STUDIA MATHEMATICA, T. LXXIV (1982)

On L^p-differentiability and difference properties of functions*

by

TORD SJÖDIN (Umeå)

Abstract. A characterization of (ordinary) differentiability for functions f by properties of their binary differences $B_h^j f$ due to J. Marcinkiewicz and A. Zygmund is generalized to the case of L_a^p-differentiability. Our main result is then that a function f is differentiable if and only if it is L^p-differentiable and satisfies certain conditions of difference type called C_s^j, for suitable j and s. These conditions are expressed in terms of the differences B_h^j or the standard differences $\Delta_h^j f$ using either L^p-norm or supremum norm. The results have applications to the study of differentiability properties of Bessel potentials of L^p-functions by the author.

1. Introduction. We study the connection between difference properties and two types of pointwise differentiability of functions defined in R^n. A function $f(x)$ is (ordinarily) differentiable at $x = a$ of order $l \geqslant 0$ if there is a polynomial

$$P(x-a) = \sum_{|\alpha| \leqslant m} C_\alpha (x-a)^\alpha$$

of degree at most m, $m \leqslant l < m+1$, such that

$$R(x) = f(x) - P(x-a) = o(|x-a|^l), \qquad x \to a. \tag{1.1}$$

Several other types of differentiability have been studied, cf. [3], [4] and [6], p. 300. We say that f is *L^p-differentiable at $x = a$ of order l* if for some P as above we have

$$\Big(|B(a,r)|^{-1} \cdot \int_{B(a,r)} |R(x)|^p\,dx\Big)^{1/p} = o(r^l), \qquad r \to 0,\ 1 \leqslant p \leqslant \infty \tag{1.2}$$

[4]. It is clear that (1.1) implies (1.2). The purpose of this paper is to find conditions on the function f which, together with (1.2), are equivalent to (1.1).

The motivation for this study is that many spaces of functions in analysis, for example Sobolev spaces, Besov spaces and Bessel potential

* This work has been supported by the Swedish Natural Science Research Council.

spaces, are known to have L^p-differentials outside some exceptional sets, cf. [2] and [4]. It is therefore natural to ask under what conditions functions in such spaces are ordinarily differentiable at certain points.

Results of this type for general functions with applications to Bessel potentials of L^p-functions were obtained by the author in [7]. The method of proof used Taylor polynomials. Compare also B.-M. Stocke [9] for the case $l = 1$. In the present paper we continue the study of these problems for general functions with a method of proof based on a type of differences introduced by J. Marcinkiewicz and A. Zygmund [5].

Let f be a function satisfying (1.1); then

$$\tag{1.3} \begin{aligned} \Delta_h^j f(a) &= c_j \cdot \sum_{|\alpha|=j} C_\alpha h^\alpha + o(|h|^j), \quad \text{when} \quad j \geqslant 1 \text{ and } j \leqslant m-1, \\ \Delta_h^m f(a) &= c_m \cdot \sum_{|\alpha|=m} C_\alpha h^\alpha + o(|h|^l), \quad \text{when} \quad m \geqslant 1, \\ \Delta_h^{m+1} f(a) &= o(|h|^l), \quad \text{when} \quad m = 0, \end{aligned}$$

as $h \to 0$, where c_j, $1 \leqslant j \leqslant m$, are constants only depending on j and $\Delta_h^j f(a)$ are the usual jth differences of f, cf. [10], p. 102. J. Marcinkiewicz and A. Zygmund [5], Lemma 1, p. 10, proved that (1.1) is equivalent to (1.3) when $\Delta_h^j f$ is replaced by a type of binary differences $B_h^j f$ (see Section 2 for definitions) in the case $n = 1$ and l integer. See also R. J. Bagby [1], where it is shown by an example that (1.3) does not imply (1.1) when $l = 3$. J. Marcinkiewicz and A. Zygmund's result extends to R^n and all $l \geqslant 0$. We begin our study with an L^p-version of their result, where L^p-differentiability (see Section 2 for definitions) is characterized by certain L^p-properties of $B_h^j f(a)$, $1 \leqslant j \leqslant m+1$ (Theorem 3.1).

We define a property of a function $f(x)$ at $x = a$, called C_s^j, which roughly says that $|B_h^j f(x)| \leqslant \varepsilon \cdot |x-a|^s$ when $|x-a|$ and $|h|/|x-a|$ are small enough. Our main result states that ordinary differentiability is equivalent to L^p-differentiability together with a condition expressed in terms of the properties C_s^j, $1 \leqslant j \leqslant m+1$ (Theorem 3.2).

2. Notation and definitions.

2.1. We use standard notation for points x in R^n and functions $f(x)$. The n-dimensional Lebesgue measure of a measurable set E is denoted by $|E|$. Integration with respect to Lebesgue measure is written $\int_E f(x)\,dx$. All functions are Lebesgue measurable.

L^p is the usual Lebesgue space with norm $\|f\|_p$, $1 \leqslant p \leqslant \infty$. A multi-index $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, where α_i are non-negative integers, has length $|\alpha| = \alpha_1 + \alpha_2 + \ldots + \alpha_n$. The differentiation operator D^α is defined by

$$D^\alpha f(x) = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \left(\frac{\partial}{\partial x_2}\right)^{\alpha_2} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n} f(x).$$

We only consider the cases where $D^\alpha f(x)$ is independent of the order of differentiation.

