SIAM J. OPTIM.
Vol. 17, No. 4, pp. 1205–1223

THE ŁOJASIEWICZ INEQUALITY FOR NONSMOOTH SUBANALYTIC FUNCTIONS WITH APPLICATIONS TO SUBGRADIENT DYNAMICAL SYSTEMS*

JÉRÔME BOLTE†, ARIS DANIILIDIS‡, AND ADRIAN LEWIS§

Abstract. Given a real-analytic function $f : \mathbb{R}^n \to \mathbb{R}$ and a critical point $a \in \mathbb{R}^n$, the Łojasiewicz inequality asserts that there exists $\theta \in [\frac{1}{2}, 1)$ such that the function $|f - f(a)|^\theta \, \|\nabla f\|^{-1}$ remains bounded around a. In this paper, we extend the above result to a wide class of nonsmooth functions (that possibly admit the value $+\infty$), by establishing an analogous inequality in which the derivative $\nabla f(x)$ can be replaced by any element x^* of the subdifferential $\partial f(x)$ of f. Like its smooth version, this result provides new insights into the convergence aspects of subgradient-type dynamical systems. Provided that the function f is sufficiently regular (for instance, convex or lower-C^2), the bounded trajectories of the corresponding subgradient dynamical system can be shown to be of finite length. Explicit estimates of the rate of convergence are also derived.

Key words. Łojasiewicz inequality, subanalytic function, nonsmooth analysis, subdifferential, dynamical system, descent method

AMS subject classifications. Primary, 26D10; Secondary, 32B20, 49K24, 49J52, 37B35, 14P15

DOI. 10.1137/050644641

1. Introduction.

Let U be a nonempty open subset of $\mathbb{R}^n$ equipped with its canonical Euclidean norm $|| \cdot ||$, and let $f : U \to \mathbb{R}$ be a real-analytic function. According to the Łojasiewicz gradient inequality [16, 17, 18], if $a \in U$ is a *critical* point of f, that is, $\nabla f(a) = 0$, then there exists $\theta \in [0, 1)$ such that the function

$$\frac{|f - f(a)|^\theta}{\|\nabla f\|} \tag{1}$$

remains bounded around the point a. (Throughout this work we set $0^0 = 1$, and we interpret $\lambda/0$ as $+\infty$ if $\lambda > 0$ and 0 if $\lambda = 0$.)

Recently, Kurdyka [13, Theorem 1] has extended the above result to C^1 functions whose graphs belong to an o-minimal structure (see [8], for example), and thus in particular to "globally subanalytic" functions. On the other hand, (1) might fail for C^∞ functions with no "adequate" geometric structure. Such functions can either satisfy a weaker condition (i.e., $\theta = 1$) or present wild oscillations around their critical point, preventing any comparison between their value and the norm of their gradient.

*Received by the editors November 9, 2005; accepted for publication (in revised form) June 6, 2006; published electronically January 12, 2007. This research was supported by NSERC.
http://www.siam.org/journals/siopt/17-4/64464.html

†Equipe Combinatoire et Optimisation (UMR 7090), Case 189, Université Pierre et Marie Curie, 4 Place Jussieu, 75252 Paris Cedex 05, France (bolte@math.jussieu.fr; http://www.ecp6.jussieu.fr/pageperso/bolte/). This author's work was partially supported by the CECM (Simon Fraser University) and the C.M.M. (Universidad de Chile).

‡Departament de Matemàtiques, C1/320, Universitat Autònoma de Barcelona, E-08193 Bellaterra (Cerdanyola del Vallès), Spain (arisd@mat.uab.es; http://mat.uab.es/∼arisd). This author's research was partially supported by the CECM (Simon Fraser University), the C.M.M. (Universidad de Chile), and the MEC grant MTM2005–08572–C03-03 (Spain).

§School of Operations Research and Industrial Engineering, Cornell University, 234 Rhodes Hall, Ithaca, NY 14853 (aslewis@orie.cornell.edu; http://www.orie.cornell.edu/∼aslewis). This author's research was partially supported by National Science Foundation grant DMS-0504032.

The following one-dimensional examples illustrate failures of these two types (around the critical point $a = 0$):

$$f(x) = \begin{cases} \exp(-1/x^2) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0, \end{cases} \qquad \text{and} \qquad g(x) = \begin{cases} \exp(-1/x^2)\sin(1/x) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

The aim of this note is to establish a *nonsmooth* version of the Łojasiewicz inequality (1) for lower semicontinuous convex subanalytic functions (Theorem 3.3) and for continuous subanalytic functions (Theorem 3.1). A first and simple illustration is given by the example of the Euclidean norm function $h(x) = \|x\|$, which satisfies (1) for every $\theta \in [0, 1)$ around zero (which is a "generalized" critical point; see Definition 2.11) but is not differentiable at 0. Behavior of this type is hereby shown to hold true for a large class of nonsmooth functions, leading to the conclusion that the Łojasiewicz inequality is more linked to the underlying geometrical structure of f than to its smoothness.

Given an extended-real-valued subanalytic function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, our approach to generalizing property (1) relies on a one-sided notion of generalized gradients called subgradients. For both a mathematical and a historical account on this notion, as well as for classical results in nonsmooth analysis, one is referred to the monographs of Clarke et al. [7] and Rockafellar and Wets [20].

Subgradients are obtained according to a two-stage process. First the equality in the definition of the usual gradient is relaxed into an inequality (Definition 2.10(i)): this gives rise to the notion of Fréchet subgradients. Then, by a closure operation, the so-called limiting subdifferential ∂f can be defined (Definition 2.10(ii)). This notion constitutes the basis for the generalization of the Łojasiewicz inequality to nonsmooth functions. Let us also mention that in this formalism Fermat's rule reads as follows: if a is a local minimizer of f, then $\partial f(a) \ni 0$; conversely, if $a \in \mathbb{R}^n$ is such that $\partial f(a) \ni 0$, the point a is called a critical point.

Variational analysis and subdifferential calculus provide a framework for the modeling of unilateral constraints in mechanics and in partial differential equations [11, 6, 9]. Such a calculus is also central in optimization. In particular it provides variational tools to treat constrained and unconstrained minimization problems on an equal theoretical level. This stems from the simple fact that minimizing f over a closed set C amounts to minimizing $f + \delta_C$ over $\mathbb{R}^n$, where δ_C is the indicator function of C; that is

$$\delta_C(x) = \begin{cases} 0 & \text{if } x \in C, \\ +\infty & \text{otherwise.} \end{cases} \tag{2}$$

Those domains have as a common topic the behavior at infinity of dynamical systems governed by subdifferential operators; see [15] for an insight in optimization. An important motivation that drove us to transpose the Łojasiewicz result into a nonsmooth context is precisely its expected consequences in the asymptotic analysis of such subgradient-type dynamical systems. Those are modeled on the following type of differential inclusion:

$$\dot{x}(t) \in -\partial f(x(t)), \quad t \geq 0, \quad x(0) \in \mathbb{R}^n,$$

where for any $x \in \mathbb{R}^n$, $\partial f(x)$ denotes the set of limiting subgradients. The above differential inclusion generalizes the classical gradient dynamical system

$$\dot{x}(t) = -\nabla f(x(t)), \quad t \geq 0, \quad x(0) \in \mathbb{R}^n. \tag{3}$$

In his pioneering work on real-analytic functions [16, 17], Łojasiewicz provided the main ingredient—namely, (1)—that allows us to derive the convergence of all bounded trajectories of (3) to critical points. As can be seen from a counterexample due to Palis and De Melo [19, p. 14], the set of cluster points of a bounded trajectory generated by the gradient of a C^∞ function is, in general, far from being a singleton. Those famous results illustrate the importance of gradient vector fields of functions satisfying the Łojasiewicz inequality. An even more striking feature is that the trajectories converge "in direction" when approaching critical points. This fact had been conjectured by Thom (around 1972; see [22]) for real-analytic functions, and established by Kurdyka, Mostowski, and Parusiński in [14]. The subanalytic generalized Thom conjecture remains open even in the smooth case (see [13, Conjecture F]).

In section 4 we extend Łojasiewicz results to a nonsmooth setting (f is a subanalytic proximal retract), by showing that all bounded trajectories have a finite length (Theorem 4.5). We also provide estimates of the asymptotic convergence rate towards the critical points (Theorem 4.7).

For related results on this topic, see [1] ; for other applications to partial differential equations, see the works of Simon [21] and Haraux [12].

2. Preliminaries.

The key ingredients for the nonsmooth extension of the Łojasiewicz inequality are subanalyticity of the function f and notions of generalized differentiation provided by variational analysis.

