Clarke critical values of subanalytic Lipschitz continuous functions.

Jérôme BOLTE, **Aris DANIILIDIS**, **Adrian LEWIS & Masahiro SHIOTA**

Abstract We prove that any subanalytic locally Lipschitz function has the Sard property. Such functions are typically nonsmooth and their lack of regularity necessitates the choice of some generalized notion of gradient and of critical point. In our framework these notions are defined in terms of the Clarke and of the convex-stable subdifferentials. The main result of this note asserts that for any subanalytic locally Lipschitz function the set of its Clarke critical values is locally finite. The proof relies on Pawlucki's extension of the Puiseux lemma. In the last section we give an example of a continuous subanalytic function which is not constant on a segment of "broadly critical" points, that is, points for which we can find arbitrarily short convex combinations of gradients at nearby points.

Key words Clarke critical point, convex-stable subdifferential, nonsmooth analysis, Morse-Sard theorem, subanalytic function.

AMS Subject Classification *Primary* 26B05 ; *Secondary* 49J52, 32B30

1 Introduction

Several Sard-type results are known in the literature using various notions of a critical point. For example, Yomdin's classical paper [18] addresses this issue for *near-critical* points and gives an evaluation of the Kolmogorov metric entropy for the set of near-critical values. In a recent work, Kurdyka-Orro-Simon [12] show that the set of *asymptotically-critical* values of a C^p-semialgebraic mapping $f : \mathbb{R}^n \to \mathbb{R}^k$ has dimension less than k provided that $p \geq \max\{1, n-k+1\}$. Concerning non-differentiable functions, Rifford [15], extending a previous result of Itoh-Tanaka [10], establishes that the set of *Clarke critical values* of the distance function to a closed submanifold of a complete Riemannian manifold has Lebesgue measure zero.

Our work relies mainly on two concepts of a critical point that we now proceed to describe. The notion of a *limiting subgradient* for a continuous function f on $\mathbb{R}^n$ can be briefly defined as follows: x^* is a limiting subgradient of f at x, written $x^* \in \partial f(x)$, if there exist sequences $x_n \to x, x_n^* \to x^*$ such that, for n fixed:

$$\liminf_{y \to x_n,\, y \neq x_n} \frac{f(y) - f(x_n) - \langle x_n^*, y - x_n \rangle}{\|y - x_n\|} \geq 0,$$

that is, each x_n^* is a *Fréchet subgradient* of f at x_n. Clearly, for C^1 functions the notion of limiting subgradient coincides with the usual derivate ∇f of f, while in general the operator $x \mapsto \partial f(x)$ is multivalued. A *limiting-critical point* x of f is therefore a point for which there exists a zero subgradient: that is $\partial f(x) \ni 0$. Concerning nonsmooth analysis and related problems of subdifferentiation, see the introductory books of Clarke [6], Clarke-Ledyaev-Stern-Wolenski [7] or Rockafellar-Wets [16].

In a recent work [4, Theorem 13], we show that any continuous subanalytic function f on $\mathbb{R}^n$ is constant on each connected component of the set of its limiting-critical points. The main motivation for proving this Sard-type result for subanalytic continuous functions was to derive a generalized Łojasiewicz inequality which in turn was used in the asymptotic analysis of subgradient-like dynamical systems ([3, Theorem 3.1]. These dynamics occur frequently in various domains such as Optimization, Mechanics and PDE's.

With this line of research in mind we adopt here a different viewpoint. The assumptions on f are enhanced – namely, f is assumed to be locally Lipschitz continuous – while the definition of a critical point is weakened. As above, this alternative notion relies on a concept of subdifferentiation: we say that x^* is a *Clarke-subgradient* of f at x if

$$x^* \in \partial^\circ f(x) := \overline{\text{co}}\, \partial f(x),$$

where $\overline{\text{co}}\, \partial f(x)$ is the closed convex hull of $\partial f(x)$. Accordingly, a point x is said to be *Clarke critical* if $\partial^\circ f(x) \ni 0$. This turns out to be equivalent to the following property

$$0 \in \overline{\text{co}} \left\{ \bigcup_{z \in B(x,\varepsilon)} \hat{\partial} f(z) \right\}, \quad \text{for every } \varepsilon > 0 \qquad (\mathcal{CR})$$

(see Proposition 9 or [5]) which reflects the idea that a point is Clarke critical if we can find short convex combinations of gradients at nearby points.[1] For instance, $x = 0$ is a Clarke critical point for the function $x \longmapsto -||x||$, but it is not a limiting-critical, since $\partial f(0) = S^{n-1}$ (the unit sphere of $\mathbb{R}^n$), while $\partial^\circ f(x) = B_{\mathbb{R}^n}(0,1)$ (the unit ball of $\mathbb{R}^n$).

Our main result asserts that any locally Lipschitz continuous subanalytic function f defined on some open subset of $\mathbb{R}^n$ is constant on each connected component of the set of its Clarke critical points. Since the latter is subanalytic, it follows directly that the set of Clarke critical values of f is locally finite. The proof of this result is based on a "path-perturbation" lemma [4, Lemma 12], which itself relies heavily on Pawlucki's extension of the Puiseux Lemma [14, Proposition 2].

