STUDIA MATHEMATICA 214 (2) (2013)

Whitney type inequality, pointwise version

by

YU. A. BRUDNYI (Haifa) and I. E. GOPENGAUZ (Moscow)

Abstract. The main result of the paper estimates the asymptotic behavior of local polynomial approximation for L_p functions at a point via the behavior of μ-differences, a generalization of the kth difference. The result is applied to prove several new and extend classical results on pointwise differentiability of L_p functions including Marcinkiewicz–Zygmund's and M. Weiss' theorems. In particular, we present a solution of the problem posed in the 30s by Marcinkiewicz and Zygmund.

1. Introduction. The main result of the present paper, Theorem 3.1, compares the asymptotics at a point of local polynomial approximations for $L_p^{\mathrm{loc}}(\mathbb{R}^d)$ functions and of a general difference characteristic named μ-difference (μ is a discrete measure, see Section 2 for definitions).

The first global result of this kind was due to Marchaud [14, p. 379] and sharpened by Whitney [19]. A multivariate result for a wide class of p-integrable functions was then established in [4] and named after Whitney. It asserts that if $f \in L_p(G)$, $1 \le p \le \infty$, and $G \subset \mathbb{R}^d$ is a convex domain then for every $k \in \mathbb{N}$ there exists a polynomial m in $x \in \mathbb{R}^d$ of degree $k-1$ such that

$$\|f - m\|_{L_p(G)} \le C \sup_h \|\Delta_h^k f\|_{L_p(G_h)}; \tag{1.1}$$

here G_h is the domain of the function $x \mapsto \Delta_h^k f$.

In the present paper, we need a more general version of this result with Δ_h^k replaced by μ-difference and the supremum over h by a spherical average over h. The required result is presented in Theorem 3.2; a special case was announced in [3] but never published.

The next key result, Theorem 3.4, compares the asymptotics at a point for the two μ-difference characteristics of a function used in Theorems 3.2 and 3.1, respectively. It first appeared in [12, Lemma 3] (announced in [11]) devoted to the positive solution to the Marcinkiewicz–Zygmund problem

2010 *Mathematics Subject Classification*: Primary 26Bxx; Secondary 41A10.
Key words and phrases: μ-difference, local polynomial approximation, Taylor classes, (k,p)-differential.

DOI: 10.4064/sm214-2-5

formulated in [20, p. 7]. Since the paper [12] is now hardly available, we briefly describe the problem and the result; see Theorem 3.6 for details.

The aforementioned problem is to find a generalization to μ-differences of the classical Marcinkiewicz–Zygmund theorem [15] that if f is a measurable function on $\mathbb{R}$ and $\limsup_{h\to 0} |\Delta_h^k f(x)|/|h|^k$ is bounded for all x from a measurable set S, then f has the (Peano) k-differential almost everywhere on S. This problem was solved in [16] for $k = 2$; in this case the μ-difference is of the form

$$\mu_h(f;x) := \sum_{j=0}^{2} \mu_j f(x + \lambda_j h)$$

where the measure $\mu = \sum_{j=0}^{2} \mu_j \delta_{\lambda_j}$ annihilates polynomials of degree 1 but $\mu(x^2) \neq 0$. The problem for $k > 2$ was formulated in [16] (see also [20]) and solved in [12].

Our main result, Theorem 3.1, allows one to reduce pointwise differentiability problems for μ-differences to analogous problems for local polynomial approximations. Since the latter problems were solved in [7, Appendix III] (announced in [6]), one obtains a new approach to the former. The results obtained are related to the more general field of so-called Taylor classes introduced in the classical paper by Calderón and Zygmund [10]; see Theorems 3.5 and 3.6 below.

The reader may wonder why the results announced long ago appear after a long period. This is explained by the peculiarities of the scientific life in the former Soviet Union; see, e.g., the letter [1].

2. Basic definitions

2.1. Classes of measures. Let $\mathcal{M}$ denote a class of discrete measures μ on $\mathbb{R}$ with support $\operatorname{supp}\mu$ satisfying

$$0 \in \operatorname{supp}\mu, \quad 1 < \operatorname{card}(\operatorname{supp}\mu) < \infty, \tag{2.1}$$

i.e., $\mu \in \mathcal{M}$ is a linear combination with nonzero coefficients of δ-measures including δ_0 (here $\delta_a f := f(a)$, $a \in \mathbb{R}$).

Moreover, $\mathcal{M}_k$ is a subclass of measures $\mu \in \mathcal{M}$ such that

$$\int_{\mathbb{R}} t^j \, d\mu(t) = \begin{cases} 0, & j = 0, 1, \dots, k-1, \\ c(\mu) \neq 0, & j = k. \end{cases} \tag{2.2}$$

The smallest closed interval containing $\operatorname{supp}\mu$ is denoted by $I(\mu)$ or $[a(\mu), b(\mu)]$, i.e.,

$$I(\mu) = \operatorname{conv}(\operatorname{supp}\mu) = [a(\mu), b(\mu)]. \tag{2.3}$$

Due to (2.1),

$$a(\mu) \leq 0, \quad b(\mu) \geq 0, \quad |I(\mu)| = |a(\mu)| + b(\mu) > 0. \tag{2.4}$$

2.2. μ-differences. A *μ-difference* is a linear operator on the space $C(\mathbb{R}^d)$ defined by

$$\mu_h(f;x) := \int_{\mathbb{R}} f(x+th)\,d\mu(t), \qquad x,h \in \mathbb{R}^d. \tag{2.5}$$

E.g., for $\mu = \Delta_1^k := \sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j}\delta_j$ the μ_h is the *k-difference* Δ_h^k, i.e.,

$$\Delta_h^k f(x) := \sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j} f(x+jh). \tag{2.6}$$

Some results use a symmetric version of (2.6) given by

$$\delta_h^k := \frac{1}{2}(\Delta_h^k - (-1)^k\Delta_{-h}^k). \tag{2.7}$$

The associated measures for (2.6) and (2.7) clearly belong to $\mathcal{M}_k$ with $c(\Delta_1^k) = c(\delta_1^k) = k!$.

More generally, let $f(t_0, t_1, \ldots, t_k)$, where $t_0 < t_1 < \cdots < t_k$, be the kth *divided difference* of $f \in C(\mathbb{R})$, i.e.

$$f(t_0, t_1, \ldots, t_k) := \sum_{j=0}^{k} \mu_j f(t_j), \tag{2.8}$$

where $\mu_j := 1/\omega'(t_j)$ with $\omega(t) := \prod_{j=0}^{k}(t-t_j)$ and suppose the associated measure $\mu := \sum_{j=0}^{k}\mu_j\delta_{t_j}$ belongs to $\mathcal{M}_k$ (see, e.g., [13, Ch. 4, §7]). The class of such measures is described by the conditions (see [17])

$$(-1)^{k-j}\mu(t_j) > 0, \qquad 0 \le j \le k,$$
$$(-1)^{k-j}\int_{t_j}^{\infty} d\mu(t) > 0, \qquad 1 \le j \le k, \qquad \int_{\mathbb{R}} d\mu(t) = 0.$$

It is easy to see that for $\mu \in \mathcal{M}_k$ and $f(x) = x^\alpha$ with $|\alpha| = k$ we have $\mu_h(f;x) = c(\mu)h^\alpha$ for all $x, h \in \mathbb{R}^d$ (1).

2.3. Local polynomial approximation. The object in the title, denoted by $E_{k,p}(f;S)$, is a function on pairs (f,S) where $f \in L_p^{\mathrm{loc}}(\mathbb{R}^d)$ and $S \subset \mathbb{R}^d$ with d-measure $0 < |S| < \infty$. It is defined by

$$E_{k,p}(f;S) := \inf_g \|f-g\|_{L_p(S)} \tag{2.9}$$

where g runs over the space $\mathcal{P}_{k-1}(\mathbb{R}^d)$ of polynomials in $x \in \mathbb{R}^d$ of degree $k-1$.

---

(1) Hereafter we use standard notations, e.g., for $x \in \mathbb{R}^d$ and $\alpha \in \mathbb{N}_0^d$, $x^\alpha := \prod_{j=1}^{d} x_j^{\alpha_j}$, $|\alpha| := \sum_{j=1}^{d}\alpha_j$ and $\alpha! := \prod_{j=1}^{d}(\alpha_j!)$.

We also use the *normalized* local polynomial approximation defined by

$$\mathcal{E}_{k,p}(f;S) := |S|^{-1/p} E_{k,p}(f;S). \tag{2.10}$$

2.4. μ-modulus of continuity and μ-oscillation. Let $\mu \in \mathcal{M}_k$ and G be a domain (an open connected set) in $\mathbb{R}^d$. The *μ-modulus of continuity* is a function on pairs (f, r) where $f \in L_p(G)$ and $r \in \mathbb{R}_+$, given by

$$\mu_p(r; f; G) := \sup_{h \in B_r} \|\mu_h(f, \cdot)\|_{L_p(G_h)}; \tag{2.11}$$

hereafter $B_r(x) := \{y \in \mathbb{R}^d;\ \|y - x\| < r\}$ stands for the Euclidean ball in $\mathbb{R}^d$ of radius r and center x and $B_r := B_r(0)$, while G_h denotes the domain of the function $x \mapsto \mu_h(f; x)$, i.e. the set

$$G_h := \{x \in G;\ \{x + ht;\ t \in \operatorname{supp} \mu\} \subset G\}. \tag{2.12}$$

This can be empty for large h, e.g., for G bounded. We stipulate $\|f\|_{L_p(\emptyset)} = 0$; therefore μ_p is constant on $[r_\mu, \infty)$ where r_μ is the smallest $r > 0$ such that $G_h = \emptyset$ if $h \notin B_r$. Note that if $\mu = \Delta_1^k$, then (2.11) becomes the classical k-modulus of continuity $\omega_{k,p}$ defined by

$$\omega_{k,p}(r; f; G) := \sup_{h \in B_r} \|\Delta_h^k(f, \cdot)\|_{L_p(G_h)}.$$

We will also use the *spherical μ-modulus of continuity* denoted by $\tilde{\mu}_p$ and defined by

$$\tilde{\mu}_p(r; f; G) := \Big\{ \int_{B_r} \|\mu_h(f, \cdot)\|_{L_p(G_h)}^p \, dh \Big\}^{1/p}. \tag{2.13}$$

Clearly, $\tilde{\mu}_p(r; f; G)|B_r|^{-1/p} \le \mu_p(f; r; G)$ for $p \in [1, \infty)$.

In what follows, two local versions of the μ-modulus of continuity will be of use. The first is defined by

$$\mu_p(f; B_r(x)) := \Big\{ \int_{B_r} |\mu_h(f; x)|^p \, dh \Big\}^{1/p}, \tag{2.14}$$

while the second, called the *μ-oscillation* on a subset S, is given by

$$\tilde{\mu}_p(f; S) := \Big\{ \frac{1}{|B_r|} \int_{B_r} \|\mu_h(f, \cdot)\|_{L_p(S_h)}^p \, dh \Big\}^{1/p}; \tag{2.15}$$

here r is the largest number such that $S_h \ne \emptyset$ for $\|h\| \le r$.

