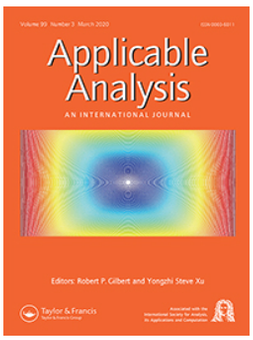

Taylor & Francis
Taylor & Francis Group

Applicable Analysis
An International Journal

ISSN: 0003-6811 (Print) 1563-504X (Online) Journal homepage: https://www.tandfonline.com/loi/gapa20

# A second-order dynamical approach with variable damping to nonconvex smooth minimization

Radu Ioan Boţ, Ernö Robert Csetnek & Szilárd Csaba László

**To cite this article:** Radu Ioan Boţ, Ernö Robert Csetnek & Szilárd Csaba László (2020) A second-order dynamical approach with variable damping to nonconvex smooth minimization, Applicable Analysis, 99:3, 361-378, DOI: 10.1080/00036811.2018.1495330

**To link to this article:** https://doi.org/10.1080/00036811.2018.1495330



Published online: 09 Jul 2018.

APPLICABLE ANALYSIS
2020, VOL. 99, NO. 3, 361–378
https://doi.org/10.1080/00036811.2018.1495330

Taylor & Francis
Taylor & Francis Group



# A second-order dynamical approach with variable damping to nonconvex smooth minimization

Radu Ioan Boţ [a], Ernö Robert Csetnek[a] and Szilárd Csaba László[b]

[a]Faculty of Mathematics, University of Vienna, Vienna, Austria; [b]Department of Mathematics, Technical University of Cluj-Napoca, Cluj-Napoca, Romania

**ABSTRACT**
We investigate a second-order dynamical system with variable damping in connection with the minimization of a nonconvex differentiable function. The dynamical system is formulated in the spirit of the differential equation which models Nesterov's accelerated convex gradient method. We show that the generated trajectory converges to a critical point, if a regularization of the objective function satisfies the Kurdyka- Lojasiewicz property. We also provide convergence rates for the trajectory formulated in terms of the Lojasiewicz exponent.

**ARTICLE HISTORY**
Received 2 May 2018
Accepted 26 June 2018

**COMMUNICATED BY**
Boris Mordukhovich

**KEYWORDS**
Second-order dynamical system; nonconvex optimization; Kurdyka–Łojasiewicz inequality; convergence rate

**AMS SUBJECT CLASSIFICATIONS**
90C26; 90C30; 65K10

## 1. Introduction

Consider the (not necessarily convex) optimization problem

$$\inf_{x\in\mathbb{R}^n} g(x), \tag{1}$$

where $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a Fréchet differentiable function with $L_g$-Lipschitz continuous gradient, i.e. there exists $L_g \geq 0$ such that $\|\nabla g(x) - \nabla g(y)\| \leq L_g\|x - y\|$ for all $x, y \in \mathbb{R}^n$. We associate to (1) the second-order dynamical system (for $t \geq t_0$)

$$\begin{aligned} &\ddot{x}(t) + \left(\frac{\alpha}{t} + \gamma\right)\dot{x}(t) + \nabla g(x(t)) = 0,\\ &x(t_0) = u_0, \quad \dot{x}(t_0) = v_0, \end{aligned} \tag{2}$$

where $t_0 > 0, u_0, v_0 \in \mathbb{R}^n, \alpha \in \mathbb{R}$ and $\gamma \in (0, +\infty)$.

The study of the system (2) is motivated by the recent developments related to the approaching of the solving of convex optimization problems from a continuous perspective.

In [1], Su, Boyd and Candes proposed the following dynamical system

$$\begin{aligned} &\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \nabla g(x(t)) = 0,\\ &x(t_0) = u_0, \quad \dot{x}(t_0) = v_0, \end{aligned} \tag{3}$$

**CONTACT** Radu Ioan Boţ radu.bot@univie.ac.at
This article was originally published with errors. This version has been corrected. Please see Erratum (http://dx.doi.org/10.1080/00036811.2015.1505361).

as the continuous counterpart of the Nesterov's accelerated gradient method [see [2]] for minimizing $g$ in the convex case. This research has been deepened by Attouch and his co-authors [see [3,4]], who proved that, if $\alpha > 3$, then the generated trajectory $x(t)$ converges to a minimizer of $g$ as $t \to +\infty$, while the convergence rate of the objective function along the trajectory is $o(1/t^2)$. The convergence of the trajectory is actually the continuous counterpart of a result due to Chambolle and Dossal [see [5]], which proves the convergence of the iterates of the modified FISTA algorithm [see [6]].

Recently, in [7], investigations have been performed concerning the convergence rate of the objective function along the trajectory in the subcritical case $\alpha \leq 3$, while some open questions related to the asymptotic properties of the trajectory have been formulated.

In this manuscript, we carry out, in the nonconvex setting, an asymptotic analysis of the dynamical system (2), which can be seen as a perturbation of the dynamical system (3) that models Nesterov's accelerated gradient method in the convex case. To the best of our knowledge, this is the first contribution addressing second-order dynamical systems with variable damping associated to nonconvex optimization problems. We show that the generated trajectory converges to a critical point of $g$ as $t \to +\infty$, provided the following regularization of $g$,

$$H : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}, \quad H(u, v) = g(u) + \tfrac{1}{2}\|u - v\|^2,$$

satisfies the Kurdyka–Łojasiewicz inequality. Moreover, we derive convergence rates in the terms of Łojasiewicz exponent, for the trajectory, its velocity and its acceleration. One of the major future goals is to study the asymptotic properties of the system (2) in case $\gamma = 0$. For other investigations of the asymptotic analysis of second-order dynamical systems with time-dependent damping, we refer to the papers of Haraux and Jendoubi [8] and Balti [9].

For $\alpha = 0$, the convergence of the trajectory generated by (2) to a critical point of $g$ has been shown by Bégout, Bolte and Jendoubi in [10] in the hypothesis that $g$ is of class $C^2$ and it satisfies the Kurdyka–Łojasiewicz property with a desingularizing function satisfying a restrictive condition [see also the papers of Haraux and Jendoubi [11] and Chill and Jendoubi [12]]. On the other hand, the dynamical system (2) is, for $\alpha = 0$, a particular instance of the second-order dynamical system of proximal-gradient type studied in [13].

The following numerical scheme, with starting points $x_0, x_1 \in \mathbb{R}^n$,

$$(\forall k \geq 1) \begin{cases} y_k = x_k + \dfrac{(1 - \gamma\sqrt{s})k - \alpha\gamma\sqrt{s}}{k + \alpha}(x_k - x_{k-1}), \\ x_{k+1} = y_k - s\nabla g(y_k), \end{cases} \tag{4}$$

where $s \leq 1/L_g$ is the step size, can be seen as a discrete counterpart of (2). One can notice that for $\gamma = 0$, this iterative scheme algorithm is similar to Nesterov's accelerated convex gradient method.

In the following, we prove that (2) can be seen in an informal way as the exact limit of (4)). We take to this end in (4) small step sizes and follow the same approach as Su, Boyd and Candes in [1, Section 2]. For this purpose, we rewrite (4) in the form

$$\frac{x_{k+1} - x_k}{\sqrt{s}} = \frac{(1 - \gamma\sqrt{s})k - \alpha\gamma\sqrt{s}}{k + \alpha} \cdot \frac{x_k - x_{k-1}}{\sqrt{s}} - \sqrt{s}\nabla g(y_k) \quad \forall\, k \geq 1 \tag{5}$$

and introduce the *Ansatz* $x_k \approx x(k\sqrt{s})$ for some twice continuously differentiable function $x : [0, +\infty) \to \mathbb{R}^n$. We let $k = t/\sqrt{s}$ and get $x(t) \approx x_k$, $x(t + \sqrt{s}) \approx x_{k+1}$, $x(t - \sqrt{s}) \approx x_{k-1}$. Then, as the step size $s$ goes to zero, from the Taylor expansion of $x$, we obtain

$$\frac{x_{k+1} - x_k}{\sqrt{s}} = \dot{x}(t) + \frac{1}{2}\ddot{x}(t)\sqrt{s} + o(\sqrt{s})$$

and

$$\frac{x_k - x_{k-1}}{\sqrt{s}} = \dot{x}(t) - \frac{1}{2}\ddot{x}(t)\sqrt{s} + o(\sqrt{s}).$$

Furthermore, since

$$\sqrt{s}\|\nabla g(y_k) - \nabla g(x_k)\| \leq \sqrt{s}L_g\|y_k - x_k\| = \sqrt{s}L_g \left|\frac{(1-\gamma\sqrt{s})k - \alpha\gamma\sqrt{s}}{k+\alpha}\right| \|x_k - x_{k-1}\| = o(\sqrt{s}),$$

it follows $\sqrt{s}\nabla g(y_k) = \sqrt{s}\nabla g(x_k) + o(\sqrt{s})$. Consequently, (5) can be written as

$$\dot{x}(t) + \frac{1}{2}\ddot{x}(t)\sqrt{s} + o(\sqrt{s}) = \frac{(1-\gamma\sqrt{s})t - \alpha\gamma s}{t+\alpha\sqrt{s}}\left(\dot{x}(t) - \frac{1}{2}\ddot{x}(t)\sqrt{s} + o(\sqrt{s})\right) - \sqrt{s}\nabla g(x(t)) + o(\sqrt{s})$$

or, equivalently

$$(t+\alpha\sqrt{s})(\dot{x}(t) + \tfrac{1}{2}\ddot{x}(t)\sqrt{s} + o(\sqrt{s})) = ((1-\gamma\sqrt{s})t - \alpha\gamma s)(\dot{x}(t) - \tfrac{1}{2}\ddot{x}(t)\sqrt{s} + o(\sqrt{s})) - \sqrt{s}(t+\alpha\sqrt{s})\nabla g(x(t)) + o(\sqrt{s}).$$

