AD-A022 744

DERIVATIVES, DIFFERENCES, MULTIPLE FOURIER KERNELS

D. B. Liu, et al

Wisconsin University

Prepared for:

Army Research Office

January 1976

DISTRIBUTED BY:

NTIS
National Technical Information Service
U. S. DEPARTMENT OF COMMERCE

100103

MRC Technical Summary Report #1578

DERIVATIVES, DIFFERENCES, MULTIPLE FOURIER KERNELS

D. B. Liu and L. C. Young

UNIVERSITY OF WISCONSIN

MATHEMATICS RESEARCH CENTER

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

January 1976

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

and

National Science Foundation
Washington, D. C. 20550

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

DERIVATIVES, DIFFERENCES, MULTIPLE FOURIER KERNELS

D. B. Liu and L. C. Young

Technical Summary Report #1578
January 1976

ABSTRACT

Identities and inequalities for Fourier kernels and for difference operators are related to a geometric series identity. The resulting machinery is applied to obtain, in the approximation theory for ordinary or partial derivatives of any order, necessary and sufficient conditions in place of classical sufficient conditions. Alternative formulations are given in terms of Tauberian Theorems, and in terms of Schwartz distributions.

The results are achieved by making use, as in L. C. Young's papers on Stochastic integrals and the like, of pairs of estimate functions in place of the classical higher moduli of continuity.

AMS(MOS) classification: 40D10, 42A04, 46F99, 60H05

Key words: Differences; multiple fourier kernels; Littlewood-Paley lemma; estimate functions; higher derivatives; Tauberian theorems; higher mixed partial derivatives; necessary and sufficient conditions; best possible approximation theory; Schwartz distributions; Stieltjes and stochastic integrals.

Sponsored by the United States Army under Contract No DAAG29-75-C-0024 and the National Science Foundation under Contract No. GP-27211.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 1578	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) DERIVATIVES, DIFFERENCES, MULTIPLE FOURIER KERNELS		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) D. B. Liu and L. C. Young		8. CONTRACT OR GRANT NUMBER(s) DAAG29-75-C-0024
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS See item 18 below		12. REPORT DATE January 1976
		13. NUMBER OF PAGES 36
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, D. C. 20550

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Differences; multiple fourier kernels; Littlewood-Paley lemma; estimate functions; higher derivatives; Tauberian theorems; higher mixed partial derivatives; necessary and sufficient conditions.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Identities and inequalities for Fourier kernels and for difference operators are related to a geometric series identity. The resulting machinery is applied to obtain, in the approximation theory for ordinary or partial derivatives of any order, necessary and sufficient conditions in place of classical sufficient conditions. Alternative formulations are given in terms of Tauberian theorems, and in terms of Schwartz distributions. The results are achieved by making use, as in L.C. Young's papers on Stochastic integrals and the like, of pairs of estimate functions in place of the classical higher moduli of continuity.

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DERIVATIVES, DIFFERENCES, MULTIPLE FOURIER KERNELS

D. B. Liu and L. C. Young

§1. Introduction. We shall be concerned, in this note, with some elementary properties of multiple Fourier kernels, and with their use in the approximation theory for derivatives of all orders, for functions of one or more variables. Some basic devices go back as far as Riemann's theory of trigonometric series and H. A. Schwarz's generalized second derivative. We take the trouble, in this connection, to establish for the first time a little lemma, conjectured by Littlewood and Paley. At various places we indicate the bearing of our methods and results on Tauberian theorems, Laurent-Schwartz distributions, and classical approximation theory. Generally speaking, our results differ from the classical ones by being necessary and sufficient conditions, rather than merely necessary, or merely sufficient. This is due to our systematic use of a pair of estimate functions, instead of a single modulus of continuity of the appropriate order. Our work is related to, and partly extends, recent work in which such pairs of estimate functions were used to study Stieltjes and Stochastic integrals (Main Theorems, Young 25, 26; "best possible" character, Young 27; see also, for instance, Young 24, Lesniewski and Orlicz 13). We provide, in this connection, an alternative approach to the Stochastic integrals, there introduced.

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024 and the National Science Foundation under Contract No. GP-27211.

§2. The elementary geometric kernel and a Littlewood-Paley lemma.

We shall require some simple identities concerning special cases of the polynomial

$$(2.1) \qquad P(z) = \sum c_\nu z^\nu = \prod \frac{z^N - a^N}{z - a} = \prod_{r=1}^{q} \frac{z^{N_r} - a_r^{N_r}}{z - a_r} .$$

Here q and the N_r denote positive integers, while z and the a_r are complex numbers $\neq 0$.

The properties of (2.1) and its special cases are the basis of methods used in various topics of analysis, such as Fourier series, summability, finite differences. We term (2.1) a geometric kernel. Equivalently we may consider a general Fourier kernel

$$(2.2) \qquad \prod_{r=1}^{q} \frac{\zeta_r^{N_r} - \zeta_r^{-N_r}}{\zeta_r - \zeta_r^{-1}} , \quad \zeta_r = z/a_r .$$

In the Fourier context it seems very clear that a general study of the polynomial (2.1), or of the kernel (2.2), is a necessary preliminary for whoever desires to see how far one can get by elementary methods. The a_r must be allowed to vary independently, at least on the unit circumference of the complex plane. This is the basic idea of the approach initiated by Littlewood and Paley (11).

The obvious starting point for such a study should be a good estimate of the coefficients c_ν . We therefore take this opportunity to establish a lemma conjectured by Littlewood and Paley (11, footnote p. 117). Our proof eliminates a weaker result, and therefore some minor Littlewood-Paley inaccuracies (e.g. loc cit, top of P. 118, where a statement in italics contradicts the Kronecker-Weyl theorem).

(2.3) Littlewood-Paley Lemma. In (2.1) suppose each $|a_r| = 1$. Then

(2.4) $$\operatorname*{Max}_{\nu} |c_\nu| \le A(q) \sum b_r$$

where $A(q)$ is a constant depending only on q, and where the b_r are defined by setting

$$N_{rs} = \operatorname{Min}(N_r, N_s), \quad a_{rs} = N_{rs}^{-1} + |a_r - a_s|, \quad b_r = \prod_{\substack{s \\ s \ne r}} a_{rs}^{-1} .$$

Proof of (2.4)†. By re-ordering we arrange that

$$a_{12} = \alpha = \operatorname*{Max}_{\substack{r,s \\ r \ne s}} a_{rs} \quad , \quad N_{12} = N_1 \le N_2 .$$

We denote by $P_1(z)$, $P_2(z)$ the functions derived from (2.1) by omitting from its right-hand side the factors corresponding respectively to $r=2$ and $r=1$. Further, in the expansions of the form $\sum c_\nu z^\nu$ for $P(z)$, $P_1(2)$, $P_2(z)$, we denote by C, C_1, C_2, respectively, the corresponding greatest $|c_\nu|$. By multiplying the expansions of

$$P_2(z) \quad \text{and} \quad \frac{z^{N_1} - a_1^{N_1}}{z - a_1}$$

we see that $N_1^{-1} C \le C_2$, while the identity

$$(a_1 - a_2) P(z) = (z^{N_2} - a_2^{N_2}) P_1(z) - (z^{N_1} - a_1^{N_1}) P_2(z)$$

evidently implies $|a_1 - a_2| C \le 2 C_1 + 2 C_2$. It follows by addition that $\alpha C \le 2C_1 + 3C_2$, and therefore, by an induction in q, that (2.4) holds with $A(q) = 5^{q-1}$.

---

† The estimate obtained by replacing a_{rs} by N_{rs}^{-1} is trivial. So is, by the partial fraction expansion of (2.1), the estimate in which a_{rs} is replaced by $|a_r - a_s|$. To the practical analyst, the existence of two estimates, $X \le A$, $X \le B$, automatically suggests $X \le C$, for some nice symmetrical expression $C \le \operatorname{Min}(A, B)$. Here (2.4) does amount to such an estimate $X \le C$.