An alternative notion of subdifferential, namely the convex-stable subdifferential, has been introduced by Burke, Lewis and Overton [5]. The corresponding critical points are precisely the points which comply with $(\mathcal{CR})$. As pointed out above, if f is a Lipschitz continuous function, one recovers exactly the notion of a Clarke critical point; however for general continuous functions the convex-stable subdifferential appears to be larger than the usual Clarke subdifferential, giving rise to another concept of a critical point: the "broadly critical points". In the last section we show that a continuous subanalytic function may fail to have the Sard property on the broadly critical set. We indeed exhibit a function $f : \mathbb{R}^3 \to \mathbb{R}$ which is not constant on some segment of points satisfying $(\mathcal{CR})$.

2 Preliminaries

In this section we recall several definitions and results necessary for further developments. For basic and fundamental results of subanalytic geometry see Bierstone-Milman [2], Łojasiewicz [13], van der Dries-Miller [9] or Shiota [17]. Concerning nonsmooth analysis some general references are Clarke [6], Clarke-Ledyaev-Stern-Wolenski [7] or Rockafellar-Wets [16].

In the first two sections, we are interested in locally Lipschitz functions: accordingly, we state the definitions and theorems of nonsmooth analysis that we use specifically for this case. The case of continuous functions is treated in Section 4.

Consequently, throughout Section 2 and 3 we make the following standing assumption:

U is a nonempty open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ is a locally Lipschitz continuous function.

We shall essentially deal with the following three notions of subdifferentiation.

Definition 1 (subdifferential) For any $x \in U$ let us define
(i) the Fréchet subdifferential $\hat{\partial} f(x)$ of f at x:

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

[1] This is no longer true for continuous functions: a point satisfying $(\mathcal{CR})$ need not be Clarke critical.

(ii) the limiting subdifferential $\partial f(x)$ of f at x:

$$x^* \in \partial f(x) \Leftrightarrow \exists x_n \in U, : \exists x_n^* \in \hat{\partial} f(x_n), : x_n \to x, : x_n^* \to x^* \text{ as } n \to \infty,$$

(iii) the Clarke subdifferential $\partial^\circ f(x)$ of f at x:

$$\partial^\circ f(x) = \overline{\text{co}}\, \partial f(x), \tag{1}$$

where $\overline{\text{co}}\, \partial f(x)$ is the closed convex hull of $\partial f(x)$.

Remark 1 (a) If $T : U \rightrightarrows \mathbb{R}^n$ is a point-to-set mapping, its domain and its graph are respectively defined by $\operatorname{dom} T := \{x \in U : T(x) \neq \emptyset\}$ and $\operatorname{Graph} T := \{(x, y) \in U \times \mathbb{R}^n : y \in T(x)\}$. Clearly $\operatorname{dom} \hat{\partial} f \subset \operatorname{dom} \partial f \subset \operatorname{dom} \partial^\circ f$. A well known result of variational analysis asserts that $\operatorname{dom} \hat{\partial} f$ is a dense subset of U (see [6], for example).
(b) Since f is locally Lipschitz continuous the point-to-set mapping $U \ni x \longmapsto \partial^\circ f(x)$ is bounded on compact subsets of U.
(c) If f is differentiable at x, then $\hat{\partial} f(x) = \{\nabla f(x)\}$.
(d) If f is a subanalytic function all the subdifferential mappings defined above have a subanalytic graph (see [4, Proposition 2.13]).

The notion of a Clarke critical point is then defined naturally.

Definition 2 (Clarke critical point) A point $a \in U$ is called *Clarke critical* for a locally Lipschitz function f if

$$\partial^\circ f(a) \ni 0,$$

or equivalently, if relation $(\mathcal{CR})$ holds (see Proposition 9).

Remark 2 (subdifferential regularity) Let us recall that a locally Lipschitz function f is called *subdifferentially regular* if

$$\hat{\partial} f = \partial f,$$

or equivalently if

$$\hat{\partial} f = \partial^\circ f.$$

For subdifferentially regular functions, the sets of Fréchet-critical and of Clarke critical points coincide and one can obtain easily the conclusion of our main result via an elementary argument (see Remark 3 for details).

Let us recall the chain-rule for subdifferentials (see [16, Theorem 10.6, page 427], for example).

Proposition 3 (subdifferential chain rule) *Let V be an open subset of $\mathbb{R}^m$ and $G : V \to U$ a C^1 mapping. Define $g : V \to \mathbb{R}$ by $g(x) = f(G(x))$ for all $x \in V$. Then*

$$\hat{\partial} g(x) \supset \nabla G(x)^T \hat{\partial} f(G(x)), \tag{2}$$

$$\partial g(x) \subset \nabla G(x)^T \partial f(G(x)), \tag{3}$$

where $\nabla G(x)^T$ denotes the transpose of the Jacobian matrix of G at x.
If in addition G is a diffeomorphism the above inclusions become equalities, thus

$$\partial g(x) = \nabla G(x)^T \partial f(G(x)), \quad \partial^\circ g(x) = \nabla G(x)^T \partial^\circ f(G(x)), \forall x \in V. \tag{4}$$

The following lemma, based on a result of Pawlucki [14], plays a key role in the proof of both Theorem 5 and Theorem 7.