2.1. Subanalytic sets and stability properties.

We recall the following definition.

Definition 2.1 (subanalyticity). (i) *A subset A of $\mathbb{R}^n$ is called* semianalytic *if each point of $\mathbb{R}^n$ admits a neighborhood V for which $A \cap V$ assumes the following form:*

$$\bigcup_{i=1}^{p} \bigcap_{j=1}^{q} \{x \in V : f_{ij}(x) = 0, g_{ij}(x) > 0\},$$

where the functions $f_{ij}, g_{ij} : V \mapsto \mathbb{R}$ are real-analytic for all $1 \le i \le p, 1 \le j \le q$.

(ii) *The set A is called* subanalytic *if each point of $\mathbb{R}^n$ admits a neighborhood V such that*

$$A \cap V = \{x \in \mathbb{R}^n : (x, y) \in B\},$$

where B is a bounded semianalytic subset of $\mathbb{R}^n \times \mathbb{R}^m$ for some $m \ge 1$.

(iii) *Given $m, n \in \mathbb{N}^*$, a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ (respectively, a point-to-set operator $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$) is called* subanalytic *if its graph is a subanalytic subset of $\mathbb{R}^n \times \mathbb{R}$ (respectively, of $\mathbb{R}^n \times \mathbb{R}^m$).*

Recall that the graphs of f and T, denoted respectively by $\mathrm{Gr}\, f$ and $\mathrm{Gr}\, T$, are defined by

$$\mathrm{Gr}\ f := \{(x, \lambda) \in \mathbb{R}^n \times \mathbb{R} : f(x) = \lambda\}, \qquad \mathrm{Gr}\ T := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : y \in T(x)\}.$$

Some of the elementary properties of subanalytic sets have been gathered below (see, e.g., [4, 10, 18]):

- Subanalytic sets are closed under locally finite union and intersection. The complement of a subanalytic set is subanalytic (Gabrielov theorem).
- If A is subanalytic, then so are its closure $\mathrm{cl}\, A$, its interior $\mathrm{int}\, A$, and its boundary $\mathrm{bd}\, A$.

– Given a subanalytic set S, the distance $d_S(x) := \inf\{\|x-a\| : a \in S\}$ is a subanalytic function.
– *Path connectedness* (see, e.g., [10, Facts 1.10–1.12]): Any subanalytic set has a locally finite number of connected components. Each component is subanalytic and subanalytically path connected; that is, every two points can be joined by a continuous subanalytic path that lies entirely in the set.
– *Curve selection lemma* (see, e.g., [4, Lemma 6.3]): If A is a subanalytic subset of $\mathbb{R}^n$ and $a \in \operatorname{bd} A$, then there exists an analytic path $z : (-1,1) \to \mathbb{R}^n$ satisfying $z(0) = a$ and $z((0,1)) \subset A$.

The image and the preimage of a subanalytic set are not in general subanalytic sets. This is essentially due to the fact that the image of an unbounded subanalytic set by a linear projection may fail to be subanalytic. Consider, for instance, the set $\{(\frac{1}{n+1}, n) : n \in \mathbb{N}\}$, whose projection onto $\mathbb{R} \times \{0\}$ is not subanalytic at 0.

To remedy to this lack of stability, let us introduce a stronger analytic-like notion called global subanalyticity (see [10] and references therein).

For each $n \in \mathbb{N}$, set $C_n = (-1,1)^n$ and define τ_n by

$$\tau_n(x_1, \dots, x_n) = \left(\frac{x_1}{1+x_1^2}, \dots, \frac{x_n}{1+x_n^2} \right) \in C_n.$$

DEFINITION 2.2 (global subanalyticity; see, e.g., [10, p. 506]). (i) *A subset S of $\mathbb{R}^n$ is called* globally subanalytic *if its image under τ_n is a subanalytic subset of $\mathbb{R}^n$.*

(ii) *An extended-real-valued function (respectively, a multivalued mapping) is called globally subanalytic if its graph is globally subanalytic.*

Globally subanalytic sets are subanalytic, and conversely any *bounded* subanalytic set is globally subanalytic. Typical examples of subanalytic sets which are not globally subanalytic are the set of integers $\mathbb{Z}$, the graph of the sinus function, the spiral $\{(t\cos t, t\sin t) \in \mathbb{R}^2 : t \geq 0\}$, etc. The class of semialgebraic sets (e.g., [3, 8]) provides an important subclass of globally subanalytic sets. Recall that a set $A \subset \mathbb{R}^n$ is called *semialgebraic* if it assumes the following form:

$$A = \bigcup_{i=1}^{p} \bigcap_{j=1}^{q} \{x \in V : f_{ij}(x) = 0, g_{ij}(x) > 0\},$$

where $f_{ij}, g_{ij} : \mathbb{R}^n \mapsto \mathbb{R}$ are polynomial functions for all $1 \leq i \leq p, 1 \leq j \leq q$. (Readers who are unfamiliar with subanalytic geometry might in a first reading replace "subanalytic" and "globally subanalytic" by "semialgebraic" in the statements that follow.)

A major fact concerning the class of globally subanalytic sets is its stability under linear projections.

THEOREM 2.3 (projection theorem; see, e.g., [10, Example 4, p. 505]). *Let $\Pi(x_1, \dots, x_{n+1}) = (x_1, \dots, x_n)$ be the canonical projection from $\mathbb{R}^{n+1}$ onto $\mathbb{R}^n$. If S is a globally subanalytic subset of $\mathbb{R}^{n+1}$, then so is $\Pi(S)$ in $\mathbb{R}^n$.*

Among the numerous consequences of the above result in terms of stability, the following properties are crucial to our main results:

– The image or the preimage of a globally subanalytic set by a globally subanalytic function (respectively, globally subanalytic multivalued operator) is globally subanalytic (see, e.g., [10, p. 504]).
– *Monotonicity lemma* (e.g., [10, Fact 4.1]): Take $\alpha < \beta$ in $\mathbb{R}$. If $\varphi : (\alpha, \beta) \to \mathbb{R}$ is a globally subanalytic function, then there is a partition $t_0 := \alpha < t_1 <$

$\cdots < t_{l+1} := \beta$ of (α, β) such that $\varphi|_{(t_i, t_{i+1})}$ is C^∞ and either constant or strictly monotone, for $i \in \{0, \ldots, l\}$. Moreover ([13], e.g.), φ admits a *Puiseux development* at $t = \alpha$; that is, there exists $\delta > 0$, a positive integer k, $l \in \mathbb{Z}$, and $\{a_n\}_{n \geq l} \subset \mathbb{R}$ such that

$$\varphi(t) = \sum_{n \geq l} a_n (t - \alpha)^{n/k} \quad \text{ for all } t \in (\alpha, \alpha + \delta).$$

– *Łojasiewicz factorization lemma* (e.g., [4, Theorem 6.4]): Let $K \subset \mathbb{R}^n$ be a compact set and $f, g : K \to \mathbb{R}$ two continuous (globally) subanalytic functions. If $f^{-1}(0) \subset g^{-1}(0)$, then there exist $c > 0$ and a positive integer r such that $|g(x)|^r \leq c\,|f(x)|$ for all $x \in K$.

2.2. Notions from nonsmooth analysis and further stability results. Throughout this paper, we essentially deal with nondifferentiable functions defined on $\mathbb{R}^n$ with values in $\mathbb{R} \cup \{+\infty\}$. We denote by $\operatorname{dom} f$ the domain of the function, that is, the subset of $\mathbb{R}^n$ on which f is finite. In a similar way, the domain of a point-to-set operator $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, denoted by $\operatorname{dom} T$, is defined as the subset of $\mathbb{R}^n$ on which T is nonempty. The epigraph and the strict epigraph of f are respectively defined by

$$\operatorname{epi} f := \{(x, \lambda) \in \mathbb{R}^n \times \mathbb{R} : \lambda \geq f(x)\}, \qquad \operatorname{epi}_s f := \{(x, \lambda) \in \mathbb{R}^n \times \mathbb{R} : \lambda > f(x)\},$$

while the *epigraphical sum* of two extended-real-valued functions $f, g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is the function defined by

$$\mathbb{R}^n \ni u \longmapsto h(u) \,=\, \inf \{f(v) + g(v - u) : v \in \mathbb{R}^n\} \in [-\infty, +\infty].$$

The terminology stems from the fact that the strict epigraph of h is the Minkowski sum of the strict epigraphs of f and g.

Even if $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is subanalytic, its domain and its epigraph may fail to be subanalytic sets.

