JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 119, No. 2, pp. 235–259, November 2003 (© 2003)

Continuous Gradient Projection Method in Hilbert Spaces

J. BOLTE[1]

Communicated by J. P. Crouzeix

Abstract. This paper is concerned with the asymptotic analysis of the trajectories of some dynamical systems built upon the gradient projection method in Hilbert spaces. For a convex function with locally Lipschitz gradient, it is proved that the orbits converge weakly to a constrained minimizer whenever it exists. This result remains valid even if the initial condition is chosen out of the feasible set and it can be extended in some sense to quasiconvex functions. An asymptotic control result, involving a Tykhonov-like regularization, shows that the orbits can be forced to converge strongly toward a well-specified minimizer. In the finite-dimensional framework, we study the differential inclusion obtained by replacing the classical gradient by the subdifferential of a continuous convex function. We prove the existence of a solution whose asymptotic properties are the same as in the smooth case.

Key Words. Gradient projection methods, dissipative dynamical systems in optimization, differential inclusions, asymptotic control, Lyapunov functions.

1. Introduction

Let H be a real Hilbert space endowed with the scalar product $\langle\, , \rangle$ and its related norm $|\cdot|$. If C is a closed, nonempty convex set in H, we denote by P_C the corresponding orthogonal projection and by $N_C(x)$ the normal cone to C at x. The indicator function of C is denoted by δ_C; let us recall that δ_C is defined on H with value 0 on C and $+\infty$ elsewhere.

[1]Doctor in Mathematics, ACSIOM-CNRS FRE 2311, Département de Mathématiques, Université Montpellier II, Montpellier, France.

Our main purpose being to minimize a convex function $\phi: H \to \mathbb{R}$ over C, we study the following type of dynamical system:

$$\text{(CGP)} \quad \dot{x}(t) + x(t) - P_C[x(t) - \mu\nabla\phi(x(t))] = 0, \qquad t \geq 0,$$
$$x(0) = x_0 \in C, \qquad \mu > 0,$$

where (CGP) stands for continuous gradient projection method.

Many primal continuous methods to perform this kind of optimization problem consist in adding some barrier function or penalty function to ϕ and then studying the new potential with a classical procedure like steepest descent. From a theoretical viewpoint, these approaches can be seen as smooth approximations of the following problem:

$$\inf_H (\phi + \delta_C). \tag{1}$$

When combining such a formulation with the steepest descent method, one is led to study

$$\dot{x}(t) + \nabla\phi(x(t)) + N_C(x(t)) \ni 0, \tag{2}$$

or equivalently [see Brézis (Ref. 1)]

$$\dot{x}(t^+) = P_{T_C(x(t))}[-\nabla\phi(x(t))],$$

where $T_C(x)$ is the tangent cone to C at x. If ϕ is a convex lower-semicontinuous proper function, Bruck (Ref. 2.) has proved that the trajectories of (2) converge to a minimizer of $\phi + \delta_C$ whenever it exists. But this solving procedure has a major drawback: the dynamics ignores the constraints until the orbit encounters the boundary of C.

This can be improved by a careful examination of the optimality condition associated to (1); indeed, the following two conditions are equivalent:

$$(\mathcal{O}) \quad \nabla\phi(x) + N_C(x) \ni 0,$$

$$(\mathcal{O}') \quad \exists \mu > 0 \text{ s.t. } x = P_C(x - \mu\nabla\phi(x)).$$

This reformulation of $\mathcal{O}$ is well known in discrete optimization; it has led to study algorithms of the type

$$x_{k+1} \in P_C(x_k - \mu\partial\phi(x_k)), \qquad x_0 \in C, \tag{3}$$

where $\partial\phi$ is the subdifferential of ϕ. For some theoretical studies in Hilbert spaces, see Polyak (Ref. 3), MacCormick and Tapia (Ref. 4), Martinet (Ref. 5), and Phelps (Ref. 6). If ϕ is only assumed proper, lower semicontinuous, and convex, the convergence of the sequence (3) is, as far as we know, an open question. In a recent work (Ref. 7), Alber, Iusem, and Solodov have obtained the weak convergence of the orbits under a local boundedness assumption on $\partial\phi$.

As a continuous dynamical system, (CGP) enjoys much stronger properties than its corresponding explicit discretizations; we will see that actually it can be considered as an interior version of (3); see Remark 6.1. The evolution problem (CGP) has been tackled by Antipin (Ref. 8) in the finite-dimensional case with a gradient Lipschtiz continuous on the whole of H. For a second-order version of (CGP), interesting results have been obtained in Ref. 9 and in Alvarez and Attouch (Ref. 9), but under the same strong assumptions on the gradient.

In this paper, the scope of the results obtained in Ref. 8 for the smooth case has been enlarged considerably. In our framework, H is an Hilbert space, ϕ is a $\mathscr{C}^1$ function not necessarily convex, and its gradient is supposed to be only Lipschtiz continuous on bounded sets. Moreover, no restriction is imposed on the stepsize μ. In Section 2, it is proved that (CGP) is a descent method generating viable trajectories; i.e., $\forall t \geq 0, x(t) \in C$.

The asymptotic behavior of the orbits when ϕ is convex or quasiconvex is a delicate matter. In Baillon (Ref. 10), one can find an example in which the trajectories of (2) do not converge strongly to an equilibrium. A key tool in the study of the convergence of the steepest descent method is the association of the Fejer monotonocity with the Opial lemma (Ref. 11); see also Section 3. To be more precise, the quadratic functionals $y \in H \rightarrow 1/2|y - x^*|^2$, where x^* is some stationary point of the studied potential, are Lyapunov functionals for the system (2), allowing us to obtain weak convergence via the Opial lemma. Due to the lack of monotonocity of the operator

$$y \rightarrow -y + P_C(y - \mu\nabla\phi(y)),$$

we propose an alternative approach of the asymptotic behavior, showing that the distance-like functions

$$y \rightarrow \mu[\phi(y) - \phi(x^*)] + (1/2)|y - x^*|^2$$

are decreasing along (CGP) trajectories. This allows us to derive the weak convergence of the solution of (CGP) to a minimizer of ϕ over C; see Section 3 and Fig. 1.

As noticed in Ref. 8, (CGP) conserves its optimizing properties even if the initial condition is not feasible; this result is extended to infinite-dimensional spaces by use of an Opial-like lemma concerning a class of Lyapunov functionals. Figure 1 gives an illustration of these results, with

$$\phi(x_1, x_2) = (1/2)(x_1 - x_2 - 5)^2 + (1/2)(2x_1 + x_2 - 4)^2 \quad \text{and} \quad C = \mathbb{R}_+ \times \mathbb{R}_+.$$

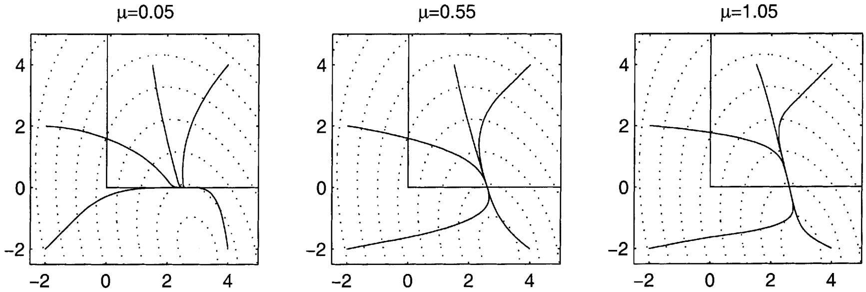

Fig. 1. Some trajectories of (CGP).

Five initial conditions have been chosen in and out of C and three different stepsizes have been used for the computations (the dashed lines delimit the lower level subsets of ϕ).