Hence,

$$\tfrac{1}{2}(2t + \alpha\sqrt{s} - \gamma t\sqrt{s} - \alpha\gamma s)\ddot{x}(t)\sqrt{s} + (\gamma t\sqrt{s} + \alpha\sqrt{s} + \alpha\gamma s)\dot{x}(t) + \sqrt{s}(t+\alpha\sqrt{s})\nabla g(x(t)) = o(\sqrt{s}).$$

After dividing by $\sqrt{s}$ and letting $s \to 0$, we obtain

$$t\ddot{x}(t) + (\gamma t + \alpha)\dot{x}(t) + t\nabla g(x(t)) = 0,$$

which, after division by $t$, gives (2), namely

$$\ddot{x}(t) + \left(\frac{\alpha}{t} + \gamma\right)\dot{x}(t) + \nabla g(x(t)) = 0.$$

## 2. Existence and uniqueness of the trajectory

We consider on the finite-dimensional space $\mathbb{R}^n$ the Euclidean topology. If $x \in \mathbb{R}^n$ is a local minimizer of $g$, then $\nabla g(x) = 0$. We denote by

$$\operatorname{crit}(g) = \{x \in \mathbb{R}^n : \nabla g(x) = 0\}$$

the set of critical points of $g$.

We are considering in the asymptotic analysis of the dynamical system (2) strong global solutions.

**Definition 2.1:** We say that $x : [t_0, +\infty) \to \mathbb{R}^n$ is a strong global solution of (2), if the following properties are satisfied:

(i) $x, \dot{x} : [t_0, +\infty) \to \mathbb{R}^n$ are locally absolutely continuous, in other words, absolutely continuous on each interval $[t_0, T]$ for $t_0 < T < +\infty$;
(ii) $\ddot{x}(t) + (\alpha/t + \gamma)\dot{x}(t) + \nabla g(x(t)) = 0$ for almost every $t \geq t_0$;
(iii) $x(t_0) = u_0$ and $\dot{x}(t_0) = v_0$.

Recall that a function $x : [t_0, +\infty) \to \mathbb{R}^n$ is absolutely continuous on an interval $[t_0, T]$, if there exists an integrable function $y : [t_0, T] \to \mathbb{R}^n$ such that

$$x(t) = x(0) + \int_{t_0}^{t} y(s)\,\mathrm{d}s \quad \forall\, t \in [t_0, T].$$

It follows from the definition that an absolutely continuous function is differentiable almost everywhere, its derivative coincides with its distributional derivative almost everywhere and one can

recover the function from its derivative $\dot{x} = y$ by the integration formula above. On the other hand, if $x : [t_0, T] \to \mathbb{R}^n$ (where $T > t_0$) is absolutely continuous and $B : \mathbb{R}^n \to \mathbb{R}^n$ is $L$-Lipschitz continuous (where $L \geq 0$), then the function $B \circ x$ is absolutely continuous, too. Moreover, $B \circ x$ is almost everywhere differentiable and the inequality $\|(\mathrm{d}/\mathrm{d}t)B(x(t))\| \leq L\|\dot{x}(t)\|$ holds for almost every $t \geq t_0$ [see [14,15]].

We prove existence and uniqueness of a strong global solution of (2) by making use of the Cauchy–Lipschitz–Picard Theorem for absolutely continues trajectories [see for example [16, Proposition 6.2.1], [17, Theorem 54]]. The key argument is that one can rewrite (2) as a particular first-order dynamical system in a suitably chosen product space [see also [18]].

**Theorem 2.1:** *For every starting points $u_0, v_0 \in \mathbb{R}^n$ there exists a unique strong global solution of the dynamical system* (2).

***Proof:*** By making use of the notation $X(t) = (x(t), \dot{x}(t))$, the system (2) can be rewritten as a first-order dynamical system:

$$\begin{aligned} \dot{X}(t) &= F(t, X(t)), \\ X(t_0) &= (u_0, v_0), \end{aligned} \tag{6}$$

where $F : [t_0, +\infty) \times \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n \times \mathbb{R}^n$, $F(t, u, v) = (v, -(\alpha/t + \gamma)v - \nabla g(u))$.

First we show that $F(t, \cdot, \cdot)$ is $L(t)$-Lipschitz continuous for every $t \geq t_0$ and that the Lipschitz constant is a function of time with the property that $L(\cdot) \in L^1_{\mathrm{loc}}([t_0, +\infty))$. Indeed, for every $(u, v), (\overline{u}, \overline{v}) \in \mathbb{R}^n \times \mathbb{R}^n$, we have

$$\begin{aligned} \|F(t, u, v) - F(t, \overline{u}, \overline{v})\| &= \sqrt{\|v - \overline{v}\|^2 + \left\| \left(\frac{\alpha}{t} + \gamma\right)(\overline{v} - v) + (\nabla g(\overline{u}) - \nabla g(u)) \right\|^2} \\ &\leq \sqrt{\left(1 + 2\left(\frac{\alpha}{t} + \gamma\right)^2\right)\|v - \overline{v}\|^2 + 2L_g^2\|u - \overline{u}\|^2} \\ &\leq \sqrt{1 + 2L_g^2 + 2\left(\frac{\alpha}{t} + \gamma\right)^2}\sqrt{\|v - \overline{v}\|^2 + \|u - \overline{u}\|^2} \\ &= \sqrt{1 + 2L_g^2 + 2\left(\frac{\alpha}{t} + \gamma\right)^2}\|(u, v) - (\overline{u}, \overline{v})\|. \end{aligned}$$

Obviously, the Lipschitz constant function $t \mapsto L(t) := \sqrt{1 + 2L_g^2 + 2(\alpha/t + \gamma)^2}$ is continuous, hence integrable, on $[t_0, T]$ for all $t_0 < T < +\infty$, consequently, $L \in L^1_{\mathrm{loc}}([t_0, +\infty))$.

Next we show that $F(\cdot, u, v) \in L^1_{\mathrm{loc}}([t_0, +\infty), \mathbb{R}^n \times \mathbb{R}^n)$ for all $u, v \in \mathbb{R}^n$. Let $u, v \in \mathbb{R}^n$ be fixed. For $t_0 < T < +\infty$, one has

$$\begin{aligned} \int_{t_0}^{T} \|F(t, u, v)\|\,\mathrm{d}t &= \int_{t_0}^{T} \sqrt{\|v\|^2 + \left\| \left(\frac{\alpha}{t} + \gamma\right)v + \nabla g(u) \right\|^2}\,\mathrm{d}t \\ &\leq \int_{t_0}^{T} \sqrt{\left(1 + 2\left(\frac{\alpha}{t} + \gamma\right)^2\right)\|v\|^2 + 2\|\nabla g(u)\|^2}\,\mathrm{d}t \\ &\leq \sqrt{\|v\|^2 + \|\nabla g(u)\|^2}\int_{t_0}^{T} \sqrt{3 + 2\left(\frac{\alpha}{t} + \gamma\right)^2}\,\mathrm{d}t \end{aligned}$$

and the conclusion follows by the continuity of $t \mapsto \sqrt{3 + 2(\alpha/t + \gamma)^2}$ on $[t_0, T]$.

The Cauchy–Lipschitz–Picard Theorem guarantees existence and uniqueness of the trajectory of the first-order dynamical system (6) and thus of the second-order dynamical system (2). ■

The next result shows that the acceleration of the trajectory generated by (2) is also locally absolutely continuous on $[t_0, +\infty)$.

**Proposition 2.1:** *For the starting points $u_0, v_0 \in \mathbb{R}^n$, let $x$ be the unique strong global solution of* (2). *Then $\ddot{x}$ is locally absolutely continuous on $[t_0, +\infty)$, hence the third-order derivative $x^{(3)}$ exists almost everywhere on $[t_0, +\infty)$.*

***Proof:*** Let $T > 0$ be fixed. According to Theorem 2.1, $X(t) := (x(t), \dot{x}(t))$ is absolutely continuous on $[t_0, T]$. We endow the product space $\mathbb{R}^n \times \mathbb{R}^n$ with the 1-norm. For arbitrary $s, t \in [t_0, T]$, we have

$$\begin{aligned}
\|\dot{X}(s) - \dot{X}(t)\|_1 &= \|F(s, X(s)) - F(t, X(t))\|_1 \\
&= \left\| \left( \dot{x}(s) - \dot{x}(t), -\left(\frac{\alpha}{s} + \gamma\right)\dot{x}(s) + \left(\frac{\alpha}{t} + \gamma\right)\dot{x}(t) - \nabla g(x(s)) + \nabla g(x(t)) \right) \right\|_1 \\
&\leq (1+\gamma)\|\dot{x}(s) - \dot{x}(t)\| + \left\| \frac{\alpha}{s}\dot{x}(s) - \frac{\alpha}{t}\dot{x}(t) \right\| + \|\nabla g(x(s)) - \nabla g(x(t))\| \\
&\leq (1+\gamma)\|\dot{x}(s) - \dot{x}(t)\| + \frac{|\alpha|}{s}\|\dot{x}(s) - \dot{x}(t)\| + \left\| \frac{\alpha}{s}\dot{x}(t) - \frac{\alpha}{t}\dot{x}(t) \right\| + L_g\|x(s) - x(t)\| \\
&\leq L_1\|\dot{x}(s) - \dot{x}(t)\| + L_2\left|\frac{\alpha}{s} - \frac{\alpha}{t}\right| + L_g\|x(s) - x(t)\|,
\end{aligned}$$

where

$$L_1 := \max_{t\in[t_0,T]} \left(1 + \gamma + \frac{|\alpha|}{t}\right) \quad \text{and} \quad L_2 := \max_{t\in[t_0,T]} \|\dot{x}(t)\|.$$

