manuscripta math. 118, 383–397 (2005)

Bogdan Bojarski · Piotr Hajłasz · Paweł Strzelecki

Sard's theorem for mappings in Hölder and Sobolev spaces

Received: 8 February 2005 / Revised version: 8 July 2005
Published online: 5 October 2005

Abstract. We prove various generalizations of classical Sard's theorem to mappings $f : M^m \to N^n$ between manifolds in Hölder and Sobolev classes. It turns out that if $f \in C^{k,\lambda}(M^m, N^n)$, then—for arbitrary k and λ—one can obtain estimates of the Hausdorff measure of the set of critical points in a typical level set $f^{-1}(y)$. The classical theorem of Sard holds true for $f \in C^k$ with sufficiently large k, i.e., $k > \max(m-n, 0)$; our estimates contain Sard's theorem (and improvements due to Dubovitskiĭ and Bates) as special cases. For Sobolev mappings between manifolds, we describe the structure of $f^{-1}(y)$.

1. Introduction

Throughout the paper we assume that M^m and N^n are smooth Riemannian manifolds of dimension m and n respectively. In 1942 A. Sard [22] (see also Sternberg's book [23]) proved the following theorem.

Theorem 1.1. *Let $f : M^m \to N^n$ be of class C^k, and let $S = \text{Crit}\, f$. If $k > \max(m-n, 0)$, then $\mathcal{H}^n(f(S)) = 0$.*

Here and in the sequel $\mathcal{H}^s$ denotes the s-dimensional Hausdorff measure (we shall follow the convention that $\mathcal{H}^s \equiv$ the counting measure for all $s \le 0$) and, for a C^1 mapping $f : M^m \to N^n$,

$$\text{Crit}\, f := \{x \in M^m \mid \text{rank}\, Df(x) < n\}$$

denotes the set of critical points of f.

All the authors were partially supported by KBN grant no 2 P03A 028 22. Piotr Hajlasz was also supported by NSF grant DMS-0500966.

B. Bojarski: Institute of Mathematics, Polish Academy of Sciences, ul. Śniadeckich 8, 00–950 Warszawa, Poland. e-mail: bojarski@impan.gov.pl

P. Hajłasz: Department of Mathematics, University of Pittsburgh, 301 Thackeray Hall, Pittsburgh, PA 15260, USA. e-mail: hajlasz@pitt.edu

P. Strzelecki: Institute of Mathematics, Warsaw University, ul. Banacha 2, 02–097 Warszawa, Poland. e-mail: pawelst@mimuw.edu.pl

Mathematics Subject Classification (2000): Primary: 46E35 Secondary: 41A63, 41A80, 41A99, 31B15

DOI: 10.1007/s00229-005-0590-1

It is well known that the assumptions of Theorem 1.1 are optimal within the scale of C^k spaces. Whitney [25] has given an example of a C^{m-1} function $f\colon (0,1)^m \to \mathbb{R}$ which is non-constant on a connected set of its critical points. Other examples of this kind can be found e.g. in [13], [15], [16]. Norton [20] explains how complicated all such examples must be: if A is a connected set of critical points of $f \in C^{k,\beta}(\mathbb{R}^m, \mathbb{R})$ and for each two points $x, y \in A$ there is a connected subset S of A such that $x, y \in S$ and the Hausdorff dimension of S is $< k+\beta$, then f is constant on A.

Now, several years after Sard's paper, A.Ya. Dubovitskiĭ [8] obtained a more general, better result. We give it here translating his notation to a more standard language.

Theorem 1.2. *Let $f\colon M^m \to N^n$ be a mapping of class C^k. Set $s = m-n-k+1$. Then*

$$\mathcal{H}^s(f^{-1}(y) \cap \operatorname{Crit} f) = 0 \qquad \textit{for } \mathcal{H}^n \textit{ a.e. } y \in N^n. \tag{1.1}$$

This theorem implies that for $\mathcal{H}^n$ almost all $y \in N^n$ the preimage $f^{-1}(y)$ consists of the regular part $f^{-1}(y) \setminus \operatorname{Crit} f$ and the critical part $f^{-1}(y) \cap \operatorname{Crit} f$. By the implicit function theorem the regular part is a C^k manifold of dimension $(m-n)$, while the critical part is a set of vanishing $\mathcal{H}^s$ measure.

It is clear that for $k > \max(m-n, 0)$ we have $s = m-n-k+1 \le 0$ and $\mathcal{H}^s$ in (1.1) is just the counting measure. Thus for such k (1.1) translates to $f^{-1}(y) \cap \operatorname{Crit} f = \emptyset$, i.e. Theorem 1.2 contains the classical Sard's theorem as a particular case.

Dubovitskiĭ, like a large number of mathematicians in the Soviet Union of that time, was isolated from the West and from the new results of western mathematics. He does not quote Sard's paper. On pages 398–402 of [8] he gives a variant of Whitney's example, and an example of a function $f \in C^k\big((0,1)^m, (0,1)^n\big)$ such that all sets $f^{-1}(y) \cap \operatorname{Crit} f$ have $(m-n-k)$-dimensional Hausdorff measure greater than zero, where m, k, n are positive integers such that $m-n-k > 0$. He attributes the first example to Menshov but gives no reference, and acknowledges Menshov, Novikov, Kronrod and Landis in his Introduction.

Taking all that into account, it seems likely that his Theorem 1.2 was proved independently from Sard's paper. The isolation mentioned above worked in fact both ways and, as far as we know, Dubovitskiĭ's work remained more or less unnoticed in the West.

Theorem 1.2 can be generalized to functions with Hölder continuous derivatives. (The definitions of $C^{k,\lambda}$ and $C^{k,\lambda+}$ are given in Section 2.)

Theorem 1.3. *Let $f \in C^{k,\lambda}(M^m, N^n)$ for some $k \ge 1$ and $\lambda \in (0,1)$, and let $s = m-n-(k+\lambda)+1$. Then for $\mathcal{H}^n$ almost all $y \in N^n$ the set $f^{-1}(y) \cap \operatorname{Crit} f$ is s-sigmafinite. In particular,*

$$\dim_H(f^{-1}(y) \cap \operatorname{Crit} f) \le s.$$

(Recall that a set E is called *s-sigmafinite* if and only if E is a countable union of E_i such that $\mathcal{H}^s(E_i) < \infty$.)

