locally absolutely continuous (indeed it can be taken locally of bounded variation). Of particular interest is the case $\lambda(t) \rightarrow 0$ as $t \rightarrow +\infty$, which provides a (nonautonomous) dynamical system whose asymptotic behavior is similar to that of the Newton method. As particular cases of operator A , by taking $A = \partial\Phi$, $\Phi : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ convex lsc. proper, (resp. $A = (\partial_x L, -\partial_y L)$, with $L : \mathcal{H} \times \mathcal{H} \rightarrow \overline{\mathbb{R}}$ closed convex-concave), one obtains Newton-like dynamics for convex minimization, (resp. convex-concave saddle value problems).

System (RN) enjoys remarkable properties:

a) Firstly, by using the Minty representation of maximal monotone operators, it can be reformulated as a classical differential equation. This representation makes use of $J_\mu^A = (I + \mu A)^{-1}$ the resolvent of index $\mu > 0$ of A , and $A_\mu = \frac{1}{\mu}(I - J_\mu^A)$ the Yosida approximation of index $\mu > 0$ of A , see [12, 13, 34] for basic facts concerning these notions. For any $t \in [0, +\infty[$ we set $\mu(t) = \frac{1}{\lambda(t)}$, and introduce the new unknown function $z : [0, +\infty[\rightarrow H$ which is defined by

$$(2) \quad z(t) = x(t) + \mu(t)v(t).$$

One can rewrite (RN) with the help of (x, z) . One first obtains

$$(3) \quad \begin{cases} x(t) = J_{\mu(t)}^A(z(t)), \\ v(t) = A_{\mu(t)}(z(t)). \end{cases}$$

In our context, this is the Minty representation of maximal monotone operators, see [25]. The second equation of (RN) can be reformulated as a classical differential equation with respect to $z(\cdot)$, which gives

$$(4) \quad \begin{cases} x(t) = J_{\mu(t)}^A(z(t)), \\ \dot{z}(t) + (\mu(t) - \dot{\mu}(t))A_{\mu(t)}(z(t)) = 0. \end{cases}$$

As a nice feature of system (4), the operators $J_\mu^A : \mathcal{H} \rightarrow \mathcal{H}$ and $A_\mu : \mathcal{H} \rightarrow \mathcal{H}$ are Lipschitz continuous, which makes this system relevant to the Cauchy-Lipschitz theorem. Indeed, in [6, Theorem 2.4], it is proved that for any Cauchy data $v_0 \in A(x_0)$ there exists a unique strong global solution $(x(\cdot), v(\cdot)) : [0, +\infty[\rightarrow \mathcal{H} \times \mathcal{H}$ of (RN) which satisfies $x(0) = x_0, v(0) = v_0$.

b) Secondly, the trajectories of system (RN) asymptotically converge under fairly general assumptions. Following [6, Theorem 3.7], suppose that $\lambda(\cdot)$ is bounded from above on $[0, +\infty[$, and satisfies

$$(5) \quad \liminf_{t \rightarrow +\infty} \frac{\dot{\lambda}(t)}{\lambda(t)} > -1.$$

Then, for any strong global solution of (RN), as t tends to $+\infty$

- i*) $v(t) \rightarrow 0$ strongly in \mathcal{H} ,
- ii*) $x(t)$ converges weakly in \mathcal{H} to a zero of A .

The condition $\frac{\dot{\lambda}(t)}{\lambda(t)} > -1$ is equivalent to $\mu(t) - \dot{\mu}(t) > 0$, and hence preserves the monotonicity of the operator governing the differential equation in (4). Condition (5) expresses that $\lambda(t)$ does not converge too rapidly to zero as $t \rightarrow +\infty$ (with, roughly speaking, as a critical size, $\lambda(t) = e^{-t}$). Note that the trajectories of system (RN) converge weakly, which makes a sharp contrast with the semigroup associated to a general maximal monotone operator, whose trajectories are solutions of the differential inclusion

$$(6) \quad \dot{x}(t) + A(x(t)) \ni 0.$$

Let us recall some classical facts concerning (6) which will enlight our study. For a general maximal monotone operator A , weak ergodic convergence of trajectories of (6) was established in [10]. Weak convergence was established in [16] for the case of demi-positive operators. This class includes the subdifferentials of closed proper convex functions $A = \partial\Phi$, as well as operators of the form $A = I - T$ with T a contraction. As the counterexample in [8] shows, even in the case of subdifferential operators, the strong convergence property may not be satisfied.

Let us now return to the (RN) dynamic and come to the central question of this paper: We know that, for a general maximal monotone operator and regularizing parameter $\lambda(\cdot)$, each trajectory of (RN) converges weakly to a zero of A . In many physical or numerical situations, obtaining strong convergence is a desirable feature of the dynamic or the algorithm, see [11] for a general discussion on this topic. Thus it is natural to ask: in what situations strong convergence of the trajectories of (RN) holds true? This turns out to be a challenging question: In the particular case $\lambda(t) = \lambda_0 e^{-t}$, one can perform a direct integration of (RN). In that case, the asymptotic convergence of $x(t)$ can be reduced to the convergence analysis of the Tikhonov approximation, as the regularization parameter goes to zero. By a classical result, see [15, 17, 30], in this situation, one obtains strong convergence to some minimal norm element.

Thus one may conjecture that strong convergence of trajectories of (RN) holds true in general. As a central result of our analysis we give a negative answer to this claim: by an adaptation to our situation of Baillon's counterexample [8], we provide a maximal monotone operator A which is the gradient of a convex C^1 function, and a trajectory of (RN), with λ constant, which converges weakly and not strongly. One may also wonder whether the case $\lambda(t) = \lambda_0 e^{-t}$, for which strong convergence holds for any maximal monotone operator, is exceptional. We shall partially answer to this question and examine some other related ones, among which: Can one provide a counterexample to the strong convergence with a parameter $\lambda(\cdot)$ which asymptotically tends to zero? (and even more precisely, in the closed-loop case?).

Let us finally mention the rich connections between (RN), integrodifferential equations (see [19, 18, 33]), and second order dynamics with inertial features (see [3, 4, 23]), for which similar questions can be raised.

The study of the continuous dynamics (RN) opens the way to algorithmic developments concerning Newton's methods for solving monotone inclusions. See [1], [5] for results in this direction. As a distinctive feature of the algorithms obtained from the time discretization of (RN), they involve a large step (proximal) condition, as proposed in [5, 26, 27]. It is reasonable to conjecture that the results obtained in this paper have a natural counterpart in the algorithmic framework. This is an interesting question for a future study.

2. THE REGULARIZED NEWTON DYNAMIC

We successively recall basic facts concerning the (RN) dynamic: the global existence and uniqueness theorem, the asymptotic convergence theorem, and some corresponding results for the closed loop-regularization approach. Then, we examine some particular situations which enlight our study.

2.1. The open-loop regularized Newton dynamic. We consider the Regularized Newton Cauchy problem ((RNC) for short)

$$(RNC) \quad \begin{cases} v(t) \in A(x(t)), \\ \lambda(t)\dot{x}(t) + \dot{v}(t) + v(t) = 0, \\ x(0) = x_0, v(0) = v_0 \in A(x_0). \end{cases}$$

Definition 2.1. We say that a pair $(x(\cdot), v(\cdot))$ is a strong global solution of (RNC) if the following properties *i*), *ii*), *iii*) and *iv*) are satisfied:

- i*) $x(\cdot), v(\cdot) : [0, +\infty[\rightarrow \mathcal{H}$ are absolutely continuous on each bounded interval $[0, b]$, $0 < b < +\infty$;
- ii*) $v(t) \in A(x(t))$ for all $t \in [0, +\infty[$;
- iii*) $\lambda(t)\dot{x}(t) + \dot{v}(t) + v(t) = 0$ for almost all $t \in [0, +\infty[$;
- iv*) $x(0) = x_0, v(0) = v_0$.

