Bull. Aust. Math. Soc. **93** (2016), 99–112
doi:10.1017/S0004972715000726

ŁOJASIEWICZ-TYPE INEQUALITIES AND GLOBAL ERROR BOUNDS FOR NONSMOOTH DEFINABLE FUNCTIONS IN O-MINIMAL STRUCTURES

DŨNG PHI HOÀNG

(Received 25 March 2015; accepted 4 April 2015; first published online 10 August 2015)

Abstract

In this paper, we give some Łojasiewicz-type inequalities for continuous definable functions in an o-minimal structure. We also give a necessary and sufficient condition for the existence of a global error bound and the relationship between the Palais–Smale condition and this global error bound. Moreover, we give a Łojasiewicz nonsmooth gradient inequality at infinity near the fibre for continuous definable functions in an o-minimal structure.

2010 *Mathematics subject classification*: primary 32B20; secondary 90C31, 49J52.

Keywords and phrases: Łojasiewicz inequalities, error bounds, o-minimal structures, nonsmooth analysis.

1. Introduction

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a real analytic function with $f(0) = 0$, $V := \{x \in \mathbb{R}^n \mid f(x) = 0\}$ and let K be a compact subset in $\mathbb{R}^n$. The (classical) Łojasiewicz inequality (see [27, 28]) asserts that there exist $c > 0$, $\alpha > 0$ such that

$$|f(x)| \geq c d(x, V)^\alpha \quad \text{for } x \in K. \tag{1.1}$$

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a real analytic function with $f(0) = 0$ and $\nabla f(0) = 0$. The Łojasiewicz gradient inequality (see [27, 28]) asserts that there exist $C > 0$, $\rho \in [0, 1)$ and a neighbourhood U of 0 such that

$$\|\nabla f(x)\| \geq C|f(x)|^\rho \quad \text{for } x \in U. \tag{1.2}$$

As a consequence of (1.1), the order of the zero of an analytic function is finite, and if $f(x)$ is close to 0 then x is close to the zero set of f. However, if K is not compact, the latter assertion is not always true and (1.1) does not always hold (see [9, Remark 3.5]). Similarly from (1.2), the order of the zero of the gradient of an analytic function is

This research was partially supported by Vietnam National Foundation for Science and Technology Development (NAFOSTED) grant 101.04-2013.07.

smaller than the order of its zero. But if U is not a bounded set, (1.2) does not always hold (see Example 3.8 in Section 3.4 below).

In the Łojasiewicz inequality (1.1), in the case $K = \mathbb{R}^n$, Hörmander (see [16]) replaced the left-hand side by a quantity greater than $|f(x)|$ to show

$$\exists c, \alpha, \beta > 0 \text{ such that } |f(x)|(1 + |x|^\beta) \geq c d(x, V)^\alpha \quad \forall x \in K.$$

In the recent papers [14] and [9], V is replaced by larger real algebraic sets, and versions of the Łojasiewicz inequalities are obtained in some noncompact cases.

Classical Łojasiewicz inequalities have many applications to dynamical systems, algebraic geometry and optimisation (see [16, 27, 28]). For example, the gradient (1.2) has an important relation with the classical gradient dynamical system $\dot{x}(t) = -\nabla f(x(t))$ (see [1, 2] and especially [4]).

Next we discuss connections with error bounds in optimisation. Error bounds have many applications, including convergence analysis in optimisation problems, variational inequalities and identifying active constraints (see [24, 25, 30]).

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous real-valued function. Set

$$S := \{x \in \mathbb{R}^n \mid f(x) \leq 0\}, \tag{1.3}$$

and set $[f(x)]_+ := \max\{0, f(x)\}$. We say that (1.3) has a global Hölderian error bound if there exist $c > 0$, $\alpha > 0$, $\beta > 0$ such that

$$d(x, S) \leq c([f(x)]_+^\alpha + [f(x)]_+^\beta) \tag{1.4}$$

for all $x \in \mathbb{R}^n$, where $d(x, S)$ denotes the Euclidean distance between x and S. If, in addition, $\alpha = \beta = 1$, then we refer to (1.4) as a global Lipschitzian error bound.

In the convex case, there are many examples of error bounds (see [3, 15, 19, 32, 33]). The existence of a (Lipschitzian) error bound usually requires convexity and the so-called Slater condition. When the Slater condition is not satisfied and the set S is defined by one or more polynomial inequalities, global Hölderian error bounds have been shown in [21, 29, 31, 36]. In the nonconvex case, a global Hölderian error bound for a polynomial of degree 2 was given in [31, Theorem 3.1], the first such result for a *nonconvex* polynomial. A global error bound with sharper exponent for a class of high degree nonconvex polynomials was obtained in [23]. Another nonconvex case with global Hölderian error bound was presented in [22].

A criterion for the existence of a global Hölderian error bound (1.4) for a polynomial of any degree is given in [13, Theorem A], without the assumption of convexity or the Slater condition. Moreover, it is shown that if a polynomial satisfies the Palais–Smale condition then there is a global Hölderian error bound.

In this paper, we will extend some results of [13] from polynomial functions to continuous definable functions in an o-minimal structure. We do not require the functions to be convex or to satisfy the Slater condition.

The Łojasiewicz gradient inequality (1.2) is proved in the case of differentiable definable functions in an o-minimal structure and U bounded in [20] and, in the case of subanalytic functions, in [4]. In [5], there are necessary and sufficient conditions for the Łojasiewicz gradient inequality in the nonsmooth case and applications. With some specific cases of o-minimal structures, other Łojasiewicz-type inequalities in compact domains are given in [26]. In this paper, we will establish the Łojasiewicz gradient inequality in a noncompact case with differentiable definable real-valued functions in an o-minimal structure.

The rest of the paper is organised as follows. In Section 2, we give a short introduction to o-minimal structures and some of their properties. Section 3 contains a criterion for the existence of Łojasiewicz-type inequalities and a necessary and sufficient condition for the existence of a global Hölderian error bound. A relation between the Palais–Smale condition and the existence of error bounds will also be established. Finally, we give a characterisation of the Łojasiewicz nonsmooth gradient inequality near the fibre for continuous definable functions.

