



Global convergence of a closed-loop regularized Newton method for solving monotone inclusions in Hilbert spaces

Hedy Attouch, Patrick Redont, Benar Fux Svaiter

► To cite this version:

Hedy Attouch, Patrick Redont, Benar Fux Svaiter. Global convergence of a closed-loop regularized Newton method for solving monotone inclusions in Hilbert spaces. *Journal of Optimization Theory and Applications*, 2012, Published online (Published online), pp.Published online 21 November 2012. 10.1007/s10957-012-0222-3 . hal-00803194

HAL Id: hal-00803194

<https://hal.science/hal-00803194>

Submitted on 21 Mar 2013

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.

Global Convergence of a Closed-loop Regularized Newton Method for Solving Monotone Inclusions in Hilbert Spaces

H. Attouch · P. Redont · B. F. Svaiter

Abstract We analyze the global convergence properties of some variants of regularized continuous Newton methods for convex optimization and monotone inclusions in Hilbert spaces. The regularization term is of Levenberg-Marquardt type and acts in an open-loop or closed-loop form. In the open-loop case the regularization term may be of bounded variation.

Keywords Newton-type Methods · Monotone Inclusions · Convex Optimization · Closed-loop Regularization · BV Controls · Levenberg-Marquardt Method.

Mathematics Subject Classification (2000) 47H05 · 49J52 · 47N10.

1 Introduction

Iterative algorithms and continuous dynamical systems have straight relationships. Discretizing continuous dynamical systems gives rise to sequences of iterates; conversely, an iterative algorithm may often be considered as a discretized version of a continuous dynamical system. Once this dynamical system is identified, its asymptotic properties (and its Liapunov functions) can give new insights into the properties of the algorithm. For example Cauchy's steepest descent algorithm is related to a “minus” gradient flow, which can be extended to a differential inclusion equation for maximal monotone operators.

H. Attouch
I3M UMR CNRS 5149, Université Montpellier II, Pl. E. Bataillon, 34095 Montpellier, France
E-mail: attouch@math.univ-montp2.fr

P. Redont
I3M UMR CNRS 5149, Université Montpellier II, Pl. E. Bataillon, 34095 Montpellier, France

B. F. Svaiter
IMPA, Estrada Dona Castorina 110, 22460-320 Rio de Janeiro, Brazil
E-mail: benar@impa.br

Newton's method, however, requires a lot of regularity both in its discrete and continuous version. To circumvent the ill-posedness of Newton's method, regularizations of some sort have to be used: regularization of the objective function, hyperbolic regularization. . .

Levenberg-Marquardt regularization is well known for the discrete Newton method. Recently, Attouch and Svaiter in [1] have introduced a Levenberg-Marquardt regularization for the continuous Newton method that enjoys nice properties: well-posedness, good asymptotic behavior, generality (it works for maximal monotone operators). Yet the dynamics in [1] is nonautonomous; the regularization coefficient is a function of the time variable and so, in control terms, is of open-loop type. Another limitation of that dynamics is the absolute continuity assumption on the regularization coefficient.

The main aim of the present work is to extend the results of [1] to a Levenberg-Marquardt regularization of closed-loop type, *i.e.* depending on the state of the system. We also extend the results of [1] to the nonautonomous case in which the regularization parameter is of bounded variation.

2 Preliminaries: Theoretical Background

Let H be a real Hilbert space endowed with the scalar product $\langle \cdot, \cdot \rangle$ and associated norm $\|x\| = \sqrt{\langle x, x \rangle}$ for $x \in H$. Let $A : H \rightrightarrows H$ be a maximal monotone operator. We are concerned with the existence, and the asymptotic behavior (as $t \rightarrow +\infty$) of strong solutions of the differential inclusion

$$\begin{cases} v(t) \in A(x(t)) \\ G(x(t), v(t))\dot{x}(t) + \dot{v}(t) + v(t) = 0 \end{cases} \quad (1)$$

where $G : H \times H \rightarrow [0, \infty[$ is, for $v \neq 0$, strictly positive. This differential inclusion may be viewed as a perturbation of the following one

$$\begin{cases} v(t) \in A(x(t)) \\ \dot{v}(t) + v(t) = 0. \end{cases} \quad (2)$$

For $A = \nabla f$, the gradient of a twice differentiable function f , (2) reduces to $\nabla^2 f(x(t))\dot{x}(t) + \nabla f(x(t)) = 0$, the continuous Newton dynamics for f . Thus, system (2) deserves to be called the continuous Newton dynamical system for operator A . Unfortunately, (2) is ill-posed. Indeed, after an integration step, it is equivalent to $v(0)e^{-t} \in A(x(t))$, and so boils down to first knowing whether the segment $(0, v(0))$ is included in the range of A , in which case, for each t , one has to solve the above equation, and hopefully find a differentiable function of t ... These are questions far too general without extra particular assumptions (see however [2, § 3] for the case of a subdifferential operator). The ill-posedness of system (2) is in sharp contrast to the richness of (1). Actually, the term $G(x(t), v(t))\dot{x}(t)$ acts as a Levenberg-Marquardt regularization, in a closed-loop way, of the continuous Newton method, and can be seen as a feed-back function of the current state $(x(t), v(t))$. This point of view may be convenient for developing numerical schemes.

Our approach relies heavily on [1], which concerns the open-loop case

$$\begin{cases} v(t) \in A(x(t)) \\ \lambda(t)\dot{x}(t) + \dot{v}(t) + v(t) = 0. \end{cases} \quad (3)$$

Let us summarize the results obtained in [1]. If λ is positive, locally absolutely continuous, then for any Cauchy data $x(0) = x_0$, $v(0) = v_0$, $v_0 \in A(x_0)$, there exists a unique strong global solution of (3). Moreover, if $\lambda(t)$ tends to zero not too fast, as $t \rightarrow \infty$ (roughly speaking no faster than e^{-t}), then $v(t) \rightarrow 0$ strongly, and $x(t)$ converges weakly to some equilibrium $x_\infty \in A^{-1}(0)$, provided the latter set is nonempty. In [1] the main concern was precisely to asymptotically stabilize the system, with $\lambda(t)$ as small as possible, and thus to keep close to the Newton dynamics. So doing, one can obtain continuous and discrete Newton-like dynamics attached to solving the equation

$$\text{find } x \in H \text{ such that } 0 \in Ax. \quad (4)$$

In this paper, we show how to pass from an open-loop $\lambda(t)$ to a closed-loop regularization coefficient $G(x(t), v(t))$. The open-loop case $\lambda(t) = e^{-t}$ is particularly convenient because (3) can then be integrated: $x(t) = (I + e^t A)^{-1}(x_0 + v_0)$. The asymptotic convergence follows from classical results for resolvents of maximal monotone operators; see [3]. By contrast, the closed-loop analysis requires a preliminary detailed study of the open-loop case for λ belonging to an appropriate class of functions. Indeed, we use a fixed point argument for an operator defined on that class.

Note that, even in the open-loop case, when choosing a particular λ , it is important for numerical reasons to study the stability of the solution with respect to perturbations of λ . This stability property is a key for our analysis. It is worth noting that system (1) or (3) is a regularized continuous Newton method; as such it allows one to handle general monotone maximal operators. In contrast, nonsmooth Newton methods deal with operators with some regularity and strive to find a substitute for the Newton direction; see [4] for recent results and the references therein.

Let us list the main points of our study:

1) In Section 3 we review some facts concerning the Cauchy problem for the open-loop regularized system (3). We first consider the case where λ is locally absolutely continuous. In Theorem 3.1 we prove the Lipschitz continuous dependence of the solution with respect to λ . This property will play a key role in the study of the closed-loop case in the next section. Moreover, the Lipschitz constant only depends on the L^1 norm of the time derivative of λ .

This naturally leads us to extend our analysis to the case where λ is a function with bounded variation (possibly involving jumps). We use a regularization by convolution method in order to reduce to the smooth case, and then pass to the limit in the equations. So doing, in Theorem 3.2 and Corollary 3.2, we prove the existence and uniqueness of a strong solution for (3) with Cauchy data.

Using discontinuous controllers and feedbacks turns out to be an important property in the stabilization of nonlinear systems; see [5] for a survey on this active research field. The bounded variation case (where the system can be stabilized by discontinuous controls) can be seen as intermediate between the continuous and the discrete dynamics and this naturally paves the way to the study of discrete algorithms.

2) In Section 4, Theorem 4.1, under general assumptions on G (essentially local Lipschitz continuity), we prove the existence and uniqueness of a strong global solution $(x, v) : [0, \infty[\rightarrow H \times H$ to the closed-loop system (1), with Cauchy data. The main argument of the proof is the Banach-Picard fixed-point theorem.

3) In Section 5, we examine the asymptotic stabilization properties of trajectories of system (1) in the particular case $G(x, v) = \alpha(\|v\|^2)$, that is

$$\begin{cases} v(t) \in A(x(t)) \\ \alpha(\|v(t)\|^2)\dot{x}(t) + \dot{v}(t) + v(t) = 0; \\ x(0) = x_0, v(0) = v_0 \in A(x_0), v_0 \neq 0. \end{cases} \quad (5)$$

We first examine the case where α is continuously differentiable. Under the sole assumption that $\alpha :]0, +\infty[\rightarrow]0, +\infty[$ is a C^1 function, we show that $\lim_{t \rightarrow \infty} \|v(t)\| = 0$. Thus, the asymptotic behavior of trajectories of system (5) depends on the behavior of α near 0. Precisely, in Theorem 5.1 we show that, if α is bounded above near 0, and

$$\limsup_{r \rightarrow 0^+} \frac{r\dot{\alpha}(r)}{\alpha(r)} < \frac{1}{2}, \quad (6)$$

then, any trajectory of (5) weakly converges to a zero of A . In particular, for $0 < \gamma < 1$, we obtain the asymptotic convergence of the system

$$\begin{cases} v(t) \in A(x(t)) \\ \|v(t)\|^\gamma \dot{x}(t) + \dot{v}(t) + v(t) = 0; \\ x(0) = x_0, v(0) = v_0 \in A(x_0), v_0 \neq 0. \end{cases} \quad (7)$$

Then, in Theorem 5.2, we examine the case where $\alpha : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is an increasing function (with possible jumps). Denoting by $[\alpha]$ the extension of α obtained by filling in the jumps of α , we obtain the existence of strong global solutions of the closed-loop regularized Newton continuous dynamics

$$\begin{cases} v(t) \in A(x(t)) \\ [\alpha](\|v(t)\|^2)\dot{x}(t) + \dot{v}(t) + v(t) \ni 0; \\ x(0) = x_0, v(0) = v_0 \in A(x_0), v_0 \neq 0, \end{cases} \quad (8)$$

and analyze some of their convergence properties.