We make the following standing Regularity assumption on function $\lambda(\cdot)$:

$$(RA) \quad \lambda : [0, +\infty[\rightarrow]0, +\infty[\text{ is absolutely continuous on each interval } [0, b], \quad 0 < b < +\infty.$$

Set $\mu(t) = \frac{1}{\lambda(t)}$. Thus $\mu(\cdot)$ is absolutely continuous on each interval $[0, b]$, $0 < b < +\infty$.

Theorem 2.2. [6, Theorem 2.4] The following properties hold:

- i*) For any Cauchy data $x(0) = x_0, v(0) = v_0 \in A(x_0)$, there exists a unique strong global solution $(x(\cdot), v(\cdot)) : [0, +\infty[\rightarrow \mathcal{H} \times \mathcal{H}$ of (RNC);
- ii*) The solution pair $(x(\cdot), v(\cdot))$ of (RNC) can be represented as: for any $t \in [0, +\infty[$,

$$(7) \quad \begin{cases} x(t) = J_{\mu(t)}^A(z(t)) \\ v(t) = A_{\mu(t)}(z(t)), \end{cases}$$

where $z(\cdot) : [0, +\infty[\rightarrow \mathcal{H}$ is the unique strong solution of the Cauchy problem

$$(8) \quad \begin{cases} \dot{z}(t) + (\mu(t) - \dot{\mu}(t))A_{\mu(t)}(z(t)) = 0, \\ z(0) = x_0 + \mu(0)v_0. \end{cases}$$

Remark 2.3. In some cases, it is worth reinforcing, resp. to weaken, the regularity assumption (RA):

- a) Suppose $\lambda(\cdot)$ Lipschitz continuous, then so are $z(\cdot), x(\cdot)$ and $v(\cdot)$.
- b) Suppose $\lambda(\cdot)$ of bounded variation on bounded sets, then Theorem 2.2 is still valid [7, Theorem 3.2].

The following result plays a central role in the asymptotical analysis.

Theorem 2.4. [6, Theorem 3.7] Suppose that $\lambda(\cdot)$ is bounded from above on $[0, +\infty[$, and

$$(9) \quad \liminf_{t \rightarrow +\infty} \frac{\dot{\lambda}(t)}{\lambda(t)} > -1.$$

Then $v(t) \rightarrow 0$ strongly, and $x(t)$ converges weakly to a zero of A , as t goes to $+\infty$.

2.2. The closed-loop regularized Newton dynamic. Given a C^1 function $\alpha :]0, +\infty[\rightarrow]0, +\infty[$, we now consider the closed-loop regularized Newton dynamic

$$(10) \quad \begin{cases} v(t) \in A(x(t)), \\ \alpha(\|v(t)\|^2) \dot{x}(t) + \dot{v}(t) + v(t) = 0, \\ x(0) = x_0, v(0) = v_0, v_0 \in A(x_0), v_0 \neq 0. \end{cases}$$

Theorem 2.5. ([7, Theorem 4.1, Theorem 5.1]) There exists a unique global strong solution $(x(\cdot), v(\cdot)) :]0, +\infty[\rightarrow \mathcal{H} \times \mathcal{H}$ to (10). Moreover

a) v satisfies

$$(11) \quad \lim_{t \rightarrow +\infty} \|v(t)\| = 0.$$

b) Let us suppose moreover that α is bounded above in a neighbourhood of 0 and

$$(12) \quad \limsup_{r \rightarrow 0^+} \frac{r\dot{\alpha}(r)}{\alpha(r)} < \frac{1}{2}.$$

Then, $x(t)$ converges weakly to a zero of A , as $t \rightarrow +\infty$.

Corollary 2.6. ([7, Corollary 5.1]) Take $\alpha(r) = r^\theta$. Then the conclusions of Theorem 2.5 hold for $\theta < \frac{1}{2}$. Equivalently, for $0 < \gamma < 1$, there is asymptotic stabilization of the system

$$(13) \quad \begin{cases} v(t) \in A(x(t)) \\ \|v(t)\|^\gamma \dot{x}(t) + \dot{v}(t) + v(t) = 0; \\ x(0) = x_0, v(0) = v_0 \in A(x_0), v_0 \neq 0. \end{cases}$$

2.3. The linear antisymmetric case. We follow [6, Section 6.3]. The anti-clockwise rotation in the plane about the origin through the angle $\frac{\pi}{2}$ provides a well-known situation where the semigroup generated by a maximal monotone operator fails to converge (it only converges in the ergodic sense). Thus, take

$$\mathcal{H} = \mathbb{R} \times \mathbb{R}, \quad A = \text{rot}\left(0, \frac{\pi}{2}\right), \quad A(x_1, x_2) = (-x_2, x_1).$$

The operator A is maximal monotone, linear, $A^* = -A$ (antisymmetric), and $\langle Ax, x \rangle = 0 \quad \forall x \in \mathcal{H}$. Moreover the origin is the unique equilibrium (zero of A). Recall that, in the linear case, convex subdifferential operators (for which convergence holds) correspond to symmetric operators. Let us describe the (RN) dynamic in this particular situation.

Setting $X(t) = x_1(t) + ix_2(t)$, (RN) is equivalent to

$$(14) \quad (\lambda(t) + i)\dot{X}(t) + iX(t) = 0.$$

Let us successively examine the different cases:

a) $\lambda > 0$ is constant. Then, direct integration of (14) gives

$$X(t) = X_0 \exp\left(-\frac{1+i\lambda}{1+\lambda^2} t\right).$$

We observe that $x(t) \rightarrow 0$ as $t \rightarrow +\infty$. This makes contrast with the ergodic convergence of the semigroup generated by A , which exhibits periodic circular orbits.

b) $\lambda(\cdot)$ is a positive function. Integration of (14) gives $\|X(t)\| = \|X(0)\| \exp\left(-\int_0^t \frac{1}{1+\lambda(s)^2} ds\right)$. Hence

$$\int_0^{+\infty} \frac{1}{1+\lambda(t)^2} dt = +\infty \Rightarrow x(t) \rightarrow 0 \text{ as } t \rightarrow +\infty.$$

The number of rotations around the origin depends on $\int_0^{+\infty} \frac{\lambda(t)}{1+\lambda(t)^2} dt$.

2.4. **The case** $\lambda(t) = \lambda_0 e^{-t}$, $\lambda_0 > 0$. For this choice of $\lambda(\cdot)$, for any $t \geq 0$

$$0 > \dot{\lambda}(t) = -\lambda(t).$$

It follows that the trajectory $x(\cdot)$ is bounded (see [6, Lemma 3.5]), and $v(t)$ converges to 0 as $t \rightarrow +\infty$. It is possible to have a closed formula for $x(t)$, $v(t)$, and to estimate how fast is the convergence of $v(t)$ to 0. Define $z(\cdot)$ by

$$z(t) = x(t) + \frac{1}{\lambda(t)}v(t) = x(t) + \frac{e^t}{\lambda_0}v(t).$$

Setting $\mu(t) = \frac{1}{\lambda(t)} = \frac{e^t}{\lambda_0}$, we have $\dot{\mu}(t) = \mu(t)$, which, by (8), implies $\dot{z}(t) = 0$ for $t \geq 0$. Hence, for $t \geq 0$,

$$x(t) + \frac{e^t}{\lambda_0}v(t) = z(0) = x_0 + \frac{1}{\lambda_0}v_0$$

which, in view of the inclusion $v(t) \in A(x(t))$ is equivalent to

$$(15) \quad x(t) = J_{e^t/\lambda_0}^A(z(0)), \quad v(t) = A_{e^t/\lambda_0}(z(0)).$$

The next proposition is a direct consequence of the above equation.

Proposition 2.7. ([6, Proposition 5.1]) Let $\lambda(t) = \lambda_0 e^{-t}$. Assume that $A^{-1}(0)$ is non-empty, and let x_0^* be the orthogonal projection of $x_0 + \lambda_0^{-1}v_0$ onto $A^{-1}(0)$. Then, the following properties are satisfied:

- i) $\forall t \geq 0$, $\|x(t) - (x_0 + \lambda_0^{-1}v_0)\| \leq \|x_0^* - (x_0 + \lambda_0^{-1}v_0)\|$;
- ii) $\forall t \geq 0$, $\|v(t)\| \leq \lambda_0 e^{-t} \|x_0^* - (x_0 + \lambda_0^{-1}v_0)\|$;
- iii) $\lim_{t \rightarrow +\infty} x(t) = x_0^*$.