Example 2.4. Consider the function $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ whose graph is given by the set $S := \{(\frac{1}{n}, n)\}$. Then the domain of f is not subanalytic, whereas its graph is. If $g : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ has $-S := \{(-\frac{1}{n}, -n) : n \in \mathbb{N}\}$ as its graph, both its domain and epigraph are not subanalytic.

Additional geometrical properties like convexity are also not sufficient to obtain regularity on the domain. This is shown in the example below.

Example 2.5. Let $\{q_n\}_{n \geq 1}$ be an enumeration of the rationals $\{q_n\}$, and define $h : \mathbb{R}^2 \to \mathbb{R} \cup \{+\infty\}$ in polar coordinates by

$$h(r, \theta) = \begin{cases} 0 & \text{if } r \in [0, 1), \\ n & \text{if } r = 1 \text{ and } \theta = q_n \pmod{2\pi}, \\ +\infty & \text{otherwise.} \end{cases}$$

Then h is convex and subanalytic, but its domain is not subanalytic.

As expected, such a behavior can be avoided by requiring the function to be globally subanalytic. The following two results are basic consequences of the projection theorem.

PROPOSITION 2.6. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a globally subanalytic function. Then the domain, the epigraph, and the strict epigraph of f are globally subanalytic sets.*

PROPOSITION 2.7. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a subanalytic function which is relatively bounded on its domain; that is, $\{f(x) : x \in dom\ f \cap B\}$ is bounded for every bounded subset B of $\mathbb{R}^n$. Then the domain, the epigraph, and the strict epigraph of f are subanalytic sets.*

Remark 2.8. Observe that Propositions 2.6 and 2.7 involve distinct assumptions and provide different results. This can be seen by considering, for instance, the subanalytic functions $f(x) = x^{-1}$ with dom $f = (0, +\infty)$ and $g := \delta_{\mathbb{N}}$.

The case for which functions under consideration are convex but not necessarily continuous requires more attention.

PROPOSITION 2.9. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous convex and subanalytic function such that $\inf_{\mathbb{R}^n} f \in \mathbb{R}$. Define $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ as the epigraphical sum of f and the square function $\frac{1}{2}||\cdot||^2$, that is,*

$$h(x) = \inf \left\{ f(u) + \frac{1}{2}||x - u||^2 : u \in \mathbb{R}^n \right\}, \quad x \in \mathbb{R}^n.$$

Then h is a C^1 subanalytic function.

Proof. The proof consists mainly of showing that the epigraphical sum of a convex function with a coercive function is a "graphically local" operation. The fact that h takes finite values and is a C^1 function is a classical result (see [20], for example). Therefore it suffices to prove that $h + \delta_B$ is subanalytic for every bounded subset B of $\mathbb{R}^n$. Let us fix some nonempty bounded set B of $\mathbb{R}^n$, and let us set $M = \sup\{h(x) : x \in B\}$. Thanks to the continuity of h we have $M < +\infty$.

The infimum in the definition of $h(x)$ is always attained at a unique point denoted $J(x)$, and the mapping $J : \mathbb{R}^n \to \mathbb{R}^n$ so defined is a nonexpansive mapping (see [6]). Moreover, the function f is bounded on the bounded set $J(B)$. Indeed, if $u = J(b)$ for some $b \in B$, the definition of J implies that

$$f(u) = f(J(b)) = h(b) - \frac{1}{2}||b - J(b)||^2 < M.$$

Let C be some ball containing the bounded set $J(B)$, and let $f^M : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be the function whose graph is given by Gr $f \cap (C \times [\inf_{\mathbb{R}^n} f, M])$. By definition the function f^M has a bounded subanalytic graph, and it is therefore globally subanalytic. According to the above considerations the values of h on B coincide with those of the function

$$\hat{h}(x) := \inf \left\{ f^M(u) + \frac{1}{2}||x - u||^2 : u \in \mathbb{R}^n \right\}, \quad x \in \mathbb{R}^n.$$

The strict epigraph of $\hat{h}$ is the sum of the strict epigraphs of the bounded subanalytic function f^M and the square function $u \mapsto \frac{1}{2}||x - u||^2$ (which is globally subanalytic for it is semialgebraic). This yields that $\hat{h}$ (and consequently $h + \delta_B$) is globally subanalytic; hence h is subanalytic. ◻

The notion of subdifferential—that is, an appropriate multivalued operator playing the role of the usual gradient mapping—is crucial for our considerations. In what follows we denote by $\langle \cdot, \cdot \rangle$ the usual Euclidean product of $\mathbb{R}^n$.

DEFINITION 2.10 (subdifferential; see, e.g., [20, Definition 8.3]). (i) *The* Fréchet subdifferential $\hat{\partial} f(x)$ *of a lower semicontinuous function f at $x \in \mathbb{R}^n$ is given by*

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

whenever $x \in dom\ f$, and by $\hat{\partial} f(x) = \emptyset$ otherwise.

(ii) *The* limiting subdifferential *at $x \in \mathbb{R}^n$, denoted by $\partial f(x)$, is the set of all cluster points of sequences $\{x_n^*\}_{n\geq 1}$ such that $x_n^* \in \hat{\partial} f(x_n)$ and $(x_n, f(x_n)) \to (x, f(x))$ as $n \to +\infty$.*

If the function f is of class C^1, the above notion coincides with the usual concept of gradient; that is, $\partial f(x) = \hat{\partial} f(x) = \{\nabla f(x)\}$. For a general lower semicontinuous function, the limiting subdifferential $\partial f(x)$ (thus, a fortiori the Fréchet subdifferential $\hat{\partial} f(x)$) can possibly be empty at several points $x \in \operatorname{dom} f$. Nevertheless (see, e.g., [20, Chapter 8]), both the domain of $\hat{\partial} f$ and (a fortiori) the domain of ∂f are dense in the domain of f.

Using the limiting subdifferential ∂f, we define the *nonsmooth slope* of f by

$$m_f(x) := \inf\{\|x^*\| : x^* \in \partial f(x)\}. \tag{4}$$

By definition, $m_f(x) = +\infty$ whenever $\partial f(x) = \emptyset$.

Let us recall that if f is continuous, the operator $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ has a closed graph. This is also the case for a lower semicontinuous convex function, where both $\partial f(x)$ and $\hat{\partial} f(x)$ coincide with the classical subdifferential of convex analysis; that is,

$$\partial f(x) = \hat{\partial} f(x) = \{x^* \in \mathbb{R}^n : f(\cdot) - \langle x^*, \cdot\rangle \text{ has a global minimum at } x\}\,. \tag{5}$$

We are ready to state the notion of generalized critical point (in the sense of variational analysis).

DEFINITION 2.11 (critical point). *A point $a \in \mathbb{R}^n$ is said to be a (generalized)* critical point *of the function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ if it belongs to the set*

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

Remark 2.12. If f is lower semicontinuous convex or if $\operatorname{dom} f$ is closed and $f\,|_{\operatorname{dom} f}$ is continuous, then the graph of ∂f is closed, which implies that the set $\operatorname{crit} f$ of the critical points of f is closed. In that case, let us also observe that the slope $m_f(x)$ is a lower semicontinuous function, and that

$$\operatorname{crit} f = m_f^{-1}(0).$$

The following result illustrates further the properties of stability of subanalytic sets recalled in subsections 2.1 and 2.2.

PROPOSITION 2.13. *Let f be an extended-real-valued function.*

(i) *If f is globally subanalytic, then the operators $\hat{\partial} f$ and ∂f, the function m_f, and the set* $\operatorname{crit} f$ *are globally subanalytic.*

(ii) *If f is subanalytic and relatively bounded on its domain, then the operators $\hat{\partial} f$ and ∂f, the function m_f, and the set* $\operatorname{crit} f$ *are subanalytic.*

Proof. The local nature of the Fréchet and the limiting subdifferential allows us to restrict our proof to the globally subanalytic function $f_B := f + \delta_B$, where B is some nontrivial ball. It suffices therefore to establish (i).

Thanks to the projection theorem (Theorem 2.3), the proof becomes a routine application of [8, Theorem 1.13], which asserts that if $\Phi(x_1, \dots, x_n)$ is a first order formula (in the language of the subanalytic structure of $\mathbb{R}^n$), then the set $\{(x_1, \dots, x_n) \in \mathbb{R}^n : \Phi(x_1, \dots, x_n)\}$ is definable, or in other words, it belongs to the structure.[1]

[1]Global subanalytic sets form a model-complete first order theory. In fact, whether or not a structure is "model complete" depends only on the theory of the structure, that is, the set of the sentences (i.e., quantifier-free formulas) of its language which are true in this theory. We refer to [23, p. 1052] for more details.

