ON THE WALSH FUNCTIONS[1]

BY

N. J. FINE

TABLE OF CONTENTS

1. **Introduction.** The system of orthogonal functions introduced by Rademacher [6][2] has been the subject of a great deal of study. This system is not a complete one. Its completion was effected by Walsh [7], who studied some of its Fourier properties, such as convergence, summability, and so on. Others, notably Kaczmarz [3], Steinhaus [4], Paley [5], have studied various aspects of the Walsh system. Walsh has pointed out the great similarity between this system and the trigonometric system.

It is convenient, in defining the functions of the Walsh system, to follow Paley's modification. The Rademacher functions are defined by

$$(1.1)\qquad \begin{aligned} &\phi_0(x) = 1 \ \ (0 \leqq x < 1/2), && \phi_0(x) = -1 \ \ (1/2 \leqq x < 1), \\ &\phi_0(x+1) = \phi_0(x), && \phi_n(x) = \phi_0(2^n x) \qquad (n = 1, 2, \cdots). \end{aligned}$$

The Walsh functions are then given by

$$(1.2)\qquad \psi_0(x) \equiv 1, \qquad \psi_n(x) = \phi_{n_1}(x)\phi_{n_2}(x) \cdots \phi_{n_r}(x)$$

for $n = 2^{n_1} + 2^{n_2} + \cdots + 2^{n_r}$, where the integers n_i are uniquely determined by $n_{i+1} < n_i$. Walsh proves that $\{\psi_n(x)\}$ form a complete orthonormal set. Every periodic function $f(x)$ which is integrable in the sense of Lebesgue on (0, 1) will have associated with it a (Walsh-) Fourier series

$$(1.3)\qquad f(x) \sim c_0 + c_1\psi_1(x) + c_2\psi_2(x) + \cdots,$$

where the coefficients are given by

Presented to the Society, August 23, 1946; received by the editors February 13, 1948.

(1) Presented to the Faculty of the Graduate School of the University of Pennsylvania in partial fulfillment of the requirements for the degree of Doctor of Philosophy, 1946.

(2) Numbers in brackets refer to the bibliography at the end of this paper.

$$c_n = \int_0^1 \psi_n(x)f(x)dx \qquad (n = 0, 1, 2, \cdots). \tag{1.4}$$

We shall set

$$s_n(x) = s_n(x; f) = \sum_{k=0}^{n-1} c_k\psi_k(x). \tag{1.5}$$

For every real number x and for every non-negative integer n we define $\alpha_n = \alpha_n(x)$ and $\beta_n = \beta_n(x)$ by

$$\alpha_n = m\cdot 2^{-n} \leqq x < (m+1)2^{-n} = \beta_n. \tag{1.6}$$

Following Paley we may write

$$s_{2^n}(x) = \sum_{k=0}^{2^n-1} c_k\psi_k(x) = \int_0^1 f(t) \sum_{k=0}^{2^n-1} \psi_k(t)\psi_k(x)dt. \tag{1.7}$$

It is easily verified that the kernel is identically

$$D_{2^n}(x, t) \equiv \sum_{k=0}^{2^n-1} \psi_k(t)\psi_k(x) \equiv \prod_{r=0}^{n-1} (1 + \phi_r(t)\phi_r(x)). \tag{1.8}$$

The expression (1.8) vanishes except on the interval $\alpha_n \leqq t < \beta_n$ (mod 1)(3). On this interval it has the value 2^n. Hence (1.7) becomes

$$s_{2^n}(x) = 2^n \int_{\alpha_n}^{\beta_n} f(t)dt = \frac{F(\beta_n) - F(\alpha_n)}{\beta_n - \alpha_n}, \tag{1.9}$$

where $F(x)$ is an integral of $f(x)$. From (1.9) follows a result of Kaczmarz [3], namely, that $s_{2^n}(x) \to f(x)$ almost everywhere. In particular, we have Walsh's result that $s_{2^n}(x) \to f(x)$ at every point of continuity of $f(x)$. Other theorems, such as the uniform convergence of $s_{2^n}(x)$ in an interval of continuity of $f(x)$, follow readily from (1.9). We shall omit the statement of other modifications.

Coming now to the question of convergence of the full sequence of partial sums, Walsh proves several theorems in which bounded variation is assumed of the given function $f(x)$. The two main results are:

(A) *If $f(x)$ is of bounded variation, and if x_0 is a point of continuity or a dyadic rational, then $s_n(x_0)$ converges.*

(B) *If $f(x)$ is of bounded variation, and if x_0 is neither a dyadic rational nor a point of continuity, then $s_n(x_0)$ diverges.*

(It is assumed, of course, that $f(x)$ has been redefined properly at removable discontinuities.) Modifications of (A) and (B) are obtained by Walsh by means of the localization theorem for Walsh-Fourier series. The failure of the analogue of Dirichlet's Theorem, as evidenced by (B), is one of the striking

(3) In the future we shall omit the obvious qualification of periodicity when there is no risk of confusion.

differences between the $\{\psi_n\}$ and the trigonometric functions.

With regard to summability, Walsh treats only the $(C, 1)$ case, for which he establishes continuity of $f(x)$ at a point x_0 as a sufficient condition for the Walsh-Fourier series of $f(x)$ to be summable to $f(x_0)$ at x_0. He points out that summability may fail for functions of bounded variation at points of jump.

Using a theorem of Haar's, and the fact that the Lebesgue constants $L_n=\int_0^1 |D_n(x, t)|dt$ are unbounded, Walsh proves that for any given x_0, there exists a continuous function whose Fourier(4) series diverges at x_0.

Walsh's final theorem concerns itself with the uniqueness of Walsh (not necessarily Fourier) expansions:

If $S(x)\equiv a_0+a_1\psi_1(x)+a_2\psi_2(x)+\cdots$ *converges to zero uniformly except in the neighborhood of a finite number of points, then* $a_n=0$ *for all* n.

The preceding has been a summary of most of the known results relevant to this paper, to the best of our knowledge. It is our purpose here to extend these results in various directions. We now give a brief résumé of the contents of the sections which follow.

In §2 we define the dyadic group G and discuss its connection with the Walsh system. A correspondence between G and the reals is set up, under which the Walsh functions may be considered the character group of G. The properties of this correspondence are studied, and it is prove that the Lebesgue integral is invariant under the transform of the group operation by the correspondence. More precisely, if $\dotplus$ denotes the group operation, μ the mapping of the reals into G, and λ the inverse mapping, then for every integrable function $f(x)$ we have

$$\int_0^1 f[\lambda(\mu(x) \dotplus \mu(a))]dx = \int_0^1 f(x)dx.$$

One application of this theorem is given in §2 (Theorem II), and many others in the later sections.

In §3 we develop the expansion of the functions $x-[x]$ and $J_k(x) =\int_0^x\psi_k(u)du$. We show that the latter have a restricted orthogonality (Theorem III).

The next section deals with the analogues of certain theorems in TFS(4), concerning the order of the Fourier coefficients. An important difference appears at this point. We prove (Theorem VIII) that for every nonconstant absolutely continuous function, $c_k\neq o(1/k)$.

A fairly complete discussion of the Lebesgue constants is given in §5. It is proved that $L_k=O(\log k)$ and that $(1/n)\sum_{k=1}^n L_k=(4 \log 2)^{-1} \log n+O(1)$. An explicit formula for L_k is given, as well as recursion formulas and a rather interesting generating function.

(4) Unless otherwise stated, by "Fourier series," "Fourier coefficients," and so on, we shall mean "Walsh-Fourier." We shall abbreviate "Walsh-Fourier series" as WFS and "trigonometric Fourier series" as TFS.

In §6 and §7 we develop some of the theory of convergence and summability; we also give another proof that the WFS of a continuous function $f(x)$ is $(C, 1)$ summable to $f(x)$ uniformly (Theorem XVII).

In §8 we consider a theorem due to Kac [1] concerning the set of functions

$$\mu_n(t) = 2^n t - [2^n t] - 1/2,$$

and compare this theorem with one we obtain concerning the functions

$$\lambda_n(t) = 2^n t - [2^n t + 1/2].$$

Since $\mu_n(t) = \mu_0(2^n t)$ and $\lambda_n(t) = \mu_0(2^n t + 1/2)$, one might expect the sums

$$S_N = \sum_{n=0}^{N} \mu_n(t), \qquad S_N^* = \sum_{n=0}^{N} \lambda_n(t)$$

to behave somewhat alike, but this is not the case. Kac's theorem states that the measure of the set of t such that $\alpha < (1/N^{1/2})S_N < \beta$ tends to the Gaussian integral

$$\frac{1}{\sigma(2\pi)^{1/2}} \int_\alpha^\beta e^{-x^2/2\sigma^2} dx, \qquad \sigma = 1/2,$$

whereas we prove that S_N^* is *uniformly bounded.*

The subjects of uniqueness and localization for general Walsh series are taken up in §9. The methods are similar to those of Riemann in trigonometric series, with the important difference that all of our results can be obtained without the use of the second formal integral. Among the main results are the following:

THEOREM XXIX. *If*

$$S(x) \equiv \sum_{k=0}^{\infty} a_k \psi_k(x)$$

converges to zero except perhaps on a denumerable set, then for all $k \geqq 0$, $a_k = 0$.

THEOREM XXVIII. *If* $S(x)$ *converges everywhere, and if the limit* $f(x)$ *is integrable, then* $S(x)$ *is the* WFS *of* $f(x)$.

The unwelcome "everywhere" condition can be somewhat weakened, but we have not been able to prove that an arbitrary denumerable set can be neglected.

We have also arrived at these results via the second integral, but here too the methods used do not suffice to prove much more than Theorem XXVIII. For this reason we shall present only the development by means of the first integral.

We wish to express our deep appreciation to Professor A. Zygmund for his kind encouragement and many helpful suggestions during the course of this work.

2. **The dyadic group.** G may be defined as the infinite direct product(5) of the group with elements 0 and 1, in which the group operation is addition modulo 2. Thus the dyadic group G is the set of all 0, 1 sequences, $\bar{t}=\{t_n\}$, $\bar{u}=\{u_n\}$, and so on, in which the group operation, which we shall denote by $\dotplus$, is addition modulo 2 for each element. Corresponding to each element $\{t_n\}$ of G there is a real number

$$t = \frac{t_1}{2} + \frac{t_2}{2^2} + \cdots$$

lying in the closed interval $[0, 1]$. This correspondence is not one-to-one, since the dyadic rationals have two representations in the dyadic scale. Even if we agree to use only the finite representation for the dyadic rationals, the ambiguity cannot be dispensed with entirely; for example,

$$\{0, 1, 0, 1, 0, 1, \cdots\} \dotplus \{1, 0, 1, 0, 1, 0, \cdots\} = \{1, 1, 1, 1, \cdots\},$$

and the undesired representations intrude anyway. Nevertheless, the connection between the elements of G and the real numbers modulo 1 has heuristic importance; it may be made exact enough for many purposes.

