CRITICAL VALUES OF SET-VALUED MAPS
WITH STRATIFIABLE GRAPHS. EXTENSIONS OF SARD
AND SMALE-SARD THEOREMS

A. D. IOFFE

(Communicated by Jonathan M. Borwein)

Abstract. We prove three theorems extending Sard’s theorem and its infinite
dimensional extension due to Smale to set-valued mappings with stratifiable
graphs. The very concept of a critical value comes from (nonsmooth) vari-
tional analysis and turns out to be perfectly compatible with the natural
condition defining “good” stratifications (e.g., Whitney stratification in the
finite dimensional case).

1. Introduction

Variational analysis needs Sard-type theorems both for theoretical (logic of de-
developments of the (metric) regularity theory) and practical (basically in connection
with well-posedness and stability problems, especially acute in the absence of
smoothness) reasons. It was strongly emphasized in recent studies (e.g. [12]) that
nonsmoothness typically appearing in applications is rather well structured. On
the other hand a number of recent results related to Sard-type phenomena show
that differentiability requirements can be substantially weakened in exchange for
some structural assumptions (e.g. [2, 9, 11]).

The two latest papers of the three deal with nondifferentiable situations. The
main result of [2] states that the collection of so-called Clarke critical values of an
l.s.c. extended-real-valued function on \( \mathbb{R}^n \) whose graph admits a Whitney stratifi-
cation (satisfying some additional and not very restrictive technical assumptions)
is a set of Lebesgue measure zero in \( \mathbb{R} \).

An extension of the Sard theorem to set-valued mappings with graphs definable
in o-minimal structures was proved in [9]. The result, although its proof uses the
classical Sard theorem, does not imply the latter (as the Sard theorem applies to
arbitrary \( C^k \) functions for sufficiently large \( k \), not necessarily definable). The frame-
works of o-minimal structures also do not allow for infinite dimensional extensions
of the result in the spirit of the Smale-Sard theorem [15].

In this paper we show that stratified structures are a natural domain for non-
smooth extensions of the Sard and Smale-Sard theorems in a way perfectly com-
patible with the basic concepts of variational analysis.
Notation. In what follows we denote spaces by script letters and maps, whether single or set-valued, by capital Latin letters. The notation \( u \to x \) (respectively, \( u \to x \)) means “\( u \to x \) and \( x \in Q \)” (respectively, “\( u \to x \) and \( f(u) \to f(x) \)”). The harpoon arrow \( \vec{\to} \) stands for the weak* convergence. For set-valued mappings we use the symbol \( F : \mathcal{X} \rightrightarrows \mathcal{Y} \); \( \text{cl}^* \) means the weak*-closure.

2. Preliminaries

2.1. Regularity. Let \( \mathcal{X} \) and \( \mathcal{Y} \) be metric spaces, and let \( F : \mathcal{X} \rightrightarrows \mathcal{Y} \). For \( (x, y) \in \text{Graph } F \) we set

\[
\text{sur}F(x|y)(\lambda) = \sup \{ r \geq 0 : y + rB_{\mathcal{Y}} \subseteq F(x + \lambda B_{\mathcal{X}}) \}
\]

and then for \( (\bar{x}, \bar{y}) \in \text{Graph } F \) define the rate of surjection of \( F \) at \( (\bar{x}, \bar{y}) \) by

\[
\text{sur}F(\bar{x}|\bar{y}) = \liminf_{(x,y,\lambda) \to (\bar{x},\bar{y},0)} \frac{1}{\lambda} \text{sur}F(x|y)(\lambda).
\]

Here, as usual, \( B_{\mathcal{X}}, \text{etc.} \), is the unit ball in \( \mathcal{X} \) and we set \( \sup \emptyset = 0 \) (or else, we can calculate the \( \liminf \) in the definition of \( \text{sur}F \) only along sequences of \( (x, y) \in \text{Graph } F \)). It follows from the definition that the function \( \text{sur}F \) is defined on the closure of \( \text{Graph } F \) and is upper semi-continuous.

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

Definition 1. It is said that \( F \) is regular at \( (x, y) \in \text{Graph } F \) if \( \text{sur}F(x|y) > 0 \). Also, \( y \) is a (proper) critical value of \( F \) if there is an \( x \) such that \( y \in F(x) \) and \( \text{sur}F(x|y) = 0 \); otherwise \( y \) is called a regular value of \( F \).

