View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

J. Math. Anal. Appl. 335 (2007) 882–901

Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS

www.elsevier.com/locate/jmaa

A Sard theorem for tame set-valued mappings

A. Ioffe

Department of Mathematics, Technion, Haifa 32000, Israel

Received 17 August 2006

Available online 15 February 2007

Submitted by B.S. Mordukhovich

Abstract

If F is a set-valued mapping from $\mathbb{R}^n$ into $\mathbb{R}^m$ with closed graph, then $y \in \mathbb{R}^m$ is a critical value of F if for some x with $y \in F(x)$, F is not metrically regular at (x, y). We prove that the set of critical values of a set-valued mapping whose graph is a definable (tame) set in an o-minimal structure containing additions and multiplications is a set of dimension not greater than $m - 1$ (respectively a σ-porous set). As a corollary of this result we get that the collection of asymptotically critical values of a set-valued mapping with a semialgebraic graph has dimension not greater than $m - 1$. We also give an independent proof of the fact that a definable continuous real-valued function is constant on components of the set of its subdifferentially critical points.

© 2007 Elsevier Inc. All rights reserved.

Keywords: o-Minimal structure; Definable set-valued mapping; Rate of surjection; Critical value

1. Introduction

The classical Sard (or Morse–Sard) theorem states that the collection of critical values of a C^k-mapping F from (an open subset of) $\mathbb{R}^n$ into $\mathbb{R}^m$ has Lebesgue measure zero, provided $k \geqslant \max\{n - m + 1, 1\}$. The fundamental role of the Sard theorem in analysis and differential geometry comes from the fact that for a regular (noncritical) value y of F the set of solutions of the equation $F(x) = y$ (if nonempty) is a nice set (a manifold) which responds to variations of the right-hand side in a stable and nonchaotic way. The Sard theorem therefore ascertains that a typical value of a sufficiently smooth mapping is regular.

E-mail address: ioffe@math.technion.ac.il.

0022-247X/$ – see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmaa.2007.01.104

Such a result would be highly welcome in variational analysis in which the stability issue is of an extreme importance. It would be highly desirable to be able to make similar statements e.g. about systems of inequalities or other relations of interests in variational analysis.

At the first glance this does not seem to be possible. The Sard theorem is sharp and there are widely known examples (Whitney [24], Yomdin [27]) showing that for a less smooth function or mapping Sard's theorem does not hold. The most precise result was proved by Bates [3]: Sard's theorem holds for $C^{n-m,1}$-mappings ($n-m$ times continuously differentiable with locally Lipschitz derivatives of order $n-m$). Here "Lipschitz" cannot be strengthened to "Hölder" as was found by Norton [18].

However, recently Kurdyka, Orro and Simon [14] proved that the collection of critical and asymptotically critical values of a semialgebraic C^1-mappings is a semialgebraic set of dimension $m-1$ or less. Several results of Morse–Sard-type were proved for real-valued functions under even more general assumptions: [22,25] (quantitative results for maximum and minimax of smooth families of functions), [19] (distance function to a C^∞-submanifold of a Riemann manifold), [1] (generalized critical values of a C^1-function definable in an o-minimal structure), [4–6] (critical points of globally subanalytic functions). The last three papers have largely stimulated this study.

These results demonstrate that the differentiability requirement can be substantially weakened in exchange for some structural restrictions on the class of mappings or functions. The main result of this paper shows that the frameworks of this trade-off can be considerably expanded. This is the statement of the main theorem.

Theorem 1. *If $F:\mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is a tame set-valued mapping with locally closed graph, then the set of critical values of F is a σ-porous*[1] *set in $\mathbb{R}^m$. In particular it has Lebesgue measure zero. Moreover, if the graph of F is a definable set, then the set of critical values of F is also a definable set of dimension not exceeding $m-1$.*

The next two sections contain all information from variational analysis and the theory of o-minimal structures which is necessary for the proof of the theorem. Here we shall only add a couple of general remarks.

First we note that the concept of a "critical value" provided by modern variational analysis is very natural. (Actually this concept seems to be defined here for the first time. However the "parent" concept of (metric) regularity has been thoroughly studied during last two decades, see e.g. [13,17,21].) We observe that restricted to single-valued continuously differentiable mappings, this definition reduces to the classical concept: y is a critical value of F if there is an x such that $F(x)=y$ and the rank of the derivative $\nabla F(x)$ is smaller than m.

Definability and tameness are fundamental concepts of the theory of o-minimal structures (see e.g. [9,11,23]) which is also being very actively developed last two decades, partly in response to Grothendieck's call for a new "tame topology" based on "a real look ... at the context at which we live, breathe and work" [12]. The important point about definable and tame objects is that they are void of "pathologies" so typical for generic objects of nonsmooth analysis (e.g. Lipschitz functions that cannot be recovered from their subdifferentials).

[1] A set Q in a metric space is called *porous* if there is $\lambda>0$ such that for any $x\in Q$ and any $r>0$ the set $B(x,r)\setminus Q$ contains a ball of radius λr. A σ-porous set is a countable union of porous sets. A σ-porous set in $\mathbb{R}^n$ is both of the first Baire category and Lebesgue measure zero.

A search for good classes of nonsmooth functions ("less subject to wildness," I would add, again quoting [12]) for which certain results could be proved, was among the dominant themes in nonsmooth analysis since practically its very beginning. Just mention lower C^2-functions [20], amenable functions, prox-regular functions [21], semismooth functions [16], minimal cuscos [8], partially smooth functions [15]. Functions of these classes well serve for the purposes they have been created, but none of the classes have structural properties compared to the properties of definable and tame functions and sets (which need not be differentiable or even continuous). Thanks to these properties, definable and tame functions and sets look like an almost ideal playground for applicable finite dimensional variational analysis.

It has to be emphasized that this paper is addressed mainly to the variational analysis community (to which the author belongs) and the last sentence of the previous paragraph carries one of the main messages. For that reason the introductory material relating to variational analysis in the next section is presented in a much more sketchy way than the information about definable sets and functions in Section 3 with which the variational analysis community (and the author) are less familiar.

Theorem 1 is proved in Section 5. In Section 4 we state and prove some preliminary results needed for the proof of the theorem. Some of them are probably new but some (e.g. definability of derivative) are known. We give short proofs of the latter as well, just for convenience. The principal results here are Proposition 1 (which builds a bridge between the two parts by showing that the "rate of surjection" which is a quantitative measure of regularity is a definable or tame function, provided the graph of the set-valued mapping is respectively definable or tame) and Proposition 6 (showing that for certain definable families of functions, uniform smallness of functions implies smallness of derivatives on big sets). The last Section 6 contains statements and proofs of some corollaries.

Specifically, as a consequence of Theorem 1, we recover the part of the theorem of Kurdyka, Orro and Simon [14] relating to the dimension of the set of asymptotic critical values of semialgebraic mappings (not the fibration part of the theorem). Actually we get an extension of this theorem to set-valued mappings with semialgebraic graphs. We get this result as a part of a more general theorem in which a "stratification" of asymptotically critical values by rates of asymptotic decline of the rates of surjection is taken into consideration. We also give a separate proof of an o-minimal extension of the recent result by Bolte, Daniilidis and Lewis [4] saying that a globally analytic function which is continuous on its domain is constant on connected components of the set of its critical points.

2. Openness, regularity and critical values

The concept defined below makes sense in every Banach, and actually in every metric space. This level of generality is not needed here, so we define everything for finite dimensional Euclidean spaces and refer the reader to [13] for the general theory.

So let again $F:\mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a set-valued mapping. For $(x, y) \in \operatorname{Graph} F$ we set

$$\operatorname{Sur} F(x, y)(\lambda) = \sup\bigl\{r \geqslant 0:\ y + rB_Y \subset F(x + \lambda B_X)\bigr\}$$

and then for $(\bar{x}, \bar{y}) \in \operatorname{cl}(\operatorname{Graph} F)$ (the closure of $\operatorname{Graph} F$) define the *rate of openness* (*surjection*) of F at $(\bar{x}, \bar{y})$ by

$$\operatorname{sur} F(\bar{x}|\bar{y}) = \liminf_{(x,y,\lambda)\to(\bar{x},\bar{y},+0)} \frac{1}{\lambda} \operatorname{Sur} F(x, y)(\lambda).$$

For single-valued F we usually write $\operatorname{sur} F(\bar{x})$ (instead of $\operatorname{sur} F(\bar{x}|F(\bar{x}))$).

The function $\lambda \to \operatorname{Sur} F(x|y)(\lambda)$ is called the *modulus of surjection* of F at $(\bar{x}, \bar{y})$. Here, as usual, B_X etc. is the unit ball in $\mathbb{R}^n$ and we set $\sup \emptyset = 0$ (or else, we can calculate the lim inf in the definition of $\operatorname{Sur} F$ only along sequences of $(x, y) \in \operatorname{Graph} F$).

