

## THE STEEPEST DESCENT DYNAMICAL SYSTEM WITH CONTROL. APPLICATIONS TO CONSTRAINED MINIMIZATION

ALEXANDRE CABOT<sup>1</sup>

**Abstract.** Let  $H$  be a real Hilbert space,  $\Phi_1 : H \rightarrow \mathbb{R}$  a convex function of class  $C^1$  that we wish to minimize under the convex constraint  $S$ . A classical approach consists in following the trajectories of the generalized steepest descent system (*cf.* Brézis [5]) applied to the non-smooth function  $\Phi_1 + \delta_S$ . Following Antipin [1], it is also possible to use a continuous gradient-projection system. We propose here an alternative method as follows: given a smooth convex function  $\Phi_0 : H \rightarrow \mathbb{R}$  whose critical points coincide with  $S$  and a control parameter  $\varepsilon : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  tending to zero, we consider the “Steepest Descent and Control” system

$$(SDC) \quad \dot{x}(t) + \nabla\Phi_0(x(t)) + \varepsilon(t) \nabla\Phi_1(x(t)) = 0,$$

where the control  $\varepsilon$  satisfies  $\int_0^{+\infty} \varepsilon(t) dt = +\infty$ . This last condition ensures that  $\varepsilon$  “slowly” tends to 0. When  $H$  is finite dimensional, we then prove that  $d(x(t), \operatorname{argmin}_S \Phi_1) \rightarrow 0$  ( $t \rightarrow +\infty$ ), and we give sufficient conditions under which  $x(t) \rightarrow \bar{x} \in \operatorname{argmin}_S \Phi_1$ . We end the paper by numerical experiments allowing to compare the (SDC) system with the other systems already mentioned.

**Mathematics Subject Classification.** 34A12, 34D05, 34G20, 34H05, 37N40.

Received February 17, 2003.

### INTRODUCTION

Let  $H$  be a real Hilbert space and  $\Phi_1 : H \rightarrow \mathbb{R}$  a smooth function that we wish to minimize on the constraint set  $S$ . More precisely, we are interested in the following problem

$$(\mathcal{P}) \quad \min \{ \Phi_1(x), \quad x \in S \},$$

and we will denote the solution set by  $\operatorname{argmin}_S \Phi_1$ . To deal with problem  $(\mathcal{P})$ , a classical approach consists in following the trajectories of a gradient-like dynamical system, hopefully converging toward some solution of  $(\mathcal{P})$ . For example, when the data  $\Phi_1$  and  $S$  are convex, one can consider the generalized steepest descent system applied to the non-smooth function  $\Phi_1 + \delta_S$ , namely

$$(S1) \quad \dot{x}(t) + \nabla\Phi_1(x(t)) \in -N_S(x(t)), \quad t \geq 0,$$

---

*Keywords and phrases.* Dissipative dynamical system, steepest descent method, constrained optimization, convex minimization, asymptotic behaviour, non-linear oscillator.

<sup>1</sup> Laboratoire LACO, Faculté des Sciences, 123 avenue Albert Thomas, 87060 Limoges Cedex, France;  
e-mail: alexandre.cabot@unilim.fr

where  $\delta_S$  is the indicator function of  $S$  and  $N_S(x(t))$  is the normal cone to  $S$  at the point  $x(t)$ . This differential inclusion has been widely studied (Brezis [5], Bruck [6], ...) and the trajectories of (S1) are proved to converge to some element  $\bar{x} \in \operatorname{argmin}_S \Phi_1$ . The condition  $\bar{x} \in \operatorname{argmin}_S \Phi_1$  can be rewritten as  $\nabla \Phi_1(\bar{x}) \in -N_S(\bar{x})$ , which in turn is equivalent to

$$\bar{x} = P_S(\bar{x} - \mu \nabla \Phi_1(\bar{x})), \quad \text{for any } \mu > 0.$$

This remark leads us to consider the following dynamical system

$$(S2) \quad \dot{x}(t) + x(t) - P_S(x(t) - \mu \nabla \Phi_1(x(t))) = 0,$$

which has been initiated by Antipin [1]. The trajectories of (S2) converge toward a minimum of  $\Phi_1$  on  $S$  and enjoy moreover a viability property: the orbits starting in  $S$  are lying in  $S$ .

We propose in this paper an alternative method to solve (P) by studying the following ‘‘Steepest Descent and Control’’ system

$$(SDC) \quad \dot{x}(t) + \nabla \Phi_0(x(t)) + \varepsilon(t) \nabla \Phi_1(x(t)) = 0, \quad t \geq 0,$$

where  $\Phi_0 : H \rightarrow \mathbb{R}$  is a  $C^1$  function whose critical points coincide with  $S$  and  $\varepsilon : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is a control parameter tending to 0 when  $t \rightarrow +\infty$ . The (SDC) system can be interpreted as a steepest descent method (applied to the function  $\Phi_0$ ) plus a control term  $\varepsilon(t) \nabla \Phi_1(x(t))$ .

Some variants of the (SDC) system have been studied by several authors. For example in [3], Attouch and Cominetti couple approximation methods with the steepest descent one. To consider only a particular case of their paper, they prove that, when  $\Phi_0$  is convex and  $\varepsilon : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is a  $C^1$  control function such that  $\int_0^{+\infty} \varepsilon(t) dt = +\infty$ , then each trajectory of the system

$$\dot{x}(t) + \nabla \Phi_0(x(t)) + \varepsilon(t)x(t) = 0 \quad ((SDC) \text{ with } \Phi_1(x) = |x|^2/2)$$

strongly converges to the point of minimal norm of  $\operatorname{argmin} \Phi_0$ . The condition  $\int_0^{+\infty} \varepsilon(t) dt = +\infty$  expresses that  $\varepsilon(t)$  tends to zero slowly enough to allow the Tikhonov regularization term  $\varepsilon(t)x(t)$  to be effective asymptotically. Attouch and Czarnecki study in [4] a second-order in time version of the previous system and obtain the same type of result.

In another direction, Cabot and Czarnecki consider in [7] a particular case of the (SDC) system where  $H = \mathbb{R}^2$ ,  $\Phi_0(x, y) = \phi(x) + \phi(y)$  and  $\Phi_1(x, y) = V(x - y)$ , thus leading to

$$\begin{cases} \dot{x}(t) + \nabla \phi(x(t)) + \varepsilon(t) \nabla V(x - y)(t) = 0 \\ \dot{y}(t) + \nabla \phi(y(t)) - \varepsilon(t) \nabla V(x - y)(t) = 0. \end{cases}$$

They focus their attention on the case where the potential  $V$  modelizes a repulsion force. Such a potential allows a better exploration of the minima of  $\phi$  than a (simple) steepest descent method. Cabot and Czarnecki show in this paper that the condition  $\int_0^{+\infty} \varepsilon(t) dt = +\infty$  plays once more an essential role in order to make the repulsion efficient.

Coming back to the general (SDC) system and assuming that  $\varepsilon$  is a slow control, *i.e.*  $\int_0^{+\infty} \varepsilon(t) dt = +\infty$ , we prove in this paper that, under convex-type conditions on  $\Phi_0$  and  $\Phi_1$ , the distance of the (SDC) trajectory  $x(\cdot)$  to the set  $\operatorname{argmin}_S \Phi_1$  tends to 0 when  $t \rightarrow +\infty$ . Moreover, we obtain sufficient conditions on  $\Phi_0$  and  $\Phi_1$  ensuring the convergence of  $x(\cdot)$  toward some solution of (P).

The paper is organized as follows. In Section 1, we give global existence results relative to (SDC) and we show the interest of a slow control  $\varepsilon$ . Section 2 deals with the minimization properties of (SDC) trajectories in the case of a slow control and we prove that they tend to minimize  $\Phi_1$  over the set  $S = \operatorname{argmin} \Phi_0$ . In Section 3, we illustrate our theoretical results by numerical experiments and we compare the (SDC) system respectively with (S1) and (S2). The technical proofs of the paper are postponed in Section 4.

