

Pointwise inequalities and approximation in fractional Sobolev spaces

by

DAVID SWANSON (College Station, TX)

Abstract. We prove that a function belonging to a fractional Sobolev space $L^{\alpha,p}(\mathbb{R}^n)$ may be approximated in capacity and norm by smooth functions belonging to $C^{m,\lambda}(\mathbb{R}^n)$, $0 < m + \lambda < \alpha$. Our results generalize and extend those of [12], [4], [14], and [11].

1. Introduction. When $\Omega \subset \mathbb{R}^n$ is an open set and $p \geq 1$, the Sobolev space $W^{k,p}(\Omega)$ consists of those functions $f \in L^p(\Omega)$ whose distributional partial derivatives of order up to and including k are also members of $L^p(\Omega)$. The norm of this space is given by

$$\|f\|_{W^{k,p}(\Omega)} = \sum_{|\sigma| \leq k} \|D^\sigma f\|_p.$$

We will assume throughout that $1 < p < \infty$.

An important property of functions $f \in W^{1,p}(\mathbb{R}^n)$ is the Lusin-type property of *quasicontinuity*—given $\varepsilon > 0$ it is possible to find a continuous function g such that $f = g$ off a set with $(1,p)$ -capacity less than ε . This notion was extended to the higher-order spaces by Bagby and Ziemer [2], who proved that for $f \in W^{k,p}(\mathbb{R}^n)$, $0 \leq m \leq k$, and $\varepsilon > 0$ there is a function $g \in C^m(\mathbb{R}^n)$ such that $f = g$ off a set whose $(k - m, p)$ -capacity does not exceed ε . It is worth mentioning that the case $m = k$ was originally treated by Calderón and Zygmund [5]. Bagby and Ziemer's result, as well as the results in the present paper, rely greatly on their techniques. Michael and Ziemer [12] later showed that the function g , in addition to satisfying the above-stated properties, may be chosen so that $\|f - g\|_{m,p} < \varepsilon$.

Recently Bojarski, Hajłasz, and Strzelecki [4] have considered the problem of obtaining a continuous scale of approximations to functions $f \in W^{k,p}(\mathbb{R}^n)$. They proved that for $\varepsilon > 0$, $0 \leq m \leq k - 1$, and $0 < \lambda < 1$ there exists $g \in C^{m,\lambda}(\mathbb{R}^n)$ such that $f = g$ off a set of $(k - m - \lambda, p)$ -capacity not exceeding ε and $\|f - g\|_{m+1,p} < \varepsilon$. As in the discrete scale of approximations there is an inverse linear relationship between the order of the approximator

and the order of the capacity. An interesting aspect of their result, and one which will be considered in this paper, is that the approximator g belongs to a higher order Sobolev space than might be expected from the condition $g \in C^{m,\lambda}(\mathbb{R}^n)$.

It turns out that if $m \leq k-1$, the approximator g in Michael and Ziemer’s result may also be chosen so that $g \in C^m(\mathbb{R}^n)$ and $\|f - g\|_{m+1,p} < \varepsilon$, extending the result of Bojarski, Hajlasz, and Strzelecki to the case $\lambda = 0$. Again this is surprising in that the condition $g \in C^m(\mathbb{R}^n)$ gives no information regarding the order $m + 1$ weak differentiability of g . Of course this higher norm approximation is no longer possible when $m = k$. Combining the above with the result of [4] we obtain the following.

THEOREM 1.1. *Assume that $1 < p < \infty$, m is an integer satisfying $0 \leq m \leq k - 1$, and $0 \leq \lambda < 1$. For a function $f \in W^{k,p}(\mathbb{R}^n)$ and $\varepsilon > 0$, there exist a function $g \in C^{m,\lambda}(\mathbb{R}^n)$ and an open set Ω with the following properties:*

- (1) $B_{k-m-\lambda,p}(\Omega) < \varepsilon$,
- (2) $D^\sigma f(x) = D^\sigma g(x)$ for all $x \in \mathbb{R}^n \setminus \Omega$ and multi-indices $|\sigma| \leq m$, and
- (3) $\|f - g\|_{m+1,p} < \varepsilon$,

where $B_{\alpha,p}$ denotes the (α, p) -capacity.

As observed in [4] the case $k = 1$ and $0 < \lambda < 1$ generalizes a theorem of Malý [11]. The case $k = 1$ and $\lambda = 0$ extends the classical concept of quasicontinuity by showing that for $f \in W^{1,p}(\mathbb{R}^n)$ and $\varepsilon > 0$ there exists $g \in C(\mathbb{R}^n)$ such that f and g not only coincide off a set with $(1, p)$ -capacity less than ε , but also $g \in W^{1,p}(\mathbb{R}^n)$ and $\|f - g\|_{1,p} < \varepsilon$.

Theorem 1.1 will be proven as a special case of a more general result phrased in terms of the fractional Sobolev spaces $L^{\alpha,p}(\mathbb{R}^n)$, the so-called “Bessel potential spaces”. Recall that $L^{\alpha,p}(\mathbb{R}^n)$ consists of functions f of the form $f = g_\alpha * u$, with norm $\|g_\alpha * u\|_{\alpha,p} := \|u\|_p$, where $u \in L^p(\mathbb{R}^n)$ and g_α , $\alpha > 0$, is the Bessel kernel whose Fourier transform is

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

When $\alpha = k$ is an integer, the spaces $W^{k,p}(\mathbb{R}^n)$ and $L^{k,p}(\mathbb{R}^n)$ are equivalent with equivalent norms. When α is not an integer, a function f belongs to $L^{\alpha,p}(\mathbb{R}^n)$ if and only if, for every non-negative integer $k < \alpha$, we have $f \in W^{k,p}(\mathbb{R}^n)$ and $D^\sigma f \in L^{\alpha-k,p}(\mathbb{R}^n)$ whenever $|\sigma| = k$, in which case the norms

$$(1.1) \quad \|f\|_{\alpha,p} \quad \text{and} \quad \|f\|_{k,p} + \sum_{|\sigma|=k} \|D^\sigma f\|_{\alpha-k,p}$$

are equivalent (cf. [13, Ch. V]). The (α, p) -capacity of an arbitrary set

$E \subset \mathbb{R}^n$ is defined as

$$(1.2) \quad B_{\alpha,p}(E) = \inf\{\|u\|_p^p : u \in L^p(\mathbb{R}^n), u \geq 0, g_\alpha * u \geq 1 \text{ on } E\}.$$

It can be shown ([18, Sect. 2.6]) that $B_{\alpha,p}$ is an outer measure satisfying $B_{\alpha,p}(E) = \inf B_{\alpha,p}(U)$ whenever E is non-empty, where the infimum is taken over all open sets containing E . We are now in a position to state the main result of this paper.

THEOREM 1.2. *Assume that $1 < p < \infty$, m is an integer satisfying $0 \leq m \leq \alpha - 1$, and $0 \leq \lambda < 1$. For a function $f \in L^{\alpha,p}(\mathbb{R}^n)$ and $\varepsilon > 0$, there exist a function $g \in C^{m,\lambda}(\mathbb{R}^n)$ and an open set Ω with the following properties:*

- (1) $B_{\alpha-m-\lambda,p}(\Omega) < \varepsilon$,
- (2) $D^\sigma f(x) = D^\sigma g(x)$ for all $x \in \mathbb{R}^n \setminus \Omega$ and multi-indices $|\sigma| \leq m$, and
- (3) $\|f - g\|_{m+1,p} < \varepsilon$.

To compare our result with the existing literature, the following remarks are in order:

1. When α is an integer and $\lambda = 0$, the conclusion that $\|f - g\|_{m+1,p} < \varepsilon$ strengthens the norm approximation in [12].
2. When α is an integer and $0 < \lambda < 1$, we obtain a new proof of the result in [4].
3. When α is not an integer and $\lambda = 0$, Theorem 1.2 was obtained by Stocke [14] with the weaker conclusion that $\|f - g\|_{m,p} < \varepsilon$. However, the result in [14] holds in the case $m < \alpha < m + 1$ not treated by Theorem 1.2.
4. The case $m = 0$ was first considered by Malý [11], who showed that a function $f \in L^{\alpha,p}(\mathbb{R}^n)$ is Hölder continuous with exponent β off a set of small (α, q) -capacity for suitable β and q . Our result contains the result in [11], with no restriction on β and a sharper capacity estimate.
5. Our result does not require that $(\alpha - m - \lambda)p < n$. Although Theorem 1.2 follows from the Sobolev theorems when $(\alpha - m - \lambda)p > n$, we obtain a new approximation in the borderline case $(\alpha - m - \lambda)p = n$.

The main idea of the proofs of the Michael–Ziemer and Stocke theorems is to show that for $f \in L^{\alpha,p}(\mathbb{R}^n)$, $m < \alpha$, and $\varepsilon > 0$ there is an open set Ω such that $B_{\alpha-m,p}(\Omega) < \varepsilon$ and

$$(1.3) \quad \left(\int_{B(x,r)} |f(y) - T_x^m f(y)|^p dy \right)^{1/p} = o(r^m) \quad \text{as } r \rightarrow 0$$

uniformly for $x \in \mathbb{R}^n \setminus \Omega$, where $T_x^m f(y)$ is the order m Taylor polynomial of f centered at x . The function g is then defined to be equal to f on $\mathbb{R}^n \setminus \Omega$ and is extended to all of \mathbb{R}^n using an L^p -version of the Whitney

extension theorem formulated in terms of Calderón–Zygmund classes (cf. [5], [18, Ch. 3]).

The proof of Theorem 1.2 uses a similar idea. In addition to considering the behavior of integral averages, we also consider pointwise inequalities describing how far $f(x)$ deviates from its Taylor polynomial centered at a nearby point. Specifically, Ω may be chosen so that $B_{\alpha-m,p}(\Omega) < \varepsilon$,

$$(1.4) \quad \int_{B(x,r)} |f(y) - T_x^m f(y)| dy = o(r^m) \quad \text{as } r \rightarrow 0,$$

and

$$(1.5) \quad |D^\beta f(x) - T_y^{m-|\beta|} D^\beta f(x)| = o(|x - y|^{m-|\beta|})$$

uniformly for $x, y \in \mathbb{R}^n \setminus \Omega$ and all multi-indices $|\beta| \leq m$. Of course, (1.4) follows from (1.3) and Hölder’s inequality, but (1.3) requires the additional assumption that $(\alpha - m)p < n$, which we do not wish to make. One can use the classical Whitney extension theorem [17] to extend f off the set $\mathbb{R}^n \setminus \Omega$ to a function g having properties (1) and (2) of Theorem 1.2. This approach does not seem satisfactory to obtain the norm estimates, however, as the Whitney extension neglects the fact that f has an *a priori* definition on the whole space. Instead we use the Calderón–Zygmund extension operator. That the function g belongs to $W^{m+1,p}(\mathbb{R}^n)$ will be verified using the following trace theorem due to the author and W. P. Ziemer [15].

PROPOSITION 1.3. *Let $\Omega \subset \mathbb{R}^n$ be an arbitrary open set and suppose that $h \in W^{k,p}(\Omega)$, $1 < p < \infty$. Then $h \in W_0^{k,p}(\Omega)$ if and only if, for each multi-index $|\beta| \leq k - 1$,*

$$\lim_{r \rightarrow 0} r^{-n} \int_{B(x,r) \cap \Omega} |D^\beta h(y)| dy = 0$$

for $(k - |\beta|, p)$ -quasi-every $x \in \partial\Omega$.