It follows from the definition that the function $\operatorname{sur} F$ is defined on the closure of Graph F. If however F is a set-valued mapping with closed graph, then it is possible to show that[2]

$$\operatorname{sur} F(\bar{x}|\bar{y}) = \liminf_{(x,y) \underset{\operatorname{Graph} F}{\longrightarrow} (\bar{x},\bar{y})} \liminf_{\lambda \to +0} \frac{1}{\lambda} \operatorname{Sur}(x, y)(\lambda).$$

In general, the quantity in the right-hand side of the equality can by greater. Consider for instance the following mapping $\mathbb{R} \rightrightarrows \mathbb{R}$:

$$F(x) = \{y\colon\ 0 < |y| < |x|\}.$$

Then $\operatorname{sur} F(0, 0) = 0$ (take $x_n = \lambda_n = n^{-1}$, $y_n = n^{-2}$) but the right-hand side quantity is equal to ∞.

The reciprocal of $\operatorname{sur} F(\bar{x}|\bar{y})$ is the *rate of metric regularity* of F at $(\bar{x}, \bar{y})$:

$$\operatorname{reg} F(\bar{x}|\bar{y}) = \big[\operatorname{sur} F(\bar{x}|\bar{y})\big]^{-1}.$$

This is a quantitative measure of stability of solution $\bar{x}$ of $y \in F(x)$ at $y = \bar{y}$ for it is precisely the lower bound of positive K such that

$$d\big(x, F^{-1}(y)\big) \leqslant K d\big(y, F(x)\big)$$

for all (x, y) of a neighborhood of $(\bar{x}, \bar{y})$.

It is said that F is *regular* at (x, y) or that (x, y) is a *regular point* of F if $\operatorname{sur} F(x|y) > 0$.

Otherwise (x, y) is a *singular point* of F. Finally, y is a *singular* or *critical* value of F if there is an x such that $y \in F(x)$ and $\operatorname{sur} F(x|y) = 0$.

Remark. Observe that the above definition of a critical value covers both the case of a "proper" critical value when (x, y) belongs to the graph of F and of a "generalized" critical value when (x, y) belongs to the closure of Graph F but not to the graph of F itself. In principle, if we impose no topological conditions on F, it may happen that a critical value of F (even proper) is a regular value of the set-valued mapping whose graph is the closure of Graph F. Consider, for instance the following set-valued mapping $F : \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$:

$$F(0) = \{0\}; \qquad F(x) = \|x\| B \setminus (\text{the } x \text{ axis}), \quad \text{if } x \neq 0.$$

Then zero is a proper critical value of F but a regular value of the mapping whose graph is the closure of Graph F.

To avoid pathologies like that, we shall mainly consider set-valued mappings with locally closed graphs. Another reason for introducing such an assumption is that the all known regularity criteria need it.

The following known facts will be used in the sequel (see [13]):

- y is a critical value of F if and only if it is a critical value of the projection $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ restricted to Graph F;

[2] This fact probably has not been explicitly mentioned earlier but it easily follows from the slope characterization of the rate of surjection given in [13].

- $\operatorname{sur} F$ is a lower semicontinuous function;
- if $F(x) = Ax$ is a linear operator, then for any x

$$\operatorname{sur} A(x) = \operatorname{Sur} A(x)(1) = \inf_{\|y^*\|=1} \|A^* y^*\| = \left\| A^{*-1} \right\|^{-1},$$

  so that we can write just $\operatorname{sur} A$, etc.;
- if F is single-valued and continuously differentiable at x, then $\operatorname{sur} F(x) = \operatorname{sur}(\nabla F(x))$, where by $\nabla F(x)$ we denote the Jacobian matrix of F or/and the corresponding linear operator $\mathbb{R}^n \to \mathbb{R}^m$;
- if $F(x) = H(x) + A(x)$, where A is a linear operator, H is a set-valued mapping with locally closed graph and $y \in H(x)$, then $\operatorname{sur} F(x|y + A(x)) \geqslant \operatorname{sur} H(x) - \|A\|$;
- if $F = H \circ G$, where G is a C^1-mapping into the domain space of H and the graph of H is locally closed, then

$$\operatorname{sur} G(x) \cdot \operatorname{sur} H\big(G(x)|y\big) \leqslant \operatorname{sur} F(x|y) \leqslant \big\|\nabla G(x)\big\| \cdot \operatorname{sur} H\big(G(x)|y\big).$$

The most general regularity criterion (actually, the precise formula for the rate of regularity) is based on the concept of slope introduced by DeGiorgi, Marino and Tosques [10]. We shall state it only for single-valued mappings as it is sufficient here. We refer to the first chapter of [13] and to [2] for more details and generalities.

So let f be an extended-real-valued function which is finite at x. The *slope* of f at x is the upper bound of $K \geqslant 0$ such that $f(x) \geqslant f(u) - K\|x - u\|$ for all u of a neighborhood of x (with the convention $\sup \emptyset = 0$). The slope is usually denoted $|\nabla f|(x)$ to emphasize that for a Fréchet differentiable function the slope coincides with the norm of the gradient. Thus,

$$|\nabla f|(x) = \limsup_{\substack{u \to x \\ u \neq x}} \frac{(f(x) - f(u))^+}{d(x, u)}.$$

Given a mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ defined and continuous in a neighborhood of a certain $\bar{x}$, we set for any $y \in \mathbb{R}^m$, $f_y(x) = \|y - F(x)\|$. Then [13, Chapter 1, Theorem 2b] $\operatorname{sur} F(\bar{x}, \bar{y})$ *is the upper bound of* $\gamma \geqslant 0$ *having the property that there is an* $\varepsilon > 0$ *such that* $|\nabla f_y|(x) \geqslant \gamma$ *for all* x *and* y *satisfying* $\|x - \bar{x}\| < \varepsilon$, $y \neq F(\bar{x})$. This is the general regularity criterion for a single-valued mapping. (The same result will be obtained if, instead of all $y \neq F(\bar{x})$ we shall take $y \in U \setminus F(\bar{x})$, where U is an arbitrary neighborhood of $F(\bar{x})$.)

Let us now define the *slope of* F at x by

$$\operatorname{Sl} F(x) = \inf_{y \neq F(x)} |\nabla f_y|(x).$$

Then the general regularity criterion can be equivalently expressed as

$$\operatorname{sur} F(\bar{x}) = \liminf_{x \to \bar{x}} \operatorname{Sl} F(x).$$

It follows in particular that

$$\operatorname{sur} F(x) \leqslant \operatorname{Sl} F(x), \quad \forall x \in \operatorname{dom} F.$$

3. *o*-Minimal structures and definable functions

We give below the statements of main definitions and facts without proofs. There are two excellent introductions to the subject: [9,11]. General *o*-minimal structures can be associated with various linearly ordered sets but we here shall consider only structures associated with the real line $\mathbb{R}$. We shall keep the notation $\mathbb{R}^n$ also for Euclidean spaces and denote the inner product by $(\cdot|\cdot)$.

Definition 1. A *structure* on $\mathbb{R}$ is a sequence $\mathcal{S} = (\mathcal{S}_n)$ $(n \in \mathbb{N})$, such that for each n

(D1) $\mathcal{S}_n$ is a boolean algebra of subsets of $\mathbb{R}^n$, that is, $\emptyset \in \mathcal{S}_n$ and $\mathcal{S}_n$ contains unions, intersections and complements of its elements;
(D2) if $A \in \mathcal{S}_n$, then $A \times \mathbb{R}$ and $\mathbb{R} \times A$ belong to $\mathcal{S}_{n+1}$;
(D3) $\{x = (x_1, \dots, x_n)\colon x_i = x_j\} \in \mathcal{S}_n$ for any $1 \leqslant i < j \leqslant n$;
(D4) if $A \in \mathcal{S}_{n+1}$ then the projection $\pi(A)$ of A to $\mathbb{R}^n$ $(\pi : (x_1, \dots, x_n, x_{n+1}) \to (x_1, \dots, x_n))$ belongs to $\mathcal{S}_n$.

A structure is called *o-minimal* (short for "order minimal") if in addition

(D5) $\{(x, y) \in \mathbb{R}^2\colon x < y\} \in \mathcal{S}_2$;
(D6) the elements of $\mathcal{S}_1$ are precisely finite unions of points and open intervals.

The elements of $\mathcal{S}_n$, $n = 1, 2, \dots,$ are called *definable* (in $\mathcal{S}$). A set Q is called *tame* if its intersection with any bounded definable set is a definable set. A (set-valued) mapping F from a subset of $\mathbb{R}^n$ into $\mathbb{R}^m$ is called *definable* (*tame*) if its graph is a definable (tame) set. Likewise, a real-valued function defined on a subset of $\mathbb{R}^n$ is definable if its graph is a definable set in $\mathbb{R}^{n+1}$.

In variational analysis it is often convenient to work with extended-real-valued functions defined on all of $\mathbb{R}^n$. The definition can easily be extended to such functions: f is definable if its graph $\{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R}\colon \alpha = f(x)\}$ is definable along with the sets $\{x\colon f(x) = \infty\}$ and $\{x\colon f(x) = -\infty\}$.

