Mathematical Modelling and Analysis Publisher: Taylor&Francis and VGTU
Volume 18 Number 5, November 2013, 654–674 http://www.tandfonline.com/TMMA
http://dx.doi.org/10.3846/13926292.2013.868842 Print ISSN: 1392-6292
Online ISSN: 1648-3510

Asymptotics for a Dissipative Dynamical System with Linear and Gradient-Driven Damping

Yuhu Wu[a] **and Xiaoping Xue**[b]

[a] *Harbin University of Science and Technology*
150080 Harbin, China
[b] *Harbin Institute of Technology*
150001 Harbin, China
E-mail(*corresp.*): wuyuhu51@gmail.com
E-mail: xiaopingxue@263.net

Received October 22, 2012; revised October 29, 2013; published online December 1, 2013

Abstract. We study, in the setting of a real Hilbert space H, the asymptotic behavior of trajectories of the second-order dissipative dynamical system with linear and gradient-driven nonlinear damping

$$\ddot{x}(t) + \lambda\dot{x}(t) + \nabla f\big(\dot{x}(t)\big) + \nabla\Phi\big(x(t)\big) = 0,$$

where $\lambda > 0$ and $f, \Phi : H \to R$ are two convex differentiable functions. It is proved that if Φ is coercive and bounded from below, then the trajectory converges weakly towards a minimizer of Φ. In particular, we state that under suitable conditions, the trajectory strongly converges to the minimizer of Φ exponentially or polynomially.

Keywords: dissipative dynamical systems, nonlinear damping, asymptotic behavior, convex minimization.

AMS Subject Classification: 37L05; 37N05; 37N40; 35B40.

1 Introduction

Let H be a real Hilbert space, with scalar product and corresponding norm respectively denoted by $\langle\cdot,\cdot\rangle$ and $|\cdot|$. The gradient-like system to be considered in this paper is the following second order evolution equation

$$(GD) \quad \ddot{x}(t) + \lambda\dot{x}(t) + \nabla f\big(\dot{x}(t)\big) + \nabla\Phi\big(x(t)\big) = 0,$$

where $\lambda > 0$ is a constant and $f, \Phi : H \to R$ are two convex differentiable continuous maps. The above dynamical system can be regarded as a nonlinear oscillator involving a linear damping $\lambda\dot{x}(t)$ with a nonlinear perturbation $\nabla f(\dot{x}(t))$, which is a gradient driven damping.

In [21], Polyak introduced the heavy ball with friction (*HBF*) dynamical system

$$(HBF) \quad \ddot{x}(t) + \lambda \dot{x}(t) + \nabla \Phi\big(x(t)\big) = 0.$$

It can be studied in the classical framework of the theory of dissipative dynamical system (Hale [16], Haraux [18]). In 2000, Attouch, Guodou and Rendont [6] employed it to deal with optimization problem. The generalized (*HBF*) system with nonsmooth potential Φ has been studied in the differential inclusion framework (Cabot and Paoli [14]). Recently, Attouch and Mainge [7] investigated asymptotic behavior of second-order dissipative evolution equations combining potential with non-potential effects. Moreover, Bhaya et al. [9] designed (*HBF*) network to solve the global nonconvex minimization problem.

It is well known that the steepest descent method is a classical optimization algorithm, which is a first-order system

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

From the optimization point of view, the main feature of the (*HBF*) system departing from the (*SD*) systems lies in the fact that, it is no more a descent method: its trajectory is allowed to go up and down along the graph of Φ. See also [1, 4, 5, 8, 13, 15] for related results in this connection.

In this paper, based on the dynamic analysis of the motion of heavy ball with a propulsive force (see Section 2), we derived the (*GD*) system, which is a rational extension of (*HBF*) system. It is worth to mention that there are many real physical and economics models, which can be described by (*GD*) system [12,23]. For example, seismic wave with strong nonlinearities [3] is given by

$$(SW) \quad \ddot{x} + \lambda \dot{x} + \mu \dot{x}^3 + n^2 x + k^2 x^3 = 0.$$

It includes a nonlinear cubic term in damping. If $k = 0$, it becomes the well-known Rayleigh nonlinear ODE (see [22]). If we set $f(\dot{x}) = \frac{1}{4}\mu \dot{x}^4$ and $\Phi(x) = \frac{1}{2}n^2x^2 + \frac{1}{4}k^2x^4$, then equation (*SW*) can be considered in the framework of system (*GD*).

This paper will be dedicated to investigate the asymptotical behavior of (*GD*) system. We first state that each trajectory will weakly converge to a minima point of Φ, if Φ is coercive, convex and bounded from below. This optimization result has been established in the case of (*HBF*) system by Alvarez [1] (also see [4, 6]), while in case of (*SD*) system it is the well known Bruck's theorem [11].

A more realistic problem is whether the trajectory strongly converge to a minima point of Φ and how fast does it converge. We first state a "relaxed" version of strongly convex condition for Φ:

$$\exists \alpha > 0,\ m > 1 \quad \text{s.t.} \quad \big\langle \nabla \Phi(x) - \nabla \Phi(y), x - y \big\rangle \geq \alpha |x - y|^m, \quad \forall x, y \in H.$$

If Φ is strongly convex (which means there exists $\langle \nabla \Phi(x) - \nabla \Phi(y), x - y\rangle \geq \alpha |x-y|^2$, $\forall x, y \in H$), the trajectories of both (*SD*) and (*HBF*) systems are known to exponentially converge [5,10]. However, as far as our knowledge goes, there is no work dealing with the case of $m \neq 2$.

We will treat the general case $m > 1$, which was not covered in the previous literatures. It is proved that the trajectory of (GD) will exponentially converge if $2 \geq m > 1$ and polynomially converge if $m > 2$. This shows that the case $m > 2$ is indeed different from the case $2 \geq m > 1$, the system does possess a trajectory which is not exponentially convergent (see Example 1). A reflection of difference between the case $m > 2$ and the case $2 \geq m > 1$ is that, in the proof process, the two cases need different method to be solved.

The rest of the paper is organized as follows. In Section 2, we give the mechanical interpretation of the (GD) system. Section 3 will be devoted to study optimization property and asymptotic weak convergence of the trajectory. In Section 4, we state the strongly convergent results and present some illustrative examples.

2 The Mechanical Interpretation

Suppose H is a Hilbert space and Φ is a continuous differentiable function on H. Consider a material point M of mass m with a propulsive force, which moves on the manifold defined by $\Sigma := Graph(\Phi) \subset H \times R$. We use $p(t) \in H \times R$ to denote the position of M at time t:

$$p(t) = \begin{pmatrix} x(t) \\ \Phi(x(t)) \end{pmatrix} \quad \text{if } x(t) \in H.$$

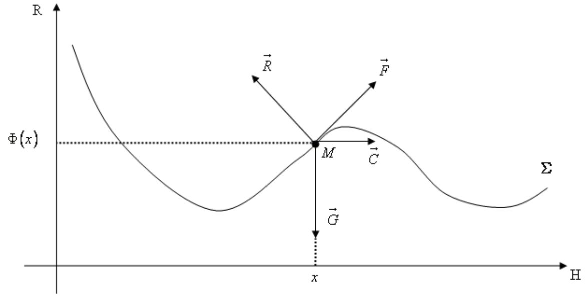

Figure 1. Force diagram of the material point M.

According to the Newton's second law, the motion of M is governed by the equation:

$$(NSL) \quad m\ddot{p} = \overrightarrow{G} + \overrightarrow{F} + \overrightarrow{R} + \overrightarrow{C}, \tag{2.1}$$

where $\ddot{p}$ denotes the acceleration of M and the right hand side of this equality is the sum of the forces which are exerted on M (see Figure 1 for the force diagram):

- The gravity force $\overrightarrow{G} = (0, -mg)^T$.
- The viscous friction force, is opposed to the movement of the particle (friction) and proportional to the speed (viscous friction): $\overrightarrow{F} = -\lambda\dot{p}(t)$, where $\lambda > 0$ is the friction coefficient.

- The reaction force $\overrightarrow{R}$ of the surface Σ, which expresses that the particle does not penetrate into Σ : $\overrightarrow{R} = R\overrightarrow{n}$, where $R > 0$ and $\overrightarrow{n}$ is the outwards unitary normal to Σ at M. Classically,

$$\overrightarrow{n} = \frac{1}{\sqrt{1+|\nabla\Phi(x)|^2}}\bigl(-\nabla\Phi(x), 1\bigr)^T.$$

- Propulsive force supplied by a thruster, which is determined by its horizontal velocity, expressed as $\overrightarrow{C} = (-\nabla f(\dot{x}(t)), 0)^T$.

