Newton-like dynamics associated to nonconvex optimization problems

Radu Ioan Boţ and Ernö Robert Csetnek

Abstract. We consider the dynamical system

$$\begin{cases} v(t) \in \partial\phi(x(t)) \\ \lambda\dot{x}(t) + \dot{v}(t) + v(t) + \nabla\psi(x(t)) = 0, \end{cases}$$

where $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a proper, convex and lower semicontinuous function, $\psi : \mathbb{R}^n \to \mathbb{R}$ is a (possibly nonconvex) smooth function and $\lambda > 0$ is a parameter which controls the velocity. We show that the set of limit points of the trajectory x is contained in the set of critical points of the objective function $\phi + \psi$, which is here seen as the set of the zeros of its limiting subdifferential. If the objective function is smooth and satisfies the Kurdyka-Łojasiewicz property, then we can prove convergence of the whole trajectory x to a critical point. Furthermore, convergence rates for the orbits are obtained in terms of the Łojasiewicz exponent of the objective function, provided the latter satisfies the Łojasiewicz property.

Mathematics Subject Classification (2000). 34G25, 47J25, 47H05, 90C26, 90C30, 65K10.

Keywords. dynamical systems, Newton-like methods, Lyapunov analysis, nonsmooth optimization, limiting subdifferential, Kurdyka-Łojasiewicz property.

1. Introduction and preliminaries

The dynamical system

$$\begin{cases} v(t) \in T(x(t)) \\ \lambda(t)\dot{x}(t) + \dot{v}(t) + v(t) = 0, \end{cases} \tag{1.1}$$

where $\lambda : [0, +\infty) \to [0, +\infty)$ and $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a (set-valued) maximally monotone operator, has been introduced and investigated in [10] as a continuous

The research of RIB has been partially supported by FWF (Austrian Science Fund), project I 2419-N32. The research of ERC has been supported by FWF (Austrian Science Fund), project P 29809-N32.

version of Newton and Levenberg-Marquardt-type algorithms. It has been shown that under mild conditions on λ the trajectory $x(t)$ converges weakly to a zero of the operator T, while $v(t)$ converges to zero as $t \to +\infty$.

These investigations have been continued in [2] in the context of solving optimization problems of the form

$$\inf_{x\in\mathbb{R}^n}\{\phi(x)+\psi(x)\}, \tag{1.2}$$

where $\phi:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ is a proper, convex and lower semicontinuous function and $\psi:\mathbb{R}^n\to\mathbb{R}$ is a convex and differentiable function with locally Lipschitz-continuous gradient. More precisely, problem (1.2) has been approached via the dynamical system

$$\left\{\begin{array}{l} v(t)\in\partial\phi(x(t)) \\ \lambda(t)\dot{x}(t)+\dot{v}(t)+v(t)+\nabla\psi(x(t))=0, \end{array}\right. \tag{1.3}$$

where $\partial\phi$ is the convex subdifferential of ϕ. It has been shown in [2] that if the set of minimizers of (1.2) is nonempty and some mild conditions on the damping function λ are satisfied, then the trajectory $x(t)$ converges to a minimizer of (1.2) as $t\to+\infty$. Further investigations on dynamical systems of similar type have been reported in [1] and [21].

The aim of this paper is to perform an asymptotic analysis of the dynamical system (1.3) in the absence of the convexity of ψ, for constant damping function λ and by assuming that the objective function of (1.2) satisfies the *Kurdyka-Łojasiewicz* property, in other words is a *KL function*. To the class of KL functions belong semialgebraic, real subanalytic, uniformly convex and convex functions satisfying a growth condition. The convergence analysis relies on methods of real algebraic geometry introduced by Łojasiewicz [30] and Kurdyka [28] and developed recently in the nonsmooth setting by Attouch, Bolte and Svaiter [7] and Bolte, Sabach and Teboulle [16].

Optimization problems involving KL functions have attracted the interest of the community since the works of Łojasiewicz [30], Simon [34], Haraux and Jendoubi [26]. The most important contributions of the last years in the field include the works of Alvarez, Attouch, Bolte and Redont [3, Section 4] and Bolte, Daniilidis and Lewis [12, Section 4]. Ever since the interest in this topic increased continuously (see [5, 6, 7, 15, 16, 20, 18, 19, 23, 24, 27, 32]).

In the first part of the paper we show that the set of limit points of the trajectory x generated by (1.3) is entirely contained in the set of critical points of the objective function $\phi+\psi$, which is seen as the set of zeros of its limiting subdifferential. Under some supplementary conditions, including the Kurdyka-Łojasiewicz property, we prove the convergence of the trajectory x to a critical point of $\phi+\psi$. Furthermore, convergence rates for the orbits are obtained in terms of the Łojasiewicz exponent of the objective function, provided the latter satisfies the Łojasiewicz property.

In the following we recall some notions and results which are needed throughout the paper. We consider on $\mathbb{R}^n$ the Euclidean scalar product and the corresponding norm denoted by $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$, respectively.

The *domain* of the function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is defined by $\operatorname{dom} f = \{x \in \mathbb{R}^n : f(x) < +\infty\}$ and we say that f is *proper*, if it has a nonempty domain. For the following generalized subdifferential notions and their basic properties we refer to [17, 31, 33]. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper and lower semicontinuous function. The *Fréchet (viscosity) subdifferential* of f at $x \in \operatorname{dom} f$ is the set

$$\hat{\partial} f(x) = \left\{ v \in \mathbb{R}^n : \liminf_{y \to x} \frac{f(y) - f(x) - \langle v, y - x \rangle}{\|y - x\|} \geq 0 \right\}.$$

If $x \notin \operatorname{dom} f$, we set $\hat{\partial} f(x) := \emptyset$. The *limiting (Mordukhovich) subdifferential* is defined at $x \in \operatorname{dom} f$ by

$$\partial_L f(x) = \{v \in \mathbb{R}^n : \exists x_k \to x, f(x_k) \to f(x) \text{ and } \exists v_k \in \hat{\partial} f(x_k), v_k \to v \text{ as } k \to +\infty\},$$

while for $x \notin \operatorname{dom} f$, we set $\partial_L f(x) := \emptyset$. Obviously, $\hat{\partial} f(x) \subseteq \partial_L f(x)$ for each $x \in \mathbb{R}^n$.

When f is convex, these subdifferential notions coincide with the *convex subdifferential*, thus $\hat{\partial} f(x) = \partial_L f(x) = \partial f(x) = \{v \in \mathbb{R}^n : f(y) \geq f(x) + \langle v, y - x \rangle \ \forall y \in \mathbb{R}^n\}$ for all $x \in \mathbb{R}^n$.

The following *closedness criterion* of the graph of the limiting subdifferential will be used in the convergence analysis: if $(x_k)_{k \in \mathbb{N}}$ and $(v_k)_{k \in \mathbb{N}}$ are sequences in $\mathbb{R}^n$ such that $v_k \in \partial_L f(x_k)$ for all $k \in \mathbb{N}$, $(x_k, v_k) \to (x, v)$ and $f(x_k) \to f(x)$ as $k \to +\infty$, then $v \in \partial_L f(x)$.