As an illustration of this standard technique, let us prove that the operator $\hat{\partial} f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is globally subanalytic. To this end, set $A = \operatorname{epi} f$, $\Gamma = \operatorname{Gr} f$, and $D = \operatorname{dom} f$, which are all globally subanalytic sets. According to Definition 2.10(ii) the graph $\operatorname{Gr} \hat{\partial} f$ of the Fréchet subdifferential $\hat{\partial} f(x)$ is the set of $(x, x^*) \in \mathbb{R}^n \times \mathbb{R}^n$ such that

$$\forall \varepsilon > 0, \exists \delta > 0, \forall (y, \beta) \in (B(x, \delta) \times \mathbb{R}) \cap A \Rightarrow (x, \beta - \langle x^*, y - x \rangle + \varepsilon \|y - x\|) \in A,$$

where $B(x, \delta)$ denotes the open ball of center x and radius $\delta > 0$. Since the above first order formula involves only globally subanalytic sets (namely, the subanalytic sets $B(x, \delta), \mathbb{R}$, and A), it follows that $\operatorname{Gr} \hat{\partial} f$ is subanalytic.

Subanalyticity of the graphs of the operator ∂f and of the function m_f can be proved similarly. Finally, $\operatorname{crit} f$ being the inverse image of (the subanalytic set) $\{0\}$ by m_f, it is a subanalytic set. $\square$

Similarly one obtains the following corollary.

COROLLARY 2.14. *Under the assumptions of Proposition* 2.13(ii), *the restrictions of the multivalued mappings $\hat{\partial} f$, ∂f, and of the slope function m_f to any bounded subanalytic subset of $\mathbb{R}^n$ are globally subanalytic.*

Remark 2.15. The assumptions (and consequently the results) of the statements (i) and (ii) of Proposition 2.13 are of different natures. For example, let us consider the lower semicontinuous convex function $f : \mathbb{R}^2 \to \mathbb{R} \cup \{+\infty\}$, defined by

$$f(x, y) = \begin{cases} x^2/y & \text{if } y > 0, \\ 0 & \text{if } x = y = 0, \\ +\infty & \text{elsewhere.} \end{cases}$$

Then Proposition 2.13(i) applies, but not (ii), since f is not relatively bounded on $\operatorname{dom} f$.

3. Main results.

3.1. The Łojasiewicz inequality for subanalytic continuous functions.

Assuming f subanalytic, and having a closed domain relative to which it is continuous, the set $\operatorname{crit} f$ is closed (Remark 2.12) and subanalytic (Proposition 2.13), so it has a locally finite number of connected components (see subsection 2.1). For any a in $\operatorname{crit} f$, let us denote by $(\operatorname{crit} f)_a$ the connected component of $\operatorname{crit} f$ containing a. In [5, Theorem 13] it has been established that

$$f \text{ is constant on } (\operatorname{crit} f)_a . \tag{6}$$

The proof of (6) relies on a fundamental structural result about subanalytic functions (stratification) and on the Pawłucki generalization of the Puiseux lemma; see [5]. Nevertheless, (6) can be easily proved for continuous functions that also satisfy

$$\hat{\partial} f(x) = \partial f(x) \quad \text{for all } x \in \mathbb{R}^n. \tag{7}$$

Indeed, given x and y in some connected component S_i of $\operatorname{crit} f$, we consider the continuous subanalytic path $z : [0, 1] \to S_i$ with $z(0) = x$ and $z(1) = y$, and the subanalytic function $h(t) = (f \circ z)(t)$ (see subsection 2.1). Since $0 \in \hat{\partial} f(z(t))$ for all $t \in [0, 1]$, from the "monotonicity lemma" and the chain rule for the Fréchet subdifferential [20, Theorem 10.6] we get $h'(t) = 0$ for almost all t. It follows that h is constant on $[0, 1]$, whence $f(x) = f(y)$.

Examples of continuous functions that satisfy (7) are C^1 functions (for which $\partial f(x) = \hat{\partial} f(x) = \{\nabla f(x)\}$), proximal retracts (or lower-C^2 functions; see [20, Definition 10.29] and section 4), or more generally subdifferentially regular functions [20, Definition 7.25].

The main result of subsection 3.1 can now be stated as follows.

THEOREM 3.1. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a subanalytic function with closed domain, and assume that $f\,|_{\mathrm{dom} f}$ is continuous. Let $a \in \mathbb{R}^n$ be a critical point of f. Then there exists an exponent $\theta \in [0, 1)$ such that the function*

$$\frac{|f - f(a)|^{\theta}}{m_f} \tag{8}$$

is bounded around a.

Note that we have adopted here the following conventions: $0^0 = 1$ and $\infty/\infty = 0/0 = 0$.

Proof. Let us set $S = \operatorname{crit} f$ and $S_a = (\operatorname{crit} f)_a$. Replacing if necessary f by $g(x) = f(x) - f(a)$, there is no loss of generality to assume $f(a) = 0$, so that (6) implies $S_a \subset f^{-1}(0)$.

We may also assume that f is globally subanalytic and that the set S_a is compact. Indeed, if this is not the case, then we replace the function f by the globally subanalytic function g defined (for some $R > 0$) by $g(x) = f(x) + \delta_{\bar{B}(a,R)}(x)$, where $\delta_{\bar{B}(a,R)}$ denotes the indicator function of the closed ball $\bar{B}(a, R)$. Then g has a closed domain relative to which it is continuous, a is a critical point for g, and $(\operatorname{crit} g)_a \cap B(a, R) = S_a \cap B(a, R)$. Establishing (8) for f is thus the equivalent of doing so for the globally subanalytic function g.

It is also sufficient to establish separately that the function $x \mapsto [m_f(x)]^{-1}\,|f(x)|^{\theta}$ is bounded when x varies inside the subanalytic set $f^{-1}((0, +\infty])$, and subsequently to do the same when x varies in $f^{-1}((-\infty, 0])$. Since this latter assertion will follow by reproducing essentially the same arguments, we may assume with no loss of generality that $f \geq 0$.

Let us choose $\Delta > 0$ so that the compact set $U = \{x \in \mathbb{R}^n : d_{S_a}(x) \leq \Delta\} \cap \mathrm{dom}\, f$ separates S_a from the other connected components of S. Note that U is a globally subanalytic set (see subsection 2.1). We claim that for all $\bar{x}$ in the boundary of S_a we have

$$\lim_{\substack{x \to \bar{x} \\ x \in U \setminus S_a}} \frac{f(x)}{m_f(x)} = 0. \tag{9}$$

If the above limit were not zero, there would exist a sequence $\{(x_p, x_p^*)\}$ in $\mathrm{Gr}\, \partial f$ and $r > 0$ with $x_p \to \bar{x}$ as $p \to +\infty$ and such that $f(x_p) > r||x_p^*|| > 0$ for all p. By the definition of the limiting subdifferential there exists a sequence $(y_p, y_p^*) \in \mathrm{Gr}\, \hat{\partial} f$ such that $f(y_p) > r||y_p^*|| > 0$, where y_p converges to $\bar{x}$. This proves that for some $r > 0$ the point $\bar{x}$ belongs to the closure of the set

$$F = \{x \in U \setminus S_a : \exists x^* \in \hat{\partial} f(x), f(x) > r\,||x^*|| > 0\}.$$

Owing to Proposition 2.13 (i) the latter set is globally subanalytic, so by the "curve selection lemma" (subsection 2.1) there exists an analytic curve $z : (-1, 1) \to \mathbb{R}^n$ with $z(0) = \bar{x}$ and $z((0, 1)) \subset F$. Hence for all small $t > 0$ there exists a nonzero subgradient $z^*(t) \in \hat{\partial} f(z(t))$ satisfying

$$f(z(t)) > r\,\|z^*(t)\| > 0. \tag{10}$$

Thanks to the continuity of $f\,|_{\mathrm{dom}f}$ at $\bar{x} = z(0)$ the subanalytic function

$$[0,1) \ni t \mapsto h(t) = (f \circ z)(t)$$

is continuous at $t = 0$, and (6) implies that $h(0) = f(\bar{x}) = 0$. Applying the "monotonicity lemma" (subsection 2.1) to the globally subanalytic function h and the chain rule calculus for the Fréchet subdifferential [20, Theorem 10.6], we get for t small enough that $|h'(t)| \leq M\,\|z^*(t)\|$, where $M = \max\{\|\dot{z}(t)\| : t \in (-1/2, 1/2)\}$. Then by applying (10), it follows that

$$\frac{h(t)}{|\,h'(t)\,|} > r\,M^{-1} > 0 \qquad \text{for all small } t > 0. \tag{11}$$

Considering the Puiseux development of h around $t = 0$ (see subsection 2.1), we conclude that for some positive rational q and some $c > 0$ we have $h(t) = ct^q + o(t^q)$ for all small $t > 0$. By differentiating the Puiseux development of h at $t = 0$ and substituting into (11), we obtain a contradiction.

