ELSEVIER

J. Math. Pures Appl. 81 (2002) 747–779

JOURNAL DE MATHÉMATIQUES PURES ET APPLIQUÉES

A second-order gradient-like dissipative dynamical system with Hessian-driven damping. Application to optimization and mechanics

F. Alvarez [a], H. Attouch [b,*], J. Bolte [b], P. Redont [b]

[a] *Departamento de Ingeniería Matemática, Centro de Modelamiento Matemático, Universidad de Chile, Blanco Encalada 2120, Santiago, Chile*

[b] *ACSIOM-CNRS FRE 2311, Département de Mathématiques, case 51, Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier cedex 5, France*

Received 23 November 2001

Abstract

Given H a real Hilbert space and $\Phi : H \to \mathbb{R}$ a smooth $\mathcal{C}^2$ function, we study the dynamical inertial system

$$\text{(DIN)} \quad \ddot{x}(t) + \alpha\dot{x}(t) + \beta\nabla^2\Phi\big(x(t)\big)\dot{x}(t) + \nabla\Phi\big(x(t)\big) = 0,$$

where α and β are positive parameters. The inertial term $\ddot{x}(t)$ acts as a singular perturbation and, in fact, regularization of the possibly degenerate classical Newton continuous dynamical system $\nabla^2\Phi(x(t))\dot{x}(t) + \nabla\Phi(x(t)) = 0$.

We show that (DIN) is a well-posed dynamical system. Due to their dissipative aspect, trajectories of (DIN) enjoy remarkable optimization properties. For example, when Φ is convex and $\operatorname{argmin}\Phi \neq \emptyset$, then each trajectory of (DIN) weakly converges to a minimizer of Φ. If Φ is real analytic, then each trajectory converges to a critical point of Φ.

A remarkable feature of (DIN) is that one can produce an equivalent system which is first-order in time and with no occurrence of the Hessian, namely

$$\begin{cases} \dot{x}(t) + c\nabla\Phi\big(x(t)\big) + ax(t) + by(t) = 0, \\ \dot{y}(t) + ax(t) + by(t) = 0, \end{cases}$$

* Corresponding author.

E-mail address: attouch@math.univ-montp2.fr (H. Attouch).

[1] Partially supported by ECOS-CONICYT (C00E05), FONDAP in Applied Mathematics and FONDECYT 1990884.

PII: S0021-7824(01)01253-3

where a, b, c are parameters which can be explicitly expressed in terms of α and β. This allows to consider (DIN) when Φ is $\mathcal{C}^1$ only, or more generally, nonsmooth or subject to constraints. This is first illustrated by a gradient projection dynamical system exhibiting both viable trajectories, inertial aspects, optimization properties, and secondly by a mechanical system with impact.

Résumé

Nous étudions le système dynamique :

$$\text{(DIN)} \quad \ddot{x}(t) + \alpha\dot{x}(t) + \beta\nabla^2\Phi\big(x(t)\big)\dot{x}(t) + \nabla\Phi\big(x(t)\big) = 0,$$

où $\Phi : H \to \mathbb{R}$ est une fonctionnelle de classe $\mathcal{C}^2$, H un espace de Hilbert réel, et α, β des paramètres > 0. Le terme inertiel $\ddot{x}(t)$ peut être vu comme une perturbation singulière mais aussi une régularisation de la méthode de Newton continue $\nabla^2\Phi(x(t))\dot{x}(t) + \nabla\Phi(x(t)) = 0$.

Le système (DIN) est bien posé. La dissipativité confère aux trajectoires des propriétés intéressantes pour l'optimisation de Φ. Par exemple, si Φ est convexe et $\operatorname{argmin}\Phi \neq \emptyset$, toute trajectoire converge faiblement vers un minimum de Φ. En dimension finie, si Φ est analytique, toute trajectoire converge vers un point critique de Φ.

De façon remarquable, (DIN) est équivalent à un système du premier ordre où le hessien $\nabla^2\Phi$ ne figure pas,

$$\begin{cases} \dot{x}(t) + c\nabla\Phi\big(x(t)\big) + ax(t) + by(t) = 0, \\ \dot{y}(t) + ax(t) + by(t) = 0, \end{cases}$$

Il est donc possible de donner un sens à (DIN) losque Φ est de classe $\mathcal{C}^1$, ou même soumise à des contraintes. Nous en donnons deux illustrations : (1) un système dynamique de type gradient projeté avec des trajectoires inertielles viables et des propriétés de minimisation ; (2) une approche du rebond inélastique en mécanique.

MSC: 37Bxx; 37Cxx; 37Lxx; 37N40; 47H06

Keywords: Continuous Newton method; Dissipative dynamical systems; Asymptotic behaviour; Gradient-like dynamical systems; Optimal control; Second-order in time dynamical system; Shocks in mechanics; Gradient-projection methods

1. Introduction

Let H be a real Hilbert space and $\Phi : H \to \mathbb{R}$ a smooth function whose gradient and Hessian are respectively denoted by $\nabla\Phi$ and $\nabla^2\Phi$. Our purpose is to study the following dynamical inertial system:

$$\text{(DIN)} \quad \ddot{x}(t) + \alpha\dot{x}(t) + \beta\nabla^2\Phi\big(x(t)\big)\dot{x}(t) + \nabla\Phi\big(x(t)\big) = 0,$$

where α and β are positive parameters. We use the following notations: t is the time variable, $x \in H$ is the state variable, trajectories in H are functions $t \mapsto x(t)$ whose first and second time derivatives are respectively denoted by $\dot{x}(t)$ and $\ddot{x}(t)$.

The above dynamical system will be referred to as the *Dynamical Inertial Newton-like* system, or (DIN) for short. This evolution problem comes naturally into play in various domains like optimization (minimization of Φ), mechanics (nonelastic shocks), control theory (asymptotic stabilization of oscillators) and PDE theory (damped wave equation). The terminology reflects the fact that (DIN) is a second-order in time dynamical system, the acceleration $\ddot{x}(t)$ being associated with inertial effects, while Newton's dynamics refers to the action of the Hessian operator $\nabla^2\Phi(x(t))$ on the velocity vector $\dot{x}(t)$ (see (CN) below).

This paper focuses on the study of (DIN) as a dissipative dynamical system; accordingly, the investigation relies on Liapounov methods (for facts on dissipative systems see [17,19,30,35]). The convergence of the trajectories of (DIN), as the time t goes to $+\infty$, is established under various assumptions on Φ: Φ analytic (Theorem 4.1), Φ convex (Theorem 5.1). Indeed, by following the trajectories of (DIN) as t goes to $+\infty$, one expects to reach local minima of Φ (global minima when Φ is convex), with clear applications to optimization and mechanics.

Let us discuss some motivations for the introduction of the (DIN) system.

In recent years, numerous papers have been devoted to the study of dynamical systems that overcome some of the drawbacks of the classical steepest descent method:

$$\text{(SD)} \quad \dot{x}(t) + \nabla\Phi\big(x(t)\big) = 0.$$

For instance, Alvarez and Pérez study in [4] the *Continuous Newton* method:

$$\text{(CN)} \quad \nabla^2\Phi\big(x(t)\big)\dot{x}(t) + \nabla\Phi\big(x(t)\big) = 0$$

as a tool in optimization and show how to combine this dynamics with an approximation of Φ by smooth functions Φ_ε, when Φ is nonsmooth. On the other hand, Attouch, Goudou and Redont study in [11] the heavy ball with friction dynamical system:

$$\text{(HBF)} \quad \ddot{x}(t) + \alpha\dot{x}(t) + \nabla\Phi\big(x(t)\big) = 0,$$

where $\alpha > 0$ can be interpreted as a viscous friction parameter. This dissipative dynamical system, which was first introduced by Polyak [31] and Antipin [6] enjoys remarkable optimization properties. For example, when Φ is convex, the trajectories of (HBF) weakly converge in H as $t \to +\infty$ to minimizers of Φ. This result, proved by Alvarez in [2], may be seen as an extension of the celebrated Bruck theorem for (SD) [16] to a second-order in time differential dynamical system; see also [3] for an implicit discrete proximal version of their result.

There is a drastic difference between (SD) and (HBF). By contrast with (SD), (HBF) is no more a descent method: the function $\Phi(x(t))$ does not decrease along the trajectories in general; it is the energy $E(t) := (1/2)|\dot{x}(t)|^2 + \Phi(x(t))$ that is decreasing. This confers to this system interesting properties for the exploration of local minima of Φ, see [11] for more details.

Both the Newton and the heavy ball with friction methods can be seen as second-order extensions of (SD), the latter in time (with $\ddot{x}$ in addition to $\dot{x}$) and the former in space (with $\nabla^2\Phi$ in addition to $\nabla\Phi$). Each one improves (SD) in some respects, but they also

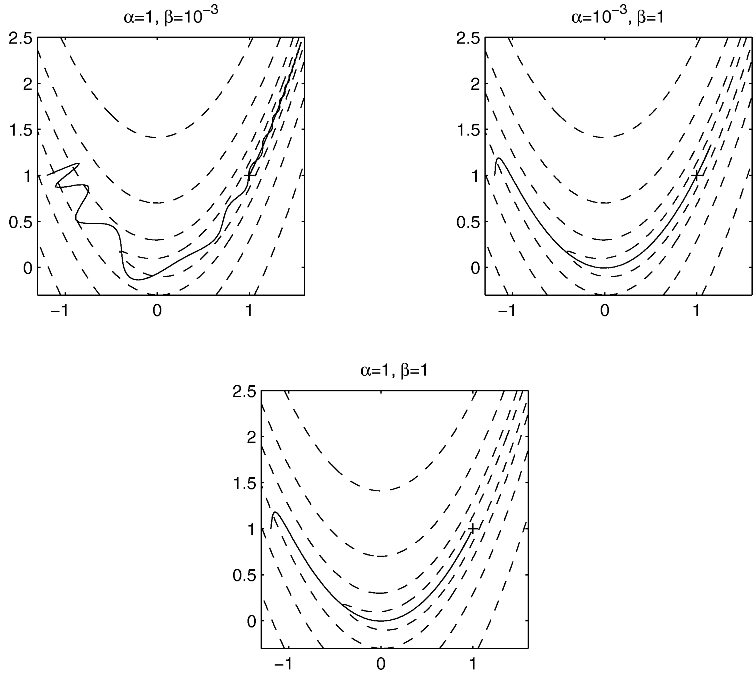

Fig. 1. Versatility of (DIN).

raise some new difficulties. In (CN), $\nabla^2\Phi(x(t))$ may be degenerate and (CN) is no more defined as a dynamical system, moreover, $\nabla^2\Phi(x(t))$ may be complicated to compute. In (HBF), the trajectories may exhibit oscillations which are not desirable for a numerical optimization purpose.

If one combines the continuous Newton dynamical system with the heavy ball with friction system, the system so obtained,

$$\text{(DIN)}\quad \ddot{x} + \alpha\dot{x} + \beta\nabla^2\Phi(x)\dot{x} + \nabla\Phi(x) = 0,$$

inherits most of the advantages of the two preceding systems and corrects both of the above-mentioned drawbacks: the term $\nabla^2\Phi(x(t))\dot{x}(t)$ is a clever geometric damping term, while the acceleration term $\ddot{x}(t)$ makes (DIN) a well-posed dynamical system, even if $\nabla^2\Phi(x(t))$ is degenerate; see Attouch and Redont [12] for a first study of this question.

The relative roles of the damping terms $\alpha\dot{x}$ and $\beta\nabla^2\Phi(x)\dot{x}$ are illustrated on Rosenbrock's function, $\Phi(x_1,x_2) = 100(x_2 - x_1^2)^2 + (1-x_1)^2$, which possesses a global minimum at point $(1,1)$ at the bottom of a flat long winding valley; see Fig. 1. When the geometric damping is low ($\beta = 10^{-3}$) the trajectory is prone to large oscillations, transversal to the valley axis, and is quite similar to a (HBF) trajectory ($\beta = 0$, see [11]). When the geometric damping is effective ($\beta = 1$), but with a low viscous damping ($\alpha = 10^{-3}$), the trajectory is forced to the bottom of the valley. While transversal oscillations are suppressed, longitudinal oscillations remain important, due to the Hessian being nearly zero in the direction of the valley. As can be seen in the lower plot, a combination of viscous and geometric damping ($\alpha = 1$, $\beta = 1$) puts down any oscillations and produces a trajectory converging regularly to the minimum.

We stress the fact that (DIN) is a second-order system both in time (because of the acceleration term $\ddot{x}(t)$) and in space ($\nabla^2\Phi(x(t))$ is the Hessian). The central point of this paper is that, surprisingly, one can "integrate" in some sense this system, and exhibit an equivalent first-order system *in time and space* in $H \times H$ which involves no Hessian (Section 6.3, Theorem 6.2):

$$\begin{cases} \dot{x}(t) + c\nabla\Phi\big(x(t)\big) + ax(t) + by(t) = 0, \\ \dot{y}(t) + ax(t) + by(t) = 0. \end{cases}$$

This result opens new interesting perspectives: it allows to consider (DIN) for nonsmooth functions, possibly only lower semicontinuous or involving constraints, with clear applications to mechanics and PDEs (wave equations, shocks). For example, when taking $H = L^2(\Omega)$ and Φ being equal to the Dirichlet integral with domain $H_0^1(\Omega)$, the system (DIN) provides the following wave equation with higher-order damping, which has been considered by Aassila in [1]:

$$\begin{cases} \dfrac{\partial^2 u}{\partial t^2} + \alpha\dfrac{\partial u}{\partial t} - \beta\Delta\left(\dfrac{\partial u}{\partial t}\right) - \Delta u = 0 & \text{in } \Omega \times]0, +\infty[, \\ u = 0 & \text{on } \partial\Omega \times]0, +\infty[, \\ u(0) = u_0, \quad \dfrac{\partial u}{\partial t}(0) = u_1 & \text{in } \Omega. \end{cases}$$

Another interesting situation corresponds to the case where Φ is proportional to the square of the distance function to a convex set K: $\Phi(x) = \Psi_{K,\lambda}(x) = (1/(2\lambda))\operatorname{dist}^2(x, K)$, $\lambda > 0$ (which is also the Moreau–Yosida approximation of the indicator function of K). In that case (DIN), written under the form

$$\ddot{x}_\lambda + 2\varepsilon\sqrt{\lambda}\,\nabla^2\Psi_{K,\lambda}(x)\dot{x}_\lambda + \nabla\Psi_{K,\lambda}(x) = -\alpha\dot{x}_\lambda,$$

is closely related to a dynamical system introduced by Paoli and Schatzman [28] to model nonelastic shocks in mechanics.