Proof. To simplify the proof, set $z_0 = x_0 + \lambda_0^{-1}v_0$.

i) To prove the first inequality, take $x^* \in A^{-1}(0)$. By monotonicity of A , and $z(t) = x(t) + \frac{1}{\lambda(t)}v(t) = z_0$, we have

$$0 \leq \frac{1}{\lambda} \langle x^* - x(t), 0 - v(t) \rangle = \langle x^* - x(t), x(t) - z_0 \rangle.$$

Thus,

$$\begin{aligned} \|x^* - z_0\|^2 &= \|x^* - x(t)\|^2 + 2\langle x^* - x(t), x(t) - z_0 \rangle + \|x(t) - z_0\|^2 \\ &\geq \|x(t) - z_0\|^2. \end{aligned}$$

This being true for any $x^* \in A^{-1}(0)$, passing to the infimum with respect to x^* establishes the formula.

ii) By (15) and item i)

$$\begin{aligned} \|v(t)\| &= \|A_{e^t/\lambda_0}(z_0)\| \\ &= \lambda_0 e^{-t} \|x(t) - z_0\| \\ &\leq \lambda_0 e^{-t} \|x_0^* - (x_0 + \lambda_0^{-1}v_0)\|. \end{aligned}$$

iii) We have $x(t) = J_{e^t/\lambda_0}^A(z_0)$, which, equivalently, can be written as

$$\lambda_0 e^{-t} (x(t) - z_0) + A(x(t)) \ni 0.$$

Equivalently, by setting $y(t) := x(t) - z_0$, and $By = A(y + z_0)$ (note that B is still maximal monotone)

$$\lambda_0 e^{-t} y(t) + B(y(t)) \ni 0.$$

Noticing that $\lambda_0 e^{-t} \rightarrow 0$ as $t \rightarrow +\infty$, by using the classical asymptotic properties of the Tikhonov approximation, see for example see [15, 17, 30], we obtain

$$\lim_{t \rightarrow +\infty} y(t) = \text{proj}_{B^{-1}(0)} 0.$$

Since $B^{-1}(0) = -z_0 + A^{-1}(0)$, we finally obtain

$$\begin{aligned} \lim_{t \rightarrow +\infty} x(t) &= z_0 + \text{proj}_{-z_0 + A^{-1}(0)} 0 \\ &= \text{proj}_{A^{-1}(0)} z_0 = x_0^*. \end{aligned}$$

□

Note that $\|v(t)\| \leq c e^{-t}$, which, as an asymptotical behavior, is almost as good as the “pure” Newton’s continuous dynamic.

3. A SITUATION WHERE THERE IS NO STRONG CONVERGENCE

Our approach is based on the adaptation to the (RN) dynamic of the counterexample of Baillon. In a different context, a similar strategy was developed in [20] to provide a sequence generated by the proximal algorithm that is weakly and not strongly convergent.

Let us recall Baillon’s result:

Theorem 3.1. ([8, Proposition 1]) There exists a closed convex proper function $\Phi : \mathcal{H} = l^2(\mathbb{N}) \rightarrow \mathbb{R}^+ \cup \{+\infty\}$, with $\partial\Phi^{-1}(0) \neq \emptyset$, such that the semigroup $S(t)$ generated by the maximal monotone operator $A = \partial\Phi$ satisfies the following property: there exists some $a \in \overline{\text{dom}\Phi}$ such that $S(t)a$ does not converge strongly to an element of $\partial\Phi^{-1}(0)$.

Bruck’s theorem [16] states that, for all $x \in \overline{\text{dom}\Phi}$, $S(t)x$ weakly converges to an element of $\partial\Phi^{-1}(0)$. Thus Baillon’s counterexample is a constructive example of a closed convex proper $\Phi : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$, and of a trajectory of the semigroup generated by $\partial\Phi$ which converges weakly and not strongly. Baillon’s thesis [9] contains an extended version of [8], with a counterexample involving a convex function Φ of class \mathcal{C}^1 . For our purpose, we need a slightly more precise statement, with $\Phi \in \mathcal{C}^{1,1}$ being taken equal to the Moreau envelope of some closed convex function. We recall the definition of the Moreau envelope (see [12] for further details): Given a closed convex proper function $\Psi : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$, its Moreau envelope of index $\lambda > 0$ is the function $\Psi_\lambda : \mathcal{H} \rightarrow \mathbb{R}$ which is defined by, for all $x \in \mathcal{H}$

$$\Psi_\lambda(x) = \inf_{\xi \in \mathcal{H}} \left\{ \Psi(\xi) + \frac{1}{2\lambda} \|x - \xi\|^2 \right\}.$$

We use the following result from [9], whose proof is reproduced here, for the convenience of the reader. It compares the asymptotic behavior of the trajectories generated respectively by $\partial\Phi$ and by its Yosida approximation $(\partial\Phi)_\lambda$.

Lemma 3.2. In a Hilbert space \mathcal{H} , let us give a closed convex proper function $\Phi : \mathcal{H} \rightarrow \mathbb{R}^+ \cup \{+\infty\}$, $\partial\Phi$ its subdifferential operator, and for any $\lambda > 0$, $(\partial\Phi)_\lambda = \nabla\Phi_\lambda$ its Yosida approximation. Let us denote by $S(t)$ and $S_\lambda(t)$ the semigroups generated respectively by $\partial\Phi$ and $(\partial\Phi)_\lambda$. Then, for any $x \in \overline{\text{dom}\Phi}$

$$\lim_{\lambda \rightarrow 0} \lim_{t \rightarrow +\infty} \|S_\lambda(t)x - S(t)x\| = 0.$$

Proof. Set $u(t) = S(t)x$ and $u_\lambda(t) = S_\lambda(t)x$. Without ambiguity, we write briefly J_λ for the resolvent of index $\lambda > 0$ of the subdifferential of Φ . From $-\dot{u}(t) \in \partial\Phi(u(t))$, and $-\dot{u}_\lambda(t) = \nabla\Phi_\lambda(u_\lambda(t)) \in \partial\Phi(J_\lambda(u_\lambda(t)))$, by monotonicity of the operator $\partial\Phi$, we obtain

$$\langle -\dot{u}_\lambda(t) + \dot{u}(t), J_\lambda(u_\lambda(t)) - u(t) \rangle \geq 0.$$

Since

$$-\dot{u}_\lambda(t) = \nabla\Phi_\lambda(u_\lambda(t)) = \frac{1}{\lambda} (u_\lambda(t) - J_\lambda(u_\lambda(t)))$$

we equivalently have

$$\langle -\dot{u}_\lambda(t) + \dot{u}(t), \lambda\dot{u}_\lambda(t) + u_\lambda(t) - u(t) \rangle \geq 0.$$

Hence

$$(16) \quad \frac{1}{2} \frac{d}{dt} \|u_\lambda(t) - u(t)\|^2 \leq \lambda \langle \dot{u}_\lambda(t), \dot{u}(t) \rangle - \lambda \|\dot{u}_\lambda(t)\|^2 \leq \frac{\lambda}{4} \|\dot{u}(t)\|^2.$$

On the other hand, by the generalized derivation chain rule (see [13, Theorem 3.2])

$$(17) \quad \frac{d}{dt} \Phi(u(t)) = -\|\dot{u}(t)\|^2.$$

By combining (16) and (17) we obtain

$$(18) \quad \frac{d}{dt} \left(\frac{1}{2} \|u_\lambda(t) - u(t)\|^2 + \frac{\lambda}{4} \Phi(u(t)) \right) \leq 0.$$