For $p = \infty$ and $\mu := \Delta_1^k$ this coincides with the classical kth oscillation

$$\operatorname{osc}_k(f; B) := \operatorname*{ess\,sup}_{x,h} \{|\Delta_h^k(f; x)|;\ x + jh \in B,\ j = 0, 1, \ldots, k\}.$$

2.5. Majorants. A continuous function $\omega : \mathbb{R}_+ \to \mathbb{R}_+$ is said to be a *majorant* if $\omega(+0) = 0$, ω is increasing and for some $C \ge 1$ and all $t > 0$,

$$\omega(2t) \le C\omega(t).$$

The well-known example of majorant is the k-modulus of continuity $\omega_{k,p}(t;f;G)$. Since $\mu_p(t;f;G)$ for $\mu\in\mathcal{M}_k$ and $G=\mathbb{R}^d$ is equivalent to $\omega_{k,p}(t;f;G)$ (see Theorem 3.3 and Remark 3.2 below), $\mu_p(t;f;G)$ is also a majorant.

3. Formulation of main results

3.1. Pointwise estimate of local approximation. In the subsequent formulations, f belongs to $L_p^{\mathrm{loc}}(\mathbb{R})$, $1\le p\le\infty$, and $S\subset\mathbb{R}^d$ is measurable and of finite measure $|S|>0$. Moreover, in the three results below, the measure μ belongs to $\mathcal{M}_k$. The first result estimates the normalized local polynomial approximation of a function via the spherical p-average of $\mu(f)$ (see (2.10) and (2.14), respectively).

THEOREM 3.1. *Let ω be a majorant. There exists a constant $C>0$ independent of f and x such that for almost every $x\in S$,*

$$\limsup_{r\to 0}\frac{\mathcal{E}_{k,p}(f;B_r(x))}{\omega(r)}\le C\limsup_{r\to 0}\frac{\mu_p(f;B_r(x))}{\omega(r)} \tag{3.1}$$

provided the right-hand side is finite on S.

3.2. Norm estimates. In the subsequent text, we write $C=C(a,b,\dots)$ for a positive constant depending **only** on the parameters in the brackets. It may change from line to line. Moreover, equivalence of functions $\varphi\approx\psi$ means that for some constants $C_1,C_2>0$,

$$C_1\varphi\le\psi\le C_2\varphi \tag{3.2}$$

for an explicitly indicated set of arguments of these functions.

THEOREM 3.2. *There exist constants $C=C(\mu,k,d)$ and $\lambda=\lambda(\mu,k,d)>1$ such that for every ball $B_r(x)$,*

$$\mathcal{E}_{k,p}(f;B_r(x))\le C\tilde{\mu}_p(f;B_{\lambda r}(x)). \tag{3.3}$$

The next result connects local polynomial approximations with the spherical μ-modulus of continuity on $\mathbb{R}^d$ (see (2.13)).

THEOREM 3.3. *Let $f\in L_p(\mathbb{R}^d)+\mathcal{P}_{k-1}(\mathbb{R}^d)$. Then*

$$\tilde{\mu}_p(r;f;\mathbb{R}^d)\approx\Big\{\int_{\mathbb{R}^d}\mathcal{E}_{k,p}(f;B_r(x))^p\,dx\Big\}^{1/p} \tag{3.4}$$

with constants independent of f and r.

From these two results one can derive the converse inequality, mentioned in Subsection 2.4, between the μ-modulus of continuity and its spherical counterpart. Actually, it asserts that for all $r>0$,

$$\mu_p(r;f;\mathbb{R}^d)\le C(\mu,k,d)\tilde{\mu}_p(r;f;\mathbb{R}^d)|B_r|^{-1/p}. \tag{3.5}$$

This inequality is known for $\omega_{k,p}$, i.e., for $\mu_h = \Delta_h^k$: see, e.g., [13, Ch. 6] where the proof is based on the identity

$$\Delta_h^k f(x) = \sum_{j=0}^{k} (-1)^j \binom{k}{j} [\Delta_{jh}^k f(x + j\tilde{h}) - \Delta_{h+j\tilde{h}} f(x)], \quad h, \tilde{h} \in \mathbb{R}^d.$$

This argument clearly cannot be used for (3.5).

REMARK 3.1. In Theorems 3.2 and 3.3, the class $\mathcal{M}_k$ can be widened by adding all measures satisfying the following conditions:

(a) μ has compact support with isolated point $\{0\}$;
(b) μ is orthogonal to $\mathcal{P}_{k-1}$;
(c) $c(\mu) \neq 0$;

cf. conditions (2.1) and (2.2).

The reader can easily verify that the proofs presented below are valid for this class.

REMARK 3.2. Theorem 3.2 implies that every μ-modulus μ_p with $G = \mathbb{R}^d$ and μ belonging, the above extension of $\mathcal{M}_k$ is equivalent to the kth modulus of continuity.

3.3. Pointwise estimate for spherical μ-oscillation. The key result of the present paper, Theorem 3.1, is a direct consequence of Theorem 3.2 and the estimate for $\tilde{\mu}_p$ via μ_p given below (see (2.15) and (2.14) for definitions). Unlike the previous results, μ now belongs to the wider class $\mathcal{M}$ (see §2.1).

THEOREM 3.4. *Let $\mu \in \mathcal{M}$ and let ω be a majorant. There exists a constant $C > 0$ independent of f and x such that for almost all $x \in S$,*

$$\limsup_{r\to 0} \frac{\tilde{\mu}_p(f; B_r(x))}{\omega(r)} \leq C \limsup_{r\to 0} \frac{\mu_p(f; B_r(x))}{\omega(r)} \tag{3.6}$$

provided the right-hand side is finite on S.

3.4. Pointwise differentiability of L_p functions. We will use the Taylor classes introduced by Calderón and Zygmund [10]. Let us recall that a function $f \in L_p^{\mathrm{loc}}(\mathbb{R}^d)$ belongs to the *Taylor class* $T_p^\lambda(x)$, $\lambda > 0$, $x \in \mathbb{R}^d$, if there exists a polynomial m of degree *strictly less than* λ such that

$$\|f - m;\, B_r(x)\|_p \leq C r^\lambda \tag{3.7}$$

for all $r \in (0, 1]$ and some $C > 0$. Hereafter $\|\cdot; S\|_p$ stands for the *normalized* p-norm given for $0 < |S| < \infty$ by

$$\|f; S\|_p := \left\{ \frac{1}{|S|} \int_S |f(x)|^p \, dx \right\}^{1/p}. \tag{3.8}$$

Further, the class $t_p^\lambda(x)$ is defined by replacing the right-hand side in (3.7) by $o(r^\lambda)$ as $r \to 0$, and assuming m to be a polynomial of degree *less than or equal to* λ. In particular, $t_p^\lambda(x)$ is a proper subclass of $T_p^\lambda(x)$.

It is easy to check that the polynomial m for $f \in t_p^\lambda(x)$ or $f \in T_p^\lambda(x)$ is unique; if $\lambda \in \mathbb{N}$, this polynomial is called the *Peano* (λ, p)*-differential* of f at x.

THEOREM 3.5. *Let* $\mu \in \mathcal{M}_k$. *A function* f *belongs to* $t_p^k(x)$ *for almost all points* x *of* S *if and only if*

$$\mu_p(f; B_r(x)) = O(r^k) \quad \text{as } r \to 0 \quad \text{for almost all } x \in S. \tag{3.9}$$

If $p = \infty$, (3.9) *can be written as*

$$\limsup_{h \to 0} \frac{|\mu_h(f; x)|}{\|h\|^k} < \infty \quad \text{for almost all } x \in S. \tag{3.10}$$

For $\mu_h = \Delta_h^k$ and measurable functions on $\mathbb{R}$ the last result gives Marcinkiewicz and Zygmund's classical theorem [15]. In [16], these authors extended their result to μ_h with $\mu \in \mathcal{M}_2$ and asked about the case of $\mu \in \mathcal{M}_k$ with $k \geq 2$ (see, e.g. [20, p. 7]). A positive answer in the general case following from Theorem 3.4 was announced in [11] and proved in [12].

Another consequence of Theorem 3.5 is the next result given by M. Weiss [18] in the one-dimensional case. To formulate it, we recall that a function $f \in L_p^{\mathrm{loc}}(\mathbb{R}^d)$ has the *symmetric* (k, p)*-differential* at x_0 if the associated symmetric function $\hat{f}(x) := \frac{1}{2}(f(x_0 + x) - (-1)^k f(x_0 - x))$ belongs to $t_p^k(0)$.

COROLLARY 3.1. *If* f *has the symmetric* (k, p)*-differential at all* $x \in S$, *then it has the Peano* (k, p)*-differential at almost all* $x \in S$.

THEOREM 3.6.

(a) *If* $0 < \lambda < k$ *and* λ *is noninteger or* $\lambda = k$, *then* $f \in T_p^\lambda(x)$ *for almost all* $x \in S$ *if and only if*

$$\mu_p(f; B_r(x)) = O(r^\lambda) \quad \text{as } r \to 0 \quad \text{for almost all } x \in S. \tag{3.11}$$

(b) *For noninteger* $\lambda \in (0, k)$, $f \in t_p^\lambda(x)$ *for almost all* $x \in S$ *if and only if*

$$\mu_p(f; B_r(x)) = o(r^\lambda) \quad \text{as } r \to 0 \quad \text{for almost all } x \in S. \tag{3.12}$$

Moreover, $f \in t_p^k(x)$ *for almost all* $x \in S$ *if and only if there exists a family* $\{m_x\}_{x \in S}$ *of homogeneous polynomials of degree* k *such that*

$$\left\{ \frac{1}{|B_r|} \int_{B_r} |\mu_h(f; x) - m_x(h)|^p \, dh \right\}^{1/p} = o(r^k) \quad \text{as } h \to 0 \tag{3.13}$$

for almost all $x \in S$.

4. Proofs. Since Theorem 3.1 is a direct consequence of Theorems 3.2 and 3.4, one begins with Theorem 3.2. The other one, unlike the other theorems, deals with the wider class of measures and therefore we postpone its proof to the last subsection.

4.1. Proof of Theorem 3.2. By shifting and rescaling the required inequality, (3.3) can be reduced to the case of $B_1 := B_1(0)$, i.e., to the inequality

$$\mathcal{E}_{k,p}(f;B_1) := |B_1|^{-1/p}E_{k,p}(f;B_1) \le C(\mu,k,d)\Big\{\frac{1}{|B_\lambda|}\int\limits_{B_\lambda}\|\mu_h(f)\|^p_{L_p((B_\lambda)_h)}\,dh\Big\}^{1/p}$$

where $(B_\lambda)_h := \{x\in B_\lambda;\ x+th\in B_\lambda \text{ for all } t\in I(\mu)\}$. We have

$$(B_\lambda)_h \supset B_{\tilde\lambda}$$

for $\tilde\lambda := \lambda - |I(\mu)|$. By the definition of λ given below, $\tilde\lambda > 0$ and $\lambda > 1$.