Lipschitz functions are differentiable a.e. and hence by the result of Whitney [26] $C^{k-1,1}$ functions coincide with C^k functions on complements of sets of arbitrarily small measure. This fact can be used to prove the following improvement of Theorem 1.3 for $\lambda = 1$.

Theorem 1.4. *Let $f \in C^{k-1,1}(M^m, N^n)$ for some $k \geq 1$, and let $s = m-n-k+1$. Then*

$$\mathcal{H}^s\big(f^{-1}(y) \cap \operatorname{Crit} f\big) = 0 \qquad \textit{for } \mathcal{H}^n \textit{ almost all } y \in N^n.$$

Remark. If $k = 1$, then f is Lipschitz, hence a.e. differentiable; in this case $\operatorname{Crit} f$ is defined as a *subset* of the set of those x for which $Df(x)$ exists.

In particular, for $k = m - n + 1$ we recover (with a different proof) the result of Bates [2] who proved that the differentiability condition in the classical Sard's theorem can be weakened from C^k to $C^{k-1,1}$.

For $\lambda < 1$ Hölder continuous functions are, in general, nowhere differentiable. Thus, in order to obtain a sharpened version of Theorem 1.3 with the conclusion as in Theorem 1.4, one has to strengthen the assumptions imposed on f.

Theorem 1.5. *Let $f \in C^{k,\lambda+}(M^m, N^n)$ for some $k \geq 1$ and $\lambda \in (0, 1)$, and let $s = m - n - (k + \lambda) + 1$. Then*

$$\mathcal{H}^s\big(f^{-1}(y) \cap \operatorname{Crit} f\big) = 0 \qquad \textit{for } \mathcal{H}^n \textit{ almost all } y \in N^n.$$

Let us now state a version of Sard's theorem which is valid for mappings in Sobolev spaces $W^{k,p}(M^m, N^n)$ of mappings between two manifolds. We write ∇ to denote the covariant derivative associated to the Riemannian metric g of M^m. Consider the subspace of $C^\infty(M^m, \mathbb{R})$ that consists of all those smooth f for which the norm

$$\|f\|_{k,p} := \left(\sum_{0\leq j\leq k} \int_{M^m} |\nabla^j f|^p \, d\,\mathrm{vol}\right)^{1/p}$$

is finite. $W^{k,p}(M^m, \mathbb{R})$ is defined as the completion of this subspace in the above norm. (If M^m is compact then this definition does not depend on the choice of g.) Next,

$$W^{k,p}(M^m, \mathbb{R}^\nu) = \{f = (f_1, \dots, f_\nu) \mid f_i \in W^{k,p}(M^m, \mathbb{R}) \quad \text{for all } i = 1, \dots, \nu\}.$$

Finally, for a manifold N^n embedded in $\mathbb{R}^\nu$ we set

$$W^{k,p}(M^m, N^n) = \{f \in W^{k,p}(M^m, \mathbb{R}^\nu) \mid f(x) \in N^n \quad \text{for a.e. } x \in M^m\}.$$

One also defines the space $W^{k,p}_{\mathrm{loc}}(M^m, N^n)$ consisting of those mappings for which the Sobolev norm is finite on every compact subset of M^m.

It is a well known result of Calderón and Zygmund [6] that functions in Sobolev spaces $W^{k,p}_{\mathrm{loc}}$ coincide with functions of class C^k on complements of sets of arbitrarily small measure. Applying Theorem 1.2 to C^k functions that agree with a given Sobolev function on large sets, we obtain the following result.

Theorem 1.6. *Assume that $f \in W^{k,p}_{\mathrm{loc}}(M^m, N^n)$. Then there exists a Borel representative of the map f such that the following is true.*

(i) For $\mathcal{H}^n$-almost all $y \in N^n$ we have

$$f^{-1}(y) = Z \cup \bigcup_{j=1}^{\infty} K_j$$

where $\mathcal{H}^{m-n-k+1}(Z) = 0$ and, for each j, $K_j \subset K_{j+1}$ and $K_j \subset S_j$ for some $(m-n)$-dimensional C^k-submanifold $S_j \subset M^m$.

(ii) If moreover $kp \geq n$ and M^m, N^n are both compact, then for $\mathcal{H}^n$-almost all $y \in N^n$ we have

$$\mathcal{H}^{m-n}\big(f^{-1}(y)\big) < \infty\,.$$

This means, in particular, that the level sets $f^{-1}(y)$ are $(\mathcal{H}^{m-n}, m-n)$-rectifiable of class C^k in the sense of Anzellotti and Serapioni [1]. It follows from (i) that in the general case the set $f^{-1}(y)$ has $\mathcal{H}^{m-n}$-sigmafinite measure, which is slightly less than the conclusion of (ii). (One can obtain variants of Theorem 1.6; we comment on this at the end of Section 4.)

A related but somewhat different result has been obtained by L. De Pascale [21]: if $f \in W^{k,p}(\mathbb{R}^m, \mathbb{R}^n)$ with $k = m-n+1 \geq 2$ and $p > m$, then $\mathcal{H}^n(f(\mathrm{Crit}\, f)) = 0$. (Note that the assumptions on k and p imply that in this case $W^{k,p}$ imbeds into C^1. Thus $\mathrm{Crit}\, f$ can be defined in a usual way.) For other results related to the Sard theorem for Sobolev mappings see [14] and [17].

For $k = 1$ Theorem 1.6 implies the following well known result, see e.g. [14], [17].

Corollary 1.7. *If $f \in W^{1,p}_{\mathrm{loc}}(M^m, N^n)$, then there exists a Borel representative of f such that $f^{-1}(y)$ is countably $(m-n)$-rectifiable for almost all $y \in N^n$.*

In differential geometry one considers, as a rule, only smooth mappings. Thus, for most applications it does not really matter that Sard's theorem can be applied only for functions or mappings with sufficiently high smoothness. On the other hand, assuming that a map belongs to the Sobolev space $W^{k,p}$ for a sufficiently large k is not natural. Sobolev mappings appear mostly as solutions to nonlinear PDE or as critical points of variational functionals. This usually imposes serious restrictions on the number of derivatives, k. Thus, it is often impossible to apply the classical Sard's theorem directly; one has to use Theorem 1.2 which allows to cope with maps that have few derivatives.

The aim of this paper is twofold. First, we give a short, self-contained proof of Dubovitskiĭ's Theorem 1.2 (his original paper is difficult, uses awkward notation and has 38 pages) and of its improvements, Theorems 1.3–1.5. Second, we combine these results with approximation theorems for Sobolev functions and apply them to mappings in Sobolev spaces to obtain Theorem 1.6.

