TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 282, Number 2, April 1984

MODULI OF CONTINUITY IN R^n AND $D \subset R^n$

BY

Z. DITZIAN[1]

ABSTRACT. The r modulus of continuity for $f \in C(R^n)$ is expressed in terms of r moduli of continuity in n independent directions. Generalizations to other spaces of functions on R^n or $D \subset R^n$ are also given.

1. For a function $f(x)$, $x \in R^n$, the r modulus of continuity is given by

$$\omega_r(f,t) = \operatorname*{Sup}_{\substack{|h| \le t \\ h,\, x \in R^n}} |\Delta_h^r f(x)| \tag{1.1}$$

where $\Delta_h f(x) \equiv f(x+h) - f(x)$ and $\Delta_h^r \equiv \Delta_h(\Delta_h^{r-1})$.

Examining $\omega_r(f,t)$, one observes that

$$\omega_r(f,t) = \operatorname*{Sup}_{e \in S_{n-1}} \operatorname*{Sup}_{\substack{0<h\le t \\ x \in R^n}} |\Delta_{he}^r f(x)| \quad \text{where } S_{n-1} = \{y \in R^n : |y| = 1\}. \tag{1.2}$$

In other words, the r modulus of continuity is the supremum on all directional moduli of continuity which are given by

$$\omega_r(f,t,e) \equiv \operatorname*{Sup}_{\substack{0 \le h \le t \\ x \in R^n}} |\Delta_{he}^r f(x)|. \tag{1.3}$$

We will show in this paper that if $\omega_r(f,t,e_i) = O(t^\alpha)$, $\alpha < r$, for n independent, e_i, $e_i \in S_{n-1}$, then $\omega_r(f,t) = O(t^\alpha)$. It is obvious that the information for $n-1$ directions is not sufficient. It will be shown by example that $\omega_r(f,t,e_i) = O(t^r)$ for n directions does not imply $\omega_r(f,t) = O(t^r)$. In fact, our "Marchaud-type" estimate will yield, in such a case, $\omega_r(f,t) = O(t^r \log 1/t)$ and the examples will show that in the $C(R^n)$ norm this is sharp. These results remain true for function spaces defined on R^n or T^n (where f defined on $T \equiv [-\pi,\pi]$ is 2π periodic) for which translation is an isometry and with appropriate choice of $e_1 \cdots e_n$ to function spaces on R_+^n for which translation is a contraction. (For $f(x)$, $x \in (R_+)^n$, or $x_i \ge 0$, we choose $e_1,\ldots,e_n \in R_+^n \cap S_{n-1}$, where S_{n-1} is the unit sphere in R^n.)

Finally the results will be proved in such a way that generalizations to $C(D)$ for some domains D will follow as well as to $L_p(D)$ and other function spaces which are lattice compatible and satisfy some mild restrictions.

Received by the editors February 16, 1983.

1980 *Mathematics Subject Classification.* Primary 26B99, 46E35, 41A25; Secondary 41A63.

Key words and phrases. Moduli of continuity in R^n smoothness of function spaces, multivariate semigroups of contractions.

[1] Supported by NSERC grant of Canada A4816.

The key to these results is a "Marchaud-type" result for mixed differences. Mixed differences were expressed in terms of directional differences in the elegant Kemperman Lemma [**3**, pp. 123–124] but there an explicit use of all directional moduli of continuity is made and here only n directions are necessary.

2. Mixed differences. The main tool of our paper is a combinatorial estimate which is of the same nature as the Marchaud inequality in R that yields

$$\omega_k(f,t) \leqslant c_k t^k \left\{ \int_t^c \frac{\omega_{k+1}(f,u)}{u^{k+1}} du + \|f\| \right\}. \tag{2.1}$$

The present result is different in content but similar in form.

For example, for $r = 2$ in R^2 the mixed difference is given by

$$\Delta_{he_1}\Delta_{ke_2} f(x) \equiv f(x + he_1 + ke_2) - f(x + he_1) - f(x + ke_2) + f(x). \tag{2.2}$$

We can now write

$$\begin{aligned} \Delta_{he_1}\Delta_{ke_2} f(x) - \tfrac{1}{4}\Delta_{2he_1}\Delta_{2ke_2} f(x) &= \tfrac{3}{8}\Delta_{he_1}^2 f(x) + \tfrac{3}{8}\Delta_{ke_2}^2 f(x) \\ &\quad - \tfrac{1}{4}\Delta_{he_1}^2 f(x + ke_2) - \tfrac{1}{4}\Delta_{ke_2}^2 f(x + he_1) \\ &\quad - \tfrac{1}{8}\Delta_{he_1}^2 f(x + 2ke_2) - \tfrac{1}{8}\Delta_{ke_2}^2 f(x + 2he_1). \end{aligned} \tag{2.3}$$

To use (2.3) and similar estimates for $L_p(D)$ where D is not R^n or R_+^n, we will use the fact that we have the difference at certain points related to x rather than the maximum, but just illustrate the type of results we recall that, in $C(R^2)$, (2.3) implies

$$\left|\Delta_{he_1}\Delta_{he_2} f(x)\right| \leqslant \tfrac{1}{4}\left|\Delta_{2he_1}\Delta_{2ke_2} f(x)\right| + \tfrac{3}{4}\omega_2(f,h,e_1) + \tfrac{3}{4}\omega_2(f,k,e_2) \tag{2.4}$$

or

$$\left|\Delta_{he_1}\Delta_{he_2} f(x)\right| \leqslant \frac{3}{4} \sum_{i=0}^{n} 4^{-i}\left\{\omega_2(f,2^i h, e_1) + \omega_2(f, 2^i k, e_2)\right\} + 4^{-n+1}\|f\|.$$

Choosing n so that $2^n h = a$, $2^n k = b$, we have

$$\left|\Delta_{he_1}\Delta_{ke_2} f(x)\right| \leqslant C\left\{ h^2 \int_h^a \frac{\omega_2(f,u,e_1)}{u^3} du + k^2 \int_k^t \frac{\omega_2(f,u,e_2)}{u^3} du + h^2\|f\| \right\}. \tag{2.5}$$

It would be nice to have a closed formula like (2.3) for higher differences but I found such formulae only for $r = 2, 3$ and 4, and the estimate for $r = 3$, 4 would not imply as good a result for a finite domain D and for R_+^n as that of $r = 2$ above or the general case below.