With these λ, $\tilde\lambda$ inequality (3.3) will follow from a stronger one,

$$E_{k,p}(f;B_1) \le C(\mu,k,d)\frac{1}{|B_\lambda|}\int\limits_{B_\lambda}\|\mu_h(f)\|_{L_p(B_{\tilde\lambda})}\,dh. \tag{4.1}$$

In fact, the mean on the right-hand side is bounded for $p\ge 1$ by

$$\Big\{\frac{1}{|B_\lambda|}\int\limits_{B_\lambda}\|\mu_h(f)\|^p_{L_p((B_\lambda)_h)}\,dh\Big\}^{1/p}.$$

The choice of λ and $\tilde\lambda$ will be indicated at the end of the proof via an intermediate constant λ_0 given by

$$\lambda_0 := 3 + 2(|a(\mu)| + b(\mu)) + (|a(\mu)| + b(\mu))^2. \tag{4.2}$$

To find the polynomial approximation giving (4.1) we need several auxiliary results.

Lemma 4.1. *Let*

$$e_k(t) := \frac{1}{(k-1)!}(t_+)^{k-1}$$

where $t_+ := \max\{0,t\}$ *and*

$$v(t) := (e_k * \mu)(t) = \int\limits_{\mathbb{R}} e_k(t-s)\,d\mu(s). \tag{4.3}$$

Then

(a) $\operatorname{supp} v = I(\mu)$*;*
(b) $v^{(k)} = \mu$ *(distributional derivative);*
(c) $\int_{\mathbb{R}} v(t)\,dt = (-1)^k c(\mu)/k! \ne 0$.

Proof. (a) If $t < a(\mu)$ then $t - s < 0$ for $s \in \operatorname{supp}\mu$; therefore $v(t) = 0$. Moreover, if $t > b(\mu)$ then $t - s > 0$ for $s \in \operatorname{supp}\mu$ and

$$v(t) = \int_{\mathbb{R}} \frac{(t-s)^{k-1}}{(k-1)!}\, d\mu(s) = 0.$$

(b) Since $(e_k)^{(k)} = \delta_0$, we obtain $v^{(k)} = \delta_0 * \mu = \mu$.

(c) We have

$$\int_{\mathbb{R}} v(t)\, dt = \int_{I(\mu)} dt \int_{\mathbb{R}} e_k(t-s)\, d\mu(s) = \int_{\mathbb{R}} d\mu(s) \int_{s}^{b(\mu)} \frac{(t-s)^{k-1}}{(k-1)!}\, dt$$
$$= \int_{\mathbb{R}} \frac{(b(\mu)-s)^k}{k!}\, d\mu(s) = \frac{(-1)^k}{k!} \int_{\mathbb{R}} s^k\, d\mu(s) = \frac{(-1)^k}{k!} c(\mu). \ \blacksquare$$

Now normalize v by setting

$$V(t) := \frac{k!}{(-1)^k c(\mu)} v(t + 1 + |a(\mu)|), \quad t \in \mathbb{R}. \tag{4.4}$$

Further, define a measure $\overline{\mu}$ as the shift of μ by $1 + |a(\mu)|$, i.e., for $f \in C(\mathbb{R})$,

$$\int_{\mathbb{R}} f(t)\, d\overline{\mu}(t) = \int_{\mathbb{R}} f(t + 1 + |a(\mu)|)\, d\mu(t).$$

In particular, due to (2.5),

$$\overline{\mu}_h(f; x) = \mu_h(f; x + (1 + |a(\mu)|)h)) \tag{4.4a}$$

and moreover

$$\operatorname{supp}\overline{\mu} = \operatorname{supp}\mu + (1 + |a(\mu)|). \tag{4.4b}$$

The previous lemma immediately implies

Lemma 4.2.

(a) $\operatorname{supp} V = [1, 1 + b(\mu) + |a(\mu)|]$.

(b) $V^{(k)} = \overline{\mu}$.

(c) $\int_{\mathbb{R}} V(t)\, dt = 1$.

Given $f \in L_1^{\mathrm{loc}}(\mathbb{R}^d)$ define a function φ_h, $h \in \mathbb{S}^{d-1} := \partial B_1$, as follows: assuming without loss of generality $\mu(\{0\}) = 1$, we write $\mu = \delta_0 - \hat{\mu}$, i.e.,

$$\hat{\mu}(\{s\}) = -\mu(\{s\}) \quad \text{if } s \neq 0 \quad \text{and} \quad \hat{\mu}(\{0\}) = 0.$$

Now set, for $x \in \mathbb{R}^d$,

$$\varphi_h(x) := \int_{\mathbb{R}} V(t) \Big(\int_{\mathbb{R}} f(x + \gamma s t h)\, d\hat{\mu}(s) \Big)\, dt \tag{4.5}$$

where

$$\gamma := \frac{\lambda_0 - 1}{(b(\mu) + |a(\mu)|)(b(\mu) + |a(\mu)| + 1)} \tag{4.6}$$

with λ_0 given by (4.2).

Clearly, the linear operator $L_h : f \mapsto \varphi_h$ acts in the space of locally integrable functions on $\mathbb{R}^d$. We now check that L_h maps $L_p(B_{\lambda_0})$ into $L_p(B_1)$.

If $s \in \operatorname{supp}\mu \subset [a(\mu), b(\mu)]$ and $t \in \operatorname{supp} V$ then for $\|h\| = 1$ and $\|x\| \le 1$ one has

$$\|x + \gamma sth\| \le 1 + \gamma|st| \le 1 + \gamma(b(\mu) + |a(\mu)|)(1 + b(\mu) + |a(\mu)|) = \lambda_0,$$

i.e., (4.5) defines φ_h on B_1 for every $f \in L_p(B_{\lambda_0})$.

LEMMA 4.3.

(a) *The norm of the restriction of* L_h *to* $L_p(B_{\lambda_0})$, *denoted by* $\|L_h\|$, *satisfies*

$$\|L_h\| \le C(\mu).$$

(b) $f - L_h f = \int_{\mathbb{R}} V(t)\mu_{\gamma th}(f)\, dt$.

(c) *The* k*th directional derivative* $D_h^k L_h(f)$, $h \in B_1$, *satisfies*

$$D_h^k L_h(f;x) = (-1)^{k-1} \sum_{s\in \operatorname{supp}\hat{\mu}} \frac{\mu(\{s\})}{(\gamma s)^k}\mu_{\gamma sh}(f; x + (1 + |a(\mu)|)h). \tag{4.7}$$

Note that by the definition $\operatorname{dist}(0, \operatorname{supp}\hat{\mu}) > 0$ and therefore s in (4.7) is separated from 0.

Proof. (a) and (b) are a matter of definitions.

(c) Write $x = x_h h + x^h$ where $h \in \mathbb{R}^d$ and x^h is orthogonal to h. Then consider $f_{x,h} : x_h \mapsto f(x_h h + x^h)$ where $x \in B_1$ and $h \in \mathbb{S}^{d-1}$.

We have

$$D_h L_h(f;x) = \frac{d}{dx_h}\Big[\int_{\mathbb{R}^2} V(t) f_{x,h}(x_h + \gamma st)\, d\hat{\mu}(s)\, dt\Big].$$

Changing variables and differentiating we get

$$\begin{aligned} D_h L_h(f;x) &= -\int_{\mathbb{R}^2} V'(t) f_{x,h}(x_h + \gamma st)\frac{d\hat{\mu}(s)}{\gamma s}\, dt \\ &= \sum_{s\in \operatorname{supp}\hat{\mu}} \frac{\mu(\{s\})}{\gamma s}\int_{\mathbb{R}} V'(t) f(x + \gamma sth)\, dt. \end{aligned}$$

Iterating the differentiation k times and using at the last step assertions (b) and then (a) of Lemma 4.2, we obtain (4.7). ∎

At the next stage, define a linear operator P_h, $\|h\| = 1$, to be the Taylor polynomial at 0 of degree $k - 1$ for the function $(L_h f)_{x,h}$.

Then $P_h f$ for $f \in L_1^{\mathrm{loc}}(\mathbb{R}^d)$ is a polynomial in x_h of degree $k - 1$ with coefficients depending on x^h; in particular, P_h acts in the space $L_1^{\mathrm{loc}}(\mathbb{R}^d)$.

Lemma 4.4.

(a) $D_h^k P_h f = 0$.

(b) *For any* $h, h' \in \mathbb{S}^{d-1}$,

$$P_h P_{h'} = P_{h'} P_h.$$

(c) *For every* $f \in L_p(B_{\lambda_0})$,

$$\|f - P_h f\|_{L_p(B_1)} \le \int_{\mathbb{R}} \|\mu_{th}(f)\|_{L_p(B_{\lambda_1})}\, d\nu(t); \tag{4.8}$$

here

$$d\nu := c(\mu,k)[\mathbf{1}_{\operatorname{supp} V}\, dt + d|\hat{\mu}|] \quad \textit{and} \quad \lambda_1 := 3 + |a(\mu)|.$$

Clearly, $\lambda_1 < \lambda_0$, see (4.2).

(d) P_h *maps* $L_p(B_{\lambda_0})$ *into* $L_p(B_1)$ *with norm*

$$\|P_h\| \le C(\mu,k).$$

Proof. Assertions (a) and (b) are clear.

(c) Apply to $(L_h)_{x,h}$ the Taylor formula with integral remainder to obtain

$$\begin{aligned}
\|(L_h - P_h)f\|_{L_p(B_1)} &= \frac{1}{(k-1)!} \Big\| \int_0^{x_h} \tau^{k-1} D_h^k f((x_h - \tau)h + x^h)\, d\tau \Big\|_{L_p(B_1)} \\
&\le \frac{1}{(k-1)!} \int_0^1 \tau^{k-1} \|D_h^k f(x - \tau h)\|_{L_p(B_1)}\, d\tau \\
&\le \left(\frac{1}{(k-1)!} \int_0^1 \tau^{k-1}\, d\tau \right) \|D_h^k f\|_{L_p(B_2)};
\end{aligned}$$

the last inequality follows from $B_1 - \tau h \subset B_2$ (as $\|\tau h\| \le 1$).