Let be $\varepsilon > 0$. Since the functions $\dot{x}(\cdot), t \mapsto \alpha/t$ and $x(\cdot)$ are absolutely continuous on $[t_0, T]$, there exists $\eta > 0$ such that for any finite family of intervals $I_k = (a_k, b_k) \subseteq [t_0, T]$, the implication

$$\begin{aligned}
&\left( I_k \cap I_j = \emptyset \text{ and } \sum_k |b_k - a_k| < \eta \right) \\
&\implies \sum_k \|\dot{x}(b_k) - \dot{x}(a_k)\| < \frac{\varepsilon}{3L_1}, \quad \sum_k \left|\frac{\alpha}{b_k} - \frac{\alpha}{a_k}\right| < \frac{\varepsilon}{3L_2} \quad \text{and} \quad \sum_k \|x(b_k) - x(a_k)\| < \frac{\varepsilon}{3L_g}
\end{aligned}$$

holds. Consequently,

$$\sum_k \|\dot{X}(a_k) - \dot{X}(b_k)\| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

hence $\dot{X}(\cdot) = (\dot{x}(\cdot), \ddot{x}(\cdot))$ is absolutely continuous on $[t_0, T]$, which shows that $\ddot{x}$ is absolutely continuous $[t_0, T]$. This proves that $\ddot{x}$ is locally absolutely continuous on $[t_0, +\infty)$, which means that the third-order derivative $x^{(3)}$ exists almost everywhere on $[t_0, +\infty)$. ■

The following results provides an estimate for the third-order derivative of the strong global solution of the dynamical system (2) in terms its first- and second-order derivatives.

**Lemma 2.1:** *For the starting points $u_0, v_0 \in \mathbb{R}^n$, let $x$ be the unique strong global solution of* (2). *Then, for almost every $t \in [t_0, +\infty)$, it holds*

$$\|x^{(3)}(t)\| \leq \left(L_g + \frac{|\alpha|}{t^2}\right)\|\dot{x}(t)\| + \left(\gamma + \frac{|\alpha|}{t}\right)\|\ddot{x}(t)\|. \tag{7}$$

***Proof:*** Let $t \in [t_0, +\infty)$ be such that $\dot{X}(t) = F(t, X(t))$. We have for almost every $h > 0$ that

$$\begin{aligned}
&\|\dot{X}(t+h) - \dot{X}(t)\|_1 = \|F(t+h, X(t+h)) - F(t, X(t))\|_1 \\
&= \left\| \left( \dot{x}(t+h) - \dot{x}(t), -\left(\frac{\alpha}{t+h} + \gamma\right)\dot{x}(t+h) + \left(\frac{\alpha}{t} + \gamma\right)\dot{x}(t) - \nabla g(x(t+h)) + \nabla g(x(t)) \right) \right\|_1 \\
&= \|\dot{x}(t+h) - \dot{x}(t)\| + \left\| -\left(\frac{\alpha}{t+h} + \gamma\right)\dot{x}(t+h) + \left(\frac{\alpha}{t} + \gamma\right)\dot{x}(t) - \nabla g(x(t+h)) + \nabla g(x(t)) \right\| \\
&\leq (1+\gamma)\|\dot{x}(t+h) - \dot{x}(t)\| + \left\| \frac{\alpha}{t+h}\dot{x}(t+h) - \frac{\alpha}{t}\dot{x}(t) \right\| + \|\nabla g(x(t+h)) - \nabla g(x(t))\| \\
&\leq (1+\gamma)\|\dot{x}(t+h) - \dot{x}(t)\| + \left\| \frac{\alpha}{t+h}\dot{x}(t+h) - \frac{\alpha}{t}\dot{x}(t) \right\| + L_g\|x(t+h) - x(t)\|.
\end{aligned}$$

Hence,

$$\begin{aligned}
\left\| \frac{\dot{X}(t+h) - \dot{X}(t)}{h} \right\|_1 \leq (1+\gamma) \left\| \frac{\dot{x}(t+h) - \dot{x}(t)}{h} \right\| + \left\| \frac{(\alpha/(t+h))\dot{x}(t+h) - (\alpha/t)\dot{x}(t)}{h} \right\| \\
+ L_g \left\| \frac{x(t+h) - x(t)}{h} \right\|.
\end{aligned}$$

By taking the limit as $h \to 0$, we obtain

$$\|\ddot{X}(t)\|_1 \leq (1+\gamma)\|\ddot{x}(t)\| + \left\| \left(\frac{\alpha}{t}\dot{x}(t)\right)' \right\| + L_g\|\dot{x}(t)\|.$$

Since $\|\ddot{X}(t)\|_1 = \|x^{(3)}(t)\| + \|\ddot{x}(t)\|$, we conclude

$$\|x^{(3)}(t)\| \leq \left(L_g + \frac{|\alpha|}{t^2}\right)\|\dot{x}(t)\| + \left(\gamma + \frac{|\alpha|}{t}\right)\|\ddot{x}(t)\|.$$

■

**Remark 2.1:** For

$$N := \max_{t \geq t_0} \left( L_g + \frac{|\alpha|}{t^2}, \gamma + \frac{|\alpha|}{t} \right),$$

we have that

$$\|x^{(3)}(t)\| \leq N(\|\ddot{x}(t)\| + \|\dot{x}(t)\|)$$

for almost every $t \in [t_0, +\infty)$.

## 3. Convergence of trajectories

In this section, we study the convergence of the trajectory of the dynamical system (2). We denote by

$$\omega(x) := \{\overline{x} \in \mathbb{R}^n : \exists t_k \longrightarrow +\infty \text{ such that } x(t_k) \longrightarrow \overline{x} \text{ as } k \longrightarrow +\infty\}$$

the set of limit points of the trajectory $x$.

Before proving a first result in this sense, we recall two technical lemmas which will play an essential role in the asymptotic analysis.

**Lemma 3.1:** *Suppose that* $F : [0, +\infty) \to \mathbb{R}$ *is locally absolutely continuous and bounded below and that there exists* $G \in L^1([0, +\infty))$ *such that for almost every* $t \in [0, +\infty)$

$$\frac{\mathrm{d}}{\mathrm{d}t}F(t) \leq G(t).$$

*Then there exists* $\lim_{t\to+\infty} F(t) \in \mathbb{R}$.

**Lemma 3.2:** *Suppose that* $F : [0, +\infty) \to [0, +\infty)$ *is locally absolutely continuous and* $F \in L^p([0, +\infty)), 1 \leq p < \infty$, *and that there exists* $G : [0, +\infty) \to \mathbb{R}, G \in L^r([0, +\infty)), 1 \leq r \leq \infty$, *such that for almost every* $t \in [0, +\infty)$

$$\frac{\mathrm{d}}{\mathrm{d}t}F(t) \leq G(t).$$

*Then it holds* $\lim_{t\to+\infty} F(t) = 0$.

**Theorem 3.1:** *Assume that g is bounded from below and, for* $u_0, v_0 \in \mathbb{R}^n$, *let x be the unique strong global solution of the dynamical system* (2). *Then the following statements are true:*

(i) $\dot{x}, \ddot{x} \in L^2([t_0, +\infty), \mathbb{R}^n)$;
(ii) *there exists* $\beta > 0$ *such that the limit* $\lim_{t\longrightarrow+\infty} g(\beta\dot{x}(t) + x(t))$ *exists and is finite;*
(iii) $\lim_{t\longrightarrow+\infty} \ddot{x}(t) = 0$ *and* $\lim_{t\longrightarrow+\infty} \dot{x}(t) = 0$;
(iv) $\omega(x) \subseteq \mathrm{crit}(g)$.