In Section 5, an asymptotic control result is obtained by considering the nonautonomous system

$$(\mathrm{CGP})_\epsilon \quad \dot{x}(t) + x(t) - P_C[x(t) - \mu\nabla\phi(x(t)) - \epsilon(t)x(t)] = 0,$$
$$x(0) = x_0 \in C.$$

where $\epsilon\colon \mathbb{R}_+ \to \mathbb{R}_+$ satisfies

$$\int_0^{+\infty} \epsilon = +\infty.$$

The Tykhonov term $t \to \epsilon(t)x(t)$ is used to force the orbits to attain a particular equilibrium for the strong topology. This work is inspired by Attouch–Cominetti and Attouch–Czarnecki (Refs. 12–13), where the authors are concerned with the steepest descent and heavy ball with friction systems.

Section 6 is devoted to the following nonsmooth version of (CGP);

$$(\mathrm{CGP})_g \quad \dot{x}(t) + x(t) - P_C[x(t) - \mu\partial\phi(x(t))] \ni 0, \qquad x(0) = x_0 \in C,$$

where ϕ is a continuous convex function on a finite-dimensional space.

Note that the multivalued vector field ruling this equation has neither the convexity property nor the regularity properties usually required in differential inclusion theory (Aubin–Cellina, Ref. 14). In order to prove the existence of a global solution, we define approximated differential systems by using the Moreau–Yosida regularization; then, obtaining estimations on the approximated trajectories, we derive by compacity arguments that $(\mathrm{CGP})_g$ actually admits a solution. This is a classical approach to solve a nonsmooth differential inclusions; see for instance Ref. 1 and Schatzman

(Ref. 15) for second-order in time systems. The asymptotic properties of $(\mathrm{CGP})_g$ are the same as in the smooth case.

2. Global Existence Results for Feasible Initial Data

In what follows, ϕ is a function from H into $\mathbb{R}$. For a given closed, nonempty convex subset C of H, we consider the following set of hypotheses:

$(\mathscr{H})$ ϕ is $\mathscr{C}^1$, bounded from below on C, and $\nabla\phi$ is Lipschitz continuous on bounded sets.

The continuous gradient projection method is given by

$$(\mathrm{CGP}) \quad \dot{x}(t) + x(t) - P_C[x(t) - \mu\nabla\phi(x(t))] = 0, \qquad x(0) = x_0 \in H,$$

where $\mu > 0$ is a positive parameter.

Theorem 2.1. Let us assume that ϕ satisfies $(\mathscr{H})$. Then, the following properties hold:

(i) For all $x_0 \in C$, there exists a unique solution x of (CGP) such that $x \in C^1([0, +\infty[; H)$.
(ii) The trajectory satisfies the following viability condition: $\forall t \geq 0$, $x(t) \in C$.
(iii) (CGP) is a descent method; more precisely, we have

$$(d/dt)\phi(x(t)) \leq -(1/\mu)|\dot{x}(t)|^2.$$

As a consequence, $\dot{x} \in L^2(0, +\infty; H)$.
(iv) If $t \to x(t)$ is bounded, then $\dot{x}(t) \to 0$ as $t \to +\infty$.

Proof. Since P_C is a Lipschitz continuous operator, the Cauchy–Lipschitz theorem yields the existence of a unique solution of (CGP) defined on some interval $[0, T]$ with $T > 0$. Let us show that, for all $t \in [0, T]$, $x(t) \in C$. (CGP) can be rewritten as

$$\dot{x}(t) + x(t) = f(t),$$

where

$$f(\cdot) = P_C[x(\cdot) - \mu\nabla\phi(x(\cdot))]$$

is a continuous function taking its values in C. A simple integration procedure gives

$$x(t) = \exp(-t)x_0 + \exp(-t)\int_0^t f(s)\exp(s)\,ds.$$

Set

$$\mu_t = \frac{\exp(s)}{\exp(t)-1} 1_{|[0,t]}\,ds.$$

Then,

$$\mu_t([0,+\infty[) = 1$$

and

$$x(t) = \exp(-t)x_0 + (1-\exp(-t))\int_0^t f(s)\,d\mu_t. \tag{4}$$

Since

$$f(s)\in C, \qquad \forall s\in[0,t],$$

it is easy to check that

$$\int_0^t f(s)\,d\mu_t \in C;$$

thus, (4) shows that

$$x(t)\in C, \qquad \forall t \text{ in } [0,T].$$

Let us now deal with (iii). For all $t\in[0,T]$, set

$$\xi(t) = x(t) - \mu\nabla\phi(x(t)).$$

Using (ii) and the optimality property of $P_C(\xi(t))$, we have

$$\langle x(t) - P_C(\xi(t)), \xi(t) - P_C(\xi(t))\rangle \le 0;$$

thus, by (CGP),

$$\langle -\dot{x}(t), -\mu\nabla\phi(x(t)) - \dot{x}(t)\rangle \le 0,$$

whence, for all t in $[0,T]$,

$$\mu(d/dt)\phi(x(t)) + |\dot{x}(t)|^2 \le 0. \tag{5}$$

Arguing by contradiction, it is now classical to prove that the trajectories are defined on the whole of $\mathbb{R}_+$. To prove (iv), observe that, if x is bounded, then (CGP) and ($\mathscr{H}$) imply that $\dot{x}$ is Lipschitz continuous. Combining this fact with (iii), it is easy to check out that

$$\lim_{t\to+\infty} \dot{x}(t) = 0. \qquad \square$$

3. Convex Minimization

3.1. Asymptotic Behavior. This section is devoted to the study of (CGP) with a convex ϕ. Let us set

$$S := \operatorname*{argmin}_C \phi = \left\{ x \in C,\ \phi(x) = \inf_C \phi \right\}.$$

Theorem 3.1. ϕ is supposed to be convex and to satisfy ($\mathscr{H}$). As before, we assume that $x_0 \in C$. Then, the following property holds:

(i) $\lim_{t\to+\infty} \phi(x(t)) = \inf_C \phi$.

Moreover, assume that $S \neq \emptyset$. Then:

(ii) There exists $M \geq 0$ such that $\phi(x(t)) - \inf_C \phi \leq M/(t+1)$, $t \geq 0$.
(iii) $x(t)$ converges weakly to some minimizer of ϕ over C as $t \to +\infty$.

Part (i) of this theorem is inspired by the Lemaire work on the steepest descent method (Ref. 16), in which it is proved that it is not necessary to suppose that $S \neq \emptyset$ to obtain a proper optimizing method. As in Refs. 1 and 17, the asymptotic analysis relies on the following lemma.

Lemma 3.1. See Opial (Ref. 11). Let H be a Hilbert space and let $x: [0, +\infty[\to H$ be a curve such that there exists a nonempty set $\mathscr{S} \subset H$ which satisfies the conditions below:
(i) All weak cluster points of x are contained in $\mathscr{S}$.
(ii) $\lim_{t\to+\infty} |x(t) - x^*|$ exists $\forall x^* \in \mathscr{S}$.
Then, $x(t)$ converges weakly to an element of $\mathscr{S}$ as $t \to +\infty$.