The subtle points of the definition are (D4) and (D6). (D4) is usually the most difficult part of the proof that a certain collection of sets is a structure. A consequence of it is that any set obtained from definable sets with the help of finitely many existential and universal quantifiers $\exists$, $\forall$ (applied to variables only, not to sets, functions and other parameters) and boolean operations is also definable.

(D6) is basically responsible for a number of remarkable structural (tameness) properties of definable sets and functions (e.g. Monotonicity, Cell Decomposition and Definable Choice theorems stated below) which exclude any possibility of "wild" behavior.

Here are several examples of o-minimal structures most useful for analysis.

(1) Let us call a set $Q \subset \mathbb{R}^n$ an *open polyhedron* if it is the intersection of finitely many open half spaces $\{x\colon f(x) < 0\}$ and hyperplanes $\{x\colon f(x) = 0\}$, where $f(x) = (a|x) + \alpha$ are affine functions. The structure $\mathcal{S}_{\text{lin}}$ of *semilinear* sets is formed by finite unions of open polyhedrons. All axioms of o-minimal structures are easy to verify in this case.

(2) If we replace affine functions by polynomials, we obtain the structure $\mathcal{S}_{\text{alg}}$ of *semialgebraic sets*. Here again all axioms are verified easily, with the exception of (D4). The latter is the subject of a deep Tarski–Seidenberg theorem (see [7]). A consequence of this fact is that semialgebraic sets admit *elimination of quantifiers*, that is any set obtained from semialgebraic sets with the help of quantifiers and boolean operations can also be obtained by means of a quantifier free formula involving only level and sublevel sets of polynomials and boolean operations.

(3) The above scheme no longer works if we make a step further and replace polynomials by real analytic function. We can define *semianalytic* sets in the same way but using arbitrary

real analytic functions. However, in this case (D4) and (D6) do not hold. Indeed, the set of zeros of $\sin x$ is an infinite collection of isolated points and there are also examples of bounded semianalytic sets whose projections are not semianalytic.

Nonetheless there exists a rich o-minimal structure in which all bounded semianalytic sets are definable. A set $Q \subset R^n$ is called *subanalytic* if locally near each of its points it is a projection of a bounded semianalytic set, that is if for any $x \in Q$ there is an open neighborhood U of x and a bounded semianalytic $S \subset \mathbb{R}^{n+k}$ such that the projection of S to $\mathbb{R}^n$ coincides with $Q \cap U$. A set $Q \subset \mathbb{R}^n$ is called *globally subanalytic* if $F(Q)$ is subanalytic whenever F is a semialgebraic homeomorphism of $\mathbb{R}^n$ onto $(-1, 1)^n$.

It turns out that globally subanalytic sets already satisfy all axioms and the corresponding o-minimal structure is denoted $\mathcal{S}_{\mathrm{an}}$.

It is possible to give an alternative "nonconstructive" description for $\mathcal{S}_{\mathrm{lin}}$ and $\mathcal{S}_{\mathrm{alg}}$, namely $\mathcal{S}_{\mathrm{lin}}$ is the minimal structure containing graphs of affine functions and $\mathcal{S}_{\mathrm{alg}}$ is the minimal structure containing graphs of all polynomials. It turns out that $\mathcal{S}_{\mathrm{an}}$ is the minimal structure containing semialgebraic sets and graphs of restrictions of real analytic functions to balls.

(4) The minimal structure $\mathcal{S}_{\mathrm{an,exp}}$ containing all globally subanalytic sets and the graph of the exponent e^x is also o-minimal. Clearly

$$\mathcal{S}_{\mathrm{lin}} \subset \mathcal{S}_{\mathrm{alg}} \subset \mathcal{S}_{\mathrm{an}} \subset \mathcal{S}_{\mathrm{an,exp}}.$$

We shall further consider only o-minimal structures satisfying the additional property:

(D7) The graphs of addition $\{(x, y, z) \in \mathbb{R}^3\colon\ z = x + y\}$ and multiplication $\{(x, y, z) \in \mathbb{R}^3\colon\ z = x \cdot y\}$ belong to $\mathcal{S}_3$.

The semilinear structure does not belong to this class but every semilinear set is semialgebraic, so all results valid for the latter are also valid for semilinear sets.

The following are some simple properties of definable and tame sets and functions which are obtained from the axioms with relative easiness:

- the closure and the interior of a definable (tame) set is a definable (tame) set;
- a function is definable (tame) if and only if its epigraph $\{(x, \lambda)\colon\ \lambda \geqslant f(x)\}$ and hypograph $\{(x, \alpha)\colon\ a \leqslant f(x)\}$ are definable (tame) sets;
- the derivative (also partial) of a definable (tame) function is a definable (tame) function;
- under a suitable agreement about operations with infinite values (e.g. $\infty - \infty = \infty$; $0 \cdot \infty = \infty$, etc.) the collection of definable (tame) extended-real-valued functions is stable under summation, subtraction, multiplication and operations of pointwise maximum and minimum; composition of definable mappings is definable;
- if $f_1, \dots, f_k$ are definable (tame) functions and $\mathbb{R}^n$ is partitioned into k definable sets $X_1, \dots, X_k$, then the function f equal to f_i on X_i is definable (tame);
- the functions

  $$\inf_y f(x, y) \quad \text{and} \quad \sup_y g(x, y)$$

  are definable, provided so are f, g;
- the image and the preimage of a definable set under a definable mapping is a definable set; the image of a tame set under a proper tame mapping is a tame set.

And now several fundamental results characterizing the tameness properties of definable sets and functions.

Monotonicity Theorem. *Let f be a definable function on $\mathbb{R}$. Then* $\operatorname{dom} f$ *is a finite union of points and* (*open*) *intervals, and on each of the intervals f is either constant or strictly monotone and continuous.*

Uniform Finiteness Theorem. *Let F be a definable set-valued mapping from $\mathbb{R}^n$ into $\mathbb{R}^m$. Suppose that every $F(x)$ contains finitely many points. Then there is a natural N such that the number of points in every $F(x)$ does not exceed N.*

The next theorem uses the concept of a *cell* whose definition we omit. Although this concept will often be used in what follows, we do not need the specific structure of cells described in the formal definition. For us it will be sufficient to think of a C^k*-cell* of dimension r as of an r-dimensional C^k-manifold which is the image of the cube $(0,1)^r$ under a definable C^k-diffeomorphism. As follows from the definition, an m-dimensional cell in $\mathbb{R}^m$ is an open set.

Cell Decomposition Theorem.

(a) *Let $Q \subset \mathbb{R}^m$ be a definable set. Then for any k, Q can be represented as a disjoint union of a finite number of cells of class C^k.*
(b) *Let F be a definable mapping from a set $Q \subset \mathbb{R}^n$ into $\mathbb{R}^m$. Then there exist a partition of Q into a finite number of cells of class C^k such that the restriction of F on each cell is a mapping of class C^k.*

The maximal dimension of the cell in a decomposition of a definable set is called the *dimension* of the set. Of course, the dimension does not depend on the choice of decomposition. The important fact concerning dimension is that *the dimension of the image of a definable set under a definable* (*single-valued*) *mapping cannot be greater than the dimension of the preimage*.

Another consequence of the cell decomposition theorem is that any definable set has a finite number of connected components. More careful analysis leads to the conclusion that *any connected definable set is definably pathwise connected*, that is any two points of the set can be joined by a definable continuous curve lying completely in the set.

For a set-valued mapping $F:\mathbb{R}^n \rightrightarrows \mathbb{R}^m$ we denote $\operatorname{dom} F = \{x \in \mathbb{R}^n\colon\ \exists y\ [y \in F(x)]\}$, that is to say, the projection of the graph of F onto the domain space. A *selection* of F is, as usual, a mapping $\varphi(x)$ from $\operatorname{dom} F$ into the image space such that $\varphi(x) \in F(x)$ for all $x \in \operatorname{dom} F$.

Definable Choice Theorem. *Any definable* (*respectively tame*) *set-valued mapping has a definable* (*respectively tame*) *selection.*

We conclude the introduction with the following simple example which demonstrates the difference between definability and tameness and shows that one must be more careful when working with tame objects.

The function $\sin t$ is a tame function and t^{-1} is a semialgebraic function but the composition $\sin t^{-1}$ is not even a tame function (its restriction to e.g. $(0,1)$ is not definable in any o-minimal structure). Thus a composition of tame functions may be not a tame function. Observe that $\sin t$ is not a proper map.

4. Some preliminary results

Proposition 1. *Let F be a definable (tame) set-valued mapping from $\mathbb{R}^n$ into $\mathbb{R}^m$. Then* $\operatorname{sur} F$ *is a definable (respectively tame) function with* $\operatorname{dom}(\operatorname{sur} F) \subset \operatorname{cl}(\operatorname{Graph} F)$.