### 1. GENERAL RESULTS

#### 1.1. Global existence

Let  $H$  be an Hilbert space and  $\Phi_0 : H \rightarrow \mathbb{R}, \Phi_1 : H \rightarrow \mathbb{R}, \varepsilon : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  functions of class  $\mathcal{C}^1$ . In the whole paper, we assume the following (rather standard) set of hypotheses ( $\mathcal{H}$ ):

$$\begin{aligned}
 (\mathcal{H}_{\Phi_0}) & \begin{cases} i - \text{the map } \Phi_0 \text{ is bounded from below on } H; \\ ii - \text{the map } \nabla\Phi_0 \text{ is Lipschitz continuous on the bounded subsets of } H. \end{cases} \\
 (\mathcal{H}_{\Phi_1}) & \begin{cases} i - \text{the map } \Phi_1 \text{ is bounded from below on } H; \\ ii - \text{the map } \nabla\Phi_1 \text{ is Lipschitz continuous on the bounded subsets of } H. \end{cases} \\
 (\mathcal{H}_\varepsilon) & \begin{cases} i - \text{the map } \varepsilon \text{ is non-increasing, i.e. } \dot{\varepsilon}(t) \leq 0 \quad \forall t \in \mathbb{R}_+; \\ ii - \text{the map } \varepsilon \text{ is Lipschitz continuous on } \mathbb{R}_+; \\ iii - \lim_{t \rightarrow +\infty} \varepsilon(t) = 0. \end{cases}
 \end{aligned}$$

Let us then consider the following dynamical system

$$(SDC) \quad \begin{cases} \dot{x}(t) + \nabla\Phi_0(x(t)) + \varepsilon(t) \nabla\Phi_1(x(t)) = 0 \\ x(0) = x_0 \end{cases}$$

which will be referred to as the ‘‘Steepest Descent and Control’’ system. We recognize the steepest descent system applied to  $\Phi_0$  plus the ‘‘control term’’  $\varepsilon(t) \nabla\Phi_1(x(t))$ . We can define along every trajectory of (SDC) the function  $E$  by

$$E(t) = \Phi_0(x(t)) - \inf \Phi_0 + \varepsilon(t) (\Phi_1(x(t)) - \inf \Phi_1).$$

By differentiating  $E$ , we obtain

$$\begin{aligned}
 \dot{E}(t) &= \langle \dot{x}(t), \nabla\Phi_0(x(t)) + \varepsilon(t) \nabla\Phi_1(x(t)) \rangle + \dot{\varepsilon}(t) (\Phi_1(x(t)) - \inf \Phi_1) \\
 &= -|\dot{x}(t)|^2 + \dot{\varepsilon}(t) (\Phi_1(x(t)) - \inf \Phi_1).
 \end{aligned}
 \tag{1.1}$$

From assumption  $(\mathcal{H}_\varepsilon - i)$ , we have  $\dot{\varepsilon}(t) \leq 0$  and hence  $\dot{E}(t) \leq 0$ : the function  $E$  is non-increasing along every trajectory, i.e. defines a Lyapounov function for the (SDC) system. The existence of a Lyapounov function for a differential equation is useful in the study of the asymptotic stability of its equilibria. Lyapounov methods and other power tools (like the Lasalle invariance principle) have been developed to study such a question. We refer to the abundant literature on this subject (Arnold [2], Haraux [8], Hirsch-Smale [9], Lasalle-Lefschetz [10], Reinhardt [12], ...). The central result of this section is given by the following proposition, whose proof mainly relies on the existence of the Lyapounov function  $E$ .

**Proposition 1.1.** *Let us assume that  $\Phi_0, \Phi_1$  and  $\varepsilon$  satisfy the assumptions  $(\mathcal{H})$ . Then, the following properties hold*

- (i) *for all  $x_0 \in H$ , there exists a unique maximal solution  $x : \mathbb{R}_+ \rightarrow H$  of (SDC), which is of class  $\mathcal{C}^1$ , and which satisfies the initial condition  $x(0) = x_0$ . Moreover, the following estimate holds:  $\dot{x} \in L^2([0, +\infty); H)$ ;*
- (ii) *assuming additionally that  $x$  is bounded (which is the case for example if  $\Phi_0$  is coercive, i.e.  $\lim_{|x| \rightarrow +\infty} \Phi_0(x) = +\infty$ ), then we have  $\lim_{t \rightarrow +\infty} \dot{x}(t) = 0$  and  $\lim_{t \rightarrow +\infty} \nabla\Phi_0(x(t)) = 0$ .*

*Proof.* (i) For  $x_0 \in H$ , the Cauchy-Lipschitz theorem and hypotheses  $(\mathcal{H}_{\Phi_0} - ii)$ ,  $(\mathcal{H}_{\Phi_1} - ii)$  ensure the existence of a unique local trajectory  $x(\cdot)$  solution of (SDC). Let  $x$  denote the maximal solution defined on the interval  $[0, T_{\max})$  with  $0 < T_{\max} \leq +\infty$ . In order to prove that  $T_{\max} = +\infty$ , let us show that  $\dot{x} \in L^2([0, T_{\max}); H)$ .

By integrating formula (1.1) on  $[0, t]$ , we deduce that  $\int_0^t |\dot{x}(s)|^2 ds \leq E(0) - E(t)$ , which combined with  $E(t) \geq 0$  implies

$$\int_0^{T_{\max}} |\dot{x}(s)|^2 ds \leq E(0) < +\infty.$$

By a standard argument, we derive that  $T_{\max} = +\infty$ . Indeed, let us assume that  $T_{\max} < +\infty$ . From the Cauchy-Schwarz inequality, we have

$$|x(t) - x(t')| = \left| \int_{t'}^t \dot{x}(s) ds \right| \leq \int_{t'}^t |\dot{x}(s)| ds \leq \left( \int_0^{T_{\max}} |\dot{x}(s)|^2 ds \right)^{1/2} \sqrt{t - t'},$$

and since  $T_{\max} < +\infty$ ,  $\lim_{t \rightarrow T_{\max}} x(t) := x_\infty$  exists. But, applying again the local existence theorem with initial data  $x_\infty$ , we can extend the maximal solution to a strictly larger interval, a contradiction. Hence,  $T_{\max} = +\infty$ , which completes the proof of (i).

(ii) We now assume that  $x(\cdot)$  is bounded. Equation (SDC) and assumptions  $(\mathcal{H}_{\Phi_0} - ii)$ ,  $(\mathcal{H}_{\Phi_1} - ii)$  clearly imply that  $\dot{x}$  is bounded, *i.e.* the map  $x$  is Lipschitz continuous on  $\mathbb{R}_+$ . Using again assumptions  $(\mathcal{H}_{\Phi_0} - ii)$  and  $(\mathcal{H}_{\Phi_1} - ii)$ , we deduce that

$$t \mapsto \nabla \Phi_0(x(t)) \quad \text{and} \quad t \mapsto \nabla \Phi_1(x(t)) \quad \text{are Lipschitz continuous and bounded on } \mathbb{R}_+.$$

From assumptions  $(\mathcal{H}_\varepsilon)$  the map  $\varepsilon$  is Lipschitz continuous and bounded. We then conclude that  $t \mapsto \dot{x}(t) = -\nabla \Phi_0(x(t)) - \varepsilon(t) \nabla \Phi_1(x(t))$  is Lipschitz continuous on  $\mathbb{R}_+$ . As a consequence the function  $h := \dot{x}$  satisfies both

$$h \in L^2([0, +\infty); H) \quad \text{and} \quad h \in \text{Lip}([0, +\infty); H).$$

According to a classical result, these two properties imply  $\lim_{t \rightarrow +\infty} h(t) = 0$ , *i.e.*  $\lim_{t \rightarrow +\infty} \dot{x}(t) = 0$ . Then, in view of  $\lim_{t \rightarrow +\infty} \varepsilon(t) = 0$ , equation (SDC) immediately gives  $\lim_{t \rightarrow +\infty} \nabla \Phi_0(x(t)) = 0$ , which ends the proof of (ii).  $\square$

## 1.2. Interest of a slow control $\varepsilon$

When  $\varepsilon \equiv 0$ , the (SDC) system reduces to the steepest descent dynamical system applied to  $\Phi_0$  and it is well-known that its trajectories weakly converge to a minimum of  $\Phi_0$  (*cf.* Bruck [6]). This last result can be generalized when  $\varepsilon$  tends to zero fast enough; indeed, we have

**Proposition 1.2.** *In addition to the hypotheses  $(\mathcal{H})$ , let us assume that the map  $\Phi_0$  is convex with  $\text{argmin} \Phi_0 \neq \emptyset$  and that  $\int_0^{+\infty} \varepsilon(t) dt < +\infty$  (fast control). Let  $x(\cdot)$  be the unique solution of (SDC). If  $x(\cdot)$  is bounded, then there exists  $x_\infty \in \text{argmin} \Phi_0$  such that*

$$\lim_{t \rightarrow +\infty} x(t) = x_\infty \quad w - H.$$

*Proof.* The central idea is to prove the weak convergence of the trajectory  $x$  by using the Opial lemma [11].

**Lemma 1.3** (Opial). *Let  $H$  be a Hilbert space and  $x : [0, +\infty[ \rightarrow H$  be a function such that there exists a non void set  $S \subset H$  which verifies:*

(i)  $\forall z \in S, \lim_{t \rightarrow +\infty} |x(t) - z|$  exists;

(ii)  $\forall t_n \rightarrow +\infty$  with  $x(t_n) \rightarrow x_\infty$  weakly in  $H$ , we have  $x_\infty \in S$ .

*Then,  $x(t)$  weakly converges as  $t \rightarrow +\infty$  to some element  $x_\infty$  of  $S$ .*

Let us apply the Opial lemma with  $S = \operatorname{argmin} \Phi_0$ .

(i) For any  $z \in \operatorname{argmin} \Phi_0 \neq \emptyset$ , let us set  $h_z(t) = \frac{1}{2}|x(t) - z|^2$ . Since  $\Phi_0$  is convex, we have

$$\begin{aligned} \dot{h}_z(t) &= \langle x(t) - z, \dot{x}(t) \rangle \\ &= -\langle x(t) - z, \nabla \Phi_0(x(t)) \rangle - \varepsilon(t) \langle x(t) - z, \nabla \Phi_1(x(t)) \rangle \\ &\leq -\varepsilon(t) \langle x(t) - z, \nabla \Phi_1(x(t)) \rangle. \end{aligned}$$

On the other hand, the boundedness of the trajectory  $x(\cdot)$  implies the existence of  $C > 0$  such that, for every  $t \geq 0$ ,  $|\langle x(t) - z, \nabla \Phi_1(x(t)) \rangle| \leq C$ . As a consequence, the following inequality holds

$$\dot{h}_z(t) \leq C \varepsilon(t).$$

Taking the positive part of the previous relation, we finally obtain  $(\dot{h}_z)_+(t) \leq C \varepsilon(t)$ . By applying the following lemma, we conclude that  $\lim_{t \rightarrow +\infty} |x(t) - z|$  exists.