***Proof:*** Choose $0 < \beta < \min(2/L_g, (\sqrt{L_g^2 + 2\gamma L_g} - L_g)/L_g)$. By using the $L_g$-Lipschitz continuity of $\nabla g$, for almost every $t \in [t_0, +\infty)$ it holds

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}g(\beta\dot{x}(t) + x(t)) &= \langle \beta\ddot{x}(t) + \dot{x}(t), \nabla g(\beta\dot{x}(t) + x(t))\rangle \\
&= \langle \beta\ddot{x}(t) + \dot{x}(t), \nabla g(\beta\dot{x}(t) + x(t)) - \nabla g(x(t))\rangle \\
&\quad + \left\langle \beta\ddot{x}(t) + \dot{x}(t), -\ddot{x}(t) - \left(\frac{\alpha}{t} + \gamma\right)\dot{x}(t)\right\rangle \\
&\leq -\beta\|\ddot{x}(t)\|^2 - \left(1 + \beta\gamma + \frac{\alpha\beta}{t}\right)\langle \ddot{x}(t), \dot{x}(t)\rangle - \left(\gamma + \frac{\alpha}{t}\right)\|\dot{x}(t)\|^2 \\
&\quad + L_g\|\beta\ddot{x}(t) + \dot{x}(t)\|\|\beta\dot{x}(t)\|.
\end{aligned}$$

Taking into account that

$$\|\beta\ddot{x}(t) + \dot{x}(t)\|\|\beta\dot{x}(t)\| \leq \beta^2\|\ddot{x}(t)\|\|\dot{x}(t)\| + \beta\|\dot{x}(t)\|^2 \leq \frac{1}{2}\beta^2\|\ddot{x}(t)\|^2 + \left(\beta + \frac{1}{2}\beta^2\right)\|\dot{x}(t)\|^2$$

and

$$-\left(1 + \beta\gamma + \frac{\alpha\beta}{t}\right)\langle \ddot{x}(t), \dot{x}(t)\rangle = -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left[\left(1 + \beta\gamma + \frac{\alpha\beta}{t}\right)\|\dot{x}(t)\|^2\right] - \frac{\alpha\beta}{2t^2}\|\dot{x}(t)\|^2,$$

we obtain for almost every $t \in [t_0, +\infty)$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(g(\beta\dot{x}(t) + x(t)) + \frac{1}{2}\left(1 + \beta\gamma + \frac{\alpha\beta}{t}\right)\|\dot{x}(t)\|^2\right)$$

$$\leq \left(-\beta + \frac{L_g\beta^2}{2}\right)\|\ddot{x}(t)\|^2 + \left(-\gamma + L_g\beta + \frac{L_g\beta^2}{2} - \frac{\alpha}{t} - \frac{\alpha\beta}{2t^2}\right)\|\dot{x}(t)\|^2. \tag{8}$$

Since $0 < \beta < \min(2/L_g, (\sqrt{L_g^2 + 2\gamma L_g} - L_g)/L_g)$, there exists $t_1 > 0$ such that for every $t \geq t_1$ it holds

$$1 + \beta\gamma + \frac{\alpha\beta}{t} > 0, \quad -\beta + \frac{L_g\beta^2}{2} < 0 \quad \text{and} \quad -\gamma + L_g\beta + \frac{L_g\beta^2}{2} - \frac{\alpha}{t} - \frac{\alpha\beta}{2t^2} < 0. \tag{9}$$

For simplicity, we denote

$$A := -\beta + \frac{L_g\beta^2}{2} \quad \text{and} \quad B(t) := -\gamma + L_g\beta + \frac{L_g\beta^2}{2} - \frac{\alpha}{t} - \frac{\alpha\beta}{2t^2} \quad \forall\, t \in [t_0, +\infty).$$

Let be $T > t_1$. Since $x \in C^2([t_1, T], \mathbb{R}^n)$, we have $x, \dot{x}, \ddot{x} \in L^2([t_1, T], \mathbb{R}^n)$. Furthermore, by the $L_g$-Lipschitz property of $\nabla g$, it holds $\nabla g \in L^2([t_1, T], \mathbb{R}^n)$. By integrating (8) on $[t_1, T]$, we obtain

$$\begin{aligned} &g(\beta\dot{x}(T) + x(T)) + \frac{1}{2}\left(1 + \beta\gamma + \frac{\alpha\beta}{T}\right)\|\dot{x}(T)\|^2 - \int_{t_1}^{T} A\|\ddot{x}(t)\|^2 \mathrm{d}t - \int_{t_1}^{T} B(t)\|\dot{x}(t)\|^2 \mathrm{d}t \\ &\leq g(\beta\dot{x}(t_1) + x(t_1)) + \frac{1}{2}\left(1 + \beta\gamma + \frac{\alpha\beta}{t_1}\right)\|\dot{x}(t_1)\|^2. \end{aligned} \tag{10}$$

Taking into account that $g$ is bounded from bellow, by letting $T \longrightarrow +\infty$, we obtain

$$\int_{t_1}^{\infty} (-A\|\ddot{x}(t)\|^2)\, \mathrm{d}t < +\infty \quad \text{and} \quad \int_{t_1}^{\infty} (-B(t)\|\dot{x}(t)\|^2)\, \mathrm{d}t < +\infty$$

Consequently $\|\ddot{x}(\cdot)\|^2,\ B(\cdot)\|\dot{x}(\cdot)\|^2 \in L^1([t_0, +\infty), \mathbb{R})$, hence

$$\ddot{x}, \quad \dot{x} \in L^2([t_0, +\infty), \mathbb{R}^n).$$

On the other hand, (8) and Lemma 3.1 ensure that the limit

$$\lim_{t \longrightarrow +\infty} \left(g(\beta\dot{x}(t) + x(t)) + \frac{1}{2}\left(1 + \beta\gamma + \frac{\alpha\beta}{t}\right)\|\dot{x}(t)\|^2\right) \tag{11}$$

exists and is finite.

As for almost every $t \in [t_0, +\infty)$

$$\frac{\mathrm{d}}{\mathrm{d}t}(\|\dot{x}(t)\|^2) = 2\langle \ddot{x}(t), \dot{x}(t)\rangle \leq \|\dot{x}(t)\|^2 + \|\ddot{x}(t)\|^2$$

and $\|\dot{x}(\cdot)\|^2 + \|\ddot{x}(\cdot)\|^2 \in L^1([t_0, +\infty))$, according to Lemma 3.2, it follows that $\lim_{t \longrightarrow +\infty} \dot{x}(t) = 0$.

As for almost every $t \in [t_0, +\infty)$

$$\frac{\mathrm{d}}{\mathrm{d}t}(\|\ddot{x}(t)\|^2) = 2\langle x^{(3)}(t), \ddot{x}(t)\rangle \leq \|x^{(3)}(t)\|^2 + \|\ddot{x}(t)\|^2$$

and $\|x^{(3)}(\cdot)\|^2 + \|\ddot{x}(\cdot)\|^2 \in L^1([t_0, +\infty))$ (see Remark 2.1 and (i)), according to Lemma 3.2, it follows that $\lim_{t \longrightarrow +\infty} \ddot{x}(t) = 0$.

Finally, by using that $\lim_{t\longrightarrow+\infty}\dot{x}(t)=0$, (11) becomes

$$\exists \lim_{t\longrightarrow+\infty}\left(g(\beta\dot{x}(t)+x(t))+\frac{1}{2}\left(1+\beta\gamma+\frac{\alpha\beta}{t}\right)\|\dot{x}(t)\|^2\right)=\lim_{t\longrightarrow+\infty}g(\beta\dot{x}(t)+x(t))\in\mathbb{R}. \quad (12)$$

Let $\overline{x}\in\omega(x)$. Then there exists a sequence $t_k\longrightarrow+\infty$, $k\longrightarrow+\infty$ such that $x(t_k)\longrightarrow\overline{x}$ as $k\longrightarrow+\infty$. By using the continuity of $\nabla g$, we have

$$0=\lim_{k\longrightarrow+\infty}\left(\ddot{x}(t_k)+\left(\frac{\alpha}{t_k}+\gamma\right)\dot{x}(t_k)+\nabla g(x(t_k))\right)=\nabla g(\overline{x}),$$

which shows that $\overline{x}\in\mathrm{crit}(g)$. ■

In the following result, we use the distance function to a set, defined for $A\subseteq\mathbb{R}^n$ as $\mathrm{dist}(x,A)=\inf_{y\in A}\|x-y\|$ for all $x\in\mathbb{R}^n$. The following result is a direct consequence of Theorem 3.1.

**Lemma 3.3:** *Assume that g is bounded from below and, for $u_0,v_0\in\mathbb{R}^n$, let x be the unique strong global solution of the dynamical system* (2)*. Define*

$$H:\mathbb{R}^n\times\mathbb{R}^n\longrightarrow\mathbb{R},\quad H(x,y)=g(x)+\tfrac{1}{2}\|x-y\|^2.$$

*Let be $0<\beta<\min(2/L_g,(\sqrt{L_g^2+2\gamma L_g}-L_g)/L_g)$ and $t_1>0$ such that for every $t\geq t_1$ the inequalities* (9) *hold. For every $t\in[t_1,+\infty)$, define*

$$u(t):=\beta\dot{x}(t)+x(t),\quad v(t):=\left(\sqrt{1+\beta\gamma+\frac{\alpha\beta}{t}}+\beta\right)\dot{x}(t)+x(t),$$

$$A=-\beta+\frac{L_g\beta^2}{2},\quad B(t):=-\gamma+L_g\beta+\frac{L_g\beta^2}{2}-\frac{\alpha}{t}-\frac{\alpha\beta}{2t^2}\quad \textit{and}$$

$$p(t):=L_g\beta+\gamma+\frac{|\alpha|}{t}+2\sqrt{1+\beta\gamma+\frac{\alpha\beta}{t}}.$$

*Then the following statements are true:*

(i) $\omega(u)=\omega(v)=\omega(x)$;
(ii) $\frac{\mathrm{d}}{\mathrm{d}t}H(u(t),v(t))\leq A\|\ddot{x}(t)\|^2+B(t)\|\dot{x}(t)\|^2\leq 0$ *for almost every* $t\geq t_1$;
(iii) *the limit* $\lim_{t\longrightarrow+\infty}H(u(t),v(t))=\lim_{t\longrightarrow+\infty}g(\beta\dot{x}(t)+x(t))$ *exists and is finite;*
(iv) *H is finite and constant on* $\omega(u,v)=\{(\overline{x},\overline{x})\in\mathbb{R}^n\times\mathbb{R}^n:\overline{x}\in\omega(x)\}$;
(v) $\|\nabla H(u(t),v(t))\|\leq\|\ddot{x}(t)\|+p(t)\|\dot{x}(t)\|$ *for almost every* $t\geq t_1$;
(vi) $\omega(u,v)\subseteq\mathrm{crit}(H)$.