Since $p(t) = (x(t), \Phi(x(t)))^T$, we have

$$\dot{p}(t) = \begin{pmatrix} \dot{x}(t) \\ \nabla\Phi(x(t))\dot{x}(t) \end{pmatrix}, \qquad \ddot{p}(t) = \begin{pmatrix} \ddot{x}(t) \\ \langle \dot{x}(t), \nabla^2\Phi(x(t))\dot{x}(t)\rangle + \nabla\Phi(x(t))\ddot{x}(t) \end{pmatrix},$$

where ∇^2 is the Hessian operator. Then equation (NSL) gives us

$$\begin{cases} m\ddot{x} = -\lambda\dot{x}(t) - RΛ(t)\nabla\Phi(x) - \nabla f\bigl(\dot{x}(t)\bigr), \\ m\bigl(\bigl\langle \dot{x}, \nabla^2\Phi(x)\dot{x}\bigr\rangle + \nabla\Phi\bigl(x(t)\bigr)\ddot{x}(t)\bigr) = -mg - \lambda\nabla\Phi\bigl(x(t)\bigr)\dot{x}(t) + R\Lambda(t), \end{cases} \tag{2.2}$$

where $\Lambda(t) = \frac{1}{\sqrt{1+|\nabla\Phi(x)|^2}}$. We can rewrite (NSL) as

$$m\ddot{p} - \overrightarrow{G} - \overrightarrow{F} - \overrightarrow{C} = R\overrightarrow{n}, \tag{2.3}$$

and project (2.3) on $\overrightarrow{n}$, then we have that $\langle m\ddot{p} - \overrightarrow{G} - \overrightarrow{F} - \overrightarrow{C}, \overrightarrow{n}\rangle = R$, i.e.

$$\begin{pmatrix} m\ddot{x} + \lambda\dot{x} + \nabla f(\dot{x}) \\ m(\langle \dot{x}, \nabla^2\Phi(x)\dot{x}\rangle + \nabla\Phi(x)\ddot{x} + g) + \lambda\nabla\Phi(x)\dot{x} \end{pmatrix}^T \begin{pmatrix} -\Lambda(t)\nabla\Phi(x) \\ \Lambda(t) \end{pmatrix} = R.$$

This leads to

$$R = \Lambda(t)\bigl(-\bigl\langle \nabla\Phi(x), \nabla f(\dot{x})\bigr\rangle + m\dot{x}\nabla^2\Phi(x)\dot{x} + mg\bigr).$$

Together with the first equation of (2.2), we obtain

$$\begin{aligned} &m\ddot{x} + \lambda\dot{x} + \Lambda^2 m\bigl(\bigl\langle \dot{x}, \nabla^2\Phi(x)\dot{x}\bigr\rangle + g\bigr)\nabla\Phi(x) \\ &\quad - \Lambda^2\bigl\langle \nabla\Phi(x), \nabla f(\dot{x})\bigr\rangle\nabla\Phi(x) + \nabla f(\dot{x}) = 0. \end{aligned} \tag{2.4}$$

Being interested in the asymptotic behavior of $x(t)$, we may assume that

$$(\mathcal{H}) \quad \begin{cases} \bullet \quad |\nabla\Phi(x)| \text{ is negligible with respect to } 1, \\ \bullet \quad \langle \dot{x}, \nabla^2\Phi(x)\dot{x}\rangle \text{ is negligible with respect to } g, \\ \bullet \quad \nabla f(\dot{x}) \text{ is bounded,} \end{cases}$$

which give us the following system:

$$m\ddot{x} + \lambda\dot{x} + mg\nabla\Phi(x) + \nabla f(\dot{x}) = 0.$$

Now, let us define $m = 1$ and $g = 1$, then the final (GD) system can be obtained:

$$\ddot{x} + \lambda\dot{x} + \nabla\Phi(x) + \nabla f(\dot{x}) = 0.$$

Remark 1. Indeed, the mechanical model of (GD) system is a development of (HBF) model. In fact, if the Propulsive force $\overrightarrow{C}$ is not valid, then the resulting force acting on the material point become sum of three forces ($\overrightarrow{G}$, $\overrightarrow{F}$ and $\overrightarrow{R}$). So from (NSL) we can reduce the (HBF) model. For more details, we refer readers to [5] and [6].

Looking into the assumption $(\mathcal{H})$, we know that the equation only possesses a mechanical sense when velocity term $\dot{x}$ is small. Despite this, the mechanical interpretation of (GD) system will help to intuitively understand its asymptotic behavior.

3 Weak Convergence Results

Let H be a real Hilbert space. Let us consider two maps $f : H \to R$ and $\Phi : H \to R$ satisfying the following assumptions respectively:

$$(\mathcal{A}_\Phi) \quad \begin{cases} (\mathcal{A}_\Phi - i) & \Phi \text{ is of class } C^1 \text{ and } \Phi \text{ is bounded from below;} \\ (\mathcal{A}_\Phi - ii) & \nabla\Phi \text{ is Lipschitz continuous on the bounded subsets of } H. \end{cases}$$

$$(\mathcal{A}_f) \quad \begin{cases} (\mathcal{A}_f - i) & f \text{ is convex, of class } C^2 \text{ on } H \text{ and } \min\limits_{x\in H} f(x) = f(0) = 0; \\ (\mathcal{A}_f - ii) & \nabla f \text{ is Lipschitz continuous on the bounded subsets of } H. \end{cases}$$

Define the mechanical energy along trajectory $x(t)$ of (GD) by

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

By differentiation of the energy function we obtain

$$\dot{E}(t) = \big\langle \dot{x}(t), \ddot{x}(t)\big\rangle + \big\langle \dot{x}(t), \nabla\Phi\big(x(t)\big)\big\rangle.$$

In view of $\ddot{x}(t) + \lambda\dot{x}(t) + \nabla f(\dot{x}(t)) + \nabla\Phi(x(t)) = 0$ and hypothesis $(\mathcal{A}_f - i)$, we deduce that

$$\begin{aligned} \dot{E}(t) &= -\lambda\big|\dot{x}(t)\big|^2 - \big\langle \nabla f\big(\dot{x}(t)\big), \dot{x}(t)\big\rangle \\ &= -\lambda\big|\dot{x}(t)\big|^2 - \big\langle \nabla f\big(\dot{x}(t)\big) - \nabla f(0), \dot{x}(t) - 0\big\rangle \le -\lambda\big|\dot{x}(t)\big|^2. \end{aligned} \tag{3.1}$$

Thus, the function $E(\cdot)$ is decreasing.

Now, using classical Cauchy–Lipschitz theorem we state the global existence result.

Theorem 1 [Existence and Uniqueness]. *Assume hypotheses $\mathcal{A}_\Phi$ and $\mathcal{A}_f$ hold. Then, the following properties hold*:

(i) *for each $(x_0, \dot{x}_0) \in H \times H$, there exists a unique global solution $x(t)$ of (GD), which is of class C^2 on $[0, +\infty)$ and satisfies the initial conditions $x(0) = x_0$ and $\dot{x}(0) = \dot{x}_0$;*

(ii) *for every trajectory $x(\cdot)$ of (GD), the corresponding energy $E(t)$ is decreasing on $[0, +\infty)$ and bounded from below, and hence converges to some real value E_∞. Moreover,*

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

Proof. (i) First we note that the system (GD) can be written as a first order autonomous system in $H \times H : \dot{Y} = G(Y)$ with

$$Y(t) = \begin{pmatrix} x(t) \\ \dot{x}(t) \end{pmatrix}, \qquad G(u,v) = \begin{pmatrix} v \\ -\lambda v(t) - \nabla f(v(t)) - \nabla \Phi(u(t)) \end{pmatrix}.$$

For $Y_0 = (x_0, \dot{x}_0)^T$ given in $H \times H$, since the maps $\nabla \Phi$ and ∇f are locally Lipschitzian, the Cauchy–Lipschitz theorem ensures the existence of a unique local solution to the problem:

$$\begin{cases} \dot{Y}(t) = G(Y), \\ Y(0) = Y_0. \end{cases}$$

Let $x(\cdot)$ be the corresponding maximal solution, defined on $[0, T_{\max})$ with $0 < T_{\max} \le +\infty$. It is clear that the solution $x(\cdot)$ is of class C^2 on $[0, T_{\max})$ by the continuity of the map $\nabla \Phi$ and ∇f. Suppose, contrary to our claim, that $T_{\max} < +\infty$. By (3.1), we have $E(t) \le E(0)$, for all $t \in [0, T_{\max})$. That is,

$$\frac{1}{2}\left|\dot{x}(t)\right|^2 + \Phi\big(x(t)\big) \le E(0). \tag{3.2}$$

Since Φ is bounded from below, we obtain that

$$\left|\dot{x}(t)\right|_{L^\infty} = \sup_{t \in [0, T_{\max})} \left|\dot{x}(t)\right| < +\infty. \tag{3.3}$$

Then, according to $|x(t_1) - x(t_2)| \le |\dot{x}(t)|_{L^\infty} |t_1 - t_2|$ and $T_{\max} < +\infty$, we know that $\lim_{t \to T_{\max}} x(t) := x_\infty$ exists. Hence, $x(\cdot)$ and $\dot{x}(\cdot)$ are bounded on $[0, T_{\max})$, and by equation (GD), $\ddot{x}(\cdot)$ is also bounded on this interval. It implies that $\lim_{t \to T_{\max}} \dot{x}(t) := \dot{x}_\infty$ exists. But, applying again the local existence theorem with initial data $(x_\infty, \dot{x}_\infty)$, we can extend the maximal solution to an interval strictly larger than $[0, T_{\max})$, which contradicts the maximality of the solution. So, $T_{\max} = +\infty$. It completes the proof of (i).