We define $E(h)f(x) = f(x+h)$ and state our result for mixed differences.

THEOREM 2.1. *For a bounded function $f(x)$ defined on $x + B$ where B is the parallelepiped spanned by $a_i e_i$ $\sum_{i=1}^n k_i = r$, $k_i \geqslant 0$, and $2^p h_i = a_i$, we have*

(2.6)

$$\left|\Delta_{h_1 e_1}^{k_1} \cdots \Delta_{h_n e_n}^{k_n} f(x)\right| \leqslant C\left(\sum_{\substack{i=1\\k_i\neq 0}} \left(\frac{h_i}{a_i}\right)^r A(a_1,\ldots,a_n)|f(x)| + \sum_{\substack{i=1\\k_i\neq 0}}^n \sum_{l=0}^{p-N} 2^{-lr} A_i(M2^l h_1,\ldots,M2^l h_n)\left|\Delta_{2^l h_i e_i}^r f(x)\right|\right)$$

where $a_i = h_i = 0$ if $k_i = 0$, N and M are integers which depend only on r and $(k_1,\ldots,k_n)$, and A_i is given by

$$A_i(\eta_1,\ldots,\eta_n) = \sum_j w_{ij} E\left(\sum_l h_{l_j} e_l\right)$$

where $w_{ij} \geqslant 0$, $\sum_j w_{ij} = 1$ and $0 \leqslant h_{l_j} \leqslant \eta_l$. Neither $A_i(\cdots)$ nor C depend on f or x.

REMARK 2.2. (a) If the interest were just in showing that $\omega_r(f, t, e_i) = O(t^\alpha)$ implies $\omega_r(f, t) = O(t^\alpha)$ for $\alpha < r$, we could prove the theorem with only 2 dimensions and the rest would follow easily.

(b) If the interest were just in R^n, T^n or R_+^n the expression would be simplified by dropping the averages (and writing maxima of the differences instead). The same is true for the theorem on $C(D)$, $D \subset R^n$, but not for $L_p(D)$, for example.

3. Mixed differences; proof of the result. In this section we will prove Theorem 2.1 in the general form that will be the basis for our results in later sections.

PROOF OF THEOREM 2.1. We recall that in proving the one-dimensional Marchaud-type inequality (see Timan [**6**, p. 105]) one has, for x and h in R,

$$\Delta_{2h}^k f(x) - 2^k \Delta_h^k f(x) = \sum_{\nu=0}^{k-1} \sum_{\mu=\nu+1}^{k} \binom{k}{\mu} \Delta_h^{k+1} f(x+\nu h). \tag{3.1}$$

We can, with no change in the proof, replace $x \in R$ by $x \in R^n$, and $h \in R$ by he, $e \in S_{n-1}$, and have

$$(3.2)\quad \Delta_{he}^{k} f(x) = 2^{-k}\Delta_{2he}^{k} f(x) - 2^{-k}\sum_{\nu=0}^{k-1}\sum_{\mu=\nu+1}^{k}\binom{k}{\mu}\Delta_{he}^{k+1} f(x+\nu he)$$

$$\text{for } 0 < h,\ e \in S_{n-1}.$$

To estimate $\Delta_{h_1e_1}^{k_1}\cdots\Delta_{h_ne_n}^{k_n} f(x)$ we may proceed using (3.2) as long as f is defined on $x + \sum \nu_i h_i e_i$ for the appropriate $\nu_i h_i e_i$ which will be guaranteed by the condition that f is defined on $x + B$. We write, for $r = k_1 + \cdots + k_n$,

$$\begin{aligned}
\left|\Delta_{h_1e_1}^{k_1}\cdots\Delta_{h_ne_n}^{k_n} f(x)\right| &\le 2^{-k_1}\left|\Delta_{2h_1e_1}^{k_1}\Delta_{h_2e_2}^{k_2}\cdots\Delta_{h_ne_n}^{k_n} f(x)\right| \\
&\quad + \frac{1}{2}k_1\sum_{\nu=0}^{k_1-1} w_\nu(1)\left|\Delta_{h_1e_1}^{k_1+1}\Delta_{h_2e_2}^{k_2}\cdots\Delta_{h_ne_n}^{k_n} f(x+\nu h_1e_1)\right| \\
&\le \cdots \le 2^{-r}\left|\Delta_{2h_1e_1}^{k_1}\cdots\Delta_{2h_ne_n}^{k_n} f(x)\right| \\
&\quad + \sum_{j=1}^{n}\frac{1}{2}k_j\sum_{\nu=0}^{k_j-1} w_\nu(j)\left(\prod_{i=1}^{j-1}2^{-k_i}\Delta_{2h_ie_i}^{k_i}\right)\left(\prod_{i=j+1}^{n}\Delta_{h_ie_i}^{k_i}\right)\Delta_{h_je_j}^{k_j+1} f(x+\nu h_je_j),
\end{aligned}$$

where $w_\nu(j) \ge 0$ and $\sum_{\nu=0}^{k_j-1} w_\nu(j) = 1$ since

$$\sum_{\nu=0}^{k_j-1}\sum_{\mu=\nu+1}^{k_j}\binom{k_j}{\mu} = k_j 2^{k_j-1}.$$

Continuing the above process by starting from $\prod_{j=1}^{n}\Delta_{2h_je}^{k_j} f(x)$ and later from $\prod_{j=1}^{n}\Delta_{2^lh_ie_i}^{k_i} f(x)$, we have

$$\begin{aligned}
(3.3)\quad \left|\Delta_{h_1e_1}^{k_1}\cdots\Delta_{he_n}^{k_n} f(x)\right| &\le \sum_{l=0}^{p-N} 2^{-lr}\sum_{j=1}^{n}\frac{1}{2}k_j\sum_{\nu=0}^{k_j-1} w_\nu(j) \\
&\quad \times\left|\left(\prod_{i=1}^{j-1}2^{-k_i}\Delta_{2^{l+1}h_ie_i}^{k_i}\right)\left(\prod_{i=j+1}^{n}\Delta_{2^lh_ie_i}^{k_i}\right)\Delta_{2^lh_je_j}^{k_j+1} f\left(x+\nu 2^l h_je_j\right)\right| \\
&\quad + 2^{-(p-N)r}2^r\|f\|.
\end{aligned}$$