Let us finally mention that the formulation of (DIN) as a first-order dynamical system which only involves the gradient of Φ, naturally suggests a way to define the second-order subdifferential $\partial^2\Phi$ of nonsmooth functions Φ. It is certainly worthwile comparing this new aproach to $\partial^2\Phi$ *via* dynamical systems, with the recent studies of R.T. Rockafellar [32], Mordukhovich–Outrata [26] and Kummer [22].

Clearly, a precise study of these quite involved questions is out of the scope of the present article. We just mention them in order to stress the importance and the versatility of the (DIN) system.

The paper is organized as follows. Section 2 gives the existence and the basic properties of the solution to (DIN). In Section 3, we justify the terminology *Dynamical Inertial Newton* method by showing that (DIN) may be considered as a perturbation of the continuous Newton method. The next two sections deal with the asymptotic behaviour of the (DIN) trajectories: convergence to a critical point is proved for an analytic function Φ (Section 4), and convergence to a minimizer is proved for a convex function (Section 5). Section 6 presents a first-order in time and space system that is equivalent to (DIN). In Section 7, constraints are introduced in that new system, which gives rise to a continuous gradient-projection system; the trajectories are shown to be viable and to enjoy optimizing properties. Section 8 concludes the paper with an illustration in impact dynamics.

2. Global existence

Throughout this paper, H is a real Hilbert space with scalar product and norm denoted by $\langle\cdot,\cdot\rangle$ and $|\cdot|$, respectively. Let $\Phi : H \to \mathbb{R}$ be a mapping satisfying:

$$(\mathcal{H})\quad \begin{cases} \Phi \text{ is bounded from below on } H, \\ \Phi \text{ is twice continuously differentiable on } H, \\ \text{the Hessian } \nabla^2\Phi \text{ is Lipschitz continuous on the bounded subsets of } H. \end{cases}$$

Given two parameters $\alpha > 0$ and $\beta > 0$, consider the following second-order in time system in H:

$$\text{(DIN)}\quad \ddot{x} + \alpha\dot{x} + \beta\nabla^2\Phi(x)\dot{x} + \nabla\Phi(x) = 0.$$

Along every trajectory of (DIN) and for $\lambda > 0$ define:

$$E_\lambda(t) = \lambda\Phi\big(x(t)\big) + \frac{1}{2}\big|\dot{x}(t) + \beta\nabla\Phi\big(x(t)\big)\big|^2. \tag{1}$$

In particular, we will write for short

$$E(t) = E_{\alpha\beta+1}(t) = (\alpha\beta+1)\Phi\big(x(t)\big) + \frac{1}{2}\big|\dot{x}(t) + \beta\nabla\Phi\big(x(t)\big)\big|^2. \tag{2}$$

Theorem 2.1. *Let Φ satisfy $(\mathcal{H})$. Then the following properties hold for* (DIN)*, provided $\alpha > 0$ and $\beta > 0$*:

(i) *For each $(x_0, \dot{x}_0) \in H \times H$, there exists a unique global solution $x(t)$ of* (DIN) *satisfying the initial conditions $x(0) = x_0$ and $\dot{x}(0) = \dot{x}_0$, with $x \in \mathcal{C}^2([0,+\infty[; H)$.*

(ii) *For every trajectory $x(t)$ of* (DIN) *and $\lambda \in [(1-\sqrt{\alpha\beta})^2, (1+\sqrt{\alpha\beta})^2]$, the scalar function E_λ defined by* (1) *is bounded from below and decreasing on* $[0,+\infty[$*, hence, it converges as $t \to +\infty$. Moreover,*
- *$\dot{x}$ and $\nabla\Phi(x)$ belong to $L^2(0,+\infty; H)$;*
- *$\lim_{t\to+\infty}\Phi(x(t))$ exists;*
- $\lim_{t\to+\infty}(\dot{x}(t) + \beta\nabla\Phi x(t)) = 0$.

(iii) *Assuming, moreover, that $x \in L^\infty(0,+\infty; H)$, we have*:
- *$\dot{x}$, $\ddot{x}$, $\nabla\Phi(x)$ and $\nabla^2\Phi(x)$ are bounded on* $[0,+\infty[$;
- $\lim_{t\to+\infty}\nabla\Phi(x(t)) = \lim_{t\to+\infty}\dot{x}(t) = \lim_{t\to+\infty}\ddot{x}(t) = 0$.

Proof. (i) For any choice of initial conditions $(x_0, \dot{x}_0) \in H \times H$, the existence and uniqueness of a classic local solution to (DIN) follow from the Cauchy–Lipschitz theorem applied to the equivalent first-order in time system in the phase space $H \times H$, $\dot{Y} = F(Y)$, with

$$Y(t) = \begin{pmatrix} x(t) \\ \dot{x}(t) \end{pmatrix} \quad \text{and} \quad F(u, v) = \begin{pmatrix} v \\ -\alpha v - \beta \nabla^2 \Phi(u) v - \nabla \Phi(u) \end{pmatrix}.$$

Let x denote the maximal solution defined on some interval $[0, T_{\max}[$ with $0 < T_{\max} \leqslant +\infty$. The regularity assumptions on Φ imply that $x \in \mathcal{C}^2([0, T_{\max}[; H)$. Suppose, contrary to our claim, that $T_{\max} < +\infty$. Differentiating $E(t)$ (see (2)) and using (DIN), we successively obtain:

$$\begin{aligned} \dot{E}(t) &= (\alpha\beta + 1)\big\langle \nabla \Phi\big(x(t)\big), \dot{x}(t)\big\rangle + \big\langle \ddot{x}(t) + \beta \nabla^2 \Phi\big(x(t)\big)\dot{x}(t), \dot{x}(t) + \beta \nabla \Phi\big(x(t)\big)\big\rangle \\ &= (\alpha\beta + 1)\big\langle \nabla \Phi\big(x(t)\big), \dot{x}(t)\big\rangle - \big\langle \alpha \dot{x}(t) + \nabla \Phi\big(x(t)\big), \dot{x}(t) + \beta \nabla \Phi\big(x(t)\big)\big\rangle \\ &= -\alpha\big|\dot{x}(t)\big|^2 - \beta\big|\nabla \Phi\big(x(t)\big)\big|^2. \end{aligned} \tag{3}$$

Hence, $E(t)$ is a Liapounov function for the trajectory x. Further, for all $t \in [0, T_{\max}[$,

$$\begin{aligned} &(\alpha\beta + 1)\Phi\big(x(t)\big) + \frac{1}{2}\big|\dot{x}(t) + \beta \nabla \Phi\big(x(t)\big)\big|^2 + \alpha \int_0^t \big|\dot{x}(\tau)\big|^2 \,\mathrm{d}\tau \\ &\quad + \beta \int_0^t \big|\nabla \Phi\big(x(\tau)\big)\big|^2 \,\mathrm{d}\tau = E(0). \end{aligned} \tag{4}$$

Since Φ is bounded from below and $\alpha, \beta > 0$, we obtain that $\dot{x}$ and $\nabla \Phi(x)$ belong to $L^2(0, T_{\max}; H)$. Therefore, for all $0 \leqslant s \leqslant t < T_{\max}$,

$$\big|x(t) - x(s)\big| \leqslant \int_s^t \big|\dot{x}(\tau)\big| \,\mathrm{d}\tau \leqslant \sqrt{t - s}\sqrt{\int_s^t \big|\dot{x}(\tau)\big|^2 \,\mathrm{d}\tau} \leqslant \sqrt{t - s}\, \|\dot{x}\|_{L^2(0, T_{\max}; H)},$$

which shows that $\lim_{t \to T_{\max}} x(t)$ exists. As a consequence, x is bounded on $[0, T_{\max}[$ and so is $\nabla^2 \Phi(x)$ in view of the Lipschitz continuity of $\nabla^2 \Phi$. Thus

$$\ddot{x} = -\alpha \dot{x} - \beta \nabla^2 \Phi(x) \dot{x} - \nabla \Phi(x)$$

belongs to $L^2(0, T_{\max}; H)$, and we have for all $0 \leqslant s \leqslant t < T_{\max}$:

$$\big|\dot{x}(t) - \dot{x}(s)\big| \leqslant \int_s^t \big|\ddot{x}(\tau)\big| \,\mathrm{d}\tau \leqslant \sqrt{t - s}\, \|\ddot{x}\|_{L^2(0, T_{\max}; H)},$$

so that $\lim_{t\to T_{\max}} \dot{x}(t)$ exists. Applying the Cauchy–Lipschitz local existence theorem to (DIN) with initial data at $T_{\max}$ given by $(\lim_{t\to T_{\max}} x(t), \lim_{t\to T_{\max}} \dot{x}(t))$, we can extend the maximal solution to an interval strictly larger than $[0, T_{\max}[$, which contradicts the maximality of the solution. Consequently, $T_{\max} = +\infty$.

(ii) The point here is to realize that there is a whole family of Liapounov functions for the trajectory x. Indeed, setting for short (recall (1))

$$E_{\pm}(t) = E_{1\pm\sqrt{\alpha\beta}} = \bigl(1 \pm \sqrt{\alpha\beta}\,\bigr)^2 \Phi\bigl(x(t)\bigr) + \frac{1}{2}\bigl|\dot{x}(t) + \beta\nabla\Phi\bigl(x(t)\bigr)\bigr|^2,$$

we obtain:

$$\dot{E}_{\pm}(t) = -\bigl|\sqrt{\alpha}\dot{x}(t) \mp \sqrt{\beta}\,\nabla\Phi\bigl(x(t)\bigr)\bigr|^2.$$

Hence, E_+ and E_- are two Liapounov functions for x, as well as any convex combination of them. As a result, for any λ in $[(1-\sqrt{\alpha\beta}\,)^2, (1+\sqrt{\alpha\beta}\,)^2]$, E_λ is decreasing on $[0, +\infty[$, (e.g., $E = E_{\alpha\beta+1} = (1/2)(E^+ + E^-)$). Further we have:

$$\begin{aligned}&\bigl(1 \pm \sqrt{\alpha\beta}\,\bigr)^2 \Phi\bigl(x(t)\bigr) + \frac{1}{2}\bigl|\dot{x}(t) + \beta\nabla\Phi\bigl(x(t)\bigr)\bigr|^2 - E_{\pm}(0)\\ &\quad= -\int_0^t \bigl|\sqrt{\alpha}\dot{x}(\tau) \mp \sqrt{\beta}\,\nabla\Phi\bigl(x(\tau)\bigr)\bigr|^2\,\mathrm{d}\tau.\end{aligned}$$

Since Φ is bounded from below, we obtain that both

$$\bigl|\sqrt{\alpha}\,\dot{x} - \sqrt{\beta}\,\nabla\Phi(x)\bigr| \quad \text{and} \quad \bigl|\sqrt{\alpha}\,\dot{x} + \sqrt{\beta}\,\nabla\Phi(x)\bigr|$$

belong to $L^2(0, +\infty)$ and hence $\dot{x}$ and $\nabla\Phi(x)$ are in $L^2(0, +\infty; H)$. Now, since E_+ and E_- are decreasing and bounded from below, $\lim_{t\to+\infty} E_+(t)$ and $\lim_{t\to+\infty} E_-(t)$ exist. Therefore, $\Phi(x(t)) = (1/(4\sqrt{\alpha\beta}\,))(E_+(t) - E_-(t))$ admits a limit as $t \to +\infty$. As a consequence, $|\dot{x}(t) + \beta\nabla\Phi(x(t))|$ has a limit as $t \to +\infty$, which is zero because $|\dot{x}(t) + \beta\nabla\Phi(x(t))| \in L^2(0, +\infty)$.

(iii) We now assume that x is in $L^\infty(0, +\infty; H)$. Then, by $(\mathcal{H})$, $\nabla^2\Phi(x)$ and $\nabla\Phi(x)$ are bounded on $[0, +\infty[$; and so are $\dot{x} = (\dot{x} + \beta\nabla\Phi(x)) - \beta\nabla\Phi(x)$ and $\ddot{x} = -\alpha\dot{x} - \beta\nabla^2\Phi(x)\dot{x} - \nabla\Phi(x)$. Set $h(t) = (1/2)|\nabla\Phi(x(t))|^2$ and note that $h \in L^1(0, +\infty)$ and $\dot{h} = \langle\nabla^2\Phi(x)\dot{x}, \nabla\Phi(x)\rangle \in L^\infty(0, +\infty)$; then, by a standard argument, $\lim_{t\to+\infty} h(t) = 0$. Likewise, if we set $k(t) = (1/2)|\dot{x}(t)|^2$ then $\lim_{t\to+\infty} k(t) = 0$. It follows that $\ddot{x}(t) \to 0$ as $t \to +\infty$. □

Corollary 2.1. *Assume that $\Phi : H \to \mathbb{R}$ satisfies $(\mathcal{H})$ and is coercive, i.e.* $\lim_{|x|\to+\infty} \Phi(x) = +\infty$*. Then the solution x of* (DIN) *is in $L^\infty(0, +\infty; H)$. In particular, the properties in Theorem* 2.1(iii) *hold.*

Proof. It suffices to observe that (4) gives $(\alpha\beta + 1)\Phi(x(t)) \leqslant E(0)$. This estimate and the coerciveness of Φ imply that the trajectory x remains bounded. □

3. (DIN) as a singular perturbation of Newton's method

In this section we assume that Φ belongs to $\mathcal{C}^2(H)$, with a Hessian Lipschitz continuous on bounded subsets, and that Φ is coercive with $\nabla\Phi$ strongly monotone on bounded subsets of H. More precisely, it is required that $\forall R > 0$, $\exists \beta_R > 0$ such that $\forall x, y \in H$,

$$\max\{|x|, |y|\} < R \quad \Rightarrow \quad \langle \nabla\Phi(x) - \nabla\Phi(y), x - y \rangle \geqslant \beta_R |x - y|^2. \tag{5}$$

In particular, Φ is strictly convex and for all $x \in H$ the Hessian operator $\nabla^2\Phi(x)$ is positive definite. Indeed, (5) yields $\forall R > 0$, $\exists \beta_R > 0$: $\forall x \in H$, if $|x| < R$ then $\forall h \in H$, $\langle \nabla^2\Phi(x)h, h\rangle \geqslant \beta_R |h|^2$. On the other hand, when $H = \mathbb{R}^n$ and $\nabla^2\Phi(x)$ is positive definite for every $x \in \mathbb{R}^n$, (5) holds with β_R being a positive lower bound for the eigenvalues of $\nabla^2\Phi(x)$ over the ball $B(0, R)$.