(ii) It has been proved that $E(\cdot)$ is decreasing. Since $E(t) \ge \Phi(x(t))$ and Φ is bounded from below, the energy $E(\cdot)$ is also bounded from below. As a consequence, $\lim_{t \to +\infty} E(t) = E_\infty$ exists, with $E_\infty \in R$. According to (3.2) and the fact that Φ is bounded from below, we get that,

$$\frac{1}{2}\left|\dot{x}(t)\right|^2 \le E(0) - \inf \Phi, \quad \forall t \ge 0.$$

It means that $\dot{x} \in L^\infty(0, +\infty; H)$.

By the inequality (3.1), we obtain that, for all $0 \le t < +\infty$,

$$\int_0^t \left|\dot{x}(s)\right|^2 ds \le \frac{1}{\lambda}\big(E(0) - E(t)\big).$$

Since $E(t)$ decrease to E_∞ as t tends to $+\infty$, we know that

$$\int_0^{+\infty} \left|\dot{x}(s)\right|^2 ds \le \frac{1}{\lambda}\big(E(0) - E_\infty\big)$$

and $\dot{x} \in L^2(0,+\infty;H)$. It completes the proof of (ii). $\square$

After establishing the existence result of solution, we then focus our attention on the asymptotic behavior of the system and its optimization property. We first give the following theorem.

Theorem 2. *Under hypothesis $\mathcal{A}_\Phi$ and $\mathcal{A}_f$, let $x(\cdot)$ be the unique solution of (GD) from a given initial state $(x_0,\dot{x}_0) \in H \times H$. Assume moreover that $x \in L^\infty(0,+\infty;H)$, then we have:*

- *$\dot{x}$, $\ddot{x}$ and $\nabla\Phi(x)$ belong to $L^\infty(0,+\infty;H)$;*
- *$\lim_{t\to+\infty} \dot{x}(t) = 0$ and $\lim_{t\to+\infty} \Phi(x(t)) = E_\infty$;*
- *if Φ is of class C^2, then $\lim_{t\to+\infty} \ddot{x}(t) = \lim_{t\to+\infty} \nabla\Phi(x(t)) = 0$.*

Proof. We already stated in Theorem 1 that $\dot{x} \in L^\infty(0,+\infty;H)$. The equation (GD), and the fact that $\nabla\Phi$ and ∇f are bounded on the bounded subsets of H, imply that $\ddot{x} \in L^\infty(0,+\infty;H)$. Let $u(t) = \dot{x}(t)$, then we have

$$u \in L^2(0,+\infty;H) \quad \text{and} \quad \dot{u} \in L^\infty(0,+\infty;H),$$

which imply that

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

By $\Phi(x(t)) = E(t) - \frac{1}{2}|\dot{x}(t)|^2$, it is clear that

$$\lim_{t\to+\infty} \Phi\big(x(t)\big) = \lim_{t\to+\infty} E(t) = E_\infty.$$

It follows from (GD) and $\nabla f(0) = 0$ that

$$\lim_{t\to+\infty} \big[\ddot{x}(t) + \nabla\Phi\big(x(t)\big)\big] = 0.$$

If we are able to prove that $\lim_{t\to+\infty} \ddot{x}(t) = 0$, then we automatically infer that $\lim_{t\to+\infty} \nabla\Phi(x(t)) = 0$.

Let us now prove that $\lim_{t\to+\infty} \ddot{x}(t) = 0$. Returning to equation (GD), since Φ is of class C^2, we have that the solution x in of class C^3. By differentiating equation (GD), we obtain

$$\dddot{x}(t) + \lambda\ddot{x}(t) + \nabla^2 f\big(\dot{x}(t)\big)\ddot{x}(t) + \nabla^2\Phi\big(x(t)\big)\dot{x}(t) = 0.$$

Let $y(t) := \ddot{x}(t)$, then we rewrite above equation as

$$\dot{y}(t) = -\big[\lambda + \nabla^2 f\big(\dot{x}(t)\big)\big]y(t) - \varphi(t) \tag{3.4}$$

with $\varphi(t) = \nabla^2\Phi(x(t))\dot{x}(t)$.

By f is of C^2 and convex, we easily deduce that the Hessian operator $\nabla^2 f$ is bilinear and positive semidefinite, that is $\langle\nabla^2 f(x)y, y\rangle \geq 0$, for all $x, y \in H$. Now set $z(t) = \frac{1}{2}|y(t)|^2$, then equation (3.4) implies

$$\begin{aligned}\dot{z}(t) &= \big\langle\dot{y}(t), y(t)\big\rangle = -\lambda\big|y(t)\big|^2 - \big\langle\nabla^2 f\big(\dot{x}(t)\big)y(t), y(t)\big\rangle - \big\langle\varphi(t), y(t)\big\rangle \\ &\leq -2\lambda z(t) + \big|\varphi(t)\big|\big|y(t)\big|.\end{aligned} \tag{3.5}$$

Since $\nabla\Phi$ is Lipschitz continuous on the bounded subsets of H and $x(t) \in L^\infty(0,\infty;H)$, we have $\nabla^2\Phi(x(t))$ is bounded, that is, there exists $d_1 > 0$ such that $|\nabla^2\Phi(x(t))| < d_1$, for $t > 0$. Hence

$$\left|\varphi(t)\right| = \left|\nabla^2\Phi\big(x(t)\big)\dot{x}(t)\right| \le d_1\left|\dot{x}(t)\right|, \quad t > 0.$$

Moreover, according to $\ddot{x}(t) \in L^\infty(0,+\infty;H)$, there exists $d_2 > 0$ such that, for all $t > 0$, $|y(t)| = |\ddot{x}(t)| < d_2$. Hence, equation (3.5) becomes

$$\dot{z}(t) \le -2\lambda z(t) + d_1 d_2\left|\dot{x}(t)\right|. \tag{3.6}$$

By comparison principle, for all $t > 0$

$$0 \le z(t) \le z(0)e^{-2\lambda t} + d_1 d_2 \int_0^t e^{2\lambda(\rho-t)}\left|\dot{x}(\rho)\right| d\rho. \tag{3.7}$$

Now we prove that

$$\lim_{t\to+\infty} \int_0^t e^{2\lambda(\rho-t)}\left|\dot{x}(\rho)\right| d\rho = 0. \tag{3.8}$$

Any given $\varepsilon > 0$, since $\dot{x} \in L^2(0,+\infty;H)$, there exists $T_1 \ge 0$ such that, for $t > T_1$

$$\int_{T_1}^t \left|\dot{x}(\rho)\right|^2 d\rho \le (2\lambda\varepsilon)^2. \tag{3.9}$$

On the other hand, there exists $T_2 \ge 0$, such that, for $t \ge T_2$

$$e^{-2\lambda t} < \lambda\varepsilon/|\dot{x}|_\infty. \tag{3.10}$$

Let $T_3 = \max\{T_1, T_2\}$. Then, for $t > 2T_3$

$$\begin{aligned}
\int_0^t e^{2\lambda(\rho-t)}\left|\dot{x}(\rho)\right| d\rho &= \int_0^{T_3} e^{2\lambda(\rho-t)}\left|\dot{x}(\rho)\right| d\rho + \int_{T_3}^t e^{2\lambda(\rho-t)}\left|\dot{x}(\rho)\right| d\rho \\
&\le |\dot{x}|_\infty \frac{e^{2\lambda(T_3-t)} - e^{-2\lambda t}}{2\lambda} + \left[\int_{T_3}^t \left|\dot{x}(\rho)\right|^2 d\rho\right]^{\frac{1}{2}} \left[\int_{T_3}^t e^{4\lambda(\rho-t)}\, d\rho\right]^{\frac{1}{2}} \\
&\le |\dot{x}|_\infty \frac{e^{-2\lambda T_3}}{2\lambda} + \frac{\varepsilon(1 - e^{\lambda(T_3-t)})}{2} < \varepsilon,
\end{aligned} \tag{3.11}$$

by equation (3.9), (3.10) and Hölder inequality. By arbitrariness of ε, we prove that (3.8).

Finally, returning to (3.7) and noticing that $z(t) = \frac{1}{2}|y(t)|^2$, we obtain

$$\lim_{t\to\infty} \ddot{x}(t) = \lim_{t\to\infty} y(t) = 0.$$

Consequently, we finish the proof of Theorem 2. □

Under the additional assumption that the map Φ is coercive, we obtain that the solution x is bounded.

Corollary 1. Under assumptions $\mathcal{A}_\Phi$ and $\mathcal{A}_f$, additionally assume that:

$$(\mathcal{A}_\Phi - iii) \quad \Phi \text{ is coercive, i.e. } \lim_{|x|\to+\infty} \Phi(x) = +\infty,$$

then the solution x is in $L^\infty([0,+\infty;H)$, hence all conclusions of Theorem 2 hold.