Now write, skipping the subscript $L_p(B_1)$,

$$\|f - P_h\| \le \|f - L_h f\| + \|(L_h - P_h)f\| \le \|f - L_h f\| + \frac{1}{k!}\|D_h^k f\|_{L_p(B_2)}.$$

Estimating the terms on the right-hand side by (b) and (c) of Lemma 4.3, respectively, we obtain the bound

$$\int_{\mathbb{R}} |V(t)|\, \|\mu_{th}(f)\|_{L_p(B_1)}\, dt + \frac{1}{k!} \sum_{s \in \operatorname{supp} \hat{\mu}} \frac{|\mu(\{s\})|}{(\gamma |s|)^k} \|\mu_{\gamma s h}(f; x + (1 + |a(\mu)|)h)\|_{L_p(B_2)}.$$

Since $B_2 + (1 + |a(\mu)|)h \subset B_{\lambda_1}$ with $\lambda_1 := 3 + |a(\mu)|$, one concludes that this is bounded by the right-hand side of (4.8), with

$$C(\mu,k) := \frac{1}{k!} \sup_{s \in \operatorname{supp} \hat{\mu}} \frac{|\mu(\{s\})|}{(\gamma|s|)^k}.$$

(d) Writing

$$\|P_h f\|_{L_p(B_1)} \le \|f - P_h f\|_{L_p(B_1)} + \|f\|_{L_p(B_1)}$$

and using (4.8) for the first summand, we get

$$\|P_h f\|_{L_p(B_1)} \le \int_{\mathbb{R}} \|\mu_{th} f\|_{L_p(B_{\lambda_1})}\, d\nu(t) + \|f\|_{L_p(B_1)}.$$

It remains to show that the norm in the integrand is bounded by $C(\mu, k)\|f\|_{L_p(B_{\lambda_0})}$. By definition of μ_h (see (2.5)),

$$\|\mu_{th} f\|_{L_p(B_{\lambda_1})} \le \sum_{s\in \operatorname{supp}\mu} |\mu(\{s\})|\, \|f\|_{L_p(B_{\lambda_1}+\, sth)}$$

and the result will follow from the embedding

$$B_{\lambda_1} + sth \subset B_{\lambda_0} \quad \text{where} \quad t \in \operatorname{supp}\nu,\ s \in \operatorname{supp}\mu \text{ and } \|h\| = 1. \tag{4.9}$$

In fact, $\operatorname{supp}\nu = (\operatorname{supp}V)\cup(\operatorname{supp}\hat{\mu}) \subset [1, 1+|a(\mu)|+b(\mu)]\cup[a(\mu), b(\mu)]$, i.e., $|t| \le 1 + |a(\mu)| + b(\mu)$. Moreover, $|s| \le |I(\mu)| = |a(\mu)| + b(\mu)$, and $\lambda_1 = 3 + |a(\mu)|$. This implies

$$\lambda_1 + \|tsh\| \le 3 + 2(|a(\mu)| + b(\mu)) + (|a(\mu)| + b(\mu))^2 = \lambda_0$$

(see (4.2)). Assertion (d) is proved. ∎

Let $E = \{e_j\}_{1\le j\le d}$ be the standard orthonormal basis of $\mathbb{R}^d$ and

$$n(k, d) := \operatorname{card}\{x^\alpha;\ |\alpha| = k\} = \binom{k+d-1}{d-1}. \tag{4.10}$$

Denote by $\mathcal{H}_k$ the class of finite subsets $H \subset \mathbb{S}^{d-1}$ satisfying the conditions

$$E \subset H \quad \text{and} \quad \operatorname{card} H = n(k, d). \tag{4.11}$$

Due to the Kemperman identity (see, e.g., [9, p. 170] for every α with $|\alpha| = k$ there exists a subset $H(\alpha) \in \mathcal{H}_k$ such that the α-derivatives satisfy

$$D^\alpha = \sum_{h\in H(\alpha)} a_h D_h^k. \tag{4.12}$$

Now set $H(k) := \bigcup_{|\alpha|=k} H(\alpha)$. Define a linear operator P by

$$P := \prod_{h\in H(k)} P_h. \tag{4.13}$$

Lemma 4.5.

(a) $Pf \in \mathcal{P}_{k-1}(\mathbb{R}^d)$; *in particular,*

$$E_{k,p}(f; B_1) \le \|f - Pf\|_{L_p(B_1)}. \tag{4.14}$$

(b) *The inequality*

$$\|f - P_h f\| \le C \sum_{h \in H(k)} \int_{\mathbb{R}} \|\mu_{th}(f)\|_{L_p(B_{\lambda_0})}\, d\nu(t) \tag{4.15}$$

holds with a constant $C = C(k, \mu, d)$ *and* ν *defined in* (4.8).

Proof. (a) As $P_h P_{h'} = P_{h'} P_h$ one can write

$$P = \Big(\prod_{h \in E} P_h\Big)\Big(\prod_{h \in H(k)\setminus E} P_h\Big).$$

Since for $g \in L_1^{\mathrm{loc}}(\mathbb{R}^d)$, the function $P_h g$ with $h = e_j$ is a polynomial of degree $k-1$ in x_j with coefficients depending only on x_i, $i \ne j$, and the P_h commute, the function $\big(\prod_{h \in E} P_h\big) g$ is a polynomial in $x \in \mathbb{R}^d$. Therefore Pf is a polynomial too.

It remains to show that $Pf \in \mathcal{P}_{k-1}(\mathbb{R}^d)$, i.e.,

$$D^\alpha P f = 0 \quad \text{for every } |\alpha| = k. \tag{4.16}$$

Writing

$$Pf = \Big(\prod_{h \in H(\alpha)} P_h\Big)\Big(\prod_{h \in H(k)\setminus H(\alpha)} P_h\Big) f$$

and using (4.12) one has

$$D^\alpha P f = \Big(\sum_{h \in H(\alpha)} a_h (D_h^k P_h)\Big)\Big(\prod_{h \in H(k)\setminus H(\alpha)} P_h\Big) f.$$

Since $D_h^k P_h = 0$ by definition of P_h, equality (4.16) is proved.

(b) Enumerate $H(k)$ as $\{h_j\}_{1 \le j \le n}$ where $n := \operatorname{card} H(k) = [\operatorname{card}\{\alpha;\ |\alpha| = k\}]^2$.

Further, let P_j denote P_h with $h = h_j$ and write

$$f - Pf = (f - P_1 f) + P_1(f - P_2 f) + \cdots + \Big(\prod_{1 \le j < n} P_j\Big)(f - P_n f).$$

Then write

$$\Big(\prod_{1 \le i < j} P_i\Big)\Big|_{L_p(B_1)} = \prod_{1 \le i < j} P_{ij} \tag{4.17}$$

where P_{ij} is the restriction of P_i to $L_p(B_{\lambda_0^{j-i-1}})$. The rescaled version of Lemma 4.4 with B_1 replaced by B_r asserts that P_{ij} acts from $L_p(B_{\lambda_0^{j-i-1}})$ into $L_p(B_{\lambda_0^{j-i}})$ with norm bounded by $C(k, \mu)$. Together with (4.17) this implies

$$\|f - Pf\|_{L_p(B_1)} \le C(k, \mu)^n \sum_{h \in H(k)} \|f - P_h f\|_{L_p(B_1)}.$$

Finally we estimate the summands above by using (4.8) to obtain the required inequality (4.15) with $C := C(k,\mu)^n$ and $n = [\operatorname{card}\{\alpha;\, |\alpha| = k\}]^2 = \binom{k+d-1}{d-1}^2$. ∎

Let us now estimate each summand of (4.15),

$$J(h) := \int_{\mathbb{R}} \|\mu_{th}(f)\|_{L_p(B_{\lambda_0})}\, d\nu(t), \tag{4.18}$$

in order to obtain the final result (4.1). To this end we need the following known fact (see, e.g., [2, Ch. III]).

Let $d\chi$ be the normalized Haar measure on the orthogonal group $O(d)$) acting on $\mathbb{R}^d$. Then

$$\int_{O(d)} g(\chi^{-1})(x)\, d\chi = \frac{1}{\sigma_d} \int_{y\in\mathbb{S}^{d-1}} g(\|x\|y)\, dy \tag{4.19}$$

where $\sigma_d := \operatorname{vol}\mathbb{S}^{d-1}$.

Note that inequality (4.15) also holds with the set $\chi^{-1}(H(k))$, $\chi \in O(d)$, substituted for $H(k)$ (with the orthonormal basis $\chi^{-1}(E) = \{\chi^{-1}(e_j)\}_{1\le j\le d}$). Hence,

$$E_{k,p}(f; B_1) \le C(k,\mu,d) \sum_{h\in H(k)} J(\chi^{-1}(h)).$$

Now integrate this over $O(d)$ using (4.19) to get

$$E_{k,p}(f; B_1) \le \frac{C(k,\mu,d)}{\sigma_j} \operatorname{card} H(k) \int_{\|y\|=1} J(y)\, dy.$$

Then use (4.18) to rewrite this as

$$E_{k,p}(f; B_1) \le C \int_{\mathbb{R}} d\nu(s) \int_{\|y\|=1} \|\mu_{sy}(f)\|_{L_p(B_{\lambda_0})}\, dy$$

with $C := (\sigma_d)^{-1} C(k\mu, d) \operatorname{card} H(k)$.

Further, change variable in the inner integral by setting $y = z/(t|s|)$ with t satisfying

$$1/\lambda_0 \le t \le 1 \tag{4.20}$$

to obtain for this t and for all $s \in \operatorname{supp}\nu$ the inequality

$$E_{k,p}(f; B_1) \le C \int_{\mathbb{R}} \frac{d\nu(s)}{(t|s|)^{d-1}} \int_{\|z\|=t|s|} \|\mu_{t^{-1}z}(f)\|_{L_p(B_{\lambda_0})}\, dz. \tag{4.21}$$

Lemma 4.6. *For $f \in L_p(B_{\lambda_0})$ and for t and s as above*

$$E_{k,p}(f; B_{1/\lambda_0}) \le C \int_{\mathbb{R}} \frac{d\nu(s)}{(t|s|)^{d-1}} \int_{\|z\|=t|s|} \|\mu_z(f)\|_{L_p(B_{\lambda_0})}\, dz. \tag{4.22}$$

Proof. One changes variables in (4.22) as follows. Due to the definition (2.5) of μ_h,

$$\mu_{t^{-1}z}(f;x) = \int_{\mathbb{R}} f(t^{-1}[tx+sz])\,d\mu(s) = \mu_z(f_t;tx)$$

where $f_t(x) := f(t^{-1}x)$, $x \in B_{t\lambda_0}$. Now change variables to get

$$\|\mu_{t^{-1}z}(f)\|_{L_p(B_{\lambda_0})} = \|\mu_z(f_t;tx)\|_{L_p(B_{\lambda_0})} = t^{-d/p}\|\mu_z(f_t)\|_{L_p(B_{t\lambda_0})}.$$

Further, by scaling one has $E_{k,p}(f;B_1) = t^{-d/p}E_{k,p}(f_t;B_t)$.

