J. Math. Pures Appl. 139 (2020) 1–34

Contents lists available at ScienceDirect

Journal de Mathématiques Pures et Appliquées

www.elsevier.com/locate/matpur

On some universal Morse–Sard type theorems

Adele Ferone [a], Mikhail V. Korobkov [b,c,*], Alba Roviello [a]

[a] *Dipartimento di Matematica e Fisica, Università degli studi della Campania "Luigi Vanvitelli", viale Lincoln 5, 81100, Caserta, Italy*
[b] *School of Mathematical Sciences, Fudan University, Shanghai 200433, China*
[c] *Sobolev Institute of Mathematics, pr-t Ac. Koptyug, 4, Novosibirsk, 630090, Russia*

ARTICLE INFO

Article history:
Received 1 September 2018
Available online 15 May 2020

Dedicated to the bright memory of Jean Bourgain, who inspired this area of research

MSC:
58C25
26B35
46E30

Keywords:
Sobolev–Lorentz mappings
Bessel potential spaces
Hölder mappings
Morse–Sard theorem
Dubovitskiĭ–Federer theorems

Mots-clés :
Fonctions de Sobolev–Lorentz
Espace potentiel de Bessel
Fonctions Hölder
Thèorème de Morse–Sard
Thèorèmes de Dubovitskiĭ–Federer

ABSTRACT

The classical Morse–Sard theorem claims that for a mapping $v\colon \mathbb{R}^n \to \mathbb{R}^{m+1}$ of class C^k the measure of critical values $v(Z_{v,m})$ is zero under condition $k \geq n-m$. Here the critical set, or m-critical set is defined as $Z_{v,m} = \{x \in \mathbb{R}^n : \operatorname{rank} \nabla v(x) \leq m\}$. Further Dubovitskiĭ in 1957 and independently Federer and Dubovitskiĭ in 1967 found some elegant extensions of this theorem to the case of other (e.g., lower) smoothness assumptions. They also established the sharpness of their results within the C^k category.

Here we formulate and prove a *bridge theorem* that includes all the above results as particular cases: namely, if a function $v : \mathbb{R}^n \to \mathbb{R}^d$ belongs to the Hölder class $C^{k,\alpha}$, $0 \leq \alpha \leq 1$, then for every $q > m$ the identity

$$\mathcal{H}^{\mu}(Z_{v,m} \cap v^{-1}(y)) = 0$$

holds for $\mathcal{H}^q$-almost all $y \in \mathbb{R}^d$, where $\mu = n - m - (k+\alpha)(q-m)$.

Intuitively, the sense of this bridge theorem is very close to *Heisenberg's uncertainty principle* in theoretical physics: the more precise is the information we receive on measure of the image of the critical set, the less precisely the preimages are described, and vice versa.

The result is new even for the classical C^k-case (when $\alpha = 0$); similar result is established for the Sobolev classes of mappings $W^k_p(\mathbb{R}^n, \mathbb{R}^d)$ with minimal integrability assumptions $p = \max(1, n/k)$, i.e., it guarantees in general only *the continuity* (not everywhere differentiability) of a mapping. However, using some N-properties for Sobolev mappings, established in our previous paper, we obtained that the sets of nondifferentiability points of Sobolev mappings are fortunately negligible in the above bridge theorem. We cover also the case of fractional Sobolev spaces.

The proofs of the most results are based on our previous joint papers with J. Bourgain and J. Kristensen (2013, 2015). We also crucially use very deep Y. Yomdin's entropy estimates of near critical values for polynomials (based on algebraic geometry tools).

---

[*] Corresponding author at: School of Mathematical Sciences, Fudan University, Shanghai 200433, China.
E-mail addresses: adele.ferone@unicampania.it (A. Ferone), korob@math.nsc.ru (M.V. Korobkov), alba.roviello@unicampania.it (A. Roviello).

https://doi.org/10.1016/j.matpur.2020.05.002
0021-7824/

RÉSUMÉ

Le théorème classique de Morse–Sard affirme que pour une fonction $v\colon \mathbb{R}^n \to \mathbb{R}^{m+1}$ de classe C^k la mesure des valeurs critiques $v(Z_{v,m})$ est nulle sous la condition $k \geq n-m$. Ici l'ensemble critique, ou m-critique, est défini comme $Z_{v,m} = \{x \in \mathbb{R}^n : \operatorname{rank} \nabla v(x) \leq m\}$. Dubovitskiĭ en 1957 et indépendamment Federer et Dubovitskiĭ en 1967 ont trouvé des élégantes extensions de ce théorème sous d'autres hypothèses de régularité. Ils ont également établi l'optimalité de leurs résultats dans l'espace C^k.
Ici, nous formulons et prouvons un *théorème pont* qui comprend tous les résultats ci-dessus comme cas particuliers. Plus précisément, si un fonction $v : \mathbb{R}^n \to \mathbb{R}^d$ appartient à la classe de Hölder $C^{k,\alpha}$, $0 \leq \alpha \leq 1$, alors pour chaque $q > m$ l'identité

$$\mathcal{H}^\mu(Z_{v,m} \cap v^{-1}(y)) = 0$$

est réalisée pour $\mathcal{H}^q$-presque tout $y \in \mathbb{R}^d$, où $\mu = n - m - (k+\alpha)(q-m)$.
Intuitivement, le sens de ce théorème pont est très proche du *principe d'incertitude de Heisenberg* en physique théorique : plus l'information que nous recevons sur la mesure de l'image de l'ensemble critique est précise, moins les préimages sont décrites avec précision, et viceversa.
Le résultat est nouveau même dans le cas classique C^k (lorsque $\alpha = 0$) ; un résultat similaire est établi pour les classes de fonctions de Sobolev $W^k_p(\mathbb{R}^n, \mathbb{R}^d)$ avec les hypothèses minimales d'intégrabilité $p = \max(1, n/k)$, i.e., il ne garantit en général *la continuité* (pas partout différentiabilité) d'une fonction. Cependant, en utilisant des N-propriétés pour les fonctions de Sobolev, établies dans un article précédent, nous avons obtenu que les ensembles des points de non différentiabilité des fonctions de Sobolev sont heureusement négligeables dans le théorème pont ci-dessus. Nous couvrons également le cas des espaces fractionnels de Sobolev.
Les preuves de la plupart des résultats sont basées sur nos précédents papiers avec J. Bourgain et J. Kristensen (2013, 2015). Nous utilisons aussi de facon cruciale les estimations d'entropie très profondes de valeurs critiques pour les polynômes (basées sur la géométrie algébrique outils).

1. Introduction

The Morse–Sard theorem in its classical form states that the image of the set of critical points of a C^{n-d+1} smooth mapping $v\colon \mathbb{R}^n \to \mathbb{R}^d$ has zero Lebesgue measure in $\mathbb{R}^d$. More precisely, assuming that $n \geq d$, the set of critical points for v is $Z_v = \{x \in \mathbb{R}^n : \operatorname{rank} \nabla v(x) < d\}$ and the conclusion is that

$$\mathcal{L}^d(v(Z_v)) = 0. \tag{1.1}$$

The theorem was proved by Morse [34] in the case $d = 1$ and subsequently by Sard [38] in the general vector–valued case. The celebrated results of Whitney [42] show that the C^{n-d+1} smoothness assumption on the mapping v is sharp. However, the following result gives valuable information also for less smooth mappings.

Theorem A *(Dubovitskiĭ 1957 [17]). Let $n, d, k \in \mathbb{N}$, and let $v\colon \mathbb{R}^n \to \mathbb{R}^d$ be a C^k–smooth mapping. Put $\nu = n - d - k + 1$. Then*

$$\mathcal{H}^\nu(Z_v \cap v^{-1}(y)) = 0 \qquad \textit{for a.a. } y \in \mathbb{R}^d, \tag{1.2}$$

where $\mathcal{H}^\nu$ denotes the ν–dimensional Hausdorff measure.

Here and in the following we interpret $\mathcal{H}^\beta$ as the counting measure when $\beta \le 0$. Thus for $k \ge n-d+1$ we have $\nu \le 0$, and $\mathcal{H}^\nu$ in (1.2) becomes simply the counting measure, so the Dubovitskiĭ theorem contains the Morse–Sard theorem as particular case.[1]

A few years later and almost simultaneously, Dubovitskiĭ [18] in 1967 and Federer [21, Theorem 3.4.3] in 1969[2] published another important generalization of the Morse–Sard theorem.

Theorem B *(Dubovitskiĭ–Federer). For* $n, k, d \in \mathbb{N}$ *let* $m \in \{0, \dots, \min(n,d)-1\}$ *and* $v\colon \mathbb{R}^n \to \mathbb{R}^d$ *be a* C^k*-smooth mapping. Put* $q_\circ = m + \frac{n-m}{k}$*. Then*

$$\mathcal{H}^{q_\circ}(v(Z_{v,m})) = 0, \tag{1.3}$$

where $Z_{v,m}$ *denotes the set of* m*–critical points of* v *defined as*

$$Z_{v,m} = \{x \in \mathbb{R}^n \,:\, \operatorname{rank} \nabla v(x) \le m\}.$$

In 2001 Moreira [33] extended the last result to the Hölder class $C^{k,\alpha}$, i.e., he proved that for a mapping $v \in C^{k,\alpha}(\mathbb{R}^n, \mathbb{R}^d)$ the equality (1.3) holds with $q_\circ = m + \frac{n-m}{k+\alpha}$.

In view of the wide range of applicability of the above results it is a natural and compelling problem to extend them to the classes of Sobolev mappings.

In the recent paper [26] by Hajłasz P., Korobkov M.V., and Kristensen J. for $k \le n$ and for Sobolev classes $W^k_p(\mathbb{R}^n, \mathbb{R}^d)$ it was proved a *bridge theorem* that includes Theorems A–B as particular cases (see below Theorem 1.3-(ii)). In the present paper we extend this result for the Hölder classes $C^{k,\alpha}$ and for Sobolev spaces $W^k_p(\mathbb{R}^n, \mathbb{R}^d)$ with arbitrary integer $k \ge 1$, and also for fractional Sobolev spaces $\mathscr{L}^{k+\alpha}_p$ (e.g., for Bessel potential spaces; see Theorems 1.1–1.3).

The integrability assumptions here are very minimal and sharp, they are of kind $p(k+\alpha) \ge n$, i.e., they guarantee in general only *the continuity* (not everywhere differentiability) of a mapping. However, we proved that the ‘bad’ set of nondifferentiability points of Sobolev mappings is fortunately negligible in the above bridge theorem (see Theorem 1.4) because of some Luzin type N–properties with respect to lower dimensional Hausdorff measures established in our previous papers [13,22,28].

Let us note, in the conclusion, that the Morse–Sard theorem for the Sobolev spaces was very fruitful in mathematical fluid mechanics, in particular, it was used in the recent solution of the so-called Leray’s problem for the steady Navier–Stokes system (see [29]).

1.1. Bridge F.-D.-theorems for the Hölder classes of mappings

We say that a mapping $v : \mathbb{R}^n \to \mathbb{R}^d$ belongs to the class $C^{k,\alpha}$ for some positive integer k and $0 < \alpha \le 1$ if there exists a constant $L \ge 0$ such that

$$|\nabla^k v(x) - \nabla^k v(y)| \le L\,|x-y|^\alpha \qquad \text{for all } x, y \in \mathbb{R}^n.$$

To simplify the notation, let us make the following agreement: for $\alpha = 0$ we identify $C^{k,\alpha}$ with usual spaces of C^k-smooth mappings. The following theorem is one of the main results of the paper.

---

[1] It is interesting to note that this first Dubovitskiĭ theorem remained almost unnoticed by West mathematicians for a long time; another proof was given in the recent paper [10] covering also some extensions to the case of Hölder spaces; see also [25] for the Sobolev case.

[2] Federer announced [20] his result in 1966, this announcement (without any proofs) was sent on 08.02.1966. For the historical details, Dubovitskiĭ sent his paper [18] (with complete proofs) a month earlier, on 10.01.1966.

Theorem 1.1. *Let* $m \in \{0, \dots, n-1\}$, $k \geq 0$, $0 \leq \alpha \leq 1$, $k+\alpha \geq 1$, $d > m$, *and* $v \in C^{k,\alpha}(\mathbb{R}^n, \mathbb{R}^d)$. *Then for any* $q \in (m, \infty)$ *the equality*

$$\mathcal{H}^{\mu_q}(Z_{v,m} \cap v^{-1}(y)) = 0 \qquad \textit{for } \mathcal{H}^q\textit{-a.a. } y \in \mathbb{R}^d$$

holds, where

$$\mu_q = n - m - (k+\alpha)(q-m),$$

and $Z_{v,m}$ *denotes the set of* m*-critical points of* v: $Z_{v,m} = \{x \in \mathbb{R}^n \,:\, \operatorname{rank} \nabla v(x) \leq m\}$.

Let us note, that for the classical C^k-case, i.e., when $\alpha = 0$, the behavior of the function μ_q is very natural:

$$\mu_q = 0 \qquad \text{for } q = q_\circ = m + \tfrac{n-m}{k} \qquad \text{(Dubovitskiĭ–Federer Theorem B);}$$

$$\mu_q < 0 \qquad \text{for } q > q_\circ \qquad \text{(Dubovitskiĭ–Federer Theorem B);}$$

$$\mu_q = \nu \qquad \text{for } q = m+1 \qquad \text{(Dubovitskiĭ Theorem A);}$$

$$\mu_q = n - m \qquad \text{for } q = m.$$

The last value cannot be improved in view of the trivial example of a linear mapping $L\colon \mathbb{R}^n \to \mathbb{R}^d$ of rank m.

Thus, Theorem 1.1 contains all the previous theorems (Morse–Sard, A, B and even the Bates theorem for $C^{k,1}$-Lipschitz functions [8]) as particular cases.

Intuitively, the sense of the Bridge Theorem 1.1 is very close to the *Heisenberg's uncertainty principle* in theoretical physics: the more precisely information we received on measure of the image of the critical set, the less precisely the preimages are described, and vice versa.

Remark 1.2. As we mentioned before, for $q = q_\circ = m + \frac{n-m}{k+\alpha}$ and $\mu_q = 0$ (as in the Dubovitskiĭ-Federer Theorem B) the assertion of Theorem 1.1 was proved in 2001 in the paper of Moreira [33]. For the minimal rank value $m = 0$ (i.e., when the gradient totally vanishes on the critical set) and $q = q_\circ = \frac{n}{k+\alpha}$, $\mu_q = 0$, the assertion of Theorem 1.1 was proved by Kucera [30] in 1972. Further, for partial case $q = m+1 = d$ (as in the Dubovitskiĭ theorem A) and under additional assumption that

$$|\nabla^k v(x) - \nabla^k v(y)| \leq \omega\big(|x-y|\big) \cdot |x-y|^\alpha \quad \text{with } \ \omega(r) \to 0 \text{ as } r \to 0, \tag{1.4}$$

the assertion of Theorem 1.1 was proved in the paper Bojarski B. et al. [10] in 2005. Under the same asymptotic assumption (1.4) the above Moreira result (i.e., when $q = q_\circ$, $\mu_q = 0$) was proved by Yomdin in the paper [43] in 1983.

1.2. Bridge F.-D.-theorems for mappings of Sobolev and fractional Sobolev spaces

Let $k \in \mathbb{N}$, $1 < p < \infty$ and $0 \leq \alpha < 1$. One of the most natural type of fractional Sobolev spaces is *(Bessel) potential spaces* $\mathscr{L}^{k+\alpha}_p$. (They are Sobolev analog of classical Hölder classes $C^{k,\alpha}$.)