*If x is bounded, then*

(vii) $\omega(u,v)$ *is nonempty and compact;*
(viii) $\lim_{t\longrightarrow+\infty}\mathrm{dist}((u(t),v(t)),\omega(u,v))=0$.

***Proof:*** (i) According to Theorem 3.1(iii),

$$\lim_{t\longrightarrow+\infty}\beta\dot{x}(t)=\lim_{t\longrightarrow+\infty}\left(\sqrt{1+\beta\gamma+\frac{\alpha\beta}{t}}+\beta\right)\dot{x}(t)=0,$$

hence $\omega(u)=\omega(v)=\omega(x)$.

(ii) and (iii) are nothing else than (8) and (12), respectively.

(iv) follows directly from (iii).

(v) Since $\nabla H(x,y) = (\nabla g(x) + x - y, y - x)$ for every $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$, by using (2), we have for almost every $t \in [t_1, +\infty)$

$$\begin{aligned}
\|\nabla H(u(t), v(t))\| &\leq \|\nabla g(u(t))\| + 2\|u(t) - v(t)\| \\
&\leq \|\nabla g(u(t)) - \nabla g(x(t))\| + \|\nabla g(x(t))\| + 2\|u(t) - v(t)\| \\
&\leq L_g\beta\|\dot{x}(t)\| + \left\|-\ddot{x}(t) - \left(\gamma + \frac{\alpha}{t}\right)\dot{x}(t)\right\| + 2\|u(t) - v(t)\| \\
&\leq \|\ddot{x}(t)\| + \left(L_g\beta + \gamma + \frac{|\alpha|}{t} + 2\sqrt{1 + \beta\gamma + \frac{\alpha\beta}{t}}\right)\|\dot{x}(t)\| = \|\ddot{x}(t)\| + p(t)\|\dot{x}(t)\|.
\end{aligned}$$

(vi) Since

$$\mathrm{crit}H = \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n : \nabla H(x,y) = (0,0)\} = \{(\overline{x}, \overline{x}) \in \mathbb{R}^n \times \mathbb{R}^n : \overline{x} \in \mathrm{crit}(g)\}$$

and (see (i))

$$\omega(u,v) = \{(\overline{x}, \overline{x}) \in \mathbb{R}^n \times \mathbb{R}^n : \overline{x} \in \omega(x)\},$$

by Theorem 3.1(iv) one obtains

$$\omega(u,v) \subseteq \mathrm{crit}(H).$$

Assume that $x$ is bounded.

(vii) Since $\dot{x}(t) \longrightarrow 0$, $t \longrightarrow +\infty$, we obtain that $u$ and $v$ are bounded, too. Thus, the set of limit points $\omega(u,v)$ is nonempty. Furthermore, since $\omega(u,v) = \{(\overline{x}, \overline{x}) \in \mathbb{R}^n \times \mathbb{R}^n : \overline{x} \in \omega(x)\}$ and $\omega(x)$ is bounded, it is enough to show that $\omega(x)$ is closed.

Let be $(\overline{x}_n)_{n\geq 1} \subseteq \omega(x)$ and assume that $\lim_{n\longrightarrow+\infty} \overline{x}_n = x^*$. We show that $x^* \in \omega(x)$. Obviously, for every $n \geq 1$, there exists a sequence $t_k^n \longrightarrow +\infty$, $k \longrightarrow +\infty$, such that

$$\lim_{k\longrightarrow+\infty} x(t_k^n) = \overline{x}_n.$$

Let be $\epsilon > 0$. Since $\lim_{n\longrightarrow+\infty} \overline{x}_n = x^*$, there exists $N(\epsilon) \in \mathbb{N}$ such that for every $n \geq N(\epsilon)$ it holds

$$\|\overline{x}_n - x^*\| < \frac{\epsilon}{2}.$$

Let $n \geq 1$ be fixed. Since $\lim_{k\longrightarrow+\infty} x(t_k^n) = \overline{x}_n$, there exists $k(n,\epsilon) \in \mathbb{N}$ such that for every $k \geq k(n,\epsilon)$ it holds

$$\|x(t_k^n) - \overline{x}_n\| < \frac{\epsilon}{2}.$$

Let be $k_n \geq k(n,\varepsilon)$ such that $t_{k_n}^n > n$. Obviously $t_{k_n}^n \longrightarrow \infty$ as $n \longrightarrow +\infty$ and for every $n \geq N(\epsilon)$

$$\|x(t_{k_n}^n) - x^*\| < \epsilon.$$

Hence,

$$\lim_{n\longrightarrow+\infty} x(t_{k_n}^n) = x^*,$$

thus $x^* \in \omega(x)$.

(viii) follows from the definition of the set $\omega(u,v)$. ■

**Remark 3.1:** Combining (iii) and (iv) in Lemma 3.3, it follows that for every $\overline{x} \in \omega(x)$ and $t_k \longrightarrow +\infty$ such that $x(t_k) \longrightarrow \overline{x}$ as $k \longrightarrow +\infty$ we have

$$\lim_{k \longrightarrow +\infty} H(u(t_k), v(t_k)) = H(\overline{x}, \overline{x}).$$

The convergence of the trajectory generated by the dynamical system (2) will be shown in the framework of functions satisfying the *Kurdyka–Łojasiewicz property*. For $\eta \in (0, +\infty]$, we denote by $\Theta_\eta$ the class of concave and continuous functions $\varphi : [0, \eta) \to [0, +\infty)$ such that $\varphi(0) = 0$, $\varphi$ is continuously differentiable on $(0, \eta)$, continuous at 0 and $\varphi'(s) > 0$ for all $s \in (0, \eta)$.

**Definition 3.1** (*Kurdyka–Łojasiewicz property*)**:** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. We say that $f$ satisfies the *Kurdyka–Łojasiewicz (KL) property* at $\overline{x} \in \mathbb{R}^n$ if there exist $\eta \in (0, +\infty]$, a neighbourhood $U$ of $\overline{x}$ and a function $\varphi \in \Theta_\eta$ such that for all $x$ in the intersection

$$U \cap \{x \in \mathbb{R}^n : f(\overline{x}) < f(x) < f(\overline{x}) + \eta\},$$

the following inequality holds

$$\varphi'(f(x) - f(\overline{x}))\|\nabla f(x))\| \geq 1.$$

If $f$ satisfies the KL property at each point in $\mathbb{R}^n$, then $f$ is called a *KL function*.

The origins of this notion go back to the pioneering work of Łojasiewicz [19], where it is proved that for a real-analytic function $f : \mathbb{R}^n \to \mathbb{R}$ and a critical point $\overline{x} \in \mathbb{R}^n$ (that is $\nabla f(\overline{x}) = 0$), there exists $\theta \in [1/2, 1)$ such that the function $|f - f(\overline{x})|^\theta \|\nabla f\|^{-1}$ is bounded around $\overline{x}$. This corresponds to the situation when $\varphi(s) = C(1-\theta)^{-1} s^{1-\theta}$. The result of Łojasiewicz allows the interpretation of the KL property as a re-parametrization of the function values in order to avoid flatness around the critical points. Kurdyka [20] extended this property to differentiable functions definable in an o-minimal structure. Further extensions to the nonsmooth setting can be found in [21–24].

To the class of KL functions belong semi-algebraic, real sub-analytic, semiconvex, uniformly convex and convex functions satisfying a growth condition. We refer the reader to [21–27] and the references therein for more details regarding all the classes mentioned above and illustrating examples.

An important role in our convergence analysis will be played by the following uniformized KL property given in [27, Lemma 6].

**Lemma 3.4:** *Let $\Omega \subseteq \mathbb{R}^n$ be a compact set and let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. Assume that $f$ is constant on $\Omega$ and $f$ satisfies the KL property at each point of $\Omega$. Then there exist $\epsilon, \eta > 0$ and $\varphi \in \Theta_\eta$ such that for all $\overline{x} \in \Omega$ and for all $x$ in the intersection*

$$\{x \in \mathbb{R}^n : \operatorname{dist}(x, \Omega) < \varepsilon\} \cap \{x \in \mathbb{R}^n : f(\overline{x}) < f(x) < f(\overline{x}) + \eta\}, \tag{13}$$

*the following inequality holds*

$$\varphi'(f(x) - f(\overline{x}))\|\nabla f(x)\| \geq 1. \tag{14}$$

The following theorem is the main result of the paper and concerns the global asymptotic convergence of the trajectory generated by (2).

**Theorem 3.2:** *Assume that $g$ is bounded from below and, for $u_0, v_0 \in \mathbb{R}^n$, let $x$ be the unique strong global solution of* (2)*. Suppose that*

$$H : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R},\ H(x, y) = g(x) + \tfrac{1}{2}\|x - y\|^2$$

*is a KL function and $x$ is bounded. Then the following statements are true:*

(*a*) $\dot{x}, \ddot{x} \in L^1([t_0, +\infty), \mathbb{R}^n)$;
(*b*) *there exists* $\overline{x} \in \operatorname{crit}(g)$ *such that* $\lim_{t \longrightarrow +\infty} x(t) = \overline{x}$.