The rest of the paper is organized as follows. In Section 2 we recall the definitions of $C^{k,\lambda}$, $C^{k,\lambda+}$, Hausdorff measure, Hausdorff dimension, and state some auxiliary lemmata (the lemma of Morse in various guises, and an inequality of Eilenberg). The necessary prerequisites from the theory of Sobolev spaces (including the theorem of Calderón and Zygmund, the co-area formula and the Nirenberg inequality) are explained briefly at the end of Section 2. Section 3 contains the proofs of Theorems 1.2–1.5. Finally, in Section 4 we prove Sard's theorem for Sobolev mappings, Theorem 1.6.

2. Prerequisites for the proofs

2.1. Spaces of Hölder continuous functions

In the sequel, two variants of Hölder spaces are used. Let $\lambda \in (0, 1]$. We say that $f \in C^{k,\lambda}(\mathbb{R}^m, \mathbb{R}^n)$ if and only if f is of class C^k and for each compact set K there exists a constant $M = M_K$ such that for each partial derivative D^α of order $|\alpha| = k$

$$|D^\alpha f(x) - D^\alpha f(y)| \leq M|x - y|^\lambda \qquad \text{for all } x, y \in K.$$

The class $C^{k,\lambda+}(\mathbb{R}^m, \mathbb{R}^n)$ is a proper subset of $C^{k,\lambda}$: a function $f \in C^{k,\lambda+}(\mathbb{R}^m, \mathbb{R}^n)$ if and only if $f \in C^k$ and for each compact set K there exists a nondecreasing continuous function $\omega\colon \mathbb{R} \to \mathbb{R}$ with $\omega(0) = 0$ such that for each α with $|\alpha| = k$,

$$|D^\alpha f(x) - D^\alpha f(y)| \leq \omega\big(|x - y|\big)|x - y|^\lambda \qquad \text{for all } x, y \in K.$$

Both definitions have obvious generalizations to mappings of Riemannian manifolds.

2.2. The inverse function theorem in Hölder spaces

We shall need the following variant of inverse function theorem. The proof is sketched e.g. in Norton's paper [20]; we repeat here his sketch for the sake of completeness.

Theorem 2.1. *Let $\Phi : \mathbb{R}^n \to \mathbb{R}^n$, $\Phi(0) = 0$, be a mapping of class $C^{k,\lambda}$ (resp., of class $C^{k,\lambda+}$), $k \geq 1$, $\lambda \in (0, 1]$, such that $D\Phi(0)$ is a linear isomorphism of $\mathbb{R}^n$. Then Φ^{-1} exists in a neighbourhood of 0 and is also of class $C^{k,\lambda}$ (resp., of class $C^{k,\lambda+}$).*

Proof. By the standard inverse function theorem, see e.g. [19], $\Phi^{-1} \in C^k$ in a neighbourhood of 0. Moreover, $D(\Phi^{-1})(y) = \big(D\Phi(\Phi^{-1}(y))\big)^{-1}$, i.e., the differential $D(\Phi^{-1})$ is given as a composition of three mappings,

$$D(\Phi^{-1}) = \mathrm{Inv} \circ D\Phi \circ \Phi^{-1}.$$

Since the inverse $\mathrm{Inv}\colon A \mapsto A^{-1}$ is of class C^∞ on the set of invertible matrices, $D\Phi$ is of class $C^{k-1,\lambda}$ (resp., $C^{k-1,\lambda+}$), and $\Phi^{-1} \in C^k$, we conclude the whole proof by applying the following Lemma. □

Lemma 2.2. *If G is of class $C^{k-1,\lambda}$ (resp., $C^{k-1,\lambda+}$) and $H \in C^k$, $k \geq 1$, then $G \circ H$ is of class $C^{k-1,\lambda}$ (resp., $C^{k-1,\lambda+}$).*

Proof. Induction on k (the case $k = 1$ is easy; for the induction step one has to write $D(G \circ H) = DG \circ H \cdot DH$ and use the induction hypothesis). □

A similar reasoning yields the following.

Lemma 2.3. *Let $k \geq 1$. If G, H are of class $C^{k,\lambda}$ (resp., $C^{k,\lambda+}$), then $G \circ H$ is of class $C^{k,\lambda}$ (resp., $C^{k,\lambda+}$).*

2.3. The lemma of Morse and its variants

The proofs of Sard's theorem and its various generalizations are usually based on a famous lemma of A.P. Morse. We recall two versions of this lemma which are suitable for our purposes.

Lemma 2.4 (A.P. Morse)**.** *Let $A \subset \mathbb{R}^m$. Fix a positive integer k. Then $A = \bigcup_{i=0}^{\infty} A_i$, where the A_i, $i \geq 0$, have the following property:*

Let $f \in C^k(V)$, where V is an open neighbourhood of A, be such that $A \subset$ Crit f. Then there exist nondecreasing functions $b_i : \mathbb{R} \to \mathbb{R}$ such that $b_i(\varepsilon) \to 0$ as $\varepsilon \to 0$ and

$$|f(x) - f(y)| \leq b_i(|x-y|)|x-y|^k \qquad \text{for all } x, y \in A_i. \tag{2.1}$$

See e.g. Sternberg's book [23, Lemma 3.3] for a proof. Norton [20, Theorem 3] obtains the following generalization to Hölder classes.

Lemma 2.5 (generalized Morse lemma)**.** *Let $A \subset \mathbb{R}^m$. Fix a positive integer k and $\lambda \in (0, 1]$. Then the following statements hold true.*

(i) There exist $A_i \subset A$ such that $A = \bigcup_{i=0}^{\infty} A_i$, where the A_i, $i \geq 0$, have the following property:

Let $f \in C^{k,\lambda}(V)$, where V is an open neighbourhood of A, be such that $A \subset$ Crit f. Then there exist constants M_i such that

$$|f(x) - f(y)| \leq M_i|x-y|^{k+\lambda} \qquad \text{for all } x, y \in A_i. \tag{2.2}$$

(ii) There exist $A_i' \subset A$ such that $A = \bigcup_{i=0}^{\infty} A_i'$, where the A_i', $i \geq 0$, have the following property:

Let $f \in C^{k,\lambda+}(V)$, where V is an open neighbourhood of A, be such that $A \subset$ Crit f. Then there exist nondecreasing functions $b_i : \mathbb{R} \to \mathbb{R}$ such that $b_i(\varepsilon) \to 0$ as $\varepsilon \to 0$ and

$$|f(x) - f(y)| \leq b_i\big(|x-y|\big)|x-y|^{k+\lambda} \qquad \text{for all } x, y \in A_i'. \tag{2.3}$$

2.4. Hausdorff measure

If X is an arbitrary metric space and $s \geq 0$, then for $E \subset X$ one defines

$$\mathcal{H}^s_\varepsilon(E) := \omega_s 2^{-s} \inf \sum_{j=1}^{\infty} (\operatorname{diam} A_j)^s ,$$

where $\omega_s = \pi^{s/2}/\Gamma(1+\frac{s}{2})$ is a normalizing constant (this is the volume of $B^s(0,1) \subset \mathbb{R}^s$ when s is a positive integer). The infimum is taken over all countable coverings $\{A_j\}_{j=1,2,\ldots}$ of E with all A_j having their diameter less than ε. It is clear that $\mathcal{H}^s_\varepsilon(E)$ is a nonincreasing function of ε; thus, the s-dimensional Hausdorff measure of E,

$$\mathcal{H}^s(E) := \lim_{\varepsilon \to 0} \mathcal{H}^s_\varepsilon(E),$$

is always well defined. $\mathcal{H}^s$ is an outer measure. All Borel sets in X are $\mathcal{H}^s$-measurable. $\mathcal{H}^0$ is the counting measure. We adopt the following notation: $\mathcal{H}^s \equiv \mathcal{H}^0$ for $s < 0$. Hausdorff dimension of a set $E \subset X$ is defined as

$$\dim_H(E) = \inf\{s \geq 0 : \mathcal{H}^s(E) = 0\} .$$

To define the Hausdorff measure on a manifold, one needs a Riemannian metric. However, sets of zero $\mathcal{H}^s$-measure are well defined on an arbitrary manifold: their definition does not depend on the choice of Riemannian structure. For $s = \dim M$ the s-dimensional Hausdorff measure on a Riemannian manifold M coincides with the standard volume form of M.

2.5. Eilenberg's inequality

We now state a general form of Eilenberg's inequality. A metric space is called boundedly compact if bounded and closed sets are compact.

Theorem 2.6. *Let $f : X \to Y$ be a Lipschitz mapping between two boundedly compact metric spaces X and Y. Let m and n be real numbers such that $0 \leq n \leq m$. Assume that a subset E of X is $\mathcal{H}^m$–measurable and $\mathcal{H}^m(E) < \infty$. Then*

(i) $f^{-1}(y) \cap E$ is $\mathcal{H}^{m-n}$–measurable for $\mathcal{H}^n$ almost all y,
(ii) $y \mapsto \mathcal{H}^{m-n}(f^{-1}(y) \cap E)$ is $\mathcal{H}^n$–measurable, and moreover

$$\int_Y \mathcal{H}^{m-n}(f^{-1}(y) \cap E)\, d\mathcal{H}^n(y) \leq (\operatorname{Lip} f)^n \frac{\omega_n \omega_{m-n}}{\omega_m} \mathcal{H}^m(E) . \tag{2.4}$$

See Federer's monograph [10], Theorem 2.10.25 and the remarks in 2.10.26.

Remark. Davies [7] has proved a stronger result: inequality (2.4) holds for Lipschitz mappings between arbitrary metric spaces (in this general case one replaces the integral by the so-called upper integral). This generalization does not play any role in our proofs, as we shall restrict ourselves to the case when X and Y are Euclidean spaces with a metric equal to some power of the classical Euclidean metric. A proof of Eilenberg's inequality in this case can also be found in Burago and Zalgaller's monograph [5].

2.6. Auxiliary facts on Sobolev spaces

For a domain $\Omega \subset \mathbb{R}^m$, $k = 1, 2, \dots$ and $1 \le p \le \infty$, the Sobolev space $W^{k,p}(\Omega)$ consists of all those functions $f \in L^p$ whose all distributional partial derivatives up to order k also belong to L^p. If $f \in W^{k,p}(U)$ for all $U \Subset \Omega$, then we say that $f \in W^{k,p}_{\mathrm{loc}}(\Omega)$. More information can be found in any monograph on Sobolev spaces, see e.g. Evans and Gariepy [9], Gilbarg and Trudinger [12, Chap. 7], or Ziemer [27].

We shall need the following theorem of Calderón and Zygmund [6] (see also [3], [4], [18], [24], [27] for more refined results).

Theorem 2.7. *Let $\Omega \subset \mathbb{R}^m$, $f \in W^{k,p}_{\mathrm{loc}}(\Omega)$ and $\varepsilon > 0$. Then there exists a closed set $F \subset \Omega$ and a function g of class $C^k(\Omega)$ such that*

$$\mathcal{H}^m(\Omega \setminus F) < \varepsilon \qquad \textit{and} \qquad g(x) = f(x) \quad \textit{for all } x \in F.$$

It is well known that $W^{k,\infty}_{\mathrm{loc}}(\Omega) = C^{k-1,1}(\Omega)$. The following special case of Theorem 2.7 was proved already by Whitney in [26].

Theorem 2.8. *Let $f \in C^{k-1,1}(\mathbb{R}^m)$, $k \ge 1$. Then for every $\varepsilon > 0$ there exists a closed set $F \subset \mathbb{R}^m$ such that $\mathcal{H}^m(\mathbb{R}^m \setminus F) < \varepsilon$ and there exists a function $g \in C^k(\mathbb{R}^m)$ such that all partial derivatives of f of order $\le k$ exist in F and equal to those of g. In particular $f = g$ in F.*

Both theorems generalize to mappings from a manifold to the Euclidean space. We will need the following special case of Nirenberg's inequality.

Theorem 2.9. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with smooth boundary. Let $k \ge 1$ be an integer and $1 \le p \le \infty$. Then $f \in W^{k,p}(\Omega) \cap L^\infty(\Omega)$ satisfies the inequality*

$$\|\nabla f\|_{L^{kp}(\Omega)} \le C(\Omega, k, p) \|f\|_{L^\infty}^{1-1/k} \|f\|_{W^{k,p}(\Omega)}^{1/k}.$$

For the proof see Friedman [11, Theorem 1.10.1]. The following result readily follows from Theorem 2.9.

Corollary 2.10. *Let M be a compact Riemannian manifold. If $f \in W^{k,p}(M) \cap L^\infty(M)$, then $\nabla f \in L^{kp}(M)$.*

We will also need a version of the co-area formula for Sobolev mappings between Riemannian manifolds.