Proof. Since the energy function $E(\cdot)$ is decreasing, so

$$\frac{1}{2}\big|\dot{x}(t)\big|^2 + \Phi\big(x(t)\big) = E(t) \le E(0), \quad \text{for all } t \ge 0,$$

which implies that $\Phi(x(t))$ is bounded from above. Together with the coerciveness of Φ, we deduce that the trajectory $x(t)$ remains bounded, i.e. $x \in L^\infty(0,+\infty;H)$. □

We are now able to present the optimization property of (GD) system, under following local growth condition near origin for f.

$$(\mathcal{A}_f - iii) \quad \text{there exist } c_1 \ge 0,\ r > 0 \text{ s.t. } \big|\nabla f(x)\big| \le c_1|x|^2, \quad \forall x \in B(0,r).$$

Theorem 3. *Under assumptions $\mathcal{A}_f$ and $\mathcal{A}_\Phi$, additionally assume that Φ is coercive, convex and f satisfies condition $(\mathcal{A}_f - iii)$. If $x \in C^2(0,+\infty;H)$ is the unique solution of (GD) from a given initial state $(x_0,\dot{x}_0) \in H \times H$, then*

$$\lim_{t\to\infty} \Phi\big(x(t)\big) = \inf \Phi. \tag{3.12}$$

Proof. In order to prove the minimizing property (3.12), it is sufficient to prove that for any $x_0 \in H$,

$$\limsup_{t\to\infty} \Phi\big(x(t)\big) \le \Phi(x_0).$$

Fix $x_0 \in H$ and define the auxiliary function $g(t) := \frac{1}{2}|x(t) - x_0|^2$. Since $x(\cdot)$ is a solution of (GD), it follows that

$$\ddot{g}(t) + \lambda\dot{g}(t) = \big\langle \nabla\Phi\big(x(t)\big), x_0 - x(t)\big\rangle + \big\langle \nabla f\big(\dot{x}(t)\big), x_0 - x(t)\big\rangle + \big|\dot{x}(t)\big|^2.$$

Noting that Φ is convex, the above equation yields

$$\ddot{g}(t) + \lambda\dot{g}(t) \le \Phi(x_0) - \Phi\big(x(t)\big) + \big\langle \nabla f\big(\dot{x}(t)\big), x_0 - x(t)\big\rangle + \big|\dot{x}(t)\big|^2. \tag{3.13}$$

Since Φ is coercive, by Corollary 1, we know that $x(t) \in L^\infty(0,+\infty;H)$. So there exists some constant $M_0 > 0$ such that $|x(t) - x_0| \le M_0$, for all $t \in [0,+\infty)$. Since $\lim_{t\to\infty} \dot{x}(t) = 0$, there exists $T > 0$, such that $|\dot{x}(t)| \le r$, for $t \ge T$. Then, by assumption $(\mathcal{A}_f - iii)$, we get $|\nabla f(\dot{x}(t))| \le c_1|\dot{x}(t)|^2$, for $t \ge T$. Hence, we obtain

$$\ddot{g}(t) + \lambda\dot{g}(t) \le \Phi(x_0) - \Phi\big(x(t)\big) + (c_1 M_0 + 1)\big|\dot{x}(t)\big|^2. \tag{3.14}$$

Furthermore, by $E(t) = \frac{1}{2}|\dot{x}(t)|^2 + \Phi(x(t))$, we rewrite (3.14) as

$$\ddot{g}(t) + \lambda\dot{g}(t) \le \Phi(x_0) - E(t) + \left(c_1 M_0 + \frac{3}{2}\right)\left|\dot{x}(t)\right|^2. \tag{3.15}$$

Given $t > 0$, since $E(\cdot)$ is nonincreasing, we get for all $\rho \in [0, t]$

$$\ddot{g}(\rho) + \lambda\dot{g}(\rho) \le \Phi(x_0) - E(t) + \left(c_1 M_0 + \frac{3}{2}\right)\left|\dot{x}(\rho)\right|^2. \tag{3.16}$$

After multiplying it by $e^{\lambda\rho}$ and integrating it over $[0, t]$, we obtain

$$\begin{aligned}\dot{g}(t) \le e^{-\lambda t}\dot{g}(0) + \frac{1 - e^{-\lambda t}}{\lambda}\big(\Phi(x_0) - E(t)\big) \\ + \left(c_1 M_0 + \frac{3}{2}\right)\int_0^t e^{-\lambda(t-\rho)}\left|\dot{x}(\rho)\right|^2 d\rho.\end{aligned} \tag{3.17}$$

Note again that $E(t)$ is nonincreasing, for any $\sigma \in [0, t]$, we have

$$\begin{aligned}\dot{g}(\sigma) \le e^{-\lambda\sigma}\dot{g}(0) + \frac{1 - e^{-\lambda\sigma}}{\lambda}\big(\Phi(x_0) - E(t)\big) \\ + \left(c_1 M_0 + \frac{3}{2}\right)\int_0^\sigma e^{-\lambda(\sigma-\rho)}\left|\dot{x}(\rho)\right|^2 d\rho.\end{aligned}$$

Integrating it once more over $[0, t]$,

$$g(t) - g(0) \le \frac{1 - e^{-\lambda t}}{\lambda}\dot{g}(0) + \frac{\lambda t - 1 + e^{-\lambda t}}{\lambda^2}\big(\Phi(x_0) - E(t)\big) + \left(c_1 M_0 + \frac{3}{2}\right)a(t),$$

where

$$a(t) = \int_0^t \int_0^\sigma e^{-\lambda(\sigma-\rho)}\left|\dot{x}(\rho)\right|^2 d\rho\, d\sigma.$$

Since $\Phi(x(t)) \le E(t)$ and $g(t)$ is non-negative, it follows that

$$\Phi\big(x(t)\big) \le \frac{\lambda^2}{\lambda t - 1 + e^{-\lambda t}}\left(g(0) + \left[c_1 M_0 + \frac{3}{2}\right]a(t)\right) + \frac{\lambda(1 - e^{-\lambda t})}{\lambda t - 1 + e^{-\lambda t}}\dot{g}(0) + \Phi(x_0). \tag{3.18}$$

On the other hand,

$$a(t) = \int_0^t \int_\rho^t e^{-\lambda(\sigma-\rho)}\left|\dot{x}(\rho)\right|^2 d\sigma\, d\rho = \frac{1}{\lambda}\int_0^t \left(1 - e^{-\lambda(t-\rho)}\right)\left|\dot{x}(\rho)\right|^2 d\rho.$$

By $\dot{x} \in L^2(0, +\infty; H)$, we get that

$$a(t) \le \frac{1}{\lambda}\int_0^t \left|\dot{x}(\rho)\right|^2 d\rho \le \frac{1}{\lambda}\int_0^{+\infty} \left|\dot{x}(\rho)\right|^2 d\rho < \infty. \tag{3.19}$$

Thus we can find some $C > 0$, such that $a(t) < C$ for all $t \in [0, +\infty)$. Therefore, it follows from (3.18) that

$$\begin{aligned}\limsup_{t\to\infty} \Phi(x(t)) &\leq \limsup_{t\to\infty} \frac{\lambda}{\lambda t - 1}\Big[\dot{g}(0) + \lambda g(0) + \Big(\frac{3}{2} + LM_0\Big)\lambda C\Big] + \Phi(x_0)\\ &= \Phi(x_0). \quad \square\end{aligned}$$

$\square$

Next we deal with asymptotic weak convergence of the trajectory in the convex case. The proof depends mainly on Opial Lemma. For the sake of completeness, let us recall the Opial Lemma [20].

Lemma 1. *Let H be a Hilbert space and $x : [0, +\infty) \to H$ be a function such that there exists a nonempty set $S \subset H$ satisfying*:

(1) $\forall t_n \to +\infty$ *with* $x(t_n) \rightharpoonup \bar{x}$ *weakly in* H, *we have* $\bar{x} \in S$;

(2) $\forall z \in S$, $\lim_{t\to+\infty} |x(t) - z|$ *exists.*

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

Theorem 4. *Under assumptions of Theorem 3, if $x \in C^2(0, +\infty; H)$ is the unique solution of (GD) from a given initial state $(x_0, \dot{x}_0) \in H \times H$, then $x(t)$ weakly converges to a minimum point of Φ as $t \to +\infty$.*

Proof. Let us apply the Opial Lemma with

$$S = \mathit{Argmin}\, \Phi = \big\{x \in H \colon \nabla\Phi(x) = 0\big\}.$$

Suppose that there exists $\widehat{x} \in H$ such that $x(t_n) \rightharpoonup \widehat{x}$ for a suitable sequence $t_n \to +\infty$. Since Φ is convex and continuous, we get that Φ is weak lower-semicontinuous. So

$$\Phi(\widehat{x}) \leq \liminf_{n\to\infty} \Phi\big(x(t_n)\big) = \lim_{n\to\infty} \Phi\big(x(t_n)\big) = \inf \Phi$$

and therefore $\widehat{x} \in \mathit{Argmin}\, \Phi = S$, which implies that the first condition of Lemma 1 is satisfied. So, we just need to prove that for any $z \in S$, the limit $\lim_{t\to+\infty} |x(t) - z|$ exists, or equivalently, $\lim_{t\to+\infty} |x(t) - z|^2$ exists. Fix $z \in S$ and define $g(t) := \frac{1}{2}|x(t) - z|^2$. Due to $\Phi(z) = \inf \Phi$, it follows from inequity (3.14) that

$$\ddot{g}(t) + \lambda \dot{g}(t) \leq (c_1 M_0 + 1)\big|\dot{x}(t)\big|^2. \tag{3.20}$$