Proof of Theorem 3.1. Let us first prove (i). Let z be an arbitrary element of C. From the convex inequality, it follows that

$$\phi(z) \geq \phi(x(t)) + \langle \nabla\phi(x(t)), z - x(t)\rangle, \qquad t \geq 0.$$

In order to use (CGP), this inequality can be rewritten in the following form:

$$\phi(x(t)) - \phi(z) \leq \langle \nabla\phi(x(t)), \dot{x}(t) + x(t) - z\rangle - \langle \nabla\phi(x(t)), \dot{x}(t)\rangle, \qquad \forall t \geq 0. \tag{6}$$

But C is a convex set. Thus,

$$\langle x(t) - \mu\nabla\phi(x(t)) - P_C[x(t) - \mu\nabla\phi(x(t))], z - P_C[x(t) - \mu\nabla\phi(x(t))]\rangle \leq 0.$$

From (CGP), we deduce that

$$\langle x(t) - \mu\nabla\phi(x(t)) - \dot{x}(t) - x(t), z - \dot{x}(t) - x(t)\rangle \leq 0;$$

therefore,

$$\langle \mu\nabla\phi(x(t)) + \dot{x}(t), \dot{x}(t) + x(t) - z\rangle \leq 0. \tag{7}$$

Coming back to the inequality (6), we obtain

$$\phi(x(t)) - \phi(z) \leq -(1/\mu)\langle \dot{x}(t), \dot{x}(t) + x(t) - z\rangle - \langle \nabla\phi(x(t)), \dot{x}(t)\rangle, \qquad t \geq 0.$$

Hence, for $t \geq 0$,

$$(1/\mu)|\dot{x}(t)|^2 + (1/\mu)\langle \dot{x}(t), x(t) - z\rangle + \langle \nabla\phi(x(t)), \dot{x}(t)\rangle + \phi(x(t)) - \phi(z) \leq 0,$$

from which we derive

$$(d/dt)[(1/2\mu)|x(t) - z|^2 + \phi(x(t))] + \phi(x(t)) - \phi(z) + (1/\mu)|\dot{x}(t)|^2 \leq 0. \tag{8}$$

Integrating (8) over $[0, t]$, we obtain

$$\begin{aligned}&(1/2\mu)|x(t) - z|^2 + \phi(x(t)) + \int_0^t [\phi(x(s)) - \phi(z)]\, ds \\ &\leq (1/2\mu)|x_0 - z|^2 + \phi(x_0).\end{aligned} \tag{9}$$

By (iii) of Theorem 2.1, we know that $\phi \circ x$ is a nonincreasing function; thus, (9) gives

$$\begin{aligned}&(1/2\mu)|x(t) - z|^2 + \phi(x(t)) + t[\phi(x(t)) - \phi(z)] \\ &\leq (1/2\mu)|x_0 - z|^2 + \phi(x_0).\end{aligned}$$

Hence,

$$\phi(x(t)) \leq \phi(z) + (1/t)\left[(1/2\mu)|x_0 - z|^2 + \phi(x_0) - \inf_C \phi\right]. \tag{10}$$

To obtain (ii), just notice that (10) is valid for all z in C. As a consequence,

$$\lim_{t\to+\infty} \phi(x(t)) = \inf_C \phi.$$

Assume now that $S \neq \emptyset$. The proof of (iii) is based on the fact that

$$t \to E(x(t), x^*) = (1/2\mu)|x(t) - x^*|^2 + \phi(x(t)) - \phi(x^*)$$

is nonincreasing for all x^* fixed in S. Indeed, (8) gives, for all $z = x^*$ in S,

$$(d/dt)E(x(t), x^*) \leq \inf_C \phi - \phi(x(t)) - (1/\mu)|\dot{x}(t)|^2 \leq 0.$$

Combining the latter result with the fact that ϕ is bounded from below implies that $t \to E(x(t), x^*)$ converges as $t \to +\infty$. Since $\phi(x(t))$ has a limit, it follows that, for all x^* in S, $|x(t) - x^*|$ converges as $t \to +\infty$. Let us now use the Opial lemma: there exist $x^* \in C$ and $t_n \to +\infty$ such that

$$w - \lim_{t_n \to +\infty} x(t_n) = x^*;$$

since ϕ is convex and continuous,

$$\phi(x^*) \leq \liminf_{n \to +\infty} \phi(x(t_n)).$$

Using (ii), it follows that

$$\phi(x^*) \leq \inf_C \phi$$

and since, by the (weak) closedness of C, $x^* \in C$, we obtain

$$x^* \in S$$

and thus

$$w - \lim_{t \to +\infty} x(t) = x_\infty, \qquad \text{with } x_\infty \in S. \qquad \square$$

3.2. Trajectories Starting Outside the Constraint Set. Our purpose in this section, is to study the trajectories of (CGP) under the following hypothesis:

($\mathscr{H}'$) $\quad \phi$ is convex and $x_0 \notin C$.

The difficult point of the following result is to cope with a dynamics which is no more a descent method (see Fig. 1); in other words, the general existence results given in Theorem 2.1 are no longer available.

Theorem 3.2. Assume that ϕ and x_0 satisfy ($\mathscr{H}$) and ($\mathscr{H}'$). Moreover, suppose that $\operatorname{argmin}_C \phi \neq \emptyset$. Then, the following results hold:

(i) The trajectory of (CGP) is defined on $\mathbb{R}_+$. For all $x^* \in \operatorname{argmin}_C \phi$, the function

$$t \geq 0 \to \mathscr{E}(t, x^*) = (1/2)|x(t) - x^*|^2 + \mu[\phi(x(t)) - \phi(x^*) - \langle \nabla\phi(x^*), x(t) - x^* \rangle]$$

is nonincreasing; more precisely,

$$\dot{\mathscr{E}}(t, x^*) = -|\dot{x}(t)|^2 - \mu\langle\nabla\phi(x(t)) - \nabla\phi(x^*), x(t) - x^*\rangle, \quad \forall t \geq 0. \tag{11}$$

(ii) The trajectory converges weakly toward an element x^* in $\operatorname{argmin}_C \phi$ and $\phi(x(t))$ converges to $\inf_C \phi$ as $t \to +\infty$.

Proof. As in Theorem 2.1, let us start by proving that the system is dissipative. Let $[0, T_{\max}[$ be the interval corresponding to the maximal solution of (CGP). For $x^* \in \operatorname{argmin}_C \phi$, the convexity of ϕ implies

$$\langle z - x^*, \nabla\phi(x^*)\rangle \geq 0, \qquad \forall z \in C. \tag{12}$$

On the other hand, the convexity of C and (CGP) yield [see (7)]

$$\langle \mu\nabla\phi(x(t)) + \dot{x}(t), z - \dot{x}(t) - x(t)\rangle \geq 0, \qquad \forall z \in C. \tag{13}$$

Take $z = x^*$ in (13) and $z = x(t) + \dot{x}(t)$ in (12). It follows that, for all t in $[0, T_{\max}[$,

$$(d/dt)((1/2)|x(t) - x^*|^2 + \mu[\phi(x(t)) - \phi(x^*) - \langle\nabla\phi(x^*), x(t) - x^*\rangle])$$
$$\leq -|\dot{x}(t)|^2 - \mu\langle\nabla\phi(x(t)) - \nabla\phi(x^*), x(t) - x^*\rangle,$$

which is precisely (11). The standard arguments evoked in Theorem 2.1 can be applied to obtain that $t \mapsto x(t)$ is defined on $[0, +\infty[$ with $\dot{x} \in L^2(0, +\infty; H)$. Besides, since

$$\phi(x(t)) - \phi(x^*) - \langle\nabla\phi(x^*), x(t) - x^*\rangle \geq 0, \qquad \text{for all } t \geq 0,$$

we can claim moreover that x is bounded, with

$$\lim_{t \to +\infty} \dot{x}(t) = 0.$$

Let us now deal with (ii). First notice that the inequality (8) remains valid; thus,

$$\phi(x(t)) - \phi(z) + (1/\mu)|\dot{x}(t)|^2 + (1/\mu)\langle\dot{x}(t), x(t) - z\rangle$$
$$+ \langle\dot{x}(t), \nabla\phi(x(t))\rangle \leq 0, \qquad \forall z \in C. \tag{14}$$