Remark. As \( \text{sur}F(x|y) \) is defined on the closure of \( \text{Graph } F \), it is possible to define a general critical value by the condition that \( (x, y) \in \text{cl}(\text{Graph } F) \). We shall relate to this concept in Corollary 9.

Let \( P = P_{\mathcal{Y}} : \mathcal{X} \times \mathcal{Y} \to \mathcal{Y} \) be the projection onto \( \mathcal{Y} \).

Proposition 1. Let \( G(x, y) \) be the restriction of \( P \) onto the graph of \( F \), that is, the (strictly speaking, set-valued) mapping which associates \( y \) with every \( (x, y) \in \text{Graph } F \) and \( \emptyset \) with \( (x, y) \notin \text{Graph } F \). Then the collections of regular points and critical values of \( G \) and \( F \) coincide.

Proof. This is an immediate consequence of [\S], Proposition 1.3. \( \square \)

If \( \mathcal{X} \) and \( \mathcal{Y} \) are Banach spaces and \( A : \mathcal{X} \to \mathcal{Y} \) is a linear bounded operator, then the rate of surjection of \( A \) does not depend on \( x \) and is given by

\[
\text{sur}A(x) = \text{Sur}A(x)(1) = \inf_{\|y^*\| = 1} \|A^*y^*\| = \|A^{-1}\|^{-1},
\]

where for any linear operator \( T \) we set

\[
\|T^{-1}\| = \sup_{\|y\| \leq 1} \inf\{\|x\| : Tx = y\} = \sup_{\|y\| \leq 1} d(0, T^{-1}(y)).
\]

Thus we can speak just about \( \text{sur}A \). If \( F \) is a mapping from a neighborhood of \( x \in \mathcal{X} \) into \( \mathcal{Y} \) which is continuously (or even strictly) differentiable at \( x \), then \( \text{sur}F(x) = \text{sur}(F'(x)) \) (a consequence of Milyutin’s theorem; see [\S], Thm. 1.6a).
2.2. Subdifferential characterization of regularity and subdifferentially critical values. From now on all spaces are Banach and separable. Given a set $Q \subseteq \mathcal{X}$, a linear functional $x^*$ is said to be Hadamard normal to $Q$ at $x \in \text{cl}Q$ if

$$
\langle x^*, u_n - x \rangle \leq o(\|u_n - x\|), \quad \text{whenever } u_n \to x, \quad \frac{u_n - x}{\|u_n - x\|} \to h.
$$

The collection of all Hadamard normals is a closed convex cone in $\mathcal{X}^*$, which we denote by $\mathcal{N}_-(Q, x)$. If $F : \mathcal{X} \rightrightarrows \mathcal{Y}$, then the Hadamard coderivative of $F$ at $(x, y) \in \text{Graph } F$ is the set-valued mapping from $\mathcal{Y}^*$ into $\mathcal{X}^*$ defined by

$$
D^*_F(x, y)(y^*) = \{x^* : (x^*, -y^*) \in \mathcal{N}_-(\text{Graph } F, (x, y))\}.
$$

If the graph of $F$ is locally closed, then ([8], Thm. 3.1a)

$$
surF(x^* y) \geq \liminf_{(x, y) \rightarrow (\bar{x}, \bar{y}) \in \text{Graph } F} \inf \{\|x^*\| : x^* \in D^*_F(\bar{x}, \bar{y})(y^*), \|y^*\| = 1\}.
$$

The equality holds if both spaces are finite dimensional ([14], Thm. 9.43). It also holds if both spaces are Asplund and the Hadamard coderivative is replaced by the Fréchet coderivative ([13], Thm. 4.5). The choice of Hadamard normals in this paper is dictated by the fact that some of the most important separable spaces (like $\mathcal{C}$ and $\mathcal{L}^1$) are not Asplund and in a non-Asplund Banach space a fairly nice set (e.g., the complement of an open convex set) may have the Fréchet normal cone identically equal to $\{0\}$. On the contrary, the Hadamard (often also called Dini-Hadamard) normal cone is densely nontrivial on the boundary of every set.

The Clarke normal cone $\mathcal{N}_C(Q, x)$ to $Q$ at $x$ is the closed convex hull of the set $\mathcal{N}(Q, x)$ (called the limiting normal cone) of all possible weak$^*$-star limits of sequences $(x^*_n)$ such that $x^*_n \in \mathcal{N}_-(Q, x_n)$ for some $x_n \in Q$ norm converging to $x$. The Clarke coderivative $D^*_C F(x, y)$ of $F$ at $(x, y)$ is defined exactly as the Hadamard coderivative if instead of the Hadamard normal cone we use the Clarke normal cone.