Recall, that a function $v : \mathbb{R}^n \to \mathbb{R}^d$ belongs to the space $\mathscr{L}^{k+\alpha}_p$, if it is a convolution of a function $g \in L_p(\mathbb{R}^n)$ with the Bessel kernel $G_{k+\alpha}$, where $\widehat{G_{k+\alpha}}(\xi) = (1 + 4\pi^2\xi^2)^{-(k+\alpha)/2}$. It is well known that for the integer exponents (i.e., when $\alpha = 0$) one has the identity

$$\mathscr{L}^k_p(\mathbb{R}^n) = W^k_p(\mathbb{R}^n) \qquad \text{if} \qquad 1 < p < \infty, \tag{1.5}$$

where $W^k_p(\mathbb{R}^n)$ is the classical Sobolev space consisting of functions whose generalized derivatives up to order $\le k$ belongs to the Lebesgue space $L_p(\mathbb{R}^n)$.

As usual, if $(k+\alpha)p > n$, then functions from the potential space $\mathscr{L}^{k+\alpha}_p(\mathbb{R}^n)$ are continuous by Sobolev Theorem. But if $(k+\alpha)p = n$, then functions from potential spaces $\mathscr{L}^{k+\alpha}_p(\mathbb{R}^n)$ are discontinuous in general. Thus for this limiting case we need to consider the Bessel–Lorentz potential space $\mathscr{L}^{k+\alpha}_{p,1}(\mathbb{R}^n)$ to have the continuity. Namely, $\mathscr{L}^{k+\alpha}_{p,1}(\mathbb{R}^n, \mathbb{R}^d)$ denotes the space of functions which could be represented as a convolution of the Bessel potential $G_{k+\alpha}$ with a function g from the Lorentz space $L_{p,1}$ (see the definition of these spaces in the section 2). Similarly to (1.5), for the integer exponents (i.e., when $\alpha = 0$) one has the identity

$$\mathscr{L}^k_{p,1}(\mathbb{R}^n) = W^k_{p,1}(\mathbb{R}^n) \qquad \text{if} \qquad 1 < p < \infty, \tag{1.6}$$

where $W^k_{p,1}(\mathbb{R}^n)$ consists of all functions $v \in W^k_p(\mathbb{R}^n)$ whose partial derivatives up to order $\le k$ belongs to the Lorentz space $\mathrm{L}_{p,1}$ (see, e.g., [22]).

Theorem 1.3. *Let $m \in \{0, \dots, n-1\}$, $k \ge 1$, $d > m$, $0 \le \alpha < 1$, $p \ge 1$ and let $v : \mathbb{R}^n \to \mathbb{R}^d$ be a mapping for which one of the following cases holds:*

(i) $\alpha = 0$, $kp > n$, and $v \in W^k_p(\mathbb{R}^n, \mathbb{R}^d)$;
(ii) $\alpha = 0$, $kp = n$, and $v \in W^k_{p,1}(\mathbb{R}^n, \mathbb{R}^d)$;
(iii) $0 < \alpha < 1$, $p > 1$, $(k+\alpha)p > n$, and $v \in \mathscr{L}^{k+\alpha}_p(\mathbb{R}^n, \mathbb{R}^d)$;
(iv) $0 < \alpha < 1$, $p > 1$, $(k+\alpha)p = n$, and $v \in \mathscr{L}^{k+\alpha}_{p,1}(\mathbb{R}^n, \mathbb{R}^d)$.

Then the mapping v is continuous and for any $q \in (m, \infty)$ the equality

$$\mathcal{H}^{\mu_q}(Z_{v,m} \cap v^{-1}(y)) = 0 \qquad \textit{for } \ \mathcal{H}^q\textit{-a.a. } y \in \mathbb{R}^d$$

holds, where again

$$\mu_q = n - m - (k+\alpha)(q-m),$$

and $Z_{v,m}$ denotes the set of m-critical points of v: $Z_{v,m} = \{x \in \mathbb{R}^n \setminus A_v : \operatorname{rank} \nabla v(x) \le m\}$.

Here A_v means the set of ‘bad’ points at which either the function v is not differentiable or which are not the Lebesgue points for ∇v. Recall, that by approximation results (see, e.g., [40] and [28]) under conditions of Theorem 1.3 the equalities

$$\begin{aligned} \mathcal{H}^\tau(A_v) &= 0 \qquad && \forall \tau > \tau_* := n - (k+\alpha-1)p \qquad && \text{in cases (i), (iii);} \\ \mathcal{H}^{\tau_*}(A_v) = \mathcal{H}^p(A_v) &= 0 && \tau_* := n - (k+\alpha-1)p = p && \text{in cases (ii), (iv)} \end{aligned}$$

are valid (in particular, $A_v = \emptyset$ if $(k+\alpha-1)p > n$). However, it was proved in [22] that the impact of the “bad” set A_v is negligible in the Bridge D.-F. Theorem 1.3, i.e., the following statement holds:

Theorem 1.4 *([22]). Let the conditions of Theorem 1.3 be fulfilled for a function $v : \mathbb{R}^n \to \mathbb{R}^d$. Then*

$$\mathcal{H}^{\mu_q}(A_v \cap v^{-1}(y)) = 0 \qquad \textit{for } \ \mathcal{H}^q\textit{–a.a. } y \in \mathbb{R}^d$$

for any $q > m$.

Remark 1.5. Note, that since $\mu_q \le 0$ for $q \ge q_\circ = m + \frac{n-m}{k+\alpha}$, the assertions of Theorems 1.3–1.4 are equivalent to the equality $0 = \mathcal{H}^q\big[v(A_v \cup Z_{v,m})\big]$ for $q \ge q_\circ$, so it is sufficient to check the assertions of Theorems 1.3–1.4 for $q \in (m, q_\circ]$ only.

Remark 1.6. Note that in the pioneering paper by De Pascale [15] the assertion of the initial Morse–Sard theorem (1.1) (i.e., when $k = n - m$, $q = q_\circ = m + 1 = d$, $\mu_q = 0$) was obtained for the Sobolev classes $W^k_p(\mathbb{R}^n, \mathbb{R}^d)$ under additional assumption $p > n$ (in this case the classical embedding $W^k_p(\mathbb{R}^n, \mathbb{R}^d) \hookrightarrow C^{k-1}$ holds, so there are no problems with nondifferentiability points). For the same Sobolev class $W^k_p(\mathbb{R}^n, \mathbb{R}^d)$ with $p > n$ the assertion of the Dubovitskiĭ Theorem A was proved in the recent paper [25] by P. Hajłasz and S. Zimmermann. Finally, the assertion of Bridge Theorem 1.3-(ii) was proved in our previous paper [26] with P. Hajłasz and J. Kristensen.[3]

In conclusion, let us comment briefly that the merge ideas for the proofs are from our previous papers [13], [27,28] and [26]. In particular, the joint papers [12,13] by one of the authors with J. Bourgain and J. Kristensen contain many of the key ideas that allow us to consider nondifferentiable Sobolev mappings. As in [13] (and subsequently in [27]) we also crucially use Y. Yomdin's (see [43]) entropy estimates of near critical values for polynomials (recalled in Theorem 2.7 below). These Yomdin's results seems to be very deep and fruitful in the topic, see, e.g., the very recent paper [7] where the Morse-Sard theorems were proved for min-type functions and for Lipschitz selections.

In addition to the above mentioned papers there is a growing number of papers on the topic, including [5,6,8,14,24,25,35,36,41].

Some words about the structure of the paper. In the second section we give some basic definitions and recall some classical results in analysis, which are very useful tools in our study. In the third sections we give the proofs of main theorems formulated above. For a reader convenience, the most technical part is moved to the last section 4 (Appendix), where we obtain estimates for the critical values on a single n-dimensional cube. These estimates are strong enough and useful for a solution of the following more general

Problem C. Let S be a subset of critical set $Z_{v,m} = \{\operatorname{rank} \nabla v \le m\}$ and the equality $\mathcal{H}^\tau(S) = 0$ (or the inequality $\mathcal{H}^\tau(S) < \infty$) holds for some $\tau > 0$. Does it imply that $\mathcal{H}^\sigma(v(E)) = 0$ for some $\sigma = \sigma(\tau)$?

The complete solution to this problem is done in our new paper [23]; this solution based on the technique developed in the present paper.

Acknowledgment. The work of M.K. is supported by Mathematical Center in Akademgorodok, the agreement with Ministry of Science and High Education of the Russian Federation number 075-15-2019-1613.

2. Preliminaries

By an *n–dimensional interval* we mean a closed cube in $\mathbb{R}^n$ with sides parallel to the coordinate axes. If Q is an n–dimensional cubic interval then we write $\ell(Q)$ for its sidelength.

For a subset S of $\mathbb{R}^n$ we write $\mathcal{L}^n(S)$ for its outer Lebesgue measure (sometimes we use the symbol meas S for the same purpose). The m–dimensional Hausdorff measure is denoted by $\mathcal{H}^m$ and the m–dimensional Hausdorff content by $\mathcal{H}^m_\infty$. Recall that for any subset S of $\mathbb{R}^n$ we have by definition

$$\mathcal{H}^m(S) = \lim_{t \searrow 0} \mathcal{H}^m_t(S) = \sup_{t>0} \mathcal{H}^m_t(S),$$

---

[3] The only technical difference is that in [26] we used the notation $Z_{v,m} = \{x \in \mathbb{R}^n \setminus A_v : \operatorname{rank} \nabla v(x) < m\}$, i.e., there $m - 1$ plays the role of the parameter m of the present article.

where for each $0 < t \le \infty$,

$$\mathcal{H}^m_t(S) = \inf\left\{\omega_m \sum_{i=1}^{\infty}\Big(\frac{\operatorname{diam} S_i}{2}\Big)^m \ :\ \operatorname{diam} S_i \le t,\quad S \subset \bigcup_{i=1}^{\infty} S_i\right\},$$

where ω_m is the volume of m-dimensional unit ball. It is well known that $\mathcal{H}^n(S) = \mathcal{H}^n_\infty(S) = \mathcal{L}^n(S)$ for sets $S \subset \mathbb{R}^n$.

To simplify the notation, we write $\|f\|_{\mathrm{L}_p}$ instead of $\|f\|_{\mathrm{L}_p(\mathbb{R}^n)}$, etc.

The Sobolev space $\mathrm{W}^k_p(\mathbb{R}^n, \mathbb{R}^d)$ is as usual defined as consisting of those $\mathbb{R}^d$-valued functions $f \in \mathrm{L}_p(\mathbb{R}^n)$ whose distributional partial derivatives of orders $l \le k$ belong to $\mathrm{L}_p(\mathbb{R}^n)$ (for detailed definitions and differentiability properties of such functions see, e.g., [19], [32], [44], [16]). Denote by $\nabla^k f$ the vector-valued function consisting of all k-th order partial derivatives of f arranged in some fixed order. However, for the case of first order derivatives $k = 1$ we shall often think of $\nabla f(x)$ as the Jacobi matrix of f at x, thus the $d \times n$ matrix whose r-th row is the vector of partial derivatives of the r-th coordinate function.

We use the norm

$$\|f\|_{\mathrm{W}^k_p} = \|f\|_{\mathrm{L}_p} + \|\nabla f\|_{\mathrm{L}_p} + \cdots + \|\nabla^k f\|_{\mathrm{L}_p},$$

and unless otherwise specified all norms on the spaces $\mathbb{R}^s$ $(s \in \mathbb{N})$ will be the usual euclidean norms.

Working with locally integrable functions, we always assume that the precise representatives are chosen. If $w \in L_{1,\mathrm{loc}}(\Omega)$, then the precise representative w^* is defined for *all* $x \in \Omega$ by

$$w^*(x) = \begin{cases} \lim\limits_{r\searrow 0} \ \fint_{B(x,r)} w(z)\,\mathrm{d}z, & \text{if the limit exists and is finite,}\\ 0 & \text{otherwise,}\end{cases}$$

where the dashed integral as usual denotes the integral mean,

$$\fint_{B(x,r)} w(z)\,\mathrm{d}z = \frac{1}{\mathcal{L}^n(B(x,r))}\int_{B(x,r)} w(z)\,\mathrm{d}z,$$

and $B(x,r) = \{y : |y-x| < r\}$ is the open ball of radius r centered at x. Henceforth we omit special notation for the precise representative writing simply $w^* = w$.

If $k < n$, then it is well-known that functions from Sobolev spaces $\mathrm{W}^k_p(\mathbb{R}^n)$ are continuous for $p > \frac{n}{k}$ and could be discontinuous for $p \le p_\circ = \frac{n}{k}$ (see, e.g., [32,44]). The Sobolev–Lorentz space $\mathrm{W}^k_{p_\circ,1}(\mathbb{R}^n) \subset \mathrm{W}^k_{p_\circ}(\mathbb{R}^n)$ is a refinement of the corresponding Sobolev space. Among other things functions that are locally in $\mathrm{W}^k_{p_\circ,1}$ on $\mathbb{R}^n$ are in particular continuous (see, e.g., [27]).

Here we only mentioned the Lorentz space $\mathrm{L}_{p,1}$, and in this case one may rewrite the norm as (see for instance [31, Proposition 3.6])

$$\|f\|_{L_{p,1}} = \int_0^{+\infty} \big[\mathcal{L}^n(\{x \in \mathbb{R}^n : |f(x)| > t\})\big]^{\frac{1}{p}}\,\mathrm{d}t.$$

Of course, we have the inequality

$$\|f\|_{L_p} \le \|f\|_{L_{p,1}}. \tag{2.1}$$

By definition we put $\|g\|_{L_{p,1}(E)} := \|1_E \cdot g\|_{L_{p,1}}$, where 1_E is the indicator function of E.

Denote by $\mathrm{W}^k_{p,1}(\mathbb{R}^n)$ the space of all functions $v \in \mathrm{W}^k_p(\mathbb{R}^n)$ whose partial derivatives up to order $\le k$ have finite Lorentz norm $\mathrm{L}_{p,1}$.

For a function $f \in L_{1,\mathrm{loc}}(\mathbb{R}^n)$ we often use the classical Hardy–Littlewood maximal function:

$$\mathcal{M}f(x) = \sup_{r>0} \fint_{B(x,r)} |f(y)|\,\mathrm{d}y. \tag{2.2}$$

2.1. On potential spaces $\mathscr{L}^\alpha_p$

To simplify our descriptions, below in the next two subsections we will write α instead of $k+\alpha$, so here we assume that $\alpha \in \mathbb{R}_+$ (i.e., here not necessarily $\alpha < 1$, as in formulations of main results).

In the paper we deal with *(Bessel)-potential space* $\mathscr{L}^\alpha_p$ with $\alpha > 0$. Recall, that function $v : \mathbb{R}^n \to \mathbb{R}^d$ belongs to the space $\mathscr{L}^\alpha_p$, if it is a convolution of the Bessel kernel G_α with a function $g \in L_p(\mathbb{R}^n)$:

$$v = \mathscr{G}_\alpha(g) := G_\alpha * g,$$

where $\widehat{G_\alpha}(\xi) = (1+4\pi^2\xi^2)^{-\alpha/2}$. In particular,

$$\|v\|_{\mathscr{L}^\alpha_p} := \|g\|_{L_p}.$$

It is well known that

$$\mathscr{L}^\alpha_p(\mathbb{R}^n) = W^\alpha_p(\mathbb{R}^n) \qquad \text{if} \quad \alpha \in \mathbb{N} \quad \text{and} \quad 1 < p < \infty.$$

Recall, that the Bessel kernel is a radial function and it could be calculated as

$$G_\alpha(x) \;=\; a_\alpha \int_0^\infty t^{\frac{\alpha-n}{2}} e^{-\frac{\pi|x|^2}{t} - \frac{t}{4\pi}} \frac{dt}{t}, \qquad \forall \alpha > 0, \tag{2.3}$$

where a_α is some constant.