For simplicity, take $\alpha = 0$ and $\beta = 1$ and, for each $\varepsilon > 0$, consider a solution $x_\varepsilon \in C^2([0, \infty[; H)$ to the initial value problem (x_ε does exist, see [12]),

$$(\varepsilon\text{-DIN}) \quad \begin{cases} \varepsilon \ddot{x}_\varepsilon + \nabla^2\Phi(x_\varepsilon)\dot{x}_\varepsilon + \nabla\Phi(x_\varepsilon) = 0, & t > 0, \\ x_\varepsilon(0) = x_0, \quad \dot{x}_\varepsilon(0) = \dot{x}_0, \end{cases}$$

where $x_0, \dot{x}_0 \in H$ are given. We are interested in the asymptotic behaviour of x_ε as $\varepsilon \to 0$. Observe that (ε-DIN) may be considered as a singular perturbation of the following evolution equation:

$$(\text{CN}) \quad \begin{cases} \nabla^2\Phi(x)\dot{x} + \nabla\Phi(x) = 0, & t > 0, \\ x(0) = x_0. \end{cases}$$

This is the *Continuous Newton* method for the minimization of Φ, which is a continuous version of the well-known Newton iteration:

$$\nabla^2\Phi(x^k)(x^{k+1} - x^k) + \nabla\Phi(x^k) = 0.$$

The unique solution $x \in C^2([0, \infty[; H)$ of (CN) satisfies:

$$\frac{\mathrm{d}}{\mathrm{d}t}\big[\nabla\Phi(x(t))\big] = -\nabla\Phi(x(t)),$$

which yields the following remarkable property of Newton's trajectories:

$$\nabla\Phi(x(t)) = \mathrm{e}^{-t}\nabla\Phi(x_0). \tag{6}$$

Moreover, since Φ is coercive, it follows from (5) and (6) that for an appropriate $\beta_R > 0$, $|x(t) - \widehat{x}| \leqslant (\mathrm{e}^{-t}/\beta_R)|\nabla\Phi(x_0)|$, where $\widehat{x}$ is the unique minimizer of Φ. We refer the reader to [4,13,34] for fuller treatments of the continuous Newton method.

Proposition 3.1. *There exists a constant $C > 0$ such that $\forall t \geqslant 0$, $|x_\varepsilon(t) - x(t)| \leqslant C\sqrt{\varepsilon}$. Therefore, $x_\varepsilon \to x$ uniformly on $[0, +\infty[$.*

Proof. Let us introduce the ε-energy

$$U_\varepsilon(t) := \frac{\varepsilon}{2}\left|\dot{x}_\varepsilon(t)\right|^2 + \Phi\left(x_\varepsilon(t)\right),$$

which satisfies

$$\dot{U}_\varepsilon(t) = -\left\langle \nabla^2\Phi\left(x_\varepsilon(t)\right)\dot{x}_\varepsilon(t), \dot{x}_\varepsilon(t)\right\rangle \leqslant 0.$$

Hence,

$$U_\varepsilon(t) \leqslant U_\varepsilon(0) = \frac{\varepsilon}{2}|\dot{x}_0|^2 + \Phi(x_0), \tag{7}$$

and consequently

$$\sup_{0<\varepsilon\leqslant 1}\sup_{t\geqslant 0}\Phi\left(x_\varepsilon(t)\right) \leqslant \frac{1}{2}|\dot{x}_0|^2 + \Phi(x_0) =: \alpha.$$

Since Φ is coercive, the sublevel set $\Gamma_\alpha(\Phi) := \{x \in H\colon\ \Phi(x) \leqslant \alpha\}$ is bounded and then $\sup_{0<\varepsilon\leqslant 1}\sup_{t\geqslant 0}|x_\varepsilon(t)| < R$ for a suitable constant $R > 0$. Similarly, we obtain that the solution $x(t)$ of (CN) satisfies $\{x(t)\colon\ t \geqslant 0\} \subset \Gamma_{\Phi(x_0)}(\Phi) \subset \Gamma_\alpha(\Phi)$, so that we may assume that $\sup_{t\geqslant 0}|x(t)| < R$. By (5), we have

$$\forall t > 0, \quad \left|x_\varepsilon(t) - x(t)\right| \leqslant \frac{1}{\beta_R}\left|\nabla\Phi\left(x_\varepsilon(t)\right) - \nabla\Phi\left(x(t)\right)\right|. \tag{8}$$

Notice that the differential equation in (ε-DIN) may be rewritten:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\varepsilon\dot{x}_\varepsilon(t) + \nabla\Phi\left(x_\varepsilon(t)\right)\right] + \nabla\Phi\left(x_\varepsilon(t)\right) = 0.$$

Setting $\omega_\varepsilon(t) := \varepsilon\dot{x}_\varepsilon(t) + \nabla\Phi(x_\varepsilon(t))$, we obtain the nonhomogeneous initial value problem:

$$\begin{cases} \dot{\omega}_\varepsilon + \omega_\varepsilon = \varepsilon\dot{x}_\varepsilon(t), & t > 0, \\ \omega_\varepsilon(0) = \varepsilon\dot{x}_0 + \nabla\Phi(x_0), \end{cases}$$

whose solution is given by:

$$\omega_\varepsilon(t) = \mathrm{e}^{-t}\left(\varepsilon\dot{x}_0 + \nabla\Phi(x_0)\right) + \varepsilon\int_0^t \mathrm{e}^{-(t-\tau)}\dot{x}_\varepsilon(\tau)\,\mathrm{d}\tau.$$

Thus

$$\nabla\Phi\left(x_\varepsilon(t)\right) = \mathrm{e}^{-t}\left(\varepsilon\dot{x}_0 + \nabla\Phi(x_0)\right) - \varepsilon\dot{x}_\varepsilon(t) + \varepsilon\int_0^t \mathrm{e}^{-(t-\tau)}\dot{x}_\varepsilon(\tau)\,\mathrm{d}\tau.$$

By (6) together with (8), we have:

$$\left|x_\varepsilon(t)-x(t)\right| \leqslant \frac{1}{\beta_R}\left(\varepsilon|\dot{x}_0|+\varepsilon\left|\dot{x}_\varepsilon(t)\right|+\int_0^t \mathrm{e}^{-(t-\tau)}\varepsilon\left|\dot{x}_\varepsilon(\tau)\right|\mathrm{d}\tau\right).$$

On the other hand, from the energy estimate (7), it follows that $\sup_{0<\varepsilon\leqslant 1}\sup_{t\geqslant 0}\varepsilon|\dot{x}_\varepsilon(t)| \leqslant \sqrt{2\varepsilon(\alpha-\inf\Phi)}$. Consequently,

$$\left|x_\varepsilon(t)-x(t)\right| \leqslant \frac{1}{\beta_R}\big(\varepsilon|\dot{x}_0|+2\sqrt{2\varepsilon(\alpha-\inf\Phi)}\big) \leqslant \frac{\sqrt{\varepsilon}}{\beta_R}\big(|\dot{x}_0|+2\sqrt{2(\alpha-\inf\Phi)}\big),$$

which completes the proof. □

4. Convergence of the trajectories: Φ analytic

Since $\lim_{t\to+\infty}\nabla\Phi(x(t))=0$, it is natural to expect that for a sufficiently smooth Φ, trajectories will converge towards a critical point of that function. Actually we show, in the finite-dimensional case, that if Φ is real analytic, x will finally converge to $x_\infty\in H$, with $\nabla\Phi(x_\infty)=0$. The proof of this convergence result relies on an inequality due to Lojasiewicz [25], linking Φ and $\nabla\Phi$ in a neighbourhood of critical points. Lojasiewicz applied it in [24] to study the asymptotic behaviour of a gradient-like system. More recently, Haraux and Jendoubi [20] showed that bounded trajectories of HBF with an analytic potential converge towards critical points. This analyticity hypothesis is also useful for infinite dimensional systems with analytic nonlinearities, see Simon's work [33] for the heat equation and Haraux [18] and Jendoubi [21] for the damped wave equation.

Let us recall the definition of a real analytic function.

Definition 4.1. Let Ω be an open subset of $\mathbb{R}^N$. A function $\Phi:\Omega\mapsto\mathbb{R}$ is real analytic (in Ω), if for every point $\xi=(\xi_1,\ldots,\xi_N)$ in Ω there exist a neighbourhood $U\subseteq\Omega$ of ξ and real coefficients $(c_{\nu_1,\ldots,\nu_N})_{(\nu_1,\ldots,\nu_N)\in\mathbb{N}^N}$ such that

$$\begin{aligned} &x=(x_1,\ldots,x_N)\in U\\ &\quad\Rightarrow\quad \Phi(x)=\sum_{(\nu_1,\ldots,\nu_N)\in\mathbb{N}^N} c_{\nu_1,\ldots,\nu_N}(x_1-\xi_1)^{\nu_1}\cdots(x_N-\xi_N)^{\nu_N}.\end{aligned}$$

Lemma 4.1 (Lojasiewicz). *Let $\Phi:\mathbb{R}^N\to\mathbb{R}$ be a function which is supposed to be analytic in a neighbourhood of a critical point a. Then, there exist $\sigma>0$ and $\theta\in\,]0,1/2[$ such that*[2]

[2] Originally [25, p. 92], the lemma states that θ lies in $]0,1[$; but it is harmless to suppose that σ satisfies $|x-a|<\sigma\Rightarrow|\Phi(x)-\Phi(a)|\leqslant 1$, which, together with $0<\theta<1$, entails $|\Phi(x)-\Phi(a)|^{1-\theta/2}\leqslant|\Phi(x)-\Phi(a)|^{1-\theta}$; this justifies the assertion $\theta\in\,]0,1/2[$.

$$|x-a|<\sigma \quad\Rightarrow\quad \bigl|\Phi(x)-\Phi(a)\bigr|^{1-\theta}\leqslant\bigl|\nabla\Phi(x)\bigr|.$$

The next corollary extends the lemma to a compact connected set of critical points.

Corollary 4.1. *Let $\Phi:\Omega\subseteq\mathbb{R}^N\to\mathbb{R}$ be a function which is supposed to be analytic in the open set Ω. Let A be a nonempty subset of Ω such that $\nabla\Phi(a)=0$, for all a in A:*

(1) *if A is connected then Φ assumes a constant value on A, say Φ_A;*
(2) *if A is connected and compact, then there exist $\sigma>0$ and $\theta\in\,]0,1/2[$ such that*

$$\operatorname{dist}(x,A)<\sigma \quad\Rightarrow\quad \bigl|\Phi(x)-\Phi_A\bigr|^{1-\theta}\leqslant\bigl|\nabla\Phi(x)\bigr|.$$

Proof. (1) Pick some a in A. After the lemma there exist $\sigma>0$ and $\theta\in\,]0,1/2[$ such that

$$|x-a|<\sigma \quad\Rightarrow\quad \bigl|\Phi(x)-\Phi(a)\bigr|^{1-\theta}\leqslant\bigl|\nabla\Phi(x)\bigr|.$$

Hence, if x belongs to $A\cap B(a,\sigma)$ where $B(a,\sigma)$ is the open ball with center a and radius σ, then $|\Phi(x)-\Phi(a)|=0$. As a consequence, the set $\{x\in A/\Phi(x)=\Phi(a)\}$ is open in A; as it is obviously closed in A and nonvoid it is equal to A.

(2) Without restriction we may assume that Φ vanishes on A. According to Lojasiewicz's lemma and owing to the compactness of A, there exists a finite family $(a_i,\sigma_i,\theta_i)_{i\in\{1,\dots,n\}}$ with $a_i\in A$, $\sigma_i>0$, $\theta_i\in\,]0,1/2[$ such that

– the balls $B(a_i,\sigma_i)$, build a finite open cover of A;
– $x\in\Omega$, $|x-a_i|<\sigma_i\Rightarrow|\Phi(x)|^{1-\theta_i}\leqslant|\nabla\Phi(x)|$.

Resorting once more to the compactness of A, and to the continuity of Φ, we assert the existence of some $\sigma>0$ such that

$$\operatorname{dist}(x,A)<\sigma \quad\Rightarrow\quad x\in\Omega,\quad x\in\bigcup_{i=1}^{n}B(a_i,\sigma_i),\quad \bigl|\Phi(x)\bigr|\leqslant 1.$$

If we set $\theta=\min\theta_i$, then any x complying with $\operatorname{dist}(x,A)<\sigma$ verifies $x\in\Omega$ and $x\in B(a_i,\sigma_i)$ for some $i\in\{1,\dots,n\}$; hence, $|\Phi(x)|^{1-\theta}\leqslant|\Phi(x)|^{1-\theta_i}\leqslant|\nabla\Phi(x)|$. □

Theorem 4.1. *Let x be a bounded solution of* (DIN) *and assume that $\Phi:\mathbb{R}^N\mapsto\mathbb{R}$ is analytic. Then $\dot{x}$ belongs to $L^1(0,+\infty;H)$ and $x(t)$ converges towards a critical point of Φ as $t\to\infty$.*

Proof. Let $\omega(x)$ denote the ω-limit set of x. Classically ([19], e.g.), $\omega(x)$ is a compact connected set which consists of critical points of Φ. Moreover, from Theorem 2.1(ii), Φ assumes a constant value on $\omega(x)$, which we may suppose to be 0. Further, $\operatorname{dist}(x(t),\omega(x))\to 0$ as $t\to\infty$.

After Corollary 4.1, there exist some $T>0$ and some $\theta \in]0,1/2[$ such that

$$t \geqslant T \quad \Rightarrow \quad \big|\Phi\big(x(t)\big)\big|^{1-\theta} \leqslant \big|\nabla\Phi\big(x(t)\big)\big|. \tag{9}$$

The proof of the convergence of x relies on the equality

$$-\frac{\mathrm{d}}{\mathrm{d}t}E(t)^{\theta} = -\dot{E}(t)E(t)^{\theta-1}$$

and on lower bounds for $-\dot{E}(t)$ and $E(t)^{\theta-1}$ involving $|\dot{x}(t)|$; recall that the energy E is defined by (2).