Multiplying (3.20) by $e^{\lambda t}$ and integrating it over $(0, t)$ we obtain

$$\dot{g}(t) \leq e^{-\lambda t}\dot{g}(0) + (c_1 M_0 + 1)\int_0^t e^{-\lambda(t-\sigma)}\big|\dot{x}(\sigma)\big|^2\, d\sigma. \tag{3.21}$$

Define $\dot{g}(t)_+ := \max\{\dot{g}(t), 0\}$. From above inequality we easily get

$$\dot{g}(t)_+ \leq e^{-\lambda t}\dot{g}(0)_+ + (c_1 M_0 + 1)\int_0^t e^{-\lambda(t-\sigma)}\big|\dot{x}(\sigma)\big|^2\, d\sigma. \tag{3.22}$$

Integrating the above inequality again over $(0,+\infty)$, it follows from (3.19) that $\int_0^{+\infty} \dot g(t)_+ dt < +\infty, i.e., \dot g(t)_+ \in L^1(0,+\infty;H)$. Now we set

$$k(t) := g(t) - \int_0^t \dot g(\sigma)_+\, d\sigma.$$

For any $0 \le s < t < +\infty$ it is true that

$$g(t) - g(s) = \int_s^t \dot g(\sigma)\, d\sigma \le \int_s^t \dot g(\sigma)_+\, d\sigma.$$

Then it follows that $k(t) = g(t) - \int_0^t \dot g(\sigma)_+\, d\sigma \le g(s) - \int_0^s \dot g(\sigma)_+\, d\sigma = k(s)$. Together with that $k(t)$ is bounded from below, we get $k(t)$ converges as $t \to +\infty$. So, finally we obtain that $\lim_{t\to+\infty} g(t) = \lim_{t\to+\infty} k(t) + \int_0^{+\infty} \dot g(t)_+\, dt$ does exist. $\square$

4 Strong Convergence Results

In this section, we devote to establish strong convergence results. First we state a simple fact.

Proposition 1. *In addition to the assumptions $\mathcal{A}_\Phi$, $\mathcal{A}_f$ and $\Phi \in C^2$, let us assume that Φ is strongly convex, that is, there exists $M > 0$, such that*

$$\big\langle \nabla\Phi(x) - \nabla\Phi(y), x-y \big\rangle \ge M|x-y|^2, \quad \forall x,y \in H. \tag{4.1}$$

Then each trajectory $x(\cdot)$ of the (GD) system is strongly convergent as $t \to +\infty$ to the unique global minimizer $\bar x$ of Φ.

Proof. Let us consider a trajectory $x(\cdot)$ of the (GD) system. Since Φ is strongly convex, it has a unique minimizer $\bar x$. Let us write the above strong monotonicity property at $\bar x$ and $x(t)$:

$$\big\langle \nabla\Phi(\bar x) - \nabla\Phi\big(x(t)\big), \bar x - x(t) \big\rangle \ge M\big|\bar x - x(t)\big|^2.$$

Since $\nabla\Phi(\bar x) = 0$, it follows that

$$M\big|\bar x - x(t)\big|^2 \le \big|\nabla\Phi\big(x(t)\big)\big|\big|x(t) - \bar x\big|.$$

From Theorem 2, we have that $\lim_{t\to+\infty} \nabla\Phi(x(t)) = 0$. So, it is easy to see that $x(t) \to \bar x$ strongly as $t \to +\infty$. $\square$

In fact, the result of Proposition 1 still holds, under the following "relaxed" version of strongly convex condition:

$$(\mathcal{A}_\Phi - iv) \quad \begin{array}{l} \text{there exist } \alpha > 0 \text{ and } m > 1 \text{ such that} \\ \big\langle \nabla\Phi(x) - \nabla\Phi(y), x-y \big\rangle \ge \alpha|x-y|^m, \quad \forall x,y \in H. \end{array} \tag{4.2}$$

Also, the conclusion can be improved, and an explicit decay rate for trajectory of (GD) system is obtained for $m > 1$.

Theorem 5. *Assume the assumptions $\mathcal{A}_\Phi$, $\mathcal{A}_f$, $(\mathcal{A}_\Phi - iv)$ and $\Phi \in C^2$. If $m > 2$, we also assume that there exists some constant $\beta \geq 0$ such that*

$$\lim_{x\to 0} \frac{|\nabla f(x)|}{|x|^l} = \beta, \quad \text{with } l = 2 - \frac{2}{m}. \tag{4.3}$$

Then, for any given initial state $(x_0, \dot{x}_0) \in H \times H$, there exist $K, r > 0$ and $t_0 \geq 0$ such that, for all $t \geq t_0$, the unique solution x of (GD) from $(x_0, \dot{x}_0)$ satisfies

$$\begin{aligned} &E(t) - \min \Phi \leq Ke^{-rt}, \quad 1 < m \leq 2, \\ &E(t) - \min \Phi \leq (K + rt)^{-2/(m-2)}, \quad m > 2. \end{aligned} \tag{4.4}$$

Proof. Without loss of generality, we can assume that the unique minimizer of Φ is $\bar{x} = 0$ and $\Phi(0) = 0$. By the argument similar to the proof of Proposition 1, we deduce that $x(t) \to 0$ strongly as $t \to +\infty$. The assumption $(\mathcal{A}_\Phi - iv)$ implies

$$\Phi(x) \geq \frac{\alpha}{m}|x|^m, \quad \text{for all } x \in X. \tag{4.5}$$

Indeed, for any given $x, y \in H$, let $\phi(t) := \Phi(y + t(x - y))$. Then

$$\phi'(t) - \phi'(0) \quad = \big\langle \nabla\Phi\big(y + t(x - y)\big) - \nabla\Phi(y), (x - y)\big\rangle \geq \alpha t^{m-1}|x - y|^m,$$

by assumption $(\mathcal{A}_\Phi - iv)$. Then

$$\begin{aligned} \Phi(x) - \Phi(y) = \phi(1) - \phi(0) &= \int_0^1 \phi'(t)\,dt \\ &\geq \int_0^1 \big(\phi'(0) + \alpha t^{m-1}|x - y|^m\big)\,dt \\ &= \big\langle \nabla\Phi(y), x - y\big\rangle + \frac{\alpha}{m}|x - y|^m. \end{aligned}$$

When $y = 0$, we obtain (4.5). We are going to analyze the asymptotic behavior of system (GD) by distinguishing two cases: Case (I) $1 < m \leq 2$; Case (II) $m > 2$. The two case need different technique.

Step 1. First we deal with Case (I). Since $\lim_{t\to+\infty} x(t) = 0$, there exists $t_0 \geq 0$, such that $|x(t)| < 1$, for $t \geq t_0$. Then (4.5) become that

$$\Phi\big(x(t)\big) \geq \frac{\alpha}{2}\big|x(t)\big|^2 \quad \text{for } t \geq t_0. \tag{4.6}$$

Similar to the proof of Theorem 3, we define $g(t) := \frac{1}{2}|x(t)|^2$, and set $w(t) = \dot{g}(t) + \lambda g(t)$, i.e.,

$$w(t) = \dot{g}(t) + \lambda g(t) = \big\langle \dot{x}(t), x(t)\big\rangle + \frac{\lambda}{2}\big|x(t)\big|^2. \tag{4.7}$$

According to inequity (3.13), and ∇f is Lipschitz continuous on the bounded set of H, we get that

$$\dot{w}(t) + E(t) \leq L\big|\dot{x}(t)\big|\big|x(t)\big| + \frac{3}{2}\big|\dot{x}(t)\big|^2, \tag{4.8}$$

where L is the Lipschitz constant for $\nabla f(\dot{x})$ on $B(0, |\dot{x}|_{L^\infty})$. We recall that the energy function satisfies the following decay law

$$\dot{E}(t) \leq -\lambda \big|\dot{x}(t)\big|^2. \tag{4.9}$$

Let us multiply the inequality (4.8) by $\eta = \frac{2\lambda}{3}$ and add to last inequality, we obtain

$$\eta \dot{w}(t) + \dot{E}(t) + \eta E(t) \leq \eta L \big|\dot{x}(t)\big|\big|x(t)\big|. \tag{4.10}$$

On the other hand, by Young's Inequality, for all $\theta > 0$,

$$\big|\dot{x}(t)\big|\big|x(t)\big| \leq \frac{1}{2\theta}\big|\dot{x}(t)\big|^2 + \frac{\theta}{2}\big|x(t)\big|^2. \tag{4.11}$$

For $t \geq t_0$, noting that $\frac{\alpha}{2}|x(t)|^2 \leq \Phi(x(t))$, we deduce that, by putting $\theta = \sqrt{\alpha}$ in above equation

$$\big|\dot{x}(t)\big|\big|x(t)\big| \leq \sqrt{\frac{1}{\alpha}} E(t). \tag{4.12}$$

