Pointwise inequalities for Sobolev functions revisited*

Bogdan Bojarski, Piotr Hajłasz, Paweł Strzelecki

Primary: 46E35 Secondary: 41A63, 41A80, 41A99, 31B15
Key words and phrases: Sobolev function, Taylor polynomial, approximation, pointwise inequalities, Bessel capacity

Abstract

The approximation theorem of Michael and Ziemer asserts that any function $u \in W^{k,p}(\mathbb{R}^n)$ can be redefined on a set of small Bessel capacity $B_{k-m,p}$ to yield a function $w \in C^m$, which moreover is close to u in the sense of Sobolev norm $W^{m,p}$. We extend this result in two ways. First, we show that it is possible to obtain the approximation in the higher order Sobolev space $W^{m+1,p}$ without changing the estimate for the capacity. Moreover, we generalize the theorem to the case of approximation by $C^{m,\lambda}$ functions. The proofs are based on a new extension formula, different from the classical one of Whitney.

1 Introduction

The well known theorem of Luzin states that for any measurable function u defined on $\mathbb{R}^n$ and any $\varepsilon > 0$, one can find a continuous function φ and a closed set F with the Lebesgue measure of $\mathbb{R}^n \setminus F$ smaller than ε, such that $u \equiv \varphi$ on F. There have been several refinements of Lusin's theorem by showing that the more regular measurable function u, the smoother approximating function φ can be choosen. Our paper is about such generalizations of the Lusin theorem for functions u in Sobolev spaces $W^{k,p}$.

Let us start with an account of the history of the problem. All the results that we are going to mention are based on the celebrated Whitney C^m-extension theorem [31] (see also [21] and Theorem 4.4 below), which provides a necessary

*This work was supported by KBN grants no. PO3A 034 08 and PO3A 055 14 and also by the EU Research Training Network HPRN–CT–1999–00118. Final version was prepared when the last author was staying at Mathematisches Institut der Universität Bonn; he gratefully acknowledges the generosity of Humboldt Foundation and hospitality of his colleagues from Bonn.

and sufficient condition for a continuous function on an arbitrary closed subset of $\mathbb{R}^n$ to be extendable to a C^m smooth function on the entire $\mathbb{R}^n$.

We postpone explanation of some of the (mostly standard) notation and terminology that will be used now till the next section, "Preliminaries".

It seems that the story has started with a result of Federer (proved implicitely in [11, p. 442]) who showed that if a function u on $\mathbb{R}^n$ is differentiable a.e., then to every $\varepsilon > 0$ there exists a function $w \in C^1(\mathbb{R}^n)$ and a closed set F such that $u \equiv w$ on F and $|\mathbb{R}^n \setminus F| < \varepsilon$ ($|A|$ denotes the Lebesgue measure of A). Observe that by the Rademacher theorem, the result applies to all Lipschitz functions. The idea of Federer's proof is the following. One first shows that there exists a closed set F with $|\mathbb{R}^n \setminus F| < \varepsilon$, such that the funtion u restricted to F satisfies the assumptions of Whitney's C^1-extension theorem. Then, the theorem follows directly from Whitney's theorem. The reduction of the problem to Whitney's extension theorem is rather easy (the argument is half-page long); what is difficult and deep here is just the Whitney theorem.

Another result is due to Whitney [32] who proved that if a function u has approximate partial derivatives, then the same claim holds as in the Federer theorem. We will not recall the definition of the approximate partial derivative, we simply note that classical partial derivatives are approximate partial derivatives as well. It is well known result of Nikodym [26] (see also [10, 4.9.1], and Lemma 4.7 below) that the Sobolev function $u \in W^{1,p}(\mathbb{R}^n)$ with first order derivatives in L^p, admits a representative which has partial derivatives a.e. and thus it follows from the Whitney theorem [32] that u coincides with a C^1 function off a set of an arbitrary small Lebesgue measure (it seems that this simple application of Whitney's theorem [32] was unnoticed in the literature).

A far reaching generalization of the last observation is due to Calderón and Zygmund [7, Theorem 13] who extended the theorem to Sobolev spaces with higher order derivatives. They proved that for u in the Sobolev space $W^{k,p}$ and arbitrary $\varepsilon > 0$ there exists a closed set F and a function $w \in C^k(\mathbb{R}^n)$ such that $|\mathbb{R}^n \setminus F| < \varepsilon$ and $u \equiv w$ on F. Again the proof was based upon the idea of reducing the problem to the Whitney C^k-extension theorem. This time, the reduction was much more difficult than in the case of first order Sobolev spaces. Actually, Calderón and Zygmund proved a modified version of Whitney's theorem, convenient for applications to Sobolev spaces [7, Theorem 9] (their theorem contains however a small gap; we comment on it later).

The next result is due to Liu [20]. He proved that for $u \in W^{k,p}$ the function $w \in C^k(\mathbb{R}^n)$ can be choosen in a way that in addition to the condition $u \equiv w$ in a closed set F with $|\mathbb{R}^n \setminus F| < \varepsilon$ from Calderón and Zygmund's theorem one obtains the estimate for the Sobolev norm $\|u - w\|_{W^{k,p}} < \varepsilon$. Again the main idea of the proof was similar to that before: reduce the problem to Whitney's C^k-extension theorem and carefully examine the norm of the Whitney extention to obtain the desired estimate for the Sobolev norm of $u - w$.

A strengthened version of the Calderón and Zygmund theorem, with additional information in terms of capacities, is due to Michael and Ziemer [25] (see also [33], [34], [6]). They proved the following result.

Theorem 1.1 *Let $\Omega \subset \mathbb{R}^n$ be open. Assume that $1 < p < \infty$, $\varepsilon > 0$ and $m \in \{0, 1, \ldots, k\}$. Then, for any $u \in W^{k,p}_{\rm loc}(\Omega)$ there exists a closed subset F of Ω and a function $w \in C^m(\Omega) \cap W^{m,p}_{\rm loc}(\Omega)$ such that*

(i) $B_{k-m,p}(\Omega \setminus F) < \varepsilon$, *where $B_{k-m,p}$ denotes the Bessel capacity;*

(ii) $D^\alpha u(x) = D^\alpha w(x)$ *for any $x \in F$ and any α with $|\alpha| \leq m$;*

(iii) $u - w \in W^{m,p}_0(\Omega)$;

(iv) $\|u - w\|_{W^{m,p}(\Omega)} < \varepsilon$.

An earlier version of this theorem, without (iii) and (iv), was proved by Bagby and Ziemer [3]. For $k = m$, the Bessel capacity $B_{0,p}$ coincides with the Lebesgue measure and hence Theorem 1.1 covers the result of Liu. If $k > m$, sets of small $B_{k-m,p}$ capacity are "smaller" than generic sets of small Lebesgue's measure. The price one has to pay a very good estimate of the size of the complement of F in (i) is the lesser degree of smoothnes of $w \in C^m$, $m < k$ and the estimate in the lower order Sobolev norm $W^{m,p}$, $m < k$ (lesser and lower than corresponding estimates in the theorem of Liu).

Let us mention that there is a small gap in the proof of Calderón and Zygmund's theorem. Namely their version of Whitney's extension theorem [7, Theorem 9] is slightly in error and it is not true without some modifications in the statement. The result of Calderón and Zygmund was employed by Bagby and Ziemer, [3], Liu, [20], and Michael and Ziemer, [25]. Fortunately Ziemer [33], [34, Chapter 3], fixed the gap thus correcting the statement of Theorem 9 in [7] and saving results mentioned above.

The theorem of Michael and Ziemer has been generalized to the case of Bessel potential spaces and Besov spaces by Stocke [29].

It is getting boring, but let us mention one more time that Michael and Ziemer employed Whitney's C^m-extension theorem (in the form proved by Calderón and Zygmund [7, Theorem 9]).

A short proof of the Michael and Ziemer theorem was obtained by Bojarski and Hajłasz, [6]. Their idea was the following. First they proved pointwise inequalities, Corollary 3.9 below. If $u \in W^{k,p}_{\rm loc} \subset W^{m,1}_{\rm loc}$ and $M^\flat_R(\nabla^m u)$ goes uniformly to 0 on a closed set F as $R \to 0$, then it immediately follows that the function u restricted to F satisfies the assumptions of classical Whitney's C^m-extension theorem (Theorem 4.4). Now the estimate for the Bessel capacity of the set where the maximal function $M^\flat_R(\nabla^m u)$ is large and the careful examination of the explicit formula for Whitney's extension of $u|_F$ give the result.

We discovered that a new technique that omits Whitney's theorem leads to a much better result: for $m < k$ one obtains a surprising higher order approximation $\|u - w\|_{W^{m+1,p}(\Omega)} < \varepsilon$ without loosing any information about capacity of the exceptional set and about the smoothness of the approximating function w. This is one of the two main results of the paper.

Theorem 1.2 *Let $\Omega \subset \mathbb{R}^n$ be open. Assume that $1 < p < \infty$, $\varepsilon > 0$ and $m \in \{0, 1, \ldots, k-1\}$. Then for any $u \in W^{k,p}_{\mathrm{loc}}(\Omega)$ there exists a closed subset $F \subset \Omega$ and a function $w \in C^m(\Omega) \cap W^{m+1,p}_{\mathrm{loc}}(\Omega)$ such that*

(i) $B_{k-m,p}(\Omega \setminus F) < \varepsilon$;

(ii) $D^\alpha u(x) = D^\alpha w(x)$ *for any* $x \in F$ *and any* α *with* $|\alpha| \le m$;

(iii) $u - w \in W^{m+1,p}_0(\Omega)$;

(iv) $\|u - w\|_{W^{m+1,p}(\Omega)} \le C(n,k,p)\|u\|_{W^{m+1,p}(\Omega\setminus F)} < \varepsilon$.

Also condition (iv) is slightly stronger than the corresponding one in Theorem 1.1.

If $m = k$, then our method leads to the approximation in $W^{k,p}_0$ only, like in theorems by Liu and by Michael and Ziemer, but, of course, in this case there is no hope to obtain an approximation in a higher order space. However we still have

$$\|u - w\|_{W^{k,p}(\Omega)} \le C\|u\|_{W^{k,p}(\Omega\setminus F)} < \varepsilon$$

like in Theorem 1.2.

Let us formulate a special case of the theorem in the case of first order derivatives. It is a classical result that a function $u \in W^{1,p}(\mathbb{R}^n)$, where $1 < p < \infty$ admits a representative which is continuous outside a set of arbitrary small Bessel capacity $B_{1,p}$. The following result which is a direct consequence of Theorem 1.2 considerably strengthens this result.

Theorem 1.3 *Let $u \in W^{1,p}(\mathbb{R}^n)$, $1 < p < \infty$. Then to every $\varepsilon > 0$ there exists a continuous function in the Sobolev space $w \in C^0 \cap W^{1,p}(\mathbb{R}^n)$ such that*

(i) $B_{1,p}(\{x \in \mathbb{R}^n : u(x) \ne w(x)\}) < \varepsilon$;

(ii) $\|u - w\|_{W^{1,p}(\mathbb{R}^n)} < \varepsilon$.

See also [18] for a direct proof and a generalization to the setting of Sobolev spaces in metric spaces introduced in [16].

It was a natural problem to ask for an approximation by functions in the class $C^{m,\lambda}$, $\lambda \in (0,1)$ in order to obtain a continuous scale of approximation. One of the ideas is to modify the proof in [6] by employing different pointwise inequalities, Corollary 3.7 below. Reasoning like this one proves that if $u \in W^{k,p}_{\rm loc}$, $p > 1$, $0 \leq m \leq k-1$, $\lambda \in (0,1)$, and $\varepsilon > 0$, then there exists $w \in C^{m,\lambda}_{\rm loc}(\Omega)$ such that $u \equiv w$ on F with $B_{k-m-\lambda,p}(\Omega \setminus F) < \varepsilon$, $\|u - w\|_{W^{m,p}(\Omega)} < \varepsilon$.

However, employing the method which leads to Theorem 1.2 we get higher order approximation in this case as well, namely $\|u - w\|_{W^{m+1,p}} < \varepsilon$ with the estimate for the capacity unchanged. This is the other main result of the paper. In order to formulate the theorem we need to define a class of smooth functions. We say that $u \in \widetilde{C}^{m,\lambda}(\Omega)$, $m \in \{0,1,2,\ldots\}$, $\lambda \in (0,1)$, if $u \in C^m(\Omega)$ and for every compact set $K \subset \Omega$

$$\lim_{\varrho \to 0} \sup_{\substack{x,y \in K \\ x \neq y \\ |x-y| \leq \varrho}} \frac{|\nabla^m u(x) - \nabla^m u(y)|}{|x-y|^\lambda} = 0 \,. \tag{1}$$

For a function $u \in C^{m,\lambda}(\Omega)$ the supremum that appears in (1) is bounded only, so (1) means that on compact sets m-th order derivatives of funtions in the class $\widetilde{C}^{m,\lambda}$ have better modulus of continuity than $C^{0,\lambda}$-Hölder continuous functions. On the other hand, a function $u \in \widetilde{C}^{m,\lambda}(\Omega)$ need not have globally Hölder continuous derivatives as the constant C in the inequality

$$|\nabla^m u(x) - \nabla^m u(y)| \leq C|x-y|^\lambda, \qquad \text{for } x,y \in K$$

may blow-up to infinity with a sequence of compact sets $K \subset \Omega$ that exorce Ω. We can now formulate the result.