2.2. Polynomials in R^n are denoted by P, Q and R. They are always written in the standard form $P(x) = \sum_{|\alpha| \leqslant m} C_\alpha \cdot x^\alpha$, where $m \geqslant 0$ is an integer. We put $P_a(x) = P(x-a)$, where $a \in R^n$. In order to make our calculations short and clear we define

$$\|f\|_{p,x,r} = \Big(|B(x,r)|^{-1} \cdot \int_{B(x,r)} |f(y)|^p\,dy\Big)^{1/p}, \quad 1 \leqslant p < \infty,$$

and

$$\|f\|_{p,x,r} = \operatorname*{ess\,sup}_{|y-x| \leqslant r} |f(y)|, \quad p = \infty,$$

for all $x \in R^n$ and $r > 0$. This means that $\|f\|_{p,x,r}$ is the L^p-norm of f relative to a certain absolutely continuous measure $\mu = \mu_{x,r}$ with total mass one, only depending on x and r, such that

$$d\mu(y) = |B(x,r)|^{-1} \cdot \chi_{B(x,r)}(y)\,dy.$$

We say that

$$\tag{2.1} f = o(r^\lambda), \quad r \to 0, \quad \text{in} \quad L^p(\mu_{x,r})$$

if $\|f\|_{p,x,r} = o(r^\lambda)$, $r \to 0$. When f depends on parameters we write $L^p(\mu_{x,r}(y))$ in (2.1) to indicate that the L^p-norm is taken with respect to the variable y.

2.3. The standard jth order differences $\Delta_h^j f(x)$ defined by

$$\Delta_h^1 f(x) = f(x+h) - f(x), \quad \Delta_h^{j+1} f(x) = \Delta_h^1(\Delta_h^j f)(x),$$

$j = 1, 2, \ldots$ and the formula

$$\tag{2.2} \Delta_h^j f(x) = \sum_{k=0}^{j} (-1)^{j-k} \cdot \binom{j}{k} \cdot f(x + k \cdot h)$$

are well known, cf. A. F. Timan [10], p. 102. The following type of binary differences $B_h^j f(x)$ were introduced in [5], p. 9. We use the notation from [1].

Definition 2.1. Let j and m be positive integers and define $T(j, m)$ by

$$T(1, m) = 1, \quad T(j+1, m) = (2^m - 2^j) \cdot T(j, m).$$

Definition 2.2. Define $B_h^j f(x)$, $j = 1, 2, 3, \ldots$, by

$$\tag{2.3} \begin{aligned} B_h^1 f(x) &= f(x+h) - f(x), \\ B_h^{j+1} f(x) &= B_{2h}^j f(x) - 2^j \cdot B_h^j f(x), \end{aligned}$$

for any $x, h \in R^n$ and any function f.

It is easily seen that for $j \geqslant 2$

$$T(j, m) = \prod_{\nu=1}^{j-1} (2^m - 2^\nu),$$

and that $T(j, m) = 0$ if and only if $j > m$.

The following elementary properties of $B_h^j f$ will be frequently used, see [1] and [5]. Let $f(x) = (x-a)^\alpha$, $|\alpha| = m \geqslant 1$. Then

(i) $B_h^j f(a) = T(j, m) \cdot h^\alpha$,

(ii) $B_h^m f(x) = T(m, m) \cdot h^\alpha$, for all $x \in R^n$,

(iii) $B_h^j f(x) = 0$, for all $x \in R^n$ when $j > m$.

Let $P(x) = \sum_{|\alpha| \leqslant m} C_\alpha \cdot x^\alpha$ be a polynomial of degree $m \geqslant 1$, then by Taylors formula and (ii) above we get for $1 \leqslant j \leqslant m$

$$B_h^j P(x) = \sum_{j \leqslant |\alpha| \leqslant m} T(j, |\alpha|) \cdot \frac{D^\alpha P(x)}{\alpha!} \cdot h^\alpha. \tag{2.4}$$

For a j times continuously differentiable function f we also have an integral formula

$$B_h^j f(x) = 2^{j(j-1)/2} \cdot \int_0^1 dt_1 \int_{t_1}^{2t_1} dt_2 \ldots \int_{t_{j-1}}^{2t_{j-1}} dt_j \cdot \sum_{|\alpha|=j} \frac{j!}{\alpha!} \cdot D^\alpha f(x + t_j h) \cdot h^\alpha, \tag{2.5}$$

analogous to [10], p. 103. It can be proved by induction over j. The proof is left to the reader.

Further properties of the differences $B_h^j f$ are found in Section 4. Here we only note that $B_h^j f(x) = \Delta_h^j f(x)$ when $j = 1$ and 2 but not when $j \geqslant 3$.

2.4. We will need the notation of an L^p-differential introduced by A. P. Calderón and A. Zygmund in [4].

Definition 2.3. Let $l \geqslant 0$, $1 \leqslant p \leqslant \infty$ and let $f(x)$ be a measurable function defined in a neighbourhood of $x = a$ in R^n. Then, if there is a polynomial $P(x) = \sum_{|\alpha| \leqslant m} C_\alpha \cdot x^\alpha$, of degree $m \leqslant l$ such that

$$\Big(|B(a, r)|^{-1} \cdot \int_{B(a,r)} |f(x) - P(x-a)|^p \, dx\Big)^{1/p} = o(r^l), \quad r \to 0, \tag{2.6}$$

we say that $f(x)$ is *L^p-differentiable at $x = a$ of order l with L^p-differential* $P(x-a)$.

We make the usual modification in (2.6) when $p = \infty$. The L^p-differential is uniquely defined in all cases ([4], p. 172).

2.5. We now define a property of a function $f(x)$ at $x = a$ expressed by the behaviour of $B_h^j f(x)$ when x is close to a, $|h| \leqslant t \cdot |x-a|$ and t is a small positive number. This property is called C_s^j and is by definition of supremum type. In Section 5 we give several equivalent definitions of this property, where $B_h^j f$ is replaced by $\Delta_h^j f$ and/or the supremum norm is replaced by an L^p-norm.

Definition 2.4. Let $s \geqslant 0$ and let j be a positive integer. Let $f(x)$ be a function defined in a neighbourhood of $x = a$ in R^n. Then f *has property C_s^j at $x = a$* if for every $\varepsilon > 0$ there are $\delta > 0$ and $0 < t < \min(\varepsilon, 1)$ such that $0 < |x-a| < \delta$ implies

$$\sup_{|h| \leqslant t|x-a|} |B_h^j f(x)| \leqslant \varepsilon |x-a|^s. \tag{2.7}$$

We prove in Lemma 5.1 that the property C_s^j is equivalently defined if, for $1 \leqslant p \leqslant \infty$, we replace (2.7) by the following requirement:

For every $\varepsilon > 0$ there are $\delta > 0$ and $0 < t < \min(\varepsilon, 1)$ such that $0 < |x-a| < \delta$ implies

$$\Big(|B(0, t|x-a|)|^{-1} \int_{B(0,t|x-a|)} |B_h^j f(x)|^p \, dh\Big)^{1/p} \leqslant \varepsilon |x-a|^s. \tag{2.8}$$

We make the usual modification in (2.8) when $p = \infty$.