Insert these in (4.21) to obtain

$$E_{k,p}(f_t;B_t) \le C\int_{\mathbb{R}} \frac{d\nu(s)}{(t|s|)^{d-1}} \int_{\|z\|=t|s|} \|\mu_z(f_t)\|_{L_p(B_{t\lambda_0})}\,dz. \tag{4.23}$$

Set $g := (f|_{B_{t\lambda_0}})_{t^{-1}}$. Then $g \in L_p(B_{\lambda_0})$ and $g_t = f$ on $B_{t\lambda_0}$. Moreover, $B_{t\lambda_0} \subset B_{\lambda_0}$ and $B_{1/\lambda_0} \subset B_t$ as $1/\lambda_0 \le t \le 1$. Finally, replace f in (4.23) by g to obtain the result. ∎

One can now complete the proof of Theorem 3.2. Rescale the inequality (4.23) to get

$$E_{k,p}(f;B_1) \le C\int_{\mathbb{R}} \frac{d\nu(s)}{(t|s|)^{d-1}} \int_{\|z\|=t|s|} \|\mu_z(f)\|_{L_p(B_{\lambda_0^2})}\,dz. \tag{4.24}$$

Multiply (4.24) by t^{2d-2} and integrate in t over $[1/\lambda_0, 1]$ to obtain

$$\frac{1}{2d-1}\left(1-\frac{1}{\lambda_0^d}\right)E_{k,p}(f;B_r) \le C\int_{\mathbb{R}} \frac{d\nu(s)}{|s|^{d-1}} \int_{1/\lambda_0}^{1} t^{d-1}\,dt \int_{\|z\|=t|s|} \|\mu_z(f)\|_{L_p(B_{\lambda_0^2})}\,dz.$$

Replace t by $u = t|s|$ and set

$$m := \min\{|s|;\, s \in \operatorname{supp}\nu\}, \quad M := \max\{|s|;\, s \in \operatorname{supp}\nu\}.$$

Then $m \le |s| \le M$ and therefore

$$\begin{aligned}
E_{k,p}(f;B_1) &\le C\frac{\operatorname{var}|\nu|}{m^d} \int_{m/\lambda_0}^{M} u^{d-1}\,du \int_{\|z\|=u} \|\mu_z(f)\|_{L_p(B_{\lambda_0^2})}\,dz \\
&\le C\int_0^{\lambda_0^2} u^{d-1}du \int_{\|z\|=u} \|\mu_z(f)\|_{L_p(B_{\lambda_0^2})}\,dz \\
&= C\int_{B_{\lambda_0^2}} \|\mu_z(f)\|_{L_p(B_{\lambda_0^2})}\,dz < \frac{C}{|B_\lambda|}\int_{B_\lambda} \|\mu_z(f)\|_{L_p(B_{\tilde\lambda})}\,dz.
\end{aligned}$$

This gives the required inequality (4.1) with $\tilde\lambda := \lambda_0^2$, $\lambda := \tilde\lambda + |I(\mu)|$ and $C = C(\mu,k,d)$. Theorem 3.2 is proved. ∎

4.2. Proof of Theorem 3.3. First, one will estimate the left-hand site of (3.4) as

$$\Big\{ \int\limits_{\mathbb{R}^d} \mathcal{E}_{k,p}(f; B_r(x))^p \, dx \Big\}^{1/p} \le C \Big\{ \int\limits_{\|h\| \le r} \|\mu_h(f)\|^p_{L_p(\mathbb{R}^d)} \, dh \Big\}^{1/p}. \tag{4.25}$$

This will be derived from inequality (3.3) that asserts that

$$\mathcal{E}_{k,p}(f; B_r(x)) \le C \left\{ \frac{1}{|B_{\hat r}|} \int\limits_{\|h\| \le \hat r} \|\mu_h(f)\|^p_{L_p((B_{\lambda r}(x))_h)} \, dh \right\}^{1/p}; \tag{4.26}$$

here $\hat r$ is the largest r such that $(B_{\lambda r}(x))_h \ne \emptyset$ if $\|h\| \le r$. Due to the definition (2.12) of G_h, $\hat r = \lambda r / |I(\mu)|$, i.e., $\hat r \approx \lambda r$ with constants depending only on inessential parameters.

Hence, $\hat r$ in (4.26) can be replaced by r. Moreover, $(B_{\lambda r}(x))_h$ can be replaced by $B_{\lambda r}$ for $f \in L_p(\mathbb{R}^d)$. Therefore, the left-hand side in (4.25) is bounded by

$$C \left\{ \frac{1}{|B_{\lambda r}|} \int\limits_{\mathbb{R}^d} dx \int\limits_{\|h\| \le \lambda r} \|\mu_h(f)\|^p_{L_p(B_{\lambda r}(x))} \, dh \right\}^{1/p}.$$

Now write the double integral as

$$\begin{aligned} \int\limits_{\mathbb{R}^d} dx \int\limits_{\|h\| \le \lambda r} dh \int\limits_{B_{\lambda r}(x)} |\mu_h(f; y)|^p \, dy &= \int\limits_{\|h\| \le \lambda r} dh \int\limits_{B_{\lambda r}} dy \int\limits_{\mathbb{R}^d} |\mu_h(f; y + x)|^p \, dx \\ &= |B_{\lambda r}| \int\limits_{\|h\| \le \lambda} dh \int\limits_{\mathbb{R}^d} |\mu_h(f; z)|^p \, dz. \end{aligned}$$

Hence, the left-hand side of (4.25) is bounded by a constant times

$$M(\lambda r) := \Big\{ \int\limits_{B_{\lambda r}} \|\mu_h(f)\|^p_{L_p(\mathbb{R}^d)} \, dh \Big\}^{1/p}.$$

To complete this part of the proof it remains to show that

$$M(\lambda r) \le C M(r) \tag{4.27}$$

with $C = C(\lambda, \mu, k, d)$. This will done later; for now, note that the reverse inequality can be proved in the same vein.

In fact, μ annihilates $\mathcal{P}_{k-1}(\mathbb{R}^d)$ and therefore

$$\left\{ \frac{1}{|B_{\hat r}|} \int\limits_{B_{\hat r}} \|\mu_h(f)\|^p_{L_p(B_r(x)_h)} \, dh \right\}^{1/p} \le \operatorname{var}|\mu| \mathcal{E}_{k,p}(f; B_r(x)).$$

Integrate the p-power of this inequality in x to obtain as above

$$M(r) \le C \Big\{ \int\limits_{\mathbb{R}^d} \mathcal{E}_{k,p}(f; B_r(x))^p \, dx \Big\}^{1/p} =: C\mathcal{E}(r). \tag{4.28}$$

Together with the previously proven result this gives

$$C_1 M(r) \le \mathcal{E}(r) \le C_2 M(\lambda r). \tag{4.29}$$

This shows that to prove (4.27) it suffices to establish an analogous fact for $\mathcal{E}(r)$ with some $\lambda > 1$, say, $\lambda = 3/2$. In the following, it will be convenient to replace in (4.29) the ball $B_r(x)$ by the cube $Q_r(x) := \{y \in \mathbb{R}^d;\ \max_{1\le i\le d} |y_i - x_i| \le r\}$. Clearly, this modification of $\mathcal{E}(r)$ denoted by $\tilde{\mathcal{E}}(r)$ satisfies

$$(\sqrt{d})^{-d/p}\mathcal{E}(r) \le \tilde{\mathcal{E}}(r) \le (\sqrt{d})^{d/p}\mathcal{E}(\sqrt{d}\, r).$$

Hence, it suffices to prove that

$$\tilde{\mathcal{E}}((3/2)r) \le C(k)\tilde{\mathcal{E}}(r). \tag{4.30}$$

For this, we need the following.

A pair S_1, S_2 of measurable subsets in $\mathbb{R}^d$ is *ε-linked*, $0 < \varepsilon < 1$, if

$$|S_1 \cap S_2| \ge \varepsilon |S_1 \cup S_2|.$$

For such S_j and some constant $C = C(k, d, \varepsilon) > 0$ we have (see [4, Theorem 2])

$$E_{k,p}(f; S_1 \cup S_2) \le C \sum_{j=1,2} E_{k,p}(f; S_j). \tag{4.31}$$

Now cover $Q_{(3/2)r}(x)$ by its subcubes $Q^j = Q_r(x + x^j)$, $1 \le j \le 2^d$, such that Q^j has one common vertex with $Q_{(3/2)r}(x)$. Then for every j,

$$\Big|Q^j \cap \Big(\bigcup_{j'\ne j} Q^{j'}\Big)\Big| \ge \frac{1}{2^d}\Big|\sum_{j=1}^{2^d} Q^j\Big|.$$

Hence, applying (4.31) 2^d times and then passing to the normalized local approximation $\mathcal{E}_{k,p}$ one gets

$$\mathcal{E}_{k,p}(f; Q_{(3/2)r}(x)) \le C \sum_{j=1}^{2^d} \mathcal{E}_{k,p}(f; Q_r(x + x^j)).$$

Finally, take the $L_p(\mathbb{R}^d)$-norm of both sides to get (4.30).

Theorem 3.3 is proved. ■

4.3. Proof of Theorem 3.5. The result will be derived from Theorem 3.1 and the following fact [7, §2, Theorem 5].

THEOREM A. *A function $f \in L_p(\mathbb{R}^d)$ belongs to the Taylor class $t_p^k(x)$ for almost all points x of a set S of positive measure if and only if for almost all $x \in S$,*

$$\limsup_{r\to 0} \frac{\mathcal{E}_{k,p}(f; B_r(x))}{r^k} < \infty. \tag{4.32}$$

Since the proof in [7, pp. 183–184] is distorted in translation, it will be repeated at the end of this subsection.

Now let

$$\mu_p(f; B_r(x)) = O(r^k) \tag{4.33}$$

for almost all $x \in S$, as assumed in Theorem 3.5. Due to Theorem 3.1 this implies (4.32) for every density point $x \in S$. Since the set of density points of S is of measure $|S|$ by the Lebesgue theorem, (4.27) holds almost everywhere on S. Hence, $f \in t_p^k(x)$ for almost all $x \in S$.

Conversely, if $f \in t_p^k(x)$ for $x \in S$, then

$$\mathcal{E}_{k,p}(f; B_r(x)) \le C(x) r^k$$

for small $r > 0$. As $\mu_p(f; B_r(x)) \le (\operatorname{var}\mu)\mathcal{E}_{k,p(f;B_r(x))}$, condition (4.33) holds for such x.

The proof of Theorem 3.5 is complete. ∎

Proof of Theorem A ([2]). We put

$$\varphi(x) := \sup_{n\in\mathbb{N}}\{2^{nk}\mathcal{E}_{k,p}(f; B_{2^{-n}}(x))\}, \qquad x \in S.$$

Being the supremum of a sequence of measurable functions, φ is measurable, and moreover it is finite almost everywhere on S by (4.32). Therefore, given $\varepsilon > 0$ there is a subset $U_\varepsilon \subset S$ such that $|S \setminus U_\varepsilon| < \varepsilon$ and $\gamma := \sup_{x\in U_\varepsilon}\varphi(x) < \infty$. Consequently, for $r = 2^{-n}$, $n = 0, 1, 2, \ldots,$

$$\sup_{x\in U_\varepsilon} \mathcal{E}_{k,p}(f; B_r(x)) \le \gamma r^k. \tag{4.34}$$

Increasing γ one can assume that (4.34) holds for $0 < r \le 1$.

