SIAM J. CONTROL OPTIM.
Vol. 38, No. 4, pp. 1102–1119

ON THE MINIMIZING PROPERTY OF A SECOND ORDER DISSIPATIVE SYSTEM IN HILBERT SPACES*

FELIPE ALVAREZ†

Abstract. We study the asymptotic behavior at infinity of solutions of a second order evolution equation with linear damping and convex potential. The differential system is defined in a real Hilbert space. It is proved that if the potential is bounded from below, then the solution trajectories are minimizing for it and converge weakly towards a minimizer of Φ if one exists; this convergence is strong when Φ is even or when the optimal set has a nonempty interior. We introduce a second order proximal-like iterative algorithm for the minimization of a convex function. It is defined by an implicit discretization of the continuous evolution problem and is valid for any closed proper convex function. We find conditions on some parameters of the algorithm in order to have a convergence result similar to the continuous case.

Key words. dissipative system, linear damping, asymptotic behavior, weak convergence, convexity, implicit discretization, iterative-variational algorithm

AMS subject classifications. 34G20, 34A12, 34D05, 90C25

PII. S0363012998335802

1. Introduction.

Consider the following differential system defined in a real Hilbert space H:

$$u'' + \gamma u' + \nabla\Phi(u) = 0, \tag{1.1}$$

where $\gamma > 0$ and $\Phi : H \to \mathbb{R}$ is differentiable. It is customary to call this equation *nonlinear oscillator with damping*. Here, the damping or friction has a linear dependence on the velocity. This is a particular case of the so-called dissipative systems. In fact, given u solution of (1.1) define $E(t) := \frac{1}{2}|u'|^2 + \Phi(u)$; it is direct to check that $E' = -\gamma|u'|^2$. Thus, the *energy* of the system is dissipated as t increases. Although (1.1) appears in various contexts with different physical interpretations, the motivation for this work comes from the dynamical approach to optimization problems.

Roughly speaking, any iterative algorithm generating a sequence $\{x_k\}_{k\in\mathbb{N}}$ may be considered as a discrete dynamical system. If it is possible to find a continuous version for the discrete procedure, one expects that the properties of the corresponding continuous dynamical system are close to those of the discrete one. This occurs, for instance, for the now classical proximal method for convex minimization: given $x_0 \in H$, solve the iterative scheme

$$\frac{x_{k+1} - x_k}{\lambda_k} + \partial f(x_{k+1}) \ni 0, \tag{Prox}$$

where $\lambda_k > 0$, $f : H \to \mathbb{R} \cup \{\infty\}$ is a closed proper convex function and ∂f denotes the usual subdifferential in convex analysis. (Prox) is an implicit discretization for the steepest descent method, which consists of solving the following differential inclusion:

$$x' + \partial f(x) \ni 0. \tag{SD}$$

*Received by the editors April 20, 1998; accepted for publication (in revised form) September 17, 1999; published electronically April 4, 2000. This work was partially supported by FONDECYT grants 1961131 and 1990884.

http://www.siam.org/journals/sicon/38-4/33580.html

†Departamento de Ingeniería Matemática, Universidad de Chile, Beauchef 850, Santiago, Chile (falvarez@dim.uchile.cl).

Under suitable conditions, both the trajectory $\{x(t) : t \to \infty\}$ defined by (SD) and the sequence $\{x_k\}$ generated by (Prox) converge toward a particular minimizer of f (see [5, 6, 7] for (SD) and [18] for (Prox); see also [12] for a survey on these and new results). The dynamical approach to iterative methods in optimization has many advantages. It provides a deep insight into the expected behavior of the method, and sometimes the techniques used in the continuous case can be adapted to obtain results for the discrete algorithm. On the other hand, a continuous dynamical system satisfying nice properties may suggest new iterative methods.

This viewpoint has motivated increasing attention in recent years; see, e.g., [1, 2, 3, 4, 8, 13, 14]. In [3], Attouch, Goudou, and Redont deal with nonconvex functions that have, a priori, many local minima. The idea is to exploit the dynamics defined by (1.1) to explore critical points of Φ (i.e., solutions of $\nabla\Phi(x) = 0$). If Φ is coercive (bounded level sets) and of class C^1 with a locally Lipschitz gradient, then it is possible to prove that for any u solution of (1.1) we have $\nabla\Phi(u(t)) \to 0$ as $t \to \infty$. The convergence of the trajectory $\{u(t) : t \to \infty\}$ is a more delicate problem. When Φ is coercive, an obvious sufficient condition for the convergence of the trajectory is that the critical points, also known as equilibrium points, are isolated. Certainly, this is not necessary. In one dimension ($H = \mathbb{R}$) and without additional conditions, the solution always converges toward an equilibrium (see, e.g., [10]). The proof relies on topological arguments that are not generalizable to higher dimensions. Indeed, this is no longer true even in two dimensions: it is possible to construct a coercive C^1 function defined on $\mathbb{R}^2$ whose gradient is locally Lipschitz and for which at least one solution of (1.1) does not converge as $t \to \infty$ (see [3]). Thus, a natural question is to find general conditions under which the trajectory converges in the degenerate case, that is, when the set of equilibrium points of Φ contains a nontrivial connected component. A positive result in this direction has recently been given by Haraux and Jendoubi [11], where convergence to an equilibrium is established when Φ is *analytic*. However, this assumption is very restrictive from the optimization point of view.

Motivated by the previous considerations, in this work we focus our attention on the asymptotic behavior as $t \to \infty$ of the solutions of (1.1) when Φ is assumed to be convex. The paper is organized as follows. In section 2 we prove that if Φ is convex and bounded from below, then the trajectory $\{u(t) : t \to \infty\}$ is minimizing for Φ. If the infimum of Φ on H is attained, then $u(t)$ converges weakly towards a minimizer of Φ. The convergence is strong when Φ is even or when the optimal set has a nonempty interior. In section 2.2 we give a localization result for the limit point, analogous to the corresponding result for the steepest descent method [13]. In section 2.4 we generalize the convergence result to cover the equation $u'' + \Gamma u' + \nabla\Phi(u) = 0$, where $\Gamma : H \to H$ is a bounded self-adjoint linear operator which we assume to be elliptic: there is $\gamma > 0$ such that for any $x \in H$, $\langle \Gamma x, x\rangle \geq \gamma|x|^2$. We refer to this equation as nonlinear oscillator with *anisotropic damping*. This equation appears to be useful to diminish oscillations or even eliminate them, and also to accelerate the convergence of the trajectory. In section 2.3 we give an heuristic motivation of the above mentioned facts, which is based on an analysis of a quadratic function. Still under the convexity condition on Φ, section 3 deals with the discretization of (1.1). Here, we consider the *implicit* scheme

$$\frac{u_{k+1} - 2u_k + u_{k-1}}{h^2} + \gamma\frac{u_{k+1} - u_k}{h} + \nabla\Phi(u_{k+1}) = 0,$$

where $h > 0$. Since Φ is convex, the latter is equivalent to the following variational problem:

$$u_{k+1} = \text{argmin}\left\{\Phi(x) + \frac{1+\gamma h}{2h^2}|x - z_k|^2 : x \in H\right\},$$

where $z_k = u_k + \frac{1}{1+\gamma h}(u_k - u_{k-1})$. This procedure does not require Φ to be differentiable and allows us to introduce the following more general iterative-variational algorithm:

$$\frac{1}{\lambda_k}(u_{k+1} - (1+\alpha_k)u_k + \alpha_k u_{k-1}) + \partial_{\epsilon_k} f(u_{k+1}) \ni 0, \tag{1.2}$$

where $\epsilon_k, \lambda_k > 0$, $\alpha_k \in [0, 1[$, $f : H \to \mathbb{R} \cup \{\infty\}$ is a closed proper convex function and $\partial_\epsilon f$ is the ϵ-approximate subdifferential in convex analysis. We call (1.2) the *inertial proximal method*. We find conditions on the parameters α_k, ϵ_k, and λ_k in order to have a convergence result similar to the continuous case. Finally, in section 4 we state some of the questions opened by this work. Let us mention that the first to consider (1.1) for finite dimensional optimization problems was B. T. Polyack [16]. He studied a two-step discrete algorithm called the "heavy-ball with friction" method, which may be interpreted as an *explicit* discretization of (1.1). Both approaches are complementary; however, the analysis and the type of results in the implicit and explicit cases are different.