4) In Section 6, we illustrate our results in the case of convex-concave saddle value problems. By taking advantage of the fact that the saddle points of a convex-concave function L are the zeroes of the associated maximal monotone operator $A = (\partial_x L, -\partial_y L)$, we are able to develop a Newton-like approach for

finding saddle points of convex-concave functions. In particular, we consider the primal-dual Lagrangian approach to linearly constrained convex minimization problems. Let us stress that our dynamics yields weakly convergent trajectories. By contrast, the approach based on the semi-group generated by A , only provides ergodic weak convergence of the trajectories.

5) Concerning Newton-type dynamics and the use of regularization methods let us mention two close companions of our dynamics:

a) In the case of convex optimization $A = \partial f$, the second-order continuous Dynamical Inertial Newton system

$$\gamma \ddot{x}(t) + \lambda \dot{x}(t) + \nabla^2 f(x(t)) \dot{x}(t) + \nabla f(x(t)) = 0 \quad (9)$$

has been introduced by Alvarez, Attouch, Bolte, and Redont in [6]; see also [7, 8]. This is the fundamental equation of (newtonian) mechanics with $\lambda \dot{x}(t)$ and $\nabla^2 f(x(t)) \dot{x}(t)$ acting respectively as viscous and geometrical Hessian-driven damping terms. This system enjoys remarkable asymptotic stabilization properties. It can be equivalently written as a first order system in time and space, which allows one to give it a sense when f is only assumed lower semicontinuous (it is different from Minty's transformation). Recently Maingé in [9] has developed (9) with a general maximal monotone operator A (instead of a subdifferential), and shown that system (3), in some instances, can be obtained as the limit of (9) when γ tends to zero. Note that (3), as a decisive advantage with respect to (9), is a first order system with respect to x , which makes it simpler for numerical purpose. This hyperbolic regularization links our dynamics with physics, and control of oscillating systems. In turn, it suggests considering new stabilization procedures for (9), based on a closed-loop control, for example $\lambda(t) = \alpha(\|\nabla f(x(t))\|^2)$.

b) Another interesting dynamics, based on the regularization of the objective function $f(\cdot, \epsilon) \rightarrow f$ as $\epsilon \rightarrow 0$, has been developed by Alvarez and Pérez in [2]:

$$\nabla^2 f(x(t), \epsilon(t)) \dot{x}(t) + \dot{\epsilon}(t) \frac{\partial^2 f}{\partial \epsilon \partial x}(x(t), \epsilon(t)) + \nabla f(x(t), \epsilon(t)) = 0. \quad (10)$$

By contrast with our approach, this dynamics involves both the regularizing parameter $\epsilon(\cdot)$ and its derivative, which may be a source of numerical difficulties. On the other hand, this approach can handle general regularization methods, like interior point methods.

3 Open-loop Regularization

In this section we consider the open-loop regularized system

$$v(t) \in A(x(t)) \quad (11a)$$

$$\lambda(t) \dot{x}(t) + \dot{v}(t) + v(t) = 0 \quad (11b)$$

$$x(0) = x_0, \quad v(0) = v_0, \quad v_0 \in A(x_0), \quad v_0 \neq 0. \quad (11c)$$

We recall the notion of solution (see [1]) that we will use all along the paper.

Definition 3.1 Given $T > 0$ and $(x_0, v_0) \in H \times H$, a function $(x, v) : [0, T] \rightarrow H \times H$ is a *strong solution* of system (11) iff (x, v) is absolutely continuous on $[0, T]$, (11a) is satisfied for all $t \in [0, T]$, (11b) is satisfied for almost every $t \in [0, T]$ and the initial condition (11c) holds true.

A function $(x, v) : [0, +\infty[\rightarrow H \times H$ is a *strong solution* of system (11) iff its restriction to $[0, T]$ is a strong solution of (11) for any $T > 0$.

We first review some facts on the open-loop regularized system (11) when λ is locally absolutely continuous. As a new result we show the Lipschitz continuous dependence of the solution with respect to λ . This property will be a key for the study of the closed-loop case in the next section. Then, we show how to extend this analysis to the case where λ is a function with bounded variation (possibly involving jumps). Here and subsequently we will often omit the time variable t and write $x, v, \dots$ for $x(t), v(t), \dots$ when no ambiguity arises.

3.1 Absolutely Continuous Regularization Coefficient $\lambda(\cdot)$

The next proposition subsumes those results of [1] to be used in the sequel. The linear space $H \times H$ is equipped with its usual Hilbertian norm $\|(\xi, \zeta)\| = \sqrt{\|\xi\|^2 + \|\zeta\|^2}$.

Proposition 3.1 *Let $c_0 > 0$, $\lambda : [0, T] \rightarrow [c_0, \infty)$. Then any strong solution (x, v) of (11) verifies:*

$$\|v(t)\| \geq e^{-t}\|v_0\|, \quad \forall t \in [0, T] \quad (12a)$$

$$\|\dot{x}(t)\| \leq \frac{\|v_0\|}{c_0}, \quad \|\dot{v}(t)\| \leq \|v(t)\| \leq \|v_0\|, \quad \text{a.e. } t \text{ in } [0, T] \quad (12b)$$

$$\|(\dot{x}(t), \dot{v}(t))\| \leq \frac{\|v_0\|}{\min\{c_0, 1\}} := L, \quad \text{a.e. } t \text{ in } [0, T] \quad (12c)$$

$$\langle v(t), \dot{v}(t) \rangle \leq 0 \quad \text{a.e. } t \text{ in } [0, T], \quad \text{hence } t \mapsto \|v(t)\| \text{ is nonincreasing.} \quad (12d)$$

If λ is absolutely continuous, then there exists a unique pair $(x, v) : [0, T] \rightarrow H \times H$ strong solution of (11) (hence (x, v) is L -Lipschitz continuous).

Corollary 3.1 *Let $\lambda : [0, +\infty[\rightarrow (0, +\infty[$ be absolutely continuous on each bounded interval $[0, T]$, $T > 0$. Then there exists a unique strong solution $(x, v) : [0, +\infty[\rightarrow H \times H$ to system (11).*

In next theorem we analyze the dependence of the pair (x, v) specified in Proposition 3.1 on the function λ and the initial point (x_0, v_0) . These results will be instrumental for proving existence and uniqueness of strong solutions of (1).

Theorem 3.1 *Suppose that $\lambda, \eta : [0, T] \rightarrow [c_0, \infty[$ are absolutely continuous functions, with $T > 0$ and $c_0 > 0$. Let $(x, v), (y, w) : [0, T] \rightarrow H \times H$ be the*

respective strong solutions of the inclusions

$$\lambda \dot{x} + \dot{v} + v = 0, \quad v \in A(x), \quad x(0) = x_0, v(0) = v_0 \quad (13)$$

$$\eta \dot{y} + \dot{w} + w = 0, \quad w \in A(y), \quad y(0) = y_0, w(0) = w_0. \quad (14)$$

Define $\varphi : [0, T] \rightarrow \mathbb{R}$ by

$$\varphi = \sqrt{c_0^2 \|x - y\|^2 + \|v - w\|^2}.$$

Then for any $s \in [0, T]$

$$\begin{aligned} \varphi(s) &\leq \int_0^s \left(\frac{(\|v_0\| + \|w_0\|)}{2c_0} |\lambda(t) - \eta(t)| + \varphi(t) \sqrt{1 + \frac{|\dot{\lambda}(t) + \dot{\eta}(t)|^2}{4c_0^2}} \right) dt \\ &\quad + \frac{\lambda(0) + \eta(0)}{2} \|x_0 - y_0\| + \|v_0 - w_0\|. \end{aligned} \quad (15)$$

In particular, if $x_0 = y_0$, $v_0 = w_0$, then

$$\|\varphi\|_{L^\infty([0, T])} \leq \frac{\|v_0\|}{c_0} \exp \left(T + \frac{\|\dot{\lambda} + \dot{\eta}\|_{L^1([0, T])}}{2c_0} \right) \|\lambda - \eta\|_{L^1([0, T])}. \quad (16)$$

Proof To simplify the exposition, define $\gamma : [0, T] \rightarrow \mathbb{R}$

$$\gamma = \frac{\lambda + \eta}{2}.$$

Using the assumptions $\lambda, \eta \geq c_0$ and the monotonicity of A (recall (13,14)) we conclude that for any $t \in [0, T]$

$$c_0 \leq \gamma, \quad \langle x - y, v - w \rangle \geq 0.$$

Therefore

$$\varphi \leq \sqrt{\gamma^2 \|x - y\|^2 + \|v - w\|^2} \leq \|\gamma(x - y) + v - w\|. \quad (17)$$

Direct algebraic manipulation of the first equalities in (13) and (14), together with the above definition of γ , shows that for almost all $t \in [0, T]$

$$\frac{d}{dt} [\gamma(x - y) + (v - w)] = \dot{\gamma}(x - y) - \frac{\lambda - \eta}{2} (\dot{x} + \dot{y}) - (v - w).$$

Since γ, x, y, v, w are absolutely continuous, the function $\gamma(x - y) + (v - w)$ is also absolutely continuous. As a consequence, integration of the above equality on $[0, s]$, for $s \in [0, T]$, yields

$$\begin{aligned} &[\gamma(x - y) + (v - w)](s) - [\gamma(x - y) + (v - w)](0) = \\ &-\frac{1}{2} \int_0^s (\lambda - \eta)(\dot{x} + \dot{y}) dt + \int_0^s (\dot{\gamma}(x - y) - (v - w)) dt. \end{aligned}$$

Hence

$$\begin{aligned} \|\gamma(x-y) + (v-w)](s)\| &\leq \gamma(0)\|x_0 - y_0\| + \|v_0 - w_0\| \\ &+ \frac{1}{2} \int_0^s |\lambda - \eta| \|\dot{x} + \dot{y}\| dt + \int_0^s \|\dot{\gamma}(x-y) - (v-w)\| dt. \end{aligned} \quad (18)$$

We know that (recall (12b))

$$\|\dot{x}\| \leq \|v_0\|/c_0, \quad \|\dot{y}\| \leq \|w_0\|/c_0, \quad (19)$$

and, besides

$$\begin{aligned} \|\dot{\gamma}(x-y) + w-v\|^2 &= |\dot{\gamma}|^2 \|x-y\|^2 + \|v-w\|^2 + 2 \left\langle c_0(x-y), \frac{\dot{\gamma}}{c_0}(w-v) \right\rangle \\ &\leq |\dot{\gamma}|^2 \|x-y\|^2 + \|v-w\|^2 + c_0^2 \|x-y\|^2 + \frac{|\dot{\gamma}|^2}{c_0^2} \|v-w\|^2 \\ &= \left(\frac{|\dot{\gamma}|^2}{c_0^2} + 1 \right) \varphi^2. \end{aligned} \quad (20)$$

Inequation (15) is then a straight consequence of (17, 18, 19, 20).