It will be recalled that a representation of a group G is a homomorphic mapping of G onto a group of matrices M: the group operation is carried over into multiplication in M. The *character* of a representation is the trace of the matrix into which a variable element of G is carried. Thus a character is a complex-valued function defined on G. Since the dyadic group G which we consider here is commutative, and each element is of order two, the characters reduce to functions on G taking only the values $1, -1$. The simplest characters that come to mind immediately may be defined by singling out a particular index n and defining $\chi_n\{t_1, t_2, \cdots\}$ as $+1$ if $t_n=0$, -1 if $t_n=1$. Clearly $\chi_n(\bar{t}\dotplus\bar{u})=\chi_n(\bar{t})\chi_n(\bar{u})$, so χ_n is a character. Since the product of a finite number of characters is also a character, we have the set $\chi_{n_1}\cdot\chi_{n_2}\cdots\chi_{n_\nu}$ to consider. We shall now show that every character χ may be represented in this way. Let $\chi(\bar{t})$ be any character. Then

$$\chi(\bar{0}) \equiv 1 \quad \textit{and} \quad \chi(\bar{t} \dotplus \bar{u}) = \chi(\bar{t})\cdot\chi(\bar{u}).$$

Write $\bar{t}=\{t_1, t_2, t_3, \cdots\}$ as $\{0, t_2, t_3, \cdots\}\dotplus\{t_1, 0, 0, \cdots\}=\bar{g}_1\dotplus\bar{h}_1$. Then $\chi(\bar{t})=\chi(\bar{g}_1)\cdot\chi(\bar{h}_1)$. Continuing in this way, we have $\chi(\bar{t})=\chi(\bar{h}_1)\cdot\chi(\bar{h}_2)\cdots\chi(\bar{h}_m)\cdot\chi(\bar{g}_m)$ where $\bar{g}_m=\{0, 0, \cdots, 0, t_{m+1}, t_{m+2}, \cdots\}$. The sequence $\bar{g}_m$ converges to the identity in the topology of G, and χ is continuous. Hence, from

(5) See L. Pontrjagin, *Topological groups*, Princeton University Press, 1939, for the theory discussed in this section.

some point on, $\chi(\bar{g}_m)$ must be identically 1. Thus we have achieved a factorization

$$\chi(\bar{t}) = \chi(\bar{h}_1) \cdots \chi(\bar{h}_m).$$

Now $\chi(\bar{h}_n)$, considered as a function of $\bar{t}$, say $\chi^{(n)}(\bar{t})$, is also a character, and is easily identified with either the identity or $\chi_n(\bar{t})$. Hence, every character may be written, by suppressing those which are identically 1, as

$$\chi(\bar{t}) = \chi_{n_1}(\bar{t}) \cdot \chi_{n_2}(\bar{t}) \cdots \chi_{n_\nu}(\bar{t}).$$

Thus we have obtained all the characters of G. These characters form a group X, in which the group operation is ordinary multiplication. It is well known that G is also the character group of X.

Suppose that we define a real-valued function $r_n(t)$ on the reals as follows: Write:

$$t = I + \sum_{n=1}^{\infty} \frac{t_n}{2^n}, \quad I \textit{ an integer}, \quad t_n = 0, 1.$$

If there is any ambiguity, choose the *finite* dyadic expansion of t. Now t corresponds to $\bar{t} = \{t_1, t_2, \cdots\} \in G$, and we write $r_n(t) \equiv \chi_n(\bar{t})$. It is easily seen that $r_n(t)$ is identical with the Rademacher function $\phi_{n-1}(t)$ defined in §1. Finally, the Walsh system may be associated, in like manner, with the full set of characters of G.

Just as the additive group of real numbers modulo 2π is fundamental in the theory of trigonometric series and classical Fourier series, so the group G is the underlying basis in the study of Walsh series and Walsh-Fourier series. We believe the analogy to be a very far-reaching one and worthy of further consideration. For the immediate purpose, however, we shall limit ourselves to a discussion of the relationships which exist between the group operation in G and the ordinary processes of analysis.

Let $\bar{x}$ be an element of G, $\bar{x} = \{x_1, x_2, \cdots\}$, $x_n = 0, 1$. We define the function

$$\lambda(\bar{x}) = \sum_{n=1}^{\infty} 2^{-n} x_n. \tag{2.1}$$

The function λ, which maps G onto the closed interval $[0, 1]$, does not have a single-valued inverse on the dyadic rationals; we shall agree to take the finite expansion in that case. Thus for all real x, if we write the inverse as μ,

$$\lambda(\mu(x)) = x - [x], \tag{2.2}$$

$[x]$ denoting the greatest integer in x. It is not necessarily true that $\mu(\lambda(\bar{x})) = \bar{x}$ for all $\bar{x} \in G$. It is true, however, provided that $\lambda(\bar{x})$ is not a dyadic rational (D.R.).

The function $\lambda(\bar{x})$ is non-negative and vanishes only for the identity $\bar{0}$. Since $\bar{y}\dot{+}\bar{z}=\bar{0}$ is true if and only if $\bar{y}=\bar{z}$, we have

$$\lambda(\bar{y} \dot{+} \bar{z}) = 0 \quad \textit{if and only if} \quad \bar{y} = \bar{z}. \tag{2.3}$$

Since $|y_n-z_n| = y_n+z_n$ (mod 2) for y_n, $z_n=0$ or 1 independently,

$$\bar{y} \dot{+} \bar{z} = \{ | y_n - z_n | \}. \tag{2.4}$$

Hence

$$\lambda(\bar{y} \dot{+} \bar{z}) \leqq \lambda(\bar{y} \dot{+} \bar{w}) + \lambda(\bar{w} \dot{+} \bar{z}). \tag{2.5}$$

Trivially $\lambda(\bar{y}\dot{+}\bar{z})=\lambda(\bar{z}\dot{+}\bar{y})$. This, together with (2.3) and (2.5), shows that λ may be used to define a metric on G. Now if $y_n' = y_n$ and $z_n' = z_n$ for $n<N$,

$$| \lambda(\bar{y}' \dot{+} \bar{z}') - \lambda(\bar{y} \dot{+} \bar{z}) | \leqq 2^{-N} + \cdots \leqq 2^{-N+1},$$

so the metric thus established is continuous on G. Conversely, if $\bar{y}$ and $\bar{z}$ are close in the metric, say $\lambda(\bar{y}\dot{+}\bar{z})<2^{-N}$, then they cannot differ in the nth place for $n\leqq N$, so they are close in G. Hence the topology induced on G by the metric is equivalent to the original topology.

An important property of the function λ is the following:

$$| \lambda(\bar{y}) - \lambda(\bar{z}) | \leqq \lambda(\bar{y} \dot{+} \bar{z}). \tag{2.6}$$

This follows from the fact that the left-hand member is

$$\left| \sum (y_n - z_n) 2^{-n} \right|$$

and the right-hand member, from (2.4), is

$$\sum | y_n - z_n | 2^{-n}.$$

We have proved above that $\phi_{n-1}(x)=\chi_n(\mu(x))$, $n\geqq 1$. We shall abbreviate $\lambda(\mu(y)\dot{+}\mu(z))$ as $y\dot{+}z$; there should be no confusion between the operation $\dot{+}$ in G and in the reals. Using this notation, it is possible to put (2.6) in a somewhat more convenient form. Let x and h be any two real numbers, and write $\mu(x)=\bar{y}$, $\mu(x)\dot{+}\mu(h)=\bar{z}$. (2.6) then becomes

$$| (x \dot{+} h) - \lambda(\bar{y}) | \leqq \lambda(\bar{y} \dot{+} \bar{z}) = \lambda(\mu(h)).$$

Taking (2.2) into account, we have

$$| (x \dot{+} h) - (x - [x]) | \leqq h - [h]. \tag{2.7}$$

In particular, if $0\leqq x<1$, $0\leqq h<1$,

$$| (x \dot{+} h) - x | \leqq h. \tag{2.8}$$

Now suppose that $f(x)$ is a continuous periodic function, with modulus of continuity $\omega(\delta; f)$, that is,

$$|f(x_1) - f(x_2)| \leqq \omega(\delta; f) \quad \textit{for} \quad |x_1 - x_2| \leqq \delta. \tag{2.9}$$

Then if $0 \leqq h \leqq \delta < 1$,

$$|f(x \dotplus h) - f(x)| \leqq \omega(\delta; f). \tag{2.10}$$

We shall next show that for each fixed y and for all z outside a certain denumerable set (depending on y), the equation

$$\phi_{n-1}(y \dotplus z) = \phi_{n-1}(y)\phi_{n-1}(z) \tag{2.11}$$

is valid. The right-hand side of (2.11) is equal to

$$\chi_n(\mu(y)) \cdot \chi_n(\mu(z)) = \chi_n(\mu(y) \dotplus \mu(z));$$

the left-hand side is

$$\chi_n(\mu(y \dotplus z)) = \chi_n(\mu(\lambda(\bar{x}))), \qquad \bar{x} = \mu(y) \dotplus \mu(z).$$

Now $\mu(\lambda(\bar{x})) = \bar{x}$ provided that $\lambda(\bar{x})$ is not a dyadic rational. Hence, if $y \dotplus z$ is not a dyadic rational, (2.11) holds. For fixed y, the exceptional set of z's consists of those whose dyadic expansions agree with that of y from some point on, or differ from some point on, and this set is clearly denumerable. From the definition of ψ_n, we have also (with the same restriction)

$$\psi_n(y \dotplus z) = \psi_n(y)\psi_n(z). \tag{2.12}$$

We turn now to considerations of measure. Let y be a fixed real number, and let z belong to a measurable set A lying in the unit interval. By $T_y(A)$ we shall mean the set $y \dotplus z$, $z \in A$. We propose to prove that T_y is a measure-preserving transformation, that is, $|T_y(A)| = |A|$. Consider the intervals

$$I(k, n):\ k \cdot 2^{-n} \leqq z < (k+1)2^{-n}, \qquad n \geqq 0, \qquad k = 0, 1, \cdots, 2^n - 1.$$

Clearly, except possibly for a denumerable set, T_y carries $I(k, n)$ into some $I(k^*, n)$, where k^* may be identical with k. Thus our statement is true for all dyadic intervals $I(k, n)$. It follows that it is true for all intervals, for we can cover any interval J with a sum Σ of non-overlapping dyadic intervals such that $|\Sigma| < |J| + \epsilon$; then $|T_y(J)| \leqq |T_y(\Sigma)| = |\Sigma| < |J| + \epsilon$, so that $|T_y(J)| \leqq |J|$; hence $|T_y(J)| = |J|$. The truth of our statement now follows for any measurable set A.

As a consequence, the function $f(y \dotplus z)$ (y fixed) is equimeasurable with $f(z)$. We have therefore proved the following theorem:

THEOREM I. *If $f(z)$ is integrable, then for every fixed y,*

$$\int_0^1 f(y \dotplus z)dz = \int_0^1 f(z)dz.$$

An example of the usefulness of this result is the following theorem, which we shall need later:

THEOREM II. *If $f(x)$ has the Fourier coefficients $\{c_n\}$, and if h is any fixed number, then the Fourier coefficients of $f(x\dot{+}h)$ are $\{c_n\psi_n(h)\}$.*

By (2.12), $\psi_n(x)=\psi_n(h)\psi_n(x\dot{+}h)$ a.e., so

$$\int_0^1 \psi_n(x)f(x \dot{+} h)dx = \psi_n(h)\int_0^1 \psi_n(x \dot{+} h)f(x \dot{+} h)dx$$

$$= \psi_n(h)\int_0^1 f(x)\psi_n(x)dx$$

$$= c_n\psi_n(h).$$

3. **Expansions of certain functions.** This section will deal with the properties of the integrals

$$J_k(x) = \int_0^x \psi_k(t)dt, \qquad k = 0, 1, 2, \cdots . \tag{3.1}$$

We have $J_0(x)=x$. We may write, using the finite expansion in case of doubt,

$$x - [x] = \sum_{n=1}^{\infty} 2^{-n}x_n. \tag{3.2}$$

It is easy to see that

$$x_{n+1} = \frac{1 - \phi_n(x)}{2}, \qquad n = 0, 1, 2, \cdots . \tag{3.3}$$

Hence

$$\begin{aligned} x - [x] &= \frac{1}{2} - \frac{1}{4}\sum_{n=0}^{\infty} 2^{-n}\phi_n(x) \\ &= \frac{1}{2} - \frac{1}{4}\sum_{n=0}^{\infty} 2^{-n}\psi_{2^n}(x). \end{aligned} \tag{3.4}$$

Since the last series converges uniformly, it is the Fourier series of the function $x-[x]$.