Since $t \mapsto x(t)$ is bounded and $\lim_{t \to +\infty} \dot{x}(t) = 0$, we infer from (14) that, $\forall z \in C$,

$$\limsup_{t \to +\infty} \phi(x(t)) - \phi(z) \leq 0;$$

thus,

$$\limsup_{t \to +\infty} \phi(x(t)) \leq \inf_C \phi. \tag{15}$$

Let t_n be an increasing sequence such that

$$\phi(x(t_n)) \mapsto \liminf_{t\to+\infty} \phi(x(t)).$$

Since $x(t_n)$ is a bounded sequence, it is weakly relatively compact in H. Therefore, there exists $t_{n_k} \to +\infty$ and x_1 in H such that

$$w - \lim_{k\to+\infty} x(t_{n_k}) = x_1.$$

Noticing that

$$\begin{aligned} w - \lim_{k\to+\infty} x(t_{n_k}) &= w - \lim_{k\to+\infty} \dot{x}(t_{n_k}) + x(t_{n_k}) \\ &= w - \lim_{k\to+\infty} P_C(x(t_{n_k}) - \mu\nabla\phi(x(t_{n_k}))), \end{aligned}$$

we see that x_1 can be obtained as a limit of a sequence in C and thus $x_1 \in C$. Using the weak lower semicontinuity of ϕ, we obtain

$$\liminf_{t\to+\infty} \phi(x(t)) = \lim_{k\to+\infty} \phi(x(t_{n_k})) \geq \phi(x_1) \geq \inf_C \phi,$$

and by (15) it ensues that

$$\lim_{t\to+\infty} \phi(x(t)) = \inf_C \phi.$$

Let us prove the weak convergence of the orbit x. Let x_1 and x_2 be two weak cluster points of x, and denote by $(t_n)_{n\in N}$ and $(\tau_n)_{n\in N}$ some corresponding real-valued subsequences, with $t_n \to \infty$, $\tau_n \to \infty$ as $n \to +\infty$. By direct algebra, we have

$$\begin{aligned} &\mathscr{E}(t, x_1) - \mathscr{E}(t, x_2) \\ &= 2\langle x(t), x_2 - x_1\rangle + |x_1|^2 - |x_2|^2 - 2\mu\langle \nabla\phi(x_1) - \nabla\phi(x_2), x(t)\rangle. \end{aligned} \tag{16}$$

Property (i) implies that $\mathscr{E}(t, x_1) - \mathscr{E}(t, x_2)$ converges as $t \to +\infty$. Replacing t successively by t_n and τ_n and passing to the limit, we obtain the equality

$$\begin{aligned} &-|x_1 - x_2|^2 - 2\mu\langle \nabla\phi(x_1) - \nabla\phi(x_2), x_1\rangle \\ &= |x_1 - x_2|^2 - 2\mu\langle \nabla\phi(x_1) - \nabla\phi(x_2), x_2\rangle. \end{aligned}$$

Equivalently,

$$2|x_1 - x_2|^2 + 2\mu\langle \nabla\phi(x_1) - \nabla\phi(x_2), x_1 - x_2\rangle = 0.$$

Resorting to the monotonicity of $\nabla\phi$, both terms of the previous equality are nonnegative, and thus $x_1 = x_2$. □

4. Convergence of the Trajectories for Quasiconvex Functions

From now on, the initial condition x_0 is supposed to be in C. Let us recall that a function $\phi: H \mapsto \mathbb{R}$ is said to be quasiconvex if its lower level sets are convex. More precisely, for all γ in $\mathbb{R}$, if we set

$$\operatorname{lev}_\gamma \phi := \{x \in H | \phi(x) \leq \gamma\},$$

then the lower level $\operatorname{lev}_\gamma \phi$ is a convex set. In addition, if ϕ is continuously differentiable, then the following property holds for all $x \in H$:

$$\langle \nabla\phi(x), z - x \rangle \leq 0, \qquad \forall z \in \operatorname{lev}_{\phi(x)}. \tag{17}$$

For some studies of dissipative systems with a quasiconvex potential, see for instance Goudou (Ref. 18) or Kiwiel–Murty (Ref. 19).

Theorem 4.1. Assume that ϕ is quasiconvex, satisfies $(\mathscr{H})$, and $\inf_C \phi$ is attained. Then, the solution of (CGP) converges weakly in H. Let x_∞ be the limit point of the trajectory and in addition assume that H is finite-dimensional. Then, x_∞ satisfies the following optimality condition:

$$\nabla\phi(x_\infty) \in -N_C(x_\infty). \tag{18}$$

Proof. If H is finite-dimensional and

$$\lim_{t \to \infty} x(t) = x_\infty,$$

we deduce from Theorem 2.1 that

$$x_\infty - P_C(x_\infty - \mu\nabla\phi(x_\infty)) = 0.$$

The inclusion (18) follows from the formula

$$P_C = (I + N_C)^{-1},$$

where I denotes the identity map of H.

To obtain the weak convergence in the general case, we need only to prove, as for the convex case, that

$$t \mapsto E(x(t), x^*) = (1/2\mu)|x(t) - x^*|^2 + \phi(x(t)) - \phi(x^*)$$

is a Lyapunov function for some well chosen x^*. This fact follows easily by considering the x^* in

$$S_m = \operatorname{lev}_m \phi \cap C, \qquad \text{where } m = \lim_{t \to \infty} \phi(x(t)). \qquad \square$$

5. Asymptotic Control Result

This section proposes an asymptotic control result involving a Tykhonov-like regularization. Consider the following dynamical system;

$$(\mathrm{CGP})_\epsilon \quad \dot{x}(t) + x(t) - P_C[x(t) - \mu\nabla\phi(x(t)) - \epsilon(t)x(t)] = 0,$$
$$x(0) = x_0 \in C,$$

where $\epsilon: \mathbb{R}_+ \to \mathbb{R}_+$ is a nonincreasing function converging to zero.

The use of Tykhonov regularization as a controlling means for differential inclusions has been proposed in Ref. 12 for the steepest descent method. Under reasonable assumptions, it allows us both to select a particular equilibrium and to obtain the strong convergence of the solutions.

Theorem 5.1. Assume that ϕ is convex, satisfies $(\mathscr{H})$, and that $S = \operatorname{argmin}_C \phi \neq \emptyset$. Let $\epsilon: \mathbb{R}_+ \to \mathbb{R}_+$ be a C^1 nonincreasing function such that $\int_0^{+\infty} \epsilon(s)\, ds = +\infty$, $\dot{\epsilon}$ is bounded and converges to zero. Then, $(\mathrm{CGP})_\epsilon$ admits a unique solution on $[0, +\infty)$ that converges strongly toward the element of minimal norm of S. Equivalently, $s - \lim_{t\to\infty} x(t) = p$, where $p = P_S(0)$.

Proof. The arguments concerning the existence and the uniqueness of a solution are very similar to those of Theorem 3.1. We give only the main lines of the proof. As before, we obtain easily that the solution x takes its values in C. From this, like in Theorem 2.1, we deduce that

$$(d/dt)E_\epsilon(t) \leq -|\dot{x}(t)|^2 + (1/2)\dot{\epsilon}(t)|x(t)|^2, \tag{19}$$

where

$$E_\epsilon(t) = \mu\phi(x(t)) + (1/2)\epsilon(t)|x(t)|^2.$$

Adapting former arguments, it follows from $(\mathscr{H})$ and (19) that the solution is defined on $[0, +\infty)$, with velocity in $L^2(0, +\infty; H)$. Moreover, if x is supposed to be bounded, it ensues that $\dot{x}(t) \to 0$ and $\phi(x(t))$ converges as $t \to \infty$. This can be summed up in the following statements:

(a) $\dot{x} \in L^2(0, +\infty; H)$; (20a)

(b) x bounded $\Rightarrow \lim_{t\to+\infty} \dot{x}(t) = 0, \ \lim_{t\to+\infty} \phi(x(t))$ exists. (20b)

Let us focus on the proof of the strong convergence of the trajectory x. The nonautonomous nature of $(\mathrm{CGP})_\epsilon$ gives rise to oscillating trajectories that prevents us from exhibiting a proper Lyapunov functional. However, an acute study of the following function allows us to overcome this difficulty. For all $t \geq 0$, set

$$F(t) = E_\epsilon(t) + (1/2)|x(t) - p|^2. \tag{21}$$

By convexity of C,

$$\langle x - \mu\nabla\phi(x) - \epsilon x - x - \dot{x}, p - x - \dot{x}\rangle \leq 0;$$

thus, after computation,

$$(d/dt)F(t) + |\dot{x}|^2 + \mu\langle\nabla\phi(x), x - p\rangle - (1/2)\dot{\epsilon}|x|^2 + \epsilon\langle x, x - p\rangle \leq 0. \tag{22}$$

The lack of monotonicity of F is due to the term $\langle x, x - p\rangle$; this leads us to consider two cases. Before going further, we need a lemma.