We need to choose N such that $f(x)$ is defined for

$$x + \left\{(k_j+1) + (k_j-1)\right\}2^l h_j e_j$$

or

$$2k_j 2^{p-N} h_j < a_j,$$

but in order to continue our estimate, we choose $2k_j2^{p-N}h_j < 2r2^{p-N}h_j < a_j/L + 1$ where $L = n(r-2)+1$ and therefore $2^{-(p-N)r}2^r \le C(h_j/a_j)^r$ where C does not depend on f, a_j or h_j. We have completed the first step in our proof that has L similar steps where $L = n(r-2)+1$. In fact, for the case $r = 2$ we can deduce our theorem from (3.3) since for $k_j \neq 0$, $k_j + 1 = 2$ and we have

$$\left|\Delta_{2\alpha h_i e_i}\Delta^2_{h_i e_i} f\right| \leqslant 2A(\cdot,\dots,\cdot)\left|\Delta^2_{h_i} f(x)\right|.$$

We observe that in (3.3) we expressed mixed differences by sums of mixed differences of higher order by 1. The induction hypothesis is that after K steps we have the estimate

$$\begin{aligned}
(3.4)\quad \left|\Delta^{k_1}_{h_1e_1}\cdots\Delta^{k_n}_{h_ne_n} f(x)\right| &\leqslant C_K \sum_{k_i\neq 0}\left(\frac{a_i}{h_i}\right)^r A(a_1,\dots,a_n)\,|f(x)|\\
&\quad + C_K\sum_{l=0}^{p-N} 2^{-lr}\sum_{S\in D(K)} A_s\left(M2^lh_1e_1,\dots,M2^lh_ne_n\right)\left|\prod_{i=1}^n \Delta^{s_i}_{2^lh_ie_i} f(x)\right|\\
&\quad + C_K\sum_{l=0}^{p-N} 2^{-lr}\sum_{\substack{i=1\\k_i\neq 0}}^{n} A_i\left(M2^lh_1e_1,\dots,M2^lh_ne_n\right)\left|\Delta^r_{2^lh_ie_i} f(x)\right|\\
&\equiv I(1)+I(2)+I(3),
\end{aligned}$$

where N is such that the points in $A_i(\ )\ |\prod_{i=1}^n \Delta^{s_i}_{2^lh_i} f(x)|$ satisfy $y \in x + (K+1)B/(n(r-2)+2)$, and $D(K)$ is all $s = (s_1,\dots,s_n)$ such that $s_i = 0$ if $k_i = 0$, $s_1 + \cdots + s_n = r + K$ and $s_i < r$. We will now show that after $K+1$ steps we have a similar estimate with $K+1$ replacing K. We have to estimate only $I(2)$. We estimate $|\prod_{i=1,\ k_i\neq 0}^n \Delta^{s_i}_{2^lh_i} f(x)|$ in very much the same way as (3.3) was estimated where s_i replaces k_i, $r+K$ replaces r, and 2^lh_i replaces h_i. Instead of the above expression at x we have really an average of the expressions of this type at finitely many points which because of previous choice belong to

$$x + (K+1)B/\left(n(r-2)+2\right).$$

We can now write

$$\begin{aligned}
\left|\prod_{\substack{i=1\\k_i\neq 0}}^{n} \Delta^{s_i}_{2^lh_ie_i} f(x)\right| &\leqslant C\sum_{m=0}^{p-l-N} 2^{-m(K+r)}\sum_{j=1}^{n}\frac{1}{2}s_j\sum_{\nu=0}^{s_j-1} w_\nu(j)\\
&\quad\times\left|\left(\prod_{i=1}^{j-1} 2^{-s_i}\Delta^{s_i}_{2^{l+m+1}h_ie_i}\right)\left(\prod_{i=j+1}^{n}\Delta^{s_i}_{2^{l+m}h_ie_i}\right)\left(\Delta^{s_j+1}_{2^{l+m}h_je_j} f(x) + \nu 2^{l+m}h_je_j\right)\right|\\
&\quad + 2^{-(p-N-l)(r+K)}\|f\|.
\end{aligned}$$

We sum the estimate above for terms of the type $A(\cdots)|\prod_{i=1,k_i\neq 0}^{n}\Delta_{2^l h_i}^{s_i} f(x)|$ for all l in (3.4) and one particular sequence $(s_1,\ldots,s_n)$ and have, using $K+r>r$, the choice of N and $t=m+l$,

$$\sum_{l=0}^{p-N} 2^{-l-r}A(M2^l h_1,\ldots,M2^l h_n)\left|\prod_{\substack{i=1\\k_i\neq 0}}^{n}\Delta_{2^l h_i}^{s_i} f(x)\right|$$

$$\leqslant \sum_{l=0}^{p-N} 2^{-lr}2^{-(p-N-l)(r+K)}A(a_1,\ldots,a_n)\,|f(x)|$$

$$+C\sum_{t=0}^{p-N}2^{-tr}\sum_{m=0}^{t}2^{-mK}\sum_{j=1}^{n}A_j(M_i 2^t h_1,\ldots,M_1 2^t h_n)$$

$$\times\left|\left(\prod_{i=1}^{j-1}2^{-s_i}\Delta_{2^{t+1}h_i e_i}^{s_i}\right)\left(\prod_{i=j+1}^{n}\Delta_{2^t h_i e_i}^{s_i}\right)\Delta_{2^t h_j e_j}^{s_j+1}f(x)\right|$$

$$\leqslant C_1\left(\frac{h_j}{a_j}\right)^r A(a_1,\ldots,a_n)\,|f(x)|$$

$$+C_1\sum_{t=0}^{p-N}2^{-tr}\sum_{j=1}^{N}A_j(M_1 2^t_{h_1},\ldots,M_1 2^t_{h_n})$$

$$\times\left|\left(\prod_{i=1}^{j-1}\Delta_{2^{t+1}h_i e_i}^{s_i}\right)\left(\prod_{i=j+1}^{n}\Delta_{2^t h_i e_i}^{s_i}\right)\Delta_{2^t h_j e_j}^{s_j+1}f(x)\right| = J(1)+J(2).$$

$J(2)$ contains elements of the type in $I(2)$ of (3.4) for $K+1$ (instead of K) but in the case $s_j+1=r$ for some j the term in question would belong to those of type in $I(3)$, we separate them accordingly and obtain (3.4) for $K+1$.