§3. Basic identities. For the kernels (2.1), (2.2), inequalities such as (2.4) would become essential here if we proposed to employ a Fourier series approach to the topics we shall consider. We prefer a direct approach.

We write ζ for the vector $(\zeta_1,\ldots,\zeta_q)$, where again $\zeta_r = z/a_r$. Similarly, a symbol such as n, which occurs as an exponent, may have components n_r $(r = 1,\ldots,q)$, and in that case ζ^n denotes the product of the n_r-th powers of the ζ_r. The following variants of Lemma A(4.1) of (26) are easy consequences of the geometric series identity.

(3.1) *Let $\mathfrak{L}$, $\mathfrak{L}^*$ denote, respectively, the sets of convex combinations*

$$\sum_{i,j} \lambda_{ij}\, Q_i\, P_j \ , \qquad \sum_{i,j} \lambda^*_{ij}\, Q^*_i\, P^*_j \ ,$$

of products $Q_i P_j$, $Q^*_i P^*_j$, where each P_j has the form ζ^n, each P^*_j the form $\frac{1}{2}(\zeta^n + \zeta^{-n})$, each Q_i the form $\zeta_r^n - 1$, and each Q^*_i the form $\frac{1}{2}(\zeta_r^n + \zeta_r^{-n} - 2)$, and where, moreover, the exponents n in each Q_i, Q^*_i and the components of the exponents n in each P_j, P^*_j are positive integers, not exceeding the corresponding integers N_r. Further let

$$F(z) = \prod_{r=1}^{q} F_r/N_r \ , \qquad F^*(z) = \prod_{r=1}^{q} F^*_r/N_r \ ,$$

$$F_r = \frac{\zeta_r^{N_r} - 1}{\zeta_r - 1} \ , \qquad F^*_r = \frac{\zeta_r^{N_r} - \zeta_r^{-N_r}}{\zeta_r - \zeta_r^{-1}} \ .$$

Then $q^{-1}(F-1) \in \mathfrak{L}$, $q^{-1}(F^* - 1) \in \mathfrak{L}^*$.

(3.2) *Let L, L^* denote, respectively, the sets of convex combinations*

$$\sum \lambda_i\, Q_i \ , \qquad \sum \lambda_i\, Q^*_i \ ,$$

#1578

of expressions Q_i, Q_i^* of the form $z^n - 1$, $\frac{1}{2}(z^n + z^{-n} - 2)$, where the exponents n are positive integers not exceeding qN. Further let $a_r = 1$, $N_r = N$ $(r = 1, \ldots, q)$, so that now

$$F(z) = \left(\frac{z^N - 1}{z - 1}\right)^q / N^q, \qquad F^*(z) = \left(\frac{z^N - z^{-N}}{z - z^{-1}}\right)^q / N^q .$$

Then $F - 1 \in L$, $F^* - 1 \in L^*$.

We shall find it more convenient to multiply up by the denominators $\zeta_r - 1$, $\zeta_r - \zeta_r^{-1}$, etc. The above assertions then take the following form.

(3.3) Let $\mathcal{X}$, $\mathcal{X}^*$ be the sets obtained from $\mathcal{L}$, $\mathcal{L}^*$ by multiplying the elements respectively by $\prod(\zeta_r - 1)$, $\prod(\zeta_r - \zeta_r^{-1})$. Similarly let K, K^* be the sets obtained from L, L^* by multiplying the elements by $(z-1)^q$, $(z - z^{-1})^q$.

Then

(3.4) $$\frac{1}{q}\prod_{r=1}^{q}(\zeta_r^{N_r} - 1)/N_r - \frac{1}{q}\prod_{r=1}^{q}(\zeta_r - 1) \in \mathcal{X},$$

(3.5) $$\frac{1}{q}\prod_{r=1}^{q}(\zeta_r^{N_r} - \zeta_r^{-N_r})/N_r - \frac{1}{q}\prod_{r=1}^{q}(\zeta_r - \zeta_r^{-1}) \in \mathcal{X}^*,$$

(3.6) $$N^{-q}(z^N - 1)^q - (z - 1)^q \in K,$$

(3.7) $$N^{-q}(z^N - z^{-N})^q - (z - z^{-1})^q \in K^* .$$

§4. Translations and differences. Let $f(x)$ be an arbitrary function defined in a Euclidean space, termed x-space, or in particular, if we write t for x, on the real line, and suppose the values of $f(x)$ lie in a vector space. We shall give to the identities of the preceding section a new interpretation in terms of the translations of $f(x)$ that arise from translations of x-space along coordinate axes (not necessarily a same coordinate axis for different translations).

To this effect, it will be convenient to define ζ_r and z in (3.4), (3.6) as translations, by setting $\zeta_r f(x) = f(x + h_r)$, $zf(t) = f(t+h)$, and to modify these definitions in (3.5), (3.7), by writing instead $\zeta_r f(x) = f(x + \frac{1}{2}h_r)$, $z f(t) = f(t + \frac{1}{2}h)$. Evidently our assertions remain valid in the algebra of such translation operators. We shall formulate them in terms of difference operators. If h is a vector along a coordinate axis, the simple difference Δ_h, and the simple symmetric difference Δ_h^*, are defined by setting

$$\Delta_h f(x) = f(x+h) - f(x), \quad \Delta_h^* f(x) = f(x + \tfrac{1}{2}h) - f(x - \tfrac{1}{2}h) .$$

(4.1) <u>Let</u> $\mathcal{K}_f, \mathcal{K}_f^*$ <u>denote, respectively, the sets of convex combinations of expressions of the form</u>

$$\text{(i)} \quad \Big(\prod_{r=1}^{q} \Delta_{h_r}\Big) \Delta_k f(x+\eta), \qquad \text{(ii)} \quad \Big(\tfrac{1}{2} \prod_{r=1}^{q} \Delta_{h_r}^*\Big)(\Delta_k^*)^2 \, \frac{f(x+\eta) + f(x-\eta)}{2},$$

<u>where</u> k <u>denotes an arbitrary positive multiple</u> $n_r h_r$ <u>of some</u> h_r, <u>and</u> η <u>an arbitrary sum</u> $\sum n_r h_r$ <u>of such positive or negative multiples, subject to the condition that each</u> n_r <u>is an integer in (i), or a half-integer in (ii) not exceeding</u> N_r. <u>Then</u>

$$q^{-1}\Big\{\prod_{r=1}^{q} (N_r^{-1} \Delta_{N_r h_r}) - \prod_{r=1}^{q} \Delta_{h_r}\Big\} f(x) \in \mathcal{K}_f ,$$

$$q^{-1}\Big\{\prod_{r=1}^{q} (N_r^{-1} \Delta_{N_r h_r}^*) - \prod_{r=1}^{q} \Delta_{h_r}^*\Big\} f(x) \in \mathcal{K}_f^* .$$

(4.2) <u>Let</u> K_f, K_f^* <u>denote, respectively, the sets of convex combinations of expression of the form</u>

$$\text{(i)} \quad (\Delta_h)^q \Delta_k f(t), \qquad \text{(ii)} \quad \tfrac{1}{2}(\Delta_h^*)^q (\Delta_k^*)^2 f(t) ,$$

<u>where</u> k <u>denotes an arbitrary multiple</u> nh <u>of</u> h <u>such that</u> n <u>is a positive integer in (i), or a positive half-integer in (ii), not exceeding</u> qN. <u>Then</u>

$$\{N^{-q}(\Delta_{Nh})^q - (\Delta_h)^q\} f(t) \in K_f , \quad \{N^{-q}(\Delta_{Nh}^*)^q - (\Delta_h^*)^q\} f(t) \in K_f^* .$$

§5. <u>A characterization theorem for distributions which reduce to continuous functions.</u> We propose to extend to continuous functions of several variables a theorem proved in (26, Appendix A) on the existence and continuity of an n-th derivative of a function $f(t)$. The straightforward extension is false: the correct extension turns out to be a theorem on Schwartz distributions, or alternatively, on a different kind of n-th derivative.