2. Dissipative differential system.

Throughout this paper, H is a real Hilbert space, $\langle \cdot, \cdot \rangle$ denotes the associated inner product, and $|\cdot|$ stands for the corresponding norm. We are interested in the behavior at infinity of $u : [0, \infty[\to H$, a solution of the following abstract evolution equation:

$$(E_\gamma; u_0, v_0) \qquad \begin{cases} u'' + \gamma u' + \nabla\Phi(u) = 0, \\ u(0) = u_0, \ u'(0) = v_0, \end{cases}$$

where $\gamma > 0$, $\Phi : H \to \mathbb{R}$, and $u_0, v_0 \in H$ are given. Note that if we assume that the gradient $\nabla\Phi$ is locally Lipschitz, then the existence and uniqueness of a local solution for $(E_\gamma; u_0, v_0)$ follow from standard results of differential equations theory. In that case, to prove that u is infinitely extendible to the right, it suffices to show that its derivative u' is bounded. Set

$$E(t) := \frac{1}{2}|u'(t)|^2 + \Phi(u(t)).$$

Since $E'(t) = -\gamma|u'(t)|^2$, the function E is nonincreasing. If we suppose that Φ is bounded from below, then u' is bounded.

2.1. Asymptotic convergence.

In that which follows, we suppose the existence of a global solution of $(E_\gamma; u_0, v_0)$. We write $\inf \Phi$ for the infimum value of Φ on H; thus, $\inf \Phi > -\infty$ will mean that Φ is bounded from below. We denote by $\text{Argmin}\,\Phi$ the set $\{x \in H : \Phi(x) = \inf \Phi\}$. On the nonlinearity we shall assume

$$(h_\Phi) \qquad \Phi \in C^1(H; \mathbb{R}) \text{ is convex and } \inf \Phi > -\infty.$$

THEOREM 2.1. *Suppose that (h_Φ) holds. If $u \in C^2([0, \infty[; H)$ is a solution of $(E_\gamma; u_0, v_0)$, then $u' \in L^2([0, \infty[; H)$, $u'(t) \to 0$ as $t \to \infty$, and*

$$\lim_{t\to\infty} \Phi(u(t)) = \inf \Phi. \tag{2.1}$$

Furthermore, if $\text{Argmin}\,\Phi \neq \emptyset$, then there exists $\widehat{u} \in \text{Argmin}\,\Phi$ such that $u(t) \rightharpoonup \widehat{u}$ weakly in H as $t \to \infty$.

We begin by noticing that u' is bounded (see the argument above). In order to prove the minimizing property (2.1), it suffices to prove that

$$\limsup_{t\to\infty} \Phi(u(t)) \le \Phi(x)$$

for any $x \in H$. Fix $x \in H$ and define the auxiliary function $\varphi(t) := \frac{1}{2}|u(t)-x|^2$. Since u is a solution of (E_γ), it follows that

$$\varphi'' + \gamma\varphi' = \langle \nabla\Phi(u), x-u\rangle + |u'|^2,$$

which together with the convexity inequality $\Phi(u) + \langle \nabla\Phi(u), x-u\rangle \le \Phi(x)$ yields

$$\varphi'' + \gamma\varphi' \le \Phi(x) - \Phi(u) + |u'|^2. \tag{2.2}$$

We do not have information on the behavior of $\Phi(u(t))$ but we know that $E(t)$ is nonincreasing. Thus, we rewrite (2.2) as

$$\varphi'' + \gamma\varphi' \le \Phi(x) - E(t) + \frac{3}{2}|u'|^2.$$

Given $t > 0$, for all $\tau \in [0, t]$ we have

$$\varphi''(\tau) + \gamma\varphi'(\tau) \le \Phi(x) - E(t) + \frac{3}{2}|u'(\tau)|^2.$$

After multiplication by $e^{\gamma\tau}$ and integration we obtain

$$\varphi'(t) \le e^{-\gamma t}\varphi'(0) + \frac{1}{\gamma}(1-e^{-\gamma t})[\Phi(x) - E(t)] + \frac{3}{2}\int_0^t e^{-\gamma(t-\tau)}|u'(\tau)|^2 d\tau.$$

We write this equation with t replaced by θ, and use the fact that $E(t)$ decreases and integrate once more to obtain

$$\varphi(t) \le \varphi(0) + \frac{1}{\gamma}(1-e^{-\gamma t})\varphi'(0) + \frac{1}{\gamma^2}(\gamma t - 1 + e^{-\gamma t})[\Phi(x) - E(t)] + h(t), \tag{2.3}$$

where

$$h(t) := \frac{3}{2}\int_0^t \int_0^\theta e^{-\gamma(\theta-\tau)}|u'(\tau)|^2 d\tau d\theta.$$

Since $E(t) \ge \Phi(u(t))$, (2.3) gives

$$\frac{1}{\gamma^2}(\gamma t - 1 + e^{-\gamma t})\Phi(u(t)) \le \varphi(0) + \frac{1}{\gamma}(1-e^{-\gamma t})\varphi'(0) + \frac{1}{\gamma^2}(\gamma t - 1 + e^{-\gamma t})\Phi(x) + h(t).$$

Dividing this inequality by $\frac{1}{\gamma^2}(\gamma t - 1 + e^{-\gamma t})$ and letting $t \to \infty$ we get

$$\limsup_{t\to\infty} \Phi(u(t)) \le \Phi(x) + \limsup_{t\to\infty} \frac{\gamma}{t} h(t).$$

It suffices to show that $h(t)$ remains bounded as $t \to \infty$. By Fubini's theorem

$$h(t) = \frac{3}{2}\int_0^t \int_\tau^t e^{-\gamma(\theta-\tau)}|u'(\tau)|^2 d\theta d\tau = \frac{3}{2\gamma}\int_0^t |u'(\tau)|^2(1-e^{-\gamma(t-\tau)})d\tau.$$

Note that from the equality $E' = -\gamma |u'|^2$ it follows that

$$\frac{1}{2}|u'|^2 + \Phi(u) + \gamma \int_0^t |u'(\tau)|^2 d\tau = E_0,$$

and in particular,

$$\int_0^t |u'(\tau)|^2 d\tau \leq \frac{E_0 - \inf \Phi}{\gamma} < \infty.$$

Then $u' \in L^2([0, \infty[; H)$, and

$$h(t) \leq \frac{3}{2\gamma} \int_0^t |u'(\tau)|^2 d\tau \leq \frac{3}{2\gamma} \int_0^\infty |u'(\tau)|^2 d\tau < \infty.$$

On the other hand, since $E(\cdot)$ is nonincreasing and bounded from below by $\inf \Phi$, it converges as $t \to \infty$. If $\lim_{t\to\infty} E(t) > \inf \Phi$, then $\lim_{t\to\infty} |u'(t)| > 0$ because of (2.1). This contradicts the fact that $u' \in L^2$. Therefore, $\lim_{t\to\infty} E(t) = \inf \Phi$, hence $u'(t) \to 0$ as $t \to \infty$.

The task now is to establish the weak convergence of $u(t)$ when $\operatorname{Argmin} \Phi \neq \emptyset$. For this purpose, we shall apply the Opial lemma [15], which holds interest in that it allows one to prove convergence without knowing the limit point. We state it as follows.