Remark. Let $f(x)$ be a polynomial of degree at most m. Then f has property C_s^j at $x = a$ for $0 \leqslant s \leqslant j$, when $1 \leqslant j \leqslant m$ and for all $s \geqslant 0$, when $j > m$. The function $g(x) = e^{-1/|x|}$, $x \neq 0$, $g(0) = 0$ has property C_s^j at $x = 0$ for all $j \geqslant 1$ and $s \geqslant 0$.

3. Main results. Our first theorem is an L^p-generalization of the result of J. Marcinkiewicz and A. Zygmund [5] mentioned in the introduction. Section 3.2 contains our main result, which is a characterization of ordinary differentiability by L^p-differentiability and properties C_s^j for $1 \leqslant j \leqslant m+1$ and suitable s.

3.1. We give the conditions in Theorem 3.1 in two equivalent forms (a)–(c) and (d)–(e).

Theorem 3.1. *Let $l \geqslant 0$, $1 \leqslant p \leqslant \infty$ and let m be the largest integer less than or equal to l. Let $f(x)$ be a measurable function defined in a neighbourhood of $x = a$ in R^n. Then $f(x)$ has an L^p-differential*

$$P(x-a) = \sum_{|\alpha| \leqslant m} C_\alpha (x-a)^\alpha, \quad C_0 = f(a),$$

of order l at a if and only if the following conditions are satisfied:

$$\text{(a)} \quad \Big(|B(0, r)|^{-1} \cdot \int_{B(0,r)} |B_h^j f(a) - T(j, j) \sum_{|\alpha|=j} C_\alpha h^\alpha|^p \, dh\Big)^{1/p} = o(r^j),$$

for $j \geqslant 1$, $j \leqslant m-1$,

(b) $$\Big(|B(0,r)|^{-1}\cdot\int_{B(0,r)}\Big|B_h^m f(a)-T(m,m)\sum_{|\alpha|=m}C_\alpha h^\alpha\Big|^p dh\Big)^{1/p}=o(r^l),\quad \text{for } m\geqslant 1,$$

(c) $$\Big(|B(0,r)|^{-1}\cdot\int_{B(0,r)}|B_h^{m+1}f(a)|^p dh\Big)^{1/p}=o(r^l),\quad \text{as}\quad r\to 0.$$

We exclude (c) *when* $l\geqslant 1$.

Theorem 3.1 holds for ordinary differentiability with all L^p-norms replaced by supremum norms. See [5] for the case $n=1$ and l integer. The general R^n-case is proved analogously.

We refer to this version of Theorem 3.1 as the case of ordinary differentiability (or the supremum norm case).

Remark. Note that the conditions (a) and (b) are empty when $0\leqslant l<2$ and $0\leqslant l<1$, respectively. It can be proved that (b) implies (c) for all $l\geqslant 0$. The proof uses (2.3) and is left to the reader.

The conditions (a)–(c) in Theorem 3.1 can be replaced by the following set of conditions:

(d)

$$\Big(|B(0,r)|^{-1}\cdot\int_{B(0,r)}|B_h^j f(a)-T(j,j)\sum_{|\alpha|=j}C_\alpha h^\alpha|^p dh\Big)^{1/p}=o(r^j),\ r\to 0,\quad \text{when}\quad 1\leqslant j\leqslant m,\tag{3.1}$$

(e)

$$\Big(|B(0,r)|^{-1}\cdot\int_{B(0,r)}|B_h^{m+1}f(a)|^p dh\Big)^{1/p}=o(r^l),\quad \text{as}\quad r\to 0.\tag{3.2}$$

In view of the remark following Theorem 3.1 it is clear that (a)–(c) implies (d)–(e). We now prove the converse.

Assume that (d) and (e) hold. It only remains to prove (b) in the case $l>m\geqslant 1$. Consider the formula

$$B_h^m f(a)=\sum_{k=1}^{N}2^{m(k-1)}B_{h/2^k}^{m+1}f(a)+2^{mN}B_{h/2^N}^m f(a),\tag{3.3}$$

which is obtained by repeated use of (2.3). We then get

$$B_h^m f(a)-T(m,m)\sum_{|\alpha|=m}C_\alpha h^\alpha = \sum_{k=1}^{N}2^{m(k-1)}B_{h/2^k}^{m+1}f(a)+2^{mN}\Big(B_{h/2^N}^m f(a)-T(m,m)\sum_{|\alpha|=m}C_\alpha(h/2^N)^\alpha\Big).\tag{3.4}$$

Taking L^p-norm over $B(0,r)$ in the variable h and using the notation from Section 2.2, we get by a simple change of variables

$$\Big\|B_h^m f(a)-T(m,m)\sum_{|\alpha|=m}C_\alpha h^\alpha\Big\|_{p,0,r}\leqslant\sum_{k=1}^{N}2^{m(k-1)}\|B_h^{m+1}f(a)\|_{p,0,r_k}+ +2^{mN}\Big\|B_h^m f(a)-T(m,m)\sum_{|\alpha|=m}C_\alpha h^\alpha\Big\|_{p,0,r_N},$$

where $r_k=r\cdot 2^{-k}$. Letting $N\to\infty$ and using (d) we get for any $r>0$

$$\Big\|B_h^m f(a)-T(m,m)\sum_{|\alpha|=m}C_\alpha h^\alpha\Big\|_{p,0,r}\leqslant\sum_{k=1}^{\infty}2^{m(k-1)}\|B_h^{m+1}f(a)\|_{p,0,r_k}.\tag{3.5}$$

It now follows from (3.5) and (e) as in [1] that (b) holds. This completes the proof that (a)–(c) in Theorem 3.1 can be replaced by (d)–(e). The same conclusion holds in the case of ordinary differentiability when we take supremum norm in (a)–(e).