We recall a definition used in the Theorem quoted. A pair of real functions φ, ψ defined on the interval $0 \le u \le 1$ are termed estimate functions of orders q, m, where q, m are positive integers, if φ is non-negative and Borel measurable, while ψ is continuous and monotone increasing, and takes the value $\psi(0) = 0$, and if further

(5.1) <u>for</u> $0 < \lambda < 1$, $\varphi(\lambda u) \ge (\frac{1}{2}\lambda)^q \varphi(u)$ <u>and</u> $\psi(\lambda u) \ge (\frac{1}{2}\lambda)^m \psi(u)$,

(5.2) $S(1) < \infty$, <u>where</u> $S(h) = \int_0^h u^{-q} \varphi(u) d\psi(u)$.

We begin by following the general line of argument of the theorem referred to. The function $f(x)$ of the preceding section will now be supposed continuous, and to have values in a Banach space, or in particular real or complex values. We fix a vector n in x-space, with non-negative integers n_i as components, where $\sum n_i = q$. We shall use the notion of the mixed derivative of order q with type n, or simply the n-th derivative $D^n f(x)$, in two senses. One of them results from n_i partial derivations in x_i, for each i. However in that case we shall always specify that the partial derivations are understood in the sense of Schwartz distributions, so that the manner in which we order them is immaterial. The other definition of $D^n f(x)$, the one we employ unless the contrary is stated, is as the unique limit of

$$\prod_{r=1}^{q} h_r^{-1} \Delta_{h_r \theta_r} f(x) ,$$

for reals $h_r \neq 0$ which tend to 0. Here $\theta_r (r = 1, \ldots, q)$ are unit vectors along coordinate axes, and we stipulate that, for each i, n_i of the θ_r are along the axis of x_i. However, we shall require also a more special definition, of what we term the symmetric n-th derivative $D^{n*} f(x)$. We define it as the limit, as $h \to 0$ of the expression $h^{-q} \Delta_h^{n*} f(x)$, where $h > 0$, where

$$\Delta_h^{n*} = \prod_{r=1}^{q} \Delta^*_{h\theta_r} ,$$

and where the θ_r are as before. We term Δ_h^{n*} the symmetric n-th difference of gauge h, or simply the $n*$ difference of gauge h. We define also an $n*$ integral of gauge h, which we write $\mathcal{J}_h^{n*}$: we do so by setting

$$\mathcal{J}_h^{n*} f(x) = \Delta_h^{n*} F(x) ,$$

where F is the n-th indefinite integral of f, which results from f by q indefinite integrations, of which, for each i, n_i are in x_i. Finally we shall have occasion to use the $n^* + 2^*$ difference

$$\Delta_h^{n*} (\Delta^*_{k\theta})^2 ,$$

where θ is an arbitrary coordinate unit vector. These various notations will on occasion, be supposed extended in the obvious way to distributions.

The function $f(x)$ will be said to admit, for its $n^* + m^*$ difference, the estimate functions φ, ψ near the point x, or at the point x, if φ, ψ are as described above, but of orders q, m and if, for all small h, k subject† to $k \geq \frac{1}{2} h$ and for all $\hat{x}$ of some neighbourhood of x, or for $\hat{x} = x$, as the case may be,

†We could avoid the unnecessary factor $\frac{1}{2}$ entirely. However, we really use the condition only for $k \geq h$, except in the case of estimate functions at the point x, where we set here $m = 2$ (Theorem 6.3).

(5.3) $$|\Delta_h^{n*}(\Delta_{k\theta}^*)^m f(\hat{x})| \leq \varphi(h)\,\psi(k),$$

for each coordinate unit vector θ. Similarly, a distribution T, expressible near x as a distribution n-th derivative of a continuous $f(x)$, will be said to admit, for its n^* integrated m^* difference the estimate functions φ,ψ, if for small h,k subject to $k \geq \frac{1}{2}h$, $f(x)$ satisfies (5.3), or equivalently if, for each coordinate unit vector θ,

$$\mathcal{J}_h^{n*}(\Delta_{k\theta}^*)^{\hat{m}}T$$

is near x a continuous function, with local supremum norm $\leq \varphi(h)\psi(k)$.

Of course every T is for some n locally such an n-th derivative. However the existence of a corresponding pair φ,ψ is a different matter. What is clear a priori is that if T satisfies the above condition for a given set of n,φ,ψ and of the associated q, it must satisfy it also for n augmented to $n+n'$, and q correspondingly to $q+q'$, if ψ is kept the same, and $\varphi(h)$ is simply replaced by $h^{q'}\varphi(h)$. For $n = 0$, the condition requires T to be near x a continuous function. (In that case a possible pair φ,ψ is obtained by setting $\varphi(u) = 1$, and by choosing for ψ the m-th modulus of continuity.)

The following is a strong converse in the case $m=2$.

(5.4) Theorem. Let n be a vector of x-space with non-negative integers n_i as components, where $\sum n_i = q$, and let T be a Schwartz distribution which admits, for its n^* integrated 2^* difference, estimate functions φ,ψ near x. Then T is, near x, a continuous function.

We shall derive Theorem (5.4) from the following result.

(5.5) Theorem. Let $f(x)$ be continuous, and suppose it admits near x^0 estimate functions φ, ψ for its $n^* + 2^*$ difference. Then $D^n f(x)$ exists near x^0 and is continuous. Moreover, near x^0,

$$(5.6) \qquad |D^n f(x) - h^{-q} \Delta_h^{n*} f(x)| \leq K S(h),$$

and the second modulus of continuity of $D^n f(x)$ is $\leq K S(h)$.

Here $S(h)$ is as in (5.2). Moreover, $D^n f(x)$ has the meaning explained above, and its continuity clearly implies that it then agrees with the corresponding n-th derivative in the sense of distributions. Thus (5.4) does reduce to a consequence of (5.5), which we prove in §§7-9 below. We repeat that $D^n f(x)$ is not here the classical n-th derivative, the existence of the latter would presuppose that of appropriate partial derivatives of lower orders, and they need not exist. This is the main difference between (5.5) and its one-dimensional version, treated in (26, Appendix A).

§6. Some one-dimensional variants, and their relation to Tauberian theorems and to approximation theory. We limit ourselves again to new results. The simplest is (6.3), in which the $n^* + 2^*$ condition is assumed only at t_0. Our first two theorems are equivalent forms of the analogue of (5.4) for the $n^* + 4^*$ difference. We have been unable to extend them further. We could also interpret them as a Tauberian theorem. We give such an interpretation in detail in the case of a one-sided variant, Theorem (6.4), with $m=1$ and n arbitrary: it constitutes Theorem (6.5). We recall that t now replaces x, and that $n = q$.

(6.1) Theorem. Let T be a Schwartz distribution on the real line, which admits, for its n^* integrated 4^* difference, estimate functions

φ,ψ near t_0. Then T is, near t_0, a continuous function, with a fourth modulus of continuity $\leq K(n)S(h)$.

(6.2). Theorem. Let $f(t)$ be continuous, and suppose it admits for its $n^* + 4^*$ difference, estimate functions φ,ψ near t_0. Then $D^n f(t)$ exists near t_0, is continuous, and is the classical n-th derivative of $f(t)$. The fourth modulus of continuity of $D^n f$ is $\leq K(n)S(h)$.