**Lemma 1.4.** *Let  $h : \mathbb{R}_+ \rightarrow \mathbb{R}$  a function of class  $\mathcal{C}^1$  which is bounded from below and such that  $\dot{h}_+ \in L^1(0, +\infty)$ . Then,  $\lim_{t \rightarrow +\infty} h(t)$  exists.*

(ii) Let us now consider a sequence  $(t_n)$  such that  $\lim_{n \rightarrow +\infty} x(t_n) = x_\infty \quad w - H$ .

We have to prove that  $x_\infty \in \operatorname{argmin} \Phi_0$ . From the convexity of  $\Phi_0$ , we have, for every  $\xi \in H$

$$\Phi_0(\xi) \geq \Phi_0(x(t_n)) + \langle \nabla \Phi_0(x(t_n)), \xi - x(t_n) \rangle.$$

By using the weak lower semicontinuity of the convex continuous function  $\Phi_0$ , and noticing that, in the duality bracket  $\langle \nabla \Phi_0(x(t_n)), \xi - x(t_n) \rangle$ , the two terms are respectively norm converging to zero and weakly convergent, we can pass to the lower limit to obtain  $\Phi_0(\xi) \geq \Phi_0(x_\infty)$ . This being true for every  $\xi \in H$ , we deduce that  $x_\infty \in \operatorname{argmin} \Phi_0$ . The Opial lemma allows us to conclude.  $\square$

When  $\varepsilon$  is a fast control, the previous result shows the convergence of the trajectories, but the limit does not depend explicitly on  $\Phi_1$ . We now show how the assumption of slow control (*i.e.*  $\int_0^{+\infty} \varepsilon(t) dt = +\infty$ ) allows to rescale conveniently the (SDC) system, then giving rise to the minimization of  $\Phi_1$  over the set  $\operatorname{argmin} \Phi_0$ . Indeed, let us divide the (SDC) equation by  $\varepsilon(t)$  so as to obtain

$$\dot{x}(t)/\varepsilon(t) + \nabla \Phi_0(x(t))/\varepsilon(t) + \nabla \Phi_1(x(t)) = 0. \tag{1.2}$$

Introducing a time-rescaling  $t = \tau(s)$ , the new function  $y(s) := x(\tau(s))$  satisfies  $\dot{y}(s) = \dot{x}(\tau(s)) \dot{\tau}(s)$ , so that we are naturally led to choose  $\tau(s)$  such that

$$\dot{\tau}(s) = 1/\varepsilon(\tau(s)) \tag{1.3}$$

in such a way that (1.2) is transformed into

$$\dot{y}(s) + \nabla \Phi_0(y(s))/\varepsilon(\tau(s)) + \nabla \Phi_1(y(s)) = 0. \tag{1.4}$$

Setting  $\mathcal{E}(t) = \int_0^t \varepsilon(s) ds$ , equality (1.3) is equivalent to

$$\dot{\mathcal{E}}(\tau(s)) \dot{\tau}(s) = 1$$

so that we obtain  $\mathcal{E}(\tau(s)) = s$ . Clearly  $\mathcal{E}$  is a strictly increasing function from  $[0, +\infty[$  onto  $[0, +\infty[$  since  $\int_0^{+\infty} \varepsilon(s) ds = +\infty$ . We then have  $\tau(s) = \mathcal{E}^{-1}(s)$  with  $\lim_{s \rightarrow +\infty} \tau(s) = +\infty$ , and then

$$\lim_{t \rightarrow +\infty} x(t) \quad \text{exists} \quad \iff \quad \lim_{s \rightarrow +\infty} y(s) \quad \text{exists}$$

in which case both limits coincide. On the other hand, equation (1.4) can be rewritten under the equivalent form

$$\dot{y}(s) + \nabla \Psi(s, y(s)) = 0,$$

where  $\Psi(s, y) = (\Phi_0(y) - \inf \Phi_0)/\varepsilon(\tau(s)) + \Phi_1(y)$ . Assuming for simplicity that  $\Phi_0$  is convex and satisfies  $S := \operatorname{argmin} \Phi_0 \neq \emptyset$ , then we have

$$\lim_{s \rightarrow +\infty} \Psi(s, y) = \delta_S(y) + \Phi_1(y),$$

where  $\delta_S$  is the indicator function of  $S$ . It is then quite natural to expect the (*SDC*) trajectories to minimize the function  $\Phi_1$  on the set  $S = \operatorname{argmin} \Phi_0$ , at least under convex-type assumptions on  $\Phi_0$  and  $\Phi_1$ . The purpose of the next paragraphs is precisely to exhibit situations in which convergence holds. As we shall see later, the proofs are much more difficult than in the case of a fast control.

## 2. MINIMIZATION OF $\Phi_1$ OVER $\operatorname{argmin} \Phi_0$ WHEN $\varepsilon$ IS A SLOW CONTROL

### 2.1. Convergence of the distance function $t \mapsto d(x(t), \operatorname{argmin}_S \Phi_1)$

We state our main theorem in the general framework of quasi-convex functions, *i.e.* functions whose sublevel sets are convex. Let us recall the following two properties about quasi-convex functions  $f : H \rightarrow \mathbb{R}$  which are differentiable:

- $f$  is quasi-convex if, and only if

$$\forall x, y \in H, \quad f(y) \leq f(x) \implies \langle \nabla f(x), y - x \rangle \leq 0;$$

- $f$  is strictly quasi-convex if, and only if it is quasi-convex and

$$\forall x, y \in H, \quad f(y) < f(x) \implies \langle \nabla f(x), y - x \rangle < 0.$$

**Theorem 2.1.** *In addition to the hypotheses ( $\mathcal{H}$ ), let us assume that  $H = \mathbb{R}^n$  and*

- $\Phi_0 : H \rightarrow \mathbb{R}$  is quasi-convex and  $S := \{x \in H, \nabla \Phi_0(x) = 0\} \neq \emptyset$ .
- $\Phi_1 : H \rightarrow \mathbb{R}$  is strictly quasi-convex and  $\emptyset \neq \operatorname{argmin}_S \Phi_1 \subset \operatorname{argmin} \Phi_0$ .
- $\int_0^{+\infty} \varepsilon(t) dt = +\infty$  (**slow control**).

*If the trajectory  $x(\cdot)$  of the (*SDC*) system is bounded, then*

- (i)  $\lim_{t \rightarrow +\infty} d(x(t), \operatorname{argmin}_S \Phi_1) = 0$ ,
- (ii)  $\lim_{t \rightarrow +\infty} \Phi_0(x(t)) = \min \Phi_0$ ,
- (iii)  $\lim_{t \rightarrow +\infty} \Phi_1(x(t)) = \min_S \Phi_1$ ,

*where  $d(\cdot, \operatorname{argmin}_S \Phi_1)$  stands for the distance function to the set  $\operatorname{argmin}_S \Phi_1$ . As a consequence, if  $\operatorname{argmin}_S \Phi_1$  is reduced to a singleton  $\{p\}$ , then the trajectory  $x(\cdot)$  converges to  $p$ .*

*Proof.* The proof is an extension of Attouch-Czarnecki [4] where the authors show the convergence toward the point of minimal norm of  $\operatorname{argmin} \Phi_0$  in the case where  $\Phi_1(x) = |x|^2/2$ . In the present situation, the proof mainly relies on the study of the function  $h$  defined by

$$h(t) = \frac{1}{2} d(x(t), C)^2,$$

where  $C := \operatorname{argmin}_S \Phi_1 \subset \operatorname{argmin} \Phi_0$ . To improve the clarity of the exposition, the proof is postponed in Section 4.  $\square$

**Corollary 2.2.** *In addition to the hypotheses  $(\mathcal{H})$ , let us assume that  $H = \mathbb{R}^n$  and*

- $\Phi_0 : H \rightarrow \mathbb{R}$  is convex and  $S = \operatorname{argmin} \Phi_0 \neq \emptyset$ .
- $\Phi_1 : H \rightarrow \mathbb{R}$  is convex and  $\operatorname{argmin}_S \Phi_1 \neq \emptyset$ .
- $\int_0^{+\infty} \varepsilon(t) dt = +\infty$  (**slow control**).

*If the trajectory  $x(\cdot)$  of the (SDC) system is bounded, then  $\lim_{t \rightarrow +\infty} d(x(t), \operatorname{argmin}_S \Phi_1) = 0$ . As a consequence, if  $\operatorname{argmin}_S \Phi_1$  is reduced to a singleton  $\{p\}$ , then the trajectory  $x(\cdot)$  converges to  $p$ .*

*Proof.* Immediate from Theorem 2.1 and the fact that any convex function is strictly quasi-convex and hence quasi-convex. □

We now give a sufficient condition which ensures that the (SDC) trajectory is bounded.