**Definition 2.** We say that $y$ is a (proper) H-critical (resp. C-critical) value of $F$ if there is an $x$ such that $y \in F(x)$ and

$$
\liminf_{(x, y) \rightarrow (\bar{x}, \bar{y}) \in \text{Graph } F} \inf \{\|x^*\| : x^* \in D^*_C F(\bar{x}, \bar{y})(y^*), \|y^*\| = 1\} = 0
$$

(respectively, $\inf \{\|x^*\| : x^* \in D^*_C F(x, y)(y^*), \|y^*\| = 1\} = 0$).

As follows from the definitions and quoted results, any critical value is H-critical and, if the spaces are finite dimensional, any H-critical value is both critical and C-critical. A simple example of a regular value which is C-critical is given by the zero value of the function $f(x_1, x_2) = |x_1| - |x_2|$ on $\mathbb{R}^2$.

2.3. Fredholm submanifolds. The building blocks for the sets to be mainly considered in the paper are submanifolds of Banach spaces with the induced metric. To be specific, let $\mathcal{L}$ be a closed subspace of $\mathcal{X}$. We shall call (following [1]) a subset $\mathcal{M} \subset \mathcal{X}$ a $C^k$-submanifold of $\mathcal{X}$ modelled on $\mathcal{L}$ if for any $x \in \mathcal{M}$ there is a neighborhood $U_x \subset \mathcal{X}$ of zero and a $C^k$-diffeomorphism $\varphi_x$ of $U_x$ onto a neighborhood $V$ of $x$ which carries $U_x \cap \mathcal{L}$ onto the intersection of $V$ with $\mathcal{M}$. We shall call $(U_x, \varphi_x)$ a local representation of $\mathcal{M}$ at $x$. Clearly, the graph of a $C^k$-mapping $F$ from an open set $U \subset \mathcal{X}$ into $\mathcal{Y}$ is a $C^k$-manifold modelled by $\mathcal{X}$ with local representations given by e.g. $U_x = (U - x) \times \mathcal{Y}$, $\varphi_x(u) = (x + u, F(x + u))$. 


So let \( M \subset X \) be a \( C^1 \)-submanifold modelled on a Banach space \( L \). We shall denote by \( T_xM \) the tangent space to \( M \) at \( x \in M \) (clearly \( T_xM = \varphi'_x(0)(L) \)) and by \( N_xM \) the normal cone to \( M \) at \( x \) (the annihilator of \( T_xM \)):

\[
N_xM = (T_xM)^\perp = \{ x^* \in X^* : \langle x^*, u \rangle = 0, \ \forall u \in T_xM \}.
\]

We observe that for a \( C^1 \)-manifold, \( N_xM = N_x(M, x) \).

From now on we use the term manifold to refer to submanifolds in a Banach space modelled on its closed subspaces.

**Proposition 2.** Let \( M_1 \) and \( M_2 \) be two \( C^1 \)-manifolds in \( X \) and \( \mathfrak{F} \in M_1 \cap M_2 \). Assume that there are neighborhoods \( U_1 \) and \( U_2 \) of \( \mathfrak{F} \) in \( \mathfrak{F} \) and a diffeomorphism \( \varphi \) of \( U_1 \cap M_1 \) onto \( U_2 \cap M_2 \) such that \( \varphi(\mathfrak{F}) = \mathfrak{F} \) and

\[
\|\varphi(x) - \varphi(x') - (x - x')\| \leq r(x, x')\|x - x'\|,
\]

where \( r(x, x') \to 0 \) when \( x, x' \to M_1 \). Finally let \( F \) be a \( C^1 \)-mapping from a neighborhood of \( X \), containing \( \mathfrak{F} \), into another Banach space \( Y \). Then

\[
\text{surf}F|_{M_1}(\mathfrak{F}) = \text{surf}F|_{M_2}(\mathfrak{F}).
\]

Here and below \( F|M \) stands for the restriction of \( F \) to \( M \).

