On damped second-order gradient systems

Pascal Bégout*, Jérôme Bolte† and Mohamed Ali Jendoubi‡

Abstract

Using small deformations of the total energy, as introduced in [31], we establish that damped second order gradient systems

$$u''(t) + \gamma u'(t) + \nabla G(u(t)) = 0,$$

may be viewed as quasi-gradient systems. In order to study the asymptotic behavior of these systems, we prove that any (nontrivial) desingularizing function appearing in KL inequality satisfies $\varphi(s) \geqslant c\sqrt{s}$ whenever the original function is definable and C^2. Variants to this result are given. These facts are used in turn to prove that a desingularizing function of the potential G also desingularizes the total energy and its deformed versions. Our approach brings forward several results interesting for their own sake: we provide an asymptotic alternative for quasi-gradient systems, either a trajectory converges, or its norm tends to infinity. The convergence rates are also analyzed by an original method based on a one-dimensional worst-case gradient system.

We conclude by establishing the convergence of solutions of damped second order systems in various cases including the definable case. The real-analytic case is recovered and some results concerning convex functions are also derived.

Contents

*TSE (Institut de Mathématiques de Toulouse, Université Toulouse I Capitole), Manufacture des Tabacs, 21 allée de Brienne, 31015 Toulouse, Cedex 06, France

†TSE (GREMAQ, Université Toulouse I Capitole), Manufacture des Tabacs, 21 allée de Brienne, 31015 Toulouse, Cedex 06, France.
Effort sponsored by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant number FA9550-14-1-0056. This research also benefited from the support of the "FMJH Program Gaspard Monge in optimization and operations research"
‡Université de Carthage, Institut préparatoire aux études scientifiques et techniques, BP 51, 2080 La Marsa, Tunisia
E-mail: Pascal.Begout@tse-fr.eu(*), Jerome.Bolte@tse-fr.eu(†), ma.jendoubi@fsb.rnu.tn(‡)
2010 Mathematics Subject Classification: 35B40, 34D05, 37N40
Key Words: dissipative dynamical systems, gradient systems, inertial systems, Kurdyka-Łojasiewicz inequality, global convergence

1 Introduction

1.1 A global view on previous results

In this paper, we develop some new tools for the asymptotic behavior as t goes to infinity of solutions $u : \mathbb{R}_+ \longrightarrow \mathbb{R}^N$ of the following second order system

$$u''(t) + \gamma u'(t) + \nabla G(u(t)) = 0, \quad t \in \mathbb{R}_+. \tag{1.1}$$

Here, $\gamma > 0$ is a positive real number which can be seen as a *damping coefficient*, $N \geqslant 1$ is an integer and $G \in C^2(\mathbb{R}^N)$ is a real-valued function. In Mechanics, (1.1) models, among other problems, the motion of an object subject to a force deriving from a potential G (e.g. gravity) and to a viscous friction force $-\gamma u'$. In particular, the above may be seen as a qualitative model for the motion of a material point subject to gravity, constrained to evolve on the graph of G and subject to a damping force, further insights and results on this view may be found in [3, 14]. This type of dynamical system has been the subject of several works in various fields and along different perspectives, one can quote for instance [4] for Nonsmooth Mechanics, [12, 11] for recent advances in Optimization and [46] for pioneer works on the topic, partial differential equations and related aspects [34, 41, 5].

The aim of this work is to provide a deeper understanding of the asymptotic behavior of such a system and of the mechanisms behind the stabilization of trajectories at infinity (making each bounded orbit approach some specific critical point). Such behaviors have been widely investigated for gradient systems,

$$u'(t) + \nabla G(u(t)) = 0,$$

for a long time now. The first decisive steps were made by Łojasiewicz for analytic functions through the introduction of the so-called gradient inequality [44, 43]. Many other works followed among which two important contributions: [13] for convex functions and [42] for definable functions. Surprisingly the asymptotic behavior of the companion dynamics (1.1) has only been "recently" analyzed. The motivation for studying (1.1) seems to come from three distinct fields PDEs, Mechanics and Optimization. Out of the convex realm [45, 1], the seminal paper is probably [31]. Like many of the works on gradient systems the main assumption, borrowed from Łojasiewicz original contributions, is the analyticity of the function – or more precisely the fact that the function satisfies the Łojasiewicz inequality. This work paved the way for many developments: convergence rates studies [33], extension to partial differential equations [47, 39, 38, 32, 37, 34, 19, 27, 26, 35, 5], use of various kind of dampings [17, 18] (see also [16, 36, 29, 40]). Despite the huge amount of subsequent works, some deep questions remained somehow unanswered; in particular it is not clear to see:

- *What are the exact connections between gradient systems and damped second-order gradient systems?*
- *Within these relationships, how central is the role of the properties/geometry of the potential function G?*

Before trying to provide some answers, we recall some fundamental notions related to these questions; they will also constitute the main ingredients in our analysis of (1.1).

Quasi-gradient fields. The notion is natural and simple: a vector field V is called *quasi-gradient* for a function L if it has the same singular point (as ∇L) and if the angle α between the field V and the gradient ∇L remains acute and bounded away from $\pi/2$. Proper definitions are recalled in Section 3.1. Of course, such systems have a behavior which is very similar to those of gradient systems (see Theorem 3.2). We refer to [6] and the references therein for further geometrical insights on the topic.

Liapunov functions for damped second order gradient systems. The most striking common point between (1.1) and gradient systems is that of a "natural" Liapunov function. In our case, it is given by the *total energy*, sum of the potential energy and the kinetic energy,

$$E_T(u, v) = G(u) + \frac{1}{2}\|v\|^2.$$

The above is a Liapunov function in the phase space, more concretely

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} E_T(u(t), u'(t)) &= \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}|u'(t)|^2 + G(u(t))\Big) \\ &= -\gamma\|u'(t)\|^2. \end{aligned}$$

Contrary to what happens for classical gradient systems the vector field associated with (1.1) is not strictly Lyapunov for E_T: it obviously degenerates on the subspace $[v = 0]$ (or $[u' = 0]$). The use of E_T is however at the heart of most results attached to this dynamical system.

KL functions. A KL function is a function whose values can be reparametrized in the neighborhood of each of its critical point so that the resulting functions become *sharp*[1]. More formally, G is called KL on the slice of level lines $[0 < G < r_0] \overset{\text{def}}{=} \{u \in \mathbb{R}^N; 0 < G(u) < r_0\}$, if there exists $\varphi \in C^0\big([0, r_0)\big) \cap C^1(0, r_0)$ concave such that $\varphi(0) = 0$, $\varphi' > 0$ and

$$\|\nabla(\varphi \circ G)(u)\| \geqslant 1, \quad \forall u \in [0 < G < r_0].$$

Proper definitions and local versions can be found in the next section. The above definition originates in [10] and is based on the fundamental work of Kurdyka [42], where it was introduced in the framework of o-minimal structure[2] as a generalization of the famous Łojasiewicz inequality.

KL functions are central in the analysis of gradient systems, the readers are referred to [10] and the references therein.

Desingularizing functions. The function appearing above, namely φ, is called a *desingularizing function*: the faster φ' tends to infinity at 0, the flatter is G around critical points. As opposed to the Łojasiewicz gradient inequality, this behavior, in the o-minimal world, is not necessarily of a "power-type". Highly degenerate functions can be met, like for instance $G(u) = \exp\big(-1/p^2(u)\big)$ where $p : \mathbb{R}^N \longrightarrow \mathbb{R}$ is any real polynomial function. This class of functions belongs to the log-exp structure, an o-minimal class that contains semi-algebraic sets and the graph of the exponential function [48]. Finally, observe that if it is obvious that φ might have an arbitrarily brutal behavior at 0, it is also pretty clear that the smoothness of G is related to a lower-control of the behavior of φ, for instance we must have $\varphi'(0) = \infty$ – which is not the case in general in the nonsmooth world (see e.g. [9]).

1.2 Main results

Several auxiliary theorems were necessary to establish our main result, we believe they are interesting for their own sake. Here they are:

– An asymptotic alternative for quasi-gradient systems: either a trajectory converges or it escapes to infinity,

– A general convergence rate result for the solutions of the gradient systems that brings forward a worst-case gradient dynamical system in dimension one,

– Lower bounds for desingularizing functions of C^2 KL functions.

We are now in position to describe the strategy we followed in that paper for the asymptotic study of the damped second order gradient system (1.1). Our method was naturally inspired by the Liapunov function provided in [31].

1. First we show that E_T can be slightly and "semi-algebraically" (respectively, definably) deformed into a smooth function E_T^{def}, so that the gradient of the new energy ∇E_T^{def} makes an uniformly acute angle with the vector field associated with (1.1) – this property only holds on bounded sets of the phase space. The system (1.1) appears therefore as a quasi-gradient system for E_T^{def}.

2. In a second step we establish/verify that the solutions of the quasi-gradient systems converge whenever they originate from a KL function.

   We also provide rates of convergence and we explain how they may be naturally and systematically derived from a one-dimensional worst-case gradient dynamics.

   At this stage it is possible to proceed abstractly to the proof of the convergence of solutions to (1.1) in several cases. For instance the definable case: we simply have to use the fact that E_T^{def} is definable whenever G is, so it is a KL function and the conclusion follows.

   Although direct and fast, this approach has an important drawback from a conceptual viewpoint since it relies on a desingularizing function attached to an auxiliary function E_T^{def} whose meaning is unclear. Whatever perspectives we may adopt (Mechanics, Optimization, PDEs), an important question is indeed to *understand what happens when G is KL and how the desingularizing function of G actually impacts the convergence of solutions to* (1.1).

[1] That is, the norms of its gradient remain bounded away from zero.

[2] A far reaching concept that generalizes semi-algebraic or (globally) subanalytic classes of sets and functions.

3. We answer to this question in the following way.
   (a) We prove that desingularizing functions of C^2 definable functions have a lower bound. Roughly speaking, we prove that for nontrivial critical points the desingularizing function has the property $\varphi(s) \geqslant c\sqrt{s}$ $\left(\text{or equivalently}^{(3)}\ \varphi'(s) \geqslant \frac{c'}{\sqrt{s}}\right)$.
   (b) We establish that if φ is definable and desingularizing for G at $\overline{u}$ then it is desingularizing for both E_T and E_T^{def} at $(\overline{u}, 0)$.
4. We conclude by combining previous results to obtain in particular the convergence of solutions to (1.1) under definability assumptions. We also provide convergence rates that depend on the desingularizing function of G, *i.e. on the geometry of the potential.*

We would like to point out and emphasize two facts that we think are of interest. First the property $\varphi(s) \geqslant c\sqrt{s}$ (see Lemma 2.9 below) is a new result and despite its "intuitive" aspect the proof is nontrivial. We believe it has an interest in its own sake.

More related to our work is the fact that (in the definable case and in many other relevant cases) our results show that the desingularizing function of G is conditioning the asymptotic behavior of solutions of the system. Within an Optimization perspective this means that the "complexity", or at least the convergence rate, of the dynamical system is entirely embodied in G when G is smooth. From a mechanical viewpoint, stabilization at infinity is determined by the conditioning of G provided the latter is smooth enough; in other words the intuition that for large time behaviors, the potential has a predominant effect on the system is correct – a fact which is of course related to the dissipation of the kinetic energy at a "constant rate".