Theorem 1.4 *Let $\Omega \subset \mathbb{R}^n$ be an open set. Assume that $1 < p < \infty$, $\lambda \in (0,1)$ and $m \in \{0,1,\ldots,k-1\}$. Then, for any $u \in W^{k,p}_{\rm loc}(\Omega)$ and any $\varepsilon > 0$ there exists a closed subset F of Ω and a function $w \in \widetilde{C}^{m,\lambda}(\Omega) \cap W^{m+1,p}_{\rm loc}(\Omega)$ such that*

(i) $B_{k-m-\lambda,p}(\Omega \setminus F) < \varepsilon$;

(ii) *In the particular case when $k = m+1$ and $(1-\lambda)p \leq n$ we also have $\mathcal{H}^{n-(1-\lambda)p}_\infty(\Omega \setminus F) < \varepsilon$, where $\mathcal{H}^{n-(1-\lambda)p}_\infty$ denotes the Hausdorff content;*

(iii) $D^\alpha u(x) = D^\alpha w(x)$ *for any $x \in F$ and any α with $|\alpha| \leq m$;*

(iv) $u - w \in W^{m+1,p}_0(\Omega)$;

(v) $\|u - w\|_{W^{m+1,p}(\Omega)} \leq C(n,k,\lambda,p)\|u\|_{W^{m+1,p}(\Omega \setminus F)} < \varepsilon$.

An abstract version of the result, which is valid for $k = 1$ and for functions in Sobolev spaces on a metric space as defined in [16], has been recently obtained by Hajłasz and Kinnunen, [17]. A careful reader should note that Theorem 1.4 strengthens and extends Malý's result [22] (thm. 7, page 252) on Hölder type quasicontinuity of Sobolev functions. Our proof seems to be more natural and, contrary to Malý's, can be easily written for Sobolev spaces with derivatives of arbitrary order: we do not use truncation at all. In our theorem we obtain sharper information on capacity even for first order derivatives, and the Hölder exponent of the approximating function can be chosen in an arbitrary way. When a preliminary version of this paper had already been completed, we received a preprint of David Swanson [30], who has extended our results to functions belonging to fractional Sobolev spaces, i.e. Bessel potential spaces $L^{\alpha,p}$.

Remarks. (1) In Theorems 1.1, 1.2 and 1.4 it is assumed that u and its derivatives are only *locally* integrable, but the choice of F allows one to control the appropriate norm of $(u-w)$ by the Sobolev norm of u on $\Omega \setminus F$ — which becomes finite (and small) if F is sufficiently "large". (2) In the theorems F is a closed subset *of* Ω, i.e. an intersection of a closed subset of $\mathbb{R}^n$ with Ω, so F does not have to be a closed subset of $\mathbb{R}^n$. In particular we can have $F=\Omega$. (3) In the supercritical case $p>n$, condition (ii) of Theorem 1.4 yields in fact an optimal imbedding of $W^{1,p}$ into $C^{0,\lambda}$. In fact, for $\lambda = 1-n/p$ condition (ii) ascertains that $\mathcal{H}^0_\infty(F)$ is small; since $\mathcal{H}^0_\infty(E) \geq 1$ for all nonempty sets E, F is empty, and (the canonical representative of) u agrees everywhere with a Hölder continuous function. Combining this with a classical inductive argument, one sees that (ii) gives in fact an optimal imbedding of $W^{k,p}$ into $C^{s,\lambda}$ in the whole supercritical range $kp>n$. Note also that the behaviour of Bessel capacity on balls (see Ziemer [34, Section 2.6]), implies that $B_{\alpha,p}(E) \leq C\mathcal{H}^{n-\alpha p}_\infty(E)$ when $\operatorname{diam} E < 1/2$ and $\alpha p < n$. Therefore, for $k=m+1$ condition (ii) of Theorem 1.1 is stronger than (i); the latter one does not imply an imbedding into $C^{0,\lambda}$ for the optimal value of Hölder exponent λ. We were not able to obtain an analogue of (ii) for $k>m+1$.

All Theorems 1.1, 1.2 and 1.4 require a very careful choice of a representative of the function $u \in W^{k,p}_{\mathrm{loc}}$ (in the class of functions which equal to u a.e.). Typically, one chooses the so-called quasicontinuous representative. This approach, however, is rather technical as it requires a good understanding of the capacity theory. In our approach we choose a *canonical Borel representative* of the function and its distributional derivatives, using the formula

$$\widetilde{u}(x) := \limsup_{r\to 0} \fint_{B(x,r)} u(y)\, dy. \tag{2}$$

This representative is well defined at each point x of the domain of u. In condition (ii) of Theorems 1.1 and 1.2 and condition (iii) of Theorem 1.4 the values of $D^\alpha u(x)$ are understood precisely in this sense.

It turns out that pointwise inequalities (Corollary 3.9 and Corollary 3.7) employed in our proof are true at *every* point of Ω. This trick, which was used for the first time in [6], simplifies the proof: no knowledge about quasicontinuous representatives is required.

However, the main novelty in our proof is the replacement of the Whitney extension theorem by a new construction that we call *Whitney's smoothing*.

When the classical Whitney extension formula is being applied to a Sobolev function u which have been restricted to some closed set F we loose *all* the information about the behaviour of the function in the complement of F. This is too much. To avoid this, we replace the Whitney extension formula by a new one which takes into account both the behaviour of the function u on F and its behaviour on $\Omega \setminus F$. The resulting function, which we denote here by $\widetilde{u}$ is, roughly speaking, defined as follows. We leave u unchanged on F i.e. $\widetilde{u} \equiv u$ on F and define $\widetilde{u}$ on $\Omega \setminus F$ by applying an approximation procedure to $u|_{\Omega\setminus F}$. We use the name 'Whitney's smoothing', because ideas related to the Whitney extension theorem are involved here. Thus in fact we do not define $\widetilde{u}$ in $\Omega \setminus F$ by extending u from F, but we pick some smooth approximation of $u|_{\Omega\setminus F}$. This leads to Theorems 1.2 and 1.4.

Though we have no formal proof of this, we are tempted to think that it is not possible to obtain our main result via an application of the classical Whitney theorem. We also believe that our results are optimal in the sense that estimate (iv) in Theorem 1.2 and estimate (v) in Theorem 1.4 cannot be replaced by the estimate $\|u-w\|_{W^{\ell,p}(\Omega)} < \varepsilon$ for $\ell > m+1$.

Let us remark that the (a priori purely analytic) Lusin theorems for Sobolev functions are closely related to the theory of the so-called C^k-rectifiable sets introduced by Anzelotti and Serapioni, see [2]. A set $M \subset \mathbb{R}^{n+k}$ is called (H^n, n)-rectifiable of class C^k (or shortly: C^k-rectifiable) if $M = M_0 \cup M_1$, where M_1 is a subset of a countable union of n-dimensional submanifolds $S_j \subset \mathbb{R}^m$, each S_j being of class C^k, and $H^n(M_0) = 0$. ($C^{s,\lambda}$-rectifiable sets are defined in [2] in an analogous way.) Such sets have H^n-a.e. "approximate tangent paraboloids" of order k. Graphs of functions $u \in W^{k,p}(\mathbb{R}^n)$ are (H^n, n)-rectifiable sets of class C^k in $\mathbb{R}^{n+1}$. Alberti [1] has shown that the theory of C^k-rectifiable sets provides a natural and optimal setting if one wants to describe singularities of convex functions and of convex surfaces. Both [2] and [1] use earlier results of Dorronsoro [9], who obtained subtle results on the existence of higher order L^q-differentials of functions in $BV^k_{\rm loc}(\mathbb{R}^n)$ (the space of functions whose distributional derivatives of order k are measures). Also both papers [2] and [1] contain the following result: if $u \in BV^k_{\rm loc}(\mathbb{R}^n)$, then it coincides with a C^k function off a set of arbitrarily small measure. Alberti applied this theorem to convex functions in $\mathbb{R}^n$ which by Alexandrov's theorem belong to $BV^2_{\rm loc}$. The results about Lusin properties of $BV^k_{\rm loc}$ can be easily put into the framework of our paper. This will be subject of a forthcoming paper.

The paper is organised as follows. In section "Preliminaries" we collect all the notation, definitions and basic results needed in the paper. In section "Pointwise inequalities" we prove pointwise inequalities which provide main estimates relating the behaviour of Sobolev functions to the behaviour of C^m-smooth functions. The last two sections are devoted to the proofs of main results, Theorem 1.2

and Theorem 1.4. In Section 4 we present a detailed proof of Theorem 1.4 and then, in Section 5, we show how to modify the proof to get Theorem 1.2.

2 Preliminaries

2.1 Notation, definitions etc.

The notation throughout the rest of this paper is either standard or selfexplanatory. The Lebesgue measure of a set A will be denoted by $|A|$. The barred integral $\fint_A f\,dx$ as well as f_A denotes the average value of a function f over a measurable set A, $\fint_A f = f_A := |A|^{-1}\int_A f\,dx$. Characteristic function of a set A will be denoted by χ_A. By B_r or $B(a,r)$ we denote the Euclidean ball in $\mathbb{R}^n$ of radius r, centered at a. The letter Q denotes a cube in $\mathbb{R}^n$ with edges parallel to coordinate axes. By kQ, $k>0$ we will denote the cube concentric with Q, with the diameter k times that of Q. By $C^{m,\lambda}(\Omega)$, where $m\in\{0,1,2,\ldots\}$ and $\lambda\in(0,1)$, we denote the class of functions u which are m times continuously differentiable on Ω with

$$|\nabla^m u(x) - \nabla^m u(y)| \leq C|x-y|^\lambda$$

for all $x,y\in\Omega$ and some constant $C>0$. Given two expressions A and B, we write $A\approx B$ if $C_1A\leq B\leq C_2A$ for some positive constants C_1 and C_2.

We write

$$\begin{aligned} T^k_x f(y) &= \sum_{|\alpha|\leq k} D^\alpha f(x)\frac{(y-x)^\alpha}{\alpha!}\,, \\ T^k_S f(y) &= \fint_S T^k_x f(y)\,dx \end{aligned}$$

to denote the Taylor polynomial and the average of the Taylor polynomial over a measurable set S, respectively.

By $\nabla^m f$ we denote the vector with the components $D^\alpha f$, $|\alpha|=m$.

C will denote a general constant which can change its value in the same string of estimates. Writing $C(n,m)$ we emphasize that the constant depends on n and m only.

If f is a locally integrable function, then we define $\widetilde{f}$ at every point by formula (2). Since by the Lebesgue differentiation theorem, [28], $f=\widetilde{f}$ a.e. in what follows, as a rule, we identify $\widetilde{f}$ with f and omit the tilda.

We say that x is a Lebesgue point of f if

$$\lim_{r\to 0}\fint_{B(x,r)} |f(y)-f(x)|\,dy = 0$$

($f(x)$ is defined by (2)). It is a well known result of Lebesgue, [28], that for $f \in L^1_{\rm loc}$ the set of points which are not the Lebesgue points of f is of the Lebesgue measure zero.

For any open $\Omega \subset \mathbb{R}^n$, $m \in \{1,2,3,\ldots\}$ and $1 \le p < \infty$ we use the following definition of the Sobolev space:

$$W^{m,p}(\Omega) = \{f \in \mathcal{D}'(\Omega) : D^\alpha f \in L^p(\Omega), \quad |\alpha| \le m\},$$

$$\|f\|_{W^{m,p}(\Omega)} = \sum_{|\alpha|\le m} \|D^\alpha f\|_p,$$

where $\|\cdot\|_p$ denotes the L^p-norm. Analogously we define the corresponding local space $W^{m,p}_{\rm loc}$. Obviously $W^{m,p} \subset W^{m,1}_{\rm loc}$. The space $W^{m,p}_0(\Omega)$ is defined as the closure of $C_0^\infty(\Omega)$ in the norm of $W^{m,p}(\Omega)$.

In the sequel, some variants of the Hardy-Littewood maximal functions are used:

$$\begin{aligned} M^\lambda_\varrho f(x) &= \sup_{r<\varrho} r^\lambda \cdot ⨍_{B(x,r)} |f(y)|dy\,, \\ M^\lambda f &= M^\lambda_\infty f\,, \\ M^\flat_\varrho f(x) &= \sup_{r<\varrho} ⨍_{B(x,r)} |f(y) - f(x)|\,dy\,, \\ M^\flat f &= M^\flat_\infty f. \end{aligned}$$

If $f = (f_1, \ldots, f_N)$ is a vector-valued function, then

$$M^\flat_\varrho f(x) := \sum_{j=1}^N M^\flat_\varrho f_j(x)\,.$$

If $\lambda = 0$ we usually omit the superscript λ. Maximal functions with the superscript $\flat$ will be called *flat maximal functions.*

We will need the following result sometimes called "$5r$-covering lemma" [28, Chapter 1], [34, Theorem 1.3.1].