2. Preliminaries

In this section, we recall some notions and results of the geometry of o-minimal structures, as found in [8, 10, 11].

DEFINITION 2.1. A structure expanding the real field $(\mathbb{R}, +, .)$ is a collection $\mathcal{O} = (\mathcal{O}_n)_{n\in\mathbb{N}}$ where each $\mathcal{O}_n$ is a set of subsets of the affine space $\mathbb{R}^n$, satisfying the following axioms.

(1) All algebraic subsets of $\mathbb{R}^n$ are in $\mathcal{O}_n$.
(2) For every n, $\mathcal{O}_n$ is closed under finite set-theoretical operations.
(3) If $A \in \mathcal{O}_n$ and $B \in \mathcal{O}_m$, then $A \times B \in \mathcal{O}_{m+n}$.
(4) If $\pi : \mathbb{R}^{n+1} \to \mathbb{R}^n$ is the projection on the first n coordinates and $A \in \mathcal{O}_{n+1}$ then $\pi(A) \in \mathcal{O}_n$.

The elements of $\mathcal{O}_n$ are called the *definable subsets* of $\mathbb{R}^n$. If $\mathcal{O}$ also satisfies the following condition (5), then $\mathcal{O}$ is called an *o-minimal structure* on $\mathbb{R}$.

(5) The elements of $\mathcal{O}_1$ are precisely the finite unions of points and intervals.

EXAMPLE 2.2. A semialgebraic set is a finite union of sets $S = \{x \in \mathbb{R}^n \mid f(x) = 0, g_j(x) < 0, j = 1, \dots, m\}$ where f, g_j are polynomials in $\mathbb{R}[x_1, \dots, x_n]$. The collection $\mathcal{O}$ of all semialgebraic sets in $\mathbb{R}^n$ for all $n \in \mathbb{N}$ is an o-minimal structure on $\mathbb{R}$.

A *first-order formula* (of the language of the o-minimal structure) is constructed according to the following rules.

(1) If $P \in \mathbb{R}[X_1, \dots, X_n]$ then $P(X_1, \dots, X_n) = 0$ and $P(X_1, \dots, X_n) > 0$ are first-order formulas.
(2) If A is a definable subset of $\mathbb{R}^n$ then $x = (x_1, \dots, x_n) \in A$ is a first-order formula.
(3) If $\Phi(x_1, \dots, x_n)$ and $\Psi(x_1, \dots, x_n)$ are first-order formulas then $\{\Phi \text{ and } \Psi\}$, $\{\Phi \text{ or } \Psi\}$, $\{\text{not } \Phi\}$ and $\{\Phi \Rightarrow \Psi\}$ are first-order formulas.

https://doi.org/10.1017/S0004972715000726 Published online by Cambridge University Press

(4) If $\Phi(y, x)$ is a first-order formula (where $y = (y_1, \ldots, y_p)$ and $x = (x_1, \ldots, x_n)$) and A is a definable subset of $\mathbb{R}^n$ then {there exists $x \in A$ such that $\Phi(y, x)$} and {for all $x \in A$, $\Phi(y, x)$} are first-order formulas.

THEOREM 2.3 [8, Theorem 1.13]. *If $\Phi(x_1, \ldots, x_n)$ is a first-order formula then the set of $(x_1, \ldots, x_n)$ in $\mathbb{R}^n$ which satisfy $\Phi(x_1, \ldots, x_n)$ is definable.*

REMARK 2.4. By rule (4) and the above theorem, the sets $\{x \in \mathbb{R}^n :$ there exists $x_{n+1}, (x, x_{n+1}) \in A\}$ (image of A by projection) and $\{x \in \mathbb{R}^n :$ for all $x_{n+1}, (x, x_{n+1}) \in A\}$ (complement of the image of the complement of A by projection) are definable.

DEFINITION 2.5. A map $f : A \to \mathbb{R}^p$ (where $A \subset \mathbb{R}^n$) is called *definable* if its graph is a definable subset of $\mathbb{R}^n \times \mathbb{R}^p$.

PROPOSITION 2.6 [10, 11].

(i) *The closure, the interior and the boundary of a definable set are definable.*
(ii) *Compositions of definable maps are definable.*
(iii) *Images and inverse images of definable sets under definable maps are definable.*
(iv) *The infimum of a bounded below definable function and the supremum of a bounded above definable function are definable functions.*

PROPOSITION 2.7. *If the function $f : \mathbb{R}^n \to \mathbb{R}$ is definable then the set $S = \{x \in \mathbb{R}^n \mid f(x) \leq 0\}$ is definable.*

PROOF. By definition, $\Gamma_f = \mathbb{R}^n \times f(\mathbb{R}^n)$ is definable. Consider the projection

$$\pi : \mathbb{R}^{n+1} \to \mathbb{R},$$
$$(x_1, \ldots, x_n, x_{n+1}) \mapsto x_{n+1}.$$

From the definition of a first-order formula, $\pi(\Gamma_f) = \{y \in \mathbb{R} \mid y = f(x), x \in \mathbb{R}^n\}$ is definable. Similarly, $\{y \in \mathbb{R} | y \leq 0\}$ is definable. So $S = \pi(\Gamma_f) \cap \{y \leq 0\}$ is definable. □

PROPOSITION 2.8. *If $S \neq \emptyset$ is a definable set, then the function $d : \mathbb{R}^n \to \mathbb{R}$ defined by*

$$d(x, S) = \inf_{y \in S} \|x - y\|$$

is well defined and is a definable function; moreover, it is a continuous function on $\mathbb{R}^n$.