In the step $L=n(r-2)+1$ we do not have any term left representing the analog of $I(2)$ in (3.4), as $\Sigma s_i = r+n(r-2)+1 = n(r-1)+1$ and at least one of the s_i has to be r, and therefore the term would belong anyway to $I(3)$, and we complete our proof.

4. Directional differences in $C(R^n)$, $C(T^n)$ **and** $C(R_+^n)$. In this section we state and prove the result in $C(R^n)$, $C(T^n)$ and $C(R_+^n)$ as corollaries of Theorem 2.1.

THEOREM 4.1. *Let e be any direction in R^n or T^n and $e_1\cdots e_n$ any n independent vectors that belong to S_{n-1}. Then for $f\in L_\infty(R^n)$ or $f\in L_\infty(T^n)$, we have*

$$\omega_r(f,t,e)\leqslant Ct^r\|f\|+C\sum_{i=1}^{n}t^r\int_t^1\frac{\omega_r(f,u,e_i)}{u^{r+1}}\,du \tag{4.1}$$

and C does not depend on f.

We observe that the condition in the theorem is $f\in L_\infty(R^n)$ not $f\in C(R^n)$ but for $\omega_r(f;u,e_i)=o(1)$, $u\to 0+$, more than just $f\in C(R^n)$ is implied.

The situation on $L_\infty(R_+^n)$ is somewhat more complicated; we define $\omega_r(f,t,e)\equiv \operatorname{Sup}_{0<h\leqslant t}\operatorname{Sup}_{x,x+rhe\in R_+^n}|\Delta_{he}^r f(x)|$ and can state

THEOREM 4.2. *Let $e \in S_{n-1}$ and $e_1, \dots, e_n$ any n independent vectors that belong to $S_{n-1} \cap R^n_+$, then*

$$\omega_r(f, t, e) \leqslant C\left(t^r \|f\| + \sum_{i=1}^{n} t^r \int_t^1 \frac{\omega_r(f, h, e_i)}{u^{r+1}} du \right) \equiv I(t). \tag{4.2}$$

Moreover,

$$\sup_{0<h_i \leqslant t} \sup_x \left| \prod_{i=1}^{n} \Delta^{k_i}_{h_i e_i} f(x) \right| \leqslant I(t) \quad \textit{where } \sum k_i = r \tag{4.3}$$

and C does not depend on f or x (but does depend on r, n and for (4.2) *also on $e_1 \cdots e_n$).*

REMARK 4.3. We prove Theorems 4.1 and 4.2 directly rather than use the result for $C(D)$, because this simpler situation is easier to generalize to other spaces, and the result is on R^n or R^n_+ rather than on $D_1 \subset D$.

PROOF OF THEOREMS 4.1 AND 4.2. Conditions of our theorems would imply that, for all x in the domain, $x + \sum a_i e_i$ is also in the domain for $a_i \geqslant 0$ (the domain being R^n, T^n or R^n_+). We choose $a_i = 1$ and proceed recalling

$$A(M2^l h_1, \dots, M2^l h_n) \,|\, \Delta^r_{2^l h_i e_i} f(x) \,| \leqslant \omega_r(f, 2^l h_i, e_i)$$

and using monotonicity of $\omega_r(f, u, e_i)$ and of u to deduce

$$\begin{aligned}\sum_{l=0}^{p-N} 2^{-lr} A_i(M2^l h_1, \dots, M2^l h_n) \left| \Delta^r_{2^l h_i e_i} f(x) \right| &\leqslant \sum_{l=0}^{p-N} 2^{-lr} \omega_r(f, 2^l h_i, e_i) \\ &= h_i^r \sum_{l=0}^{p-N} \frac{\omega_r(f, 2^l h_i, e_i)}{(2^l h_i)^{r+1}} 2^l h_i = h_i^r 2^{r+1} \int_{h_i}^1 \frac{\omega_r(f, u, e_i)}{u^{r+1}} du.\end{aligned}$$

Therefore (2.6) implies (4.3) for $L_\infty(R^n)$, $L_\infty(T^n)$ and $L_\infty(R^n_+)$. To prove (4.1) we write $e = \sum \alpha_i e_i$ and recall $E(y)f(x) = f(x + y)$.

$$\begin{aligned}\Delta^r_{he} f(x) &= \left(E(h(\alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n)) - I\right)^r f(x) \\ &= \{(E(h\alpha_1 e_1) - I) + E(h\alpha_1 e_1)(E(h\alpha_2 e_2) - I) \\ &\qquad + \cdots + E(h(\alpha_1 e_1 + \cdots + \alpha_{n-1} e_{n-1}))(E(h\alpha_n e_n) - I)\}^r f(x)\end{aligned}$$

which is a finite sum of elements of the type $E(y)\prod_i \Delta^{k_i}_{\alpha_i e_i}$ where $\sum k_i = r$ and $y - x = \sum \beta_i e_i$ with β_i between zero and $\alpha_i r h_i$, which concludes the proof of Theorem 4.1.