Lemma 4 (path perturbation lemma) ([4, Lemma 12]) *Let F be a nonempty subanalytic subset of $\mathbb{R}^n$, $\gamma : [0,1] \to \mathrm{cl}F$ a one-to-one continuous subanalytic path and $\eta > 0$.*

Then there exists a continuous subanalytic path $z : [0,1] \to \mathrm{cl}F$ such that

(i) $\|\dot{z}(t) - \dot{\gamma}(t)\| < \eta$ *for almost all* $t \in (0,1)$,

(ii) *the (subanalytic) set*

$$\Delta := \{t \in [0,1] : z(t) \in \mathrm{cl}F \diagdown F\} \tag{5}$$

has a Lebesgue measure less than η ,

(iii) $z(t) = \gamma(t)$, *for all* $t \in \Delta \cup \{0,1\}$.

Let us recall the following Sard-type result concerning the limiting-critical points of continuous subanalytic functions.

Theorem 5 (Sard theorem for limiting-critical points) ([4, Theorem 13]) *Let $g : U \to \mathbb{R}$ be a subanalytic continuous function. Then f is constant on each connected component of the set of its limiting-critical points*

$$(\partial f)^{-1}(0) := \{x \in U : \partial f(x) \ni 0\}.$$

Unless the function is subdifferentially regular, the above theorem is obviously not appropriate for the study of locally Lipschitz functions with the Clarke-subdifferential. Typical examples are given by functions whose epigraph have "inward corners", such as for instance $f(x) = -||x||$. Sharp saddle points also provide some elementary illustrations. For example if one sets

$$f : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^p \ni (x,y,z) \longmapsto ||x|| - ||y||,$$

points of the type $(0,0,z)$ are Clarke critical but they are not limiting-critical. This follows from the straightforward computations: $\partial f(0,0,z) = B_{\mathbb{R}^m}(0,1) \times S^{n-1} \times \{0\}_p$ and $\partial^\circ f(0,0,z) = B_{\mathbb{R}^m}(0,1) \times B_{\mathbb{R}^n}(0,1) \times \{0\}_p$.

3 A Sard theorem for subanalytic Lipschitz continuous functions

For the proof of the central result of this note we will need the following lemma.

Lemma 6 *Set $e := (1,0,\ldots,0) \in \mathbb{R}^n$ and assume that $[0,1]e \subset U$, with $\partial^\circ f(te) \ni 0$ for all $t \in [0,1]$. Then f is constant on $[0,1]e$.*

Proof. Let us provisionally set $S_L := \{x \in [0,1]e : 0 \in \partial f(x)\}$, where ∂f denotes the limiting subdifferential of f (Definition 1(ii)). By Remark 1 (d), the set S_L is subanalytic thus, being a (closed) subset of $[0,1]e$, it is a finite union of segments. By using Theorem 5 we conclude that f is constant on each one of them. Owing to the continuity of f, it is therefore sufficient to prove that f is also constant on each non-trivial segment of $[0,1]e \setminus S_L$. This shows that there is no loss of generality to assume that S_L is empty, that is:

$$0 \notin \partial f(te), \quad t \in [0,1].$$

Let us now fix some $\delta > 0$ and let us define

$$\Gamma_\delta = \{x \in [0,1]e : \forall x^* \in \partial f(x), \, |\langle x^*, e\rangle| > \delta\}. \tag{6}$$

We observe that (6) defines a subanalytic subset of $\mathbb{R}^n$. Let us prove by contradiction that this set is finite.

Indeed, if this were not the case, then by using the subanalyticity of Γ_δ, there would exist $a < b$ in $[0,1]$ such that $(a,b)e \subset \Gamma_\delta$. Let V be an open bounded subset of U such that $[0,1]e \subset V \subset \mathrm{cl}\, V \subset U$ and define

$$\hat{\Gamma}_\delta^+ = \{x \in \mathrm{cl}\, V : \exists x^* \in \hat{\partial} f(x), \langle x^*, e\rangle > \delta\}, \quad \hat{\Gamma}_\delta^- = \{x \in \mathrm{cl}\, V : \exists x^* \in \hat{\partial} f(x), \langle x^*, e\rangle < -\delta\},$$

where $\hat{\partial} f$ denotes the Fréchet subdifferential of f (Definition 1(i)). Since $0 \in \partial^\circ f(x) = \overline{\mathrm{co}}\, \partial f(x)$ for every $x \in \Gamma_\delta$, we have that

$$\max\{\langle x^*, e\rangle : x^* \in \partial f(x)\} > \delta \quad \text{and} \quad \min\{\langle x^*, e\rangle : x^* \in \partial f(x)\} < -\delta.$$

So using the definition of the limiting subdifferential we obtain that $(a,b)e \subset \mathrm{cl}\, \hat{\Gamma}_\delta^+$ and $(a,b)e \subset \mathrm{cl}\, \hat{\Gamma}_\delta^-$.