***Proof:*** Let be $0 < \beta < \min(2/L_g, (\sqrt{L_g^2 + 2\gamma L_g} - L_g)/L_g)$ and $t_1 > 0$ such that for every $t \geq t_1$ the inequalities (9) hold. Furthermore, we will use the notations made in Lemma 3.3, according to which we can choose $(\tilde{x}, \tilde{x}) \in \omega(u, v)$. It holds

$$\lim_{t \longrightarrow +\infty} H(u(t), v(t)) = H(\tilde{x}, \tilde{x}).$$

*Case I*. There exists $\bar{t} \geq t_1$ such that $H(u(\bar{t}), v(\bar{t})) = H(\tilde{x}, \tilde{x})$. From Lemma 3.3(ii), we have

$$\frac{\mathrm{d}}{\mathrm{d}t} H(u(t), v(t)) \leq A\|\ddot{x}(t)\|^2 + B(t)\|\dot{x}(t)\|^2 \leq 0 \text{ for almost every } t \geq t_1,$$

hence

$$H(u(t), v(t)) \leq H(\tilde{x}, \tilde{x}) \text{ for every } t \geq \bar{t}.$$

On the other hand,

$$H(u(t), v(t)) \geq \lim_{t \longrightarrow +\infty} H(u(t), v(t)) = H(\tilde{x}, \tilde{x}) \text{ for every } t \geq t_1,$$

hence

$$H(u(t), v(t)) = H(\tilde{x}, \tilde{x}) \text{ for every } t \geq \bar{t}.$$

Hence, $(\mathrm{d}/\mathrm{d}t)H(u(t), v(t)) = 0$, which leads to

$$0 \leq A\|\ddot{x}(t)\|^2 + B(t)\|\dot{x}(t)\|^2 \leq 0 \text{ for almost every } t \geq \bar{t}.$$

Since $A < 0$ and $B(t) < 0$ for every $t \geq t_1$, the latter inequality can hold only if

$$\dot{x}(t) = \ddot{x}(t) = 0 \text{ for almost every } t \geq \bar{t}.$$

Consequently, $\dot{x}, \ddot{x} \in L^1([t_0, +\infty), \mathbb{R}^n)$ and $x$ is constant on $[\bar{t}, +\infty)$.

*Case II*. We assume that $H(u(t), v(t)) > H(\tilde{x}, \tilde{x})$ holds for every $t \geq t_1$. Let $\Omega := \omega(u, v)$. According to Lemma 3.3, $\Omega$ is nonempty and compact and $H$ is constant on $\Omega$. Since $H$ is a KL function, according to Lemma 3.4, there exist $\varepsilon, \eta > 0$ and $\varphi \in \Theta_\eta$ such that for every $(\tilde{z}, \tilde{w})$ in the intersection

$$\{(z, w) \in \mathbb{R}^n \times \mathbb{R}^n : \operatorname{dist}((z, w), \Omega) < \varepsilon\} \cap \{(z, w) \in \mathbb{R}^n \times \mathbb{R}^n : H(\overline{x}, \overline{x}) < H(z, w) < H(\tilde{x}, \tilde{x}) + \eta\}$$

one has

$$\varphi'(H(\tilde{z}, \tilde{w}) - H(\tilde{x}, \tilde{x}))\|\nabla H(\tilde{z}, \tilde{w})\| \geq 1.$$

Since $\lim_{t \longrightarrow +\infty} \operatorname{dist}(u(t), v(t), \Omega) = 0$, there exists $t_2 \geq t_1$ such that $\operatorname{dist}(u(t), v(t)), \Omega) < \epsilon$ for every $t \geq t_2$. Since $\lim_{t \longrightarrow +\infty} H(u(t), v(t)) = H(\overline{x}, \overline{x})$, there exists $t_3 \geq t_1$ such that $H(\overline{x}, \overline{x}) < H(u(t), v(t)) < H(\overline{x}, \overline{x}) + \eta$ for every $t \geq t_3$. Hence, for every $t \geq T := \max(t_2, t_3)$, we have

$$\varphi'(H(u(t), v(t)) - H(\overline{x}, \overline{x})) \cdot \|\nabla H(u(t), v(t))\| \geq 1.$$

According to Lemma 3.3 (ii) and (v), we have $(\mathrm{d}/\mathrm{d}t)H(u(t), v(t)) \leq A\|\ddot{x}(t)\|^2 + B(t)\|\dot{x}(t)\|^2 \leq 0$ and $\|\nabla H(u(t), v(t))\| \leq \|\ddot{x}(t)\| + p(t)\|\dot{x}(t)\|$, hence

$$\frac{\mathrm{d}}{\mathrm{d}t} \varphi(H(u(t), v(t)) - H(\tilde{x}, \tilde{x})) = \varphi'(H(u(t), v(t)) - H(\tilde{x}, \tilde{x})) \frac{\mathrm{d}}{\mathrm{d}t} H(u(t), v(t))$$

$$\leq \frac{A\|\ddot{x}(t)\|^2 + B(t)\|\dot{x}(t)\|^2}{\|\ddot{x}(t)\| + p(t)\|\dot{x}(t)\|}$$

for almost every $t \in [T, +\infty)$.

By integrating on the interval $[T, \overline{T}]$, for $\overline{T} > T$, we obtain

$$\varphi(H(u(\overline{T}), v(\overline{T})) - H(\tilde{x}, \tilde{x})) - \int_T^{\overline{T}} \frac{A\|\ddot{x}(s)\|^2 + B(s)\|\dot{x}(s)\|^2}{\|\ddot{x}(s)\| + p(s)\|\dot{x}(s)\|} \, \mathrm{d}s \leq \varphi(H(u(T), v(T)) - H(\tilde{x}, \tilde{x})).$$

Since $\varphi$ is bounded from below, $A < 0, B(s) < 0$ and $p(s) > 0$ for every $s \geq T$ and $\overline{T}$ was arbitrarily chosen, we obtain that

$$0 \leq \int_T^{+\infty} \frac{-A\|\ddot{x}(s)\|^2 - B(s)\|\dot{x}(s)\|^2}{\|\ddot{x}(s)\| + p(s)\|\dot{x}(s)\|} \, \mathrm{d}s \leq \varphi(H(u(T), v(T)) - H(\overline{x}, \overline{x})),$$

which leads to

$$t \mapsto \frac{\|\ddot{x}(t)\|^2}{\|\ddot{x}(t)\| + p(t)\|\dot{x}(t)\|}, \quad t \mapsto \frac{\|\dot{x}(t)\|^2}{\|\ddot{x}(t)\| + p(t)\|\dot{x}(t)\|} \in L^1([t_0, +\infty), \mathbb{R}^n)$$

and further to

$$t \mapsto \frac{\|\ddot{x}(t)\|\|\dot{x}(t)\|}{\|\ddot{x}(t)\| + p(t)\|\dot{x}(t)\|} \in L^1([t_0, +\infty), \mathbb{R}^n).$$

Since $p$ is bounded from above on $[t_0, +\infty)$ and

$$\|\dot{x}(t)\| + \|\ddot{x}(t)\| = \frac{\|\ddot{x}(t)\|^2}{\|\ddot{x}(t)\| + p(t)\|\dot{x}(t)\|} + \frac{p(t)\|\dot{x}(t)\|^2}{\|\ddot{x}(t)\| + p(t)\|\dot{x}(t)\|} + \frac{(1 + p(t))\|\ddot{x}(t)\|\|\dot{x}(t)\|}{\|\ddot{x}(t)\| + p(t)\|\dot{x}(t)\|},$$

we obtain that

$$\dot{x}, \ddot{x} \in L^1([t_0, +\infty), \mathbb{R}^n).$$

Finally, since $\dot{x} \in L^1([t_0, +\infty), \mathbb{R}^n)$, the limit $\lim_{t \longrightarrow +\infty} x(t)$ exists and it is finite. In conclusion, there exists $\overline{x} \in \mathrm{crit}(g)$ such that

$$\lim_{t \longrightarrow +\infty} x(t) = \overline{x}.$$

■

**Remark 3.2:** According to Remark 2.1, there exists $N > 0$ such that $\|x^{(3)}(t)\| \leq N(\|\ddot{x}(t)\| + \|\dot{x}(t)\|)$ for almost every $t \geq t_0$, hence, under the hypotheses of Theorem 3.2, one has

$$x^{(3)} \in L^1([t_0, +\infty), \mathbb{R}^n).$$

**Remark 3.3:** Since the class of semi-algebraic functions is closed under addition [see, for example, [27]] and $(x, y) \mapsto \frac{1}{2}\|x - y\|^2$ is semi-algebraic, the conclusion of the previous theorem holds, if, instead of asking that $H$ is a KL function, we ask that $g$ is semi-algebraic.

**Remark 3.4:** Assume that $g$ is coercive, that is

$$\lim_{\|u\| \to +\infty} g(u) = +\infty.$$

For $u_0, v_0 \in \mathbb{R}^n$, let $x \in C^2([0, +\infty), \mathbb{R}^n)$ be the unique global solution of (2). Then $x$ is bounded.

Indeed, notice that $g$ is bounded from below, being a continuous and coercive function [see [28]]. From (10), it follows that $\beta\dot{x}(T) + x(T)$ is contained for every $T \geq t_1$ in a lower level set of $g$, which is bounded. According to Theorem 3.1, $\beta\dot{x}(t) \longrightarrow 0$, $t \longrightarrow +\infty$, which implies that $x$ is bounded.