When $0 < \lambda < 1$, the set Ω may be chosen so that in addition to (1.4) and (1.5) above, it satisfies $B_{\alpha-m-\lambda,p}(\Omega) < \varepsilon$,

$$(1.6) \quad \int_{B(x,r)} |f(y) - T_x^m f(y)| dy = O(r^{m+\lambda}) \quad \text{as } r \rightarrow 0$$

uniformly for $x \in \mathbb{R}^n \setminus \Omega$ and

$$(1.7) \quad |D^\beta f(x) - D^\beta f(y)| = O(|x - y|^\lambda)$$

uniformly for $x, y \in \mathbb{R}^n \setminus \Omega$ and $|\beta| = m$, although in this case Ω must be enlarged so that $B_{\alpha-m-\lambda,p}(\Omega) < \varepsilon$ is the sharpest capacity estimate possible. The approximator g is constructed in the same way as before, and all that is necessary is to show that the highest order partial derivatives of g are Hölder continuous. This follows from the construction of g and properties (1.6) and (1.7).

It is common in proofs of this nature to assume that the function f has compact support and then use a partition of unity argument to prove the general case. This is not sufficient, however, to conclude that the order m partial derivatives of f are Hölder continuous on the whole space, so we do not use this type of argument.

Since our result involves the notion of capacity, we must be careful to consider appropriate representatives of functions $f \in L^{\alpha,p}(\mathbb{R}^n)$. We use the pointwise definition $f(x) = g_\alpha * u(x)$ at those points x for which $g_\alpha * |u|(x) < \infty$, whenever $u \in L^p(\mathbb{R}^n)$ satisfies $f = g_\alpha * u$ almost everywhere. The definition of f on the set $\{x : g_\alpha * |u|(x) = \infty\}$, which has (α, p) -capacity 0, is not considered.

The paper is organized as follows. Sections 2–4 consist of preliminary material. Section 2 reviews the notation and basic results required in what follows. In Section 3 we give a simple proof, based on an analysis of local Riesz potentials, of the uniform L^1 differentiability property (1.4). Although this result is well known we include the proof to emphasize the irrelevance of the condition $(\alpha - m)p < n$. The Calderón–Zygmund smoothing operator used to construct the approximator g is given in Section 4. Lemmas 4.1 and 4.2 are used in what follows to verify the norm and Hölder continuity properties of g .

Sections 5–8 contain the proof of Theorem 1.2. The case $\lambda = 0$ is proven in Section 5. In Section 6 we derive pointwise Hölder-type estimates with exponent λ for functions in the fractional Sobolev spaces. The basic estimate in this section is Lemma 6.5, which gives a Hölder-type estimate valid (α, p) -quasi-everywhere for functions $f \in L^{\alpha,p}(\mathbb{R}^n)$. We digress briefly from the main argument in Theorem 6.7, where we use our estimates to give a particularly simple proof of the fact that if $\alpha > 0$ and $0 < \lambda < \min(\alpha, 1)$, then for any $f \in L^{\alpha,p}(\mathbb{R}^n)$ and $\varepsilon > 0$ there exists $g \in C^{0,\lambda}(\mathbb{R}^n)$ with the property that f and g coincide outside an open set whose $(\alpha - \lambda, p)$ -capacity is smaller than ε . This property of Bessel potentials could rightly be called “Hölder quasicontinuity”. In Section 7 we derive higher order estimates involving the order m Taylor polynomials of f as corollaries of the inequalities in Section 6. These are used to finish the proof of Theorem 1.2 in Section 8.

Finally, in Section 9 we prove a weaker version of Theorem 1.2 which is valid without the restriction $m \leq \alpha - 1$, where the Sobolev $(m + 1, p)$ norm is replaced by a suitable Besov norm.

2. Preliminaries. The dimension n remains fixed and $\omega(n)$ denotes the Lebesgue measure of the unit ball in \mathbb{R}^n . A multi-index is a vector $\sigma = (\sigma_1, \dots, \sigma_n) \in \mathbb{R}^n$ whose components are non-negative integers, with length $|\sigma| = \sigma_1 + \dots + \sigma_n$. When $x \in \mathbb{R}^n$, x^σ means $x_1^{\sigma_1} \dots x_n^{\sigma_n}$. Observe that

$|x^\sigma| \leq |x|^{|\sigma|}$. The notation $\beta \leq \sigma$ means that $\beta_i \leq \sigma_i$ for each coordinate i , and $\sigma! = \sigma_1! \dots \sigma_n!$.

The partial derivatives of a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ are denoted by $D^\sigma f = \partial_{x_1}^{\sigma_1} \dots \partial_{x_n}^{\sigma_n} f$. The m th-order Taylor polynomial of a function $f \in C^m(\mathbb{R}^n)$ is defined in the usual fashion:

$$(2.1) \quad T_x^m f(y) = \sum_{|\sigma| \leq m} D^\sigma f(x) \frac{(y-x)^\sigma}{\sigma!}.$$

When $E \subset \mathbb{R}^n$, we denote by $T_E^m f(y)$ the integral average of $T_x^m f(y)$ (as a function of x) over E :

$$T_E^m f(y) = \int_E T_x^m f(y) dx = \sum_{|\beta| \leq m} \int_E D^\beta f(x) \frac{(y-x)^\beta}{\beta!} dx.$$

Here the integral average is defined by

$$\bar{u}_E = \int_E u = \frac{1}{|E|} \int_E u.$$

The family $\{g_\alpha\}_{\alpha>0}$ of Bessel kernels is a semigroup with respect to convolution in the sense that

$$(2.2) \quad g_\alpha * g_\beta = g_{\alpha+\beta} \quad \text{for all } \alpha, \beta > 0.$$

In light of the fact that $\|g_\alpha\|_1 = 1$, (2.2) and the definition of capacity imply that $B_{\alpha_1, p} \leq B_{\alpha_2, p}$ whenever $\alpha_1 \leq \alpha_2$. Note also that $|E| \leq B_{\alpha, p}(E)$, $\alpha > 0$, whenever E is measurable. If a property holds for all $x \in \mathbb{R}^n$ except possibly for a set of (α, p) -capacity 0, we say the property holds (α, p) -quasi-everywhere, or simply (α, p) -q.e.

If $u \in L^p(\mathbb{R}^n)$, the definition of capacity gives

$$(2.3) \quad B_{\alpha, p}(\{x : g_\alpha * |u|(x) > t\}) \leq \frac{1}{t^p} \|u\|_p^p$$

for all $t > 0$. Passing to the limit as $t \rightarrow \infty$ we have

$$(2.4) \quad B_{\alpha, p}(\{x : g_\alpha * |u|(x) = \infty\}) = 0,$$

and hence $g_\alpha * |u|(x) < \infty$ for (α, p) -q.e. $x, y \in \mathbb{R}^n$.

The maximal function Mf of a locally integrable function f is defined as

$$Mf(x) = \sup_{r>0} \int_{B(x,r)} |f(y)| dy,$$

its local version $M_R f$ as

$$M_R f(x) = \sup_{0 < r < R} \int_{B(x,r)} |f(y)| dy,$$

and the local sharp maximal function $M_R^\# f$ as

$$M_R^\# f(x) = \sup_{0 < r < R} \int_{B(x,r)} |f(x) - f(y)| dy.$$

By the observation that $M(g_\alpha * u) \leq g_\alpha * Mu$, (2.3) above and the Hardy–Littlewood–Wiener maximal theorem imply that

$$(2.5) \quad B_{\alpha,p}(\{x : M(g_\alpha * u)(x) > t\}) \leq \frac{1}{t^p} \|Mu\|_p^p \leq \frac{C}{t^p} \|u\|_p^p$$

whenever $u \in L^p(\mathbb{R}^n)$, where $C = C_{n,p}$. As for the local sharp maximal function we have the following.

PROPOSITION 2.1 ([3, Thm. 8]). *Let $f \in L^{\alpha,p}(\mathbb{R}^n)$. For every $\varepsilon > 0$ there exists an open set U with $B_{\alpha,p}(U) < \varepsilon$ such that*

$$M_R^\# f \rightarrow 0 \quad \text{uniformly as } R \rightarrow 0 \text{ on } \mathbb{R}^n \setminus U.$$

Recall that the precise representative of a function $f = g_\alpha * u \in L^{\alpha,p}(\mathbb{R}^n)$ is defined by $f(x) = g_\alpha * u(x)$ whenever $g_\alpha * |u|(x) < \infty$. It turns out that this representative agrees with the representative defined by integral averages.

PROPOSITION 2.2 ([1, Thm. 6.2.1]). *Let $f = g_\alpha * u \in L^{\alpha,p}(\mathbb{R}^n)$. If $g_\alpha * |u|(x) < \infty$, then*

$$\lim_{r \rightarrow 0} \int_{B(x,r)} f(y) dy = g_\alpha * u(x).$$

REMARK 2.3. Assume that $f \in L^{\alpha,p}(\mathbb{R}^n)$, $\alpha \geq 1$, and that m is an integer with $0 < m \leq \alpha$. Then (1.1) states that $f \in W^{m,p}(\mathbb{R}^n)$ and $D^\sigma f \in L^{\alpha-|\sigma|,p}(\mathbb{R}^n)$ for all $|\sigma| \leq m$. If we identify each $D^\sigma f$ with its precise representative, the order k Taylor polynomial $T_x^m f$ of f may be defined formally for all x outside a set of $(\alpha - m, p)$ -capacity zero. For two multi-indices β, σ with $|\beta| + |\sigma| \leq m$, Proposition 2.2 implies that

$$(2.6) \quad D^\beta D^\sigma f(x) = D^\sigma D^\beta f(x) = D^{\beta+\sigma} f(x)$$

for $(\alpha - |\beta| - |\sigma|, p)$ -quasi-all, hence $(\alpha - m, p)$ -quasi-all $x \in \mathbb{R}^n$ since in general the mixed weak partial derivatives are equal almost everywhere. At those points x satisfying (2.6) for all $|\sigma| + |\beta| \leq m$ it follows that the classical formula (see [10])

$$(2.7) \quad D^\beta T_x^m f(y) = T_x^{m-|\beta|} D^\beta f(y)$$

holds for all $y \in \mathbb{R}^n$ and $|\beta| \leq m$. Since a polynomial of degree m agrees with its degree m Taylor polynomial centered at any point $z \in \mathbb{R}^n$, (2.7) implies that

$$(2.8) \quad T_x^m f(y) = T_z(T_x^m f)(y) = \sum_{|\sigma| \leq m} T_x^{m-|\sigma|} D^\sigma f(z) \frac{(y-z)^\sigma}{\sigma!}$$

holds for $(\alpha - m, p)$ -quasi-every $x \in \mathbb{R}^n$ and all $y, z \in \mathbb{R}^n$.

The space $W_0^{k,p}(\Omega)$ is defined as the closure of $C_0^\infty(\Omega)$ with respect to the $W^{k,p}(\Omega)$ norm. If $f \in W_0^{k,p}(\Omega)$ and

$$\tilde{f}(x) = \begin{cases} f(x), & x \in \Omega, \\ 0, & \text{otherwise,} \end{cases}$$

then $\tilde{f} \in W^{k,p}(\mathbb{R}^n)$.

3. A quick proof of L^1 differentiability. In this section we prove the claim in (1.4) without restriction on p . The proof requires the following known results. Proposition 3.1 is due to Bojarski and Hajlasz [3, Thm. 1] and Proposition 3.2 is due to Hedberg [9].

PROPOSITION 3.1. *Let $f \in L^{\alpha,p}(\mathbb{R}^n)$ and let x be a Lebesgue point of f . For each integer $0 < m < \alpha$ there is a constant C depending only on m and n such that*

$$|f(x) - T_B^m f(x)| \leq C \sum_{|\sigma|=m} \int_B \frac{|a_\sigma - D^\sigma f(z)|}{|x - z|^{n-m}} dz$$

for every ball B containing x and any family $(a_\sigma)_{|\sigma|=m}$ of real numbers.