According to (4.11), for all $\theta > 0$, it is true that

$$w(t) = \big\langle \dot{x}(t), x(t)\big\rangle + \frac{\lambda}{2}\big|x(t)\big|^2 \leq \frac{1}{2\theta}\big|\dot{x}(t)\big|^2 + \frac{\theta}{2}\big|x(t)\big|^2 + \frac{\lambda}{2}\big|x(t)\big|^2. \tag{4.13}$$

Together with (4.6) and $E(t) = \frac{1}{2}|\dot{x}(t)|^2 + \Phi(x(t))$, we have that

$$w(t) \leq \frac{\theta+\lambda}{\alpha} E(t) - \left(\frac{\theta+\lambda}{2\alpha} - \frac{1}{2\theta}\right)\big|\dot{x}(t)\big|^2, \quad t \geq t_0.$$

Choose θ_0 sufficiently large to satisfies $\frac{\theta_0+\lambda}{2\alpha} \geq \frac{1}{\theta_0}$. Then if $C_1 \geq \frac{\theta_0+\lambda}{\alpha}$, the last inequality reduces to

$$w(t) \leq C_1 E(t), \quad t \geq t_0. \tag{4.14}$$

For the convenience of analysis, take $\theta_1 = \max\{\sqrt{\frac{1}{\alpha}}, \frac{1}{\eta}\}$. Then by virtue of (4.12) and (4.14), we obtain that

$$w(t) \leq (C_1 + C_1\theta_1)E(t) - C_1\big|\dot{x}(t)\big|\big|x(t)\big|, \quad t \geq t_0.$$

Dividing this inequality by $C_1(\theta_1 + 1)$, we get

$$E(t) \geq \frac{1}{C_1(\theta_1+1)} w(t) + \frac{1}{(\theta_1+1)}\big|\dot{x}(t)\big|\big|x(t)\big|, \quad t \geq t_0. \tag{4.15}$$

Moreover, for all $s \in (0,1)$, multiplying the inequality (4.15) by $(1-s)$, we have that, for all $t \geq t_0$,

$$E(t) \geq sE(t) + \frac{1-s}{C_1(\theta_1+1)} w(t) + \frac{(1-s)}{\theta_1+1}\big|\dot{x}(t)\big|\big|x(t)\big|.$$

Now, we fix $s = \frac{1}{2}$ and choose $K_0 = \frac{2}{\eta}$, $\eta_0 = \min\{\frac{1}{2(\theta_1+1)LK_0}, \frac{\eta}{2}\}$, and $\eta_1 = \eta - \eta_0$. Then, the above inequality becomes

$$E(t) \geq \frac{1}{2}\left(E(t) + \frac{1}{C_1(\theta_1+1)} w(t)\right) + K_0\eta_0 L\big|\dot{x}(t)\big|\big|x(t)\big|. \tag{4.16}$$

By virtue of (4.6), (4.9) and (4.11), for any $\theta_2 > 0$,

$$\eta_1 L|\dot{x}(t)||x(t)| \le \frac{\eta_1 L}{2\theta_2}|\dot{x}(t)|^2 + \frac{\theta_2\eta_1 L}{2}|x(t)|^2 \le -\frac{\eta_1 L}{2\theta_2\lambda}\dot{E}(t) + \frac{\theta_2\eta_1 L}{\alpha}E(t).$$

Combining (4.10) and the above inequality,

$$\eta\dot{w}(t) + \left[1 + \frac{\eta_1 L}{2\theta_2\lambda}\right]\dot{E}(t) + \left[\eta - \frac{\theta_2\eta_1 L}{\alpha}\right]E(t) \le \eta_0 L|\dot{x}(t)||x(t)|. \tag{4.17}$$

Fix $C_1 = \max\{1, \frac{\theta_0+\lambda}{\alpha}, \frac{\eta_1^2 L^2}{\lambda\alpha\eta^2\theta_1}, \frac{1}{\lambda}\}$, and $\theta_2 = \frac{\eta_1 L}{2\lambda[\eta(\theta_1+1)C_1-1]} > 0$. By choosing of θ_2, we get $\eta = [1 + \frac{\eta_1 L}{2\theta_2\lambda}]\frac{1}{C_1(\theta_1+1)}$. Hence, by virtue of (4.16) and (4.17), we obtain for all $t \ge t_0$

$$\begin{aligned} E(t) \ge{}& \frac{1}{2}\left(E(t) + \frac{1}{C_1(\theta_1+1)}w(t)\right) \\ &+ K_0\left[1 + \frac{\eta_1 L}{2\theta_2\lambda}\right]\left(\dot{E}(t) + \frac{1}{C_1(\theta_1+1)}\dot{w}(t)\right) + K_0\left[\eta - \frac{\theta_2\eta_1 L}{\alpha}\right]E(t). \end{aligned} \tag{4.18}$$

Noticing $C_1 \ge \max\{1, \frac{\eta_1^2 L^2}{\lambda\alpha\eta^2\theta_1}\}$ and $\theta_1 \ge \frac{1}{\eta}$, it is true that

$$\frac{\eta}{2} \ge \frac{\eta_1^2 L^2}{2\lambda\alpha\eta\theta_1 C_1} > \frac{\eta_1^2 L^2}{2\lambda[\eta(\theta_1+1)C_1 - 1]\alpha} = \frac{\theta_2\eta_1 L}{\alpha},$$

which implies $K_0[\eta - \frac{\theta_2\eta_1 L}{\alpha}] > K_0\frac{\eta}{2} = 1$. Hence, by setting $r = K_0[2 + \frac{\eta_1 L}{\theta_2\lambda}]$ and $C_2 = \frac{1}{C_1(\theta_1+1)}$, we deduce from (4.18) that

$$\big(E(t) + C_2 w(t)\big) + r\big(\dot{E}(t) + C_2\dot{w}(t)\big) \le 0, \quad t \ge t_0.$$

Consequently, we deduce that

$$E(t) + C_2 w(t) \le K_1 e^{-rt}, \quad t \ge t_0 \tag{4.19}$$

with $K_1 = E(t_0) + C_2 w(t_0)$. By Young's Inequality, it is true that, for $t \ge t_0$

$$w(t) = \langle\dot{x}(t), x(t)\rangle + \frac{\lambda}{2}|x(t)|^2 \ge -\frac{1}{2\lambda}|\dot{x}(t)|^2 \ge -\frac{1}{\lambda}E(t).$$

Then it follows that, for $t \ge t_0$

$$E(t) + C_2 w(t) \ge \left(1 - \frac{C_2}{\lambda}\right)E(t)$$

and $1 - \frac{C_2}{\lambda} = 1 - \frac{1}{C_1(\theta_1+1)\lambda} > 0$, by recalling $C_1 \ge \frac{1}{\lambda}$. Therefore, finally we get, for $t \ge t_0$

$$0 \le E(t) - \min\Phi \le Ke^{-rt}, \quad \text{with } K = \frac{\lambda}{\lambda - C_2}K_1. \tag{4.20}$$

Step 2. Based on the perturbed energy method, we deal with Case (II). We have established that $\lim_{t\to+\infty}\dot{x}(t)=0$ (Theorem 2). Together with assumption (4.3), we can find $t_0\geq 0$ such that, for all $t\geq t_0$, the following property holds:

$$\big|\nabla f\big(\dot{x}(t)\big)\big|\leq(\beta+1)\big|\dot{x}(t)\big|^l. \tag{4.21}$$

Define for $\varepsilon>0$

$$E_\varepsilon(t)=E(t)+\varepsilon\bigg(\big\langle x(t),\dot{x}(t)\big\rangle+\frac{\lambda}{2}\big|x(t)\big|^2\bigg). \tag{4.22}$$

We recall that $E(t)=\frac{1}{2}|\dot{x}(t)|^2+\Phi(x(t))$. Then, from (4.5) and Hölder inequality, for all $t\geq t_0$

$$\begin{aligned}
E_\varepsilon(t)&=\frac{1}{2}\big|\dot{x}(t)\big|^2+\Phi\big(x(t)\big)+\varepsilon\bigg(\big\langle x(t),\dot{x}(t)\big\rangle+\frac{\lambda}{2}\big|x(t)\big|^2\bigg)\\
&\leq\frac{1}{2}\big|\dot{x}(t)\big|^2+\Phi\big(x(t)\big)+\frac{\varepsilon}{2}\big|\dot{x}(t)\big|^2+\frac{\varepsilon(\lambda+1)}{2}\big|x(t)\big|^2\\
&\leq(1+\varepsilon)E(t)+\frac{\varepsilon(\lambda+1)}{2}\big(\big|x(t)\big|^m\big)^{\frac{2}{m}}\\
&\leq(1+\varepsilon)E(0)^{\frac{m-2}{m}}E(t)^{\frac{2}{m}}+\frac{\varepsilon(\lambda+1)}{2}\bigg(\frac{m}{\alpha}\Phi\big(x(t)\big)\bigg)^{\frac{2}{m}}\\
&\leq B_\varepsilon E(t)^{\frac{2}{m}},
\end{aligned} \tag{4.23}$$

where $B_\varepsilon=(1+\varepsilon)E(0)^{\frac{m-2}{m}}+\frac{\varepsilon(\lambda+1)}{2}\left(\frac{m}{\alpha}\right)^{\frac{2}{m}}$.