PROOF. The set $\{\|x - y\| \mid y \in S\}$ is an image of S by the definable function $y \mapsto \|x - y\|$, so it is definable subset. Since $S \neq \emptyset$, d is well defined. Its graph Γ_d is the set of $(x, t) \in \mathbb{R}^{n+1}$ defined by the first-order formulas

$$t \geq 0, \ \forall y \in S : t^2 \leq \|x - y\|^2$$

and

$$\forall \epsilon \in \mathbb{R}, \ \epsilon > 0 \Rightarrow \exists y \in S : t^2 + \epsilon > \|x - y\|^2$$

so Γ_d is definable. Hence $d(x, S)$ is a definable function. By the triangle inequality, $|d(x, S) - d(x_0, S)| \leq d(x, x_0)$. Therefore $x \to x_0$ implies $d(x, S) \to d(x_0, S)$. Hence $d(x, S)$ is a continuous function. □

https://doi.org/10.1017/S0004972715000726 Published online by Cambridge University Press

Proposition 2.9. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable definable function in some o-minimal structure. Then the functions $\partial f/\partial x_j$ for $j = 1, \ldots, n$ are definable and $\nabla f(x)$ (gradient of f) is a definable mapping.*

Proof. By the definition of partial derivatives,

$$\frac{\partial f}{\partial x_j}(a) = \lim_{x_j \to a_j} \frac{f(x_1, \ldots, x_j, \ldots, x_n) - f(a_1, \ldots, a_j, \ldots, a_n)}{x_j - a_j} \quad a \in \mathbb{R}^n,$$

so we have the first-order formula

$$-\epsilon < \frac{f(x_1, \ldots, x_j + h, \ldots, x_n) - f(x_1, \ldots, x_n)}{h} - \frac{\partial f}{\partial x_j} < \epsilon \quad \forall \epsilon > 0, h > 0, j = 1, \ldots, n.$$

By Theorem 2.3, $\partial f/\partial x_j$ is a definable function and so $\nabla f(x)$ is definable. □

The following useful result is a property of semialgebraic functions in one variable.

Lemma 2.10 (Growth dichotomy lemma [11]). *Let $f : (0, \epsilon) \to \mathbb{R}$ be a semialgebraic function with $f(s) \neq 0$ for all $s \in (0, \epsilon)$. Then there exist constants $c \neq 0$ and $q \in \mathbb{Q}$ such that $f(s) = cs^q + o(s^q)$ as $s \to 0^+$.*

The following property is important for our purpose.

Theorem 2.11 (Monotonicity theorem [8, 10, 11]). *Let $f : (a, b) \to \mathbb{R}$ be a definable function with $-\infty \leq a < b \leq +\infty$. Then there exist $a_0, a_1, \ldots, a_{k+1}$ with $a = a_0 < a_1 < \cdots < a_k < a_{k+1} = b$ such that f is continuous on each interval (a_i, a_{i+1}); moreover, f is either strictly monotone or constant on each interval $(a_i, a_{i+1}), i = 1, \ldots, k$.*

We now recall the notion of the subdifferential of a continuous function. This notion plays the role of the usual gradient map.

Definition 2.12 [6, 34].

(i) The *Fréchet subdifferential* $\hat{\partial} f(x)$ of a continuous function $f : \mathbb{R}^n \to \mathbb{R}$ at $x \in \mathbb{R}^n$ is given by

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

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

Remark 2.13.

(i) It is easy to show that for a continuous function f on $\mathbb{R}^n$, the set $\{x : \hat{\partial} f(x) \neq \emptyset\}$ is dense in $\mathbb{R}^n$.

(ii) It is not hard to show that if f is a definable function then $\hat{\partial} f(x)$ and $\partial f(x)$ are definable sets [18, Proposition 3.1].

https://doi.org/10.1017/S0004972715000726 Published online by Cambridge University Press

DEFINITION 2.14. The *nonsmooth slope* of the function f is given by

$$\mathfrak{m}_f(x) := \inf\{\|v\| : v \in \partial f(x)\}.$$

By definition, $\mathfrak{m}_f(x) = +\infty$ whenever $\partial f(x) = \emptyset$.

DEFINITION 2.15. The *strong nonsmooth slope* of the function f is given by

$$|\nabla f|(x) := \lim_{h\to 0} \sup_{h\neq 0} \frac{[f(x) - f(x+h)]_+}{\|h\|},$$

with $[a]_+ = \max\{a, 0\}$.

From [17], the nonsmooth slope, strong nonsmooth slope and subdifferential are connected by

$$\inf\{\|y\| : y \in \hat{\partial} f(x)\} \geq |\nabla f|(x) \geq \mathfrak{m}_f(x).$$

REMARK 2.16.

(i) It is not hard to show that if f is a definable function then $\mathfrak{m}_f(x)$ and $|\nabla f|(x)$ are definable [18, Proposition 3.1].

(ii) If f is a differentiable function then the above notions coincide with the usual gradient; that is, $\partial f(x) = \hat{\partial} f(x) = \{\nabla f(x)\}$ and $\mathfrak{m}_f(x) = |\nabla f|(x) = \|\nabla f(x)\|$.

3. Main results

3.1. Łojasiewicz-type inequalities. The following results extend the results of [13] (see also [9]) from polynomial functions to continuous definable functions. The main difficulty in extending the proof of [13] is that we do not have the growth dichotomy lemma in general o-minimal structures. We use the monotonicity theorem instead.

PROPOSITION 3.1 (Łojasiewicz-type inequality 'near the set S'). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous definable function. Assume that $S := \{x \in \mathbb{R}^n \mid f(x) \leq 0\} \neq \emptyset$. Let $[f(x)]_+ := \max\{f(x), 0\}$. Then the following two statements are equivalent.*

(i) *For any sequence $x^k \in \mathbb{R}^n \setminus S$ with $x^k \to \infty$,*

$$f(x^k) \to 0 \quad \Longrightarrow \quad d(x^k, S) \to 0.$$

(ii) *There exist $\delta > 0$ and a function $\mu : [0, \delta] \to \mathbb{R}$ which is definable, continuous and strictly increasing on $[0, \delta)$ with $\mu(0) = 0$ such that*

$$\mu([f(x)]_+) \geq d(x, S) \quad \forall x \in f^{-1}((-\infty, \delta]).$$

PROOF. (ii) $\Rightarrow$ (i): Assume that $x^k \notin S$, $x^k \to \infty$ and $f(x^k) \to 0$. We have $[f(x^k)]_+ = f(x^k)$. By the continuity of μ at 0, we get $\mu(f(x^k)) \to 0$. Note that $0 < f(x^k) < \delta$ if $k \gg 1$. Then it follows from the inequality in (ii) that $d(x^k, S) \to 0$.