It is well known, that $G_\alpha(x) < a_\alpha\,|x|^{\alpha-n}$ for $0 < \alpha < n$ (see, e.g., [4, page 10]) and we need some simple technical extension of this fact to the derivatives.

Lemma 2.1. *If $0 < \alpha < n+2$, then for any integer $j \in \mathbb{N}$ the estimate*

$$\left|\nabla^j G_\alpha(x)\right| \le C\,|x|^{\alpha-n-j} \tag{2.4}$$

holds, where the constant C depends on α, n, j only.

Proof. Denote

$$f_\alpha(r) = \int_0^\infty t^{\frac{\alpha-n}{2}} e^{-\frac{\pi r^2}{t} - \frac{t}{4\pi}} \frac{dt}{t}.$$

Then by direct calculation

$$f'_\alpha(r) = -2\pi r \int_0^\infty t^{\frac{\alpha-n-2}{2}} e^{-\frac{\pi^2}{t} - \frac{t}{4\pi}} \frac{dt}{t} = -2\pi r^{\alpha-n-1} \int_0^\infty t^{\frac{\alpha-n-2}{2}} e^{-\frac{\pi}{t} - \frac{tr^2}{4\pi}} \frac{dt}{t}$$

(see, e.g., [4, page 13]). This finishes the proof for $j = 1$. The proof for $j > 1$ could be produced the same way by induction. □

2.2. On Lorentz potential spaces $\mathscr{L}^{\alpha}_{p,1}$

To cover some other limiting cases, denote by $\mathscr{L}^{\alpha}_{p,1}(\mathbb{R}^n, \mathbb{R}^d)$ the space of functions which could be represented as a convolution of the Bessel potential G_α with a function g from the Lorentz space $L_{p,1}$; respectively,

$$\|v\|_{\mathscr{L}^{\alpha}_{p,1}} := \|g\|_{L_{p,1}}.$$

Because of inequality (2.1), we have an evident inclusion

$$\mathscr{L}^{\alpha}_{p,1}(\mathbb{R}^n) \subset \mathscr{L}^{\alpha}_{p}(\mathbb{R}^n).$$

Theorem 2.2 *(see, e.g., Theorem 2.2 in [22], cf. with Lemma 3 on page 136 in [39]). Let $\alpha \geq 1$ and $1 < p < \infty$. Then $f \in \mathscr{L}^{\alpha}_{p,1}(\mathbb{R}^n)$ iff $f \in \mathscr{L}^{\alpha-1}_{p,1}(\mathbb{R}^n)$ and $\frac{\partial f}{\partial x_j} \in \mathscr{L}^{\alpha-1}_{p,1}(\mathbb{R}^n)$ for every $j = 1, \dots, n$.*

(Here for convenience we use the agreement that $\mathscr{L}^{\alpha}_{p}(\mathbb{R}^n) = L_p(\mathbb{R}^n)$ when $\alpha = 0$.)

Corollary 2.3. *Let $k \in \mathbb{N}$ and $1 < p < \infty$. Then $\mathscr{L}^{k}_{p,1}(\mathbb{R}^n) = W^{k}_{p,1}(\mathbb{R}^n)$, where $W^{k}_{p,1}(\mathbb{R}^n)$ is the space of functions such that all its distributional partial derivatives of order $\leq k$ belong to $L_{p,1}(\mathbb{R}^n)$.*

Note, that *locally* the space $W^{k}_{p,1,\mathrm{loc}}(\mathbb{R}^n)$ admits also a simpler description: it consists of functions f from the usual Sobolev space $W^{k}_{p,\mathrm{loc}}(\mathbb{R}^n)$ satisfying the additional condition $\nabla^k f \in L_{p,1,\mathrm{loc}}(\mathbb{R}^n)$ (i.e., this condition is on the highest derivatives only), see, e.g., [31].

2.3. Approximation of Sobolev functions by polynomials

For a mapping $u \in \mathrm{L}_1(Q, \mathbb{R}^d)$, $Q \subset \mathbb{R}^n$, $m \in \mathbb{N}$, define the polynomial $P_{Q,m}[u]$ of degree at most m by the following rule:

$$\int\limits_Q y^\gamma \left(u(y) - P_{Q,m}[u](y)\right) \mathrm{d}y = 0$$

for any multi-index $\gamma = (\gamma_1, \dots, \gamma_n)$ of length $|\gamma| = \gamma_1 + \dots + \gamma_n \leq m$.

The following well–known bounds will be used on several occasions.

Lemma 2.4 *(see, e.g., [27]). Suppose $v \in \mathrm{W}^k_1(\mathbb{R}^n, \mathbb{R}^d)$ with $k \geq n$. Then v is a continuous mapping and for any n-dimensional cubic interval $Q \subset \mathbb{R}^n$ the estimates*

$$\big\|v - P\big\|_{L_\infty(Q)} \leq C\ell(Q)^{k-n}\, \|\nabla^k v\|_{\mathrm{L}_1(Q)};$$

$$\big\|\nabla\big(v - P\big)\big\|_{L_\infty(Q)} \leq C\ell(Q)^{k-1-n}\, \|\nabla^k v\|_{\mathrm{L}_1(Q)} \qquad \text{if } \ k \geq n+1;$$

hold, where $P = P_{Q,k-1}[v]$ and C is a constant depending on n, d, k only. Moreover, the mapping $v_Q(y) = v(y) - P(y)$, $y \in Q$, can be extended from Q to the entire $\mathbb{R}^n$ such that the extension (denoted again) $v_Q \in \mathrm{W}^k_1(\mathbb{R}^n, \mathbb{R}^d)$ and

$$\|\nabla^k v_Q\|_{\mathrm{L}_1(\mathbb{R}^n)} \leq C_0 \|\nabla^k v\|_{\mathrm{L}_1(Q)}, \tag{2.5}$$

where C_0 also depends on n, d, k only.

2.4. Approximation of fractional Sobolev functions by polynomials

We need the following natural estimate whose analogs for Sobolev case are well-known (see, e.g., [32]).

Theorem 2.5. *Let $k \geq 1$, $0 \leq \alpha < 1$, $1 < p < \infty$, and $v \in \mathscr{L}^{k+\alpha}_p(\mathbb{R}^n, \mathbb{R}^d)$, i.e., $v = \mathscr{G}_{k+\alpha}(g) := G_{k+\alpha} * g$ for some $g \in L_p(\mathbb{R}^n)$. Suppose in addition that*

$$1 < k + \alpha < n + 2.$$

Then for any n-dimensional interval $Q \subset \mathbb{R}^n$ there exists a polynomial $P = P_Q$ of degree k such that the difference $v_Q = v - P$ satisfies the estimate

$$|\nabla v_Q(x)| \leq C \int\limits_Q \frac{\mathcal{M}g(y)}{|x-y|^{n-k-\alpha+1}}\, dy \qquad \forall x \in Q, \tag{2.6}$$

where $r = \ell(Q)$, the constant C depends on n, k, α, d, p only, and $\mathcal{M}g$ is the usual Hardy—Littlewood maximal function for g (see (2.2)).

Proof. Really, this theorem in essence was proved in the paper [22]. Let us recall some arguments from there. Fix an n-dimensional interval $Q \subset \mathbb{R}^n$ and denote by $2Q$ the double cube with the same center as Q of size $\ell(2Q) = 2\ell(Q)$. We have

$$v(x) = \int\limits_{\mathbb{R}^n} G_{k+\alpha}(x-y)\, g(y)\, dy.$$

Split the function v into the sum

$$v = v_1 + v_2, \tag{2.7}$$

where

$$v_1(x) := \int\limits_{2Q} g(y)\, G_{k+\alpha}(x-y)\, dy, \qquad v_2(x) := \int\limits_{\mathbb{R}^n \setminus 2Q} g(y)\, G_{k+\alpha}(x-y)\, dy.$$

From [22, Lemma A.1] and from the estimate

$$|\nabla G_{k+\alpha}(z)| \leq C|z|^{k+\alpha-n-1} \tag{2.8}$$

(see Lemma 2.1) it follows immediately that

$$|\nabla v_1(x)| \leq C \int\limits_Q \frac{\mathcal{M}g(y)}{|x-y|^{n-k-\alpha+1}}\, dy \qquad \forall x \in Q. \tag{2.9}$$

Analogously, from the similar estimate

$$|\nabla^j G_{k+\alpha}(z)| \leq C|z|^{k+\alpha-n-j} \tag{2.10}$$

[22, Lemma A.1] and from Lemma A.2 of the paper [22] and its proof, applying to the function $\nabla^k v_2$ with[4] parameter $\theta = 1-\alpha$, we obtain

$$\operatorname{diam}\bigl[\nabla^k v_2(Q)\bigr] \le C\, r^{\alpha-n} \int\limits_Q \mathcal{M}\, g(y)\, dy. \tag{2.11}$$

Take the corresponding approximate polynomial $P = P_Q$ of degree k, then for the difference $\tilde{v} = v_2 - P$ we obtain the following estimates:

$$\sup_{x\in Q} |\nabla^k \tilde{v}(x)| \le C\, r^{\alpha-n} \int\limits_Q \mathcal{M}\, g(y)\, dy, \tag{2.12}$$

$$\sup_{x\in Q} |\nabla \tilde{v}(x)| \le C\, r^{\alpha-n+k-1} \int\limits_Q \mathcal{M}\, g(y)\, dy. \tag{2.13}$$

Evidently,

$$r^{\alpha-n+k-1} \int\limits_Q \mathcal{M}\, g(y)\, dy \;\le\; C \int\limits_Q \frac{\mathcal{M}g(y)}{|x-y|^{n-k-\alpha+1}}\, dy \qquad \forall x \in Q.$$

From the last formula and inequalities (2.13), (2.9) the required estimate (2.6) follows directly. □

Remark 2.6. If under above conditions we have in addition $(k+\alpha-1)p > n$, then by Hölder inequality the estimate (2.6) implies

$$\sup_{x\in Q} |\nabla v_Q(x)| \le C\, r^{k+\alpha-1-\frac{n}{p}} \|\mathcal{M}g\|_{L_p(Q)}. \tag{2.14}$$

2.5. On Yomdin's entropy estimates for the nearcritical values of polynomials

For a subset A of $\mathbb{R}^d$ and $\varepsilon > 0$ the ε–entropy of A, denoted by $\operatorname{Ent}(\varepsilon, A)$, is the minimal number of closed balls of radius ε covering A. Further, for a linear map $L\colon \mathbb{R}^n \to \mathbb{R}^d$ we denote by $\lambda_j(L)$, $j = 1, \dots, d$, its singular values arranged in decreasing order: $\lambda_1(L) \ge \lambda_2(L) \ge \dots \ge \lambda_d(L)$. Geometrically the singular values are the lengths of the semiaxes of the ellipsoid $L(\partial B(0,1))$. We recall that the singular values of L coincide with the eigenvalues repeated according to multiplicity of the symmetric nonnegative linear map $\sqrt{LL^*}\colon \mathbb{R}^d \to \mathbb{R}^d$. Also for a mapping $f\colon \mathbb{R}^n \to \mathbb{R}^d$ that is approximately differentiable at $x \in \mathbb{R}^n$ put $\lambda_j(f,x) = \lambda_j(d_x f)$, where by $d_x f$ we denote the approximate differential of f at x. The next result is the basic ingredient of our proof.

Theorem 2.7 *([43]). Let $m \in \{0, \dots, n-1\}$ and $m < d$. Then for any polynomial $P\colon \mathbb{R}^n \to \mathbb{R}^d$ of degree at most k, for each n-dimensional cube $Q \subset \mathbb{R}^n$ of size $\ell(Q) = r > 0$, and for any number $\varepsilon > 0$ we have that*

$$\begin{aligned}\operatorname{Ent}\bigl(\varepsilon r, \{P(x) : x \in Q,\ \lambda_1 \le 1+\varepsilon, \dots, \lambda_m \le 1+\varepsilon, \lambda_{m+1} \le \varepsilon, \dots, \lambda_d \le \varepsilon\}\bigr)\\ \le C_Y\bigl(1 + \varepsilon^{-m}\bigr),\end{aligned}$$

where the constant C_Y depends on n, d, k, m only and for brevity we wrote $\lambda_j = \lambda_j(P,x)$.

---

[4] That means, that now our function $\nabla^k v_2$ plays the role of mapping v from arguments of [22, proof of Lemma A.2].

2.6. On Choquet type integrals

Recall the following classical theorem referred to D.R. Adams, see, e.g., [1]–[2] or [3].

Theorem 2.8. *Let $\beta > 0$, $n - \beta p > 0$, and $s > p > 1$. Then for any $g \in L_p(\mathbb{R}^n)$ the estimate*

$$\int\limits_0^\infty \mathcal{H}^\tau_\infty\big(\{x \in \mathbb{R}^n : \mathcal{M}\big(I_\beta g\big)(x) \geq t^{\frac{1}{s}}\}\big)\,\mathrm{d}t \leq C\|g\|^s_{\mathrm{L}_p} \tag{2.15}$$

holds with $\tau = \frac{s}{p}(n - \beta p)$, where C depends on $n,\ p,\ s,\ \beta$ only.

Here

$$I_\beta g(x) := \int\limits_{\mathbb{R}^n} \frac{g(y)}{|y - x|^{n-\beta}}\,\mathrm{d}y$$

is the classical Riesz potential of order β.

The above estimate (2.15) fails for the limiting case $s = p$. Namely, there exist functions $g \in L_p(\mathbb{R}^n)$ such that $|I_\beta g|(x) = +\infty$ on some set of positive $(n - \beta p)$–Hausdorff measure. One possible way to cover this limiting case $s = p$ is using the Lorentz norm instead of Lebesgue norm in the right hand side of (2.15). Such possibility was proved in the recent paper [28].

Theorem 2.9 *(see Theorem 0.2 in [28]). Let $\beta > 0$, $n - \beta p > 0$, and $p > 1$. Then for any $g \in L_p(\mathbb{R}^n)$ the estimate*

$$\int\limits_0^\infty \mathcal{H}^\tau_\infty\big(\{x \in \mathbb{R}^n : \mathcal{M}\big(I_\beta g\big)(x) \geq t^{\frac{1}{p}}\}\big)\,\mathrm{d}t \leq C\|g\|^p_{\mathrm{L}_{p,1}} \tag{2.16}$$

holds with $\tau = n - \beta p$, where C depends on $n,\ p,\ \beta$ only.

The above theorems are not fulfilled in general for $p = 1$. However, similar results hold in case $p = 1$ for derivatives of Sobolev mappings. Namely, the following Theorem was proved by D.R. Adams [2].

Theorem 2.10. *Let $k, l \in \{1, \dots, n\}$, $l < k$. Then for any function f from the Sobolev space $\mathrm{W}^k_1(\mathbb{R}^n)$ the estimates*

$$\int\limits_0^\infty \mathcal{H}^\tau_\infty\big(\{x \in \mathbb{R}^n : \mathcal{M}\big(\nabla^l f\big)(x) \geq t\}\big)\,\mathrm{d}t \leq C\|\nabla^k f\|_{\mathrm{L}_1} \tag{2.17}$$

hold, where $\tau = n - k + l$ and the constant C depends on n, k, l.