Let us now establish (8). To this end, let us consider the globally subanalytic function

$$\varphi(t) = \inf\{m_f(x) : x \in U \cap f^{-1}(t)\} \qquad \text{if } t \in \mathbb{R}_+.$$

Clearly $\varphi(0) = 0$, while from the definition of U, it ensues that $0 < \varphi(t) \leq +\infty$ for all small $t > 0$. If for every $\delta > 0$ the function φ assumes at least one infinite value in the interval $(0, \delta)$, then the subanalyticity of dom φ guarantees that 0 is an isolated point in dom φ. In this case (8) holds trivially. We may thus assume that φ is finite around 0. Evoking again the "monotonicity lemma" (subsection 2.1), we deduce that

$$l = \lim_{t \to 0^+} \varphi \in [0, +\infty].$$

In case $l \neq 0$, equation (8) follows easily (with $\theta = 0$), so we may assume $l = 0$ and φ continuous. In this case, we consider the Puiseux expansion of φ, which has the form

$$\varphi(t) = \sum_{n=0}^{+\infty} a_n t^{\frac{n}{k}} \quad \text{for all small } t > 0,$$

where k is a positive integer. Let $n_0 \in \mathbb{N}^*$ be the first integer such that $a_{n_0} \neq 0$, and let us set $\eta = \frac{n_0}{k}$. Then

$$\varphi(t) = ct^\eta + o(t^\eta), \tag{12}$$

where $c := a_{n_0} > 0$. Unless (8) holds trivially, we may assume by (6) that there exists a sequence $\{x_\nu\}_\nu \subset U \diagdown S_a$ such that $x_\nu \to a$, $m_f(x_\nu) \to 0$, and $f(x_\nu) \geq 0$. Let us consider the globally subanalytic set

$$A = \{x \in U \diagdown S_a : m_f(x) = \varphi(f(x)),\ f(x) \geq 0\} \neq \emptyset.$$

We claim that

$$\mathrm{cl}\, A \cap S_a \neq \emptyset. \tag{13}$$

Indeed, if (13) were not true, then by a standard compactness argument, there would exist an open neighborhood V around S_a such that $S_a \subset V \cap \mathrm{dom}\, f \subset U$

and $A \cap V = \emptyset$. Setting $t_\nu = f(x_\nu)$ (for the sequence $\{x_\nu\}_\nu$ mentioned above) and considering $y_\nu \in U$ such that $m_f(y_\nu) = \varphi(t_\nu)$ (by Remark 2.12, m_f is lower semicontinuous) and $f(y_\nu) = t_\nu$, we would obtain $\{y_\nu\}_\nu \subset U \diagdown V$. By compactness, we could then assume that $y_\nu \to y \in U \diagdown V$, which would yield (by continuity of φ) that $m_f(y) = 0$, that is, $y \in S_a$, and a contradiction follows.

Thus (13) holds, and there exists an analytic curve $z : (-1, 1) \to \mathbb{R}^n$ with $z(0) := b \in S_a$ and $z((0, 1)) \subset A$. As $s \searrow 0^+$ we get (by continuity of f and φ) that $f(z(s)) \to 0$ and $m_f(z(s)) = \varphi(f(z(s))) \to 0$. We deduce from (12) that

$$m_f(z(s)) = c(f(z(s)))^\eta + o((f(z(s))^\eta),$$

so (9) implies that $\eta < 1$. Take $\theta \in (\eta, 1)$ and apply (12) to obtain the existence of $t_0 > 0$ such that $\varphi(t) \geq ct^\theta$ for all $t \in [0, t_0)$. By using the continuity of $f\,|_{\operatorname{dom} f}$ at a, it follows that there exists $\mu > 0$ such that $|f(x)| < t_0$ for all $x \in \operatorname{dom} f \cap B(a, \mu)$. Finally, to obtain (8), we simply observe that

$$m_f(x) \geq \varphi(f(x)) \geq cf(x)^\theta \quad \text{for all } x \in B(a, \mu).$$

The proof is complete. □

Remark 3.2. Let us note that (8) still holds around any point $a \in \operatorname{dom} f \diagdown \operatorname{crit} f$. Indeed, if $a \notin \operatorname{crit} f$, then $m_f(x)$ is bounded below away from 0 in a neighborhood of a, so (8) follows from the continuity of f. In this case, the assumption of subanalyticity is obviously not needed.

3.2. The Łojasiewicz inequality for subanalytic lower semicontinuous convex functions. In this subsection we are interested in lower semicontinuous convex subanalytic functions $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ which are somewhere finite, that is, convex functions for which $\operatorname{dom} f \neq \emptyset$. In this case, in view of (5), the set of critical points $\operatorname{crit} f$ is closed and convex and coincides with the set of minimizers of f.

Before proceeding let us recall classical facts from convex analysis (e.g., [20]). Let us denote by g the epigraphical sum of f and $\frac{1}{2}\|\cdot\|^2$ (see Proposition 2.9). The function $g : \mathbb{R}^n \to \mathbb{R}$ is finite-valued, and C^1 and enjoys the following properties:

(a) $g \leq f$.

(b) The set of critical points of g is exactly the set of critical points of f.

(c) The infimum values of f and g coincide; i.e., $\inf_{\mathbb{R}^n} f = \inf_{\mathbb{R}^n} g$.

The properties of g are related to the so-called Moreau regularizing process; for more details and further results, see [20].

We are ready to state the main result of this subsection.

THEOREM 3.3. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous convex subanalytic function with* $\operatorname{crit} f \neq \emptyset$. *For any bounded set K there exists an exponent $\theta \in [0, 1)$ such that the function*

$$\frac{|f - \min f|^\theta}{m_f} \tag{14}$$

is bounded on K.

Proof. By Proposition 2.9, the function g defined above is subanalytic and continuous. Applying (b) and the results of the preceding section, we see that $S := \operatorname{crit} f$ is subanalytic. Let us show how g may be used to derive a growth condition for f. For any $x \in K$, the equivalence

$$d_S(x) = 0 \Longleftrightarrow |g(x) - \min g| = 0,$$

combined with the Łojasiewicz factorization lemma (subsection 2.1) for the continuous subanalytic functions $|g - \min g|$ and d_S (restricted to the bounded set K), yields the existence of $r > 1$ and $c > 0$ such that

$$c\,[d_S(x)]^r \le |g(x) - \min g| \quad \text{for all } x \in K.$$

On the other hand, the properties (a), (b), (c) of g imply that

$$|f(x) - \min f| \ge |g(x) - \min g| \quad \text{for all } x \in \mathbb{R}^n,$$

so that

$$[d_S(x)] \le c^{-1/r}\,|f(x) - \min f|^{1/r}. \tag{15}$$

Moreover, since f is convex we get for all a in S and all $(x, x^*) \in \operatorname{Gr}\partial f$

$$f(a) \ge f(x) + \langle x^*, a - x\rangle.$$

Thus for all $(x, x^*) \in \operatorname{Gr}\partial f$ it follows that $|f(x) - f(a)| \le \|x^*\|\,\|x - a\|$, and by taking the infimum over all $a \in S$, we obtain

$$|f(x) - \min f| \le \|x^*\|\,d_S(x). \tag{16}$$

We therefore deduce from (15) that for all $x \in K$ and all $(x, x^*) \in \operatorname{Gr}\partial f$

$$|f(x) - \min f| \le c^{-1/r}\,\|x^*\| \cdot |f(x) - \min f|^{1/r}.$$

By setting $\theta = 1 - r^{-1}$, the latter inequality implies $|f(x) - \min f|^{\theta} \le c^{-1/r}\,m_f(x)$ for all $x \in K$, and (14) follows. □

Remark 3.4. The lower semicontinuous convex function f considered in Remark 2.15 provides an example where Theorem 3.3 applies while Theorem 3.1 does not.