In the proof of Theorem 4.2 we have already shown that (4.3) follows Theorem 2.1 and the argument above, but for the estimate (4.2) we have to be more careful as the above procedure may take us out of R^n_+. (That could not happen if R^n_+ is formed by $\sum \alpha_i e_i$ for $\alpha_i \geqslant 2$.) Let us, therefore, rearrange e_i so that in $\sum \alpha_i e_i = e$, $\alpha_i \geqslant 0$ for $i \leqslant j$, and $\alpha_i < 0$ for $i > j$. Then $y = x + kh\sum_{i=1}^{j} \alpha_i e_i + \sum_{i=j+1}^{n} h l_i \alpha_i e_i$ for $l_i \leqslant k$ is in R^n_+ since $x + khe \in R^n_+$ and

$$y = x + khe - \sum_{i=j+1}^{n} (k - l_i)\alpha_i h e_i = x + khe + \sum_{i=j+1}^{n} \beta_i h e_i, \qquad \beta_i > 0.$$

We now set

$$(E(he)-I)^r=\left(\left[E(he)-E(he_*)\right]+\left[E(he_*)-I\right]\right)^r$$

where $e_*=\sum_{i=1}^{j}\alpha_i e_i$ and

$$(E(he)-I)^r=\sum_{k=0}^{r}\binom{r}{k}\left[E(he_*)-I\right]^{k-r}\left[E\left(h\sum_{i=j+1}^{n}\alpha_i e_i\right)-I\right]^k E(khe_*).$$

On each of the terms $(E(he_*)-I)^{k-r}$ and $(E(h\sum_{i=j+1}^{n}\alpha_i e_i)-I)^k E(khe_*)$, we follow the earlier procedure, but since x and $x+khe$ belong to R_+^n so does $x+khe^*+h\sum_{i=j+1}^{n}\alpha_i l_i e_i$ for $l_i\leq k$.

REMARK 4.4. One could have used integrals in a similar way to that of [**3** or **4**], as will be explained in §7, to overcome the above minor combinatorial difficulty. This would be more standard but somewhat more complicated. If e_i are those forming R_+^n, that difficulty would not have arisen. In §7 we use the present technique where we cannot use Stekelov-type integrals and vice versa.

5. Moduli of continuity in other spaces and for semigroups. The theorem of the last section can be extended to multivariate semigroups of contractions on a Banach space. $T(t)$ is a semigroup of contractions on a Banach space B if $T(t)f$ and f belong to B for $t\in R_+^n$, $t\in R^n$ or $t\in T^n$, $T(t_1+t_2)f=T(t_1)T(t_2)f$, and $\|T(t)f\|\leq\|f\|$. (From the last mentioned it follows that if $t\in R^n$ or $t\in T^n$, we have a group of isometries.)

We define the r directional modulus of continuity for $e\in R_+^n\cap S_{n-1}$ (or in case of a group of isometries no restriction on e) by

$$\omega_r(T(\cdot)f,u,e,B)=\operatorname*{Sup}_{t}\operatorname*{Sup}_{0\leq v\leq u}\left\|(T(ve)-I)^rT(t)f\right\|_B. \tag{5.1}$$

For t and $t+rhe$ in R_+^n, and $(T(he)-I)^rT(t)$ defined by

$$(T(he)-I)^rT(t)\equiv\sum_{k=0}^{r}(-1)^{r-k}\binom{r}{k}T(t+khe), \tag{5.2}$$

the definition of a semigroup implies the right-hand side exists. (We could have defined it using an argument similar to that used in proving (4.2). This would have been equivalent but (5.2) is somewhat easier to handle.)

We can now state our theorem.

THEOREM 5.1. *Let $T(t)$ be a multivariate semigroup of contractions on a Banach space B, $t\in R_+^n$, $t\in R^n$ or $t\in T^n$, then for $f\in B$ and both t_0 and t_0+rhe belonging to R_+^n (for a group on R^n or T^n the last condition is dropped), we have*

$$\left\|(T(he)-I)^rT(t_0)f\right\|_B\leq C\left\{h^r\|f\|+h^r\sum_{i=1}^{n}\int_h^1\frac{\omega_r(Tf,u,e_i,B)}{u^{r+1}}du\right\} \tag{5.3}$$

where $(T(he)-I)^rT(t_0)$ is defined by (5.2), *$e_1\cdots e_n$ are independent vectors in $S_{n-1}\cap R_+^n$, or in S_{n-1} in case of a group on R^n or T^n, and C depends only on r, n and e_i. (If $e=\sum_{i=1}^{n}\alpha_i e_i$ and $M=\max_i|\alpha_i|$, C depends only on r, n and M.)*

For a Banach space B of functions, measures or generalized functions on R_+^n for which translation is a contraction, we can write

$$(5.4)\qquad \omega_r(f, u, e, B) = \operatorname*{Sup}_{0\leqslant h\leqslant u} \operatorname*{Sup}_{\substack{x_0\\ x_0,\, x_0+rhe\in R_+^n}} \|\Delta_{he}^r f(\cdot + x_0)\|_B,$$

and in case R^n or T^n replace R_+^n we have,

$$(5.5)\qquad \omega_r(f, u, e, B) = \operatorname*{Sup}_{0\leqslant h\leqslant u} \|\Delta_{he}^r f(\cdot)\|_B.$$

The above formulae will yield the result moduli of continuity for $L_p(R^n)$ or $L_p(R_+^n)$, $1 \leqslant p \leqslant \infty$.

We have the following corollary of Theorem 5.1.

COROLLARY 5.2. *Let B be a Banach space of functions, measures or generalized functions on R_+^n (or R^n or T^n) for which translation is a contraction (or an isometry). Then for n independent, $e_i \in S_{n-1} \cap R_+^n$ ($e_i \in S_{n-1}$) and a unit vector e, we have*

$$(5.6)\qquad \omega_r(f, h, e, B) \leqslant C\left\{h^r\|f\| + h^r \sum_{i=1}^n \int_h^1 \frac{\omega_r(f, u, e_i, B)}{u^{r+1}}\,du\right\}$$

where C depends only on r, n and $e_1 \cdots e_n$.