The Fermat rule reads in this nonsmooth setting as follows: if $x \in \mathbb{R}^n$ is a local minimizer of f, then $0 \in \partial_L f(x)$. We denote by

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

the set of *(limiting)-critical points* of f.

When f is continuously differentiable around $x \in \mathbb{R}^n$ we have $\partial_L f(x) = \{\nabla f(x)\}$. We will also make use of the following subdifferential sum rule: if $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is proper and lower semicontinuous and $h : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function, then $\partial_L (f + h)(x) = \partial_L f(x) + \nabla h(x)$ for all $x \in \mathbb{R}^n$.

Further, we recall the notion of a locally absolutely continuous function and state two of its basic properties.

Definition 1.1. (see [10, 2]) A function $x : [0, +\infty) \to \mathbb{R}^n$ is said to be locally absolutely continuous, if it absolutely continuous on every interval $[0, T]$ for $T > 0$.

Remark 1.2. (a) 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 integration.

(b) If $x : [0, T] \to \mathbb{R}^n$ is absolutely continuous for $T > 0$ and $B : \mathbb{R}^n \to \mathbb{R}^n$ is L-Lipschitz continuous for $L \geq 0$, then the function $z = B \circ x$ is absolutely continuous, too. Moreover, z is differentiable almost everywhere on $[0, T]$ and the inequality $\|\dot{z}(t)\| \leq L\|\dot{x}(t)\|$ holds for almost every $t \in [0, T]$.

The following two results, which can be interpreted as continuous versions of the quasi-Fejér monotonicity for sequences, will play an important role in the asymptotic analysis of the trajectories of the dynamical system (1.3). For their proofs we refer the reader to [2, Lemma 5.1] and [2, Lemma 5.2], respectively.

Lemma 1.3. *Suppose that* $F : [0, +\infty) \to \mathbb{R}$ *is locally absolutely continuous and bounded from below and that there exists* $G \in L^1([0, +\infty))$ *such that for almost every* $t \in [0, +\infty)$

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

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

Lemma 1.4. *If* $1 \leq p < \infty$, $1 \leq r \leq \infty$, $F : [0, +\infty) \to [0, +\infty)$ *is locally absolutely continuous,* $F \in L^p([0, +\infty))$, $G : [0, +\infty) \to \mathbb{R}$, $G \in L^r([0, +\infty))$ *and for almost every* $t \in [0, +\infty)$

$$\frac{d}{dt}F(t) \leq G(t),$$

then $\lim_{t\to+\infty} F(t) = 0$.

The following result, which is due to Brézis ([22, Lemme 3.3, p. 73]; see also [8, Lemma 3.2]), provides an expression for the derivative of the composition of convex functions with absolutely continuous trajectories.

Lemma 1.5. *Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *be a proper, convex and lower semicontinuous function. Let* $x \in L^2([0, T], \mathbb{R}^n)$ *be absolutely continuous such that* $\dot{x} \in L^2([0, T], \mathbb{R}^n)$ *and* $x(t) \in \operatorname{dom} f$ *for almost every* $t \in [0, T]$. *Assume that there exists* $\xi \in L^2([0, T], \mathbb{R}^n)$ *such that* $\xi(t) \in \partial f(x(t))$ *for almost every* $t \in [0, T]$. *Then the function* $t \mapsto f(x(t))$ *is absolutely continuous and for almost every* t *such that* $x(t) \in \operatorname{dom} \partial f$ *we have*

$$\frac{d}{dt}f(x(t)) = \langle \dot{x}(t), h \rangle \ \forall h \in \partial f(x(t)).$$

2. Asymptotic analysis

In this paper we investigate the dynamical system

$$\begin{cases} v(t) \in \partial\phi(x(t)) \\ \lambda\dot{x}(t) + \dot{v}(t) + v(t) + \nabla\psi(x(t)) = 0 \\ x(0) = x_0, v(0) = v_0 \in \partial\phi(x_0), \end{cases} \tag{2.1}$$

where $x_0, v_0 \in \mathbb{R}^n$ and $\lambda > 0$. We assume that $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is proper, convex and lower semicontinuous and $\psi : \mathbb{R}^n \to \mathbb{R}$ is possibly nonconvex and

Fréchet differentiable with L-Lipschitz continuous gradient, for $L > 0$; in other words, $\|\nabla\psi(x) - \nabla\psi(y)\| \leq L\|x - y\|$ for all $x, y \in \mathbb{R}^n$.

In the following we specify what we understand under a solution of the dynamical system (2.1).

Definition 2.1. Let $x_0, v_0 \in \mathbb{R}^n$ and $\lambda > 0$ be such that $v_0 \in \partial\phi(x_0)$. We say that the pair (x, v) is a strong global solution of (2.1) if the following properties are satisfied:

(i) $x, v : [0, +\infty) \to \mathbb{R}^n$ are locally absolutely continuous functions;
(ii) $v(t) \in \partial\phi(x(t))$ for every $t \in [0, +\infty)$;
(iii) $\lambda\dot{x}(t) + \dot{v}(t) + v(t) + \nabla\psi(x(t)) = 0$ for almost every $t \in [0, +\infty)$;
(iv) $x(0) = x_0, v(0) = v_0$.

The existence and uniqueness of the trajectories generated by (2.1) has been investigated in [2]. A careful look at the proofs in [2] reveals the fact that the convexity of ψ is not used in the mentioned results on the existence, but the Lipschitz-continuity of its gradient.

We start our convergence analysis with the following technical result.

Lemma 2.2. *Let $x_0, v_0 \in \mathbb{R}^n$ and $\lambda > 0$ be such that $v_0 \in \partial\phi(x_0)$. Let $(x, v) : [0, +\infty) \to \mathbb{R}^n \times \mathbb{R}^n$ be the unique strong global solution of the dynamical system* (2.1)*. Then the following statements are true:*

(i) $\langle\dot{x}(t), \dot{v}(t)\rangle \geq 0$ *for almost every* $t \in [0, +\infty)$*;*
(ii) $\frac{d}{dt}\phi(x(t)) = \langle\dot{x}(t), v(t)\rangle$ *for almost every* $t \in [0, +\infty)$*.*

Proof. (i) See [10, Proposition 3.1]. The proof relies on the first relation in (2.1) and the monotonicity of the convex subdifferential.