PROPOSITION 3.2. *Let $s > 0$ and let $f \in L_{\text{loc}}^1(\mathbb{R}^n)$. There is a constant $C = C_{s,n}$ such that*

$$\int_{B(x,r)} \frac{|f(y)|}{|x - y|^{n-s}} dy \leq Cr^s M_r f(x)$$

for all $x \in \mathbb{R}^n$ and $r > 0$.

PROPOSITION 3.3 (Uniform L^1 differentiability). *Let $f \in W^{\alpha,p}(\mathbb{R}^n)$, $\alpha > 0$, $1 < p < \infty$, and let m be an integer with $0 \leq m < \alpha$. Then for every $\varepsilon > 0$ there is an open set Ω with $B_{\alpha-m,p}(\Omega) < \varepsilon$ such that*

$$r^{-m} \int_{B(x,r)} |f(y) - T_x^m f(y)| dy \rightarrow 0$$

uniformly for $x \in \mathbb{R}^n \setminus \Omega$.

Proof. Let $f \in L^{\alpha,p}(\mathbb{R}^n)$, $0 \leq m < \alpha$, and $\varepsilon > 0$. If $m = 0$ then the statement of the proposition is implied by Proposition 2.1 so we assume $m \geq 1$. Moreover, applying Proposition 2.1 to each $D^\sigma f$ with $|\sigma| = m$ we see that there is an open set Ω with $(\alpha - m, p)$ -capacity less than ε such that $T_x^m f$ is defined for all $x \in \mathbb{R}^n \setminus \Omega$ and $\sum_{|\sigma|=m} M_R^\# D^\sigma f \rightarrow 0$ uniformly as $R \rightarrow 0$ on $\mathbb{R}^n \setminus \Omega$. Fix $r > 0$ and let $B = B(x, r)$. By Proposition 3.1 we have

$$|f(y) - T_B^m f(y)| \leq C \sum_{|\sigma|=m} \int_B \frac{|D^\sigma f(x) - D^\sigma f(z)|}{|y - z|^{n-m}} dz$$

for almost all $y \in B$. Integrating over B and applying Tonelli's theorem and Proposition 3.2 gives

$$\begin{aligned}
 (3.1) \quad \int_B |f(y) - T_B^m f(y)| dy &\leq C \sum_{|\sigma|=m} \int_B \int_B \frac{|D^\sigma f(x) - D^\sigma f(z)|}{|y-z|^{n-m}} dz dy \\
 &\leq Cr^m \sum_{|\sigma|=m} \int_B |D^\sigma f(x) - D^\sigma f(z)| dz \\
 &\leq Cr^m \sum_{|\sigma|=m} M_r^\# D^\sigma f(x).
 \end{aligned}$$

On the other hand, we may use formula (2.8) to write

$$T_x^m f(y) - T_z^m f(y) = \sum_{|\beta| \leq m} (D^\beta f(x) - T_z^{m-|\beta|} D^\beta f(x)) \frac{(y-x)^\beta}{\beta!}$$

for $(\alpha-m, p)$ -quasi-every, hence almost every $z \in \mathbb{R}^n$. Integrating this over B we make the estimate

$$(3.2) \quad |T_x^m f(y) - T_B^m f(y)| \leq \sum_{|\beta| \leq m} r^{|\beta|} |D^\beta f(x) - T_B^{m-|\beta|} D^\beta f(x)|.$$

When $|\beta| = m$, the expression on the right side of (3.1) is

$$|D^\beta f(x) - \overline{(D^\beta f)}_B| \leq \int_B |D^\beta f(x) - D^\beta f(z)| dz \leq M_r^\# (D^\sigma f)(x),$$

and when $|\beta| < m$ we apply Propositions 3.1 and 3.2 with $a_\sigma = D^{\sigma+\beta} f(x)$:

$$\begin{aligned}
 |D^\beta f(x) - T_B^{m-|\beta|} D^\beta f(x)| &\leq C \sum_{|\sigma|=m-|\beta|} \int_B \frac{|D^{\sigma+\beta} f(x) - D^\sigma D^\beta f(z)|}{|x-z|^{n-m+|\beta|}} dz \\
 &\leq C \sum_{|\sigma|=m} \int_B \frac{|D^\sigma f(x) - D^\sigma f(z)|}{|x-z|^{n-m+|\beta|}} dz \\
 &\leq Cr^{m-|\beta|} \sum_{|\sigma|=m} M_r^\# (D^\sigma f)(x).
 \end{aligned}$$

Substituting these inequalities into (3.2) we get

$$|T_x^m f(y) - T_B^m f(y)| \leq Cr^m \sum_{|\sigma|=m} M_r^\# (D^\sigma f)(x),$$

and comparison with (3.1) implies that

$$(3.3) \quad r^{-m} \int_{B(x,r)} |f(y) - T_x^m f(y)| dy \leq C \sum_{|\sigma|=m} M_r^\# (D^\sigma f)(x)$$

for all $x \in \mathbb{R}^n \setminus \Omega$ and $r > 0$. That the right-hand side of (3.3) tends to 0 uniformly follows from the definition of Ω . ■

4. The Calderón–Zygmund extension operator. The material in this section is adapted from [5]. We will show that if Ω is an open set with $\text{dist}(x, \partial\Omega) < 1$ for all $x \in \Omega$, then for every integer $k \geq 1$ there exists a linear mapping $u \mapsto \bar{u}$ from $L^1_{\text{loc}}(\Omega)$ to $C^\infty(\Omega)$ which is bounded on $W^{k,p}(\Omega)$, $1 < p < \infty$. The definition of \bar{u} is simple. Let $\delta \in C^\infty(\Omega)$ satisfy

$$C_n \text{dist}(x, \partial\Omega) \leq \delta(x) \leq \text{dist}(x, \partial\Omega)$$

and $|D^\sigma \delta(x)| \leq C_{n,\sigma} \delta(x)^{1-|\sigma|}$ for all $x \in \Omega$ and $|\sigma| \geq 0$. For $k \geq 1$, let $\varphi \in C^\infty_0(B(0, 1))$ have the property that $P = P * \varphi_\varepsilon$ for all $\varepsilon > 0$ whenever P is a polynomial whose degree does not exceed k , where $\varphi_\varepsilon(x) = \varepsilon^{-n} \varphi(x/\varepsilon)$. For $x \in \Omega$ and $z \in \mathbb{R}^n$ define

$$\psi_z(x) = \delta(x)^{-n} \varphi\left(\frac{x-z}{\delta(x)}\right).$$

Then $\psi_z \in C^\infty(\Omega)$ for fixed z and for each multi-index $\beta \geq 0$ there is a constant $C = C_{\beta,k,n}$ such that

$$|D^\beta \psi_z(x)| \leq C \delta(x)^{-n-|\beta|}$$

whenever $x \in \Omega$ and $z \in \mathbb{R}^n$. See [5], [18, Ch. 3] for justifications.

Assuming $u \in L^1_{\text{loc}}(\Omega)$, we define the smoothing \bar{u} of u by

$$\bar{u}(x) = \int_{\mathbb{R}^n} \psi_z(x) u(z) dz$$

for all $x \in \Omega$. Since δ and φ are smooth, it is evident that $\bar{u} \in C^\infty(\Omega)$.

LEMMA 4.1. *Let $u \in L^p(\Omega)$, P a polynomial with degree $\leq k$, and β a multi-index with $|\beta| \leq k$. There is a constant $C = C_{n,k}$ such that*

$$|D^\beta \bar{u}(x) - D^\beta P(x)| \leq C \delta(x)^{-|\beta|} \int_{B(x,\delta(x))} |u(z) - P(z)| dz$$

for all $x \in \Omega$.

Proof. Writing $u(z) = P(z) + (u(z) - P(z))$, we have

$$\bar{u}(x) = P(x) + \int_{\mathbb{R}^n} \psi_z(x) (u(z) - P(z)) dz.$$

for all $x \in \Omega$. Differentiating gives

$$D^\beta \bar{u}(x) = D^\beta P(x) + \int_{\mathbb{R}^n} D^\beta \psi_z(x) (u(z) - P(z)) dz.$$

Since each ψ_z vanishes off $B(x, \delta(x))$, this is equivalent to

$$D^\beta \bar{u}(x) = D^\beta P(x) + \int_{B(x,\delta(x))} D^\beta \psi_z(x) (u(z) - P(z)) dz,$$

and so there is a constant $C = C_{k,n}$ satisfying

$$|D^\beta \bar{u}(x) - D^\beta P(x)| \leq C\delta(x)^{-n-|\beta|} \int_{B(x,\delta(x))} |u(z) - P(z)| dz$$

for all $x \in \Omega$. ■

It is easy to see that this smoothing process maps $L^p(\Omega)$ into $L^p(\Omega)$. Applying the preceding lemma with $\beta = 0$ and $P = 0$ we have

$$|\bar{u}(x)| \leq C \int_{B(x,\delta(x))} |u(z)| dz \leq CMu(x)$$

whenever $x \in \Omega$, and so the Hardy–Littlewood–Wiener theorem implies

$$\|\bar{u}\|_{L^p(\Omega)} \leq C\|u\|_{L^p(\Omega)},$$

where $C = C_{k,n,p}$. In fact, even more is true—if $m \leq k$, then the smoothing maps $W^{m,p}(\Omega)$ to $W^{m,p}(\Omega)$.

LEMMA 4.2. *Assume that $m \leq k$ and $u \in W^{m,p}(\Omega)$. Then $\bar{u} \in W^{m,p}(\Omega)$ and there is a constant $C = C_{k,n,p}$ such that*

$$\|\bar{u}\|_{W^{m,p}(\Omega)} \leq C\|u\|_{W^{m,p}(\Omega)}.$$

Proof. Fix $x \in \Omega$ and apply the preceding lemma with $P(z) = T_B^m u(z)$, where $B = B(x, \delta(x))$. There is a constant $C = C_{k,n}$ such that

$$|D^\beta \bar{u}(x) - D^\beta T_B^m u(x)| \leq C\delta(x)^{-|\beta|} \int_B |u(z) - T_B^m u(z)| dz$$

for each multi-index $|\beta| \leq m$. Proposition 3.1 (with $a_\sigma = 0$) and Proposition 3.2 imply that

$$\int_B |u(z) - T_B^m u(z)| dz \leq C\delta(x)^m \sum_{|\sigma|=m} \int_B |D^\sigma u(z)| dz,$$

hence that

$$|D^\beta \bar{u}(x) - D^\beta T_B^m u(x)| \leq C\delta(x)^{m-|\beta|} \sum_{|\sigma|=m} M(D^\sigma u)(x).$$

From the definition of $T_B^m u$ we have

$$\begin{aligned} |D^\beta T_B^m u(x)| &\leq \sum_{\substack{|\sigma| \leq m \\ \sigma \geq \beta}} \delta(x)^{|\sigma|-|\beta|} \int_B |D^\sigma u(z)| dz \\ &\leq \sum_{\substack{|\sigma| \leq m \\ \sigma \geq \beta}} \delta(x)^{|\sigma|-|\beta|} M(D^\sigma u)(x). \end{aligned}$$

Since $\delta(x) < 1$, we conclude that

$$|D^\beta \bar{u}(x)| \leq C \sum_{|\sigma| \leq m} M(D^\sigma u)(x)$$

for all $x \in \Omega$, and so the Hardy–Littlewood–Wiener theorem implies

$$\|D^\beta \bar{u}\|_{L^p(\Omega)} \leq C \sum_{|\sigma| \leq m} \|D^\sigma u\|_{L^p(\Omega)} = C \|u\|_{W^{m,p}(\Omega)},$$

where the constant in general depends on k, n , and p . The result follows by summing over all $|\beta| \leq k$. ■

5. The proof of Theorem 1.2, $\lambda = 0$. Assume that $1 < p < \infty$. In addition to the results of the preceding sections, the proof of Theorem 1.2 will use the following pointwise inequality of Bojarski and Hajlasz [3, Thm. 3].