3.2. We are now in a position to state our main result. Recall the property C_s^j defined in Section 2.5.

THEOREM 3.2. *Let $l\geqslant 0$ and let m be the largest integer $\leqslant l$. Let $f(x)$ be a function defined in a neighbourhood of $x=a$ in R^n. Then $f(x)$ is ordinarily differentiable at $x=a$ of order l if and only if f is L^p-differentiable at $x=a$ of order l with an L^p-differential whose constant term is $f(a)$ and f has property C_s^j at $x=a$, $1\leqslant j\leqslant m+1$, where*

$$s=\begin{cases} j, & \textit{when}\quad j\geqslant 1\ \textit{and}\ j\leqslant m, \qquad (3.6)\\ l, & \textit{when}\quad j=m+1, \qquad (3.7)\end{cases}$$

Remark. It will be clear from the proof of Theorem 3.2 that conditions (3.6) and (3.7) correspond to (3.1) and (3.2), respectively. Note that (3.6) is empty when $0\leqslant l<1$.

The proof of Theorem 3.2 makes use of the characterization of L^p-differentiability given by Theorem 3.1 (conditions (3.1) and (3.2)) and its analogous version for ordinary differentiability described in Section 3.1.

Before we can go into the details of the proof we need more information about the differences $B_h^m f$. The proof of Theorem 3.2 is given in Section 6.

4. Some lemmas on the differences $B_h^m f$ and $\Delta_h^m f$. We begin this section with a simple lemma about $B_h^m f$. The more technical Lemma 4.2 seems to be new. It will be used in Section 5 (proving Lemma 5.1) and in the proof of Theorem 3.2. The second part of this section studies the relationship between the differences $B_h^m f$ and $\Delta_h^m f$.

4.1. Lemma 4.1 is well known for the differences $\Delta_h^m f$, cf. [10], p. 102. The present case (except for (4.3)) was proved in [5], p. 9.

LEMMA 4.1. *Let m be a positive integer. Then there are unique integers $a_i = a(i, m)$, $0 \leqslant i \leqslant m-1$, and $\beta = \beta(m)$, only depending on i and m, such that*

$$B_h^m f(x) = \sum_{i=0}^{m-1} a_i f(x+2^i h) + \beta f(x), \tag{4.1}$$

for all $x, h \in R^n$ and all functions f.

Further,

$$\sum_{i=0}^{m-1} a_i + \beta = 0, \tag{4.2}$$

$$\operatorname{sgn} a(i, m) = (-1)^{i+1-m}, \quad \operatorname{sgn} \beta(m) = (-1)^m, \quad \text{and} \quad a_{m-1} = 1. \tag{4.3}$$

Proof of Lemma 4.1. The formulas (4.1) and (4.2) were proved in [5]. Combining (4.1) with (2.3) we get the relations

$$\begin{aligned} a(0, m+1) &= -2^m \cdot a(0, m), \\ a(i, m+1) &= a(i-1, m) - 2^m \cdot a(i, m), \qquad 1 \leqslant i \leqslant m-1, \\ a(m, m+1) &= a(m-1, m), \\ \beta(m+1) &= (1-2^m) \cdot \beta(m), \end{aligned}$$

and $a(0, 1) = 1$, $\beta(1) = -1$, for $m \geqslant 1$. The proof of (4.3) is now finished by an induction argument. We omit the details.

Remark. Lemma 4.1 implies that $B_h^m f(x) = \Delta_h^m f(x)$ only when $m = 1$ and 2.

Our next lemma is of a more technical nature. It gives a connection between $B_h^m f(x)$ and $B_k^m f(x)$ for arbitrary h and k. The terms on the right side of (4.4) are linear combinations of the values of f at certain points in R^n. These points make up a simple geometrical figure (a triangle subdivided by a set of parallel lines and a set of lines through a corner of the triangle).

LEMMA 4.2. *Let m be a positive integer. There are unique non-zero integers u_i, $0 \leqslant i \leqslant m-1$, v_i, $0 \leqslant i \leqslant m-2$ and z such that*

$$B_k^m f(x) = \sum_{i=0}^{m-1} u_i \cdot B^m_{2^{i+1-m}(k-h)} f(x+2^i h) + \sum_{i=0}^{m-2} v_i \cdot B^m_{h+2^{i+1-m}(k-h)} f(x) + z \cdot B_h^m f(x), \tag{4.4}$$

for all $x, h, k \in R^n$ and all functions f.

(The middle term on the right-hand side is excluded when $m = 1$.) In fact $u_i = a(i, m)$, $v_i = -a(i, m)$ and $z = -\beta(m)$, the numbers defined in Lemma 4.1.

Proof of Lemma 4.2. The case $m = 1$ is easily settled. We let $m \geqslant 2$ and $h \neq k$. We shall prove that u_i, v_i and z are uniquely determined by the requirement that (4.4) holds for all f. Inserting (4.1) into (4.4), (4.4) becomes a linear equation in the values of f at $(m(m+1)+1)$ different points with coefficients depending on u_i, v_i, z, $a_i = a(i, m)$ and $\beta = \beta(m)$. Thus (4.4) holds for all functions f if and only if all these coefficients are zero. This gives the following system of linear equations in the unknowns u_i, v_i and z:

$$\beta \cdot \sum_{j=0}^{m-2} v_j + \beta \cdot z = \beta, \tag{4.5}$$

$$a_i = u_i \cdot a_{m-1}, \qquad 0 \leqslant i \leqslant m-1, \tag{4.6}$$

$$\beta \cdot u_i + z \cdot a_i = 0, \qquad 0 \leqslant i \leqslant m-1, \tag{4.7}$$

$$u_i \cdot a_j + v_j \cdot a_i = 0, \qquad 0 \leqslant i \leqslant m-1, \quad \text{and} \quad 0 \leqslant j \leqslant m-2. \tag{4.8}$$

Since a_i and β are non-zero by Lemma 4.1, we find that (4.6)–(4.8) has the unique solution stated in the lemma. These values of u_i and v_i satisfies (4.5) because of (4.2). When $h = k$, (4.4) reduces to an identity. The proof of Lemma 4.2 is now complete.