**Proof.** We have \( \|F(x) - F(x') - F'(\mathfrak{F})(x - x')\| \leq r(x, x')\|x - x'\| \), with \( r(x, x') \to 0 \) when \( x, x' \to \mathfrak{F} \) (no loss of generality follows from writing here the same \( r(x, x') \) as above). Set \( F_1 = F|_{U_1 \cap M_1} \) and consider also the mappings \( G = F \circ \varphi \) and \( H = F_1 - G \). Then for \( x, x' \in M_1 \) sufficiently close to \( \mathfrak{F} \) we have

\[
\|H(x) - H(x')\| \leq \|F'(\mathfrak{F})(x - x') - (\varphi(x) - \varphi(x'))\|
+ (r(x, x') + r(\varphi(x), \varphi(x'))\|x - x'\| \leq \rho(x, x')\|x - x'\|
\]

where \( \rho(x, x') \to 0 \) as \( x, x' \to \mathfrak{F}, x, x' \in M_1 \).

Take a small \( \varepsilon > 0 \) and choose a \( \delta > 0 \) such that \( \rho(x, x') \leq \varepsilon \) if \( x, x' \) are within \( \delta \) of \( \mathfrak{F} \) (and belong to \( M_1 \)). Then the above inequality means that \( H \) satisfies the Lipschitz condition with constant \( \varepsilon \) in a small neighborhood of \( \mathfrak{F} \). This in turn means (see [S], Thm. 1.6) that the rates of surjection of \( F_1 \) and \( G \) differ by at most \( \varepsilon \). As \( \varepsilon \) can be chosen arbitrarily small, this implies that the rates are equal.

By (1), \( \|\varphi(x) - \varphi(x')\| \leq (1 + \varepsilon)\lambda \) if \( x, x' \in \mathfrak{B}(\mathfrak{F}, \delta) \cap M_1 \) and \( \|x - x'\| \leq \lambda \). So for \( x \in M_1 \) close to \( \mathfrak{F} \) and \( \lambda > 0 \) sufficiently small, \( \varphi(\mathfrak{B}(x, \lambda)) \subset \mathfrak{B}(\varphi(x), (1 + \varepsilon)\lambda) \), so that \( \text{Sur}F_2(\varphi(x), (1 + \varepsilon)\lambda) \geq \text{Sur}G(x)(\lambda) \) for all such \( x \) and \( \lambda \), whence

\[
(1 + \varepsilon)\text{Sur}F_2(\varphi(x)) \geq \text{Sur}G(x) \text{ for all } x \text{ in a neighborhood of } \mathfrak{F}.
\]

As we can choose such a neighborhood for any \( \varepsilon > 0 \), we finally get \( \text{Sur}F_2(\mathfrak{F}) \geq \text{Sur}G(\mathfrak{F}) = \text{Sur}F_1(\mathfrak{F}) \).

To get the opposite inequality, we change the roles of \( M_1 \) and \( M_2 \) by replacing \( \varphi \) by \( \psi = \varphi^{-1} \). Clearly (1) implies the analogous property for \( \psi \).

**Corollary 3.** Let \( M \subset X \) be a \( C^1 \)-manifold, let \( \mathfrak{F} \in M \), let \( L = T_{\mathfrak{F}}M \) and let, as above, \( F : X \to Y \) be a \( C^1 \)-mapping defined in a neighborhood of \( \mathfrak{F} \). Then

\[
\text{surf}F'(\mathfrak{F})(L) = \text{surf}F|_{\mathfrak{F} + L}(x) = \text{surf}F|M(\mathfrak{F}).
\]

**Proof.** Set \( M_1 = \mathfrak{F} + L \), \( M_2 = M \) and use \( L \) as the model space for the manifold. Take a local representation \((U, \psi)\) for \( M \) at \( \mathfrak{F} \) and set \( \varphi(x) = \mathfrak{F} + \psi(x - \mathfrak{F}) \).

Let \( M \subset X \times Y \) be a \( C^1 \)-manifold. We can consider it as the graph of a set-valued mapping \( F : X \to Y \). Denote by \( M_Y \) the restriction to \( M \) of the projection \( P : X \times Y \to Y \). By Proposition 1, \((\bar{x}, \bar{y}) \in M \) is a regular point of \( F \) if and only if it is a regular point of \( M_Y \).
Definition 3. Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces, and let $\mathcal{M}$ be a $C^1$-manifold in $\mathcal{X} \times \mathcal{Y}$ modelled on a closed subspace $\mathcal{L} \subset \mathcal{X} \times \mathcal{Y}$. We say that $\mathcal{M}$ is a Fredholm manifold with respect to $\mathcal{Y}$ (or just Fredholm for short) if $P \circ \varphi$ is a Fredholm mapping (from $\mathcal{L}$ into $\mathcal{Y}$) for any local representation $(U, \varphi)$ of $\mathcal{M}$. (Recall (see e.g. [15, 1]) that a $C^1$-mapping from a smooth manifold $\mathcal{M}$ into a Banach space is Fredholm if its derivative at every point is a Fredholm linear operator.)