PROPOSITION 5.1. *Let $\alpha > 0$ and let m be an integer with $0 \leq m < \alpha$. If $f \in L^{\alpha,p}(\mathbb{R}^n)$ then*

$$|f(x) - T_y^m f(x)| \leq C|x - y|^m \sum_{|\sigma|=m} (M_{|x-y|}^\# D^\sigma f(x) + M_{|x-y|}^\# D^\sigma f(y))$$

for $(\alpha - m, p)$ -quasi-every $x, y \in \mathbb{R}^n$, where $C = C_{n,k}$.

Observe that the case $m = 0$ follows directly from the definition of $M^\#$ and the fact that f is defined (α, p) -quasi-everywhere by integral averages. Applying Proposition 5.1 to $D^\beta f \in L^{\alpha-|\beta|,p}(\mathbb{R}^n)$, we also have

$$\begin{aligned} |D^\beta f(x) - T_y^{m-|\beta|} D^\beta f(x)| \\ \leq C|x - y|^{m-|\beta|} \sum_{|\sigma|=m} (M_{|x-y|}^\# D^\sigma f(x) + M_{|x-y|}^\# D^\sigma f(y)) \end{aligned}$$

for $(\alpha - k, p)$ -quasi-every $x, y \in \mathbb{R}^n$ whenever β is a multi-index with $|\beta| \leq m$. In light of Proposition 2.1, there is an open set $U \subset \mathbb{R}^n$ with $B_{\alpha-m,p}(U) < \varepsilon$ such that

$$\frac{|D^\beta f(x) - T_y^{m-|\beta|} D^\beta f(x)|}{|x - y|^{m-|\beta|}} \rightarrow 0$$

uniformly as $|x - y| \rightarrow 0$ for $x, y \in \mathbb{R}^n \setminus U$ and all multi-indices $|\beta| \leq m$.

The proof of Theorem 1.2 is divided into several steps. We assume that $\alpha > 0, f \in L^{\alpha,p}(\mathbb{R}^n), \lambda = 0, m$ is an integer with $0 \leq m \leq \alpha - 1$, and we fix $\varepsilon > 0$.

STEP 1. There exists an open set $\Omega \subset \mathbb{R}^n$ such that $B_{\alpha-m,p}(\Omega) < \varepsilon, T_x^m f$ is defined and satisfies (2.7), (2.8) for all $x \in \mathbb{R}^n \setminus \Omega$,

$$(5.1) \quad r^{-m} \int_{B(x,r)} |f(y) - T_x^m f(y)| dy \rightarrow 0$$

uniformly as $r \rightarrow 0$ for $x \in \mathbb{R}^n \setminus \Omega$, and

$$(5.2) \quad \frac{|D^\beta f(x) - T_y^{m-|\beta|} D^\beta f(x)|}{|x - y|^{m-|\beta|}} \rightarrow 0$$

uniformly as $|x - y| \rightarrow 0$ for $x, y \in \mathbb{R}^n \setminus \Omega$, $x \neq y$. This follows from Remark 2.3, Propositions 3.3 and 5.1, and the fact that $B_{\alpha-m,p}$ is an outer capacity. As $|\Omega| \leq B_{\alpha-m,p}(\Omega)$ we may assume that Ω is sufficiently small that $\text{dist}(x, \partial\Omega) < 1$ for all $x \in \Omega$. Define $K = \mathbb{R}^n \setminus \Omega$.

STEP 2. Since $m \leq \alpha - 1$ we have $f \in W^{m+1,p}(\mathbb{R}^n)$, hence $f \in W^{m+1,p}(\Omega)$. We may assume without loss of generality that Ω has been chosen sufficiently small that $\|f\|_{W^{m+1,p}(\Omega)} < \delta$ for any particular choice of $\delta > 0$ (see Step 5). Let \bar{f} be the smoothing of f on Ω satisfying $\|\bar{f}\|_{W^{m+1,p}(\Omega)} \leq C\|f\|_{W^{m+1,p}(\Omega)}$ as constructed in Section 4. Define the approximator g by

$$g(x) = \begin{cases} \bar{f}(x), & x \in \Omega, \\ f(x), & x \in K, \end{cases}$$

and for each multi-index $|\beta| \leq m$ define

$$g^\beta(x) = \begin{cases} D^\beta \bar{f}(x), & x \in \Omega, \\ D^\beta f(x), & x \in K. \end{cases}$$

Let $\delta(x)$ denote the regularized distance function on Ω .

STEP 3. Let $y \in K$ and let $|\beta| \leq m$. Then

$$|g^\beta(x) - D^\beta T_y^m f(x)| = o(|x - y|^{m-|\beta|})$$

as $x \rightarrow y$. To verify this, we must show that for arbitrary $\eta > 0$ we have $|g^\beta(x) - D^\beta T_y^m f(x)| \leq \eta|x - y|^{m-|\beta|}$ provided $|x - y|$ is sufficiently small. When $x \in K$ this follows from (5.2) and the definition of g^β . We therefore fix $\eta > 0$ and consider $x \in \Omega$.

Let $x^* \in K$ satisfy $|x - x^*| = \text{dist}(x, \Omega)$. We estimate $|g^\beta(x) - D^\beta T_{x^*}^m f(x)|$ and $|D^\beta T_{x^*}^m f(x) - D^\beta T_y^m f(x)|$ separately and use the triangle inequality. Define the order m polynomial $P(z) = T_{x^*}^m f(z)$. Lemma 4.1 implies that

$$|g^\beta(x) - D^\beta T_{x^*}^m f(x)| \leq C\delta(x)^{-|\beta|} \int_{B(x,\delta(x))} |f(z) - T_{x^*}^m f(z)| dz,$$

and since $C|x - x^*| \leq \delta(x) \leq |x - x^*|$ it follows that

$$|g^\beta(x) - D^\beta T_{x^*}^m f(x)| \leq C|x - x^*|^{-|\beta|} \int_{B(x^*,2|x-x^*|)} |f(z) - T_{x^*}^m f(z)| dz.$$

As $x^* \in K$, the uniformity of (5.1) implies

$$(5.3) \quad |g^\beta(x) - D^\beta T_{x^*}^m f(x)| \leq C\eta|x - x^*|^{m-|\beta|}$$

provided $|x - x^*|$ is sufficiently small regardless of the particular values of x and x^* . On the other hand, since $x^*, y \in K$ we may use (2.7) and (2.8) to write the difference $D^\beta T_y^m f(x) - D^\beta T_{x^*}^m f(x)$ as

$$\sum_{|\gamma| \leq m-|\beta|} (D^{\gamma+\beta} f(y) - T_{x^*}^{m-|\beta|-|\gamma|} D^{\gamma+\beta} f(y)) \frac{(x - y)^\gamma}{\gamma!},$$

implying that $|D^\beta T_y^m f(x) - D^\beta T_{x^*}^m f(y)|$ does not exceed

$$\sum_{|\gamma| \leq m - |\beta|} |D^{\gamma + \beta} f(y) - T_{x^*}^{m - |\beta| - |\gamma|} D^{\gamma + \beta} f(y)| \cdot |x - y|^{|\gamma|}.$$

In this case the uniformity of (5.2) shows that

$$(5.4) \quad |D^\beta T_y^m f(x) - D^\beta T_{x^*}^m f(x)| \leq \sum_{|\gamma| \leq m - |\beta|} \eta |x^* - y|^{m - |\beta| - |\gamma|} |x - y|^{|\gamma|}$$

provided that $|x^* - y|$ is sufficiently small. The choice of x^* implies $|x^* - y| \leq |x - y|$ and therefore $|x^* - y| \leq |x^* - y| + |x - y| \leq 2|x - y|$, so comparing (5.3) and (5.4) we conclude that $|g^\beta(x) - D^\beta T_y^m f(x)| \leq C\eta |x - y|^{m - |\beta|}$ provided that $|x - y|$ is sufficiently small. As $\eta > 0$ is arbitrary, this proves that $|g^\beta(x) - D^\beta T_y^m f(x)| = o(|x - y|^{m - |\beta|})$, as desired.

STEP 4. $g \in C^k(\mathbb{R}^n)$ and $D^\beta g = g^\beta$ for all multi-indices $|\beta| \leq k$. Assume first that $|\beta| = k$. Since \tilde{f} is smooth on Ω , g^β is continuous on Ω . On the other hand, for $y \in K$ we have

$$|g^\beta(x) - g^\beta(y)| = |g^\beta(x) - D^\beta f(y)| = |g^\beta(x) - D^\beta T_y^m f(x)| = o(1)$$

by Step 3. Thus g^β is continuous on \mathbb{R}^n .

Now assume $|\beta| \leq k - 1$ and let δ_j denote the multi-index of length 1 with j th coordinate 1. The equation $D_j^\delta g^\beta(y) = g^{\delta_j + \beta}(y)$ holds by definition for all $y \in \Omega$. For $y \notin K$ we expand the polynomial $D^\beta T_y^m f(x)$ as

$$D^\beta f(y) + \sum_{j=1}^n D^{\beta + \delta_j} f(y) \cdot (x_j - y_j) + \sum_{2 \leq |\gamma| \leq m - |\beta|} D^{\beta + \gamma} D^{\beta + \gamma} f(y) \frac{(x - y)^\gamma}{\gamma!}$$

for all $x \in \mathbb{R}^n$, and since $g^\beta(y) = D^\beta f(y)$ it follows that

$$\begin{aligned} & \left| g^\beta(x) - g^\beta(y) - \sum_{j=1}^n D^{\beta + \delta_j} f(y) \cdot (x_j - y_j) \right| \\ & \leq |g^\beta(x) - T_y^{m - |\beta|} D^\beta f(x)| + \sum_{2 \leq |\gamma| \leq m - |\beta|} |D^{\beta + \gamma} f(y)| \cdot |x - y|^{|\gamma|} \\ & = o(|x - y|) \end{aligned}$$

by Step 3. Thus g^β has first order partial derivatives at y satisfying $D_j^\delta g^\beta(y) = D^{\beta + \delta_j} f(y) = g^{\beta + \delta_j}(y)$ for all j . It follows that g^β is continuous at all points $y \in \mathbb{R}^n$, and a simple induction argument shows that $g \in C^m(\mathbb{R}^n)$ and $D^\beta g = g^\beta$ for all $|\beta| \leq m$. In particular, we have $D^\beta g(y) = D^\beta f(y)$ for all $y \in K$ and $|\beta| \leq m$, which completes the proof of parts (1) and (2) of Theorem 1.2.

STEP 5. It only remains to show that the approximator g belongs to $W^{m+1,p}(\mathbb{R}^n)$ and that $\|f - g\|_{m+1,p} < \varepsilon$. From Step 2 we deduce that $g \in W^{m+1,p}(\Omega)$ and satisfies $\|g\|_{W^{m+1,p}(\Omega)} \leq C\|f\|_{W^{m+1,p}(\Omega)}$, and therefore $h = g - f$ belongs to $W^{m+1,p}(\Omega)$ and satisfies $\|h\|_{W^{m+1,p}(\Omega)} \leq C\|f\|_{W^{m+1,p}(\Omega)}$.