For $k\geqq 1$, write $k=2^n+k'$, where $0\leqq k'<2^n$. Let $I(p, n)$ denote the interval $p\cdot 2^{-n}\leqq x<(p+1)2^{-n}$. The function $\psi_k(x)$ is constant on each interval $I(p, n+1)$, and $\psi_{k'}(x)$ on each $I(p, n)$. Since $\psi_k(x)=\psi_{2^n}(x)\psi_{k'}(x)$ and $\psi_{2^n}(x)$ changes sign between $I(2p, n+1)$ and $I(2p+1, n+1)$, so does $\psi_k(x)$. It follows that $J_k(p\cdot 2^{-n})=0$; furthermore, on each $I(p, n)$, $J_k(x)$ is a reproduction of J_1 with reduced scale and a possible change of sign. It is easily verified that

$$J_k(x) = 2^{-n}\psi_{k'}(x)J_1(2^n x) = \psi_{k'}(x)J_{2^n}(x). \tag{3.5}$$

Now $J_1(x)=1/2+\psi_1(x)\cdot(x-[x]-1/2)$. Using (3.4), we have

$$J_1(x) = \frac{1}{2} + \psi_1(x)\left(-\frac{1}{4}\sum_{r=0}^{\infty} 2^{-r}\psi_{2^r}(x)\right)$$
$$= \frac{1}{4} - \frac{1}{4}\sum_{r=1}^{\infty} 2^{-r}\psi_{2^r+1}(x).$$

From (3.5), with $k=2^n$, $k'=0$,

$$J_{2^n}(x) = 2^{-(n+2)}\left\{1 - \sum_{r=1}^{\infty} 2^{-r}\psi_{2^r+1}(2^n x)\right\}$$
$$= 2^{-(n+2)}\left\{1 - \sum_{r=1}^{\infty} 2^{-r}\psi_{2^{n+r}+2^n}(x)\right\}.$$

In general, therefore,

$$J_k(x) = 2^{-(n+2)}\left\{\psi_{k'}(x) - \sum_{r=1}^{\infty} 2^{-r}\psi_{2^{n+r}+2^n+k'}(x)\right\}$$
$$= 2^{-(n+2)}\left\{\psi_{k'}(x) - \sum_{r=1}^{\infty} 2^{-r}\psi_{2^{n+r}+k}(x)\right\}. \tag{3.6}$$

We shall now obtain another useful formula for $J_k(x)$. Let us define α_n, β_n by (1.6), and let $\gamma_n=\gamma_n(x)$ be that one of α_n, β_n which is nearer to x; if x is the midpoint of (α_n, β_n), set $\gamma_n=\beta_n$. With this definition we find that

$$J_{2^n}(x) = \pm (x - \gamma_n),$$

and the sign is determined as $+$ if x lies in the left half of (α_n, β_n), $-$ if x lies in the right half. Hence

$$J_{2^n}(x) = \psi_{2^n}(x)(x - \gamma_n), \tag{3.7}$$

$$J_k(x) = \psi_k(x)(x - \gamma_n). \tag{3.8}$$

It may be of interest to note that the $J_k(x)$ exhibit a restricted orthogonality, as shown by Theorem III. We shall not give the proof of this theorem here, since it is somewhat involved, but shall merely remark that it depends on the Fourier expansion of $J_k(x)$ given by (3.6). In order to avoid a discussion of cases, we shall adopt the following definitions:

DEFINITIONS. If k is a non-negative integer, we define the set of integers (k', k'', n_1, n_2) by

(i) $[2^{n_1}]\leq k<[2^{n_1+1}]$,
(ii) $k'=k-[2^{n_1}]$,
(iii) $[2^{n_2}]\leq k'<[2^{n_2+1}]$,
(iv) $k''=k'-[2^{n_2}]=k-[2^{n_1}]-[2^{n_2}]$.

Let the set (s', s'', m_1, m_2) correspond to s, and let $\delta(a, b)=0$ if $a\neq b$, $\delta(a, b)=1$

if $a=b$. With these conventions we may now state our theorem.

THEOREM III. *If k and s are any non-negative integers, then*

$$\int_0^1 J_k(x)J_s(x)dx = \frac{1}{16}\left(1+\frac{1}{3}\delta(k, s)\right)(2\delta(k', s') - 1)F(k, s),$$

where

$$\begin{aligned} F(k, s) = {} & 2^{-(n_1+n_2)}\delta(k'', s'')\delta(k'', s) \\ & + 2^{-(n_1+m_1)}\delta(k', s')\delta(k', s'')(1-\delta(k', s)) \\ & + 2^{-(m_1+m_2)}\delta(k, s'')(1-\delta(k, s')). \end{aligned}$$

4. **Fourier coefficients.** Walsh has proved that the Fourier coefficients of an integrable function converge to zero. In this section we shall discuss the rapidity of convergence for several classes of functions. For all theorems in this section we shall assume that $f(x)$ is periodic, integrable, and has the Fourier expansion

$$f(x) \sim c_0 + c_1\psi_1(x) + c_2\psi_2(x) + \cdots .$$

THEOREM IV. *Let*

$$\omega(\delta; f) = \underset{|h|\leqq\delta,\, 0\leqq x<1}{\text{l.u.b.}} |f(x+h) - f(x)|,$$

$$\omega_1(\delta; f) = \underset{|h|\leqq\delta}{\text{l.u.b.}} \int_0^1 |f(x+h) - f(x)|\, dx.$$

Then for $k>0$,
(i) $|c_k| \leqq \omega(1/k; f)/2$,
(ii) $|c_k| \leqq \omega_1(1/k; f)$.

Suppose that $2^n \leqq k < 2^{n+1}$. Then, by Theorem II,

$$f(x \dot{+} 2^{-(n+1)}) \sim \sum_{k=0}^{\infty} c_k\psi_k(2^{-(n+1)})\psi_k(x).$$

Hence the kth Fourier coefficient of $f(x) - f(x\dot{+}2^{-(n+1)})$ is $c_k(1-\psi_k(2^{-(n+1)}))$. Now

$$\psi_k(2^{-(n+1)}) = \psi_{2^n}(2^{-(n+1)})\psi_{k'}(2^{-(n+1)}),$$

where $k'<2^n$. For $m<n$, $\psi_{2^m}(2^{-(n+1)}) = \psi_1(2^{m-n-1}) = 1$, so $\psi_{k'}(2^{-(n+1)}) = 1$. But $\psi_{2^n}(2^{-(n+1)}) = \psi_1(1/2) = -1$. Hence $\psi_k(2^{-(n+1)}) = -1$. Thus

$$2c_k = \int_0^1 \psi_k(x)\{f(x) - f(x \dot{+} 2^{-(n+1)})\}dx.$$

Using (2.10), we find

$$\left| c_k \right| = \frac{1}{2} \left| \int_0^1 \psi_k(x) \{ f(x) - f(x \dotplus 2^{-(n+1)}) \} dx \right|$$
$$\leqq \omega(2^{-(n+1)}; f)/2 \leqq \omega(1/k; f)/2.$$

To prove (ii), we observe that $x \dotplus 2^{-(n+1)}$ takes the values $x + 2^{-(n+1)}$, $x - 2^{-(n+1)}$ on two mutually disjoint sets E_+ and E_-, each of measure 1/2. Hence

$$\frac{1}{2} \int_0^1 \left| f(x) - f(x \dotplus 2^{-(n+1)}) \right| dx$$
$$= \frac{1}{2} \int_{E+} \left| f(x) - f(x + 2^{-(n+1)}) \right| dx + \frac{1}{2} \int_{E_-} \left| f(x) - f(x - 2^{-(n+1)}) \right| dx$$
$$\leqq \int_0^1 \left| f(x) - f(x + 2^{-(n+1)}) \right| dx \leqq \omega_1(2^{-(n+1)}; f)$$
$$\leqq \omega_1(1/k; f).$$

As an immediate corollary, we obtain the following theorem:

THEOREM V. *If $f(x)$ satisfies a Lipschitz condition of order α, $0 < \alpha \leqq 1$, then $c_k = O(k^{-\alpha})$.*

Our next theorem, dealing with functions of bounded variation, has an exact analogue in the classical theory of Fourier series.

THEOREM VI. *If $f(x)$ is of bounded variation, and V is its total variation over* (0, 1), *then $|c_k| \leqq V/k$ for $k > 0$.*

Write

$$f(x) = P(x) - N(x) + f(0),$$

where $P(x)$ and $N(x)$ are the positive and negative variations of $f(x)$; each is non-negative and nondecreasing. Applying the second mean-value theorem,

$$\int_0^1 \psi_k(x) P(x) dx = P(1 - 0) \int_\xi^1 \psi_k(x) dx = - P(1 - 0) J_k(\xi);$$

similarly,

$$\int_0^1 \psi_k(x) N(x) dx = - N(1 - 0) J_k(\xi').$$

Now by (3.8),

$$\left| J_k(x) \right| = \left| \psi_k(x)(x - \gamma_n) \right| \leqq 2^{-(n+1)} < 1/k;$$

it follows that

$$|c_k| \leqq \frac{P(1-0)+N(1-0)}{k} = \frac{V}{k}.$$

We now come to an important difference between trigonometric Fourier series and Walsh-Fourier series. Let $f(x)$ have mean-value zero over $(0, 2\pi)$. Then the periodic function

$$F(x) = \int_0^x f(t)dt$$

has (trigonometric) Fourier coefficients B_k satisfying

$$B_k = o(1/k).$$

If $F(x)$ is a pth integral, then $B_k = o(k^{-p})$. We shall now show that the order of the coefficients of a WFS *cannot* be improved by smoothness conditions.

THEOREM VII. *Let*

$$f(x) \sim c_1\psi_1(x) + c_2\psi_2(x) + \cdots,$$

$$F(x) = \int_0^x f(t)dt \sim b_0 + b_1\psi_1(x) + b_2\psi_2(x) + \cdots.$$

Then for fixed $k' \geqq 0$ *and* $n \to \infty$,

(i) $$b_{2^n+k'} = -2^{-(n+2)}c_{k'} + o(2^{-n}).$$

Write $k = 2^n + k'$, and take $2^n > k'$. Then, integrating by parts,

$$b_{2^n+k'} = \int_0^1 \psi_k(x)F(x)dx = \left[J_k(x)F(x)\right]_0^1 - \int_0^1 J_k(x)f(x)dx.$$

The first term on the right vanishes; using the Fourier expansion obtained in (3.6), we have

$$\begin{aligned} b_{2^n+k'} &= -2^{-(n+2)}\int_0^1 \psi_{k'}(x)f(x)dx + 2^{-(n+2)}\sum_{r=1}^{\infty} 2^{-r}\int_0^1 \psi_{2^{n+r}+k}(x)f(x)dx \\ &= -2^{-(n+2)}c_{k'} + 2^{-(n+2)}\sum_{r=1}^{\infty} 2^{-r}c_{2^{n+r}+k}. \end{aligned}$$

The series is dominated by

$$\max_{p>2^n} |c_p|,$$

which tends to zero as $n \to \infty$. This proves (i).

THEOREM VIII. *The only absolutely continuous functions whose Fourier coefficients satisfy* $b_k = o(1/k)$ *are the constants.*

If there is one k' for which $c_{k'} \neq 0$, then

$$-2^n b_{2^n+k'} = c_{k'}/4 + o(1) \neq o(1).$$

But if all $c_{k'}$ vanish, $f(x) \equiv 0$.

The theorem just proved is heuristically reasonable: if the coefficients converge to zero too rapidly, then the jumps of the functions $\psi_k(x)$ cannot be smoothed out in time. We might suspect, then, the existence of restrictions on the order of Fourier coefficients for arbitrary continuous functions. For example, if $\sum_{k=2}^{\infty} |b_k| < |b_1|$, the discontinuity at the point $x = 1/2$ must remain. The question seems worthy of further study, although we shall not pursue it here.