LEMMA (Opial). *Let H be a Hilbert space, let $\{u(t) : t \to \infty\} \subset H$ be a trajectory, and denote by W the set of its weak limit points*

$$W := \{y \in H : \exists t_k \to \infty \ s.t. \ u(t_k) \rightharpoonup y\}.$$

If there exists $\emptyset \neq S \subset H$ such that

$$\forall z \in S, \quad \lim_{t\to\infty} |u(t) - z| \ \text{exists}, \tag{2.4}$$

then $W \neq \emptyset$. Moreover, if $W \subset S$, then $u(t)$ converges weakly toward $\widehat{u} \in S$ as $t \to \infty$.

In order to apply the above result, we must find an adequate set S. Suppose that there exists $\widehat{u} \in H$ such that $u(t_k) \rightharpoonup \widehat{u}$ for a suitable sequence $t_k \to \infty$. The function Φ is weak lower-semicontinuous, because Φ is convex and continuous; hence

$$\Phi(\widehat{u}) \leq \liminf_{k\to\infty} \Phi(u(t_k)) = \lim_{t\to\infty} \Phi(u(t)) = \inf \Phi,$$

and therefore $\widehat{u} \in \operatorname{Argmin} \Phi$. According to the Opial lemma, we have only to prove that

$$\forall z \in \operatorname{Argmin} \Phi, \quad \lim_{t\to\infty} |u(t) - z| \ \text{ exists.}$$

For this, fix $z \in \operatorname{Argmin} \Phi$ and define $\varphi(t) := \frac{1}{2}|u(t) - z|^2$. The following lemma provides a sufficient condition on $[\varphi']_+$, the positive part of the derivative, in order to ensure convergence for φ.

LEMMA 2.2. *Let $\theta \in C^1([0, \infty[; \mathbb{R})$ be bounded from below. If $[\theta']_+ \in L^1([0, \infty[; \mathbb{R})$, then $\theta(t)$ converges as $t \to \infty$.*

Proof. Set

$$w(t) := \theta(t) - \int_0^t [\theta'(\tau)]_+ d\tau.$$

Since $w(t)$ is bounded from below and $w'(t) \leq 0$, then $w(t)$ converges as $t \to \infty$, and consequently $\theta(t)$ converges as $t \to \infty$. ∎

On account of this result, it suffices to prove that $[\varphi']_+$ belongs to $L^1(0,\infty)$. Of course, to obtain information on φ' we shall use the fact that $u(t)$ is solution of (E_γ). Due to the optimality of z, it follows from (2.2) that

$$\varphi'' + \gamma\varphi' \leq |u'|^2. \tag{2.5}$$

LEMMA 2.3. *If $\omega \in C^1([0,\infty[;\mathbb{R})$ satisfies the differential inequality*

$$\omega' + \gamma\omega \leq g(t) \tag{2.6}$$

with $\gamma > 0$ and $g \in L^1([0,\infty[;\mathbb{R})$, then $[\omega]_+ \in L^1([0,\infty[;\mathbb{R})$.

Proof. We can certainly assume that $g \geq 0$, for if not, we replace g by $|g|$. Multiplying (2.6) by $e^{\gamma t}$ and integrating we get

$$\omega(t) \leq e^{-\gamma t}\omega(0) + \int_0^t e^{-\gamma(t-\tau)} g(\tau)d\tau.$$

Thus

$$[\omega(t)]_+ \leq e^{-\gamma t}[\omega(0)]_+ + \int_0^t e^{-\gamma(t-\tau)} g(\tau)d\tau,$$

and Fubini's theorem gives $\int_0^\infty \int_0^t e^{-\gamma(t-\tau)} g(\tau)d\tau dt = \frac{1}{\gamma}\int_0^\infty g(\tau)d\tau < \infty$. ∎

Recalling that $|u'|^2 \in L^1([0,\infty[;\mathbb{R})$, the proof of the theorem is completed by applying Lemma 2.3 to (2.5). ∎

We say that $\nabla\Phi$ is *strongly monotone* if there exists $\beta > 0$ such that for any $x, y \in H$ we have

$$\langle \nabla\Phi(x) - \nabla\Phi(y), x - y\rangle \geq \beta|x-y|^2.$$

A weaker condition is the strong monotonicity over bounded sets, that is to say, for all $K > 0$ there exists $\beta_K > 0$ such that for any $x, y \in B[0,K]$ we have

$$\langle \nabla\Phi(x) - \nabla\Phi(y), x - y\rangle \geq \beta_K|x-y|^2. \tag{2.7}$$

If the latter property holds, then we have strong convergence for $u(t)$ when the infimum of Φ is attained. The argument is standard: let $\widehat{u}$ be the (unique) minimum point for Φ and set $K := \max\{\sup_{t\geq 0}|u(t)|, |\widehat{u}|\}$; then from (2.7) we deduce

$$\Phi(\widehat{u}) + \frac{\beta_K}{2}|u(t) - \widehat{u}|^2 \leq \Phi(u(t)). \tag{2.8}$$

Since we have proven that $\lim_{t\to\infty}\Phi(u(t)) = \inf\Phi = \Phi(\widehat{u})$, estimate (2.8) implies $u(t) \to \widehat{u}$ strongly in H. Note that we do not need to apply the Opial lemma.

The latter is the case of a nondegenerate minimum point. When Φ admits multiple minima, it is not possible to obtain strong convergence without additional assumptions on Φ or the space H. For instance, we have the following.

THEOREM 2.4. *Under the hypotheses of Theorem* 2.1, *if either*

(i) $\operatorname{Argmin}\Phi \neq \emptyset$ *and* Φ *is even*

or

(ii) $\text{int}(\text{Argmin}\,\Phi) \neq \emptyset$,
then

$$u(t) \to \widehat{u} \textit{ strongly in } H \textit{ as } t \to \infty,$$

where $\widehat{u} \in \text{Argmin}\,\Phi$.

Proof. The proof is adapted from the corresponding results for the steepest descent method; see [7] for the analogous hypothesis of (i) and [6] for (ii).

(i) Fix $t_0 > 0$ and define $g : [0, t_0] \to \mathbb{R}$ by

$$g(t) := |u(t)|^2 - |u(t_0)|^2 - \frac{1}{2}|u(t) - u(t_0)|^2.$$

Then $g'(t) = \langle u'(t), u(t) + u(t_0)\rangle$ and $g''(t) = \langle u''(t), u(t) + u(t_0)\rangle + |u'(t)|^2$. Consequently

$$g''(t) + \gamma g'(t) = \langle -\nabla\Phi(u(t)), u(t) + u(t_0)\rangle + |u'(t)|^2.$$

Since $E(t) = \frac{1}{2}|u'(t)|^2 + \Phi(u(t))$ is decreasing and Φ is even, we deduce that

$$E(t) \geq \frac{1}{2}|u'(t_0)|^2 + \Phi(-u(t_0))$$

for all $t \in [0, t_0]$. By the convexity of Φ we conclude that

$$E(t) \geq \frac{1}{2}|u'(t_0)|^2 + \Phi(u(t)) + \langle \nabla\Phi(u(t)), -u(t) - u(t_0)\rangle$$

and hence that

$$\frac{1}{2}|u'(t)|^2 \geq \langle -\nabla\Phi(u(t)), u(t) + u(t_0)\rangle.$$

Thus

$$g''(t) + \gamma g'(t) \leq \frac{3}{2}|u'(t)|^2.$$

The standard integration procedure yields

$$g(t_0) - g(t) \leq \frac{1}{\gamma}(e^{-\gamma t} - e^{-\gamma t_0})g'(0) + \frac{3}{2}\int_t^{t_0}\int_0^{\theta} e^{-\gamma(\theta-\tau)}|u'(\tau)|^2 d\tau d\theta.$$

Therefore, for all $t \in [0, t_0]$ we have that

$$\frac{1}{2}|u(t) - u(t_0)|^2 \leq |u(t)|^2 - |u(t_0)|^2 + \frac{1}{\gamma}(e^{-\gamma t} - e^{-\gamma t_0})g'(0) + h(t_0) - h(t), \tag{2.9}$$

where

$$h(t) = \frac{3}{2}\int_0^{t}\int_0^{\theta} e^{-\gamma(\theta-\tau)}|u'(\tau)|^2 d\tau d\theta.$$

On the other hand, in the proof of Theorem 2.1 we have shown that $h(t)$ is convergent as $t \to \infty$. We also proved that for all $z \in \text{Argmin}\,\Phi$ the $\lim_{t\to\infty} |u(t) - z|$ exists. Since Φ is convex and even, we have $0 \in \text{Argmin}\,\Phi$ whenever the infimum is realized.