(ii) The proof makes use of Lemma 1.5. This relation has been already stated in [2, relation (51)] without making use in its proof of the convexity of ψ. □

Lemma 2.3. *Let $x_0, v_0 \in \mathbb{R}^n$ and $\lambda > 0$ be such that $v_0 \in \partial\phi(x_0)$. Let $(x, v) : [0, +\infty) \to \mathbb{R}^n \times \mathbb{R}^n$ be the unique strong global solution of the dynamical system* (2.1)*. Suppose that $\phi + \psi$ is bounded from below. Then the following statements are true:*

(i) $\frac{d}{dt}(\phi + \psi)(x(t)) + \lambda\|\dot{x}(t)\|^2 + \langle\dot{x}(t), \dot{v}(t)\rangle = 0$ *for almost every* $t \geq 0$*;*
(ii) $\dot{x}, \dot{v}, v+\nabla\psi(x) \in L^2([0, +\infty); \mathbb{R}^n)$, $\langle\dot{x}(\cdot), \dot{v}(\cdot)\rangle \in L^1([0, +\infty); \mathbb{R})$ *and* $\lim_{t\to+\infty} \dot{x}(t) = \lim_{t\to+\infty} \dot{v}(t) = \lim_{t\to+\infty} \big(v(t) + \nabla\psi(x(t))\big) = 0$*;*
(iii) *the function* $t \to (\phi+\psi)(x(t))$ *is decreasing and* $\exists \lim_{t\to+\infty}(\phi+\psi)\big(x(t)\big) \in \mathbb{R}$*.*

Proof. (i) The statement follows by inner multiplying the both sides of the second relation in (2.1) by $\dot{x}(t)$ and by taking afterwards into consideration Lemma 2.2(ii).

(ii) After integrating the relation (i) and by taking into account that $\phi + \psi$ is bounded from below, we easily derive $\dot{x} \in L^2([0, +\infty); \mathbb{R}^n)$ and $\langle\dot{x}(\cdot), \dot{v}(\cdot)\rangle \in L^1([0, +\infty); \mathbb{R})$ (see also Lemma 2.2(i)). Further, by using the second relation in

(2.1), Remark 1.2(b) and Lemma 2.2(i), we obtain for almost every $t \geq 0$:

$$\begin{aligned}
\frac{d}{dt}\left(\frac{1}{2}\|v(t)+\nabla\psi(x(t))\|^2\right) &= \left\langle \dot{v}(t)+\frac{d}{dt}\nabla\psi(x(t)), v(t)+\nabla\psi(x(t))\right\rangle \\
&= \left\langle \dot{v}(t)+\frac{d}{dt}\nabla\psi(x(t)), -\lambda\dot{x}(t)-\dot{v}(t)\right\rangle \\
&= -\lambda\langle\dot{v}(t),\dot{x}(t)\rangle - \|\dot{v}(t)\|^2 - \lambda\left\langle\frac{d}{dt}\nabla\psi(x(t)),\dot{x}(t)\right\rangle \\
&\quad -\left\langle\frac{d}{dt}\nabla\psi(x(t)),\dot{v}(t)\right\rangle \\
&\leq -\|\dot{v}(t)\|^2 - \lambda\left\langle\frac{d}{dt}\nabla\psi(x(t)),\dot{x}(t)\right\rangle \\
&\quad -\left\langle\frac{d}{dt}\nabla\psi(x(t)),\dot{v}(t)\right\rangle \\
&\leq -\|\dot{v}(t)\|^2 + \lambda L\|\dot{x}(t)\|^2 + L\|\dot{x}(t)\|\cdot\|\dot{v}(t)\| \\
&\leq -\|\dot{v}(t)\|^2 + \lambda L\|\dot{x}(t)\|^2 + L^2\|\dot{x}(t)\|^2 + \frac{1}{4}\|\dot{v}(t)\|^2,
\end{aligned}$$

hence

$$\frac{d}{dt}\left(\frac{1}{2}\|v(t)+\nabla\psi(x(t))\|^2\right) + \frac{3}{4}\|\dot{v}(t)\|^2 \leq L(\lambda+L)\|\dot{x}(t)\|^2. \tag{2.2}$$

Since $\dot{x} \in L^2([0,+\infty);\mathbb{R}^n)$, by a simple integration argument we obtain $\dot{v} \in L^2([0,+\infty);\mathbb{R}^n)$. Considering the second equation in (2.1), we further obtain that $v+\nabla\psi(x) \in L^2([0,+\infty);\mathbb{R}^n)$. This fact combined with Lemma 1.4 and (2.2) implies that $\lim_{t\to+\infty}\big(v(t)+\nabla\psi(x(t))\big)=0$. From the second equation in (2.1) we obtain

$$\lim_{t\to+\infty} \lambda\dot{x}(t)+\dot{v}(t)=0. \tag{2.3}$$

Further, from Lemma 2.2(i) we have for almost every $t \geq 0$

$$\|\dot{v}(t)\|^2 \quad\leq\quad \lambda^2\|\dot{x}(t)\|^2 + 2\lambda\langle\dot{x}(t),\dot{v}(t)\rangle + \|\dot{v}(t)\|^2 = \|\lambda\dot{x}(t)+\dot{v}(t)\|^2,$$

hence from (2.3) we get $\lim_{t\to+\infty}\dot{v}(t)=0$. Combining this with (2.3) we conclude that $\lim_{t\to+\infty}\dot{x}(t)=0$.

(iii) From (i) and Lemma 2.2(i) it follows that

$$\frac{d}{dt}(\phi+\psi)(x(t)) \leq 0 \tag{2.4}$$

for almost every $t \geq 0$. The conclusion follows by applying Lemma 1.3. □

Lemma 2.4. *Let $x_0, v_0 \in \mathbb{R}^n$ and $\lambda > 0$ be such that $v_0 \in \partial\phi(x_0)$. Let $(x,v) : [0,+\infty) \to \mathbb{R}^n \times \mathbb{R}^n$ be the unique strong global solution of the dynamical system (2.1). Suppose that $\phi+\psi$ is bounded from below. Let $(t_k)_{k\in\mathbb{N}}$ be a sequence such that $t_k \to +\infty$ and $x(t_k) \to \overline{x} \in \mathbb{R}^n$ as $k \to +\infty$. Then*

$$0 \in \partial_L(\phi+\psi)(\overline{x}).$$

Proof. From the first relation in (2.1) and the subdifferential sum rule of the limiting subdifferential we derive for any $k \in \mathbb{N}$

$$v(t_k) + \nabla\psi(x(t_k)) \in \partial\phi(x(t_k)) + \nabla\psi(x(t_k)) = \partial_L(\phi + \psi)(x(t_k)). \tag{2.5}$$

Further, we have

$$x(t_k) \to \overline{x} \text{ as } k \to +\infty \tag{2.6}$$

and (see Lemma 2.3(ii))

$$v(t_k) + \nabla\psi(x(t_k)) \to 0 \text{ as } k \to +\infty. \tag{2.7}$$

According to the closedness property of the limiting subdifferential, the proof is complete as soon as we show that

$$(\phi + \psi)(x(t_k)) \to (\phi + \psi)(\overline{x}) \text{ as } k \to +\infty. \tag{2.8}$$

From (2.6), (2.7) and the continuity of $\nabla\psi$ we get

$$v(t_k) \to -\nabla\psi(\overline{x}) \text{ as } k \to +\infty. \tag{2.9}$$

Further, since $v(t_k) \in \partial\phi(x(t_k))$, we have

$$\phi(\overline{x}) \geq \phi(x(t_k)) + \langle v(t_k), \overline{x} - x(t_k)\rangle \ \forall k \in \mathbb{N}.$$

Combining this with (2.6) and (2.9) we derive

$$\limsup_{k \to +\infty} \phi(x(t_k)) \leq \phi(\overline{x}).$$

A direct consequence of the lower semicontinuity of ϕ is the relation

$$\lim_{k \to +\infty} \phi(x(t_k)) = \phi(\overline{x}),$$

which combined with (2.6) and the continuity of ψ yields (2.8). □

We define the *limit set of* x as

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

We use also the *distance function* to a set, defined for $A \subseteq \mathbb{R}^n$ as $\operatorname{dist}(x, A) = \inf_{y \in A} \|x - y\|$ for all $x \in \mathbb{R}^n$.