**Proposition 2.3.** *Under the assumptions of Theorem 2.1, let  $x(\cdot)$  be the unique solution of the (SDC) system. If the following condition holds*

$$(C) \quad \text{for every } M > 0, \quad \text{the set } \{x \in H, \quad \Phi_0(x) \leq M\} \cap \{x \in H, \quad \Phi_1(x) \leq \min_S \Phi_1\} \quad \text{is bounded,}$$

*then the trajectory  $x(\cdot)$  is bounded and hence, in view of Theorem 2.1,  $\lim_{t \rightarrow +\infty} d(x(t), \operatorname{argmin}_S \Phi_1) = 0$ .*

*Moreover, condition (C) is satisfied in each of the following cases*

- (a)  $\Phi_0$  is coercive, i.e.  $\lim_{|x| \rightarrow +\infty} \Phi_0(x) = +\infty$ ;
- (b)  $\Phi_1$  is coercive, i.e.  $\lim_{|x| \rightarrow +\infty} \Phi_1(x) = +\infty$ .

*Proof.* It is postponed in Section 4. □

### 2.2. Convergence of the trajectory when $\operatorname{argmin} \Phi_0 \cap \operatorname{argmin} \Phi_1 \neq \emptyset$

When  $\operatorname{argmin} \Phi_0 \cap \operatorname{argmin} \Phi_1$  is non empty, the following convergence result holds

**Proposition 2.4.** *In addition to the hypotheses  $(\mathcal{H})$ , let us assume that  $H = \mathbb{R}^n$  and*

- $\Phi_0 : H \rightarrow \mathbb{R}$  and  $\Phi_1 : H \rightarrow \mathbb{R}$  are convex.
- The set  $\operatorname{argmin} \Phi_0 \cap \operatorname{argmin} \Phi_1$  is non empty.
- $\int_0^{+\infty} \varepsilon(t) dt = +\infty$  (**slow control**).

*Then, the trajectory  $x(\cdot)$  of the (SDC) system converges to some  $\bar{x} \in \operatorname{argmin} \Phi_0 \cap \operatorname{argmin} \Phi_1$ .*

*Proof.* Given any  $z \in \operatorname{argmin} \Phi_0 \cap \operatorname{argmin} \Phi_1$ , let us define the function  $h_z$  by

$$h_z(t) := \frac{1}{2} |x(t) - z|^2.$$

By differentiation of  $h$ , we obtain

$$\dot{h}_z(t) = \langle \dot{x}(t), x(t) - z \rangle = -\langle \nabla \Phi_0(x(t)), x(t) - z \rangle - \varepsilon(t) \langle \nabla \Phi_1(x(t)), x(t) - z \rangle. \tag{2.1}$$

From the convexity of  $\Phi_0$  and  $\Phi_1$  and since  $z \in \operatorname{argmin} \Phi_0 \cap \operatorname{argmin} \Phi_1$ , we have  $\langle \nabla \Phi_0(x(t)), x(t) - z \rangle \geq 0$  and  $\langle \nabla \Phi_1(x(t)), x(t) - z \rangle \geq 0$ , which combined with (2.1) yields  $\dot{h}_z(t) \leq 0$  and hence  $h_z$  is non increasing; thus

$$\lim_{t \rightarrow +\infty} |x(t) - z| \quad \text{exists for any } z \in \operatorname{argmin} \Phi_0 \cap \operatorname{argmin} \Phi_1. \tag{2.2}$$

Hence, in particular the trajectory  $x(\cdot)$  is bounded and we can extract a converging subsequence: there exist  $\bar{x} \in H$  and  $x(t_n)$  such that  $\lim_{n \rightarrow +\infty} x(t_n) = \bar{x}$ . Since  $x(\cdot)$  is bounded, Corollary 2.2 applies and we have  $\bar{x} \in \operatorname{argmin}_S \Phi_1 = \operatorname{argmin} \Phi_0 \cap \operatorname{argmin} \Phi_1$ . Taking  $z = \bar{x}$  in (2.2), we deduce that  $\lim_{t \rightarrow +\infty} |x(t) - \bar{x}|$  exists, which combined with  $\lim_{n \rightarrow +\infty} |x(t_n) - \bar{x}| = 0$ , finally yields  $\lim_{t \rightarrow +\infty} |x(t) - \bar{x}| = 0$ . □

It is interesting to notice that the result of the previous proposition remains true in infinite dimensional spaces if one strengthens the assumption on  $\varepsilon$ . More precisely, we have the following result

**Proposition 2.5.** *In addition to the hypotheses  $(\mathcal{H})$ , let us assume that*

- $\Phi_0 : H \rightarrow \mathbb{R}$  and  $\Phi_1 : H \rightarrow \mathbb{R}_+$  are convex.
- The set  $\operatorname{argmin} \Phi_0 \cap \operatorname{argmin} \Phi_1$  is non empty.
- There exists  $m > 0$  such that  $\varepsilon(t) \geq m/t$  for  $t$  large enough.

Let  $x(\cdot)$  be the unique trajectory of the (SDC) dynamical system. Then we have

- (i)  $\Phi_0(x(t)) - \min \Phi_0 = o(1/t)$  ( $t \rightarrow +\infty$ ).
- (ii)  $\Phi_1(x(t)) - \min \Phi_1 = o(1/(t\varepsilon(t)))$  ( $t \rightarrow +\infty$ ).
- (iii) The trajectory  $x(\cdot)$  weakly converges to some  $\bar{x} \in \operatorname{argmin} \Phi_0 \cap \operatorname{argmin} \Phi_1$ .

*Proof.* We adopt here the same notations as in the proof of Proposition 2.4. The starting point lies in equality (2.1). From the convexity of  $\Phi_0$  and  $\Phi_1$  and since  $z \in \operatorname{argmin} \Phi_0 \cap \operatorname{argmin} \Phi_1$ , we have  $\langle \nabla \Phi_0(x(t)), x(t) - z \rangle \geq \Phi_0(x(t)) - \min \Phi_0$  and  $\langle \nabla \Phi_1(x(t)), x(t) - z \rangle \geq \Phi_1(x(t)) - \min \Phi_1$ , which combined with (2.1) yields

$$\dot{h}_z(t) + \Phi_0(x(t)) - \min \Phi_0 + \varepsilon(t) (\Phi_1(x(t)) - \min \Phi_1) \leq 0. \quad (2.3)$$

This implies in particular that  $\dot{h}_z(t) \leq 0$  and hence  $h_z$  is non increasing; thus

$$\lim_{t \rightarrow +\infty} |x(t) - z| \text{ exists for any } z \in \operatorname{argmin} \Phi_0 \cap \operatorname{argmin} \Phi_1. \quad (2.4)$$

Coming back to inequality (2.3) and integrating on  $[0, t]$ , we obtain

$$h_z(t) - h_z(0) + \int_0^t \Phi_0(x(s)) - \min \Phi_0 + \varepsilon(s) (\Phi_1(x(s)) - \min \Phi_1) ds \leq 0$$

and therefore

$$\int_0^{+\infty} \Phi_0(x(s)) - \min \Phi_0 + \varepsilon(s) (\Phi_1(x(s)) - \min \Phi_1) ds \leq h_z(0) < +\infty.$$

The energy function  $t \mapsto \Phi_0(x(t)) - \min \Phi_0 + \varepsilon(t) (\Phi_1(x(t)) - \min \Phi_1)$  is non increasing, and from a classical result we deduce that it is negligible compared with  $1/t$  in the neighbourhood of  $+\infty$ . Therefore, we have

$$\Phi_0(x(t)) - \min \Phi_0 = o(1/t) \quad (t \rightarrow +\infty), \quad (2.5)$$

which proves (i), and moreover

$$\Phi_1(x(t)) - \min \Phi_1 = o(1/(t\varepsilon(t))) \quad (t \rightarrow +\infty),$$

which proves (ii). Taking into account the assumption  $\varepsilon(t) \geq m/t$  for  $t$  large enough, we then obtain

$$\lim_{t \rightarrow +\infty} \Phi_1(x(t)) = \min \Phi_1. \quad (2.6)$$

In order to prove the weak convergence of the trajectory, we use the Opial lemma with the set  $\operatorname{argmin} \Phi_0 \cap \operatorname{argmin} \Phi_1$ .