In that case, $|u(t)|$ is convergent as $t \to \infty$ and we infer from (2.9) that $\{u(t) : t \to \infty\}$ is a Cauchy net. Hence $u(t)$ converges strongly as $t \to \infty$ and, by Theorem 2.1, the limit belongs to $\operatorname{Argmin} \Phi$.

(ii) Let $z_0 \in \operatorname{int}(\operatorname{Argmin} \Phi)$. There exists $\rho > 0$ such that for every $z \in H$ with $|z - z_0| \le \rho$, then $z \in \operatorname{int}(\operatorname{Argmin} \Phi)$. In particular, if $|z - z_0| \le \rho$, then $\nabla\Phi(z) = 0$. Consequently,

$$\langle \nabla\Phi(x), x - z_0\rangle \ge \langle \nabla\Phi(x), z - z_0\rangle$$

for every $x \in H$ and z with $|z - z_0| \le \rho$. Hence,

$$\langle \nabla\Phi(x), x - z_0\rangle \ge \rho|\nabla\Phi(x)|$$

for every $x \in H$. Applying this inequality to $x = u(t)$ we deduce that

$$-\langle u'' + \gamma u', u - z_0\rangle \ge \rho|u'' + \gamma u'|.$$

Set $\varphi(t) := \frac{1}{2}|u(t) - z_0|^2$. We thus obtain

$$-\varphi'' + |u'|^2 - \gamma\varphi' \ge \rho|u'' + \gamma u'|.$$

Integrating this inequality yields

$$\varphi'(0) - \varphi'(t) + \int_0^t |u'(\tau)|^2 d\tau + \gamma(\varphi(0) - \varphi(t)) \ge \rho \int_0^t |u''(\tau) + \gamma u'(\tau)| d\tau.$$

We have already proved that the $\lim_{t\to\infty} \varphi(t)$ exists and $\lim_{t\to\infty} \varphi'(t) = 0$. Moreover, $u' \in L^2(0, \infty; H)$. As a conclusion, $u'' + \gamma u' \in L^1(0, \infty; H)$. We deduce that the $\lim_{t\to\infty} u'(t) + \gamma u(t)$ exists, which finishes the proof because $u'(t) \to 0$ as $t \to \infty$. □

2.2. Localization of the limit point. In the proof of Theorem 2.1 we have used the differential inequality (2.2), which in some sense measures the evolution of the system. A simpler but analogous inequality appears in the asymptotic analysis for the steepest descent inclusion (SD). This was used by B. Lemaire in [13] to *locate* the limit point of the trajectories of (SD). Following this approach, in this section we give a localization result of the limit point of the solutions of (E_γ). For simplicity of notation, set $S := \operatorname{Argmin} \Phi$ and we denote by $\operatorname{proj}_S : H \to S$ the projection operator onto the closed convex set S.

PROPOSITION 2.5. *Let u be solution of $(E_\gamma; u_0, v_0)$ and $\widehat{u} \in S$ be such that $u(t) \rightharpoonup \widehat{u}$ weakly as $t \to \infty$. Then, for all $x \in S$*

$$|\widehat{u} - x| \le |u_0 + \frac{1}{\gamma}v_0 - x| + \frac{1}{\gamma}\delta(u_0), \tag{2.10}$$

where $\delta(u_0) = \sqrt{2}[\Phi(u_0) - \inf \Phi]^{1/2}$. Consequently

(i) $$|\widehat{u} - \operatorname{proj}_S(u_0 + \frac{1}{\gamma}v_0)| \le d(u_0 + \frac{1}{\gamma}v_0, S) + \frac{1}{\gamma}\delta(u_0),$$

where $d(u_0, S)$ is the distance between u_0 and the set S.

(ii) *If S is an affine subspace of H, then*

$$|\widehat{u} - \operatorname{proj}_S(u_0 + \frac{1}{\gamma}v_0)| \le \frac{1}{\gamma}\delta(u_0).$$

If, moreover, Φ is a quadratic form, then

$$u(t) \to \mathrm{proj}_S(u_0 + \frac{1}{\gamma}v_0) \text{ strongly in } H \text{ as } t \to \infty.$$

Proof. Let $x \in S$ and set $\varphi(t) := \frac{1}{2}|u(t) - x|^2$. The inequality (2.2) and the optimality of x give $\varphi'' + \gamma\varphi' \leq |u'|^2$. Hence

$$\varphi(t) \leq \varphi(0) + \frac{1}{\gamma}(1 - e^{-\gamma t})\varphi'(0) + \int_0^t \int_0^\theta e^{-\gamma(\theta-\tau)}|u'(\tau)|^2 d\tau d\theta.$$

Due to the weak lower-semicontinuity of the norm and Fubini's theorem, we can let $t \to \infty$ to obtain

$$\frac{1}{2}|\widehat{u} - x|^2 \leq \frac{1}{2}|u_0 - x|^2 + \frac{1}{\gamma}\langle v_0, u_0 - x\rangle + \frac{1}{\gamma}\int_0^\infty |u'(\tau)|^2 d\tau. \tag{2.11}$$

On the other hand, from the energy equation

$$\frac{1}{2}|u'|^2 + \Phi(u) + \gamma\int_0^t |u'(\tau)|^2 d\tau = \frac{1}{2}|v_0|^2 + \Phi(u_0),$$

it follows that

$$\int_0^\infty |u'(\tau)|^2 d\tau \leq \frac{1}{\gamma}\Big[\frac{1}{2}|v_0|^2 + \Phi(u_0) - \inf \Phi\Big].$$

Replacing the last estimate in (2.11), it easy to show that (2.10) holds.

For (i), it suffices to take $x = \mathrm{proj}_S(u_0 + \frac{1}{\gamma}v_0)$ in (2.10).

For (ii), let $e := \widehat{u} - \mathrm{proj}_S(u_0 + \frac{1}{\gamma}v_0)$. If $e \neq 0$, then set

$$x_r := \mathrm{proj}_S\Big(u_0 + \frac{1}{\gamma}v_0\Big) - rd\Big(u_0 + \frac{1}{\gamma}v_0, S\Big)\frac{e}{|e|},$$

which belongs to S. An easy computation shows that

$$\Big|u_0 + \frac{1}{\gamma}v_0 - x_r\Big| - |\widehat{u} - x_r| = \Big(\sqrt{1 + r^2} - r\Big)d\Big(u_0 + \frac{1}{\gamma}v_0 - x, S\Big) - |e|,$$

which together with (2.10) yields

$$|e| \leq \Big(\sqrt{1 + r^2} - r\Big)d\Big(u_0 + \frac{1}{\gamma}v_0 - x, S\Big) + \frac{1}{\gamma}\delta(u_0).$$

Letting $r \to \infty$ we get the result.