We have the following counterpart of Lemma 4.2 for the differences $\Delta_h^m f$.

LEMMA 4.3. *Let m be a positive integer; then it holds that*

$$\Delta_h^m f(x) = \sum_{i=1}^{m} (-1)^{m-i} \binom{m}{i} \Delta^m_{(i/m)(k-h)} f(x+ih) - \sum_{j=1}^{m} (-1)^{m-j} \binom{m}{j} \Delta^m_{h+(j/m)(k-h)} f(x),$$

for all $x, h, k \in R^n$ and all functions f.

The proof of Lemma 4.3 is straightforward using formula (2.2). We omit the details.

4.2. We remarked in Section 4.1 that $B_h^m f(x)$ and $\Delta_h^m f(x)$ are not identical when $m \geqslant 3$. In spite of this fact we can prove that either of these two differences can be expressed in terms of the other, as described by our next two lemmas. A variant of Lemma 4.4 is proved in [5], p. 11 (except for the uniqueness of the coefficients $u(i, m)$).

LEMMA 4.4. *For every positive integer m, there are unique integers*

$u(i,m)$, $0 \leqslant i \leqslant N_m$, $N_m = 2^{m-1}-m$, such that

$$B_h^m f(x) = \sum_{i=0}^{N_m} u(i,m)\cdot \Delta_h^m f(x+i\cdot h), \tag{4.9}$$

for all $x, h \in R^n$ and all functions f.

Proof of Lemma 4.4. The cases $m=1$ and $m=2$ are easily settled by the remark following Lemma 4.1, since $N_1 = N_2 = 0$.

We prove the existence of the integers $u(i,m)$ by induction over m. Assume that there exist integers $u(i,m)$, $0 \leqslant i \leqslant N_m$, such that (4.9) holds for all x, h and f. Then by (2.3) and the induction hypothesis we get

$$\begin{aligned} B_h^{m+1} f(x) &= B_{2h}^m f(x) - 2^m\cdot B_h^m f(x) \\ &= \sum_{i=0}^{N_m} u(i,m) \sum_{j=1}^{m}\sum_{k=0}^{j-1}\binom{m}{j}\Delta_h^{m+1} f(x+2ih+kh) + \\ &\qquad + \sum_{i=0}^{N_m} u(i,m)\cdot 2^m\cdot\big(\Delta_h^m f(x+2ih) - \Delta_h^m f(x+ih)\big) \\ &= I + II, \end{aligned}$$

where we have used the formula

$$\Delta_{2h}^m f(x) - 2^m \Delta_h^m f(x) = \sum_{j=1}^{m}\sum_{k=0}^{j-1}\binom{m}{j}\Delta_h^{m+1} f(x+kh),$$

c.f. [10], p. 105. In the term II we write

$$\Delta_h^m f(x+2ih) - \Delta_h^m f(x+ih) = \sum_{j=i}^{2i-1}\Delta_h^{m+1} f(x+jh).$$

It is now clear that

$$B_h^{m+1} f(x) = \sum_{i=0}^{N'_m} u(i,m)'\cdot \Delta_h^{m+1} f(x+ih)$$

where $u(i,m)'$ are integers and $N'_m \leqslant 2\cdot N_m + (m-1) = N_{m+1}$. This completes the proof of the existence of the integers $u(i,m)$, $0 \leqslant i \leqslant N_m$, for all $m \geqslant 1$.

In order to prove uniqueness of the integers $u(i,m)$, $0 \leqslant i \leqslant N_m$ in (4.9), we assume that

$$\sum_{i=0}^{N_m} y_i\cdot \Delta_h^m f(x+ih) = 0, \tag{4.10}$$

holds for all $x, h \in R^n$ and all functions f. Now (4.10) is equivalent to

$$\sum_{i=0}^{N_m}\sum_{j=0}^{m} y_i\cdot\binom{m}{j}\cdot(-1)^{m-j}\cdot f\big(x+(i+j)h\big) = 0$$

which holds for all functions f if and only if

$$\sum_{i+j=k} y_i\cdot\binom{m}{j}\cdot(-1)^{m-j} = 0, \tag{4.11}$$

for every $0 \leqslant k \leqslant 2^{m-1}$. It is however easy to prove that (4.11) implies $y_i = 0$, $0 \leqslant i \leqslant N_m$. We leave the details to the reader. This completes the proof of Lemma 4.4.

LEMMA 4.5. *For every positive integer m, there are real numbers $c(i,j) = c(i,j,m)$ such that*

$$\Delta_{2^{m-1}h}^m f(x) = \sum_A c(i,j)\cdot B_{ih}^m f(x+jh) \tag{4.12}$$

holds for all $x, h \in R^n$ and all functions f, where

$$A = \{(i,j);\ i \geqslant 1, j \geqslant 0 \text{ and } 2^{m-1}\cdot i + j \leqslant 2^{m-1}\cdot m\}.$$

Remark. Note that $i \leqslant m$ and $j \leqslant 2^{m-1}\cdot(m-1)$, where $(i,j)\in A$.