- Let  $z \in \operatorname{argmin} \Phi_0 \cap \operatorname{argmin} \Phi_1$ . From (2.4),  $\lim_{t \rightarrow +\infty} |x(t) - z|$  exists.
- Let us now assume that a subsequence  $(x(t_n))$  weakly converges to some  $\bar{x}$  and prove that  $\bar{x} \in \operatorname{argmin} \Phi_0 \cap \operatorname{argmin} \Phi_1$ . Since  $\Phi_0$  and  $\Phi_1$  are convex and continuous, they are weakly lower semicontinuous and therefore

$$\Phi_0(\bar{x}) \leq \liminf_{n \rightarrow +\infty} \Phi_0(x(t_n)) \quad \text{and} \quad \Phi_1(\bar{x}) \leq \liminf_{n \rightarrow +\infty} \Phi_1(x(t_n)).$$

In view of (2.5) and (2.6), we conclude that  $\Phi_0(\bar{x}) \leq \min \Phi_0$  and  $\Phi_1(\bar{x}) \leq \min \Phi_1$ , i.e.  $\bar{x} \in \operatorname{argmin} \Phi_0 \cap \operatorname{argmin} \Phi_1$ . Then the Opial lemma applies and there exists  $\bar{x} \in \operatorname{argmin} \Phi_0 \cap \operatorname{argmin} \Phi_1$  such that  $x(t) \rightharpoonup \bar{x}$  ( $t \rightarrow +\infty$ ), which proves (iii).  $\square$

### 2.3. Cases of convergence: Conclusion

Unless otherwise specified, we assume in this section that  $H$  is finite dimensional, that  $\Phi_0 : H \rightarrow \mathbb{R}$  and  $\Phi_1 : H \rightarrow \mathbb{R}$  are convex functions and  $\varepsilon : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is a slow control. We assume moreover that  $S := \operatorname{argmin} \Phi_0$  and  $\operatorname{argmin}_S \Phi_1$  are non empty. We are going to explore the situations in which the (SDC) trajectories are known to converge to some point  $\bar{x} \in \operatorname{argmin}_S \Phi_1$ .

- First assume that  $\operatorname{argmin}_S \Phi_1 \cap \operatorname{int}(S) \neq \emptyset$  and let  $x \in \operatorname{argmin}_S \Phi_1 \cap \operatorname{int}(S)$ . Since  $x \in \operatorname{argmin}_S \Phi_1$ , we have  $\nabla \Phi_1(x) \in -N_S(x)$ . On the other hand,  $x \in \operatorname{int}(S)$  implies that  $N_S(x) = \{0\}$  and finally  $\nabla \Phi_1(x) = 0$ , i.e.  $x \in \operatorname{argmin} \Phi_1$ . Since  $x \in \operatorname{argmin} \Phi_0$ , we deduce that  $\operatorname{argmin} \Phi_0 \cap \operatorname{argmin} \Phi_1 \neq \emptyset$ . In such a situation, from Proposition 2.4, the trajectory converges to some point  $\bar{x} \in \operatorname{argmin} \Phi_0 \cap \operatorname{argmin} \Phi_1$ .
- Conversely, assume that  $\operatorname{argmin}_S \Phi_1 \subset \operatorname{bd}(S) = S \setminus \operatorname{int}(S)$ . In such a case, the convex set  $\operatorname{argmin}_S \Phi_1$  has an empty interior, which classically implies that  $\operatorname{argmin}_S \Phi_1$  is contained in some hyperplane  $\mathcal{H}$ . When  $\operatorname{argmin}_S \Phi_1$  is reduced to a singleton  $\{p\}$ , then the (SDC) trajectory converges to  $p$  in view of Corollary 2.2. This is the case for example when
  - the boundary  $\operatorname{bd}(S)$  of  $S$  contains no segment line other than trivial ones;
  - the function  $\Phi_1$  is strictly convex.

In the general case, the question of the convergence of the (SDC) trajectory is still an open problem.

## 3. NUMERICAL ASPECTS

### 3.1. Convergence rate

Since we have numerical purposes in mind, it is crucial to evaluate the convergence rate of the (SDC) trajectory toward its limit. In this direction, estimates (i) and (ii) of Proposition 2.5 give a first answer in the case where  $\operatorname{argmin} \Phi_0 \cap \operatorname{argmin} \Phi_1 \neq \emptyset$ . Finding the expression of the convergence rate in the general case is a difficult and open problem. In the sequel, we focus on the particular situation where the function  $\Phi_0$  is strongly convex. We then show that the distance (at time  $t$ ) between the (SDC) trajectory and its limit is majorized by  $C\varepsilon(t)$ , for some  $C > 0$ . More precisely, we have the following result:

**Proposition 3.1.** *In addition to the hypotheses ( $\mathcal{H}$ ), let us assume that  $\Phi_0$  is convex and admits  $a \in H$  as a strong minimum, i.e. there exists some positive  $\alpha > 0$  such that*

$$\forall x \in H, \quad \Phi_0(x) \geq \Phi_0(a) + \alpha |x - a|^2. \tag{3.1}$$

*We assume moreover that  $\lim_{t \rightarrow +\infty} \dot{\varepsilon}(t)/\varepsilon(t) = 0$ . Then the trajectory  $x(\cdot)$  of the (SDC) system converges toward  $a$  and the following convergence rate holds:*

$$|x(t) - a| = O(\varepsilon(t)) \quad (t \rightarrow +\infty). \tag{3.2}$$

*Proof.* For every  $p \in ]1, 2]$ , let us define the function  $h$  by  $h(t) := \frac{1}{p}|x(t) - a|^p$ . The function  $h$  is differentiable and its first derivative is given by

$$\dot{h}(t) = \langle \dot{x}(t), x(t) - a \rangle |x(t) - a|^{p-2}$$

(observe that the function  $h$  is no more differentiable for  $p = 1$ ). In view of (SDC), the previous expression of  $\dot{h}$  becomes

$$\dot{h}(t) = -\langle \nabla \Phi_0(x(t)), x(t) - a \rangle |x(t) - a|^{p-2} - \varepsilon(t) \langle \nabla \Phi_1(x(t)), x(t) - a \rangle |x(t) - a|^{p-2}. \tag{3.3}$$

From (3.1), the function  $\Phi_0$  is coercive and hence the trajectory  $x(\cdot)$  is bounded. As a consequence, there exists a constant  $C \in \mathbb{R}_+$  (independent of  $p$ ) such that

$$|\langle \nabla \Phi_1(x(t)), x(t) - a \rangle| |x(t) - a|^{p-2} \leq |\nabla \Phi_1(x(t))| |x(t) - a|^{p-1} \leq C. \tag{3.4}$$

On the other hand, from the convexity of  $\Phi_0$ , we have  $\langle \nabla \Phi_0(x(t)), x(t) - a \rangle \geq \Phi_0(x(t)) - \Phi_0(a)$ , which combined with (3.1) yields

$$\langle \nabla \Phi_0(x(t)), x(t) - a \rangle \geq \alpha |x(t) - a|^2. \quad (3.5)$$

Taking into account (3.3), (3.4) and (3.5), we finally obtain

$$\dot{h}(t) + p\alpha h(t) \leq C \varepsilon(t).$$

Let us multiply this last inequality by  $e^{p\alpha t}$  and integrate on  $[0, t]$  to find  $e^{p\alpha t} h(t) - h(0) \leq C \int_0^t \varepsilon(s) e^{p\alpha s} ds$ , which can be rewritten as

$$h(t) \leq e^{-p\alpha t} h(0) + C e^{-p\alpha t} \int_0^t \varepsilon(s) e^{p\alpha s} ds. \quad (3.6)$$

To compute the integral of the right member, let us perform an integration by parts

$$\int_0^t \varepsilon(s) e^{p\alpha s} ds = \frac{1}{p\alpha} (\varepsilon(t) e^{p\alpha t} - \varepsilon(0)) - \frac{1}{p\alpha} \int_0^t \dot{\varepsilon}(s) e^{p\alpha s} ds.$$

Since  $\dot{\varepsilon}(t) = o(\varepsilon(t))$  when  $t \rightarrow +\infty$ , we also have  $\int_0^t \dot{\varepsilon}(s) e^{p\alpha s} ds = o\left(\int_0^t \varepsilon(s) e^{p\alpha s} ds\right)$  and we deduce

$$\int_0^t \varepsilon(s) e^{p\alpha s} ds \sim \frac{1}{p\alpha} (\varepsilon(t) e^{p\alpha t} - \varepsilon(0)) \quad (t \rightarrow +\infty).$$

In view of (3.6), we infer the existence of  $C_1, C_2 \in \mathbb{R}_+$  (independent of  $p \in ]1, 2]$ ) such that, for  $t$  large enough

$$h(t) \leq C_1 e^{-p\alpha t} + C_2 \varepsilon(t). \quad (3.7)$$

Using again the assumption  $\dot{\varepsilon}(t) = o(\varepsilon(t))$ , it is immediate to verify that, for every  $\eta > 0$ ,  $e^{-\eta t} = o(\varepsilon(t))$  in the neighbourhood of  $+\infty$ . This remark combined with (3.7) implies the existence of  $C_3 \in \mathbb{R}_+$  (independent of  $p \in ]1, 2]$ ) such that, for  $t$  large enough,  $h(t) \leq C_3 \varepsilon(t)$ , *i.e.*

$$|x(t) - a|^p \leq p C_3 \varepsilon(t).$$

This being true for every  $p \in ]1, 2]$ , we can pass to the limit when  $p \rightarrow 1$  to obtain  $|x(t) - a| \leq C_3 \varepsilon(t)$ , which concludes the proof.  $\square$

**Remark 3.2.** When  $\Phi_1 \equiv 0$ , it is easy to verify (directly or by using the previous proof) that the map  $t \mapsto |x(t) - a|$  exponentially decreases to 0:  $|x(t) - a| = O(e^{-\alpha t})$ . This speed of convergence is faster than the one given by the previous proposition (especially if the control  $\varepsilon$  is slow). As a consequence, the term  $\varepsilon(t) \nabla \Phi_1(x)$  has a slowing down effect on the convergence rate (when  $\Phi_1 \neq 0$ ).