(i) $\Rightarrow$ (ii): Without loss of generality, we can suppose that $S \neq \mathbb{R}^n$. Then there exists $t_0 > 0$ such that $f^{-1}(t_0) \neq \emptyset$. Because f is continuous, $f^{-1}(t) \neq \emptyset$ for all $0 \leq t \ll 1$.

Let $\mu(t) := \sup_{x\in f^{-1}(t)} d(x, S), t \geq 0$. We will show that there exists $\delta > 0$ sufficiently small such that $\mu(t)$ has the desired properties. Clearly, $\mu(0) = 0$.

We now show that there exists $\delta > 0$ such that $\mu(t) < +\infty$ for all $t \in [0, \delta)$. Assume to the contrary that there exists a sequence $t_k > 0, t_k \to 0$ such that $\mu(t_k) = \infty$ for all k. This implies the existence of a sequence $x^k \in f^{-1}(t_k)$ such that $d(x^k, S) \to +\infty$ as $k \to \infty$, and hence $x^k \to \infty$. This is a contradiction.

So $\mu(t) < +\infty$ on $[0, \delta]$ with $\delta > 0$. By Propositions 2.8 and 2.9(iv), $\mu(t)$ is definable on $[0, \delta]$. Using the monotonicity theorem, the function μ is continuous and monotone on $(0, \delta]$ if $0 < \delta \ll 1$.

We now show that μ is continuous at 0. If μ is not continuous at 0, there exists a sequence $t_k \to 0$ such that $\mu(t_k) = \sup_{x\in f^{-1}(t_k)} d(x, S) \nrightarrow 0$. Hence there exists a sequence $x^k \in f^{-1}(t_k)$ such that $t_k = f(x^k) \to 0$ and $d(x^k, S) \nrightarrow 0$. On the other hand, $x^k \to \infty$. Indeed, if there exists $x < \infty$ such that $x^k \to x$ then by the continuity of f, $f(x^k) \to f(x)$, which implies $f(x) = 0$. That means $d(x^k, S) \to 0$, a contradiction. So we have a sequence $x^k \to \infty$, $f(x^k) \to 0$ and $d(x^k, S) \nrightarrow 0$. This contradicts (i).

Hence, μ is continuous and monotone on $[0, \delta]$. Since $\mu(0) = 0$ and $\mu(t) > 0$, for all $t \in (0, \delta)$, if δ is sufficiently small then $\mu(t)$ is strictly increasing on $[0, \delta]$.

For $0 < t < \delta$, let $x \in f^{-1}(t)$, then $\mu(t) = \sup_{a\in f^{-1}(t)} d(a, S) \geq d(x, S)$. Hence $\mu([f(x)]_+) \geq d(x, S)$ for all $x \in f^{-1}((-\infty, \delta])$. □

Remark 3.2. Note that the conditions in (ii) that μ is continuous at 0 and that $\mu(0) = 0$ are necessary. Let us consider the function $f : \mathbb{R} \to \mathbb{R}, x \mapsto x/(1 + x^2)$. The function f is a differentiable semialgebraic function because its graph is the set $\{(x, y) \in \mathbb{R}^2 \mid (1 + x^2)y = x\}$. Thus f is a definable function. We have $S = (-\infty, 0]$. Choose $\mu(t) := \sup_{x/(1+x^2)=t} d(x, S)$ on $0 < t < 1/2$. This function is definable and continuous on $(0, 1/2)$ but is not continuous at 0. Moreover, $x^k \to +\infty$ satisfies $f(x^k) \to 0$ but $d(x^k, S) \to +\infty$, so statement (i) fails.

Proposition 3.3 (Łojasiewicz-type inequality 'far from the set S'). *Suppose that for any sequence $x^k \in \mathbb{R}^n \setminus S$ with $x^k \to \infty$ and $d(x^k, S) \to \infty$ we have $f(x^k) \to \infty$. Then there exist $r > 0$ and a function $\mu : [r, +\infty) \to \mathbb{R}$ which is definable, increasing and continuous on $[r, +\infty)$ such that*

$$\mu([f(x)]_+) \geq d(x, S) \quad \forall x \in f^{-1}([r, +\infty)).$$

Proof. Let us consider two cases.

Case 1. The function f is bounded from above, that is, $r := \sup_{x\in\mathbb{R}^n} f(x) < +\infty$. By the assumption, there exists $M > 0$ such that $d(x, S) \leq M$ for all $x \in \mathbb{R}^n$. For all $x \in f^{-1}([r', r))$ $(0 < r' < r)$,

$$f(x) \geq r' = \frac{r'}{M} M \geq \frac{r'}{M} d(x, S).$$

Thus the function $\mu(t) := (M/r')t$ with $t \geq r'$ has the required properties.

https://doi.org/10.1017/S0004972715000726 Published online by Cambridge University Press

Case 2. The function f is not bounded from above. By continuity of f and $S \neq \emptyset$ we have $f^{-1}(t) \neq \emptyset$ for all $t \geq 0$. Set $\mu(t) = \sup_{x \in f^{-1}(t)} d(x, S)$.

We claim that there exists $r \gg 1$ such that $\mu(t) = \sup_{x \in f^{-1}(t)} d(x, S) < \infty$ for all $t \geq r$. Assume to the contrary that $\mu(t) = \infty$ for some $t \gg 1$. Then there exists a sequence $x^k \in f^{-1}(t)$ such that $d(x^k, S) \to \infty$. But $x^k \to \infty$, a contradiction.

So $\mu(t) < +\infty$ for all $t \in [r, +\infty)$. This implies that μ is a definable function on $[r, +\infty)$. By the monotonicity theorem, μ is continuous and monotone on $[r, +\infty)$ for $r \gg 1$.

Let

$$M := \sup_{t \in [r, +\infty)} \mu(t).$$

We have two subcases.