Proof. We can represent the graph of $(x, y, \lambda) \to \operatorname{Sur} F(x|y)(\lambda)$ (considered a function on $\operatorname{Graph} F \times (0, \infty)$) as $(P \cup Q) \cap S$, where

$$P = \operatorname{Graph} F \times (0, \infty) \times \{0\};$$
$$Q = \big\{(x, y, \lambda, r)\colon \forall 0 \leqslant \rho < r,\ \forall v\ \big[\|v - y\| \leqslant \rho \Rightarrow \exists u,\ \|u - x\| \leqslant \lambda,\ v \in F(u)\big]\big\};$$

and

$$S = \big\{(x, y, \lambda, r)\colon \forall \varepsilon > 0\ \exists v\ \big[\|v - y\| < r + \varepsilon \text{ and } \forall u\ \big[\|u - x\| \leqslant \lambda \Rightarrow v \notin F(u)\big]\big]\big\}.$$

If F is a definable mapping, then P is clearly definable and the other two sets are also definable by (D4) as was explained in the previous section. Hence so is the graph of $\operatorname{Sur} F(x|y)(\lambda)$. If F is a tame mapping, then for any $K > 0$ the intersections of three sets with $\{(x, y, \lambda, r)\colon \|x\| \leqslant K,\ \|y\| \leqslant K,\ 0 \leqslant \lambda, r \leqslant K\}$ is a definable set, so the intersection of the graph of $\operatorname{Sur} F$ with any such set is a definable set.

We note further the epigraph of $\operatorname{sur} F$ is the intersection of the closure of the epigraph of the function $(x, y, \lambda) \to \lambda^{-1} \operatorname{Sur}(x|y)(\lambda)$ with the set $\{(x, y, \lambda, \alpha)\colon \lambda = 0\}$, so $\operatorname{sur} F$ is a definable (or tame) function and its domain lies in the closure of $\operatorname{Graph} F$ by definition. □

Proposition 2. *Let F be a continuous definable (single-valued) mapping from an open definable subset of $\mathbb{R}^n$ into $\mathbb{R}^n$. Then the dimension of the set of critical values of F is not greater than $n - 1$.*

Proof. By the cell decomposition theorem, $\operatorname{dom} F$ is the union of C^1-cells and the restriction of F onto each of them is C^1. Then the set of critical values of the restriction of F to any cell of dimension n has measure zero by the Sard theorem, hence, by definability, its dimension cannot be greater than $n - 1$. On the other hand the image of the union of F-images of all other cells is also a set of dimension not greater than $n - 1$. □

Proposition 3. *Let U be an open definable subset of $\mathbb{R}^n$ and F a continuous definable (single-valued) mapping from U into $\mathbb{R}^m$. Assume that* $\operatorname{sur} F(x) = 0$ *for every $u \in U$. Then* $\dim F(U) \leqslant m - 1$.

Proof. Assume the contrary: $\dim F(U) = m$. Then there is an m-dimensional cell $Q \subset F(U)$ which by definition is an open subset of $\mathbb{R}^m$.

Applying the definable choice theorem, we shall find a definable mapping $G : Q \to \mathbb{R}^n$ such that $\{(x, y)\colon x = G(y),\ y \in Q\} \subset \operatorname{Graph} F$. This means that $F \circ G = Id|_Q$.

As G is definable, there is a smaller definable cell $\tilde{Q} \subset Q$ such that G is C^1 on $\tilde{Q}$. We have $(F \circ G)(\tilde{Q}) = \tilde{Q}$. Therefore the set $G(\tilde{Q})$ also has dimension m. This set is also definable as a definable image of a definable set. Therefore there is an m-dimensional cell $D \subset G(\tilde{Q})$ such that the restriction of F to D is C^1. But then $G \circ F|_D = Id|_D$ and as F is continuously differentiable at every $x \in D$, G is continuously differentiable at every $y = F(x)$, $x \in D$. Without loss of generality we can identify D with $(-1, 1)^m$. So we have $\nabla G(y) \circ \nabla(F|_D)(x) = I$, that is

$\nabla(F|_D)(x)$ is a nonsingular operator $\mathbb{R}^m \to \mathbb{R}^m$ which by the Lusternik–Graves theorem means that $\mathrm{sur}(F|_D)(x) > 0$. As $\mathrm{sur}\, F(x) \geqslant \mathrm{sur}(F|_D)$ (because $F|_D(x) \subset F(x)$ for every x), we get a contradiction. $\square$

Remark. Observe that a continuously differentiable mapping into $\mathbb{R}^m$, $m \geqslant 2$, with the rate of surjection identically zero can be surjective (see [26]).

Proposition 4 *(Differentiability theorem). If f is a definable function on $\mathbb{R}$, then f' is also definable and*

(a) $\mathrm{dom}\, f'$ *is an open set*;
(b) f' *is continuous on* $\mathrm{dom}\, f'$;
(c) $\mathrm{dom}\, f \setminus \mathrm{dom}\, f'$ *is a finite set.*

Proof. We have

$$\mathrm{Graph}\, f' = \big\{(t,\alpha) \in \mathrm{dom}\, f \times \mathbb{R}\colon\ \forall \varepsilon > 0\ \exists \delta > 0 \\ \big[|\tau - t| < \delta \Rightarrow \big|f(\tau) - f(t) - \alpha(\tau - t)\big| < \varepsilon|\tau - t|\big]\big\}.$$

Hence $\mathrm{Graph}\, f'$ is a definable set.

(a) Let $t \in \mathrm{dom}\, f'$. Then for any $\delta > 0$ the sets $(t, t+\delta) \cap \mathrm{dom}\, f$ and $(t-\delta, t) \cap \mathrm{dom}\, f$ are nonempty. The monotonicity theorem now implies (as every monotone function is almost everywhere differentiable) that also the sets $(t, t+\delta) \cap \mathrm{dom}\, f'$ and $(t-\delta, t) \cap \mathrm{dom}\, f'$ are nonempty for any positive δ, hence they are infinite. As these sets are definable, by (D6) there must be a $\delta > 0$ such that $(t, t+\delta) \subset \mathrm{dom}\, f'$ and $(t-\delta, t) \subset \mathrm{dom}\, f'$, that is $(t-\delta, t+\delta) \subset \mathrm{dom}\, f'$.

(b) As f' is a definable function, it obeys the monotonicity theorem. Thus we only need to observe that if there are $\tau < \tau' < \tau''$ such that $(\tau, \tau'') \subset \mathrm{dom}\, f'$ and f' is monotone on (τ, τ') and (τ', τ''), then

$$\lim_{t \to \tau' - 0} f'(t) = \lim_{t \to \tau' + 0} f'(t)$$

for otherwise f would not be differentiable at τ'.

(c) The set $\mathrm{dom}\, f \setminus \mathrm{dom}\, f'$ must have Lebesgue measure zero due to almost everywhere differentiability of monotone functions. Therefore it cannot contain intervals and, being definable, must be finite. $\square$

Corollary. *Let f be a definable function on an interval (a,b). Then there is a finite number of points $a = t_0 < t_1 < \cdots < t_k = b$ such that f is continuously differentiable and f' either strictly positive, or strictly negative or identically equal to zero on each interval (t_i, t_{i+1}).*

Proof. Let t_i be either the points of nondifferentiability of f or isolated points of $\{t\colon\ f'(t) = 0\}$ or the ends of intervals in the decomposition of the set according to (D6).

Proposition 5 *(Uniform integrability lemma). Let $Q \subset \mathbb{R}^n$ be a definable set, and let $\varphi(x,t,s)$ be a definable function on $Q \times (a,b) \times (0,1)$, where $-\infty < a < b < \infty$. Suppose that $\varphi(x,t,s) \to 0$ as $s \to 0$ uniformly on $Q \times (a,b)$. Then*

$$\int_a^b \big|\varphi_t(x,t,s)\big|\, dt \to 0, \quad \textit{as } s \to 0,\ \textit{uniformly on } Q.$$

Here φ_t is the derivative of φ with respect to t.

Proof. By the preceding corollary for any $x \in Q$, $s \in (0,1)$ there are finitely many, say $N(x,s)+1$ points $\tau_i = \tau_i(x,s)$ (with $\tau_0 = a$, $\tau_{N(x,s)} = b$) on (a,b) such that on each interval bounded by a pair of adjacent points, φ is continuously differentiable with respect to t and the derivative is either identical zero or does not change the sign. The set-valued mapping which associates with every (x,s) the collection of these points $\tau_i(x,s)$ is definable.