PROOF OF THEOREM 5.1. We choose $g \in B^*$, $\|g\|_{B^*} = 1$, and obtain $\langle T(t)f, g\rangle = F(t)$, and $F(t)$ is a function on R_+^n (or R^n or T^n) which is bounded by $\|f\|_B \cdot \|g\|_{B^*} = \|f\|_B$. Therefore, $F(t)$ satisfies the conditions of Theorems 4.1 and 4.2,

$$\begin{aligned}\omega_r(F(t), h, e) &\leqslant Ch^r\|F\| + C\sum_{i=1}^n h^r \int_h^1 \frac{\omega_r(F, u, e_i)}{u^{r+1}}\,du\\ &\leqslant Ch^r\|f\| + C\sum_{i=1}^n h^r \int_0^1 \frac{\omega_r(T(t)f, u, e_i, B)}{u^{r+1}}\,du\end{aligned}$$

since

$$\omega_r(F, u, e_i) = \operatorname*{Sup}_{\substack{\eta\leqslant u\\ t}} |\Delta_{he_i}^r F(t)| \leqslant \operatorname*{Sup}_{\substack{\eta\leqslant u\\ t}} \|(T(\eta e_i) - I)^r T(t) f\|_B.$$

We choose t_0 and $\eta \leqslant h$ such that $t_0, t_0 + \eta re \in R_+^n$ and $\|\Delta_{\eta e}^r T(t_0) f\| \geqslant \omega_r(T(t)f, h, e) - \varepsilon$. We choose $g \in B^*$, $\|g\| = 1$ such that

$$|\Delta_{\eta e}^r F(t_0)| = |\langle \Delta_{\eta e}^r T(t_0)f, g\rangle| \geqslant \|\Delta_{\eta e}^r T(t_0) f\|_B - \varepsilon \geqslant \omega_r(T(t)f, h, e) - 2\varepsilon$$

but $|\Delta_{\eta e}^r F(t_0)| \leqslant \omega_r(F(t), h, e)$ and ε can be arbitrarily small, which completes the proof.

REMARK 5.3. (a) In the case of function spaces we have $T(t)f(x) = f(x + t)$ and $\langle T(t)f, g\rangle = \int f(x + t)g(x)\,dx$ as was done in **[1]**.

(b) We did not require $T(t)$ to be strongly continuous.

6. The case $\alpha = r$. In an earlier section it was proved that in R^n (and, in fact, for other domains), $\omega_r(f, h, e_i) \leqslant Mh^\alpha$ for n independent e_i, and $\alpha < r$ implies that $\omega_r(f, h, e) \leqslant Kh^\alpha$ for any direction e. In the case $\alpha = r$ the result proved implied

$\omega_r(f, h, e) \leqslant Kh^r \log(1/h)$. In the following example it will be shown that this estimate is sharp.

EXAMPLE 6.1. Let $\psi(x, y) \in C_0^\infty$; that is, ψ has compact support and all derivatives, and let $\psi(x, y) = 1$ for $x^2 + y^2 \leqslant 1$. We define

$$f(x, y) = \psi(x, y)xy\log(x^2 + y^2) \quad \text{and} \quad f(0,0) = 0.$$

It is not hard to see that $|\Delta_{he_i}^2 f(x, y)| \leqslant Mh^2$ for $e_1 = (1,0)$ and $e_2 = (0,1)$, while $\sup|\Delta_{he}^2 f(x, y)| \sim Kh^2 \log(1/h)$ in other directions. Also the mixed difference $\sup|\Delta_{he_i}\Delta_{he_2} f(x, y)| \sim Kh^2 \log(1/h)$.

EXAMPLE 6.2. For R^n and the standard basis e_i and a given r, we have

$$f(x_1,\dots,x_n) = \psi(x_1,\dots,x_n)\left(\prod_{i=1}^n x_i^{k_i}\right)\log\left(\sum_{\substack{i=1\\ k_i\neq 0}}^n x_i^2\right)^{-1}$$

and $f(0,0,\dots,0) = 0$ where $\psi(x_1,\dots,x_n) \in C_0^\infty$ and $\psi(x_1,\dots,x_n) = 1$ for $|(x_1,\dots,x_n)| \leqslant 1$ and $\sum k_i = r$. If at least two k_i are different from 0, then all $\omega_r(f, h, e_i) = O(h^r)$ in the $C(R^n)$ norm but for the vector $e = (1/\sqrt{n})(1,1,\dots,1)$, for example, $\omega_r(f, h, e) = O(h^r \log(1/h))$ in $C(R^n)$.

The above situation does not hold in L_p, $1 < p < \infty$. In fact, one can deduce from known results the following

THEOREM 6.3. *If f has compact support, $f \in L_p(R^n)$ for $1 < p < \infty$ and $\omega_r(f, h, e_i, L_p) = O(h^r)$ for n independent e_i, then $\omega_r(f, h, e, L_p) = O(h^r)$ for any direction e. (Also all mixed differences of order r can be estimated by $O(h^r)$.)*

PROOF. For $f \in L_p(R^n)$, $1 < p < \infty$, and $\omega_r(f, h, e_i, L_p) = O(h^r)$, we actually have f, $\partial f/\partial x_i, \dots, (\partial/\partial x_i)^r f$ in L_p and $\int (\partial/\partial x_i)^j f\, dx_i = (\partial/\partial x_i)^{j-1} f$ for $j \leqslant r$ (where the derivatives are the strong L_p derivatives) and $\|(\partial/\partial x_i)^r f\|_p \leqslant \sup_h h^{-r}\omega_r(f, h, e_i, L_p)$. This is a classical result in essence. However, appropriate minor modification of the results of Hardy-Littlewood and of Berens-Butzer to adjust to R^n and the present situation are much more cumbersome than the following proof.