Lemma 2.5. *Let $x_0, v_0 \in \mathbb{R}^n$ and $\lambda > 0$ be such that $v_0 \in \partial\phi(x_0)$. Let $(x, v) : [0, +\infty) \to \mathbb{R}^n \times \mathbb{R}^n$ be the unique strong global solution of the dynamical system* (2.1)*. Suppose that $\phi + \psi$ is bounded from below and x is bounded. Then the following statements are true:*

(i) $\omega(x) \subseteq \operatorname{crit}(\phi + \psi)$*;*
(ii) $\omega(x)$ *is nonempty, compact and connected;*
(iii) $\lim_{t \to +\infty} \operatorname{dist}\big(x(t), \omega(x)\big) = 0$*;*
(iv) $\phi + \psi$ *is finite and constant on* $\omega(x)$*.*

Proof. Statement (i) is a direct consequence of Lemma 2.4.

Statement (ii) is a classical result from [25]. We also refer the reader to the proof of Theorem 4.1 in [3], where it is shown that the properties of $\omega(x)$ of being nonempty, compact and connected hold for bounded trajectories fulfilling $\lim_{t\to+\infty}\dot{x}(t)=0$.

Statement (iii) follows immediately since $\omega(x)$ is nonempty.

(iv) According to Lemma (2.3)(iii), there exists $\lim_{t\to+\infty}(\phi+\psi)\big(x(t)\big)\in\mathbb{R}$. Let us denote by $l\in\mathbb{R}$ this limit. Take $\overline{x}\in\omega(x)$. Then there exists $t_k\to+\infty$ such that $x(t_k)\to\overline{x}$ as $k\to+\infty$. From the proof of Lemma 2.4 we have that $(\phi+\psi)(x(t_k))\to(\phi+\psi)(\overline{x})$ as $k\to+\infty$, hence $(\phi+\psi)(\overline{x})=l$. □

Remark 2.6. Suppose that $\phi+\psi$ is coercive, in other words,

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

Let $x_0,v_0\in\mathbb{R}^n$ and $\lambda>0$ be such that $v_0\in\partial\phi(x_0)$. Let $(x,v):[0,+\infty)\to\mathbb{R}^n\times\mathbb{R}^n$ be the unique strong global solution of the dynamical system (2.1). Then $\phi+\psi$ is bounded from below and x is bounded.

Indeed, since $\phi+\psi$ is a proper, lower semicontinuous and coercive function, it follows that $\inf_{u\in\mathbb{R}^n}[\phi(u)+\psi(u)]$ is finite and the infimum is attained. Hence $\phi+\psi$ is bounded from below. On the other hand, from Lemma 2.3(iii) it follows

$$(\phi+\psi)(x(T))\leq(\phi+\psi)(x_0)\ \forall T\geq 0.$$

Since $\phi+\psi$ is coercive, the lower level sets of $\phi+\psi$ are bounded, hence the above inequality yields that x is bounded. Notice that in this case v is bounded too, due to the relation $\lim_{t\to+\infty}\big(v(t)+\nabla\psi(x(t))\big)=0$ (Lemma 2.3(ii)) and the Lipschitz continuity of $\nabla\psi$.

3. Convergence of the trajectory when the objective function satisfies the Kurdyka-Łojasiewicz property

In order to enforce the convergence of the whole trajectory $x(t)$ to a critical point of the objective function as $t\to+\infty$ more involved analytic features of the functions have to be considered.

A crucial role in the asymptotic analysis of the dynamical system (2.1) is played by the class of functions satisfying the *Kurdyka-Łojasiewicz* property. For $\eta\in(0,+\infty]$, we denote by Θ_η the class of concave and continuous functions $\varphi:[0,\eta)\to[0,+\infty)$ such that $\varphi(0)=0$, φ 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 $f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ be a proper and lower semicontinuous function. We say that f satisfies the *Kurdyka-Łojasiewicz (KL) property* at $\overline{x}\in\operatorname{dom}\partial_L f=\{x\in\mathbb{R}^n:\partial_L f(x)\neq\emptyset\}$, if there exist

$\eta \in (0, +\infty]$, a neighborhood 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

$$\varphi'(f(x) - f(\overline{x}))\,\mathrm{dist}(0, \partial_L f(x)) \geq 1.$$

If f satisfies the KL property at each point in $\mathrm{dom}\,\partial_L f$, then f is called *KL function.*

The origins of this notion go back to the pioneering work of Łojasiewicz [30], 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 $\varphi(s) = Cs^{1-\theta}$ for $C > 0$. 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 [28] extended this property to differentiable functions definable in o-minimal structures. Further extensions to the nonsmooth setting can be found in [12, 6, 13, 14].

One of the remarkable properties of the KL functions is their ubiquity in applications (see [16]). We refer the reader to [12, 6, 14, 16, 13, 7, 5] and the references therein for more properties of the KL functions and illustrating examples.

In the analysis below the following uniform KL property given in [16, Lemma 6] will be used.

Lemma 3.2. *Let $\Omega \subseteq \mathbb{R}^n$ be a compact set and let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper and lower semicontinuous function. Assume that f is constant on Ω and that it satisfies the KL property at each point of Ω. Then there exist $\varepsilon, \eta > 0$ and $\varphi \in \Theta_\eta$ such that for all $\overline{x} \in \Omega$ and all x in the intersection*

$$\{x \in \mathbb{R}^n : \mathrm{dist}(x, \Omega) < \varepsilon\} \cap \{x \in \mathbb{R}^n : f(\overline{x}) < f(x) < f(\overline{x}) + \eta\} \tag{3.1}$$

the inequality

$$\varphi'(f(x) - f(\overline{x}))\,\mathrm{dist}(0, \partial_L f(x)) \geq 1. \tag{3.2}$$

holds.

Due to some reasons outlined in Remark 3.6 below, we prove the convergence of the trajectory $x(t)$ generated by (2.1) as $t \to +\infty$ under the assumption that $\phi : \mathbb{R}^n \to \mathbb{R}$ is convex and differentiable with ρ^{-1}-Lipschitz continuous gradient for $\rho > 0$. In these circumstances the dynamical system (2.1) reads

$$\begin{cases} v(t) = \nabla\phi(x(t)) \\ \lambda\dot{x}(t) + \dot{v}(t) + \nabla\phi(x(t)) + \nabla\psi(x(t)) = 0 \\ x(0) = x_0, v(0) = v_0 = \nabla\phi(x_0), \end{cases} \tag{3.3}$$

where $x_0, v_0 \in \mathbb{R}^n$ and $\lambda > 0$.