Using (4.34) one verifies that $f \in t_p^k(x)$ for almost all $x \in V_\varepsilon \subset U_\varepsilon$ where V_ε is such that $|U_\varepsilon \setminus V_\varepsilon| < \varepsilon$. Since ε is arbitrary, it then follows that $f \in t_p^k(x)$ for almost all $x \in S$.

By the extension theorem of §4 in [5] (see (55) there) and Theorem 7 of the same section we deduce from (4.34) that on some subset $V_\varepsilon \subset U_\varepsilon$ such that $|U_\varepsilon \setminus V_\varepsilon| < \varepsilon$ the function f coincides with the trace of a function $F \in C^k(\mathbb{R}^d)$.

Now we interrupt the derivation to explain the results just referred to.

Comments. The cited extension theorem asserts (see [5] and [9] for a more general result):

If (4.34) *holds on a d-regular subset $V \subset \mathbb{R}^d$, then f can be extended to a function from $C^{k-1,1}(\mathbb{R}^d)$ or what is the same, from the Sobolev space $W_\infty^k(\mathbb{R}^d)$.*

([2]) For the convenience of the reader, the results referred to within [7] are explained in more detail.

A measurable subset $V \subset \mathbb{R}^d$ is said to be (Ahlfors) *d-regular* if for some constant $\gamma > 0$ and every ball $B_r(x)$ with $x \in V$ and $0 < r \le 1$,

$$|V \cap B_r(x)| \ge \gamma |B_r(x)|. \tag{4.35}$$

The V_ε from the derivation of Theorem A is in fact a d-regular subset of U_ε. Its existence follows from the Lebesgue density point theorem asserting, in particular, that for almost all $x \in U_\varepsilon$, (4.35) holds with some $\gamma = \gamma(x) > 0$.

Theorem 7 of [7, §7] asserts that if $f \in W^k_\infty(B_1)$, then for every $\varepsilon > 0$ there exists a function $f_\varepsilon \in C^k(B_1)$ such that

$$|\{x \in B_1;\ f(x) \ne f_\varepsilon(x)\}| < \varepsilon.$$

As without loss of generality the set S can be assumed to be a subset of B_1, these two theorems imply the stated result.

Now we put $g := f - F$ and verify that this function belongs to $t^k_p(x)$ for almost all $x \in V_\varepsilon$. In fact, by Theorem 1 of [7] (see Comments below) the Taylor polynomial $m_x \in \mathcal{P}_{k-1}(\mathbb{R}^d)$ of f at $x \in U_\varepsilon$ exists and satisfies

$$\|f - m_x\|_{L_p(B_r(x))} = O(r^{k+d/p}) \quad \text{if } 0 < r \le 1.$$

But for $x \in V_\varepsilon$ this polynomial is also the Taylor polynomial at x for the extension F. Hence,

$$\|g\|_{L_p(B_r(x))} = O(r^{k+d/p}) \quad \text{for all } x \in V_\varepsilon \quad \text{and} \quad 0 < r \le 1. \tag{4.36}$$

Due to the Calderón–Zygmund result [10, Theorem 10], (4.36) implies that for almost every $x \in V_\varepsilon$,

$$\|g\|_{L_p(B_r(x))} = o(r^{k+d/p}) \quad \text{as } r \to 0.$$

Then for the Taylor polynomial of F at x of degree k, denoted by $\tilde{m}_x$, one gets

$$\|f - \tilde{m}_x\|_{L_p(B_r(x))} \le \|g\|_{L_p(B_r(x))} + \|F - \tilde{m}_x\|_{L_p(B_r(x))} = o(r^{k+d/p}) \text{ as } r \to 0.$$

This means that f belongs to $t^k_p(x)$ for almost all $x \in V_\varepsilon$, as required in Theorem 3.5.

Comments. The cited Theorem 1 asserts, in particular, that if (4.34) holds, then f belongs to $T^k_p(x)$ for almost all $x \in U_\varepsilon$. This implies the existence of the Taylor polynomial $m_x \in \mathcal{P}_{k-1}(\mathbb{R}^d)$ for f at almost all $x \in U_\varepsilon$.

As the converse to the result just obtained is evident, Theorem A is proved. ∎

4.4. Proof of Theorem 3.6. The result will follow from [7, §1, Theorems 3 and 4], giving the following description of Taylor classes via local polynomial approximation.

Theorem B. (a) *Let $0 < \lambda < k$ be noninteger or $\lambda = k$. Then $f \in T_p^k(x)$ if and only if*

$$\mathcal{E}_{k,p}(f; Q_r(x)) = O(r^\lambda) \quad \text{as } r \to 0. \tag{4.37}$$

(b) *Let $0 < \lambda < k$ be noninteger. Then $f \in t_p^\lambda(x)$ if and only if*

$$\mathcal{E}_{k,p}(f; Q_r(x)) = o(r^\lambda) \quad \text{as } r \to 0. \tag{4.38}$$

(c) *If $\lambda = k$, then $f \in t_p^k(x)$ if and only if there exists a family of polynomials*

$$M_{r,x}(y) = \sum_{|\alpha| \le k} C_\alpha(r,x)(y-x)^\alpha, \qquad y \in \mathbb{R}^d,\ 0 < r \le 1, \tag{4.39}$$

such that

$$\left\{ \frac{1}{|B_r|} \int_{B_r(x)} |f(y) - M_{r,x}(y)|^p \, dy \right\}^{1/p} = o(r^k) \quad \text{as } r \to 0 \tag{4.40}$$

and moreover the limits $\lim_{r\to 0} D^\alpha M_{r,x}$ $(= \lim_{r\to 0} C_\alpha(r,x))$ exist for all $|\alpha| = k$.

To derive assertion (a) of Theorem 3.6 we use first Theorem 3.1 for $f \in L_p^{\mathrm{loc}}(\mathbb{R}^d)$ satisfying the assumption

$$\limsup_{r\to 0}\ r^{-\lambda} \mu_p(f; B_r(x)) < \infty \quad \text{a.e. on } S. \tag{4.41}$$

This gives, for those f,

$$\limsup_{r\to 0}\ r^{-\lambda} \mathcal{E}_{k,p}(f; B_r(x)) < \infty \quad \text{a.e. on } S.$$

Then assertion (a) of Theorem B implies that $f \in T_p^\lambda(x)$ a.e. on S. Assertion (b) of Theorem 3.6 can be proved in the same vein using statement (b) of Theorem B and the analog of (4.41) with $o(1)$ as $r \to 0$ on the right-hand side.

Let now $f \in L_p^{\mathrm{loc}}(\mathbb{R}^d)$ satisfy the assumption of Theorem 3.6(c), i.e., for almost all $x \in S$ there exists a family $\{m_x\}_{x\in S}$ of homogeneous polynomials of degree k such that

$$\left\{ \frac{1}{|B_r|} \int_{B_r(x)} |\mu_h(f; y) - m_x(y)|^p \, dy \right\}^{1/p} = o(r^k) \quad \text{as } r \to 0. \tag{4.42}$$

Set $P_\alpha(x) := x^\alpha$, $x \in \mathbb{R}^d$. As μ is orthogonal to $\mathcal{P}_{k-1}(\mathbb{R}^d)$ and $c(\mu) := \int_{\mathbb{R}} t^k \, d\mu(t) \ne 0$, one gets

$$\mu_h(P_\alpha; x) = \int_{\mathbb{R}} (x + th)^\alpha \, d\mu(t) = c(\mu) h^\alpha.$$

Hence, for

$$m_x(y) := \frac{1}{c(\mu)} \sum_{|\alpha|=k} C_\alpha(x) y^\alpha, \qquad y \in \mathbb{R}^d,$$

we have

$$\mu_h(m_x) = \sum_{|\alpha|=k} C_\alpha(x) h^\alpha = m_x(h).$$

This and (4.42) imply that for almost all $x \in S$,

$$\left\{ \frac{1}{|B_r|} \int\limits_{B_r(x)} |\mu_h(f;x) - m_x(h))|^p \, dh \right\} = o(r^k) \qquad \text{as } r \to 0.$$

By Theorem 3.2 this, in turn, gives

$$\mathcal{E}_{k,p}(f - m_x; B_r(x)) = o(r^k) \qquad \text{as } r \to 0 \tag{4.43}$$

for almost all $x \in S$.

Let $P_{r,x}$ be a polynomial of degree $k-1$ such that

$$\left\{ \frac{1}{|B_r|} \int\limits_{B_r(x)} |(f(y) - m_x(y)) - P_{r,x}(y)|^p \, dy \right\}^{1/p} = \mathcal{E}_{k,p}(f - m_x; B_1(x)).$$

Set $M_{r,x} := P_{r,x} + m_x$; then by (4.42) and (4.43) the family $\{M_{r,x}\}_{x \in S}$ satisfies

$$\left\{ \frac{1}{|B_r|} \int\limits_{B_r(x)} |f(y) - M_{r,x}(y)|^p \, dy \right\}^{1/p} = o(r^k) \qquad \text{as } r \to 0,$$

and moreover

$$\lim_{r \to 0} D^\alpha M_{r,x} = D^\alpha m_x = C_\alpha(x) \frac{\alpha!}{c(\mu)}, \qquad |\alpha| = k,$$

for almost all $x \in S$.

Hence, the assertion of Theorem B(c) holds for f at almost all points of S and therefore $f \in t_p^k(x)$ at those points. The proof of Theorem 3.6 is complete. ■

4.5. Proof of Corollary 3.1. If f has the symmetric (k,p)-differential at x, then by definition there exists a (Taylor) polynomial, say, $T_x(y) := \sum_{|\alpha| \le k} C_\alpha(x)(y-x)^\alpha$, $y \in \mathbb{R}^d$, such that

$$\left\{ \frac{1}{|B_r|} \int\limits_{B_r(x)} |\hat{f}_x(y) - T_x(y)|^p \, dy \right\}^{1/p} = o(r^k) \qquad \text{as } r \to 0; \tag{4.44}$$

here $\hat{f}_x(y) := \frac{1}{2}(f(x+y) - (-1)^k f(x-y))$ for $y \in \mathbb{R}^d$.

Let $\mu := \frac{1}{2}(\Delta_1^k - (-1)^k \Delta_{-1}^k)$. i.e., $\mu_h(f;x) = \frac{1}{2}(\Delta_h^k f(x) - (-1)^k \Delta_{-h}^k f(x))$. Since $\mu_h(f;x) = \Delta_h^k(\hat{f}_x; y)|_{y=0}$ and

$$m_x(h) := \Delta_h^k(T_x(y))|_{y=0} = k! \sum_{|\alpha|=k} C_\alpha(x) h^\alpha,$$

one gets

$$\left\{ \frac{1}{|B_r|} \int_{B_r(x)} |\mu_h(f;x) - m_x(h)|^p \, dh \right\}^{1/p} = \left\{ \frac{1}{|B_r|} \int_{B_r(x)} |\Delta_h^k((\hat{f}_x; y) - T_x(y))|_{y=0}|^p \, dh \right\}^{1/p}.$$

Due to (4.44) and the assumption of the corollary, the right-hand side is $o(r^k)$ as $r \to 0$ for almost all $s \in S$.