Now, if $x_0 = y_0$ and $v_0 = w_0$, we derive from (15)

$$\varphi(s) \leq \frac{\|v_0\|}{c_0} \|\lambda - \eta\|_{L^1([0,T])} + \int_0^s \varphi \sqrt{\frac{|\dot{\gamma}|^2}{c_0^2} + 1} dt.$$

Applying Gronwall's inequality (see *e.g.* [3, Appendice 4]) yields

$$\varphi(s) \leq \frac{\|v_0\|}{c_0} \|\lambda - \eta\|_{L^1([0,T])} \exp \left(\int_0^s \sqrt{\frac{|\dot{\gamma}|^2}{c_0^2} + 1} dt \right).$$

With $\sqrt{\frac{|\dot{\gamma}|^2}{c_0^2} + 1} \leq 1 + |\dot{\gamma}|/c_0$, we finally have, for any $s \in [0, T]$

$$\varphi(s) \leq \frac{\|v_0\|}{c_0} \|\lambda - \eta\|_{L^1([0,T])} \exp \left(T + \frac{\|\dot{\gamma}\|_{L^1([0,T])}}{c_0} \right)$$

whence (16). $\square$

Remark 3.1 It is worth noticing that, besides the Lipschitz continuous dependence with respect to λ of the solution (x, v) of (11), Theorem 3.1 also provides its continuous dependence with respect to the initial data (x_0, v_0) . More precisely, if (x_n, v_n) (resp. (x, v)) is the solution of (11) corresponding to the Cauchy data (x_{0n}, v_{0n}) (resp. (x_0, v_0)), as a direct consequence of (15) and the Gronwall lemma we obtain for all $T > 0$

$$(A \ni (x_{0n}, v_{0n}) \rightarrow (x_0, v_0)) \Rightarrow (x_n, v_n) \rightarrow (x, v) \text{ uniformly on } [0, T].$$

Note also that, by contrast with the steepest descent dynamics [3], there is no regularizing effect on the initial condition: there is no way to define a solution

of (11) for $x_0 \in \overline{\text{dom } A} \setminus \text{dom } A$ since in that case, for any approximating sequence $x_{0n} \in \text{dom } A$ and $v_{0n} \in A(x_{0n})$, one has $\lim_n \|v_{0n}\| = +\infty$, which by (12a) would imply the blow-up of the sequence (x_n, v_n) as $n \rightarrow +\infty$, on any finite time interval.

The next proposition gives a condition on λ yielding $\lim_{t \rightarrow +\infty} v(t) = 0$ for the strong solution (x, v) of system (11). It will prove useful in Section 5.

Proposition 3.2 *Let $\lambda : [0, +\infty[\rightarrow]0, +\infty[$ be absolutely continuous on each bounded interval $[0, T]$, $T > 0$; let $(x, v) : [0, +\infty[\rightarrow H \times H$ be the strong solution of system (11). Suppose $A^{-1}(0) \neq \emptyset$. Then $\|v\| \in L^2(0, +\infty)$, $\|v(t)\| = O(1/\sqrt{t})$ as $t \rightarrow +\infty$ and $\lim_{t \rightarrow +\infty} v(t) = 0$ under any of the two conditions*

- a) λ is nonincreasing;
- b) $\dot{\lambda} \in L^1(0, +\infty)$ and there exists $\lambda_\infty = \lim_{t \rightarrow +\infty} \lambda(t) > 0$.

Proof The existence and uniqueness of the strong solution (x, v) follows from Corollary 3.1. Take $\|v\| \in L^2(0, +\infty)$ for granted; since $\|v\|$ is nonincreasing (recall (12d)) we have

$$\int_0^{+\infty} \|v(s)\|^2 ds \geq \int_0^t \|v(s)\|^2 ds \geq t\|v(t)\|^2, \quad \forall t \geq 0.$$

Hence $\|v(t)\| = O(1/\sqrt{t})$ as $t \rightarrow +\infty$ and $\lim_{t \rightarrow +\infty} v(t) = 0$.

a) $\|v\| \in L^2(0, +\infty)$ is proved in [1, Lemma 3.4].

b) Set $\dot{\lambda}^+ = \max(\dot{\lambda}, 0)$, $\dot{\lambda}^- = \max(-\dot{\lambda}, 0)$, so that $\dot{\lambda} = \dot{\lambda}^+ - \dot{\lambda}^-$ and $\dot{\lambda}^+$, $\dot{\lambda}^-$ belong to $L^1(0, +\infty)$. Define $\Lambda(t) = \frac{\lambda_\infty}{2} + \int_t^\infty \dot{\lambda}^-(s) ds$ and note $\lambda(t) =$

$\left(\frac{\lambda_\infty}{2} - \int_t^\infty \dot{\lambda}^+(s) ds\right) + \Lambda(t)$. The following facts are straightforward:

i) $\dot{\Lambda} = -\dot{\lambda}^-$;

ii) Λ is bounded below by $\frac{\lambda_\infty}{2}$ and λ is bounded above by $\lambda(0) + \int_0^\infty |\dot{\lambda}(s)| ds$; hence there exists some constant l such that $0 < l \leq \frac{\Lambda(t)}{\lambda(t)}$, $\forall t \geq 0$;

iii) there exists some $\tau \geq 0$ such that $\frac{\Lambda(t)}{\lambda(t)} \leq 1$, $\forall t \geq \tau$.

Fix $\hat{x} \in A^{-1}(0)$. Define $g(t) = \frac{1}{2}\|(x(t) - \hat{x}) + \frac{v(t)}{\lambda(t)}\|^2$ and $H(t) = \Lambda^2(t)g(t)$. We have

$$g = \frac{1}{2}\|x - \hat{x}\|^2 + \frac{1}{\lambda}\langle v, x - \hat{x} \rangle + \frac{1}{2\lambda^2}\|v\|^2.$$

Using (11b) we have, after some calculations

$$\dot{g} = -\left(\frac{1}{\lambda} + \frac{\dot{\lambda}}{\lambda^2}\right)\langle v, x - \hat{x} \rangle - \left(\frac{1}{\lambda^2} + \frac{\dot{\lambda}}{\lambda^3}\right)\|v\|^2.$$

Hence we deduce

$$\begin{aligned}\dot{H} &= \Lambda^2 \dot{g} + 2\Lambda \dot{A}g \\ &= -\|v\|^2 \left\{ \frac{\Lambda^2}{\lambda^2} + \frac{\Lambda^2 \dot{\lambda}^+}{\lambda^3} + \frac{\Lambda \dot{\lambda}^-}{\lambda^2} \left(1 - \frac{\Lambda}{\lambda} \right) \right\} \\ &\quad - \langle v, x - \hat{x} \rangle \left\{ \frac{\Lambda^2}{\lambda} + \frac{\Lambda^2 \dot{\lambda}^+}{\lambda^2} + \frac{\Lambda \dot{\lambda}^-}{\lambda} \left(2 - \frac{\Lambda}{\lambda} \right) \right\} - \|x - \hat{x}\|^2 \Lambda \dot{\lambda}^-.\end{aligned}$$

Noting $\langle v, x - \hat{x} \rangle \geq 0$ (because $v \in Ax$, $0 \in A\hat{x}$ and A is monotone) and taking into account the above-mentioned facts on λ and Λ , we deduce first

$$\dot{H} \leq -\|v\|^2 \frac{\Lambda^2}{\lambda^2},$$

and next

$$\begin{aligned}\frac{dH}{dt}(t) + l^2 \|v(t)\|^2 &\leq 0, \text{ for a.e. } t \geq \tau \\ l^2 \int_{\tau}^t \|v(s)\|^2 ds &\leq H(\tau) - H(t) \leq H(\tau), \quad \forall t \geq \tau.\end{aligned}$$

Hence $\int_{\tau}^{\infty} \|v(s)\|^2 ds < +\infty$. □

3.2 Bounded Variation Regularization Coefficient $\lambda(\cdot)$

Let us now suppose that $\lambda : [0, T] \rightarrow]0, \infty[$ is of bounded variation on $[0, T]$, where $T > 0$. That is $TV(\lambda, [0, T]) < +\infty$, where $TV(\lambda, [0, T])$ is the total variation of λ on $[0, T]$:

$$TV(\lambda, [0, T]) = \sup \sum_{i=1}^p |\lambda(\tau_i) - \lambda(\tau_{i-1})|$$

the supremum being taken over all $p \in \mathbb{N}$ and all strictly increasing finite sequences $\tau_0 < \tau_1 < \dots < \tau_p$ of points of $[0, T]$. Function λ may involve jumps. We also suppose that λ is bounded away from 0:

$$\inf \lambda([0, T]) > 0.$$

The following lemma gathers some classical facts concerning the approximation of functions of bounded variation by smooth functions (see for example [10, Theorem 10.1.2.], [11, 12]) together with some technical results useful for the sequel. For convenience of the reader we give a self-contained proof.

Lemma 3.1 *Let $\lambda : [0, T] \rightarrow]0, \infty[$ be of bounded variation on $[0, T]$. Then, there exists a sequence $(\lambda_n)_{n \in \mathbb{N}}$, with $\lambda_n \in \mathcal{C}^\infty([0, T])$ for each $n \in \mathbb{N}$, such that*

(i) $\inf \lambda([0, T]) \leq \lambda_n(t) \leq \sup \lambda([0, T]), \forall t \in [0, T], \forall n \in \mathbb{N}$.

In particular, $\lambda_n \geq 0$ if $\lambda \geq 0$;

(ii) $\lambda_n \rightarrow \lambda$ in $L^p(0, T)$ for any $1 \leq p < \infty$;

(iii) $TV(\lambda_n, [0, T]) = \int_0^T |\dot{\lambda}_n(t)| dt \leq TV(\lambda, [0, T])$.

Proof Extend λ to the function $\bar{\lambda}$ defined on $\mathbb{R}$ by

$$\begin{aligned} t \leq 0 : \bar{\lambda}(t) &= \lambda(0) \\ 0 \leq t \leq T : \bar{\lambda}(t) &= \lambda(t) \\ T \leq t : \bar{\lambda}(t) &= \lambda(T). \end{aligned}$$

Function $\bar{\lambda}$ is of bounded variation on $\mathbb{R}$. Let $(\rho_n)_{n \in \mathbb{N}}$ be a mollifying sequence ($\rho_n \in \mathcal{C}_c^\infty(\mathbb{R})$, $\rho_n \geq 0$, $\int \rho_n = 1$ and $\text{support}(\rho_n) \subseteq [-s_n, s_n]$ with $s_n > 0$, $s_n \rightarrow 0$). Classically, the convolution $\lambda_n = \rho_n * \bar{\lambda}$ belongs to $\mathcal{C}^\infty(\mathbb{R})$ and its restriction to $[0, T]$ satisfies (i) and (ii).