Remark 3.3. We notice that we do not require second order assumptions for ϕ. However, we want to notice that if ϕ is a twice continuously differentiable function, then the dynamical system (3.3) can be equivalently written as

$$\begin{cases} \lambda\dot{x}(t) + \nabla^2\phi(x(t))(\dot{x}(t)) + \nabla\phi(x(t)) + \nabla\psi(x(t)) = 0 \\ x(0) = x_0, v(0) = v_0 = \nabla\phi(x_0), \end{cases} \tag{3.4}$$

where $x_0, v_0 \in \mathbb{R}^n$ and $\lambda > 0$. This is a differential equation with a Hessian-driven damping term. We refer the reader to [3] and [9] for more insights into dynamical systems with Hessian-driven damping terms and for motivations for considering them. Moreover, as in [9], the driving forces have been split as $\nabla\phi + \nabla\psi$, where $\nabla\psi$ stands for classical smooth driving forces and $\nabla\phi$ incorporates the contact forces.

In this context, an improved version of Lemma 2.2(i) can be stated.

Lemma 3.4. *Let $x_0, v_0 \in \mathbb{R}^n$ and $\lambda > 0$ be such that $v_0 = \nabla\phi(x_0)$. Let $(x, v) : [0, +\infty) \to \mathbb{R}^n \times \mathbb{R}^n$ be the unique strong global solution of the dynamical system* (3.3)*. Then:*

$$\langle \dot{x}(t), \dot{v}(t)\rangle \geq \rho\|\dot{v}(t)\|^2 \text{ for almost every } t \in [0, +\infty). \tag{3.5}$$

Proof. Take an arbitrary $\delta > 0$. For $t \geq 0$ we have

$$\begin{aligned} \langle v(t+\delta) - v(t), x(t+\delta) - x(t)\rangle &= \langle \nabla\phi(x(t+\delta)) - \nabla\phi(x(t)), x(t+\delta) - x(t)\rangle \\ &\geq \rho\|\nabla\phi(x(t+\delta)) - \nabla\phi(x(t))\|^2 \\ &= \rho\|v(t+\delta) - v(t)\|^2, \end{aligned} \tag{3.6}$$

where the inequality follows from the Baillon-Haddad Theorem [11, Corollary 18.16]. The conclusion follows by dividing (3.6) by δ^2 and by taking the limit as δ converges to zero from above. □

We are now in the position to prove the convergence of the trajectories generated by (3.3).

Theorem 3.5. *Let $x_0, v_0 \in \mathbb{R}^n$ and $\lambda > 0$ be such that $v_0 = \nabla\phi(x_0)$. Let $(x, v) : [0, +\infty) \to \mathbb{R}^n \times \mathbb{R}^n$ be the unique strong global solution of the dynamical system* (3.3)*. Suppose that $\phi + \psi$ is a KL function which is bounded from below and x is bounded. Then the following statements are true:*

(i) $\dot{x}, \dot{v}, \nabla\phi(x) + \nabla\psi(x) \in L^2([0, +\infty); \mathbb{R}^n)$, $\langle \dot{x}(\cdot), \dot{v}(\cdot)\rangle \in L^1([0, +\infty); \mathbb{R})$ *and* $\lim_{t\to+\infty} \dot{x}(t) = \lim_{t\to+\infty} \dot{v}(t) = \lim_{t\to+\infty} \big(\nabla\phi(x(t)) + \nabla\psi(x(t))\big) = 0$;

(ii) *there exists* $\overline{x} \in \operatorname{crit}(\phi+\psi)$ *(that is* $\nabla(\phi+\psi)(\overline{x}) = 0$*) such that* $\lim_{t\to+\infty} x(t) = \overline{x}$.

Proof. According to Lemma 2.5, we can choose an element $\overline{x} \in \operatorname{crit}(\phi + \psi)$ (that is $\nabla(\phi + \psi)(\overline{x}) = 0$) such that $\overline{x} \in \omega(x)$. According to Lemma 2.3(iii), the proof of Lemma 2.4 and the proof of Lemma 2.5(iv), we have

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

We consider the following two cases.

I. There exists $\overline{t} \geq 0$ such that

$$(\phi + \psi)(x(\overline{t})) = (\phi + \psi)(\overline{x}).$$

From Lemma 2.3(iii) we obtain for every $t \geq \overline{t}$ that

$$(\phi + \psi)(x(t)) \leq (\phi + \psi)(x(\overline{t})) = (\phi + \psi)(\overline{x})$$

Thus $(\phi + \psi)(x(t)) = (\phi + \psi)(\overline{x})$ for every $t \geq \overline{t}$. According to Lemma 2.3(i) and (3.5), it follows that $\dot{x}(t) = \dot{v}(t) = 0$ for almost every $t \in [\overline{t}, +\infty)$, hence x and v are constant on $[\overline{t}, +\infty)$ and the conclusion follows.

II. For every $t \geq 0$ it holds $(\phi + \psi)(x(t)) > (\phi + \psi)(\overline{x})$. Take $\Omega := \omega(x)$.

By using Lemma 2.5(ii), (iv) and the fact that $\phi + \psi$ is a KL function, by Lemma 3.2, there exist positive numbers ϵ and η and a concave function $\varphi \in \Theta_\eta$ such that for all u belonging to the intersection

$$\{u \in \mathbb{R}^n : \operatorname{dist}(u, \Omega) < \epsilon\} \cap \{u \in \mathbb{R}^n : (\phi + \psi)(\overline{x}) < (\phi + \psi)(u) < (\phi + \psi)(\overline{x}) + \eta\}, \tag{3.7}$$

one has

$$\varphi'\Big((\phi + \psi)(u) - (\phi + \psi)(\overline{x})\Big) \cdot \|\nabla\phi(u) + \nabla\psi(u)\| \geq 1. \tag{3.8}$$

Let $t_1 \geq 0$ be such that $(\phi + \psi)(x(t)) < (\phi + \psi)(\overline{x}) + \eta$ for all $t \geq t_1$. Since $\lim_{t\to+\infty} \operatorname{dist}\big(x(t), \Omega\big) = 0$ (see Lemma 2.5(iii)), there exists $t_2 \geq 0$ such that for all $t \geq t_2$ the inequality $\operatorname{dist}\big(x(t), \Omega\big) < \epsilon$ holds. Hence for all $t \geq T := \max\{t_1, t_2\}$, $x(t)$ belongs to the intersection in (3.7). Thus, according to (3.8), for every $t \geq T$ we have

$$\varphi'\Big((\phi + \psi)(x(t)) - (\phi + \psi)(\overline{x})\Big) \cdot \|\nabla\phi(x(t)) + \nabla\psi(x(t))\| \geq 1. \tag{3.9}$$

From the second equation in (3.3) we obtain for almost every $t \in [T, +\infty)$

$$(\lambda\|\dot{x}(t)\| + \|\dot{v}(t)\|) \cdot \varphi'\Big((\phi + \psi)(x(t)) - (\phi + \psi)(\overline{x})\Big) \geq 1. \tag{3.10}$$