First, we have (recall (3)),

$$-\dot{E}(t) \geqslant \frac{1}{2}\min(\alpha,\beta)\big\{\big|\dot{x}(t)\big| + \big|\nabla\Phi\big(x(t)\big)\big|\big\}^2. \tag{10}$$

Further, for $C = \max(\alpha\beta+1, \beta^2)$, we have (recall (2)),

$$E(t) \leqslant C\big\{\big|\Phi\big(x(t)\big)\big| + \big|\dot{x}(t)\big|^2 + \big|\nabla\Phi\big(x(t)\big)\big|^2\big\}.$$

Hence (using the inequality $(r+s)^{1-\theta} \leqslant r^{1-\theta} + s^{1-\theta}$),

$$E(t)^{1-\theta} \leqslant C^{1-\theta}\big\{\big|\Phi\big(x(t)\big)\big|^{1-\theta} + \big|\dot{x}(t)\big|^{2(1-\theta)} + \big|\nabla\Phi\big(x(t)\big)\big|^{2(1-\theta)}\big\}.$$

Using (9), we have for $t \geqslant T$:

$$E(t)^{1-\theta} \leqslant C^{1-\theta}\big\{\big|\nabla\Phi\big(x(t)\big)\big| + \big|\dot{x}(t)\big|^{2(1-\theta)} + \big|\nabla\Phi\big(x(t)\big)\big|^{2(1-\theta)}\big\}.$$

Since $|\nabla\Phi(x(t))|$ and $|\dot{x}(t)|$ tend to zero as $t \to \infty$ and since $2(1-\theta) > 1$, the quantities $|\nabla\Phi(x(t))|^{2(1-\theta)}$ and $|\dot{x}(t)|^{2(1-\theta)}$ are negligible with respect to $|\nabla\Phi(x(t))|$ and $|\dot{x}(t)|$. Therefore, there is some constant $D>0$ such that, for $t \geqslant T$,

$$E(t)^{1-\theta} \leqslant D\big\{\big|\nabla\Phi\big(x(t)\big)\big| + \big|\dot{x}(t)\big|\big\}. \tag{11}$$

If $|\nabla\Phi(x(t))| + |\dot{x}(t)|$ happens to vanish at some time $t_1 \geqslant T$, then owing to the unicity of the solution to (DIN), $x(t)$ is equal to $x(t_1)$ for $t \geqslant t_1$, and the theorem is proved.

Else from (10) and (11) we obtain for $t \geqslant T$:

$$-\frac{\mathrm{d}}{\mathrm{d}t}E(t)^{\theta} \geqslant \frac{1}{2D}\min(\alpha,\beta)\big\{\big|\nabla\Phi\big(x(t)\big)\big| + \big|\dot{x}(t)\big|\big\}.$$

Since $\lim_{t\to\infty} E(t)$ exists, $|\dot{x}|$ belongs to $L^1([0,+\infty[)$ and consequently $\lim_{t\to\infty} x(t)$ exists. □

5. Convergence of the trajectories: Φ convex

5.1. Weak convergence in the general convex case

The proof of the asymptotic convergence in the convex case relies on the following lemma, which is essentially due to Opial [27].

Lemma 5.1 (Opial). *Let H be a Hilbert space and $x : [0, +\infty[\mapsto H$ a function such that there exists a nonempty set $S \subseteq H$ verifying*:

(a) *if $x(t_n) \rightharpoonup \bar{x}$ weakly in H for some $t_n \to +\infty$ then $\bar{x} \in S$*;
(b) $\forall z \in S$, $\lim_{t\to+\infty} |x(t) - z|$ *exists.*

Then, $x(t)$ weakly converges as $t \to +\infty$ to an element of S.

Theorem 5.1. *Let Φ be a convex function satisfying* $(\mathcal{H})$ *and assume that* $\operatorname{Argmin} \Phi \neq \emptyset$. *Let x be a solution of* (DIN). *Then for all $z \in \operatorname{Argmin} \Phi$, $\lim_{t\to+\infty} |x(t) - z|$ exists, and $x(t)$ weakly converges to a minimum point of Φ as $t \to +\infty$.*

Proof. Write $S = \operatorname{Argmin} \Phi$ and pick some z in S. In order to prove the existence of $\lim_{t\to+\infty} |x(t) - z|$, we introduce an auxiliary energy:

$$E_\varepsilon(t) = E(t) + \varepsilon\left(\frac{\alpha}{2}\big|x(t) - z\big|^2 + \big\langle \dot{x}(t) + \beta\nabla\Phi\big(x(t)\big), x(t) - z\big\rangle\right), \tag{12}$$

where E is the energy defined by (2) and ε is a positive parameter. Let us show that, by choosing ε small enough, E_ε is a Liapounov function for (DIN). Using (DIN) and (3), we have:

$$\begin{aligned}\dot{E}_\varepsilon(t) \;=\; &-(\alpha - \varepsilon)\big|\dot{x}(t)\big|^2 - \beta\big|\nabla\Phi\big(x(t)\big)\big|^2 \\ &- \varepsilon\big\langle\nabla\Phi\big(x(t)\big), x(t) - z\big\rangle + \varepsilon\big\langle\beta\nabla\Phi\big(x(t)\big), \dot{x}(t)\big\rangle.\end{aligned}$$

Using the Young inequality for the last term, we obtain:

$$\begin{aligned}\dot{E}_\varepsilon(t) \;\leqslant\; &-\left(\alpha - \frac{3\varepsilon}{2}\right)\big|\dot{x}(t)\big|^2 - \beta\left(1 - \frac{\varepsilon\beta}{2}\right)\big|\nabla\Phi\big(x(t)\big)\big|^2 \\ &- \varepsilon\big\langle\nabla\Phi\big(x(t)\big), x(t) - z\big\rangle.\end{aligned} \tag{13}$$

Take ε so small that each term in the previous expression is nonpositive (for the last term, use the fact that $\nabla\Phi$ is monotone and $z \in S$); then E_ε is nonincreasing and we readily obtain:

$$\big\langle\dot{x}(t) + \beta\nabla\Phi\big(x(t)\big), x(t) - z\big\rangle + \frac{\alpha}{2}\big|x(t) - z\big|^2 \leqslant \frac{1}{\varepsilon}\big(E_\varepsilon(0) - E(t)\big).$$

Since $E(t)$ is bounded from below, because so is Φ, there exists some constant M such that

$$\bigl\langle \dot{x}(t) + \beta\nabla\Phi\bigl(x(t)\bigr), x(t) - z\bigr\rangle + \frac{\alpha}{2}\bigl|x(t) - z\bigr|^2 \leqslant M.$$

As $\dot{x} + \beta\nabla\Phi(x)$ is bounded by Theorem 2.1(ii), $|x(t) - z|$ is bounded. Hence, $E_\varepsilon(t)$, which is bounded from below and decreasing, admits a limit as $t \to +\infty$. Moreover, Theorem 2.1(ii)–(iii) asserts the following: $\lim_{t\to+\infty} E(t)$ exists and $\lim_{t\to+\infty} \dot{x}(t) = \lim_{t\to+\infty} \nabla\Phi(x(t)) = 0$; hence, after (12), $\lim_{t\to+\infty} |x(t) - z|$ exists.

In order to apply the Opial lemma we need to prove that the weak cluster points of the trajectory x are in S. Let $\bar{x} \in H$ and $t_n \to +\infty$ be such that $x(t_n) \rightharpoonup \bar{x}$. Using the convexity inequality, we have for any $z \in S$,

$$\Phi(z) = \min\Phi \geqslant \Phi\bigl(x(t_n)\bigr) + \bigl\langle \nabla\Phi\bigl(x(t_n)\bigr), z - x(t_n)\bigr\rangle.$$

Since $\nabla\Phi(x(t_n)) \to 0$ and Φ is lower semicontinuous, we obtain:

$$\min\Phi \geqslant \liminf_{n\to+\infty} \Phi\bigl(x(t_n)\bigr) \geqslant \Phi(\bar{x}),$$

which means that $\bar{x} \in S$. The Opial lemma then applies, ensuring the weak convergence of x, and we also deduce that $\Phi(x(t)) \to \min\Phi$ as $t \to \infty$.

5.2. *Strong convergence under* int(Argmin Φ) $\neq \emptyset$

A counterexample due to Baillon [14] for the steepest descent equation $\dot{x} + \nabla\Phi(x) = 0$ suggests that, likely, convexity alone is not sufficient for the trajectories of (DIN) to converge strongly in H. Nevertheless, a result of Brézis [15, Theorem 3.13] shows that the steepest descent trajectories do strongly converge under the additional hypothesis int(Argmin Φ) $\neq \emptyset$. This property also holds for (DIN) trajectories.

Proposition 5.1. *Under the hypotheses of Theorem* 5.1, *if, moreover,* int(Argmin Φ) $\neq \emptyset$ *then every trajectory of* (DIN) *converges to a minimizer of Φ with respect to the strong topology of H.*

Proof. Fix $z \in$ int(Argmin Φ) so that there exists $\rho > 0$ such that for every $z' \in H$ with $|z' - z| < \rho$ then $z' \in$ int(Argmin Φ) and consequently $\nabla\Phi(z') = 0$. By monotonicity of $\nabla\Phi$, we have:

$$\bigl\langle \nabla\Phi(y), y - z\bigr\rangle \geqslant \bigl\langle \nabla\Phi(y), z' - z\bigr\rangle$$

for all $y \in H$ and $z' \in H$ with $\nabla\Phi(z') = 0$. Thus, for every $y \in H$,

$$\bigl\langle \nabla\Phi(y), y - z\bigr\rangle \geqslant \rho\bigl|\nabla\Phi(y)\bigr|.$$

Specialize y to $x(t)$ to obtain for all $t \geqslant 0$ and all $z \in \operatorname{int}(\operatorname{Argmin} \Phi)$:

$$\big\langle \nabla\Phi\big(x(t)\big), x(t) - z\big\rangle \geqslant \rho\big|\nabla\Phi\big(x(t)\big)\big|. \tag{14}$$

Now, for $\varepsilon > 0$ small enough, the inequality (13) may be simplified to

$$0 \leqslant \varepsilon\big\langle \nabla\Phi\big(x(t)\big), x(t) - z\big\rangle \leqslant -\dot{E}_\varepsilon(t);$$

integrating the latter yields

$$0 \leqslant \varepsilon \int_0^t \big\langle \nabla\Phi\big(x(s)\big), x(s) - z\big\rangle \, \mathrm{d}s \leqslant E_\varepsilon(0) - E_\varepsilon(t).$$

Since $\lim_{t\to+\infty} E_\varepsilon(t)$ exists, after the proof of Theorem 5.1, we deduce that $\langle \nabla\Phi(x), x - z\rangle$ belongs to $L^1(0, +\infty)$, and so does $|\nabla\Phi(x)|$ in view of (14). If we now integrate (DIN),

$$\dot{x}(t) + \alpha x(t) + \beta\nabla\Phi\big(x(t)\big) + \int_0^t \nabla\Phi\big(x(s)\big)\,\mathrm{d}s = \dot{x}_0 + \alpha x_0 + \beta\nabla\Phi(x_0),$$

we see that $\lim_{t\to+\infty} x(t)$ exists in H, since $\lim_{t\to+\infty} \dot{x}(t) = \lim_{t\to+\infty} \nabla\Phi(x(t)) = 0$, after Theorem 2.1(iii). □

5.3. *Strong convergence under the symmetry property $\Phi(y) = \Phi(-y)$*

Bruck [16] has shown that the convexity of Φ together with the symmetry assumption $\Phi(y) = \Phi(-y)$ entails the strong convergence of the steepest descent trajectories. This result has been extended by Alvarez [2] to (HBF) trajectories and we extend it now to (DIN) trajectories.

Proposition 5.2. *Under the hypotheses of Theorem* 5.1*, if, moreover, Φ is supposed to be even, i.e. $\forall y \in H$, $\Phi(y) = \Phi(-y)$, then every trajectory of* (DIN) *converges to a minimizer of Φ with respect to the strong topology of H.*

Proof. Let us successively consider the case $\alpha\beta \leqslant 1$ and the case $\alpha\beta > 1$.

1. Case $\alpha\beta \leqslant 1$. Fix $t_0 > 0$ and define $g_{t_0} : [0, t_0] \mapsto \mathbb{R}$ by

$$g_{t_0}(t) = \big|x(t)\big|^2 - \big|x(t_0)\big|^2 - \frac{1}{2}\big|x(t) - x(t_0)\big|^2.$$

We have $\dot{g}_{t_0}(t) = \langle \dot{x}(t), x(t) + x(t_0)\rangle$ and $\ddot{g}_{t_0}(t) = \langle \ddot{x}(t), x(t) + x(t_0)\rangle + |\dot{x}(t)|^2$. From this we obtain:

$$\begin{aligned}\ddot{g}_{t_0}(t)+\alpha\dot{g}_{t_0}(t) &= \big\langle -\beta\nabla^2\Phi\big(x(t)\big)\dot{x}(t)-\nabla\Phi\big(x(t)\big), x(t)+x(t_0)\big\rangle+\big|\dot{x}(t)\big|^2\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\big\langle -\beta\nabla\Phi\big(x(t)\big), x(t)+x(t_0)\big\rangle+\big\langle\beta\nabla\Phi\big(x(t)\big),\dot{x}(t)\big\rangle\\ &\quad+\frac{1}{\beta}\big\langle -\beta\nabla\Phi\big(x(t)\big), x(t)+x(t_0)\big\rangle+\big|\dot{x}(t)\big|^2\\ &= \mathrm{e}^{-(1/\beta)t}\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{(1/\beta)t}\big\langle -\beta\nabla\Phi\big(x(t)\big), x(t)+x(t_0)\big\rangle\\ &\quad+\big\langle\dot{x}(t)+\beta\nabla\Phi\big(x(t)\big),\dot{x}(t)\big\rangle.\end{aligned}$$