Choose points $\tau_0 < \tau_1 < \dots < \tau_p$ in $[0, T]$:

$$\begin{aligned} \sum_{i=1}^p |\lambda_n(\tau_i) - \lambda_n(\tau_{i-1})| &\leq \sum_{i=1}^p \int_{\mathbb{R}} |\bar{\lambda}(\tau_i - t) - \bar{\lambda}(\tau_{i-1} - t)| \rho_n(t) dt \\ &= \int_{-s_n}^{s_n} \left(\sum_{i=1}^p |\bar{\lambda}(\tau_i - t) - \bar{\lambda}(\tau_{i-1} - t)| \right) \rho_n(t) dt. \end{aligned}$$

If $|t| \leq s_n$, we have

$$\begin{aligned} \sum_{i=1}^p |\bar{\lambda}(\tau_i - t) - \bar{\lambda}(\tau_{i-1} - t)| &\leq TV(\bar{\lambda}, [-s_n, T + s_n]) \\ &= TV(\bar{\lambda}, [-s_n, 0]) + TV(\bar{\lambda}, [0, T]) + TV(\bar{\lambda}, [T, T + s_n]) \\ &= TV(\lambda, [0, T]). \end{aligned}$$

Hence

$$\sum_{i=1}^p |\lambda_n(\tau_i) - \lambda_n(\tau_{i-1})| \leq TV(\lambda, [0, T])$$

and consequently $TV(\lambda_n, [0, T]) \leq TV(\lambda, [0, T])$. Owing to λ_n being continuously differentiable, $TV(\lambda_n, [0, T]) = \int_0^T |\dot{\lambda}_n| dt$. $\square$

Lemma 3.2 *Let $z_n, z \in \mathcal{C}([0, T], H)$ be such that $z_n \rightarrow z$ uniformly and $(z_n)_{n \in \mathbb{N}}$ is L -Lipschitz continuous for some positive constant L independent of $n \in \mathbb{N}$. Let $\lambda_n \rightarrow \lambda$ in $L^2(0, T)$. Then $\lambda_n \dot{z}_n$ converges weakly to $\lambda \dot{z}$ in $L^2(0, T; H)$.*

Proof Classically, $\dot{z}_n \rightarrow \dot{z}$ in the distribution sense, and the sequence $(\dot{z}_n)$ remains bounded in $L^\infty(0, T; H)$, hence in $L^2(0, T; H)$. Therefore $\dot{z}_n$ converges weakly to $\dot{z}$ in $L^2(0, T; H)$. Now, given any $\zeta \in \mathcal{C}_c([0, T[, H)$, $\lambda_n \zeta \rightarrow \lambda \zeta$ strongly in $L^2(0, T; H)$; as a consequence

$$\int_0^T \langle \zeta, \lambda_n \dot{z}_n \rangle = \int_0^T \langle \zeta \lambda_n, \dot{z}_n \rangle \rightarrow \int_0^T \langle \zeta \lambda, \dot{z} \rangle = \int_0^T \langle \zeta, \lambda \dot{z} \rangle.$$

Hence $\lambda_n \dot{z}_n \rightharpoonup \lambda \dot{z}$ weakly in $L^2(0, T; H)$, since $(\lambda_n \dot{z}_n)_{n \in \mathbb{N}}$ is bounded in $L^2(0, T; H)$, and $\mathcal{C}_c([0, T[, H)$ is dense in $L^2(0, T; H)$. $\square$

Theorem 3.2 *Let $\lambda : [0, T] \rightarrow (0, \infty)$ be of bounded variation on $[0, T]$, and suppose $c_0 = \inf \lambda([0, T]) > 0$. Then there exists a unique $(x, v) : [0, T] \rightarrow H \times H$ strong solution of the differential inclusion*

$$v(t) \in A(x(t)), \quad 0 \leq t \leq T \quad (21a)$$

$$\lambda(t)\dot{x}(t) + \dot{v}(t) + v(t) = 0, \quad \text{a.e. } 0 \leq t \leq T \quad (21b)$$

$$x(0) = x_0, \quad v(0) = v_0, \quad v_0 \in A(x_0), \quad v_0 \neq 0 \quad (21c)$$

Moreover (x, v) is L -Lipschitz continuous, with $L = \|v_0\| / \min(c_0, 1)$.

Proof Existence: According to Lemma 3.1, there exists a sequence $(\lambda_n)_{n \in \mathbb{N}}$ in $\mathcal{C}^\infty([0, T], [c_0, +\infty])$ which converges to λ in $L^p(0, T)$ for $p \geq 1$ and satisfies the additional condition

$$\int_0^T |\dot{\lambda}_n(t)| dt \leq TV(\lambda). \quad (22)$$

For each λ_n there exists a unique (x_n, v_n) solution of the differential inclusion

$$v_n(t) \in A(x_n(t)), \quad 0 \leq t \leq T \quad (23a)$$

$$\lambda_n(t)\dot{x}_n(t) + \dot{v}_n(t) + v_n(t) = 0, \quad \text{a.e. } 0 \leq t \leq T \quad (23b)$$

$$x_n(0) = x_0, \quad v_n(0) = v_0. \quad (23c)$$

We will show that (x_n, v_n) converges uniformly to a solution of (21).

Consider, in $C([0, T], H \times H)$, the norm

$$\|(z, w)\|_{c_0} = \max_{t \in [0, T]} \sqrt{c_0^2 \|z(t)\|^2 + \|w(t)\|^2}.$$

It is equivalent to the *sup* norm in $C([0, T], H \times H)$. Using Theorem 3.1(16) and (22) we deduce that for any n, m and for any $t \in [0, T]$ we have

$$\begin{aligned} \|(x_n, v_n)(t) - (x_m, v_m)(t)\|_{c_0} &\leq \frac{\|v_0\|}{c_0} \exp\left(T + \frac{\|\dot{\lambda}_n + \dot{\lambda}_m\|_1}{2c_0}\right) \|\lambda_n - \lambda_m\|_1 \\ &\leq \frac{\|v_0\|}{c_0} \exp\left(T + \frac{TV(\lambda)}{c_0}\right) \|\lambda_n - \lambda_m\|_1, \end{aligned}$$

where $\|\cdot\|_1 = \|\cdot\|_{L^1(0,T)}$. Hence (x_n, v_n) is a Cauchy sequence with respect to the *sup* norm. Therefore, (x_n, v_n) converges uniformly to some continuous $(x, v) : [0, T] \rightarrow H \times H$. Using Proposition 3.1 we have that (x_n, v_n) is Lipschitz continuous with constant $L = \|v_0\| / \min(c_0, 1)$. Therefore, (x, v) is L -Lipschitz continuous.

Now Lemma 3.2 ensures that $\lambda_n \dot{x}_n$ and $\dot{v}_n$ converge weakly to $\lambda \dot{x}$ and $\dot{v}$ in $L^2(0, T; H)$; hence letting $n \rightarrow \infty$ in (23b), we obtain (21b). Finally, since the graph of A is closed, (23a) entails (21a).

Uniqueness. We adapt the proof of Theorem 3.1, using differential and integral calculus for BV functions which involves differential measures. A convenient reference is [13].

Define $\lambda^- : [0, T] \mapsto [c_0, +\infty[$ by

$$\begin{aligned} \lambda^-(0) &= \lambda(0) \\ 0 < t \leq T : \lambda^-(t) &= \lim_{\varepsilon > 0, \varepsilon \rightarrow 0} \lambda(t - \varepsilon). \end{aligned}$$

Let $(x, v), (y, w) : [0, T] \rightarrow H \times H$ be two strong solutions of (21). Explicitly:

$$\begin{aligned} \lambda \dot{x} + \dot{v} + v &= 0 \text{ a.e.; } v(t) \in A(x(t)) \forall t; x(0) = x_0, v(0) = v_0 \\ \lambda \dot{y} + \dot{w} + w &= 0 \text{ a.e.; } w(t) \in A(y(t)) \forall t; y(0) = x_0, w(0) = v_0. \end{aligned}$$

Since $\lambda = \lambda^-$ a.e., we also have

$$\lambda^- \dot{x} + \dot{v} + v = 0 \quad \text{and} \quad \lambda^- \dot{y} + \dot{w} + w = 0 \text{ a.e.}$$

and consequently

$$\lambda^- (\dot{x} - \dot{y}) + (\dot{v} - \dot{w}) + v - w = 0 \text{ a.e.} \quad (24)$$

In terms of differential measures on $[0, T]$ we have ([13, Proposition 11.1])

$$d[\lambda^-(x - y) + (v - w)] = \lambda^- d(x - y) + (x - y) d\lambda^- + d(v - w). \quad (25)$$

Integrating the left hand term on $[0, s]$ and taking the initial condition into account, we obtain for $s \in [0, T]$ ([13, Corollary 8.2])

$$\int_{[0, s[} d[\lambda^-(x - y) + (v - w)] = \lambda^-(s)(x(s) - y(s)) + (v(s) - w(s)). \quad (26)$$

Now integrating the right hand term of (25) on $[0, s]$ and taking (24) into account, we get

$$\begin{aligned} & \int_{[0, s[} [\lambda^- d(x - y) + (x - y) d\lambda^- + d(v - w)] \\ &= \int_{[0, s[} [\lambda^- d(x - y) + d(v - w)] + \int_{[0, s[} (x - y) d\lambda^- \\ &= \int_{[0, s[} [\lambda^- (\dot{x} - \dot{y}) + (\dot{v} - \dot{w})] dt + \int_{[0, s[} (x - y) d\lambda^- \\ &= - \int_{[0, s[} (v - w) dt + \int_{[0, s[} (x - y) d\lambda^-. \end{aligned} \quad (27)$$

From (25, 26, 27) we deduce

$$\lambda^-(s)(x(s) - y(s)) + (v(s) - w(s)) = \int_{[0,s[} (x - y) d\lambda^- - \int_{[0,s[} (v - w) dt.$$

Whence

$$\|\lambda^-(s)(x(s) - y(s)) + (v(s) - w(s))\| \leq \int_{[0,s[} \|x - y\| |d\lambda^-| + \int_{[0,s[} \|v - w\| dt. \quad (28)$$

Define $\varphi(s) = (c_0^2 \|x(s) - y(s)\|^2 + \|v(s) - w(s)\|^2)^{1/2}$. The same reasoning as in Theorem 3.1 (eq. (17)) yields

$$\varphi(s) \leq \|\lambda^-(s)(x(s) - y(s)) + (v(s) - w(s))\|, \quad \forall s \in [0, T].$$

Besides we also obviously have $\varphi(s) \geq c_0 \|x(s) - y(s)\|$ and $\varphi(s) \geq \|v(s) - w(s)\|$. Hence, with (28)

$$\varphi(s) \leq \frac{1}{c_0} \int_{[0,s[} \varphi |d\lambda^-| + \int_{[0,s[} \varphi dt = \int_{[0,s[} \varphi d\mu, \quad (29)$$

where $d\mu$ denotes the nonnegative measure $\frac{1}{c_0} |d\lambda^-| + dt$.