Lemma 2.1 *Let $E \subset \bigcup_{i=1}^\infty B(x_i, r_i)$ be a covering of a set $E \subset \mathbb{R}^n$ by balls, with $\sup_i r_i < \infty$. Then we can select a subfamily (possibly finite) of pairwise disjoint balls $B(x_{i_j}, r_{i_j})$, $j = 1,2,\ldots$ such that $E \subset \bigcup_{j=1}^\infty B(x_{i_j}, 5r_{i_j})$.*

2.2 Potentials, capacity and content

Let G_α, $\alpha > 0$, be the kernel of $(I - \Delta)^{-\alpha/2}$. Its Fourier transform is given by

$$\widehat{G}_\alpha(\xi) = (2\pi)^{-n/2}(1 + |\xi|^2)^{-\alpha/2}\,.$$

It is well known that G_α is positive, integrable and

$$G_\alpha * G_\beta = G_{\alpha+\beta} \tag{3}$$

for all $\alpha, \beta > 0$. If $0 < \alpha < n$, then the kernel G_α has the following estimate near the origin:

$$|x|^{\alpha-n} \le C(n,\alpha) G_\alpha(x) \qquad \text{for } |x| \le 1. \tag{4}$$

For $0 < \alpha < \infty$ and $1 < p < \infty$ the *space of Bessel potentials* is defined by

$$L^{\alpha,p}(\mathbb{R}^n) = \{G_\alpha * g : g \in L^p(\mathbb{R}^n)\}.$$

It is a Banach space with the norm $\|f\|_{\alpha,p} = \|g\|_p$, where $f = G_\alpha * g$. It is a well known result of Calderón and Lizorkin that for k a positive integer

$$W^{k,p}(\mathbb{R}^n) = L^{k,p}(\mathbb{R}^n)$$

as sets, and the norms are equivalent, see [28, Chapter 5].

The *Bessel capacity* is defined for any set $E \subset \mathbb{R}^n$ as

$$B_{\alpha,p}(E) = \inf \|f\|_p^p \, ,$$

the infimum being taken over the set of those nonnegative $f \in L^p(\mathbb{R}^n)$ for which $G_\alpha * f(x) \ge 1$ for all $x \in E$.

All sets of small capacity have small Lebesgue's measure. In fact, $B_{\alpha,p}(E) \ge C|E|^{(n-\alpha p)/n}$ when $\alpha p < n$. This follows from the Sobolev embedding and the definition of the Bessel capacity. In some sense sets of small capacity are "smaller" than generic sets of small Lebesgue's measure as for $1 < p < n/\alpha$ the estimates for the capacity are related to the estimates for the Hausdorff measure and dimension. Namely, $B_{\alpha,p}(E) = 0$ if $\mathcal{H}^{n-\alpha p}(E) < \infty$, and $B_{\alpha,p}(E) = 0$ implies $\mathcal{H}^{n-\alpha p+\varepsilon}(E) = 0$ for any $\varepsilon > 0$. If $\alpha p = n$, then a similar result holds with a "logarithmic Hausdorff measure". Finally if $\alpha p > n$, $B_{\alpha,p}(E) \ge C > 0$, whenever $E \ne \emptyset$. This means any set of small Bessel capacity $B_{\alpha,p}$ is empty when $\alpha p > n$.

For more details on these topics, see Ziemer's monograph [34, Chapter 2] and the original paper of Meyers [24].

The concepts of Bessel capacity is strictly related to that of *Hausdorff content* $\mathcal{H}^s_\infty$, where $s \ge 0$, which is defined by

$$\mathcal{H}^s_\infty(E) = \inf \sum_{i=1}^\infty r_i^s \, ,$$

the infimum being taken over *all* countable coverings of E with balls $B(x_i, r_i)$. It is almost obvious that $\mathcal{H}^s(E) = 0$ if and only if $\mathcal{H}^s_\infty(E) = 0$. On the other hand, $\mathcal{H}^s_\infty(E)$ is finite for all bounded $E \subset \mathbb{R}^n$.

It is well known, [34, Theorem 2.6.13], that

$$C_1 r^{n-\alpha p} \le B_{\alpha,p}(B(x,r)) \le C_2 r^{n-\alpha p}$$

for all $x \in \mathbb{R}^n$ and $r < 1$, whenever $\alpha p < n$. This and the $5r$-covering lemma (Lemma 2.1) imply that

$$B_{\alpha,p}(E) \le C\mathcal{H}^{n-\alpha p}_{\infty}(E) \tag{5}$$

for any set $E \subset \mathbb{R}^n$ with $\operatorname{diam} E \le 1$, provided $\alpha p < n$. It follows from this inequality that when $k = m+1$ and $(1-\lambda)p \le n$, then the estimate (ii) in Theorem 1.4 is stronger than the estimate from (i).

We close this section with a few basic estimates of the Bessel capacity and the Hausdorff content that will be used in the proof of the main result. The results below are variants of known results.

Lemma 2.2 *Let $f \in L^{\alpha,p}(\mathbb{R}^n)$, $n > \lambda > 0$ and $p > 1$. Then*

$$B_{\alpha+\lambda,p}(\{x :\ M_1^\lambda f(x) > t\}) \le Ct^{-p}\|f\|^p_{\alpha,p},$$

where $C = C(\alpha, p, \lambda, n)$.

Remark. The lemma is true for any $\lambda \ge 0$ with a slightly more technical, but otherwise similar proof.

Proof. Let $f = G_\alpha * g$, $\|f\|_{\alpha,p} = \|g\|_p$ and let $\omega_r^\lambda = r^\lambda |B(0,r)|^{-1}\chi_{B(0,r)}$. Then for all $r < 1$ we have

$$\begin{aligned} r^\lambda \fint_{B(x,r)} |f(y)|\,dy &= \omega_r^\lambda * |f|(x) \le \omega_r^\lambda * G_\alpha * |g|(x) \\ &= G_\alpha * \omega_r^\lambda * |g|(x) \le G_\alpha * M_1^\lambda |g|(x)\,, \end{aligned}$$

which implies $M_1^\lambda f(x) \le G_\alpha * M_1^\lambda g(x)$. Invoking (4), we get

$$\begin{aligned} r^\lambda \fint_{B(x,r)} |g(y)|\,dy &\le C(n) \int_{B(x,r)} \frac{|g(y)|}{|x-y|^{n-\lambda}}\,dy \\ &\le C(n,\lambda) \int_{B(x,r)} G_\lambda(x-y)|g(y)|\,dy. \end{aligned}$$

Hence $M_1^\lambda g(x) \le CG_\lambda * g(x)$. This and (3) give in turn

$$M_1^\lambda f(x) \le G_\alpha * M_1^\lambda g(x) \le CG_\alpha * G_\lambda * |g|(x) = CG_{\alpha+\lambda} * |g|(x).$$

We conclude as follows.

$$B_{\alpha+\lambda,p}(\{M_1^\lambda f > t\}) \le B_{\alpha+\lambda,p}(\{G_{\alpha+\lambda} * (C|g|/t) > 1\}) \le \|Cg/t\|^p_p = Ct^{-p}\|f\|^p_{\alpha,p}.$$

The proof is complete now. □

The following lemma is one of the key tools in the proof of our main theorem.

Lemma 2.3 *Assume that $u \in W^{k,p}(\mathbb{R}^n)$, where $p > 1$. Let $0 < \lambda < n$. Then for any $\varepsilon > 0$ there exists an open set V such that*

(i) $B_{k+\lambda,p}(V) < \varepsilon$,

(ii) *as $\varrho \to 0$, $M_\varrho^\lambda u(x) \to 0$ uniformly on the set $\mathbb{R}^n \setminus V$.*

Proof. Take $\varepsilon > 0$. Since compactly supported smooth functions are dense in $W^{k,p}(\mathbb{R}^n)$, we can pick $h \in C_0^\infty(\mathbb{R}^n)$ which satisfies the condition

$$\|u - h\|_{W^{k,p}} < \varepsilon^{(p+1)/p}\,. \tag{6}$$

As h is bounded, one has $M_\varrho^\lambda h(x) \le \varepsilon$ for all $x \in \mathbb{R}^n$, if $\varrho = \varrho(\varepsilon)$ is taken to be sufficiently small (e.g., $\varrho = (\varepsilon\,\|h\|_\infty^{-1})^{1/\lambda}$). By the subadditivity of the maximal function, we obtain $M_\varrho^\lambda u \le M_\varrho^\lambda h + M_\varrho^\lambda(u-h) \le \varepsilon + M_1^\lambda(u-h)$. Hence, by Lemma 2.2, (6) and the fact that $W^{k,p} = L^{k,p}$, we are led to

$$\begin{aligned} B_{k+\lambda,p}\left(\{x\colon\ M_\varrho^\lambda u(x) > 2\varepsilon\}\right) &\le B_{k+\lambda,p}\left(\{x\colon\ M_1^\lambda(u-h)(x) > \varepsilon\}\right) \\ &\le C\varepsilon^{-p}\|u-h\|_{W^{k,p}}^p < C\varepsilon\,. \end{aligned} \tag{7}$$

Set $\varepsilon_i = \varepsilon/(C\cdot 2^i)$ and $\varrho_i = \varrho(\varepsilon_i)$, where $i = 1, 2, \ldots$ Applying inequality (7), we check that the set

$$V := \bigcup_{i=1}^\infty \{x\colon\ M_{\varrho_i}^\lambda u(x) > 2\varepsilon_i\}$$

has small Bessel capacity. Indeed,

$$B_{k+\lambda,p}(V) \le \sum_{i=1}^\infty B_{k+\lambda,p}\left(\{M_{\varrho_i}^\lambda u > 2\varepsilon_i\}\right) \le \sum_{i=1}^\infty C\frac{\varepsilon}{C\cdot 2^i} = \varepsilon\,.$$

Moreover, V is open (as the maximal function $M_\varrho^\lambda u$ is always lower semicontinuous), and, of course, on its complement $M_\varrho^\lambda u$ converges uniformly to zero as ϱ goes to 0. □

Let us state a simple but useful corollary.

Corollary 2.4 *Let $(k+\lambda)p > n$. Then, for any $u \in W^{k,p}(\mathbb{R}^n)$, the maximal functions $M_\varrho^\lambda u$ tend to zero uniformly on $\mathbb{R}^n$ as ϱ goes to 0.*

If $u \in W^{k,p}$ and $\lambda \in (0,1)$, then it follows from Lemma 2.2 that for $m+1 \le k$

$$B_{k-m-\lambda,p}(\{M_1^{1-\lambda}|\nabla^{m+1}u| > t\}) \le Ct^{-p}\|u\|_{W^{k,p}}^p. \tag{8}$$

This inequality will be employed in the proof of the main result.

Similar estimates can be proved for the Hausdorff content in place of the Bessel capacity.

Lemma 2.5 *Let $f \in L^p(\mathbb{R}^n)$ with $p \in [1,\infty)$. Assume that $\lambda > 0$ and $\lambda p \leq n$. Then,*

$$\mathcal{H}^{n-\lambda p}_\infty\Big(\{x \in \mathbb{R}^n \mid M^\lambda_1 f(x) > t\}\Big) \leq C(n,\lambda,p) t^{-p} \|f\|^p_{L^p(\mathbb{R}^n)}\,.$$

Proof. Set $E_t = \{x \in \mathbb{R}^n \mid M^\lambda_1 f(x) > t\}$. For every point $x \in E_t$ there exists a number $r = r_x \in (0,1)$ such that

$$r^\lambda \fint_{B(x,r)} |f(y)|\, dy > t\,.$$

By the Hölder inequality, this implies

$$r^{n-\lambda p} \leq t^{-p} \int_{B(x,r)} |f(y)|^p\, dy\,.$$

Applying $5r$-covering lemma (Lemma 2.1) to the covering $(B(x,r_x))_{x\in E_t}$, we obtain a disjointed subfamily $(B(x_i,r_i))_{i=1,2,\ldots}$ such that $(B(x_i,5r_i))_{i=1,2,\ldots}$ forms a covering of E_t. Summing appropriate powers of the radii, we conclude the proof. □

A counterpart of Lemma 2.3 also holds true.

Lemma 2.6 *Let $f \in L^p(\mathbb{R}^n)$ with $p \in [1,\infty)$. Assume that $\lambda > 0$ and $\lambda p \leq n$. Then for any $\varepsilon > 0$ there exists an open set V such that*

(i) $\mathcal{H}^{n-\lambda p}_\infty(V) < \varepsilon$,

(ii) *as $\varrho \to 0$, $M^\lambda_\varrho f(x) \to 0$ uniformly on the set $\mathbb{R}^n \setminus V$.*

We omit the proof since it is almost identical to the proof of Lemma 2.3.

Let $u \in W^{k,p}(\mathbb{R}^n)$, where $k = m+1$ and let $\lambda \in (0,1)$ be such that $(k-m-\lambda)p \leq n$. Then it immediately follows from Lemma 2.5 that

$$\mathcal{H}^{n-(k-m-\lambda)p}_\infty\Big(\{x \in \mathbb{R}^n \mid M^{1-\lambda}_1|\nabla^{m+1}u|(x) > t\}\Big) \leq C t^{-p} \|\nabla^k u\|^p_{L^p(\mathbb{R}^n)}\,. \tag{9}$$

By (5), this inequality is stronger than (8). It would be nice to have (9) also for $k > m+1$, but we do not know if it holds true in that case. Such an estimate would improve the statement of Theorem 1.4: one could replace the Bessel capacity $B_{k-m-\lambda,p}$ in (i) in Theorem 1.4 by the Hausdorff content $\mathcal{H}^{n-(k-m-\lambda)p}_\infty$ (as in (ii) in Theorem 1.4) for all values of m.