Notation. The finite-dimensional space $\mathbb{R}^N$ $(N \geqslant 1)$ is endowed with the canonical scalar product $\langle\, .\,,\, .\,\rangle$ whose norm is denoted by $\|\,.\,\|$. The product space $\mathbb{R}^N \times \mathbb{R}^N$ is endowed with the natural product metric which we still denote by $\langle\, .\,,\, .\,\rangle$. We also define for any $\overline{u} \in \mathbb{R}^N$ and $r > 0$, $B(\overline{u}, r) = \{u \in \mathbb{R}^N; \|u - \overline{u}\| < r\}$. When S is a subset of $\mathbb{R}^N$ its interior is denoted by $\operatorname{int} S$ and its closure by $\overline{S}$. If $F : \mathbb{R}^N \longrightarrow \mathbb{R}$ is a differentiable function, its gradient is denoted by ∇F. When F is a twice differentiable function, its Hessian is denoted by $\nabla^2 F$. The set of critical points of F is defined by

$$\operatorname{crit} F = \Big\{u \in \mathbb{R}^N; \nabla F(u) = 0\Big\}.$$

This paper is organized as follows. In Section 2, we provide a lower bound for desingularizing function of C^2 functions under various assumptions, like definability (Proposition 2.8 and Lemma 2.9). In Section 3, we recall the behavior of a first order system having a quasi-gradient structure for some KL function and we provide an asymptotic alternative (Theorem 3.2). In Theorem 3.7, the convergence rate of any solution to a first order system having a quasi-gradient structure is proved to be better than that of a one-dimensional worst-case gradient dynamics (various known results are recovered in a transparent way). Finally, we establish that any function which desingularizes G in (1.1) also desingularizes the total energy and various relevant deformation of the latter (Proposition 3.11). In Section 4, we study the asymptotic behavior of solutions to (1.1) (Theorem 4.1) while in Section 5, we describe several consequences of our main results. Appendix provides, for the comfort of the reader, some elementary facts on o-minimal structures.

2 Structural results: lower bounds for desingularizing functions of C^2 functions

To keep the reading smooth and easy, we will not formally define here o-minimal structure. The definition is postponed in Appendix. Let us however recall, at this stage, that the simplest o-minimal structure (containing the graph of the real product) is given by the class of real semi-algebraic sets and functions. A semi-algebraic set is the finite union of sets of the form

$$\Big\{u \in \mathbb{R}^N;\ p(u) = 0,\ p_i(u) < 0, \forall i \in I\Big\}, \tag{2.1}$$

[3] Recall that φ is definable.

where I is a finite set and $p, \{p_i\}_{i\in I}$ are real polynomial functions.

Let us recall a fundamental concept for dissipative dynamical systems of gradient type.

Definition 2.1 (**Kurdyka-Łojasiewicz property and desingularizing function**).
Let $G : \mathbb{R}^N \longrightarrow \mathbb{R}$ be a differentiable function.

(i) We shall say that G has the *KL property* at $\overline{u} \in \mathbb{R}^N$ if there exist $r_0 > 0$, $\eta > 0$ and $\varphi \in C([0, r_0); \mathbb{R}_+)$ such that

1. $\varphi(0) = 0$, $\varphi \in C^1((0, r_0); \mathbb{R}_+)$ concave and φ' positive on $(0, r_0)$,
2. $u \in B(\overline{u}, \eta) \implies |G(u) - G(\overline{u})| < r_0$; and for each $u \in B\big(\overline{u}, \eta\big)$, such that $G(u) \neq G(\overline{u})$,

$$\big\|\nabla(\varphi \circ |G(\,.\,) - G(\overline{u})|)(u)\big\| \geqslant 1. \tag{2.2}$$

Such a function φ is called a *desingularizing function* of G at $\overline{u}$ on $B(\overline{u}, \eta)$.

(ii) The function G is called a *KL function* if it has the KL property at each of its points.

The following result is due to Łojasiewicz in its real-analytic version (see e.g. [43, 44]), it was generalized to o-minimal structures and considerably simplified by Kurdyka in [42] (see Appendix).

Theorem 2.2 (**Kurdyka-Łojasiewicz inequality** [42][(4)]). *Let $\mathcal{O}$ be an o-minimal structure and let $G \in C^1(\mathbb{R}^N; \mathbb{R})$ be a definable function. Then G is a KL function.*

Remark 2.3. (a) Theorem 2.2 is of course trivial when $\overline{u} \notin \operatorname{crit} G$ – take indeed, $\varphi(s) = cs$ where $c = \frac{1+\varepsilon}{\|\nabla G(\overline{u})\|}$ and $\varepsilon > 0$.
(b) Restrictions of real-analytic functions to compact sets included in their (open) domain belong to the o-minimal structure of globally analytic sets [25]. They are therefore KL functions (see indeed Example A.2). In some o-minimal structures there are nontrivial functions for which all derivatives vanish on some nonempty set, like $G(u) = \exp(-1/f^2(u))$ where $f \neq 0$ is any smooth semi-algebraic function achieving the value 0[(5)] (see also Example A.2). For these cases, φ is not of power-type – as it is the case when G is semi-algebraic or real-analytic. Other types of functions satisfying the KL property in various contexts are provided in [2] (see also Corollary 5.5).
(c) Desingularizing functions of definable functions can be chosen to be definable, strictly concave and C^k (where k is arbitrary).

The following trivial notion is quite convenient.

Definition 2.4 (**Trivial critical points**). A critical point u of a differentiable function $G : \mathbb{R}^N \longrightarrow \mathbb{R}$ is called *trivial* if $u \in \operatorname{int} \operatorname{crit} G$. It is *nontrivial* otherwise. Observe that u is nontrivial if, and only if, there exists $u_n \xrightarrow{n\to\infty} u$ such that $G(u_n) \neq G(u)$, for any $n \in \mathbb{N}$.

When $\overline{u}$ is a trivial critical point of G, any concave function $\varphi \in C^0\big([0, r_0)\big) \cap C^1(0, r_0)$ such that $\varphi' > 0$ and $\varphi(0) = 0$ is desingularizing at $\overline{u}$.

An immediate consequence of the KL inequality is a local and strong version of Sard's theorem.

Remark 2.5 (**Local finiteness of critical values**). Let $G \in C^1(\mathbb{R}^N; \mathbb{R})$ and $\overline{u} \in \mathbb{R}^N$. Assume that G satisfies the KL property at $\overline{u}$ on $B(\overline{u}, \eta)$. Then

$$u \in B(\overline{u}, \eta) \text{ and } \nabla G(u) = 0 \implies G(u) = G(\overline{u}).$$

The simplest functions we can think of with respect to the behavior of the solutions to (1.1) are given by functions with linear gradients, that is quadratic forms

$$G(u) = \frac{1}{2}\langle Au, u\rangle,\ u \in \mathbb{R}^N,\ \text{where } A \in \mathscr{M}_N(\mathbb{R}),\ A^T = A.$$

When $A \neq 0$, it is easy to establish directly that $\varphi(s) = \sqrt{\frac{1}{|\lambda|}s}$ (where λ is a nonzero eigenvalue with smallest absolute value) provides a desingularizing function. In the subsections to come, we show that the best we can hope in general for a desingularizing function φ attached to a C^2 function G is precisely a quantitative behavior of square-root type.

[4] See comments in Appendix.
[5] This function is definable in the log-exp structure of Wilkie [48].

2.1 Lower bounds for desingularizing functions of potentials having a simple critical point structure

Our first assumption, formally stated below, asserts that points having critical value must be critical points. The assumption is rather strong in general but it will be complemented in the next section by a far more general result for definable functions.

$$\begin{cases} \text{Let } \overline{u} \in \operatorname{crit} G. \\ \text{There exists } \eta > 0 \text{ such that for any } u \in B(\overline{u}, \eta), \\ \big(G(u) = G(\overline{u}) \implies u \in \operatorname{crit} G\big). \end{cases} \tag{2.3}$$

Example 2.6. (a) When $N = 1$ and $G \in C^1$ is KL then assumption (2.3) holds.
[If the result does not hold then there exists a sequence $(x_n)_{n\in\mathbb{N}}$ such that $x_n \xrightarrow{n\longrightarrow\infty} \overline{u}$ and

$$G(x_n) = G(\overline{u}), \tag{2.4}$$
$$G'(x_n) \neq 0, \tag{2.5}$$

for any $n \in \mathbb{N}$. Without loss of generality, we may assume that $(x_n)_{n\in\mathbb{N}}$ is monotone, say decreasing. From (2.4)–(2.5) and Rolle's Theorem, there exists a sequence $(u_n)_{n\in\mathbb{N}}$ such that $x_{n+1} < u_n < x_n$, $G'(u_n) = 0$, $G(u_n) \neq G(\overline{u})$, for any $n \in \mathbb{N}$. Thus $G(u_n)$ are critical values distinct from $G(\overline{u})$ such that $G(u_n) \longrightarrow G(\overline{u})$; this contradicts the local finiteness of critical values – see Remark 2.5.]
(b) Of course, the result in (a) cannot be extended to higher dimensions. Consider for instance

$$G : \mathbb{R}^2 \longrightarrow \mathbb{R}, \quad G(u_1, u_2) = u_1^2 - u_2^2,$$

which is obviously KL. One has $\nabla G(u) = 0$ if, and only if, $u = 0$, yet $G(t, -t) = 0$ for any t in $\mathbb{R}$.
(c) If G is convex, (2.3) holds globally, *i.e.*, with $\eta = \infty$. [This follows directly from the well-known fact that $G(u) = \min G$ if, and only if, $\nabla G(u) = 0$.]

Lemma 2.7 (Comparing values growth with gradients growth).
Let $G \in C^{1,1}_{\mathrm{loc}}(\mathbb{R}^N; \mathbb{R})$ and $\overline{u} \in \operatorname{crit} G$. Assume there exists $\varepsilon > 0$ such that

$$u \in B(\overline{u}, 2\varepsilon) \text{ and } G(u) = G(\overline{u}) \implies u \in \operatorname{crit} G,$$

in other words (2.3) holds (with $\eta = 2\varepsilon$). Then there exists $c > 0$ such that

$$|G(u) - G(\overline{u})| \geqslant c\|\nabla G(u)\|^2, \tag{2.6}$$

for any $u \in B(\overline{u}, \varepsilon)$.

Proof. Working if necessary with $\widetilde{G}(u) = G(u) - G(\overline{u})$, we may assume, without loss of generality, that $G(\overline{u}) = 0$. Let us proceed in two steps.
Step 1. Let $H \in C^{1,1}\big(\overline{B}(\overline{u}, 2\varepsilon); \mathbb{R}\big)$ with $\overline{u} \in \operatorname{crit} H$ and assume further that $H \geqslant 0$. We claim that there exists $c > 0$ such that

$$\forall u \in B(\overline{u}, \varepsilon),\ H(u) \geqslant c\|\nabla H(u)\|^2. \tag{2.7}$$

Denote by L_2 the Lipschitz constant of ∇H on $\overline{B}(\overline{u}, 2\varepsilon)$, let $L_1 = \max\limits_{u\in\overline{B}(\overline{u},2\varepsilon)} \|\nabla H(u)\|$ and set $L = L_1 + L_2$. Since,

$$\big(L_1 = 0 \text{ or } L_2 = 0\big) \implies \nabla H_{|B(\overline{u},\varepsilon)} \equiv 0 \implies (2.7),$$

we may assume that $L_2 > 0$ and $L_1 > 0$. Let $u \in B(\overline{u}, \varepsilon)$. We have for any $v \in B(0, 2\varepsilon)$,

$$H(v) - H(u) = \int_0^1 \langle \nabla H\big((1-t)u + tv\big), v - u\rangle \mathrm{d}t$$
$$= \int_0^1 \langle \nabla H\big((1-t)u + tv\big) - \nabla H(u), v - u\rangle \mathrm{d}t + \langle \nabla H(u), v - u\rangle,$$

so that for any $v \in B(0,2\varepsilon)$,

$$\Big|H(v) - H(u) - \langle \nabla H(u), v-u\rangle\Big| \leqslant \frac{L_2}{2}\|v-u\|^2. \tag{2.8}$$

Note that $\big\|\big(u - \frac{\varepsilon}{L}\nabla H(u)\big) - \overline{u}\big\| \leqslant \|u-\overline{u}\| + \frac{\varepsilon}{L}\|\nabla H(u)\| < \varepsilon + \varepsilon\frac{L_1}{L} < 2\varepsilon$. By convexity, we infer that $\big[u, u - \frac{\varepsilon}{L}\nabla H(u)\big] \subset B(\overline{u}, 2\varepsilon)$. It follows that $v = u - \frac{\varepsilon}{L}\nabla H(u)$ is an admissible choice in (2.8). Without loss of generality, we may assume that $\varepsilon \leqslant 1$. This leads to

$$0 \leqslant H(v) \leqslant H(u) - \frac{\varepsilon}{2L}\|\nabla H(u)\|^2.$$

Whence the claim.