Lemma 5.1. Under the assumptions of Theorem 5.1 and if the solution x is bounded, then all the weak limit points of x are minimizers of $\phi_{|C}$.

Proof of Lemma 5.1. x is bounded; hence, $\dot{x}(t) \to 0$ as $t \to +\infty$ and $\nabla\phi(x(\cdot))$ is bounded. Thus,

$$(d/dt)F(t) = \langle\dot{x}(t), \nabla\phi(x(t))\rangle + \epsilon(t)\langle\dot{x}(t), x(t)\rangle + (1/2)\dot{\epsilon}(t)|x(t)|^2 + \langle\dot{x}(t), x(t) - p\rangle$$

tends to zero, which combined with inequality (22) gives

$$\limsup_{t \to +\infty} \mu\langle\nabla\phi(x(t)), x(t) - p\rangle \leq 0. \tag{23}$$

The above number being nonnegative as well, it follows that

$$\lim_{t \to +\infty} \langle\nabla\phi(x(t)), x(t) - p\rangle = 0.$$

Let x^* be a weak limit point of x relatively to an increasing sequence of positive real numbers τ_n. Using the convexity of ϕ, we obtain

$$\phi(p) \geq \phi(x(\tau_n)) + \langle\nabla\phi(x(\tau_n)), p - x(\tau_n)\rangle.$$

Passing to the inf limit and according to the lower semicontinuity of ϕ, one obtains

$$\phi(p) \geq \phi(x^*),$$

which means exactly that $x^* \in S$. □

Case 1. In this part, we assume that there exists $t_0 \in \mathbb{R}_+$ such that

$$\langle x(t), x(t) - p\rangle \geq 0, \qquad \forall t \geq t_0.$$

For simplicity, we assume that $t_0 = 0$. F becomes in this case a nonincreasing function, and since E_ϵ is bounded, so is F [see (21)]. Hence, applying (19), it follows that the functions F and $|x-p|$ have a limit as $t \to +\infty$.

Let us argue by contradiction and assume that

$$|x(t)-p| \to l > 0, \qquad \text{as } t \to +\infty.$$

First, we notice that

$$\liminf_{t\to+\infty} \langle p, x(t)-p\rangle \geq 0.$$

Indeed, if t_n is a sequence of real numbers realizing the inf limit, by boundedness of x and Lemma 5.1 there exists a subsequence t_{n_k} of t_n such that

$$x(t_{n_k}) \to x^* \in S.$$

Therefore,

$$\liminf_{t\to+\infty} \langle p, x(t)-p\rangle = \langle p, x^*-p\rangle = -\langle 0-P_S(0), x^*-P_S(0)\rangle \geq 0.$$

Since

$$\langle x, x-p\rangle = |x-p|^2 + \langle p, x-p\rangle,$$

we can assume that there exists some $T>0$ such that $t \geq T$ implies

$$\langle x(t), x(t)-p\rangle \geq 1/2.$$

From (22), it ensues that

$$(d/dt)F(t) + \epsilon(t)\langle x(t), x(t)-p\rangle \leq 0, \qquad \text{for all } t \text{ in } \mathbb{R}_+.$$

Integrating over (T, t), $t \geq T$, the above inequality becomes

$$F(t) - F(T) + (1/2)\int_T^t \epsilon \leq 0.$$

But we know that F converges, whereas $\int_{\mathbb{R}_+} \epsilon = +\infty$. This yields a contradiction.

Case 2. In this part, the function $\langle x, x-p\rangle$ is allowed to reach real negative values as time elapses, which leads us naturally to introduce the set

$$B = \{y \in H | \langle y, y-p\rangle \leq 0\}.$$

Observing that

$$\langle y, y-p\rangle = \langle y-p/2+p/2, y-p/2-p/2\rangle = |y-p/2|^2 - |p/2|^2,$$

we see that B is the closed ball of radius $|p/2|$ centered at the point $p/2$. Set

$$I = \{t \in \mathbb{R}_+, x(t) \in B\}, \qquad J = \{t \in \mathbb{R}_+, x(t) \notin B\}$$

and assume that I is unbounded. I, J are respectively closed and open in $\mathbb{R}$; hence, there exists a nondecreasing sequence of real numbers t_k such that

$$I = \bigcup_{k \in N} [t_{2k}, t_{2k+1}], \qquad J = \bigcup_{k \in N}]t_{2k+1}, t_{2k+2}[.$$

Note that we have assumed implicitly that $x_0 \in B$, which is not restrictive in our study.

In order to tackle the most difficult problem first, we assume J to be unbounded and we start by proving that x is bounded. If $t \in I$, then by definition of I, $x(t) \in B$, which implies that $x_{|I}$ is bounded. For t in J, there exists $k(t) := k$ in N such that t belongs to $]t_{2k+1}, t_{2k+2}[$. Coming back to (22), we deduce that $F_{|]t_{2k+1}, t_{2k+2}[}$; thus, the functions $F_{|[t_{2k+1}, t_{2k+2}]}$ are nonincreasing. Since

$$F(t) \leq F(t_{2k+1}) \quad \text{and} \quad x(t_{2k+1}) \in B,$$

it follows that

$$\begin{aligned} F(t) &\leq F(t_{2k+1}) \\ &= E_\epsilon(t_{2k+1}) + (1/2)|x(t_{2k+1}) - p|^2 \\ &\leq E_\epsilon(0) + \max_{y \in B} (1/2)|y - p|^2. \end{aligned}$$

But since E_ϵ is bounded, so is $x_{|J}$.

Let us focus on the limit points of $x_{|I}(t)$ when $t \to +\infty, t \in I$. Observe that

$$B \cap S = \{p\};$$

thus, by Lemma 5.1, $x_{|I}(t)$ has to converge weakly to its unique limit point p as $t \to +\infty, t \in I$. To obtain strong convergence, one has to notice that $y \in B$ implies $|y| \leq |p|$; thus,

$$\limsup_{t \to +\infty} |x_{|I}(t)|^2 \leq |p|^2.$$

Now, we prove an equivalent result for $x_{|J}$. If $t \in J$, let $k(t)$ be as above and set $\tau_t = t_{2k(t)+1}$. Observe that τ_t belongs to I; therefore, $x(\tau_t) \to p$ strongly as $t \to +\infty, t \in J$. Besides, we know that $F(t) \leq F(\tau_t)$; thus,

$$\begin{aligned} &\mu\phi(x(t)) + (1/2)\epsilon(t)|x(t)|^2 + (1/2)|x(t) - p|^2 \\ &\leq \mu\phi(x(\tau_t)) + (1/2)\epsilon(\tau_t)|x(\tau_t)|^2 + (1/2)|x(\tau_t) - p|^2. \end{aligned}$$

Passing to the sup limit, (20) yields

$$\mu \lim_{t\to+\infty,\, t\in J} \phi(x(t)) + \limsup_{t\to+\infty,\, t\in J} (1/2)|x(t)-p|^2$$
$$\leq \mu \lim_{t\to+\infty,\, t\in J} \phi(x(\tau_t)). \tag{24}$$

Finally, we deduce from (24) that $x(t), t\in J$, converges strongly to p as $t\to+\infty$.