Let β be a multi-index with $|\beta| \leq m$. Proposition 2.1 implies that

$$\lim_{r \rightarrow 0} \int_{B(x,r)} |D^\beta f(y) - D^\beta f(x)| dy = 0$$

for $(m + 1 - |\beta|, p)$ -quasi-every $x \in \mathbb{R}^n$ since $f \in W^{m+1,p}(\mathbb{R}^n)$, and

$$\lim_{r \rightarrow 0} \int_{B(x,r)} |D^\beta g(y) - D^\beta g(x)| dy = 0$$

for all $x \in \mathbb{R}^n$ since $g \in C^m(\mathbb{R}^n)$. Since $D^\beta f$ and $D^\beta g$ coincide on K , it follows that

$$\lim_{r \rightarrow 0} r^{-n} \int_{B(x,r) \cap \Omega} |D^\beta h(y)| dy = \lim_{r \rightarrow 0} r^{-n} \int_{B(x,r) \cap \Omega} |D^\beta f(y) - D^\beta g(y)| dy = 0$$

for $(m + 1 - |\beta|, p)$ -quasi-every $x \in \partial\Omega$ and all $|\beta| \leq m$. We apply [15, Thm. 3.1] to conclude that $h \in W_0^{m+1,p}(\mathbb{R}^n)$, and therefore that the function \tilde{h} which coincides with h on Ω and vanishes on K belongs to $W^{m+1,p}(\mathbb{R}^n)$. Since f and g coincide on K we have $g = \tilde{h} + f$, hence $g \in W^{m+1,p}(\mathbb{R}^n)$ and

$$\|f - g\|_{m+1,p} = \|\tilde{h}\|_{m+1,p} = \|h\|_{W^{m+1,p}(\Omega)} \leq C\|f\|_{W^{m+1,p}(\Omega)} < \varepsilon$$

provided that Ω has been chosen sufficiently small. ■

6. Hölder-type estimates. We assume throughout this section and the next that $0 < m + \lambda < \alpha$, where $m \geq 0$ is an integer and $0 < \lambda < 1$. The first four propositions will be used to prove the inequalities of this section. Proposition 6.1 is a simple but elegant result apparently first used by Frostman [7], Proposition 6.2 is a variant of [16, Prop. IV.2.3], and Propositions 6.3 and 6.4 are elementary.

PROPOSITION 6.1. *Let $0 < s < n$. There is a constant $C = C_{s,n}$ such that*

$$\int_{B(x,r)} |y|^{s-n} dy \leq C|x|^{s-n}$$

for all $x \in \mathbb{R}^n$ and $r > 0$.

PROPOSITION 6.2. *Let $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ be non-negative, integrable, radial, and decreasing as a function of $|x|$. There is a constant $C = C_n$ such that*

$$\int_{\mathbb{R}^n} \delta^{-n} \varphi\left(\frac{x-z}{\delta}\right) |f(z)| dz \leq C\|\varphi\|_1 Mf(x)$$

for all $\delta > 0$, $x \in \mathbb{R}^n$, and $f \in L^1_{\text{loc}}(\mathbb{R}^n)$.

PROPOSITION 6.3. Let $x, y, z, \zeta \in \mathbb{R}^n$ satisfy $|x - z| \geq \delta$, $|y - z| \geq \delta$, and $\zeta = tx + (1 - t)y$, where $\delta = \frac{3}{4}|x - y|$ and $0 < t < 1$. Then

$$(6.1) \quad \frac{|\zeta - z|}{|x - z|} \geq \frac{1}{5}.$$

Proof. Observe that $|x - \zeta| + |y - \zeta| = |x - y|$, so either $|x - \zeta| \leq \frac{1}{2}|x - y|$ or $|y - \zeta| \leq \frac{1}{2}|x - y|$. It follows that $|\zeta - z| \geq \frac{1}{4}|x - y|$, since one of the inequalities

$$|\zeta - z| \geq |x - z| - |x - \zeta| \geq \frac{1}{4}|x - y|$$

and

$$|\zeta - z| \geq |y - z| - |y - \zeta| \geq \frac{1}{4}|x - y|$$

must hold. Now we obtain the inequality in (6.1) by considering two separate cases. If $|x - z| \leq \frac{5}{4}|x - y|$, we have

$$\frac{|\zeta - z|}{|x - z|} \geq \frac{\frac{1}{4}|x - y|}{\frac{5}{4}|x - y|} = \frac{1}{5},$$

and if $|x - z| > \frac{5}{4}|x - y|$ we get the inequality

$$\frac{|\zeta - z|}{|x - z|} \geq \frac{|x - z| - |x - \zeta|}{|x - z|} = 1 - \frac{|x - \zeta|}{|x - z|} \geq 1 - \frac{|x - y|}{\frac{5}{4}|x - y|} = \frac{1}{5}. \blacksquare$$

PROPOSITION 6.4. Let $u \in L^p(\mathbb{R}^n)$. Then

$$|g_\alpha * u(x) - g_\alpha * u(y)| \leq \int_{\mathbb{R}^n} |g_\lambda(x - z) - g_\lambda(y - z)| g_{\alpha-\lambda} * |u|(z) dz$$

whenever $g_\alpha * |u|(x)$ and $g_\alpha * |u|(y)$ are both finite.

Proof. The semigroup property (2.2) implies that

$$\begin{aligned} g_\alpha * u(\zeta) &= \int_{\mathbb{R}^n} g_\alpha(\zeta - w)u(w) dw \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} g_\lambda(\zeta - z)g_{\alpha-\lambda}(z - w)u(w) dz dw \end{aligned}$$

for any $\zeta \in \mathbb{R}^n$. Thus if $g_\alpha * |u|(x)$ and $g_\alpha * |u|(y)$ are both finite, we have the estimate

$$\begin{aligned} |g_\alpha * u(x) - g_\alpha * u(y)| &\leq \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |g_\lambda(x - z) - g_\lambda(y - z)| g_{\alpha-\lambda}(z - w)|u(w)| dz dw \\ &= \int_{\mathbb{R}^n} |g_\lambda(x - z) - g_\lambda(y - z)| \int_{\mathbb{R}^n} g_{\alpha-\lambda}(z - w)|u(w)| dw dz \\ &= \int_{\mathbb{R}^n} |g_\lambda(x - z) - g_\lambda(y - z)| g_{\alpha-\lambda} * |u|(z) dz \end{aligned}$$

by an application of Tonelli's theorem. \blacksquare

The following Lemmas 6.5 and 6.6 give Hölder-type estimates for functions $f \in L^{\alpha,p}(\mathbb{R}^n)$ in which the quantity $|f(x) - f(y)|/|x - y|^\lambda$ is bounded quasi-everywhere by an appropriate maximal function. The awkward bound in Lemma 6.6 is useful in proving the integral estimates which follow.

LEMMA 6.5. *Let $f = g_\alpha * u \in L^{\alpha,p}(\mathbb{R}^n)$. There is a constant $C = C_{\lambda,n}$ such that*

$$(6.2) \quad |f(x) - f(y)| \leq C|x - y|^\lambda(M(g_{\alpha-\lambda} * |u|)(x) + M(g_{\alpha-\lambda} * |u|)(y))$$

whenever $g_\alpha * |u|(x)$ and $g_\alpha * |u|(y)$ are both finite.

Proof. Throughout the proof C will denote a constant whose value depends only on λ and n . Assume that $g_\alpha * |u|(x)$ and $g_\alpha * |u|(y)$ are both finite. In light of Proposition 6.4 it suffices to prove (6.2) with $|f(x) - f(y)|$ replaced by

$$\int_{\mathbb{R}^n} |g_\lambda(x - z) - g_\lambda(y - z)|g_{\alpha-\lambda} * |u|(z) dz.$$

Let $\delta = \frac{3}{4}|x - y|$. If $z \in B(y, \delta)$, then

$$|x - z| \geq |x - y| - |y - z| \geq \frac{4}{3}\delta - \delta \geq \frac{1}{3}|y - z|,$$

hence $|x - z|^{\lambda-n} \leq C|y - z|^{\lambda-n}$. Since $g_\lambda(w) \leq C|w|^{\lambda-n}$ for all $w \in \mathbb{R}^n$ ([13, Ch. V.2]), we have

$$|g_\lambda(x - z) - g_\lambda(y - z)| \leq C(|x - z|^{\lambda-n} + |y - z|^{\lambda-n}) \leq C|y - z|^{\lambda-n}.$$

Therefore

$$(6.3) \quad \int_{B(y,\delta)} |g_\lambda(x - z) - g_\lambda(y - z)|g_{\alpha-\lambda} * |u(z)| dz \leq C \int_{B(y,\delta)} \frac{g_{\alpha-\lambda} * |u|(z)}{|y - z|^{n-\lambda}} dz,$$

where

$$\int_{B(y,\delta)} \frac{g_{\alpha-\lambda} * |u|(z)}{|y - z|^{n-\lambda}} dz \leq C|x - y|^\lambda M(g_{\alpha-\lambda} * |u|)(y)$$

by Proposition 3.2 above. An identical computation leads to

$$\int_{B(x,\delta)} |g_\lambda(x - z) - g_\lambda(y - z)|g_{\alpha-\lambda} * |u(z)| dz \leq C|x - y|^\lambda M(g_{\alpha-\lambda} * |u|)(x),$$

so in order to verify (6.2) above it suffices to show that

$$(6.4) \quad \int_E |g_\lambda(x - z) - g_\lambda(y - z)|g_{\alpha-\lambda} * |u(z)| dz \leq C|x - y|^\lambda M(g_{\alpha-\lambda} * |u|)(x)$$

where $E = \mathbb{R}^n \setminus (B(x, \delta) \cup B(y, \delta))$.

We use the fact that g_λ is differentiable away from 0 and satisfies $|Dg_\lambda(w)| \leq C|w|^{\lambda-n-1}$ for all $w \neq 0$ ([13, Ch. V.2]). Let $z \in E$. It is evident that the line segment connecting $x - z$ and $y - z$ does not pass through the origin, so there is a point $\zeta = tx + (1 - t)y$, $0 < t < 1$, such that

$$|g_\lambda(x - z) - g_\lambda(y - z)| \leq |Dg_\lambda(\zeta - z)| \cdot |x - y| \leq C|\zeta - z|^{\lambda-n-1}|x - y|.$$

Applying Proposition 6.3 to x, y, z , and ζ , we have $|\zeta - z| \geq \frac{1}{5}|x - z|$, so

$$|g_\lambda(x - z) - g_\lambda(y - z)| \leq C|x - z|^{\lambda-n-1}|x - y|.$$

Now, $|x - z| \geq \delta$ implies

$$|x - z|^{\lambda-n-1} \leq 2^{n+1-\lambda}\delta^{\lambda-n-1} \left(1 + \frac{|x - z|}{\delta}\right)^{\lambda-n-1},$$

and therefore

$$|g_\lambda(x - z) - g_\lambda(y - z)| \leq C|x - y|^\lambda \delta^{-n} \left(1 + \frac{|x - z|}{\delta}\right)^{\lambda-n-1}.$$

Multiplying this quantity by $g_{\alpha-\lambda} * |u|(z)$ and integrating over the set E we have

$$\int_E |g_\lambda(x - z) - g_\lambda(y - z)| g_{\alpha-\lambda} * |u|(z) dz \leq C|x - y|^\lambda \|\varphi\|_1 M(g_{\alpha-\lambda} * |u|)(x)$$

by Proposition 6.2, taking $\varphi(z) = (1 + |z|)^{\lambda-n-1}$. A simple integration shows that

$$\|\varphi\|_1 \leq \omega_n \left(\frac{n}{1 - \lambda}\right),$$

which gives (6.4) and proves the lemma. ■

LEMMA 6.6. *In the conclusion of Lemma 6.5 we also have*

$$|f(x) - f(y)| \leq C \left(|x - y|^\lambda M(g_{\alpha-\lambda} * |u|)(x) + \int_{B(y, |x-y|)} \frac{g_{\alpha-\lambda} * |u|(z)}{|y - z|^{n-\lambda}} dz \right)$$

whenever $g_\alpha * |u|(x)$ and $g_\alpha * |u|(y)$ are both finite.