Step 2. Define for any $u \in \overline{B}(\overline{u}, 2\varepsilon)$, $H(u) = |G(u)|$. Since $\big(G(u) = 0 \implies \nabla G(u) = 0\big)$, we easily deduce that $H \in C^1_{\rm b}\big(B(\overline{u}, 2\varepsilon); \mathbb{R}\big)$ and for any $u \in B(\overline{u}, 2\varepsilon)$, $\nabla H(u) = \operatorname{sign}\big(G(u)\big)\nabla G(u)$. Denote by L_2 the Lipschitz constant of ∇G on $\overline{B}(\overline{u}, 2\varepsilon)$. We claim that,

$$\|\nabla H(u) - \nabla H(v)\| \leqslant L_2\|u-v\|, \tag{2.9}$$

for any $(u,v) \in B(\overline{u}, 2\varepsilon) \times B(\overline{u}, 2\varepsilon)$. Let $(u,v) \in B(\overline{u}, 2\varepsilon) \times B(\overline{u}, 2\varepsilon)$. Estimate (2.9) being clear if $G(u)G(v) \geqslant 0$, we may assume that $G(u)G(v) < 0$. By the Mean Value Theorem and the assumptions on G, it follows that there exists $t \in (0,1)$ such that for $w = (1-t)u + tv$, $G(w) = 0$ and $\nabla G(w) = 0$. We then infer,

$$\begin{aligned}
&\|\nabla H(u) - \nabla H(v)\| = \|\nabla G(u) + \nabla G(v)\| \leqslant \|\nabla G(u)\| + \|\nabla G(v)\| \\
&= \|\nabla G(u) - \nabla G(w)\| + \|\nabla G(w) - \nabla G(v)\| \\
&\leqslant L_2\|u-w\| + L_2\|w-v\| = L_2\|u-v\|.
\end{aligned}$$

Hence (2.9). It follows that $H \in C^{1,1}\big(\overline{B}(\overline{u}, 2\varepsilon); \mathbb{R}\big)$ and H satisfies the assumptions of Step 1. Applying (2.7) to H, we get (2.6). This concludes the proof. □

Proposition 2.8 (**Lower bound for desingularizing functions**). *Let $G \in C^{1,1}_{\rm loc}(\mathbb{R}^N; \mathbb{R})$ and let $\overline{u}$ be a nontrivial critical point, i.e. $\overline{u} \in \operatorname{crit} G \setminus \operatorname{int}\operatorname{crit} G$. Assume that G satisfies the KL property at $\overline{u}$ and that assumption (2.3) holds at $\overline{u}$.*
Then there exists $\beta > 0$ such that for any desingularizing function φ of G at $\overline{u}$,

$$\varphi'(s) \geqslant \frac{\beta}{\sqrt{s}}, \tag{2.10}$$

for any small positive s.

Proof. We may assume $G(\overline{u}) = 0$. Combining (2.2) and (2.6), we deduce that $\varphi'(|G(u)|) \geqslant \frac{1}{\|\nabla G(u)\|} \geqslant \frac{\beta}{\sqrt{|G(u)|}}$, for any $u \in B(\overline{u}, \varepsilon)$ such that $G(u) \neq G(\overline{u})$ (Remark 2.5). Changing G into $-G$ if necessary, there is no loss of generality to assume that there exists u_n such that $u_n \longrightarrow \overline{u}$ with $G(u_n) > 0$ (recall $\overline{u}$ is a nontrivial critical point). Since G is continuous, this implies by a connectedness argument that for some ρ there exists $r > 0$ such that $\big|G\big(B(\overline{u}, \rho)\big)\big| \supset (0, r)$. Using the parametrization $s \in (0, r)$ we conclude that $\varphi'(s) \geqslant \frac{\beta}{\sqrt{s}}$, for any s sufficiently small. □

2.2 Lower bounds for desingularizing functions of definable C^2 functions

This part makes a strong use of definability arguments (these are recalled in the last section).

Lemma 2.9 (**Lower bounds for desingularizing functions of C^2 definable functions**). *Let $G : \Omega \longrightarrow \mathbb{R}$ be a C^2 definable function on an open subset $\Omega \ni 0$ of $\mathbb{R}^N$. We assume that 0 is a nontrivial critical point*[6] *and that $G(0) = 0$.*

[6] Equivalently, we assume that there exists $u_n \xrightarrow{n\to\infty} 0$ such that $G(u_n) \neq 0$.

Since G is definable it has the KL property[7] *that is, there exist $\eta, r_0 > 0$ and $\varphi : [0, r_0) \longrightarrow \mathbb{R}$ as in Definition 2.1 such that*

$$\|\nabla\big(\varphi \circ |G|\big)(u)\| \geqslant 1, \tag{2.11}$$

for any u in $B(0,\eta)$ such that $G(u) \neq 0$.
Then there exists $c > 0$ such that

$$\varphi'(s) \geqslant \frac{c}{\sqrt{s}}, \tag{2.12}$$

so that $\varphi(s) \geqslant 2c\sqrt{s}$, for any small $s > 0$.

Proof. Let us outline the ideas of the proof: after a simple reduction step, we show that the squared norm of a/the smallest gradient on a level line increases at most linearly with the function values. In the second step, we show that this estimate is naturally linked to the increasing rate of φ itself and to property (2.12). Let $\varphi : [0, r_0) \longrightarrow \mathbb{R}$ be any desingularizing function of G at 0 on $B(0,\eta)$, as in Definition 2.1.

Changing G in $-G$ if necessary, we may assume by Definition 2.4, without loss of generality, that there exists a sequence $(u_n)_n$ such that $u_n \xrightarrow{n\to\infty} 0$ and $G(u_n) > 0$, for any $n \in \mathbb{N}$. Let us proceed with the proof in three steps.

Step 1. We first modify the function G as follows. Let $\rho \in C^2(\mathbb{R}^N; [0,1])$ be a semi-algebraic function such that

$$\begin{cases} \operatorname{supp}\rho \subset B(0,\eta) \subset \Omega, \\ \rho(x) = 1, \text{ if } x \in B\left(0, \frac{\eta}{2}\right). \end{cases}$$

Let us define $\widehat{G}$ on $\mathbb{R}^N$ by

$$\widehat{G}(u) = \begin{cases} \rho(u)G(u) + \operatorname{dist}\left(u, B\left(0, \frac{\eta}{2}\right)\right)^3, & \text{if } u \in \Omega, \\ 0, & \text{if } u \in \mathbb{R}^N \setminus \Omega. \end{cases}$$

It follows that $\widehat{G} \in C^2(\mathbb{R}^N; \mathbb{R})$, leaves the set of desingularizing functions at 0 unchanged, has compact lower level sets and is definable in the same structure (recall Definition A.1 (iii)). Finally, we obviously have,

$$u_n \xrightarrow{n\to\infty} 0 \text{ with } \widehat{G}(u_n) > 0, \ \forall n \in \mathbb{N}. \tag{2.13}$$

Without loss of generality, we may assume that $\eta \leqslant 1$ and $r_0 \leqslant \frac{\eta^3}{8}$. Let $u \in \mathbb{R}^N \setminus B(0,\eta)$. One has,

$$\widehat{G}(u) = \operatorname{dist}\left(u, B\left(0, \frac{\eta}{2}\right)\right)^3 = \left(\|u\| - \frac{\eta}{2}\right)^3 \geqslant \frac{\eta^3}{8} \geqslant r_0.$$

It follows that,

$$\inf_{u \in B(0,\eta) \cap [\widehat{G}=r]} \|\nabla\widehat{G}(u)\| = \min_{u \in [\widehat{G}=r]} \|\nabla\widehat{G}(u)\|, \ \forall r \in (0, r_0). \tag{2.14}$$

Step 2. For $r > 0$, we introduce

$$(P_r) \qquad \psi(r) = \min\left\{\frac{1}{2}\|\nabla\widehat{G}(u)\|^2;\ u \in \mathbb{R}^N,\ \widehat{G}(u) = r\right\}.$$

Since the set of critical values of a definable function is finite and since the level sets are compact, we may choose, if necessary, r_0 so that $\psi > 0$ on $(0, r_0)$ (the fact that 0 is a nontrivial critical point excludes the case when ψ vanishes around 0). If we denote by $S(r)$ the nonempty compact set of solutions to (P_r), one easily sees that

$$S : (0, r_0) \rightrightarrows \mathbb{R}^N,$$

[7] See Theorem 2.2.

is a definable point-to-set mapping – this follows by a straightforward use of quantifier elimination (*i.e.*, by the use of Definition A.1). Using the Definable Selection Lemma (Lemma A.4), one obtains a definable curve $u : (0, r_0) \longrightarrow \mathbb{R}^N$ such that $u(r) \in S(r)$, for any $r \in (0, r_0)$. Finally, using the Monotonicity Lemma (Lemma A.3) repeatedly on the coordinates u_i of u, one can shrink r_0 so that u is actually in $C^1((0, r_0); \mathbb{R}^N)$.

Fix now r in $(0, r_0)$. Since r is noncritical the problem (P_r) is qualified and we can apply Lagrange's Theorem for constrained problems. This yields the existence of a real multiplier $\lambda(r)$ such that

$$\nabla^2\widehat{G}(u(r))\nabla\widehat{G}(u(r)) - \lambda(r)\nabla\widehat{G}(u(r)) = 0, \tag{2.15}$$

with of course $\widehat{G}(u(r)) = r$.

Note that for any $r \in (0, r_0)$, $\nabla\widehat{G}(u(r)) \neq 0$ (as seen at the beginning of this step) so that $\lambda(r)$ is an actual eigenvalue of $\nabla^2\widehat{G}(u(r))$. Since $\widehat{G}$ is C^2, the curve $\nabla^2\widehat{G}(u(r))$ is bounded in the space of matrices $\mathscr{M}_N(\mathbb{R})$. Since eigenvalues depend continuously on operators, one deduces from the previous remarks that there exists $\overline{\lambda} \geqslant 0$ such that

$$|\lambda(r)| \leqslant \overline{\lambda},\ \forall r \in (0, r_0).$$

Multiplying (2.15) by $u'(r)$ gives $\langle\nabla^2\widehat{G}(u(r))\nabla\widehat{G}(u(r)), u'(r)\rangle = \lambda(r)\langle\nabla\widehat{G}(u(r)), u'(r)\rangle$, which is nothing else than

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}r}\|\nabla\widehat{G}(u(r))\|^2 = \lambda(r)\frac{\mathrm{d}}{\mathrm{d}r}\widehat{G}(u(r)).$$

Since $\widehat{G}(u(r)) = r$, one has

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}r}\|\nabla\widehat{G}(u(r))\|^2 = \lambda(r),$$

so after integration on $[s, r] \subset (0, r_0)$, one obtains

$$\left|\|\nabla\widehat{G}(u(r))\|^2 - \|\nabla\widehat{G}(u(s))\|^2\right| = 2\left|\int_s^r \lambda(\tau)d\tau\right| \leqslant 2\overline{\lambda}|r - s| \xrightarrow{r,s\to 0} 0. \tag{2.16}$$

It follows that $\left(\|\nabla\widehat{G}(u(r))\|^2\right)_{s>0}$ is a Cauchy's family, so that the limit ℓ of $\|\nabla\widehat{G}(u(s))\|^2$ as s goes to zero exists in $[0, \infty)$. We recall that by assumption (2.13), $u_n \xrightarrow{n\to\infty} 0$, $\widehat{G}(u_n) > 0$ and $\nabla\widehat{G}(u_n) \xrightarrow{n\to\infty} 0$. Now, setting $r_n = \widehat{G}(u_n)$, one has by definition of $u(r_n)$, $\|\nabla\widehat{G}(u_n)\| \geqslant \|\nabla\widehat{G}(u(r_n))\|$. This implies that $\ell = 0$ and as a consequence (2.16) yields

$$\frac{1}{2}\|\nabla\widehat{G}(u(r))\|^2 = \int_0^r \lambda(\tau)d\tau \leqslant \overline{\lambda}r, \tag{2.17}$$

in other words

$$\psi(r) \leqslant \overline{\lambda}r,\ \ \forall r \in (0, r_0). \tag{2.18}$$

Step 3. Let us now conclude. By KL inequality one has for any $r \in (0, r_0)$,

$$\varphi'(r) \geqslant \frac{1}{\|\nabla\widehat{G}(u)\|}, \quad \forall u \in B(0, \eta) \cap [G = r]. \tag{2.19}$$