The case for which J is bounded can be solved with similar ideas. □

6. Gradient Projection Method for a Continuous Convex Criterion

In the sequel, H is supposed to be finite dimensional and ϕ to be convex continuous on H. Classically, the subdifferential of ϕ at $y_0 \in H$ is the convex subset $\partial\phi(y_0)$ of H characterized by the following property:

$$z\in\partial\phi(y_0) \Leftrightarrow \forall y\in H,\ \phi(y)\geq\phi(y_0)+\langle z, y-y_0\rangle. \tag{25}$$

We propose to establish the existence of an absolutely continuous solution to the following differential inclusion:

$$(\mathrm{CGP})_g \quad \dot{x}(t)+x(t)-P_C[x(t)-\mu\partial\phi(x(t))]\ni 0, \qquad \text{a.e. in } (0,+\infty),$$

with $x(0)=x_0$. Defining $A\colon H^2\to H$ by

$$A_x(v)=x-P_C(x-\mu v), \qquad \text{for all } (x,v) \text{ in } H^2,$$

it is easy to see that, if x is fixed, then $A_x\colon H\to H$ is a maximal monotone operator. Therefore, $(\mathrm{CGP})_g$ can be rewritten in the following form:

$$A^{-1}_{x(t)}(\dot{x}(t))+\partial\phi(x(t))\ni 0, \qquad \text{a.e. in } (0,+\infty).$$

This formulation is akin to the doubly nonlinear problems arising in PDEs; see Colli–Visintin (Ref. 20) and the references therein. As in Ref. 20, where $B\dot{x}(t)+\partial\phi(x(t))\ni 0$ was considered, with B maximal monotone, we have not been able to prove the uniqueness of the solution and, as far as we know, this is still an open question.

To obtain a solution of $(\mathrm{CGP})_g$, let us define the following approximated problems for any positive λ:

$$(\mathrm{CGP})_\lambda \quad \dot{x}_\lambda(t)+x_\lambda(t)-P_C[x_\lambda(t)-\mu\nabla\phi_\lambda(x_\lambda(t))]=0,$$
$$x_\lambda(0)=x_{0\lambda}\in C, \tag{26}$$

where $x_{0\lambda}$ is a sequence in C such that

$$\lim_{\lambda\to 0} x_{0\lambda}=x_0$$

and ϕ_λ is the Moreau–Yosida approximation of ϕ. The general results concerning the Moreau–Yosida approximate can be found in Ref. 1 or in Rockafellar-Wets (Ref. 21). Let us recall that, for any positive λ, ϕ_λ is defined as the episum of ϕ and the quadratic kernel $y \in H \to (1/2\lambda)|y|^2$; that is,

$$\phi_\lambda(y) = \inf_{z \in H} \{\phi(z) + (1/2\lambda)|y - z|^2\}, \qquad \forall y \in H.$$

ϕ_λ is a C^1 function from H into $\mathbb{R}$, whose gradient $\nabla\phi_\lambda$ is Lipschitz continuous. Moreover, for any y in H,

$$\sup_{\lambda > 0} \phi_\lambda(y) = \lim_{\lambda \to 0} \phi_\lambda(y) = \phi(y). \tag{27}$$

Set

$$\partial\phi^o(y) = \inf_{z \in \partial\phi(y)} |z|,$$

where

$$y \in \operatorname{dom} \partial\phi.$$

Then,

$$|\nabla\phi_\lambda(y)| \leq \partial\phi^o(y). \tag{28}$$

Let us state the central result of this section.

Theorem 6.1. Assume that ϕ is convex, continuous on H, and bounded from below. Then, there exists an absolutely continuous solution $t \in [0, +\infty) \to x(t) \in H$ satisfying $(\mathrm{CGP})_g$. Moreover, the following properties hold:

(i) x takes its values in C.

(ii) The function $t \in [0, +\infty) \to \phi(x(t))$ is absolutely continuous with

$$\mu(d/dt)\phi(x(t)) \leq -|\dot{x}(t)|^2, \qquad \text{a.e. in } (0, +\infty),$$

and therefore $\dot{x} \in L^2(0, +\infty; H)$.

(iii) $\lim_{t \to +\infty} \phi(x(t)) = \inf_C \phi$.

(iv) If $\operatorname{argmin}_C \phi \neq \emptyset$, then $x(t)$ converges to some minimizer of ϕ over C.

Besides, there exists a nonnegative constant M such that

$$\left[\phi(x(t)) - \inf_C \phi\right] \leq M/t + 1, \qquad \forall t \geq 0.$$

Fix $T > 0$ and denote by $\mathscr{D}(0, T)$ the set of $\mathscr{C}^\infty$ real functions with compact support in $]0, T[$. The following two lemmas are classical and very useful; their proofs may be found respectively in Rockafellar (Ref. 22, Theorem 24.7, page 237) and in Ref. 1, Lemma 3.3, page 73.

Lemma 6.1. If ϕ is a continuous convex function on a finite-dimensional space H, then $\partial\phi$ is bounded on bounded sets. More precisely, if B is bounded in H, then there exists some $M > 0$ such that

$$|z| \leq M, \qquad \forall y \in B, \forall z \in \partial\phi(y). \tag{29}$$

Lemma 6.2. Let $t \in [0, +\infty) \to u(t) \in H$ be an absolutely continuous function, and assume that $t \in [0, +\infty) \to \phi(u(t))$ is also absolutely continuous. Let D be the subset of $\mathbb{R}_+$ on which $t \to \phi(u(t))$ and $t \to u(t)$ are derivable. Then, dt being the Lebesgue measure on $\mathbb{R}$, $dt(\mathbb{R}_+ \backslash D) = 0$ and

$$(d/dt)\phi(u(t)) = \langle \dot{u}(t), z \rangle, \qquad \forall t \in D, \forall z \in \partial\phi(u(t)).$$

Proof of Theorem 6.1. First, some uniform estimations relying on the solutions of $(\mathrm{CGP})_\lambda$ are established on a bounded time interval $[0, T]$. Then, arguing by compacity, we pass to the limit to obtain that a solution of $(\mathrm{CGP})_g$ on $[0, T]$. When no confusion can occur, the time variable t will be omitted. For the sake of simplicity, all subsequences of $x_\lambda, \dot{x}_\lambda \ldots$ are still denoted $x_\lambda, \dot{x}_\lambda \ldots$

(a) Estimations. Owing to Theorem 2.1, the solution x_λ of the approximation scheme $(\mathrm{CGP})_\lambda$ satisfies, for all t in $[0, T]$,

$$\phi_\lambda(x_\lambda(t)) - \phi_\lambda(x_{0\lambda}) + \mu \int_0^t |\dot{x}_\lambda|^2 \leq 0. \tag{30}$$

By (27) and the fact that ϕ is bounded from below, we obtain that $\dot{x}_\lambda$ is a bounded sequence in $L^2(0, T; H)$. Thus, one can extract from $\dot{x}_\lambda$ a subsequence that converges weakly in $L^2(0, T; H)$ to some function $v: (0, T) \to H$.

(b) From the formula

$$x_\lambda(t) - x_\lambda(\tau) = \int_\tau^t \dot{x}_\lambda,$$

we deduce that, for all $t \geq \tau$ in $[0, T]$,

$$|x_\lambda(t) - x_\lambda(\tau)| \leq \sqrt{t - \tau} \sqrt{\int_\tau^t |\dot{x}_\lambda|^2}.$$

It ensues that x_λ is an equicontinuous bounded sequence in $C([0, T], H)$ equipped with the supremum norm; therefore, the Ascoli theorem gives the existence of a cluster point $x \in C([0, T], H)$ to the sequence x_λ. Moreover, by Theorem 2.1, $x_\lambda([0, T]) \subset C$; therefore, C being closed,

$$x(t) \in C, \qquad \forall t \in [0, T]. \tag{31}$$

(c) The preceding two points yield the existence of a subsequence x_λ such that

$$x_\lambda \to x, \qquad \text{in } C([0, T], H), \tag{32}$$

$$\dot{x}_\lambda \to \dot{x}, \qquad \text{in } w - L^2(0, T; H), \tag{33}$$

where x belongs to $W^{1,2}(0, T; H)$.