Set $f(t)=\langle\dot{x}(t)+\beta\nabla\Phi(x(t)),\dot{x}(t)\rangle$. Since $\dot{x}$ and $\nabla\Phi(x)$ are in $L^2(0,+\infty;H)$, f belongs to $L^1(0,+\infty)$. We have:

$$\frac{\mathrm{d}}{\mathrm{d}t}\big[\mathrm{e}^{\alpha t}\dot{g}_{t_0}(t)\big]=\mathrm{e}^{(\alpha-1/\beta)t}\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{(1/\beta)t}\big\langle -\beta\nabla\Phi\big(x(t)\big), x(t)+x(t_0)\big\rangle+\mathrm{e}^{\alpha t}f(t)$$

and so, for every $t\in\,]0,t_0]$,

$$\mathrm{e}^{\alpha t}\dot{g}_{t_0}(t)-\dot{g}_{t_0}(0)=\int_0^t \mathrm{e}^{(\alpha-1/\beta)\tau}\frac{\mathrm{d}}{\mathrm{d}s}\big[\beta\mathrm{e}^{s/\beta}\omega_{t_0}(s)\big]_{s=\tau}\,\mathrm{d}\tau+\int_0^t \mathrm{e}^{\alpha\tau}f(\tau)\,\mathrm{d}\tau,$$

with $\omega_{t_0}(s)=\langle -\nabla\Phi(x(s)), x(s)+x(t_0)\rangle$. An integration by parts yields

$$\begin{aligned}&\int_0^t \mathrm{e}^{(\alpha-1/\beta)\tau}\frac{\mathrm{d}}{\mathrm{d}s}\big[\beta\mathrm{e}^{s/\beta}\omega_{t_0}(s)\big]_{s=\tau}\,\mathrm{d}\tau\\ &\quad=\beta\mathrm{e}^{\alpha t}\omega_{t_0}(t)-\beta\omega_{t_0}(0)+(1-\alpha\beta)\int_0^t \mathrm{e}^{\alpha\tau}\omega_{t_0}(\tau)\,\mathrm{d}\tau.\end{aligned}$$

We conclude that

$$\begin{aligned}\dot{g}_{t_0}(t) &= \big\langle\dot{x}_0+\beta\nabla\Phi(x_0), x_0+x(t_0)\big\rangle\mathrm{e}^{-\alpha t}+\beta\omega_{t_0}(t)\\ &\quad+\int_0^t \mathrm{e}^{-\alpha(t-\tau)}\big[(1-\alpha\beta)\omega_{t_0}(\tau)+f(\tau)\big]\,\mathrm{d}\tau.\end{aligned}$$

Set $F(t)=(1/2)|\dot{x}(t)|^2+\Phi(x(t))$, which is nonincreasing because Φ is convex (in fact, $\dot{F}(t)=-\alpha|\dot{x}(t)|^2-\beta\langle\nabla^2\Phi(x(t))\dot{x}(t),\dot{x}(t)\rangle\leqslant 0$). Then, for all $t\in[0,t_0]$,

$$F(t)\geqslant F(t_0)=\frac{1}{2}\big|\dot{x}(t_0)\big|^2+\Phi\big(x(t_0)\big)=\frac{1}{2}\big|\dot{x}(t_0)\big|^2+\Phi\big(-x(t_0)\big).$$

By convexity of Φ,

$$\Phi\big(-x(t_0)\big) \geqslant \Phi\big(x(t)\big) + \big\langle \nabla\Phi\big(x(t)\big), -x(t_0) - x(t)\big\rangle$$

and, consequently,

$$\omega_{t_0}(t) = \big\langle -\nabla\Phi\big(x(t)\big), x(t) + x(t_0)\big\rangle \leqslant \frac{1}{2}\big|\dot{x}(t)\big|^2.$$

Therefore,

$$\dot{g}_{t_0}(t) \leqslant \big\langle \dot{x}_0 + \beta\nabla\Phi(x_0), x_0 + x(t_0)\big\rangle \mathrm{e}^{-\alpha t} + \frac{\beta}{2}\big|\dot{x}(t)\big|^2 + \int_0^t \mathrm{e}^{-\alpha(t-\tau)} h(\tau)\,\mathrm{d}\tau,$$

where $h(t) = ((1-\alpha\beta)/2)|\dot{x}(t)|^2 + |f(t)| \in L^1(0,\infty)$. Hence, for all $t \in [0, t_0]$,

$$\begin{aligned} g_{t_0}(t_0) - g_{t_0}(t) \;\leqslant\; & \frac{1}{\alpha}\big\langle \dot{x}_0 + \beta\nabla\Phi(x_0), x_0 + x(t_0)\big\rangle\big(\mathrm{e}^{-\alpha t} - \mathrm{e}^{-\alpha t_0}\big) \\ & + \frac{\beta}{2}\int_t^{t_0} \big|\dot{x}(\tau)\big|^2\,\mathrm{d}\tau + \int_t^{t_0}\int_0^{\theta} \mathrm{e}^{-\alpha(\theta-\tau)} h(\tau)\,\mathrm{d}\tau\,\mathrm{d}\theta \end{aligned}$$

which gives

$$\begin{aligned} \frac{1}{2}\big|x(t_0) - x(t)\big|^2 \;\leqslant\; & \big|x(t)\big|^2 - \big|x(t_0)\big|^2 \\ & + \frac{1}{\alpha}\big\langle \dot{x}_0 + \beta\nabla\Phi(x_0), x_0 + x(t_0)\big\rangle\big(\mathrm{e}^{-\alpha t} - \mathrm{e}^{-\alpha t_0}\big) + \int_t^{t_0} p(\theta)\,\mathrm{d}\theta, \end{aligned}$$

where $p \in L^1(0,\infty)$. We know that $x(t) \rightharpoonup x_\infty$ as $t \to \infty$ where $x_\infty \in \operatorname{Argmin}\Phi$. Moreover, for all $z \in \operatorname{Argmin}\Phi$ there exists some $l_z \in \mathbb{R}$ such that $|x(t) - z|^2 \to l_z$, as $t \to \infty$ (see Theorem 5.1). Since Φ is even, 0 is a minimizer of Φ so that there is some $l_0 \in \mathbb{R}$ such that $\lim_{t\to\infty} |x(t)|^2 = l_0$. From the inequality above it follows that $\{x(t) : t \to \infty\}$ is a Cauchy net in H, hence, $x(t) \to x_\infty$ strongly in H.

2. Case $\alpha\beta > 1$. The conclusion follows in this case from a well-known result of Bruck [16] applied to an equivalent gradient-type first-order system defined on $H \times H$ (see Section 6.3). □

Remark. If $\Phi(x) = (1/2)\langle Ax, x\rangle$ where $A : H \mapsto H$ is a positive self-adjoint and bounded linear operator, then $\operatorname{Argmin}\Phi = \operatorname{Ker} A = \{z \in H\colon\ Az = 0\}$ and $x(t)$ strongly converges in H to the projection of $x_0 + (1/\alpha)\dot{x}_0$ on $\operatorname{Ker} A$. Indeed, for every $z \in \operatorname{Ker} A$ and $t > 0$, we have:

$$\begin{aligned}\big\langle \dot{x}(t)+\alpha x(t)-\dot{x}_0-\alpha x_0, z\big\rangle &= \int_0^t \big\langle -\beta\nabla^2\Phi\big(x(\tau)\big)\dot{x}(\tau)-\nabla\Phi\big(x(\tau)\big), z\big\rangle\,\mathrm{d}\tau\\ &= \int_0^t \big\langle -\beta A\dot{x}(\tau)-Ax(\tau), z\big\rangle\,\mathrm{d}\tau\\ &= \int_0^t \big\langle -\beta\dot{x}(\tau)-x(\tau), Az\big\rangle\,\mathrm{d}\tau = 0.\end{aligned}$$

Since $\dot{x}(t)\to 0$ and $x(t)\to x_\infty \in \operatorname{Ker} A$ strongly, we deduce that $\langle x_\infty - x_0 - (1/\alpha)\dot{x}_0, z\rangle = 0$ for all $z\in \operatorname{Ker} A$, which proves our claim.

6. (DIN) as a first-order in time gradient-like system

This part is devoted to establishing two remarkable properties of (DIN):

- actually (DIN) proves to be equivalent to a system of first-order in time with no occurrence of the Hessian of Φ;
- further, if the positive parameters α and β satisfy $\alpha\beta > 1$, then (DIN) is a gradient system.

6.1. *(DIN) as a system of first-order in time and with no occurrence of the Hessian of Φ*

In this section, the requirements on the constants α, β and on the function Φ in (DIN) may be relaxed to $\beta \neq 0$ and $\Phi\in\mathcal{C}^2(H)$ only.

Let x be a solution of (DIN), and define the function y by:

$$\dot{x}+\beta\nabla\Phi(x)+\left(\alpha-\frac{1}{\beta}\right)x+\frac{1}{\beta}y=0. \tag{15}$$

Differentiate (15) to obtain:

$$\beta\left[\ddot{x}+\beta\nabla^2\Phi(x)\dot{x}+\left(\alpha-\frac{1}{\beta}\right)\dot{x}\right]+\dot{y}=0,$$

which, in view of (DIN), yields

$$\beta\left[-\nabla\Phi(x)-\frac{1}{\beta}\dot{x}\right]+\dot{y}=0. \tag{16}$$

Adding (15) and (16) gives:

$$\left(\alpha-\frac{1}{\beta}\right)x+\dot{y}+\frac{1}{\beta}y=0. \tag{17}$$

Collecting (15) and (17) gives the first-order system:

$$\begin{cases} \dot{x} + \beta \nabla \Phi(x) + \left(\alpha - \dfrac{1}{\beta}\right)x + \dfrac{1}{\beta} y = 0, \\ \dot{y} + \left(\alpha - \dfrac{1}{\beta}\right)x + \dfrac{1}{\beta} y = 0. \end{cases} \tag{18}$$

Conversely, let (x, y) be a solution of (18). Combining the two lines of (18) yields $\dot{y} = \dot{x} + \beta \nabla \Phi(x)$, while differentiating the first equation yields

$$\ddot{x} + \beta \nabla^2 \Phi(x)\dot{x} + \left(\alpha - \frac{1}{\beta}\right)\dot{x} + \frac{1}{\beta}\dot{y} = 0.$$

Substituting the value of $\dot{y}$ in the above equation gives (DIN) again. Thus (DIN) is equivalent to (18).

It is natural now to introduce the following first-order system (where g stands for *generalized*)

$$\text{(g-DIN)} \qquad \begin{cases} \dot{x} + \beta \nabla \Phi(x) + ax + by = 0, \\ \dot{y} + ax + by = 0, \end{cases}$$

which is a slight generalization of (18); indeed (g-DIN) is (18) if we set:

$$a = \alpha - \frac{1}{\beta}, \quad b = \frac{1}{\beta}. \tag{19}$$

The following theorem summarizes the above computation, and emphasizes the equivalence of (DIN), which is of second-order in time and involves the Hessian of Φ, with a system which is of first-order in time and with no occurrence of the Hessian.

Theorem 6.1. *Suppose $\Phi \in \mathcal{C}^2(H)$, and let the constants α, β, a, b satisfy $\beta \neq 0$ and* (19). *The systems* (DIN) *and* (g-DIN) *are equivalent in the sense that x is a solution of* (DIN) *if and only if there exists $y \in \mathcal{C}^2([0, +\infty[, H)$ such that (x, y) is a solution of* (g-DIN).

6.2. Existence and asymptotic behaviour of the solutions of (g-DIN)

Beyond being of first-order in time, the system (g-DIN) is interesting because it does not involve the Hessian of Φ. As a first consequence, the numerical solution of (DIN) is highly simplified, since it may be performed on (g-DIN) and only requires approximating the gradient of Φ. As a second consequence, (g-DIN) allows to give a sense to (DIN) when Φ is of class $\mathcal{C}^1$ only, or when Φ is nonsmooth or involves constraints, provided that a notion of generalized gradient is available (e.g., the subdifferential set for a convex function Φ). But that remark would be of little utility if (g-DIN) did not have good existence and convergence properties under the sole assumption $\Phi \in \mathcal{C}^1(H)$; recall that (DIN), as studied in the previous sections, requires $\Phi \in \mathcal{C}^2(H)$. Actually (g-DIN) enjoys the same properties as (DIN) does, at least if $\Phi \in \mathcal{C}^{1,1}(H)$, and theorems similar to Theorems 2.1 and 5.1 can be stated about (g-DIN).

Theorem 6.2. *Assume that $\Phi : H \mapsto \mathbb{R}$ is bounded from below, differentiable with $\nabla\Phi$ Lipschitz continuous on the bounded subsets of H; assume further $\beta > 0$, $b > 0$, $b + a > 0$ in* (g-DIN). *Then the following properties hold*:

(i) *For each (x_0, y_0) in $H \times H$, there exists a unique solution (x, y) of* (g-DIN) *defined on the whole interval $[0, +\infty[$, which belongs to $\mathcal{C}^1(0, \infty; H) \times \mathcal{C}^2(0, \infty; H)$ and satisfies the initial conditions $x(0) = x_0$ and $y(0) = y_0$.*

(ii) *For any $\lambda \in [\beta(\sqrt{a+b} - \sqrt{b})^2, \beta(\sqrt{a+b} + \sqrt{b})^2]$ the function*

$$F_\lambda : (x, y) \in H \times H \mapsto \lambda\Phi(x) + (1/2)|ax + by|^2$$

is a Liapounov function of (g-DIN); *for every solution (x, y) the energy $F_\lambda(x(t), y(t))$ is decreasing on $[0, +\infty[$, bounded from below and hence, it converges to some real value as $t \to +\infty$. Moreover,*
- *$\dot{x}$ and $\nabla\Phi(x)$ belong to $L^2(0, +\infty; H)$;*
- *$\lim_{t\to+\infty} \Phi(x(t))$ exists;*
- *$\lim_{t\to+\infty}(\dot{x}(t) + \beta\nabla\Phi x(t)) = 0$.*

(iii) *Assuming moreover that x is in $L^\infty(0, +\infty; H)$, then we have*:
- *$\dot{x}$, $\nabla\Phi(x)$ are bounded on $[0, +\infty[$;*
- *$\lim_{t\to+\infty} \nabla\Phi(x(t)) = \lim_{t\to+\infty} \dot{x}(t) = 0$.*

Theorem 6.3. *In addition to the hypotheses of Theorem* 6.2 *assume that Φ is convex and that* Argmin Φ*, the set of minimizers of Φ on H, is nonempty. Then for any solution (x, y) of* (g-DIN)*, $x(t)$ weakly converges to a minimizer of Φ on H as t goes to infinity.*

The proof follows the lines of Theorems 2.1 and 5.1 and will not be given. Besides, a more general situation will be examinated in Section 7 (cf. Theorems 7.1 and 7.2).