**Remark 3.3.** The convergence rate given by Proposition 3.1 is optimal in the sense that it cannot be improved in general. Indeed, consider the one-dimensional case  $H = \mathbb{R}$  and given two distinct real numbers  $a$  and  $b$ , take  $\Phi_0(x) = \frac{1}{2}(x - a)^2$  and  $\Phi_1(x) = \frac{1}{2}(x - b)^2$ . With such data, the (SDC) system reduces to a linear differential equation. An elementary computation then leads to the following estimate when  $t \rightarrow +\infty$ :  $x(t) - a \sim (b - a) \varepsilon(t)$ . Hence, the upper bound (3.2) for the convergence rate is achieved, which proves its optimality.

### 3.2. Numerical illustrations

Given a convex set  $S \subset H$  and a convex function  $\Phi_1 : H \rightarrow \mathbb{R}$ , consider the problem of minimizing  $\Phi_1$  over  $S$ :

$$(\mathcal{P}) \quad \min\{\Phi_1(x), \quad x \in S\}.$$

A classical method consists in using the generalized steepest descent system applied to the non-smooth function  $\Phi_1 + \delta_S$ , namely

$$(S1) \quad \dot{x}(t) + \nabla\Phi_1(x(t)) \in -N_S(x(t)).$$

This differential inclusion has the following remarkable property: at each  $t > 0$ , the velocity  $\dot{x}(t)$  selects the element of minimal norm in the set  $-\nabla\Phi_1(x(t)) - N_S(x(t))$ , so that (S1) can be equivalently interpreted as a gradient-projection method

$$\dot{x}(t) = P_{T_S(x(t))}(-\nabla\Phi_1(x(t))),$$

where  $T_S(x(t))$  is the tangent cone to  $S$  at  $x(t)$ . From a numerical point of view, the system (S1) is difficult to handle due to the presence of the multivalued object  $N_S(x)$ .

Antipin [1] has initiated another interesting dynamical system

$$(S2) \quad \dot{x}(t) + x(t) - P_S(x(t) - \mu\nabla\Phi_1(x(t))) = 0,$$

whose trajectories are known to converge toward some minimum of  $\Phi_1$  on the set  $S$ . In the numerical treatment of (S2), the main difficulty comes from the projection operator  $P_S$ . If the structure of the set  $S$  is complex, the mapping  $P_S$  may be quite complicated to compute. In the (SDC) system, the constraint  $S$  arises through the function  $\Phi_0$ . The only request for  $\Phi_0$  is to satisfy  $\text{argmin}\Phi_0 = S$ , which allows a certain latitude in the choice of  $\Phi_0$ . In theoretical problems, it is convenient to take  $\Phi_0(x) = \frac{1}{2}d(x, S)^2$ , whose gradient is given by  $\nabla\Phi_0(x) = x - P_S(x)$ . When the set  $S$  has a particular structure, it is also possible to adapt the function  $\Phi_0$ . For example, consider the convex program

$$\min \Phi_1(x) \quad \text{under the constraint} \quad S := \{x \in H, \theta_1(x) \leq 0, \dots, \theta_n(x) \leq 0\},$$

where  $n \in \mathbb{N}$  and  $\theta_1, \dots, \theta_n$  are convex functions defined on  $H$ . In such a situation, it is pertinent to take

$$\Phi_0(x) = \sum_{i=1}^n (\theta_i^+(x))^2,$$

where  $\theta^+ := \max\{\theta, 0\}$  denotes the non-negative part of  $\theta$ . The gradient of such a function  $\Phi_0$  is easier to compute than the corresponding projection mapping  $P_S$ .

We are now going to compare (SDC) with (S1) and (S2) on a simple example of quadratic programming. We want to minimize the function  $\Phi_1(x, y) = \frac{1}{2}(4(x+y+1)^2 + (x-y-1)^2)$  on the non-negative orthant  $S = \mathbb{R}_+^2$ :

$$(P) \quad \min \left\{ \frac{1}{2}(4(x+y+1)^2 + (x-y-1)^2), \quad x \geq 0, \quad y \geq 0 \right\},$$

whose solution is the point  $(0, 0)$ . The respective trajectories of (SDC), (S1) and (S2) are shown in Figure 1 (left). The points of the different trajectories are computed by an explicit discretization technique. The (SDC) system is performed with  $\Phi_0(x, y) = 1/2 d((x, y), S)^2$  and  $\varepsilon(t) = 1/t$  for  $t \geq 1$ . The (S2) trajectory is drawn with the value  $\mu = 0.1$ . The initial condition for the three curves is  $(x_0, y_0) = (5, 2)$ . We observe that the (SDC) system acts as a penalization method whereas (S2) is an interior point method. The direction of the (S1) trajectory is a "pure" gradient one before hitting the boundary of  $S$  and then follows the boundary. It is interesting to test the influence of the control parameter  $\varepsilon$  on the (SDC) trajectories. Figure 1 (right) shows the (SDC) trajectories corresponding respectively to  $\varepsilon(t) = 1/(t \ln t)$ ,  $1/t$ ,  $1/\sqrt{t}$  for  $t \geq 2$ . We notice that, the more the control  $\varepsilon$  is slow, the more the (SDC) trajectory escapes from the "attraction" of the set  $S$ . Figure 2 shows the evolution of the distance  $\sqrt{x^2 + y^2}$  between the trajectory (SDC) (resp. (S1), (S2)) and the point  $(0, 0)$ . We observe that the convergence of the (SDC) (resp. (S1), (S2)) system is approximately achieved for  $t = 10$  (resp.  $t = 20$ ,  $t = 5$ ). This numerical experiment seems to indicate that the convergence rate of (SDC) is comprised between the other two ones. It is not surprising that the convergence speed of (SDC) is rather slow.

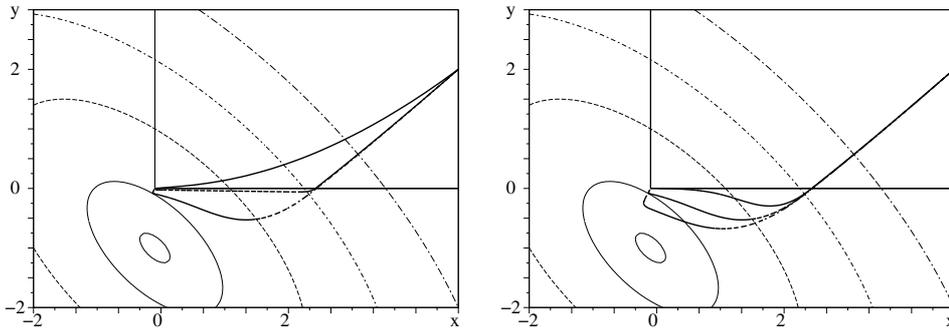


FIGURE 1. Left: comparison of the  $(SDC)$ ,  $(S1)$  and  $(S2)$  trajectories. Right: influence of the control parameter  $\varepsilon$  on the  $(SDC)$  trajectories:  $\varepsilon(t) = 1/(t \ln t)$ ,  $1/t$ ,  $1/\sqrt{t}$  for  $t \geq 2$ .

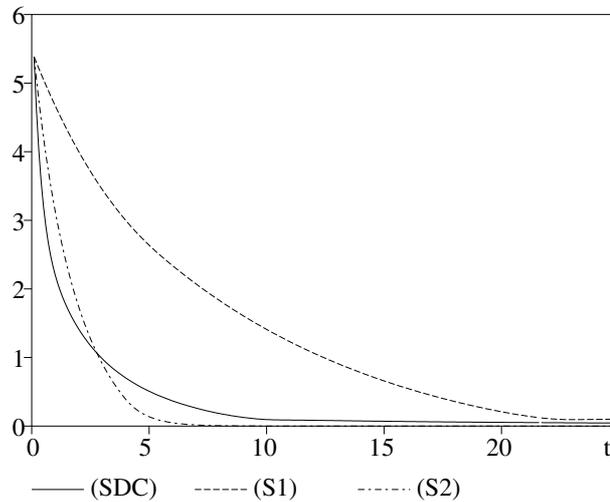


FIGURE 2. Comparison of the convergence rates for the  $(SDC)$ ,  $(S1)$  and  $(S2)$  trajectories.

Indeed, in view of Paragraph 3.1 it is majorized by  $\varepsilon(t)$ , which is itself a slow control. Notice that, if the control map  $\varepsilon$  tends to 0 too slowly, the convergence of  $(SDC)$  will be very long to obtain. On the other hand, if one chooses a too fast control  $\varepsilon$ , the trajectory  $(SDC)$  may not converge to some element of  $\text{argmin}_S \Phi_1$ . Therefore, in numerical applications, one has to find a suitable balance between these two extremes.

#### 4. PROOF OF THE MAIN RESULTS

In this section, we give the technical proofs of the main results of the paper.

##### 4.1. Proof of Theorem 2.1

Let us first remark, that if we are able to prove  $(i)$ , then we immediately infer  $(ii)$  and  $(iii)$ . Indeed, let  $(\Phi_0(x(t_n)))_{n \geq 0}$  (resp.  $(\Phi_1(x(t_n)))_{n \geq 0}$ ) a converging subsequence of the bounded map  $(\Phi_0(x(t)))_{t \geq 0}$  (resp.  $(\Phi_1(x(t)))_{t \geq 0}$ ). Since  $x(t_n)$  is bounded, there is a subsequence of  $x(t_n)$ , still denoted by  $x(t_n)$  which converges to  $\bar{x} \in H$ . From  $(i)$ , we have  $\lim_{n \rightarrow +\infty} d(x(t_n), \text{argmin}_S \Phi_1) = 0$  and hence  $\bar{x} \in \text{argmin}_S \Phi_1 \subset \text{argmin}_S \Phi_0$ .