Let us set $l = b - a$ and $M := \sup\{||x^*|| : x^* \in \partial^\circ f(x), x \in \mathrm{cl}\, V\}$. The finiteness of M comes from the Lipschitz continuity property of f (see Remark 1 (b) for instance) and the compactness of cl V. The function $t \to f(te)$ is subanalytic and continuous, hence absolutely continuous ([4, Lemma 5]). Thus by using relation (3) of Proposition 3 (subdifferential chain rule), we obtain that

$$\int_u^v |\frac{d}{dt} f(te)| dt \leq (v-u) \sup\{|\langle e, x^*\rangle| : t \in [u,v], x^* \in \partial f(te)\} \leq (v-u)M, \quad \text{for all} \quad 0 \leq u \leq v \leq 1.$$

Take $\eta > 0$ and apply Lemma 5 (path perturbation lemma) for $F = \hat{\Gamma}_\delta^+$, and $\gamma(t) = te$, $t \in (a,b)$. Since $\dot{\gamma}(t) = e$, for all $t \in [0,1]$, it follows that there exists a subanalytic continuous curve

$$z : [a,b] \to \mathrm{cl}\, \hat{\Gamma}_\delta^+$$

such that

- $||\dot{z}(t) - e|| < \eta$ for almost all $t \in (a,b)$,
- the (subanalytic) set $\Delta := \left\{t \in [a,b] : z(t) \in \mathrm{cl}\hat{\Gamma}_\delta^+ \setminus \hat{\Gamma}_\delta^+\right\}$ has a Lebesgue measure less than η,
- $z(t) = \gamma(t)$, for all $t \in \Delta \cup \{a,b\}$.

The continuous function $g(t) = f(z(t))$ is also subanalytic so for all but finitely many t's in $(a,b) \setminus \Delta$ we conclude from relation (2) of Proposition 3 and Remark 1 (c) that

$$\{g'(t)\} = \hat{\partial} g(t) \supset \langle \dot{z}(t), \hat{\partial} f(z(t))\rangle \supset \{\langle \dot{z}(t), z_+^*(t)\rangle\},$$

where $z_+^*(t) \in \hat{\partial} f(z(t))$ can be chosen in order to satisfy $\langle e, z_+^*(t)\rangle > \delta$ (since $z(t) \in \hat{\Gamma}_\delta^+$). Thus for almost all t in $[a,b] \setminus \Delta$ we have

$$g'(t) = \langle e, z_+^*(t)\rangle + \langle \dot{z}(t) - e, z_+^*(t)\rangle \geq \delta - ||\dot{z}(t) - e||M \geq \delta - \eta M,$$

so that

$$\begin{aligned} f(be) - f(ae) &= \int_a^b \frac{d}{dt} f(z(t)) dt \\ &\geq \int_{[a,b]\setminus\Delta} g'(t) dt - \int_\Delta |\frac{d}{dt} f(z(t))| dt \\ &\geq (l - \eta)(\delta - \eta M) - \eta M. \end{aligned}$$

By choosing η small enough, the above quantity can be made positive so that $f(be) > f(ae)$. It suffices to repeat the argument with $\hat{\Gamma}_\delta^-$ to obtain $f(be) < f(ae)$, which yields a contradiction.

Thus the set Γ_δ is finite. We further set

$$\Gamma_0 = \{x \in [0,1]e : \exists x^* \in \partial f(te), \langle x^*, e\rangle = 0\}.$$

It follows easily from Definition 1 (ii) that the limiting subdifferential ∂f has closed values. Thus, the set $\partial f(te)$ is closed for every $t \in [0,1]$, which yields

$$[0,1]e = \Gamma_0 \cup \left(\cup_{i\geq 1}\Gamma_{1/i}\right).$$

Note that $\cup_{i\geq 1}\Gamma_{1/i}$ is countable and equal to the subanalytic set $[0,1]e\diagdown\Gamma_0$. It follows that $\cup_{i\geq 1}\Gamma_{1/i}$ is finite and so $\{t \in [0,1], te \in \Gamma_0\}$ is a finite union of intervals with a finite complement in $[0,1]$. Using the continuity of f, it suffices to prove that f is constant on each segment of Γ_0.

Let $(a,b)e \subset \Gamma_0$ with $0 \leq a < b \leq 1$. For any $\epsilon > 0$ we define

$$\hat{\Gamma}_0^\epsilon := \{x \in \operatorname{cl} V : \exists x^* \in \hat{\partial} f(x), |\langle x^*, e\rangle| < \epsilon\}.$$

By definition of the limiting subdifferential, $(a,b)e \subset \operatorname{cl}\hat{\Gamma}_0^\epsilon$. Applying Lemma 4 for the set $\hat{\Gamma}_0^\epsilon$, for $\eta < \epsilon$ and for the path $\gamma(t) = te$, we obtain a curve $z : [a,b] \to \hat{\Gamma}_0^\epsilon$ and a set $\Delta \subset [a,b]$ satisfying (i), (ii), (iii) of Lemma 4. Set $h(t) = f(z(t))$. As before, for all but finitely many t's in $[a,b]\setminus\Delta$: $\{h'(t)\} = \{\langle \dot{z}(t), z_\epsilon^*(t)\rangle\}$, where $z_\epsilon^*(t) \in \hat{\partial} f(z(t))$ can be taken such that $|\langle z_\epsilon^*(t), e\rangle| < \epsilon$. Therefore for almost all t in $[a,b]\setminus\Delta$ we have

$$|h'(t)| = |\langle e, z_+^*(t)\rangle + \langle \dot{z}(t) - e, z_+^*(t)\rangle| \leq \epsilon + \eta M,$$

so that

$$\begin{aligned} |f(be) - f(ae)| &\leq \int_a^b |\frac{d}{dt} f(z(t))|dt \\ &\leq |\int_{[a,b]\setminus\Delta} |h'(t)| + \int_\Delta |\frac{d}{dt} f(z(t))dt| \\ &\leq (l-\eta)(\epsilon + \eta M) + \eta M. \end{aligned}$$