Remark 3.5. A careful examination of the proof of Theorem 3.3 shows that the important assumption is not subanalyticity of the function, but rather the growth condition near critical values that subanalyticity implies. Indeed, let K be a compact set and f be any lower semicontinuous convex function f that satisfies

$$|f(x) - \min f| \ge c\,d_S(x)^r \qquad \text{for all } x \in K, \tag{17}$$

where $c > 0$, $r \ge 1$ and with $S = \operatorname{crit} f \ne \emptyset$. The argument of Theorem 3.3 may be then slightly modified in order to derive a Łojasiewicz inequality around any critical point a belonging to the interior of K.

Remark 3.6. From relation (16), which is true for all lower semicontinuous convex functions, a weaker version of (14) can be deduced. Indeed, if f is convex (but not necessarily subanalytic), then the function

$$\frac{|f - \min f|}{m_f}$$

is bounded around any critical point of f.

Remark 3.7. By using elementary arguments it can be shown that f satisfies the Łojasiewicz inequality around any point $a \in \operatorname{dom} f$ (cf. Remark 3.2).

4. Applications to dynamical systems. Throughout this section, unless otherwise stated, we make the following assumptions:

($\mathcal{H}1$) f is either lower semicontinuous convex or lower-C^2 with $\operatorname{dom} f = \mathbb{R}^n$.

($\mathcal{H}2$) f is somewhere finite ($\operatorname{dom} f \neq \emptyset$) and bounded from below.

We recall (see [20, Definition 10.29], for example) that a function f is called *lower-C^2* if for every $x_0 \in \operatorname{dom} f$ there exist a neighborhood U of x_0, a compact topological space S, and a jointly continuous function $F : U \times S \to \mathbb{R}$ satisfying $f(x) = \max_{s \in S} F(x, s)$ for all $x \in U$ and such that the (partial) derivatives $\nabla_x F(\cdot, \cdot)$ and $\nabla_x^2 F(\cdot, \cdot)$ exist and are jointly continuous.

A lower-C^2 function f is locally Lipschitz and locally representable as a difference of a convex continuous and a convex quadratic function [20, Theorem 10.33]. In particular, it satisfies

$$\partial f = \hat{\partial} f. \tag{18}$$

Note that (18) is also true for a lower semicontinuous convex function (see relation (5)).

As mentioned in the introduction, an important motivation for establishing the Łojasiewicz inequality for classes of nonsmooth functions is the expected asymptotic properties of the corresponding subgradient dynamical systems. This latter term refers to differential inclusions of the form

$$\dot{x}(t) + \partial f(x(t)) \ni 0,$$

where $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is the limiting subdifferential of f. A *trajectory* of the above dynamical system is any absolutely continuous curve $x : [0, T) \to \mathbb{R}^n$ that satisfies

$$(\mathcal{G}) \qquad \begin{cases} \dot{x}(t) + \partial f(x(t)) \ni 0 \quad \text{a.e. on } (0, T), \\ \\ \partial f(x(t)) \neq \emptyset \quad \text{for all } t \in [0, T), \end{cases}$$

where the notation "a.e." stands for "almost everywhere" in the sense of the Lebesgue measure of $\mathbb{R}$. Let us also recall that an absolutely continuous function (or curve) $x(t)$ is a.e. differentiable and can be entirely determined, up to a constant, by integration of its classical derivative. A trajectory $x(t)$ is called *maximal* if there is no possible extension of its domain compatible with ($\mathcal{G}$).

The following existence-uniqueness result is known to hold (see [6, Theorem 3.2, p. 57] or [2, Chapter 3.4] for the convex case, and [6, Proposition 3.12, p. 106] for the convex case with Lipschitz perturbation; see also [9] for related work).

Existence of trajectories. Under the assumptions ($\mathcal{H}1$) and ($\mathcal{H}2$), for every $x_0 \in \mathbb{R}^n$ such that $\partial f(x_0) \neq \emptyset$, there exists a unique trajectory $x : [0, T) \to \mathbb{R}^n$ of ($\mathcal{G}$) satisfying

$$(\mathcal{T}) \qquad x(0) = x_0.$$

In addition, the function $h := f \circ x$ is absolutely continuous.

Let us now recall some classical consequences of (18) and of the above existence result. For the sake of completeness, some elementary proofs are provided.

Corollary 4.1. *Let $x : [0, T) \to \mathbb{R}^n$ be a trajectory of $(\mathcal{G})$ satisfying $(\mathcal{T})$.*
(i) *For almost all $t \in (0, T)$*

$$\frac{d}{dt}(f \circ x)(t) = \langle \dot{x}(t), x^* \rangle \quad \text{for all } x^* \in \partial f(x(t)).$$

(ii) *For almost all $t \in (0, T)$, the function $x^* \mapsto \langle \dot{x}(t), x^* \rangle$ is constant on $\partial f(x(t))$.*

(iii) *The trajectory x can be extended to a maximal trajectory $\hat{x} \in W^{1,2}(\mathbb{R}_+ ; \mathbb{R}^n)$.*

Proof. Set $h = f \circ x$ and note that the absolutely continuous functions h and x are simultaneously differentiable on $(0, T) \diagdown N$, where N is a set of measure zero. Let $t \in (0, T) \diagdown N$. Since $x(t) \in \operatorname{dom} \partial f$ and $\partial f(x(t)) = \hat{\partial} f(x(t))$, one may adapt the ideas of [6, Lemma 3.3, p. 73] (chain rule) to obtain

$$\partial h(t) = \{h'(t)\} = \left\{ \frac{d}{dt}(f \circ x)(t) \right\} = \{ \langle \dot{x}(t),\, x^* \rangle,\ x^* \in \partial f(x(t)) \}.$$

Thus (i) and (ii) follow.

To establish (iii) let us first prove that $x \in W^{1,2}\left((0, T) ; \mathbb{R}^n\right)$. Thanks to $(\mathcal{G})$, we deduce from (i) that

$$\frac{d}{dt}(f \circ x)(t) = -\|\dot{x}(t)\|^2 \quad \text{for all } (0, T).$$

Hence f is a Lyapunov function of the dynamical system $(\mathcal{G})$, and

$$\int_0^T \|\dot{x}(t)\|^2 dt = f(x_0) - f(x(T)) < +\infty;$$

that is, $\dot{x} \in L^2((0, T) ; \mathbb{R}^n)$. Note that $\dot{x}(t)$ remains bounded as t converges to T. (For a lower semicontinuous convex function f this is a classical result (see [2, p. 147], for example); if f is lower-C^2, this follows from $(\mathcal{G})$ and the fact that ∂f is locally bounded around T.) Since the graph of ∂f is closed (Remark 2.12) we get $x(T) \in \operatorname{dom} \partial f$. Thus, thanks to the existence result $(\mathcal{T})$, the initial trajectory is in fact extendible to a semiopen interval $[0, T + \delta)$, for some $\delta > 0$, containing $[0, T]$. A standard argument shows that the maximal extension of $x(t)$ is defined in $(0, +\infty)$. ∎

An interesting hidden property of $(\mathcal{G})$ is the following.

Corollary 4.2. *Let x be a maximal trajectory of $(\mathcal{G})$ satisfying $(\mathcal{T})$. Then for almost all $t \in \mathbb{R}_+$*

$$\|\dot{x}(t)\| = m_f(x(t)) \qquad \textit{and} \qquad \frac{d}{dt}(f \circ x)(t) = -\left[m_f(x(t))\right]^2.$$

Proof. From $(\mathcal{G})$, we obtain the existence of a curve $t \mapsto g(t) \in \partial f(x(t))$ such that

$$\dot{x}(t) = -g(t) \quad \text{a.e. on } \mathbb{R}_+.$$

Combining this with Corollary 4.1(ii), we get that for almost all t in $\mathbb{R}_+$

$$\|g(t)\|^2 = \langle g(t),\, x^* \rangle \quad \text{for all } x^* \in \partial f(x(t)),$$

which yields via a standard argument that $\|g(t)\| = m_f(x(t))$. Now evoking Corollary 4.1(i) finishes the proof. ∎

Remark 4.3. Corollary 4.2 says that the trajectories of $(\mathcal{G})$ (the existence of which is guaranteed under the assumptions $(\mathcal{H}1)$ and $(\mathcal{H}2)$) are necessarily “slow solutions” (see [2, p. 139]) of the differential inclusion ($\mathcal{G}$). In particular, if the trajectory $x(t)$ meets a critical point of f, that is, if there exists $t_0 > 0$ such that $m_f(x(t_0)) = 0$, then Corollary 4.2 guarantees that the trajectory stops there; that is, $x(t) = x(t_0)$ for all $t \geq t_0$. In this case, the trajectory has a finite length equal to $\int_0^{t_0} \|\dot{x}(s)\| ds$.