3 Pointwise inequalities

To render our exposition self-contained, we repeat here some of the computations and proofs from Section 2 of [6].

To begin with, recall a well known inequality (see e.g. [13, Lemma 7.16]): there exists a constant $C = C(n)$ such that, for any cube $Q \subset \mathbb{R}^n$, any measurable set $S \subset Q$ and for any function $f \in C^1(Q)$, one has

$$|f(x) - f_S| \le C\frac{|Q|}{|S|}\int_Q \frac{|\nabla f(y)|}{|x-y|^{n-1}}\, dy \qquad \text{for all } x \in Q\,. \tag{10}$$

Extending a trick, used e.g. by Bojarski in [5] or Reshetnyak in [27], one can obtain a stronger inequality, involving the derivatives of any order m, which is more sophisticated than (10) even in the simple case $m = 1$.

Lemma 3.1 *If $f \in C^m(Q)$ and $a = (a_\alpha)_{|\alpha|=m}$, with all a_α being real numbers, then for any measurable set $S \subset Q$*

$$|f(x) - T_S^{m-1} f(x)| \le C\frac{|Q|}{|S|}\int_Q \frac{|\nabla^m f(y)|}{|x-y|^{n-m}} dy, \tag{11}$$

and

$$|f(x) - T_S^{m} f(x)| \le C\frac{|Q|}{|S|}\int_Q \frac{|\nabla^m f(y) - a|}{|x-y|^{n-m}}\, dy. \tag{12}$$

Both constants C depend on n and m only.

Proof. To prove inequality (12), compute first order partial derivatives (with respect to y) of

$$\varphi_x(y) = \sum_{|\alpha|\le m-1} D^\alpha f(y)\frac{(x-y)^\alpha}{\alpha!} + \sum_{|\alpha|=m} a_\alpha \frac{(x-y)^\alpha}{\alpha!}\,,$$

check that $|\nabla\varphi_x(y)| \le C(m,n)|\nabla^m f(y) - a||x-y|^{m-1}$, and write

$$f(x) - T_S^m f(x) = \left(\varphi_x(x) - \fint_S \varphi_x(y)\, dy\right) - \fint_S \sum_{|\alpha|=m} \big(D^\alpha f(y) - a_\alpha\big)\frac{(x-y)^\alpha}{\alpha!}\, dy$$

to obtain (12) as a direct consequence of (10); inequality (11) follows easily, by substituting $a = 0$. $\square$

Lemma 3.1 has an extension, which is valid for all Sobolev functions $f \in W^{m,1}_{\rm loc}$. For such f, we choose Borel representatives of the function defined at *every* point by the formula

$$f(x) := \limsup_{r\to 0} \fint_{B(x,r)} f(y)\, dy\,. \tag{13}$$

We will also need the following elementary lemma (see e.g. Lemma 2 in [6])

Lemma 3.2 *Let $\alpha > 0$, then there exists a constant $C = C(\alpha, n)$, such that for all $x, z \in \mathbb{R}^n$ and all $r > 0$*

$$\fint_{B(x,r)} |y - z|^{\alpha - n}\, dy \le \begin{cases} C|x - z|^{\alpha - n} & \text{if } \alpha \le n, \\ C(r + |x - z|)^{\alpha - n} & \text{if } \alpha > n. \end{cases}$$

Proof. We assume that $z = 0$ and consider two cases: $r < |x|/2$ and $r \ge |x|/2$. In the first case, $|y| \approx |x|$ for all $y \in B(x, r)$. In the second case, $B(x, r) \subset B(0, 3r)$; one increases the domain of integration and computes the integral explicitly. $\Box$

Theorem 3.3 *There exists a constant $C_{m,n}$ such that if $f \in W^{m,1}(Q)$ is defined at every point by formula (13) and $a = (a_\alpha)_{|\alpha|=m}$ is an arbitrary family of real numbers, then for every measurable set $S \subset Q$ the following inequalities hold at* each *point $x \in Q$:*

$$|f(x) - T_S^{m-1} f(x)| \le C_{m,n} \frac{|Q|}{|S|} \int_Q \frac{|\nabla^m f(y)|}{|x - y|^{n-m}} dy, \tag{14}$$

$$|f(x) - T_S^m f(x)| \le C_{m,n} \frac{|Q|}{|S|} \int_Q \frac{|\nabla^m f(y) - a|}{|x - y|^{n-m}} dy. \tag{15}$$

Remark. In most of the cases we will apply Theorem 3.3 for $S = Q$.

Proof. As in the proof of Lemma 3.1, the first inequality follows from the second one by substituting $a = 0$; therefore, we only sketch the proof of the second inequality. A standard approximation argument shows that for every $f \in W^{m,1}(Q)$ inequality (15) holds a.e. Next, we average both sides over the ball $B(x, r)$, apply Fubini theorem to the right-hand side, and estimate the integrand using Lemma 3.2. Upon passing to the limit $r \to 0$, the theorem follows. $\Box$

In the sequel we need the following version of Hedberg's lemma [19].

Lemma 3.4 *If $\mu_1 \ge \mu_2 > 0$, then there exists a constant $C = C(n, \mu_1, \mu_2)$ such that, for all integrable u and all $x \in Q$,*

$$\int_Q \frac{|u(y)|}{|x - y|^{n-\mu_1}} dy \le C(\operatorname{diam} Q)^{\mu_2} M_{\operatorname{diam} Q}^{\mu_1 - \mu_2} u(x)\,.$$

Proof. Break the integral which stands in the left hand side into the sum of the integrals over "rings" $Q \cap \big(B(x, \operatorname{diam} Q/2^k) \setminus B(x, \operatorname{diam} Q/2^{k+1})\big)$, where $k = 0, 1, 2, \ldots$ In each "ring", we have $|x - y|^\mu \approx \big(\operatorname{diam} Q/2^k\big)^\mu$ for any exponent μ. Now, estimate the integral over the "ring" by the integral over the ball $B(x, \operatorname{diam} Q/2^k)$, note that this integral does not exceed the appropriate maximal function, and compute the sum of a geometric series to conclude the argument. $\Box$

In the remaining part of the paper we assume that the values of all functions $f \in W^{m,p}_{\mathrm{loc}}$ and all their derivatives of order less than or equal to m are defined at *every* point x by formula (13) i.e.

$$D^\alpha f(x) := \limsup_{r\to 0} \fint_{B(x,r)} D^\alpha f(y)\, dy\,, \qquad 0 \le |\alpha| \le m\,. \tag{16}$$

We now turn to inequalities satisfied by the difference of $f(y)$ and the *non-averaged* Taylor polynomial, $T^s_x f(y)$, for $s \in \{m-1, m\}$.

Take $f \in W^{m,1}_{\mathrm{loc}}(\mathbb{R}^n)$. Fix $x, y \in \mathbb{R}^n$ and a cube Q containing both these points. By triangle inequality,

$$|f(y) - T^{m-1}_x f(y)| \le |f(y) - T^{m-1}_Q f(y)| + |T^{m-1}_Q f(y) - T^{m-1}_x f(y)|\,. \tag{17}$$

The first term on the right hand side can be estimated by a direct application of Theorem 3.3 with $S = Q$. To estimate the second one, note an obvious fact: Taylor's expansion up to order r of any polynomial P of degree r is identically equal to P. Hence, we have the identity

$$T^{m-1}_Q f(y) \equiv \sum_{|\alpha|\le m-1} D^\alpha T^{m-1}_Q f(x) \frac{(y-x)^\alpha}{\alpha!}\,.$$

Since $D^\alpha T^{m-1}_Q f(x) = T^{m-1-|\alpha|}_Q D^\alpha f(x)$, we can write

$$\begin{aligned} &|T^{m-1}_Q f(y) - T^{m-1}_x f(y)| \\ &= \left| \sum_{|\alpha|\le m-1} \Big(D^\alpha f(x) - T^{m-1-|\alpha|}_Q D^\alpha f(x) \Big) \frac{(y-x)^\alpha}{\alpha!} \right| \\ &\le \sum_{|\alpha|\le m-1} (\operatorname{diam} Q)^{|\alpha|} |D^\alpha f(x) - T^{m-1-|\alpha|}_Q D^\alpha f(x)|\,. \end{aligned} \tag{18}$$

Now, Theorem 3.3 can be employed to estimate all terms on the right hand side. To shorten the notation, let, for $\gamma > 0$,

$$I^\gamma_Q g(x) := \int_Q \frac{g(y)}{|x-y|^{n-\gamma}}\, dy$$

denote the local Riesz potential of a function g. By inequality (14) of Theorem 3.3 with $S = Q$, we have

$$|D^\alpha f(x) - T^{m-1-|\alpha|}_Q D^\alpha f(x)| \le C I^{m-|\alpha|}_Q |\nabla^m f|(x)\,,$$

and, since $I^\gamma_Q g(x) \le (\operatorname{diam} Q)^{\gamma-1} I^1_Q g(x)$ for any $\gamma \ge 1$, any nonnegative function g, and all $x \in Q$, we finally arrive at the following result.

Theorem 3.5 *Assume that $f \in W^{m,1}_{\rm loc}(\mathbb{R}^n)$ has the derivatives $D^\alpha f$ defined pointwise by formula (16). Then there exists a constant $C = C(n,m)$ such that for any cube Q, and all $x, y \in Q$, we have*

$$|f(y) - T^{m-1}_x f(y)| \leq C(\operatorname{diam} Q)^{m-1}(I^1_Q|\nabla^m f|(x) + I^1_Q|\nabla^m f|(y))\,. \qquad (19)$$

Remark. If for some points $x, y \in Q$ we have an indefinite expression like e.g. $|\infty - \infty + \ldots|$ on the left hand side of (19), then we assume that the left hand side equals infinity for those x and y. Hence, the left hand side of inequality (19) is always well defined, and the inequality holds true for all $x, y \in Q$. In the indefinite case the inequality follows from the fact that if $|D^\alpha f(z)| = \infty$ for some $|\alpha| \leq m-1$ and $z \in Q$, then $I^1_Q|\nabla^m f|(z) = \infty$.

Theorem 3.6 *Let $\lambda \in (0,1]$. For $f \in W^{m,1}_{\rm loc}(\mathbb{R}^n)$ with all derivatives defined by (16), and for any $x \neq y \in \mathbb{R}^n$ we have*

$$\frac{|f(y) - T^{m-1}_x f(y)|}{|x-y|^{m-1}} \leq C|x-y|^\lambda(M^{1-\lambda}_{|x-y|}|\nabla^m f|(x) + M^{1-\lambda}_{|x-y|}|\nabla^m f|(y))\,. \qquad (20)$$

Remark. A comment similar to that following Theorem 3.5 applies here. We assume $x \neq y$ to avoid 0 in the denominator.

Proof. Apply Theorem 3.5 for an arbitrary cube Q which contains both points x and y. Next, use Lemma 3.4 with $\mu_1 := 1$ and $\mu_2 := \lambda$ to estimate Riesz potentials by maximal functions. Finally, take infimum over $\operatorname{diam} Q$ and observe that $\inf \operatorname{diam} Q = |x-y|$ (rotate the cube if necessary!). $\square$

Writing down the inequality of Theorem 3.6 with f replaced by $D^\alpha f$, we immediately obtain the following.

Corollary 3.7 *Let $\lambda \in (0,1]$. For $f \in W^{m,1}_{\rm loc}(\mathbb{R}^n)$ with all derivatives defined by (16), and for any $x \neq y \in \mathbb{R}^n$ we have*

$$\begin{aligned} &\frac{|D^\alpha f(y) - T^{m-1-|\alpha|}_x D^\alpha f(y)|}{|x-y|^{m-1-|\alpha|}} \\ &\quad \leq \quad C|x-y|^\lambda(M^{1-\lambda}_{|x-y|}|\nabla^m f|(x) + M^{1-\lambda}_{|x-y|}|\nabla^m f|(y)) \end{aligned} \qquad (21)$$

for every α with $|\alpha| \leq m-1$.

Remark. Ineqality (21) in the particular case of first order derivatives has been employed in [15] in the study of boundary behaviour of conformal and quasiconformal mappings.

Note that in order to obtain the last corollary and the previous two theorems, we have used only the first inequality of Theorem 3.3. The second one can be used to produce a pointwise estimate which resembles Theorem 3.6, with flat maximal functions $M^\flat$ appearing on the right hand side of the inequality. To this end, one estimates $|f(y)-T_x^m f(y)|$ as in the proof of Theorem 3.5, with one slight change: all the terms containing m-th order derivatives of f have to be estimated directly, without resorting to Theorem 3.3. This computation yields

$$\begin{aligned} |f(y)-T_x^m f(y)| \quad &\leq \quad C(\operatorname{diam} Q)^{m-1}\Big(I_Q^1|\nabla^m f - a|(x) + I_Q^1|\nabla^m f - b|(y)\Big) \\ &\qquad + C|x-y|^m \fint_Q |\nabla^m f(z) - \nabla^m f(x)|\, dz\,, \end{aligned} \tag{22}$$

where $a=(a_\alpha)_{|\alpha|=m}$ and $b=(b_\alpha)_{|\alpha|=m}$ are arbitrary constant vectors, and Q is an arbitrary cube which contains x and y. Next, applying Lemma 3.4 with $\mu_1=\mu_2=1$ and putting $a=\nabla^m f(x)$, $b=\nabla^m f(y)$, one obtains the following theorem and its obvious corollary.