Theorem 2.1 is a mere corollary of Theorems 6.1 and 6.2. Indeed suppose that Φ and α, β meet the assumptions of Theorem 2.1: Φ satisfies $(\mathcal{H})$ and $\alpha > 0$, $\beta > 0$. Then $\nabla\Phi$ is Lipschitz continuous on the bounded subsets of H, and the constants $a = \alpha - 1/\beta$ and $b = 1/\beta$ satisfy $a + b > 0$, $b > 0$. So the assumptions of Theorem 6.2 are met; in view of the equivalence between (DIN) and (g-DIN) given by Theorem 6.1, the conclusions of Theorem 6.2 apply to (DIN).

Further, if $\Phi \in \mathcal{C}^2(H)$ meets the assumptions of Theorem 6.2, the system (DIN) makes sense but Theorem 2.1 does not apply since $\nabla^2\Phi$ need not be Lipschitz continuous. Yet we can resort to Theorems 6.1 and 6.2 to assert the existence of a solution to (DIN) enjoying the properties stated in Theorem 6.2. Consequently, the assumptions of Theorem 2.1 may be weakened, while its conclusions remain valid, as far as $\ddot{x}$ and $\nabla^2\Phi$ are not concerned.

Likewise Theorem 5.1 is a corollary of Theorems 6.1 and 6.3 and its hypotheses may be weakened.

6.3. (DIN) as a gradient system if $\alpha\beta > 1$

Suppose $\Phi \in \mathcal{C}^1(H)$ and $a > 0$, $b > 0$ in (g-DIN). Rescaling the variable y by $y = \sqrt{a/b}\,z$ transforms (g-DIN) into the equivalent system:

$$\begin{cases} \dot{x} + \beta\nabla\Phi(x) + ax + \sqrt{ab}\,z = 0, \\ \dot{z} + \sqrt{ab}\,x + bz = 0. \end{cases} \tag{20}$$

We note that (20) is exactly the gradient system

$$\dot{X} + \nabla\mathcal{E}(X) = 0, \tag{21}$$

where $X = (x, z)$ and $\mathcal{E} : H \times H \mapsto \mathbb{R}$ is defined by:

$$\mathcal{E}(X) = \beta\Phi(x) + \frac{1}{2}\left|\sqrt{a}\,x + \sqrt{b}\,z\right|^2.$$

Suppose now that Φ belongs to $\mathcal{C}^2(H)$ and let us turn to (DIN) which we know is equivalent to (g-DIN) with $a = \alpha - 1/\beta$, $b = 1/\beta$. If α, β satisfy $\alpha\beta > 1$ in addition to $\alpha > 0$, $\beta > 0$, then a, b satisfy $a > 0$, $b > 0$. As a consequence, (DIN) is equivalent to the gradient system (20); using the parameters α, β the expression of $\mathcal{E}$ is

$$\mathcal{E}(X) = \mathcal{E}(x, z) = \beta\Phi(x) + \frac{1}{2\beta}\left|\sqrt{\alpha\beta - 1}\,x + z\right|^2. \tag{22}$$

We state as a proposition that remarkable property of (DIN).

Proposition 6.1. *Suppose $\Phi \in \mathcal{C}^2(H)$, $\alpha > 0$, $\beta > 0$ and $\alpha\beta > 1$. The system* (DIN) *is equivalent to the gradient system* (21) *with $\mathcal{E}$ given by* (22).

Since the functional $\mathcal{E}$ equals $\beta\Phi$ plus a positive quadratic form, it inherits most of the eventual properties of Φ: boundedness from below, coercivity, regularity, analyticity, convexity... Moreover, if $(\bar{x}, \bar{z})$ is a critical (or minimum) point of $\mathcal{E}$ then $\bar{x}$ is a critical (or minimum) point of Φ. Thus the equivalence of (DIN) with the gradient system (21) allows properties of gradient systems to pass to (DIN).

For example, if Φ is analytic then so is $\mathcal{E}$. Further, if x is a bounded solution of (DIN) then $\dot{x}$ is bounded (Theorem 2.1(iii)) and (x, z) is a bounded solution of (21) which is known to converge to a critical point of $\mathcal{E}$ [33,24]. Hence, x converges to a critical point of Φ.

Likewise in the convex case, Theorem 5.1 and Propositions 5.1 and 5.2 are consequences of theorems of Bruck [16] and Brézis [15]; that remark completes the proof of Proposition 5.2 where the case $\alpha\beta > 1$ was pending.

6.4. Remarks

6.4.1. Structure of (DIN) when $\alpha\beta < 1$

Suppose $\Phi \in \mathcal{C}^1(H)$ and $a < 0$, $b > 0$ in (g-DIN). Rescaling the variable y by $y = \sqrt{-a/b}z$ transforms (g-DIN) into the equivalent system:

$$\begin{cases} \dot{x} + \beta\nabla\Phi(x) + ax + \sqrt{-ab}\, z = 0, \\ \dot{z} - \sqrt{-ab}\, x + bz = 0. \end{cases} \tag{23}$$

Set $X = (x, z)$ and define the functional $\mathcal{F}: H \times H \mapsto \mathbb{R}$ by $\mathcal{F}(X) = \beta\Phi(x) + (1/2)(a|x|^2 + b|z|^2)$, and the linear operator $J : H \times H \mapsto H \times H$ by $J(X) = \sqrt{-ab}(z, -x)$. Then (23) can be written

$$\dot{X} + \nabla\mathcal{F}(X) + J(X) = 0 \tag{24}$$

which appears as a gradient system perturbed by the monotone operator J. Unfortunately, properties such as convexity or boundedness from below do not pass from Φ to $\mathcal{F}$ since the quadratic form $(1/2)(a|x|^2 + b|z|^2)$ is not positive.

As to (DIN), if we suppose $\Phi \in \mathcal{C}^2(H)$, $\alpha > 0$, $\beta > 0$ and $\alpha\beta < 1$, then the equivalent (g-DIN) system verifies $a < 0$, $b > 0$, and (DIN) turns to be equivalent to (24) too.

The system (g-DIN) can be given another equivalent form if we suppose $a < 0$ and $a + b > 0$.[3] Indeed make the change of variable $y = (1/b)(\sqrt{-a(a+b)}\, z - ax)$; then (g-DIN) becomes:

$$\begin{cases} \dot{x} + \beta\nabla\Phi(x) + \sqrt{-a(a+b)}\, z = 0, \\ \dot{z} - \beta\sqrt{\dfrac{-a}{a+b}}\, \nabla\Phi(x) + (a+b)z = 0. \end{cases} \tag{25}$$

Introduce the function $\mathcal{G}(X) = \mathcal{G}(x, z) = \beta\Phi(x) + (1/2)|z|^2$ and the linear monotone operator $J(x, z) = \sqrt{-a/(a+b)}(z, -x)$, then (25) becomes

$$\dot{X} + (1 + J)\nabla\mathcal{G}(X) = 0. \tag{26}$$

Turning back to (DIN), if we suppose $\Phi \in \mathcal{C}^2(H)$, $\alpha > 0$, $\beta > 0$ and $\alpha\beta < 1$, then we have $a < 0$ and $a + b > 0$ in the system (g-DIN) associated *via* (19), and, hence, (DIN) is equivalent to (26).

Unfortunately, systems (24) and (26) are not easy to deal with, and when $\alpha\beta < 1$ in (DIN) (or $a < 0$ in (g-DIN)) the only results remain those given in Sections 2, 4, 5 (or by Theorems 6.2 and 6.3).

[3] We are indebted to our colleague X. Goudou for pointing out this fact to us.

6.4.2. The change of coordinates in (15), which allows to transform (DIN) into the first-order system (g-DIN), may appear as a trick. Yet, when investigating the minimum (or critical) points of Φ, there often appears a function of the form $\Psi(x,y)=\Phi(x)+(1/2)|ax+by|^2$ (x,y in H and a,b real) the decrease of which lies at the root of the analysis. One recognizes in Ψ the energy functional of (DIN) or (HBF), and perhaps more subtly the function $(x,y)\mapsto\Phi(x)+(1/(2\lambda))|x-y|^2$ ($\lambda>0$) which occurs in the minimization of Φ by the proximal algorithm [23]:

$$x_{n+1}=\operatorname*{argmin}_{x\in H}\left\{\Phi(x)+\frac{1}{2\lambda}|x-x_n|^2\right\}.$$

Applying the continuous steepest descent method to Ψ is then tempting; it yields a first-order system such as (g-DIN), and eliminating y gives (DIN). Performing the computations backward and generalizing them leads to the developments of Sections 6.1 and 6.2.

6.4.3. (DIN) can be written as an integro–differential equation:

$$\dot{x}(t)+\beta\nabla\Phi\big(x(t)\big)=(\alpha\beta-1)\int_0^t\nabla\Phi\big(x(s)\big)\exp\big(\alpha(s-t)\big)\,\mathrm{d}s\\ +\big(\dot{x}_0+\beta\nabla\Phi(x_0)\big)\exp(-\alpha t).$$

Thus, if $\alpha\beta=1$, one obtains the nonautonomous first-order gradient system:

$$\dot{x}(t)+\beta\nabla\Phi\big(x(t)\big)=\big(\dot{x}_0+\beta\nabla\Phi(x_0)\big)\exp(-\alpha t).$$

7. Application to constrained optimization

The equivalence between (DIN) and (g-DIN) suggests a method to solve constrained optimization problems with the help of a dynamical system like (g-DIN); that is the subject of this section.

Fix C a nonempty closed convex set of H. In the following we suppose that Φ is $\mathcal{C}^1$ with $\nabla\Phi$ Lipschitz continuous on bounded sets and we consider the following problem

$$(\mathcal{P})\qquad \inf_C \Phi.$$

When we want to solve $(\mathcal{P})$ with a second-order in time dynamical system, we have to face a major difficulty: how can we both force the orbits starting in C to lie in C and to keep their inertial aspects? In many practical cases such a *viability property* is of interest. Those problems of viability are easier to handle when we deal with first-order systems. If we consider, for example, the following system initiated by Antipin [5,6]:

$$(\mathrm{S1})\qquad \begin{cases}\dot{x}(t)+x(t)-P_C\big[x(t)-\mu\nabla\Phi\big(x(t)\big)\big]=0,\\ x(0)=x_0\in C,\end{cases}$$

where P_C is the projection on C and $\mu > 0$, then the viability property is obvious since the corresponding vector field enters the set of constraints. This dynamics provides moreover orbits that enjoy nice asymptotic properties: if we suppose Φ to be convex then trajectories weakly converge towards a minimum of Φ on C, even if we only assume $x_0 \in C$. This system has also been studied in its second-order in time form, namely:

$$\text{(S2)} \qquad \begin{cases} \ddot{x}(t) + \alpha\dot{x}(t) + x(t) - P_C\big[x(t) - \mu\nabla\Phi\big(x(t)\big)\big] = 0, \\ x(0) = x_0 \in C, \quad \dot{x}(0) = \dot{x}_0 \in H, \end{cases}$$

but in that case the viability property is no longer maintained. This naturally leads to strong hypotheses on the potential Φ to obtain a proper optimizing system, see, for example, [6–8].

We propose in the following theorem to combine (g-DIN) and (S1) to solve ($\mathcal{P}$). More precisely, given real parameters β, a and b such that $\beta > 0$, $a \neq 0$, $b > 0$ and $b + a > 0$, we consider the first-order system in $H \times H$:

$$\text{(c-DIN)} \qquad \begin{cases} \dot{x}(t) + x(t) - P_C\big[x(t) - \beta\nabla\Phi\big(x(t)\big) - ax(t) - by(t)\big] = 0, \\ \dot{y}(t) + ax(t) + by(t) = 0, \end{cases}$$

with initial conditions

$$x(0) = x_0 \in C, \qquad y(0) = y_0 \in H. \tag{27}$$

Of course, (c-DIN) reduces to (g-DIN) if $C = H$. The functional Φ is required to satisfy the following hypotheses:

$$(\mathcal{H}\text{-c}) \qquad \begin{cases} \Phi \text{ is defined and continuously differentiable} \\ \quad \text{on an open neighbourhood of the closed convex set } C, \\ \Phi \text{ is bounded from below on } C, \\ \text{the gradient } \nabla\Phi \text{ is Lipschitz continuous} \\ \quad \text{on the bounded subsets of } C. \end{cases}$$

If (x, y) is a solution to (c-DIN) and for $\lambda > 0$, let us define:

$$E_\lambda(t) = \lambda\Phi\big(x(t)\big) + \frac{1}{2}\big|ax(t) + by(t)\big|^2. \tag{28}$$

A theorem similar to Theorem 2.1 can be stated and proved for (c-DIN).

Theorem 7.1. *Let Φ satisfy the hypotheses* ($\mathcal{H}$-c) *and assume $\beta > 0$, $a \neq 0$, $b > 0$ and $b + a > 0$. Then the following properties hold*:

(i) *For each $(x_0, y_0) \in C \times H$, there exists a unique solution $(x(t), y(t))$ of* (c-DIN) *defined on the whole interval $[0, +\infty[$ which satisfies the initial conditions $x(0) = x_0$, $y(0) = y_0$; (x, y) belongs to $\mathcal{C}^1(0, +\infty; H) \times \mathcal{C}^2(0, +\infty; H)$ and x is viable, that is $x(t)$ lies in C for all $t \geqslant 0$.*

(ii) *For every trajectory* $(x(t), y(t))$ *of* (c-DIN) *and for* $\lambda \in [\beta(\sqrt{b} - \sqrt{b+a})^2, \beta(\sqrt{b} + \sqrt{b+a})^2]$*, the energy* E_λ *is decreasing on* $[0, +\infty[$*, bounded from below and, hence, converges to some real value as* $t \to +\infty$*. Moreover,*
- $\dot{x}$ *and* $\dot{y}$ *belong to* $L^2(0, +\infty; H)$;
- $\lim_{t\to+\infty} \Phi(x(t))$ *exists*;
- $\lim_{t\to+\infty} \dot{y}(t) = 0$.

(iii) *Assuming in addition that* x *is in* $L^\infty(0, +\infty; H)$*, we have*:
- $\nabla\Phi(x)$*,* y*,* $\dot{x}$ *are bounded on* $[0, +\infty[$;
- $\lim_{t\to+\infty} \dot{x}(t) = 0$.

The proof essentially goes along the same lines as in Theorem 2.1. The nonlinearity caused by the projection P_C is compensated by the characteristic inequality $\langle v - P_C u, u - P_C u\rangle \leqslant 0$ for all (u, v) in $H \times C$. The natural quantities upon which the calculations rely are $\dot{x}$ and $\dot{y}$ (rather than $\dot{x}$ and $\nabla\Phi(x)$ in the proof of Theorem 2.1).