In particular, a closed subspace $\mathcal{L}$ of $\mathcal{X} \times \mathcal{Y}$ is a Fredholm subspace if the restriction to $\mathcal{L}$ of the projection onto $\mathcal{Y}$ is a Fredholm operator from $\mathcal{L}$ into $\mathcal{Y}$. We shall call the index of the restriction the index of $\mathcal{L}$. It is clear that $\mathcal{M}$ is a Fredholm manifold if for any $(x, y) \in \mathcal{M}$ the tangent space $\mathcal{T}_{(x,y)} \mathcal{M}$ is Fredholm.

Recall that a linear bounded operator $A : X \to Y$ is Fredholm if $\dim(\ker A) < \infty$ and $\text{Im } A$ is a closed subspace of finite codimension. The index of a Fredholm operator is $\text{ind } A = \dim(\ker A) - \text{codim}(\text{Im } A)$. The set of Fredholm operators is open, and the index is a continuous function of the operator (e.g. [10, Thm. 5.22]). Hence for any local representation $(U, \varphi)$ the index of $P \circ \varphi$ is constant on $U$. Thus, if $\mathcal{M}$ is a connected Fredholm manifold, we can define the index $\text{ind } \mathcal{M}$ of $\mathcal{M}$ as the common index of tangent spaces to $\mathcal{M}$.

Definition 4. Let $\mathcal{M}$ be a $C^1$-manifold in $\mathcal{X} \times \mathcal{Y}$. We shall say that $\mathcal{M}$ is $\mathcal{Y}$-normally compact if the following property holds for any (norm) converging sequence $(x_n, y_n) \subset \mathcal{M}$:

$$(x^*, y^*) \in \mathcal{N}_{(x_n, y_n)} \mathcal{M}, \quad ||x_n^*|| \to 0, \quad y_n^* \to 0 \Rightarrow ||y_n^*|| \to 0.$$  

Proposition 4. If $\mathcal{L} \subset \mathcal{X} \times \mathcal{Y}$ is a Fredholm subspace, then $\mathcal{L}$ is $\mathcal{Y}$-normally compact.

Proof. The projection $P : \mathcal{L} \to \mathcal{Y}$ is by definition a Fredholm operator. Therefore (see e.g. [7, Thm. 6]) $||y_n^*|| \to 0$ if $y_n^* \to 0$ and $||P^* y_n^*|| \to 0$. Assume now that $(x_n^*, y_n^*) \in \mathcal{L}^\perp$ and $||x_n^*|| \to 0$, $y_n^* \to 0$. Set $l_n^* = P^* y_n^*$ (that is, $(l_n^*, (x, y)) = (y_n^*, y)$ for all $(x, y) \in \mathcal{L}$). We have

$$||l_n^*|| = \inf \{||u^*|| + ||v^*-y_n^*|| : (u^*, v^*) \in \mathcal{L}^\perp \} \leq ||x_n^*|| \to 0.$$  

Hence $||y_n^*|| \to 0$. 

Thus the implication in Definition 4 automatically holds for a Fredholm manifold if the limit point of the sequence belongs to the manifold. This means that $\mathcal{Y}$-normal compactness relates to the behavior of a Fredholm manifold near the boundary. The property automatically holds also if both spaces are finite dimensional.

2.4. Stratification. Usually a stratification of a set is defined as a locally finite partition into manifolds (strata) which fit together regularly in the sense that the frontier of any stratum $(\text{cl } \mathcal{M}_i) \setminus \mathcal{M}_i$ is a union of some other strata (the frontier condition). The latter property is not needed in this paper, so we (as some other authors do in a similar situation; see e.g. [9]) mean by a $C^k$-stratification of a subset of a Banach space just a locally finite partition of the set into $C^k$-submanifolds. However, we do require that strata meet each other in a certain regular way.