As a consequence,

$$(19) \quad t \mapsto \left(\|u_\lambda(t) - u(t)\|^2 + \frac{\lambda}{2} \Phi(u(t)) \right)$$

is a nonincreasing function, which implies that, for all $s \geq 0$

$$(20) \quad \lim_{t \rightarrow +\infty} \left(\|u_\lambda(t) - u(t)\|^2 + \frac{\lambda}{2} \Phi(u(t)) \right) \leq \|u_\lambda(s) - u(s)\|^2 + \frac{\lambda}{2} \Phi(u(s)).$$

On the other hand, by (17), $t \mapsto \Phi(u(t))$ is a nonincreasing function, and Φ has been assumed to be nonnegative. Hence $\Phi(u(t))$ converges to a finite value, as well as $\|u_\lambda(t) - u(t)\|$, which allows us to rewrite (20) as follows:

$$(21) \quad \lim_{t \rightarrow +\infty} \|u_\lambda(t) - u(t)\|^2 + \frac{\lambda}{2} \lim_{t \rightarrow +\infty} \Phi(u(t)) \leq \|u_\lambda(s) - u(s)\|^2 + \frac{\lambda}{2} \Phi(u(s)).$$

Since $\Phi(u(t))$ is nonnegative we deduce that for all $s \geq 0$

$$(22) \quad \lim_{t \rightarrow +\infty} \|u_\lambda(t) - u(t)\|^2 \leq \|u_\lambda(s) - u(s)\|^2 + \frac{\lambda}{2} \Phi(u(s)),$$

and hence

$$(23) \quad \lim_{t \rightarrow +\infty} \|u_\lambda(t) - u(t)\| \leq \|u_\lambda(s) - u(s)\| + \left(\frac{\lambda}{2} \Phi(u(s)) \right)^{\frac{1}{2}}.$$

Since $u_\lambda(s)$ converges to $u(s)$ as λ goes to zero (see [13, Theorem 3.2]), and $\Phi(u(s))$ is finitely valued for $s > 0$, we obtain

$$(24) \quad \limsup_{\lambda \rightarrow 0} \lim_{t \rightarrow +\infty} \|u_\lambda(t) - u(t)\| \leq 0$$

which gives our claim

$$(25) \quad \lim_{\lambda \rightarrow 0} \lim_{t \rightarrow +\infty} \|u_\lambda(t) - u(t)\| = 0.$$

□

We can now state the following reinforced version of Baillon's counterexample:

Proposition 3.3. There exists a closed convex proper function $\Phi : \mathcal{H} = l^2(\mathbb{N}) \rightarrow \mathbb{R}^+ \cup \{+\infty\}$, with $S = \partial\Phi^{-1}(0) \neq \emptyset$, such that the semigroup $S_\lambda(t)$ generated by the Yosida approximation $A_\lambda = \nabla\Phi_\lambda$ of the maximal monotone operator $A = \partial\Phi$ satisfies the following property: there exists some $a \in \text{dom}\partial\Phi$, and $\lambda_0 > 0$, such that for any $0 < \lambda < \lambda_0$, $S_\lambda(t)a$ does not converge strongly to an element of $\partial\Phi^{-1}(0)$.

Proof. Let $\Phi : \mathcal{H} \rightarrow \mathbb{R}^+ \cup \{+\infty\}$, and $a \in \overline{\text{dom}\Phi}$ provided by Baillon's counterexample. Thus $S(t)a$ does not converge strongly. Set

$$S(t)a \rightharpoonup \xi \text{ weakly as } t \rightarrow +\infty.$$

Let us show the existence of some $\lambda_0 > 0$ such that, for any $0 < \lambda < \lambda_0$, $S_\lambda(t)a$ does not converge strongly. Let us argue by contradiction. Thus there exists a sequence (λ_n) with $\lambda_n \rightarrow 0$ such that $S_{\lambda_n}(t)a$ converge strongly. Set

$$(26) \quad S_{\lambda_n}(t)a \rightarrow \xi_n \text{ strongly as } t \rightarrow +\infty.$$

By the lower semicontinuity property of the norm for the weak topology

$$\|\xi_n - \xi\| \leq \liminf_{t \rightarrow +\infty} \|S_{\lambda_n}(t)a - S(t)a\|.$$

By Lemma 3.2

$$\lim_{\lambda \rightarrow 0} \lim_{t \rightarrow +\infty} \|S_\lambda(t)a - S(t)a\| = 0.$$

By combining the two last equations we obtain

$$(27) \quad \lim_{n \rightarrow +\infty} \|\xi_n - \xi\| = 0.$$

By the triangle inequality

$$\|S(t)a - \xi\| \leq \|S(t)a - S_{\lambda_n}(t)a\| + \|S_{\lambda_n}(t)a - \xi_n\| + \|\xi_n - \xi\|.$$

Let $t \rightarrow \infty$ in the above inequality. By using the strong convergence property (26), we obtain

$$\limsup_{t \rightarrow +\infty} \|S(t)a - \xi\| \leq \lim_{t \rightarrow +\infty} \|S(t)a - S_{\lambda_n}(t)a\| + \|\xi_n - \xi\|.$$

This being true for any $n \in \mathbb{N}$, by letting $n \rightarrow +\infty$, and by using (27)

$$\limsup_{t \rightarrow +\infty} \|S(t)a - \xi\| \leq \lim_{n \rightarrow +\infty} \lim_{t \rightarrow +\infty} \|S(t)a - S_{\lambda_n}(t)a\|.$$

Using again Lemma 3.2, we deduce that

$$\lim_{t \rightarrow +\infty} \|S(t)a - \xi\| = 0$$

a clear contradiction with the weak but not strong asymptotic convergence of $S(t)a$.

Passing from $a \in \overline{\text{dom}\Phi}$ to $a \in \text{dom}\partial\Phi$ is immediate. One just need to take an arbitrary $t_0 > 0$, and as a new initial point $S(t_0)a$, which, by the regularizing effect, belongs to $\text{dom}\partial\Phi$. \square

We can now provide an example where the regularized Newton dynamic has a trajectory which converges weakly but not strongly.

Theorem 3.4. There exists a closed convex proper function $\Phi : \mathcal{H} = l^2(\mathbb{N}) \rightarrow \mathbb{R}^+ \cup \{+\infty\}$, with $\partial\Phi^{-1}(0) \neq \emptyset$, and $\lambda_0 > 0$, such that for any $\lambda > \lambda_0$, there exists $(x_0, v_0) \in \partial\Phi$ such that the solution trajectory $(x(\cdot), v(\cdot))$ of

$$(28) \quad \begin{cases} v(t) \in \partial\Phi(x(t)) \\ \lambda x(t) + \dot{v}(t) + v(t) = 0; \\ x(0) = x_0, v(0) = v_0 \in \partial\Phi(x_0) \end{cases}$$

verifies: $x(t)$ converges weakly but not strongly to an element of $\partial\Phi^{-1}(0)$.