The application of above estimates on maximal functions is facilitated through the following simple Lipschitz type inequality (see for instance Lemma 2 in [16], cf. with [9]).

Lemma 2.11. *Let $u \in \mathrm{W}^1_1(\mathbb{R}^n, \mathbb{R}^d)$. Then for any ball $B \subset \mathbb{R}^n$ of radius $r > 0$ and for any number $\varepsilon > 0$ the estimate*

$$\operatorname{diam}(\{u(x) : x \in B,\ (\mathcal{M}\nabla u)(x) \leq \varepsilon\}) \leq C_M \varepsilon r$$

holds, where C_M is a constant depending on n, d only.

Using the similar calculations, one could obtain the following refinement of the above Lemma.

Lemma 2.12. *Let $u \in \mathrm{W}^1_1(Q, \mathbb{R}^d)$, where Q is an n-dimensional interval. Then for any ball $B \subset \mathbb{R}^n$ of radius $r > 0$ and for any number $\varepsilon > 0$ the estimate*

$$\operatorname{diam}(\{u(x) : x \in B \cap Q,\ (\mathcal{M}_Q \nabla u)(x) \le \varepsilon\}) \le C_M \varepsilon r$$

holds, where C_M is a constant depending on n, d only, and

$$\mathcal{M}_Q f := \mathcal{M}(1_Q \cdot f),$$

i.e.,

$$\mathcal{M}_Q f(x) = \sup_{r>0} \frac{1}{|B(x,r)|} \int\limits_{Q \cap B(x,r)} |f(y)|\,\mathrm{d}y. \tag{2.18}$$

2.7. On Fubini type theorems for graphs of continuous functions

Recall that by usual Fubini theorem, if a set $E \subset \mathbb{R}^2$ has a zero plane measure, then for $\mathcal{H}^1$-almost all straight lines L parallel to coordinate axes we have $\mathcal{H}^1(L \cap E) = 0$. The next result could be considered as functional Fubini type theorem.

Theorem 2.13 *(see Theorem 5.3 in [26]). Let $\mu \ge 0$, $q > 0$, and $v : \mathbb{R}^n \to \mathbb{R}^d$ be a continuous function. For a set $E \subset \mathbb{R}^n$ define the set function*

$$\Phi(E) = \inf_{E \subset \bigcup_j D_j} \sum_j \bigl(\operatorname{diam} D_j\bigr)^{\mu} \bigl[\operatorname{diam} v(D_j)\bigr]^q,$$

where the infimum is taken over all countable families of compact sets $\{D_j\}_{j \in \mathbb{N}}$ such that $E \subset \bigcup_j D_j$. Then $\Phi(\cdot)$ is a countably subadditive and the implication

$$\Phi(E) = 0 \;\Rightarrow\; \Bigl[\mathcal{H}^{\mu}\bigl(E \cap v^{-1}(y)\bigr) = 0 \quad \textit{for } \mathcal{H}^q\textit{-almost all } y \in \mathbb{R}^d\Bigr]$$

holds.

2.8. On local properties of considered potential spaces

Let $\mathcal{B}$ be some space of functions defined on $\mathbb{R}^n$. For a set $\Omega \subset \mathbb{R}^n$ define the space $\mathcal{B}_{\mathrm{loc}}(\Omega)$ in the following standard way:

$$\mathcal{B}_{\mathrm{loc}}(\Omega) := \{f : \Omega \to \mathbb{R} :\ \text{for any compact set } E \subset \Omega\ \exists g \in \mathcal{B} \text{ such that } f(x) = g(x)\ \forall x \in E\}.$$

Put for simplicity $\mathcal{B}_{\mathrm{loc}} = \mathcal{B}_{\mathrm{loc}}(\mathbb{R}^n)$.

It is well known that for $q > p > 1$ the inclusions

$$L_{q,\mathrm{loc}} \subset L_{p,1,\mathrm{loc}} \subset L_{p,\mathrm{loc}},$$

hold (see, e.g., [31]). Respectively, it is easy to see that for $\alpha > 0$ one has

$$\mathscr{L}^\alpha_{q,\mathrm{loc}} \subset \mathscr{L}^\alpha_{p,1,\mathrm{loc}} \subset \mathscr{L}^\alpha_{p,\mathrm{loc}}.$$

Since the Morse–Sard type theorems have a local nature, if we prove some of these theorems for $\mathscr{L}^\alpha_p$, then the same result will be valid for the spaces $\mathscr{L}^\alpha_{p,1}$ and $\mathscr{L}^\alpha_q$ for all $q > p$. Similarly, if we prove some Morse–Sard type theorems for $\mathscr{L}^\alpha_{p,1}$, then the same result will be valid for the spaces $\mathscr{L}^\alpha_q$ with $q > p$, etc.

2.9. Approximation by Hölder–smooth functions

We need also the following approximation result.

Theorem 2.14 *(see, e.g., Chapter 3 in [44] or [11]). Let $p > 1$, $k \in \mathbb{N}$, $\alpha \in (0, 1)$. Then for any $f \in \mathscr{L}^{k+\alpha}_p(\mathbb{R}^n)$ and for each $\varepsilon > 0$ there exist an open set $U \subset \mathbb{R}^n$ and a function $h \in \mathrm{C}^{k,\alpha}(\mathbb{R}^n)$ such that*

(i) $\mathcal{L}^n(U) < \varepsilon$;
(ii) each point $x \in \mathbb{R}^n \setminus U$ is an Lebesgue point for f and ∇f;
(iii) $f \equiv h$ and $\nabla f \equiv \nabla h$ on $\mathbb{R}^n \setminus U$.

Note, that in the cited references the approximation property is discussed for the case of Sobolev spaces W^k_p, but the proof for the $\mathscr{L}^{k+\alpha}_p(\mathbb{R}^n)$ space easily follows from the just mentioned Sobolev case and some standard arguments on real analysis concerning approximation limits and Whitney-type extension theorems for Hölder classes (see, e.g., Theorem 4 in [39, §2.3, Chapter 6]).

3. Proofs of the main results

3.1. Bridge Federer–Dubovitskiĭ theorem for Sobolev mappings

Recall, that *bridge Dubovitskiĭ–Federer Theorem* 1.3 for the case (ii) was proved in our previous paper [26]. The purpose here is to prove Theorem 1.3 (i). But of course, the case (i) with $k \le n$ follows immediately from the case (ii) (see the section 2.8). So we need to consider here only the situation Theorem 1.3 (i) with

$$k > n \quad \text{and} \quad p = 1.$$

Fix integers $m \in \{0, \dots, n-1\}$, $d > m$, $k > n$, and a mapping $v \in \mathrm{W}^k_1(\mathbb{R}^n, \mathbb{R}^d)$. Then, by Lemma 2.4 the function v is C^1-smooth.

Denote $Z_{v,m} = \{x \in \mathbb{R}^n : \operatorname{rank} \nabla v(x) \le m\}$. Fix a number $q > m$. Denote in this subsection

$$\mu = \mu_q = n - m - k(q - m).$$

Recall, that we need to consider only the case

$$q \in (m, q_\circ],$$

where $q_\circ = m + \frac{n-m}{k}$ (see Remark 1.5). Then by direct calculation we have $\mu \ge 0$.

The required assertion of the *bridge Dubovitskiĭ–Federer Theorem* 1.3-(i) is equivalent (by virtue of Theorem 2.13) to the identity

$$\Phi(Z_{v,m}) = 0,$$

where by definition

$$\Phi(E) := \inf_{E\subset\bigcup_j D_j} \sum_j \bigl(\operatorname{diam} D_j\bigr)^{\mu}\bigl[\operatorname{diam} v(D_j)\bigr]^{q}. \tag{3.1}$$

As indicated the infimum is taken over all countable families of compact sets $\{D_j\}_{j\in\mathbb{N}}$ such that $E\subset\bigcup_j D_j$.

Before embarking on the detailed proof we make some preliminary observations that allow us to make a few simplifying assumptions. We could assume without loss of generality that

$$|\nabla v(x)|\le 1 \qquad \forall x\in\mathbb{R}^n.$$

Denote $Z_v=Z_{v,m}$. Then from Theorem 4.6 (i), applied to our values of $q,\mu=\mu_q$ we obtain immediately

Lemma 3.1. *Let $q\in(m,q_\circ]$. Then for any sufficiently small n-dimensional interval $Q\subset\mathbb{R}^n$ the estimate*

$$\Phi(Z_v\cap Q)\le C\,\ell(Q)^{n(m-q+1)}\cdot\|\nabla^k v\|^{q-m}_{\mathrm{L}_1(Q)} \tag{3.2}$$

holds, where the constant C depends on n,m,k,d only.

(Indeed, we have by direct elementary calculation, that now the exponent $q+\mu+(k-1-n)(q-m)$ from the formula (4.49) coincides with $n(m-q+1)$ from (3.2).)

Note, that by our assumptions $k>n$, therefore

$$q\le q_\circ=m+\frac{n-m}{k}<m+1.$$

Corollary 3.2. *Let $q\in(m,q_\circ]$. Then for any $\varepsilon>0$ there exists $\delta>0$ such that for any subset E of $\mathbb{R}^n$ we have $\Phi(Z_v\cap E)<\varepsilon$ provided $\mathcal{L}^n(E)<\delta$. In particular, $\Phi(Z_v\cap E)=0$ whenever $\mathcal{L}^n(E)=0$.*

Proof. Let $\mathcal{L}^n(E)<\delta$. Then we can find a family of nonoverlapping n-dimensional dyadic intervals Q_j such that $E\subset\bigcup_j Q_j$ and $\sum\limits_j \ell^n(Q_j)<\delta$. Of course, for sufficiently small δ the estimates

$$\|\nabla^k v\|_{\mathrm{L}_1(Q_j)}<1, \qquad \ell(Q_j)\le\delta^{\frac{1}{n}}$$

are fulfilled for every j. Denote

$$r_j=\ell(Q_j), \qquad \sigma_j=\|\nabla^k v\|_{\mathrm{L}_1(Q_j)}, \qquad \sigma=\|\nabla^k v\|_{\mathrm{L}_1}. \tag{3.3}$$

In view of Lemma 3.1 we have

$$\Phi(E)\le C\sum_j r_j^{n(m-q+1)}\,\sigma_j^{q-m}.$$

Since by our assumptions

$$0<m-q_\circ+1\le m-q+1<1,$$

we have

$$\begin{aligned}\sum_j r_j^{n(m-q+1)}\,\sigma_j^{q-m} &\overset{\text{Hölder ineq.}}{\le} C\Big(\sum_j r_j^n\Big)^{m-q+1}\cdot\Big(\sum_j \sigma_j\Big)^{q-m}\\ &\le C'\delta^{m-q+1}\cdot\sigma^{q-m}.\end{aligned}$$

The lemma is proved. □

By the classical approximation results (see, e.g., Chapter 3 in [44] or [11]), our mapping v coincides with a mapping $g\in \mathrm{C}^k(\mathbb{R}^n,\mathbb{R}^d)$ off an exceptional set of small n–dimensional Lebesgue measure. So we need to check the assertion of Theorem 1.3-(i) for C^k–smooth mappings now.

Lemma 3.3. *Let $q\in(m,q_\circ]$ and $g\in \mathrm{C}^k(\mathbb{R}^n,\mathbb{R}^d)$, $k>n$. Then*

$$\Phi_g(Z_{g,m})=0, \tag{3.4}$$

where Φ_g is calculated by the same formula (3.1) with g instead of v and $Z_{g,m}=\{x\in\mathbb{R}^n : \operatorname{rank}\nabla g(x)\le m\}$.

Proof. We can assume without loss of generality that g has compact support and that $|\nabla g(x)|\le 1$ for all $x\in\mathbb{R}^n$. We then clearly have that $g\in \mathrm{W}^k_1(\mathbb{R}^n,\mathbb{R}^d)$, hence we can in particular apply the above results to g. The following assertion plays the key role:
$(*)$ *For any n-dimensional interval $Q\subset\mathbb{R}^n$ the estimate*

$$\Phi(Z_{g,m}\cap Q)\le C\,\ell(Q)^{n(m-q+1)}\,\|\nabla^k\bar g_Q\|^{q-m}_{\mathrm{L}_1(Q)}$$

holds, where the constant C depends on n,m,k,d only, and we denoted

$$\nabla^k\bar g_Q(x)=\nabla^k g(x)-\fint_Q\nabla^k g(y)\,\mathrm{d}y.$$

The proof of $(*)$ is almost the same as that of Lemma 3.1, with evident modifications (we need to take the approximation polynomial $P_Q(x)$ of degree k instead of $k-1$, etc.).

By elementary facts of the Lebesgue integration theory, for an arbitrary family of nonoverlapping n-dimensional intervals Q_j one has

$$\sum_j\|\nabla^k\bar g_{Q_j}\|_{\mathrm{L}_1(Q_j)}\to 0\qquad\text{as}\quad\sup_j\ell(Q_j)\to 0 \tag{3.5}$$

The proof of this estimate is really elementary since now $\nabla^k g$ is a continuous and compactly supported function, and, consequently, is uniformly continuous and bounded.

From $(*)$ and (3.5), repeating the arguments of Corollary 3.2, using the assumptions on g and taking

$$\sigma_j=\|\nabla^k\bar g_{Q_j}\|_{\mathrm{L}_1(Q_j)},\qquad \sigma=\sum_j\sigma_j$$

in definitions (3.3), we obtain that $\Phi_g(Z_{g,m})<\varepsilon$ for any $\varepsilon>0$, hence the sought conclusion (3.4) follows. □

By the above–mentioned approximation results, the investigated mapping v equals a mapping $g\in \mathrm{C}^k(\mathbb{R}^n,\mathbb{R}^d)$ off an exceptional set of small n–dimensional Lebesgue measure. This fact together with Lemma 3.3 readily implies

Corollary 3.4 *(cp. with [15]). Let $q \in (m, q_\circ]$. Then there exists a set $\widetilde{Z}_v \subset Z_v$ of n-dimensional Lebesgue measure zero such that $\Phi(Z_v \setminus \widetilde{Z}_v) = 0$. In particular, $\Phi(Z_v) = \Phi(\widetilde{Z}_v)$.*

From Corollaries 3.2 and 3.4 we conclude that $\Phi(Z_v) = 0$, and this finishes the proof of Theorem 1.3-(i) for the required case $k > n$, $p = 1$.

Remark 3.5. As we could see from the above proofs, the assertion of Theorem 1.3-(i) is valid also under assumption $v \in \mathrm{BV}_k(\mathbb{R}^n, \mathbb{R}^d)$ (instead of W_1^k) with the same k. Here BV_k means the space of functions $v \in W_1^{k-1}$ such that its k-th (distributional) derivatives are Radon measures.

Remark 3.6. Thus, the assertions of Theorem 1.3-(i)-(ii) are proved. Of course, these assertions imply the fulfillment of the statement of Theorem 1.1 for the extreme borderline cases $\alpha = 0$ and $\alpha = 1$, because the corresponding Hölder spaces C^k and $C^{k,1}$ are contained in the Sobolev spaces W_p^k and W_p^{k+1} (with $p > n$) respectively.

3.2. Bridge F.-D. theorem for Hölder classes of mappings

This subsection is devoted to the proof of Theorem 1.1. It is sufficient to consider the general case $0 < \alpha < 1$ (see Remark 3.6).