Finally, suppose that $\Phi(x) = \frac{1}{2}\langle Ax, x\rangle$ where $A : H \to H$ is a positive and self-adjoint bounded linear operator. Then $S = \{x \in H \mid Ax = 0\}$ the null space of A. Let $z \in S$; for all $t \geq 0$ we have that

$$\begin{aligned}\langle u'(t) - v_0, z\rangle + \gamma\langle u(t) - u_0, z\rangle &= \int_0^t \langle u''(\tau) + \gamma u'(\tau), z\rangle d\tau \\ &= \int_0^t \langle -Au(\tau), z\rangle d\tau \\ &= \int_0^t -\langle u(\tau), Az\rangle d\tau = 0.\end{aligned}$$

Since $u'(t) \to 0$ and $u(t) \to \widehat{u} \in S$ strongly (Φ is even) as $t \to \infty$, we can deduce that

$$\Big\langle \widehat{u} - \Big(u_0 + \frac{1}{\gamma}v_0\Big), z\Big\rangle = 0$$

for all $z \in S$, which completes the proof. ∎

2.3. Linear system: Heuristic comparison. Before proceeding further, it is interesting from the optimization viewpoint to compare the behavior of the trajectories defined by

$$(E_\gamma) \qquad u'' + \gamma u' + \nabla\Phi(u) = 0,$$

with the steepest descent equation

$$\text{(SD)} \qquad u' + \nabla\Phi(u) = 0,$$

and with the continuous Newton's method

$$\text{(N)} \qquad u' + \nabla^2\Phi(u)^{-1}\nabla\Phi(u) = 0.$$

For simplicity, in this section we restrict ourselves to the associated linearized systems in a finite dimensional space. We shall consider $H = \mathbb{R}^N$ and assume that $\Phi \in C^2(\mathbb{R}^N;\mathbb{R})$. Related to (SD), we have the linearized system around some $x_0 \in \mathbb{R}^N$, which is defined by

$$\text{(LSD)} \qquad x' + \nabla^2\Phi(x_0)(x - x_0) + \nabla\Phi(x_0) = 0.$$

We assume that the Hessian matrix $\nabla^2\Phi(x_0)$ is positive definite. An explicit computation shows that $x(t) \to \widehat{x} := x_0 - \nabla^2\Phi(x_0)^{-1}\nabla\Phi(x_0)$ as $t \to \infty$. In fact, the solutions of (LSD) are of the form $y(t) = \widehat{x} + \eta(t)$, where η solves the homogeneous equation $\eta' + \nabla^2\Phi(x_0)\eta = 0$. Take a matrix P such that $P^{-1}\nabla^2\Phi(x_0)P = \text{diag}(\lambda_1, \dots, \lambda_N)$, where $\lambda_i > 0$, and set $P\xi = \eta$. We obtain the system $\xi_i' + \lambda_i\xi_i = 0$, whose solutions are $\xi_i(t) = C_i e^{-\lambda_i t}$. Generally speaking, if there is a $\lambda_i << 1$, we will have a relative slow convergence towards the solution; on the other hand, when dealing with large λ_i's the numerical integration by an approximate method will present stability problems. Thus we see that the numerical performance of (SD) is strongly determined by the *local geometry* of the function Φ.

We turn now to the linearized version of (N), given by

$$\text{(LN)} \qquad y' + y - \widehat{x} = 0.$$

The solutions are of the form $y(t) = \widehat{x} + e^{-t}y(0)$, which are much better than the previous ones. The major properties are (1) the straight-line geometry of the trajectories; (2) that the rate of convergence is independent of the quadratic function to be minimized. Certainly, this is just a local approximation of the original function and the global behavior of the trajectory may be complicated. Nevertheless, this outstanding *normalization* property of Newton's system makes it effective in practice, due to the fact that the associated trajectories are easy to follow by a discretization method. Of course, an important disadvantage of (N) is the computation of the inverse of the Hessian matrix, which may be involved for a numerical algorithm.

Finally, we consider

$$(LE_\gamma) \qquad z'' + \gamma z' + \nabla^2\Phi(x_0)(z - x_0) + \nabla\Phi(x_0) = 0.$$

For this equation we have $z(t) = \widehat{x} + \epsilon(t)$, where ϵ solves the homogeneous problem $\epsilon'' + \gamma\epsilon' + \nabla^2\Phi(x_0)\epsilon = 0$. Setting $P\delta = \epsilon$ with P as above, then δ_i satisfies $\delta_i'' + \gamma\delta_i' + \lambda_i\delta_i = 0$. It is a simple matter to show that $|\delta_i(t)| \le C_i e^{-\mu_i(\gamma)t}$ with $\mu_i :]0,\infty[\to]0,\infty[$ continuous and C_i a constant independent of γ. In fact, $\mu_i(\gamma) = \frac{\gamma}{2}$ if $\gamma \in]0, 2\sqrt{\lambda_i}]$ and

$\mu_i(\cdot)$ is nonincreasing on $]2\sqrt{\lambda_i},\infty[$. Moreover, if $\gamma \geq 2\sqrt{\lambda_i}$, then the corresponding $\delta_i(t)$ does not present oscillations. Thus the choice $\gamma = 2\sqrt{\lambda_i}$ gives $\mu_i = \sqrt{\lambda_i}$, the greatest rate that can be obtained. But we can get any value in the interval $]0,\sqrt{\lambda_i}]$; for instance, when $\lambda_i > 1$ we obtain $\mu_i = 1$ either with $\gamma = 2$ or $\gamma = \lambda_i + 1$. The last choice has the advantage that the associated trajectory is not oscillatory, which is interesting by numerical reasons. Note that we should take a different parameter γ according to the corresponding eigenvalue λ_i.

Therefore, the presence of the damping parameter γ gives us a control on the behavior of the solutions of (E_γ) and, in particular, on some qualitative properties of the associated trajectories. For a general Φ we must take into account that (a) a careful selection of the damping parameter γ should depend on the local geometry of the function Φ, leading to a nonautonomous damping; (b) this selection could give a different value of γ for some particular directions, leading to an anisotropic damping. No attempt has been made here to develop a theory in order to guide these choices.

2.4. Linear and anisotropic damping. In the preceding section we have seen that it may be of interest to consider an anisotropic damping. With the aim of contributing to this issue, in this section we establish the asymptotic convergence for the solutions of the following system:

$$(E_\Gamma; u_0, v_0) \qquad \begin{cases} u'' + \Gamma u' + \nabla\Phi(u) = 0, \\ u(0) = u_0,\ u'(0) = v_0, \end{cases}$$

where $\Gamma : H \to H$ is a bounded self-adjoint linear operator, which we assume to be elliptic:

(h_Γ) there exists $\gamma > 0$ such that for any $x \in H$, $\langle \Gamma x, x\rangle \geq \gamma |x|^2$.

THEOREM 2.6. *Suppose (h_Φ) and (h_Γ) hold. If $u \in C^2([0,\infty[; H)$ is a solution of $(E_\Gamma; u_0, v_0)$, then it satisfies $u' \in L^2([0,\infty[; H)$, $u'(t) \to 0$ as $t \to \infty$, and*

$$\lim_{t\to\infty} \Phi(u(t)) = \inf \Phi. \tag{2.12}$$

Furthermore, if $\operatorname{Argmin}\Phi \neq \emptyset$, *then there exists* $\widehat{u} \in \operatorname{Argmin}\Phi$ *such that* $u(t) \rightharpoonup \widehat{u}$ *weakly in H as $t \to \infty$.*

Proof. We only need to adapt the proof of Theorem 2.1. First, note that the properties of existence, uniqueness, and infinite extendibility to the right of the solution follow by similar arguments. Likewise, the energy $E(t) := \frac{1}{2}|u'(t)|^2 + \Phi(u(t))$ satisfies $E' = -\langle \Gamma u', u'\rangle$, and we can deduce that $u' \in L^2$.