Proof. Let us specialize Theorem 2.2 to the particular case $\lambda(\cdot)$ constant. Setting $\mu = \frac{1}{\lambda}$ the solution pair $(x(\cdot), v(\cdot))$ of (RNC) can be represented as: for any $t \in [0, +\infty)$,

$$(29) \quad \begin{cases} x(t) = J_\mu^A(z(t)) \\ v(t) = A_\mu(z(t)), \end{cases}$$

where J_μ^A and A_μ denote respectively the resolvent and the Yosida approximation of the maximal monotone operator A (we will specify it later), and $z(\cdot) : [0, +\infty[\rightarrow \mathcal{H}$ is the unique strong solution of the Cauchy problem

$$(30) \quad \begin{cases} \dot{z}(t) + \mu A_\mu(z(t)) = 0, \\ z(0) = x_0 + \mu v_0. \end{cases}$$

The change of time scale

$$w(t) := z\left(\frac{t}{\mu}\right)$$

in (30) gives

$$(31) \quad \begin{cases} \dot{w}(t) + A_\mu(w(t)) = 0, \\ w(0) = x_0 + \mu v_0. \end{cases}$$

Take $A = \partial\Phi$, $a \in \text{dom}\partial\Phi$, and $\mu_0 > 0$ provided by Proposition 3.3. Thus $\Phi : \mathcal{H} = l^2(\mathbb{N}) \rightarrow \mathbb{R}^+ \cup \{+\infty\}$ is a closed convex proper function with $S = \partial\Phi^{-1}(0) \neq \emptyset$ such that, for any $0 < \mu < \mu_0$, $S_\mu(t)a$ does not converge strongly to an element of $\partial\Phi^{-1}(0)$. Fix $0 < \mu < \mu_0$. By Minty's theorem $R(I + \mu A) = \mathcal{H}$, which gives the existence of $(x_{0\mu}, v_{0\mu}) \in A$ such that $x_{0\mu} + \mu v_{0\mu} = a$. Let us notice that

$$w(t) = S_\mu(t)a.$$

Hence $w(t)$ does not converge strongly as $t \rightarrow +\infty$. Clearly the same property holds for $z(t) = w(\mu t)$. On the other hand, the solution $z(\cdot)$ of the classical differential equation (30) satisfies

$$(32) \quad \dot{z}(t) \rightarrow 0 \text{ strongly as } t \rightarrow +\infty.$$

This is a classical property of the semigroup generated by the gradient of a \mathcal{C}^1 convex, minorized function. Indeed, this is a direct consequence of the energy estimate $\int_0^\infty \|\dot{z}(t)\|^2 dt < +\infty$ (which itself is a consequence of (17)), and of the nonincreasing property of $t \mapsto \|\dot{z}(t)\|$ (valid for any semigroup of contractions). Noticing that (30) can be equivalently written as

$$(33) \quad \dot{z}(t) + z(t) - J_\mu(z(t)) = 0$$

we deduce from (32) and (33) that

$$(34) \quad z(t) - J_\mu(z(t)) \rightarrow 0 \text{ strongly as } t \rightarrow +\infty.$$

Since $z(t)$ does not converge strongly as $t \rightarrow +\infty$, we deduce that

$$(35) \quad x(t) = J_\mu(z(t)) \text{ converges weakly and not strongly as } t \rightarrow +\infty.$$

This argument is valid for any $\mu = \frac{1}{\lambda} < \mu_0$, and hence for any $\lambda > \lambda_0 =: \frac{1}{\mu_0}$, which completes the proof. \square

Remark 3.5. a) The argument of Theorem 3.4 can be easily adapted so as to provide a counterexample with a \mathcal{C}^1 convex function.

b) The counterexample to the strong convergence of (RN) has been obtained with $\lambda(\cdot)$ constant. For numerical purpose it would be of great interest to analyze the weak versus strong convergence properties in the case where $\lambda(t) \rightarrow 0$ as $t \rightarrow +\infty$. Similarly, the closed-loop case should be analyzed from this perspective.

4. ON THE STRONG CONVERGENCE PROPERTY

4.1. $\lambda(t) = \lambda_0 e^{-t}$ as a particular case. In section 2.4, in the case $\lambda(t) = \lambda_0 e^{-t}$, we showed that for any maximal monotone operator A , strong convergence is satisfied by any trajectory generated by (RN). Thus it is a natural question to ask whether this result holds true only for this particular choice of $\lambda(\cdot)$, or if there is some latitude in the choice of $\lambda(\cdot)$. We will give some answers to this question.

In this section we assume that $\lambda(\cdot)$ is a positive function which satisfies a) and b):

$$a) \lambda(t) \rightarrow 0 \text{ as } t \rightarrow +\infty, \text{ and hence } \mu(t) = \frac{1}{\lambda(t)} \rightarrow +\infty \text{ as } t \rightarrow +\infty.$$

b) $\lambda(t) + \dot{\lambda}(t) \geq 0$ for all $t \geq 0$, and hence $\mu(t) - \dot{\mu}(t) \geq 0$ for all $t \geq 0$.

Let us recall that the solution of (RN) verifies

$$(36) \quad x(t) = J_{\mu(t)}^A(z(t))$$

where J_{μ}^A is Yosida approximation of index $\mu > 0$ of the maximal monotone operator A , and $z(\cdot) : [0, +\infty[\rightarrow \mathcal{H}$ is the unique strong solution of the Cauchy problem

$$(37) \quad \begin{cases} \dot{z}(t) + (\mu(t) - \dot{\mu}(t))A_{\mu(t)}(z(t)) = 0, \\ z(0) = x_0 + \mu(0)v_0. \end{cases}$$

Note first that it suffices to prove that $z(t)$ converges strongly, as shown in the following lemma.

Lemma 4.1. Suppose that $z(t) \rightarrow z_{\infty}$ strongly as $t \rightarrow +\infty$. Then $x(t) \rightarrow x^*$ strongly as $t \rightarrow +\infty$, where x^* is the orthogonal projection of z_{∞} onto $A^{-1}(0)$.

Proof. By (36), and the contraction property of the resolvent operator $J_{\mu(t)}$ we have

$$(38) \quad \|x(t) - J_{\mu(t)}z_{\infty}\| = \|J_{\mu(t)}z(t) - J_{\mu(t)}z_{\infty}\| \leq \|z(t) - z_{\infty}\|.$$

By the same argument as in section 2.4 (see also [12, Theorem 23.47]), we have

$$(39) \quad \lim_{t \rightarrow +\infty} J_{\mu(t)}z_{\infty} = \text{proj}_{A^{-1}(0)}z_{\infty} = x^*.$$

The conclusion follows from (38), (39), and the assumption $z(t) \rightarrow z_{\infty}$ strongly as $t \rightarrow +\infty$. \square

Theorem 4.2. Set $\psi(t) = \frac{\dot{\mu}(t)}{\mu(t)}$. Let us assume that the two following conditions hold ¹, where M is a nonnegative real number:

$$(40) \quad 1 - \frac{M}{t} \leq \psi(t) \leq 1$$

$$(41) \quad \limsup_{s, t \rightarrow +\infty} \left(\psi(t) - \psi(s) + \int_s^t (d\psi)^-(\tau) \right) (t - s) \leq 0.$$

Then, for any maximal monotone operator A and any trajectory generated by (RN), $x(t)$ converges strongly to an element of $A^{-1}(0)$.

Proof. Let us fix some $\xi \in A^{-1}(0)$. We thus have $(\mu(t) - \dot{\mu}(t))A_{\mu(t)}\xi = 0$. On the other hand by (37) $-\dot{z}(t) = (\mu(t) - \dot{\mu}(t))A_{\mu(t)}(z(t))$. Using the monotonicity property of $A_{\mu(t)}$ and the nonnegativity of $\mu(t) - \dot{\mu}(t)$ we obtain

$$(42) \quad \langle -\dot{z}(t), z(t) - \xi \rangle \geq 0,$$

which expresses that

$$(43) \quad \|z(t) - \xi\| \downarrow \quad (\text{nonincreasing function}).$$

Let us further analyze the differential equation which is satisfied by $z(\cdot)$. We have

$$(44) \quad \dot{z}(t) + \frac{\mu(t) - \dot{\mu}(t)}{\mu(t)}(z(t) - \xi) = \frac{\mu(t) - \dot{\mu}(t)}{\mu(t)}(J_{\mu(t)}z(t) - \xi).$$

Taking the square norm of the two members of this equality yields

$$(45) \quad \|\dot{z}(t)\|^2 + 2\frac{\mu(t) - \dot{\mu}(t)}{\mu(t)} \langle \dot{z}(t), z(t) - \xi \rangle + \left(\frac{\mu(t) - \dot{\mu}(t)}{\mu(t)} \right)^2 \|z(t) - \xi\|^2 = \left(\frac{\mu(t) - \dot{\mu}(t)}{\mu(t)} \right)^2 \|J_{\mu(t)}z(t) - \xi\|^2.$$