Indeed the set

$$\bigl\{(x,t,s) \in Q \times (a,b) \times (0,1)\colon\ \varphi(x,\cdot,s) \text{ is discontinuous at } t\bigr\}$$

is definable as it is equal to the intersection of

$$\bigl\{(x,t,s)\colon\ \exists r > 0,\ \varepsilon > 0\ \forall \delta \in (0,\varepsilon)\ \bigl[\varphi(x,t+\delta,s) - r > \varphi(x,t-\delta,s)\bigr]\bigr\}$$

and

$$\bigl\{(x,t,s)\colon\ \exists r > 0,\ \varepsilon > 0\ \forall \delta \in (0,\varepsilon)\ \bigl[\varphi(x,t+\delta,s) + r < \varphi(x,t-\delta,s)\bigr]\bigr\}.$$

The same is true for the points at which the derivative of φ with respect to t is discontinuous. Equally simple arguments lead to the conclusion that the set of (x,t,s) such that $\varphi_t(x,t,s) = 0$ (the derivative with respect to t) and t is either an isolated point of the zero set of $\varphi_t(x,\cdot,s)$ or an end of the interval at which the function is equal to zero, is also definable.

By the uniform finiteness theorem the numbers $N(x,s)$ are uniformly bounded, that is there is N such that $N \geqslant N(x,s)$ for all $x \in Q$, $s \in (0,1)$.

Fix $\varepsilon > 0$ and choose $\delta > 0$ such that $|\varphi(x,t,s)| \leqslant \varepsilon/2(N+1)$ if $0 < s < \delta$. Therefore for any $x \in Q$, $s \in (0,\delta)$

$$\int_{\tau_i}^{\tau_{i+1}} \bigl|\varphi_t(x,t,s)\bigr|\, dt = \bigl|\varphi(x,\tau_{i+1},s) - \varphi(x,\tau_i,s)\bigr| \leqslant \varepsilon/(N+1)$$

and therefore,

$$\int_a^b \bigl|\varphi_t(x,t,s)\bigr|\, dt \leqslant \varepsilon. \qquad \square$$

We shall use this result as a starting point for obtaining uniform estimates for the norms of derivatives of "small" definable mappings. Let F be a C^1-mapping from a neighborhood of $x \in \mathbb{R}^n$ into $\mathbb{R}^m$. We shall denote by $\nabla F(x)$ the Jacobian matrix of F at x, namely

$$\bigl(\nabla F(x)\bigr)^i_j = \frac{\partial F^i}{\partial x_j}(x),$$

where $x = (x_1, \ldots, x_n)$ and $F(x) = (F^1(x), \ldots, F^m(x))$. Recall also that the norm of a linear operator defined by matrix $A = (a^i_j)$ in the standard basis of $\mathbb{R}^n$ is

$$\|A\| = \left(\sum_{i=1}^m \sum_{j=1}^n \bigl(a^i_j\bigr)^2\right)^{1/2}.$$

Proposition 6. *Let $Q = (-1, 1)^n$ and let $F : Q \times (0, 1) \to \mathbb{R}^m$ be a definable mapping having the property that $\|F(x, s)\| \to 0$ uniformly on Q when $s \to 0$. Then for any $\varepsilon > 0$ there is $\delta > 0$ such that, whenever $s < \delta$, there is an open set $\Omega = \Omega(s) \subset Q$ such that $\|\nabla_x F(x, s)\| < \varepsilon$ for all $x \in \Omega$.*

Here the subscript x refers to differentiation with respect to x only, not to s.

Proof. Set for simplicity $\varphi_i(x, s) = (F(x, s))^i$. By the assumption every φ_i goes to zero uniformly in $x \in Q$ as $s \to 0$. Set $Q_j = \{(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)\colon |x_k| < 1\}$. Applying Proposition 5 for each φ_i and treating consecutively every x_j as t, we conclude that for every i, j

$$\int_{-1}^{1} \left| \frac{\partial \varphi_i(x, s)}{\partial x_j} \right| dx_j \to 0$$

uniformly on Q_j as $s \to 0$. As an consequence we get

$$\int_Q \left\| \nabla_x F(x, s) \right\| dx \to 0, \quad \text{as } s \to 0.$$

Indeed by the cell decomposition theorem we can partition Q into a finite number of cells such that F is continuously differentiable on each of them. Clearly, continuous differentiability on a cell of dimension n is the same as continuous differentiability without any restrictions. Equally clear is that the union of all cells of dimension n is a set of full measure in Q. This means that the integral above makes sense. On the other hand

$$\int_Q \left\| \nabla_x F(x, s) \right\| dx \leqslant \sum_{ij} \int_{Q_j} \left(\int_{-1}^{1} \left| \frac{\partial \varphi_i(x, s)}{\partial x_j} \right| dx_j \right) dx_1 \dots dx_{j-1}\, dx_{j+1} \dots dx_n.$$

We can choose $\delta > 0$ so small that the above integral is not greater than $\varepsilon/2$ if $s < \delta$. But then for any such s the Lebesgue measure of $\{x \in Q\colon \|\nabla_x F(x, s)\| < \varepsilon\}$ must be at least $1/2$. This is a definable set, hence it has nonempty interior as its measure is positive. $\square$

5. Proof of Theorem 1

Step 1. There is no loss of generality in assuming that the graph of F is closed (so that any critical value is "proper"). Indeed, as the graph of F is locally closed, the closure operation does not add points to a small neighborhood of any point of the graph, so that any critical value of F, no matter "proper" or "generalized" remains a critical value of the mapping whose graph is cl(Graph F).

Furthermore, as we mentioned in Section 2, y is a critical value of F if and only if it is a critical value of the restriction of the projection $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ to the closure of the graph of F. It follows that it is sufficient to prove the theorem for a single-valued mapping F which is a restriction of a linear operator $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ to a tame set dom $F \subset \mathbb{R}^n$.

We also observe that any cell in $\mathbb{R}^m$ of dimension smaller than m is a porous set (which is immediate from definitions), hence any definable set in $\mathbb{R}^m$ of dimension smaller than m is σ-porous. It follows that the theorem will be proved if we show that *for any definable mapping which is a restriction of a linear operator to a definable set the dimension of the set of its critical values cannot exceed $m - 1$.*

Indeed, the mapping F restricted to the set $\{x \in \operatorname{dom} F\colon\ \|x\| \leqslant N\}$ is definable by definition and any critical value of F is a critical value of the restriction if N is sufficiently big.

Step 2. By Proposition 2 the theorem is true if the dimension of dom F coincides with the dimension of the image space. Assume now that for a given $m \geqslant 1$ the theorem holds for any definable mapping whose domain has dimension not greater than $r \geqslant m$ and let F be a mapping from $\mathbb{R}^n$ into $\mathbb{R}^m$ with $\dim(\operatorname{dom} F) = r + 1$ (of course $n \geqslant r + 1$).

By the cell decomposition theorem dom F can be partitioned into finitely many C^k-cells C_i ($k \geqslant r + 2 - m$). Denote by $\mathcal{N} = \{\bigcup C_i\colon\ \dim C_i = r + 1\}$ the union of all $(r + 1)$-dimensional cells of the partition, $\mathcal{K} = \{\bigcup(\operatorname{cl}(C_i) \cap (\operatorname{dom} F)),\ \dim C_i \leqslant r\}$ the union of intersections of the closures of cells of dimensions $\leqslant r$ with the domain of F.

Then

(a) the collection of critical values of $F|_{\mathcal{K}} = A|_{\mathcal{K}}$ is a set of dimension $\leqslant m - 1$ by the induction assumption;
(b) for any cell C_i the collection of critical values of $F|_{C_i} = A|_{C_i}$ is a set of dimension $\leqslant m - 1$ by the classical Sard theorem. Indeed, if $\dim C_i = s$, then C_i is the image of the open s-cube $(-1, 1)^s$ under a C^k-mapping G. Therefore singular points of $A|_{C_i}$ are among singular points of $A \circ G$. The latter is a C^k-mapping from the cube into $\mathbb{R}^m$ and as $k \geqslant s + 1 - m$ as $s \leqslant r + 1$, Sard's theorem applies.

Let $x \in \operatorname{dom} F$ be a singular point of F which does not belong to any of the two above mentioned types. This means that

(c) x is a regular point of the restriction of F to the cell of the partition containing x. We claim that for some i there is a sequence $(x^\nu) \subset C_i$ converging to x such that

$$\lim_{\nu \to +0} \operatorname{sur} F|_{C_i}\bigl(x^\nu\bigr) = 0. \tag{1}$$

Indeed, as $\operatorname{sur} F(x) = 0$, there is a sequence (x^ν) converging to x and such that $\operatorname{Sl} F(x^\nu) \to 0$ and there is no loss of generality in assuming that all x^ν belong to the same cell, call it C. Furthermore, it follows from the definition of the slope and the inequality at the end of Section 2 that

$$\operatorname{Sl} F\bigl(x^\nu\bigr) \geqslant \operatorname{Sl} F|_C\bigl(x^\nu\bigr) \geqslant \operatorname{sur} F|_C\bigl(x^\nu\bigr)$$

which implies (1).

We note further that x^ν cannot belong to $\mathcal{K}$ for in this case x also belongs to $\mathcal{K}$ as the latter is closed. Thus x is a singular point of $F|_{\mathcal{K}}$ which is the case of (a). It follows that $x^\nu \in \mathcal{N}$ for all ν. The limiting point x cannot belong to the same cell as x^ν since we assume that x is a regular point of the restriction of F to the cell.