Next, define the operator $A : H \to H$ by $Ax := \Gamma x - \gamma x$, with $\gamma > 0$ given by (h_Γ). Fix $x \in H$ and set $\varphi(t) := \frac{1}{2}|u(t) - x|^2$ and $\rho(t) := \frac{1}{2}\langle A(u(t)-x), u(t)-x\rangle$, in such a way that

$$\varphi'' + \gamma\varphi' + \rho' = \langle \nabla\Phi(u), x - u\rangle + |u'|^2. \tag{2.13}$$

As in the proof of Theorem 2.1, (2.13) gives

$$\varphi'(t) + \int_0^t e^{-\gamma(t-\tau)}\rho'(\tau)d\tau \leq e^{-\gamma t}\varphi'(0) + \frac{1}{\gamma}(1 - e^{-\gamma t})[\Phi(x) - E(t)] + r(t), \tag{2.14}$$

with

$$r(t) := \frac{3}{2}\int_0^t e^{-\gamma(t-\tau)}|u'(\tau)|^2 d\tau,$$

the only difference being the term $\int_0^t e^{-\gamma(t-\tau)}\rho'(\tau)d\tau$. An integration by parts yields

$$\int_0^t e^{-\gamma(t-\tau)}\rho'(\tau)d\tau = \rho(t) - e^{-\gamma t}\rho(0) - \gamma \int_0^t e^{-\gamma(t-\tau)}\rho(\tau)d\tau.$$

Setting $f(t) := \int_0^t e^{-\gamma(t-\tau)}\rho(\tau)d\tau$, we have

$$\int_0^t e^{-\gamma(t-\tau)}\rho'(\tau)d\tau = f'(t) - e^{-\gamma t}\rho(0).$$

Thus, we can rewrite (2.14) as

$$\varphi'(t) + f'(t) \leq e^{-\gamma t}(\varphi'(0) + \rho(0)) + \frac{1}{\gamma}(1 - e^{-\gamma t})[\Phi(x) - E(t)] + r(t).$$

We leave it to the reader to verify that the minimizing property (2.12) can now be established as in Theorem 2.1. The proof of $u'(t) \to 0$ as $t \to \infty$ is analogous.

When $\operatorname{Argmin}\Phi \neq \emptyset$, we fix $z \in \operatorname{Argmin}\Phi$ and consider the corresponding functions φ and ρ as above (with x replaced by z). Using the optimality of z, it follows that

$$\varphi'(t) + f'(t) \leq e^{-\gamma t}(\varphi'(0) + \rho(0)) + \int_0^t e^{-\gamma(t-\tau)}|u'(\tau)|^2 d\tau, \tag{2.15}$$

with f associated with ρ as above. Integrating this inequality we conclude that $\varphi(t)$ stays bounded as $t \to \infty$, but we cannot deduce its convergence. Then, we rewrite (2.15) in the form

$$\varphi'(t) + \int_0^t e^{-\gamma(t-\tau)}\rho'(\tau)d\tau \leq e^{-\gamma t}\varphi'(0) + \int_0^t e^{-\gamma(t-\tau)}|u'(\tau)|^2 d\tau,$$

and we conclude that $[\varphi'(t) + \int_0^t e^{-\gamma(t-\tau)}\rho'(\tau)d\tau]_+ \in L^1([0,\infty[;\mathbb{R})$. We note that

$$\varphi'(t) + \int_0^t e^{-\gamma(t-\tau)}\rho'(\tau)d\tau = \mu'(t) + \xi'(t),$$

where

$$\mu(t) := \frac{1}{2\gamma}\langle \Gamma(u(t) - z), u(t) - z\rangle,$$

and

$$\xi(t) := -\frac{1}{\gamma}\int_0^t e^{-\gamma(t-\tau)}\rho'(\tau)d\tau.$$

By virtue of Lemma 2.2, if we show that $\xi(t)$ is bounded from below, then $\mu(t)+\xi(t)$ converges as $t \to \infty$. Since $\rho'(t) = \langle Au'(t), u(t) - z\rangle$, there exists a constant $M > 0$ independent of t such that $|\rho'(t)| \leq M|u'(t)|\sqrt{\varphi(t)}$ for any $t > 0$. We conclude that $\rho'(t) \to 0$ as $t \to \infty$. From this fact it follows easily that $\xi(t) \to 0$ as $t \to \infty$. Therefore, $\mu(t) + \xi(t)$ converges as $t \to \infty$, hence $\mu(t)$ converges as well.

The proof is completed by applying the Opial lemma to the trajectory $\{u(t) : t \to \infty\}$, where the Hilbert space H is endowed with the inner product $\langle\langle \cdot, \cdot\rangle\rangle : H \times H \to \mathbb{R}$ defined by $\langle\langle x, y\rangle\rangle := \frac{1}{\gamma}\langle \Gamma x, y\rangle$ and its associated norm. ∎

Remark 1. In Theorems 2.1, 2.4, and 2.6 we do not require any coerciveness assumption on Φ. When $\operatorname{Argmin}\Phi \neq \emptyset$, the dissipativeness in the dynamics suffices for the convergence of the solutions. If the infimum value is not realized, the trajectory may be unbounded as in the one-dimensional equation $u'' + \gamma u' + e^u = 0$, whose solutions $u \in C^2([0,\infty[;\mathbb{R})$ are so that $u(t) \to -\infty$ and $u'(t) \to 0$ as $t \to \infty$. In any case, our results assert that the dynamical system defined by (E_γ) (or more generally by (E_Γ)) is dissipative in the sense that every trajectory evolves towards a minimum of the energy. Certainly, there is a strong connection with the concept of point dissipativeness or ultimately boundedness in the theory of dynamical systems, where the Lyapunov function associated with the semigroup is usually supposed to be coercive (cf. [9, gradient systems]).

Remark 2. To ensure local existence and uniqueness of a classical solution for the differential equation, it suffices to require a local Lipschitz property on $\nabla\Phi$. Actually, in some situations this hypothesis is not necessary and the existence may be established by other arguments. For instance, that is the case of the Hille–Yosida theorem for evolution equations governed by monotone operators and the theory of linear and nonlinear semigroups for partial differential equations. Note that such a Lipschitz condition on the gradient is not used in the asymptotic analysis of the trajectories. Therefore, the previous asymptotic results remain valid for other classes of infinite dimensional dissipative systems provided the existence of a global solution. It is not our purpose to develop this point here for the continuous system because it exceeds the scope of this paper. However, in the next section we consider an implicit discretization of the continuous system. As we will see, the existence of the discrete trajectory is ensured by variational arguments. This will allow us to apply the discrete scheme to nonsmooth convex functions and to adapt the asymptotic analysis to this case.

3. Discrete approximation method. Once we have established the existence of a solution of an initial value problem, we are interested in its numerical values. We must accept that most differential equations cannot be solved explicitly; we are thus led to work with approximate methods. An important class of these methods is based on the approximation of the exact solution over a discrete set $\{t_n\}$: associated with each point t_n we compute a value u_n, which approximates $u(t_n)$ the exact solution at t_n. Generally speaking, these procedures have the disadvantage that a large number of calculations has to be done in order to keep the discretization error $e_n := u_n - u(t_n)$ sufficiently small. In addition to this, the estimates for the errors strongly depend on the length of the discretization range for the t variable. It turns out that these methods are not well adapted to the approximation of the exact solution on an unbounded domain.

Nevertheless, there is an important point to note here. If our objective is the asymptotic behavior of the solutions as t goes to ∞, then the accurate approximation of the whole trajectory becomes immaterial. We present a discrete method whose feature is that no attempt is made to approximate the exact solution over a set of points but that the discrete values are sought only to preserve the asymptotic behavior of the solutions.