Definition 5. Let $Q \subset \mathcal{X} \times \mathcal{Y}$ admit a $C^1$-stratification $(\mathcal{M}_j)$. We shall say that the stratification is normally regular if for any $i \neq j$ and any sequence $(x_n, y_n) \subset \mathcal{M}_i$ converging to a certain $(x, y) \in \mathcal{M}_j$ the following holds:

$$(\text{NR}) \quad (x_n^*, y_n^*) \in \mathcal{N}_{(x_n, y_n)} \mathcal{M}_i, \quad ||x_n^*-x^*|| \to 0, \quad y_n^* \to y^* \Rightarrow (x^*, y^*) \in \mathcal{N}_{(x,y)} \mathcal{M}_j.$$
We note some asymmetry in the type of convergence considered in the definition. But this is precisely the type of convergence that plays a central part in infinite dimensional variational analysis. The asymmetry disappears if both spaces are finite dimensional. In this case normal regularity is equivalent to the condition (a) of Whitney [16]. We emphasize that this condition is different from the frontier condition mentioned in the beginning of the subsection. Here is a simple example.

Let \( G \subset \mathbb{R}^2 \) be a closed convex set with \( \text{int} \, G \neq \emptyset \) and countably many points of nondifferentiability of the boundary having e.g. \((\bar{x}, \bar{y})\) an accumulation point. Set \( g(x, y) = (d((x, y), G))^2 \), and let \( Q \subset \mathbb{R}^3 \) be the union of the graph of \( g \) and the plane \( z = 0 \). Taking \( M_1 \) equal to the graph of \( g \) and \( M_2 = \{(x, y, z) : (x, y) \notin G, \; z = 0 \} \), we get a normally regular \( C^1 \)-stratification of \( Q \). On the other hand, any partition of the boundary of \( G \) into smooth manifolds will not be locally finite at \((\bar{x}, \bar{y})\).

A stratification of a finite dimensional set satisfying both the Whitney (a) condition and the frontier condition is usually referred to as a Whitney stratification.

The following simple proposition plays a crucial role in the proofs.

**Proposition 5.** Let \( Q \subset X \times Y \) admit a normally regular \( C^1 \)-stratification \((M_i)\). Let \((x_n, y_n) \to (x, y) \in M_j \) and assume that \((x_n^*, y_n^*) \in N_-(Q, (x_n, y_n))\), \( \|x_n^* - x^*\| \to 0 \), and \( y_n^* \to y^* \). Then \((x^*, y^*) \in N_{(x, y)}M_j\).

**Proof.** The local finiteness condition allows us to assume that all \((x_n, y_n)\) belong to the same stratum \( M_i \). Now the result follows from normal regularity and the inclusion \( N_-(Q, (x, y)) \subset N_{(x, y)}M_i \), which is immediate from the definitions. \( \square \)

### 2.5. Definable sets and mappings

We refer to [3, 4] for the definition of an o-minimal structure. For its use in this paper it is important to know that an o-minimal structure is a sequence of families of sets \( S_n \), each being a collection of subsets of \( \mathbb{R}^n \), having a number of good properties. Elements of these families are called definable sets (in the given o-minimal structure). A (set-valued) mapping \( F \) from \( \mathbb{R}^n \) to \( \mathbb{R}^m \) is definable if its graph belongs to \( S_{n+m} \). The main properties of definable sets and mappings needed here are:

(i) The closure of a definable set and the preimage and the image of a definable set under a definable mapping are definable sets. A consequence of the latter is that if a set-valued mapping \( F : \mathbb{R}^n \to \mathbb{R}^n \) is definable, then so is \( x \to \text{cl}(\text{conv} \, F(x)) \).

(ii) A set obtained from a finite family of definable sets by means of finitely many quantifiers \( \exists \) and \( \forall \) (applied only to variables) and Boolean operations is definable. To illustrate this remarkable property we observe that, given a closed definable set \( Q \subset \mathbb{R}^n \), the set-valued mappings \( x \to N_-(Q, x) \) and \( x \to N'(Q, x) \) are also definable. Indeed, \((x, x^*) \in \text{Graph} \, N_-(Q, \cdot)\) if and only if \( x \in Q \) and

\[
\forall \; \varepsilon > 0, \; \exists \delta > 0, \forall u \in Q \cap B(x, \delta), \quad \left( x^*, u - x \right) \leq \varepsilon \|u - x\|, \]

and \((x, x^*) \in \text{Graph} \, N'(Q, \cdot)\) if and only if \( x \in Q \) and

\[
\forall \; \varepsilon > 0, \; \exists (u, u^*) \in Q \times \mathbb{R}^n, \quad \left( \|u - x\| < \varepsilon, \; \|u^* - x^*\| < \varepsilon, \; u^* \in N_-(Q, u) \right). \]

(iii) A definable set admits a Whitney \( C^k \)-stratification for any \( k \) (see e.g. [3]). A set is tame (in an o-minimal structure) if its intersection with any ball is a definable set. A set-valued mapping is tame if its graph is tame.
3. Main theorems

Theorem 6. Assume that \( \mathcal{X} \) and \( \mathcal{Y} \) are finite dimensional spaces and \( F \) is a set-valued mapping from \( \mathcal{X} \) into \( \mathcal{Y} \) with locally closed graph. Suppose that the graph of \( F \) admits a \( C^k \)-stratification \( (\mathcal{M}_i) \) satisfying (NR). If \( k > \max_i (\dim \mathcal{M}_i) - \dim \mathcal{Y} \), then the set of proper \( C \)-critical values of \( F \) has Lebesgue measure zero in \( \mathcal{Y} \).