(6.3) Theorem. Let $f(t)$ be continuous, and suppose it admits at t_0 estimate functions φ,ψ for its $n^* + 2^*$ difference. Then the symmetric n-th derivative $D^{n*} f(t_0)$ exists, and

$$|D^{n*} f(t_0) - h^{-n} \Delta_h^{n*} f(t_0)| \leq K(n)S(h).$$

We shall formulate a one-sided variant only in the simplest case. The function $f(t)$ is then real-valued and has a continuous n-th derivative $g(t)$, and we write by definition $I_h^n \Delta_k g(t)$ to mean $\Delta_h^n \Delta_k f(t)$. We say that g admits, for its n-th integrated first difference for large t, one-sided estimate functions φ,ψ, if the latter are of orders $n,1$, and if, for large t and small k, where $k \geq h > 0$, we have $\Delta_h^n \Delta_k f(t) \geq -\varphi(h)\psi(h)$. In particular, if $n = 0$, $\varphi = 1$, $f = g$, the condition is $\Delta_k g(t) \geq -\psi(k)$, which simply means $\Delta_k g(t) \geq -\varepsilon_k$ where $\varepsilon_k \to 0$ as $k \to 0$. This will be termed uniform one-sided continuity, and we consider it for large t.

(6.4) Theorem. Let $g(t)$ be real and continuous, and suppose it admits, for its n-th integrated first difference for large t, one-sided estimate functions φ,ψ. Then $g(t)$ has uniform one-sided continuity for large t, in fact

$$\Delta_k g(t) \geq K(n)S(k).$$

(6.5) Theorem. Let $g(t)$ be continuous, and suppose it is either real and subject to the hypotheses of (6.4), or else that it is uniformly bounded and satisfies for large t the hypotheses of (6.1) for the distribution there denoted by T. In addition, suppose that, as $t \to \infty$,

$$G(t) = e^{-t} \int_0^t g(u) e^u du \to L .$$

Then $g(t) \to L$.

We observe that in the real case, with the one-sided hypotheses of (6.4), this last result reduces to triviality in the case of a function g which is monotone for $\nu-1 \leq t \leq \nu$, $\nu = 1, 2, \ldots$. In fact, if we write $s_\nu = g(\nu)$, $a_\nu = s_\nu - s_{\nu-1}$, the assumption $G(t) \to L$ becomes

$$s_\nu - \theta a_\nu = s_{\nu-1} + (1-\theta) a_\nu \to L ,$$

where θ is some θ_ν subject to $0 < \theta < 1$. However, if we first change variables by substituting t for e^t, a corresponding specialization of g provides, in the case $n=0$, one of the most elementary of known Tauberian theorems: Suppose $\sum a_\nu$ converges (C,1). Then in order that it converges in the ordinary sense, it is necessary and sufficient that it satisfy the one-sided Landau condition, namely that, as function of $\delta > 0$, the quantity

$$\operatorname*{Sup}_{\mu,\nu} (s_\nu - s_\mu) \text{ for } \mu < \nu < (1+\delta)\mu$$

be $\geq -\varepsilon_\delta$, where $\varepsilon_\delta \to 0$ as $\delta \to 0$. Clearly, for a series which converges (C,1) this condition is then necessary, as well as sufficient, for convergence in the ordinary sense. It follows from a theorem of N. Wiener (11, Theorem 4, p. 73), that if we strengthen the condition by requiring s_ν to be bounded, it becomes necessary and sufficient for ordinary convergence of a series which converges in the Abel sense. This is then the long-known one-sided version of a theorem of Landau (10, §3, p. 270), which is also necessary and sufficient. It includes the famous one-sided Tauberian theorem of Littlewood, which assumes $a_\nu \geq -K/\nu$, a condition merely sufficient. Our theorem (6.5), on the other hand, while still necessary and sufficient in this context, weakens the one-sided Landau condition by n integrations with respect to the old variable t which is now $\log t$.

Similar comments apply when we compare our theorems to the results of classical approximation theory. The classical theorems, like the Tauberian theorem of Littlewood, are best possible of their kind, but they provide only sufficient conditions for their conclusions. Our results are best possible, as proved in (27), that is to say if φ, ψ are subject to (5.1) and not to (5.2), there is an f for which the corresponding assertions become false. However our conditions are necessary as well as sufficient, whereas the classical ones, which go back to Marchaud (13) and which set $h = k$, instead of $\frac{1}{2}h \leq k$, are not necessary. Marchaud established the existence and continuity of $D^n f(t)$, if the $n+1$-st modulus of continuity of f is, in our terminology, $\leq \varphi(h)\psi(h)$, provided φ, ψ are subject to a stronger condition than (5.2), namely

$$(6.6) \qquad \int_0^h u^{-n-1}\varphi(u)\psi(u)du < \infty .$$

In fact, if $D^n f$ exists and is continuous, our theorems allow us to take for $\psi(h)$ its modulus of continuity and to set $\varphi(h) = c\,h^n$, while the classical condition would require further that $\int u^{-1}\psi(u)du < \infty$.

The assumption (6.6), which we do not make here, is precisely what distinguishes from our results the classical ones requiring only the case $h = k$

of (5.3). With this assumption, our proofs would not only be correspondingly simpler, they would also allow us rather easily to extend the results to $m > 4$.

We could have simplified our theorems, alternatively, by setting throughout $m = 1$, and by replacing symmetrical differences by their translations, and therefore by ordinary differences. Then (6.1) and (6.2), the most difficult of our theorems, would have become more special cases of the correspondingly simplified form of Theorem (5.5), where the final part of the conclusion would now concern the first, instead of the second, modulus of continuity of $D^n f$. The theorems would be weaker, mainly because, by setting $m = 1$, we would strengthen our hypotheses in (5.1), and therefore affect the possible orders of magnitude of $S(u)$. The resulting restrictions on φ,ψ would be such as to render it immaterial in practice, whether the modulus of continuity of $D^n f$ considered is the first or the second. This last can be seen by applying classical equivalence theorems for the higher moduli of continuity, theorems which may be considered to assert, in most practical cases, the "near-futility" of higher moduli of continuity when they do not tend to 0 rapidly enough. Thus, if $\chi(h)$ denotes an order of magnitude, and ε,θ are between 0 and 1, the relation

$$\chi(\theta h) < K\,\theta^{\varepsilon}\chi(h)$$

ensures the equivalence of the types of relations

$$\omega_{q+r} = O(h^q/\chi) \text{ and } \omega_q = O(h^q/\chi)\,, \tag{6.7}$$

for moduli of continuity $\chi_p(h)$ of orders $q+r$ and q of uniformly bounded functions where $r > 0$. See for instance Timan (17).

We observe that conditions of the type (5.3), which involve two kinds of differences, one in h, the other in k, are most natural in the case of a function f expressible as a convolution $f_1 * f_2$, since we then have

$$\Delta_h^{n*}(\Delta_{k\theta}^*)^m f = (\Delta_h^{n*} f_1) * (\Delta_{k\theta}^*)^m f_2 \quad .$$

It is in this form that our methods were first used in (23,24), to establish the existence of certain Stieltjes integrals, and in (25,26) to define corresponding stochastic integrals. In the case where $\varphi(u)$, $\psi(u)$ are powers u^α, u^β $(0 < \alpha,\beta < 1)$ the Stieltjes integrals were found in (12) to reduce to classical ones by fractional integration by parts.

Generally, fractional integration and derivation has been used also in the classical approximation theorems. We shall indicate briefly in §10 the effect of these operations on our results.

§7. The lemmas for $m = 4$. It is convenient to begin with these, as all our theorems can then be obtained by a single method of proof. Our lemmas concern symmetric differences.

(7.1) Lemma. The quantity

$$\frac{1}{3}(\Delta_{k'}^*)^2(\Delta_{k''}^*)^2 f(t)$$

can be expressed as a convex combination of terms of the form

(7.2) $$(\Delta_{\frac{1}{2}k'\pm\frac{1}{2}k''}^*)^4 \frac{f(t+\eta)+f(t-\eta)}{2}, \quad -(\Delta_{k'}^*)^4 f(t), -(\Delta_{k''}^*)^4 f(t),$$

where the η have the form $\frac{1}{2}r(k'\pm k'')$, $r = 0,1,2,3,4$.