Fix $m \in \{0, \dots, n-1\}$, $k \geq 1$, $d > m$, $0 < \alpha < 1$, and $v \in C^{k,\alpha}(\mathbb{R}^n, \mathbb{R}^d)$. Take also a parameter $q > m$. Of course, as in the previous subsection, it is sufficient to consider the case

$$q \in (m, q_\circ],$$

where $q_\circ = m + \frac{n-m}{k+\alpha}$. By definition of the space $C^{k,\alpha}$ and since the result has the local nature, we may assume without loss of generality that

$$|\nabla^k v(x) - \nabla^k v(y)| \leq |x-y|^\alpha \quad \text{for all } x, y \in \mathbb{R}^n, \tag{3.6}$$

$$|\nabla v(x)| \leq 1 \qquad \text{for all } x \in \mathbb{R}^n, \tag{3.7}$$

$$v(x) \equiv 0 \qquad \text{if } \ |x| > 1. \tag{3.8}$$

Denote again $Z_v = Z_{v,m} = \{x \in \mathbb{R}^n : \operatorname{rank} \nabla v(x) \leq m\}$. Now the parameter μ is different from the previous subsection:

$$\mu = n - m - (k+\alpha)(q-m). \tag{3.9}$$

As before, the assertion of the Bridge Dubovitskiĭ–Federer Theorem 1.1 is equivalent (by virtue of Theorem 2.13) to

$$\Phi(Z_v) = 0,$$

where we denoted

$$\Phi(E) = \inf_{E \subset \bigcup_j D_j} \sum_j (\operatorname{diam} D_j)^\mu \big[\operatorname{diam} v(D_j)\big]^q.$$

As indicated the infimum is taken over all countable families of compact sets $\{D_j\}_{j \in \mathbb{N}}$ such that $E \subset \bigcup_j D_j$.

The proof of Theorem 1.1 consists of several steps.

Step I. Applying Theorem 4.1 to the present case μ defined by (3.9), we obtain immediately the following assertion, which is the main tool for further arguments:

(∗∗) *Under above assumptions on v, for an arbitrary sufficiently small n-dimensional cube Q of size $r = \ell(Q)$ the estimate*

$$\Phi(Z_v \cap Q) \le C\, r^n \tag{3.10}$$

holds, where the constant C depends on n, m, k, d, α only.

Of course, the last estimate implies

$$\Phi(Z_v \cap F) \le \Psi(Z_v \cap F) \le C \cdot \mathcal{L}^n(F) \tag{3.11}$$

for any measurable set $F \subset \mathbb{R}^n$, where the countably subadditive set function Ψ is defined as

$$\Psi(F) = \lim_{\delta \to 0} \inf_{\substack{F \subset \bigcup_j D_j,\\ \operatorname{diam} D_j \le \delta}} \sum_j \bigl(\operatorname{diam} D_j\bigr)^{\mu} \bigl[\operatorname{diam} v(D_j)\bigr]^{q}. \tag{3.12}$$

Here the infimum is taken over all countable families of compact sets $\{D_j\}_{j \in \mathbb{N}}$ such that $F \subset \bigcup_j D_j$ and $\operatorname{diam} D_j \le \delta$ for all j.

Step II: the case $m = 0$.

Suppose now that $m = 0$. In other words, now $Z_v = \{x \in \mathbb{R}^n : \nabla v(x) = 0\}$. Thus $\nabla^k v(x) \equiv 0$ for almost all $x \in Z_v$. Then we have the decomposition:

$$Z_v = E_0 \cup E_1,$$

where $\mathcal{L}^n(E_0) = 0$, and every $x \in E_1$ is a density point for the set $\{x \in \mathbb{R}^n : \nabla^k v(x) = 0\}$. It implies, by elementary arguments, that

$$\lim_{y \to x} \frac{|\nabla^k v(y) - \nabla^k v(x)|}{|x - y|^{\alpha}} \to 0 \qquad \forall x \in E_1.$$

Then, checking the proof of the basic estimate (3.10) (see also (4.11)), we see that for any point $x \in E_1$ the identity

$$\lim_{r \to 0} \frac{\Phi(Z_v \cap Q(x, r))}{r^n} = 0 \tag{3.13}$$

holds, where $Q(x, r)$ denotes the cube centered at x with $\ell(Q) = r$. By usual elementary facts of real analysis and by subadditivity of $\Phi(\cdot)$, the convergence (3.13) implies that $\Phi(E_1) = 0$. The equality $\Phi(E_0) = 0$ follows from the condition $\mathcal{L}^n(E_0) = 0$ and (3.11). So $\Phi(Z_v) = 0$ as required. The case $m = 0$ is finished completely.

Step III: the case $m > 0$.

From this point, for all the steps below we assume that $m \ge 1$. By definitions, we have

$$Z_v = E_0 \cup E_1, \tag{3.14}$$

where

$$E_0 = \{x \in \mathbb{R}^n : \operatorname{rank} \nabla v(x) < m\}, \qquad E_1 := \{x \in \mathbb{R}^n : \operatorname{rank} \nabla v(x) = m\}.$$

By construction, $E_0 \subset Z_{v,m-1}$, so we can apply the previous estimate (3.11) for $m' = m-1$ instead of m to obtain

$$\widetilde{\Psi}(E_0) < \infty, \tag{3.15}$$

where $\widetilde{\Psi}$ is defined as Ψ (see (3.12)) with μ replaced by

$$\tilde{\mu} = n - m' - (k+\alpha)(q-m') = \mu - (k+\alpha-1) < \mu.$$

Of course, the inequalities $\widetilde{\Psi}(E_0) < \infty$ and $\tilde{\mu} < \mu$ imply $\widetilde{\Psi}(E_0) \geq \Psi(E_0) = 0$, and, consequently,

$$\Phi(E_0) = 0. \tag{3.16}$$

Step IV. Now we have to estimate the last term $\Phi(E_1)$ with $\operatorname{rank} \nabla v|_{E_1} \equiv m$. By Implicit Function Theorem and by the local nature of the considered results, we can assume without loss of generality that

$$v(x) = v(y,z) = (y, f(y,z)) \qquad \forall x \in E_1, \tag{3.17}$$

where for $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ we denote $y = (x_1, \dots, x_m)$, $z = (x_{m+1}, \dots, x_n)$, in particular, $x = (y,z)$, and $f : \mathbb{R}^{n-m} \to \mathbb{R}^d$ is some $C^{k,\alpha}$-Hölder mapping.

Now the condition $\operatorname{rank} \nabla v|_{E_1} \equiv m$ can be rewritten in the following equivalent form:

$$\nabla_z v(y,z) \equiv 0 \qquad \forall (y,z) \in E_1, \tag{3.18}$$

where we denote $\nabla_z v = \frac{\partial v}{\partial z} = \big(\frac{\partial v}{\partial x_{m+1}}, \dots, \frac{\partial v}{\partial x_n}\big)$.

From formula (3.18) using Fubini's Theorem and standard facts about density points of Lebesgue measurable sets it is easy to deduce that there exists a decomposition $E_1 = E_* \cup E_{**}$ such that

$$\mathcal{L}^n(E_*) = 0, \tag{3.19}$$

$$\nabla_z^l v(y,z) \equiv 0 \qquad \forall (y,z) \in E_{**} \quad \forall l = 1, \dots, k; \tag{3.20}$$

$$\lim_{z \to z_0} \frac{|v(y_0,z) - v(y_0,z_0)|}{|z-z_0|^{k+\alpha}} = 0 \qquad \forall (y_0,z_0) \in E_{**}. \tag{3.21}$$

Then from Egoroff's Theorem on uniform convergence of measurable functions, for any $\delta > 0$ there exists a decomposition

$$E_{**} = F \cup E \tag{3.22}$$

such that

$$\mathcal{L}^n(F) < \delta, \tag{3.23}$$

and the **uniform** convergence

$$\frac{|v(y,z+h) - v(y,z)|}{|h|^{k+\alpha}} \underset{h \to 0}{\to} 0 \qquad \forall (y,z) \in E \tag{3.24}$$

holds. Here $0 \neq h \in \mathbb{R}^{n-m}$, and convergence is uniform with respect to h, i.e., for any $\xi > 0$ there exists $\rho_\xi > 0$ such that for every $(y,z) \in E$ and for each $h \in \mathbb{R}^{n-m}$ satisfying $0 < |h| < \rho_\xi$ the inequality

$$|v(y,z+h)-v(y,z)|<\xi\,|h|^{k+\alpha} \tag{3.25}$$

holds.

Step V. We claim that for any $x_0=(y_0,z_0)\in E$ the convergence

$$\lim_{r\to 0}\frac{\Phi(E\cap Q(x_0,r))}{r^n}=0 \tag{3.26}$$

holds, where we denote by $Q(x,r)$ the n-dimensional cube centered at x with side length r.

This claim is proved by direct elementary calculations. Indeed, fix arbitrary $\xi\in(0,1)$ and take $\rho_\xi<1$ from the previous Step IV. Let $0<r<\rho_\xi$. The cube $Q=Q(x_0,r)$ can be represented as $Q=Q_y\times Q_z$, where Q_y and Q_z are the corresponding cubes of the same size in spaces $\mathbb{R}^m$ and $\mathbb{R}^{n-m}$ respectively.

Denote $\varepsilon=\xi\,r^{k+\alpha-1}$. Assume without loss of generality that $\frac{1}{\varepsilon}$ is an integer number, put

$$N=\varepsilon^{-m}, \tag{3.27}$$

and consider the decomposition

$$Q_y=\bigcup_{j=1}^{N}Q_j,$$

where every Q_j is the corresponding cube of size εr in the "y-space" $\mathbb{R}^m$. Denote further

$$J=\{j=1,\dots,N: E\cap(Q_j\times Q_z)\neq\emptyset\}.$$

By construction, every rectangle $Q_j\times Q_z$, $j\in J$, has the size εr and 1-Lipschitz condition (see (3.7)) along y-coordinates, and respectively size r and the strong Lipschitz conditions (3.25) along z-coordinates, that imply

$$\operatorname{diam} v(Q_j\times Q_z)\le 2\big(\varepsilon r+\xi r^{k+\alpha}\big)=4\varepsilon r=4\xi r^{k+\alpha}. \tag{3.28}$$

Finally we have

$$n^{-\frac{\mu}{2}}\Phi(E\cap Q)\le\sum_{j\in J}r^{\mu}\big[\operatorname{diam} v(Q_j\times Q_z)\big]^q\le(4\xi)^q N r^{(k+\alpha)q+\mu}\overset{(3.27),(3.9)}{=}4^q\xi^{q-m}r^n. \tag{3.29}$$

Since ξ could be taken arbitrary small, the proof of the Claim (3.26) is complete. Of course, this claim and the countable subadditivity of Φ imply immediately

$$\Phi(E)=0. \tag{3.30}$$

Let us summarize what was done on the previous steps III–V. For arbitrary $\delta>0$ we obtain the decomposition

$$Z_v=E_0\cup E_*\cup F\cup E,$$

where $\Phi(E_0)=\Phi(E)=0$, $\mathcal{L}(E_*)=0$, and $\mathcal{L}(F)<\delta$. By virtue of estimate (3.11) this implies

$$\Phi(Z_v)<C\,\delta,$$

consequently,

$$\Phi(Z_v) = 0.$$

Thus Theorem 1.1 is proved completely.

3.3. Bridge F.-D. theorem for mappings of potential spaces $\mathscr{L}_p^{k+\alpha}(\mathbb{R}^n, \mathbb{R}^d)$: case $k \le n$

This subsection is devoted to the proof of Theorem 1.3, case (iii). Here we consider the situation when $1 \le k \le n$ and $0 < \alpha < 1$. Then $k + \alpha - 1 < n$. Since our results have a local nature (see Subsection 2.8 for more precise explanations), it is sufficient to consider only the case

$$n - p < (k + \alpha - 1)p < n \tag{3.31}$$

i.e., when v is a continuous function by Sobolev Imbedding Theorems, but the gradient ∇v could be discontinuous.

Fix $m \in \{0, \dots, n-1\}$, $d > m$, and $v \in \mathscr{L}_p^{k+\alpha}(\mathbb{R}^n, \mathbb{R}^d)$. In other words,

$$v = \mathscr{G}_{k+\alpha}(g) := G_{k+\alpha} * g$$

fore some $g \in L_p(\mathbb{R}^n)$.

Take a parameter $q > m$. Of course, as in previous subsection, it is sufficient to consider the case

$$q \in (m, q_\circ],$$

where $q_\circ = m + \frac{n-m}{k+\alpha}$. In particular, we have

$$q - m \le q_\circ - m = \frac{n - m}{k + \alpha} < p. \tag{3.32}$$

Now the parameter μ is the same as in the previous subsection:

$$\mu = n - m - (k + \alpha)(q - m). \tag{3.33}$$

Denote again $Z_{v,m} = \{x \in \mathbb{R}^n \setminus A_v : \operatorname{rank} \nabla v(x) \le m\}$. Here A_v means the set of 'bad' points, where v is not differentiable or $\mathcal{M}\nabla v = \infty$. Recall that this set has a small size, namely,

$$\mathcal{H}^\tau(A_v) = 0 \qquad \forall \tau > \tau_* = n - (k + \alpha - 1)p.$$

As before, the assertion of the Bridge Dubovitskiĭ–Federer Theorem 1.3 is equivalent (by virtue of Theorem 2.13) to

$$\Phi(Z_{v,m}) = 0,$$

where

$$\Phi(E) := \inf_{E \subset \bigcup_j D_j} \sum_j \big(\operatorname{diam} D_j\big)^\mu \big[\operatorname{diam} v(D_j)\big]^q.$$

Since Φ is a subadditive set-function, it is sufficient to check only the equality

$$\Phi(Z'_v) = 0,$$

where we denote

$$Z'_v := \{x \in Z_{v,m} : |\nabla v(x)| \le 1\}.$$

From the inequality $q - m < p$ (see (3.32)) we conclude that

$$q + \mu \overset{(3.33)}{=} n - (k + \alpha - 1)(q - m) > n - (k + \alpha - 1)p = \tau_*, \tag{3.34}$$

so the key assumption (4.20) of Theorem 4.3 is fulfilled now.

Step I: estimates on a single cube. Applying the just mentioned Theorem 4.3 to the present case $\mu = \mu_q = n - m - (k + \alpha)(q - m)$, we obtain immediately the following assertion, which is the main tool for further arguments.

Lemma 3.7. *Under above assumptions on v, for an arbitrary n-dimensional cube Q of size $r = \ell(Q)$ the estimate*

$$\Phi(Z'_v \cap Q) \le C \left(\sigma^q r^{(k+\alpha-\frac{n}{p})q+\mu} \; + \; \sigma^{q-m} r^{n(1-\frac{q-m}{p})} \right), \tag{3.35}$$

holds, where

$$\sigma = \|\mathcal{M}g\|_{L_p(Q)}, \tag{3.36}$$

and the constant C depends on n, m, k, α, d, p only.

Step II: estimates on sets of small n-Lebesgue measure.