Theorem 3.8 *For $f\in W^{m,1}_{\rm loc}(\mathbb{R}^n)$ with all derivatives defined by (16), and for any $x\neq y\in\mathbb{R}^n$ we have*

$$\frac{|f(y)-T_x^m f(y)|}{|x-y|^m} \leq C\Big(M^\flat_{|x-y|}(\nabla^m f)(x) + M^\flat_{|x-y|}(\nabla^m f)(y)\Big)\,. \tag{23}$$

Corollary 3.9 *For $f\in W^{m,1}_{\rm loc}(\mathbb{R}^n)$ with all derivatives defined by (16), and for any $x\neq y\in\mathbb{R}^n$ we have*

$$\frac{|D^\alpha f(y) - T_x^{m-|\alpha|}D^\alpha f(y)|}{|x-y|^{m-|\alpha|}} \leq C\Big(M^\flat_{|x-y|}(\nabla^m f)(x) + M^\flat_{|x-y|}(\nabla^m f)(y)\Big) \tag{24}$$

for every α with $|\alpha|\leq m$.

Remark. A verion of this inequality was proved in the monograph [4, Corollary 5.8] by Bennett and Sharpley and applied in the interpolation theory. Namely the ineqality was employed in a simplified proof (basically due to Calderón and Milman [8]) of DeVore and Scherer's theorem which gives an explicit formula for the so called K-functional for the couple $(W^{k,1},W^{k,\infty})$ of Sobolev spaces.

The next two colloraries are well known. The first one a direct consequence of Theorem 3.3 and the Hardy-Littlewood-Sobolev theorem for Riesz potentials.

Corollary 3.10 *If $f\in W^{m,p}_{\rm loc}(\mathbb{R}^n)$ and $p>1$, $mp<n$, then the inequality*

$$\left(\fint_Q |f(x)-T_Q^{m-1}f(x)|^{p^*}\,dx\right)^{1/p^*} \leq C(\operatorname{diam} Q)^m \left(\fint_Q |\nabla^m f(y)|^p dy\right)^{1/p}$$

holds for each cube $Q\subset\mathbb{R}^n$ with the constant C depending on m,n,p only.

A weaker version of this corollary, a Poincaré-type inequality, is valid for all $p \geq 1$ in any dimension n.

Corollary 3.11 *If $f \in W^{m,p}_{\mathrm{loc}}(\mathbb{R}^n)$, then for each cube $Q \subset \mathbb{R}^n$ we have*

$$\left(\fint_Q |f(x) - T_Q^{m-1} f(x)|^p dx \right)^{1/p} \leq C(\operatorname{diam} Q)^m \left(\fint_Q |\nabla^m f(y)|^p dy \right)^{1/p},$$

with the constant C depending only on m, n, and p.

4 Proof of Theorem 1.4

In this section, we present a detailed proof of Theorem 1.4. We begin with a variant of Theorem 1.4 which holds for functions with compact support. The general case can be then easily obtained by a standard partition of unity argument, as in the proof of Meyers-Serrin theorem.

Theorem 4.1 *Assume that $u \in W^{k,p}(\mathbb{R}^n)$, where $1 < p < \infty$, has compact support contained in some cube Q, having the edge of unit length. Fix $m \in \{0, 1, \ldots, k-1\}$ and $\lambda \in (0,1)$. Then, for any $\varepsilon > 0$ there exists a closed set $F \subset \mathbb{R}^n$ and a function $w \in C^{m,\lambda}(\mathbb{R}^n) \cap W^{m+1,p}(\mathbb{R}^n)$ with compact support contained in $3Q$ such that*

(i) $B_{k-m-\lambda,p}(\mathbb{R}^n \setminus F) < \varepsilon$;

(ii) *In the particular case when $k = m+1$ and $(1-\lambda)p \leq n$ we also have $\mathcal{H}^{n-(1-\lambda)p}_\infty(\mathbb{R}^n \setminus F) < \varepsilon$;*

(iii) *$D^\alpha u(x) = D^\alpha w(x)$ for all $x \in F$ and all α with $|\alpha| \leq m$;*

(iv) $\|u - w\|_{W^{m+1,p}(\mathbb{R}^n)} \leq C\|u\|_{W^{m+1,p}(\mathbb{R}^n \setminus F)}$;

(v) *For any α with $|\alpha| = m$, the modulus of continutity of $D^\alpha w$ goes to zero faster than t^λ, i.e.*

$$\lim_{\varrho \to 0} \left(\sup_{\substack{x \neq y \\ |x-y| \leq \varrho}} \frac{|D^\alpha w(x) - D^\alpha w(y)|}{|x-y|^\lambda} \right) = 0\,.$$

Remarks. (1) As usual we assume that u and all its distributional derivatives up to order k are defined *everywhere* by the formula

$$D^\alpha u(x) = \limsup_{r \to 0} \fint_{B(x,r)} D^\alpha u(y)\, dy\,, \qquad 0 \leq |\alpha| \leq k.$$

(2) The complement of F has small Lebesgue measure (since it has small capacity). Therefore, the fourth condition of Theorem 4.1 implies that in fact the norm $\|u - w\|_{W^{m+1,p}(\mathbb{R}^n)}$ can be made arbitrarily small.

Proof. For a fixed $\varepsilon > 0$, we select a closed set F satisfying the two following conditions:

$$M_1^{1-\lambda}|\nabla^{m+1}u|(x) \le t = t(\varepsilon) \qquad \text{for all } x \in F, \tag{25}$$

$$M_\varrho^{1-\lambda}|\nabla^{m+1}u| \overset{\varrho=0}{\longrightarrow} 0 \qquad \text{uniformly on } F. \tag{26}$$

We choose the number $t = t(\varepsilon)$ in (25) sufficiently large to have $B_{k-m-\lambda,p}(\mathbb{R}^n \setminus F) < \varepsilon$. By the results of Section 2.2, this is always possible. With no loss of generality one can assume that (ii) is also satisfied.

Moreover, when $(k - m - \lambda)p > n$, Corollary 2.4 implies that we can simply take $F = \mathbb{R}^n$ in (25) and (26). In the latter case $M_1^{1-\lambda}|\nabla^{m+1}u|$ is a bounded function on $\mathbb{R}^n$. It follows then from (25), (26) and Corollary 3.7 that the function u coincides with a $C^{m,\lambda}$ function which, in addition, satisfies (v) and there is nothing more to prove.

Hence, from now on we suppose that $(k - m - \lambda)p \le n$ and that $\mathbb{R}^n \setminus F$ is nonempty.

Since u has compact support, the set $\mathbb{R}^n \setminus F$ is bounded. In fact, it is contained in $3Q$, since $M_1^{1-\lambda}|\nabla^{m+1}u|(x) \equiv 0$ for $x \notin 3Q$.

Now, take the Whitney cube decomposition of $\mathbb{R}^n \setminus F$,

$$\mathbb{R}^n \setminus F = \bigcup_{i \in I} Q_i\,,$$

where all the cubes Q_i are dyadic, and select an associated smooth partition of unity $\{\varphi_i\}_{i \in I}$. Recall the standard conditions satisfied by Q_i and φ_i:

(i) $\operatorname{dist}(2Q_i, F) \le \operatorname{diam} 2Q_i \le 4\operatorname{dist}(2Q_i, F)$;

(ii) Every point of $\mathbb{R}^n \setminus F$ is covered by at most $C(n) = 4^n$ different cubes $2Q_i$;

(iii) For each $i \in I$, $\operatorname{supp} \varphi_i \subset 2Q_i \subset \mathbb{R}^n \setminus F$;

(iv) $\sum_{i \in I} \varphi_i(x) \equiv 1$ on $\mathbb{R}^n \setminus F$, and, for every $i \in I$, and every α, $|D^\alpha \varphi_i| \le C_\alpha (\operatorname{diam} Q_i)^{-|\alpha|}$.

To leave u unchanged on F and make it sufficiently smooth on the whole space (without changing its Sobolev norm too much), we introduce the function w defined by

$$w(x) = \begin{cases} u(x) & \text{for } x \in F, \\ \displaystyle\sum_{i \in I} \varphi_i(x) T^m_{2Q_i} u(x) & \text{for } x \in \mathbb{R}^n \setminus F. \end{cases} \tag{27}$$

Here, as before,

$$T^m_{2Q_i}u(x) = \fint_{2Q_i} \left(\sum_{|\alpha|\le m} D^\alpha u(z) \frac{(x-z)^\alpha}{\alpha!} \right) dz \,.$$

Remark. To construct the classical Whitney extension (which, up to now, has been used in all proofs of Michael and Ziemer's theorem), one takes in the above formula $T^m_{a_i}u(x)$ instead of $T^m_{2Q_i}u(x)$, with $a_i \in F$ minimizing the distance from F to $\operatorname{supp}\varphi_i$. In contrast with Whitney's extension theorem, the function w in (27) is not an extension of u from F, but it is defined by taking a suitable smooth approximation of $u|_{\mathbb{R}^n\setminus F}$.

The rest of the proof will be divided into three independent parts. Each of them is contained in a seperate subsection. First, we prove that the Sobolev norm of w can be controlled on $\mathbb{R}^n \setminus F$ by the Sobolev norm of u. Next, we prove that the extension w belongs in fact to $C^{m,\lambda}$ and that its modulus of continuity decreases at 0 faster than t^λ. This is the most tedious (and longest) part of the proof. Finally, in the last part, we prove that the extension belongs to the appropriate Sobolev space on the whole of $\mathbb{R}^n$.

4.1 Sobolev norm estimates

We claim that

$$\|w\|_{W^{m+1,p}(\mathbb{R}^n\setminus F)} \le C\|u\|_{W^{m+1,p}(\mathbb{R}^n\setminus F)} \,. \tag{28}$$

Since $w \in C^\infty(\mathbb{R}^n\setminus F)$ it suffices to compute derivatives of w of order $|\alpha| \le m+1$ and estimate their L^p norms. To check this, fix α with $|\alpha| \le m+1$, and apply Leibniz formula to obtain

$$D^\alpha \left(\sum_{i\in I} \varphi_i(x) T^m_{2Q_i}u(x) \right) = \sum_{\substack{\beta+\gamma=\alpha \\ |\gamma|\le m}} \binom{\alpha}{\beta} S_{\beta,\gamma}(x) \,,$$

where

$$S_{\beta,\gamma}(x) = \sum_{i\in I} D^\beta\varphi_i(x) T^{m-|\gamma|}_{2Q_i} D^\gamma u(x) \,. \tag{29}$$

To estimate the L^p norm of $S_{\beta,\gamma}$, we shall consider separately the cases $|\beta| = 0$ and $|\beta| \ne 0$.

Case 1. If $|\beta| = 0$, then necessarily $\gamma = \alpha$ and $|\alpha| \le m$ (otherwise $S_{\beta,\gamma}$ is identically equal to zero). Since each point in $\mathbb{R}^n \setminus F$ belongs to at most $C(n)$ cubes $2Q_i$, we have

$$\int_{\mathbb{R}^n\setminus F} |S_{0,\alpha}(x)|^p \, dx \quad \le \quad C \sum_{i\in I} \int_{2Q_i} \left| T^{m-|\alpha|}_{2Q_i} D^\alpha u(x) \right|^p dx$$

$$
\begin{aligned}
&\leq C \sum_{i\in I} \int_{2Q_i} \sum_{l=0}^{m-|\alpha|} |\nabla^{l+|\alpha|} u(z)|^p \left(\fint_{2Q_i} |x-z|^{lp}\, dx \right) dz \\
&\leq C \sum_{i\in I} \sum_{l=0}^{m-|\alpha|} (\operatorname{diam} Q_i)^{lp} \int_{2Q_i} |\nabla^{l+|\alpha|} u(z)|^p\, dz \\
&\leq C \|u\|_{W^{m+1,p}(\mathbb{R}^n\setminus F)} .
\end{aligned}
$$

In the last inequality, we have used the fact that the diameters of Q_i are uniformly bounded — this follows from the assumption $\operatorname{supp} u \subset Q$, which forces the boundedness of $\mathbb{R}^n \setminus F$.