Thus, we arrive to the following situation: there is a cell C of dimension $r + 1$ or higher and a sequence (x^ν) converging to x such that $x^\nu \in C$ for all ν, $x \in (\operatorname{dom} F) \setminus C$ and (1) holds.

Let $\mathcal{M}$ denote the collection of such points x (associated with the same cell C). We have to show that

$$\dim F(\mathcal{M}) \leqslant m - 1. \tag{2}$$

Step 3. Thus we reduce the problem to the following. Given

- a definable mapping F from $\mathbb{R}^n$ into $\mathbb{R}^m$ which is a restriction of a linear operator A to a definable set $\operatorname{dom} F \subset \mathbb{R}^n$;
- a cell $C \subset \operatorname{dom} F$ of class C^1 (and dimension $\geqslant r+1$);
- a nonempty set

$$\mathcal{M} = \big\{x \in (\operatorname{dom} F) \setminus C\colon\ \forall s > 0\ \exists y \in C\ \big[\|x-y\| < s,\ \operatorname{sur} F(y) < s\big]\big\}.$$

It is clear from the definition of $\mathcal{M}$ that it is a definable set lying completely in the closure of C, that is in the boundary of C, as $\mathcal{M}$ and C do not meet.

We have to prove that (2) holds. Assume by way of contradiction that $\dim A(\mathcal{M}) = m$. Then $\mathcal{M}$ contains a q-dimensional ($q \geqslant m$) cell of class C^1 whose A-image has dimension m. This means that there is a diffeomorphism G of $Q = (-1,1)^q$ into $\mathcal{M}$ such that $\dim(A \circ G)(Q) = m$. If $\operatorname{sur}(A \circ G)(u) = 0$ for all $u \in Q$ then by Proposition 3 $\dim(A \circ G)(Q) \leqslant m-1$, so the rate of surjection of $A \circ G$ must be positive at certain points of Q. Since $\operatorname{sur}(A \circ G)$ is a lower semicontinuous function, we can assume, taking a smaller cube if necessary, that

$$\operatorname{sur}(A \circ G)(u) \geqslant \alpha > 0, \quad \forall u \in Q, \tag{3}$$

and also that G satisfies the Lipschitz condition in Q.

Step 4. Consider the set

$$\varGamma = \big\{(x,y,s)\colon\ x \in \mathcal{M},\ y \in C,\ s \in (0,1),\ \|y-x\| < s,\ \operatorname{sur} F(y) < s\big\}.$$

This is a definable set and its projection onto the x-component space is $\mathcal{M}$. By the definable choice theorem there is a definable mapping $y(x,s)$ from $\mathcal{M} \times (0,1)$ into $\mathbb{R}^n$ such that $(x, y(x,s), s) \in \varGamma$ for all (x,s). We have: $\|x - y(x,s)\| < s$ for all $x \in \mathcal{M}$ and all $s \in (0,1)$. Set

$$\varPsi(x,s) = x - y(x,s); \qquad \varPhi(u,s) = \varPsi\big(G(u), s\big).$$

Then $\varPhi(u,s) \to 0$ uniformly on Q as $s \to 0$.

Applying Proposition 6 to $\varPhi$, we conclude that for each given $\varepsilon > 0$ there is an $s = s(\varepsilon) \leqslant \varepsilon$ and an open set $\varOmega(\varepsilon) \subset Q$ such that $\|\nabla_u \varPhi(u,s)\| < \varepsilon$ for all $u \in \varOmega(\varepsilon)$. This means that

$$\big\|\nabla G(u) - \nabla(y_\varepsilon \circ G)(u)\big\| < \varepsilon, \tag{4}$$

where $y_\varepsilon(x) = y(x, s(\varepsilon))$.

Now for any $\varepsilon > 0$ choose $u = u(\varepsilon) \in \varOmega(\varepsilon)$ such that y_ε be differentiable at u. Then, the inequality above implies that

$$\nabla(A \circ G)(u) = \nabla(A \circ y_\varepsilon \circ G)(u) + T(\varepsilon), \tag{5}$$

where $T(\varepsilon) = (A \circ (\nabla G - \nabla(y_\varepsilon \circ G)))(u)$ is a linear operator from $\mathbb{R}^q$ into $\mathbb{R}^m$ with $\|T(\varepsilon)\| \to 0$ as $\varepsilon \to 0$. We have by (3)

$$\operatorname{sur}\big(\nabla(A \circ G)(u)\big) = \operatorname{sur}\big(\nabla(F \circ G)(u)\big) = \operatorname{sur}(F \circ G)(u) \geqslant \alpha. \tag{6}$$

On the other hand, by (4)

$$\begin{aligned}\operatorname{sur}\big(\nabla(A \circ y_\varepsilon \circ G)\big)(u) &= \operatorname{sur}\big(\nabla\big(F \circ (y_\varepsilon \circ G)\big)\big)(u)\\ &= \operatorname{sur}\big(F \circ (y_\varepsilon \circ G)\big)(u) \leqslant (K+\varepsilon)\operatorname{sur} F\big(y_\varepsilon\big(G(u)\big)\big),\end{aligned}$$

where $K = \|\nabla G(u)\|$.

The last three relations are however contradictory, as $\operatorname{sur} F(y_\varepsilon(G(u))) \to 0$ and from (5), (6) we get

$$0 < \alpha \leqslant \operatorname{sur}\big(\nabla(A \circ G)(u)\big) \leqslant \operatorname{sur}\big(\nabla(A \circ y_\varepsilon \circ G)(u)\big) + \big\|T(\varepsilon)\big\| \to 0.$$

This completes the proof of the theorem.

6. Some corollaries

In [14] Kurdyka, Orro and Simon proved that the set of asymptotically critical values of a continuously differentiable semialgebraic mapping $\mathbb{R}^n \to \mathbb{R}^m$ has dimension less than m. Their proof was based on calculation of an estimate for the m-dimensional measure of the set. Theorem 1 allows to avoid these calculations and to get some information of asymptotically critical values of definable set-valued mappings in general. First we note the following

Proposition 7. *Let $H : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, and let $\rho(x)$ be a continuously differentiable strictly positive function. Set $L(x) = H(\rho(x)x)$. Then for any $(x, y) \in \operatorname{Graph} L$*

$$\operatorname{sur} L(x|y) \leqslant \big(\rho(x) + \big\|\rho'(x)\big\| \cdot \|x\|\big) \operatorname{sur} H\big(\rho(x)x|y\big).$$

Proof. Indeed, the norm of the derivative of the mapping $x \to \rho(x)x$ at x is not greater than $\rho(x) + \|\rho'(x)\|\|x\|$. $\square$

Let $\eta(t)$ be a strictly positive continuous function on $[0, \infty)$ such that

$$\int_0^\infty \frac{1}{\eta(t)} < \infty. \tag{7}$$

Given a set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, we call $y \in \mathbb{R}^m$ an *asymptotically η-critical value* of F if there is a sequence of pairs (x^ν, y^ν) such that $y^\nu \in F(x^\nu)$, $\|x^\nu\| \to \infty$, $y^\nu \to y$ and $\eta(\|x^\nu\|) \cdot \operatorname{sur} F(x^\nu|y^\nu) \to 0$. We shall denote by $K_\infty(F, \eta)$ the set of asymptotically η-critical values of F.

Theorem 2. *Let F be a definable set-valued mapping with locally closed graph, and let $\varphi(t)$ be strictly increasing continuously differentiable positive definable function on $[0, \infty)$, bounded from above and equal to zero at 0. Set*

$$\eta(t) = \frac{1}{\varphi'(t)}.$$

Then $K_\infty(F, \eta)$ is a definable set with $\dim(K_\infty(F, \eta)) < m$.

Proof. Clearly, η satisfies (7). We obviously have

$$\begin{aligned} K_\infty(F) = \big\{ & y \in \mathbb{R}^m \colon\ \forall N \in (0, \infty),\ \exists (u, v) \in \mathbb{R}^n \times \mathbb{R}^m \\ & \big[v \in F(u),\ \|v - y\| < N^{-1},\ \|u\| \geqslant N,\ \eta\big(\|u\|\big) \cdot \operatorname{sur} F(u|v) < N^{-1}\big]\big\} \end{aligned}$$

which shows that the set is definable by Proposition 1.

Without loss of generality we may assume that $\varphi(t) \to 1$ as $t \to \infty$. Let $\psi(\cdot)$ be the inverse of $\varphi(\cdot)$, that is $\psi(\varphi(t)) \equiv t$. Set

$$G(u) = F\left(\psi\big(\|u\|\big)\frac{u}{\|u\|}\right).$$

In other words, we consider the following pair of mutually inverse changes of variables:

$$x = \psi\big(\|u\|\big)\frac{u}{\|u\|}, \qquad u = \varphi\big(\|x\|\big)\frac{x}{\|x\|}$$

which transfer the open unit ball into the entire space and vice versa. We obviously have for each pair of corresponding x and u:

$$\|x\| = \psi\big(\|u\|\big), \qquad \|u\| = \varphi\big(\|x\|\big).$$

(Here and below we consider the Euclidean norm in $\mathbb{R}^n$.)