## 4. Convergence rates

In this section, we will assume that the regularized function $H$ satisfies the Lojasiewicz property, which, as noted in the previous section, corresponds to a particular choice of the desingularizing function $\varphi$ [see [19,22,25]].

**Definition 4.1:** Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a differentiable function. The function $f$ is said to fulfil the Łojasiewicz property, if for every $\overline{x} \in \operatorname{crit} f$ there exist $K, \epsilon > 0$ and $\theta \in (0, 1)$ such that

$$|f(x) - f(\overline{x})|^{\theta} \leq K\|\nabla f(x)\| \text{ for every } x \text{ fulfilling } \|x - \overline{x}\| < \epsilon.$$

The number $\theta$ is called the Łojasiewicz exponent of $f$ at the critical point $\overline{x}$.

In the following theorem, we provide convergence rates for the trajectory generated by (2), its velocity and acceleration in terms of the Łojasiewicz exponent of $H$ [see, also, [22,25]].

**Theorem 4.1:** *Assume that $g$ is bounded from below and, for $u_0, v_0 \in \mathbb{R}^n$, let $x$ be the unique strong global solution of* (2). *Suppose that $x$ is bounded, let $\overline{x} \in \operatorname{crit}(g)$ be such that $\lim_{t \longrightarrow +\infty} x(t) = \overline{x}$ and suppose that*

$$H : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}, \quad H(x, y) = g(x) + \tfrac{1}{2}\|x - y\|^2$$

*fulfils the Łojasiewicz property at $(\overline{x}, \overline{x}) \in \operatorname{crit} H$ with Łojasiewicz exponent $\theta$. Let be* (see Remark 2.1)

$$N := \max_{t \geq t_0} \left( L_g + \frac{|\alpha|}{t^2}, \gamma + \frac{|\alpha|}{t} \right).$$

*Then, there exist $a_1, a_2, a_3, a_4 > 0$ and $T > 0$ such that for almost every $t \in [T, +\infty)$, the following statements are true:*

(a) *if $\theta \in (0, \frac{1}{2})$, then $x$ converges in finite time;*
(b) *if $\theta = \frac{1}{2}$, then $\|x(t) - \overline{x}\| \leq a_1 e^{-a_2 t}$, $\|\dot{x}(t)\| \leq a_1 e^{-a_2 t}$ and $\|\ddot{x}(t)\| \leq N a_1 e^{-a_2 t}$;*
(c) *if $\theta \in (\frac{1}{2}, 1)$, then $\|x(t) - \overline{x}\| \leq (a_3 t + a_4)^{-(1-\theta)/(2\theta-1)}$, $\|\dot{x}(t)\| \leq (a_3 t + a_4)^{-(1-\theta)/(2\theta-1)}$ and $\|\ddot{x}(t)\| \leq N(a_3 t + a_4)^{-(1-\theta)/(2\theta-1)}$.*

***Proof:*** Let be $0 < \beta < \min(2/L_g, (\sqrt{L_g^2 + 2\gamma L_g} - L_g)/L_g)$ and $t_1 > 0$ such that for every $t \geq t_1$ the inequalities (9) hold. We define for every $t \in [t_1, +\infty)$

$$\sigma(t) := \int_t^{+\infty} (\|\dot{x}(s)\| + \|\ddot{x}(s)\|)\, \mathrm{d}s.$$

Let $t \in [t_1, +\infty)$ be fixed. For $T \geq t$, we have

$$\|x(t) - \overline{x}\| = \left\| x(T) - \overline{x} - \int_t^T \dot{x}(s)\, \mathrm{d}s \right\| \leq \|x(T) - \overline{x}\| + \int_t^T \|\dot{x}(s)\|\, \mathrm{d}s.$$

By taking the limit as $T \longrightarrow +\infty$, we obtain

$$\|x(t) - \overline{x}\| \leq \int_t^{+\infty} \|\dot{x}(s)\|\, \mathrm{d}s \leq \sigma(t). \tag{15}$$

Furthermore, for $T > t$, we have

$$\|\dot{x}(t)\| = \left\| \dot{x}(T) - \int_t^T \ddot{x}(s)\, \mathrm{d}s \right\| \leq \|\dot{x}(T)\| + \int_t^T \|\ddot{x}(s)\|\, \mathrm{d}s.$$

By taking the limit as $T \longrightarrow +\infty$, we obtain

$$\|\dot{x}(t)\| \leq \int_t^{+\infty} \|\ddot{x}(s)\| \, \mathrm{d}s \leq \sigma(t). \tag{16}$$

According to Remark 2.1, it holds $\|x^{(3)}(t)\| \leq N(\|\ddot{x}(t)\| + \|\dot{x}(t)\|)$ for almost every $t \geq t_1$,

$$\|\ddot{x}(t)\| = \left\| \ddot{x}(T) - \int_t^T x^{(3)}(s) \, \mathrm{d}s \right\| \leq \|\ddot{x}(T)\| + \int_t^T \|x^{(3)}(s)\| \, \mathrm{d}s \leq \|\ddot{x}(T)\| + N \int_t^T (\|\ddot{x}(s)\| + \|\dot{x}(s)\|) \, \mathrm{d}s \quad \forall \, T > t.$$

By taking the limit as $T \longrightarrow +\infty$, we obtain

$$\|\ddot{x}(t)\| \leq N\sigma(t). \tag{17}$$

Next, we show that there exists $m < 0$ such that

$$\frac{A\|\ddot{x}(t)\|^2 + B(t)\|\dot{x}(t)\|^2}{\|\ddot{x}(t)\| + p(t)\|\dot{x}(t)\|} \leq m(\|\dot{x}(t)\| + \|\ddot{x}(t)\|). \tag{18}$$

Indeed,

$$(\|\ddot{x}(t)\| + p(t)\|\dot{x}(t)\|)(\|\dot{x}(t)\| + \|\ddot{x}(t)\|) = \|\ddot{x}(t)\|^2 + (1 + p(t))\|\dot{x}(t)\|\|\ddot{x}(t)\| + p(t)\|\dot{x}(t)\|^2$$
$$\leq \left(\frac{3}{2} + \frac{p(t)}{2}\right)\|\ddot{x}(t)\|^2 + \left(\frac{1}{2} + \frac{3p(t)}{2}\right)\|\dot{x}(t)\|^2 \leq \frac{A}{m}\|\ddot{x}(t)\|^2 + \frac{B(t)}{m}\|\dot{x}(t)\|^2,$$

where

$$m := \max\left(\max_{t \geq t_1} \frac{A}{3/2 + p(t)/2}, \max_{t \geq t_1} \frac{B(t)}{(3/2)p(t) + 1/2}\right).$$

It is an easy verification that $m < 0$.

As we have seen in the proof of Theorem 3.2, if there exists $\bar{t} \geq t_1$ such that $H(u(\bar{t}), v(\bar{t})) = H(\overline{x}, \overline{x})$, then $x$ is constant on $[\bar{t}, +\infty)$ and the conclusion follows.

On the other hand, if for every $t \geq t_1$ one has that $H(u(t), v(t)) > H(\overline{x}, \overline{x})$, then according to the proof of Theorem 3.2, there exist $\epsilon > 0$ and $T \geq t_1$ such that

$$\|(u(t), v(t)) - (\overline{x}, \overline{x})\| < \epsilon$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}(H(u(t), v(t)) - H(\overline{x}, \overline{x}))^{1-\theta} \leq \frac{A\|\ddot{x}(t)\|^2 + B(t)\|\dot{x}(t)\|^2}{\|\ddot{x}(t)\| + p(t)\|\dot{x}(t)\|} \text{ for almost every } t \geq T.$$

Busing (18), we obtain that

$$M(\|\dot{x}(t)\| + \|\ddot{x}(t)\|) + \frac{\mathrm{d}}{\mathrm{d}t}(H(u(t), v(t)) - H(\overline{x}, \overline{x}))^{1-\theta} \leq 0 \text{ for almost every } t \geq T,$$

where $M := -m > 0$.

For $t \geq T$, we integrate the last relation on the interval $[t, \widetilde{T}]$, where $\widetilde{T} > t$, which yields

$$M\int_t^{\widetilde{T}} (\|\dot{x}(s)\| + \|\ddot{x}(s)\|)\,\mathrm{d}s + (H(u(\widetilde{T}), v(\widetilde{T})) - H(\overline{x}, \overline{x}))^{1-\theta} \leq (H(u(t), v(t)) - H(\overline{x}, \overline{x}))^{1-\theta}.$$

By taking the limits as $\widetilde{T} \longrightarrow +\infty$, we get

$$M\sigma(t) \leq (H(u(t), v(t)) - H(\overline{x}, \overline{x}))^{1-\theta}.$$

On the other hand, according to the KL property for $H$ and Lemma 3.3 (v), we have

$$(H(u(t), v(t)) - H(\overline{x}, \overline{x}))^{\theta} \leq K\|\nabla H(u(t), v(t))\| \leq K(\|\ddot{x}(t)\| + p(t)\|\dot{x}(t)\|) \text{ for almost every } t \geq T,$$

hence

$$M\sigma(t) \leq K^{(1-\theta)/\theta}(\|\ddot{x}(t)\| + p(t)\|\dot{x}(t)\|)^{(1-\theta)/\theta} \text{ for almost every } t \geq T.$$

By denoting $a := \max_{t\geq T} p(t) \in (0, +\infty)$, one can easily see that $a > 1$ and so

$$M\sigma(t) \leq (aK)^{(1-\theta)/\theta}(\|\dot{x}(t)\| + \|\ddot{x}(t)\|)^{(1-\theta)/\theta} \text{ for almost every } t \geq T.$$

Taking into account that $\|\dot{x}(t)\| + \|\ddot{x}(t)\| = -\dot{\sigma}(t)$, the previous inequality is nothing else than

$$-c\sigma^{\theta/(1-\theta)}(t) \geq \dot{\sigma}(t) \text{ for almost every } t \geq T, \tag{19}$$

where $c := M^{\theta/(1-\theta)}/aK > 0$.