As a consequence, we can use (2.14) in (2.19) and the linear estimate (2.18) above to conclude as follows:

$$\begin{aligned}
\varphi'(r) &\geqslant \frac{1}{\inf\left\{\|\nabla\widehat{G}(u)\|;\ u \in B(0, \eta) \cap [\widehat{G} = r]\right\}} \\
&= \frac{1}{\min\left\{\|\nabla\widehat{G}(u)\|;\ u \in [\widehat{G} = r]\right\}} \\
&\geqslant \frac{1}{\sqrt{2\psi(r)}} \\
&\geqslant \frac{c}{\sqrt{r}},
\end{aligned}$$

for any $r \in (0, r_0)$, with $c = \left(\sqrt{2\overline{\lambda}}\right)^{-1}$. Hence (2.12). □

Remark 2.10. (a) Note that if $G \notin C^2$ then (2.12) does not hold. Indeed, take $G(u) = u^{\frac{3}{2}}$ and $\varphi(s) = s^{\frac{2}{3}}$ as a (semi-algebraic) counter-example.
(b) When we omit the assumption that 0 is a nontrivial critical point, *i.e.* $0 \in \operatorname{int} \operatorname{crit} G$, then G vanishes in a neighborhood of 0. In that case, the result is not true in general since any concave increasing function adequately regular is desingularizing for G. However a function $\varphi(s) = c\sqrt{s}$ can still be chosen as a desingularizing function. Hence, *for an arbitrary C^2 definable function, we can always assume that for any critical point, the corresponding desingularizing function satisfies $\varphi'(s) \geqslant c\frac{1}{\sqrt{s}}$ (locally for some positive constant c).*

3 Damped second order gradient systems

3.1 Quasi-gradient structure and KL inequalities

Definition 3.1. Let Γ be a nonempty closed subset of $\mathbb{R}^N$ and let $F : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ be a locally Lipschitz continuous mapping.

(i) We say that the first order system

$$u'(t) + F\big(u(t)\big) = 0,\ t \in \mathbb{R}_+, \tag{3.1}$$

has a *quasi-gradient structure for E on Γ*, if there exist a differentiable function $E : \mathbb{R}^N \longrightarrow \mathbb{R}$ and $\alpha_\Gamma = \alpha > 0$ such that

$$\textbf{(angle condition)} \qquad \big\langle \nabla E(u), F(u)\big\rangle \geqslant \alpha \,\|\nabla E(u)\|\,\|F(u)\|, \text{ for any } u \in \Gamma, \tag{3.2}$$

$$\textbf{(rest-points equivalence)} \qquad \operatorname{crit} E \cap \Gamma = F^{-1}(\{0\}) \cap \Gamma. \tag{3.3}$$

(ii) Equivalently a vector field F having the above properties is said to be *quasi-gradient for E on Γ*.

The following result involves classical material and ideas, yet, the fact that an asymptotic alternative can be derived in this setting does not seem to be well-known (see however [2] in a discrete context).

Theorem 3.2 (Asymptotic alternative for quasi-gradient fields). *Let $F : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ be a locally Lipschitz mapping that defines a quasi-gradient vector field for E on $\mathbb{R}^N$, for some differentiable function $E : \mathbb{R}^N \longrightarrow \mathbb{R}$. Assume further that the function E is KL. Let u be any solution to (3.1). Then,*

(i) either $\|u(t)\| \xrightarrow{t\to\infty} \infty$,

(ii) or u converges to a singular point u_∞ of F as $t \longrightarrow \infty$.

When (ii) holds then $u' \in L^1\big((0,\infty);\mathbb{R}^N\big)$ and $u'(t) \xrightarrow{t\to\infty} 0$. Moreover, we have the following estimate,

$$\|u(t) - u_\infty\| \leqslant \frac{1}{\alpha}\varphi\big(E(u(t)) - E(u_\infty)\big), \tag{3.4}$$

where φ is a desingularizing function of E at u_∞ and α is the constant in (3.2).

Proof. We assume that (i) does not hold, so there exist $u_\infty \in \mathbb{R}^N$ and a sequence $s_n \nearrow \infty$ such that $u(s_n) \xrightarrow{n\to\infty} u_\infty$. Note that by continuity of E, one has $E\big(u(s_n)\big) \xrightarrow{n\to\infty} E(u_\infty)$. Observe also that from equation (3.1) and the angle condition (3.2), one has for any $t \geqslant 0$,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\big(E \circ u\big)(t) &= \big\langle \nabla E\big(u(t)\big), u'(t)\big\rangle \\
&= -\big\langle \nabla E\big(u(t)\big), F\big(u(t)\big)\big\rangle \\
&\leqslant -\alpha\|\nabla E(u(t))\|\,\|F(u(t))\|,
\end{aligned} \tag{3.5}$$

and thus the mapping $t \longmapsto E(u(t))$ is nonincreasing, which implies

$$\lim_{t\to\infty} E(u(t)) = E(u_\infty).$$

Note that if $E(u(\bar t)) = E(u_\infty)$ for some $\bar t$, one would have $\frac{\mathrm{d}}{\mathrm{d}t}\big(E\circ u\big)(t) = 0$ for any $t > \bar t$, which would in turn imply, by (3.5), that $\|\nabla E(u(t))\|\,\|F(u(t))\| = 0$, for any such t. In view of the rest point equivalence (3.3), this would mean that $F(u(t)) = 0$, hence by uniqueness of solution curves, that $u(t) = u_\infty$ for any $t \geqslant 0$. We can thus assume without loss of generality that

$$E(u(t)) > E(u_\infty),\ \forall t \geqslant 0. \tag{3.6}$$

Let $t_0 > 0$ be such that $u(t_0) \in B\left(u_\infty, \frac{\eta}{2}\right)$ and $\varphi\big(E\big(u(t_0)\big) - E(u_\infty)\big) \in \left(0, \frac{\eta\alpha}{2}\right)$, where $\alpha > 0$ is the constant in (3.2) [in view of our preliminary comments and of the continuity of E such a t_0 exists]. By continuity of u, there exists $\tau > 0$ such that for any $t \in [t_0, t_0 + \tau)$, $u(t) \in B(u_\infty, \eta)$. So we may define $T \in (t_0, \infty]$ as

$$T = \sup\Big\{t > t_0\,;\, \forall s \in [t_0, t),\ u(s) \in B(u_\infty, \eta)\Big\}.$$

By (3.5), the Kurdyka-Łojasiewicz inequality (2.2) and equation (3.1), we have for any $t \in (t_0, T)$,

$$\begin{aligned}
&-\frac{\mathrm{d}}{\mathrm{d}t}\Big(\varphi\circ\Big(E\big(u(\,.\,)\big) - E\big(u_\infty\big)\Big)\Big)(t)\\
&= -\,\varphi'\big(E\big(u(t)\big) - E(u_\infty)\big)\frac{\mathrm{d}}{\mathrm{d}t}\big(E\circ u\big)(t)\\
&\geqslant \alpha\,\varphi'\big(E\big(u(t)\big) - E(u_\infty)\big)\,\big\|\nabla E\big(u(t)\big)\big\|\,\big\|F\big(u(t)\big)\big\|\\
&= \alpha\,\big\|F\big(u(t)\big)\big\|\,\big\|\nabla\big(\varphi\circ\big(E(\,.\,) - E(u_\infty)\big)\big(u(t)\big)\big\|\\
&\geqslant \alpha\|u'(t)\|.
\end{aligned} \tag{3.7}$$

It follows from the above estimate that

$$\|u(t) - u(t_0)\| \leqslant \int_{t_0}^{t}\|u'(s)\|\mathrm{d}s \leqslant \frac{\varphi\big(E\big(u(t_0)\big) - E(u_\infty)\big)}{\alpha} < \frac{\eta}{2}, \tag{3.8}$$

for any $t \in (t_0, T)$. We claim that $T = \infty$. Indeed, otherwise $T < \infty$ and (3.8) applies with $t = T$. Hence,

$$\|u(T) - u_\infty\| \leqslant \|u(T) - u(t_0)\| + \|u(t_0) - u_\infty\| < \eta.$$

Then $u(T) \in B(u_\infty, \eta)$, which contradicts the definition of T. As a consequence the curve u' belongs to $L^1\big((t_0,\infty);\mathbb{R}^N\big)$ by (3.8) and the curve u converges to u_∞ by Cauchy's criterion. Finally since 0 must be a cluster point of u' (recall indeed $\int_0^\infty \|u'(t)\|\mathrm{d}t < \infty$ and u' is uniformly continuous by (3.1)), one must have $F(u_\infty) = 0$. The announced estimate follows readily from (3.8) and the fact that $T = \infty$. □

Corollary 3.3. *Let $F : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ be locally Lipschitz continuous and assume that for any $R > 0$ the mapping F defines a quasi-gradient vector field for some differentiable function $E_R : \mathbb{R}^N \longrightarrow \mathbb{R}$ on $\overline{B}(0, R)$. Assume further that each of the functions E_R is KL.*
Let u be any bounded solution to (3.1). Then u converges to a singular point u_∞ of F, u' is integrable and converges to 0. In particular, if we take $R \geqslant \sup\big\{\|u(t)\|;\ t \in [0,\infty)\big\}$, we have the following estimate,

$$\|u(t) - u_\infty\| \leqslant \frac{1}{\alpha_R}\varphi\Big(E_R(u(t)) - E_R(u_\infty)\Big), \tag{3.9}$$

where φ is a desingularizing function of E_R at u_∞ and α_R is the constant in (3.2), for the ball $\overline{B}(0, R)$.

Proof. Take $R \geqslant \sup\big\{\|u(t)\|;\ t \in [0,\infty)\big\}$ and observe that the previous proof may be reproduced as it is: just replace E by E_R. □

3.2 Convergence rate of quasi-gradient systems and worst-case dynamics

To simplify our presentation we consider first a proper gradient system:

$$u'(t) + \nabla E(u(t)) = 0, \tag{3.10}$$

where $E : \mathbb{R}^N \longrightarrow \mathbb{R}$ is a twice continuously differentiable KL function. We assume that u is bounded so, by virtue of our previous considerations, the curve converges to some critical point u_∞ of E. Observe that if u_∞ is a trivial critical point, one actually has $u(0) = u_\infty$ and the asymptotic study is trivial.

We thus assume u_∞ to be nontrivial, and we denote by φ a desingularizing function of E at u_∞. We set

$$\psi = \varphi^{-1},$$

whose domain is denoted by $[0, a)$, (with $a \in (0, \infty]$) and we consider the *one-dimensional worst-case gradient dynamics* (see [8]):

$$\nu'(t) + \psi'(\nu(t)) = 0, \quad \nu(0) = \nu_0 \in (0, a). \tag{3.11}$$

We shall assume that

$$\varphi'(s) \geqslant \frac{c}{\sqrt{s}}, \text{ on } (0, r_0), \tag{3.12}$$

which implies that solutions ν to (3.11) are globally defined on $[0, \infty)$ and satisfy $\lim\limits_{t \nearrow \infty} \nu(t) = 0$ with $\nu(t) \geqslant \nu_0 e^{-c_0 t}$, for any $t \geqslant 0$ (and for some $c_0 > 0$). Uniqueness holds by concavity of φ. Finally, note that if E is a C^2 definable function then φ can be chosen to be C^2, strictly concave and satisfying (3.12) (Remark 2.3 (c) and Lemma 2.9).

Radial functions and worst-case dynamics. A full justification of the terminology "worst-case dynamics" is to be given further, but at this stage one can observe that E could be taken of the form

$$E_{\text{rad}}(u) = \varphi^{-1}(\|u - u_\infty\|), \text{ with } u \in B(u_\infty, \eta) \ (\eta > 0),$$

provided that φ^{-1} is smooth enough. In that case φ is clearly desingularizing and the solutions of the gradient system (3.10) are radial in the sense that they are of the form[8]

$$u(t) = u_\infty + \nu(t) \frac{u_0 - u_\infty}{\|u_0 - u_\infty\|}, \tag{3.13}$$

where ν is a solution to (3.11). In this case, the dynamics (3.11) exactly measures the convergence rates for (3.10), since one has for any $t \geqslant 0$ and any u_0 such that $\nu(0) = \|u_0 - u_\infty\|$,

$$E_{\text{rad}}(u(t)) = \psi(\nu(t)), \tag{3.14}$$

$$\|u(t) - u_\infty\| = \nu(t). \tag{3.15}$$

We are about to see that this behavior in terms of convergence rate is actually the worst we can expect.