Lemma 3.8. *Under above assumptions on v, for any $\varepsilon > 0$ there exists $\delta > 0$ such that for any subset E of $\mathbb{R}^n$ we have $\Phi(Z'_v \cap E) \le \varepsilon$ provided $\mathcal{L}^n(E) \le \delta$. In particular, $\Phi(Z'_v \cap E) = 0$ whenever $\mathcal{L}^n(E) = 0$.*

Proof. Let $\mathcal{L}^n(E) < \delta$, then we can find a family of nonoverlapping n-dimensional dyadic intervals Q_β such that $E \subset \bigcup_\beta Q_\beta$ and $\sum\limits_\beta \ell^n(Q_\beta) < \delta$. Of course, for sufficiently small δ the estimates

$$\|\nabla^k v\|_{\mathrm{L}_p(Q_\beta)} < 1, \qquad \ell(Q_\beta) \le \delta^{\frac{1}{n}} < 1 \tag{3.37}$$

are fulfilled for every β. Denote

$$r_\beta = \ell(Q_\beta), \qquad \sigma_\beta = \|\nabla^k v\|_{\mathrm{L}_p(Q_\beta)}, \qquad \sigma = \|\nabla^k v\|_{\mathrm{L}_p(\cup_\beta Q_\beta)}. \tag{3.38}$$

In view of Lemma 3.7 we have

$$\Phi(E) \le C \sum_\beta \sigma_\beta^{q-m} r_\beta^{n(1-\frac{q-m}{p})} + C \sum_\beta \sigma_\beta^q r_\beta^{(k+\alpha-\frac{n}{p})q+\mu}. \tag{3.39}$$

Now let us estimate the first sum. Since by our assumptions $q - m < p$ (see (3.32)), we have

$$\begin{aligned}\sum_\beta \sigma_\beta^{q-m} r_\beta^{n(1-\frac{q-m}{p})} &\overset{\text{Hölder ineq.}}{\le} C\Big(\sum_\beta \sigma_\beta^p\Big)^{\frac{q-m}{p}} \cdot \Big(\sum_\beta r_\beta^n\Big)^{\frac{p-q+m}{p}} \\ &\le C\sigma^{q-m} \cdot \Big(\mathcal{L}^n(E)\Big)^{\frac{p-q+m}{p}}. \end{aligned} \tag{3.40}$$

The estimates of the second sum are again handled by consideration of two separate cases.

Case I. $q \ge p$. Then

$$\sum_\beta \sigma_\beta^q r_\beta^{(k+\alpha-\frac{n}{p})q+\mu} \overset{(3.37)}{\le} \sum_\beta \sigma_\beta^p \le \sigma^p. \tag{3.41}$$

Case II. $q < p$. From the definition (3.33) we have the following elementary identity:

$$\big(k+\alpha-\frac{n}{p}\big)q+\mu = n(1-\frac{q}{p}) + m(k+\alpha-1) > n\big(1-\frac{q}{p}\big). \tag{3.42}$$

Then

$$\begin{aligned}\sum_\beta \sigma_\beta^q r_\beta^{(k+\alpha-\frac{n}{p})q+\mu} \le \sum_\beta \sigma_\beta^q r_\beta^{n(1-\frac{q}{p})} &\overset{\text{Hölder ineq.}}{\le} \Big(\sum_\beta \sigma_\beta^p\Big)^{\frac{q}{p}} \cdot \Big(\sum_\beta r_\beta^n\Big)^{\frac{p-q}{p}} \\ &\le \sigma^q \delta^{\frac{p-q}{p}}. \end{aligned} \tag{3.43}$$

Now for both cases (I) and (II) we have by (3.39)–(3.43) that $\Phi(E) \le \omega(\delta)$, where the function $\omega(\delta)$ satisfies the condition $\omega(\delta) \searrow 0$ as $\delta \searrow 0$. The lemma is proved. $\square$

Step III (finishing of the proof of Theorem 1.3 (iii) for the case $k \le n$).

From Theorem 2.14 it follows that for any $\varepsilon > 0$ there exists a decomposition $\mathbb{R}^n = U \cup E$, $E = \mathbb{R}^n \setminus U$, where $\operatorname{meas}(U) < \varepsilon$ and the identities $v = h$ and $\nabla v = \nabla h$ hold on the set E, where the mapping h belongs to the class $C^{k,\alpha}(\mathbb{R}^n, \mathbb{R}^d)$. By Theorem 1.1, proved in the previous subsection, we have $\Phi(Z'_v \cap E) = \Phi_h(Z'_h \cap E) = 0$. On the other hand, by Lemma 3.8 the value $\Phi(Z'_v \cap U)$ could be made arbitrary small. Therefore, $\Phi(Z'_v) = 0$ as required.

3.4. Bridge F.-D. theorem for mappings of potential spaces $\mathscr{L}_p^{k+\alpha}(\mathbb{R}^n, \mathbb{R}^d)$: case $k > n$

The proof for this case can be done almost word by word the same way as in the previous subsection with the following evident difference: on the Step 1 for the estimates on a single cube we need to use Theorem 4.2 instead of Theorem 4.3. Thus the proof now is even much easier since on the right hand side of the estimate of Theorem 4.2 we have only one term (instead of two terms in the estimate of Theorem 4.3).

3.5. Bridge F.-D. theorem for mappings of Sobolev–Lorentz potential spaces $\mathscr{L}_{p,1}^{k+\alpha}(\mathbb{R}^n, \mathbb{R}^d)$ with $(k+\alpha)p = n$

The proof for this case could be done almost word by word the same way as in previous subsection 3.3 with the following evident difference: on the Step 1 for the estimates on a single cube we need to use Theorem 4.4 (ii) instead of Theorem 4.3. Also we need to use the following well-known supadditive property for Lorentz norms:

Suppose that $1 \le p < \infty$ and $E = \bigcup_{j\in\mathbb{N}} E_j$, where E_j are measurable and mutually disjoint subsets of $\mathbb{R}^n$. Then

$$\sum_j \|f\|^p_{\mathrm{L}_{p,1}(E_j)} \le \|f\|^p_{\mathrm{L}_{p,1}(E)},$$

where by definition $\|f\|_{L_{p,1}(E)} := \|f \cdot 1_E\|_{L_{p,1}}$, *and* 1_E *denotes the indicator function of the set* E. (See, e.g., [37] or [31].)

Summarizing the results, obtained in above subsections 3.1, 3.3 – 3.5, we conclude, that the proof of Theorem 1.3 is finished completely.

4. Appendix: estimates of the critical values on cubes

Let $\mu \ge 0$, $q > 0$, and $v : \mathbb{R}^n \to \mathbb{R}^d$ be a continuous function. For a set $E \subset \mathbb{R}^n$ as before define the set function

$$\Phi(E) = \inf_{E \subset \bigcup_j D_j} \sum_j \bigl(\operatorname{diam} D_j\bigr)^\mu \bigl[\operatorname{diam} v(D_j)\bigr]^q, \tag{4.1}$$

where the infimum is taken over all countable families of compact sets $\{D_j\}_{j\in\mathbb{N}}$ such that $E \subset \bigcup_j D_j$. Then $\Phi(\cdot)$ is a countably subadditive and the implication

$$\Phi(E) = 0 \;\Rightarrow\; \Bigl[\mathcal{H}^\mu\bigl(E \cap v^{-1}(y)\bigr) = 0 \quad \text{for } \mathcal{H}^q\text{-almost all } y \in \mathbb{R}^d\Bigr] \tag{4.2}$$

holds (see Theorem 2.13).

Our purpose here is to estimate Φ for subsets of critical set in cubes for different classes of mappings.

For all the following four subsections fix $m \in \{0, \dots, n-1\}$ and $d \ge m$. Take also a positive parameter $q \ge m$ and nonnegative $\mu \ge 0$ required in the definition of the set–function Φ.

For a regular (in a sense) mapping $v : \mathbb{R}^n \to \mathbb{R}^d$ denote as before

$$Z_{v,m} = \{x \in \mathbb{R}^n \setminus A_v : \operatorname{rank} \nabla v(x) \le m\}.$$

Here A_v means the set of ‘bad’ points, where v is not differentiable or which are not Lebesgue points for ∇v (of course, $A_v = \emptyset$ if the gradient ∇v is a continuous function). It is convenient (and sufficient for our purposes) to restrict our attention on the following subset of critical points

$$Z'_v = \{x \in Z_{v,m} : |\nabla v(x)| \le 1\}. \tag{4.3}$$

4.1. Estimates on cubes for Hölder classes of mappings

Fix $k \ge 1$, $0 \le \alpha \le 1$, and $v \in C^{k,\alpha}(\mathbb{R}^n, \mathbb{R}^d)$. By definition of the space $C^{k,\alpha}$ and since the result has the local nature, we may assume without loss of generality that

$$|\nabla^k v(x) - \nabla^k v(y)| \le |x-y|^\alpha \quad \text{for all } x, y \in \mathbb{R}^n, \tag{4.4}$$

$$|\nabla v(x)| \le 1 \qquad \text{for all } x \in \mathbb{R}^n. \tag{4.5}$$

The main result of this subsection is contained in the following

Theorem 4.1. *Under above assumptions, for any sufficiently small n-dimensional interval $Q \subset \mathbb{R}^n$ the estimate*

$$\Phi(Q \cap Z_{v,m}) \le C\,\ell(Q)^{q+\mu+(k+\alpha-1)(q-m)} \tag{4.6}$$

holds, where the constant C depends on n, m, k, α, d only.

Proof. Let the assumptions in the beginning of this subsection be fulfilled. Fix an n-dimensional interval $Q \subset \mathbb{R}^n$ of size $\ell(Q) < \frac{1}{\sqrt{n}}$. Without loss of generality we may assume that the origin $0 \in Q$. Take the polynomial P of degree k such that

$$\nabla^j v(0) = \nabla^j P(0) \qquad \forall j = 0, 1, \dots, k. \tag{4.7}$$

Denote $v_Q = v - P$. Then from assumption (4.4) we have

$$|\nabla^k v_Q(x)| \le r^\alpha \qquad \text{for all } x \in Q. \tag{4.8}$$

$$|\nabla v_Q(x)| \le r^{k+\alpha-1} \qquad \text{for all } x \in Q, \tag{4.9}$$

where we denote for convenience $r = \sqrt{n}\,\ell(Q)$. Put $\varepsilon = r^{k+\alpha-1}$. Since by our choice $\ell(Q) < \frac{1}{\sqrt{n}}$, we have in particular that

$$\varepsilon < 1. \tag{4.10}$$

Denote $Z_v = Q \cap Z_{v,m}$. Since $\nabla P_Q(x) = \nabla v(x) - \nabla v_Q(x)$, $|\nabla v_Q(x)| \le \varepsilon$, $|\nabla v(x)| \le 1$, and $\lambda_{m+1}(v, x) = 0$ for $x \in Z_v$, we have[5]

$$Z_v \subset \big\{x \in Q : \lambda_1(P_Q, x) \le 1 + \varepsilon, \dots, \lambda_m(P_Q, x) \le 1 + \varepsilon,\ \lambda_{m+1}(P_Q, x) \le \varepsilon\big\}.$$

Applying Theorem 2.7 to polynomial P and using the Lipschitz condition $\operatorname{diam} v_Q(Q) \le \varepsilon\, r$, we find a finite family of balls $T_j \subset \mathbb{R}^d$, $j = 1, \dots, N$ with $N \le C_Y(1 + \varepsilon^{-m})$, each of radius $2\varepsilon r$, such that

$$\bigcup_{j=1}^{N} T_j \supset v(Z_v).$$

Therefore, we have

$$\Phi(Z_v) \le C\, N \varepsilon^q r^{q+\mu} \le C_1 (1 + \varepsilon^{-m})\, \varepsilon^q\, r^{q+\mu} \overset{(4.10)}{\le} C\, \varepsilon^{q-m}\, r^{q+\mu}.$$

The last formula, because of definition of ε, implies the required estimate (4.6). The Theorem is proved. □

The analysis of this simple proof shows, that if we replace the condition (4.4) by more general assumption

$$|\nabla^k v(x) - \nabla^k v(y)| \le A\, \ell(Q)^\alpha \quad \text{for all } x, y \in Q,$$

then instead of (4.6) the modified estimate

$$\Phi(Q \cap Z_{v,m}) \le C\, A^{q-m}\, \ell(Q)^{q+\mu+(k+\alpha-1)(q-m)} \tag{4.11}$$

holds, where again the constant C depends on n, m, k, α, d only.

---

[5] Here we use the following elementary fact: for any linear maps $L_1 \colon \mathbb{R}^n \to \mathbb{R}^d$ and $L_2 \colon \mathbb{R}^n \to \mathbb{R}^d$ the estimates $\lambda_l(L_1 + L_2) \le \lambda_l(L_1) + \|L_2\|$ hold for all $l = 1, \dots, d$, see, e.g., [43, Proposition 2.5 (ii)].

4.2. Estimates on cubes for Sobolev classes of mappings. Case I: $(k+\alpha-1)p>n$

Fix $k\ge 1$, $0\le\alpha<1$, $1<p<\infty$, and $v\in\mathscr{L}_p^{k+\alpha}(\mathbb{R}^n,\mathbb{R}^d)$. In this subsection we consider the case, when

$$(k+\alpha-1)p>n, \tag{4.12}$$

i.e., when the gradient ∇v is a continuous and uniformly bounded function by Sobolev Imbedding Theorems. For convenience, in this section we assume that

$$\sup_{x\in\mathbb{R}^n}|\nabla v(x)|\le 1.$$

Theorem 4.2. *Under above assumptions, there exists a function $h\in L_p(\mathbb{R}^n)$ (depending on v only) such that for any sufficiently small n-dimensional interval $Q\subset\mathbb{R}^n$ the estimate*

$$\Phi(Q\cap Z_{v,m})\le C\,\sigma^{q-m}\ell(Q)^{q+\mu+(k+\alpha-1-\frac{n}{p})(q-m)} \tag{4.13}$$

holds, where

$$\sigma=\|h\|_{L_p(Q)}, \tag{4.14}$$

and the constant C depends on n,m,k,α,d,p only.

Proof. The proof here is very similar to the proof of Theorem 4.1 from the previous subsection, but in the beginning we have to obtain the uniform estimates for the gradient of the difference between the function and some polynomial.

Let the assumptions in the beginning of this subsection be fulfilled. Put $l=\min\{i\in\mathbb{Z}_+: k+\alpha-n-2<i\}$. In particular, $l\ge 0$. Denote $\tilde{k}=k-l$. Then by construction

$$\tilde{k}+\alpha-n-2<0. \tag{4.15}$$

We claim also, that

$$(\tilde{k}+\alpha-1)p>n. \tag{4.16}$$

Indeed, if $k+\alpha-n-2<0$, then $l=0$, $\tilde{k}=k$, and the inequality (4.16) follows immediately from the assumption (4.12). On the other hand, if $k+\alpha-n-2\ge 0$, then $l>0$ and by construction we have $l-1\le k+\alpha-n-2$, that is equivalent $\tilde{k}+\alpha-n-1\ge 0$, and the inequality (4.16) follows from the assumption $p>1$.

Put $u=\nabla^l v$. From the inclusion $u\in\mathscr{L}_p^{\tilde{k}+\alpha}(\mathbb{R}^n,\mathbb{R}^d)$ it follows that $u=\mathscr{G}_{\tilde{k}+\alpha}(g):=G_{\tilde{k}+\alpha}*g$ for some $g\in L_p(\mathbb{R}^n)$. We will use the Hardy–Littlewood maximal function $\mathcal{M}g$ (see (2.2)). The well-known properties of maximal functions (see, e.g., [39]) imply

$$h:=\mathcal{M}g\in L_p(\mathbb{R}^n).$$

Fix an n-dimensional interval $Q\subset\mathbb{R}^n$ of size $r=\ell(Q)\le 1$. From Remark 2.6 (see formula (2.14)), applied to the function $u=G_{\tilde{k}+\alpha}*g$, and from assumption (4.16), it follows that there exists a polynomial $\tilde{P}$ of degree $\tilde{k}$ such that

$$\sup_{x\in Q}|\nabla u_Q(x)|\le C\,\|\mathcal{M}g\|_{L_p(Q)}\,r^{\tilde{k}+\alpha-1-\frac{n}{p}}, \tag{4.17}$$

where $u_Q = u - \tilde{P}$. Since $u = \nabla^l v$, we have that for some polynomial P of degree k the estimate

$$\sup_{x\in Q} |\nabla v_Q(x)| \le r^l \sup_{x\in Q} |\nabla u_Q(x)| \le C\, \|\mathcal{M}g\|_{L_p(Q)}\, r^{k+\alpha-1-\frac{n}{p}}. \tag{4.18}$$

Denote the right hand side of the last inequality by ε. Of course, taking $r = \ell(Q)$ sufficiently small, we could assume without loss of generality that $\varepsilon < 1$.