5. **The Lebesgue constants.** Let $\{f_k(x)\}$, $k \geqq 0$, be an arbitrary orthonormal system on an interval (a, b). By analogy with TFS, we may define the "Dirichlet kernel"

$$D_k(x, u) = f_0(x)f_0(u) + \cdots + f_{k-1}(x)f_{k-1}(u). \tag{5.1}$$

The *Lebesgue functions* of the system are then given by

$$L_k(x) = \int_a^b |D_k(x, u)| du. \tag{5.2}$$

If, as in TFS, the functions $L_k(x)$ are independent of x, we speak of the Lebesgue constants L_k. It is clear that the expressions (5.1) and (5.2) must play an important part in the theory of the system $\{f_k\}$. For example, the partial sums of the Fourier development of a given function $g(x)$ are

$$\begin{aligned} s_k(x; g) \equiv s_k(x) &= f_0(x) \int_a^b g(u)f_0(u)du + \cdots \\ &\quad + f_{k-1}(x) \int_a^b g(u)f_{k-1}(u)du \\ &= \int_a^b g(u)D_k(x, u)du. \end{aligned} \tag{5.3}$$

In this section we shall prove that the $L_k(x)$ for the system $\{\psi_k\}$ are independent of x; we shall evaluate the L_k explicitly and obtain recursion formulas for them; we shall determine their average order as well as their maximum order; finally, we shall derive a generating function for the L_k. It might be well to recall at this time the fact that in the trigonometric system we have

$$L_k = \frac{4}{\pi^2} \log k + O(1).$$

By (2.12), we have for fixed x and almost all u,

$$\psi_n(x)\psi_n(u) = \psi_n(x \dotplus u).$$

Hence, setting $D_k(0, u) \equiv D_k(u)$,

$$D_k(x, u) = \sum_{n=0}^{k-1} \psi_n(x)\psi_n(u) = D_k(x \dotplus u), \qquad (x \text{ fixed, a.a.}u).$$

$$L_k(x) = \int_0^1 | D_k(x \dotplus u) | \, du = \int_0^1 | D_k(u) | \, du.$$

Thus the $L_k(x)$ are in fact independent of x. Now let $k = 2^n + k'$, $0 \leqq k' < 2^n$, $n \geqq 0$. If we agree to set $D_0(u)$ equal to zero, we have

$$\begin{aligned} D_k(u) &= D_{2^n}(u) + \sum_{r=0}^{k'-1} \psi_{2^n+r}(u) \\ &= D_{2^n}(u) + \psi_{2^n}(u) \sum_{r=0}^{k'-1} \psi_r(u) \\ &= D_{2^n}(u) + \psi_{2^n}(u) D_{k'}(u). \end{aligned}$$

On the interval $(0, 2^{-(n+1)})$, $D_{2^n}(u) = 2^n$, $\psi_{2^n}(u) = 1$, and $D_{k'}(u) = k'$; on the interval $(2^{-(n+1)}, 2^{-n})$, $D_{2^n}(u) = 2^n$, $\psi_{2^n}(u) = -1$, and $D_{k'}(u) = k'$; on $(2^{-n}, 1)$, $D_{2^n}(u) = 0$. Hence

$$\begin{aligned} | D_k(u) | &= 2^n + k' & (0 \leqq u < 2^{-(n+1)}) \\ &= 2^n - k' & (2^{-(n+1)} \leqq u < 2^{-n}) \\ &= D_{k'}(u) & (2^{-n} \leqq u < 1). \end{aligned}$$

We find, therefore, that

$$\begin{aligned} L_k &= \int_0^{2^{-(n+1)}} + \int_{2^{-(n+1)}}^{2^{-n}} + \int_{2^{-n}}^1 | D_k(u) | \, du \\ &= 2^{-(n+1)}((2^n + k') + (2^n - k')) + \int_{2^{-n}}^1 | D_{k'}(u) | \, du \qquad (5.4) \\ &= 1 + L_{k'} - \int_0^{2^{-n}} | D_{k'}(u) | \, du \\ &= 1 + L_{k'} - k'/2^n. \end{aligned}$$

If $k = 2^{n_1} + 2^{n_2} + \cdots + 2^{n_\nu}$, $n_1 > n_2 > \cdots > n_\nu$, and $k^{(p)} = k^{(p-1)} - 2^{n_p}$, successive applications of (5.4) yield

$$(5.5) \quad L_k = \sum_{p=1}^{\nu} (1 - 2^{-n_p} k^{(p)}) = \nu - \sum_{p=1}^{\nu-1} 2^{-n_p} \sum_{r=p+1}^{\nu} 2^{n_r} = \nu - \sum_{1 \leqq p < r \leqq \nu} 2^{n_r - n_p}.$$

The effect of replacing k by $2k$ is to leave ν unaltered and to increase each

n_r by 1, thus leaving $n_r - n_p$ unchanged. Hence

$$L_{2k} = L_k. \tag{5.6}$$

If, however, we consider $2k+1$, we have $n_{\nu+1}=0$, so

$$\begin{aligned} L_{2k+1} &= \nu + 1 - \sum_{1 \leqq p < r \leqq \nu} 2^{(n_r+1)-(n_p+1)} - \sum_{p=1}^{\nu} 2^{n_{\nu+1}-(n_p+1)} \\ &= L_k + 1 - (2^{-n_1} + 2^{-n_2} + \cdots + 2^{-n_\nu})/2 \\ &= L_k + 1 - g(k)/2. \end{aligned} \tag{5.7}$$

We observe that $g(2k) = g(k)/2$, $g(2k+1) = 1 + g(2k)$. It is not difficult to prove, now, that

$$L_{2k+1} = (1 + L_k + L_{k+1})/2. \tag{5.8}$$

Let us define, for real x, the function

$$T(x) = \sum_{k=1}^{[x]} L_k; \qquad \text{if } x < 1, \qquad T(x) = 0.$$

We now need an estimate for L_k from (5.5):

$$L_k \leqq \nu.$$

Since $n_\nu \geqq 0$, $n_{\nu-1} \geqq 1, \cdots, n_1 \geqq \nu - 1$, we have $k \geqq 1 + 2 + 4 + \cdots + 2^{\nu-1} = 2^\nu - 1$, so

$$\nu \leqq \frac{\log(k+1)}{\log 2},$$

$$L_k = O(\log k). \tag{5.9}$$

Now if $x = 2m+1$, m a positive integer, we find

$$\begin{aligned} T(x) &= L_1 + \{L_2 + L_4 + \cdots + L_{2m}\} + \{L_3 + L_5 + \cdots + L_{2m+1}\} \\ &= L_1 + T\left(\frac{x}{2}\right) + \frac{1}{2}\{(L_1 + L_2 + 1) + (L_2 + L_3 + 1) + \cdots \\ &\qquad + (L_m + L_{m+1} + 1)\} \\ &= L_1 + T\left(\frac{x}{2}\right) + \frac{1}{2}T\left(\frac{x}{2}\right) + \frac{1}{2}\{L_2 + L_3 + \cdots + L_{m+1}\} + \frac{m}{2} \\ &= 2T\left(\frac{x}{2}\right) + \frac{1}{2}\left[\frac{x}{2}\right] + \frac{1}{2}(L_1 + L_{m+1}); \end{aligned}$$

$$T(x) = 2T\left(\frac{x}{2}\right) + \frac{x}{4} + O(\log x). \tag{5.10}$$

It is easy to see that (5.10) is valid for all $x>1$. If we set $A(x)=T(x)/x$ and apply (5.10) repeatedly, we obtain

$$
\begin{aligned}
A(x) &= A\left(\frac{x}{2}\right)+\frac{1}{4}+O\left(\frac{\log x}{x}\right),\\
A\left(\frac{x}{2}\right) &= A\left(\frac{x}{4}\right)+\frac{1}{4}+O\left(\frac{\log x/2}{x/2}\right),\\
&\cdots\cdots\cdots\cdots\cdots\cdots,\\
A\left(\frac{x}{2^{r-1}}\right) &= A\left(\frac{x}{2^r}\right)+\frac{1}{4}+O\left(\frac{\log x/2^{r-1}}{x/2^{r-1}}\right).
\end{aligned}
\tag{5.11}
$$

Summing (5.11), we have

$$
A(x) = A\left(\frac{x}{2^r}\right)+\frac{r}{4}+O\left(\frac{\log x}{x}+\cdots+\frac{\log x/2^{r-1}}{x/2^{r-1}}\right). \tag{5.12}
$$

Choose r so that $2^{r-1}<x\leqq 2^r$. Then (5.12) becomes

$$
A(x) = \frac{r}{4}+O(1) = \frac{\log x}{4\log 2}+O(1). \tag{5.13}
$$

At this point we should like to present a short proof of a theorem communicated to us by Professor Rademacher. This theorem was proved by him in an unpublished paper written in 1921, in which he effected the completion of the Rademacher functions. The theorem deals not with the *average* value of the Lebesgue constants but with their *maximal* value. We present our proof rather than Professor Rademacher's because of the simplification resulting from the use of the recurrence formulas for $L(k)$. (Throughout this proof we shall write $L(k)$ for L_k.)

THEOREM. $\text{Lim sup}_{k\to\infty}\ \{L(k)-(4/9+(1/3)\log_2 3k)\}=0$.

Proof. Define the sequence of integers

$$
t_0 = 1,\ t_1 = 3,\ \cdots,\ t_{n+1} = 2t_n + (-1)^n. \tag{5.14}
$$

Using the recurrence formulas

$$
\begin{aligned}
L(2k) &= L(k),\\
L(2k+1) &= (1+L(k)+L(k+1))/2,
\end{aligned}
\tag{5.15}
$$

it is easy to prove by induction (considering separately the cases n even and n odd) that

$$
L(t_{n+1}) = (1+L(t_n)+L(t_{n-1}))/2,
$$

from which it follows readily that

$$L(t_n) \geqq L(k) \qquad \text{for } 2^n \leqq k \leqq 2^{n+1}$$

and that

$$L(t_n) = \frac{n}{3} + \frac{10}{9} - \frac{1}{9}\left(-\frac{1}{2}\right)^n.$$

Therefore

$$\lim_{n\to\infty} \left\{L(t_n) - \left(\frac{10}{9} + \frac{n}{3}\right)\right\} = 0. \tag{5.16}$$

Now from (5.14),

$$t_n = \frac{4}{3}\,2^n - \frac{1}{3}(-1)^n,$$

and therefore

$$n = \log_2(3t_n) - 2 + o(1). \tag{5.17}$$

Combining (5.16) and (5.17), and defining

$$e(k) = L(k) - \left(\frac{4}{9} + \frac{1}{3}\log_2 3k\right), \tag{5.18}$$

we see that

$$\limsup_{k\to\infty} e(k) \geqq 0. \tag{5.19}$$

From (5.15) and (5.18),

$$e(2k) = e(k) - 1/3. \tag{5.20}$$

Also

$$\begin{aligned} e(2k+1) &= \frac{1}{2}(1 + e(k) + e(k+1)) - \frac{1}{6}\log_2\left\{\frac{(2k+1)^2}{k(k+1)}\right\} \\ &< \frac{1}{2}(e(k) + e(k+1)) + \frac{1}{6}. \end{aligned}$$

Considering the two cases separately,

$$\begin{aligned} e(4k+1) &< \frac{1}{2}(e(2k) + e(2k+1)) + \frac{1}{6} \\ &= \frac{1}{2}(e(k) + e(2k+1)), \end{aligned} \tag{5.21}$$

$$e(4k+3) < \frac{1}{2}(e(2k+1)+e(2k+2)) + \frac{1}{6} \tag{5.22}$$
$$= \frac{1}{2}(e(2k+1)+e(k+1)).$$

By direct computation we find that for $2 \leqq k \leqq 8$,

$$e(k) < 0. \tag{5.23}$$

Assuming that (5.23) is true for $2^{n-2} \leqq k \leqq 2^n$, (5.20), (5.21), and (5.22) assure us that it is true for $2^n \leqq k \leqq 2^{n+2}$ also. Hence $e(k) < 0$ for all $k > 1$ and

$$\limsup_{k\to\infty} e(k) \leqq 0. \tag{5.24}$$

Taking (5.19) and (5.24) together, and recalling (5.18), we see that the theorem is proved.