Remark 3.4. (a) As can be seen below, the worst-case gradient system is introduced to measure the rate of convergence of solutions for large t. Since nontrivial solutions to (3.11) have the same asymptotic behavior (they are, indeed, all of the form $\nu_1(t) = \nu(t + t_0)$ where t_0 is some real number), the choice of the initial condition $\nu(0)$ in $(0, a)$ can be made arbitrarily.
(b) The above rewrites $\nu'(t)\varphi'\big(\varphi^{-1}(\nu(t))\big) = -1$. Thus if μ denotes an antiderivative of $\varphi' \circ \varphi^{-1}$, one has $\nu(t) = \mu^{-1}(-t + a_0)$ (where a_0 is a constant), for any $t > 0$ large enough.
(c) In general, the explicit integration of such a system depends on the integrability properties of ψ and on the fact that $\varphi' \circ \varphi^{-1}$ admits an antiderivative in a closed form.
For instance if $\varphi(s) = (\frac{s}{c})^\theta$, with $c > 0$ and $\theta \in \left(0, \frac{1}{2}\right)$, then $\psi(s) = cs^{\frac{1}{\theta}}$ and

$$\nu'(t) + \frac{c}{\theta} \nu(t)^{\frac{1-\theta}{\theta}} = 0, \quad \nu(0) \in (0, a).$$

[8] Just use the formula in (3.10).

Thus by integration

$$\frac{\mathrm{d}}{\mathrm{d}t}\nu^{1-\frac{1-\theta}{\theta}}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\nu^{\frac{-1+2\theta}{\theta}}(t) = c_1,$$

with $c_1 > 0$. As a consequence,

$$\nu(t) = \big(c_2 + c_1 t\big)^{-\frac{\theta}{1-2\theta}},$$

with $c_2 > 0$. When $\theta = \frac{1}{2}$ one easily sees that $\nu(t) = \nu(0)\exp\left(-2ct\right)$.

Theorem 3.5 (**The worst-case rate and worst-case one-dimensional gradient dynamics**).
Let $E \in C^2(\mathbb{R}^N;\mathbb{R})$ be a KL function, let u be a bounded solution to (3.10) and let $u_\infty \in \operatorname{crit} E$ satisfying $u(t) \xrightarrow{t\to\infty} u_\infty$ (such a u_∞ exists by Theorem 3.2). Then for any t large enough,

$$E(u(t)) - E(u_\infty) \leqslant \psi(\nu(t)), \tag{3.16}$$

and

$$\|u(t) - u_\infty\| \leqslant \nu(t), \tag{3.17}$$

where ν is a solution to (3.11).

Proof. Without loss of generality, we may assume that $E(u_\infty) = 0$. From the previous results, we know that for any $t \geqslant t_0$, we have $u(t) \in B(u_\infty, \eta)$ and $E(u(t)) \in (0, r_0)$, so that the KL inequality gives (see Theorem 3.2 and (3.7)):

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(\varphi \circ E(u)\big)(t) \geqslant \|u'(t)\|.$$

Set $z(t) = E(u(t))$. Since $\frac{\mathrm{d}}{\mathrm{d}t}(E\circ u)(t) = -\|u'(t)\|^2$, one has $-\frac{\mathrm{d}}{\mathrm{d}t}(\varphi\circ z)(t) \geqslant \sqrt{-z'(t)}$, or equivalently

$$\varphi'\big(z(t)\big)^2 z'(t) \leqslant -1.$$

Consider now the worst-case gradient system with initial condition $\nu(t_0) = \varphi\big(E(u(t_0))\big)$ and set $z_a(t) = \psi(\nu(t)) = \varphi^{-1}(\nu(t))$, for $t \geqslant t_0$. The system (3.11) becomes $\varphi'(z_a(t))z_a'(t) + \frac{1}{\varphi'(z_a(t))} = 0$, *i.e.*, $\varphi'(z_a(t))^2 z_a'(t) = -1$. If μ is an antiderivative of φ'^2 on $(0, r_0)$, it is an increasing function and one has

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mu\circ z)(t) = \varphi'\big(z(t)\big)^2 z'(t) \leqslant -1 = \varphi'\big(z_a(t)\big)^2 z_a'(t) = \frac{\mathrm{d}}{\mathrm{d}t}(\mu\circ z_a)(t),$$

and $\mu(z(t_0)) = \mu(z_a(t_0))$. As a consequence, $\mu(z(t)) \leqslant \mu(z_a(t))$, hence $z(t) \leqslant z_a(t)$ for any $t \geqslant t_0$, which is exactly (3.16). Using (3.4), we conclude by observing that

$$\|u(t) - u_\infty\| \leqslant \varphi(E(u(t))) \leqslant \varphi(z_a(t)) = \nu(t).$$

The theorem is proved. □

Remark 3.6. Observe that in the case of a desingularizing function of power type (see Remark 3.4 (c)), we recover well-known estimates [33].

Theorem 3.7 (**The worst-case one-dimensional gradient dynamics for quasi-gradient systems**).
Let $F : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ be a locally Lipschitz continuous mapping that defines a quasi-gradient vector field for some function $E \in C^2(\mathbb{R}^N;\mathbb{R})$ on $\overline{B}(0,R)$, for any $R > 0$. Assume further that the function E is KL and that for any $R > 0$, there exists a positive constant $b > 0$ such that

$$\|\nabla E(u)\| \leqslant b\|F(u)\|, \tag{3.18}$$

for any $u \in \overline{B}(0,R)$. Assume further that for a given initial data $u_0 \in \mathbb{R}^N$ the solution u to (3.1) converges to some rest point u_∞. Denote by φ some desingularizing function for E at u_∞.
Then there exist some constants $c, d > 0, t_0 \in \mathbb{R}$ such that

$$\|u(t) - u_\infty\| \leqslant d\nu\left(ct + t_0\right), \tag{3.19}$$

where ν is a solution to (3.11).

Proof. Combining the techniques used in Theorems 3.2 and 3.5, the proof is almost identical to that of Theorem 3.5. Without loss of generality, we may assume that $E(u_\infty) = 0$. We simply need to check the following inequality which is itself a consequence of the assumption (3.18) applied with $R = \sup_{t>0} \|u(t)\|$.

$$\begin{aligned} -\frac{\mathrm{d}}{\mathrm{d}t}(E \circ u)(t) &= -\langle u'(t), \nabla E(u(t))\rangle \\ &\leqslant \|F(u(t))\| \, \|\nabla E(u(t))\| \\ &\leqslant b \, \|F(u(t))\|^2 \\ &\leqslant b \, \|u'(t)\|^2. \end{aligned}$$

From (3.7) one has $-\frac{\mathrm{d}}{\mathrm{d}t}(\varphi \circ E)(u(t)) \geqslant \alpha \|u'(t)\|$, for any t sufficiently large. Setting $z(t) = E(u(t))$, one obtains $-\frac{\mathrm{d}}{\mathrm{d}t}(\varphi \circ z)(t) \geqslant \frac{\alpha}{\sqrt{b}} \sqrt{-z'(t)}$. The conclusion follows as before by using a reparametrization of (3.11). □

Remark 3.8. Assumption (3.18) is of course necessary and simply means that the vector field F drives solutions to their rest points at least "as fast as ∇E" (see also [20]).

3.3 Damped second order systems are quasi-gradient systems

As announced earlier our approach to the asymptotic behavior of damped second order gradient system is based on the observation that (1.1) can be written as a system having a quasi-gradient structure. For $G \in C^2(\mathbb{R}^N; \mathbb{R})$, let us define $\mathcal{F} : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ by

$$\mathcal{F}(u, v) = \big(-v, \, \gamma v + \nabla G(u) \big).$$

Then (1.1) is equivalent to

$$U'(t) + \mathcal{F}\big(U(t)\big) = 0, \quad t \in \mathbb{R}_+, \text{ with } U = (u, v). \tag{3.20}$$

As explained in the introduction the total energy function $E_T(u, v) = G(u) + \frac{1}{2}\|v\|^2$ (sum of the potential energy and the kinetic energy) is a Liapunov function for our dynamical system (1.1). Formally

$$\big\langle \nabla E_T(u, v), \mathcal{F}(u, v) \big\rangle = \gamma \|v\|^2.$$

From the above we see, that the damped system (1.1) is not quasi-gradient for E_T since one obviously has a degeneracy phenomenon

$$\big\langle \nabla E_T(u, v), \mathcal{F}(u, v) \big\rangle = 0 \ \text{ whenever } \ v = 0, \tag{3.21}$$

where in general $\nabla E_T(u, v) \neq 0$ and $\mathcal{F}(u, v) \neq 0$.
The idea that follows consists in continuously deforming the level sets of E_T, through a family of functions:

$$\mathcal{E}_\lambda : \mathbb{R}^N \times \mathbb{R}^N \longrightarrow \mathbb{R} \ \text{ with } \ \mathcal{E}_0 = E_T \ \ (\lambda \text{ denotes here a positive parameter}),$$

so that the angle formed between each of the gradients of the resulting functions $\mathcal{E}_\lambda, \lambda > 0$ and the vector $\mathcal{F}$ remains far away from $\pi/2$. In other words we seek for functions making $\circ F$ a quasi-gradient vector field.

Proposition 3.9 (The second order gradient systems are quasi-gradient systems). *Let $G \in C^2(\mathbb{R}^N; \mathbb{R})$ and let $\gamma > 0$. For $\lambda > 0$, define $\mathcal{E}_\lambda \in C^1(\mathbb{R}^N \times \mathbb{R}^N; \mathbb{R})$ by*

$$\mathcal{E}_\lambda(u, v) = \left(\frac{1}{2}\|v\|^2 + G(u) \right) + \lambda \langle \nabla G(u), v \rangle.$$

For any $R > 0$, there exists $\lambda_0 > 0$ satisfying the following property. For any $\lambda \in (0, \lambda_0]$, there exists $\alpha > 0$ such that

$$\big\langle \nabla \mathcal{E}_\lambda(u, v), \mathcal{F}(u, v) \big\rangle \geqslant \alpha \, \|\nabla \mathcal{E}_\lambda(u, v)\| \, \|\mathcal{F}(u, v)\|, \tag{3.22}$$

for any $(u, v) \in \overline{B}(0, R) \times \mathbb{R}^N$. Furthermore,

$$\operatorname{crit} \mathcal{E}_\lambda \cap \big(\overline{B}(0, R) \times \mathbb{R}^N \big) = \mathcal{F}^{-1}(\{0\}) \cap \big(\overline{B}(0, R) \times \mathbb{R}^N \big), \tag{3.23}$$

for any $\lambda \in [0, \lambda_0]$.

Proof. For each $(u,v)\in\mathbb{R}^N\times\mathbb{R}^N$, we have $\nabla\mathcal{E}_\lambda(u,v)=\big(\nabla G(u)+\lambda\,\nabla^2G(u)v,\,v+\lambda\nabla G(u)\big)$. Let $R>0$ be given and let $M=\max\big\{\|\nabla^2G(u)\|;\ u\in\overline{B}(0,R)\big\}$. Choose $\lambda_0>0$ small enough to have

$$\gamma-\left(M+\frac{\gamma^2}{2}\right)\lambda_0>0.$$

Let $\lambda\in(0,\lambda_0]$. Then for any $(u,v)\in\overline{B}(0,R)\times\mathbb{R}^N$, we obtain by Young's inequality,

$$\begin{aligned}\big\langle\nabla\mathcal{E}_\lambda(u,v),\mathcal{F}(u,v)\big\rangle&=\gamma\|v\|^2-\lambda\,\langle\nabla^2G(u)v,v\rangle+\lambda\,\langle\nabla G(u),\gamma v\rangle+\lambda\,\|\nabla G(u)\|^2\\&\geqslant\left(\gamma-M\lambda_0-\frac{\lambda_0}{2}\gamma^2\right)\|v\|^2+\frac{\lambda}{2}\,\|\nabla G(u)\|^2\\&\geqslant\alpha_0\,(\|v\|^2+\|\nabla G(u)\|^2),\end{aligned}\tag{3.24}$$

where $\alpha_0=\min\left\{\gamma-\left(M+\frac{\gamma^2}{2}\right)\lambda_0,\ \frac{\lambda}{2}\right\}>0$. Moreover,

$$\|\nabla\mathcal{E}_\lambda(u,v)\|\,\|\mathcal{F}(u,v)\|\leqslant\frac{1}{2}\|\nabla\mathcal{E}_\lambda(u,v)\|^2+\frac{1}{2}\|\mathcal{F}(u,v)\|^2\leqslant C(\|v\|^2+\|\nabla G(u)\|^2).\tag{3.25}$$

Combining (3.25) with (3.24), we deduce that the angle condition (3.22) is satisfied with $\alpha=\frac{\alpha_0}{C}$. Finally, the rest point equivalence (3.23) follows from (3.24). □

Remark 3.10. Note that for $\lambda=0$, we recover the *total energy* $E_T(u,v)=\mathcal{E}_0(u,v)=\frac{1}{2}\|v\|^2+G(u)$.