If $\varphi \not\equiv 0$ on $[0, T]$, define $t_0 = \inf\{t \in [0, T], \varphi(t) > 0\}$. Note $t_0 < T$ and $\varphi(t_0) = 0$, since φ is continuous. With (29) we then have

$$\varphi(s) \leq \int_{]t_0,s[} \varphi d\mu, \quad t_0 < s \leq T. \quad (30)$$

In view of $\int_{]t_0,t_0]} d\mu = 0$ and of the right continuity at t_0 of $t \mapsto \int_{]t_0,t]} d\mu$ ([13, Proposition 9.1]), there exists some $t_1 \in]t_0, T]$ such that $\int_{]t_0,t_1]} d\mu < 1/2$. Let M be an upper bound of φ on $[0, t_1]$; from (30) we deduce, for $s \in]t_0, t_1]$

$$\varphi(s) \leq M \int_{]t_0,s[} d\mu \leq M \int_{]t_0,t_1]} d\mu \leq \frac{M}{2}.$$

Hence $M/2$ is also an upper bound of φ on $[0, t_1]$, which necessarily entails $M = 0$ and $\varphi \equiv 0$ on $[0, t_1]$. But this is a contradiction with the very definition of t_0 . Hence $\varphi \equiv 0$ and $(x, v) \equiv (y, w)$ on $[0, T]$. $\square$

Theorem 3.2 has a natural global version formulation:

Corollary 3.2 *Suppose that $\lambda : [0, \infty[\rightarrow]0, \infty[$ is of bounded variation on $[0, T]$ and $\inf \lambda([0, T]) > 0$ for any $T < \infty$. Let $(x_0, v_0) \in A$ and $v_0 \neq 0$. Then there exists a unique $(x, v) : [0, \infty[\rightarrow H \times H$ strong solution of the differential inclusion*

$$\lambda \dot{x} + \dot{v} + v = 0, \quad v \in A(x), \quad x(0) = x_0, \quad v(0) = v_0 \quad (31)$$

where the first equality holds for almost all $t \in [0, \infty[$, and the inclusion holds for all $t \in [0, \infty[$.

4 A Closed-loop Regularized Newton Dynamics: Existence and Uniqueness

We come to the central topic of the paper which is the existence and uniqueness of a global solution of the Cauchy problem for (1). We consider the case where G is Lipschitz continuous on bounded sets. In the next section we study convergence properties of the trajectories for a particular G , possibly discontinuous.

Let us state some assumptions, definitions and facts in force in this paragraph:

- (i) $v_0 \in Ax_0$, $v_0 \neq 0$;
- (ii) $G : H \times (H \setminus \{0\}) \rightarrow]0, \infty[$ satisfies a) and b):
 - a) for any $0 < r < R$, G is Lipschitz continuous on the bounded set

$$\{(x, v) \in H \times H, \|x\| \leq R; r \leq \|v\| \leq R\};$$

- b) for any $\delta > 0$,

$$\inf \{G(x, v) \mid (x, v) \in H \times H, \|v\| \geq \delta\} > 0.$$

We are concerned by the strong solutions of

$$G(x, v)\dot{x} + \dot{v} + v = 0, \quad v \in A(x), \quad x(0) = x_0, \quad v(0) = v_0. \quad (32)$$

Specifically, (x, v) is a strong solution of (32) iff it is a strong solution of (11) with $\lambda = G(x, v)$.

Theorem 4.1 *Suppose (i), (ii). Then, for any $(x_0, v_0) \in A$, $v_0 \neq 0$, the system (32) has a unique strong global solution $(x, v) : [0, \infty[\rightarrow H \times H$. Moreover $t \mapsto G(x(t), v(t))$ is Lipschitz continuous on any bounded time interval.*

Proof Let us first prove local existence and uniqueness of a strong solution (x, v) to (32). To that end, we fix $T_0 > 0$, and look for some time interval $[0, T]$, $T < T_0$ with T small enough, in order to be able to apply the Banach-Picard fixed point theorem in $\mathcal{C}([0, T]; H \times H)$ equipped with the *sup* norm. In view of (ii)(b) we have

$$\exists c_0 > 0 / (\xi, \zeta) \in H \times H, \|\zeta\| \geq \|v_0\|e^{-T_0} \Rightarrow G(\xi, \zeta) > c_0. \quad (33)$$

Without restriction of generality we suppose $c_0 \leq 1$; note also $G(x_0, v_0) > c_0$. For any $T \in (0, T_0]$ define $\mathcal{S}_T$ as the set of functions $(x, v) : [0, T] \mapsto H \times H$ satisfying

- i) (x, v) is Lipschitz continuous;
- ii) $x(0) = x_0$, $v(0) = v_0$;
- iii) $\|v(t)\| \geq \|v_0\|e^{-t}$ for all $t \in [0, T]$;
- iv) $\|\dot{v}(t)\| \leq \|v_0\|$, $\|\dot{x}(t)\| \leq \frac{\|v_0\|}{c_0}$ for almost every $t \in [0, T]$.

Equipped with the distance induced by the *sup* norm,

$$\|(x, v)\|_\infty = \sup_{0 \leq t \leq T} \|(x(t), v(t))\|,$$

$\mathcal{S}_T$ is a closed subset of the Banach space $(\mathcal{C}([0, T]; H \times H), \|(x, v)\|_\infty)$. Hence, $\mathcal{S}_T$ is a complete metric space. Any $(x, v) \in \mathcal{S}_T$ takes its values in

$$B = \left\{ (\xi, \zeta) \in H \times (H \setminus \{0\}) / \|(\xi, \zeta) - (x_0, v_0)\| \leq T \frac{\|v_0\|}{c_0}; \|\zeta\| \geq \|v_0\| e^{-T_0} \right\}.$$

According to assumption (ii)(a), G is Lipschitz continuous on B . Let L_0 denote its Lipschitz constant on B .

Let $(x, v) \in \mathcal{S}_T$; v does not vanish on $[0, T]$. Hence $\lambda = G(x, v)$ is defined on $[0, T]$ and verifies $\lambda > c_0$, in view of (33); moreover it is absolutely continuous as the composition of two Lipschitz continuous functions. Hence, with Proposition 3.1, there exists a unique strong solution (X, V) to (11), which further belongs to $\mathcal{S}_T$. Let us examine the continuity of the map

$$\Psi_T : (x, v) \in \mathcal{S}_T \rightarrow (X, V) \in \mathcal{S}_T$$

we have just defined. Let $((x, v), (y, w)) \in \mathcal{S}_T \times \mathcal{S}_T$, and set $(X, V) = \Psi_T(x, v)$, $(Y, W) = \Psi_T(y, w)$, $\lambda = G(x, v)$, $\eta = G(y, w)$. We have (note that λ, η are $L_0 \frac{\|v_0\|}{c_0}$ -Lipschitz continuous):

$$\begin{aligned} \|\lambda - \eta\|_{L^1(0, T)} &\leq T \|\lambda - \eta\|_{L^\infty(0, T)} \leq T L_0 \|(x, v) - (y, w)\|_\infty; \\ \|\dot{\lambda}\|_{L^1(0, T)} &\leq T \|\dot{\lambda}\|_{L^\infty(0, T)} \leq T L_0 \frac{\|v_0\|}{c_0}; \\ \|\dot{\eta}\|_{L^1(0, T)} &\leq T L_0 \frac{\|v_0\|}{c_0}; \\ \|\dot{\lambda} + \dot{\eta}\|_{L^1(0, T)} &\leq 2 T L_0 \frac{\|v_0\|}{c_0}. \end{aligned}$$

Hence, in view of (16)

$$c_0 \|(X, V) - (Y, W)\|_\infty \leq \frac{\|v_0\|}{c_0} \exp \left(T + \frac{T L_0 \|v_0\|}{c_0^2} \right) T L_0 \|(x, v) - (y, w)\|_\infty.$$

So, for any $T \in]0, T_0]$ small enough, say $0 < T \leq T_1$, $\Psi_T : \mathcal{S}_T \rightarrow \mathcal{S}_T$ is a contraction and admits a unique fixed point, which is moreover a strong solution to (32). This proves existence.

Now, in view of Proposition 3.1, for $T \leq T_1$ any strong solution of (32) necessarily belongs to $\mathcal{S}_T$ and must coincide with *the* fixed point of Ψ_T . This proves uniqueness.

Let us now consider a maximal solution of (32)

$$(x, v) : [0, T_m[\rightarrow H \times H, \quad v(t) \neq 0.$$

The existence of a maximal solution follows from a classical argument relying on the local existence and uniqueness result. Let us prove that the maximal solution is a global solution, i.e., $T_m = +\infty$. Define

$$\lambda : [0, T_m[\rightarrow [0, \infty[, \quad \lambda(t) = G(x, v).$$

Then, as already observed in the above argument, $t \mapsto \lambda(t)$ is locally Lipschitz continuous and hence absolutely continuous on bounded closed sub-intervals of $[0, T_m[$. Since $t \mapsto (x(t), v(t))$ is also a strong solution of (11), using (12a) of Proposition 3.1 we have

$$\|v(t)\| \geq \|v_0\|e^{-t}, \quad \forall t \in [0, T_m[.$$

If $T_m < \infty$, define

$$\delta = \frac{\|v_0\|}{2}e^{-T_m}.$$

Using the two equations above we conclude that

$$(x, v) \in H \times (H \setminus \overline{B_\delta(0)}), \quad \forall t \in [0, T_m[.$$

Therefore, there exists $\varepsilon > 0$ such that

$$\lambda(t) = G(x, v) > \varepsilon, \quad \forall t \in [0, T_m[.$$

Using Proposition 3.1 again, we conclude that $t \mapsto (x(t), v(t))$ is Lipschitz continuous on $[0, T_m[$. Therefore, there exists

$$(\bar{x}, \bar{v}) = \lim_{t \rightarrow T_m} (x(t), v(t)).$$

Moreover, since the graph of A is closed, $\bar{v} \in A(\bar{x})$ and from the preceeding argument, $\bar{v} \neq 0$. Therefore, the differential inclusion

$$G(z, w)\dot{z} + \dot{w} + w = 0, \quad w \in A(z), \quad z(0) = \bar{x}, \quad w(0) = \bar{v}$$

has a strong solution $(z, w) : [0, T'] \rightarrow H \times H$, for some $T' > 0$, which allows one to extend (x, v) to $[0, T_m + T']$, contradicting the maximality of (x, v) . $\square$

5 Convergence and Optimization Properties of a Closed-loop Regularized Newton Dynamics

Let $A : H \rightrightarrows H$ be a maximal monotone operator. We assume that the set of equilibria $S = A^{-1}(0)$ is nonempty. We consider the differential inclusion

$$v(t) \in A(x(t)), \tag{34a}$$

$$\alpha(\|v(t)\|^2) \dot{x}(t) + \dot{v}(t) + v(t) = 0, \tag{34b}$$

$$x(0) = x_0, \quad v(0) = v_0, \quad v_0 \in A(x_0), \quad v_0 \neq 0, \tag{34c}$$

where $\alpha : [0, +\infty[\rightarrow [0, +\infty[$ is a given function. For this particular choice of the feed-back regularization function $G(x, v) = \alpha(\|v\|^2)$, we shall be able to

analyze the convergence properties of the system. In the following paragraphs we successively examine the case where α is continuously differentiable and the case where α is only a BV function.