We shall now turn our attention to the generating function

$$F(z) = \sum_{k=1}^{\infty} L_k z^k \qquad (|z| < 1). \tag{5.25}$$

Equations (5.6), (5.8), and (5.25) yield the functional equation

$$F(z) = \frac{(1+z)^2}{2z} F(z^2) + \frac{1}{2}\,\frac{z}{1-z^2}\,. \tag{5.26}$$

Writing $F_0 = F(z)$, $F_n = F(z^{2^n})$, we have, for $n \geqq 0$,

$$F_n = \frac{(1+z^{2^n})^2}{2z^{2^n}} F_{n+1} + \frac{1}{2}\,\frac{z^{2^n}}{1-z^{2^{n+1}}}\,. \tag{5.27}$$

Now set

$$P_n = \frac{\{(1+z)(1+z^2)\cdots(1+z^{2^{n-1}})\}^2}{2^n z^{2^n-1}}, \qquad H_1 = \frac{1}{2}\,\frac{z}{1-z^2}\,.$$

We shall define H_n recursively so that the following equation, which is true for $n=1$, shall hold for $n \geqq 1$:

$$F_0 = P_n F_n + H_n. \tag{5.28}$$

If we substitute for F_n in (5.28) its value in (5.27), we obtain

$$F_0 = \frac{(1+z^{2^n})^2}{2z^{2^n}} P_n F_{n+1} + \frac{1}{2}\,\frac{z^{2^n}}{1-z^{2^{n+1}}} P_n + H_n \tag{5.29}$$
$$= P_{n+1}F_{n+1} + H_n + \frac{1}{2}\,\frac{z^{2^n}}{1-z^{2^{n+1}}} P_n.$$

Comparing (5.28) and (5.29), we see that the desired recursion is

$$H_{n+1} = H_n + \frac{1}{2}\,\frac{z^{2^n}}{1 - z^{2^{n+1}}}\,P_n, \qquad n \geqq 1. \tag{5.30}$$

Summing (5.30) for $n=1, 2, \cdots, N$, we have

$$H_{N+1} = \frac{1}{2}\left[\frac{z}{1-z^2} + \frac{z^2}{1-z^4}\cdot\frac{(1+z)^2}{2z} + \cdots + \frac{z^{2^N}}{1-z^{2^{N+1}}}\cdot\frac{\{(1+z)\cdots(1+z^{2^{N-1}})\}^2}{2^N z^{2^N-1}}\right]$$

$$= \frac{1}{2}\cdot\frac{z}{(1-z)^2}\left[\frac{1-z}{1+z} + \frac{1}{2}\,\frac{1-z^2}{1+z^2} + \cdots + \frac{1}{2^N}\,\frac{1-z^{2^N}}{1+z^{2^N}}\right].$$

Now $F(z)=O(z)$ for $|z|$ small, and $F_n(z)=O(z^{2^n})$; the numerator of P_n is bounded and the denominator exceeds $2^n|z|^{2^n}$ in absolute value. Therefore $\lim P_nF_n=0$, and

$$F(z) = \frac{1}{2}\,\frac{z}{(1-z)^2}\sum_{n=0}^{\infty}\frac{1}{2^n}\,\frac{1-z^{2^n}}{1+z^{2^n}}\,. \tag{5.31}$$

We collect the results of this section:

THEOREM IX. *The Lebesgue constants* L_k *of the Walsh system satisfy*

(i) $$L_k = \nu - \sum_{1\leqq p<r\leqq \nu} 2^{n_r - n_p},$$

where

$$k = 2^{n_1} + 2^{n_2} + \cdots + 2^{n_\nu} \qquad (n_1 > n_2 > \cdots > n_\nu \geqq 0).$$

(ii) $L_{2k} = L_k; \quad L_{2k+1} = (1 + L_k + L_{k+1})/2.$

(iii) $L_k = O(\log k).$

(iv) $$\frac{1}{n}\sum_{k=1}^{n} L_k = \frac{\log n}{4\log 2} + O(1).$$

(v) $$\limsup_{k\to\infty}\left\{L_k - \left(\frac{4}{9} + \frac{1}{3}\log_2 3k\right)\right\} = 0.$$

(vi) $$F(z) \equiv \sum_{k=1}^{\infty} L_k z^k = \frac{1}{2}\,\frac{z}{(1-z)^2}\sum_{n=0}^{\infty}\frac{1}{2^n}\,\frac{1-z^{2^n}}{1+z^{2^n}} \qquad (|z|<1).$$

6. **Convergence.** We first prove a lemma concerning the size of $D_k(u)$.

LEMMA 1. *For all u such that* $0<u<1$, $|D_k(u)| <2/u$.

Suppose that $2^{-n} \leqq u < 2^{-n+1}$, and write k in the form $k=p\cdot 2^n+q$, $0\leqq q \leqq 2^n$.

We assert that for all u,

$$D_k(u) = D_{2^n}(u)D_p(2^n u) + \psi_p(2^n u)D_q(u). \tag{6.1}$$

For $\psi_r(u)\psi_s(u) = \psi_{r+s}(u)$ provided that r and s have dyadic expansions without any exponents common to both; hence

$$\begin{aligned} D_k(u) &= \sum_{r=0}^{k-1} \psi_r(u) = \sum_{r=0}^{p-1} \sum_{t=0}^{2^n-1} \psi_{r\cdot 2^n+t}(u) + \sum_{t=0}^{q-1} \psi_{p\cdot 2^n+t}(u) \\ &= \sum_{r=0}^{p-1} \sum_{t=0}^{2^n-1} \psi_{r\cdot 2^n}(u)\psi_t(u) + \sum_{t=0}^{q-1} \psi_{p\cdot 2^n}(u)\psi_t(u) \\ &= D_{2^n}(u) \sum_{r=0}^{p-1} \psi_r(2^n u) + \psi_p(2^n u)D_q(u) \\ &= D_{2^n}(u)D_p(2^n u) + \psi_p(2^n u)D_q(u). \end{aligned}$$

Since $D_{2^n}(u)$ contains $1+\psi_{2^{n-1}}(u) = 1+\psi_1(2^{n-1}u)$ as a factor, and $1/2 \leqq 2^{n-1}u < 1$, $D_{2^n}(u) = 0$; hence

$$|D_k(u)| = |D_q(u)| \leqq q \leqq 2^n < 2/u.$$

An immediate consequence of Lemma 1 is the following theorem.

THEOREM X. *Let $\{a_k\}$ be of bounded variation, that is,*

$$\sum_{k=0}^{\infty} |a_k - a_{k+1}| < \infty.$$

Furthermore, let $a_k \to 0$. Then $\sum_{k=0}^{\infty} a_k\psi_k(u)$ converges uniformly in $\delta \leqq u < 1$, $\delta > 0$.

To prove this theorem we need only apply Abel's summation formula and observe that $1+\psi_1(u)+\psi_2(u)+ \cdots$ has partial sums uniformly bounded in $\delta \leqq u < 1$.

We are now in a position to give an easy proof of the localization theorem for Fourier series. (Cf. Walsh [7].)

THEOREM XI. *Let $f(x)$ and $h(x)$ coincide in some neighborhood of the point x_0. Then the Fourier series of these functions are equi-convergent at the point x_0.*

Let $g(x) = f(x) - h(x)$. The partial sums of the WFS for $g(x)$ at the point x_0 are, by (5.3), (2.12), and Theorem I,

$$\begin{aligned} s_k(x_0) &= \int_0^1 g(u)D_k(x_0, u)du = \int_0^1 g(u)D_k(x_0 \dotplus u)du \\ &= \int_0^1 g(x_0 \dotplus u)D_k(u)du. \end{aligned} \tag{6.2}$$

We are assuming that $g(u)=0$ for $|x_0-u|<\delta$, $\delta>0$. Since $|(x_0\dotplus u)-x_0|\leqq u$ by (2.8), $g(x_0\dotplus u)=0$ for $|u|<\delta$. Hence

$$s_k(x_0)=\int_\delta^1 g(x_0\dotplus u)D_k(u)du. \tag{6.3}$$

Choose n so that $2^{-n}\leqq\delta$. By (6.1), $D_k(u)=\psi_{p\cdot 2^n}(u)+\psi_{p\cdot 2^n+1}(u)+\cdots+\psi_{p\cdot 2^n+q-1}(u)$ in $\delta\leqq u<1$, so that (6.3) reduces to a sum of a bounded number of Fourier coefficients for the function $g(x_0\dotplus u)$; these coefficients tend to zero as $p\to\infty$, hence as $k\to\infty$. Thus $s_k(x_0)\to 0$ and the theorem is established.

REMARK. If x_0 is a dyadic rational, the point $x_0\dotplus u$ lies to the right of x_0 for u sufficiently small. Thus Theorem XI is still valid if we restrict ourselves to a right-hand neighborhood of a dyadic rational. This reflects clearly the topology of the group G, in which the two representatives of a dyadic rational are distinct. Unless there is an essential point involved, we shall not call attention to the special character of the dyadic rationals in the future theorems.

Our next two theorems also have their counterparts in the theory of TFS.

THEOREM XII. *Let $(f(u)-c)/(u-x_0)$ be absolutely integrable in $|u-x_0|<\delta$ for some $\delta>0$; then the* WFS *of $f(u)$ converges to c at the point x_0.*

We may write

$$\begin{aligned} s_k(x_0;f)-c &= \int_0^1 (f(u)-c)D_k(x_0\dotplus u)du \\ &= \int_{|u-x_0|<\delta}(f(u)-c)D_k(x_0\dotplus u)du+\int_{|u-x_0|>\delta} \end{aligned}$$

By Lemma 1 and (2.6),

$$|D_k(x_0\dotplus u)|<\frac{2}{x_0\dotplus u}\leqq\frac{2}{|u-x_0|};$$

therefore

$$\begin{aligned} |s_k(x_0;f)-c| \leqq{}& 2\int_{|u-x_0|<\delta}\left|\frac{f(u)-c}{u-x_0}\right|du \\ &+\left|\int_{|u-x_0|>\delta}(f(u)-c)D_k(x_0\dotplus u)du\right|. \end{aligned}$$

Given $\epsilon>0$, we can choose δ small enough so that the first integral does not exceed $\epsilon/2$; with δ thus chosen and fixed, the second integral can be made less than $\epsilon/2$ for $k>k_0(\delta,\epsilon)=k_0(\epsilon)$. This completes the proof of Theorem XII.