Proof. This follows immediately from (6.3) and (6.4) of the proof of Lemma 6.5. ■

THEOREM 6.7. *Assume that $m = 0$. Then for $f \in L^{\alpha,p}(\mathbb{R}^n)$ and $\varepsilon > 0$ there exist a function $g \in C^{0,\lambda}(\mathbb{R}^n)$ and an open set Ω such that $B_{\alpha-\lambda,p}(\Omega) < \varepsilon$ and $f(x) = g(x)$ for all $x \notin \Omega$.*

Proof. Note that the assumptions imply $0 < \lambda < \min(1, \alpha)$. Let $u \in L^p(\mathbb{R}^n)$ satisfy $f = g_\alpha * u$ and let E_t denote the set $\{M(g_{\alpha-\lambda} * |u|) > t\}$. By (2.5) above there is a constant C so that $B_{\alpha-\lambda,p}(E_t) \leq Ct^{-p} \|u\|_p^p$, and therefore $B_{\alpha-\lambda,p}(E_t) < \varepsilon$ for sufficiently large t . Fix such a t and define $E = E_t \cap \{x : g_\alpha * |u|(x) < \infty\}$. Then $B_{\alpha-\lambda,p}(E) < \varepsilon$ and by Lemma 6.5,

$|f(x) - f(y)| \leq Ct|x - y|^\lambda$ for all $x \in \mathbb{R}^n \setminus E$. The desired function g may then be given by

$$g(x) = \inf_{y \in E} \{g(y) + Ct|x - y|^\alpha\}$$

for all $x \in \mathbb{R}^n$ (cf. [6, 2.10.44]). Moreover, since $B_{\alpha-\lambda}$ is an outer capacity, one may choose an open set Ω containing $\{f \neq g\}$ such that $B_{\alpha-\lambda,p}(\Omega) < \varepsilon$. ■

The following two lemmas are integral estimates which measure the variation of a function $f \in L^{\alpha,p}(\mathbb{R}^n)$ over a ball B in terms of appropriate maximal functions.

LEMMA 6.8. *Assume $s > 0$ and let $f = g_\alpha * u \in L^{\alpha,p}(\mathbb{R}^n)$. There is a constant $C = C_{\lambda,s,n}$ such that*

$$\int_{B(x,r)} \frac{|f(x) - f(y)|}{|x - y|^{n-s}} dy \leq Cr^{s+\lambda} M(g_{\alpha-\lambda} * |u|)(x)$$

for all $x \in \mathbb{R}^n$ satisfying $g_\alpha * |u|(x) < \infty$ and $r > 0$.

Proof. Choose x with $g_\alpha * |u|(x) < \infty$ and let $B = B(x, r)$. We divide the inequality in Lemma 6.6 (which is valid for almost all $y \in B$) by $|x - y|^{n-s}$ and integrate over B to get

$$\int_B \frac{|f(x) - f(y)|}{|x - y|^{n-s}} dy \leq C(I(x) + J(x)),$$

where

$$I(x) = M(g_{\alpha-\lambda} * |u|)(x) \int_B \frac{1}{|x - y|^{n-s-\lambda}} dy$$

and

$$J(x) = \int_B \int_{B(y, |x-y|)} \frac{g_{\alpha-\lambda} * |u|(z)}{|y - z|^{n-\lambda} |x - y|^{n-s}} dz dy.$$

Proposition 3.2 implies $I(x) \leq Cr^{s+\lambda} M(g_{\alpha-\lambda} * |u|)(x)$. As for $J(x)$, writing $B(x, 2r) = 2B$ we have

$$\begin{aligned} J(x) &\leq \int_B \int_{2B} \frac{g_{\alpha-\lambda} * |u|(z)}{|y - z|^{n-\lambda} |x - y|^{n-s}} dz dy \\ &= \int_{2B} g_{\alpha-\lambda} * |u|(z) \int_B \frac{1}{|y - z|^{n-\lambda} |x - y|^{n-s}} dy dz \\ &= \int_{2B} g_{\alpha-\lambda} * |u|(z) \int_{B(x-z,r)} \frac{|y|^{\lambda-n}}{|(x - z) - y|^{n-s}} dy dz \end{aligned}$$

by Tonelli's theorem and the change of variable $y \mapsto y + z$. Writing $v(y) =$

$|y|^{\lambda-n}$ we apply Propositions 3.2 and 6.1 to get

$$\int_{B(x-z, 2r)} \frac{v(y)}{|(x-z) - y|^{n-s}} dy \leq Cr^s Mv(x-z) \leq Cr^s |x-z|^{\lambda-n},$$

which implies that

$$J(x) \leq Cr^s \int_{2B} \frac{g_{\alpha-\lambda} * |u|(z)}{|x-z|^{n-\lambda}} dz \leq Cr^{s+\lambda} M(g_{\alpha-\lambda} * |u|)(x)$$

by a third application of Proposition 3.2, where $C = C_{\lambda, s, n}$. ■

The following is a substitute result for Lemma 6.8 in the case $s = 0$.

LEMMA 6.9. *Let $f = g_\alpha * u \in L^{\alpha, p}(\mathbb{R}^n)$. There is a constant $C = C_{\lambda, n}$ such that*

$$\int_{B(x, r)} |f(x) - f(y)| dy \leq Cr^\lambda M(g_{\alpha-\lambda} * |u|)(x)$$

for all $x \in \mathbb{R}^n$ satisfying $g_\alpha * |u|(x) < \infty$ and all $r > 0$. In particular, this implies that

$$\sup_{r>0} r^{-\lambda} M_r^\# f(x) \leq CM(g_{\alpha-\lambda} * |u|)(x)$$

for (α, p) -quasi-every $x \in \mathbb{R}^n$.

Proof. As in the proof of Lemma 6.8 we integrate the inequality in Lemma 6.6 over $B = B(x, r)$ and write

$$\int_B |f(x) - f(y)| dy \leq C(I(x) + J(x)),$$

where

$$I(x) = \int_B |x - y|^\lambda M(g_{\alpha-\lambda} * |u|)(x) dy \leq r^\lambda M(g_{\alpha-\lambda} * |u|)(x)$$

and

$$\begin{aligned} J(x) &= \int_B \int_{B(y, |x-y|)} \frac{g_{\alpha-\lambda} * |u|(z)}{|y-z|^{n-\lambda}} dz dy \\ &\leq \int_{2B} g_{\alpha-\lambda} * |u|(z) \int_{B(x, r)} \frac{1}{|y-z|^{n-\lambda}} dy dz \\ &\leq C \int_{2B} \frac{g_{\alpha-\lambda} * |u|(z)}{|x-z|^{n-\lambda}} dz \\ &\leq Cr^\lambda M(g_{\alpha-\lambda} * |u|)(x) \end{aligned}$$

by Propositions 6.1 and 3.2. ■

7. Taylor estimates. The results in this section extend the lemmas of the previous section to Hölder-type estimates involving Taylor polynomials of f . The main result of the section, Theorem 7.2, is used to verify the Hölder continuity of the approximator g in the proof of Theorem 1.2. The result of Theorem 7.3, although not used in what follows, is of independent interest as it gives a fractional analogue to the pointwise inequality of Bojarski and Hajlasz (Proposition 5.1 above).

Given a function $f = g_\alpha * u \in L^{\alpha,p}(\mathbb{R}^n)$ and a multi-index $|\sigma| \leq m$ we use the convention that u_σ is the function in $L^p(\mathbb{R}^n)$ satisfying $D^\sigma f = g_{\alpha-|\sigma|} * u_\sigma \in L^{\alpha-|\sigma|,p}(\mathbb{R}^n)$.

LEMMA 7.1. *Let $f = g_\alpha * u \in L^{\alpha,p}(\mathbb{R}^n)$ and assume that $m \geq 1$. There is a constant $C = C_{\lambda,m,n}$ such that*

$$|f(x) - T_{B(x,r)}^m f(x)| \leq Cr^{m+\lambda} \sum_{|\sigma|=m} M(g_{\alpha-m-\lambda} * |u_\sigma|)(x)$$

and

$$\int_{B(x,r)} |f(y) - T_{B(x,r)}^m f(y)| dy \leq Cr^{m+\lambda} \sum_{|\sigma|=m} M(g_{\alpha-m-\lambda} * |u_\sigma|)(x)$$

for $(\alpha - m, p)$ -quasi-every $x \in \mathbb{R}^n$ and all $r > 0$.

Proof. Let x be a Lebesgue point of f and of each of its partial derivatives with the property that $g_{\alpha-|\sigma|} * |u_\sigma| < \infty$ for all $|\sigma| \leq m$. Observe that $(\alpha - m, p)$ -quasi-every $x \in \mathbb{R}^n$ has this property. Let $r > 0$ and write $B = B(x, r)$. Fixing $|\sigma| = m$ we apply Lemma 6.8 to $D^\sigma f$ to get

$$\int_B \frac{|D^\sigma f(x) - D^\sigma f(z)|}{|x - z|^{n-m}} dw \leq Cr^{m+\lambda} M(g_{\alpha-m-\lambda} * |u_\sigma|)(x).$$

The first statement follows from Proposition 3.1 with $a_\sigma = D^\sigma f(x)$. To prove the second statement we integrate the inequality in Proposition 3.1 over B to get

$$\int_B |f(y) - T_B^m f(y)| dy \leq C \sum_{|\sigma|=m} \int_B \int_B \frac{|a_\sigma - D^\sigma f(z)|}{|y - z|^{n-m}} dz dy$$

for any family $(a_\sigma)_{|\sigma|=m}$ of real numbers. Tonelli's theorem and Proposition 3.2 imply

$$\int_B |f(y) - T_B^m f(y)| dx \leq Cr^m \sum_{|\sigma|=m} \int_B |a_\sigma - D^\sigma f(z)| dz,$$

and the result follows from Lemma 6.9 with $a_\sigma = D^\sigma f(x)$. ■

THEOREM 7.2. *Let $f = g_\alpha * u \in L^{\alpha,p}(\mathbb{R}^n)$. There is a constant $C = C_{\lambda,m,n}$ such that*

$$r^{-m-\lambda} \int_{B(x,r)} |f(y) - T_x^m f(y)| dy \leq C \sum_{|\sigma|=m} M(g_{\alpha-m-\lambda} * |u_\sigma|)(x)$$

for $(\alpha - m, p)$ -quasi-every $x \in \mathbb{R}^n$ and all $r > 0$.