From this step we could repeat almost word by word the last part of the proof of Theorem 4.1 to obtain the required estimate (4.13). □

4.3. Estimates on cubes for Sobolev classes of mappings. Case II: $n - p < (k+\alpha-1)p < n$

Fix $k \ge 1$, $0 \le \alpha < 1$, $1 < p < \infty$, and $v \in \mathscr{L}^{k+\alpha}_p(\mathbb{R}^n, \mathbb{R}^d)$. In this subsection we consider the case, when $k + \alpha > 1$ and

$$n - p < (k+\alpha-1)p < n, \tag{4.19}$$

i.e., when v a continuous function by Sobolev Imbedding Theorems, but the gradient ∇v could be discontinuous.

From the inclusion $v \in \mathscr{L}^{k+\alpha}_p(\mathbb{R}^n, \mathbb{R}^d)$ it follows that $\nabla^k v = \mathscr{G}_\alpha(g) := G_\alpha * g$ for some $g \in L_p(\mathbb{R}^n)$. We will use the Hardy–Littlewood maximal function $\mathcal{M}g$. The well-known properties of maximal functions (see, e.g., [39]) imply

$$\mathcal{M}g \in L_p(\mathbb{R}^n).$$

Theorem 4.3. *Under above assumptions, if an addition*

$$q + \mu > \tau_* := n - (k+\alpha-1)p, \tag{4.20}$$

then for any n-dimensional interval $Q \subset \mathbb{R}^n$ of size $r = \ell(Q)$ the estimate

$$\Phi(Z'_v \cap Q) \le C\left(\sigma^q r^{(k+\alpha-\frac{n}{p})q+\mu} \; + \; \sigma^{q-m} r^{q+\mu+(k+\alpha-1-\frac{n}{p})(q-m)} \right), \tag{4.21}$$

holds, where

$$\sigma = \|\mathcal{M}g\|_{L_p(Q)}, \tag{4.22}$$

and the constant C depends on n, m, k, α, d, p only.

Proof. Let the assumptions in the beginning of this subsection be fulfilled. Fix an n-dimensional interval $Q \subset \mathbb{R}^n$ of size $r = \ell(Q)$. Take the polynomial P of degree k from Lemma 2.5 such that

$$|\nabla v_Q(x)| \le C \int_Q \frac{\mathcal{M}g(y)}{|x-y|^{n-k-\alpha+1}}\, dy \qquad \forall x \in Q, \tag{4.23}$$

where $v_Q = v - P$.

Put

$$\varepsilon_* = \|\mathcal{M}g\|_{L_p(Q)}\, r^{k+\alpha-1-\frac{n}{p}}. \tag{4.24}$$

We emphasise that now we can not assume that ε_* is small since the exponent $k+\alpha-1-\frac{n}{p}$ is *strictly less* than zero.

Denote

$$\sigma = \|\mathcal{M}g\|_{\mathrm{L}_p(Q)},$$

and for each $j \in \mathbb{Z}$ define

$$E_j = \bigl\{x \in Q : \mathcal{M}_Q|\nabla v_Q|(x) \in (2^{j-1}, 2^j]\bigr\} \quad \text{and} \quad \delta_j = \mathcal{H}^{q+\mu}_\infty(E_j),$$

where $\mathcal{M}_Q$ is the modified Hardy–Littlewood maximal function defined by formula (2.18). Put

$$s = \frac{(q+\mu)p}{n-(k+\alpha-1)p}, \qquad \tau = \frac{s}{p}\bigl(n-(k+\alpha-1)p\bigr) = q+\mu.$$

From these definitions it follows, in particular, that

$$\left(\frac{\sigma}{\varepsilon_*}\right)^s = r^{q+\mu}, \tag{4.25}$$

further, since $n < (k+\alpha)p$, we have

$$s > q+\mu, \tag{4.26}$$

moreover, since by Theorem assumption $q+\mu > n-(k+\alpha-1)p$, we have

$$s > p. \tag{4.27}$$

Then by Theorem 2.8 (applied for the case $\beta = (k+\alpha-1)$),

$$\sum_{j=-\infty}^{\infty} \delta_j 2^{js} \le C\sigma^s \tag{4.28}$$

for a constant C depending on n, d, β, τ, s only. By the definition of the Hausdorff measure, for each $j \in \mathbb{Z}$ there exists a family of balls $B_{ij} \subset \mathbb{R}^n$ of radii r_{ij} such that

$$E_j \subset \bigcup_{i=1}^{\infty} B_{ij} \quad \text{and} \quad \sum_{i=1}^{\infty} r_{ij}^{q+\mu} \le c\,\delta_j. \tag{4.29}$$

Denote

$$Z_j = Z'_v \cap E_j \quad \text{and} \quad Z_{ij} = Z_j \cap B_{ij}.$$

By construction $Z'_v \cap Q = \bigcup_j Z_j$ and $Z_j = \bigcup_i Z_{ij}$.

Take an integer value j_* such that $\varepsilon_* \in (2^{j_*-1}, 2^{j_*}]$. Denote $Z_* = \bigcup_{j<j_*} Z_j$, $Z_{**} = \bigcup_{j\ge j_*} Z_j$. Then by construction

$$Z'_v \cap Q = Z_* \cup Z_{**}, \quad Z_* \subset \{x \in Z'_v \cap Q : \mathcal{M}_Q|\nabla v_Q|)(x) < \varepsilon_*\}.$$

Further, since $\nabla P_Q(x) = \nabla v(x) - \nabla v_Q(x)$, $|\nabla v_Q(x)| \le 2^j$, $|\nabla v(x)| \le 1$, and $\lambda_m(v,x) = 0$ for $x \in Z_{ij}$, we have[6]

$$Z_{ij} \subset \left\{x \in B_{ij} : \lambda_1(P_Q, x) \le 1 + 2^j, \dots, \lambda_m(P_Q, x) \le 1 + 2^j,\ \lambda_{m+1}(P_Q, x) \le 2^j\right\}.$$

From definition of E_j and from Lemma 2.12 (applying to the function v_Q) we have

$$|v_Q(x) - v_Q(y)| \le C_M 2^j r_{ij} \qquad \forall x, y \in Z_{ij}.$$

From this fact, applying Theorem 2.7 to polynomial P_Q with $B = B_{ij}$ and $\varepsilon = \varepsilon_j = 2^j$, we find a finite family of balls $T_\nu \subset \mathbb{R}^d$, $\nu = 1, \dots, \nu_j$ with $\nu_j \le C_Y(1 + \varepsilon_j^{-m})$, each of radius $(1 + C_M)\varepsilon_j r_{ij}$, such that

$$\bigcup_{\nu=1}^{\nu_j} T_\nu \supset v(Z_{ij}).$$

Therefore, for every $j \ge j_*$ we have

$$\Phi(Z_{ij}) \le C_1 \nu_j \varepsilon_j^q r_{ij}^{q+\mu} = C_2(1 + \varepsilon_j^{-m}) 2^{jq} r_{ij}^{q+\mu} \le C_2(1 + \varepsilon_*^{-m}) 2^{jq} r_{ij}^{q+\mu}, \tag{4.30}$$

where all the constants C_μ above depend on n, m, k, d only. By the same reasons, but this time applying Theorem 2.7 and Lemma 2.12 with $\varepsilon = \varepsilon_*$ and instead of the balls B_{ij} taking the big ball $B \supset Q$ with radius $\sqrt{n}r$, we have

$$\Phi(Z_*) \le C_3(1 + \varepsilon_*^{-m}) \varepsilon_*^q r^{q+\mu}. \tag{4.31}$$

The right hand side of the last formula, from the definition of $\varepsilon_* = \sigma\, r^{k+\alpha-1-\frac{n}{p}}$, is equivalent to the right hand side of the required estimate (4.21).

From (4.30) we get immediately

$$\begin{aligned}
\Phi(Z_{**}) &\le C_2(1 + \varepsilon_*^{-m}) \sum_{j \ge j_*} \sum_i 2^{jq} r_{ij}^{q+\mu} \overset{(4.29)}{\le} C_3(1 + \varepsilon_*^{-m}) \sum_{j \ge j_*} 2^{jq} \delta_j \\
&\overset{(4.26)}{\le} C_3(1 + \varepsilon_*^{-m}) 2^{j_*(q-s)} \sum_{j \ge j_*} 2^{js} \delta_j \overset{(4.28)}{\le} C_4(1 + \varepsilon_*^{-m}) \varepsilon_*^{(q-s)} \sigma^s \\
&\overset{(4.25)}{\le} C_5(1 + \varepsilon_*^{-m}) \varepsilon_*^q r^{q+\mu}.
\end{aligned} \tag{4.32}$$

The right hand side of the last formula, by the same reasons (see the commentary after (4.31)) is equivalent to the right hand side of the required estimate (4.21).

Thus from (4.31)–(4.32) the required estimate (4.21) follows. The Lemma is proved. □

4.4. Estimates on cubes for Sobolev–Lorentz classes of mappings: the general case $(k+\alpha)p \ge n$

Fix $k \ge 1$, $0 \le \alpha < 1$, $1 < p < \infty$, and $v \in \mathscr{L}^{k+\alpha}_{p,1}(\mathbb{R}^n, \mathbb{R}^d)$. In this subsection we consider the case, when $k + \alpha > 1$ and

$$(k+\alpha)p \ge n, \tag{4.33}$$

---

[6] Here we use the following elementary fact: for any linear maps $L_1 \colon \mathbb{R}^n \to \mathbb{R}^d$ and $L_2 \colon \mathbb{R}^n \to \mathbb{R}^d$ the estimates $\lambda_l(L_2 + L_2) \le \lambda_l(L_1) + \|L_2\|$ hold for all $l = 1, \dots, d$, see, e.g., [43, Proposition 2.5 (ii)].

i.e., when v is a continuous function (see, e.g., [27]), but the gradient ∇v could be discontinuous in general (if $(k+\alpha-1)p<n$).

Theorem 4.4. *Under above assumptions, there exists a function $h \in L_{p,1}(\mathbb{R}^n)$ (depending on v) such that the following statements are fulfilled:*

(i) if $(k+\alpha-1)p \geq n$, then gradient ∇v is continuous and uniformly bounded function, and for any sufficiently small n-dimensional interval $Q \subset \mathbb{R}^n$ the estimate

$$\Phi(Z'_v \cap Q) \leq C\, \sigma^{q-m} r^{q+\mu+(k+\alpha-1-\frac{n}{p})(q-m)} \tag{4.34}$$

holds, where again

$$r=\ell(Q), \qquad \sigma=\|h\|_{L_{p,1}(Q)}, \tag{4.35}$$

and the constant C depends on n, m, k, α, d, p only.

(ii) if $n-p \leq (k+\alpha-1)p<n$, then under additional assumption

$$q+\mu \geq \tau_* := n-(k+\alpha-1)p \tag{4.36}$$

for any n-dimensional interval $Q \subset \mathbb{R}^n$ the estimate

$$\Phi(Z'_v \cap Q) \leq C\left(\sigma^q r^{(k+\alpha-\frac{n}{p})q+\mu}+\sigma^{q-m} r^{q+\mu+(k+\alpha-1-\frac{n}{p})(q-m)}\right) \tag{4.37}$$

holds with the same σ, r.

Remark 4.5. Formally estimates in Theorem 4.4 are the same as in Theorems 4.2–4.3, the only difference is in the definition of σ (using the Lorentz norm instead of Lebesgue one, cf. formulas (4.14) and (4.35)). However, Theorem 4.4 is 'stronger' in a sense than the previous Theorems 4.2–4.3. Namely, there are some important (limiting) cases, which are not covered by Theorem 4.3, but one could still apply the Theorem 4.4 for these cases. It happens for the following values of the parameters:

$$(k+\alpha)p=n, \tag{4.38}$$

or

$$(k+\alpha-1)p=n, \tag{4.39}$$

or

$$q+\mu=\tau_*. \tag{4.40}$$

It means that the Lorentz norm is a sharper and more accurate tool here than the Lebesgue norm.

Proof of Theorem 4.4. As in the previous subsection 4.2 put $l=\min\{i \in \mathbb{Z}_+ : k+\alpha-n-2<i\}$. In particular, $l \geq 0$. Denote $\tilde{k}=k-l$ and $u=\nabla^l v$. Then by construction

$$u \in \mathscr{L}_p^{\tilde{k}}(\mathbb{R}^n), \qquad \tilde{k}+\alpha-n-2<0, \tag{4.41}$$

moreover,

$$(\tilde{k}+\alpha-1)p \geq n \qquad \text{if} \quad (k+\alpha-1)p \geq n \tag{4.42}$$

(see the above discussion around (4.15)–(4.16)). From the inclusion $u \in \mathscr{L}^{\tilde{k}+\alpha}_{p,1}(\mathbb{R}^n,\mathbb{R}^d)$ it follows that $u = \mathscr{G}_{\tilde{k}+\alpha}(g) := G_{\tilde{k}+\alpha} * g$ for some $g \in L_p(\mathbb{R}^n)$. We will use the Hardy–Littlewood maximal function $\mathcal{M}g$. Recall, that by properties of Lorentz spaces, the standard estimate

$$\|\mathcal{M}g\|_{L_{p,1}} \leq C\,\|g\|_{L_{p,1}}$$

holds for the considered case $1 < p < \infty$ (see, e.g., [31, Theorem 4.4]). Put $h = \mathcal{M}g$.

The proof of Theorem 4.4 is very similar to that one of Theorems 4.2–4.3: the main differences concern the limiting cases (4.38)–(4.40) mentioned above.