Our condition that $h^{-r}\Delta_{he_i}^r f$ is in a ball in L_p, and therefore so is $h^{-j}\Delta_{he_i}^j f$ for $j \leqslant r$ which implies $h^{-j}\Delta_{he_i}^j f$ for $j \leqslant r$, has a weak* accumulation point which we call ϕ_j. For $1 < p < \infty$ the weak* closure is in L_p and $\|\phi_j\| \leqslant \operatorname{Sup}\|h^{-j}\Delta_{he_i}^j f\|_{L_p}$. For $\psi \in C_{00}^\infty$ (C^∞ functions with compact support)

$$\left\langle h^{-j}\Delta_{he_i}^j f, \psi \right\rangle = \left\langle f, h^{-j}\Delta_{-he_i}^j \psi \right\rangle \to (-1)^j \left\langle f, \left(\frac{\partial}{\partial x_i}\right)^j \psi \right\rangle = \left\langle \phi_j, \psi \right\rangle$$

or, in other words, since C_{00}^∞ is dense in L_q, $q^{-1} + p^{-1} = 1$, $\phi_j = (\partial/\partial x_i)^j f$ where the derivative is taken in the distributional sense (and therefore ϕ_j is unique as an accumulation point). Define $(g(x))_h = h^{-1}\int_0^h g(x+ue_i)\,du$ and recall that for $g \in L_p$, $\lim_{h\to 0+} g_h(x) = g(x)$ in L_p and a.e. For $1 \leqslant j \leqslant r$ we have

$$\left\langle (\phi_j)_h, \psi \right\rangle = \left\langle \phi_j, (\psi)_{-h} \right\rangle = -\left\langle \phi_{j-1}, \left(\frac{\partial}{\partial x_i}\right)(\psi)_{-h} \right\rangle = -\left\langle \phi_{j-1}, \left(\frac{\partial}{\partial x_i}\psi\right)_{-h} \right\rangle$$
$$= -\left\langle (\phi_{j-1})_h, \frac{\partial}{\partial x_i}\psi \right\rangle = \left\langle \frac{\partial}{\partial x_i}(\phi_{j-1})_h, \psi \right\rangle = \left\langle \frac{1}{h}\Delta_{he_i}\phi_{j-1}, \psi \right\rangle$$

or $(\phi_j)_h = h^{-1}\Delta_{he_i}\phi_{j-1}$ but $(\phi_j)_h \to \phi_j$ in L_p and therefore so does $h^{-1}\Delta_{he_i}\phi_{j-1}$.

Using a result by Ilin [**2**, p. 301], we have $\|D^\nu f\|_p \leqslant C\{\sum_{i=1}^n \|(\partial/\partial x_i)^r f\|_p + \|f\|\}$ where $\nu = (\nu,\ldots,\nu_n)$ and $\nu_1 + \cdots + \nu_r \leqslant r$. This would yield an estimate of all mixed derivatives and therefore a derivative in any direction e satisfies $\|(\partial/\partial\xi)^r f\|_p \leqslant K\{\sum_{i=1}^n \|(\partial/\partial x_i)^r f\|_p + \|f\|_p\}$ and hence $\omega_r(f, h, e, L_p) = O(h^r)$. One should note that when r is even a more accessible source for the above is Stein's text [**5**, p. 114] as $\Sigma(\partial/\partial x_i)^r = P(D)$ is elliptic and $\|P(D)f\|_p \leqslant \sum_{i=1}^n \|(\partial/\partial x_i)^r f\|_p$.

REMARK 6.4. It is not known to me whether for $L_1(R^n)$ we have the sharp estimate by $O(h^r \log(1/h))$. (We cannot hope for an analog to Theorem 6.3.)

7. The result for $C(D)$, $L_p(D)$ and other spaces. We define for a domain D, $D \subset R^n$, the constant a, and n independent vectors $e_1,\ldots,e_n$, the domain $D_a(e_1,\ldots,e_n)$ by

$$\begin{aligned} &D_a \equiv D_a(e_1,\ldots,e_n) = \{x;\, x + B_a(i_1,\ldots,i_n) \subset D \text{ for some vector of} \\ (7.1)\quad &\text{integers } (i_1,\ldots,i_n) \text{ where} \\ &\qquad B_a(i_1,\ldots,i_n) = [0,(-1)^{i_1}ae_1] \times \cdots \times [0,(-1)^{i_n}ae_n]. \end{aligned}$$

We always have $D_a \subset D$. If D is a box with sides parallel to e_i, and a small enough then, $D_a = D$ and the same is true if $D = \{\Sigma\,\xi_i e_i,\ \xi_i > 0\}$ but in most cases D_a is strictly smaller than D.

We will observe a few properties common to many Banach spaces that will be needed in the theorem below.

DEFINITION 7.1. A Banach space will satisfy condition A if

(a) $f(x) \in B$ implies $f(x+h) \in B$, $x, h \in R^n$,

(b) $f \in B$ and E measurable implies $\chi(E)f(x) \in B$,

(c) $|f| \leqslant |g|$ (and $f, g \in B$) implies $\|f\| \leqslant \|g\|$.

DEFINITION 7.2. A Banach space will satisfy condition B if it satisfies condition A, its elements are locally Lebesgue integrable and $\|f(\cdot+h) - f(\cdot)\|_B = o(1)$, $|h| \to 0$, $h \in R^n$. (And therefore $\|(1/h)\int_0^h f(\cdot+ue)\,du - f(\cdot)\|_B = o(1)$ as $h \to 0$, $h \in R$.)

It is clear that $C(R^n)$ does not satisfy condition A but $L_p(R^n)$, $1 \leqslant p \leqslant \infty$, does and so do many other spaces. The following results will apply to $C(R^n)$ or $C(D)$ through $L_\infty(R^n)$ or $L_\infty(D)$. ($L_\infty(R^n)$ does not satisfy condition B while L_p for $1 \leqslant p < \infty$ does.) We also define

$$(7.2)\qquad D(he) = \bigcap_{0\leqslant\eta\leqslant h} \{-\eta e + D\}.$$

DEFINITION 7.3. The r directional modulus of continuity is given by

$$\omega_r(f, h, e, B(D)) = \sup_{\eta \leqslant h} \|\chi(D(r\eta e))\Delta^r_{\eta e} f\|_B. \tag{7.3}$$

We are now able to state and prove our theorem.

THEOREM 7.1. *For a Banach space B satisfying condition* A, *n independent unit vectors $e_1, \dots, e_n$, and integers k_i satisfying $k_1 + \cdots + k_n = r$, we have*

(7.4)

$$\left\|\chi(D_a)\Delta^{k_1}_{he_1} \cdots \Delta^{k_n}_{he_n} f\right\|_B \leqslant C\left(\|\chi(D)f\|_B h^r + \sum_{i=1}^n h^r \int_h^a \frac{\omega_r(f, u, e_i, B(D))}{u^{r+1}} du\right)$$

where C depends only on r, n and a (and not on Bh, f, D and e_i). For $e = \sum \alpha_i e_i$, $|\alpha_i| \leqslant M$, $|e| = 1$ we have

$$\|\chi(D^*_{rMh})\Delta^r_{he} f\|_B \leqslant C\left(\|\chi(D)f\|_B h^r + \sum_{i=1}^n h^r \int_h^a \frac{\omega_r(f, u, e_i, B(D))}{u^{r+1}} du\right) \tag{7.5}$$

where $D^*_{rMh} = \{x;\ x + B^*_{rMh} \subset D_a\}$ where

$$B^*_{rMh} \equiv \left\{y;\ y = \sum \beta_i e_i \text{ and } |\beta_i| \leqslant rMh\right\}$$

and C depends only on M, a and r.