Theorem 7. Let \( \mathcal{X} \) and \( \mathcal{Y} \) be separable Banach spaces and \( F \) a set-valued mapping from \( \mathcal{X} \) into \( \mathcal{Y} \) with locally closed graph. Suppose that the graph of \( F \) admits a \( C^k \)-stratification \( (\mathcal{M}_i) \) into Fredholm \( \mathcal{Y} \)-normally compact manifolds satisfying (NR). If \( k \) is greater than the index of any \( \mathcal{M}_i \), then the collection of proper \( H \)-critical values of \( F \) is a set of the first Baire category in \( \mathcal{Y} \).

Proof. The proofs of both theorems follow basically the same pattern. So we shall give a unified proof with the stipulation of differences between the theorems whenever necessary. Let \( (\mathcal{M}_i) \) be the stratification. Denote by \( M_i \) the restriction of \( P \) (the projection \( \mathcal{X} \times \mathcal{Y} \to \mathcal{Y} \)) to \( \mathcal{M}_i \). Fix a certain \( \mathcal{M}_i \). As \( k > \dim \mathcal{M}_i - \dim \mathcal{Y} \), the set of critical values of \( M_i \) is a set of Lebesgue measure zero in \( \mathcal{Y} \) in the situation of the first theorem (by Sard’s theorem) or it is a set of first Baire category in the situation of Theorem 7 (by Smale’s theorem). As the partition into \( \mathcal{M}_i \) is locally finite and the spaces are separable, it follows that the union of all critical values of \( M_i \) is either a set of measure zero or a set of the first category in \( \mathcal{Y} \).

Thus, to prove the theorems it is sufficient to show that any \( C \)-critical value of \( F \) (in Theorem 6) or \( H \)-critical value (in Theorem 7) is a critical value of some of the \( \mathcal{M}_i \). In both cases the verification is rather simple. Consider first the situation of Theorem 6. If \( \mathcal{y} \) is a \( C \)-critical value of \( F \), then there exist \( \mathcal{x} \) and \( y^* \neq 0 \) such that \( \mathcal{y} \in F(\mathcal{x}) \) and \( (0,y^*) \in \mathcal{N}_C(\text{Graph } F, (\mathcal{x}, \mathcal{y})) \). As follows from the definition of Clarke’s normal cone, there are finitely many pairs \( (x_k^*, y_k^*) \in \mathcal{N}(\text{Graph } F, (\mathcal{x}, \mathcal{y})) \) (limiting normal cone) such that \( (0, y^*) \) is a convex combination of \( (x_k^*, y_k^*) \). By Proposition 5, \( (x_k^*, y_k^*) \in \mathcal{N}(\mathcal{x}, \mathcal{y}) \mathcal{M}_i \) (where \( \mathcal{M}_i \) is the stratum containing \( (\mathcal{x}, \mathcal{y}) \)); hence \( (0, y^*) \in \mathcal{N}(\mathcal{x}, \mathcal{y}) \mathcal{M}_i \), as the latter is a subspace.

In the case of Theorem 7, if \( \mathcal{y} \) is an \( H \)-critical value of \( F \), then there is an \( \mathcal{x} \) and a sequence of quadruples \( (x_n, y_n, x_n^*, y_n^*) \) such that \( (x_n, y_n) \in \text{Graph } F \) converge to \( (\mathcal{x}, \mathcal{y}) \), \( x_n^* \in D^*_F(x_n, y_n)(y_n^*) \), \( \|y_n^*\| = 1 \) and \( \|x_n^*\| \to 0 \). The local finiteness of the stratification allows us to assume that all \( (x_n, y_n) \) belong to some \( \mathcal{M}_j \) with \( j \neq i \).