On the other hand, since $J_{\mu(t)}$ is a contraction and $J_{\mu(t)}\xi = \xi$, we have

$$(46) \quad \|J_{\mu(t)}z(t) - \xi\|^2 = \|J_{\mu(t)}z(t) - J_{\mu(t)}\xi\|^2 \leq \|z(t) - \xi\|^2.$$

¹when ψ is continuously differentiable, $(d\psi)^-$ is the negative part of the measure with density the derivative of ψ

Combining (45) and (46) we obtain

$$(47) \quad \|\dot{z}(t)\|^2 + 2\frac{\mu(t) - \dot{\mu}(t)}{\mu(t)} \langle \dot{z}(t), z(t) - \xi \rangle + \left(\frac{\mu(t) - \dot{\mu}(t)}{\mu(t)}\right)^2 \|z(t) - \xi\|^2 \leq \left(\frac{\mu(t) - \dot{\mu}(t)}{\mu(t)}\right)^2 \|z(t) - \xi\|^2,$$

which, after simplification, gives

$$(48) \quad \|\dot{z}(t)\|^2 + (1 - \psi(t)) \frac{d}{dt} (\|z(t) - \xi\|^2) \leq 0.$$

For any $0 < s \leq t < +\infty$, by integration between s and t of the above inequality, and using Cauchy-Schwarz inequality, we deduce that

$$(49) \quad \|z(t) - z(s)\|^2 \leq (t - s) \int_s^t \|\dot{z}(\tau)\|^2 d\tau$$

$$(50) \quad \leq (t - s) \int_s^t (\psi(\tau) - 1) \frac{d}{d\tau} (\|z(\tau) - \xi\|^2) d\tau.$$

After integration by part of this last expression we obtain

(51)

$$(52) \quad \begin{aligned} \|z(t) - z(s)\|^2 &\leq (t - s) \left((\psi(t) - 1) \|z(t) - \xi\|^2 - (\psi(s) - 1) \|z(s) - \xi\|^2 - \int_s^t \frac{d\psi}{d\tau}(\tau) \|z(\tau) - \xi\|^2 d\tau \right) \\ &\leq (t - s) \left((\psi(t) - 1) \|z(t) - \xi\|^2 - (\psi(s) - 1) \|z(s) - \xi\|^2 + \int_s^t \left(\frac{d\psi}{d\tau}\right)^-(\tau) \|z(\tau) - \xi\|^2 d\tau \right). \end{aligned}$$

Since $\tau \mapsto \|z(\tau) - \xi\|$ is nonincreasing, we deduce that

$$(53) \quad \|z(t) - z(s)\|^2 \leq (t - s) \left((\psi(t) - 1) \|z(t) - \xi\|^2 - (\psi(s) - 1) \|z(s) - \xi\|^2 + \|z(s) - \xi\|^2 \int_s^t (d\psi)^- \right).$$

Equivalently

$$\begin{aligned} \|z(t) - z(s)\|^2 &\leq (t - s) (1 - \psi(t)) (\|z(s) - \xi\|^2 - \|z(t) - \xi\|^2) \\ &\quad + \|z(s) - \xi\|^2 (t - s) \left(\psi(t) - \psi(s) + \int_s^t (d\psi)^- \right) \\ &\leq t(1 - \psi(t)) (\|z(s) - \xi\|^2 - \|z(t) - \xi\|^2) + \|z(s) - \xi\|^2 (t - s) \left(\psi(t) - \psi(s) + \int_s^t (d\psi)^- \right). \end{aligned}$$

Using assumption (40) on ψ , we obtain

$$(54) \quad \|z(t) - z(s)\|^2 \leq M (\|z(s) - \xi\|^2 - \|z(t) - \xi\|^2) + \|z(s) - \xi\|^2 (t - s) \left(\psi(t) - \psi(s) + \int_s^t (d\psi)^- \right).$$

By using the convergence property of $t \mapsto \|z(t) - \xi\|$, and assumption (41), we deduce that $z(t)$ satisfies the Cauchy property (as $t \rightarrow +\infty$), which implies the strong convergence of $z(t)$. \square

Remark 4.3. The assumption (40), $1 - \frac{M}{t} \leq \psi(t) = \frac{\dot{\mu}(t)}{\mu(t)} \leq 1$, implies that $\lim_{t \rightarrow +\infty} \psi(t) = 1$, we means that asymptotically $\mu(t)$ behaves like $\mu_0 e^t$, for some $\mu_0 > 0$.

The second assumption (41) $\limsup_{s, t \rightarrow +\infty} \left(\psi(t) - \psi(s) + \int_s^t (d\psi)^-(\tau) \right) (t - s) \leq 0$ is even more stringent. Assuming $\psi(\cdot)$ to be increasing, it is equivalent to

$$\lim_{s, t \rightarrow +\infty, s < t} (\psi(t) - \psi(s)) (t - s) = 0,$$

which itself is equivalent to

$$\forall \epsilon > 0 \exists s_\epsilon : (\psi(t) - \psi(s_\epsilon)) (t - s_\epsilon) \leq \epsilon \quad \forall t \geq s_\epsilon.$$

Hence, the function $t \mapsto (\psi(t) - \psi(s_\epsilon))(t - s_\epsilon)$ remains bounded as $t \rightarrow +\infty$. Since $\lim_{t \rightarrow +\infty} \psi(t) = 1$, this forces $\psi(s_\epsilon) = 1$. Hence, $\psi(t) = 1$ for t large enough, which implies $\mu(t) = \mu_0 e^t$, for some $\mu_0 > 0$.

4.2. Strongly monotone operators. Let us recall that a maximal monotone operator A is said to be strongly monotone, if there exists some $\alpha > 0$ such that for any $(x, v) \in A$, $(y, w) \in A$

$$\langle w - v, y - x \rangle \geq \alpha \|y - x\|^2.$$

Proposition 4.4. (see also [6, Theorem 3.14]) Suppose that A is a maximal monotone operator which is strongly monotone, and that $\lambda(\cdot)$ is a nonincreasing function which satisfies: for almost all $t > 0$

$$\lambda(t) + \dot{\lambda}(t) \geq 0.$$

Then for any trajectory $(x(\cdot), v(\cdot))$ of the (RN) dynamic,

- a) $x(t)$ converges strongly to the unique zero of A ,
- b) $v(t)$ converges strongly to 0.

Proof. By [6, Corollary 3.6], $v(t) \rightarrow 0$ as $t \rightarrow +\infty$ (this is a consequence of the energy estimate $v \in L^2(0, +\infty; \mathcal{H})$, and of the decreasing property of $\|v(t)\|$). Let \bar{x} be the unique zero of A . From $0 \in A(\bar{x})$, $v(t) \in A(x(t))$ and the strong monotonicity of A , we have

$$\langle v(t), x(t) - \bar{x} \rangle \geq \alpha \|x(t) - \bar{x}\|^2.$$

By Cauchy-Schwarz inequality, we deduce that

$$\|x(t) - \bar{x}\| \leq \frac{1}{\alpha} \|v(t)\|.$$

Since $v(t) \rightarrow 0$, we obtain $\|x(t) - \bar{x}\| \rightarrow 0$ as $t \rightarrow +\infty$. □

4.3. Subdifferential of a convex even function.

Proposition 4.5. Suppose that $A = \partial\Phi$ is the subdifferential of a closed convex even function $\Phi : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$. Suppose that $\lambda(\cdot)$ is constant. Then for any trajectory $(x(\cdot), v(\cdot))$ of the (RN) dynamic,

- a) $x(t)$ converges strongly to a zero of A ,
- b) $v(t)$ converges strongly to 0.