(7.3) Lemma. Let $R = R_h = R_{h,N} = \{(Nh)^{-q}(\Delta_{Nh}^*)^q - h^{-q}(\Delta_h^*)^q\}f(t)$, and let $\hat{R} = R_{\hat{h}}$. Then $R - M^{-2}\hat{R}$ is a convex combination of terms of the form

$$-\tfrac{1}{2}h^{-q}(\Delta_h^*)^q(\Delta_{k'}^*)^2(\Delta_{k''}^*)^2 f(t)$$

and of similar terms multiplied by M^{-2}, where k', k'' are quarter integer multiples sh of h, $s \le 2qMN$.

Proofs. From the identity

$$12u^2v^2 = (u+v)^4 + (u-v)^4 - 2u^4 - 2v^4,$$

we derive, by setting $u = \sin 2\alpha$, $v = \sin 2\beta$, an expression of

$$\tfrac{1}{3}(2\sin 2\alpha)^2(2\sin 2\beta)^2$$

as a convex combination of the terms

$$(2\sin(\alpha+\beta)\cos(\alpha-\beta))^4,\ (2\sin(\alpha-\beta)\cos(\alpha+\beta))^4,\ -(2\sin 2\alpha)^4,\ -(2\sin 2\beta)^4.$$

Here, by setting $y = e^{i\alpha}$, $z = e^{i\beta}$, we express

$$\tfrac{1}{3}(y^2-y^{-2})^2(z^2-z^{-2})^2$$

as a convex combination of terms of the form

$$(yz - y^{-1}z^{-1})^4 P,\ (yz^{-1}-y^{-1}z)^4 P,\ -(y^2-y^{-2})^4,\ -(z^2-z^{-2})^4,$$

where P is of the form

$$\tfrac{1}{2}(p^r + p^{-r}),\ r = 0,1,2,3,4,\quad p = yz \text{ or } p = yz^{-1}.$$

From the identity in y, z thus obtained, we derive (7.3) by interpreting y, z as translation-operators $t \to t+\frac{1}{4}k'$, $t \to t+\frac{1}{4}k''$, on $f(t)$.

To prove (7.3), we subtract from (4.2), M^{-q-2} times the corresponding identity with h replaced by $\hat{h} = Mh$. In this way $h^q(R-M^{-2}\hat{R})$ becomes a convex combination of expressions

$$\tfrac{1}{2}(\Delta_h^*)^q(\Delta_{k'}^*)^2 f(t) - \tfrac{1}{2}M^{-q-2}(\Delta_{Mh}^*)^q(\Delta_{Mk}^*)^2 f(t),$$

each of which is the mean (with weights $\frac{1}{2}$) of

$$-(\Delta_h^*)^q\{M^{-2}(\Delta_{Mk}^*)^2 - (\Delta_k^*)^2\}f(t) \quad\text{and}\quad -M^{-2}\{M^{-q}(\Delta_{Mh}^*)^q-(\Delta_h^*)^q\}(\Delta_{Mk}^*)^2 f(t).$$

By applying (4.2), with M in place of N, to the two curly brackets, the second time with 2 in place of q, we find that the last two expressions become, respectively, convex combinations of terms of the form required.

§8. Characteristic inequalities along carefully thinned binary sequences. Let $\{h\}$ be a subsequence of the binary sequence $h_0 2^{-\nu}$, $\nu = 0,1,2,\ldots, h_0 > 0$. [A ternary sequence $h_0 3^{-\nu}$ would be slightly more convenient--the ratio of successive terms of a subsequence is then odd, and an odd N in (4.2) eliminates half-integers.] We choose this subsequence so that the characteristic property of a geometric series in practice--that its remainder be of the order of its difference--should hold, not for its terms h_ν $\nu = 0,1,2,\ldots$, but for the corresponding $\psi(h_\nu)$. We term it a carefully thinned binary sequence, if

(8.1) $$2\psi(h_\nu) \leq \psi(h_{\nu-1}) \leq 2^{2m+1}\psi(h_\nu) .$$

The existence of such an $\{h\}$--elsewhere termed one subject to the conditions C(1) and C(2)--is proved in (26, Appendix A). We associate with it the series

(8.2) $$\sum_{\nu=1}^{\infty} h_\nu^{-q}\varphi(h_\nu)\psi(h_{\nu-1}) ,$$

or "φ, ψ estimate series", whose sum $S_{\{h\}}$ is shown in the same reference to satisfy the relations

(8.3) $$2^{-q}S(h_0) \leq S_{\{h\}} \leq 2^{q+4n+2} S(h_1) ,$$

where h_0, h_1 are the initial terms of $\{h\}$, and where $S(h)$ is the integral defined in (5.2).

To avoid keeping precise track of factors $K(q,m)$ such as occur in (8.3), we term generally K-estimate series, and denote by $\sum b_\nu$, any series derived from (8.2) by at most K operations, consisting of multiplying by at most K, or of adding term by term the remainder series after at most K terms. For instance, on account of (8.1), the general term of (8.2) is increased at most by a factor K, if in it $\psi(h_{\nu-1})$ is replaced by $\psi(Mh_{\nu-1})$ $M \leq 2^r$, or even by $\psi(h_{\nu-1-r})$, where $r < K$. Similarly it becomes at most K times the corresponding $(\nu+r)$-th term, if in it we replace $\varphi(h_\nu)$ by $\varphi(2^{-r}h_\nu)$. With this notation, the proofs of our theorems will depend on estimates of the form

(8.4) (i) $\alpha_\nu, \alpha_\nu \leq b_\nu$, or (ii) $\sum \alpha_\nu \leq \sum b_\nu + 2\alpha_0$, $\sum a_\nu \leq \sum b_\nu + 2a_0$,

or, in the one-sides case,

(8.5) $$R_\nu^- \leq b_\nu, \quad R_\nu^+ \geq -b_\nu .$$

On the left of these "characteristic inequalities", the various quantities are defined as follows, R_ν^-, R_ν^+ as functions of t, and α_ν, a_ν as the maxima in x (or in particular in t) of functions $\alpha_\nu(x)$, $a_\nu(x)$ in a closed neighbourhood of x_0 or t_0, except in connection with Theorem (6.3) when we take α_ν, a_ν to be the corresponding values for $t = t_0$. We first write, in agreement with the preceding section, $R_{h,N}$ for

$$(Nh)^{-q}\Delta_{Nh}^{n*} f(x) - h^{-q}\Delta_h^{n*} f(x) ,$$

and, when $x = t$, $h > 0$, $R_{h,N}^+$ for the corresponding expression without stars, $R_{h,N}^-$ for the similar expression with $-h$ for h. We then define R_ν, R_ν^+, R_ν^- as the values of these quantities when $h = h_\nu$, $Nh = h_{\nu-1}$, and we set (with $x = t$ when relevant)

$$\alpha_\nu(x) = |R_{h,M}| \text{ when } h = h_\nu \text{ and } M \text{ is fixed,}$$

$$a_\nu(x) = \mathrm{Sup}_{h,N}|R_{h,N}| \text{ when } h_\nu \leq h,\ Nh \leq h_{\nu-1},$$

and when h, Nh are further restricted to belong to the complete binary sequence $h_0 2^{-s}$, $s = 0,1,2,\ldots$.

With the hypotheses of Theorems (5.5), (6.3) or (6.4), the verification of (8.4) (i), or in the case of (6.4) that of (8.5), follows at once from (5.3) and the identities (4.1) or (4.2), or in the case of (6.4) the one-sided analogue of (5.3) together with these identities. [This is where we use the unnecessary factor $\frac{1}{2}$ in the relation $k \geq \frac{1}{2}h$ which affects (5.3).]