We can apply Proposition 7 to get estimates of the rate of surjection of G: for each $u \neq 0$ with $\|u\| < 1$ and each $y \in G(u)$

$$\operatorname{sur} G(u|y) \leqslant \left(\frac{\psi(\|u\|)}{\|u\|} + \left\|\left[\frac{\psi(\|u\|)}{\|u\|}\right]'\right\| \cdot \|u\|\right) \operatorname{sur} F(x|y).$$

After a simple calculation we get:

$$\left\|\left[\frac{\psi(\|u\|)}{\|u\|}\right]'\right\| \cdot \|u\| \leqslant \psi'\big(\|u\|\big) + \frac{\psi(\|u\|)}{\|u\|},$$

so that

$$\operatorname{sur} G(u|y) \leqslant 2\left(\psi'\big(\|u\|\big) + \frac{\psi(\|u\|)}{\|u\|}\right) \operatorname{sur} F(x|y) = 2\left(\frac{1}{\varphi'(\|x\|)} + \frac{\|x\|}{\varphi(\|x\|)}\right) \operatorname{sur} F(x|y).$$

We now recall that $\eta(\cdot)$ is reciprocal of $\varphi'(\cdot)$ and that by (7) $\eta(t)$ grows to infinity faster than t. Thus, we can be sure that for sufficiently large x

$$\operatorname{sur} G(u|y) \leqslant 3\eta\big(\|x\|\big) \cdot \operatorname{sur} F(x|y). \tag{8}$$

Now let $\tilde{G}$ be a set-valued mapping whose graph is the closure of $\operatorname{Graph} G$. Clearly, if $\|x^\nu\| \to \infty$ then the sequence of the corresponding u^ν contains a subsequence converging to an element of the unit sphere. Therefore as follows from (8) every asymptotical η-critical value of F is a critical value of $\tilde{G}$. The latter is a definable set-valued mapping, so by Theorem 1 the entire set of its critical values, including $K_\infty(F, \eta)$ has the dimension smaller than m. $\square$

Let us call, following [14], a point $y \in \mathbb{R}^m$ an *asymptotically critical value* of F if there is a sequence of pairs (x^ν, y^ν) such that $y^\nu \in F(x^\nu)$, $\|x^\nu\| \to \infty$, $y^\nu \to y$ and $\|x^\nu\| \operatorname{sur} F(x^\nu|y^\nu) \to 0$. We shall denote by $K_\infty(F)$ the set of asymptotically critical values of F.

Of course, any asymptotically critical value is asymptotically η-critical for any $\eta(t)$ satisfying the requirements in the definition but we cannot, in principle, expect an asymptotically η-critical value to belong to $K_\infty(F)$. However, Kurdyka, Otto and Simon showed in [14, Lemma 3.1] that in case when F is a semialgebraic C^1-mapping on $\mathbb{R}^n$ (or an open semialgebraic subset of $\mathbb{R}^n$), there is an $\gamma > 0$ (depending on F) such that $K_\infty(F, \eta) \subset K_\infty(F)$ if $\eta(t) = t^{1+\alpha}$ with $0 < \alpha < \gamma$.

The proof of this fact given in [14] extends without change to arbitrary semialgebraic set-valued mappings. Actually, the only use of differentiability in the proof is in the definition of the distance from the derivative of F to the set of singular operators playing the same role in the definition of asymptotically critical value given in [14] as the rate of surjection in the definition above. But this distance is precisely the rate of surjection of F at the corresponding point—the fact well known in variational analysis and actually also proved in [14, Propositions 2.1, 2.2].

Thus, combining the quoted result of [14] with Theorem 2, we get an extension of (the first part of) the main theorem of [14] to arbitrary semialgebraic set-valued mappings.

Theorem 3. *Let F be a semialgebraic set-valued mapping from $\mathbb{R}^n$ into $\mathbb{R}^m$ whose graph is locally closed. Then $K_\infty(F)$ is a closed semialgebraic set of dimension smaller than m.*

Our final result is an extension to definable functions of a recent theorem of Bolte, Daniilidis and Lewis [4] stating that a continuous globally subanalytic function is constant on every connected component of its critical points.[3]

In [4] this fact is used to prove that such function has only finitely many critical points of that sort. Both are direct consequences of Theorem 1. But we give for Theorem 4 below an independent proof because, unlike the proof of Theorem 1, it does not use the argument ad absurdum. This implication can be partly reversed since Theorem 4 can be used to get a direct proof of Theorem 1 for single-valued locally Lipschitz tame mappings using the fact that for such mappings sur $F(x) = 0$ is equivalent to x being a critical point of $(y^* \circ F)(x) = 0$ for some y^* with $\|y^*\| = 1$.

Theorem 4. *Let f be a definable function which is continuous on its domain* dom f. *Then f is constant on every connected component of the set of its critical points.*

Proof. Let u and w be two different points belonging to the same connected component of the set of critical points of f. This set is definable since so is the function sur f by Proposition 1. Hence there is a definable curve $x(t)$, $0 \leqslant t \leqslant 1$, joining u and w and lying completely in the set of critical points of f. By definition for any $\varepsilon > 0$ and any $t \in [0,1]$ there is a z and a $\lambda \in (0, \varepsilon)$ such that $\|z - x(t)\| < \varepsilon$, and $\mathrm{Sur}\, f(z)(\lambda) < \varepsilon\lambda$. By the definable choice theorem there are functions $z_\varepsilon(t)$ and $\lambda_\varepsilon(t)$ (definably depending on both variables) such that $\|x(t) - z_\varepsilon(t)\| < \varepsilon$, $0 < \lambda_\varepsilon(t) < \varepsilon$ and $\mathrm{Sur}\, f(z_\varepsilon(t))(\lambda_\varepsilon(t)) < \varepsilon\lambda_\varepsilon(t)$ for all t and all ε. If we set

$$\mu_\varepsilon^+(t) = \sup\bigl\{f(z) - f\bigl(z_\varepsilon(t)\bigr)\colon \bigl\|z - z_\varepsilon(t)\bigr\| < \lambda_\varepsilon(t),\ z \in \operatorname{dom} f\bigr\},$$
$$\mu_\varepsilon^-(t) = \sup\bigl\{f\bigl(z_\varepsilon(t)\bigr) - f(z)\colon \bigl\|z - z_\varepsilon(t)\bigr\| < \lambda_\varepsilon(t),\ z \in \operatorname{dom} f\bigr\},$$

the latter amounts to

$$\mu_\varepsilon(t) = \min\bigl\{\mu_\varepsilon^+(t), \mu_\varepsilon^-(t)\bigr\} < \varepsilon\lambda_\varepsilon(t). \tag{9}$$

Applying Proposition 5 to each component of $x(t) - z_\varepsilon(t)$, we conclude that

$$\int_0^1 \bigl\|\dot{x}(t) - \dot{z}_\varepsilon(t)\bigr\|\, dt \to 0, \quad \text{as } \varepsilon \to 0. \tag{10}$$

[3] It has to be mentioned that in [4] another definition of a critical point is used, namely that a point is lower critical if zero belongs to the limiting Fréchet subdifferential of the function at the point. It can be shown, and actually follows from one of the central facts of the finite dimensional calculus of subdifferentials [13,17,21], that x is a critical point of f (that is considered as a mapping into $\mathbb{R}$) if and only if it is a lower critical point of either f or $-f$.

The functions $z_\varepsilon(\cdot)$ and $\lambda_\varepsilon(\cdot)$ may have points of discontinuity but by the uniform finiteness theorem the number of such points is bounded by the same constant for all ε. Note also that as $x(t)$ is continuous, the size of each jump of z_ε does not exceed 2ε.

We observe further that $\mu_\varepsilon^\pm$ and μ_ε are definable functions. Therefore for any ε there are finitely many points on [0, 1] such that between any pair of adjacent points either $\mu_\varepsilon(t) = \mu_\varepsilon^+(t)$ or $\mu_\varepsilon(t) = \mu_\varepsilon^-(t)$.

Thus there is a natural N such that for any $\varepsilon > 0$ there are points τ_i, $i = 1, \dots, k$ $(0 = \tau_0 \leqslant \tau_1 < \cdots < \tau_k \leqslant \tau_{k+1} = 1,\ k \leqslant N)$, such that on every interval (τ_i, τ_{i+1}), $z_\varepsilon(t)$ and $\lambda_\varepsilon(t)$ are continuous and either $\mu_\varepsilon(t) = \mu_\varepsilon^+(t)$ or $\mu_\varepsilon(t) = \mu_\varepsilon^-(t)$ for all t in the interval.