Proof. Following Lemma 4.1 it suffices to prove that $z(t)$ converges strongly where $z(\cdot) : [0, +\infty[\rightarrow \mathcal{H}$ is a solution of the differential equation

$$(55) \quad \dot{z}(t) + \mu A_\mu(z(t)) = 0.$$

As a key property note that, if $\Phi : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ is an even function, then so is Φ_μ . Following [16, Theorem 5] we have that that $z(t)$ converges strongly, which proves our claim. □

4.4. Linear monotone operator. In the linear case, we can perform explicit calculations, so as to obtain an accurate view of the asymptotic behavior of the trajectories of (RN). Take $\mathcal{H} = \ell^2(\mathbb{N})$, and denote by $x = (x_k)$ a generic element of \mathcal{H} . Let us give (a_k) a sequence of positive numbers which is bounded from above, i.e., there exists some $M > 0$ such that for all $k \in \mathbb{N}$

$$(56) \quad 0 < a_k \leq M.$$

Let us define the operator $A : \mathcal{H} \rightarrow \mathcal{H}$ by

$$(57) \quad Ax = (a_k x_k).$$

Clearly, A is a linear continuous monotone operator on \mathcal{H} , and hence is a maximal monotone operator. Let us consider the (RN) equation with $\lambda(\cdot)$ a positive locally absolutely continuous function which is just assumed to be bounded from above, i.e.,

$$(58) \quad 0 < \lambda(t) \leq \Lambda$$

for some $\Lambda > 0$. Set $x(t) = (x_k(t))$ and $x(0) = (x_{0k})$. For each $k \in \mathbb{N}$, we have

$$(59) \quad (\lambda(t) + a_k)\dot{x}_k(t) + a_k x_k(t) = 0.$$

By integration of (59) we obtain

$$(60) \quad x_k(t) = x_{0k} e^{-\int_0^t \frac{a_k}{a_k + \lambda(\tau)} d\tau}.$$

By (58), for each $k \in \mathbb{N}$, and all $\tau \geq 0$

$$(61) \quad \frac{a_k}{a_k + \lambda(\tau)} \geq \frac{a_k}{a_k + \Lambda}$$

which implies

$$(62) \quad |x_k(t)| \leq |x_{0k}| e^{-t \frac{a_k}{a_k + \Lambda}}.$$

Since $a_k > 0$, we obtain $x_k(t) \rightarrow 0$ as $t \rightarrow +\infty$, which gives the weak convergence $x(t) \rightharpoonup 0$.

Let us now consider the strong convergence property. We have

$$(63) \quad \|x(t)\|^2 = \sum_k |x_{0k}|^2 e^{-2 \int_0^t \frac{a_k}{a_k + \lambda(\tau)} d\tau}.$$

In order to obtain that $\|x(t)\| \rightarrow 0$, let us apply the Lebesgue dominated convergence theorem: On the one hand, by (61), for each $k \in \mathbb{N}$

$$(64) \quad \lim_{t \rightarrow +\infty} \int_0^t \frac{a_k}{a_k + \lambda(\tau)} d\tau = \int_0^{+\infty} \frac{a_k}{a_k + \lambda(\tau)} d\tau = +\infty$$

$$(65) \quad \geq \int_0^{+\infty} \frac{a_k}{a_k + \Lambda} d\tau = +\infty.$$

Hence for each $k \in \mathbb{N}$

$$(66) \quad \lim_{t \rightarrow +\infty} |x_{0k}|^2 e^{-2 \int_0^t \frac{a_k}{a_k + \lambda(\tau)} d\tau} = 0.$$

On the other hand

$$(67) \quad |x_{0k}|^2 e^{-2 \int_0^t \frac{a_k}{a_k + \lambda(\tau)} d\tau} \leq |x_{0k}|^2.$$

Since $\sum_k |x_{0k}|^2 < +\infty$, by the dominated convergence theorem we obtain that $x(t)$ converges strongly to 0, which is the unique zero of A .

4.5. A Tikhonov-like result. Let us make the following assumptions on $\lambda(\cdot)$:

- i) $\lambda(t) = e^{-t}(1 + \epsilon(t))$.
- ii) $\dot{\epsilon}(t) \geq 0$.
- iii) $\epsilon(t) \uparrow 0$ as $t \rightarrow +\infty$.

Note that under these assumptions

$$(68) \quad \lambda(t) + \dot{\lambda}(t) = e^{-t} \dot{\epsilon}(t) \geq 0.$$

As a model example of function $\lambda(\cdot)$ verifying i) ii) iii) one can consider

$$(69) \quad \lambda(t) = e^{-t}(1 - e^{-2t})$$

with $\epsilon(t) = -e^{-2t}$.

Given A a general maximal monotone operator, let us consider a trajectory of the (RNC) dynamic which is associated to A and $\lambda(\cdot)$

$$(70) \quad (\text{RNC}) \quad \begin{cases} v(t) \in A(x(t)), \\ \lambda(t)\dot{x}(t) + \dot{v}(t) + v(t) = 0, \\ x(0) = x_0, v(0) = v_0 \in A(x_0). \end{cases}$$

After multiplication by e^t of the differential equation in (70) we obtain

$$(71) \quad \lambda(t)e^t\dot{x}(t) + e^t(\dot{v}(t) + v(t)) = 0,$$

that is

$$(72) \quad \lambda(t)e^t\dot{x}(t) + \frac{d}{dt}(e^tv)(t) = 0.$$

By integration between 0 and t

$$(73) \quad \int_0^t \lambda(\tau)e^\tau\dot{x}(\tau)d\tau + e^tv(t) = v_0.$$

Let us integrate by part the integral term in (73)

$$(74) \quad \lambda(t)e^tx(t) - \int_0^t e^\tau(\lambda(\tau) + \dot{\lambda}(\tau))x(\tau)d\tau + e^tv(t) = v_0 + \lambda(0)x_0.$$

Then use the particular form of $\lambda(\cdot)$ and the corresponding relation (68) to obtain

$$(75) \quad (1 + \epsilon(t))x(t) - \int_0^t \dot{\epsilon}(\tau)x(\tau)d\tau + e^tv(t) = v_0 + \lambda(0)x_0.$$

Set $z_0 = v_0 + \lambda(0)x_0$ and $\nu(t) = \frac{e^t}{1+\epsilon(t)}$. After division of (75) by $1 + \epsilon(t)$ we obtain

$$(76) \quad x(t) + \nu(t)v(t) = \frac{1}{1 + \epsilon(t)}z_0 + \frac{1}{1 + \epsilon(t)} \int_0^t \dot{\epsilon}(\tau)x(\tau)d\tau,$$

i.e.,

$$(77) \quad x(t) = J_{\nu(t)}^A \left(\frac{1}{1 + \epsilon(t)}z_0 + \frac{1}{1 + \epsilon(t)} \int_0^t \dot{\epsilon}(\tau)x(\tau)d\tau \right).$$

Thus we have

$$(78) \quad x(t) = J_{\nu(t)}^A \eta(t)$$

with

$$(79) \quad \eta(t) := \frac{1}{1 + \epsilon(t)}z_0 + \frac{1}{1 + \epsilon(t)} \int_0^t \dot{\epsilon}(\tau)x(\tau)d\tau.$$

Let us examine the asymptotic behavior of $\eta(t)$. By (68) we have $\lambda(t) + \dot{\lambda}(t) \geq 0$. By using [6, Lemma 3.5] this implies that the trajectory $x(\cdot)$ is bounded. Set $M = \sup_t \|x(t)\|$.

Since $\dot{\epsilon}(t) \geq 0$ and $\epsilon(t) \uparrow 0$ as $t \rightarrow +\infty$

$$(80) \quad \begin{aligned} \int_0^t \|\dot{\epsilon}(\tau)x(\tau)\|d\tau &= \int_0^t \dot{\epsilon}(\tau)\|x(\tau)\|d\tau \\ &\leq M \int_0^t \dot{\epsilon}(\tau)d\tau \\ &\leq M|\epsilon(0)|. \end{aligned}$$

Hence this integral is convergent (for the norm topology) and

$$(81) \quad \int_0^t \dot{\epsilon}(\tau)x(\tau)d\tau \rightarrow \int_0^\infty \dot{\epsilon}(\tau)x(\tau)d\tau.$$

Since $\frac{1}{1+\epsilon(t)} \rightarrow 1$ we obtain

$$(82) \quad \lim_{t \rightarrow \infty} \eta(t) = \eta := z_0 + \int_0^\infty \dot{\epsilon}(\tau)x(\tau)d\tau.$$

Since $J_{\nu(t)}^A$ is a contraction, by (77)

$$(83) \quad \|x(t) - J_{\nu(t)}^A \eta\| \leq \|\eta(t) - \eta\|$$

which tends to zero. The strong convergence of $x(t)$ is a consequence of the strong convergence of $J_{\nu(t)}^A \eta$, and of the fact that $\nu(t)$ tend to infinity (indeed like e^t). This is a consequence of the classical asymptotic properties of the Tikhonov approximation, see for example see [15, 17, 30].