Case 2. For β of nonzero length, we have $\sum_{i\in I} D^\beta \varphi_i(x) \equiv 0$ on $\mathbb{R}^n \setminus F$, and one can write

$$
-S_{\beta,\gamma}(x) = \sum_{i\in I} D^\beta \varphi_i(x) \left(D^\gamma u(x) - T^{m-|\gamma|}_{2Q_i} D^\gamma u(x) \right) .
$$

Now, by the property (iv) of Whitney decomposition and its associated partition of unity, for every $x \in \mathbb{R}^n \setminus F$ we have

$$
|S_{\beta,\gamma}(x)| \leq \sum_{i\in I} (\operatorname{diam} Q_i)^{-|\beta|} |D^\gamma u(x) - T^{m-|\gamma|}_{2Q_i} D^\gamma u(x)| \chi_{2Q_i}(x) .
$$

The sum on the right hand side contains at most 4^n nonzero terms (recall that the cubes $2Q_i$ do not overlap "too much"). Therefore,

$$
|S_{\beta,\gamma}(x)|^p \leq C \sum_{i\in I} (\operatorname{diam} Q_i)^{-|\beta|p} |D^\gamma u(x) - T^{m-|\gamma|}_{2Q_i} D^\gamma u(x)|^p \chi_{2Q_i}(x) .
$$

Hence, by Poincaré inequality (see Corollary 3.11 in Section 3), keeping in mind that $\beta + \gamma = \alpha$, we obtain

$$
\begin{aligned}
&\int_{\mathbb{R}^n\setminus F} |S_{\beta,\gamma}(x)|^p\, dx \\
&\quad\leq C \sum_{i\in I} (\operatorname{diam} Q_i)^{-|\beta|p} \int_{2Q_i} |D^\gamma u(x) - T^{m-|\gamma|}_{2Q_i} D^\gamma u(x)|^p\, dx \\
&\quad\leq C \sum_{i\in I} (\operatorname{diam} Q_i)^{-|\beta|p+(m-|\gamma|+1)p} \int_{2Q_i} |\nabla^{m-|\gamma|+1} D^\gamma u(x)|^p\, dx \\
&\quad\leq C \sum_{i\in I} (\operatorname{diam} Q_i)^{(m+1-|\alpha|)p} \int_{2Q_i} |\nabla^{m+1} u(x)|^p\, dx \\
&\quad\leq C \int_{\mathbb{R}^n\setminus F} |\nabla^{m+1} u(x)|^p\, dx .
\end{aligned}
$$

As before, the last inequality uses the boundedness of the diameters of Q_i (note that the exponent $(m+1-|\alpha|)p$ is nonnegative).

This completes the proof of inequality (28), which in turn, once we know that $w \in W^{m+1,p}(\mathbb{R}^n)$, implies the condition (iv) of Theorem 4.1.

4.2 Smoothness of the extension

We claim that

$$w \in C^{m,\lambda}(\mathbb{R}^n) \tag{30}$$

and moreover, for any β with $|\beta| = m$,

$$\lim_{\varrho \to 0} \left(\sup_{\substack{x \neq y \\ |x-y| \le \varrho}} \frac{|D^\alpha w(x) - D^\alpha w(y)|}{|x-y|^\lambda} \right) = 0\,. \tag{31}$$

Set, for $|\alpha| \le m$,

$$\widetilde{w}^\alpha(x) = \begin{cases} D^\alpha u(x) & \text{for } x \in F, \\ D^\alpha w(x) & \text{for } x \in \mathbb{R}^n \setminus F. \end{cases} \tag{32}$$

As before, let Q denote a cube with edge of unit length such that $\operatorname{supp} u \subset Q$ and let $D^\alpha u(x)$ be defined at every point of F by formula (16). The following lemma is the key estimate of this part of the proof.

Lemma 4.2 *There exists a constant C such that for any α with $|\alpha| \le m$, any $x \in \mathbb{R}^n$, and any point $a \in F$ we have*

$$|\widetilde{w}^\alpha(x) - D^\alpha T_a^m u(x)| \le C\omega(|x-a|) \cdot |x-a|^{m-|\alpha|}\,, \tag{33}$$

where the function $\omega : [0,\infty) \to [0,\infty)$ is concave, increasing, $\omega(0) = 0$ and moreover

$$\lim_{t=0} \omega(t)/t^\lambda = 0.$$

To prove this lemma, one is forced to rewrite important parts of the proof of Whitney's extension theorem. Before we shall start the detailed and somewhat lengthy computations, let us explain briefly the rough idea. We express the difference $\widetilde{w}^\alpha(x) - D^\alpha T_a^m u(x)$ using Leibniz formula, as in the previous subsection, and estimate separately various terms of the resulting sum. One of the important steps is to estimate the difference between $D^\gamma u(b)$ and $T^r_{2Q_i} D^\gamma u(b)$, where b is a point of F "not too far from $2Q_i$" — we estimate this difference by a Riesz potential over a cube centered at b, with edge comparable with $\operatorname{diam}(2Q_i)$. Such a potential can be controlled by the value of an appropriate maximal function at b, which in turn does not exceed a certain constant (by the very definition of the set F).

Those readers who are not interested in all the details might skip Subsections 4.2.1 and 4.2.2 now, and jump directly to the next lemma.

4.2.1 Whitney jets and their properties

For the remaining readers and for the sake of completeness, we shall recall now the notion of a Whitney jet and other terminology which is usually employed to formulate and prove Whitney's extension theorem and which will be used in the proof of Lemma 4.2. To a large extent, our exposition follows [21, Chapter 1].

Let K be a compact set in $\mathbb{R}^n$. By a *jet of order* m on K we mean here a family $f = (f^\alpha)_{|\alpha|\le m}$ of continuous functions on K. The space of all jets is denoted by $J^m(K)$. We write $f(x) = f^0(x)$, and the "Taylor polynomial" of f is defined by the familiar formula

$$T_a^m f(x) = \sum_{|\alpha|\le m} f^\alpha(a)\frac{(x-a)^\alpha}{\alpha!}\,.$$

For a fixed $a \in K$ and $f \in J^m(K)$, $T_a^m f$ is a polynomial of variable $x \in \mathbb{R}^n$.

If $|\beta| \le m$, then $D^\beta\colon J^m(K) \to J^{m-|\beta|}(K)$ is a linear map defined by

$$D^\beta\colon (f^\alpha)_{|\alpha|\le m} \longmapsto (f^{\alpha+\beta})_{|\alpha|\le m-|\beta|}\,.$$

Any function $g \in C^m(\mathbb{R}^n)$ gives rise to the jet $J^m(g) = (D^\alpha g|_K)_{|\alpha|\le m}$. Here, D^α denotes the standard partial derivative. In all the following computations of this subsection, we shall identify $T_a^m f$ with the jet $J^m(T_a^m f)$. The formal *Taylor remainder* of f is defined by

$$R_a^m f := f - J^m(T_a^m f) \in J^m(K)\,.$$

Finally, by a *modulus of continuity* we mean any concave, increasing and continuous function $\omega : [0,\infty) \to [0,\infty)$ with $\omega(0) = 0$. A typical example is $\omega(s) = s^\lambda$ for some fixed $\lambda \in (0,1]$. The following theorem gives three equivalent versions of the inequalities which are satisfied by jets obtained from smooth function by restricting them to a compact set.

Proposition 4.3 *Let $f \in J^m(K)$. The following three conditions are equivalent:*

(i) $(R_x^m f)^\alpha(y) = \mathrm{o}(|x-y|^{m-|\alpha|})$ *for $x, y \in K$ and $|\alpha| \le m$, as $|x-y| \to 0$.*

(ii) *There exists a modulus of continuity ω such that*

$$|(R_x^m f)^\alpha(y)| \le \omega(|x-y|)\cdot|x-y|^{m-|\alpha|} \quad \textit{for } x,y \in K \textit{ and } |\alpha| \le m.$$

(iii) *There exists a modulus of continuity ω_1 such that*

$$|T_x^m f(z) - T_y^m f(z)| \le \omega_1(|x-y|)\cdot(|x-z|^m + |y-z|^m)$$

for $x, y \in K$, $z \in \mathbb{R}^n$.

Moreover, if (ii) holds, then we can choose $\omega_1 = C\omega$, and if (iii) holds, then we can choose $\omega = C\omega_1$ (in both cases C depends only on m and n).

Proof. See, e.g., [21, Chapter 1]. $\Box$

The space $\mathcal{E}^m(K)$ of *Whitney functions of class C^m*, or *Whitney jets of order m*, consists of those jets $f \in J^m(K)$ for which one of the equivalent conditions of the above proposition is satisfied. It is a Banach space with the norm

$$\|f\|_m^K = \sup_{\substack{x\in K\\ |\alpha|\le m}} |f^\alpha(x)| + \sup_{\substack{x,y\in K\\ x\neq y\\ |\alpha|\le m}} \frac{|(R_x^m f)^\alpha(y)|}{|x-y|^{m-|\alpha|}}\,.$$

This space of jets is closely connected to the famous Whitney extension theorem (see [31], [21]). Although we will not use this theorem in the paper we will formulate it to show how our constructions are related to that result. Whitney's theorem reads as follows.

Theorem 4.4 *Given a jet $f \in J^m(K)$, where $K \subset \mathbb{R}^n$ is compact. Then there exists a function $g \in C^m(\mathbb{R}^n)$ such that $J^m(g) = f$ if and only if $f \in \mathcal{E}^m(K)$.*

The necessity of being a Whitney jet is an obvious consequence of Taylor's formula, but the sufficiency is difficult.

Remark. In Section 4, we have seen that a Sobolev function $u \in W^{k,p}(\mathbb{R}^n)$ gives rise to the jet $(D^\alpha u)_{|\alpha|\le m}$ on an appropriate set F, which roughly speaking consists of those points where the fractional maximal function of $|\nabla^m u|$ was not too large. This F was closed but not compact; however, all the components of the jet given by u were identically zero outside a fixed cube. It follows from Corollary 3.7 and (26) that $(D^\alpha u)_{|\alpha|\le m}$ is a Whitney jet of order m and hence it follows from Whitney's theorem that $u|_F$ is a restriction of a $C^m(\mathbb{R}^n)$ function to F. However, we shall not use this observation in our proof.

4.2.2 Hölder estimates, part 1

We are now ready to give the

Proof of Lemma 4.2. First we consider an easy case when $x \in F$ and $a \in F$. Applying Corollary 3.7 we have

$$\begin{aligned}|\widetilde{w}^\alpha(x) - D^\alpha T_a^m u(x)| &= \left|D^\alpha u(x) - T_a^{m-|\alpha|} D^\alpha u(x)\right| \\ &\le C|x-a|^{m-|\alpha|+\lambda}\left(M_{|x-a|}^{1-\lambda}|\nabla^{m+1}u|(x) + M_{|x-a|}^{1-\lambda}|\nabla^{m+1}u|(a)\right). \quad (34)\end{aligned}$$

Now, set

$$\eta(t) := 2t^\lambda \cdot \sup_{\substack{z\in F\\ \varrho\le t}} M_\varrho^{1-\lambda}|\nabla^{m+1}u|(z)\,. \quad (35)$$

By the definition of F, we have $\eta(t) \le Ct^\lambda$ and $\eta(t)/t^\lambda \to 0$ as $t \to 0$. Moreover, η is bounded, continuous, increasing and $\eta(0) = 0$. It is an easy exercise to show that there is a continuous, increasing and concave function $\omega : [0,\infty) \to [0,\infty)$, such that $\eta(t) \le \omega(t) \le Ct^\lambda$ and $\omega(t)/t^\lambda \to 0$ as $t \to 0$ (i.e., ω is a modulus of continuity). Hence inequality (34) leads to

$$|\widetilde{w}^\alpha(x) - D^\alpha T_a^m u(x)| \le C\omega(|x-a|)|x-a|^{m-|\alpha|}\,.$$

Thus in what follows we may assume that $x \in \mathbb{R}^n \setminus F$. Since

$$T_a^m u(x) = \sum_{i\in I} \varphi_i(x) T_a^m u(x),$$

for any $x \in \mathbb{R}^n \setminus F$ we have

$$\begin{aligned} \widetilde{w}^\alpha(x) - D^\alpha T_a^m u(x) &= D^\alpha \left(\sum_{i\in I} \varphi_i(x) \left(T_{2Q_i}^m u(x) - T_a^m u(x)\right)\right) \\ &= \sum_{\beta+\gamma=\alpha} \binom{\alpha}{\beta} S_{\beta,\gamma}(x)\,, \end{aligned}$$

where

$$S_{\beta,\gamma}(x) = \sum_{i\in I} D^\beta \varphi_i(x) \left(T_{2Q_i}^{m-|\gamma|} D^\gamma u(x) - T_a^{m-|\gamma|} D^\gamma u(x)\right)\,.$$

If $|\beta| > 0$, then $\sum_{i\in I} D^\beta \varphi_i(x) \equiv 0$ and hence

$$S_{\beta,\gamma}(x) = \sum_{i\in I} D^\beta \varphi_i(x) \left(T_{2Q_i}^{m-|\gamma|} D^\gamma u(x) - T_b^{m-|\gamma|} D^\gamma u(x)\right)$$

for any point b. Let $K = 5Q \cap F$, where Q is the unit cube containing $\operatorname{supp} u$. Choose $b \in K$ such that $|x-b| = \operatorname{dist}(x,K)$ $(= \operatorname{dist}(x,F))$. With

$$b_\beta = \begin{cases} a & \text{if } |\beta| = 0, \\ b & \text{if } |\beta| > 0 \end{cases}$$

we may write

$$S_{\beta,\gamma}(x) = \sum_{i\in I} D^\beta \varphi_i(x) \left(T_{2Q_i}^{m-|\gamma|} D^\gamma u(x) - T_{b_\beta}^{m-|\gamma|} D^\gamma u(x)\right)\,.$$

Let $x \in \mathbb{R}^n \setminus F = 5Q \setminus F$ and $a \in K$. For each $i \in I$, choose a point $b_i \in K$ such that

$$\operatorname{dist}(b_i, 2Q_i) = \operatorname{dist}(K, 2Q_i)\,.$$

Observe that $\operatorname{dist}(K, 2Q_i) = \operatorname{dist}(F, 2Q_i)$. By the triangle inequality,

$$\begin{aligned}\left|T_{2Q_i}^{m-|\gamma|} D^\gamma u(x) - T_{b_\beta}^{m-|\gamma|} D^\gamma u(x)\right| &\leq \left|T_{2Q_i}^{m-|\gamma|} D^\gamma u(x) - T_{b_i}^{m-|\gamma|} D^\gamma u(x)\right| \\ &+ \left|T_{b_i}^{m-|\gamma|} D^\gamma u(x) - T_{b_\beta}^{m-|\gamma|} D^\gamma u(x)\right| \\ &\equiv H_i(x) + J_i(x)\,.\end{aligned}$$