The following result is of primary importance: roughly speaking it shows that functions which desingularize the potential G at some critical point $\overline{u}$, also desingularize the energy function E_T and more generally the family of deformed functions $\mathcal{E}_\lambda$ at the corresponding critical point $(\overline{u},0)$. This result implies in turn that the decay rate of the energy is essentially conditioned by the geometry of G as one might expect from a mechanical or an intuitive perspective.

In the proposition below one needs the kinetic energy to be desingularized by φ. This explains our main assumption.

Proposition 3.11 (Desingularizing functions of the energy). *Let $G\in C^2(\mathbb{R}^N;\mathbb{R})$, $\overline{u}\in\operatorname{crit}G$ and assume that there exists a desingularizing function $\varphi\in C^1\big((0,r_0);\mathbb{R}_+\big)$ of G at $\overline{u}$ on $B(\overline{u},\eta)$ such that $\varphi'(s)\geqslant\frac{c}{\sqrt{s}}$, for any $s\in(0,r_0)$.*
Then there exist $\lambda_1>0$, $\eta_1>0$ and $c>0$ such that

$$\left\|\nabla\left(\varphi\circ\frac{1}{2}\,|\mathcal{E}_\lambda(\,.\,,\,.\,)-\mathcal{E}_\lambda(\overline{u},0)|\right)(u,v)\right\|\geqslant c,\tag{3.26}$$

for any $\lambda\in[0,\lambda_1]$ and any $(u,v)\in B(\overline{u},\eta_1)\times B(0,\eta_1)$ such that $\mathcal{E}_\lambda(u,v)\neq\mathcal{E}_\lambda(\overline{u},0)$.

Proof. By standard translation arguments, we may assume without loss of generality that $G(\overline{u})=0$ and $\overline{u}=0$. Then $\mathcal{E}_\lambda(0,0)=0$ and (3.26) consists in showing that for some constant $c>0$,

$$\varphi'\left(\frac{1}{2}\,|\mathcal{E}_\lambda(u,v)|\right)\geqslant\frac{c}{\|\nabla\mathcal{E}_\lambda(u,v)\|},$$

for any $\lambda\in[0,\lambda_1]$ and any $(u,v)\in B(0,\eta_1)\times B(0,\eta_1)$ such that $\mathcal{E}_\lambda(u,v)\neq0$. Recall that $0\in\operatorname{crit}G$. Let $M=\max\left\{\|\nabla^2G(u)\|;\ u\in\overline{B}(0,\eta)\right\}$ and define $\lambda_1=\min\left\{\frac{1}{4}\,,\,\frac{1}{2(M^2+1)}\right\}$. We have,

$$\begin{aligned}\|\nabla\mathcal{E}_\lambda(u,v)\|^2&=\|\nabla G(u)+\lambda\nabla^2G(u)v\|^2+\|v+\lambda\nabla G(u)\|^2\\&\geqslant\|\nabla G(u)\|^2+\|v\|^2-\lambda_1(M^2+1)\|v\|^2-2\lambda_1\|\nabla G(u)\|^2\\&\geqslant\frac{1}{2}\big(\|v\|^2+\|\nabla G(u)\|^2\big),\end{aligned}\tag{3.27}$$

and in particular,

$$\|\nabla G(u)\| \leqslant 2\|\nabla \mathcal{E}_\lambda(u,v)\|, \tag{3.28}$$

for any $\lambda \in [0,\lambda_1]$ and any $(u,v) \in \mathbb{R}^N \times \mathbb{R}^N$. Let now $(\lambda,u,v) \in [0,\lambda_1] \times B(0,\eta) \times \mathbb{R}^N$ be such that $\mathcal{E}_\lambda(u,v) \neq 0$. Since φ' is nonincreasing, we have

$$\begin{aligned}\varphi'\left(\frac{1}{2}\left|\mathcal{E}_\lambda(u,v)\right|\right) &\geqslant \varphi'\left(\frac{1}{2}\left|\mathcal{E}_\lambda(u,v)-\mathcal{E}_\lambda(u,0)\right| + \frac{1}{2}\left|\mathcal{E}_\lambda(u,0)\right|\right)\\ &\geqslant \varphi'\left(\max\left\{\left|\mathcal{E}_\lambda(u,v)-\mathcal{E}_\lambda(u,0)\right|, \left|\mathcal{E}_\lambda(u,0)\right|\right\}\right).\end{aligned} \tag{3.29}$$

Let us first find a lower bound on $\varphi'(|\mathcal{E}_\lambda(u,0)|)$. Observe that necessarily $\mathcal{E}_\lambda(u,0) = G(u) \neq 0$. In particular, $\nabla G(u) \neq 0$ (Remark 2.5). We then have by (2.2) and (3.28), $\nabla \mathcal{E}_\lambda(u,v) \neq 0$ and

$$\varphi'\big(|\mathcal{E}_\lambda(u,0)|\big) = \varphi'\big(|G(u)|\big) \geqslant \frac{1}{\|\nabla G(u)\|} \geqslant \frac{1}{2\|\nabla \mathcal{E}_\lambda(u,v)\|}, \tag{3.30}$$

for any $\lambda \in [0,\lambda_1]$ and any $(u,v) \in B(0,\eta) \times \mathbb{R}^N$ such that $\mathcal{E}_\lambda(u,0) \neq 0$.

Let us now estimate $\varphi'(|\mathcal{E}_\lambda(u,v) - \mathcal{E}_\lambda(u,0)|)$ in (3.29) under the assumption $\mathcal{E}_\lambda(u,v) \neq \mathcal{E}_\lambda(u,0)$. Cauchy-Schwarz' inequality implies that for any $\lambda \in [0,\lambda_1]$,

$$|\mathcal{E}_\lambda(u,v) - \mathcal{E}_\lambda(u,0)| \leqslant \frac{1}{2}\big(\|v\|^2 + \lambda_1\|v\|^2 + \lambda_1\|\nabla G(u)\|^2\big). \tag{3.31}$$

Combining (3.31) with (3.27), we deduce that for any $\lambda \in [0,\lambda_1]$ and any $(u,v) \in \mathbb{R}^N \times \mathbb{R}^N$,

$$|\mathcal{E}_\lambda(u,v) - \mathcal{E}_\lambda(u,0)| \leqslant (1+\lambda_1)\|\nabla \mathcal{E}_\lambda(u,v)\|^2. \tag{3.32}$$

By continuity of ∇G, there exists $\eta_1 \in (0,\eta)$ such that

$$\sup\Big\{(1+\lambda_1)\|\nabla \mathcal{E}_\lambda(u,v)\|^2;\ (\lambda,u,v) \in [0,\lambda_1] \times B(0,\eta_1) \times B(0,\eta_1)\Big\} < r_0.$$

Using successively the fact that φ' is nonincreasing and $\varphi'(s) \geqslant \frac{c}{\sqrt{s}}$, it follows from (3.32) that if $(u,v) \in B(0,\eta_1) \times B(0,\eta_1)$ with $\mathcal{E}_\lambda(u,v) \neq \mathcal{E}_\lambda(u,0)$ then $\nabla \mathcal{E}_\lambda(u,v) \neq 0$ and

$$\varphi'\big(|\mathcal{E}_\lambda(u,v) - \mathcal{E}_\lambda(u,0)|\big) \geqslant \varphi'\big((1+\lambda_1)\|\nabla \mathcal{E}_\lambda(u,v)\|^2\big) \geqslant \frac{c_1}{\|\nabla \mathcal{E}_\lambda(u,v)\|}, \tag{3.33}$$

where $c_1 > 0$ is a constant. Finally, inequalities (3.30) and (3.33) together with (3.29) yield the existence of a constant $c > 0$ such that for any $\lambda \in [0,\lambda_1]$ and any $(u,v) \in B(0,\eta_1) \times B(0,\eta_1)$ such that $\mathcal{E}_\lambda(u,v) \neq 0$, there holds $\nabla \mathcal{E}_\lambda(u,v) \neq 0$ and $\varphi'\left(\frac{1}{2}|\mathcal{E}_\lambda(u,v)|\right)\|\nabla \mathcal{E}_\lambda(u,v)\| \geqslant c$, which is the desired result. □

4 Convergence results

Before providing our last results, we would like to recall to the reader that a bounded trajectory of (1.1) may not converge to a single critical point; finite-dimensional counterexamples for $N = 2$ are provided in [4, 41], in each case the trajectory of (1.1) ends up circling indefinitely around a disk.

We now proceed to establish a central result whose specialization to various settings will provide us with several extensions of Haraux-Jendoubi's initial work [31].

Theorem 4.1. *Let $G \in C^2(\mathbb{R}^N;\mathbb{R})$ and $(u_0,u_0') \in \mathbb{R}^N \times \mathbb{R}^N$ be a set of initial conditions for (1.1). Denote by $u \in C^2\big([0,\infty);\mathbb{R}^N\big)$ the unique regular solution to (1.1) with initial data (u_0,u_0'). Assume that the following holds.*

1. ***(The trajectory is bounded)*** $\sup\limits_{t>0}\|u(t)\| < \infty$.

2. (***Convergence to a critical point***) *G is a KL function. Each desingularizing function φ of G satisfies*

$$\varphi'(s) \geqslant \frac{\beta}{\sqrt{s}}, \tag{4.1}$$

for any $s \in (0, \eta_0)$, where β and η_0 are positive constants (see Definition 2.1).

Then,

(i) u' and u'' belong to $L^1\big((0,\infty);\mathbb{R}^N\big)$ and in particular u converges to a single limit u_∞ in $\operatorname{crit} G$.

(ii) When u converges to u_∞, we denote by φ the desingularizing function of G at u_∞. One has the following estimate

$$\|u(t) - u_\infty\| \leqslant c\nu(t),$$

where ν is the solution of the worst-case gradient system

$$\nu'(t) + (\varphi^{-1})'(\nu(t)) = 0,\ \nu(0) > 0.$$

Proof of Theorem 4.1. Let $G \in C^2(\mathbb{R}^N;\mathbb{R})$, let $(u_0, u_0') \in \mathbb{R}^N \times \mathbb{R}^N$, let $u \in C^2\big([0,\infty);\mathbb{R}^N\big)$ and let $\overline{u} \in \mathbb{R}^N$. Set $U(t) = \big(u(t), u'(t)\big)$, $U_0 = (u_0, u_0')$ and $\overline{U} = (\overline{u}, 0)$. Let $\mathcal{F}$ and let $\mathcal{E}_\lambda$ be defined as in Subsection 3.1 and Proposition 3.9, respectively. Note that if $\overline{u} \notin \operatorname{crit} G$ then $\overline{U} \notin \operatorname{crit} \mathcal{E}_\lambda$ and $\varphi(t) = ct$ desingularizes $\mathcal{E}_\lambda$ at $\overline{U}$, for any $\lambda \geqslant 0$ (Remark 2.3 (a) and (3.23)). Otherwise, $\overline{u} \in \operatorname{crit} G$ and we shall apply Proposition 3.11. Since $\sup_{t>0}\|u(t)\| < \infty$, $u''(t) + \gamma u'(t) = A(t)$ where A is bounded. Thus, $u'(t) = u'(0)e^{-\gamma t} + \int_0^t \exp(-\gamma(t-s))A(s)\mathrm{d}s$, and by a straightforward calculation, $\sup_{t>0}\|u'(t)\| < \infty$. It follows that $\sup_{t>0}\|U(t)\| < \infty$. Let $R = \sup_{t>0}\|U(t)\|$. Let $\lambda_0 > 0$ and $0 < \lambda_1 < \lambda_0$ be given by Propositions 3.9 and 3.11, respectively. Let us fix $0 < \lambda_\star < \lambda_1$ and let $\alpha > 0$ be given by Proposition 3.9 for such $\mathcal{E}_{\lambda_\star}$ and R. By Proposition 3.9, the first order system

$$U'(t) + \mathcal{F}\big(U(t)\big) = 0, \quad t \in \mathbb{R}_+, \tag{4.2}$$

has a quasi-gradient structure for $\mathcal{E}_{\lambda_\star}$ on $\overline{B}(0,R)$ (Definition 3.1). Finally, since G has the KL property at $\overline{u}$, $\mathcal{E}_{\lambda_\star}$ also has the KL property at $\overline{U}$ (Proposition 3.11). It follows that Theorem 3.2 applies to U, from which (i) follows.