Proof. When $m = 0$ the statement of Theorem 7.2 is simply Lemma 6.9, so we assume $m \geq 1$. Let $x \in \mathbb{R}^n$ be a point satisfying the conclusion of Lemma 7.1 for f and for all $D^\beta f \in L^{\alpha-|\beta|,p}(\mathbb{R}^n)$, $|\beta| \leq m$. Notice that $(\alpha - m, p)$ -quasi-every $x \in \mathbb{R}^n$ has this property. Let $r > 0$ and $B = B(x, r)$. In light of the second conclusion of Lemma 7.1 it suffices to show that

$$\int_{B(x,r)} |T_x^m f(y) - T_B^m f(y)| dy \leq Cr^{m+\lambda} \sum_{|\sigma|=m} M(g_{\alpha-m-\lambda} * |u_\sigma|)(x).$$

As in (3.2) above we use (2.8) to obtain

$$(7.1) \quad |T_x^m f(y) - T_B^m f(y)| \leq \sum_{|\beta| \leq m} r^{|\beta|} |D^\beta f(x) - T_B^{m-|\beta|} D^\beta f(x)|$$

for $(\alpha - m, p)$ -quasi-every $y \in B$, and integrating over B we deduce from this that

$$\int_B |T_x^m f(y) - T_B^m f(y)| dy \leq \sum_{|\beta| \leq m} r^{|\beta|} |D^\beta f(x) - T_B^{m-|\beta|} D^\beta f(x)|.$$

It suffices therefore to prove that

$$|D^\beta f(x) - T_B^{m-|\beta|} D^\beta f(x)| \leq Cr^{m-|\beta|+\lambda} \sum_{|\sigma|=m} M(g_{\alpha-m-\lambda} * |u_\sigma|)(x)$$

for all $|\beta| \leq m$. When $|\beta| < m$ this follows from Lemma 7.1 applied to $D^\beta f \in L^{\alpha-|\beta|,p}(\mathbb{R}^n)$ and when $|\beta| = m$ we use the estimate

$$|D^\beta f(x) - T_B^{m-|\beta|} D^\beta f(x)| = |D^\beta f(y) - \overline{(D^\beta f)}_B| \leq \int_B |D^\beta f(y) - D^\beta f(z)| dz$$

and Lemma 6.9. ■

THEOREM 7.3. *Let $f = g_\alpha * u \in L^{\alpha,p}(\mathbb{R}^n)$. There is a constant $C = C_{\lambda,m,n}$ such that $|f(x) - T_y^m f(x)|$ does not exceed*

$$C|x - y|^{m+\lambda} \sum_{|\sigma|=m} (M(g_{\alpha-m-\lambda} * |u_\sigma|)(x) + M(g_{\alpha-m-\lambda} * |u_\sigma|)(y))$$

for $(\alpha - m, p)$ -quasi-every $x, y \in \mathbb{R}^n$.

Proof. When $m = 0$ the statement of Theorem 7.3 is simply Lemma 6.5, so we assume $m \geq 1$ and let $x, y \in \mathbb{R}^n$ be two points satisfying the conclusion

of Lemma 7.1. Let B be a ball with radius $|x - y|$ containing x and y . By the first result of Lemma 7.1 it suffices to prove that

$$|T_y^m f(x) - T_B^m f(x)| \leq C|x - y|^{m+\lambda} \sum_{|\sigma|=m} M(g_{\alpha-m-\lambda} * |u_\sigma|)(y).$$

This follows from (7.1) above (with x and y interchanged), the second conclusion of Lemma 7.1, and the proof of Theorem 7.2. ■

8. The proof of Theorem 1.2, $0 < \lambda < 1$. We assume that $\alpha > 0$, m is an integer with $0 \leq m \leq \alpha - 1$, and that $0 < \lambda < 1$. Let $f \in L^{\alpha,p}(\mathbb{R}^n)$ and fix $\varepsilon > 0$.

The set Ω is constructed as in Section 5 with only a minor modification. Let Ω_1 be an open set with $B_{\alpha-m,p}(\Omega_1) < \varepsilon/4$ satisfying conditions (5.1) and (5.2) above. Apply Theorem 7.2 to find an open set Ω_2 with $B_{\alpha-m,p}(\Omega_2) < \varepsilon/4$ such that

$$\int_{B(x,r)} |f(y) - T_x^m f(y)| dy \leq Cr^{m+\lambda} \sum_{|\sigma|=m} M(g_{\alpha-m-\lambda} * |u_\sigma|)(x)$$

for all $x \in \Omega_2$ and $r > 0$. By (2.4) there is an open set Ω_3 with $B_{\alpha-m,p}(\Omega_3) < \varepsilon/4$ such that $g_{\alpha-|\sigma|} * |u_\sigma|(x) < \infty$ for all $x \in \mathbb{R}^n \setminus \Omega_3$ and $|\sigma| \leq m$. Observe that Lemma 6.5 implies that

$$|D^\sigma f(x) - D^\sigma f(y)| \leq C|x - y|^\lambda (M(g_{\alpha-m-\lambda} * |u_\sigma|)(x) + M(g_{\alpha-m-\lambda} * |u_\sigma|)(y))$$

for all $x, y \in \mathbb{R}^n \setminus \Omega_3$. Finally, we define

$$E_t = \left\{ x : \sum_{|\sigma|=m} M(g_{\alpha-m-\lambda} * |u_\sigma|)(x) > t \right\}$$

for $t > 0$. Then (2.5) implies that $B_{\alpha-m-\lambda,p}(E_t) \leq Ct^{-p}$ for all t , where C depends on f but not on t . Fix t sufficiently large that $B_{\alpha-m-\lambda,p}(E_t) < \varepsilon/4$, and let $\Omega_4 = E_t$ (E_t is open owing to the lower semicontinuity of the maximal operator). Then $\sum_{|\sigma|=m} M(g_{\alpha-m-\lambda} * |u_\sigma|)(x) \leq t$ for all $x \in \mathbb{R}^n \setminus \Omega_4$.

Let $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3 \cup \Omega_4$. The monotonicity of the Bessel capacity implies that $B_{\alpha-m-\lambda,p}(\Omega) < \varepsilon$. Define $K = \mathbb{R}^n \setminus \Omega$. The construction of Ω implies that

$$(8.1) \quad r^{-m-\lambda} \int_{B(x,r)} |f(y) - T_x^m f(y)| dy \leq Ct$$

for all $x \in K$ and $r > 0$, and

$$(8.2) \quad |D^\sigma f(x) - D^\sigma f(y)| \leq Ct|x - y|^\lambda$$

for all $x, y \in K$ and $|\sigma| = m$. Since $K \subset \mathbb{R}^n \setminus \Omega_1$ we may follow Steps 2–5 in Section 5 to conclude that there is a function $g \in C^m(\mathbb{R}^n) \cap W^{m+1,p}(\mathbb{R}^n)$ such that

1. $D^\sigma g(x) = D^\sigma f(x)$ for all $x \in K$ and $|\sigma| \leq m$, and
2. $\|f - g\|_{m+1,p} < \varepsilon$

provided that $|\Omega|$ has been chosen sufficiently small. Thus to establish Theorem 1.2 it suffices to prove that the order m partial derivatives of g inherit Hölder continuity from f . Recall from Lemma 4.1 and the definition of g that we may assume

$$|D^\beta g(x) - D^\beta P(x)| \leq C\delta(x)^{-|\beta|} \int_{B(x,\delta(x))} |f(z) - P(z)| dz$$

for all $x \in \Omega$, multi-indices $|\beta| \leq m + 1$, and polynomials P whose degree does not exceed $m + 1$. Here $\delta(x)$ is the smooth distance function introduced in Section 4 satisfying $C\text{dist}(x, \partial\Omega) \leq \delta(x) \leq \text{dist}(x, \partial\Omega)$.

Fix a multi-index $|\sigma| = m$. Following the method in [4] we will prove that $D^\sigma g$ is Hölder continuous.

STEP 1. If $x, y \in K$, then

$$(8.3) \quad |D^\sigma g(x) - D^\sigma g(y)| = |D^\sigma f(x) - D^\sigma f(y)| \leq Ct|x - y|^\lambda$$

by (8.2) and the definition of g .

STEP 2. If $x \in \Omega$ and $y \in K$, let $x^* \in K$ satisfy $\text{dist}(x, K) = |x - x^*|$. Then $|x^* - y| \leq |x - y| + |x - x^*| \leq 2|x - y|$, and therefore

$$(8.4) \quad |D^\sigma g(x^*) - D^\sigma g(y)| \leq Ct|x^* - y|^\lambda \leq Ct|x - y|^\lambda$$

by Step 1. Define the order m polynomial $P(z) = T_{x^*}^m f(z)$. Then $D^\sigma P(z) = D^\sigma g(x^*)$, hence

$$\begin{aligned} |D^\sigma g(x) - D^\sigma g(x^*)| &= |D^\sigma g(x) - D^\sigma T_{x^*}^m f(x)| \\ &\leq C\delta(x)^{-|\sigma|} \int_{B(x,\delta(x))} |f(z) - T_{x^*}^m f(z)| dz. \end{aligned}$$

Since $C|x - x^*| \leq \delta(x) \leq |x - x^*|$ this implies

$$(8.5) \quad |D^\sigma g(x) - D^\sigma g(x^*)| \leq C|x - x^*|^{-m} \int_{B(x^*,2|x-x^*|)} |f(z) - T_{x^*}^m f(z)| dz,$$

and since $x^* \in K$ it follows from (8.1) that

$$\int_{B(x^*,2|x-x^*|)} |f(z) - T_{x^*}^m f(z)| dz \leq Ct|x - x^*|^{m+\lambda}.$$

Therefore

$$(8.5) \quad |D^\sigma g(x) - D^\sigma g(x^*)| \leq Ct|x - x^*|^\lambda \leq Ct|x - y|^\lambda,$$

and by combining (8.4) and (8.5) we conclude that $|D^\sigma g(x) - D^\sigma g(y)| \leq Ct|x - y|^\lambda$ whenever $x \in \Omega$ and $y \in K$.

STEP 3. To complete the proof of the theorem it only remains to verify (8.3) when $x, y \in \Omega$. Let $x, y \in \Omega$ and let $x^* \in K$ satisfy $\text{dist}(x, K) = |x - x^*|$. If $|x - x^*| \leq 2|x - y|$, then

$$\begin{aligned} |D^\sigma g(x) - D^\sigma g(y)| &\leq |D^\sigma g(x) - D^\sigma g(x^*)| + |D^\sigma g(y) - D^\sigma g(x^*)| \\ &\leq Ct(|x - x^*| + |y - x^*|) \end{aligned}$$

by Step 2 since $x^* \in K$. But $|x - x^*| \leq 2|x - y|$ and $|y - x^*| \leq |x - y| + |x - x^*| \leq 3|x - y|$, implying that $|D^\sigma g(x) - D^\sigma g(y)| \leq Ct|x - y|^\lambda$. On the other hand, when $|x - x^*| > 2|x - y|$, the line segment joining x and y must lie entirely within Ω . Since $g \in C^\infty(\Omega)$ we use the mean value theorem to find a point $w \in \Omega$ satisfying

$$(8.6) \quad |D^\sigma g(x) - D^\sigma g(y)| \leq |DD^\sigma g(w)| \cdot |x - y|$$

and $|x - y| = |x - w| + |w - y|$. Let $w^* \in K$ satisfy $\text{dist}(w, K) = |w - w^*|$ and define the order m polynomial $P(z) = T_{w^*}^m f(z)$. Then, since $D^{\sigma+\delta_j} P \equiv 0$, Lemma 4.1 implies

$$|D^{\delta_j} D^\sigma g(w)| \leq C\delta(w)^{-m-1} \int_{B(w, \delta(w))} |f(z) - T_{w^*}^m f(z)| dz$$

for all $j = 1, \dots, n$. As above it follows that

$$|D^{\delta_j} D^\sigma g(w)| \leq C|w - w^*|^{-m-1} \int_{B(w^*, 2|w - w^*|)} |f(z) - T_{w^*}^m f(z)| dz,$$

and since $w^* \in K$, (8.1) implies that the value of the integral above does not exceed $Ct|w - w^*|^{m+\lambda}$. Thus $|D^{\delta_j} D^\sigma g(w)| \leq Ct|w - w^*|^{\lambda-1}$ for all $j = 1, \dots, n$, hence $|DD^\sigma g(w)| \leq Ct|w - w^*|^{\lambda-1}$. Now we use the geometry of the situation. Since $2|x - y| < |x - x^*|$ and $|x - w| \leq |x - y|$ we have $|x - y| = 2|x - y| - |x - y| \leq |x - x^*| - |x - w|$, where $|x - x^*| \leq |x - w^*| \leq |x - w| + |w - w^*|$. Therefore $|x - y| \leq |w - w^*|$, implying that $|DD^\sigma g(w)| \leq Ct|x - y|^{\lambda-1}$. Substituting this into (8.6), we conclude that

$$|D^\sigma g(x) - D^\sigma g(y)| \leq Ct|x - y|^\lambda$$

whenever $x, y \in \Omega$. Therefore $D^\sigma g$ is Hölder continuous for all $|\sigma| = m$, implying that $g \in C^{m, \lambda}(\mathbb{R}^n)$ and completing the proof of the theorem.