As f is continuous, the theorem will be proved if we show that

$$\bigl|f\bigl(z_\varepsilon(0)\bigr) - f\bigl(z_\varepsilon(1)\bigr)\bigr| \to 0. \tag{11}$$

So fix $\varepsilon > 0$ and let τ_i, $i = 1, \dots, k$ $(0 = \tau_0 \leqslant \tau_1 < \cdots < \tau_k \leqslant \tau_{k+1} = 1,\ k \leqslant N)$, be the points specified above. For any i we set

$$z_\varepsilon(\tau_i)^+ = \lim_{t \to \tau_i + 0} z_\varepsilon(t), \qquad z_\varepsilon(\tau_i)^- = \lim_{t \to \tau_i - 0} z_\varepsilon(t).$$

Take a certain interval (τ_i, τ_{i+1}) and assume for instance that $\mu_\varepsilon(t) = \mu_\varepsilon^+(t)$ on this interval. This means for any $t \in (\tau_i, \tau_{i+1})$ there is a $t' \in (t, \tau_{i+1})$ such that $\|z_\varepsilon(t') - z_\varepsilon(t)\| < \lambda_\varepsilon(t)$. Fix a t and let t^+ be the upper bound of such t'. Then in the inequality above we either get an equality at t^+ (as z_ε and λ_ε are continuous on the interval) or t^+ coincides with the right end of the interval. In the last case

$$\bigl|f\bigl(z_\varepsilon(\tau_{i+1})^-\bigr) - f\bigl(z_\varepsilon(t)\bigr)\bigr| \leqslant \varepsilon^2.$$

If we get an equality at some $t^+ \leqslant \tau_{i+1}$, then by (9) (recall that $\mu_\varepsilon(t) = \mu_\varepsilon^+(t)$), we have

$$\bigl|f\bigl(z_\varepsilon(t^+)\bigr) - f\bigl(z_\varepsilon(t)\bigr)\bigr| \leqslant \varepsilon\lambda_\varepsilon(t^+) = \varepsilon\bigl\|z_\varepsilon(t^+) - z_\varepsilon(t)\bigr\| \leqslant \varepsilon \int_t^{t^+} \bigl\|\dot z_\varepsilon(s)\bigr\|\, ds.$$

We can ask about the upper bound τ^+ of t^+ for which the last inequality holds. The standard argument shows that either this upper bound is τ_{i+1} or $\|z_\varepsilon(\tau_{i+1}) - z_\varepsilon(\tau^+)\| < \lambda_\varepsilon(\tau^+) < \varepsilon$. Indeed, if the opposite inequality holds, then there is $\tau \in (\tau^+, \tau_{i+1}]$ such that $\|z_\varepsilon(\tau) - z_\varepsilon(\tau^+)\| = \lambda_\varepsilon(\tau^+)$ and therefore

$$\bigl|f\bigl(z_\varepsilon(\tau)\bigr) - f\bigl(z_\varepsilon(\tau^+)\bigr)\bigr| \leqslant \varepsilon \int_{\tau^+}^{\tau} \bigl\|\dot z_\varepsilon(s)\bigr\|\, ds$$

and we arrive to a contradiction with the definition of τ^+. Thus we can conclude by stating that for any t in the interval

$$\bigl|f\bigl(z_\varepsilon(\tau_{i+1})^-\bigr) - f\bigl(z_\varepsilon(t)\bigr)\bigr| \leqslant \varepsilon\left(\int_t^{\tau_{i+1}} \bigl\|\dot z_\varepsilon(s)\bigr\|\, ds + \varepsilon\right)$$

and consequently, by continuity

$$\bigl|f\bigl(z_\varepsilon(\tau_{i+1})^-\bigr) - f\bigl(z_\varepsilon(\tau_i)^+\bigr)\bigr| \leqslant \varepsilon\left(\int_{\tau_i}^{\tau_{i+1}} \bigl\|\dot z_\varepsilon(t)\bigr\|\, dt + \varepsilon\right). \tag{12}$$

The same argument, with an obvious change, applies to intervals on which $\mu_\varepsilon(t) = \mu_\varepsilon^-(t)$, and we can be sure that (12) holds for each interval of the partition.

Let finally $\omega(r)$ be the modulus of continuity of f in a neighborhood of $x(\cdot)$. Then taking into account (12) along with the fact that we have at most N points of discontinuity of z_ε and the jumps cannot exceed 2ε, we find that

$$\left|f\big(z_\varepsilon(1)\big) - f\big(z_\varepsilon(0)\big)\right| \leqslant \varepsilon(N+1)\left(\int_0^1 \left\|\dot{x}(t)\right\| dt + \varepsilon + \int_0^1 \left\|\dot{x}(t) - \dot{z}_\varepsilon(t)\right\| dt\right) + N\omega(2\varepsilon)$$

from which (11) immediately follows in view of (10). □

Acknowledgments

This paper was written while I was on sabbatical in the Department of Computer Science of Dalhousie University. I wish to express my gratitude to the department and especially to Jon Borwein for their hospitality and excellent working conditions I was provided with. I am also thankful to Adrian Lewis for inspiring discussions and to the reviewer for careful reading.

References

[1] D. d'Acunto, Valeurs critiques asymptotiques de fonctions définissables dans un structure o-minimale, Ann. Polon. Math. LXXV (2000) 35–45.

[2] D. Azé, J.-N. Corvellec, Characterization of error bounds for lower semicontinuous functions on metric spaces, ESAIM Control Optim. Calc. Var. 10 (2004) 409–425.

[3] S.M. Bates, Toward a precise smoothness hypothesis in Sard's theorem, Proc. Amer. Math. Soc. 117 (1993) 279–283.

[4] J. Bolte, A. Daniilidis, A. Lewis, A nonsmooth Morse–Sard theorem for subanalytic functions, J. Math. Anal. Appl. 321 (2) (2006) 729–740.

[5] J. Bolte, A. Daniilidis, A. Lewis, The Lojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems, SIAM J. Optim., in press.

[6] J. Bolte, A. Daniilidis, A. Lewis, M. Shiota, A Sard-type theorem for Clarke critical values of subanalytic Lipschitz continuous functions, Ann. Polon. Math., in press.

[7] J. Bochnak, M. Coste, M.-F. Roy, Real Algebraic Geometry, Springer, 1998.

[8] J.M. Borwein, Q.J. Zhu, Techniques of Variational Analysis, Springer, 2005.

[9] M. Coste, An introduction to o-minimal geometry, Inst. Rech. Math., Univ. de Rennes, http://name.math.univ-rennes1.fr/michel.coste/polyens/OMIN.pdf, 1999.

[10] E. De Giorgi, A. Marino, M. Tosques, Problemi di evoluzione in spazi metrici e curve di massima pendenza, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. 68 (1980) 180–187.

[11] L. van den Dries, Tame Topology and o-Minimal Structures, Cambridge Univ. Press, 1998.

[12] A. Grothendieck, Sketch of a proposal, in: L. Schneps, P. Lochak (Eds.), Geometric Galois Actions, Cambridge Univ. Press, 1997.

[13] A.D. Ioffe, Metric regularity and subdifferential calculus, Uspekhi Mat. Nauk 55 (3) (2000) 103–162, English translation: Russian Math. Surveys 55 (2000) 501–558.

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

[15] A.S. Lewis, Active sets, nonsmoothness and sensitivity, SIAM J. Optim. 13 (2003) 702–725.

[16] R. Mifflin, Semismooth and semiconvex functions in constrained optimization, Math. Oper. Res. 2 (1977) 191–207.

[17] B.S. Mordukhovich, Variational Analysis and Generalized Differentiation, vol. 1, Springer, 2005.

[18] A. Norton, Functions not constant on fractal quasi-arcs of critical points, Proc. Amer. Math. Soc. 106 (1989) 397–405.

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

[20] R.T. Rockafellar, Favorable classes of Lipschitz continuous functions in subgradient optimization, in: E. Nurminski (Ed.), Progress in Non-Differentiable Optimization, Pergamon Press, 1981.
[21] R.T. Rockafellar, R.J.B. Wets, Variational Analysis, Springer, 1998.
[22] A. Rohde, On Sard's theorem for nonsmooth functions, Numer. Funct. Anal. Optim. 9–10 (1997) 1023–1039.
[23] B. Teissier, Tame and stratified objects, in: L. Schneps, P. Lochak (Eds.), Geometric Galois Actions, Cambridge Univ. Press, 1997, pp. 231–242.
[24] H. Whitney, A function not constant on a connected set of critical points, Duke Math. J. 1 (1935) 514–517.
[25] Y. Yomdin, Maxima of smooth families III: Morse–Sard theorem, preprint, MOI, Bonn, 1984.
[26] Y. Yomdin, G. Comte, Tame Geometry with Applications to Smooth Analysis, Lecture Notes in Math., vol. 1834, Springer, 2004.
[27] Y. Yomdin, Surjective mapping whose differential is nowhere surjective, Proc. Amer. Math. Soc. 111 (1991) 267–270.