3.1. Implicit iterative scheme. Dealing with the discretization of a first order differential equation $y' = F(y)$, it is classical to consider the implicit iterative scheme

$$\frac{y_{k+1} - y_k}{h} = F(y_{k+1}), \tag{3.1}$$

where $h > 0$ is a parameter called step size. In the case of equation (E_γ), or more precisely its first order equivalent system, (3.1) corresponds to recursively solve

$$\frac{u_{k+1} - 2u_k + u_{k-1}}{h^2} + \gamma \frac{u_{k+1} - u_k}{h} + \nabla\Phi(u_{k+1}) = 0. \tag{3.2}$$

Since Φ is convex, (3.2) is equivalent to the following variational problem:

$$u_{k+1} = \operatorname{argmin}\left\{\Phi(x) + \frac{1+\gamma h}{2h^2}|x - z_k|^2 : x \in H\right\},$$

where $z_k = u_k + \frac{1}{1+\gamma h}(u_k - u_{k-1})$. This motivates the introduction of the more general iterative procedure

$$u_{k+1} = \operatorname{argmin}\left\{\Phi(x) + \frac{1}{2\lambda}|x - z_k|^2 : x \in H\right\},$$

where $z_k = u_k + \alpha(u_k - u_{k-1})$, λ and α are positive. Note that when $\alpha = 0$, we recover the standard (Prox) iteration. If $\alpha > 0$, the starting point for the next iteration is computed as a development in terms of the *velocity* of the already generated sequence. Therefore, this iterative scheme defines a second order dynamics, while (Prox) is actually of a first order nature.

We have been working under the assumption that Φ is differentiable. However, for the above iterative variational method this regularity is no longer necessary. Thus, in that which follows $f : H \to \mathbb{R} \cup \{\infty\}$ denotes a closed proper convex function (see [17]), which eventually realizes the value ∞, and we consider

$$u_{k+1} = \operatorname{argmin}\left\{f(x) + \frac{1}{2\lambda}|x - z_k|^2 : x \in H\right\}, \tag{3.3}$$

where $z_k = u_k + \alpha(u_k - u_{k-1})$. In terms of the stationary condition, (3.3) is equivalent to

$$\frac{1}{\lambda}(u_{k+1} - (1+\alpha)u_k + \alpha u_{k-1}) + \partial f(u_{k+1}) \ni 0,$$

where ∂f is the standard convex subdifferential [17].

3.2. Convergence for the variational algorithm. By numerical reasons, it is natural to consider the following approximate iterative scheme:

$$\frac{1}{\lambda_k}(u_{k+1} - (1+\alpha_k)u_k + \alpha_k u_{k-1}) + \partial_{\epsilon_k} f(u_{k+1}) \ni 0, \tag{3.4}$$

where α_k is nonnegative, λ_k is positive, and $\partial_\epsilon f$ is the ϵ-subdifferential. Note that a sequence $\{u_k\} \subset H$ satisfying (3.4) always exists. Indeed, given $u_{k-1}, u_k \in H$, we can take u_{k+1} as the unique solution of the strongly convex problem $\min\{f(x) + \frac{1}{2\lambda_k}|x - z_k|^2 : x \in H\}$ with z_k as above.

THEOREM 3.1. *Assume that f is closed proper convex and bounded from below. Let $\{u_k\} \subset H$ be a sequence generated by (3.4), where*

(i) $0 \le \alpha_k \le 1$ *and $\{\lambda_k\}$ is bounded from below by a positive constant,*

(ii) *the sequence $\{\alpha_k/\lambda_k\}$ is nonincreasing and $\sum \lambda_k \epsilon_k < \infty$.*

Then

$$\lim_{k\to\infty} \frac{1}{\lambda_k}(u_{k+1} - (1+\alpha_k)u_k + \alpha_k u_{k-1}) = 0, \tag{3.5}$$

and in particular $\lim_{k\to\infty} d(0, \partial_{\epsilon_k} f(u_{k+1})) = 0$.

When $\operatorname{Argmin} f \neq \emptyset$, *assume in addition that*

(iii) *there exists* $\bar{\alpha} \in]0,1[$ *such that* $0 \leq \alpha_k \leq \bar{\alpha}$, *and* $\{\lambda_k\}$ *is bounded from above if there is at least one* $\alpha_k > 0$.

Then, there exists $\widehat{u} \in \operatorname{Argmin} f$ *such that* $u_k \rightharpoonup \widehat{u}$ *weakly as* $k \to \infty$.

Proof. The proof consists of adapting the analysis done for the differential equation (E_γ). We begin by defining the discrete energy by

$$E_{k+1} = \frac{\alpha_k}{2\lambda_k}|u_{k+1} - u_k|^2 + f(u_{k+1}),$$

and we study the successive difference $E_{k+1} - E_k$. Since $\alpha_k/\lambda_k \leq \alpha_{k-1}/\lambda_{k-1}$,

$$E_{k+1} - E_k \leq \frac{\alpha_k}{2\lambda_k}\left(|u_{k+1} - u_k|^2 - |u_k - u_{k-1}|^2\right) + f(u_{k+1}) - f(u_k).$$

By definition of $\partial_{\epsilon_k} f$, (3.4) yields

$$f(u_{k+1}) - f(u_k) \leq -\frac{1}{\lambda_k}\langle u_{k+1} - (1+\alpha_k)u_k + \alpha_k u_{k-1}, u_{k+1} - u_k\rangle + \epsilon_k.$$

As we can write

$$\langle u_{k+1} - (1+\alpha_k)u_k + \alpha_k u_{k-1}, u_{k+1} - u_k\rangle = |u_{k+1} - u_k|^2 - \alpha_k\langle u_k - u_{k-1}, u_{k+1} - u_k\rangle,$$

we have

$$E_{k+1} - E_k \leq -\frac{\alpha_k}{2\lambda_k}|u_{k+1} - 2u_k + u_{k-1}|^2 - \frac{1-\alpha_k}{\lambda_k}|u_{k+1} - u_k|^2 + \epsilon_k,$$

and consequently

$$\sum_{k=1}^{N}\left[\frac{\alpha_k}{2\lambda_k}|u_{k+1} - 2u_k + u_{k-1}|^2 + \frac{1-\alpha_k}{\lambda_k}|u_{k+1} - u_k|^2\right] \leq E_1 - E_{N+1} + \sum_{k=1}^{N}\epsilon_k.$$

Noting that

$$E_1 - E_{N+1} + \sum_{k=1}^{N}\epsilon_k \leq E_1 - \inf f + \sum \epsilon_k < \infty,$$

and because $0 \leq \alpha_k \leq 1$, we deduce that

$$\sum \frac{\alpha_k}{2\lambda_k}|u_{k+1} - 2u_k + u_{k-1}|^2 < \infty$$

and

$$\sum \frac{1-\alpha_k}{\lambda_k}|u_{k+1} - u_k|^2 < \infty. \tag{3.6}$$

As $0 \leq \alpha_k \leq 1$ and λ_k is bounded from below by a positive constant, we have

$$\lim_{k\to\infty}\frac{\alpha_k}{\lambda_k}|u_{k+1} - 2u_k + u_{k-1}| = \lim_{k\to\infty}\frac{(1-\alpha_k)}{\lambda_k}|u_{k+1} - u_k| = 0.$$

Writing

$$u_{k+1} - (1+\alpha_k)u_k + \alpha_k u_{k-1} = \alpha_k(u_{k+1} - 2u_k + u_{k-1}) + (1-\alpha_k)(u_{k+1} - u_k),$$

we conclude that (3.5) holds.