- Case $(k+\alpha-1)p > n$. The required assertion in (i) follows immediately from Theorem 4.2 and from the well-known inequality

  $$\|f\|_{L_p(Q)} \leq \|f\|_{L_{p,1}(Q)}.$$

- Case $(k+\alpha-1)p = n$. Then by above notations $l = 0$ and $v = u = G_{k+\alpha} * g$. Fix an n-dimensional interval $Q \subset \mathbb{R}^n$ of size $r = \ell(Q)$. Take the polynomial P of degree k from Theorem 2.5 such that

  $$|\nabla v_Q(x)| \leq C \int\limits_Q \frac{\mathcal{M}g(y)}{|x-y|^{n-k-\alpha+1}}\,dy \qquad \forall x \in Q, \tag{4.43}$$

  where $v_Q = v - P$. Put $\sigma = \|\mathcal{M}g\|_{L_{p,1}(Q)}$.
  Put $\beta = (k+\alpha-1)$. From the generalized Hölder inequality for Lorentz norms

  $$\int\limits_Q \frac{\mathcal{M}g(y)}{|y-x|^{n-\beta}}\,dy \leq \|\mathcal{M}g\|_{L_{p,1}(Q)} \cdot \left\|\frac{1_Q}{|\cdot - x|^{n-\beta}}\right\|_{L_{\frac{p}{p-1},\infty}} = C\,\|g\|_{L_{p,1}}$$

  (see, e.g., [31, Theorem 3.7]) and from (4.43) it follows immediately that

  $$\sup_{x\in Q} |\nabla v_Q(x)| \leq C\,\|\mathcal{M}g\|_{L_{p,1}(Q)} = C\,\sigma. \tag{4.44}$$

  Denote the right hand side of the last inequality by ε. Of course, taking $r = \ell(Q)$ sufficiently small, we can assume without loss of generality that $\varepsilon < 1$. Thus to finish the proof for this case we could repeat almost word by word the last part of the proof of Theorem 4.1 to obtain the inequality

  $$\Phi(Z'_v \cap Q) \leq C\,\sigma^{q-m} r^{q+\mu}, \tag{4.45}$$

  which is equivalent to the required estimate (4.34) for the considered values of k, α, p.
- Case $n-p \leq (k+\alpha-1)p < n$. Then again $l = 0$, $u = v$, and for this last case the required assertion (ii) can be proved repeating almost "word by word" the same arguments as in the previous Theorem 4.3 with the following evident modification: now it is possible that $s = p$ (respectively, $q+\mu = \tau_*$), and for this situation one has to apply Theorem 2.9 instead of previous Theorem 2.8 (where $s > p$). □

4.5. Estimates on cubes for Sobolev classes of mappings $W^k_1(\mathbb{R}^n)$, $k \geq n$

As usual, by $W^k_p(\mathbb{R}^n, \mathbb{R}^d)$ we denote the space of functions $v : \mathbb{R}^n \to \mathbb{R}^d$ such that all its generalized derivatives $D^j v$, $j = 1, \ldots, k$ up to order k belong to the space L_p. It is well known that

$$W^k_p(\mathbb{R}^n) = \mathscr{L}^k_p(\mathbb{R}^n)$$

for $p > 1$. In this subsection we consider the limiting case $p = 1$ for Sobolev spaces W^k_1. It is well known that functions from the Sobolev space $W^k_1(\mathbb{R}^n, \mathbb{R}^d)$ are continuous if

$$k \geq n, \tag{4.46}$$

so we assume this condition below: fix $k \geq n$, and $v \in W^k_1(\mathbb{R}^n, \mathbb{R}^d)$.

Denote again $Z_{v,m} = \{x \in \mathbb{R}^n \setminus A_v : \operatorname{rank} \nabla v(x) \leq m\}$. Here A_v means the set of 'bad' points, at which either the function v is not differentiable or which are not Lebesgue points for ∇v. Recall, that the set A_v is small in a sense: the equality

$$\mathcal{H}^{\tau_*}(A_v) = 0 \qquad \text{for} \quad \tau_* := n - k + 1 \tag{4.47}$$

holds (see, e.g., [13]); in particular, $A_v = \emptyset$ if $k - 1 \geq n$.

As in the previous section, it is convenient (and sufficient for our purposes) to restrict our attention on the following subset of critical points

$$Z'_v = \{x \in Z_{v,m} : |\nabla v(x)| \leq 1\}. \tag{4.48}$$

Take also a positive parameter $q \geq m$ and nonnegative $\mu \geq 0$.

Theorem 4.6. *Under above assumptions, the following statements hold:*

(i) if $k-1 \geq n$, then the gradient ∇v is continuous and uniformly bounded function, and for any sufficiently small n-dimensional interval $Q \subset \mathbb{R}^n$ the estimate

$$\Phi(Z'_v \cap Q) \leq C\, \sigma^{q-m} r^{q+\mu+(k-1-n)(q-m)} \tag{4.49}$$

holds, where again

$$r = \ell(Q), \qquad \sigma = \|\nabla^k v\|_{L_1(Q)}, \tag{4.50}$$

and the constant C depends on n, m, k, d only.

(ii) if $k = n$, then under additional assumption

$$q + \mu \geq 1 \tag{4.51}$$

for any n-dimensional interval $Q \subset \mathbb{R}^n$ the estimate

$$\Phi(Z'_v \cap Q) \leq C \left(\sigma^q r^\mu + \sigma^{q-m} r^{\mu+m} \right), \tag{4.52}$$

holds with the same r, σ, and with C depending on n, m, k, d only.

Remark 4.7. Estimates in Theorem 4.6 are very close to the estimates in Theorem 4.4 (formally, we could obtain these estimates, taking $\alpha = 0$ and $p = 1$ in Theorem 4.4 and using the corresponding another definition for σ, cf. (4.35) and (4.50)). But in Theorem 4.6 we could not use the maximal function as before: it is well known, that in general

$$\mathcal{M}\nabla^k v \notin L_1(\mathbb{R}^n).$$

Proof of Theorem 4.6. Let the assumptions in the beginning of this subsection be fulfilled. Fix an n-dimensional interval $Q \subset \mathbb{R}^n$. Take the approximating polynomial $P = P_Q$ from the subsection 2.3 and denote $v_Q = v - P$, $r = \ell(Q)$. Then we have

$$\begin{aligned} \big\|v_Q\big\|_{L_\infty(Q)} &\le C\, r^{k-n}\, \|\nabla^k v\|_{\mathrm{L}_1(Q)}; \\ \big\|\nabla v_Q\big\|_{L_\infty(Q)} &\le C\, r^{k-1-n}\, \|\nabla^k v\|_{\mathrm{L}_1(Q)} \qquad \text{if } \ k \ge n+1, \end{aligned} \tag{4.53}$$

where C is a constant depending on n, d, k only. Moreover, the mapping v_Q can be extended from Q to the entire $\mathbb{R}^n$ such that the extension (denoted again) $v_Q \in \mathrm{W}_1^k(\mathbb{R}^n, \mathbb{R}^d)$ and

$$\|\nabla^k v_Q\|_{\mathrm{L}_1(\mathbb{R}^n)} \le C_0 \|\nabla^k v\|_{\mathrm{L}_1(Q)}, \tag{4.54}$$

where C_0 also depends on n, d, k only.

The rest of the proof of Theorem is very similar to that one of Theorems 4.2–4.4. Consider two different cases.

- Case $k > n$. Then from (4.53) it follows immediately that

  $$\big\|\nabla v_Q\big\|_{L_\infty(Q)} \le C\, \sigma r^{k-1-n}, \tag{4.55}$$

  where $\sigma = \|\nabla^k v\|_{\mathrm{L}_1(Q)}$. Denote the right hand side of (4.55) by ε. Now to finish the proof for this case we can repeat almost word by word the last part of the proof of Theorem 4.1 to obtain the required estimate (4.49).
- Case $k = n$. Now from Theorem 2.10 and (4.54) we have

  $$\int_0^\infty \mathcal{H}^1_\infty(\{x \in Q : \mathcal{M}(\nabla v_Q)(x) \ge t\})\, \mathrm{d}t \le C\, \|\nabla^k v\|_{\mathrm{L}_1(Q)} = C\, \sigma. \tag{4.56}$$

  Then for this last case the required assertion (ii) can be proved repeating almost "word by word" the same arguments as in the previous Theorem 4.3 with many obvious simplifications (now we have parameters $\alpha = 0$, $s = p = 1$, etc.), and with the following evident modification: one has to apply formula (4.56) instead of previous Theorem 2.8. □

References

[1] D.R. Adams, A trace inequality for generalized potentials, Stud. Math. 48 (1973) 99–105.
[2] D.R. Adams, A note on Choquet integrals with respect to Hausdorff capacity, in: S. Janson (Ed.), Function Spaces and Applications, Lund 1986, in: Lecture Notes in Mathematics, Springer, New York, 1986, pp. 115–124.
[3] D.R. Adams, Morrey Spaces, Lecture Notes in Applied and Numerical Harmonic Analysis, Birkhäuser/Springer, Cham, 2015.
[4] D.R. Adams, L. Hedberg, Function Spaces and Potential Theory, Grundlehren der Mathematischen Wissenschaften (Fundamental Principles of Mathematical Sciences), vol. 314, Springer-Verlag, Berlin, 1996.
[5] G. Alberti, Generalized N-property and Sard theorem for Sobolev maps, Rend. Lincei Mat. Appl. 23 (2012) 477–491.

[6] J. Azzam, R. Schul, Hard Sard: quantitative implicit function and extension theorems for Lipschitz maps, Geom. Funct. Anal. 22 (2012) 1062–1123.

[7] L. Barbet, M. Dambrine, A. Daniilidis, L. Rifford, Sard theorems for Lipschitz functions and applications in optimization, Isr. J. Math. 212 (2) (2016) 757–790.

[8] S.M. Bates, Toward a precise smoothness hypothesis in Sard's theorem, Proc. Am. Math. Soc. 117 (1) (1993) 279–283.

[9] B. Bojarski, P. Hajłasz, Pointwise inequalities for Sobolev functions and some applications, Stud. Math. 106 (1993) 77–92.

[10] B. Bojarski, P. Hajłasz, P. Strzelecki, Sard's theorem for mappings in Hölder and Sobolev spaces, Manuscr. Math. 118 (2005) 383–397.

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

[12] J. Bourgain, M.V. Korobkov, J. Kristensen, On the Morse–Sard property and level sets of Sobolev and BV functions, Rev. Mat. Iberoam. 29 (1) (2013) 1–23.

[13] J. Bourgain, M.V. Korobkov, J. Kristensen, On the Morse–Sard property and level sets of $\mathrm{W}^{n,1}$ Sobolev functions on $\mathbb{R}^n$, J. Reine Angew. Math. (Crelles J.) 2015 (700) (2015) 93–112, https://doi.org/10.1515/crelle-2013-0002, Online first 2013.

[14] D. Bucur, A. Giacomini, P. Trebeschi, Whitney property in two dimensional Sobolev spaces, Proc. Am. Math. Soc. 136 (7) (2008) 2535–2545.

[15] L. De Pascale, The Morse–Sard theorem in Sobolev spaces, Indiana Univ. Math. J. 50 (2001) 1371–1386.

[16] J.R. Dorronsoro, Differentiability properties of functions with bounded variation, Indiana Univ. Math. J. 38 (4) (1989) 1027–1045.

[17] Dubovitskiĭ A. Ya, Structure of level sets for differentiable mappings of an n-dimensional cube into a k-dimensional cube (Russian), Izv. Akad. Nauk SSSR, Ser. Mat. 21 (3) (1957) 371–408.

[18] Dubovitskiĭ A. Ya, On the set of degenerate points (Russian), Izv. Akad. Nauk SSSR, Ser. Mat. 31 (1) (1967) 27–36, https://doi.org/10.1070/IM1967v001n01ABEH000545, English transl.: Math. USSR, Izv. 1 (1) (1967) 25–33.

[19] L.C. Evans, R.F. Gariepy, Measure Theory and Fine Properties of Functions, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1992.

[20] H. Federer, Two theorems in geometric measure theory, Bull. Am. Math. Soc. 72 (1966) 719.

[21] H. Federer, Geometric Measure Theory, Springer-Verlag, New York, Heidelberg, Berlin, 1969.

[22] A. Ferone, M.V. Korobkov, A. Roviello, On the Luzin N-property and uncertainty principle for Sobolev mappings, Anal. PDE 12 (5) (2019) 1149–1175, https://doi.org/10.2140/apde.2019.12.1149.

[23] A. Ferone, M.V. Korobkov, A. Roviello, Morse–Sard theorem and the Luzin N-property: a new synthesis for smooth and Sobolev spaces, Sib. Math. J. 60 (5) (2019) 916–926, https://doi.org/10.1134/S0037446619050148.

[24] A. Figalli, A simple proof of the Morse–Sard theorem in Sobolev spaces, Proc. Am. Math. Soc. 136 (2008) 3675–3681.

[25] P. Hajłasz, S. Zimmermann, The Dubovitskii-Sard theorem in Sobolev spaces, Indiana Univ. Math. J. 66 (2) (2017) 705–723, https://doi.org/10.1512/iumj.2017.66.6011.

[26] P. Hajłasz, M.V. Korobkov, J. Kristensen, A bridge between Dubovitskii-Federer theorems and the coarea formula, J. Funct. Anal. 272 (3) (2017) 1265–1295, https://doi.org/10.1016/j.jfa.2016.10.031.

[27] M.V. Korobkov, J. Kristensen, On the Morse-Sard Theorem for the sharp case of Sobolev mappings, Indiana Univ. Math. J. 63 (6) (2014) 1703–1724, https://doi.org/10.1512/iumj.2014.63.5431.

[28] M.V. Korobkov, J. Kristensen, The trace theorem, the Luzin N- and Morse–Sard properties for the sharp case of Sobolev–Lorentz mappings, J. Geom. Anal. 28 (3) (2018) 2834–2856, https://doi.org/10.1007/s12220-017-9936-7.

[29] M.V. Korobkov, K. Pileckas, R. Russo, Solution of Leray's problem for stationary Navier-Stokes equations in plane and axially symmetric spatial domains, Ann. Math. 181 (2) (2015) 769–807, https://doi.org/10.4007/annals.2015.181.2.7.

[30] M. Kučera, Hausdorff measures of the set of critical values of functions of the class $C^{k,\lambda}$, Comment. Math. Univ. Carol. 13 (2) (1972) 333–350.

[31] J. Malý, Advanced theory of differentiation – Lorentz spaces, preprint, available at https://www.scribd.com/document/291534463/Advanced-Diff-Theory-Jan-Maly, 2003.

[32] V.G. Maz'ya, Sobolev Spaces with Applications to Elliptic Partial Differential Equations, Grundlehren der Mathematischen Wissenschaften (Fundamental Principles of Mathematical Sciences), vol. 342, Springer, Heidelberg, 2011.

[33] C.G.T. De A. Moreira, Hausdorff measures and the Morse–Sard theorem, Publ. Mat. 45 (2001) 149–162.

[34] A.P. Morse, The behavior of a function on its critical set, Ann. Math. 40 (1939) 62–70.

[35] A. Norton, A critical set with nonnull image has large Hausdorff dimension, Trans. Am. Math. Soc. 296 (1986) 367–376.

[36] D. Pavlica, L. Zajíček, Morse–Sard theorem for d.c. functions and mappings on $\mathbb{R}^2$, Indiana Univ. Math. J. 55 (3) (2006) 1195–1207.

[37] A.S. Romanov, Absolute continuity of the Sobolev type functions on metric spaces, Sib. Math. J. 49 (5) (2008) 911–918.

[38] A. Sard, The measure of the critical values of differentiable maps, Bull. Am. Math. Soc. 48 (1942) 883–890.

[39] E.M. Stein, Singular Integrals and Differentiability Properties of Functions, Princeton Mathematical Series, vol. 30, Princeton University Press, Princeton, N.J., 1970.

[40] D. Swanson, Pointwise inequalities and approximation in fractional Sobolev spaces, Bull. Am. Math. Soc. 149 (2) (2002) 147–174.

[41] R. Van der Putten, The Morse-Sard theorem in $W^{n,n}(\Omega)$: a simple proof, Bull. Sci. Math. 136 (2012) 477–483.

[42] H. Whitney, A function not constant on a connected set of critical points, Duke Math. J. 1 (1935) 514–517.

[43] Y. Yomdin, The geometry of critical and near-critical values of differentiable mappings, Math. Ann. 264 (4) (1983) 495–515.

[44] W.P. Ziemer, Weakly Differentiable Functions. Sobolev Spaces and Functions of Bounded Variation, Graduate Texts in Math., vol. 120, Springer-Verlag, New York, 1989.