As a model example of our study, we will consider

$$\alpha(s) = cs^\theta \quad (35)$$

for some $c, \theta > 0$. The point is to determine for which values of θ the trajectories of system (34) converge to an element of S .

When $A = \nabla f$ and $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a convex, twice differentiable function, system (34) becomes

$$\alpha(\|\nabla f(x(t))\|^2)\dot{x}(t) + \nabla^2 f(x(t))\dot{x}(t) + \nabla f(x(t)) = 0. \quad (36)$$

In that case S is the set of minimizers of f , and the asymptotic stabilization property is equivalent to the asymptotic optimization of f .

5.1 Smooth Feed-back Function

Note that (34) is a particular case of the differential inclusion (32) with $G(x, v) = \alpha(\|v\|^2)$. This particular G satisfies condition (ii)(a) of Section 4, but it may fail to satisfy (ii)(b), if $\liminf_{r \rightarrow +\infty} \alpha(r) = 0$. In that case, v_0 being fixed, define $\beta(r) = \alpha(r)$ if $0 \leq r \leq \|v_0\|^2$ and $\beta(r) = \alpha(\|v_0\|^2)$ if $r \geq \|v_0\|^2$. Any strong solution (x, v) of (34) is also a strong solution of (34) with β in place of α , and vice versa (this is a consequence of $\|v\|$ being decreasing). Since β satisfies (ii)(a) and (ii)(b), function $G(x, v) = \alpha(\|v\|^2)$ actually is relevant to Theorem 4.1, and system (34) has a unique solution $(x, v) : [0, \infty[\rightarrow H \times H$. Define

$$\lambda(t) = \alpha(\|v(t)\|^2).$$

Then $\lambda : [0, \infty[\rightarrow [0, \infty[$ is absolutely continuous on bounded sets and (x, v) is the unique strong global solution of the system

$$\begin{aligned} v &\in A(x); \\ \lambda\dot{x} + \dot{v} + v &= 0; \\ v(0) &= v_0 \in A(x_0), \quad v_0 \neq 0. \end{aligned}$$

In order to obtain convergence of the trajectories of system (34), we now appeal to Theorem 3.7 of [1], which tells us that if $A^{-1}(0) \neq \emptyset$, if $\lambda(\cdot)$ is bounded from above on $[0, +\infty[$ and if

$$\liminf_{t \rightarrow +\infty} \frac{\dot{\lambda}(t)}{\lambda(t)} > -1, \quad (37)$$

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

We are now in a position to state our main result.

Theorem 5.1 *Let us suppose $A^{-1}(0) \neq \emptyset$ and that $\alpha :]0, +\infty[\rightarrow]0, +\infty[$ is a C^1 function. Let (x, v) be the unique global strong solution to (34). Then*

a) *v satisfies*

$$\lim_{t \rightarrow +\infty} \|v(t)\| = 0. \quad (38)$$

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

$$\limsup_{r \rightarrow 0^+} \frac{r\dot{\alpha}(r)}{\alpha(r)} < \frac{1}{2}. \quad (39)$$

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

Proof The existence and uniqueness of the strong global solution (x, v) to (34) is clear from the previous discussion.

a) Since $\|v\|$ is decreasing, $p = \lim_{t \rightarrow +\infty} \|v(t)\|$ exists. We reason by contradiction and suppose $p \neq 0$.

We have $\dot{\lambda} = 2\dot{\alpha}(\|v\|^2) \langle v, \dot{v} \rangle$ and since $\langle v, \dot{v} \rangle \leq 0$ (recall (12c)) we have further

$$\begin{aligned} |\dot{\lambda}| &\leq 2|\dot{\alpha}(\|v\|^2)|(-\langle v, \dot{v} \rangle) \\ \int_0^t |\dot{\lambda}(s)| ds &\leq \max_{p^2 \leq r \leq \|v_0\|^2} |\dot{\alpha}(r)| [\|v_0\|^2 - \|v(t)\|^2], \quad \forall t \geq 0. \end{aligned}$$

Hence $\dot{\lambda} \in L^1(0, \infty)$. Since we have moreover $\lim_{t \rightarrow +\infty} \lambda(t) = \alpha(p^2) > 0$ (and also $A^{-1}(0) \neq \emptyset$), applying Proposition 3.2, we obtain $p = 0$, a contradiction.

b) In view of $\langle v, \dot{v} \rangle \leq 0$ and of $\|\dot{v}\| \leq \|v\|$ (see (12b)) we have

$$\dot{\lambda} \geq 2\dot{\alpha}^+(\|v\|^2) \langle v, \dot{v} \rangle \geq -2\dot{\alpha}^+(\|v\|^2) \|v\| \|\dot{v}\| \geq -2\dot{\alpha}^+(\|v\|^2) \|v\|^2.$$

Since $\lim_{t \rightarrow +\infty} \|v(t)\| = 0$, a sufficient condition ensuring (37) is $\liminf_{r \rightarrow 0^+} -2r\dot{\alpha}^+(r)/\alpha(r) > -1$, or equivalently

$$\limsup_{r \rightarrow 0^+} \frac{r\dot{\alpha}^+(r)}{\alpha(r)} < 1/2. \quad (40)$$

Since $\dot{\alpha} \leq \dot{\alpha}^+$ this condition obviously implies (39). But negating (40) implies the existence of a sequence $r_n > 0$ tending to 0 as $n \rightarrow +\infty$ and such that

$$\frac{r_n \dot{\alpha}^+(r_n)}{\alpha(r_n)} > \frac{1}{2} - \frac{1}{n}.$$

Then $\dot{\alpha}(r_n) = \dot{\alpha}^+(r_n) > 0$, and we obtain the negation of (39). So (39) and (40) are equivalent. Moreover, λ is bounded on $[0, +\infty[$, since $\lim_{t \rightarrow +\infty} \|v(t)\| = 0$ and α is bounded near 0. With [1, th. 3.7], this property, together with $A^{-1}(0) \neq \emptyset$ and (39) $\Rightarrow$ (37), implies the weak convergence of $x(t)$ to a zero of A , as $t \rightarrow +\infty$. $\square$

Let us return to our model example.

Corollary 5.1 *Take $\alpha(r) = r^\theta$. Then the conclusion of Theorem 5.1 holds for $\theta < \frac{1}{2}$. Equivalently, for $0 < \gamma < 1$, there is asymptotic stabilization of the system*

$$\begin{cases} v(t) \in A(x(t)) \\ \|v(t)\|^\gamma \dot{x}(t) + \dot{v}(t) + v(t) = 0; \\ x(0) = x_0, v(0) = v_0 \in A(x_0), v_0 \neq 0. \end{cases} \quad (41)$$

Note that, when A is the subdifferential of a convex function f , the convergence of the values $f(x(t))$ to $\inf_H f$ is obtained under the sole assumption that α is nondecreasing.

Proposition 5.1 *Suppose that $A = \partial f$ where $f : H \rightarrow \mathbb{R} \cup \{+\infty\}$ is a convex lower semicontinuous function. Assume $\alpha :]0, +\infty[\rightarrow]0, +\infty[$ nondecreasing. Let (x, v) be the global strong solution of (34). Then*

- (i) $f(x(t))$ decreases to $\inf_H f$ as $t \rightarrow +\infty$;
- (ii) if f is bounded from below, then $\|v(\cdot)\| \in L^2(0, +\infty)$ and $v(t) \rightarrow 0$ as $t \rightarrow +\infty$.

Proof Since $\|v\|$ is nonincreasing, $\lambda = \alpha(\|v\|^2)$ is also nonincreasing, and [1, Theorem 4.1] applies. $\square$

5.2 Bounded Variation Feed-back Function

In this paragraph, H is a finite-dimensional Hilbert space. We consider the closed-loop regularized continuous Newton dynamics

$$v \in A(x), \quad \alpha(\|v(t)\|^2) \dot{x}(t) + \dot{v}(t) + v(t) = 0, \quad x(0) = x_0, v(0) = v_0, \quad (42)$$

where $v_0 \in A(x_0)$, $v_0 \neq 0$ and $\alpha : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is only supposed to be a nondecreasing function (with possible jumps). By using regularization by convolution, $\alpha_n = \rho_n * \alpha \rightarrow \alpha$, we are reduced to the situation α_n smooth, studied in the previous paragraph. The question is to pass to the limit as $n \rightarrow +\infty$, because of the possible discontinuities of α .

Let us introduce some notations: for $r > 0$ we set $\alpha^-(r) = \lim_{s \rightarrow r, s < r} \alpha(s)$, the left limit of α at r , and $\alpha^+(r) = \lim_{s \rightarrow r, s > r} \alpha(s)$, the right limit of α at r . Since α is nondecreasing these limits exist, and $\alpha^-(r) \leq \alpha^+(r)$ for each $r > 0$. We set

$$[\alpha](r) = [\alpha^-(r), \alpha^+(r)].$$

In the following theorem, we establish existence of a strong (*i.e.* absolutely continuous on compact subintervals of $[0, +\infty)$) solution to system (42) in a relaxed sense, and provide some information concerning its asymptotical behavior. Note that we do not claim uniqueness of the solution.