(d) By (28), we have that, for all t in $[0, T]$,

$$|\nabla \phi_\lambda(x_\lambda(t))| \leq |\partial \phi^o(x(t))|.$$

Now, Lemma 6.1 and the continuity property of x imply that $\nabla \phi_\lambda(x_\lambda(\cdot))$ is a bounded sequence in $L^\infty(0, T; H)$. In particular, it is relatively compact in $w - L^2(0, T; H)$, with at least some cluster point, say $g \in L^2(0, T; H)$. Therefore, after extraction, we have

$$\nabla \phi_\lambda(x_\lambda(\cdot)) \to g, \qquad \text{in } w - L^2(0, T; H). \tag{34}$$

(e) Let us study the sequence $\phi_\lambda(x_\lambda(\cdot))$. For all $t \geq \tau$ in $[0, T]$,

$$\begin{aligned} |\phi_\lambda(x_\lambda(t)) - \phi_\lambda(x_\lambda(\tau))| &\leq \int_\tau^t |< \nabla \phi_\lambda(x_\lambda), \dot{x}_\lambda >| ds \\ &\leq M \sqrt{t - \tau} \sqrt{\int_\tau^t |\dot{x}_\lambda|^2}, \end{aligned} \tag{35}$$

where M is a bound of $|\nabla \phi_\lambda(x_\lambda(\cdot))|$ on $[0, T]$. By the Ascoli theorem, this shows that a subsequence of $\phi_\lambda(x_\lambda(\cdot))$ converges uniformly on $[0, T]$ to an element ψ of $C([0, T], H)$.

Let us prove that, for all $t \in [0, T]$,

$$\psi(t) = \phi(x(t)).$$

Take t in $[0, T]$. If $\lambda_0 > \lambda > 0$, it follows from (27) that

$$\phi_{\lambda_0}(x_\lambda(t)) \leq \phi_\lambda(x_\lambda(t));$$

letting $\lambda \to 0$, (32) yields

$$\phi_{\lambda_0}(x(t)) \leq \psi(t).$$

Using (27) again, it ensues that

$$\phi(x(t)) \leq \psi(t).$$

Let $M > 0$ be a bound of $\nabla\phi_\lambda(x_\lambda(\cdot))$. The convex inequality gives

$$\phi_\lambda(x(t)) \geq \phi_\lambda(x_\lambda(t)) - M|x_\lambda(t) - x(t)|. \tag{36}$$

From (27) and (32), we deduce finally that

$$\phi(x(t)) \geq \psi(t);$$

thus,

$$\phi_\lambda(x_\lambda(t)) \to \phi(x(t)), \qquad \text{in } C([0, T], H). \tag{37}$$

Besides, $\dot{x}_\lambda$ and $\nabla\phi_\lambda(x_\lambda(\cdot))$ being respectively bounded sequences in $L^2(0, T; H)$ and $L^\infty(0, T; H)$, it follows that $(d/dt)\phi_\lambda(x_\lambda(\cdot))$ is bounded in $L^2(0, T; H)$. This implies that the first derivative in the sense of distributions of $t \in (0, T) \to \phi(x(t))$ is in $L^2(0, T; H)$ and in particular that $\phi(x(\cdot))$ is absolutely continuous.

(f) Let us identify g. Fix $\theta \geq 0$ in $\mathscr{D}(0, T)$. Integrating the convex inequality, we obtain that, for all $y \in H$,

$$\int_0^T \theta(t)[\phi_\lambda(y) - \phi_\lambda(x_\lambda(t)) - \langle\nabla\phi_\lambda(x_\lambda(t)), y - x_\lambda(t)\rangle]\, dt \geq 0. \tag{38}$$

From (32), (34), (27), and (37), we obtain

$$\int_0^T \theta(t)[\phi(y) - \phi(x(t)) - \langle g(t), y - x(t)\rangle]\, dt \geq 0.$$

The latter being true for all $\theta \geq 0$ in $\mathscr{D}(0, T)$, it follows that

$$\phi(y) \geq \phi(x(t)) + \langle g(t), y - x(t)\rangle, \qquad \text{a.e. in } [0, T].$$

By definition of the subdifferential, this implies that

$$g(t) \in \partial\phi(x(t)), \qquad \text{a.e. in } [0, T]. \tag{39}$$

(g) Passing to the Limit. The (sub)sequence x_λ verifies

$$\dot{x}_\lambda(t) + x_\lambda(t) - P_C[x_\lambda(t) - \mu\nabla\phi_\lambda(x_\lambda(t))] = 0, \qquad \forall t \in [0, T],$$

$$x_\lambda(0) = x_{0\lambda},$$

which is equivalent to

$$\langle -\mu\nabla\phi_\lambda(x(t)) - \dot{x}_\lambda(t), \xi - x_\lambda(t) - \dot{x}_\lambda(t)\rangle \leq 0,$$

$$\forall t \in [0, T], \qquad \forall \xi \in C, \qquad x_\lambda(0) = x_{0\lambda}.$$

Therefore, we have

$$\int_0^T [\langle \mu\nabla\phi_\lambda(x_\lambda) + \dot{x}_\lambda, x_\lambda - \xi\rangle + \mu(d/dt)\phi_\lambda(x_\lambda) + |\dot{x}_\lambda|^2]\theta \leq 0,$$

$$\forall \theta \geq 0 \in \mathscr{D}(0, T), \qquad \forall \xi \in C,$$

with $x_\lambda(0) = x_{0\lambda}$. Now, in order to take the inf limit of each term in the previous inequality, fix $\theta \geq 0$ and ξ in C. By (weak) lower semicontinuity of the seminorm $\int_0^T |\cdot|^2\theta(s)\,ds$ on $L^2(0, T; H)$ and (33), we have

$$\int_0^T |\dot{x}|^2\theta \leq \liminf_{\lambda\to 0} \int_0^T |\dot{x}_\lambda|^2\theta.$$

By combining (32), (33), and (34), we obtain also that

$$\lim_{\lambda\to 0} \int_0^T \theta\langle \mu\nabla\phi_\lambda(x_\lambda) + \dot{x}_\lambda, x_\lambda - \xi\rangle = \int_0^T \theta\langle \mu g + \dot{x}, x - \xi\rangle.$$

From Lemma 6.2 and (37), we deduce that

$$\lim_{\lambda\to 0} \int_0^T \theta(s)(d/dt)\phi_\lambda(x_\lambda(s))\,ds = \int_0^T \theta(s)(d/dt)\phi(x(s))\,ds$$

$$= \int_0^T \theta\langle \dot{x}, g\rangle.$$

Combining the last three limits yields

$$\int_0^T \langle \mu g + \dot{x}, x - \xi\rangle\theta + \int_0^T \mu\langle g, \dot{x}\rangle + \int_0^T |\dot{x}|^2\theta \leq 0, \qquad x(0) = x_0,$$

for all $\theta \geq 0 \in \mathscr{D}(0, T)$, $\xi \in C$; thus, after rearranging terms, we obtain

$$\int_0^T \theta\langle x - \mu g - x - \dot{x}, \xi - x - \dot{x}\rangle \leq 0,$$

$$x(0) = x_0, \qquad \forall \theta \in \mathscr{D}(0, T), \qquad \forall \xi \in C. \tag{40}$$