THEOREM XIII (DINI-LIPSCHITZ]). *If $f(x)$ is continuous, and if its modulus of continuity satisfies $\omega(\delta;f)=o(\log\delta^{-1})^{-1}$ as $\delta\to 0$, then its* WFS *converges to* *f(x) uniformly.*

Since $s_{2^n}(x)\to f(x)$ uniformly (see §1), it is sufficient to consider, for any given $k=2^n+k'$ $(0\leqq k'<2^n)$, the difference

$$(6.4)\qquad s_{2^n+k'}(x) - s_{2^n}(x) = \int_0^1 \psi_{2^n}(x)\psi_{2^n}(u)D_{k'}(x \dotplus u)f(u)du.$$

We recall that for $p<2^n$, $\psi_p(2^{-(n+1)})=1$, so that $D_{k'}(z\dotplus 2^{-(n+1)})=D_{k'}(z)$; also $\psi_{2^n}(u\dotplus 2^{-(n+1)})=-\psi_{2^n}(u)$. Hence, using the invariance of integration, we may write

$$(6.5)\quad s_{2^n+k'}(x) - s_{2^n}(x) = -\int_0^1 \psi_{2^n}(x)\psi_{2^n}(u)D_{k'}(x \dotplus u)f(u \dotplus 2^{-(n+1)})du.$$

Let us add (6.4) and (6.5):

$$2(s_{2^n+k'}(x) - s_{2^n}(x)) = \int_0^1 \psi_{2^n}(x \dotplus u)D_{k'}(x \dotplus u)\{f(u) - f(u \dotplus 2^{-(n+1)})\}du;$$

$$\begin{aligned} 2\,|\,s_{2^n+k'}(x) - s_{2^n}(x)\,| &\leqq \max_{0\leqq u<1} |\,f(u) - f(u \dotplus 2^{-(n+1)})\,| \int_0^1 |\,D_{k'}(x \dotplus u)\,|\,du \\ &\leqq \omega(2^{-(n+1)};f)L_{k'} \\ &\leqq \omega(2^{-n};f)\log(2^n) \\ &= o(1), \end{aligned}$$

by virtue of (2.10) and Theorem IX (iii).

As immediate corollaries, we have

THEOREM XIV. *If $f(x)\in\text{Lip}\,\alpha$, $0<\alpha\leqq 1$, its* WFS *converges to $f(x)$ uniformly.*

THEOREM XV. *If $f(x)$ is continuous, $s_k(x;f)=o$ (log k) uniformly in x.*

Our next theorem gives a sufficient condition for the absolute convergence of a WFS. The analogue for TFS was proved by S. Bernstein(6).

THEOREM XVI. *If $f(x)\in\text{Lip}\,\alpha$, $\alpha>1/2$, then the* WFS *of $f(x)$ converges absolutely.*

By Theorem II, if

$$f(x) \sim \sum_{k=0}^{\infty} c_k\psi_k(x),$$

then

(6) *Sur la convergence absolue des series trigonometriques,* C. R. Acad. Sci. Paris. vol. 158 (1914) pp. 1661–1664. Cf. also Zygmund, *Trigonometrical series,* pp. 135–136.

$$f(x \dot{+} h) \sim \sum_{k=0}^{\infty} c_k \psi_k(x) \psi_k(h).$$

By Parseval's equality,

$$\int_0^1 [f(x \dot{+} h) - f(x)]^2 dx = \sum_{k=0}^{\infty} c_k^2 (1 - \psi_k(h))^2.$$

Set $h = 2^{-(n+1)}$. Then $\psi_k(2^{-(n+1)}) = -1$ for $2^n \leqq k < 2^{n+1}$. Hence

$$\sum_{k=2^n}^{2^{n+1}-1} c_k^2 \leqq \int_0^1 [f(x \dot{+} 2^{-(n+1)}) - f(x)]^2 dx.$$

Again using (2.10), we have

$$\sum_{k=2^n}^{2^{n+1}-1} c_k^2 \leqq [\omega(2^{-(n+1)}; f)]^2 \leqq A 2^{-n\alpha},$$

where A is a constant. By Schwarz's inequality,

$$\sum_{k=2^n}^{2^{n+1}-1} |c_k| \leqq \left(\sum_{k=2^n}^{2^{n+1}-1} c_k^2 \right)^{1/2} \left(\sum_{k=2^n}^{2^{n+1}-1} 1^2 \right)^{1/2} \leqq A^{1/2} 2^{-n(\alpha-1/2)}. \tag{6.6}$$

Since $\alpha > 1/2$, the right-hand side of (6.6) is the nth term of a convergent series, so Theorem XVI is established.

It is possible to generalize Theorem XVI, but we shall not do so here.

7. **Summability.** The kernel for $(C, 1)$ summation (Fejér's kernel) is defined by

$$K_k(x, u) = \frac{1}{k} \sum_{r=1}^{k} D_r(x, u), \tag{7.1}$$

which may be written, setting $K_k(u) = K_k(0, u)$, as

$$K_k(x, u) = K_k(x \dot{+} u) = \frac{1}{k} \sum_{r=1}^{k} D_r(x \dot{+} u) \qquad (x \dot{+} u \neq \text{D.R.}). \tag{7.2}$$

The $(C, 1)$ mean of order k for the WFS of a function $f(x)$ is then

$$\sigma_k(x; f) = \sigma_k(x) = \int_0^1 K_k(x \dot{+} u) f(u) du. \tag{7.3}$$

Similarly we define the Abel kernel by

$$P(x, u; r) = \sum_{k=0}^{\infty} \psi_k(x) \psi_k(u) r^k \qquad (0 \leqq r < 1). \tag{7.4}$$

Setting $P(u; r) = P(0, u; r)$ and making the usual factorization, we obtain

$$(7.5)\qquad \begin{aligned} P(x, u; r) &= P(x \dotplus u; r) \\ &= \prod_{n=0}^{\infty} (1 + \psi_{2^n}(x \dotplus u) r^{2^n}) \qquad (x \dotplus u \neq \text{D.R.}). \end{aligned}$$

The Abel mean of $f(x)$ is then

$$(7.6)\qquad f(x; r) = \int_0^1 P(x \dotplus u; r) f(u) du.$$

Walsh has given a proof of the theorem that $\sigma_k(x; f)$ tends to $f(x)$ uniformly if $f(x)$ is continuous. We shall offer a different proof, in the course of which some of the properties of the $(C, 1)$ kernels will be brought to light. These properties we now state as a sequence of lemmas.

LEMMA 2. *For* $n \geqq 0$, $0 \leqq k' \leqq 2^n$,

$$(2^n + k') K_{2^n+k'}(u) = 2^n K_{2^n}(u) + k' D_{2^n}(u) + \psi_{2^n}(u) k' K_{k'}(u).$$

By definition, and by (6.1),

$$\begin{aligned} (2^n + k') K_{2^n+k'}(u) &= \sum_{r=1}^{2^n} D_r(u) + \sum_{q=1}^{k'} D_{2^n+q}(u) \\ &= 2^n K_{2^n}(u) + \sum_{q=1}^{k'} \{D_{2^n}(u) + \psi_{2^n}(u) D_q(u)\} \\ &= 2^n K_{2^n}(u) + k' D_{2^n}(u) + \psi_{2^n}(u) \sum_{q=1}^{k'} D_q(u) \\ &= 2^n K_{2^n}(u) + k' D_{2^n}(u) + \psi_{2^n}(u) k' K_{k'}(u). \end{aligned}$$

LEMMA 3. $K_{2^n}(u) \geqq 0$ for $n \geqq 0$.

Take $k' = 2^n$ in the preceding lemma:

$$2^{n+1} K_{2^{n+1}}(u) = (1 + \psi_{2^n}(u)) 2^n K_{2^n}(u) + 2^n D_{2^n}(u).$$

Since $D_{2^n}(u) \geqq 0$, $1 + \psi_{2^n} \geqq 0$, $K_1(u) \equiv 1$, induction shows that Lemma 3 is true for all n.

LEMMA 4. *Let* $k = 2^{n_1} + 2^{n_2} + \cdots + 2^{n_\nu}$, $n_1 > n_2 > \cdots > n_\nu \geqq 0$; *let* $k' = k - 2^{n_1}$, $k^{(i)} = k^{(i-1)} - 2^{n_i}$, $i = 2, 3, \cdots, \nu$. *Then*

$$k K_k(u) = \sum_{i=1}^{\nu} 2^{n_i} \psi_{k - k^{(i)}}(u) K_{2^{n_i}}(u) + \sum_{i=1}^{\nu} k^{(i)} D_{2^{n_i}}(u).$$

Lemma 4 follows immediately by iteration of Lemma 2.

LEMMA 5. *If u is not a dyadic rational*, $\lim_{k\to\infty} K_k(u) = 0$.

We may assume that $2^{-n} < u < 2^{-n+1}$. Write $k = p \cdot 2^n + q$, $0 \leqq q < 2^n$. For all

u, by (6.1), we have

$$D_{p_1 \cdot 2^n + q_1}(u) = D_{2^n}(u) D_{p_1}(2^n u) + \psi_{p_1}(2^n u) D_{q_1}(u).$$

Using the fact that $2^{-n} < u < 2^{-n+1}$, $D_{2^n}(u) = 0$; so

$$D_{p_1 \cdot 2^n + q_1}(u) = \psi_{p_1}(2^n u) D_{q_1}(u). \tag{7.7}$$

Summing (7.7),

$$\begin{aligned} kK_k(u) &= \sum_{p_1=0}^{p-1} \sum_{q_1=1}^{2^n} \psi_{p_1}(2^n u) D_{q_1}(u) + \sum_{q_1=1}^{q} \psi_p(2^n u) D_{q_1}(u) \\ &= 2^n K_{2^n}(u) D_p(2^n u) + \psi_p(2^n u) q K_q(u). \end{aligned}$$

Now $|K_r(u)| < 2/u$ for all r; $|D_p(2^n u)| < 2/(2^n u - 1)$; $q < 2^n < 2/u$. Hence

$$| kK_k(u) | < \frac{4}{u(u - 2^{-n})} + \frac{4}{u^2}, \tag{7.8}$$

from which the lemma follows. We observe that we cannot obtain an estimate of the form $K_k(u) = O(k^{-1}u^{-2})$ such as holds in the trigonometric case.

Remark. The failure of the estimate for Fejér's kernel just mentioned leads to an interesting question. In the theory of TFS, one has the following theorems (cf. Zygmund [8, pp. 61–62]):

(7.9) *If* $f(x) \in \text{Lip } \alpha$, $0 < \alpha \leqq 1$, *then*

$$s_k(x; f) - f(x) = O(k^{-\alpha} \log k).$$

(7.10) *If* $f(x) \in \text{Lip } \alpha$, $0 < \alpha < 1$, *then*

$$\sigma_k(x; f) - f(x) = O(k^{-\alpha}).$$

(7.11) *There exists a function* $f(x) \in \text{Lip } 1$, *such that*

$$\sigma_k(x; f) - f(x) \neq o(1/k);$$

in fact

$$\sigma_k(0; f) - f(0) \cong c\left(\frac{\log k}{k}\right).$$

For the Walsh functions, it is easy to establish (7.9), and (7.11) is true for the function

$$J_1(x) = \int_0^x \psi_1(u) du.$$

Whether (7.10) holds for WFS is still an open question. One method of proof in the trigonometric case depends strongly on the estimate for Fejér's kernel mentioned above.

LEMMA 6. *If $f(x)$ is continuous, $\sigma_{2^n}(x;f)\to f(x)$ uniformly.*

The argument is of standard type:

$$\sigma_{2^n}(x;f)-f(x)=\int_0^1 K_{2^n}(x\dot{+}u)\{f(u)-f(x)\}du$$
$$=\int_0^1 K_{2^n}(u)\{f(x\dot{+}u)-f(x)\}du$$
$$=\int_0^\delta K_{2^n}(u)\{f(x\dot{+}u)-f(x)\}du+\int_\delta^1,$$
$$|\sigma_{2^n}(x;f)-f(x)|\leqq\int_0^\delta |f(x\dot{+}u)-f(x)|K_{2^n}(u)du+\int_\delta^1$$
$$\leqq \omega(\delta;f)+M\int_\delta^1 K_{2^n}(u)du,$$

where $M=2\max|f(u)|$. Fix δ such that $\omega(\delta;f)<\epsilon/2$. $K_{2^n}(u)\to 0$ almost everywhere and is majorized by $2/\delta$, so the integral can be made less than $\epsilon/2$ by choosing n sufficiently large.