Hence, f satisfies the condition of Theorem 3.6(a) and therefore $f \in t_p^k(x)$ for almost every $x \in S$. ■

4.6. Proof of Theorem 3.4. One should compare the behavior at points of S with $0 < |S| \le \infty$ of two μ-characteristics of $f \in L_p^{\mathrm{loc}}(\mathbb{R}^d)$, $1 \le p \le \infty$, namely

$$\tilde{\mu}_p(f; B_r(x)) := \left\{ \frac{1}{|B_r|} \int_{B_r} \|\mu_h(f, \cdot)\|_{L_p(B_r(x)_h)}^p \, dh \right\}^{1/p} \tag{4.45a}$$

and

$$\mu_p(f; B_r(x)) := \left\{ \int_{B_r} |\mu_h(f;x)|^p \, dh \right\}^{1/p}; \tag{4.45b}$$

see (2.15) and (2.14), respectively. Recall that μ now belongs to the class $\mathcal{M}$ of discrete measures on $\mathbb{R}$ satisfying only the conditions

$$0 \in \operatorname{supp} \mu \quad \text{and} \quad 1 < \operatorname{card}(\mu) < \infty. \tag{4.46}$$

Since Theorem 3.4 is invariant with respect to dilation $h \mapsto \lambda h$, $\lambda > 0$, assume without loss of generality that $|I(\mu)| > 1$.

We will also use an intermediate difference characteristic of f given by

$$\mu_p(f; S; x; r, r') := \left\{ \frac{1}{|B_r|\,|B_{r'}|} \int_{S_r(x)} dy \int_{B_{r'}} |\mu_h(f;y)|^p \, dh \right\}^{1/p} \tag{4.47}$$

where $S_r(x) := S \cap B_r(x)$. We will omit S in (4.47) if $S = \mathbb{R}^d$, and x if $x = 0$.

LEMMA 4.7. *There is a positive constant $C = C(d,p)$ such that*

$$\mu_p(f; S; x; r, r') \le C \sup_{y \in S_r(x)} \mu_p(f; S; y; r', r') \quad \text{for } 0 < r' \le r. \tag{4.48}$$

Proof. It suffices to consider the case of $x = 0$; therefore x will be omitted from the corresponding notations. Let $\{x^i\}$ be a maximal r'-separated set in $B_{r'}$. Since the distances between the points x^i are r' or more, the open balls $B_{r'/2}(x^i)$ are pairwise disjoint. Moreover, every such ball is contained in $B_{r+r'/2} \subset B_{(3/2)r}$ and therefore

$$N := \operatorname{card}\{B_{r'/2}(x^i)\}_i \le \frac{|B_{(3/2)r}|}{|B_{r'}|} = \left(\frac{3}{2}\right)^d \frac{|B_r|}{|B_{r'}|}.$$

Finally, due to maximality of the r'-separated set, the doubled balls $B_{r'}(x^i)$ cover B_r. Therefore

$$\begin{aligned}
&[\mu_p(f;S;r,r')]^p \\
&= \frac{1}{|B_r|\,|B_{r'}|} \int_{S_r} dy \int_{B_{r'}} |\mu_h(f;y)|^p\,dh \le \frac{1}{|B_r|\,|B_{r'}|} \sum_{i=1}^{N} \int_{S_{r'}(x^i)} dy \int_{B_{r'}} |\mu_h(f;y)|^p\,dh \\
&\le \frac{|B_{r'}|}{|B_r|} \sum_{i=1}^{N} [\mu_p(f;S;x^i;r',r')]^p \le \left(\frac{3}{2}\right)^d \sup_{y\in S_r} [\mu_p(f;S;y;r',r')]^p.
\end{aligned}$$

This proves the inequality (4.48). ∎

LEMMA 4.8. *Let $\mu,\nu \in \mathcal{M}$ and let x be a density point of S. There are positive constants $r_0 = r_0(x)$ and $C = C(\mu,\nu)$, $C' = C'(p,d,\mu,\nu)$ such that*

$$\nu_p(f,x;r,r') \le C'\{\nu_p(f;S;x;Cr,r') + \mu_p(f;S;x;Cr,r)\} \tag{4.49}$$

for $0 < r' \le r \le r_0$.

Note that the left-hand side, in contrast to the right-hand one, is independent of S.

Proof. Once again assume that $x = 0$ and omit x from notation. Let $\operatorname{supp}\mu := \{t_1,\ldots,t_m\}$, and $\operatorname{supp}\nu := \{s_1,\ldots,s_m\}$, $m,n > 1$. For the composition of the difference operators ν_h and μ_g, $h,g \in \mathbb{R}^d$, one then gets the evident identity

$$\nu_h(f;y) = \frac{1}{\mu(\{t_m\})}(\mu_g\nu_h - \nu_h\tilde{\mu}_g)(f;y - t_m g) \tag{4.50}$$

with $\tilde{\mu} := \mu - \mu(\{t_m\})\delta_{t_m}$. This will be used later in the proof.

Define the following subsets of B_r:

$$\begin{aligned}
F_i &:= F_i(y,h,r) = \big(t_m^{-1}(y + s_i h) - t_m^{-1}S\big) \cap B_r, && 1 \le i \le n,\\
H_i &:= H_i(y,h,r) = \big((t_m - t_i)^{-1}y - (t_m - t_i)^{-1}S\big) \cap B_r, && 1 \le i \le m-1.
\end{aligned}$$

We claim that the measure of each of F_i, H_i is equivalent as $r \to 0$ to the measure of B_r uniformly in $y,h \in B_r$.

In fact, since $x = 0$ is a density point for S, for $S^c := \mathbb{R}^d \setminus S$ and some function $\varphi(r) \to 0$ as $r \to 0$ one has

$$|S^c \cap B_r| = \varphi(r)|B_r|.$$

Each F_i and each H_i is of the form $D = (\alpha y + \beta h + kS) \cap B_r$. As $D^c \cap B_r = (\alpha y + \beta h + kS^c) \cap B_r$, for $y, h \in B_r$ one has $D^c \cap B_r - (\alpha y + \beta h) \subset (kS^c) \cap B_{lr}$ where $l := |\alpha| + |\beta| + 1$.

Therefore, $|D^c \cap B_r| \le |k|^d \varphi(lr)|B_{lr}| = (|k|l)^d \varphi(lr)|B_r|$. Since the right-hand side tends to 0 as $r \to 0$ uniformly in y, h, the claim follows.

Consider further the set

$$J = J(y,h,r) := \Big(\bigcap_{i=0}^{n} F_i\Big) \cap \Big(\bigcap_{i=0}^{m-1} H_i\Big).$$

It follows from the claim above that for some $r_0 > 0$ and all $0 < r \le r_0$,

$$\tfrac{1}{2}|B_r| \le |J| \le |B_r|.$$

Due to the definition of J, for some $C = C(\nu, \mu) > 0$, and every $y \in B_r$ and $z \in J$,

$$y - t_m z + s_i h \in S_{Cr}, \quad i \le n, \quad \text{and} \quad y - (t_m - t_i)z \in S_{Cr}, \ \ i \le m-1. \tag{4.51}$$

Now we return to the proof of the lemma, first for $1 \le p < \infty$.

Applying the Hölder inequality to (4.50) one gets

$$\begin{aligned} |\nu_h(f;y)|^p \le (C_1)^p \Big\{ &\sum_{i=1}^{n} |\mu_t(f; y - t_m z + s_i h)|^p \\ &+ \sum_{i=1}^{m-1} |\nu_h(f; (y - t_m - t_i)z)|^p \Big\} \end{aligned} \tag{4.52}$$

with some $C_1 = C_1(\mu, \nu)$. Integration in z over J gives

$$\begin{aligned} \frac{1}{2}|B_r|\,|\nu_h(f;y)|^p \le (C_1)^p \Big(&\sum_{s \in \operatorname{supp}\nu} \int_J |\mu_z(f; y - t_m z + sh)|^p\, dz \\ &+ \sum_{t \in \operatorname{supp}\tilde{\mu}} \int_J |\nu_h(f; y - (t_m - t)z)|^p\, dz \Big). \end{aligned}$$

The change of variables $u = y - t_m z + s_i h$ (with Jacobian 1) gives, for $0 < r' \le r \le r_0$,

$$\begin{aligned} &\int_{B_{r'}} dh \Big(\int_{B_r} dy \int_J |\mu_z((f; y - t_m z + sh)|^p\, dz \Big) \\ &\quad \le \int_{B_{r'}} dh \Big(\int_{S_{Cr}} du \int_{B_r} |\mu_z(f;u)|^p\, dz \Big) = |B_{r'}| \{ |B_{Cr}|\,|B_r| [\mu_p(f; S; Cr, r)]^p \}. \end{aligned}$$

Similarly,

$$\int\limits_{B_{r'}} dh \Big(\int\limits_{B_r} dy \int\limits_J |\nu_h((f; y-(t-sh)z)|^p\, dz\Big)$$

$$\le \int\limits_{B_{r'}} dh \Big(\int\limits_{S_{Cr}} du \int\limits_{B_r} |\nu_h(f;u)|^p\, dz\Big) = |B_r|\{|B_{r'}|\,|B_{Cr}|[\nu_p(f;S;Cr,r')]^p\}.$$

Therefore, for $0 < r' \le r \le r_0$ and $p < \infty$ one has the required inequality

$$\nu_p(f;r,r') \le C'\{\nu_p(f;S;Cr,r') + \mu_p(f;S;Cr,r)\}$$

with $C' = [2(m+n-1)]^{1/p} C_1 C^{d/p}$.

Now let $p = \infty$. If $y \in B_r$, $h \in B_{r'}$ and $0 < r' \le r \le r_0$, then for $z \in F_i$ the point $y - t_m z + s_i h$ is in S_{Cr} (see (4.51)); therefore,

$$|\mu_z(f; y - t_m z + s_i h)| \le \mu_\infty(f;S;Cr,r).$$

Similarly, if $t \in H_i$ then $y - (t_m - t_i) \in S_{Cr}$ and so

$$|\nu_h(f; y - (t_m - s_i)z)| \le \nu_\infty(f;S;Cr,r').$$

Inequality (4.52) for $p = 1$, together with the two just proved, gives, for $z \in J$,

$$\nu_\infty(f;r,r') \le C'\{\nu_\infty(f;S;Cr,r') + \mu_\infty(f;S;Cr,r)\}.$$

The proof of Lemma 4.8 is complete. ∎

Now we return to the proof of Theorem 3.4. Since ω is a majorant, one gets, for some $C = C(\omega) > 1$,

(4.53)
$$\limsup_{\substack{k\in\mathbb{N}\\ k\to\infty}} \frac{\mu_p(f;B_{2^{-k}}(x))}{\omega(2^{-k})} \le \limsup_{r\to\infty} \frac{\mu_p(f;B_r(x))}{\omega(r)} \le C \limsup_{\substack{k\in\mathbb{N}\\ k\to\infty}} \frac{\mu_p(f;B_{2^{-k}}(x))}{\omega(2^{-k})}.$$

Hence, in the proof one can replace the upper limit for $r \to 0$ by that for $k \to \infty$ ($k \in \mathbb{N}$).