On the other hand, take a constant $\delta=\frac{\lambda}{4}$, then

$$\begin{aligned}
E_\varepsilon(t)&\geq E(t)-\varepsilon\bigg(\frac{1}{4\delta}\big|\dot{x}(t)\big|^2+\delta\big|x(t)\big|^2\bigg)+\frac{\varepsilon\lambda}{2}\big|x(t)\big|^2\\
&=E(t)-\frac{\varepsilon}{4\delta}\big|\dot{x}(t)\big|^2+\varepsilon\bigg(\frac{\lambda}{2}-\delta\bigg)\big|x(t)\big|^2\\
&\geq E(t)-\frac{\varepsilon}{\lambda}\big|\dot{x}(t)\big|^2.
\end{aligned} \tag{4.24}$$

By choosing ε so small that $1-\frac{2\varepsilon}{\lambda}>0$, then

$$\bigg(1-\frac{2\varepsilon}{\lambda}\bigg)E(t)\leq E_\varepsilon(t). \tag{4.25}$$

Computing the derivative of (4.22), and using inequality (3.1) and convexity of Φ and f, we have

$$\begin{aligned}
\dot{E}_\varepsilon(t)&=\dot{E}(t)+\varepsilon\big(\big|\dot{x}(t)\big|^2+\big\langle\ddot{x}(t),x(t)\big\rangle+\lambda\big\langle\dot{x}(t),x(t)\big\rangle\big)\\
&\leq-\lambda\big|\dot{x}(t)\big|^2+\varepsilon\big|\dot{x}(t)\big|^2-\varepsilon\big\langle\nabla f\big(\dot{x}(t),x(t)\big\rangle-\varepsilon\big\langle\nabla\Phi(x(t),x(t)\big\rangle\\
&\leq(\varepsilon-\lambda)\big|\dot{x}(t)\big|^2-\varepsilon\big\langle\nabla f(\dot{x}(t),x(t)\big\rangle-\varepsilon\Phi\big(x(t)\big).
\end{aligned} \tag{4.26}$$

From Young inequality, it follows that, for all $\mu > 0$, $t \geq t_0$

$$
\begin{aligned}
\big|\langle \nabla f(\dot{x}(t)), x(t)\rangle\big| &\leq \frac{\mu}{m}\big|x(t)\big|^m + \frac{m-1}{m}\mu^{-\frac{m'}{m}}\big|\nabla f(\dot{x}(t))\big|^{m'} \\
&\leq \frac{\mu}{\alpha m}\Phi(x(t)) + \frac{m-1}{m}\mu^{-\frac{m'}{m}}\big|\nabla f(\dot{x}(t))\big|^{m'}
\end{aligned} \tag{4.27}
$$

with $m' = \frac{m}{m-1}$. Returning to (4.26), and using inequality (4.21) and condition (4.3), we get, for all $t \geq t_0$

$$
\begin{aligned}
\dot{E}_\varepsilon(t) &\leq (\varepsilon - \lambda)\big|\dot{x}(t)\big|^2 + \varepsilon\left(\frac{\mu}{\alpha m} - 1\right)\Phi(x(t)) \\
&\quad + \varepsilon\frac{m-1}{m}\mu^{-\frac{m'}{m}}\big|\nabla f(\dot{x}(t))\big|^{m'} \\
&\leq \left[\varepsilon\left(1 + \frac{m-1}{m}(\beta+1)^{m'}\mu^{-\frac{m'}{m}}\right) - \lambda\right]\big|\dot{x}(t)\big|^2 + \varepsilon\left(\frac{\mu}{\alpha m} - 1\right)\Phi(x(t)).
\end{aligned} \tag{4.28}
$$

Now we choose μ small sufficiently such that $\frac{\mu}{\alpha m} - 1 < 0$. Once μ is chosen we then take ε small sufficiently such that $\varepsilon(1 + \frac{m-1}{m}(\beta+1)^{m'}\mu^{-\frac{m'}{m}}) - \lambda < 0$, and inequality (4.25) remains valid. Consequently, (4.23) and (4.28) yields, for all $t \geq t_0$

$$
\dot{E}_\varepsilon(t) \leq -k_1 E(t) \leq -k_1 B_\varepsilon^{-\frac{m}{2}} E_\varepsilon(t)^{\frac{m}{2}} \tag{4.29}
$$

with

$$
k_1 = \min\left\{2\left[\lambda - \varepsilon\left(1 + \frac{m-1}{m}(\beta+1)^{m'}\mu^{-\frac{m'}{m}}\right)\right], \varepsilon\left(1 - \frac{\mu}{\alpha m}\right)\right\}.
$$

Thus, a simple integration of (4.29) leads to, for all $t \geq t_0$,

$$
E_\varepsilon(t) \leq \left(E_\varepsilon(0)^{-\frac{m-2}{2}} + r_0 t\right)^{-\frac{2}{m-2}}, \tag{4.30}
$$

where $r_0 = \frac{m-2}{2}B_\varepsilon^{-\frac{m}{2}}k_1$. According to inequality (4.25), we get that

$$
E(t) \leq (K + rt)^{-\frac{2}{m-2}}, \tag{4.31}
$$

with

$$
K = \left[\left(1 - \frac{\varepsilon}{2\delta}\right)E_\varepsilon^{-1}(0)\right]^{\frac{m-2}{2}}, \qquad r = \left(1 - \frac{\varepsilon}{2\delta}\right)^{\frac{m-2}{2}} r_0.
$$

Finally, by virtue of (4.20) and (4.31), we completes the proof. □

Remark 2. Under the assumption of above theorem, one easily deduces that the trajectory $x(t)$ of (GD) strongly converges to minimum point $\bar{x}$. More precisely, for all $t \geq t_0$, the trajectory $x(t)$ satisfies

$$
\begin{aligned}
\big|x(t) - \bar{x}\big| &\leq \sqrt{K}e^{-\frac{r}{2}t}, \quad 1 < m \leq 2, \\
\big|x(t) - \bar{x}\big| &\leq (K + rt)^{-\frac{2}{m(m-2)}}, \quad m > 2.
\end{aligned} \tag{4.32}
$$

Now, we present some examples to illustrate the previous results. As far as we know, the polynomial decay result is very rare in this field. Firstly, we give an example to show that the polynomial decay, in case of $m > 2$, cannot improve to exponential decay in general.

Example 1. We consider a second order ODE in one dimensional case $H = R^1$ as following

$$\begin{cases} \ddot{x}(t) + \dot{x}(t) + \nabla f(\dot{x}(t)) + \nabla \Phi(x(t)) = 0, \\ f(x) = \frac{4}{5}|x|^{\frac{5}{2}}, \qquad \Phi(x) = \frac{1}{3}|x|^3. \end{cases} \tag{4.33}$$

Let $\bar{x} = 0$. Obviously the above equation satisfies condition (4.2) with $\alpha = 1$ and $m = 3$. Moreover $l = 2 - \frac{2}{m} = \frac{4}{3}$, and $\frac{|\nabla f(x)|}{|x|^l} = 2|x|^{\frac{1}{6}}$, which implies that the equation (4.33) satisfies condition (4.3) with $\beta = 0$. Hence, applying Theorem 5, for any initial condition $(x(0), \dot{x}(0)) \in R^2$, we obtain that the decay rate of energy $E(t)$ is $-\frac{2}{m-2} = -2$. Indeed, if we choose initial condition $x(0) = 1$ and $\dot{x}(0) = -1$, the unique global solution of (4.33) is given by

$$x(t) = (1 + t)^{-1}$$

and the associated energy function is

$$E(t) = \frac{1}{2}\big|\dot{x}(t)\big|^2 + \Phi(x(t)) = \frac{1}{2}(1 + t)^{-4} + \frac{1}{3}(1 + t)^{-3}.$$

It coincides with our estimation.

Example 2. Consider the following nonlinear strongly coupled oscillator system with two degrees of freedom (see [19] and [23] for the model of this kind):

$$\begin{cases} u_{tt} + \lambda u_t + 2\big(u_t^4 + v_t^2\big)^2 u_t^3 + 2su + 8\big(u^4 + |v|^q\big)u^3 = 0, \\ v_{tt} + \lambda v_t + \big(u_t^4 + v_t^2\big)^2 v_t + 2sv + 2q\, sgn(v)\big(u^4 + |v|^q\big)|v|^{q-1} = 0, \end{cases} \tag{4.34}$$

where $\lambda > 0$, $s > 0$, $q > 1$, and

$$sgn(v) = \begin{cases} 1 & \text{if } v > 0, \\ 0 & \text{if } v = 0, \\ -1 & \text{if } v < 0. \end{cases}$$

The differential system (4.34) can be rewritten as

$$\ddot{X}(t) + \lambda \dot{X}(t) + \nabla f(\dot{X}(t)) + \nabla \Phi(X(t)) = 0,$$

where

$$X(t) = \begin{pmatrix} u(t) \\ v(t) \end{pmatrix}, \quad \text{and} \quad \begin{cases} f(u, v) = \frac{1}{6}\big(u^4 + v^2\big)^3, \\ \Phi(u, v) = s\big(u^2 + v^2\big) + \big(u^4 + |v|^q\big)^2. \end{cases}$$

Obviously, f and Φ satisfy assumption $(\mathcal{A}_f)$ and $(\mathcal{A}_\Phi)$, so all solutions are global from Theorem 1. Since $s > 0$ and $q > 1$, one easily check that $\Phi(u, v) : R^2 \to$

R^+ is coercive and convex. Moreover, since, for any $(u,v) \in R^2$, $\nabla\Phi(u,v) = 2(su + 4(u^4+|v|^q)u^3, sv + q(u^4+|v|^q)|v|^{q-2}v)^T$, so Φ is strongly convex and $(0,0)$ is the strong minimum of Φ. Hence, by Theorem 5, we know that the energy

$$E(t) := \frac{1}{2}\left(u_t^2 + v_t^2\right) + \Phi(u,v)$$

associated to the system (4.34) exponentially decay to zero, and the trajectory $(u(t), v(t))$ exponentially converges to the minimum point $(0,0)$.