THEOREM XVII. *If $f(x)$ is continuous, $\sigma_k(x;f)\to f(x)$ uniformly.*

Let $g(u)=f(x\dot{+}u)-f(x)$. Then

$$\sigma_k(x;f)-f(x)=\sigma_k(0;g)=\int_0^1 K_k(u)g(u)du.$$

By Lemma 4,

$$|\sigma_k(0;g)|\leqq\sum_{i=1}^{\nu}\left(\frac{2^{n_i}}{k}\right)\int_0^1 K_{2^{n_i}}(u)|g(u)|du$$
$$+\sum_{i=1}^{\nu}\left(\frac{k^{(i)}}{k}\right)\int_0^1 D_{2^{n_i}}(u)|g(u)|du.$$

By definition, $k^{(i)}<2^{n_i}$; and $2^{n_i}<2^{n_{i-1}}$. Denote by ϵ_n the greater of the two integrals,

$$\int_0^1 K_{2^n}(u)|g(u)|du,\qquad \int_0^1 D_{2^n}(u)|g(u)|du.$$

By Lemma 6 and the fact that the second integral is

$$2^n\int_0^{2^{-n}}|g(u)|du\leqq\omega(2^{-n};f),$$

we see that $\epsilon_n\to 0$ as $n\to\infty$. Therefore

$$\frac{1}{2}\left| \sigma_k(0; g) \right| \leqq \sum_{i=1}^{\nu} (2^{n_i}/k)\epsilon_{n_i}.$$

The right-hand side is a weighted average of a null-sequence; it is easy to see that the weights are distributed in such a manner as to make the average converge to zero with increasing k. This proves the theorem.

When we come to functions of bounded variation, the results are of a highly negative nature, just as in the case of convergence. Walsh has proved that convergence cannot take place at a point of discontinuity which is not a dyadic rational (§1). Since the Fourier coefficients are $O(1/k)$, Littlewood's theorem shows that the series cannot be Cesàro summable to any order; the same remark also applies to Abel summability.

We end this section by stating without proof a few results concerning the Abel kernel $P(u; r)$ and its partial sums.

$$P_k(u; r) = \sum_{s=0}^{k-1} \psi_s(u) r^s.$$

THEOREM XVIII.

(i) $P(u; r) \geqq 0, \qquad 0 \leqq r < 1.$

(ii) $P_{2^n}(u; r) \geqq 0, \qquad 0 \leqq r < 1, \; n \geqq 0.$

(iii) $P_k(u; r) \geqq 0, \qquad 0 \leqq r \leqq r_0 = (5^{1/2} - 1)/2, \; k \geqq 0.$

(iv) $P_k(u; r) \geqq 0, \qquad 0 \leqq u < 1, \; 0 \leqq r \leqq \alpha^{2u} (\alpha = 1/2), \; k \geqq 0.$

(v) $P(u; r) < 2/u, \qquad 0 < u < 1, \; 0 \leqq r < 1.$

We remark that the constant $(5^{1/2}-1)/2$ in (iii) is best possible. The constant $\alpha = 1/2$ in (iv) is not known to be best possible; the problem of determining the exact region of positivity for all $P_k(u; r)$ is open.

8. **On certain sums.** Let $g_t(x)$ be the periodic function of x defined by

$$g_t(x) = \begin{cases} 1 & \text{for } 0 \leqq x < t, \\ 0 & \text{for } t \leqq x < 1 \end{cases} \tag{8.1}$$

and let $\lambda_n'(t)$ be the 2^nth partial sum of the WFS of $g_t(x)$ evaluated at the point $x = t$. In other words, if

$$g_t(x) \sim c_0(t) + c_1(t)\psi_1(x) + c_2(t)\psi_2(x) + \cdots, \tag{8.2}$$

then

$$\lambda_n'(t) = \sum_{k=0}^{2^n-1} c_k(t)\psi_k(t). \tag{8.3}$$

Appealing to (1.9), we have

$$(8.4) \qquad \lambda_n'(t) = 2^n \int_{\alpha_n(t)}^{\beta_n(t)} g_t(x)dx = 2^n \int_{\alpha_n(t)}^{t} dx = 2^n(t - \alpha_n(t)).$$

Since $0 \leqq t - \alpha_n(t) < 2^{-n}$,

$$(8.5) \qquad 0 \leqq \lambda_n'(t) < 1.$$

Now we have

$$(8.6) \qquad c_k(t) = \int_0^1 g_t(x)\psi_k(x)dx = \int_0^t \psi_k(x)dx = J_k(t).$$

If $k = 2^m + r$, $0 \leqq r < 2^m$, (3.5) yields

$$\psi_k(t)J_k(t) = \psi_{2^m}(t)J_{2^m}(t).$$

Hence, from (8.3),

$$(8.7) \qquad \begin{aligned} \lambda_n'(t) &= t + \sum_{m=0}^{n-1} \sum_{r=0}^{2^m-1} \psi_{2^m}(t)J_{2^m}(t) \\ &= t + \sum_{m=0}^{n-1} 2^m \psi_{2^m}(t)J_{2^m}(t). \end{aligned}$$

Now $2^m J_{2^m}(t) = J_1(2^m t)$ and $\psi_{2^m}(t) = \psi_1(2^m t)$. Furthermore $\psi_1(x)J_1(x) = x - [x + 1/2]$, so that

$$(8.8) \qquad \begin{aligned} \lambda_n'(t) &= t + \sum_{m=0}^{n-1} \psi_1(2^m t)J_1(2^m t) \\ &= t + \sum_{m=0}^{n-1} \left(2^m t - \left[2^m t + \frac{1}{2}\right]\right). \end{aligned}$$

If we define

$$\begin{aligned} \mu_0(t) &= t - [t] - 1/2, \\ \mu_n(t) &= \mu_0(2^n t), \\ \lambda_n(t) &= \mu_0(2^n t + 1/2), \end{aligned}$$

equation (8.8) becomes, by virtue of (8.4),

$$(8.9) \qquad \sum_{m=0}^{n-1} \lambda_m(t) = \lambda_n'(t) - t = 2^n(t - \alpha_n) - t,$$

and using (8.5) and the fact that $0 \leqq t < 1$,

$$(8.10) \qquad \left| \sum_{m=0}^{n-1} \lambda_m(t) \right| < 1.$$

We have therefore proved the following theorem:

THEOREM XIX. *For all t, and for all $n \geqq 1$,*

$$\left| \sum_{m=0}^{n-1} \mu_0\left(2^m t + \frac{1}{2}\right) \right| < 1.$$

Now it is easy to see that

$$\begin{aligned} \mu_0(x + 1/2) &= \mu_0(x) + 1/2 && (\text{for } 0 \leqq x < 1/2) \\ &= \mu_0(x) - 1/2 && (\text{for } 1/2 \leqq x < 1) \\ \mu_0(x + 1/2) &= \mu_0(x - 1/2) && (\text{for all } x). \end{aligned}$$

Hence $\mu_0(x+1/2) = \mu_0(x) + \psi_1(x)/2$, and

$$(8.11) \qquad \begin{aligned} \sum_{m=0}^{n-1} \mu_0\left(2^m t + \frac{1}{2}\right) &= \sum_{m=0}^{n-1} \mu_0(2^m t) + \frac{1}{2} \sum_{m=0}^{n-1} \psi_1(2^m t) \\ &= \sum_{m=0}^{n-1} \mu_m(t) + \frac{1}{2} \sum_{m=0}^{n-1} \psi_{2^m}(t). \end{aligned}$$

If we define the random variables X_m by

$$\text{Prob}\ \{X_m < x\} = \left| \underset{t,\, 0 \leqq t < 1}{E} \{\psi_{2^m}(t) < x\} \right|,$$

it is easy to see that $X_1, X_2, \cdots$ are mutually independent and identically distributed, with mean value 0 and standard deviation 1. By the central limit theorem the variable

$$Z_n = \frac{X_0 + X_1 + \cdots + X_{n-1}}{n^{1/2}}$$

has a distribution which tends to the normal. In other words

$$(8.12) \qquad \lim_{n \to \infty} \left| \underset{t}{E} \left\{ \alpha < \frac{1}{n^{1/2}} \sum_{m=0}^{n-1} \psi_{2^m}(t) < \beta \right\} \right| = \frac{1}{(2\pi)^{1/2}} \int_\alpha^\beta e^{-x^2/2} dx.$$

It follows readily from (8.11) and Theorem XIX that

$$(8.13) \qquad \lim_{n \to \infty} \left| \underset{t}{E} \left\{ \alpha < \frac{2}{n^{1/2}} \sum_{m=0}^{n-1} \mu_m(t) < \beta \right\} \right| = \frac{1}{(2\pi)^{1/2}} \int_\alpha^\beta e^{-x^2/2} dx,$$

and finally, by an elementary transformation, we arrive at a theorem of Kac [1]:

$$\lim_{n \to \infty} \left| \underset{t}{E} \left\{ \alpha < \frac{2^{1/2}}{n^{1/2}} \sum_{m=0}^{n-1} \mu_m(t) < \beta \right\} \right| = \frac{1}{\pi^{1/2}} \int_\alpha^\beta e^{-y^2} dy.$$

In another paper [2] Kac points out that the set of functions $f(2^n t)$ may in some cases (for example, $f(t)=t-[t]-1/2$) exhibit a certain degree of statistical dependence. In the case of complete independence one would expect

$$(8.14)\quad \lim_{n\to\infty}\left| \underset{t}{E}\left\{\alpha < \frac{1}{n^{1/2}}\sum_{m=0}^{n-1} f(2^m t) < \beta\right\}\right| = \frac{1}{\sigma(2\pi)^{1/2}}\int_\alpha^\beta e^{-u^2/2\sigma^2}du,$$

where

$$(8.15)\quad \sigma^2 = \int_0^1 f^2(t)dt - \left(\int_0^1 f(t)dt\right)^2.$$

Kac's theorem shows that (8.14) is valid for $f(t)=t-[t]-1/2$ but σ^2 is not given by (8.15). To show that this discrepancy is not easily remedied, Theorem XIX indicates that even a simple translate of $f(t)$, for which the corresponding random variable is identically distributed with that of $f(t)$, need not behave like $f(t)$ with respect to sums of the type considered here. We may still think of (8.15) as being valid for $f(t)=\mu_0(t+1/2)=t-[t+1/2]$ if we interpret the right side as the singular normal distribution with $\sigma=0$.

It might be of interest to determine the behavior of the sums

$$\sum_{m=0}^{n-1} \mu_0(2^m t + \xi)$$

for values of ξ other than 0 and 1/2.

9. **Uniqueness and localization theorems.** Let us consider a series

$$(9.1)\quad S(x) \equiv a_0 + a_1\psi_1(x) + a_2\psi_2(x) + \cdots,$$

not necessarily a Walsh-Fourier series. We shall be concerned with finding conditions under which $S(x)$ is actually a WFS. The analogous problem in the theory of trigonometric series has long been a part of classical mathematics. (Cf. Zgymund [8, chap. 11].) Riemann employed the device of twice integrating a given trigonometrical series (formally) and studying the properties of the resulting function, relating it to the original series by means of a generalized second derivative. In attacking the problem for Walsh series, we find a rather surprising state of affairs. It turns out that the *first* integral is the natural weapon to use; we shall show (Theorem XX) that the first integral converges wherever $S(x)$ does. That this is not necessarily true of trigonometric series is seen from the example $\sum(\sin nx)/\log n$.

THEOREM XX. *Let* $S(x)\equiv\sum_{k=0}^\infty a_k\psi_k(x)$ *converge at the point* x, *and let*

$$(9.2)\quad L(x) \equiv a_0 x + \sum_{k=1}^{\infty} a_k J_k(x).$$

Then the series (9.2) *also converges at the point* x.