By using Lemma 2.3(i), that $\varphi' > 0$ and

$$\frac{d}{dt}\varphi\Big((\phi+\psi)(x(t)) - (\phi+\psi)(\overline{x})\Big) = \varphi'\Big((\phi+\psi)(x(t)) - (\phi+\psi)(\overline{x})\Big)\frac{d}{dt}(\phi+\psi)(x(t)),$$

we further deduce that for almost every $t \in [T, +\infty)$ it holds

$$\frac{d}{dt}\varphi\Big((\phi + \psi)(x(t)) - (\phi + \psi)(\overline{x})\Big) \leq -\frac{\lambda\|\dot{x}(t)\|^2 + \langle \dot{x}(t), \dot{v}(t)\rangle}{\lambda\|\dot{x}(t)\| + \|\dot{v}(t)\|}. \tag{3.11}$$

We invoke now Lemma 3.5 and obtain

$$\frac{d}{dt}\varphi\Big((\phi + \psi)(x(t)) - (\phi + \psi)(\overline{x})\Big) \leq -\frac{\lambda\|\dot{x}(t)\|^2 + \rho\|\dot{v}(t)\|^2}{\lambda\|\dot{x}(t)\| + \|\dot{v}(t)\|}. \tag{3.12}$$

Let $\alpha > 0$ (not depending on t) be such that

$$-\frac{\lambda\|\dot{x}(t)\|^2 + \rho\|\dot{v}(t)\|^2}{\lambda\|\dot{x}(t)\| + \|\dot{v}(t)\|} \leq -\alpha\|\dot{x}(t)\| - \alpha\|\dot{v}(t)\| \ \forall t \geq 0. \tag{3.13}$$

One can for instance chose $\alpha > 0$ such that $2\alpha \max(\lambda, 1) \leq \min(\lambda, \rho)$. From (3.12) we derive the inequality

$$\frac{d}{dt}\varphi\Big((\phi+\psi)(x(t)) - (\phi+\psi)(\overline{x})\Big) \leq -\alpha\|\dot{x}(t)\| - \alpha\|\dot{v}(t)\|, \tag{3.14}$$

which holds for almost every $t \geq T$. Since φ is bounded from below, by integration it follows $\dot{x}, \dot{v} \in L^1([0,+\infty);\mathbb{R}^n)$. From here we obtain that $\lim_{t\to+\infty} x(t)$ exists and the conclusion follows from the results obtained in the previous section. □

Remark 3.6. Taking a closer look at the above proof, one can notice that the inequality (3.11) can be obtained also when $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a (possibly nonsmooth) proper, convex and lower semicontinuous function. Though, in order to conclude that $\dot{x} \in L^1([0,+\infty);\mathbb{R}^n)$ the inequality obtained in Lemma 2.2(i) is not enough. The improved version stated in Lemma 3.4 is crucial in the convergence analysis.

If one attempts to obtain in the nonsmooth setting the inequality stated in Lemma 3.4, from the proof of Lemma 3.4 it becomes clear that one would need the inequality

$$\langle \xi_1^* - \xi_2^*, x_1 - x_2 \rangle \geq \rho\|\xi_1^* - \xi_2^*\|^2$$

for all $(x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^n$ and all $(\xi_1^*, \xi_2^*) \in \mathbb{R}^n \times \mathbb{R}^n$ such that $\xi_1^* \in \partial\phi(x_1)$ and $\xi_2^* \in \partial\phi(x_2)$. This is nothing else than (see for example [11])

$$\langle \xi_1^* - \xi_2^*, x_1 - x_2 \rangle \geq \rho\|\xi_1^* - \xi_2^*\|^2$$

for all $(x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^n$ and all $(\xi_1^*, \xi_2^*) \in \mathbb{R}^n \times \mathbb{R}^n$ such that $x_1 \in \partial\phi^*(\xi_1^*)$ and $x_2 \in \partial\phi^*(\xi_2^*)$. Here $\phi^* : \mathbb{R}^n \to \overline{\mathbb{R}}$ denotes the Fenchel conjugate of ϕ, defined for all $x^* \in \mathbb{R}^n$ by $\phi^*(x^*) = \sup_{x\in\mathbb{R}^n}\{\langle x^*, x\rangle - \phi(x)\}$. The latter inequality is equivalent to $\partial\phi^*$ is ρ-strongly monotone, which is further equivalent (see [35, Theorem 3.5.10] or [11]) to ϕ^* is is strongly convex. This is the same with asking that ϕ is differentiable on the whole $\mathbb{R}^n$ with Lipschitz-continuous gradient (see [11, Theorem 18.15]). In conclusion, the smooth setting provides the necessary prerequisites for obtaining the result in Lemma 3.4 and, finally, Theorem 3.5.

4. Convergence rates

In this subsection we investigate the convergence rates of the trajectories $(x(t), v(t))$ generated by the dynamical system (3.3) as $t \to +\infty$. When solving optimization problems involving KL functions, convergence rates have been proved to depend on the so-called Łojasiewicz exponent (see [30, 12, 5, 24]). The main result of this subsection refers to the KL functions which satisfy Definition 3.1 for $\varphi(s) = Cs^{1-\theta}$, where $C > 0$ and $\theta \in (0,1)$. We recall the following definition considered in [5].

Definition 4.1. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper and lower semicontinuous function. The function f is said to have the Łojasiewicz property, if for every

$\overline{x} \in \operatorname{crit} f$ there exist $C, \varepsilon > 0$ and $\theta \in (0,1)$ such that

$$|f(x) - f(\overline{x})|^{\theta} \leq C\|x^*\| \text{ for every } x \text{ fulfilling } \|x - \overline{x}\| < \varepsilon \text{ and every } x^* \in \partial_L f(x). \tag{4.1}$$

According to [6, Lemma 2.1 and Remark 3.2(b)], the KL property is automatically satisfied at any noncritical point, fact which motivates the restriction to critical points in the above definition. The real number θ in the above definition is called *Łojasiewicz exponent* of the function f at the critical point $\overline{x}$.

The convergence rates obtained in the following theorem are in the spirit of [12] and [5].

Theorem 4.2. *Let $x_0, v_0 \in \mathbb{R}^n$ and $\lambda > 0$ be such that $v_0 = \nabla\phi(x_0)$. Let $(x, v) : [0, +\infty) \to \mathbb{R}^n \times \mathbb{R}^n$ be the unique strong global solution of the dynamical system* (3.3)*. Suppose that x is bounded and $\phi+\psi$ is a function which is bounded from below and satisfies Definition* 3.1 *for $\varphi(s) = Cs^{1-\theta}$, where $C > 0$ and $\theta \in (0,1)$. Then there exists $\overline{x} \in \operatorname{crit}(\phi + \psi)$ (that is $\nabla(\phi + \psi)(\overline{x}) = 0$) such that $\lim_{t\to+\infty} x(t) = \overline{x}$ and $\lim_{t\to+\infty} v(t) = \nabla\phi(\overline{x}) = -\nabla\psi(\overline{x})$. Let θ be the Łojasiewicz exponent of $\phi+\psi$ at $\overline{x}$, according to the Definition* 4.1*. Then there exist $a_1, b_1, a_2, b_2 > 0$ and $t_0 \geq 0$ such that for every $t \geq t_0$ the following statements are true:*