Theorem 2.11. *Assume that $m \ge n$. If $f \in W^{1,1}_{\mathrm{loc}}(M^m, N^n)$, then there exists a Borel representative of f such that for every measurable set $E \subset M^m$*

$$\int_E |J_f(x)|\, d\mathcal{H}^m(x) = \int_{N^n} \mathcal{H}^{m-n}(f^{-1}(y) \cap E)\, d\mathcal{H}^n(y)\,, \tag{2.5}$$

where $|J_f(x)| = \sqrt{\det(Df_x) \circ (Df_x)^T}$.

The case of Lipschitz mappings is due to Federer [10, Theorem 3.2.11] and the case of Sobolev mappings was proved in [14] (see also [17]). If f is Lipschitz, then formula (2.5) holds for the continuous representative of f i.e. we do not have to modify f on a set of measure zero (like in the case $f \in W^{1,1}_{\mathrm{loc}}$) in order for (2.5) to be satisfied.

3. Dubovitskiĭ's theorem and its variants

Since Theorems 1.2–1.5 are of purely local nature, we may assume that $M^m = \mathbb{R}^m$, $N^n = \mathbb{R}^n$. We use these assumptions throughout this section. The whole reasoning is split into several lemmata.

Throughout this section, for a mapping $f = (f_1, \dots, f_n)\colon \mathbb{R}^m \to \mathbb{R}^n$ of class C^1 we set

$$K_r : = \{x \in \mathbb{R}^m : \operatorname{rank} Df(x) = r\}, \qquad r = 0, 1, 2, \dots$$

Sometimes we write $K_r = K_r(f)$, to make the dependence on f explicit. Note that $x \in K_0(f)$ if and only if $x \in \operatorname{Crit} f_i$ for all $i = 1, \dots, n$. Moreover,

$$\operatorname{Crit} f = \bigcup_{i=0}^{n-1} K_i \,. \tag{3.1}$$

3.1. Points of rank zero in level sets

We begin with an estimate of Hausdorff measure of $f^{-1}(y) \cap K_0$.

Lemma 3.1. *Let $f \in C^k(\mathbb{R}^m, \mathbb{R}^n)$, $k \geq 1$. Then*

$$\mathcal{H}^{m-kn}\big(f^{-1}(y) \cap K_0\big) = 0 \qquad \textit{for } \mathcal{H}^n\textit{-a.e. } y \in \mathbb{R}^n. \tag{3.2}$$

Proof. With no loss of generality assume that $K_0 \subset (0,1)^m$. We use the Morse lemma to decompose K_0 into the union of sets A_i. Next, we split each A_i,

$$A_i = \bigcup_{j=0}^{\infty} B_{ij}, \qquad \text{where } b_i(\operatorname{diam} B_{ij}) < \varepsilon \text{ for } j = 0, 1, 2, \dots$$

Thus

$$K_0 = \bigcup_{i,j=0}^{\infty} B_{ij}$$

and

$$|f(x) - f(y)| \leq \varepsilon |x - y|^k \qquad \text{for all } x, y \in B_{ij}. \tag{3.3}$$

Of course we can assume that the sets B_{ij} are pairwise disjoint. If we define a new metric in $\mathbb{R}^n$ by

$$d_k(p, q) = |p - q|^{1/k},$$

then (3.3) yields

$$d_k(f(x), f(y)) \leq \varepsilon^{1/k} |x - y| \qquad \text{for all } x, y \in B_{ij}, \tag{3.4}$$

i.e., $f : B_{ij} \to (\mathbb{R}^n, d_k)$ is an $\varepsilon^{1/k}$-Lipschitz mapping. Assume now that $m \geq kn$. By Eilenberg's inequality with n replaced by kn we obtain

$$\int_{\mathbb{R}^n} \mathcal{H}^{m-kn}(f^{-1}(y) \cap B_{ij})\, d\mathcal{H}^{kn}_{d_k}(y) \leq (\varepsilon^{1/k})^{kn} \frac{\omega_{kn}\omega_{m-kn}}{\omega_m} \mathcal{H}^m(B_{ij}), \quad (3.5)$$

where $\mathcal{H}^{m-kn}$ and $\mathcal{H}^m$ are standard Hausdorff measures in $\mathbb{R}^m$, while $\mathcal{H}^{kn}_{d_k}$ is the Hausdorff measure in $\mathbb{R}^n$ with respect to the metric d_k. Since $\mathcal{H}^{kn}_{d_k}$ coincides with the standard $\mathcal{H}^n$ measure on $\mathbb{R}^n$ up to the constant factor $\omega_{kn}2^{-kn}/(\omega_n 2^{-n})$, we conclude from (3.5) that

$$\int_{\mathbb{R}^n} \mathcal{H}^{m-kn}(f^{-1}(y) \cap B_{ij})\, d\mathcal{H}^n(y) \leq \varepsilon^n C(k, m, n)\mathcal{H}^m(B_{ij}), \quad (3.6)$$

and, upon adding up all B_{ij},

$$\int_{\mathbb{R}^n} \mathcal{H}^{m-kn}(f^{-1}(y) \cap K_0)\, d\mathcal{H}^n(y) \leq \varepsilon^n C(k, m, n)\mathcal{H}^m(K_0).$$

Since $\mathcal{H}^m(K_0) < \infty$ and ε can be made arbitrarily small, we conclude that the left hand side is zero. Thus, the integrand $\mathcal{H}^{m-kn}\big(f^{-1}(y) \cap K_0\big)$ must be zero for $\mathcal{H}^n$-a.e. $y \in \mathbb{R}^n$. If $m < kn$, then it easily follows from (3.3) that $\mathcal{H}^{m/k}(f(B_{ij})) < \infty$ and hence $\mathcal{H}^n(f(K_0)) = \sum_{i,j} \mathcal{H}^n(f(B_{ij})) = 0$. This yields (3.2) because in this case (3.2) means that $f^{-1}(y) \cap K_0 = \emptyset$ for $\mathcal{H}^n$-a.e. $y \in \mathbb{R}^n$. $\square$

For mappings $f \in C^{k,\lambda}(\mathbb{R}^m, \mathbb{R}^n)$, resp. $f \in C^{k,\lambda+}(\mathbb{R}^m, \mathbb{R}^n)$, a similar reasoning based on the generalized Morse lemma yields the following two lemmata. We will sketch the proof of the first lemma only. The proof of the second lemma is almost the same as that for Lemma 3.1. Later we will prove Lemma 3.4 which is an improvement of Lemma 3.2 for $\lambda = 1$.