Theorem 5.2 *Suppose that H is finite-dimensional and that $\alpha : [0, +\infty[\rightarrow [0, +\infty[$ is a nondecreasing function (possibly discontinuous). Then*

(i) *there exists a global strong solution $(x, v) : [0, \infty[\rightarrow H \times H$ of system (42) in the following sense*

$$v(t) \in A(x(t)) \text{ for all } t \geq 0 \quad (43a)$$

$$[\alpha](\|v(t)\|^2)\dot{x}(t) + \dot{v}(t) + v(t) \ni 0 \text{ for almost all } t \geq 0 \quad (43b)$$

$$x(0) = x_0, \quad v(0) = v_0 \in A(x_0), \quad v_0 \neq 0. \quad (43c)$$

Moreover (x, v) is Lipschitz continuous on any bounded interval of $[0, +\infty[$. Precisely equation (43b) means that there exists a nonincreasing function $\lambda : [0, +\infty[\rightarrow]0, +\infty[$ such that

$$\lambda(t) \in [\alpha](\|v(t)\|^2) \text{ for all } t \geq 0$$

$$\lambda(t)\dot{x}(t) + \dot{v}(t) + v(t) = 0 \text{ for almost all } t \geq 0.$$

(ii) $t \mapsto \|v(t)\|^2$ is a nonincreasing function and $\alpha(\|v\|^2)\dot{x} + \dot{v} + v = 0$ almost everywhere on each interval where $t \mapsto \|v(t)\|^2$ is strictly decreasing.

(iii) if $A^{-1}(0) \neq \emptyset$, then $\|v(t)\| \rightarrow 0$ as $t \rightarrow +\infty$.

Proof Extend α to the function $\bar{\alpha}$ defined on $\mathbb{R}$ by

$$\begin{aligned} t \leq 0 & : \bar{\alpha}(t) = 0; \\ t > 0 & : \bar{\alpha}(t) = \alpha(t). \end{aligned}$$

Function $\bar{\alpha}$ is nondecreasing. Let $(\rho_n)_{n \in \mathbb{N}}$ be a mollifying sequence (see Lemma 3.1). Classically, the convolution $\alpha_n = \rho_n * \bar{\alpha}$ belongs to $\mathcal{C}^\infty(\mathbb{R})$, α_n is nondecreasing, and $\alpha_n(r) > 0$ for $r > 0$.

Let us apply Theorem 4.1 with $G(x, v) = \alpha_n(\|v\|^2)$. Clearly $G : H \times (H \setminus \{0\}) \rightarrow]0, \infty[$ is locally Lipschitz continuous and for any $\delta > 0$

$$\inf \{G(x, v) \mid x, v \in H, \|v\| \geq \delta\} \geq \alpha_n(\delta^2) > 0.$$

Thus, for each $n \in \mathbb{N}$ there exists a unique global strong solution $(x_n, v_n) : [0, +\infty[\rightarrow H \times H$ of

$$v_n \in A(x_n), \quad \alpha_n(\|v_n\|^2)\dot{x}_n + \dot{v}_n + v_n = 0, \quad x_n(0) = x_0, \quad v_n(0) = v_0 \quad (44)$$

Let us first restrict t to $[0, T]$ for some arbitrary fixed $T > 0$ and gather some facts about x_n, v_n and $\lambda_n = \alpha_n(\|v_n\|^2)$.

From (12a, 12c) we deduce for any $t \in [0, T]$ and any $n \in \mathbb{N}$

$$e^{-T}\|v_0\| \leq \|v_n(t)\| \leq \|v_0\|.$$

Since α_n is nondecreasing, we infer for any $t \in [0, T]$ and any $n \in \mathbb{N}$

$$\alpha_n(e^{-2T}\|v_0\|^2) \leq \lambda_n(t) \leq \alpha_n(\|v_0\|^2). \quad (45)$$

By the definition of $\alpha_n = \rho_n * \bar{\alpha}$ (recall $\text{support}(\rho_n) \subseteq [-s_n, s_n]$, $s_n > 0$, $s_n \rightarrow 0$) and due to $\bar{\alpha}$ being nondecreasing we have

$$\alpha_n(e^{-2T}\|v_0\|^2) \geq \bar{\alpha}(e^{-2T}\|v_0\|^2 - s_n) \geq \alpha\left(\frac{1}{2}e^{-2T}\|v_0\|^2\right) = c_0 > 0 \quad (46)$$

$$\alpha_n(\|v_0\|^2) \leq \bar{\alpha}(\|v_0\|^2 + s_n) \leq \alpha(\|v_0\|^2 + 1) = d_0 \quad (47)$$

where the second inequalities are true as soon as n is so large, say $n \geq N_0$, that $s_n \leq \min(1, 1/2e^{-2T}\|v_0\|^2)$. From (45, 46, 47) we obtain for any $t \in [0, T]$ and $n \geq N_0$

$$0 < c_0 \leq \lambda_n(t) \leq d_0 \quad (48)$$

whence we also derive, since λ_n is nonincreasing

$$\text{TV}(\lambda_n, [0, T]) \leq d_0. \quad (49)$$

Applying Proposition 3.1 with $\lambda = \lambda_n$, we infer, for all $t \in [0, T]$ and $n \geq N_0$

$$\|(\dot{x}_n, \dot{v}_n)\| \leq \frac{\|v_0\|}{\min\{c_0, 1\}} = L. \quad (50)$$

Hence, the sequence $(x_n, v_n) : [0, T] \rightarrow H \times H$ is equi-Lipschitz continuous, with values in a bounded subset of $H \times H$, as $(x_n(0), v_n(0)) = (x_0, v_0)$. Since H is finite-dimensional, the sequence (x_n, v_n) satisfies the assumptions of the Ascoli-Arzelà theorem, and consequently is relatively compact in $\mathcal{C}([0, T], H)$ for the topology of the uniform convergence on $[0, T]$.

Moreover, with (48) and (49), the sequence λ_n satisfies the assumptions of Helly's first theorem ([14, p. 222]) and, thus, it admits a subsequence that converges pointwise on $[0, T]$ to a nonincreasing function¹.

We now allow t to vary in $[0, +\infty[$. After considering successive intervals of the form $[0, pT]$, $p \in \mathbb{N}$, and after extraction of a diagonal sequence, we infer the existence of a subsequence $(x_{n_k}, v_{n_k}, \lambda_{n_k}) : [0, \infty[\rightarrow H \times H \times]0, +\infty[$ and of functions $(x, v) \in \mathcal{C}([0, +\infty[, H \times H)$ and $\lambda : [0, +\infty[\rightarrow]0, +\infty[$ nonincreasing such that

$$\begin{aligned} (x_{n_k}, v_{n_k}) &\rightarrow (x, v) \text{ uniformly on } [0, T] \text{ for all } T > 0, \\ \lambda_{n_k}(t) &\rightarrow \lambda(t) \text{ for all } t \geq 0. \end{aligned}$$

Without ambiguity we now omit the subscript k . From (50) we deduce that (x, v) is L -Lipschitz continuous on each $[0, T]$. By the closedness property of A and $v_n \in A(x_n)$, we immediately infer (43a).

The most delicate point is to pass to the limit ($n \rightarrow \infty$) on the equation

$$\lambda_n \dot{x}_n + \dot{v}_n + v_n = 0, \quad (51)$$

so as to prove (43b).

(i) Lebesgue's dominated convergence theorem ensures $\lambda_n \rightarrow \lambda$ in $L^2(0, T)$ for any $T > 0$. Lemma 3.2 allows one to pass to the limit in the sense of distributions on (51)

$$\lambda(t)\dot{x}(t) + \dot{v}(t) + v(t) = 0 \text{ for almost all } t \geq 0.$$

¹ The compact imbedding of $\text{BV}([0, T])$, the space of real-valued functions of bounded variation on $[0, T]$ equipped with the topology of the intermediate convergence, into $L^1(0, T)$ (see e.g. [10, Theorem 10.1.4], [11, 12]) yields a subsequence of λ_n converging *almost* everywhere on $[0, T]$.

Let us prove $\lambda(t) \in [\alpha](\|v(t)\|^2)$ for all $t \geq 0$. Since $v_n(t) \rightarrow v(t)$, there exists $\epsilon_n > 0$, $\epsilon_n \rightarrow 0$ such that

$$\|v(t)\|^2 - \epsilon_n \leq \|v_n(t)\|^2 \leq \|v(t)\|^2 + \epsilon_n. \quad (52)$$

Fix $t > 0$. By the definition of α_n , for n large enough, we have

$$\alpha_n(\|v_n(t)\|^2) = \int_{-s_n}^{s_n} \alpha(\|v_n(t)\|^2 - s) \rho_n(s) ds.$$

Since α is nondecreasing and with (52) we deduce that

$$\alpha(\|v(t)\|^2 - \epsilon_n - s_n) \leq \alpha_n(\|v_n(t)\|^2) \leq \alpha(\|v(t)\|^2 + \epsilon_n + s_n).$$

By passing to the limit as $n \rightarrow \infty$ we obtain

$$\alpha^-(\|v(t)\|^2) \leq \lambda(t) \leq \alpha^+(\|v(t)\|^2),$$

whence $\lambda(t) \in [\alpha](\|v(t)\|^2)$.

(ii) Since $\|v_n\|$ is nonincreasing and v_n converges to v uniformly on bounded intervals, $\|v\|$ is nonincreasing. Suppose that $\tau \mapsto \|v(\tau)\|^2$ is strictly decreasing on some interval $[s, t]$. Then the set $\{\tau \in [s, t] : \alpha \text{ is discontinuous at } \|v(\tau)\|^2\}$ is at most countable. On the complement of this set, the same argument as above yields

$$\alpha(\|v(\tau)\|^2) = \lim \alpha_n(\|v_n(\tau)\|^2)$$

whence $\lambda(\tau) = \alpha(\|v(\tau)\|^2)$.