Taking ϵ (and thus η) sufficiently small, we see that the function f is constant on $[0,1]e$ and the proof is complete. □

Theorem 7 (main result) *Let U be a nonempty open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ a locally Lipschitz subanalytic mapping. Let S denote the set of Clarke critical points of f, that is*

$$S := \{x \in U : \partial^\circ f(x) \ni 0\}.$$

Then f is constant on each connected component of S.

Proof. Let x, y belong to the same connected component of S. It is sufficient to prove that $f(x) = f(y)$. Since $S = (U \times \{0\}_n) \cap \operatorname{Graph}\partial^\circ f$, we conclude by Remark 1(d) that it is a subanalytic set, so every connected component of S is also path-connected (see [1], [2] or [8], for example). Thus, there exists a continuous subanalytic path $\gamma : [0,1] \to S$ joining x to y. To prove that $f(x) = f(y)$ it suffices to prove that f is constant on $\gamma(0,1)$. By using the subanalyticity of γ together with the continuity of f, we can assume that

- $\gamma(0,1)$ is a subanalytic submanifold of U,

[Indeed, since $\gamma(0,1)$ is a finite union of subanalytic manifolds, we can deal with each one separately obtaining (as will be described below) that f is constant on each such manifold. Then the same conclusion will follow for $\gamma(0,1)$ by a continuity argument.]

- there exists a subanalytic diffeomorphism G from a neighbourhood V of $\gamma(0,1)$ into an open subset of $\mathbb{R}^n$ such that $G(\gamma(0,1)) = (0,1)e$; see [2] for instance.

In view of relation (4) of Proposition 3 we have that

$$\gamma(0,1) \subset (\partial^\circ f)^{-1}(0) \quad \text{if and only if} \quad (0,1)e \subset [\partial^\circ (f \circ G^{-1})]^{-1}(0).$$

This is indeed a consequence of the equivalence

$$\partial^\circ f(x) \ni 0 \Leftrightarrow \partial^\circ [f \circ G^{-1}](G(x)) \ni 0, \quad \text{for all } x \in V.$$

As a consequence f is constant on $\gamma(0,1)$ if and only if $f \circ G^{-1}$ is constant on $(0,1)e$. The conclusion follows then from Lemma 6. $\square$

Corollary 8 (Sard theorem for Clarke critical points) *Under the assumptions of Theorem 7 the set $f(S)$ of the Clarke critical values of f is countable (and hence has measure zero).*

Proof. This follows from Theorem 7 and the fact that the set S, being subanalytic, has at most a countable number of connected components (a finite number on each compact subset of U). $\square$

Let us finally conclude with the following remark.

Remark 3 (an easy proof for the case of subdifferential regularity) If f is assumed to be subdifferentially-regular (see Remark 1(b)) then Theorem 7 follows via a straightforward application of [16, Theorem 10.6].

Let us recall the simple argument (see also [3, Remark 3.2]): Assume that x, y are in the same connected component of S. Let $z : [0,1] \to S$ be a continuous subanalytic path with $z(0) = x$ and $z(1) = y$ and define the subanalytic function $h(t) = (f \circ z)(t)$. From the "monotonicity lemma" (see [9, Fact 4.1], or [11, Lemma 2], for example) we get $h'(t) = 0$, for all $t \in [0,1] \diagdown F$ where F is a finite set. Since $0 \in \hat{\partial} f(z(t))$ for all $t \in [0,1]$, using the chain rule for the Fréchet subdifferential we obtain

$$\{h'(t)\} = \hat{\partial} h(t) \sqsupseteq z'(t) \hat{\partial} f(z(t)) \sqsupseteq \{0\},$$

for all $t \in [0,1] \diagdown F$. It follows that h is constant on $[0,1]$, whence $f(x) = f(y)$.

4 An example of a continuous subanalytic function which is not constant on the set of its broadly critical points

In this section we assume that $f : \mathbb{R}^n \to \mathbb{R}$ is continuous. In such a case the definition of the Clarke subdifferential (1) of f at $x \in \mathbb{R}^n$ is as follows:

$$\partial^\circ f(x) = \overline{\text{co}} \{\partial f(x) + \partial^\infty f(x)\} \tag{7}$$

where $\partial^\infty f(x)$ is the *asymptotic limiting subdifferential* of f at x, that is the set of all $y^* \in \mathbb{R}^n$ such that there exists $\{t_n\}_n \subset \mathbb{R}_+$ with $\{t_n\} \searrow 0_+$, $\{y_n\}_n \subset \mathbb{R}^n$, $y_n^* \in \hat{\partial} f(x_n)$ such that $y_n \to x$ and $t_n y_n^* \to y^*$. When f is locally Lipschitz continuous, the local boundedness of the limiting-subgradients (Remark 1 (b)) entails $\partial^\infty f(x) = 0$, and so the above definition is - of course - compatible with Definition 1 (iii).