As a consequence,  $\lim_{n \rightarrow +\infty} \Phi_0(x(t_n)) = \min \Phi_0$  (resp.  $\lim_{n \rightarrow +\infty} \Phi_1(x(t_n)) = \min_S \Phi_1$ ). Since  $\min \Phi_0$  (resp.  $\min_S \Phi_1$ ) is the limit of every converging subsequence of  $(\Phi_0(x(t)))_{t \geq 0}$  (resp.  $(\Phi_1(x(t)))_{t \geq 0}$ ), we deduce that  $\lim_{t \rightarrow +\infty} \Phi_0(x(t)) = \min \Phi_0$  (resp.  $\lim_{t \rightarrow +\infty} \Phi_1(x(t)) = \min_S \Phi_1$ ), which proves (ii) (resp. (iii)).

Let us now come back to the proof of (i). It relies on the study of the function  $h$  defined by

$$h(t) = \frac{1}{2}d(x(t), C)^2,$$

where  $C := \operatorname{argmin}_S \Phi_1 \subset \operatorname{argmin} \Phi_0$ . We have to prove that  $h$  converges to 0. The following claim tells us that  $C$  is convex and hence, that  $h$  is differentiable.

**Claim 4.1.** Under the assumptions of Theorem 2.1, we have

$$\operatorname{argmin}_S \Phi_1 = \operatorname{argmin} \Phi_0 \cap [\Phi_1 \leq \min_S \Phi_1],$$

and hence the set  $C := \operatorname{argmin}_S \Phi_1$  is convex as an intersection of convex sets.

*Proof.* Let us remark that  $\operatorname{argmin}_S \Phi_1 \subset [\Phi_1 \leq \min_S \Phi_1]$ , which combined with the assumption  $\operatorname{argmin}_S \Phi_1 \subset \operatorname{argmin} \Phi_0$  gives

$$\operatorname{argmin}_S \Phi_1 \subset \operatorname{argmin} \Phi_0 \cap [\Phi_1 \leq \min_S \Phi_1].$$

On the other hand, since  $\operatorname{argmin} \Phi_0 \subset S$ ,

$$\operatorname{argmin} \Phi_0 \cap [\Phi_1 \leq \min_S \Phi_1] \subset S \cap [\Phi_1 \leq \min_S \Phi_1] = \operatorname{argmin}_S \Phi_1,$$

which proves the second inclusion. □

By differentiating  $h$ , we find

$$\dot{h}(t) = -\langle x(t) - P_C(x(t)), \nabla \Phi_0(x(t)) \rangle - \varepsilon(t) \langle x(t) - P_C(x(t)), \nabla \Phi_1(x(t)) \rangle.$$

From the quasi-convexity of  $\Phi_0$ , the inequality  $\min \Phi_0 = \Phi_0(P_C(x(t))) \leq \Phi_0(x(t))$  implies

$$\langle P_C(x(t)) - x(t), \nabla \Phi_0(x(t)) \rangle \leq 0,$$

and hence,

$$\dot{h}(t) \leq \varepsilon(t) \langle P_C(x(t)) - x(t), \nabla \Phi_1(x(t)) \rangle. \tag{4.1}$$

The main idea of the proof is now to respectively distinguish the cases where  $\langle P_C(x(t)) - x(t), \nabla \Phi_1(x(t)) \rangle > 0$  and  $\langle P_C(x(t)) - x(t), \nabla \Phi_1(x(t)) \rangle \leq 0$ . Precisely, we distinguish the three cases:

- (a)  $\exists T \geq 0, \forall t \geq T \quad \langle P_C(x(t)) - x(t), \nabla \Phi_1(x(t)) \rangle \leq 0;$
- (b)  $\exists T \geq 0, \forall t \geq T \quad \langle P_C(x(t)) - x(t), \nabla \Phi_1(x(t)) \rangle > 0;$
- (c)  $\forall T \geq 0, \exists t \geq T \quad \langle P_C(x(t)) - x(t), \nabla \Phi_1(x(t)) \rangle > 0.$

Case (c) obviously contains case (b), but the main points of the proof are made clearer with this distinction. In the whole proof, we use the following notation

$$E_C := \{x \in H, \quad \langle P_C(x) - x, \nabla \Phi_1(x) \rangle \geq 0\}.$$

Let us state a first claim which will be useful in the sequel.

**Claim 4.2.** Under the assumptions of Theorem 2.1, we have  $E_C \cap S = C$ .

*Proof.* First of all, it is immediate that  $C \subset E_C \cap S$ . Let us prove the reverse inclusion  $E_C \cap S \subset C$ . Let  $x \in E_C \cap S$ ; from the definition of  $E_C$ , we have

$$\langle P_C(x) - x, \nabla\Phi_1(x) \rangle \geq 0,$$

which combined with the strict quasi-convexity of  $\Phi_1$ , yields  $\Phi_1(P_C(x)) \geq \Phi_1(x)$ , *i.e.*  $\min_S \Phi_1 \geq \Phi_1(x)$ . Since  $x \in S$ , we conclude that  $x \in C = \operatorname{argmin}_S \Phi_1$ .  $\square$

Case (a). We assume that there is  $T \geq 0$ , such that, for every  $t \geq T$ ,  $\langle P_C(x(t)) - x(t), \nabla\Phi_1(x(t)) \rangle \leq 0$ , hence from (4.1), we deduce that, for every  $t \geq T$ ,  $\dot{h}(t) \leq 0$ . This implies that  $\lim_{t \rightarrow +\infty} h(t)$  exists and hence

$$\lim_{t \rightarrow +\infty} d(x(t), C) \text{ exists.} \tag{4.2}$$

We now prove that  $\lim_{t \rightarrow +\infty} d(x(t), C) = 0$ . Let us argue by contradiction and assume that  $\lim_{t \rightarrow +\infty} d(x(t), C) > 0$ . Let us first prove that  $\limsup_{t \rightarrow +\infty} \langle P_C(x(t)) - x(t), \nabla\Phi_1(x(t)) \rangle < 0$ . Indeed, if it is not true, there exists a sequence  $(t_n)$  such that  $\lim_{n \rightarrow +\infty} t_n = +\infty$  and  $\lim_{n \rightarrow +\infty} \langle P_C(x(t_n)) - x(t_n), \nabla\Phi_1(x(t_n)) \rangle = 0$ . Since, the function  $x$  is bounded, without any loss of generality, we may assume that there is  $\bar{x} \in H$  such that  $x(t_n)$  converges to  $\bar{x}$ . At the limit when  $n \rightarrow +\infty$ , we obtain that

$$\langle P_C(\bar{x}) - \bar{x}, \nabla\Phi_1(\bar{x}) \rangle = 0,$$

which implies  $\bar{x} \in E_C$ . From Proposition 1.1 (ii), we get that  $\lim_{n \rightarrow +\infty} \nabla\Phi_0(x(t_n)) = 0$ , and hence  $\bar{x} \in S$ . Finally  $\bar{x} \in E_C \cap S = C$ . On the other hand, we have  $d(\bar{x}, C) = \lim_{n \rightarrow +\infty} d(x(t_n), C) > 0$ , a contradiction. As a consequence,  $\limsup_{t \rightarrow +\infty} \langle P_C(x(t)) - x(t), \nabla\Phi_1(x(t)) \rangle < 0$ , *i.e.* there exists  $l > 0$  and  $t_1 > 0$  such that, for every  $t \geq t_1$ ,

$$\langle P_C(x(t)) - x(t), \nabla\Phi_1(x(t)) \rangle \leq -l.$$

Hence  $\dot{h}(t) \leq -\varepsilon(t)l$ . By integrating this last inequality between  $t_1$  and  $t$  and passing to the limit when  $t \rightarrow +\infty$ , we get

$$\lim_{t \rightarrow +\infty} h(t) + l \int_{t_1}^{+\infty} \varepsilon(s) \, ds \leq h(t_1)$$

which contradicts the fact that  $\varepsilon \notin L^1(0, +\infty)$ . Hence we conclude that  $\lim_{t \rightarrow +\infty} d(x(t), C) = 0$ .

Case (b). We assume that there is  $T \geq 0$ , such that, for every  $t \geq T$ ,  $\langle P_C(x(t)) - x(t), \nabla\Phi_1(x(t)) \rangle > 0$  and hence for every  $t \geq T$ ,  $x(t) \in E_C$ .

Since  $x(\cdot)$  is bounded, from Proposition 1.1 (ii), we get that  $\lim_{t \rightarrow +\infty} \nabla\Phi_0(x(t)) = 0$ . Considering a subsequence  $x(t_n)$ , we then have  $\lim_{n \rightarrow +\infty} \nabla\Phi_0(x(t_n)) = 0$ . Applying the following claim to the sequence  $(x(t_n))$ , we obtain that  $\lim_{n \rightarrow +\infty} d(x(t_n), C) = 0$ , which concludes the proof of case (b).