(iii) Let us fix $\hat{x} \in A^{-1}(0)$. Following the argument of [1, Lemma 3.4], for each $n \in \mathbb{N}$, we introduce the Liapunov function

$$h_n(t) := \frac{1}{2} \|\lambda_n(t)(x_n(t) - \hat{x}) + v_n(t)\|^2.$$

Differentiating h_n , and using (44) we obtain

$$\begin{aligned} \frac{d}{dt} h_n(t) &= \langle \lambda_n(t)(x_n(t) - \hat{x}) + v_n(t), \lambda_n(t)\dot{x}_n(t) + \dot{v}_n(t) \rangle \\ &\quad + \dot{\lambda}_n(t) \langle \lambda_n(t)(x_n(t) - \hat{x}) + v_n(t), x_n(t) - \hat{x} \rangle \\ &= - \langle \lambda_n(t)(x_n(t) - \hat{x}) + v_n(t), v_n(t) \rangle \\ &\quad + \dot{\lambda}_n(t) [\lambda_n(t) \|x_n(t) - \hat{x}\|^2 + \langle v_n(t), x_n(t) - \hat{x} \rangle]. \end{aligned}$$

By monotonicity of A and $0 \in A(\hat{x})$, $v_n(t) \in A(x_n(t))$, we have

$$\langle x_n(t) - \hat{x}, v_n(t) \rangle \geq 0.$$

Moreover, λ_n being a nonincreasing function, we deduce

$$\frac{d}{dt} h_n(t) + \|v_n(t)\|^2 \leq 0.$$

After integration with respect to t of the above inequality, using that $h_n(t)$ is nonnegative, and that $\sup_n h_n(0) = \sup_n \frac{1}{2} \|\lambda_n(0)(x_0 - \hat{x}) + v_0\|^2 < +\infty$ we deduce

$$\sup_n \int_0^\infty \|v_n(t)\|^2 dt < +\infty.$$

This immediately implies that $\|v(\cdot)\|^2 \in L^1(0, +\infty)$. Since $t \mapsto \|v(t)\|$ is decreasing, we conclude $\lim_{t \rightarrow +\infty} v(t) = 0$. $\square$

6 A continuous Newton-like Dynamics for Convex-concave Saddle Value Problems

The theory of maximal monotone operators allows one to treat in a unified way convex optimization, convex-concave saddle value problems, and fixed point theory of nonexpansive mappings. Let us recall that, given a general closed convex-concave bivariate function L , the operator $A = (\partial_x L, -\partial_y L)$ is maximal monotone; see [15, 16]. Let us take advantage of this flexible framework, and illustrate the results of the preceding sections by considering the Lagrangian (convex-concave) approach to convex linearly constrained minimization problems:

Given

- X, Y, Z Hilbert spaces,
- $f : X \rightarrow \mathbb{R}$ and $g : Y \rightarrow \mathbb{R}$ convex, $\mathcal{C}^2$ functions,
- $A : X \rightarrow Z$ and $B : Y \rightarrow Z$ continuous linear operators,

consider the following (primal) minimization problem (see [17] for various applications):

$$\min \{f(x) + g(y) : Ax - By = 0\}. \quad (53)$$

This is equivalent to the convex-concave saddle value problem

$$\min_{x,y} \max_z \{f(x) + g(y) + \langle z, Ax - By \rangle\}, \quad (54)$$

which is associated to the Lagrangian function $L : X \times Y \times Z \rightarrow \mathbb{R}$

$$L(x, y, z) = f(x) + g(y) + \langle z, Ax - By \rangle. \quad (55)$$

L is convex with respect to (x, y) and linear (hence concave) with respect to z . The maximal monotone operator $\mathcal{A} : X \times Y \times Z \rightarrow X \times Y \times Z$ which is associated to L is given by

$$\mathcal{A}(x, y, z) = (\nabla f(x) + {}^t A z, \nabla g(y) - {}^t B z, B y - A x) \quad (56)$$

where ${}^t A$ and ${}^t B$ are respectively the transpose operators of A and B . Solutions of (54) are the zeros of $\mathcal{A}$, and we suppose that this solution set $S = \mathcal{A}^{-1}(0)$ is non empty.

Take $H = X \times Y \times Z$. The open-loop regularized Newton dynamics (11a), (11b) can be written as

$$\begin{cases} \lambda(t)\dot{x} + \nabla^2 f(x)\dot{x} + {}^t A\dot{z} + \nabla f(x) + {}^t A z = 0 \\ \lambda(t)\dot{y} + \nabla^2 g(y)\dot{y} - {}^t B\dot{z} + \nabla g(y) - {}^t B z = 0 \\ \lambda(t)\dot{z} + B\dot{y} - A\dot{x} + B y - A x = 0. \end{cases} \quad (57)$$

Theorem 3.7 of [1] tells us that, if $\lambda(\cdot)$ is bounded from above on $[0, +\infty[$ and

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

then $(x(t), y(t), z(t))$ converges weakly to a zero of $\mathcal{A}$, as t goes to $+\infty$. Note that we obtain weak convergence of the primal and dual trajectories respectively to the primal and dual solutions. This makes contrast with the semi-group approach (semi-group of contractions generated by $\mathcal{A}$) which only provides weak ergodic convergent trajectories.

Set $v = (\nabla f(x) + {}^t A z, \nabla g(y) - {}^t B z, B y - A x)$, and consider the closed-loop dynamics

$$\begin{cases} \alpha(\|v\|^2)\dot{x} + \nabla^2 f(x)\dot{x} + {}^t A\dot{z} + \nabla f(x) + {}^t A z = 0 \\ \alpha(\|v\|^2)\dot{y} + \nabla^2 g(y)\dot{y} - {}^t B\dot{z} + \nabla g(y) - {}^t B z = 0 \\ \alpha(\|v\|^2)\dot{z} + B\dot{y} - A\dot{x} + B y - A x = 0. \end{cases} \quad (58)$$

Take $\alpha(r) = r^\theta$. By Theorem 5.1, for any $\theta \geq 0$, we have $\lim_{t \rightarrow +\infty} \|v(t)\| = 0$ (indeed this holds true for any real θ). By Corollary 5.1, the asymptotic convergence of $(x(t), y(t), z(t))$ to a zero of $\mathcal{A}$ (which is assumed to be nonempty) is obtained by taking $\theta < \frac{1}{2}$. Equivalently

$$\lambda(t) = (\|\nabla f(x) + {}^t A z\|^2 + \|\nabla g(y) - {}^t B z\|^2 + \|B y - A x\|^2)^\theta. \quad (59)$$

7 Conclusion

The analysis of the continuous dynamics may serve as a guideline and an incentive to study some related questions for discrete dynamics.

a) Our approach can be traced back to the Levenberg-Marquardt regularization of the Newton method; see [18, 19]. For equations governed by monotone operators, it gives an original and rather simple hint to the delicate question: at each step $k = 1, 2, \dots$ how to choose the regularizing parameter λ_k ? In accordance with the continuous case, it is a natural, and open question, in the discrete case, to try to take, in an adaptive way, at step k , $\lambda_k = a \|\nabla f(x_{k-1})\|^\theta$, with θ strictly less than 1. Among the rich literature devoted to this question for general non-linear problems, one can consult [20–22], [23, Chapter 10].

b) Note that systems (57) and (58) naturally suggest corresponding discrete dynamics and algorithms. The result would be new splitting algorithms involving in a decomposable way regularized Newton steps with respect to the functions $x \mapsto f(x)$ and $y \mapsto g(y)$, with, at each step, updating of the dual variable z . Clearly, this is a topic for further research.

Acknowledgements H. A. and P. R. are partially supported by ANR-08-BLAN-0294-03; B. F. S. is partially supported by CNPq grants 302962/2011-5, 474944/2010-7, FAPERJ grant E-26/102.940/2011 and PRONEX-Optimization.

References

1. Attouch, H., Svaiter, B.F., A continuous dynamical Newton-like approach to solving monotone inclusions, *SIAM J. Control Optim.*, **49**, 574-598 (2011).
2. Alvarez, A., Perez, J.M., A dynamical system associated with Newton's method for parametric approximations of convex minimization problems, *Appl. Math. and Optim.*, **38**, 193-217 (1998).
3. Brezis, H., *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, Elsevier, New York (1973).
4. Hoheisel, T., Kanzow, C., Mordukhovich, B.S., Phan, H., Generalized Newton's method based on graphical derivatives, *Nonlinear Analysis: Theory, Methods & Applications*, **75**, 1324-1340 (2012).
5. Clarke, F., Discontinuous feedback and nonlinear systems, In: *Proc. IFAC Conf. Nonlinear Control (NOLCOS)*, Bologna, pp. 1-29 (2010).
6. Alvarez, F., Attouch, H., Bolte, J., Redont, P., A second-order gradient-like dissipative dynamical system with Hessian-driven damping. Application to optimization and mechanics, *J. Math. Pures Appl.*, **81**, 747-779 (2002).
7. Attouch, H., Maingé, P.E., Redont, P., A second-order differential system with Hessian-driven damping; Application to non-elastic shock laws, *Differential Equations and Applications*, **4**, 27-65 (2012).
8. Attouch, H., Redont, P., The second-order in time continuous Newton method, In: *Approximation, optimization and mathematical economics*, Physica, Heidelberg, 25-36 (2001).
9. Maingé, P.E., First order continuous Newton-like systems for monotone inclusions, preprint, Université des Antilles-Guyane, Martinique (2011).
10. Attouch, H., Buttazzo, G., Michaille, G., *Variational analysis in Sobolev and BV spaces*, SIAM series on optimization, Philadelphia (2006).
11. Ziemer, W.P., *Weakly Differentiable Functions*, Springer, New York (1989).
12. Evans, L.C., Gariepy, R.F., *Measure Theory and Fine Properties of Functions*, CRC Press, Boca Raton (1992).
13. Moreau, J.J., Bounded variation in time, In: Moreau, J.J., Panagiotopoulos, P.D., Strang, G. (eds.): *Topics in nonsmooth mechanics*, pp. 1-76. Birkhauser Verlag, Basel-Boston-Berlin (1988).
14. Natanson, I.P., *Theory of functions of a real variable*, Frederick Ungar, New York (1955).
15. Rockafellar, R.T., Monotone operators associated with saddle-functions and mini-max problems, In: Browder, F. (ed.): *Nonlinear Functional Analysis, Proceedings of Symposia in Pure Mathematics*, vol. 18, pp. 241-250. American Mathematical Society (1976).
16. Attouch, H., Soueycatt, M., Augmented Lagrangian and proximal alternating direction methods of multipliers in Hilbert spaces. Applications to games, PDE's and control, *Pacific Journal of Optimization*, **5**, 17-37 (2009).
17. Attouch, H., Briceno, L., Combettes, P., A parallel splitting method for coupled monotone inclusions, *SIAM J. Control Optim.*, **48**, 3246-3270 (2010).
18. Levenberg, K., A method for the solution of certain non-linear problems in the least squares, *Quarterly of Applied Mathematics*, **2**, 164-168 (1944).
19. Marquardt, D.W., An algorithm for least squares estimation of non-linear parameters, *SIAM J. Appl. Math.*, **11**, 431-441 (1963).
20. Dennis, J.E., Schnabel, R.B., *Numerical Methods for Unconstrained Minimization*, Prentice-Hall, Englewood Cliffs (1983). Reprinted by SIAM publications (1993).
21. Kanzow, C., Yamashita, N., Fukushima, M., Levenberg-Marquardt methods for constrained nonlinear equations with strong local convergence properties, *Journal of Computational and Applied Mathematics*, **172**, 375-397 (2004).

-
22. Madsen, K., Nielsen, H.B., Tingleff, O., Methods for non-linear least squares problems, IMM Technical University of Denmark (2004), http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/3215/pdf/imm3215.pdf.
 23. Nocedal, J., Wright, S., Numerical Optimization, Springer Series in Operations Research, Springer-Verlag, New-York (1999).