Proof of Theorem 7.1. (i) Since the projection P_C is a Lipschitz continuous operator, the local existence and the uniqueness of a solution to (c-DIN) with initial conditions (27) follow from the Cauchy–Lipschitz theorem. Let (x, y) denote the maximal solution defined on some interval $[0, T_{\max}[$ with $0 \leqslant T_{\max} \leqslant +\infty$.

First let us show that x is viable for $t \in [0, T_{\max}[$. Define $p : [0, T_{\max}[\mapsto C$ by $p(t) = P_C[x(t) - \beta\nabla\Phi(x(t)) - ax(t) - by(t)]$ and integrate the equation $\dot{x} + x = p$ on $[0, t] \subset [0, T_{\max}[$:

$$x(t) = \int_0^t \mathrm{e}^{-(t-s)} p(s)\,\mathrm{d}s + \mathrm{e}^{-t} x_0.$$

Observe that $\xi(t) = \int_0^t \mathrm{e}^{-(t-s)}/(1 - \mathrm{e}^{-t})\, p(s)\,\mathrm{d}s$ belongs to C, as the weight function $s \mapsto \mathrm{e}^{-(t-s)}/(1 - \mathrm{e}^{-t})$ is positive and its integral over $[0, t]$ is 1. Now writing $x(t) = (1 - \mathrm{e}^{-t})\xi(t) + \mathrm{e}^{-t}x_0$ shows that $x(t)$ belongs to C.

Next, the viability of x and the convexity of C are used to derive the following inequality on $[0, T_{\max}[$:

$$\bigl\langle x - P_C\bigl(x - \beta\nabla\Phi(x) + \dot{y}\bigr), x - \beta\nabla\Phi(x) + \dot{y} - P_C\bigl(x - \beta\nabla\Phi(x) + \dot{y}\bigr)\bigr\rangle \leqslant 0,$$

which, in view of (c-DIN), successively reduces to

$$\bigl\langle -\dot{x}, -\dot{x} - \beta\nabla\Phi(x) + \dot{y}\bigr\rangle \leqslant 0, \qquad \beta\bigl\langle \dot{x}, \nabla\Phi(x)\bigr\rangle \leqslant -|\dot{x}|^2 + \langle \dot{x}, \dot{y}\rangle. \tag{29}$$

Further, in order to apply classical energy arguments, we show that E_λ defined by (28) is decreasing along the trajectory (x, y), at least for some value of λ. Indeed, we have (using the second equation in (c-DIN)):

$$\dot{E}_\lambda = \lambda\bigl\langle \dot{x}, \nabla\Phi(x)\bigr\rangle - b|\dot{y}|^2 - a\langle \dot{x}, \dot{y}\rangle.$$

Taking (29) into account, we obtain:

$$\dot{E}_\lambda \leqslant -\frac{\lambda}{\beta}|\dot{x}|^2 - b|\dot{y}|^2 + \left(\frac{\lambda}{\beta} - a\right)\langle \dot{x}, \dot{y}\rangle. \tag{30}$$

In particular, if we choose $\lambda = \beta(a + 2b)$ (this last quantity is positive), we have:

$$\dot{E}_{\beta(a+2b)} \leqslant -(a+b)|\dot{x}|^2 - b|\dot{x} - \dot{y}|^2. \tag{31}$$

Integrating this inequality over $[0, t] \subset [0, T_{\max}[$, we obtain:

$$\beta(a+2b)\Phi\big(x(t)\big) + \frac{1}{2}\big|ax(t) + by(t)\big|^2 + (a+b)\int_0^t \big|\dot{x}(\tau)\big|^2 \,\mathrm{d}\tau + b\int_0^t \big|\dot{x}(\tau) - \dot{y}(\tau)\big|^2 \,\mathrm{d}\tau$$
$$\leqslant \beta(a+2b)\Phi(x_0) + \frac{1}{2}|ax_0 + by_0|^2. \tag{32}$$

Finally, to prove that (x, y) is defined over $[0, +\infty[$, we suppose that $T_{\max} < +\infty$ and argue by contradiction. Since x is viable and Φ is bounded from below, (32) shows that $\dot{y} = -(ax + by)$ is bounded on $[0, T_{\max}[$; hence, $\lim_{t\to T_{\max}} y(t)$ exists. As a consequence, y and $x = -(1/a)(\dot{y} + by)$ are bounded, and so is $\nabla\Phi(x)$ in view of ($\mathcal{H}$-c). Then (c-DIN) shows that $\dot{x}$ is bounded too. Hence, $\lim_{t\to T_{\max}} x(t)$ exists. This classically yields a contradiction, and $T_{\max}$ must be equal to $+\infty$.

The last assertion, $(x, y) \in \mathcal{C}^1(0, +\infty; H) \times \mathcal{C}^2(0, +\infty; H)$, immediately follows from (c-DIN).

(ii) Set $q(\lambda) = -(\lambda/\beta)|\dot{x}|^2 - b|\dot{y}|^2 + ((\lambda/\beta) - a)\langle \dot{x}, \dot{y}\rangle$, $\lambda_{\min} = \beta(\sqrt{b} - \sqrt{b+a}\,)^2$, and $\lambda_{\max} = \beta(\sqrt{b} + \sqrt{b+a}\,)^2$. The inequality (30) yields:

$$\dot{E}_{\lambda_{\min}} \leqslant q(\lambda_{\min}) = -\big|\big(\sqrt{b} - \sqrt{b+a}\,\big)\dot{x} + \sqrt{b}\,\dot{y}\big|^2,$$
$$\dot{E}_{\lambda_{\max}} \leqslant q(\lambda_{\max}) = -\big|\big(\sqrt{b} + \sqrt{b+a}\,\big)\dot{x} - \sqrt{b}\,\dot{y}\big|^2.$$

Since q is an affine function of λ for every $\lambda \in [\lambda_{\min}, \lambda_{\max}]$, $\dot{E}_\lambda$ lies between $q(\lambda_{\min})$ and $q(\lambda_{\max})$ and hence, is nonpositive. The energy E_λ is then decreasing on $[0, +\infty[$ and converges since Φ is bounded from below on C.

The inequality (32) shows that $\dot{x}$ and $\dot{y}$ belong to $L^2(0, +\infty; H)$.

Now, considering two different values λ, λ' in $[\lambda_{\min}, \lambda_{\max}]$ shows that $\Phi(x) = (1/(\lambda' - \lambda))(E_{\lambda'} - E_\lambda)$ admits a limit as $t \to +\infty$.

Hence, $|\dot{y}|^2 = |ax + by|^2 = 2(E_\lambda - \lambda\Phi(x))$ also admits a limit which necessarily is zero since $|\dot{y}|$ belongs to $L^2(0, +\infty; H)$.

(iii) If x is bounded, then $\nabla\Phi(x)$ is bounded (after ($\mathcal{H}$-c)), and $y = -(1/b)(ax + \dot{y})$ is bounded (recall $\dot{y} \to 0, t \to +\infty$). Further $\dot{x}$ is bounded in view of (c-DIN). Since $\dot{x}$ and $\dot{y}$ are bounded, x and y are Lipschitz continuous, which shows, in view of (c-DIN), that $\dot{x}$ itself is Lipschitz continuous. But $\dot{x}$ belongs to $L^2(0, +\infty; H)$, hence, according to a classical argument, $\dot{x}(t) \to 0$ as $t \to +\infty$. $\square$

Theorem 7.2. *In addition to the hypotheses of Theorem* 7.1*, assume that Φ is convex and that* $\operatorname{Argmin}_C \Phi$*, the set of minimizers of Φ on C, is nonempty. Then for any solution* $(x(t), y(t))$ *of* (c-DIN)*, $x(t)$ weakly converges to a minimizer of Φ on C as t goes to infinity.*

Proof. First, let us establish some useful inequalities. Let x^* be a minimizer of Φ on C. Use the characteristic inequality for P_C to write (it is implicit that the time variable t varies in $[0, +\infty[$ in the following):

$$\big\langle x^* - P_C\big(x - \beta\nabla\Phi(x) + \dot{y}\big), x - \beta\nabla\Phi(x) + \dot{y} - P_C\big(x - \beta\nabla\Phi(x) + \dot{y}\big)\big\rangle \leqslant 0.$$

In view of (c-DIN) we derive

$$\begin{gathered}\big\langle x^* - x - \dot{x}, -\dot{x} - \beta\nabla\Phi(x) + \dot{y}\big\rangle \leqslant 0,\\ \langle x^* - x, \dot{y} - \dot{x}\rangle + \beta\big\langle \dot{x}, \nabla\Phi(x)\big\rangle \leqslant \big\langle x^* - x, \beta\nabla\Phi(x)\big\rangle - |\dot{x}|^2. \end{gathered}\tag{33}$$

But $\langle x^* - x, \nabla\Phi(x^*) - \nabla\Phi(x)\rangle$ is nonnegative since Φ is convex; and $\langle x^* - x, -\nabla\Phi(x^*)\rangle$ is nonnegative because x^* is a minimizer of Φ on C. Hence, $\langle x^* - x, -\nabla\Phi(x)\rangle$ is nonnegative and (33) entails

$$\langle x^* - x, \dot{y} - \dot{x}\rangle + \beta\big\langle \dot{x}, \nabla\Phi(x)\big\rangle \leqslant -|\dot{x}|^2. \tag{34}$$

Our aim now is to introduce an energy functional involving the term $|x^* - x|$. Set

$$F(t) = \big\langle x^* - x(t), ax(t) + by(t)\big\rangle + \frac{1}{2}(b + a)\big|x^* - x(t)\big|^2 + b\beta\Phi\big(x(t)\big).$$

We have

$$\dot{F} = b\big(\langle x^* - x, \dot{y} - \dot{x}\rangle + \big\langle \dot{x}, \beta\nabla\Phi(x)\big\rangle\big) + \langle \dot{x}, \dot{y}\rangle,$$

and in view of (34) we obtain:

$$\dot{F} \leqslant \langle \dot{x}, \dot{y}\rangle - b|\dot{x}|^2 \leqslant -\left(b - \frac{3}{2}\right)|\dot{x}|^2 + \frac{1}{2}|\dot{y} - \dot{x}|^2. \tag{35}$$

In view of (31) and (35) we may fix some $\varepsilon > 0$ so small that the function $\mathcal{E} : \mathbb{R} \mapsto H$ defined by:

$$\begin{aligned}\mathcal{E} = E_{a+2b} + \varepsilon F \ &= (a + 2b + \varepsilon b\beta)\Phi(x) + \frac{1}{2}|ax + by|^2 \\ &\quad + \varepsilon\langle x^* - x, ax + by\rangle + \frac{\varepsilon}{2}(a + b)|x^* - x|^2\end{aligned}$$

is decreasing and, hence, bounded above. Since $\Phi(x)$ is bounded from below on C, the quantity

$$-|ax+by||x^*-x|+\frac{1}{2}(b+a)|x-x^*|^2,$$

which is less than $\langle x^*-x, ax+by\rangle+(1/2)(b+a)|x^*-x|^2$, is bounded from above; hence, $|x^*-x|$ is bounded because $\dot y = ax+by$ is bounded (Theorem (7.1)(ii)). From that we deduce that $\mathcal{E}$ is bounded below and admits a limit as $t\to+\infty$. Now in the expression of $\mathcal{E}$ the first three terms are known to have a limit, as $t\to+\infty$, hence, $|x^*-x|$ has a limit.

In order to apply Opial's lemma, we now show that any weak limit point x_∞ of x belongs to $\mathrm{Argmin}_C\,\Phi$. Let x^* be an element of $\mathrm{Argmin}_C\,\Phi$. Invoking the convexity of Φ and inequality (33), we have:

$$\Phi(x^*)\geqslant\Phi\big(x(t)\big)+\big\langle x^*-x,\nabla\Phi(x)\big\rangle,$$

$$\Phi(x^*)\geqslant\Phi\big(x(t)\big)+\frac{1}{\beta}\langle x^*-x,\dot y-\dot x\rangle+\frac{1}{\beta}\big\langle\dot x,\dot x+\beta\nabla\Phi(x)\big\rangle.$$

Since $|\dot x|+|\dot y|\to 0$ as $t\to+\infty$, and since (x^*-x) and $(\dot x+\beta\nabla\Phi(x))$ are bounded, we have:

$$\langle x^*-x,\dot y-\dot x\rangle+\big\langle\dot x,\dot x+\beta\nabla\Phi(x)\big\rangle\to 0,\quad t\to+\infty.$$

So, if t_n is a sequence going to infinity such that $x(t_n)$ weakly converges to x_∞, we have $\Phi(x^*)\geqslant\liminf\Phi(x(t_n))\geqslant\Phi(x_\infty)$. Hence, x_∞ is a minimizer of Φ on C, and Opial's lemma entails that $x(t)$ weakly converges to x_∞. □

The inertial aspect and the effect of the constraints in (c-DIN) are illustrated by a two-dimensional example (Fig. 2): $\Phi(x_1,x_2)=(1/2)\{(x_1+x_2+1)^2+4(x_1-x_2-1)^2\}$, $C=\mathbb{R}^{+2}$;

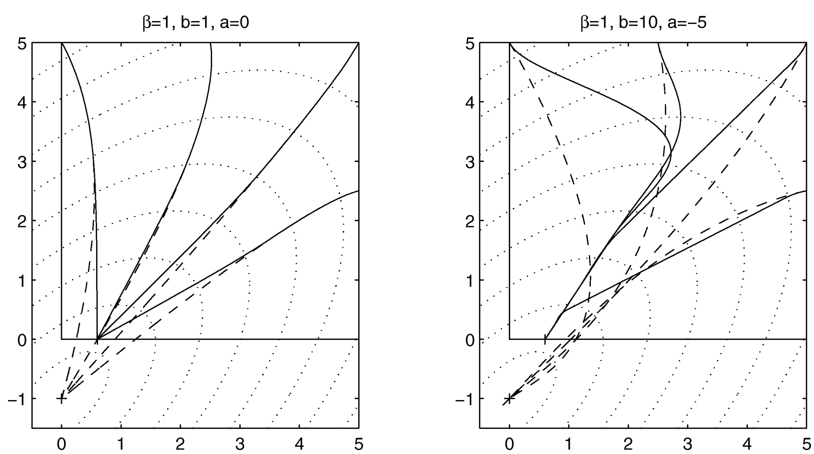

Fig. 2. A few trajectories of (c-DIN).