Suppose now that $\operatorname{Argmin} f \neq \emptyset$. We apply the Opial lemma to prove the weak convergence of $\{u_k\}$. On account of (3.5), it is sufficient to show that for any $z \in \operatorname{Argmin} f$, the sequence of positive numbers $\{|u_k - z|\}$ is convergent. Fix $z \in \operatorname{Argmin} f$; since u_{k+1} satisfies (3.4), we have

$$f(u_{k+1}) - \frac{1}{\lambda_k}\langle u_{k+1} - (1+\alpha_k)u_k + \alpha_k u_{k-1}, z - u_{k+1}\rangle \leq f(z) + \epsilon_k,$$

and by the optimality of z

$$\langle u_{k+1} - u_k, u_{k+1} - z\rangle - \alpha_k\langle u_k - u_{k-1}, u_{k+1} - z\rangle \leq \lambda_k \epsilon_k. \tag{3.7}$$

Set $\varphi_k := \frac{1}{2}|u_k - z|^2$. It is direct to check that for any $k \in \mathbb{N}$

$$\varphi_{k+1} = \varphi_k + \langle u_{k+1} - u_k, u_{k+1} - z\rangle - \frac{1}{2}|u_{k+1} - u_k|^2.$$

Since $\langle u_k - u_{k-1}, u_{k+1} - z\rangle = \langle u_k - u_{k-1}, u_k - z\rangle + \langle u_k - u_{k-1}, u_{k+1} - u_k\rangle$, (3.7) shows that

$$\varphi_{k+1} - \varphi_k - \alpha_k\left(\varphi_k - \varphi_{k-1} + \frac{1}{2}|u_k - u_{k-1}|^2 + \langle u_k - u_{k-1}, u_{k+1} - u_k\rangle\right) \leq \lambda_k\epsilon_k,$$

and therefore

$$\varphi_{k+1} - (1+\alpha_k)\varphi_k + \alpha_k\varphi_{k-1} \leq \delta_k,$$

where $\delta_k = \alpha_k|u_k - u_{k-1}|^2 + \frac{\alpha_k}{2}|u_{k+1} - u_k|^2 + \lambda_k\epsilon_k$. Using (iii) and (3.6) it follows that $\sum |u_{k+1} - u_k|^2 < \infty$, thus $\sum \delta_k < \infty$. Set $\theta_k := \varphi_k - \varphi_{k-1}$; the above inequality implies

$$[\theta_{k+1}]_+ \leq \bar{\alpha}[\theta_k]_+ + \delta_k.$$

Thus

$$[\theta_{k+1}]_+ \leq \bar{\alpha}^k[\theta_1]_+ + \sum_{j=0}^{k-1}\bar{\alpha}^j\delta_{k-j},$$

which yields

$$\sum_{k=0}^{\infty}[\theta_{k+1}]_+ \leq \frac{1}{1-\bar{\alpha}}\left([\theta_1]_+ + \sum_{k=1}^{\infty}\delta_k\right) < \infty.$$

Set $w_k := \varphi_k - \sum_{j=1}^{k}[\theta_j]_+$. Since $\varphi_k \geq 0$ and $\sum[\theta_j]_+ < \infty$, w_k is bounded from below. As $\{w_k\}$ is nonincreasing we have that it converges. Hence $\{\varphi_k\}$ converges, which completes the proof of the theorem. ∎

For simplicity, we have considered in this section the isotropic damping system. However, a similar analysis can be done for the anisotropic damping associated with an elliptic self-adjoint linear operator $\Gamma : H \to H$. The variational problem associated with the implicit discretization is

$$u_{k+1} = \operatorname{argmin} \left\{ \Phi(x) + \frac{1}{2h^2} \left| x - z_k \right|^2_{(I+h\Gamma)} : x \in H \right\},$$

where $z_k = u_k + (I + h\Gamma)^{-1}(u_k - u_{k-1})$ and for any $y \in H$,

$$|y|_{(I+h\Gamma)} := \sqrt{\langle (I + h\Gamma)y, y \rangle}.$$

For a function $f : H \to \mathbb{R} \cup \{\infty\}$ closed proper and convex, the latter motivates the scheme

$$R(u_{k+1} - (I + S)u_k + Su_{k-1}) + \partial f(u_{k+1}) \ni 0,$$

where $R : H \to H$ is a linear positive definite operator and $S : H \to H$ is linear and positive semidefinite. If we assume both R and $I - S$ are elliptic, it is possible to obtain a convergence result like the previous one. It suffices to adapt the main arguments. Since the basic ideas are contained in the proof of Theorems 2.6 and 3.1, we shall go no further in this matter.

4. Some open problems. In the case of multiple optimal solutions, our convergence results do not provide additional information on the point attained in the limit. A possible approach to overcome this disadvantage may be to couple the dissipative system with approximation techniques such as regularization, interior-barrier or globally defined penalizations, and viscosity methods. In the continuous case, this alternative has been considered with success for the steepest descent equation in [2] and for Newton's method in [1], giving a characterization for the limit point under suitable assumptions on the approximate scheme. On account of these results, one may conjecture that this can be done for the equations considered in the present work.

On the other hand, we have seen that the behavior of the trajectories depends on a relation between the damping and the local geometry of the function we wish to minimize. This remark leads us to the obvious problem of the choice of the damping parameter, made in order to have a better control on the trajectory. This is also a problem in the discrete algorithm. Usually we have an incomplete knowledge of the objective function, which makes the question more difficult. We think that a first step in this direction may be the study of more general damped equations, with nonlinear and/or nonautonomous damping.

Acknowledgments. I wish to express my gratitude to the Laboratoire d'Analyse Convexe de l'Université Montpellier II for their hospitality and support, and especially to Professor Hedy Attouch. I gratefully acknowledge financial support through a French Foreign Scholarship grant from the French Ministry of Education and a Chilean National Scholarship grant from the CONICYT of Chile. I wish to thank a referee for helpful comments concerning Theorems 2.2 and 3.1.

REFERENCES

[1] F. Alvarez and J. M. Pérez, *A dynamical system associated with Newton's method for parametric approximations of convex minimization problems*, Appl. Math. Optim., 38 (1998), pp. 193–217.

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

[3] H. Attouch, X. Goudou, and P. Redont, *A dynamical method for the global exploration of stationary points of a real-valued mapping: The heavy ball method*, Commun. Contemp. Math., to appear.

[4] D. Bayer and J. C. Lagarias, *The nonlinear geometry of linear programming* (*parts* I *and* II), Trans. Amer. Math. Soc., 314 (1989), pp. 527–581.

[5] H. Brezis, *Monotonicity methods in Hilbert spaces and some applications to nonlinear partial differential equations*, in Contributions to Nonlinear Analysis, E. H. Zarantonello, ed., Academic Press, New York, 1971, pp. 101–151.

[6] H. Brezis, *Opérateurs Maximaux Monotones*, North-Holland Math. Stud. 5, North-Holland, Amsterdam, 1973.

[7] R. E. Bruck, *Asymptotic convergence of nonlinear contraction semi-groups in Hilbert spaces*, J. Funct. Anal., 18 (1975), pp. 15–26.

[8] R. Cominetti, *Asymptotic convergence of the steepest descent method for the exponential penalty in linear programming*, J. Convex Anal., 2 (1995), pp. 145–152.

[9] J. K. Hale, *Asymptotic Behavior of Dissipative Systems*, Math. Surveys Monogr. 25, American Mathematical Society, Providence, RI, 1988.

[10] A. Haraux, *Systèmes dynamiques dissipatifs et applications*, Rech. Math. Appl. 17, Masson, Paris, 1991.

[11] A. Haraux and M. A. Jendoubi, *Convergence of solutions of second-order gradient-like systems with analytic nonlinearities*, J. Differential Equations, 144 (1998), pp. 313–320.

[12] B. Lemaire, *About the convergence of the proximal method*, in Advances in Optimization, Proceedings Lambrecht, 1991, Lecture Notes in Econ. and Math. Systems 382, Springer-Verlag, New York, 1992, pp. 39–51.

[13] B. Lemaire, *An asymptotical variational principle associated with the steepest descent method for a convex function*, J. Convex Anal., 3 1996, pp. 63–70.

[14] G. P. McCormick, *The projective SUMT method for convex programming*, Math. Oper. Res., 14 (1989), pp. 203–223.

[15] Z. Opial, *Weak convergence of the sequence of successive approximations for nonexpansive mappings*, Bull. Amer. Math. Soc., 73 (1967), pp. 591–597.

[16] B. T. Polyack, *Some methods of speeding up the convergence of iterative methods*, Zh. Vychisl. Mat. Mat. Fiz., 4 (1964), pp. 1–17.

[17] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[18] R. T. Rockafellar, *Monotone operators and the proximal point algorithm*, SIAM J. Control Optim., 14 (1976), pp. 877–898.