Proof of Lemma 4.5. The set A contains all pairs (i,j) for which $1 \leqslant i \leqslant m$ and $0 \leqslant j \leqslant 2^{m-1}\cdot(m-i)$. Hence we have $\operatorname{Card} A = 2^{m-2}\times\times m(m-1)+m$. Let $h \neq 0$ and define

$$E = \{x + k\cdot h;\ k \text{ integer and } 0 \leqslant k \leqslant m\cdot 2^{m-1}\}.$$

Then the left hand side of (4.12) is a linear combination of the values of f at the points in E with known coefficients given by (2.2). The right hand side of (4.12) is by Lemma 4.1 a linear combination of the values of f at the points of E whose coefficients are linear forms in the variables $c(i,j)$. Now assume that (4.12) holds for all functions f. Then (4.12) is equivalent to a system L of linear equations in the variables $c(i,j)$. The number of variables is $\operatorname{Card} A$ and the number of equations is $\operatorname{Card} E$. Conversely, any solution $c(i,j)$, $(i,j)\in A$, of the system L implies that (4.12) holds for all functions f. It therefore suffices to prove that the system L has a solution. Lemma 4.5 clearly holds when $m=1$ and $m=2$ since then $B_h^m f = \Delta_h^m f$. When $m \geqslant 3$ we have $\operatorname{Card} A = 2^{m-2}\cdot m(m-1)+m > 2^{m-1}\cdot m+1 = \operatorname{Card} E$. Now it follows from the general theory of systems of linear equations that L has a (not necessarily unique) solution. This completes the proof of Lemma 4.5.

5. Some lemmas on property C_s^j. In this section we prove that property C_s^j can be equivalently defined with the supremum norm replaced by an L^p-norm, $1 \leqslant p \leqslant \infty$ and/or the differences $B_h^j f$ replaced by the differences $\Delta_h^j f$. This is done in a series of lemmas. In Section 5.1 we show how to pass from an L^p-norm to supremum norm and in Section 5.2 we prove that the differences $B_h^j f$ and $\Delta_h^j f$ are equivalent in a certain sense.

5.1. We start with the case of the differences $B_h^j f$.

LEMMA 5.1. *Let $s \geqslant 0$ $1 \leqslant p \leqslant \infty$ and let j be a positive integer. Let $f(x)$ be a function defined in a neighbourhood of $x = a$ in R^n. Then $f(x)$ has property C_s^j at $x = a$ if and only if* (2.8) *holds.*

Proof of Lemma 5.1. It is evident that property C_s^j implies (2.8). Assume that (2.8) holds, $p = 1$ and that ε is an arbitrary positive number. Choose δ and t according to (2.8). We now use Lemma 4.2 to split $B_h^j f(x)$ into a linear combination of the following three types of terms:

$$B_k^j f(x),\ B^j_{2^{i+1-j}\cdot(k-h)} f(x+2^i h) \quad \text{and} \quad B^j_{h+2^{i+1-j}\cdot(k-h)} f(x). \tag{5.1}$$

Considering (5.1) together with (2.7) we are led to define the following sets

$$E_i = \{k;\ 2^{i+1-j}\cdot|k-h| \leqslant t\cdot|x-a+2^i h|\}, \quad 0 \leqslant i \leqslant j-1,$$
$$F_i = \{k;\ |h+2^{i+1-j}\cdot(k-h)| \leqslant t\cdot|x-a|\}, \quad 0 \leqslant i \leqslant j-2,$$
$$G = \{k;\ |k| \leqslant t\cdot|x-a|\}.$$

Let H be the intersection of all sets E_i, F_i and G. We are going to prove that

$$\text{all sets } E_i, F_i \text{ and } G \text{ have measure} \leqslant c(j, n)\cdot|H|, \tag{5.2}$$

provided $|h| < 2^{-1-j}\cdot t\cdot|x-a|$ and $0 < t \leqslant 1$. First we note that all sets E_i, F_i and G have measure less than $c(j, n)\cdot(t\cdot|x-a|)^n$. The proof of (5.2) is completed by showing that

$$\{k;\ |k| < 2^{-1-j}\cdot t\cdot|x-a|\} \subset H. \tag{5.3}$$

The proof of (5.3) is straightforward and is omitted. This proves (5.2). Now by Lemma 4.2 we get that

$$|B_h^j f(x)| = |H|^{-1}\cdot \int_H |B_h^j f(x)|\, dk$$

is majorized by $c(j, n)$ times a sum of terms of the three types

$$|G|^{-1}\cdot \int_G |B_k^j f(x)|\, dk, \tag{5.4}$$

$$|E_i|^{-1}\cdot \int_{E_i} |B^j_{2^{i+1-j}(k-h)} f(x+2^i h)|\, dk, \quad 0 \leqslant i \leqslant j-1, \tag{5.5}$$

and

$$|F_i|^{-1}\cdot \int_{F_i} |B^j_{h+2^{i+1-j}(k-h)} f(x)|\, dk, \quad 0 \leqslant i \leqslant j-2. \tag{5.6}$$

By our assumptions and a simple change of variables in (5.5) and (5.6), we find that there is $\delta > 0$ such that (5.4), (5.5) and (5.6) are less than $c(j, n)\cdot\varepsilon\cdot|x-a|^s$, provided $0 < |x-a| < \delta/2$ and $|h| < 2^{-1-j}\cdot t\cdot|x-a|$. This proves that $f(x)$ has property C_s^j at $x = a$ when $p = 1$. The case $p > 1$ now follows from Hölders inequality. This completes the proof of Lemma 5.1.

Our next lemma and its proof are the analogue of Lemma 5.1 with $B_h^j f$ replaced by $\Delta_h^j f$ and using Lemma 4.3 instead of Lemma 4.2 in the proof. The proof is therefore omitted.

LEMMA 5.2. *Let $s \geqslant 0, 1 \leqslant p \leqslant \infty$ and let j be a positive integer. Let $f(x)$ be a function defined in a neighbourhood of $x = a$ in R^n. Then the following two properties of f are equivalent:*

(a) *f satisfies Definition 2.4 with $B_h^j f$ replaced by $\Delta_h^j f$,*

(b) *f satisfies (2.8) with $B_h^j f$ replaced by $\Delta_h^j f$.*

5.2. Here we prove that property C_s^j can be equivalently defined with the difference $B_h^j f$ replaced by $\Delta_h^j f$.