With the hypotheses of Theorem (6.2), we shall establish only (8.4) (ii). For this, we show that both $\sum \alpha_\nu$ and $\sum a_\nu$ are $\leq K \sum \bar{b}_\nu$, where $\bar{b}_\nu = b_\nu + b_{\nu+3}$. We first observe that $K\bar{b}_\nu$ is not less than the norm of the expression

(8.6) $\quad h^{-q} \Delta_h^{q*} \Delta_k^{4*} f(t) \quad$ for $h_\nu \leq h \leq h_{\nu-1}$, $\frac{1}{8} h \leq k \leq K h_{\nu-1}$.

By (5.1) this norm is $\leq K b_\nu$ as long as $k \geq h$, since f satisfies (5.3). However, if $k < h$, the norm is at most K times a convex combination of translations of the function of t given by the norm of the quantity

$$\left(\frac{h}{8}\right)^{-q} \Delta_{\frac{h}{8}}^{q*} \Delta_k^{4*} f(t), \qquad \frac{h}{8} \leq k \leq h,$$

which is at most $b_{\nu+3}$ in an appropriate neighbourhood of t_0.

Next we apply Lemma (7.3), with the roles of M, N reversed, and with $N = h_{\nu-1}/h_\nu$, $h = h_\nu$, and we transform the result by Lemma (7.1). We find that

$$R_{h,M} - N^{-2} R_{Nh,M}$$

is a convex combination of multiples by $\pm K$ of translations of (8.6). Since $N \geq 2$, we may write, by identifying $K\bar{b}_\nu$ with b_ν,

$$\alpha_\nu - \frac{1}{4}\alpha_{\nu-1} \leq b_\nu,$$

from which it evidently follows that $\frac{3}{4} \sum \alpha_\nu \leq \frac{5}{4}\alpha_0 + \sum b_\nu$, or as we prefer to write, that $\sum \alpha_\nu \le 2\alpha_0 + \sum b_\nu$, by a K-operation. Similarly, if the sums are from ν_0 to ∞, α_0 is here replaced by the value of α_ν for $\nu = \nu_0$.

It remains to be shown that $\sum a_\nu \le 2a_0 + \sum b_\nu$. We apply for this purpose Lemma (7.3) with $M = 2$. If we again transform the result by Lemma (7.1), the difference expressed as a convex combination is

$$a_\nu - \frac{1}{4}\hat{a}_\nu$$

where, for some t, h, N which provide the maximum a_ν for $R_{h,N}$ under the conditions stated in its definition, the quantity $\hat{a}_\nu$ denotes the corresponding value of $R_{2h,N}$. The terms of the convex combination have as before the common majorant b_ν. Further, for $\nu \ge 1$, we clearly have $\hat{a}_\nu \le a_\nu + a_{\nu-1}$. By combining this with the relation

$$a_\nu - \frac{1}{4}\hat{a}_\nu \le b_\nu ,$$

we find that, for $\nu \ge 1$,

$$a_\nu + a_{\nu-1} \ge 4(a_\nu - b_\nu) ,$$

and hence that

$$2\sum a_\nu \le 4\sum b_\nu + 3a_0 \le 4\sum b_\nu + 4a_0 ,$$

which, by a K-operation, can be written in the desired form.

§9. Proofs of theorems of §§5,6. The verification of (6.5), subject to (6.4) and to its hypotheses, is elementary. Now g is real and has uniform one-sided continuity. The assertion becomes obvious, if the upper limits of $L - g(t)$ and $g(t+h) - L$ as $t \to \infty$ are $\ge -\varepsilon_h$, where $\varepsilon_h \to 0$ as $h \to 0$. However this follows at once from the relations

$$e^h G(t+h) - G(t) - (e^h - 1)g(t) = \int_t^{t+h} (g(u) - g(t))e^{u-t}du \ge -(e^h - 1)\varepsilon_h ,$$

$$-e^h G(t+h) + G(t) + (e^h - 1)g(t+h) = \int_t^{t+h} (g(t+h) - g(u))e^{u-t}du \ge -(e^h - 1)\varepsilon_h ,$$

by dividing by $e^h - 1$ and making $t \to \infty$.

We next dispose of (6.4). In this case g is the continuous n-th derivative of its indefinite n-th integral f. Hence $D^n f = g$, so that $\sum R_\nu^+$ converges and has for sum the value for $h = h_0$ of the difference

$$h^{-n} \Delta_h^n f(t) - g(t) .$$

Similarly $\sum R_\nu^-$ converges to a similar expression, in which h is changed to $-h$, and in it we change the variable from t to $t + nh$. Since the difference ratios

$$h^{-n} \Delta_h^n f(t), \ (-h)^{-n} \Delta_{-h}^n f(t + nh)$$

are identical, we find by subtraction that $g(t + nh) - g(t)$ is a difference $\sum R_\nu^+ - \sum R_\nu^-$, where the first sum is at the point t and the second at $t + nh$. Hence by (8.5) and (8.3)

$$g(t + n h_0) - g(t) \geq -KS(h_1) \geq -KS(nh_0) ,$$

so that Theorem (6.4) follows by taking for $n h_0$ an arbitrary small $k > 0$.

We pass on to (5.5), (6.2) and (6.3). From (8.4) (i) or (ii), it then follows first that $\sum a_\nu$, and therefore $\sum R_\nu$, converge. This means that the difference ratio

$$h^{-q} \Delta_h^{n*} f(x) , \tag{9.1}$$

or in particular the corresponding expression with t or t_0 for x, has a limit L_0 as h describes the sequence $\{h\}$. In fact $\sum R_\nu$ has the value of (9.1) at $h = h_0$, diminished by L_0. The same limit L_0 is clearly then also that of (9.1) as h describes the complete binary sequence $h_0 2^{-\nu}$, $\nu = 0,1,2,\dots$, since $a_\nu \to 0$. Moreover, since $\alpha_\nu \to 0$, L_0 does not alter when h_0 is replaced by any multiple $M h_0$ by a positive integer M, nor therefore when h_0 is replaced by a rational multiple.

With the hypotheses of (5.5) or (6.3), we have, by (8.4) (i) and (8.3),

$$(9.2) \qquad |h^{-q}\Delta_h^{n*} f(x) - L_0| \le K\,S(h)\,,$$

for $h = h_0$. Therefore, if h_0 is rational, this holds also for any rational h below a certain bound, without altering L_0. Evidently it must then hold equally, when h is irrational. By definition $D^{n*}f(x)$ therefore exists and equals L_0, which we now see has to be independent of h_0. This incidentally disposes of (6.3).

We proceed with the proofs of (5.5) and (6.2). In the latter we have still not quite got to the same stage, and we go back to the point where $L_0 = L_0(x, h_0)$ was found to be the limit of (9.1) along $\{h\}$. Clearly this limit is uniform, by the Weierstrass M-test, so that L_0 is continuous near x_0 (or t_0) in x. We have moreover, for $h = h_0$, with the hypotheses of (6.2),

$$(9.3) \qquad |h^{-q}\Delta_h^{n*} f(t) - L_0| \le \sum a_\nu \le 2a_0 + \sum b_\nu\,.$$

If we replace in this $f(t)$ by $\Delta_h^{4*} f(t)$, a_0 evidently becomes $\le 2\,b_0$, and the remaining a_ν are multiplied by at most K, so that the extreme right-hand side becomes $\le KS(h)$. Similarly the norm of (9.1) becomes $\le b_0$. We thus find that, for $h = h_0$,

$$(9.4) \qquad |\Delta_{h\theta}^{m*} L_0| \le K\,S(h)$$

with $m = 4$, $\theta = 1$ in the case relevant to (6.2). The same is true with $m = 2$, θ a unit vector along any of the axes, in the case relevant to (5.5), by a simplification of the above. It remains only to show that L_0 is $D^{n*}f(x)$.