If $\theta = \frac{1}{2}$, then (19) becomes $c\sigma(t) + \dot{\sigma}(t) \leq 0$ for almost every $t \geq T$. By multiplying with $e^{ct}$ and integrating on $[T, t]$, it follows that there exists $a_1 > 0$ such that

$$\sigma(t) \leq a_1 e^{-a_2 t} \text{ for every } t \geq T,$$

where $a_2 = c$. Using (15)–(17), we obtain

$$\|x(t) - \overline{x}\| \leq a_1 e^{-a_2 t}, \|\dot{x}(t)\| \leq a_1 e^{-a_2 t} \quad \text{and} \quad \|\ddot{x}(t)\| \leq N a_1 e^{-a_2 t} \text{ for every } t \geq T,$$

which proves (b).

Assume now that $0 < \theta < \frac{1}{2}$. In this case, (19) leads to

$$\frac{\mathrm{d}}{\mathrm{d}t}\sigma^{(1-2\theta)/(1-\theta)}(t) = \frac{1-2\theta}{1-\theta}\sigma^{-\theta/(1-\theta)}(t)\dot{\sigma}(t) \leq -c\frac{1-2\theta}{1-\theta} \text{ for almost every } t \geq T.$$

By integrating on $[T, t]$ we obtain

$$\sigma^{(1-2\theta)/(1-\theta)}(t) \leq -\overline{\alpha}t + \overline{\beta}, \text{ for every } t \geq T,$$

where $\overline{\alpha} > 0$. Then there exists $\widehat{T} \geq T$ such that $\sigma(t) \leq 0$ for every $t \geq \widehat{T}$, thus, $x$ is constant on $[\widehat{T}, +\infty)$ and (a) follows.

Assume now that $\frac{1}{2} < \theta < 1$. In this case, (19) leads to

$$\frac{\mathrm{d}}{\mathrm{d}t}\sigma^{(1-2\theta)/(1-\theta)}(t) = \frac{1-2\theta}{1-\theta}\sigma^{-\theta/(1-\theta)}(t)\dot{\sigma}(t) \geq c\frac{2\theta-1}{1-\theta} \text{ for almost every } t \geq T.$$

By integrating on $[T, t]$ we obtain

$$\sigma^{(1-2\theta)/(1-\theta)}(t) \geq a_3 t + a_4 \text{ for every } t \geq T,$$

where $a_3, a_4 > 0$, or, equivalently,

$$\sigma(t) \leq (a_3 t + a_4)^{-(1-\theta)/(2\theta-1)} \text{ for every } t \geq T.$$

Using again (15)–(17), we obtain

$$\|x(t) - \overline{x}\| \leq (a_3 t + a_4)^{-(1-\theta)/(2\theta-1)}, \quad \|\dot{x}(t)\| \leq (a_3 t + a_4)^{-(1-\theta)/(2\theta-1)} \quad \text{and}$$

$$\|\ddot{x}(t)\| \leq N(a_3 t + a_4)^{-(1-\theta)/(2\theta-1)} \text{ for every } t \geq T,$$

which proves (c). ■

## Acknowledgements

The authors are thankful to an anonymous reviewer for comments and remarks which were helpful to gain a better insight into the asymptotic behaviour of the trajectories of the studied dynamical system.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

Radu Ioan Boţ's research was partially supported by FWF (Austrian Science Fund) [project I 2419-N32] ; Ernö Robert Csetnek's research was supported by FWF (Austrian Science Fund) [project P 29809-N32] and Szilárd Csaba László's research was supported by a grant of Ministry of Research and Innovation - Unitatea Executiva pentru Finantarea Invatamantului Superior, a Cercetarii, Dezvoltarii si Inovarii (CNCS-UEFISCDI) [project number PN-III-P1-1.1-TE-2016-0266] within PNCDI III.

## ORCID

*Radu Ioan Boţ* http://orcid.org/0000-0002-4469-314X

## References

[1] Su W, Boyd S, Candes EJ. A differential equation for modeling Nesterov's accelerated gradient method: theory and insights. J Mach Learn Res. 2016;17:1–43.
[2] Nesterov YE. A method for solving the convex programming problem with convergence rate $O(1/k^2)$. (Russian) Dokl Akad Nauk SSSR. 1983;269(3):543–547.
[3] Attouch H, Chbani Z, Peypouquet J, et al. Fast convergence of inertial dynamics and algorithms with asymptotic vanishing viscosity. Math Program. 2018;168(1–2, Ser. B):123–175.
[4] Attouch H, Peypouquet J, Redont P. Fast convex optimization via inertial dynamics with Hessian driven damping. J Differ Equat. 2016;261(10):5734–5783.
[5] Chambolle A, Dossal C. On the convergence of the iterates of the "fast iterative shrinkage/thresholding algorithm". J Optim Theory Appl. 2015;166(3):968–982.
[6] Beck A, Teboulle M. A fast iterative shrinkage- thresholding algorithm for linear inverse problems. SIAM J Imag Sci. 2009;2(1):183–202.
[7] Attouch H, Chbani Z, Riahi H. Rate of convergence of the Nesterov accelerated gradient method in the subcritical case $\alpha \leq 3$. 2017. arXiv:1706.05671.
[8] Haraux A, Jendoubi MA. Asymptotics for a second order differential equation with a linear, slowly time-decaying damping term. Evol Equat Contr Theor. 2013;2(3):461–470.
[9] Balti M. Asymptotic behavior for second-order differential equations with nonlinear slowly time-decaying damping and integrable source. Electron J Differ Equat. 2015;302:11pp.
[10] Bégout P, Bolte J, Jendoubi MA. On damped second-order gradient systems. J Differ Equat. 2015;259:3115–3143.
[11] Haraux A, Jendoubi MA. Convergence of solutions of second-order gradient-like systems with analytic nonlinearities. J Differ Equat. 1998;144(2):313–320.
[12] Chill R, Jendoubi MA. Asymptotics for a second order differential equation with a linear, slowly time-decaying damping term. Evol Equat Contr Theor. 2013;2(3):461–470.
[13] Boţ RI, Csetnek ER, László SC. Approaching nonsmooth nonconvex minimization through second-order proximal-gradient dynamical systems. J Evol Equat. 2018. doi:10.1007/s00028-018-0441-7.
[14] Abbas B, Attouch H, Svaiter BF. Newton-like dynamics and forward-backward methods for structured monotone inclusions in Hilbert spaces. J Optim Theory Appl. 2014;161(2):331–360.
[15] Attouch H, Svaiter BF. A continuous dynamical Newton-like approach to solving monotone inclusions. SIAM J Control Optim. 2011;49(2):574–598.
[16] Haraux A. Systèmes Dynamiques Dissipatifs et Applications, Recherches en Mathématiques Appliquées 17, Masson, Paris, 1991.
[17] Sontag ED. Mathematical control theory. Deterministic finite-dimensional systems. 2nd ed. New York: Springer-Verlag; 1998. (Texts in Applied Mathematics 6).

[18] Alvarez F, Attouch H, Bolte J, et al. A second-order gradient-like dissipative dynamical system with Hessian-driven damping. Application to optimization and mechanics. J Math Pures et Appl. 2002;81(8):747–779.
[19] Łojasiewicz S. Une propriété topologique des sous-ensembles analytiques réels, Les Équations aux Dérivées Partielles, Éditions du Centre National de la Recherche Scientifique Paris, 87–89, 1963.
[20] Kuryka K. On gradients of functions definable in o-minimal structures. Ann l Fourier (Grenoble). 1998;48(3):769–783.
[21] Attouch H, Bolte J, Redont P, et al. Proximal alternating minimization and projection methods for nonconvex problems: an approach based on the Kurdyka–Łojasiewicz inequality. Math Oper Res. 2010;35(2):438–457.
[22] Bolte J, Daniilidis A, Lewis A. The Łojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems. SIAM J Optim. 2006;17(4):1205–1223.
[23] Bolte J, Daniilidis A, Lewis A, et al. Clarke subgradients of stratifiable functions. SIAM J Optim. 2007;18(2):556–572.
[24] Bolte J, Daniilidis A, Ley O, et al. Characterizations of Łojasiewicz inequalities: subgradient flows, talweg, convexity. Trans Am Math Soc. 2010;362(6):3319–3363.
[25] Attouch H, Bolte J. On the convergence of the proximal algorithm for nonsmooth functions involving analytic features. Math Program. 2009;116(1–2, Series B):5–16.
[26] Attouch H, Bolte J, Svaiter BF. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods. Math Program. 2013;137(1–2, Series A):91–129.
[27] Bolte J, Sabach S, Teboulle M. Proximal alternating linearized minimization for nonconvex and nonsmooth problems. Math Program Ser A. 2014;146(1–2):459–494.
[28] Rockafellar RT, Wets RJ-B. Variational analysis, fundamental principles of mathematical sciences, 317. Berlin: Springer-Verlag; 1998.