In order to use the variational characterization of $P_C(x(\cdot) - \mu g(\cdot))$ in (40), let us prove that

$$\dot{x}(t) + x(t) \in C, \qquad \text{a.e. on } [0, T].$$

Consider the following subset of $L^2(0, T; H)$;

$$\mathscr{C} = \{ f \in L^2(0, T; H) \,|\, f(t) \in C, \qquad \text{a.e. on } (0, T).$$

Clearly, $\mathscr{C}$ is closed in $L^2(0, T; H)$ for the strong topology; since $\mathscr{C}$ is convex, it is also closed for the weak topology. By $(\mathrm{CGP})_\lambda$, we have

$$\dot{x}_\lambda + x_\lambda \in \mathscr{C},$$

whence from (32), (33), and the weak closedness property of $\mathscr{C}$, it follows that

$$\dot{x} + x \in \mathscr{C}.$$

Using (40), we obtain that

$$\dot{x}(t) + x(t) - P_C[x(t) - \mu g(t)] = 0, \qquad \text{a.e. on } (0, T);$$

thus, by (39), it follows that $(\mathrm{CGP})_g$ is satisfied on $[0, T)$. To obtain a solution of $(\mathrm{CGP})_g$ defined on $[0, +\infty[$, let us observe that (27) and (30) imply that the sequence $\dot{x}_\lambda$ is actually bounded in $L^2(0, +\infty; H)$. Combining this fact with those obtained above, gives the existence of a global solution x satisfying the announced properties. The viability property (i) of x is guaranteed by (31). The proofs of (ii), (iii), and (iv) rely on the absolute continuity of x and $\phi(x(\cdot))$. By use of Lemma 6.2, this allows us to reproduce the arguments of Theorems 2.1 and 3.1 with nearly no change. $\square$

Remark 6.1

(a) Denote by ri C the relative interior of C. By adapting the argument of Theorem 2.1(ii), it follows easily that

$$x_0 \in \mathrm{ri}\, C \Rightarrow x(t) \in \mathrm{ri}\, C, \qquad \forall t \geq 0.$$

In other words $(\mathrm{CGP})_g$ is an interior method as soon as the initial condition is strictly feasible.

(b) Implementation of the Method. Given some sequences $\mu_k, \Delta t_k > 0$, an explicit discretization of $(\mathrm{CGP})_g$ gives

$$(x_{k+1} - x_k)/\Delta t_k + x_k - P_C[x_k - \mu_k \partial\phi(x_k)] \ni 0, \qquad k \in N,$$

which can be reformulated as

$$x_{k+1} \in (1 - \Delta t_k)\, x_k + \Delta t_k P_C[x_k - \mu_k \partial\phi(x_k)], \qquad k \in N. \tag{41}$$

This approach for solving approximatively dynamical systems is well known; of course, for $\mu_k = \mu$ and $\Delta t_k = 1$, the usual gradient-projected method (3) is recovered. By the above remark (a), we know that $(\mathrm{CGP})_g$ is an interior method whenever x_0 belongs to ri C; this suggests that a good discrete approximation should also enjoy this property. Very simple examples show that it is not the case of (3); however, if we assume that the steptime parameters of (41) satisfy $\Delta t_k < 1$, an easy induction implies that the sequences x_k, $k \in N$, complying with (41) also verify

$$x_0 \in \mathrm{ri}\, C \Rightarrow x_k \in \mathrm{ri}\, C, \qquad \forall k \in N.$$

Besides, much like as in convex feasibility problems (see Ref. 23 and references therein), the form of (41) suggests to interpret Δt_k, $k \in N$, as a sequence of relaxation parameters and to study (41) within that perspective. Such a study is out of the scope of the present paper, but it is certainly a matter for future research.

References

1. Brézis, H., *Opérateurs Maximaux Monotones*, North-Holland, Amsterdam, Holland, 1973.
2. Bruck, R. E., *Asymptotic Convergence of Nonlinear Contraction Semigroups in Hilbert Space*, Journal of Functional Analysis, Vol. 18, pp. 15–26, 1974.
3. Polyak, B. T., *Introduction to Optimization*, Optimization Software, New York, NY, 1987.
4. MacCormick, G. P, and Tapia, R. A, *The Gradient Projection Method under Mild Differentiability Conditions*, SIAM Journal on Control and Optimization, Vol. 10, pp. 93–98, 1972.
5. Martinet, B., *Perturbation des Méthodes d'Optimisation*: *Applications*, RAIRO Analyse Numérique, Vol. 12, pp. 153–171, 1978.
6. Phelps, R. R., *The Gradient Projection Method Using Curry's Steplength*, SIAM Journal on Control and Optimization, Vol. 24, pp. 692–699, 1986.
7. Alber, Y. I., Iusem, A. N., and Solodov, M. V., *On the Projected Subgradient Method for Nonsmooth Convex Optimization in a Hilbert Space*, Mathematical Programming, Vol. 81, pp. 23–25, 1998.
8. Antipin, A. S., *Minimization of Convex Functions on Convex Sets by Means of Differential Equations*, Differential Equations, Vol. 30, pp. 1365–1375, 1994.
9. Attouch, H., and Alvarez, F., *The Heavy Ball with Friction Dynamical System for Convex Constrained Minimization Problems*, Optimization (Namur, 1998); Lecture Notes in Economic and Mathematical Systems, Springer, Berlin, Germany, Vol. 481, pp. 23–35, 2000.
10. Baillon, J. B., *Un Exemple Concernant le Comportement Asymptotique de la Solution du Problème $du/dt + \partial\phi(u) \ni 0$*, Journal of Functional Analysis, Vol. 28, pp. 369–376, 1978.
11. Opial, Z., *Weak Convergence of the Sequence of Successive Approximations for Nonexpansive Mappings*, Bulletin of the American Mathematical Society, Vol. 73, pp. 591–597, 1967.
12. Attouch, H., and Cominetti, R., *A Dynamical Approach to Convex Minimization Coupling Approximation with the Steepest Descent Method*, Journal of Differential Equations, Vol. 128, pp. 519–540, 1996.

13. Attouch, H., and Czarnecki, M. O., *Asymptotic Control and Stabilization of Nonlinear Oscillators with Nonisolated Equilibria*, Journal of Differential Equations (to appear).
14. Aubin, J. P., and Cellina, A., *Differential Inclusions*, Springer Verlag, New York, NY, 1984.
15. Schatzman, M., *A Class of Nonlinear Differential Equations of Second-Order in Time*, Nonlinear Analysis, Vol. 2, pp. 355–373, 1978.
16. Lemaire, B., *An Asymptotical Variational Principle Associated with the Steepest Descent Method for a Convex Function*, Journal of Convex Analysis, Vol. 3, pp. 63–70, 1996.
17. Alvarez, F., *On the Minimizing Property of a Second-Order Dissipative System in Hilbert Space*, SIAM Journal on Control and Optimization, Vol. 38, pp. 1102–1119, 2000.
18. Goudou, X., *The Heavy Ball with Friction System for a Smooth Quasiconvex Function*, Working Paper, Université Montpellier II, Montpellier, France.
19. Kiwiel, K. C., and Murty, K., *Convergence of the Steepest Descent Method for Minimizing Quasiconvex Functions*, Journal of Optimization Theory and Applications, Vol. 89, pp. 221–226, 1996.
20. Colli, P., and Visintin, A., *On a Class of Doubly Nonlinear Evolution Problems*, Communications in Partial Differential Equations, Vol. 15, pp. 737–756, 1990.
21. Rockafellar, R. T., and Wets, R., *Variational Analysis*, Springer, New York, NY, 1998.
22. Rockafellar, R. T., *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.
23. Censor, Y., Eggermont, P. P. B., and Gordon, D., *Strong Underrelaxation in Kaczmarz's Method for Inconsistent Systems*, Numerische Mathematik, Vol. 41, pp. 83–92, 1983.