We will estimate $H_i(x)$ and $J_i(x)$ for $x \in 2Q_i$. We first estimate the second term. Let $a, y \in K$. Take an arbitrary multiindex μ with $|\mu| \leq m - |\gamma|$. We apply Corollary 3.7 to estimate the (formal) Taylor remainder of $D^\gamma u$; this gives

$$\begin{aligned}|(R_a^{m-|\gamma|} D^\gamma u)^\mu(y)| &\equiv \left|D^\mu \left(D^\gamma u(y) - T_a^{m-|\gamma|} D^\gamma u(y)\right)\right| \\ &\leq C|a-y|^{m-(|\gamma|+|\mu|)+\lambda} \left(M_{|a-y|}^{1-\lambda}|\nabla^{m+1} u|(a) + M_{|a-y|}^{1-\lambda}|\nabla^{m+1} u|(y)\right) \\ &\leq C\omega(|a-y|)|a-y|^{m-(|\gamma|+|\mu|)}\,.\end{aligned}$$

This is the condition (ii) in Lemma 4.3. Hence by Lemma 4.3, condition (iii) must also be satisfied. Therefore

$$J_i(x) \leq C\omega(|b_\beta - b_i|)\left(|x-b_\beta|^{m-|\gamma|} + |x-b_i|^{m-|\gamma|}\right)$$

for any $x \in \mathbb{R}^n \setminus F$. Let $x \in 2Q_i$. We estimate $J_i(x)$ when $|\beta| > 0$ and $|\beta| = 0$ separately. Put $d_i = \operatorname{diam} Q_i$. Assume first that $|\beta| > 0$. Since $|b - b_i| \leq |x - b| + |x - b_i|$ and by the properies of Whitney cubes Q_i we have $|x - b| \approx d_i$, $|x - b_i| \approx d_i$, we conclude that $|b - b_i| \leq C(n)d_i$. Hence, invoking concavity of ω and equality $b_\beta = b$, we obtain

$$J_i(x) \leq C\omega(d_i)d_i^{m-|\gamma|} \qquad \text{for } x \in 2Q_i. \tag{36}$$

When $\beta = 0$ we have $\gamma = \alpha$ and hence

$$J_i(x) \leq C\omega(|a-b_i|)\left(|x-a|^{m-|\gamma|} + |x-b_i|^{m-|\gamma|}\right)$$

For $x \in 2Q_i$, the choice of b_i implies that

$$|x - b_i| \approx d_i \approx \operatorname{dist}(x, F) \leq |x - a|$$

and

$$|a - b_i| \leq |a - x| + |x - b_i| \leq C|x - a|\,.$$

Hence employing concavity of ω we conclude

$$J_i(x) \leq C\omega(|x-a|)|x-a|^{m-|\alpha|} \tag{37}$$

for $x \in 2Q_i$. One can check that it is possible to take here a constant C which depends on n and m only.

Note that since $d_i \le C|x-a|$ and $m - |\alpha| \le m - |\gamma|$, estimate (37) holds also in the case $|\beta| > 0$ by (36) and concavity of ω, but then it is weaker than (36).

To obtain a similar estimate for $H_i(x)$, observe that any polynomial f of degree s is identical to its Taylor polynomial $T^s_a f$ for any choice of a. An application of this fact to the polynomial

$$T^{m-|\gamma|}_{2Q_i} D^\gamma u(x) - T^{m-|\gamma|}_{b_i} D^\gamma u(x)$$

yields the estimate

$$H_i(x) \le \sum_{|\eta| \le m-|\gamma|} \left| T^{m-(|\gamma|+|\eta|)}_{2Q_i} D^{\gamma+\eta} u(b_i) - D^{\gamma+\eta} u(b_i) \right| \frac{|x-b_i|^{|\eta|}}{\eta!} \,.$$

Denote by $\widetilde{Q}_i$ the smallest cube centered at b_i, such that $2Q_i \subset \widetilde{Q}_i$. It easily follows that $\operatorname{diam} \widetilde{Q}_i \approx C d_i$. Now inequality (14), Hedberg's lemma (Lemma 3.4) and concavity of ω yield

$$\begin{aligned}
&\left| D^{\gamma+\eta} u(b_i) - T^{m-(|\gamma|+|\eta|)}_{2Q_i} D^{\gamma+\eta} u(b_i) \right| \\
&\le \; C \frac{|\widetilde{Q}_i|}{|2Q_i|} \int_{\widetilde{Q}_i} \frac{|\nabla^{m+1} u(y)|}{|b_i - y|^{n-(m-(|\gamma|+|\eta|)+1)}} \, dy \\
&\le \; C(\operatorname{diam} \widetilde{Q}_i)^{m-(|\gamma|+|\eta|)+\lambda} M^{1-\lambda}_{\operatorname{diam} \widetilde{Q}_i} |\nabla^{m+1} u|(b_i) \\
&\le \; C(\operatorname{diam} \widetilde{Q}_i)^{m-(|\gamma|+|\eta|)} \omega(\operatorname{diam} \widetilde{Q}_i) \le C(n,m) d_i^{m-(|\gamma|+|\eta|)} \omega(d_i) \,.
\end{aligned}$$

Collecting all the above estimates we arrive at the inequality

$$\begin{aligned}
H_i(x) \;&\le\; C(n,m) \cdot \sum_{|\eta| \le m-|\gamma|} \omega(d_i) \cdot d_i^{m-(|\gamma|+|\eta|)} \cdot d_i^{|\eta|} \\
&=\; C(n,m) \omega(d_i) d_i^{m-|\gamma|} \,,
\end{aligned} \tag{38}$$

valid for $x \in 2Q_i$. Since in this case we have $d_i \approx |x - b_i| \le C|x-a|$, and every point x in $\mathbb{R}^n \setminus F$ belongs to at most $C(n)$ different cubes $2Q_i$, we finally obtain

$$\begin{aligned}
|S_{\beta,\gamma}(x)| &\le \sum_{i \in I} |D^\beta \varphi_i(x)| (H_i(x) + J_i(x)) \\
&\le C \sum_{i \in I} d_i^{-|\beta|} \left(\omega(d_i) d_i^{m-|\gamma|} + J_i(x) \right) \chi_{2Q_i}(x) \le C \omega(|x-a|) |x-a|^{m-|\alpha|} \,,
\end{aligned}$$

and therefore $|\widetilde{w}^\alpha(x) - D^\alpha T_a^m u(x)| \le C\omega(|x-a|)|x-a|^{m-|\alpha|}$ for $x \in \mathbb{R}^n \setminus F$ and $a \in K$.

This inequality implies, in particular, that $\widetilde{w}^\alpha$ is a bounded function on $\mathbb{R}^n \setminus F \subset 3Q$. On the other hand $\omega(|x-a|)|x-a|^{m-|\alpha|} \ge C > 0$ for all $x \in \mathbb{R}^n \setminus F$ and $a \in F \setminus 5Q$. Thus in the remaining case $x \in \mathbb{R}^n \setminus F$ and $a \in F \setminus 5Q$ we also have

$$|\widetilde{w}^\alpha(x) - D^\alpha T_a^m u(x)| = |\widetilde{w}^\alpha(x)| \le C < \widetilde{C}\omega(|x-a|)|x-a|^{m-|\alpha|}\,.$$

This concludes the proof of Lemma 4.2. □

The following result is a direct consequence of the above proof. We will need it in the sequel.

Corollary 4.5 *Let $x \in 2Q_i$ and let $b \in F$ be such that $|x-b| = \operatorname{dist}(x,F)$. Then for any multiindex γ with $|\gamma| \le m$ we have*

$$\left|T_{2Q_i}^{m-|\gamma|} D^\gamma u(x) - T_b^{m-|\gamma|} D^\gamma u(x)\right| \le C\omega(d_i) d_i^{m-|\gamma|}\,,$$

where $d_i = \operatorname{diam} Q_i$.

4.2.3 Hölder estimates, part 2

Let us now show that Lemma 4.2 implies smoothness of w, or, more precisely, both claims, (30) and (31), stated at the beginning of Subsection 4.2.

Lemma 4.6 *The function w defined by (27), i.e.*

$$w(x) = \begin{cases} u(x) & \text{for } x \in F, \\ \displaystyle\sum_{i\in I} \varphi_i(x) T_{2Q_i}^m u(x) & \text{for } x \in \mathbb{R}^n \setminus F \end{cases}$$

is m times differentiable in the classical sense. Its derivatives of order m are Hölder continuous with exponent λ. Moreover, the modulus of continuity of m-th order derivatives behaves like $\mathrm{o}(t^\lambda)$ for $t \to 0$.

Proof. We set $\delta_j = (0,\dots,0,1,0,\dots,0)$ (the j-th component is equal to 1). Recall the notation

$$\widetilde{w}^\alpha(x) = \begin{cases} D^\alpha u(x) & \text{for } x \in F, \\ D^\alpha w(x) & \text{for } x \in \mathbb{R}^n \setminus F. \end{cases}$$

For $a \in F$, $x \in \mathbb{R}^n$, and $|\alpha| < m$ we have, by the triangle inequality,

$$\left|\widetilde{w}^\alpha(x) - \widetilde{w}^\alpha(a) - \sum_{j=1}^n (x_j - a_j) D^{\alpha+\delta_j} u(a)\right|$$

$$
\begin{aligned}
&\leq \quad |\widetilde{w}^\alpha(x) - D^\alpha T_a^m u(x)| + \left| \sum_{2\leq|\gamma|\leq m-|\alpha|} D^{\alpha+\gamma}u(a)\frac{(x-a)^\gamma}{\gamma!} \right| \\
&= \quad \mathrm{o}(|x-a|) \qquad \text{for } x \to a.
\end{aligned}
$$

Thus, $\widetilde{w}^\alpha$ is differentiable in F and

$$
\frac{\partial \widetilde{w}^\alpha}{\partial x_j}(a) = \widetilde{w}^{\alpha+\delta_j}(a) \qquad \text{for any } a \in F\,.
$$

As w is obviously (infinitely) smooth on $\mathbb{R}^n \setminus F$, this proves, by a simple induction, that w is m–times differentiable everywhere in $\mathbb{R}^n$ and $D^\alpha w = \widetilde{w}^\alpha$ in $\mathbb{R}^n$ for all $|\alpha| \leq m$.

We now show, that inequality (33) of Lemma 4.2 implies also Hölder continuity of the highest order derivatives of w, including appropriate estimates of their modulus of continuity.

Note first that for $|\alpha| = m$ inequality (33) takes the form

$$
|D^\alpha w(x) - D^\alpha w(a)| \leq C\omega(|x-a|) \qquad \text{for all } a \in F \text{ and all } x \in \mathbb{R}^n, \tag{39}
$$

with $\omega(s) = g(s)s^\lambda$, where $g(s) \to 0$ as $s \to 0$. Hence, it is enough to estimate $|D^\alpha w(x) - D^\alpha w(y)|$ for $x, y \notin F$. Consider now two cases.

Case 1. Assume that $\operatorname{dist}(x,F) \leq 2|x-y|$. Pick $a, b \in F$ such that $\operatorname{dist}(x,F) = |x-a|$ and $\operatorname{dist}(y,F) = |y-b|$. Then,

$$
|y-b| \leq |y-a| \leq 3|x-y|\,,
$$

and

$$
|a-b| \leq |a-y| + |y-b| \leq 6|x-y|\,.
$$

Hence, by (39), using concavity of ω, we have

$$
\begin{aligned}
|D^\alpha w(x) - D^\alpha w(y)| \quad &\leq \quad |D^\alpha w(x) - D^\alpha w(a)| + |D^\alpha w(a) - D^\alpha w(b)| \\
&\qquad + |D^\alpha w(b) - D^\alpha w(y)| \\
&\leq \quad C\omega(|x-a|) + \omega(|a-b|) + \omega(|b-y|)) \leq C\omega(|x-y|).
\end{aligned}
$$

Case 2. Assume that $\operatorname{dist}(x,F) > 2|x-y|$. We shall use the estimate

$$
|D^\alpha w(x)| \leq C\omega(\operatorname{dist}(x,F))\operatorname{dist}(x,F)^{-1} \qquad \text{for } |\alpha| = m+1 \text{ and } x \notin F, \tag{40}
$$

where ω denotes the modulus of continuity introduced Lemma 4.2. To verify this estimate, fix α with $|\alpha| = m+1$ and write

$$
D^\alpha w(x) = \sum_{\substack{\beta+\gamma=\alpha \\ |\beta|\geq 1}} \binom{\alpha}{\beta} S_{\beta,\gamma}(x)\,,
$$

where

$$\begin{aligned}
S_{\beta,\gamma}(x) &= \sum_{i\in I} D^\beta\varphi_i(x) D^\gamma T^m_{2Q_i}u(x) \\
&= \sum_{i\in I} D^\beta\varphi_i(x)\left(T^{m-|\gamma|}_{2Q_i}D^\gamma u(x) - T^{m-|\gamma|}_b D^\gamma u(x)\right),
\end{aligned}$$

and $b \in F$ is chosen so that $|x-b| = \mathrm{dist}\,(x,F)$. We employed here the fact that $D^\alpha T^m_{2Q_i}u \equiv 0$, so the term with $\beta = 0$ does not appear.