Now set

$$S^0 := \Big\{x \in S;\ \lim_{k\to\infty} \frac{\mu_p(f;B_{2^{-k}}(x))}{\omega(2^{-k})} = 0\Big\},$$

$$S^1 := \Big\{x \in S;\ 0 < \limsup_{k\to\infty} \frac{\mu_p(f;B_{2^{-k}}(x))}{\omega(2^{-k})} < \infty\Big\}.$$

By the assumption of the theorem, $S = S^0 \cup S^1$.

These two subsets can be represented as follows. First given $m, n \in \mathbb{N}$, define

$$S^0_{m,n} := \left\{ x \in S;\ \sup_{k \ge n} \frac{\mu_p(f; B_{2^{-k}}(x))}{\omega(2^{-k})} < \frac{1}{m} \right\},$$

$$S^1_{m,n} := \left\{ x \in S;\ 2^m \le \limsup_{k \to \infty} \frac{\mu_p(f; B_{2^{-k}}(x))}{\omega(2^{-k})} \text{ and } \sup_{k \ge n} \frac{\mu_p(f; B_{2^{-k}}(x))}{\omega(2^{-k})} < 2^{m+1} \right\}.$$

Then it is a matter of definition that for some functions $\mathbb{N} \ni m \mapsto n(m) \in \mathbb{N}$,

$$S^0 = \bigcap_{m \in \mathbb{N}} \bigcup_{n \ge n(m)} S^0_{m,n} \quad \text{and} \quad S^1 = \bigcup_{m \in \mathbb{Z}} \bigcap_{n \ge n(m)} S^1_{m,n}. \tag{4.54}$$

Now we show that for density points x of S^0,

$$\lim_{r \to 0} \frac{\tilde{\mu}_p(f; B_r(x))}{\omega(r)} = \lim_{r \to 0} \frac{\mu_p(f; B_r(x))}{\omega(r)} = 0, \tag{4.55}$$

i.e., the required inequality (3.6) is true for those x. To this end take $y \in S^0_{m,n}$ and $k \ge n$. By definition, $\mu_p(f; B_{2^{-k}}(y)) < \frac{1}{m}\omega(2^{-k})$. Raising this to the power p, integrating in y over $B_r(x) \cap S^0_{m,n}$ with $x \in S^0_{m,n}$ and using (4.47) one obtains

$$\mu_p(f; S^0_{m,n}; x; 2^{-k}, 2^{-k}) \le \frac{1}{m}\omega(2^{-k})|B_{2^{-k}}|^{-1/p}.$$

Together with Lemma 4.7 this implies that for $0 < r' \le r \le 2^{-n}$ with sufficiently large n,

$$\mu_p(f; S^0_{m,n}; x; r, r') \le C_1 \frac{1}{m}\omega(r')|B_{r'}|^{-1/p}$$

where C_1 is independent of r, r' and f.

Now let x be a density point of $S^0_{m,n}$. Lemma 4.8 applied to the case $\nu = \mu$ gives, for such x, r, r',

$$\begin{aligned}\mu_p(f, x; r, r') &\le C_2\{\mu_p(f; S^0_{m,n}; x; Cr, r') + \mu_p(f : S^0_{m,n}; x; Cr, r)\}\\ &\le 2C_1C_2\omega(r)\frac{1}{m}|B_{r'}|^{-1/p}.\end{aligned}$$

Recall that C_1 and C_2 depend only on μ, d, p. Hence, for almost all $x \in S^0_{m,n}$,

$$\limsup_{r \to 0} \frac{\mu_p(f, x; r, r)|B_r|^{1/p}}{\omega(r)} \le 2C_1C_2\frac{1}{m}. \tag{4.56}$$

Moreover, since $c := |I(\mu)| \ge 1$,

$$\tilde{\mu}_p(f; B_r(x)) \le c^{d/p}|B_r|^{1/p}\mu_p(f; x; r, r), \tag{4.57}$$

In fact, by (4.45b), (4.47) and the equality $(B_r(x))_h = \emptyset$ for $\|h\| > r/c$, one gets

$$\tilde{\mu}_p(f; B_r(x)) = \left\{ \frac{1}{|B_r|} \int\limits_{B_{r/c}} \|\mu_h(f;\cdot)\|^p_{L_p((B_r(x))_h)}\, dh \right\}^{1/p}$$

$$\le \left\{ \frac{1}{|B_r|} \int\limits_{B_{r/c}} dh \int\limits_{B_r(x)} |\mu_h(f;y)|^p\, dy \right\}^{1/p}$$

$$\le c^{d/p} \left\{ \frac{1}{|B_r|} \int\limits_{B_r} dh \int\limits_{B_r(x)} |\mu_h(f;y)|^p\, dy \right\}^{1/p}$$

$$= c^{d/p} |B_r|^{1/p} \mu_p(f, x; r, r).$$

Combining inequalities (4.56), (4.57) and letting $m \to \infty$ one then obtains, for a.e. $x \in S^0$,

$$\lim_{r\to 0} \frac{\tilde{\mu}_p(f; B_r(x))}{\omega(r)} = \lim_{r\to 0} \frac{\mu_p(f; B_r(x))}{\omega(r)} = 0.$$

This proves (4.55).

Further one proves the required inequality (3.6) for almost all points of S^1. Let $x, y \in S^1_{m,n}$ and $k \ge n$. By definition of this subset,

$$\mu_p(f; B_{2^{-k}}(y)) \le 2^{m+1}\omega(2^{-k}) \le 2\omega(2^{-k}) \limsup_{i\to\infty} \frac{\mu_p(f; B_{2^{-i}}(x))}{\omega(2^{-i})}.$$

Integrating the power p of this inequality in y over $B_{2^{-k}}(x) \cap S^1_{m,n}$ we get

$$\mu_p(f; S^1_{m,n}; x; 2^{-k}, 2^{-k}) \le 2^{m+1}\omega(2^{-k})$$

$$\le 2\omega(2^{-k}) |B_{2^{-k}}|^{-1/p} \cdot \limsup_{i\to\infty} \frac{\mu_p(f; B_{2^{-i}}(x))}{\omega(2^{-i})}.$$

Now we use the argument for (4.55) to obtain, for a density point x of $S^1_{m.n}$ and sufficiently small $0 < r' \le r$, the inequality

$$\mu_p(f, x; r, r') \le C\omega(r) |B_{r'}|^{-1/p} \cdot \limsup_{r\to 0} \frac{\mu_p(f; B_r(x))}{\omega(r)}$$

with C independent of f and r. Finally, using (4.57) we get

$$\limsup_{r\to 0} \frac{\tilde{\mu}_p(f, B_r(x))}{\omega(r)} \le C \limsup_{r\to 0} \frac{\mu_p(f; B_r(x))}{\omega(r)}$$

for almost all $x \in S^1_{m,n}$, hence, for almost all $x \in S^1$ as well (see (4.54)).

The proof of Theorem 3.4 is complete. ∎

References

[1] M. Artin et al., *The situation in Soviet mathematics*, Notices Amer. Math. Soc., November 1978, 495–497.

[2] N. Bourbaki, *Integration I*, Springer, Berlin, 2004.
[3] Yu. Brudnyi, *On the best local approximation of functions by polynomials*, Dokl. Akad. Nauk SSSR 161 (1965), 746–749 (in Russian).
[4] Yu. Brudnyi, *A multidimensional analog of a theorem of Whitney*, Mat. Sb. 62 (124) (1970), 175–191 (in Russian); English transl.: Math. USSR-Sb. 11 (1970), 157–170.
[5] Yu. Brudnyi, *On an extension theorem*, Funktsional. Anal. i Prilozhen. 4 (1970), no. 3, 96–97 (in Russian); English transl.: Funct. Anal. Appl. 4 (1970), 252–253.
[6] Yu. Brudnyi, *Local approximation and differential properties of functions of several variables*, Uspekhi Mat. Nauk 29 (1974), no. 4, 163–164 (in Russian).
[7] Yu. Brudnyi, *Adaptive approximation of functions with singularities*, Trudy Moskov. Mat. Obshch. 55 (1994), 149–242 (in Russian); English transl.: Transl. Moscow Math. Soc. 1994, 123–186.
[8] A. Brudnyi and Yu. Brudnyi, *Remez type inequality and Morrey–Campanato spaces on Ahlfors regular sets*, in: Contemp. Math. 445, Amer. Math. Soc., 2007, 19–44.
[9] A. Brudnyi and Yu. Brudnyi, *Methods of Geometric Analysis in Extension and Trace Problems*, Birkhäuser, Basel, 2011.
[10] A. P. Calderón and A. Zygmund, *Local properties of solutions of elliptic partial differential equations*, Studia Math. 20 (1961), 171–225.
[11] I. E. Gopengauz, *Difference properties and differentiability of functions of several variables*, Uspekhi Mat. Nauk 24 (1973), no. 4, 215–216 (in Russian).
[12] I. E. Gopengauz, *Conditions of the differentiability of functions of several variables*, in: Investigations in the Theory of Functions of Several Real Variables, Yaroslavl' State Univ., Yaroslavl', 1978, 64–87 (in Russian).
[13] R. A. DeVore and G. G. Lorentz, *Constructive Approximation.* Springer, Berlin, 1993.
[14] A. Marchaud, *Sur les dérivées et sur les différences des fonctions de variables réelles*, J. Math. Pures Appl. 6 (1927), 337–425.
[15] J. Marcinkiewicz and A. Zygmund, *On the differentiability of functions and summability of trigonometrical series*, Fund. Math. 26 (1936), 1–43.
[16] J. Marcinkiewicz et A. Zygmund, *Sur la dérivée seconde généralisée*, Bull. Sém. Math. Univ. Wilno 2 (1939), 35–40.
[17] D. J. Newman and T. J. Rivlin, *A characterization of weights in a divided difference*, Pacific J. Math. 89 (1981), 407–413.
[18] M. Weiss, *On symmetric derivatives in L^p*, Studia Math. 24 (1964), 89–100.
[19] H. Whitney, *On functions of bounded nth differences*, J. Math. Pures Appl. 36 (1957), 67–95.
[20] A. Zygmund, *Józef Marcinkiewicz*, preface to: Collected Papers of J. Marcinkiewicz, Państwowe Wydawnictwo Naukowe, Warszawa, 1964.

Yu. A. Brudnyi (corresponding author)
Department of Mathematics
Technion
Haifa, 32000, Israel
E-mail: ybrudnyi@math.technion.ac.il

I. E. Gopengauz
Department of Mathematics
National Research Technical University (MISA)
Moscow, 119049, Russia
E-mail: iegopengauz@mail.ru

Received September 22, 2012
Revised version March 6, 2013 (7625)