Case 2.1. $M = +\infty$. Then $\lim_{t \to +\infty} \mu(t) = +\infty$. This means that for $r \gg 1$, the function μ is strictly increasing on $[r, +\infty)$. Furthermore,

$$\mu([f(x)]_+) = \mu(f(x)) \geq d(x, S) \quad \forall x \in f^{-1}([r, +\infty)).$$

Case 2.2. $M < +\infty$. For all x with $f(x) \geq r$ we have $d(x, S) \leq M$, so

$$f(x) \geq r = \frac{r}{M} M \geq \frac{r}{M} d(x, S).$$

The function $\mu := (M/r)t, t \geq r$ has the required properties. □

Remark 3.4. Note that the converse of the above theorem is false. Consider the function $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x/\sqrt{1 + x^2}$. The function f is a differentiable semialgebraic function since its graph is the set $\{(x, y) \in \mathbb{R}^2 | (1 + x^2)y^2 = x^2\} \cap \{xy > 0\}$. We have $S = (-\infty, 0]$. We choose $0 < r < 1$ and let

$$\mu(t) := \begin{cases} \sup_{x/\sqrt{1+x^2}=t} d(x, S) & \text{on } [r, 1), \\ +\infty & \text{on } [1, +\infty). \end{cases}$$

This function is definable, increasing and continuous. However, we have $x^k \to +\infty$, $d(x^k, S) \to +\infty$ and $f(x^k) \to 1$.

3.2. Global Hölderian error bound for continuous definable functions in o-minimal structures. The following criterion extends the error bound result of [13] from polynomial functions to definable functions in o-minimal structures.

Theorem 3.5. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous definable function. Assume that $S := \{x \in \mathbb{R}^n \mid f(x) \leq 0\} \neq \emptyset$ and set $[f(x)]_+ := \max\{f(x), 0\}$. Then the following two statements are equivalent.*

(i) *For any sequence $x^k \in \mathbb{R}^n \backslash S, x^k \to \infty$:*

  (a) *if $f(x^k) \to 0$ then $d(x^k, S) \to 0$;*
  (b) *if $d(x^k, S) \to \infty$ then $f(x^k) \to \infty$.*

https://doi.org/10.1017/S0004972715000726 Published online by Cambridge University Press

(ii) *There exists a function* $\mu : [0, +\infty) \to \mathbb{R}$ *which is definable, strictly increasing and continuous on* $[0, +\infty)$ *with* $\mu(0) = 0, \lim_{t\to+\infty} \mu(t) = +\infty$ *such that*

$$d(x, S) \le \mu([f(x)]_+) \quad \forall x \in \mathbb{R}^n.$$

Proof. The implication (ii) $\Rightarrow$ (i) is straightforward. We prove that (i) $\Rightarrow$ (ii).

By Propositions 3.1 and 3.3, there are two continuous, strictly increasing, definable functions μ_1 on $[0, \delta]$ with $0 < \delta \ll 1$ and μ_2 on $[r, +\infty)$ with $r \gg 1$ such that

$$d(x, S) \le \mu_1([f(x)]_+) \quad \forall x \in f^{-1}((-\infty, \delta])$$

and

$$d(x, S) \le \mu_2([f(x)]_+) \quad \forall x \in f^{-1}([r, +\infty)).$$

By assumption (b), there is $M > 0$ with $d(x, S) \le M$ for all $x \in f^{-1}([\delta, r])$. Thus

$$f(x) \ge \delta = \frac{\delta}{M} M \ge \frac{\delta}{M} d(x, S)$$

for all $x \in f^{-1}([\delta, r])$. Put $\mu_3(t) := (M/\delta)t$ for $t \in [\delta, r]$. Then $\mu_3(t) \ge d(x, S)$ and μ_3 is an increasing function on $[\delta, r]$.

By definition of μ_3 and since $\lim_{t\to 0} \mu_1(t) = 0$ (Theorem 3.1), we may choose δ such that $\mu_1(t) \le \mu_3(\delta) = M$ for all $t \in [0, \delta]$. Indeed, if there is $t \in [0, \delta]$ such that $\mu_1(t) > \mu_3(\delta)$, then we put

$$M' := \max\Big\{\sup_{t\in[0,\delta]} \mu_1(t), M\Big\} \quad \text{and} \quad \mu_3(t) := \frac{M'}{\delta} t,$$

so $\mu_1(t) \le M' = \mu_3(\delta)$ for all $t \in [0, \delta]$.

Similarly, by definition of $\mu_3(t)$ and $\mu_2(t)$, we may choose r such that $\mu_3(r) = (M/\delta)r \le \mu_2(t)$ for all $t \in [r, +\infty)$. Indeed, if there is $t \in [r, +\infty)$ such that $\mu_3(r) > \mu_2(t)$, then we may choose $\mu_2'(t) \ge \mu_2(t) + C$ with $C = \mu_3(r)$, so $d(x, S) \le \mu_2(t) < \mu_2'(t)$ for $t \in [r, +\infty)$ and $\mu_3(r) \le \mu_2'(t)$ for all $t \in [r, +\infty)$. Moreover, by the definition of μ_2, we may choose $\mu_2'(t)$ as above such that if $r \gg 1$ then $\mu_2'(t)$ is strictly increasing on $[r, +\infty)$ and $\lim_{t\to+\infty} \mu_2'(t) = +\infty$.

By choosing suitable δ, r and M, the function μ formed from μ_1, μ_2, μ_3 by

$$\mu(t) = \begin{cases} \mu_1(t) & \text{if } t \in [0, \delta], \\ \mu_3(t) & \text{if } t \in [\delta, r], \\ \mu_2'(t) & \text{if } t \in [r, +\infty) \end{cases}$$

is definable, strictly increasing and continuous and satisfies (ii). □

3.3. The Palais–Smale condition and the existence of error bounds. In this section, we consider continuous functions in an o-minimal structure. It is well known via the Ekeland principle that the Palais–Smale condition implies the existence of an error bound (see, for example, [7] for general results with continuous functions on metric spaces). We will prove such results for definable functions.

https://doi.org/10.1017/S0004972715000726 Published online by Cambridge University Press

DEFINITION 3.6. Given a continuous function $f : \mathbb{R}^n \to \mathbb{R}$ and a real number t, we say that f satisfies the *Palais–Smale condition at level t* if every sequence $\{x^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ such that $f(x^k) \to t$ and $\mathfrak{m}_f(x^k) \to 0$ as $k \to \infty$ possesses a convergent subsequence.