Following the terminology of [5], let us now introduce the convex-stable subdifferential. For every $x \in \mathbb{R}^n$ set

$$T_f(x) = \bigcap_{\varepsilon > 0} \overline{\text{co}} \left\{ \bigcup_{x \in B(x_0, \varepsilon)} \hat{\partial} f(x) \right\}. \tag{8}$$

A point $x_0 \in \mathbb{R}^n$ is called *broadly critical point* for f if

$$0 \in T_f(x_0). \tag{9}$$

Proposition 9 *(i) For any continuous function f we have*

$$\partial^\circ f(x) \subset T_f(x), \quad \textit{for all } x \in \mathbb{R}^n.$$

(ii) If f is a locally Lipschitz function, then

$$\partial^\circ f(x) = T_f(x), \quad \textit{for all } x \in \mathbb{R}^n.$$

Consequently, in this case Clarke critical and broadly critical points coincide.

Proof. (i) Since for every $x \in \mathbb{R}^n$ the set $T_f(x)$ is closed and convex, it is clearly sufficient to show that

$$\partial f(x) + \partial^\infty f(x) \subset T_f(x).$$

To this end, fix $\varepsilon > 0$ and let $x^* \in \partial f(x)$ and $y^* \in \partial^\infty f(x)$. Then there exist $(x_n, x_n^*) \in \operatorname{Graph} \hat{\partial} f$ $(y_n, y_n^*) \in \operatorname{Graph} \hat{\partial} f(x_n)$, $\{t_n\}_n \subset \mathbb{R}_+$ with $\{t_n\} \searrow 0_+$, such that $x_n \to x$, $y_n \to x$, $x_n^* \to x^*$ and $t_n y_n^* \to y^*$. For n sufficiently large, we have $0 < t_n < 1$ and $x_n, y_n \in B(x, \varepsilon)$. It follows that

$$p_n := (1 - t_n) x_n^* + t_n y_n^* \in T_\varepsilon(x) = \overline{\text{co}} \left\{ \bigcup_{x' \in B(x,\varepsilon)} \hat{\partial} f(x') \right\},$$

thus $x^* + y^* = \lim_n p_n \in T_\varepsilon(x)$. It follows that $x^* + y^* \in T_f(x)$ and the assertion follows.

(ii) It remains to show that if f is locally Lipschitz then $T_f(x) \subset \partial^\circ f(x)$. Set

$$H_\varepsilon(x) = \left\{ \bigcup_{x' \in B(x,\varepsilon)} \hat{\partial} f(x') \right\}$$

and note that since f is locally Lipschitz, $\overline{H_\varepsilon(x)}$ is (nonempty and) compact. Thus, $T_\varepsilon(x) = \overline{\text{co}}\, H_\varepsilon(x) = \text{co}\, \overline{H_\varepsilon(x)}$. Let now any $p \in T_f(x)$. Then by the Caratheodory theorem, for every $\varepsilon > 0$ there exist $\{x_{1,\varepsilon}, \ldots, x_{n+1,\varepsilon}\} \subset B(x, \varepsilon)$, $x_{i,\varepsilon}^* \in \hat{\partial} f(x_{i,\varepsilon})$ and $\{\lambda_{1,\varepsilon}, \ldots \lambda_{n+1,\varepsilon}\} \subset \mathbb{R}_+$ with $\sum_i \lambda_{i,\varepsilon} = 1$ such that $p = \lim_{\varepsilon \searrow 0^+} p_\varepsilon$ where $p_\varepsilon = \sum_i \lambda_{i,\varepsilon} x_{i,\varepsilon}^*$. Using a standard compactness argument, we may assume as $\varepsilon \to 0^+$ that for every $i \in \{1, \ldots, n+1\}$ we have $x_{i,\varepsilon} \to x, x_{i,\varepsilon}^* \to x_i^* \in \partial f(x)$ and $\lambda_{i,\varepsilon} \to \lambda_i$. It follows that $p = \sum_i \lambda_i x_i^* \in T_f(x)$ and the assertion follows. □

We now provide an example showing that the conclusion of Theorem 7 (Main result) is no more valid for the set of broadly critical points of a continuous subanalytic function. More precisely, **there exists a continuous subanalytic function** $f : \mathbb{R}^3 \to \mathbb{R}$ **which is not constant on a segment of broadly critical points** (see Fact 1-3 below).