Lemma 3.2. *Let $f \in C^{k,\lambda}(\mathbb{R}^m, \mathbb{R}^n)$, $k \geq 1$, $\lambda \in (0, 1]$. Then the set $f^{-1}(y) \cap K_0$ is $(m - (k + \lambda)n)$-sigmafinite for $\mathcal{H}^n$-a.e. $y \in \mathbb{R}^n$.*

Lemma 3.3. *Let $f \in C^{k,\lambda+}(\mathbb{R}^m, \mathbb{R}^n)$, $k \geq 1$, $\lambda \in (0, 1]$. Then*

$$\mathcal{H}^{m-(k+\lambda)n}\big(f^{-1}(y) \cap K_0\big) = 0 \qquad \text{for } \mathcal{H}^n\text{-a.e. } y \in \mathbb{R}^n. \quad (3.7)$$

To prove Lemma 3.2 we decompose the set K_0 into the A_i's using the generalized Morse lemma. If $B \subset \mathbb{R}^m$ is an arbitrary set, then the mapping $f : A_i \cap B \to (\mathbb{R}^n, d_{k+\lambda})$ is $M_i^{1/(k+\lambda)}$-Lipschitz and hence Eilenberg's inequality yields

$$\int_{\mathbb{R}^n} \mathcal{H}^{m-(k+\lambda)n}\Big(f^{-1}(y) \cap A_i \cap B\Big)\, d\mathcal{H}^n(y) \leq M_i^n C(m, n, k, \lambda)\mathcal{H}^m(A_i \cap B). \quad (3.8)$$

This inequality implies that $\mathcal{H}^{m-(k+\lambda)n}(f^{-1}(y) \cap A_i \cap B) < \infty$ for each cube B and $\mathcal{H}^n$-a.e. $y \in \mathbb{R}^n$. Hence the set $f^{-1}(y) \cap K_0$ is $(m - (k + \lambda)n)$-sigmafinite for $\mathcal{H}^n$-a.e. $y \in \mathbb{R}^n$.

Lemma 3.4. *Let $f \in C^{k-1,1}(\mathbb{R}^m, \mathbb{R}^n)$, $k \geq 2$. Then*

$$\mathcal{H}^{m-kn}\big(f^{-1}(y) \cap K_0\big) = 0 \qquad \text{for } \mathcal{H}^n\text{-a.e. } y \in \mathbb{R}^n. \tag{3.9}$$

Proof. Applying Whitney's theorem, Theorem 2.8, we select a sequence of closed sets $F_i \subset \mathbb{R}^m$ and of functions $g_i \in C^k(\mathbb{R}^m, \mathbb{R}^n)$ such that

$$\mathcal{H}^m(\mathbb{R}^m \setminus F_i) \leq \frac{1}{i} \quad \text{and} \quad D^\alpha f\big|_{F_i} = D^\alpha g_i\big|_{F_i} \quad \text{for all } |\alpha| \leq k,\ i = 1, 2, \ldots \tag{3.10}$$

In particular

$$K_0(f) \cap F_i = K_0(g_i) \cap F_i \qquad \text{for } i = 1, 2, \ldots$$

Let $B = \mathbb{R}^m \setminus \bigcup_{i=1}^\infty F_i$. To conclude the proof, it is enough to show that $\mathcal{H}^{m-kn}(f^{-1}(y) \cap K_0 \cap F_i) = 0$, $i = 1, 2, \ldots$ and $\mathcal{H}^{m-kn}(f^{-1}(y) \cap K_0 \cap B) = 0$ for $\mathcal{H}^n$-a.e. $y \in \mathbb{R}^n$. The first family of equalities (for $i = 1, 2, \ldots$) follows from Lemma 3.1 applied to g_i. Since $\mathcal{H}^m(B) = 0$, the last equality follows from (3.8) for $\lambda = 1$, with k replaced by $k - 1$.

3.2. Points of rank $r > 0$ in level sets

To prove Theorems 1.2–1.5, we need now to estimate the size of $f^{-1}(y) \cap K_r$ for $r > 0$. To this end, we apply (as one does in the proof of classical Sard's theorem) the inverse function theorem and reduce the problem to the case $r = 0$ considered above.

Lemma 3.5. *Let $f \in C^k(\mathbb{R}^m, \mathbb{R}^n)$, $k \geq 1$. Assume that $r \leq n - 1$. Then*

$$\mathcal{H}^{m-r-k(n-r)}\big(f^{-1}(y) \cap K_r\big) = 0 \qquad \text{for } \mathcal{H}^n\text{-a.e. } y \in \mathbb{R}^n.$$

Proof. Fix an arbitrary point $p \in K_r$. It is enough to prove that for some open neighbourhood U of p we have

$$\mathcal{H}^{m-r-k(n-r)}\big(f^{-1}(y) \cap K_r \cap U\big) = 0 \qquad \text{for } \mathcal{H}^n\text{-a.e. } y \in \mathbb{R}^n.$$

By an easy argument involving the inverse function theorem, we can assume with no loss of generality that $p = 0 \in \mathbb{R}^m$ and moreover that in some open cube $U = \{x : |x_i| < \delta, i = 1, \ldots, m\}$ centered at p the mapping f satisfies

$$f_i(x_1, \ldots, x_m) = x_i \qquad \text{for all } i = 1, \ldots, r. \tag{3.11}$$

Indeed, by renumbering the coordinates if necessary, we can assume that

$$\det\big[\partial f_i/\partial x_j(p)\big]_{i,j=1,\ldots,r} \neq 0.$$

Hence the mapping $\Phi : \mathbb{R}^m \to \mathbb{R}^m$ defined by

$$\Phi(x_1, \ldots, x_m) = (f_1(x_1, \ldots, x_m), \ldots, f_r(x_1, \ldots, x_m), x_{r+1}, \ldots, x_m)$$

satisfies $\det D\Phi(p) \neq 0$. Thus by the inverse mapping theorem Φ^{-1} is of class C^k in a neighbourhood of $\Phi(p)$. Now it is easy to see that

$$(f \circ \Phi^{-1})_i(x_1, \dots, x_m) = x_i \qquad \text{for all } i = 1, \dots, r.$$

Since the diffeomorphism Φ maps $K_r(f)$ onto $K_r(f \circ \Phi^{-1})$ it suffices to prove the lemma for f replaced by $f \circ \Phi^{-1}$. That means we can assume (3.11).