Note that the strong limit x_∞ of $x(t)$ satisfies:

$$(84) \quad x_\infty = \text{proj}_{A^{-1}0} \left(z_0 + \int_0^\infty \dot{\epsilon}(\tau)x(\tau)d\tau \right).$$

When $\epsilon = 0$ one recovers the classical formula.

REFERENCES

- [1] B. Abbas, H. Attouch and B.F. Svaiter, *Newton-like dynamics and forward-backward methods for structured monotone inclusions in Hilbert spaces*, J. Optim. Theory Appl., **161** (2014), No. 2, pp. 331-360.
- [2] F. Alvarez and J.M. Pérez, *A dynamical system associated with Newton's method for parametric approximations of convex minimization problems*, Appl. Math. Optim., **38** (1998), pp. 193-217.
- [3] F. Alvarez, H. Attouch, J. Bolte and P. Redont, *A second-order gradient-like dissipative dynamical system with Hessian-driven damping. Application to optimization and mechanics*, J. Math. Pures Appl., **81** (2002), pp. 747-779.
- [4] H. Attouch, P.E. Maingé, and P. Redont, *A second-order differential system with Hessian-driven damping; Application to non-elastic shock laws*, Differential Equations and Applications, **4** (2012), No. 1, pp. 27-65.
- [5] H. Attouch, M. Marques Alves and B.F. Svaiter, *A dynamic approach to a proximal-Newton method for monotone inclusions in Hilbert Spaces, with complexity $\mathcal{O}(1/n^2)$* , Journal of Convex Analysis, **23** (2016), No. 1, pp. 139-180.
- [6] H. Attouch, and B.F. Svaiter, *A continuous dynamical Newton-like approach to solving monotone inclusions*, SIAM J. Control Optim., **49** (2011), No. 2, pp. 574-598.
- [7] H. Attouch, P. Redont, and B.F. Svaiter, *Global convergence of a closed-loop regularized Newton method for solving monotone inclusions in Hilbert spaces*, J. Optim. Theory Appl., **157** (2013), No. 3, pp. 624-650.
- [8] J.-B. Baillon, *Un exemple concernant le comportement asymptotique de la solution du problème $\frac{du}{dt} + \partial\Phi(u) \ni 0$* , J. Funct. Anal., **28** (1978), pp. 369-376.
- [9] J.-B. Baillon, *Thèse, Université Paris VI*, 1978.
- [10] J.-B. Baillon and H. Brézis, *Une remarque sur le comportement asymptotique des semi-groupes non linéaires*, Houston J. Math., **2** (1976), pp. 5-7.
- [11] H.H. Bauschke and P.L. Combettes, *A weak-to-strong convergence principle for Fejer-monotone methods in Hilbert spaces*, Mathematics of Operations Research, **26** (2001), pp. 248-264.
- [12] H.H. Bauschke and P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, (2011).
- [13] H. Brézis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland/Elsevier, New-York, 1973.
- [14] F.E. Browder, *Nonlinear Operators and Nonlinear Equations of Evolution in Banach Spaces*, Proc. Sympos. Pure Math., **18** (part 2), Amer. Math. Soc., Providence, RI, 1976.
- [15] F.E. Browder, *Existence and approximation of solutions of nonlinear variational inequalities*, Proc. Nat. Acad. Sci. USA, **56** (1966), pp. 1080-1086.
- [16] R.E. Bruck, *Asymptotic convergence of nonlinear contraction semigroups in Hilbert spaces*, J. Funct. Anal., **18** (1975), pp. 15-26.
- [17] R. Cominetti, J. Peypouquet and S. Sorin, *Strong asymptotic convergence of evolution equations governed by maximal monotone operators with Tikhonov regularization*, J. of Diff. Equation, **245** (2008), pp. 3753-3763.
- [18] M.G. Crandall, S.O. Londen and J.A. Nohel, *An abstract nonlinear Volterra integrodifferential equation*, J. of Math. Anal. Appl., **64** (1978), pp. 701-735.
- [19] X. Goudou and J. Munier, *Asymptotic behavior of solutions of a gradient like integrodifferential Volterra inclusion*, Advances in Math. Sciences and Applications, **15** (2005), pp. 509-525.
- [20] O. Güler, *On the convergence of the proximal point algorithm for convex minimization*, SIAM J. Control Optim., **29** (1991), 403-419.
- [21] A. Haraux, *Systèmes dynamiques dissipatifs et applications*, RMA 17, Masson, Paris, 1991.
- [22] K. Levenberg, *A method for the solution of certain non-linear problems in the least squares*, Quarterly of Applied Mathematics, **2** (1944), pp. 164-168.

- [23] P.-E. Maingé, *First order continuous Newton-like systems for monotone inclusions*, SIAM Journal on Control and Optimization, **51** (2) (2013), pp. 1615–1638.
- [24] D.W. Marquardt, *An algorithm for least squares estimation of non-linear parameters*, SIAM J. Appl. Math., **11** (1963), pp. 431–441.
- [25] G.J. Minty, *Monotone (nonlinear) operators in Hilbert spaces*, Duke Mathematical Journal, **29** (1962), pp. 341–346.
- [26] R. D. C. Monteiro and B. F. Svaiter, *On the complexity of the hybrid proximal extragradient method for the iterates and the ergodic mean*, SIAM J. Optim., **20** (2010), No. 6, pp. 2755–2787.
- [27] R. D. C. Monteiro and B. F. Svaiter, *Iteration-complexity of a Newton proximal extragradient method for monotone variational inequalities and inclusion problems*, SIAM J. Optim., **22** (2012), No. 3, pp. 914–935.
- [28] Z. Opial, *Weak convergence of the sequence of successive approximations for nonexpansive mappings*, Bull. Amer. Math. Soc., **73** (1967), pp. 591–597.
- [29] J. Peypouquet and S. Sorin, *Evolution equations for maximal monotone operators: asymptotic analysis in continuous and discrete time*, J. of Convex Analysis, **17** (2010), pp. 1113–1163.
- [30] S. Reich, *Nonlinear evolution equations and nonlinear ergodic theorems*, Nonlinear Anal., **1** (1976), pp. 319–330.
- [31] R.T. Rockafellar, *Maximal monotone relations and the second derivatives of nonsmooth functions*, Ann. Inst. Henri Poincaré, **2** (1985), pp. 167–184.
- [32] M.V. Solodov and B.F. Svaiter, *A globally convergent inexact Newton method for systems of monotone equations*, Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, Edited by M. Fukushima and L. Qi, Kluwer Academic Publishers, (1999), pp. 355–369.
- [33] V. Vergara and R. Zacher, *Lyapunov functions and convergence to steady state for differential equations of fractional order*, Math. Z., **259** (2008), pp. 287–309.
- [34] E. Zeidler, *Nonlinear functional analysis and its applications, Part II: Monotone operators*, Springer-Verlag, New-York, 1990.

INSTITUT MONTPELLIÉRAIN ALEXANDER GROTHENDIECK, UMR 5149 CNRS, UNIVERSITÉ MONTPELLIER, PLACE EUGÈNE BATAILLON, 34095 MONTPELLIER CEDEX 5, FRANCE

E-mail address: hedy.attouch@univ-montp2.fr

UNIVERSITÉ PARIS 1 PANTHÉON-SORBONNE, CERMSEM CNRS-UMR 8095, UFR 27 MATHÉMATIQUES ET INFORMATIQUE

E-mail address: Jean-Bernard.Baillon@univ-paris1.fr