(i) *if $\theta \in (0, \frac{1}{2})$, then x and v converge in finite time;*
(ii) *if $\theta = \frac{1}{2}$, then $\|x(t) - \overline{x}\| + \|v(t) - \nabla\phi(\overline{x})\| \leq a_1 \exp(-b_1 t)$;*
(iii) *if $\theta \in (\frac{1}{2}, 1)$, then $\|x(t) - \overline{x}\| + \|v(t) - \nabla\phi(\overline{x})\| \leq (a_2 t + b_2)^{-\left(\frac{1-\theta}{2\theta-1}\right)}$.*

Proof. According to the proof of Theorem 3.5, $\dot{x}, \dot{v} \in L^1([0, +\infty); \mathbb{R}^n)$ and there exists $\overline{x} \in \operatorname{crit}(\phi+\psi)$, in other words $\nabla(\phi+\psi)(\overline{x}) = 0$, such that $\lim_{t\to+\infty} x(t) = \overline{x}$ and $\lim_{t\to+\infty} v(t) = \nabla\phi(\overline{x}) = -\nabla\psi(\overline{x})$. Let θ be the Łojasiewicz exponent of $\phi+\psi$ at $\overline{x}$, according to the Definition 4.1.

We define $\sigma : [0, +\infty) \to [0, +\infty)$ by (see also [12])

$$\sigma(t) = \int_t^{+\infty} \|\dot{x}(s)\|ds + \int_t^{+\infty} \|\dot{v}(s)\|ds \quad \text{for all } t \geq 0.$$

It is immediate that

$$\|x(t) - \overline{x}\| \leq \int_t^{+\infty} \|\dot{x}(s)\|ds \ \forall t \geq 0. \tag{4.2}$$

Indeed, this follows by noticing that for $T \geq t$

$$\begin{aligned} \|x(t) - \overline{x}\| &= \left\| x(T) - \overline{x} - \int_t^T \dot{x}(s)ds \right\| \\ &\leq \|x(T) - \overline{x}\| + \int_t^T \|\dot{x}(s)\|ds, \end{aligned}$$

and by letting afterwards $T \to +\infty$.

Similarly, we have

$$\|v(t)-\nabla\phi(\overline{x})\| \leq \int_t^{+\infty}\|\dot{v}(s)\|ds \ \forall t\geq 0. \tag{4.3}$$

From (4.2) and (4.3) we derive

$$\|x(t)-\overline{x}\|+\|v(t)-\nabla\phi(\overline{x})\| \leq \sigma(t) \ \forall t\geq 0. \tag{4.4}$$

We assume that for every $t \geq 0$ we have $(\phi+\psi)(x(t)) > (\phi+\psi)(\overline{x})$. As seen in the proof of Theorem 3.5 otherwise the conclusion follows automatically. Furthermore, by invoking again the proof of Theorem 3.5 , there exist $\varepsilon>0$, $t_0\geq 0$ and $\alpha>0$ such that for almost every $t\geq t_0$ (see (3.14))

$$\alpha\|\dot{x}(t)\|+\alpha\|\dot{v}(t)\|+\frac{d}{dt}\Big[(\phi+\psi)(x(t))-(\phi+\psi)(\overline{x})\Big]^{1-\theta}\leq 0 \tag{4.5}$$

and

$$\|x(t)-\overline{x}\|<\varepsilon.$$

We derive by integration for $T\geq t\geq t_0$

$$\alpha\int_t^T\|\dot{x}(s)\|ds+\alpha\int_t^T\|\dot{v}(s)\|ds+\Big[(\phi+\psi)(x(T))-(\phi+\psi)(\overline{x})\Big]^{1-\theta}$$
$$\leq\Big[(\phi+\psi)(x(t))-(\phi+\psi)(\overline{x})\Big]^{1-\theta},$$

hence

$$\alpha\sigma(t)\leq\Big[(\phi+\psi)(x(t))-(\phi+\psi)(\overline{x})\Big]^{1-\theta} \ \forall t\geq t_0. \tag{4.6}$$

Since θ is the Łojasiewicz exponent of $\phi+\psi$ at $\overline{x}$, we have

$$|(\phi+\psi)(x(t))-(\phi+\psi)(\overline{x})|^{\theta}\leq C\|\nabla(\phi+\psi)(x(t))\|$$

for every $t\geq t_0$. From the second relation in (3.3) we derive for almost every $t\in[t_0,+\infty)$

$$|(\phi+\psi)(x(t))-(\phi+\psi)(\overline{x})|^{\theta}\leq C\lambda\|\dot{x}(t)\|+C\|\dot{v}(t)\|,$$

which combined with (4.6) yields

$$\alpha\sigma(t)\leq\big(C\lambda\|\dot{x}(t)\|+C\|\dot{v}(t)\|\big)^{\frac{1-\theta}{\theta}}\leq(C\max(\lambda,1))^{\frac{1-\theta}{\theta}}\cdot(\|\dot{x}(t)\|+\|\dot{v}(t)\|)^{\frac{1-\theta}{\theta}}. \tag{4.7}$$

Since

$$\dot{\sigma}(t)=-\|\dot{x}(t)\|-\|\dot{v}(t)\|, \tag{4.8}$$

we conclude that there exists $\alpha'>0$ such that for almost every $t\in[t_0,+\infty)$

$$\dot{\sigma}(t)\leq-\alpha'\big(\sigma(t)\big)^{\frac{\theta}{1-\theta}}. \tag{4.9}$$

If $\theta=\frac{1}{2}$, then

$$\dot{\sigma}(t)\leq-\alpha'\sigma(t)$$

for almost every $t\in[t_0,+\infty)$. By multiplying with $\exp(\alpha' t)$ and integrating afterwards from t_0 to t, it follows that there exist $a_1,b_1>0$ such that

$$\sigma(t)\leq a_1\exp(-b_1 t) \ \forall t\geq t_0$$

and the conclusion of (b) is immediate from (4.4).

Assume that $0 < \theta < \frac{1}{2}$. We obtain from (4.9)

$$\frac{d}{dt}\left(\sigma(t)^{\frac{1-2\theta}{1-\theta}}\right) \leq -\alpha' \frac{1-2\theta}{1-\theta}$$

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

By integration we obtain

$$\sigma(t)^{\frac{1-2\theta}{1-\theta}} \leq -\overline{\alpha}t + \overline{\beta} \ \forall t \geq t_0,$$

where $\overline{\alpha} > 0$. Thus there exists $T \geq 0$ such that

$$\sigma(T) \leq 0 \ \forall t \geq T,$$

which implies that x and y are constant on $[T, +\infty)$.

Finally, suppose that $\frac{1}{2} < \theta < 1$. We obtain from (4.9)

$$\frac{d}{dt}\left(\sigma(t)^{\frac{1-2\theta}{1-\theta}}\right) \geq \alpha' \frac{2\theta-1}{1-\theta}$$

for almost every $t \in [t_0, +\infty)$. By integration we derive

$$\sigma(t) \leq (a_2 t + b_2)^{-\left(\frac{1-\theta}{2\theta-1}\right)} \ \forall t \geq t_0,$$

where $a_2, b_2 > 0$. Statement (c) follows from (4.4). □

References

[1] B. Abbas, *An asymptotic viscosity selection result for the regularized Newton dynamic*, arXiv:1504.07793v1, 2015.