The following theorem extends [13, Theorem B] from polynomial functions to continuous definable functions. In the case of continuous definable functions, we use the subdifferential instead of the gradient in [13].

THEOREM 3.7. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous definable function. Assume that $S := \{x \in \mathbb{R}^n \mid f(x) \le 0\} \neq \emptyset$. If f satisfies the Palais–Smale condition at each level $t \ge 0$, then there exists a function $\mu : [0, +\infty) \to \mathbb{R}$ which is definable, strictly increasing and continuous with $\mu(0) = 0$ and $\lim_{t\to\infty} \mu(t) = \infty$ such that*

$$d(x, S) \le \mu([f(x)]_+) \quad \forall x \in \mathbb{R}^n.$$

PROOF. By Theorem 3.5, it is enough to show that if f satisfies the Palais–Smale condition at each level $t \ge 0$, there is no sequence $x^k \to \infty$, $x^k \in \mathbb{R}^n \backslash S$ such that

$$f(x^k) \to 0 \quad \text{but } d(x, S) > \delta > 0$$

or

$$d(x^k, S) \to \infty \quad \text{but } 0 \le f(x^k) \le M$$

for some $\delta > 0$ and $M > 0$.

Proceeding by contradiction, assume that for a sequence $x^k \to \infty$, $x^k \in \mathbb{R}^n \backslash S$ we have $f(x^k) \to 0$ and $d(x^k, S) \ge \delta > 0$. Arguing as in [13, Theorem B], by the Ekeland Variational Principle [12] we obtain a sequence y^k such that

$$\frac{1}{\|h\|}(f(y^k + h) - f(y^k)) \ge -\sqrt{\epsilon_k}$$

with $h \in \mathbb{R}^n$, $0 < \|h\| < \frac{1}{2}\delta$ and $\epsilon_k = f(x^k)$. This implies that

$$\frac{1}{\|h\|}(f(y^k) - f(y^k + h)) \le \sqrt{\epsilon_k}$$

or

$$\frac{1}{\|h\|}[f(y^k) - f(y^k + h)]_+ \le \sqrt{\epsilon_k}.$$

By the definition of the strong slope,

$$0 \le |\nabla f|(y^k) = \limsup_{h\to 0, h\neq 0} \frac{[f(y^k) - f(y^k + h)]_+}{\|h\|} \le \sqrt{\epsilon_k}.$$

Thus,

$$0 \le \mathfrak{m}_f(y^k) \le |\nabla f|(y^k) \le \sqrt{\epsilon_k}.$$

Letting $k \to \infty$ we get $\mathfrak{m}_f(y^k) \to 0$. So we have found a sequence $y^k \to \infty$, $y^k \in \mathbb{R}^n \backslash S$, with $\mathfrak{m}_f(y^k) \to 0$ and $f(y^k) \to 0$. This means that f does not satisfy the Palais–Smale condition at the value $t = 0$, a contradiction. So we get (a) of Theorem 3.5.

https://doi.org/10.1017/S0004972715000726 Published online by Cambridge University Press

Now, suppose there is a sequence $x^k \in \mathbb{R}^n \setminus S$ with $x^k \to \infty$ such that

$$d(x^k, S) \to \infty \quad \text{and} \quad f(x^k) \nrightarrow \infty.$$

Without loss of generality, we may assume $f(x^k) \to t_0$ with $t_0 \in [0, +\infty)$. Again arguing as in [13, Theorem B], we find a sequence y^k such that $0 < f(y^k) \le f(x^k)$ and

$$\frac{1}{\|h\|}(f(y^k + h) - f(y^k)) \ge -\epsilon_k \cdot \lambda_k$$

with $h \in \mathbb{R}^n$, $0 < \|h\| < \frac{1}{2}\delta$, $\epsilon_k = f(x^k)$ and $\lambda_k = 2/d(x^k, S)$. This implies that

$$\frac{1}{\|h\|}[f(y^k) - f(y^k + h)]_+ \le \epsilon_k \lambda_k.$$

By the definition of the strong slope,

$$0 \le \mathfrak{m}_f(y^k) \le |\nabla f|(y^k) \le \epsilon_k \lambda_k = \frac{2\epsilon_k}{d(x^k, S)}.$$

If $k \to \infty$ then $\epsilon_k = f(x^k) \to t_0$ and $d(x^k, S) \to \infty$ and so $\mathfrak{m}_f(y^k) \to 0$. Since $0 < f(y^k) \le f(x^k)$, y^k has a subsequence y'^k with $f(y'^k) \to t_1$ for $0 \le t_1 \le t_0$ and

$$y'^k \to \infty, \quad \mathfrak{m}_f(y'^k) \to 0, \quad f(y'^k) \to t_1.$$

This means that f does not satisfy the Palais–Smale condition at t_1, a contradiction. So we have (b) of Theorem 3.5 and the theorem is proved. □

3.4. The Łojasiewicz nonsmooth slope inequality near the fibre for continuous definable functions in an o-minimal structure. The Łojasiewicz gradient inequality is a useful tool in many problems, such as evolution equations [35] and minimisation algorithms [2, 4, 5]. Characterising this inequality is an important problem. In the nonsmooth case, [5] examines the relationships between this inequality and objects such as the length of subgradient curves and metric regularity under the assumption of compactness of the neighbourhood U. The classical Łojasiewicz gradient inequality (1.2) is not always true when U is unbounded, as shown in the following example.