Construction of the example

Let us consider the function $\theta_0 : [0, \pi) \to [0, \pi/2]$ defined by

$$\theta_0(z) := \begin{cases} z, & \text{if} \quad 0 \le z \le \pi/2, \\ \pi - z, & \text{if} \quad \pi/2 < z < \pi. \end{cases}$$

We extend the domain of θ_0 from $[0, \pi)$ to $\mathbb{R}$ in the following way:

$$z \longmapsto \tilde{\theta}_0(z) := \theta_0(z \,(\text{mod}\ \pi)).$$

Then for every $(\theta, z) \in [0, \frac{\pi}{2}] \times \mathbb{R}$ we define:

$$\sigma(\theta, z) := \begin{cases} 1, & \text{if } \theta \geq \tilde{\theta}_0(z), \\ -1, & \text{if } \theta < \tilde{\theta}_0(z). \end{cases}$$

Finally, for every $(\rho, \theta, z) \in \mathbb{R}_+^* \times [0, \frac{\pi}{2}] \times \mathbb{R}$ we set:

$$\Phi_1(\rho, \theta, z) = \begin{cases} (\frac{2}{\pi})\, \tilde{\theta}_0(z) + \sigma(\theta, z)\, \rho, & \text{if } \rho \leq (\frac{2}{\pi})\, |\theta - \tilde{\theta}_0(z)|, \\ \\ (\frac{2}{\pi})\, \theta, & \text{if } \rho > (\frac{2}{\pi})\, |\theta - \tilde{\theta}_0(z)|. \end{cases} \tag{10}$$

Now for $(\rho, \theta, z) \in \mathbb{R}_+^* \times [0, \pi) \times \mathbb{R}$ we set:

$$\Phi_2(\rho, \theta, z) = \begin{cases} \Phi_1(\rho, \theta, z), & \text{if } 0 \leq \theta \leq \pi/2 \\ \Phi_1(\rho, \pi - \theta, z), & \text{if } \pi/2 < \theta \leq \pi. \end{cases}$$

Finally we define $\Phi : \mathbb{R}_+^* \times [0, 2\pi) \times \mathbb{R} \to [0, 1]$ by

$$\Phi(\rho, \theta, z) = \begin{cases} \Phi_2(\rho, \theta, z), & \text{if } 0 \leq \theta \leq \pi, \\ \Phi_2(\rho, \theta - \pi, z), & \text{if } \pi < \theta < 2\pi. \end{cases} \tag{11}$$

Let us define $f : \mathbb{R}^3 \to [0, 1]$ as the function whose graph in cartesian coordinates is the one of Φ in cylindrical coordinates. For instance, for any $x, y > 0$ we have

$$f(x, y, z) \;=\; \Phi(\sqrt{x^2 + y^2},\, \arctan{(\frac{y}{x})}\,,\, z).$$

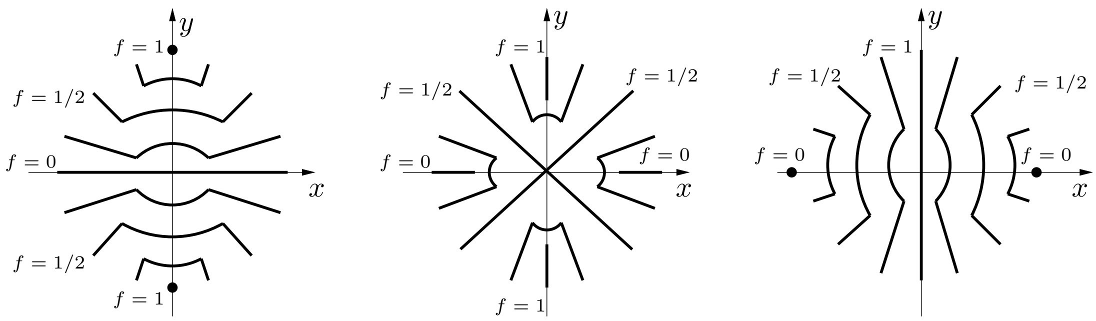

Figure 1: Level sets of the function $f(\cdot, \cdot, z)$, for $z = 0, \frac{\pi}{4}, \frac{\pi}{2}$

Fact 1. The function f is continuous and subanalytic.

Fact 2. The restriction of f on the set $Z = \{(0, 0, z) : z \in \mathbb{R}\}$ is not constant.

Fact 3. Every point of Z is broadly critical, that is, $Z \subset \{u \in \mathbb{R}^3 : T_f(u) \ni 0\}$.

Proof. Facts 1 and 2 follow readily from the definition of f. To establish the third point it is sufficient to prove that if $0 < z_0 < (\pi/2)$, then $0 \in \partial^\circ f((0, 0, z_0))$.

To this end, set $u_0 = (0, 0, z_0)$, $\theta_0 = \tilde{\theta}_0(z_0)$ (so that $0 < \theta_0 < \pi/2$) and let

$$\theta_n = \theta_0 + \frac{\pi}{2^{n+2}} \tag{12}$$

(so that $\theta_n \searrow \theta_0$). Then set $a_n = \tan\theta_n$ and

$$x_n = \frac{1}{2^n\sqrt{1+a_n^2}} \tag{13}$$

(so that $x_n \searrow 0$), $y_n := a_n x_n$ and thus

$$\rho_n = \sqrt{x_n^2 + y_n^2} = (\sqrt{1+a_n^2})x_n = \frac{1}{2^n}. \tag{14}$$