To this effect, we denote by f_1 the n-th indefinite integral of the continuous function derived from L_0 by multiplying by a suitable localizing factor, and we write $f_0 = f - f_1$. If now λ denotes a C^∞ function with support in the neighbourhood of x_0, the convolution $f_0 * \lambda$ has an n-th derivative which is the uniform limit along $\{h\}$ of the finite difference ratio corresponding to (9.1). This n-th derivative is identically zero, obviously. Hence $f_0 * \lambda$ has a vanishing n-th difference, and by choosing λ so that $f_0 * \lambda$ converges uniformly to f_0, we see that the n-th difference of f_0 vanishes identically near x_0. Thus near x_0 we have $D^n f(x) = D^n f_1(x)$ and since $D_n f_1(x) = L_0$, we find that $D^n f(x) = L_0$. The relation (9.4) then shows that the m-th modulus of continuity of $D^n f(x)$ near x_0 is $\leq K S(h)$, while (9.2) gives the approximation to $D^n f(x)$ asserted in (5.5). Finally, in the one-dimensional case, f_0 is evidently a polynomial of degree $< n$ near x_0, and L_0 is the ordinary derivative of f_1, and therefore of f. This completes the proofs of (5.5) and (6.2). Clearly (5.4) and (6.1) are immediate consequences.

Finally we dispose of the part of (6.5) which relates to hypotheses of the type of (6.2). By the arguments used in the one-sided version, it is enough to show that g is uniformly continuous. This follows from the classical futility Theorem (1 :), quoted in connection with (6.7) above, since g is bounded and has by (6.2) its fourth modulus of continuity $\leq K\,S(h)$.

§10. Fractional derivatives, derivatives in L^2. By contrast with the formula of fractional integration by parts, which, as stated in §6, provides information about the entities concerned, the extension of the results of §§5,6 to fractional derivatives, merely repeats, in a different notation, results already obtained. Nevertheless, the notation is convenient.

We shall consider only the Riesz definition of an α-th integral, and we restrict ourselves to real α. What matters here is not the choice of such a definition, but the manner in which this choice then determines both the notion of $(q-\alpha)$th - derivative and that of $(q-\alpha)$-th difference. Thus an intrinsic definition of $(q-\alpha)$-th derivative would be useless to us, unless attached to a correspondingly intrinsic definition of $(q-\alpha)$-th difference.

We define, for $0 < \alpha < 1$; $D^{q-\alpha}f(t)$ to mean $D^q f_\alpha(t)$, where f_α denotes the α-th integral of f. Similarly the differences $\Delta_h^{q-\alpha}$, $\Delta_h^{*q-\alpha}$ of $f(t)$ mean $h^{-\alpha}$ times the corresponding differences Δ_h^q, Δ_h^{*q} of f_α. The modulus of continuity $\omega_{q-\alpha}(f,u)$ is $u^{-\alpha}$ times the supremum in t,h, for $|h| \leq u$ and arbitrary t, of the quantity $|\Delta_h^{q-\alpha}f(t)\, h^\alpha|$. In these definitions, we restrict ourselves to functions $f(t)$ of one real variable; and we suppose each f periodic, of period 2π and mean value 0, and moreover either continuous, or p-th power integrable in the period for a suitable p.

The notions of estimate functions φ, ψ of orders λ, μ, where λ, μ are positive reals, are the formal analogues of those given previously for $\lambda, \mu = q, m$. We shall, however, restrict ourselves to $\mu = m = 1, 2$, or 4. What matters is that if φ, ψ are estimate functions of orders $q-\alpha, m$ then the functions $u^\alpha \varphi(u)$, $\psi(u)$ are estimate functions of orders q, m.

The fractional analogues of Theorems (5.5) and (6.2) are automatic.

(10.1) <u>Theorem.</u> <u>Let</u> $0 \leq \alpha < 1$, <u>let</u> $m = 2$ <u>or</u> 4, <u>and let</u> q <u>be a positive integer. Further let</u> $f(t)$ <u>be continuous, of period</u> 2π <u>and mean-value</u> 0, <u>and suppose that, for</u> $0 < h \leq 1$, $\frac{1}{2}h \leq k \leq 1$,

$$|\Delta_h^{*q-\alpha}\, \Delta_k^{*m}\, f(t)| \leq \varphi(h)\, \psi(k),$$

where φ,ψ *are estimate functions of orders* $q-\alpha, m$, *and where in particular*

$$S(h) = \int_0^h u^{-(q-\alpha)}\varphi(u)d\psi(u)$$

converges. Then $f(t)$ *has a continuous* $(q-\alpha)$*-th derivative* $D^{q-\alpha}f(t)$, *of modulus of continuity of order* m *not exceeding* $K(q)S(u)$. *Moreover, in the case* $m = 2$,

$$|D^{q-\alpha}f(t) - h^{-(q-\alpha)}\Delta_h^{*q-\alpha}f(t)| \leq K(p)S(h) .$$

With regard to derivatives in L^2, we recall the fact that, as described in (26, Appendix A, pp. 162-166), the theorems concerning them are theorems of the type (10.1) for skew self-convolutions

$$F = f * g \quad \text{where} \quad f(t) = \bar{f}(-t) .$$

Here f is taken to be complex-valued, and $\bar{f}(-t)$ is the complex conjugate of $f(t)$. In addition f is again supposed periodic, and this time square integrable over a period. The convolution is taken over a period.

In passing from (10.1) to the corresponding result for derivatives in L^2, the values of q,α,m are, however, halved. It is therefore necessary in them to suppose now $0 \leq \alpha < \frac{1}{2}$, and q can be a half-integer, while m is now 1 or 2. This is one place where the artificialities of the definitions of fractional derivatives and differences becomes apparent, since for $\alpha = 1$ the expression used to define $\Delta_h^{q-\alpha}f$ leads, not to $\Delta_h^{q-1}f$, but to $h^{-1}\Delta_h^q f_1$, where f_1 is the indefinite integral of f.

§11. (Added by L. C. Young) *Stochastic integrals.* Let $X(t)$ be Hilbert-valued, let ΔX denote its difference at the ends of a t-interval Δ, and let $\kappa(\Delta,\Delta^*)$ denote the scalar product $(\Delta X, \Delta^*\bar{X})$. Here $\bar{X}$ is the complex conjugate of X : the vector-valued function $X(t)$ is understood to mean the same as the complex valued function $x(t,\omega)$, whose values for constant t are the points, or vectors, X of a Hilbert space $L^2(\Omega)$; and $\bar{X}(t)$ is understood to arise similarly from the complex conjugate $\bar{x}(t,\omega)$ of $x(t,\omega)$. We suppose $x(t,\omega)$ defined on $T\times\Omega$, where T is an interval of the real line, and Ω is a measure space with a unit measure $d\omega$. We speak of $X(t)$ as a stochastic process, and of $\kappa(\Delta,\Delta^*)$ as its covariance.

We shall suppose that, for intervals Δ, Δ^* of equal length $\leq h$, $|\kappa|$ has the supremum $\varphi^2(h)$ when Δ,Δ^* are non-overlapping, and the supremum $q^2(h)$ when $\Delta = \Delta^*$. By Schwarz's inequality $\varphi(h) \leq q(h)$, and we are mainly interested in the case where φ is "much less" than q, i.e. when $\varphi = o(q)$ as $h \to 0$. In that case $\Delta_h X(t)$ and $\Delta_h^2 X(t)$ are of the same order. In fact:

(11.1) <u>Theorem.</u> <u>If</u> $\Delta_h^2 X(t) = o(q)$, <u>then</u> φ,q <u>are of the same order.</u>

(11.2) <u>Corollary.</u> <u>If</u> $X(t)$ <u>has a non-vanishing derivative, then</u> $\varphi = q(1-o(1))$.