Example 3.8. Set $f(x, y) = (xy - 1)^2 + (x - 1)^2$, $U = \mathbb{R}^2$ and $x^k = ((1 + k)/(1 + k^2), k)$. As $x^k \to \infty$, $\nabla f(x^k) = (0, 2(k^2 - 1)/(1 + k^2)^2) \to 0$, but

$$f(x^k) = \left(\frac{(1 + k)k}{1 + k^2} - 1\right)^2 + \left(\frac{1 + k}{1 + k^2} - 1\right)^2 \to 1.$$

We prove that there is no $\delta > 0$, $C > 0$ and $\rho \in \mathbb{R}$ such that $\|\nabla f(x)\| \ge C|f(x)|^\rho$ for $x \in f^{-1}(D_\delta)$ where $D_\delta = \{t : |t| < \delta\}$. Assume that there are $\delta > 0$, $C > 0$ and $\rho \in \mathbb{R}$ such that the Łojasiewicz gradient inequality holds. On the one hand,

$$\nabla f\left(\frac{1}{k} + 1, \frac{k}{k + 1}\right) \to 0, \quad f\left(\frac{1}{k} + 1, \frac{k}{k + 1}\right) \to 0$$

so $\rho > 0$. On the other hand,

$$\nabla f\left(\frac{1 + k}{1 + k^2}, k\right) \to 0, \quad f\left(\frac{1 + k}{1 + k^2}, k\right) \to 1$$

so $\rho \le 0$, a contradiction.

We give a criterion for there to be a Łojasiewicz nonsmooth slope inequality on $f^{-1}(D_\epsilon)$ with $D_\epsilon = (-\epsilon, \epsilon)$. Let

$$\widetilde{K}_\infty(f) := \{t \in \mathbb{R} \mid \exists x^k \to \infty, \mathfrak{m}_f(x^k) \to 0, f(x^k) \to t\}$$

denote the set of asymptotic critical values at infinity and let

$$\widetilde{K}(f) := \{t \in \mathbb{R} \mid \exists x^k, \mathfrak{m}_f(x^k) \to 0, f(x^k) \to t\}$$

be the set of asymptotic critical values. Note that $\widetilde{K}(f) = \widetilde{K}_0(f) \cup \widetilde{K}_\infty(f)$, where $\widetilde{K}_0(f)$ is the set of critical values.

Theorem 3.9. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous definable function in some o-minimal structure and suppose that $\widetilde{K}(f) \cap D_\epsilon = \{0\}$. Then the following two statements are equivalent.*

(i) *For any sequence $x^k \to \infty$, $\mathfrak{m}_f(x^k) \to 0$ implies $f(x^k) \to 0$.*
(ii) *There exists a function $\varphi : (0, \delta) \to \mathbb{R}$ which is definable, monotone and continuous such that $\mathfrak{m}_f(x) \geq \varphi(|f(x)|)$ for all $x \in f^{-1}(D_\delta)$.*

Proof. We note that (ii) $\Rightarrow$ (i) is straightforward. We prove (i) $\Rightarrow$ (ii).

Let $\varphi(t) := \inf\{\mathfrak{m}_f(x) : |f(x)| = t\}$. It is easy to see that φ is a definable function (see Remark 2.16 and Proposition 2.6). We have to prove that $\varphi(t) \not\equiv 0$.

We claim that there exists δ_1 such that $\varphi(t) > 0$ for all $t \in (0, \delta_1)$. Indeed, by the assumption $\widetilde{K}(f) \cap D_\epsilon = \{0\}$, $(0, \epsilon)$ contains no critical point of f. Assume that for any δ' there exists a value $t \in (0, \delta')$ such that $\varphi(t) = 0$. Then there exists a sequence t_k such that $t_k \to t$ implies $\varphi(t_k) \to 0$. Therefore there exists a sequence x^k such that $|f(x^k)| = t_k$ and $\mathfrak{m}_f(x^k) \to 0$. So we have $\mathfrak{m}_f(x^k) \to 0$ but $f(x_k) \to t \neq 0$, which contradicts (i). This proves the claim.

On the other hand, by the monotonicity theorem, $\varphi(t)$ is continuous and monotone on $(0, \delta_2)$ for $0 < \delta_2 \ll 1$. We have $\varphi(0) = 0$ and $\varphi(t) > 0$ for $t \in (0, \delta_1)$, so φ is strictly monotone on $(0, \delta)$ (with $\delta = \min\{\delta_1, \delta_2\} < \epsilon$). By the definition of φ, $\varphi(t) \leq \mathfrak{m}_f(x)$ for $x \in f^{-1}(D_\delta)$, which means that $\mathfrak{m}_f(x) \geq \varphi(|f(x)|)$ for all $x \in f^{-1}(D_\delta)$. □

Remark 3.10. The cardinality of the set $\widetilde{K}_\infty(f)$ can be infinite. Indeed, consider $f(x, y) = x/(1 + y^2)$ in the o-minimal structure of all semialgebraic sets. Any $t \in \mathbb{R}$ belongs to $\tilde{K}_\infty(f)$ via the sequence $x^k = (t(1 + k^2), k)$. It is easy to see that $x^k \to \infty$, $\|\nabla f(x^k)\| = \sqrt{(1/(1 + k^2))^2 + (2tk/(1 + k^2))^2} \to 0$ and $f(x^k) = t$.

Remark 3.11. In Theorem 3.9, if f is a polynomial then $\varphi(t)$ is a semialgebraic function in one variable. By the growth dichotomy lemma, there exist $a > 0$ and $u > 0$ such that $\varphi(t) = at^u + o(t^u)$ for $t \in (0, \epsilon)$ with $\epsilon \ll 1$. This implies $\varphi(t) \geq ct^u$ for all $t \in (0, \epsilon)$. By the definition of φ we have $\|\nabla f(x)\| \geq \varphi(t) \geq ct^u$. Note that $t = |f(x)|$, so we get the Łojasiewicz gradient inequality on $f^{-1}(D_\delta)$.

Remark 3.12. In preparation of this article, we received advice that Professor Ta Le Loi has also obtained some results similar to Theorem 3.9.

https://doi.org/10.1017/S0004972715000726 Published online by Cambridge University Press

Acknowledgements

The author would like to thank Professor Ha Huy Vui for proposing the problem, Associate Professor Pham Tien Son for his suggestions, Dr Dinh Si Tiep for useful discussions and Professor Ta Le Loi for information about his unpublished results. The author would also like to thank the referee for valuable suggestions and comments.

References

[1] P. A. Absil and K. Kurdyka, 'On the stable equilibrium points of gradient systems', *Systems Control Lett.* **55** (2006), 573–577.