If condition B is satisfied or our Banach space is $C(D)$ and the interval $[x, x + rhe] \subset D_a$, then

$$\|\chi(D_a)\Delta^r_{he} f\|_B \leqslant C\left(\|\chi(D)f\|_B h^r + \sum_{i=1}^n h^r \int_h^a \frac{\omega_r(f, h, e_i, B(D))}{u^{r+1}} du\right), \tag{7.6}$$

where C depends on a, r, and $e_1 \cdots e_n$ only.

PROOF. Using (2.6) for $x \in D_a$ and recalling that for such x,

$$\left\|\chi(D_a)\left|\Delta^r_{ue_i} f(\cdot + 2^l h_1 e_1 + \cdots + 2^l h_n e_n)\right|\right\| \leqslant \omega_r(f, u, e_i B(D))$$

and, therefore, simply using the lattice compatibility and the triangle inequality, we obtain (7.4). Actually Theorem 3.1 is proved in a way that is amenable to proving (7.4). To prove (7.5) we recall that $x \in D^*_{rMh}$ guarantees that the process of writing the directional difference in terms of mixed differences following Theorem 4.1 does not take x out of D_a. In fact, in view of Theorem 4.2, we can use a somewhat less restrictive domain than D^*_{rMh}. However, for most spaces in question that would not matter as one can prove the more general (7.6) if our Banach space satisfies condition B as L_p, $1 \leqslant p < \infty$, do. If sets of the type $\{x, x + B_a(i_1, \dots, i_n) \subset D\}$ cover D_a, so do $\{x, x + B_{a/2}(i_1, \dots, i_n) \subset D\}$ and in those the isolated points of D_a are not isolated. We treat the domains $\{x;\ x + B_{a/2}(i_1, \dots, i_n) \subset D\} \cap D_a$ one at a time which we may because of lattice compatibility and which we, for convenience, call $E = E_1 \cap D_a$. But in each we may define

$$F_h(x) = \frac{(-r)^{rn}}{h^{rn}} \int_0^{h/r} \cdots \int_0^{h/r} \Big\{ (-1)^{rn} \Delta^r_{(u_1 + \cdots + u_r)e_1} \Delta^r_{(u_{r+1} + \cdots + u_{2r})e_2} \cdots \Delta^r_{(u_{rn-n+1} + \cdots + u_{rn})e_n} - 1 \Big\} \cdot f(x)\, du_1 \cdots du_{rn}.$$

Obviously, $\|\chi(E)F_h(x) - f(x)\|$ can be estimated by $w_r(f, h, e, B(D))$. Moreover, F_h being just a Stekelov-type average, its mixed derivative of order r will be h^{-r} times a finite combination of mixed differences of order r that were already estimated in (7.4). (This is the step in which we use condition B or our Banach spaces is $C(D)$.) This yields an estimate for the r derivative in the e direction of F_h and we complete our theorem. We have to observe that we actually use the earlier part of the theorem with $a/2$ instead of a and that at least one E_1 contains x and part of $[x, x + rhe]$ i.e., $[x, x + \delta e]$.

Sometimes for a particular domain it is useful to use a finite but bigger collection of vectors e_i. For instance, if we discuss the simplex $(0,0)$, $(0,1)$ and $(1,0)$ for D, it is useful to have the vectors $(0,1)$, $(1,0)$ and $(1/\sqrt{2})(1,1)$, and D is the union of D_a generated by two of the above 3 vectors (3 different D_a with small enough a, say $a = \frac{1}{4}$). This and similar situations are important for actual approximation problems while not crucial for the present question. Of particular use will be a situation in which D is covered by finitely many D_a which are generated by subsets of $e_1 \cdots e_k$; but while for a very general D (see Sharpley [**4**]) there exists such a collection $e_1 \cdots e_k$, our problem here is, however: given the fixed collection, on what part of D is the result valid? If the result is valid on $D_1 \cdots D_l$, then it is valid on their union.

REMARK. We can choose a to be small and D_a will be close to D. Moreover, when $\omega_r(f, h, e_i, B(D)) \leqslant Mh^\alpha$, $\alpha < r$, we can choose $a \leqslant Mh^{1-\alpha/r}$ and the constants will depend on M rather than a.

REFERENCES

1. Z. Ditzian, *Some remarks on approximation theorems on various Banach spaces*, J. Math. Anal. Appl. **77** (1980), 567–576.

2. V. P. Ilin, *Conditions of validity of inequalities between L_p norms of partial derivatives of functions of several variables*, Proc. Steklov Inst. Math. **96** (1968), 259–305.

3. H. Johnen and K. Scherer, *On equivalence of K functional and moduli of continuity and some applications*, Constructive Theory of Functions of Several Variables (Proc. Conf. Oberwolfbach, 1976), Lecture Notes in Math., vol. 571, Springer-Verlag, Berlin and New York, 1976, pp. 119–140.

4. R. C. Sharpley, *Cone conditions and the modulus of continuity* (Proc. 2nd Edmonton Conf. on Approximation Theory), CMS Conference Proceedings, vol. 3, Amer. Math. Soc., Providence, R. I., 1983, pp. 341–351.

5. E. Stein, *Singular integrals and differentiability properties of functions*, Princeton Univ. Press, Princeton, N. J., 1970.

6. A. F. Timan, *Theory of approximation of functions of a real variable* (English transl.), Pergamon Press, Macmillan, New York, 1963.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ALBERTA, EDMONTON, ALBERTA, CANADA T6G 2G1