We may assume that \( y_{n}^* \) weak* converges to some \( y^* \) (as \( \mathcal{Y} \) is separable). Then \( y^* \neq 0 \) by \( \mathcal{Y} \)-normal compactness. Applying again Proposition 5 we arrive at the same conclusion: there is a \( y^* \neq 0 \) such that \( (0, y^*) \in \mathcal{N}(\mathcal{x}, \mathcal{y}) \mathcal{M}_i \). This means that

\[
\mathcal{T}(\mathcal{x}, \mathcal{y}) \mathcal{M}_i = (\mathcal{N}(\mathcal{x}, \mathcal{y}) \mathcal{M}_i)_{\perp} \subset \mathcal{X} \times \{y^*\}_{\perp}
\]

(where \( \{y^*\}_{\perp} = \{y \in \mathcal{Y} : \langle y^*, y \rangle = 0 \} \)). Let \( \mathcal{L} \) stand for the subspace on the right-hand side. Then the projection of \( \mathcal{L} \) onto \( \mathcal{Y} \) is a proper subspace of \( \mathcal{Y} \). All the more, this applies to \( \mathcal{T}(\mathcal{x}, \mathcal{y}) \mathcal{M}_i \). Thus by Corollary 3, \( (\mathcal{x}, \mathcal{y}) \) is a singular point of \( \mathcal{M}_i \).

The following two corollaries of Theorem 6 strengthen the results of [2, 9] mentioned in the introduction.

Let \( f(x) \) be an l.s.c. function on a Banach space \( X \), finite at \( x \). Clarke’s generalized gradient of \( f \) at \( x \) is

\[
\partial_c f(x) = \{x^* \in X^* : \langle x^*, -1 \rangle \in \mathcal{N}_C(\text{epi } f, (x, f(x)))\}.
\]
A point $x$ is Clarke critical if $0 \in \partial_c f(x)$.

**Corollary 8.** Let $f(x)$ be a lower semicontinuous function on $\mathbb{R}^n$. Suppose that the epigraph of $f$: $\text{epi } f = \{(x, \alpha) \in X \times \mathbb{R} : \alpha \geq f(x)\}$ admits a $C^k$-stratification $(\mathcal{M}_i)$ satisfying (NR) (for $Y = \mathbb{R}$). Then the collection of proper $C$-critical values of $f$ is a set of Lebesgue measure zero, provided that $k \geq n$.

The proof consists in the observation that the points of a stratum $\mathcal{M}_i$ whose dimension is $n + 1$ cannot be critical. The theorem extends the result of [2] (Corollary 5) in one respect: it is assumed in [2] that there is a Whitney stratification with strata without vertical tangent spaces.

**Corollary 9.** Let $X$, $Y$ and $F$ be as in Theorem 6. Suppose the graph of $F$ is a locally closed tame set in a certain o-minimal structure. Then the collection of general $C$-critical values of $F$ is a $\sigma$-porous subset of $Y$. Moreover, if $\text{Graph } F$ is definable in an o-minimal structure, then the collection of general critical values of $F$ is a definable set with dimension strictly smaller than the dimension of $Y$.

The corollary is proved by applying the theorem to the set-valued mapping whose graph is the closure of $\text{Graph } F$, taking into account that the set of $C$-critical values of a definable set-valued mapping is definable (by 2.5). So the second part of the corollary follows from the fact that any definable set of measure zero must have smaller dimension. To prove the first part, it is enough to take into account that any set of dimension smaller than the dimension of $Y$ is porous in $Y$. In [9] the same result was proved for critical values.

**Concluding remarks.** 1) Theorem 6 contains Sard’s theorem: If $F : \mathbb{R}^n \to \mathbb{R}^m$ is $C^k$, then the graph of $F$ is a $C^k$-manifold of dimension $n$. Moreover, it is enough to assume $F$ single-valued on its domain to get the estimate $k > n - m$. Theorem 7 likewise contains the theorem of Smale but (as was observed by the reviewer) does not offer any a priori finite lower estimate for the order of smoothness.

2) It is possible to extend Theorem 6 to an infinite dimensional setting in the spirit of Theorem 7, but it would require a stronger version of (NR) with weak*-convergence of $x_n^*$.

**Acknowledgement**

I wish to express my thanks to the referee for thoughtful and helpful comments on the first version of the paper.

**References**


---

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


A.D. Ioffe, Directional compactness, scalarization and nonsmooth semi-Fredholm mappings, Nonlinear Analysis, TMA 29 (1997), 201-219. MR1446225 (98c:49043)


Department of Mathematics, Technion, Haifa 32000, Israel