Now the estimate of $S_{\beta,\gamma}$ follows from Corollary 4.5, the fact that any $x \in \mathbb{R}^n \setminus F$ belongs to at most $C(n)$ supports $\mathrm{supp}\,\varphi_i$, the fact that $d_i \approx \mathrm{dist}\,(x,F)$ for $x \in \mathrm{supp}\,\varphi_i$ and the concavity of ω. Namely,

$$|S_{\beta,\gamma}| \leq \sum_{i:\, x\in \mathrm{supp}\,\varphi_i} C d_i^{-|\beta|}\omega(d_i)d_i^{m-|\gamma|} \leq C\omega(\mathrm{dist}\,(x,F))\mathrm{dist}\,(x,F)^{m-(|\beta|+|\gamma|)}.$$

Since $|\beta| + |\gamma| = m+1$ inequality (40) follows.

Now, let $|\beta| = m$. Using the mean value theorem, we find, in the interval joining x and y, a point z which satisfies $|x-y| < \mathrm{dist}\,(z,F) \approx \mathrm{dist}\,(x,F)$ and

$$\begin{aligned}
|D^\beta w(x) - D^\beta w(y)| &= \left|\sum_{j=1}^n (x_j - y_j)\frac{\partial(D^\beta w)}{\partial x_j}(z)\right| \\
&\leq C|x-y|\omega(\mathrm{dist}\,(z,F))\mathrm{dist}\,(z,F)^{-1} \\
&\leq \omega(|x-y|)
\end{aligned}$$

(the first inequality follows from (40), and the second from concavity of ω). This completes the proof of the lemma. □

4.3 Distributional derivatives of the extension

We check now that $w \in W^{m+1,p}(\mathbb{R}^n)$. We need the following variant of Nikodym's theorem, see [10, 4.9.2].

Lemma 4.7 *Assume* $1 \leq p < \infty$.

(i) If $u \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n)$ *is defined everywhere by formula (2), then for each* $k = 1, 2, \ldots, n$ *the functions*

$$t \mapsto u(x_1, \ldots, x_{k-1}, t, x_{k+1}, \ldots, x_n) \tag{41}$$

are absolutely continuous on bounded intervals in $\mathbb{R}$ *for almost every point* $x' = (x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \in \mathbb{R}^{n-1}$. *Moreover the partial derivatives of* u*, which exist a.e. (since absolutely continuous functions are differentiable a.e.) coincide with distributional derivatives of* u *and hence belong to* $L^p_{\mathrm{loc}}(\mathbb{R}^n)$.

(ii) Conversly, suppose $u \in L^p_{\mathrm{loc}}(\mathbb{R}^n)$ is such that for each $k = 1, 2, \ldots, n$ the functions (41) are absolutely continuous on bounded intervals in $\mathbb{R}$ for a.e. point $x' = (x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \in \mathbb{R}^{n-1}$ and partial derivatives of u, which exist, a.e. belong to $L^p_{\mathrm{loc}}(\mathbb{R}^n)$, then $u \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n)$.

Taking into account the previous parts of the proof, in order to prove that $w \in W^{m+1,p}(\mathbb{R}^n)$ it is enough to show the following.

Lemma 4.8 *If $w \in C^m(\mathbb{R}^n) \cap W^{m+1,p}(\mathbb{R}^n \setminus F)$, where F is a closed set, $u \in W^{m+1,p}(\mathbb{R}^n)$, and $D^\alpha w(x) = D^\alpha u(x)$ for any $x \in F$ and any α with $|\alpha| \le m$, then $w \in W^{m+1,p}(\mathbb{R}^n)$.*

Proof. Obviously w is of class $W^{m,p}(\mathbb{R}^n)$, and it suffices to prove that $D^\alpha w \in W^{1,p}(\mathbb{R}^n)$ for any α with $|\alpha| = m$. Fix such a multiindex α and set $v = D^\alpha w - D^\alpha u$.

(i) It follows from the previous lemma that $D^\alpha u$ is absolutely continuous on almost all lines paralell to one of coordinate axes, and has L^p-integrable derivative along these lines.

Since $D^\alpha w$ is continuous on $\mathbb{R}^n$, it follows that $v = D^\alpha w - D^\alpha u$ is continuous on all lines described in (i). Since $D^\alpha w \in W^{1,p}(\mathbb{R}^n \setminus F)$ we see that

(ii) for almost all lines ℓ paralell to coordinate axes $D^\alpha w|_\ell$ is absolutely continuous on compact intervals contained in $\ell \cap (\mathbb{R}^n \setminus F)$, with the derivative of class $L^p(\ell \cap (\mathbb{R}^n \setminus F))$.

Thus for almost all lines paralell to coordinate axe both conditions (i) and (ii) are satisfied. Pick such a line ℓ, paralell to the x_i axis. Then v is continuous on ℓ, absolutely continuous on compact intervals in $\ell \cap (\mathbb{R}^n \setminus F)$ with derivative in $L^p(\ell \cap (\mathbb{R}^n \setminus F))$ and $v \equiv 0$ in F. Now it easily follows that $v|_\ell$ coincides with the integral of a function $v' \in L^p \cap L^1_{\mathrm{loc}}(\ell)$ which is identically zero in $\ell \cap F$ and is equal to $\partial v / \partial x_i$ on $\ell \cap (\mathbb{R}^n \setminus F)$. Another application of Lemma 4.7 yields the desired result. $\square$

The whole Theorem 4.1 easily follows from the results of Subsections 4.1, 4.2 and 4.3.

The general case of Theorem 1.1 can be reduced in a standard and easy way to the one considered above, via a partition of unity and an "$\varepsilon/2^j$-argument", as in the proof of Meyers–Serrin theorem. See [6], [14] for a related trick. This concludes the proof of Theorem 1.4.

5 Proof of Theorem 1.2

To obtain the proof of Theorem 1.2 one should use the same extension formula (27). However, in order to define the set F, one should use the maximal functions $M^\flat$. This forces some technical changes in the proof (in particular, one should use different pointwise estimates), but the overall idea remains the same. Here is the sketch of most important steps.

Step 1. It suffices to prove the theorem for functions u with compact support contained in the unit cube Q. For such u, take a closed set F which satisfies the two following conditions:

$$M_1^\flat(\nabla^m u)(x) \le t = t(\varepsilon) \qquad \text{for all } x \in F, \tag{42}$$

$$M_\varrho^\flat(\nabla^m u) \overset{\varrho=0}{\longrightarrow} 0 \qquad \text{uniformly on } F. \tag{43}$$

By Section 4 of [6] we can choose $t(\varepsilon)$ sufficiently large to have $B_{k-m,p}(\mathbb{R}^n \setminus F) < \varepsilon$.

Next, take the Whitney cube decomposition $\mathbb{R}^n \setminus F = \bigcup_i Q_i$ and its associated smooth partition of unity $\{\varphi_i\}$. Set, as before,

$$w(x) = \begin{cases} u(x) & \text{for } x \in F, \\ \displaystyle\sum_{i\in I} \varphi_i(x) T^m_{2Q_i} u(x) & \text{for } x \in \mathbb{R}^n \setminus F. \end{cases} \tag{44}$$

Step 2 (Sobolev norm estimates). Exactly the same computations as in Subsection 4.1, give

$$\|w\|_{W^{m+1,p}(\mathbb{R}^n\setminus F)} \le C\|u\|_{W^{m+1,p}(\mathbb{R}^n\setminus F)}\,. \tag{45}$$

Step 3. Adapt the proof presented in Section 4.2 and check that in fact $w \in C^m(\mathbb{R}^n)$ and $D^\alpha w = D^\alpha u$ on F for all α such that $|\alpha| \le m$. This is rather tedious, but only minor changes of purely technical nature are necessary. In particular to estimate the formal Taylor remainders one should use Corollary 3.9 instead of Corollary 3.7. Actually the proof is slightly easier than that in Section 4.2 as we prove less (we do not prove Hölder continuity of m-th order derivatives). Now the theorem follows directly from Lemma 4.8.

We leave the missing details of this reasoning as an exercise for interested readers. □

References

[1] Alberti, G.: On the structure of singular sets of convex functions. *Calc. Var. Partial Differential Equations* 2 (1994), no. 1, 17–27.

[2] ANZELLOTTI, G., SERAPIONI, R.: C^k-rectifiable sets. *J. Reine Angew. Math.* 453 (1994), 1–20.
[3] BAGBY, T., ZIEMER, W. P.: Pointwise differentiability and absolute continuity. *Trans. Amer. Math. Soc.* 191 (1974), 129–148.
[4] BENNETT, C., SHARPLEY, R.: *Interpolation of operators.* Pure and Applied Mathematics, 129. Academic Press, Inc., Boston, MA, 1988.
[5] BOJARSKI, B.: Remarks on local function spaces. In: Lecture Notes in Math. 1302, Springer-Verlag, 1988, 137–152.
[6] BOJARSKI, B., HAJŁASZ, P.: Pointwise inequalities for Sobolev functions. *Studia Math.* 106 (1993), 77–92.
[7] CALDERÓN, A.P., ZYGMUND, A.: Local properties of solutions of elliptic partial differential equations. *Studia Math.* 20 (1961), 171-225.
[8] CALDERÓN, C. P., MILMAN, M.: Interpolation of Sobolev spaces. The real method. *Indiana Univ. Math. J.* 32 (1983), no. 6, 801–808.
[9] DORRONSORO, J. R.: Differentiability properties of functions with bounded variation. *Indiana Univ. Math. J.* 38 (1989), no. 4, 1027–1045.
[10] EVANS, L. C., AND GARIEPY, R. F.: *Measure Theory and Fine Properties of Functions.* Studies in Advanced Mathematics, 1992.
[11] FEDERER, H.: Surface area II. *Trans. Amer. Math. Soc.* 55 (1944), 438–456.
[12] FEDERER, H.: *Geometric measure theory.* Springer-Verlag, 1969.
[13] GILBARG, D., TRUDINGER, N.: *Elliptic partial differential equations of second order.* Springer-Verlag, 1983.
[14] HAJŁASZ, P.: Note on Meyers–Serrin's theorem. *Expo. Math.*, 11 (1993), 377–379.
[15] HAJŁASZ, P.: Boundary behaviour of Sobolev mappings. *Proc. Amer. Math. Soc.* 123 (1995), 1145–1148.
[16] HAJŁASZ, P.: Sobolev spaces on an arbitrary metric space. *Potential Analysis* 5 (1996), 403–415.
[17] HAJŁASZ, P., KINNUNEN, J.: Hölder quasicontinuity of Sobolev functions on metric spaces. *Revista Math. Iberoamericana* 14 (1998), 601–622.
[18] HAJŁASZ, P., STRZELECKI, P.: In preparation.
[19] HEDBERG, L.: On certain convolution inequalities. *Proc. Amer. Math. Soc.* 36 (1972), 505-510.
[20] LIU, FON-CHE: A Lusin type property of Sobolev functions. *Indiana Univ. Math. J.* 26 (1977), 645-651.
[21] MALGRANGE, B.: *Ideals of differentiable functions.* Oxford Univ. Press, London 1966.
[22] MALÝ, J.: Hölder type quasicontinuity. *Potential Analysis* 2 (1993), 249–254.
[23] MAZ'YA, V.G.: *Sobolev spaces.* Springer-Verlag 1985.
[24] MEYERS, N.: A theory of capacities for potentials of functions in Lebesque classes. *Math. Scand.* 26 (1970), 255-292.
[25] MICHAEL, J., ZIEMER, W.P.: A Lusin type approximation of Sobolev functions by smooth functions. *Contemp. Math.* 42 (1985), 135-167.
[26] NIKODYM, O.: Sur une classe de fonctions considérées dans le probleme de Dirichlet. *Fundam. Math.* 21 (1933), 129–150.
[27] RESHETNYAK, YU. G.: Remarks on integral representations, *Sibirsk. Mat. Zh.* 25, no. 5, (1984), 198–200 (in Russian).
[28] STEIN, E.: *Singular integrals and differentiability properties of functions*, Princeton Univ. Press, 1970.
[29] STOCKE, B. M.: A Lusin type approximation of Bessel potentials and Besov functions by smooth functions. *Math. Scand.* 77 (1995), 60–70.
[30] SWANSON, D.: Approximation in fractional Sobolev spaces. Preprint, Indiana University.
[31] WHITNEY, H.: Analytic extensions of differentiable functions defined in closed sets. *Trans. Amer. Math Soc.* 36 (1934), 63-89.
[32] WHITNEY, H.: On totally differentiable and smooth functions. *Pac. Journ. Math.* 1 (1951), 143–159.
[33] ZIEMER, W.P.: Uniform differentiability of Sobolev functions. *Indiana Univ. Math. J.* 37 (1988), 789-799.

[34] Ziemer, W.P.: *Weakly differentiable functions*. Springer-Verlag, 1989.

Bogdan Bojarski:
Institute of Mathematics, Polish Academy of Sciences, ul. Śniadeckich 8, 00–950 Warszawa, Poland.

Piotr Hajłasz:
Institute of Mathematics, Warsaw University, ul. Banacha 2, 02–097 Warszawa, Poland. E-mail: `hajlasz@mimuw.edu.pl`

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