[2] P. A. Absil, R. Mahony and B. Andrews, 'Convergence of the iterates of descent methods or analytic cost functions', *SIAM J. Optim.* **16** (2005), 531–547.

[3] A. Auslander and J.-P. Crouzeix, 'Global regularity theorem', *Math. Oper. Res.* **13** (1988), 243–253.

[4] J. Bolte, A. Daniilidis and A. S. Lewis, 'The Łojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems', *SIAM J. Optim.* **17**(4) (2007), 1205–1223.

[5] J. Bolte, A. Daniilidis, O. Ley and L. Mazet, 'Characterizations of Łojasiewicz inequalities and applications', *Trans. Amer. Math. Soc.* **362**(6) (2010), 3319–3363.

[6] F. H. Clarke, *Optimization and Nonsmooth Analysis* (John Wiley & Sons, New York, 1983).

[7] J. N. Corvellec and V. V. Montreanu, 'Nonlinear error bounds for lower semi-continuous functions on metric spaces', *Math. Program.* **114** (2008), 291–319.

[8] M. Coste, *An Introduction to O-minimal Geometry* (Instituti Editoriali e poligrafici internazionali, Università di Pisa, Pisa, 1999).

[9] S. T. Dinh, H. V. Hà and T. T. Nguyen, 'Łojasiewicz inequality for polynomial functions on non-compact domains', *Int. J. Math.* **23** (2012), 1250033, doi:10.1142/S0129167X12500334.

[10] L. van den Dries, *Tame Topology and O-minimal Structures* (Cambridge University Press, Cambridge, 1998).

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

[12] I. Ekeland, 'Nonconvex minimization problems', *Bull. Amer. Math. Soc. (N.S.)* **1** (1974), 443–474.

[13] H. V. Hà, 'Global Hölderian error bound for non-degenerate polynomials', *SIAM J. Optim.* **23**(2) (2013), 917–933.

[14] H. V. Hà and H. D. Nguyen, 'Łojasiewicz inequality at infinity for polynomial in two real variables', *Math. Z.* **266** (2010), 243–264.

[15] A. J. Hoffman, 'On approximate solutions of linear inequalities', *J. Res. Natl. Bur. Stand.* **49** (1952), 263–265.

[16] L. Hörmander, 'On the division of distributions by polynomials', *Ark. Mat.* **3**(53) (1958), 555–568.

[17] A. D. Ioffe, 'Metric regularity and subdifferential calculus', *Uspekhi Mat. Nauk.* **55** (2000), 103–162 (in Russian); English transl. *Russian Math. Surveys* **55** (2000), 501–588.

[18] A. D. Ioffe, 'An invitation to tame optimization', *SIAM J. Optim.* **19**(4) (2009), 1894–1917.

[19] D. Klatte and A. Li, 'Asymptotic constraint qualifications and global error bounds for convex inequalities', *Math. Program.* **84** (1999), 137–140.

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

[21] G. Li, 'On the asymptotic well behaved functions and global error bound for convex polynomials', *SIAM J. Optim.* **20**(4) (2010), 1923–1943.

[22] G. Li, 'Global error bounds for piecewise convex polynomials', *Math. Program.* **137**(1–2) (2013), 37–64.

[23] G. Li, B. S. Mordukhovich and T. S. Pham, 'New error bounds for polynomial systems with applications to Hölderian stability in optimization and spectral theory of tensors', *Math. Program.* (2014), doi:10.1007/s10107-014-0806-9.

[24] G. Li and K. F. Ng, 'Error bounds of generalized D-gap functions for nonsmooth and nonmonotone variational inequality problems', *SIAM. J. Optim.* **20**(2) (2009), 667–690.

[25] G. Li, C. Tang and Z. X. Wei, 'Error bound results for generalized D-gap functions of nonsmooth variational inequality problems', *Int. J. Comput. Appl. Math.* **233**(11) (2010), 2795–2806.

[26] T. L. Loi, 'Łojasiewicz inequalities for sets definable in the structure $\mathbb{R}_{\exp}$', *Ann. Inst. Fourier (Grenoble)* **45** (1995), 951–957.

[27] S. Łojasiewicz, 'Sur le problème de la division', *Studia Math.* **18** (1959), 87–136.

[28] S. Łojasiewicz, *Ensembles Semi-Analytiques* (Publications Mathématiques Institut des Hautes Études Scientifiques, Bures-sur-Yvette, France, 1964).

[29] X. D. Luo and Z. Q. Luo, 'Extensions of Hoffman's error bound to polynomial systems', *SIAM J. Optim.* **4** (1994), 383–392.

[30] Z. Q. Luo, 'New error bounds and their applications to convergence analysis of iterative algorithms', *Math. Program.* **88**(2, Ser. B) (2000), 341–355.

[31] Z. Q. Luo and J. F. Sturm, 'Error bound for quadratic systems', in: *High Perfomance Optimization* (eds. H. Frenk, K. Roos, T. Terlaky and S. Zhang) (Kluwer, Dordrecht, The Netherlands, 2000), 383–404.

[32] O. L. Mangasarian, 'A condition number for differentiable convex inequalities', *Math. Oper. Res.* **10** (1985), 175–179.

[33] S. Robinson, 'An application of error bounds for convex programming in a linear space', *SIAM J. Control Optim.* **13** (1975), 271–273.

[34] R. T. Rockafellar and R. Wets, *Variational Analysis*, Grundlehren der Mathematischen Wissenschaften, 317 (Springer, New York, 1998).

[35] L. Simon, 'Asymptotics for a class of non-linear evolution equations, with applications to geometric problems', *Ann. of Math.* (2) **118** (1983), 525–571.

[36] W. H. Yang, 'Error bounds for convex polynomials', *SIAM J. Optim.* **19** (2009), 1633–1647.

DŨNG PHI HOÀNG, Laboratory of Applied Mathematics and Computing,
Posts and Telecommunications Institute of Technology (PTIT), Vietnam
and
Department of Scientific Fundamentals, PTIT,
Km10 Nguyen Trai Road, Ha Dong District, Hanoi, Vietnam
e-mail: dunghp@ptit.edu.vn