The estimate part of the proof of (ii) will follow from Theorem 3.7, if we establish that for any $R > 0$, there exists $b > 0$ such that for any $(u,v) \in \overline{B}(0,R) \times \overline{B}(0,R)$,

$$\|\nabla \mathcal{E}_{\lambda_\star}(u,v)\| \leqslant b\|\mathcal{F}(u,v)\|.$$

First we observe that for each $R > 0$ and for any $(u,v) \in \overline{B}(0,R) \times \overline{B}(0,R)$, there exists $k_1 \geqslant 0$ such that

$$\|\nabla \mathcal{E}_{\lambda_\star}(u,v)\|^2 \leqslant k_1\big(\|\nabla G(u)\|^2 + \|v\|^2\big). \tag{4.3}$$

This follows trivially by Cauchy-Schwarz' inequality and the fact that $\nabla^2 G$ is continuous hence bounded on bounded sets. Fix $\sigma > 0$ and recall the inequality $2ab \leqslant \sigma^2 a^2 + \frac{b^2}{\sigma^2}$ for all real numbers a, b. By Cauchy-Schwarz' inequality and the previous inequality

$$\begin{aligned}
\|\mathcal{F}(u,v)\|^2 &= \|v\|^2 + \|\gamma v + \nabla G(u)\|^2 \\
&\geqslant (1+\gamma^2)\|v\|^2 + \|\nabla G(u)\|^2 - 2\|\gamma v\|\|\nabla G(u)\| \\
&\geqslant (1+\gamma^2)\|v\|^2 + \|\nabla G(u)\|^2 - \sigma^2\|\gamma v\|^2 - \frac{1}{\sigma^2}\|\nabla G(u)\|^2 \\
&= (1-(\sigma^2-1)\gamma^2)\|v\|^2 + \left(1 - \frac{1}{\sigma^2}\right)\|\nabla G(u)\|^2.
\end{aligned}$$

Choosing $\sigma > 1$ so that $1-(\sigma^2-1)\gamma^2 > 0$ yields $k_2 > 0$ such that $\|\mathcal{F}(u,v)\|^2 \geqslant k_2\big(\|\nabla G(u)\|^2 + \|v\|^2\big)$, for any u, v in $\mathbb{R}^N$. Combining this last inequality with (4.3), we obtain $\|\nabla \mathcal{E}_{\lambda_\star}(u,v)\|^2 \leqslant \frac{k_1}{k_2}\|\mathcal{F}(u,v)\|^2$, for any $(u,v) \in \overline{B}(0,R) \times \overline{B}(0,R)$. Hence the result. □

Remark 4.2. (a) As announced previously convergence rates depend directly on the geometry of G through φ.
(b) The fact that the length of the velocity curve u' is finite suggests that highly oscillatory phenomena are unlikely.

5 Consequences

In the following corollaries, the mapping $\mathbb{R}_+ \ni t \longmapsto u(t)$ is a solution curve of (1.1).

Corollary 5.1 (**Convergence theorem for real-analytic functions** [31]). *Assume that $G : \mathbb{R}^N \longrightarrow \mathbb{R}$ is real-analytic and let u be a bounded solution to (1.1). Then we have the following result.*

(i) (u, u') has a finite length. In particular u converges to a critical point u_∞.

(ii) When u converges to u_∞, we denote by $\varphi(s) = cs^\theta$ $\left(\text{with } c > 0 \text{ and } \theta \in \left(0, \frac{1}{2}\right]\right)$ the desingularizing function of G at u_∞ – the quantity θ is the Łojasiewicz exponent associated with u_∞. One has the following estimates.

(a) $\|u(t) - u_\infty\| \leqslant ct^{-\frac{\theta}{1-2\theta}}$, with $c > 0$, when $\theta \in \left(0, \frac{1}{2}\right)$.

(b) $\|u(t) - u_\infty\| \leqslant c' \exp(-c't)$, with $c', c'' > 0$, when $\theta = \frac{1}{2}$.

Proof. The proof follows directly from the original Łojasiewicz inequality [44, 43] and the fact that desingularizing functions for real-analytic functions are indeed of the form $\varphi(s) = cs^\theta$ with $\theta \in (0, \frac{1}{2}]$. Hence (2.10) holds and Theorem 4.1 applies, see also Remark 3.4 (c). □

Corollary 5.2 (**Convergence theorem for definable functions**). *Let $\mathcal{O}$ be an o-minimal structure that contains the collection of semi-algebraic sets. Assume $G : \mathbb{R}^N \longrightarrow \mathbb{R}$ is C^2 and definable in $\mathcal{O}$. Let u be a bounded solution to (1.1). Then we have the following result.*

(i) u' and u'' belong to $L^1\big((0, \infty); \mathbb{R}^N\big)$ and in particular u converges to a single limit u_∞ in $\operatorname{crit} G$.

(ii) When u converges to u_∞ we denote by φ the desingularizing function of G at u_∞. One has the following estimate

$$\|u(t) - u_\infty\| \leqslant c\nu(t),$$

where ν is a solution of the worst-case gradient system

$$\nu'(t) + (\varphi^{-1})'(\nu(t)) = 0, \ \nu(0) > 0.$$

Proof. G is a KL function by Kurdyka's version of the Łojasiewicz inequality. The fact that $\varphi'(s) \geqslant \frac{c}{\sqrt{s}}$ comes from Lemma 2.9. So, Theorem 4.1 applies. □

Corollary 5.3 (**Convergence theorem for the one-dimensional case** [30]). *Let $G \in C^2(\mathbb{R}; \mathbb{R})$ and let u be a bounded solution to (1.1). Then u converges to a single point and we have the same type of rate of convergence as in the previous corollary.*

Proof. We proceed as in [49]. Argue by contradiction and assume that $\omega(u_0, u_0')$, the ω-limit set of (u_0, u_0'), is not a singleton. Since $\omega(u_0, u_0')$ is connected in $\mathbb{R}$, it is an interval and has a nonempty interior. Take $\overline{u}$ in the interior of $\omega(u_0, u_0')$ The Łojasiewicz inequality trivially holds at $\overline{u}$ for $G \equiv 0$ with $\varphi(s) = \sqrt{s}$ (recall $\overline{u}$ is interior). Apply then Theorem 4.1. □

Remark 5.4. In the one-dimensional case, convergence can be obtained with much more general forms of damping, see [15].

Corollary 5.5 (**Convergence theorem for convex functions satisfying growth conditions**). *Let $G \in C^2(\mathbb{R}^N; \mathbb{R})$ be a convex function such that*

$$\operatorname{argmin} G \stackrel{def}{=} \big\{u \in \mathbb{R}^N; G(u) = \min G\big\},$$

is nonempty (note that $\operatorname{argmin} G = \operatorname{crit} G$). Assume further that, for each minimizer x^*, there exists $\eta > 0$, such that G satisfies

$$\forall u \in B(x^*, \eta), \ G(u) \geqslant \min G + c \operatorname{dist}(u, \operatorname{argmin} G)^r, \tag{5.1}$$

with $r \geqslant 1$ and $c > 0$. Then the solution curve $t \longmapsto (u(t), u'(t))$ has a finite length. In particular u converges to a minimizer u_∞ of G as t goes to ∞.

Proof. A general result of Alvarez [1] ensures that u is bounded (and even converges). On the other hand it has been shown in [10] that functions satisfying the growth assumption (5.1), also satisfy the Łojasiewicz inequality with desingularizing functions of the form $s \longmapsto c's^{1-1/r}$ with $c' > 0$. Combining the previous arguments, the conclusion follows readily. □

Remark 5.6. An alternative and more general approach to establish that trajectories have a finite length has been developed for convex functions in [45, 22].

A Appendix: some elements on o-minimal structures

Some references for o-minimal structures are [21, 25, 42, 23]. We only collect in this appendix the elements that are necessary to follow our main developments.

Definition A.1 (o-minimal structure [21, Definition 1.5]). An *o-minimal* structure on $(\mathbb{R}, +, .)$ is a sequence of Boolean algebras[9] $\mathcal{O} = \{\mathcal{O}_n\}_{n\in\mathbb{N}}$ of subsets of $\mathbb{R}^n$ such that for each $n \in \mathbb{N}$,

(i) if A belongs to $\mathcal{O}_n$ then $A \times \mathbb{R}$ and $\mathbb{R} \times A$ belong to $\mathcal{O}_{n+1}$;

(ii) if $\Pi : \mathbb{R}^{n+1} \longrightarrow \mathbb{R}^n$ is the canonical projection onto $\mathbb{R}^n$ then for any $A \in \mathcal{O}_{n+1}$, the set $\Pi(A)$ belongs to $\mathcal{O}_n$;

(iii) $\mathcal{O}_n$ contains the family of real algebraic subsets of $\mathbb{R}^n$, that is, every set of the form

$$\Big\{x \in \mathbb{R}^n;\, p(x) = 0\Big\},$$

where $p : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a real polynomial function;

(iv) the elements of $\mathcal{O}_1$ are exactly the finite unions of intervals and points.

Being given an o-minimal structure $\mathcal{O}$, a set $A \subset \mathbb{R}^n$ is called *definable* (in $\mathcal{O}$) if $A \in \mathcal{O}_n$. A mapping $F : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is said to be *definable in* $\mathcal{O}$ if its graph is definable in Ω as a subset of $\mathbb{R}^n \times \mathbb{R}^m$. A point-to-set mapping

$$S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m,$$

maps each point x in $\mathbb{R}^n$ to a subset $S(x)$ of $\mathbb{R}^m$. The *domain of* S, denoted by $\operatorname{dom} S$, is given by the set of elements x in $\mathbb{R}^n$ such that $S(x)$ is nonempty. The graph of S is defined by

$$\operatorname{graph} S = \Big\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m; y \in S(x)\Big\}.$$

As previously a point-to-set mapping is called *definable* (*in* $\mathcal{O}$) if its graph is definable in $\mathbb{R}^n \times \mathbb{R}^m$.

Example A.2. (a) **Semi-algebraic sets.** The first and simplest example of o-minimal structure is given by the class of semi-algebraic objects (see (2.1)). Tarski-Seidenberg principle (see [7]) asserts that linear projections of semi-algebraic sets are semi-algebraic sets, in other words item (ii) of Definition A.1 holds for the class of semi-algebraic sets. The other items of the definition are easy to establish.
(b) **Globally subanalytic sets.** There exists an o-minimal structure that contains semi-algebraic sets and sets of the form $\big\{(x, t) \in [-1, 1]^n \times \mathbb{R};\, f(x) = t\big\}$, where $f : [-1, 1]^n \longrightarrow \mathbb{R}$ $(n \in \mathbb{N})$ is a real analytic function that can be extended analytically on a neighborhood of the square $[-1, 1]^n$ – these are sometimes called restricted analytic functions. This result is essentially due to Gabrielov [28]; sets belonging to this structure are called *globally subanalytic sets* (see [24] and the references therein).
(c) **Log-exp structure.** There exists an o-minimal structure containing the globally subanalytic sets and the graph of $\exp : \mathbb{R} \longrightarrow \mathbb{R}$, see [24].