LEMMA 5.3. *Let $s \geqslant 0$ and let j be a positive integer. Let $f(x)$ be a function defined in a neighbourhood of $x = a$ in R^n. Then f has property C_s^j at a if and only if (2.7) holds with $B_h^j f$ replaced by $\Delta_h^j f$.*

Proof of Lemma 5.3. First suppose that f has property C_s^j at a. Let ε be an arbitrary positive number and choose $\delta > 0$ and $0 < t < \min(\varepsilon, 1)$ according to (2.7). By Lemma 4.5 we have

$$\Delta_h^j f(x) = \sum_A c(i, k)\cdot B^j_{2^{1-j}\cdot ih} f(x+k\cdot 2^{1-j}\cdot h). \tag{5.7}$$

Easy calculations show that for all x and h satisfying $0 < |x-a| < \frac{2}{3}\delta$ and $|h| \leqslant \min\left(1/2j, (1/j)\cdot 2^{j-2}t\right)\cdot|x-a|$, (5.7) implies that

$$|\Delta_h^j f(x)| \leqslant \sum_A |c(i, k)|\cdot\varepsilon\cdot|x+k\cdot 2^{1-j}h-a|^s \leqslant 2^s \sum_A |c(i, k)|\cdot\varepsilon\cdot|x-a|^s.$$

This proves that (2.7) holds with $B_h^j f$ replaced by $\Delta_h^j f$. The converse is proved analogously and the proof is therefore omitted.

5.3. Lemmas 5.1–5.3 now give the following result.

LEMMA 5.4. *Let j be a positive integer and let $s \geqslant 0$. Then property C_s^j is equivalently defined when the supremum norm in (2.7) is replaced by an L^p-norm, $1 \leqslant p \leqslant \infty$, (as in (2.8)) and/or the difference $B_h^j f$ is replaced by $\Delta_h^j f$.*

Lemma 5.4 implies that Theorem 3.2 can be formulated using the differences $\Delta_h^j f$ instead of $B_h^j f$ (in the definition of property C_s^j) although our proof of Theorem 3.2 requires the differences $B_h^j f$ (via Theorem 3.1).

Roughly speaking, Lemma 5.4 means that the differences $B_h^j f$ and $\Delta_h^j f$ are equivalent in the sense of property C_s^j expressed in supremum form (Definition 2.4) or in L^p-form ((2.8)).

6. Proof of Theorems 3.1 and 3.2.

6.1. Proof of Theorem 3.1. Theorem 3.1 was proved in the case of ordinary differentiability, where $n = 1$ and l are integers in [5], p. 10. See also [1]. The same method of proof can be used in our L^p-case and therefore the details are left to the reader.

6.2. Proof of Theorem 3.2. The necessity part of Theorem 3.2 is obvious in view of the remark following Definition 2.4. We omit the details. The idea of proof of the sufficiency part of the theorem is to use the characterizations of L^p-differentiability and ordinary differentiability (conditions (3.1) and (3.2)) given by the two versions of Theorem 3.1 described in Section 3.1.

Let $P(x-a) = \sum_{|a|\leqslant m} C_a(x-a)^a$ be the L^p-differential of $f(x)$ at $x = a$. Lemma 4.2 gives for $1 \leqslant j \leqslant m$

$$\begin{aligned}(6.1)\qquad \beta\Big(B_h^j f(a) - T(j,j)\cdot \sum_{|a|=j} C_a h^a\Big) &= \sum_{r=0}^{j-1} a_r B^j_{2^{r+1-j}(k-h)} f(a+2^r h) - \\ &- \sum_{r=0}^{j-1} a_r\Big(B^j_{h+2^{r+1-j}(k-h)} f(a) - T(j,j) \sum_{|a|=j} C_a\big(h+2^{r+1-j}(k-h)\big)^a\Big) + \\ &+ T(j,j)\sum_{r=0}^{j-1} a_r \cdot \sum_{|a|=j} C_a\Big(h^a - \big(h+2^{r+1-j}(k-h)\big)^a\Big) \\ &= I + II + III.\end{aligned}$$

Let $\varepsilon > 0$ and choose $\delta > 0$ and $0 < t < \min(\varepsilon, 1)$ according to the definition of property C_j^j. Then for any fixed h, $|h| < 2^{1-j}\cdot\delta$, we get

$$\sup_{|k-h|\leqslant t|h|} |I| \leqslant \varepsilon \sum_{r=0}^{j-1} |a_r|\cdot(2^r\cdot|h|)^j \leqslant \varepsilon\cdot \sum_{r=0}^{j-1} |a_r|\cdot 2^{(j-1)j}\cdot|h|^j = c(j,n)\cdot\varepsilon\cdot|h|^j,$$

and

$$\begin{aligned}\sup_{|k-h|\leqslant t|h|} |III| &\leqslant c(j,n)|T(j,j)|\cdot\sum_{r=0}^{j-1}|a_r|\cdot\sum_{|a|=j}|C_a|\cdot\varepsilon\cdot|h|^j \\ &= c(j,n)\cdot\sum_{|a|=j}|C_a|\cdot\varepsilon|h|^j.\end{aligned}$$

We take L^p-norm in (6.1) w.r.t. k over $|k-h| \leqslant t|h|$. Then we make the change of variables $h+2^{r+1-j}(k-h) = z$ in each term in II and note

that $|z| \leqslant (1+t)\cdot|h|$. This yields the estimate

$$\begin{aligned}(6.2)\qquad |\beta|\cdot\Big|B_h^j f(a) - T(j,j)\sum_{|a|=j} C_a h^a\Big| &\leqslant 2^{(j-1)/p}\sum_{r=0}^{j-1}|a_r|\cdot(1+1/t)^{n/p}\cdot\big\|B_z^j f(a) - T(j,j)\cdot\sum_{|a|=j} C_a z^a\big\|_{p,0,(1+t)|h|} + \\ &+ c(j,n)\cdot\Big(1+\sum_{|a|=j}|C_a|\Big)\cdot\varepsilon\cdot|h|^j,\end{aligned}$$

for every fixed h, $|h| < 2^{1-j}\cdot\delta$, where we have used the notation from Section 2.2. The L^p-norm in the first term in (6.2) is by assumption $o(|h|^j)$ as $h\to 0$.

It now follows that for $1 \leqslant j \leqslant m$

$$B_h^j f(a) - T(j,j)\sum_{|a|=j} C_a\cdot h^a = o(|h|^j),$$

as $h \to 0$.