For every $n \geq 1$ we define

$$u_n := (x_n, y_n, z_0) \quad \text{and} \quad \bar{u}_n = (-x_n, -y_n, z_0).$$

In view of (10), (13) and (14), the sequences $\{u_n\}_{n\geq 1}$, $\{\bar{u}_n\}_{n\geq 1} \subset \mathbb{R}^3$ converge to u_0 and satisfy

$$f(u_n) = f(\bar{u}_n) = \Phi(\rho_n, \theta_n, z_0) = (\frac{2}{\pi})\theta_n.$$

By (11) and (10) it is easily seen that f is differentiable at u_n (respectively, at $\bar{u}_n$). Precisely, we have

$$\frac{\partial\Phi}{\partial\rho}(u_n) = \frac{\partial\Phi}{\partial z}(u_n) = 0$$

and

$$\frac{\partial\Phi}{\partial\theta}(u_n) = \frac{2}{\pi},$$

so we conclude that

$$\nabla f(u_n) = \frac{2}{\pi}(\frac{-y_n}{x_n^2+y_n^2}, \frac{x_n}{x_n^2+y_n^2}, 0).$$

Repeating the above for the sequence $\{\bar{u}_n\}_{n\geq 1}$ we obtain

$$\nabla f(\bar{u}_n) = -\nabla f(u_n),$$

or in other words,

$$0 \in \bigcap_{\varepsilon>0} \mathrm{co}\,\{\nabla f(u) : u \in B(u_0, \varepsilon) \cap D_f\}$$

where D_f denotes the points of differentiability of f. This shows that the point u_0 is broadly critical. $\square$

Acknowledgement. The first two authors wish to thank the CMM (Santiago of Chile) for its hospitality and financial support. The second author wishes also to thank the Universities of Nagoya and Saitama (Japan), the University of Savoie (France) for hospitality and financial support and K. Kurdyka, P. Orro, T. Fukui, L. Rifford, and N. Hadjisavvas for useful discussions.

References

[1] Benedetti, R. & Risler, J.-J., *Real Algebraic and Semialgebraic Sets*, Hermann, Éditeurs des sciences et des Arts, (Paris, 1990).

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

[3] Bolte, J., Daniilidis, A. & Lewis, A.S., The Łojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems, preprint, 22 pages, 2004.
Available electronically at: `http://pareto.uab.es/~adaniilidis/`

[4] Bolte, J., Daniilidis, A. & Lewis, A.S., The Morse-Sard theorem for non-differentiable subanalytic functions, preprint, 12 pages, 2004.
Available electronically at: http://pareto.uab.es/~adaniilidis/

[5] Burke, J.V., Lewis A.S., & Overton M.L., Approximating subdifferentials by random sampling of gradients, *Math. Oper. Res.* **27**, N. 3, (2002), 567–584.

[6] Clarke, F.-H., *Optimization and Nonsmooth Analysis*, Wiley-Interscience, (New York, 1983).

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

[8] Coste, M., *An Introduction to O-minimal Geometry*, RAAG Notes, 81 pages, Institut de Recherche Mathématiques de Rennes, November 1999.

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

[10] Itoh, J.-I. & Tanaka, M., A Sard theorem for the distance function, *Math. Ann.*, **320**, (2001), 1–10.

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

[12] Kurdyka, K., Orro, P. & Simon, S., Semialgebraic Sard theorem for generalized critical values, *J. Differential Geom.* **56** (2000), 67–92.

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

[14] Pawłucki, W., Le théorème de Puiseux pour une application sous-analytique, *Bull. Polish Acad. Sci. Math.*, **32** (1984), 555-560.

[15] Rifford, L., A Morse-Sard theorem for the distance function on Riemannian manifolds, *Manuscripta Math.* **113** (2004), 251-265.

[16] Rockafellar, R.T. & Wets, R., *Variational Analysis*, Grundlehren der Mathematischen, Wissenschaften, Vol. **317**, (Springer, 1998).

[17] Shiota, M, *Geometry of Subanalytic and Semialgebraic Sets*, Progress in Mathematics 150, Birkhäuser, (Boston, 1997).

[18] Yomdin, Y., The geometry of critical and near-critical values of differentiable mappings. *Math. Ann.* **264** (1983), 495–515

Jérôme BOLTE (bolte@math.jussieu.fr ; http://www.ecp6.jussieu.fr/pageperso/bolte/)
Equipe Combinatoire et Optimisation (UMR 7090), Case 189, Université Pierre et Marie Curie
4 Place Jussieu, 75252 Paris Cedex 05.

Aris DANIILIDIS (arisd@mat.uab.es ; http://mat.uab.es/~arisd)
Departament de Matemàtiques, C1/320
Universitat Autònoma de Barcelona
E-08193 Bellaterra (Cerdanyola del Vallès), Spain.

Adrian LEWIS (aslewis@orie.cornell.edu; http://www.orie.cornell.edu/~aslewis)
School of Operations Research and Industrial Engineering
Cornell University
234 Rhodes Hall, Ithaca, NY 14853, United States.

Masahiro SHIOTA (shiota@math.nagoya-u.ac.jp)
Department of Mathematics
Nagoya University (Furocho, Chikusa)
Nagoya 464-8602, Japan.