Another consequence of Corollary 4.2 is that ($\mathcal{G}$) defines a descent method in the sense that f decreases along any trajectory. Although compactness implies that bounded trajectories have at least one cluster point as $t \to +\infty$, those might not converge towards one of them—and a fortiori, have an infinite length. The next result shows that this cannot happen if f is assumed subanalytic (or more generally, if f satisfies the Łojasiewicz inequality). Indeed, via a "Łojasiewicz-type" argument (e.g., [14]) we establish successively that the tail of the trajectory is trapped inside a convenient ball of its cluster point, that this tail necessarily has a finite length, and finally that the trajectory converges to this cluster point. In the remainder, in addition to ($\mathcal{H}1$) and ($\mathcal{H}2$), the following is also assumed:

($\mathcal{H}3$) f is a *subanalytic* function.

Let us give some examples of subanalytic functions related with optimization problems.

Example 4.4. - (supremum operations) Let $g : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$ be an analytic function, and let K be a compact subanalytic subset of $\mathbb{R}^p$. Then

$$f(x) = \sup_{y \in K} g(x, y)$$

is a lower-C^2 subanalytic function (see [4], for example). If in addition $x \mapsto g(x, y)$ is convex for all y, then f is convex.

- (constraints sets) Let $g_i : \mathbb{R}^n \to \mathbb{R}, i \in \{1, \ldots, m\}$, be a family of analytic functions. The feasible set

$$C := \{x \in \mathbb{R}^n : g_i(x) \leq 0, \, \forall i \in \{1, \ldots, m\}\, \}$$

together with its indicator function are subanalytic objects.

- (Barrier and penalty functions) Those can be used to minimize convex functions via parametric versions of ($\mathcal{G}$). Typical examples on $\mathbb{R}$ are the functions $h_1 : x > 0 \mapsto x^{-p}\,(p \geq 1)$, $h_2 : x \geq 0 \mapsto -x^{\nu}$, $(\nu \in (0, 1))$, $h_3(x) = x^2$ if $x \leq 0$ and $h_3(x) = 0$ otherwise.

We are now ready to state the following result.

THEOREM 4.5. *Assume that a function f satisfies* ($\mathcal{H}1$)–($\mathcal{H}3$). *Then any bounded maximal trajectory of* ($\mathcal{G}$) *has a finite length and converges to some critical point of f.*

Proof. Let $\{x(t)\}_{t \geq 0}$ be a bounded maximal trajectory of ($\mathcal{G}$). By Corollary 4.1, the trajectory is defined over all $\mathbb{R}_+$. Using ($\mathcal{H}2$) and Corollary 4.2(iii), we conclude that there exists $\beta \in \mathbb{R}$ such that $\lim_{t \to +\infty} f(x(t)) = \beta$. Replacing f by $f - \beta$ and using the basic rules of subdifferential calculus, we may assume that $\beta = 0$.

In view of Remark 4.3, we may also assume that $f(x(t)) \neq 0$ for all $t > 0$. Consequently, the function $t \mapsto (f \circ x)(t)$ is positive and strictly decreasing to 0 as $t \to +\infty$. Moreover, by compactness, there exists some cluster point $a \in \mathbb{R}^n$ for the trajectory $x(t)$. So there exists an increasing sequence $(t_n)_{n \geq 1}$ with $t_n \to +\infty$ such that

$$\lim_{t_n \to +\infty} x(t_n) = a. \tag{19}$$

By continuity of $(f \circ x)$ we deduce that $f(a) = 0$. Using $(\mathcal{T})$, (19), and the fact that ∂f has a closed graph (see Remark 2.12), we deduce that $a \in \operatorname{dom} \partial f$. We do not know yet whether a is critical or not, but nevertheless, the Łojasiewicz inequality holds around a. Indeed, if $a \in \operatorname{crit} f$, then use Theorem 3.1 or Theorem 3.3, and if $a \notin \operatorname{crit} f$, then just recall Remarks 3.2 and 3.7. It follows that there exist $c > 0$, $\theta \in [0, 1)$, and $\varepsilon > 0$ (defining an open neighborhood $B(a, \varepsilon)$ of a) such that

$$|f(x)|^{\theta} \leq c\, m_f(x) \quad \text{for all } x \in B(a, \varepsilon). \tag{20}$$

Let us consider the (positive, absolutely continuous) function $\tilde{h} = (f \circ x)^{1-\theta}$. Since $x(t) \to a$ and since the function $\tilde{h}$ is strictly decreasing and converges to 0 (as $t \to +\infty$), there exists $t_0 > 0$ such that for all $t \geq t_0$

$$\frac{|\tilde{h}(t) - \tilde{h}(t_0)|}{c^{-1}(1-\theta)} \leq \frac{\varepsilon}{3}, \tag{21}$$

with

$$\|x(t_0) - a\| \leq \frac{\varepsilon}{3}. \tag{22}$$

Let us set

$$T_{\text{out}} := \inf \{\, t \geq t_0,\, x(t) \notin B(a, \varepsilon) \,\}. \tag{23}$$

By continuity of the trajectory we have $t_0 < T_{\text{out}} \leq +\infty$.

Claim $T_{\text{out}} = +\infty$ (that is, the tail of the trajectory remains trapped in $B(a, \varepsilon)$).

Proof of the claim. For almost all $t \in [t_0, T_{\text{out}})$ we have

$$\begin{aligned} \frac{d}{dt}\tilde{h}(t) &= (1-\theta)\, f(x(t))^{-\theta}\, \frac{d}{dt}(f \circ x)(t) \\ &\leq -(1-\theta)\, f(x(t))^{-\theta}\, [m_f(x(t))]^2 \\ &\leq -(1-\theta)\, c^{-1}\, m_f(x(t)), \end{aligned}$$

where we have successively used Corollary 4.2 and (20). By integration, we obtain for all $t \in [t_0, T_{\text{out}})$

$$\int_{t_0}^{t} m_f(x(s))ds \leq -\left[\frac{\tilde{h}(t) - \tilde{h}(t_0)}{c^{-1}(1-\theta)}\right], \tag{24}$$

which according to (21) and Corollary 4.2, yields

$$\int_{t_0}^{t} \|\dot{x}(s)\| ds \leq \frac{\varepsilon}{3} \quad \text{ for all } t \in [t_0, T_{\text{out}}]. \tag{25}$$

To see that $T_{\text{out}} = +\infty$, we just argue by contradiction. If $T_{\text{out}} < +\infty$, then using (22) and (25), we obtain

$$\|x(T_{\text{out}}) - a\| \leq \left| \left(x(t_0) + \int_{t_0}^{T_{\text{out}}} \|\dot{x}(s)\| ds \right) - a \right| \leq \frac{2\varepsilon}{3}.$$

The latter obviously contradicts (23). Thus $T_{\text{out}} = +\infty$, and the claim is proved.

Resorting to (25) again, we conclude that $\int_{t_0}^{+\infty} \|\dot{x}(s)\| ds \leq \frac{\varepsilon}{3}$, so $x(t)$ has a finite length and hence converges. Thus $\lim_{t\to+\infty} x(t) = a$, and $m_f(x(t))$ admits 0 as a limit point. By using the closedness of $\operatorname{Gr}\partial f$, we conclude that a is a critical point of f. $\square$

Remark 4.6 (generalized gradient conjecture). The "gradient conjecture" of Thom [22] can obviously be reformulated in this nonsmooth setting. For any bounded trajectory $x(\cdot)$ of $(\mathcal{G})$, let us set $x_\infty := \lim_{t\to+\infty} x(t)$. Is it true that

$$t \mapsto \frac{x(t) - x_\infty}{||x(t) - x_\infty||}$$

has a limit as t goes to infinity? For real-analytic functions this conjecture has been proved by Kurdyka, Mostowski, and Parusiński [14].

Before we proceed to an estimate of the rate of convergence, let us introduce some terminology.

- We define

$$\sigma(t) = \int_t^{+\infty} \|\dot{x}(s)\| ds \quad \text{for all } t \in \mathbb{R}_+ \tag{26}$$

  to be the tail length function for the trajectory $x(t)$.
- A *Łojasiewicz exponent* of the function f at a point $a \in \mathbb{R}^n$ of its domain is any number $\theta \in [0, 1)$ for which the Łojasiewicz inequality holds around a.