Remark 3. Let $w = \dot{u}$, $y = \dot{v}$ and $Z = (w,u,y,v)^T$, then the differential system (4.34) can be expressed as an "abstract form" first order system:

$$\dot{Z} = CZ + H(Z) \tag{4.35}$$

with

$$C = \begin{pmatrix} -\lambda & -2s & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -\lambda & -2s \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

and

$$H(Z) = \begin{pmatrix} -2(w^4+y^2)^2w^3 - 8(u^4+|v|^q)u^3 \\ 0 \\ -(w^4+y^2)^2y - 2q\,sgn(v)(v^4+|v|^q)|v|^{q-1} \\ 0 \end{pmatrix}.$$

Denote by $\mathcal{E}$ the set of equilibrium points of the equation (4.35), then it is obvious that $\mathcal{E}$ is a single point set $\{(0,0,0,0)^T\}$. If $q > 2$, from the general result of Hale and Raugel [17, Th2.4], one can also deduce the convergence (but not exponential convergence) of trajectory. If $2 \geq q > 1$, because the differentiability of $H(Z)$ is not valid, so the result of [17] cannot be applied to deduce the convergence of trajectory. However, as mentioned above, one obtain the exponential convergence for any $q > 1$ via the results in this work.

Finally, let us emphasize that, combine the asymptotic techniques developed in this paper with appropriate evolution triple $V \hookrightarrow H \hookrightarrow V'$, one can analyze the convergence of global solutions of some hyperbolic systems. For example, let us consider nonlinear Klein-Gordon equation

$$\begin{cases} u_{tt} + u_t + \alpha|u_t|^2u_t - \Delta u = -Cu^p - \beta u & \text{in } \Omega \times [0,+\infty), \\ u(\cdot,0) = u_0, \quad u_t(\cdot,0) = v_0 & \text{in } \Omega, \\ u(x,t) = 0 & \text{on } \partial\Omega \times [0,+\infty), \end{cases} \tag{4.36}$$

where p is an odd integer, $C > 0$, $\alpha \geq 0$, $\beta \geq 0$, and $\Omega \subset R^n$ be an open bounded domain with boundary $\partial\Omega$ sufficiently regular. By Poincaré inequality and Theorem 5, we obtain the following convergence result for the evolution problem (4.36). The detailed proof is omitted here. For related work, we refer to [2], [24].

Proposition 2. *For any $u_0 \in H^1(\Omega)$ and $v_0 \in L^2(\Omega)$, there exists a unique weak solution of to hyperbolic problem* (4.36) *such that*:

(i) *the map $t \longmapsto u_t(\cdot,t)$ belongs to $L^2(0,+\infty;L^2(\Omega)) \cap L^\infty(0,+\infty;L^2(\Omega))$.*

(ii) *the solution $t \longmapsto u(\cdot,t)$ strongly exponential converges in $H_0^1(\Omega)$, as $t \to +\infty$, toward a solution of the degenerate Logistic equation*

$$\begin{cases} -\Delta u + Cu^p + \beta u = 0 & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

Acknowledgement

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions to improve the manuscript. This work is supported by the National Natural Science Foundation of China (Grant 11226128, 11271099).

References

[1] F. Alvarez. On the minimizing property of a second order dissipative system in Hilbert space. *SIAM J. Control Optim.*, **38**(4):1102–1119, 2000. http://dx.doi.org/10.1137/S0363012998335802.

[2] F. Alvarez and H. Attouch. Convergence and asymptotic stabilization for some damped hyperbolic equations with non-isolated equilibria. *ESAIM Control Optim. Calc. Var.*, **6**:539–552, 2001. http://dx.doi.org/10.1051/cocv:2001100.

[3] P.N. Andriotaki. Propagation of surface (seismic) waves: Ordinary differential equations with strongly nonlinear damping. *Theoret. Appl. Fracture Mechanics*, **43**:308–323, 2005. http://dx.doi.org/10.1016/j.tafmec.2005.03.002.

[4] H. Attouch and M.-O. Czarnecki. Asymptotic control and stabilization of nonlinear oscillators with non isolated equilibria. *J. Differential Equations*, **179**:278–310, 2002. http://dx.doi.org/10.1006/jdeq.2001.4034.

[5] H. Attouch and M.-O. Czarnecki. Motion with friction of a heavy particle on a manifold applications to optimization. *Math. Model. Numer. Anal.*, **36**(3):505–516, 2002. http://dx.doi.org/10.1051/m2an:2002023.

[6] H. Attouch, X. Goudou and P. Redont. The heavy ball with friction method. I. the continuous dynamical system. *Comm. Contemp. Math.*, **2**(1):1–34, 2000. http://dx.doi.org/10.1142/S0219199700000025.

[7] H. Attouch and M. Mainge. Asymptotic behavior of second-order dissipative evolution equations combining potential with non-potential effects. *ESAIM Control Optim. Calc. Var.*, **17**(03):836–857, 2011. http://dx.doi.org/10.1051/cocv/2010024.

[8] I. Ben Hassen and A. Haraux. Convergence and decay estimates for a class of second order dissipative equations involving a non-negative potential energy. *J. Funct. Anal.*, **260**(10):2933–2963, 2011. http://dx.doi.org/10.1016/j.jfa.2011.02.010.

[9] A. Bhaya, F. Pazos and E. Kaszkurewicz. Comparative study of the CG and HBF ODEs used in the global minimization of nonconvex functions. *Artificial Neural Networks–ICANN 2009*, pp. 668–677, 2009.

[10] H. Brézis. *Opérateurs maximaux monotones et semigroupes de contractions dans les espaces de Hilbert*, volume 5 of *Math. Stud.* North-Holland, Amsterdam, 1973.

[11] R.E. Bruck. Asymptotic convergence of nonlinear contraction semigroups in Hilbert space. *J. Funct. Anal.*, **18**:15–26, 1975. http://dx.doi.org/10.1016/0022-1236(75)90027-0.

[12] P.B. Burt. *Quantem Mechanics and Nonlinear Wave.* Harwood Academic, New York, 1981.

[13] A. Cabot. Stabilization of oscillators subject to dry friction: finite time convergence versus exponential decay results. *Trans. Amer. Math. Soc.*, **360**(1):103, 2008. http://dx.doi.org/10.1090/S0002-9947-07-03990-6.

[14] A. Cabot and L. Paoli. Asymptotics for some vibro-impact problems with a linear dissipation term. *J. Math. Pures Appl.*, **87**(3):291–323, 2007. http://dx.doi.org/10.1016/j.matpur.2007.01.003.

[15] X. Goudou and J. Munier. The gradient and heavy ball with friction dynamical systems: the quasiconvex case. *Math. Program. (Ser. B)*, **116**:173–191, 2009.

[16] J. Hale. *Asymptotic Behavior of Dissipative Systems*, volume 25 of *Math. Surveys Monogr.* Amer. Math. Soc., Providence, 1988.

[17] J. Hale and G. Raugel. Convergence in gradient-like systems with applications to PDE. *Z. Angew. Math. Phys.*, **43**:63–124, 1992. http://dx.doi.org/10.1007/BF00944741.

[18] A. Haraux. *Systèmes dynamiques dissipatifs et applications*, volume 17 of *RMA*. Masson, Paris, 1991.

[19] R.D. Neilson, A. Ivanović, A. Starkey, J. Rodger and A. Albert. Design and dynamic analysis of a pneumatic impulse generating device for the non-destructive testing of ground anchorages. *Mech. Syst. Signal Processing*, **21**(6):2523–2545, 2007. http://dx.doi.org/10.1016/j.ymssp.2006.11.011.

[20] Z. Opial. Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bull. Amer. Math. Soc.*, **73**:591–597, 1967. http://dx.doi.org/10.1090/S0002-9904-1967-11761-0.

[21] B.T. Polyak. *Introduction to Optimization, Optimization Software.* New York, 1988.

[22] L. Rayleigh. *The Theory of Sound, vol. 1, second ed.* Dover Reprint, London, 1894.

[23] W.T. Thomson. *Theory of Vibration With Applications.* 4th ed. Englewood Cliffs, NJ, Prentice-Hall, 1993.

[24] E. Zuazua. Stability and decay for a class of nonlinear hyperbolic problems. *Asymptot. Anal.*, **1**:161–185, 1998.