[2] B. Abbas, H. Attouch, B.F. Svaiter, *Newton-like dynamics and forward-backward methods for structured monotone inclusions in Hilbert spaces*, Journal of Optimization Theory and its Applications **161(2)** (2014), 331–360.

[3] F. Alvarez, H. Attouch, J. Bolte, P. Redont, *A second-order gradient-like dissipative dynamical system with Hessian-driven damping. Application to optimization and mechanics*, Journal de Mathématiques Pures et Appliquées **81(8)** (2002), 747–779.

[4] H. Attouch, G. Buttazzo, G. Michaille, *Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization, Second Edition*, MOS-SIAM Series on Optimization, Philadelphia, 2014.

[5] H. Attouch, J. Bolte, *On the convergence of the proximal algorithm for nonsmooth functions involving analytic features*, Mathematical Programming **116(1-2)** (2009) 5–16.

[6] H. Attouch, J. Bolte, P. Redont, A. Soubeyran, *Proximal alternating minimization and projection methods for nonconvex problems: an approach based on the Kurdyka-Łojasiewicz inequality*, Mathematics of Operations Research **35(2)** (2010), 438–457.

[7] H. Attouch, J. Bolte, B.F. Svaiter, *Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods*, Mathematical Programming **137(1-2)** (2013), 91–129.

[8] H. Attouch, M.-O. Czarnecki, *Asymptotic behavior of coupled dynamical systems with multiscale aspects*, Journal of Differential Equations **248(6)** (2010), 1315–1344.

[9] H. Attouch, P.-E. Maingé, P. Redont, *A second-order differential system with Hessian-driven damping; application to non-elastic shock laws*, Differential Equations and Applications **4(1)** (2012), 27–65.

[10] H. Attouch, B.F. Svaiter, *A continuous dynamical Newton-like approach to solving monotone inclusions*, SIAM Journal on Control and Optimization **49(2)** (2011), 574–598.

[11] H.H. Bauschke, P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics, Springer, New York, 2011.

[12] J. Bolte, A. Daniilidis, A. Lewis, *The Lojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems*, SIAM Journal on Optimization **17(4)** (2006), 1205–1223.

[13] J. Bolte, A. Daniilidis, A. Lewis, M. Shiota, *Clarke subgradients of stratifiable functions*, SIAM Journal on Optimization **18(2)** (2007), 556–572.

[14] J. Bolte, A. Daniilidis, O. Ley, L. Mazet, *Characterizations of Lojasiewicz inequalities: subgradient flows, talweg, convexity*, Transactions of the American Mathematical Society **362(6)** (2010), 3319–3363.

[15] J. Bolte, T.P. Nguyen, J. Peypouquet, B.W. Suter, *From error bounds to the complexity of first-order descent methods for convex functions*, Mathematical Programming **165(2)** (2017), 471–507.

[16] J. Bolte, S. Sabach, M. Teboulle, *Proximal alternating linearized minimization for nonconvex and nonsmooth problems*, Mathematical Programming **146(1–2)** (2014), 459–494.

[17] J.M. Borwein, Q.J. Zhu, *Techniques of Variational Analysis*, Springer, New York, 2005.

[18] R.I. Boţ, E.R. Csetnek, S. László, *An inertial forward-backward algorithm for the minimization of the sum of two nonconvex functions*, EURO Journal on Computational Optimization **4** (2016), 3–25.

[19] R.I. Boţ, E.R. Csetnek, *Approaching nonsmooth nonconvex optimization problems through first order dynamical systems with hidden acceleration and Hessian driven damping terms*, Set-Valued and Variational Analysis **26(2)** (2018), 227–245.

[20] R.I. Boţ, E.R. Csetnek, *A forward-backward dynamical approach to the minimization of the sum of a nonsmooth convex with a smooth nonconvex function*, ESAIM: Control, Optimisation and Calculus of Variations **24(2)** (2018), 463–477.

[21] R.I. Boţ, E.R. Csetnek, *Levenberg-Marquardt dynamics associated to variational inequalities*, Set-Valued and Variational Analysis **25** (2017), 569–589.

[22] H. Brézis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland Mathematics Studies No. 5, Notas de Matemática (50), North-Holland/Elsevier, New York, 1973.

[23] E. Chouzenoux, J.-C. Pesquet, A. Repetti, *Variable metric forward-backward algorithm for minimizing the sum of a differentiable function and a convex function*, Journal of Optimization Theory and its Applications **162(1)** (2014), 107–132.

[24] P. Frankel, G. Garrigos, J. Peypouquet, *Splitting methods with variable metric for Kurdyka-Lojasiewicz functions and general convergence rates*, Journal of Optimization Theory and its Applications **165(3)** (2015), 874–900.

[25] A. Haraux, *Systèmes Dynamiques Dissipatifs et Applications*, Recherches en Mathématiques Appliquées 17, Masson, Paris, 1991.

[26] A. Haraux, M. Jendoubi, *Convergence of solutions of second-order gradient-like systems with analytic nonlinearities*, Journal of Differential Equations **144(2)** (1998), 313–320.

[27] R. Hesse, D.R. Luke, S. Sabach, M.K. Tam, *Proximal heterogeneous block input-output method and application to blind ptychographic diffraction imaging*, SIAM Journal on Imaging Sciences **8(1)** (2015), 426–457.

[28] K. Kurdyka, *On gradients of functions definable in o-minimal structures*, Annales de l'institut Fourier (Grenoble) **48(3)** (1998), 769–783.

[29] G. Li, T.K. Pong, *Calculus of the exponent of Kurdyka-Łojasiewicz inequality and its applications to linear convergence of first-order methods*, Foundations of Computational Mathematics, DOI https://doi.org/10.1007/s10208-017-9366-8

[30] S. Łojasiewicz, *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.

[31] B. Mordukhovich, *Variational Analysis and Generalized Differentiation, I: Basic Theory, II: Applications*, Springer-Verlag, Berlin, 2006.

[32] P. Ochs, Y. Chen, T. Brox, T. Pock, *iPiano: Inertial proximal algorithm for nonconvex optimization*, SIAM Journal on Imaging Sciences **7(2)** (2014), 1388–1419.

[33] R.T. Rockafellar, R.J.-B. Wets, *Variational Analysis*, Fundamental Principles of Mathematical Sciences 317, Springer-Verlag, Berlin, 1998.

[34] L. Simon, *Asymptotics for a class of nonlinear evolution equations, with applications to geometric problems*, Annals of Mathematics **118(2)** (1983), 525–571.

[35] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific, Singapore, 2002.

Acknowledgment

The authors are thankful to an anonymous referee for carefully reading the paper and suggestions.

Radu Ioan Boţ
University of Vienna
Faculty of Mathematics
Oskar-Morgenstern-Platz 1
A-1090 Vienna
Austria
e-mail: `radu.bot@univie.ac.at`

Ernö Robert Csetnek
University of Vienna
Faculty of Mathematics
Oskar-Morgenstern-Platz 1
A-1090 Vienna
Austria
e-mail: ernoe.robert.csetnek@univie.ac.at