Let us finally point out some facts arising from the proof of Theorem 4.5. Replacing $\tilde{h}(t)$ by $[f(x(t)]^{1-\theta}$ and $m_f(x(s))$ by $\|\dot{x}(s)\|$ (see Corollary 4.2) in (24) and letting $t \to +\infty$, we deduce

$$\int_{t_0}^{+\infty} \|\dot{x}(s)\| ds \leq \frac{c}{(1-\theta)} f(x(t_0))^{1-\theta}.$$

The above inequality remains true for every $t \geq t_0$ (in view of the Claim). Thus assuming $\theta > 0$ and evoking again (20) and Corollary 4.2, we obtain (for $k = c^{1/\theta}$)

$$\int_t^{+\infty} \|\dot{x}(s)\| ds \leq \frac{k}{(1-\theta)} \|\dot{x}(t)\|^{\frac{1-\theta}{\theta}} \qquad \text{for all } t \geq t_0. \tag{27}$$

We are now ready to state the following result.

THEOREM 4.7. *Under the assumptions $(\mathcal{H}1)$–$(\mathcal{H}3)$, let $x(t)$ be a bounded maximal trajectory of $(\mathcal{G})$. Then $x(t)$ converges to some critical point $a \in \mathbb{R}^n$ of f. Let $\theta \in [0, 1)$ be a Łojasiewicz exponent at this point. Then there exist $k > 0$, $k' > 0$, and $t_0 \geq 0$ such that for all $t \geq t_0$ the following estimates hold:*

– *If $\theta \in (\frac{1}{2}, 1)$, then $\|x(t) - a\| \leq k(t+1)^{-(\frac{1-\theta}{2\theta-1})}$.*
– *If $\theta = \frac{1}{2}$, then $\|x(t) - a\| \leq k \exp(-k't)$.*
– *If $\theta \in [0, \frac{1}{2})$, then $x(t)$ converges in finite time.*

Proof. We can always assume that $\theta > 0$. (If $\theta = 0$, we replace it by some $\theta' \in (0, 1/2)$, and we proceed as below.)

Let U be a neighborhood of a in which the Łojasiewicz inequality holds. Since $x(t)$ converges to a there exists $t_0 \geq 0$ such that $x(t) \in U$ for every $t \geq t_0$. In particular, (27) holds. Let us now consider the tail length function $\sigma(t)$ defined in (26). Note that

$$\|x(t) - a\| \leq \sigma(t). \tag{28}$$

Since $\dot{\sigma}(t) = -\|\dot{x}(s)\|$ for all $t \geq t_0$, inequality (27) yields

$$\sigma(t) \leq \frac{k}{(1-\theta)} \left[-\dot{\sigma}(t)\right]^{\frac{1-\theta}{\theta}}. \tag{29}$$

Thus $\sigma(t)$ is an absolutely continuous function and satisfies the following differential inequality:

$$\dot{\sigma}(t) \leq -L\left[\sigma(t)\right]^{\frac{\theta}{1-\theta}} \qquad \text{for all } t \geq t_0, \tag{30}$$

where L is a positive constant. To obtain the announced estimates it suffices to solve the following differential equation—considering separately the cases $\theta \in (1/2, 1)$, $\theta = 1/2$, and $\theta \in (0, 1/2)$:

$$\begin{cases} \dot{y}(t) = -L\left[y(t)\right]^{\frac{\theta}{1-\theta}} & \text{for all } t \geq t_0, \\ y(t_0) = \sigma(t_0). \end{cases} \tag{31}$$

The announced estimates then follow from (28) and the fact that $\sigma(t) \leq y(t)$ for all $t \geq t_0$. (Indeed, if $\sigma(\bar{t}) = y(\bar{t})$ for some $\bar{t} \geq t_0$, then a comparison of (30) and (31) shows that $\dot{\sigma}(\bar{t}) \leq \dot{y}(\bar{t})$.) The proof is complete. $\square$

Remark 4.8. The results of this section can be generalized to a wider setting as follows. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function complying with the following requirements:

(i) $\operatorname{dom} f \neq \emptyset$ and $\hat{\partial} f = \partial f$.

(ii) either f is convex or $f\,|_{\operatorname{dom} f}$ is continuous.

(iii) f has the Łojasiewicz property; that is, property (8) holds around any critical point.

If we assume in addition that, for all initial conditions $x_0 \in \operatorname{dom} \partial f$, the differential inclusion $(\mathcal{G})$ has a (unique) global solution x such that $f \circ x$ is absolutely continuous, then both Theorems 4.5 and 4.7 can be extended in this wider setting.

Prominent examples of functions meeting the above-mentioned conditions are continuous subanalytic ϕ-convex functions [9], or lower semicontinuous convex functions satisfying some growth condition of the type (17).

Acknowledgments. The second author wishes to thank K. Kurdyka, M. Quincampoix, and P. Cardaliaguet for useful discussions. A major part of this work was accomplished during a research visit of the first and the second authors to the Center for Experimental and Constructive Mathematics (Simon Fraser University, Canada). These two authors wish to thank their hosts at Simon Fraser University for their hospitality. The authors would also like to thank the anonymous referees for their useful comments.

REFERENCES

[1] P.-A. Absil, R. Mahony, and B. Andrews, *Convergence of the iterates of descent methods for analytic cost functions,* SIAM J. Optim., 16 (2006), pp. 531–547.

[2] J.-P. Aubin and A. Cellina, *Differential Inclusions*, Grundlehren Math. Wiss. 264, Springer, New York, 1984.

[3] R. Benedetti and J.-J. Risler, *Real Algebraic and Semialgebraic Sets*, Hermann, Paris, 1990.

[4] E. Bierstone and P. Milman, *Semianalytic and subanalytic sets*, IHES Publ. Math., 67 (1988) pp. 5–42.

[5] J. Bolte, A. Daniilidis, and A.S. Lewis, *A Sard theorem for non-differentiable functions,* J. Math. Anal. Appl, 321 (2006), pp 729–740.

[6] H. Brézis, *Opérateurs maximaux monotones et semi-groupes de contraction dans des espaces de Hilbert*, North-Holland Math. Stud. 5, North–Holland, Amsterdam, 1973.

[7] F.H. Clarke, Yu. Ledyaev, R.I. Stern, and P.R. Wolenski, *Nonsmooth Analysis and Control Theory*, Graduate Texts in Math. 178, Springer-Verlag, New York, 1998.

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

[9] M. Degiovanni, A. Marino, and M. Tosques, *Evolution equations with lack of convexity,* Nonlinear Anal., 9 (1985), pp 1401–1443.

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

[11] M. Fremond, J. Haslinger, J.-J. Moreau, P.M. Suquet, and J.J. Telega, *Nonsmooth Mechanics and Applications*, J.-J. Moreau and P.D. Panagiotopoulos, eds., CISM Courses and Lectures 302, Springer-Verlag, Vienna, 1988.

[12] A. Haraux, *A hyperbolic variant of Simon's convergence theorem*, in Evolution Equations and Their Applications in Physical and Life Sciences (Bad Herrenalb, Germany, 1998), Lecture Notes in Pure and Appl. Math. 215, Dekker, New York, 2001, pp. 255–264.

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

[14] K. Kurdyka, T. Mostowski, and A. Parusiński, *Proof of the gradient conjecture of R. Thom*, Ann. Math., 152 (2000), pp. 763–792.

[15] B. Lemaire, *The proximal algorithm,* in New Methods in Optimization and Their Industrial Uses (Pau/Paris, 1987), Internat. Schriftenreihe Numer. Math. 87, Birkhäuser, Basel, 1989, pp. 73–87.

[16] S. Łojasiewicz, *Une propriété topologique des sous-ensembles analytiques réels*, in Les Équations aux Dérivées Partielles, Éditions du Centre National de la Recherche Scientifique, Paris, 1963, pp. 87–89.

[17] S. Łojasiewicz, *Sur les trajectoires de gradient d'une fonction analytique*, Seminari di Geometria 1982-1983 (lecture notes), Dipartemento di Matematica, Universita di Bologna, 1984, pp. 115–117.

[18] S. Łojasiewicz, *Sur la géométrie semi- et sous-analytique,* Ann. Inst. Fourier, 43 (1993), pp. 1575–1595.

[19] J. Palis and W. De Melo, *Geometric Theory of Dynamical Systems. An Introduction,* Translated from the Portuguese by A.K. Manning, Springer-Verlag, New York/Berlin, 1982.

[20] R.T. Rockafellar and R. Wets, *Variational Analysis*, Grundlehren Math. Wiss. 317, Springer, New York, 1998.

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

[22] R. Thom, *Problèmes rencontrés dans mon parcours mathématique: Un bilan,* IHES Publ. Math., 70 (1989), pp. 199–214.

[23] A. 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 (1996), pp. 1051–1094.