9. Further results. In this section we prove a variant of Theorem 1.2 valid without the restriction $m \leq \alpha - 1$. Here the Sobolev norm $\|f\|_{m+1, p}$ is replaced by the Besov norm $\|f\|_{m+\lambda, p, \infty}$, defined as the sum of $\|f\|_{m, p}$ and

$$\sum_{|\sigma|=m} \left(\sup_{\substack{h \in \mathbb{R}^n \\ |h|>0}} |h|^{-\lambda} \left(\int_{\mathbb{R}^n} |D^\sigma f(x+h) - D^\sigma f(x)|^p dx \right)^{1/p} \right),$$

where $\|f\|_{0, p} = \|f\|_p$ in case $m = 0$. When $\lambda = 0$ observe that $\|f\|_{m, p, \infty} \approx \|f\|_{m, p}$.

THEOREM 9.1. *Assume that $1 < p < \infty$ and that $0 \leq m + \lambda < \alpha$, where $m \geq 0$ is an integer and $0 \leq \lambda < 1$. For a function $f \in L^{\alpha,p}(\mathbb{R}^n)$ and $\varepsilon > 0$, there exist a function $g \in C^{m,\lambda}(\mathbb{R}^n)$ and an open set Ω with the following properties:*

- (1) $B_{\alpha-m-\lambda,p}(\Omega) < \varepsilon$,
- (2) $D^\sigma f(x) - D^\sigma g(x)$ for all $x \in \mathbb{R}^n \setminus \Omega$ and multi-indices $|\sigma| \leq m$, and
- (3) $\|f - g\|_{m+\lambda,p,\infty} < \varepsilon$.

The case $\lambda = 0$ of Theorem 9.1 was proven by Stocke [14] so we shall assume $0 < \lambda < 1$. When $m \leq \alpha - 1$, Theorem 9.1 is weaker than Theorem 1.2. Specifically, there is a constant C independent of f such that

$$(9.1) \quad \|f\|_{m+\lambda,p,\infty} \leq C\|f\|_{m+1,p}$$

for all $f \in W^{m+1,p}(\mathbb{R}^n)$. This follows from well known inclusion relations between the Sobolev and Besov spaces ([13, Ch. V.5]), but it is possible to establish (9.1) using the estimates of Section 6. Namely, if we write $D^\sigma f = g_1 * u_\sigma \in L^{1,p}(\mathbb{R}^n)$ for $f \in W^{m+1,p}(\mathbb{R}^n)$ and $|\sigma| = m$, Lemma 6.5 implies that

$$|D^\sigma f(x+h) - D^\sigma f(x)| \leq C|h|^\lambda (M(g_{1-\lambda} * |u_\sigma|)(x) + M(g_{1-\lambda} * |u_\sigma|)(y))$$

for almost all $x \in \mathbb{R}^n$. Thus integrating over \mathbb{R}^n , we get

$$|h|^{-\lambda} \left(\int_{\mathbb{R}^n} |D^\sigma f(x+h) - D^\sigma f(x)|^p dx \right)^{1/p} \leq C \|M(g_{1-\lambda} * |u_\sigma|)\|_p$$

for all $h \neq 0$. By the Hardy–Littlewood–Wiener theorem and (1.1) we have

$$\|M(g_{1-\lambda} * |u_\sigma|)\| \leq C \|g_{1-\lambda} * |u_\sigma|\|_p \leq C \|u_\sigma\|_p \leq C \|D^\sigma f\|_{1,p}$$

for each $|\sigma| = m$, and therefore

$$\|f\|_{m+\lambda,p,\infty} \leq \|f\|_{m,p} + C \sum_{|\sigma|=m} \|D^\sigma f\|_{1,p} \leq C\|f\|_{m+1,p}.$$

We now prove Theorem 9.1 using the notation of Section 8. Since $m + \lambda < \alpha$ we may proceed as above to find an open set Ω with $B_{\alpha-m-\lambda,p}(\Omega) < \varepsilon$ and a function $g \in C^{m,\lambda}(\mathbb{R}^n)$ such that $D^\sigma g(x) = D^\sigma f(x)$ for all $x \in \mathbb{R}^n \setminus \Omega$ and $|\sigma| \leq m$ and $\|f - g\|_{m,p} < \varepsilon/2$. This is the best possible approximation in Sobolev norm since, in general, $f \notin W^{m+1,p}(\mathbb{R}^n)$. The Hölder condition on the order m partial derivatives of g is given by

$$|D^\sigma g(x) - D^\sigma g(y)| \leq Ct|x - y|^\lambda,$$

where t was chosen sufficiently small that $B_{\alpha-m-\lambda,p}(E_t) < \varepsilon/4$. We assume without loss of generality that

$$\frac{1}{t^{p+1}} < \frac{\varepsilon}{4}$$

and that each $\Omega_i, i = 1, 2, 3$, was chosen so as to satisfy $B_{\alpha-m,p}(\Omega_i) < t^{-p-1}$.

For convenience we write $w_\sigma = M(g_{\alpha-m-\lambda} * |u_\sigma|)$ and $\Omega = \Omega_t$ to emphasize the dependence on t . Since the Lebesgue measure of a set does not exceed its Bessel capacity we have

$$|\Omega_t| \leq |\Omega_1| + |\Omega_2| + |\Omega_3| + |E_t| \leq 3t^{-p-1} + \left| \left\{ x : \sum_{|\sigma|=m} w_\sigma(x) > t \right\} \right|.$$

Since each function w_σ belongs to $L^p(\mathbb{R}^n)$ this implies $\lim_{t \rightarrow \infty} t^p |\Omega_t| = 0$, hence

$$(9.2) \quad \lim_{t \rightarrow \infty} \left(\sum_{|\sigma|=m} \|w_\sigma\|_{L^p(\Omega_t)} + t|\Omega_t|^{1/p} \right) = 0.$$

Fix a non-zero $h \in \mathbb{R}^n$ and let $|\sigma| = m$. Then

$$|D^\sigma f(x+h) - D^\sigma f(x)| \leq C|h|^\lambda (w_\sigma(x+h) + w_\sigma(x))$$

for almost every $x \in \mathbb{R}^n$ by Lemma 6.5 and

$$|D^\sigma g(x+h) - D^\sigma g(x)| \leq Ct|h|^\lambda$$

for all $x \in \mathbb{R}^n$ by the Hölder condition on g . Set $v_\sigma(x) = D^\sigma f(x) - D^\sigma g(x)$. Then

$$|v_\sigma(x+h) - v_\sigma(x)| \leq C|h|^\lambda (w_\sigma(x+h) + w_\sigma(x) + t)$$

for almost all $x \in \mathbb{R}^n$. Since $v_\sigma(x) = 0$ and $w_\sigma(x) \leq t$ for all $x \in \mathbb{R}^n \setminus \Omega_t$, this implies

$$\begin{aligned} |v_\sigma(x+h) - v_\sigma(x)| &\leq C|h|^\lambda (w_\sigma(x+h) + 2t)\chi_{\Omega_t}(x+h) + (w_\sigma(x) + 2t)\chi_{\Omega_t}(x). \end{aligned}$$

Integrating over \mathbb{R}^n and using the translation invariance of the integral implies

$$\left(\int_{\mathbb{R}^n} |v_\sigma(x+h) - v_\sigma(x)|^p dx \right)^{1/p} \leq C|h|^\lambda (\|w_\sigma\|_{L^p(\Omega_t)} + t|\Omega_t|^{1/p}).$$

In light of (9.2) it is possible to have chosen t sufficiently large that

$$\sum_{|\sigma|=m} |h|^{-\lambda} \left(\int_{\mathbb{R}^n} |v_\sigma(x+h) - v_\sigma(x)|^p dx \right)^{1/p} < \frac{\varepsilon}{2},$$

in which case we conclude that $\|f - g\|_{m+\lambda,p,\infty} < \varepsilon$, as desired.

References

- [1] D. R. Adams and L. I. Hedberg, *Function Spaces and Potential Theory*, Grundlehren Math. Wiss. 314, Springer, 1996.
- [2] T. Bagby and W. P. Ziemer, *Pointwise differentiability and absolute continuity*, Trans. Amer. Math. Soc. 191 (1974), 129–148.
- [3] B. Bojarski and P. Hajłasz, *Pointwise inequalities for Sobolev functions and some applications*, Studia Math. 106 (1993), 77–92.

- [4] B. Bojarski, P. Hajłasz and P. Strzelecki, *Pointwise inequalities for Sobolev functions revisited*, preprint.
- [5] A. P. Calderón and A. Zygmund, *Local properties of solutions of elliptic partial differential equations*, *Studia Math.* 20 (1961), 171–225.
- [6] H. Federer, *Geometric Measure Theory*, Springer, 1969.
- [7] O. Frostman, *Potentiel d'équilibre et capacité des ensembles avec quelques applications à la théorie des fonctions*, *Medd. Lunds Univ. Mat. Sem.* 3 (1935), 1–118.
- [8] P. Hajłasz, *Sobolev spaces on an arbitrary metric space*, *Potential Anal.* 5 (1996), 403–415.
- [9] L. I. Hedberg, *On certain convolution inequalities*, *Proc. Amer. Math. Soc.* 36 (1972), 505–510.
- [10] B. Malgrange, *Ideals of Differentiable Functions*, Oxford Univ. Press, 1966.
- [11] J. Malý, *Hölder type quasicontinuity*, *Potential Anal.* 2 (1993), 249–254.
- [12] J. L. Michael and W. P. Ziemer, *A Lusin type approximation of Sobolev functions by smooth functions*, in: *Contemp. Math.* 42, Amer. Math. Soc., 1985, 135–167.
- [13] E. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton Univ. Press, 1970.
- [14] B. M. Stocke, *A Lusin type approximation of Bessel potentials and Besov functions by smooth functions*, *Math. Scand.* 77 (1995), 60–70.
- [15] D. Swanson and W. P. Ziemer, *Sobolev functions whose inner trace at the boundary is zero*, *Ark. Mat.* 37 (1999), 373–380.
- [16] A. Torchinsky, *Real-Variable Methods in Harmonic Analysis*, Academic Press, 1986.
- [17] H. Whitney, *Analytic extensions of differentiable functions defined in closed sets*, *Trans. Amer. Math. Soc.* 36 (1934), 63–89.
- [18] W. P. Ziemer, *Weakly Differentiable Functions*, Grad. Texts in Math. 120, Springer, 1989.

Department of Mathematics
Texas A&M University
College Station, TX 77843, U.S.A.
E-mail: dswanson@math.tamu.edu

Received January 22, 2001
Revised version September 3, 2001

(4672)