The case $j = m+1$ is treated analogously. We leave the details to the reader. The proof of Theorem 3.2 is now finished by an application of the supremum norm case of Theorem 3.1.

7. A counterexample.

R. J. Bagby [1] considered the function $f(x) = (-1)^{j+k}\cdot x^3$, when $x = \pm 2^j\cdot 3^k$ and $f(x) = 0$ elsewhere. Then $\Delta_h^j f(0) = o(|h|^j)$, $j = 1, 2, 3$ (in fact, $\Delta_h^3 f(0) \equiv 0$) but f is not differentiable of order 3 at $x = 0$. This proves that Theorem 3.1 is false for ordinary differentiability with $B_h^j f$ replaced by $\Delta_h^j f$ when $n = 1$ and $l = 3$.

It remains an open question if there is an analogous counter-example for the L^p-case of Theorem 3.1.

Bagby's example indicates that the differences $B_h^j f(a)$ are better adopted than $\Delta_h^j f(a)$ to describe differentiability of $f(x)$ at $x = a$. However it do es not rule out that there is a formulation of Theorem 3.1 only using the differences $\Delta_h^j f(a)$ (and not $B_h^j f(a)$).

Acknowledgement. I would like to thank Professor R. J. Bagby for pointing out an error in my attempt to extend his example in Section 7 to the L^p-case.

References

[1] R. J. Bagby, *Taylor polynomials and difference quotients*, Amer. Math. Monthly 86 8 (1979), 681–684.

[2] T. Bagby, W. P. Ziemer, *Pointwise differentiability and absolute continuity*, Trans. Amer. Math. Soc. 191 (1974), 129–148.

[3] A. M. Bruckner, *Differentiation of real functions*, Lecture Notes in Mathematics, vol. 659, Springer-Verlag, Berlin, Heidelberg 1978.
[4] A. P. Calderón and A. Zygmund, *Local properties of solutions of elliptic partial differential equations*, Studia Math. 20 2 (1961), 171–225.
[5] J. Marcinkiewicz and A. Zygmund, *On the differentiability of functions and summability of trigonometrical series*, Fund. Math. 26 (1936), 1–43.
[6] S. Saks, *Theory of the integral*, Monografie Matematyczne 7, Warszawa–Lwów 1937; Dover Publications, New York 1964.
[7] T. Sjödin, *On ordinary differentiability of Bessel potentials*, Department of Math., Univ. of Umeå, Report No 6, 1980.
[8] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Math. Series, No 30, Princeton Univ. Press, Princeton, N. J. 1970.
[9] B-M. Stocke, *Differentiability of Bessel potentials*, Department of Math., Univ. of Umeå, Report No 3, 1980.
[10] A. F. Timan, *Theory of approximation of functions of a real variable*, International Series of Monographs in Pure and Applied Mathematics, Pergamon Press, 1963. [Translated from Russian.]

Received March 4, 1981 (1675)

STUDIA MATHEMATICA, T. LXXIV. (1982)

On Morrey–Besov inequalities

by

DAVID R. ADAMS and JOHN L. LEWIS (Lexington, Kentucky)

Abstract. The authors consider functions u which satisfy a local integrability condition of Morrey–Besov type. They prove that u is locally in a certain Lorentz space. Examples are given to show this result is best possible.

1. Introduction. Let x be a point in n dimensional Euclidean space $\boldsymbol{R}^n$, $|x|$ the Euclidean norm of x, dx Lebesgue measure on $\boldsymbol{R}^n$, and $|E|$ the Lebesgue measure of a measurable set $E \subset \boldsymbol{R}^n$. If Q_0 is a cube in $\boldsymbol{R}^n$, let $L^p(Q_0)$, $1 \leqslant p \leqslant \infty$, denote the usual Lebesgue space of functions that are pth power integrable on Q_0. Then in [8], Ross shows that any function $u \in L^p(Q_0)$ which satisfies

$$\int_Q |u(x+t)-u(x)|^p\,dx \leqslant A^p |t|^{\alpha p} |Q|^{1-\lambda/n}, \tag{1.1}$$

whenever Q and $Q+t = \{x+t\colon x \in Q\}$ are parallel subcubes of Q_0, also belongs to $L^r(Q_0)$ for all $r < p^* = \lambda p/(\lambda - \alpha p)$. Here $1 \leqslant p < \lambda/\alpha$, $0 < \alpha \leqslant 1$, $0 < \lambda \leqslant n$, and A is a positive constant independent of $t \in \boldsymbol{R}^n$ and Q. (1.1) can be viewed as a mixture of the usual Nikol'skii condition ([6], p. 153) when $\lambda = n$, and a Morrey type condition involving the αth difference quotient. When $\lambda = n$, it is known [4], that (1.1) implies $u \in$ weak-$L^{p^*}(Q_0)$, whereas when $\alpha = 1$, it is known [1] that condition (1.1) implies $u \in L^{p^*}(Q_0)$. Thus we are motivated to determine, in all cases, the "best integrability class" for functions satisfying (1.1) as well as for a related Morrey–Besov condition, (1.2) below. To be more specific, let $L(a, b)$, $1 \leqslant a < \infty$, $1 \leqslant b \leqslant \infty$, denote the Lorentz space of measurable functions f on $\boldsymbol{R}^n$ with

$$\|f\|_{a,b} = \Big(\int_0^\infty \big[s\cdot|\{x\colon |f(x)| > s\}|^{1/a}\big]^b s^{-1}\,ds\Big)^{1/b} < \infty.$$

Note that $L(a, a) = L^a(\boldsymbol{R}^n)$, $L(a, \infty) =$ weak-$L^a(\boldsymbol{R}^n)$, and $L(a, b_1) \subseteq L(a, b_2)$ when $b_1 \leqslant b_2$. Given a function u, let $\Delta_t u(x) = u(x+t) - u(x)$, whenever x and $x+t$ are in the domain of definition of u. Inductively, define $\Delta_t^k u$ for $k \geqslant 2$, a positive integer, by $\Delta_t^k u = \Delta_t(\Delta_t^{k-1} u)$. We then