**Claim 4.3.** Let  $(x_n)$  be a bounded sequence in  $E_C$ , such that  $\lim_{n \rightarrow +\infty} \nabla\Phi_0(x_n) = 0$ . Then  $\lim_{n \rightarrow +\infty} d(x_n, C) = 0$ .

*Proof.* Let  $(d(x_{\sigma(n)}, C))_{n \geq 0}$  a converging subsequence of  $(d(x_n, C))_{n \geq 0}$ . Since  $(x_n)$  is bounded, there exist  $\bar{x} \in E_C$  and a subsequence of  $(x_{\sigma(n)})$ , still denoted by  $(x_{\sigma(n)})$  such that  $\lim_{n \rightarrow +\infty} x_{\sigma(n)} = \bar{x}$ . In view of  $\lim_{n \rightarrow +\infty} \nabla\Phi_0(x_{\sigma(n)}) = 0$ , we get that  $\bar{x} \in S$ . On the other hand, from Claim 4.2,  $E_C \cap S = C$ , so that we obtain  $\bar{x} \in C$  and hence  $\lim_{n \rightarrow +\infty} d(x_{\sigma(n)}, C) = 0$ . Since 0 is the limit of every convergent subsequence of  $(d(x_n, C))_{n \geq 0}$ , we deduce that the sequence  $(d(x_n, C))_{n \geq 0}$  converges to 0.  $\square$

Case (c). We now assume that, for every  $T \geq 0$ , there exists some  $t \geq T$  such that  $\langle P_C(x(t)) - x(t), \nabla\Phi_1(x(t)) \rangle > 0$ . Take any sequence  $(t_n) \subset \mathbb{R}_+$  such that  $\lim_{n \rightarrow +\infty} t_n = +\infty$  and let us prove that  $\lim_{n \rightarrow +\infty} d(x(t_n), C) = 0$ . First assume that there is a subsequence  $(t'_n)$  of  $(t_n)$  such that  $x(t'_n) \in E_C$ . Since the map  $x$  is bounded, from

Proposition 1.1 (ii), we have  $\lim_{t \rightarrow +\infty} \nabla \Phi_0(x(t)) = 0$  and hence  $\lim_{n \rightarrow +\infty} \nabla \Phi_0(x(t'_n)) = 0$ . Since  $x(t'_n) \in E_C$ , from Claim 4.3, we deduce that  $\lim_{n \rightarrow +\infty} d(x(t'_n), C) = 0$ .

We now assume that there is a subsequence  $(t''_n)$  of  $(t_n)$  such that  $x(t''_n) \notin E_C$  and we prove that  $\lim_{n \rightarrow +\infty} d(x(t''_n), C) = 0$ . We need the following claim.

**Claim 4.4.** Let  $t \geq 0$  such that  $x(t) \notin E_C$ , and let

$$\tau(t) = \inf \left\{ u \in [0, t] \mid x([u, t]) \cap E_C = \emptyset \right\}.$$

Then we have  $d(x(t), C) \leq d(x(\tau(t)), C)$ .

*Proof.* For every  $u \in ]\tau(t), t]$ ,  $x(u) \notin E_C$ , that is,  $\langle P_C(x(u)) - x(u), \nabla \Phi_1(x(u)) \rangle \leq 0$ . From (4.1), we deduce that  $\dot{h}(u) \leq 0$ , which immediately yields the expected inequality.  $\square$

We now come back to the proof of case (c). Since  $x(t''_n) \notin E_C$ , let  $\tau(t''_n)$  be defined by Claim 4.4. We first notice that  $\lim_{n \rightarrow +\infty} \tau(t''_n) = +\infty$  and  $x(\tau(t''_n)) \in E_C$  for  $n$  large enough. Let  $n$  be large enough. Since  $x(\tau(t''_n)) \in E_C$ , from Claim 4.3, we have  $\lim_{n \rightarrow +\infty} d(x(\tau(t''_n)), C) = 0$ . Hence, in view of Claim 4.4, we deduce that  $\lim_{n \rightarrow +\infty} d(x(t''_n), C) = 0$ , which concludes the proof of case (c).

### 4.2. Proof of Proposition 2.3

We keep here the notations of the proof of Theorem 2.1. Setting  $C := \operatorname{argmin}_S \Phi_1$ , we still use the function  $h$  defined by  $h(t) = \frac{1}{2}d(x(t), C)^2$  and the set  $E_C = \{x \in H, \langle P_C(x) - x, \nabla \Phi_1(x) \rangle \geq 0\}$ . From Claim 4.1, the set  $C$  equals to  $\operatorname{argmin} \Phi_0 \cap [\Phi_1 \leq \min_S \Phi_1]$ , which is bounded in view of condition (C).

From the energy decay, there exists  $E_0 \in \mathbb{R}$  such that, for every  $t \geq 0$ ,  $\Phi_0(x(t)) \leq E_0$ , i.e.

$$\{x(t), t \geq 0\} \subset [\Phi_0 \leq E_0]. \tag{4.3}$$

On the other hand, from the strict quasi-convexity of  $\Phi_1$ , we have for every  $x \in H$ ,

$$\langle P_C(x) - x, \nabla \Phi_1(x) \rangle \geq 0 \implies \Phi_1(x) \leq \Phi_1(P_C(x)) = \min_S \Phi_1$$

and therefore

$$E_C \subset \left[ \Phi_1 \leq \min_S \Phi_1 \right]. \tag{4.4}$$

We now distinguish the same cases (a), (b) and (c) as in the proof of Theorem 2.1.

Case (a). From (4.2),  $\lim_{t \rightarrow +\infty} d(x(t), C)$  exists, which implies that the map  $t \mapsto d(x(t), C)$  is bounded and since  $C$  is bounded, we conclude that the map  $x$  is bounded too.

Case (b). The trajectory  $\{x(t), t \geq T\}$  is contained in  $E_C$ , hence from (4.3) and (4.4) in the set  $[\Phi_1 \leq \min_S \Phi_1] \cap [\Phi_0 \leq E_0]$ , which is bounded in view of condition (C); so the map  $x$  is bounded on  $[T, +\infty)$  hence, since it is continuous, it is bounded on  $[0, +\infty)$ .

Case (c). Let  $T \geq 0$  such that  $x(T) \in E_C$  and consider  $t \geq T$ . If  $x(t) \in E_C$ , then  $d(x(t), C) \leq \rho$ , where  $\rho$  is the diameter of the bounded set  $[\Phi_1 \leq \min_S \Phi_1] \cap [\Phi_0 \leq E_0]$ . If  $x(t) \notin E_C$ , let  $\tau(t)$  be defined by Claim 4.4. Clearly  $T \leq \tau(t) < t$  and  $x(\tau(t)) \in E_C$ , which implies that  $d(x(\tau(t)), C) \leq \rho$ . In view of Claim 4.4, we deduce that  $d(x(t), C) \leq \rho$ . This proves that the map  $t \mapsto d(x(t), C)$  is bounded on  $[T, +\infty)$ . Since it is continuous, it is bounded on the interval  $[0, +\infty)$ . Using again the boundedness of  $C$ , this implies that the map  $x$  is bounded.

*Acknowledgements.* The author would like to express his gratitude to Prof. H. Attouch and M.-O. Czarnecki for fruitful discussions and to the anonymous referees for some comments which were very useful to improve the presentation of the paper.

## REFERENCES

- [1] A.S. Antipin, Minimization of convex functions on convex sets by means of differential equations. *Differ. Equ.* **30** (1994) 1365-1375 (1995).
- [2] V. Arnold, *Equations différentielles ordinaires*. Éditions de Moscou (1974).
- [3] H. Attouch and R. Cominetti, A dynamical approach to convex minimization coupling approximation with the steepest descent method. *J. Differ. Equ.* **128** (1996) 519-540.
- [4] H. Attouch and M.-O. Czarnecki, Asymptotic control and stabilization of nonlinear oscillators with non isolated equilibria. *J. Differ. Equ.* **179** (2002) 278-310.
- [5] H. Brézis, Opérateurs maximaux monotones dans les espaces de Hilbert et équations d'évolution. *Lect. Notes* **5** (1972).
- [6] R.E. Bruck, Asymptotic convergence of nonlinear contraction semigroups in Hilbert space. *J. Funct. Anal.* **18** (1975) 15-26.
- [7] A. Cabot and M.-O. Czarnecki, Asymptotic control of pairs of oscillators coupled by a repulsion, with non isolated equilibria. *SIAM J. Control Optim.* **41** (2002) 1254-1280.
- [8] A. Haraux, *Systèmes dynamiques dissipatifs et applications*. RMA 17, Masson, Paris (1991).
- [9] W. Hirsch and S. Smale, *Differential equations, dynamical systems and linear algebra*. Academic Press, New York (1974).
- [10] J.P. Lasalle and S. Lefschetz, *Stability by Lyapounov's Direct Method with Applications*. Academic Press, New York (1961).
- [11] Z. Opial, Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bull. Amer. Math. Soc.* **73** (1967) 591-597.
- [12] H. Reinhardt, *Equations différentielles. Fondements et applications*. Dunod, Paris, 2<sup>e</sup> edn. (1989).
- [13] A.N. Tikhonov and V.Ya. Arsenine, *Méthodes de résolution de problèmes mal posés*. MIR (1976).