– the trajectories of (c-DIN) (continuous lines) converge to point $(3/5, 0)$, the minimum of Φ on C;
– in the absence of constraints, the trajectories (dashed lines) converge to $(0, -1)$, the minimum of Φ on $\mathbb{R}^2$.

8. Application to impact dynamics

In [28], Paoli and Schatzman have studied the system:

$$\begin{cases} \ddot{x}(t) + \partial\Psi_K\big(x(t)\big) \ni f\big(t, x(t), \dot{x}(t)\big), \\ \dot{x}(t^+) = -e\dot{x}_N(t^-) + \dot{x}_T(t^-) \quad \text{for any } t \text{ such that } x(t) \in \partial K, \end{cases} \tag{36}$$

where K is a closed convex subset of a finite-dimensional Hilbert space H, and $\partial\Psi_K$ is the subgradient set of the indicator function Ψ_K ($\Psi_K(x) = 0$ if $x \in K$ and $\Psi_K(x) = +\infty$ elsewhere). The first equation models the evolution of a mechanical system under the action of the force f, with state $x(t)$ subject to remain in K. The second equation models the instantaneous change in the system whenever its representative point $x(t)$ hits the boundary of K: the tangential velocity is conserved, while the normal velocity is reversed and multiplied by the *restitution coefficient* $e \in]0, 1]$; this rule accounts for a possible loss of energy at the impact.

Owing to Ψ_K being a definitely nonsmooth function, Paoli and Schatzman have to define a notion of solution to (36), and in order to prove the existence they introduce a regularized version obtained by a penalty method:

$$\ddot{x}_\lambda(t) + \frac{2\varepsilon}{\sqrt{\lambda}} G\big(\nabla\Psi_{K,\lambda}\big(x_\lambda(t)\big), \dot{x}_\lambda(t)\big) + \nabla\Psi_{K,\lambda}\big(x_\lambda(t)\big) = f\big(t, x_\lambda(t), \dot{x}_\lambda(t)\big). \tag{37}$$

The function $\Psi_{K,\lambda}(x) = (1/(2\lambda))\mathrm{dist}^2(x, K)$ is the usual Moreau–Yosida regularization of Ψ_K with parameter $\lambda > 0$, and the operator $G : H \times H \mapsto H$ is defined by $G(w, 0) = 0$ and $G(w, v) = \langle w, v/|v|\rangle v/|v|$ if $v \neq 0$. The constant $\varepsilon \in [0, +\infty[$ is related to e by $\varepsilon = -\log e/\sqrt{\pi^2 + \log^2 e}$. Passing to the limit $\lambda \to 0$ in (37) then yields a solution to (36).

We propose below a slightly different, and hopefully simpler, approach to (36). If K is a whole half-space, then it is not difficult to realize that $(1/\lambda)G(\nabla\Psi_{K,\lambda}(x), v)$ is exactly the Hessian $\nabla^2\Psi_{K,\lambda}(x)$ applied to v, except if x belongs to ∂K in which case $\nabla^2\Psi_{K,\lambda}(x)$ is not defined. When K is arbitrary, a formal, and bold, linearization of the boundary of K leads to replacement $G(\nabla\Psi_{K,\lambda}(x_\lambda(t)), \dot{x}_\lambda(t))$ in (37) by $\lambda\nabla^2\Psi_{K,\lambda}(x_\lambda(t))\dot{x}_\lambda(t)$, which gives:

$$\ddot{x}_\lambda(t) + 2\varepsilon\sqrt{\lambda}\,\nabla^2\Psi_{K,\lambda}\big(x_\lambda(t)\big)\dot{x}_\lambda(t) + \nabla\Psi_{K,\lambda}\big(x_\lambda(t)\big) = f\big(t, x_\lambda(t), \dot{x}_\lambda(t)\big).$$

For simplicity, assume henceforth that the exterior force reduces to a viscous friction: $f(t, x_\lambda(t), \dot{x}_\lambda(t)) = -\alpha\dot{x}_\lambda(t)$, $\alpha \geqslant 0$. The preceding equation becomes:

$$\ddot{x}_\lambda + \alpha\dot{x}_\lambda + 2\varepsilon\sqrt{\lambda}\,\nabla^2\Psi_{K,\lambda}(x)\dot{x}_\lambda + \nabla\Psi_{K,\lambda}(x) = 0.$$

This is (DIN) with $\beta = 2\varepsilon\sqrt{\lambda}$. But this equation has to be given a sense since $\Psi_{K,\lambda}$ is not twice differentiable everywhere. The cure is to write it in the form (g-DIN) which is of first-order in time and space (recall $\beta = 2\varepsilon\sqrt{\lambda}$):

$$\begin{cases} \dot{x}_\lambda + \beta\nabla\Psi_{K,\lambda}(x_\lambda) + \left(\alpha - \dfrac{1}{\beta}\right)x_\lambda + \dfrac{1}{\beta}y_\lambda = 0, \\ \dot{y}_\lambda + \left(\alpha - \dfrac{1}{\beta}\right)x_\lambda + \dfrac{1}{\beta}y_\lambda = 0. \end{cases} \tag{38}$$

This system is numerically solvable as it stands. A few numerical experiments are reported in Fig. 3: K is the unit disk, $\alpha = 0$, $\lambda = 10^{-4}$, the system representative point starts from position $(0.5, 0)$ with velocity $(0, 0.1)$; the coefficient $\beta = 2\varepsilon\sqrt{\lambda}$ runs through $\{0.02, 0.01, 0.008, 0.006, 0.004, 0.002, 0.001, 0.0001, 10^{-7}\}$, and correspondingly the restitution coefficient e runs through $\{0, 0.16, 0.25, 0.37, 0.53, 0.73, 0.85, 0.98, 0.99998\}$.

The experiments display the whole range of possible shocks:

– completely anelastic shocks for $\beta = 0.02$: after the first shock the trajectory follows the boundary;

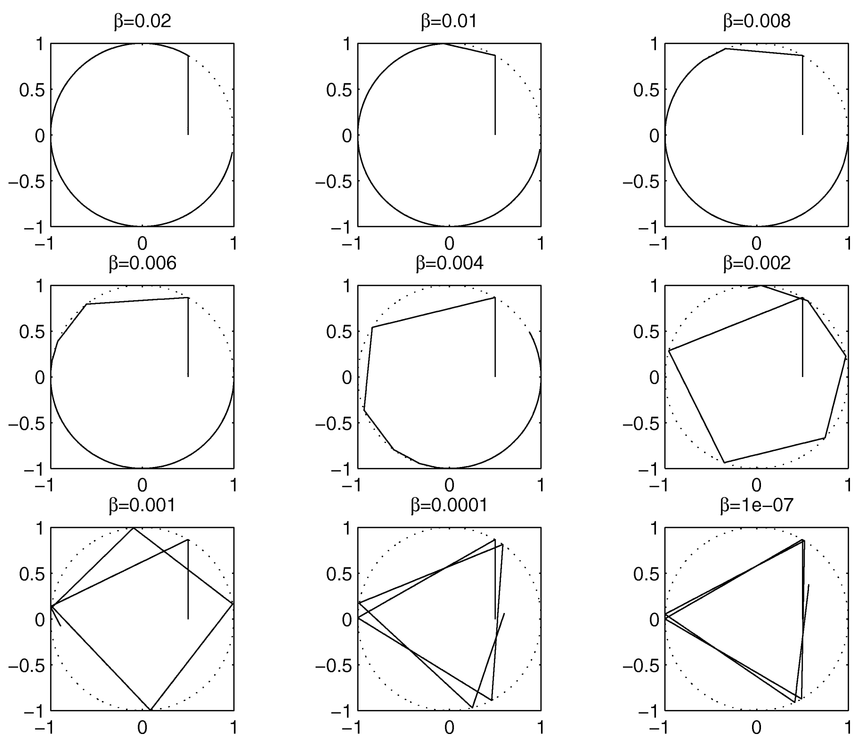

Fig. 3. Impacts in a disk.

- nearly perfectly elastic shocks for $\beta = 10^{-7}$ (the theoretical trajectory in the disk – without penalization – is an equilateral triangle);
- shocks with partial restitution of energy for intermediate values of β.

The purpose of these experiments is to illustrate the behaviour of the solutions of (38) and to suggest the latter as a theoretical regularization of (36). The numerical solution of (38) is prone to stiffness as λ becomes smaller (see [29] in this respect).

Additional literature [9,10].

References

[1] M. Aassila, A new approach of strong stabilization of distributed systems, Differential Integral Equations 11 (2) (1998) 369–376.
[2] F. Alvarez, On the minimizing property of a second-order dissipative system in Hilbert space, SIAM J. Control Optim. 38 (4) (2000) 1102–1119.
[3] F. Alvarez, H. Attouch, An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping, Set-Valued Anal. 9 (1/2) (2001) 3–11.
[4] F. Alvarez, J.M. Pérez, A dynamical system associated with Newton's method for parametric approximations of convex minimization problems, Appl. Math. Optim. 38 (1998) 193–217.
[5] A.S. Antipin, Linearization method, in nonlinear dynamic systems: Qualitative analysis and control (in Russian), Proc. ISA Russ. Acad. Sci., (2) (1994) 4–20; English translation: Comput. Math. Model. 8 (1) (1997) 1–15.
[6] A.S. Antipin, Minimization of convex functions on convex sets by means of differential equations (in Russian), Differ. Uravn. 30 (9) (1994) 1475–1486; English translation: Differential Equations 30 (9) (1994) 1365–1375.
[7] A.S. Antipin, A. Nedich, Continuous second-order linearization method for convex programming problems, Comput. Math. Cybern., Moscow University 2 (1996) 1–9.
[8] H. Attouch, F. Alvarez, The heavy ball with friction dynamical system for convex constrained minimization problems, in: Optimization, Namur, 1998, 25–35, in: Lecture Notes in Econom. and Math. Systems, Vol. 481, Springer-Verlag, Berlin, 2000.
[9] H. Attouch, A. Cabot, P. Redont, The dynamics of elastic shocks via epigraphical regularization of a differential inclusion. Barrier and penalty approximations, Adv. Math. Sci. Appl. 12 (1) (2002), in print.
[10] H. Attouch, R. Cominetti, A dynamical approach to convex minimization coupling approximation with the steepest descent method, J. Differential Equations 128 (2) (1996) 519–540.
[11] H. Attouch, X. Goudou, P. Redont, The heavy ball with friction method, I. The continuous dynamical system: global exploration of the global minima of a real-valued function by asymptotic analysis of a dissipative dynamical system, Commun. Contemp. Math. 2 (1) (2000) 1–34.
[12] H. Attouch, P. Redont, The second-order in time continuous Newton method, in: M. Lassonde (Ed.), Proceedings of the Fifth International Conference on Approximation and Optimization in the Caribbean, in: Approx. Optim. Math. Econ., Physica-Verlag, 2001, pp. 25–36.
[13] J.P. Aubin, A. Cellina, Differential Inclusions: Set-Valued Maps and Viability Theory, Springer-Verlag, Berlin, 1984.
[14] J.-B. Baillon, Un exemple concernant le comportement asymptotique de la solution du problème $du/dt + \partial\varphi(u) \ni 0$, J. Funct. Anal. 28 (1978) 369–376.
[15] H. Brézis, Opérateurs Maximaux Monotones, Math. Studies, Vol. 5, North-Holland, Amsterdam, 1973.
[16] R.E. Bruck, Asymptotic convergence of nonlinear contraction semigroups in Hilbert space, J. Funct. Anal. 18 (1975) 15–26.
[17] J.K. Hale, Asymptotic behavior of dissipative systems, Math. Surveys Monogr., Vol. 25, AMS, Providence, RI, 1987.
[18] A. Haraux, Semilinear hyperbolic problems in bounded domains, in: J. Dieudonné (Ed.), Mathematical Reports 3, Part I, Harwood Academic, Reading, UK, 1987.

[19] A. Haraux, Systèmes Dynamiques Dissipatifs et Applications, RMA, Vol. 17, Masson, Paris, 1991.

[20] A. Haraux, M. Jendoubi, Convergence of solutions of second-order gradient-like systems with analytic nonlinearities, J. Differential Equations 144 (2) (1998).

[21] M. Jendoubi, Convergence of bounded and global solutions of the wave equation with linear dissipation and analytic nonlinearity, J. Differential Equations 144 (1998) 302–312.

[22] B. Kummer, Generalized Newton and NCP Methods: convergence, regularity, actions, Discuss. Math. Differential Incl., Control Optim. 20 (2000) 209–244.

[23] C. Lemaréchal, J.-B. Hiriart-Urruty, Convex Analysis and Minimization Algorithms II: Advanced Theory and Bundle Methods, Springer-Verlag, Berlin, 1993.

[24] S. Lojasiewicz, Une propriété topologique des sous-ensembles analytiques réels, in: Colloques Internationaux du CNRS. Les équations aux dérivées partielles, Vol. 117, 1963.

[25] S. Lojasiewicz, Ensembles semi-analytiques, IHES Notes (1965).

[26] B. Mordukhovich, J. Outrata, On second-order subdifferentials and their applications, SIAM J. Optim., 12, (1) 139–169.

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

[28] L. Paoli, M. Schatzman, Mouvement à un nombre fini de degrés de liberté avec contraintes unilatérales: cas avec perte d'énergie, Mod. Math. Anal. Num. 27 (1993) 673–717.

[29] L. Paoli, M. Schatzman, Ill-posedness in vibro-impact and its numerical consequences, ECOMAS conference 2000, Proceedings, to appear.

[30] J. Palis, W. de Melo, Geometric Theory of Dynamical Systems, Springer-Verlag, 1982.

[31] B.T. Polyak, Introduction to Optimization, Optimization Software, New York, 1987.

[32] R. Rockfellar, R. Wets, Variational Analysis, Springer-Verlag, Berlin, 1998.

[33] L. Simon, Asymptotics for a class of non-linear evolution equations, with applications to geometric problems, Ann. Math. 118 (1983) 525–571.

[34] S. Smale, A convergent process of price adjustment and global Newton methods, J. Math. Econ. 3 (1976) 107–120.

[35] R. Temam, Infinite dimensional dynamical systems in mechanics and physics, Appl. Math. Sci. 68 (1997).