<u>Proof of (11.1).</u> We have

$$|\Delta_{2h}^{*2} X(t)|^2 = |\Delta_h X(t) - \Delta_{-h} X(t)|^2 .$$

Hence

$$|\Delta_h X(t)|^2 \leq o(q^2) + O(\varphi^2) - |\Delta_{-h} X(t)|^2 \leq o(q^2) + O(\varphi^2) ,$$

and, by taking the supremum in h, we derive (11.1) since $\varphi < q$.

<u>Proof of (11.2).</u> In this case $q^2 = Ah^2$, where $A = \operatorname{Sup}|\dot{X}(t)|^2 \neq 0$. In this case $|\Delta_{\pm h} X(t)|^2$ can be chosen simultaneously close to $A\,h^2$ and their sum is at most $o(h^2) + 2|\kappa|$, where κ is the scalar product of $\Delta_h X(t)$ and $\Delta_{-h} X(t)$. Thus $2Ah^2 \leq o(h^2) + 2\varphi^2(h)$, and (11.2) follows since $\varphi^2 \leq A\,h^2$.

The above results suggest that it is only for non-differentiable $X(t)$ that is is worthwhile to distinguish φ from q. For these reasons, the existence of a derivative cannot be expected to follow from any weaker assumptions than those considered in §6, which involve q and not φ. The question therefore arises as to whether the existence of stochastic integrals, such as those of [25, 26], can be derived directly from theorems such as those of §6, and not from hypothetical refinements involving the covariance of non-overlapping intervals.

For convenience we now write F in place of f in the theorems of §6, and we suppose F Hilbert-valued. Moreover we weaken the results by setting $n = 1$, $m = 1$. In the theorems referred to m was mainly 2 or 4, but their hypotheses will be satisfied a fortiori, apart from a constant K in (5.3). Actually we need to change (5.3) slightly in a different way. We shall suppose there are two pairs of estimate functions, φ, ψ and ρ, σ, and that, for $k \geq h > 0$,

$$|\Delta_h \Delta_k F(t)| \leq \rho(h)\,\sigma(k) + \varphi(h)\,\psi(k)\,. \tag{11.3}$$

Here moreover, σ and ψ will be supposed connected by the condition Λ of [26]. This ensures that a carefully thinned binary sequence for σ makes not only the ρ, σ estimate series of the same order as the integral $\int u^{-1}\rho d\sigma$, but also, at the same time, the φ, ψ estimate series of the same order as the integral $\int u^{-1}\varphi d\psi$. With these small changes, it is easy to derive from the arguments used to establish the theorems of §6 that a continuous $F(t)$ subject to (11.3) has a continuous derivative $\dot{F}(t)$ and that

$$|\dot{F}(t) - h^{-1}\Delta_h F(t)| \leq K\,S(h),$$

where $S(h)$ is now the sum of the integrals from 0 to h of $u^{-1}\rho\, d\sigma$ and $u^{-1}\varphi\, d\psi$.

Let now $X(t)$ be a stochastic process with φ and q defined as before, and let $\rho(u) = q(u)\sqrt{u}$. We denote by $f(t)$ a deterministic function, i.e. a complex-valued (scalar) function subject to the integrated Lipschitz conditions

$$\int_{T^+} |f(t+h) - f(t)|\, dt \leq \psi(h),$$

(11.4)

$$\int_{T^+} |f(t+h) - f(t)|^2 dt \leq \sigma^2(h),$$

and we write F for the convolution $f * X$. (Here T^+ is an interval with T in its interior.) We shall suppose the intervals of definition slightly expanded, if necessary, so that F is defined near 0 at least. In these circumstances, the stochastic integral

$$\int_T f(t)\, d\, X(t)$$

is defined as the derivative $-\dot{F}(0)$ of the function $-F(t)$. To show that this derivative exists, it is sufficient to verify (11.3) for the function F, and this turns out to be an elementary exercise in multiple integration.

REFERENCES

1. Bruneau, Michel, Fonctions d'une variable réele; caractérisation des ponts extrémaux de la boule unité de l'espace des fonctions à p-variation bornée. C.R. Acad. Sci. Paris Sér. A-B 274 (1972), 51-54.

2. ______________, Fonctions d'une variable réele, p-variation fine d'une fonction à p-variation bornée. C. R. Acad. Sci. Paris Sér. A-B 270 (1970), 585-588.

3. ______________, Fonctions d'une variable réele, fonctions p, α-fines et fonctions p-fines. C. R. Acad. Sci. Paris Sér. A-B 270 (1970), 1420-1423.

4. Burkill, J.C., Differential properties of the Young-Stieltjes integral, J. London Math. Soc. 23 (1948), 22-28.

5. Butzer, P. L. & Nessel, R.J., Fourier Analysis and approximation, Vol. I, Academic Press, 1971.

6. L. Carleson On convergence and growth of partial sums of Fourier series. Acta Math. 116 (1966), 135-157.

7. Erdélyi, A., Transformation of hypergeometric integrals by means of fractional integration by parts, Quart. J. Math. Oxford Serv., (1939), 176-189.

8. Fefferman, C., Pointwise convergence of Fourier series, Ann. of Math. 98, (1973), 551-571.

9. Gehring, F. W., A study in α-variation, Trans. Amer. Math. Scc., 76 , (1954), 420-443.

10. Landau, E., Über einen Satz des Herrn Littlewood, Rend. Circ. Matem. Palermo, XXXV, 1913.

11. Littlewood, J.E. & Paley, R.E.A., Theorems of Fourier series and power series III, Proc. London Math. Soc. Ser. 3 (1938), 105-126.

12. Love, E.R. & Young, L.C., Fractional integration by parts, Proc. London Math. Soc. Ser 2, 44, (1938), 1-35.

13. Lésniewicz, R. & Orliz, W., On generalized variation II, Studia Math. (1973), 71-109.

14. Marchaud, A., Sur les dérivées et sur les différences des fonctions des variables réeles. J. Math.pures Appl. 9 (1927), 337-425.

15. Musielak, J. & Orlicz, W., On generalized Variation I, Studia Math. 18 (1959), 11-41.

16. Orlicz, W., Généralization d'un théorème de Banach et Mazur Ann. Soc. Pol. Math. 9 (1946-47), 62-65.

17. Timan, A.F., Theory of approximation of functions of a real variable, Translated by J. Berry. The Macmillan Company, 1963.

18. Wiener, N., The Fourier Integral and certain of its applications, Cambridge, 1933.

#1578

19. Young, L.C., An inequality of the Hölder type connected with Stieltjes integration, Acta Math. 67 (1936), 251-282.

20. ____________, Inequalities connected with bounded p-th power variation in the Wiener sense and with integrated Lipschitz conditions, Proc. London Math. Soc. Ser. 2 43, (1937), 449-467.

21. ____________, Integrales de Stieltjes et convergence des séries de Fourier, Soc. Math. de France, conférences de la Réunion internat. de Math. Paris (1937).

22. ____________, Inequalities connected with binary derivation, integrated Lipschitz conditions and Fourier series, Math. Z. 43 (1938), 255-270.

23. ____________, General inequalities of Stieltjes integrals and the convergence of Fourier series, Math. Ann. 115 (1938), 581-612.

24. ____________, A further inequality for Stieltjes integrals, J. London Math. Soc. 18 (1943), 78-82.

25. ____________, Some new stochastic integrals and Stieltjes integrals, Part I. Analogues of Hardy-Littlewood classes, Advances in probability, 1970, Vol. II, 161-240.

26. ____________, Some new stochastic integrals, part II, Nigh-Martingales, ibid, 1974, Vol. III, 101-178.

27. Young, L. C., A converse to some inequalities and approximations in the theory of Stieltjes and stochastic integrals; n-th derivatives. Studia Math., to appear.

28. Zygmund, A., Trigonometric series I,II. Cambridge University Press, 1968.

Current address of the first author:

Instituto de Mathematica
Universidade Estadual de Campinas
Campinas-Sp. Brasil

Current address of the second author:

Department of Mathematics
University of Wisconsin
Madison, Wisconsin 53706
U.S.A.