[9] Recall that a Boolean algebra is stable by finite union, finite intersection and contains the empty set and the total space; here $\emptyset \in \mathcal{O}_n$ and $\mathbb{R}^n \in \mathcal{O}_n$.

There are other results on o-minimal structures and the field is still very active, but the above examples give a good idea of the power of the concept.

We now describe some stability/regularity results that we used in this paper.

Let $\mathcal{O}$ be an o-minimal structure on $(\mathbb{R}, +, .)$.

Lemma A.3 (Monotonicity Lemma [25, Theorem 4.1]). *Let $f : I \subset \mathbb{R} \longrightarrow \mathbb{R}$ be a definable function and $k \in \mathbb{N}$. Then there exists a finite partition of I into p intervals $I_1, \dots, I_p$, such that f restricted to each nontrivial interval I_j, $j \in \{1, \dots, p\}$, is C^k and either strictly monotone or constant. Observe that some I_j can be reduced to a singleton.*

Lemma A.4 (Definable Selection Lemma [21]). *Let $S : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a definable point-to-set mapping. Then there exists a definable mapping $F : \operatorname{dom} S \longrightarrow \mathbb{R}^m$ such that*

$$F(x) \in S(x),\ \forall x \in \operatorname{dom} S.$$

We recall the following theorem as stated in Kurdyka's original work [42].

Theorem A.5. *Let Ω be a nonempty open bounded subset of $\mathbb{R}^n$ and $f : \Omega \to \mathbb{R}$ a differentiable definable function with $f > 0$ on Ω. Then there exist $r_0 > 0$ and a continuous definable function $\varphi : [0, r_0) \to \mathbb{R}_+$ such that $\varphi(0) = 0$, $\varphi \in C^1(0, r_0)$ and $\varphi' > 0$ such that*

$$\|\nabla (\varphi \circ f)(x)\| \geqslant 1,\ \forall x \in \Omega.$$

Remark A.6. Let us show how to recover the form of KL inequality given in Theorem 2.2.
We adopt the notation of Theorem 2.2. Fix $\mu > 0$. Apply first, the above result to $G - G(\overline{u})$ (respectively, to $G(\overline{u}) - G$) on $\Omega_1 = B(\overline{u}, \mu) \cap [G - G(\overline{u}) > 0]$ (respectively, on $\Omega_2 = B(\overline{u}, \mu) \cap [G(\overline{u}) - G > 0]$). This gives $\varphi_1 : [0, r_1) \longrightarrow \mathbb{R}_+$ and $\varphi_2 : [0, r_2) \longrightarrow \mathbb{R}_+$, as in Kurdyka's Theorem. Let us now build a "global" φ as in Theorem 2.2. First recall that the derivative of a differentiable definable function is definable in the same structure, see [21]. Set $p(s) = (\varphi_1' - \varphi_2')(s)$. By definability, p is positive, negative or null on an interval of the form $(0, \varepsilon)$. This yields the existence of r in $(0, \min\{r_1, r_2\})$ such that, for instance, $\varphi_1' > \varphi_2'$ on $(0, r)$. Set then $\varphi = \varphi_1$ and observe that

$$\|\nabla (\varphi \circ |G(\,\cdot\,) - G(\overline{u})|)(u)\| \geqslant 1,\ \forall u \in B(0, \eta) \setminus [G \neq G(\overline{u})],$$

when η is sufficiently small.

Acknowledgements. We are grateful to the referees for their very careful reading and their constructive input.

References

[1] F. Alvarez. On the minimizing property of a second order dissipative system in Hilbert spaces. *SIAM J. Control Optim.*, 38(4):1102–1119 (electronic), 2000.

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

[3] H. Attouch, A. Cabot, and P. Redont. The dynamics of elastic shocks via epigraphical regularization of a differential inclusion. Barrier and penalty approximations. *Adv. Math. Sci. Appl.*, 12(1):273–306, 2002.

[4] H. Attouch, X. Goudou, and P. Redont. The heavy ball with friction method. I. The continuous dynamical system: global exploration of the local minima of a real-valued function by asymptotic analysis of a dissipative dynamical system. *Commun. Contemp. Math.*, 2(1):1–34, 2000.

[5] T. Bárta. Convergence to equilibrium of relatively compact solutions to evolution equations. *Electron. J. Differential Equations*, No. 81, pp. 1–9, 2014.

[6] T. Bárta, R. Chill, and E. Fašangová. Every ordinary differential equation with a strict Lyapunov function is a gradient system. *Monatsh. Math.*, 166(1):57–72, 2012.

[7] J. Bochnak, M. Coste, and M.-F. Roy. *Real algebraic geometry*, volume 36 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1998. Translated from the 1987 French original, Revised by the authors.

[8] J. Bolte. Sur quelques principes de convergence en Optimisation. Habilitation à diriger des recherches, Université Pierre et Marie Curie, 2008.

[9] J. Bolte, A. Daniilidis, A. Lewis, and M. Shiota. Clarke subgradients of stratifiable functions. *SIAM J. Optim.*, 18(2):556–572 (electronic), 2007.

[10] J. Bolte, A. Daniilidis, O. Ley, and L. Mazet. Characterizations of Łojasiewicz inequalities: subgradient flows, talweg, convexity. *Trans. Amer. Math. Soc.*, 362(6):3319–3363, 2010.

[11] R. I. Boţ, E. R. Csetnek, and S. László. An inertial forward-backward algorithm for the minimization of the sum of two nonconvex functions. arXiv:1410.0641 [math.OC], 2014.

[12] T. Brox, Y. Chen, P. Ochs, and T. Pock. iPiano: inertial proximal algorithm for nonconvex optimization. *SIAM J. Imaging Sci.*, 7(2):1388–1419, 2014.

[13] R. E. Bruck, Jr. Asymptotic convergence of nonlinear contraction semigroups in Hilbert space. *J. Funct. Anal.*, 18:15–26, 1975.

[14] A. Cabot. Motion with friction of a heavy particle on a manifold—applications to optimization. *M2AN Math. Model. Numer. Anal.*, 36(3):505–516, 2002.

[15] A. Cabot, H. Engler, and S. Gadat. On the long time behavior of second order differential equations with asymptotically small dissipation. *Trans. Amer. Math. Soc.*, 361(11):5983–6017, 2009.

[16] A. Cabot and P. Frankel. Asymptotics for some semilinear hyperbolic equations with non-autonomous damping. *J. Differential Equations*, 252(1):294–322, 2012.

[17] L. Chergui. Convergence of global and bounded solutions of a second order gradient like system with nonlinear dissipation and analytic nonlinearity. *J. Dynam. Differential Equations*, 20(3):643–652, 2008.

[18] L. Chergui. Convergence of global and bounded solutions of the wave equation with nonlinear dissipation and analytic nonlinearity. *J. Evol. Equ.*, 9(2):405–418, 2009.

[19] R. Chill. On the Łojasiewicz-Simon gradient inequality. *J. Funct. Anal.*, 201(2):572–601, 2003.

[20] R. Chill, A. Haraux, and M. A. Jendoubi. Applications of the Łojasiewicz-Simon gradient inequality to gradient-like evolution equations. *Anal. Appl. (Singap.)*, 7(4):351–372, 2009.

[21] M. Coste. An introduction to o-minimal geometry. RAAG Notes, Institut de Recherche Mathématiques de Rennes, 1999.

[22] A. Daniilidis, O. Ley, and S. Sabourau. Asymptotic behaviour of self-contracted planar curves and gradient orbits of convex functions. *J. Math. Pures Appl. (9)*, 94(2):183–199, 2010.

[23] L. van den Dries. *Tame topology and o-minimal structures*, volume 248 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1998.

[24] L. van den Dries, A. Macintyre, and D. Marker. The elementary theory of restricted analytic fields with exponentiation. *Ann. of Math. (2)*, 140(1):183–205, 1994.

[25] L. van den Dries and C. Miller. Geometric categories and o-minimal structures. *Duke Math. J.*, 84(2):497–540, 1996.

[26] E. Feireisl, F. Issard-Roch, and H. Petzeltová. A non-smooth version of the Łojasiewicz-Simon theorem with applications to non-local phase-field systems. *J. Differential Equations*, 199(1):1–21, 2004.

[27] E. Feireisl and F. Simondon. Convergence for semilinear degenerate parabolic equations in several space dimensions. *J. Dynam. Differential Equations*, 12(3):647–673, 2000.

[28] A. Gabrielov. Complements of subanalytic sets and existential formulas for analytic functions. *Invent. Math.*, 125(1):1–12, 1996.

[29] S. Gadat and F. Panloup. Long time behaviour and stationary regime of memory gradient diffusions. *Ann. Inst. Henri Poincaré Probab. Stat.*, 50(2):564–601, 2014.

[30] A. Haraux. Asymptotics for some nonlinear O.D.E. of the second order. *Nonlinear Anal.*, 10(12):1347–1355, 1986.

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

[32] A. Haraux and M. A. Jendoubi. Convergence of bounded weak solutions of the wave equation with dissipation and analytic nonlinearity. *Calc. Var. Partial Differential Equations*, 9(2):95–124, 1999.

[33] A. Haraux and M. A. Jendoubi. Decay estimates to equilibrium for some evolution equations with an analytic nonlinearity. *Asymptot. Anal.*, 26(1):21–36, 2001.

[34] A. Haraux and M. A. Jendoubi. On the convergence of global and bounded solutions of some evolution equations. *J. Evol. Equ.*, 7(3):449–470, 2007.

[35] A. Haraux and M. A. Jendoubi. The Łojasiewicz gradient inequality in the infinite-dimensional Hilbert space framework. *J. Funct. Anal.*, 260(9):2826–2842, 2011.

[36] A. Haraux and M. A. Jendoubi. Asymptotics for a second order differential equation with a linear, slowly time-decaying damping term. *Evol. Equ. Control Theory*, 2(3):461–470, 2013.

[37] K.-H. Hoffmann and P. Rybka. Convergence of solutions to Cahn-Hilliard equation. *Comm. Partial Differential Equations*, 24(5-6):1055–1077, 1999.

[38] M. A. Jendoubi. Convergence of global and bounded solutions of the wave equation with linear dissipation and analytic nonlinearity. *J. Differential Equations*, 144(2):302–312, 1998.

[39] M. A. Jendoubi. A simple unified approach to some convergence theorems of L. Simon. *J. Funct. Anal.*, 153(1):187–202, 1998.

[40] M. A. Jendoubi and R. May. On an asymptotically autonomous system with Tikhonov type regularizing term. *Arch. Math. (Basel)*, 95(4):389–399, 2010.

[41] M. A. Jendoubi and P. Poláčik. Non-stabilizing solutions of semilinear hyperbolic and elliptic equations with damping. *Proc. Roy. Soc. Edinburgh Sect. A*, 133(5):1137–1153, 2003.

[42] K. Kurdyka. On gradients of functions definable in o-minimal structures. *Ann. Inst. Fourier (Grenoble)*, 48(3):769–783, 1998.

[43] S. Łojasiewicz. Ensembles semi-analytiques. Preprint, I.H.E.S. Bures-sur-Yvette, 1965.

[44] S. Łojasiewicz. Une propriété topologique des sous-ensembles analytiques réels. In *Les Équations aux Dérivées Partielles (Paris, 1962)*, pages 87–89. Éditions du Centre National de la Recherche Scientifique, Paris, 1963.

[45] P. Manselli and C. Pucci. Maximum length of steepest descent curves for quasi-convex functions. *Geom. Dedicata*, 38(2):211–227, 1991.

[46] B. Poljak. Some methods of speeding up the convergence of iteration methods. *U.S.S.R. Comput. Math. Math. Phys.*, 4(5):1–17, 1964.

[47] L. Simon. Asymptotics for a class of nonlinear evolution equations, with applications to geometric problems. *Ann. of Math. (2)*, 118(3):525–571, 1983.

[48] A. J. Wilkie. Model completeness results for expansions of the ordered field of real numbers by restricted Pfaffian functions and the exponential function. *J. Amer. Math. Soc.*, 9(4):1051–1094, 1996.

[49] T. I. Zelenjak. Stabilization of solutions of boundary value problems for a second-order parabolic equation with one space variable. *Differencial′nye Uravnenija*, 4:34–45, 1968.