We write $U = U' \times U''$ and $x = (x', x'')$, where

$$x' = (x_1, \dots, x_r), \qquad x'' = (x_{r+1}, \dots, x_n),$$

and U', U'' denote cubes of dimensions r and $m-r$, respectively. For each $x' \in U'$ consider an auxiliary mapping $F_{x'} : \mathbb{R}^{m-r} \supset U'' \to \mathbb{R}^{n-r}$ defined by

$$F_{x'}(x'') := (f_{r+1}(x), \dots, f_n(x)). \tag{3.12}$$

With this notation $f(x', x'') = (x', F_{x'}(x''))$. Note that $(x', x'') \in K_r$ if and only if $\operatorname{rank} DF_{x'}(x'') = 0$ i.e., $x'' \in K_0(F_{x'})$, where the Jacobi matrix $DF_{x'}(x'')$ is computed with respect to the variables x''. We have $F_{x'}$ in C^k. Let Z denote the set of all those $y \in \mathbb{R}^n$ for which

$$\mathcal{H}^{m-r-k(n-r)}\big(f^{-1}(y) \cap K_r \cap U\big) > 0.$$

We write $y = (y', \tilde{y})$ where $y' = (y_1, \dots, y_r) \in \mathbb{R}^r$ and $\tilde{y} = (y_{r+1}, \dots, y_n) \in \mathbb{R}^{n-r}$. To show that $\mathcal{H}^n(Z) = 0$ we shall show that for each y' the slice

$$Z_{y'} = \{\tilde{y} \in \mathbb{R}^{n-r} : y = (y', \tilde{y}) \in Z\}.$$

satisfies $\mathcal{H}^{n-r}(Z_{y'}) = 0$. The whole lemma will follow then from Fubini's theorem.

Now, fix $y' \in \mathbb{R}^r$. For each $y = (y', \tilde{y}) \in \mathbb{R}^n$ (3.11) yields

$$f^{-1}(y) \cap K_r \cap U = \big(\{y'\} \times F_{y'}^{-1}(\tilde{y})\big) \cap K_r \cap U.$$

Thus a point $x = (x', x'') \in \mathbb{R}^r \times \mathbb{R}^{m-r}$ belongs to $f^{-1}(y) \cap K_r \cap U$ if and only if $x' = y'$, $x'' \in F_{y'}^{-1}(\tilde{y}) \cap U''$ and $\operatorname{rank} DF_{y'}(x'') = 0$. In other words, the slice $Z_{y'}$ is equal to the set of those $\tilde{y} \in \mathbb{R}^{n-r}$ for which

$$\mathcal{H}^{m-r-k(n-r)}\big(F_{y'}^{-1}(\tilde{y}) \cap K_0(F_{y'}) \cap U''\big) > 0\,.$$

Therefore, Lemma 3.1 gives $\mathcal{H}^{n-r}(Z_{y'}) = 0$, and the argument is complete. $\square$

Theorem 1.2 follows immediately from the last lemma. Just observe that

$$\max_{0 \le r \le n-1} \big(m - r - k(n-r)\big) = m - n - k + 1.$$

To obtain Theorems 1.3, 1.5, and Theorem 1.4 for $k \geq 2$, we replace Lemma 3.5 by the following result.

Lemma 3.6. *Let $f \in C^{k,\lambda}(\mathbb{R}^m, \mathbb{R}^n)$, $k \geq 1$. Assume that $r \leq n-1$ and set $s = m - r - (k+\lambda)(n-r)$.*

(i) The set $f^{-1}(y) \cap K_r$ is s-sigmafinite for $\mathcal{H}^n$-a.e. $y \in \mathbb{R}^n$.

(ii) If $\lambda = 1$ or if $f \in C^{k,\lambda+}(\mathbb{R}^m, \mathbb{R}^n)$, then $\mathcal{H}^s(f^{-1}(y) \cap K_r) = 0$ for $\mathcal{H}^n$-a.e. $y \in \mathbb{R}^n$.

The proof is identical to the proof of Lemma 3.5. One only needs to: (a) adapt the construction of good coordinates, i.e., use Theorem 2.1 and Lemma 2.3 to prove that $f \circ \Phi^{-1}$ is of class $C^{k,\lambda}$ or $C^{k,\lambda+}$; (b) quote an appropriate lemma from Section 3.1 in the last step.

Theorem 1.4 for $k = 1$ follows directly from the co-area formula (2.5) with $E = \operatorname{Crit} f$ and the fact that $|J_f(x)| = 0$ for $x \in \operatorname{Crit} f$.

4. Sard's theorem for Sobolev mappings

In this section we prove Theorem 1.6.

Part (i). Assume that N^n is a submanifold of $\mathbb{R}^\nu$. With no loss of generality we may assume that $f(x) \in N^n$ for all $x \in M^m$. Locally, the manifold N^n has a tubular neighbourhood. Adding up the local tubular neighbourhoods we obtain an open set $U \subset \mathbb{R}^\nu$ such that $N^n \subset U$ and that the nearest point projection $\pi : U \to N^n$ is well defined and smooth. Applying Theorem 2.7, we pick $g_j \in C^k(M^m, \mathbb{R}^\nu)$ such that

$$\mathcal{H}^m \left(\left\{x \in M^m : f(x) \neq g_j(x)\right\}\right) < \frac{1}{2^j}, \qquad j = 1, 2, \ldots \tag{4.1}$$

Set

$$W_j := g_j^{-1}(U), \qquad h_j := \pi \circ g_j|_{W_j} : W_j \to N^n.$$

Obviously h_j is a mapping of class C^k from an open set $W_j \subset M^m$ into N^n. Applying Theorem 1.2, we write

$$h_j^{-1}(y) = M_j(y) \cup Z_j(y),$$

where, for almost all $y \in N^n$, $M_j(y)$ is an $(m-n)$-dimensional C^k submanifold of M^m, and $\mathcal{H}^{m-n-k+1}(Z_j(y)) = 0$. Set

$$D_s = \bigcap_{j=s}^{\infty} \left\{x \in M^m : f(x) = g_j(x)\right\}.$$

Obviously, $D_s \subset D_{s+1}$. Moreover, $\mathcal{H}^m \left(M^m \setminus \bigcup_{s=1}^\infty D_s\right) = 0$. Since the three mappings f, g_j and $h_j = \pi \circ g_j$ coincide on D_j, for $\mathcal{H}^n$-a.e. $y \in N^n$ we obtain

$$\begin{aligned} f^{-1}(y) \cap D_j &= g_j^{-1}(y) \cap D_j \\ &= h_j^{-1}(y) \cap D_j \\ &= \left[M_j(y) \cup Z_j(y)\right] \cap D_j. \end{aligned}$$

To complete the proof of part (i), it suffices now to define

$$Z := \bigcup_{j=1}^{\infty} \Big(Z_j(y) \cap D_j\Big), \qquad K_j := (M_j(y) \setminus Z) \cap D_j,$$

and to redefine f on $M^m \setminus \bigcup_{j=1}^{\infty} D_j$, making it constant on this set.

Part (ii). Since the manifold M^m is compact, $W^{k,p}_{\mathrm{loc}}(M^m, N^n) = W^{k,p}(M^m, N^n)$. Then compactness of N^n implies that $f \in W^{k,p}(M^m, N^n)$ is bounded as a mapping into $\mathbb{R}^\nu$. Hence Corollary 2.10 gives $Df \in L^{kp}(M^m) \subset L^n(M^m)$. Now, by Hölder's inequality, the Jacobian

$$|J_f(x)| = \sqrt{\det(Df_x) \circ (Df_x)^T}$$

is integrable, and the co-area formula for Sobolev mappings, Theorem 2.11, with $E = M^m$ yields

$$\int_{N^n} \mathcal{H}^{m-n}(f^{-1}(y))\, d\mathcal{H}^n < +\infty\,.$$

This completes the proof. □

Let us remark that in the proof of Theorem 1.6 one can replace the result of Calderón and Zygmund, Theorem 2.7, by more refined theorems, ascertaining that on complements of sets of arbitrarily small Bessel capacity quasicontinuous representatives of Sobolev functions $f \in W^{k,p}$ coincide with functions of class $C^{r,\alpha}$, $r + \alpha < k$, see [4] and [24]. This change yields slightly worse information on the differentiability class of S_j's and Hausdorff measure of the exceptional set Z in Theorem 1.6 (i). However, the advantage is that one needs to modify f not on an unspecified set of measure zero, but only on a set having appropriate capacity zero, provided that the representative of f is quasicontinuous.

References

[1] Anzellotti, G., Serapioni, R.: C^k-rectifiable sets. J. Reine Angew. Math. **453**, 1–20 (1994)

[2] Bates, S. M.: Towards a precise smoothness hypothesis in Sard's theorem. Proc. Am. Math. Soc. **117**, 279–283 (1993)

[3] Bojarski, B., Hajłasz, P.: Pointwise inequalities for Sobolev functions. Studia Math. **106**, 77–92 (1993)

[4] Bojarski, B., Hajłasz, P., Strzelecki, P.: Improved $C^{k,\lambda}$ approximation of higher order Sobolev functions in norm and capacity. Indiana Univ. Math. J. **51** (3), 507–540 (2002)

[5] Burago, Yu.D., Zalgaller, A.V.: Geometric Inequalities. Grundlehren der Math. Wissenschaften 285. Springer, Berlin, 1988

[6] Calderón, A.P., Zygmund, A.: Local properties of solutions of elliptic partial differential equations. Studia Math. **20**, 171–225 (1961)

[7] Davies, R.O.: Increasing sequences of sets and Hausdorff measure. Proc. London Math. Soc. **20**, 222–236 (1970)

[8] Dubovickiĭ, A. Ya.: On the structure of level sets of differentiable mappings of an n-dimensional cube into a k-dimensional cube. (Russian) Izv. Akad. Nauk SSSR. Ser. Mat. **21**, 371–408 (1957)

[9] Evans, L.C., Gariepy, R.F.: Measure Theory and Fine Properties of Functions. Studies in Advanced Mathematics, 1992

[10] Federer, H.: Geometric Measure Theory. Springer, Berlin, 1969
[11] Friedman, A.: Partial Differential Equations. Holt, Rinehart and Winston, Inc., New York-Montreal, Que.-London, 1969
[12] Gilbarg, D., Trudinger, N.: Elliptic Partial Differential Equations of Second Order. Springer, Berlin, 1983
[13] Grinberg, E.L.: On the smoothness hypothesis in Sard's theorem. Am. Math. Monthly **92** (10), 733–734 (1985)
[14] Hajłasz, P.: Sobolev mappings, co-area formula and related topics. In: Proceedings on Analysis and Geometry (Novosibirsk Akademgorodok, 1999), Izdat. Ross. Akad. Nauk Sib. Otd. Inst. Mat., Novosibirsk, 2000, pp. 227–254
[15] Hajłasz, P.: Whitney's example by way of Assouad's embedding. Proc. Am. Math. Soc. **131**, 3463–3467 (2003)
[16] Kaufman, R.: A singular map of a cube onto a square. J. Differential Geom. **14**, 593–594 (1979)
[17] Malý, J., Swanson, D., Ziemer, W.P.: The co-area formula for Sobolev mappings. Trans. Am. Math. Soc. **355** (2), 477–492 (2003)
[18] Michael, J., Ziemer, W.P.: A Lusin type approximation of Sobolev functions by smooth functions. Contemp. Math. **42**, 135–167 (1985)
[19] Narasimhan, R.: Analysis on Real and Complex Manifolds. Advanced Studies in Pure Mathematics. North Holland, Amsterdam, 1968
[20] Norton, A.: A critical set with nonnull image has large Hausdorff dimension. Trans. Am. Math. Soc. **296** (1), 367–376 (1986)
[21] De Pascale, L.: The Morse-Sard theorem in Sobolev spaces. Indiana Univ. Math. J. **50** (3), 1371–1386 (2001)
[22] Sard, A.: The measure of the critical values of differentiable maps. Bull. Am. Math. Soc. **48**, 883–890 (1942)
[23] Sternberg, S: Lectures On Differential Geometry. Prentice Hall, 1964
[24] Swanson, D.: Pointwise inequalities and approximation in fractional Sobolev spaces. Studia Math. **149**, 147–174 (2002)
[25] Whitney, H.: A function not constant on a connected set of critical points. Duke Math. J. **1**, 514–517 (1935)
[26] Whitney, H.: On totally differentiable and smooth functions. Pacific J. Math. **1**, 143–159 (1951)
[27] Ziemer, W.P.: Weakly Differentiable Functions. Springer, Berlin, 1989