Consider the partial sums of order 2^N, $N\geqq 0$:

$$\begin{aligned} L_{2^N}(x) &= a_0 x + \sum_{k=1}^{2^N-1} a_k J_k(x) \\ &= a_0 x + \sum_{n=0}^{N-1} \sum_{r=0}^{2^n-1} a_{2^n+r} J_{2^n+r}(x) \\ &= a_0 x + \sum_{n=0}^{N-1} \sum_{r=0}^{2^n-1} a_{2^n+r} \psi_r(x) J_{2^n}(x) \\ &= a_0 x + \sum_{n=0}^{N-1} \psi_{2^n}(x) J_{2^n}(x) \sum_{r=0}^{2^n-1} a_{2^n+r} \psi_{2^n+r}(x) \\ &= a_0 x + \sum_{n=0}^{N-1} \psi_{2^n}(x) J_{2^n}(x)(s_{2^{n+1}}(x) - s_{2^n}(x)), \end{aligned}$$

where $s_{2^n}(x) = \sum_{k=0}^{2^n-1} a_k \psi_k(x)$. Since $S(x)$ converges the terms of the last series are $o(2^{-n})$. (We recall, from §3, that $|J_{2^n}(x)| = J_{2^n}(x) < 2^{-n}$.) Now the partial sum of order $2^N + q$, $0 \leq q < 2^N$, may be written

$$L_{2^N+q}(x) = L_{2^N}(x) + \sum_{r=0}^{q-1} a_{2^N+r} J_{2^N+r}(x).$$

The convergence of $S(x)$ implies that $a_k \to 0$, so the last sum is $o(q \cdot 2^{-N}) = o(1)$ and the theorem is proved.

It will be observed that the convergence of (9.2) can be established under much weaker assumptions. For example, if $a_k \to 0$ and $\sum 2^{-n} |s_{2^n}(x)|$ converges the theorem is still true.

If $S(x)$ is the WFS of $f(x)$, then $s_{2^n}(x) \to f(x)$ almost everywhere, so $L(x)$ converges almost everywhere. But much more than this can be said.

THEOREM XXI. *Let*

$$S(x) \equiv \sum_{k=0}^{\infty} a_k \psi_k(x) \sim f(x).$$

Then

$$L(x) = \int_0^x f(u) du. \tag{9.3}$$

Let $g(u) \equiv g_x(u)$ be the characteristic function of the interval $(0, x)$, mod 1, and let

$$g(u; r) = \int_0^1 P(t \dotplus u; r) g(t) dt$$

be its Abel mean. We have

$$|g(u;r)| \leqq \max |g(t)| \int_0^1 |P(t+u;r)|\,du = 1.$$

Also $g(u;r) \to g(u)$ almost everywhere. Hence

$$\int_0^1 f(u)g(u;r)du \to \int_0^1 f(u)g(u)du = \int_0^x f(u)du.$$

But

$$g(u;r) = \sum_{k=0}^{\infty} \psi_k(u) J_k(x) r^k. \tag{9.4}$$

By the uniform convergence of (9.4),

$$\int_0^1 f(u)g(u;r)du = \sum_{k=0}^{\infty} a_k J_k(x) r^k.$$

Hence the series for $L(x)$ is Abel summable to $\int_0^x f(u)du$. But the terms of that series are $o(1/k)$, from which follows ordinary convergence, and (9.3) is established.

Our next theorem shows that $L(x)$, when it exists almost everywhere, behaves very much like an integral (see Theorem VII).

THEOREM XXII. *If $S(x)$ converges almost everywhere, $L(x)-a_0[x]$ is integrable and has Fourier coefficients*

$$c_0 = \frac{1}{4}\sum_{r=0}^{\infty} 2^{-r} a_{2^r}, \tag{9.5}$$

$$c_k = c_{2^n+k'} = -2^{-(n+2)}\left\{a_{k'} - \sum_{r=1}^{\infty} 2^{-r} a_{2^{n+r}+k}\right\} \quad (n \geqq 0,\ 0 \leqq k' < 2^n). \tag{9.6}$$

The function $L(x)-a_0[x]$ exists almost everywhere, by Theorem XX, and is periodic. From §3, we have

$$x - [x] = \sum_{m=0}^{\infty} \mu_0^{(m)} \psi_m(x),$$

$$J_k(x) = \sum_{m=0}^{\infty} \mu_k^{(m)} \psi_m(x) \qquad (k > 0),$$

where

$$\begin{aligned} \mu_0^{(m)} &= 1/2 && \text{if } m = 0, \\ \mu_0^{(m)} &= -2^{-(n+2)} && \text{if } m = 2^n, \\ \mu_0^{(m)} &= 0 && \text{otherwise,} \end{aligned} \tag{9.7}$$

$$
(9.8)\qquad \begin{aligned}
\mu_k^{(m)} &= 2^{-(n+2)} && \text{if } k = 2^n + m,\ 0 \leqq m < 2^n,\\
\mu_k^{(m)} &= -\,2^{-(p+2)} && \text{if } m = 2^p + k,\ 0 < k < 2^p,\\
\mu_k^{(m)} &= 0 && \text{otherwise.}
\end{aligned}
$$

It follows that

$$
\begin{aligned}
F_N(x) \equiv a_0(x - [x]) + \sum_{k=1}^{N} a_k J_k(x)
&= \sum_{k=0}^{N} a_k \sum_{m=0}^{\infty} \mu_k^{(m)} \psi_m(x)\\
&= \sum_{m=0}^{\infty} \psi_m(x) \sum_{k=0}^{N} a_k \mu_k^{(m)}\\
&= \sum_{m=0}^{\infty} c_m^{(N)} \psi_m(x),
\end{aligned}
$$

where

$$
c_m^{(N)} = \sum_{k=0}^{N} a_k \mu_k^{(m)}.
$$

From (9.7) and (9.8), we have

$$
(9.9)\qquad \sum_{k=0}^{\infty} \left| \mu_k^{(0)} \right| = \frac{1}{2} + \sum_{n=0}^{\infty} 2^{-(n+2)} = 1,
$$

$$
(9.10)\qquad \begin{aligned}
\sum_{k=0}^{\infty} \left| \mu_k^{(m)} \right| &= 2^{-(p+2)} + \sum_{n=p+1}^{\infty} 2^{-(n+2)}\\
&= 2^{-(p+1)} < 1/m, \qquad \text{if } m = 2^p + k',\ 0 \leqq k' < 2^p.
\end{aligned}
$$

If max $|a_k| = A$, we deduce that $c_m^{(N)} \to c_m$ uniformly, and that $|c_m^{(N)}| \leqq A/m$ for $m>0$; also $|c_m| \leqq A/m$ for $m>0$. By the Riesz-Fischer theorem, there exists a function $F(x) \in L^2$ with Fourier coefficients $\{c_m\}$. The series

$$
\sum_{m=0}^{\infty} (c_m^{(N)} - c_m)^2 = \int_0^1 (F_N(x) - F(x))^2 dx
$$

converges uniformly in N, and the terms tend to zero. Hence $F_N(x)$ converges to $F(x)$ in the mean of L^2. But $F_N(x)$ converges to $L(x) - a_0[x]$ almost everywhere, so $L(x) - a_0[x] = F(x) \sim \{c_m\}$ almost everywhere.

The theorem just proved enables us to deduce properties of the coefficients a_k from those of the function $L(x)$.

THEOREM XXIII. *If $a_k \to 0$ and $L(x) \equiv \sum_{k=0}^{\infty} a_k J_k(x)$ is absolutely continuous, then $\sum_{k=0}^{\infty} a_k \psi_k(x)$ is the Fourier series of $L'(x)$.*

In establishing (9.5) and (9.6) it was necessary merely to ensure the existence of $L(x)$ and the boundedness of the a_k, and these are supplied by our hypotheses. Hence if $L(x)-a_0[x]\sim\{c_m\}$, for each fixed $k\geqq 0$,

$$a_k = -2^{n+2}c_{2^n+k} + o(1). \tag{9.11}$$

If $L'(x)\sim b_0+b_1\psi_1(x)+\cdots$, Theorem VII assures us that b_k also satisfies (9.11), so $a_k=b_k$.

For our next theorem we recall the definitions of $\alpha_n\equiv\alpha_n(x)$ and $\beta_n\equiv\beta_n(x)$:

$$\alpha_n = p\cdot 2^{-n} \leqq x < (p+1)2^{-n} = \beta_n. \tag{9.12}$$

THEOREM XXIV. *If $a_k\to 0$, then for every x and for every $n\geqq 0$, $L(\alpha_n)$ and $L(\beta_n)$ exist and*

$$L(\beta_n) - L(\alpha_n) = (\beta_n - \alpha_n)s_{2^n}(x) = 2^{-n}s_{2^n}(x), \tag{9.13}$$

$$L(\beta_n)-L(\alpha_n)=o(1) \text{ uniformly in } x. \tag{9.14}$$

For $k\geqq 2^n$, $J_k(\alpha_n)=J_k(\beta_n)=0$. Hence

$$\begin{aligned} L(\beta_n) - L(\alpha_n) &= \sum_{k=0}^{2^n-1} a_k(J_k(\beta_n) - J_k(\alpha_n)) \\ &= \sum_{k=0}^{2^n-1} a_k \int_{\alpha_n}^{\beta_n} \psi_k(u)du \\ &= \int_{\alpha_n}^{\beta_n} \sum_{k=0}^{2^n-1} a_k\psi_k(u)du \\ &= \int_{\alpha_n}^{\beta_n} s_{2^n}(u)du. \end{aligned}$$

But $s_{2^n}(u)$ is constant for $\alpha_n\leqq u<\beta_n$, and the point x lies in this interval. (9.14) follows from (9.13) and the fact that $s_{2^n}(x)=o(2^n)$ uniformly, since $a_k\to 0$.

Let us now define $\alpha_n'\equiv\alpha_n'(x)$:

$$\begin{aligned} \alpha_n' &= \alpha_n && \text{if } \alpha_n < x, \\ \alpha_n' &= \alpha_n - 2^{-n} && \text{if } \alpha_n = x. \end{aligned} \tag{9.15}$$

THEOREM XXV. *If $a_k\to 0$ and $L(x)$ exists at the point x_0, then $L(\alpha_n')\to L(x_0)$, $L(\beta_n)\to L(x_0)$.*

If x_0 is a dyadic rational the results follow from (9.14); otherwise, we have

$$\begin{aligned} L_{2^n}(x_0) - L_{2^n}(\alpha_n') &= L_{2^n}(x_0) - L(\alpha_n') \\ &= \int_{\alpha_n}^{x_0} s_{2^n}(u)du = (x_0 - \alpha_n)s_{2^n}(x_0). \end{aligned} \tag{9.16}$$

But $L_{2^n}(x_0) \to L(x_0)$ and $(x_0 - \alpha_n)s_{2^n}(x_0) \to 0$, and the first result follows; (9.14) establishes the second.

The following lemma is well known(7).

LEMMA 7. *Let $f(x)$, $a \leqq x \leqq b$, be an integrable function, $F(x)$ the indefinite integral of $f(x)$, and $\epsilon > 0$ an arbitrary number. Then there exist two functions $\phi(x)$ and $\psi(x)$ such that*

(i) *$\phi(x)$ and $\psi(x)$ are continuous,*

(ii) $|\phi(x) - F(x)| < \epsilon$, $|\psi(x) - F(x)| < \epsilon$,

(iii) *At every point where $f(x) \neq +\infty$, all the derivates of $\psi(x)$ exceed $f(x)$; at every point where $f(x) \neq -\infty$, all the derivates of $\phi(x)$ are less than $f(x)$.*

Our next lemma (or even a weaker form of it) enables us to deduce the main theorem for a Walsh series which converges everywhere to an integrable function.

LEMMA 8. *Let $G(x)$, $a < x < b$, satisfy*

(i) $G(\alpha_n') \to G(x)$;

(ii) *For all x in (a, b),*

$$\liminf_{n\to\infty} 2^n(G(\beta_n) - G(\alpha_n)) \leqq B(x) < +\infty;$$

(iii) *For all x in (a, b), except perhaps for a denumerable set E,*

$$\liminf_{n\to\infty} 2^n(G(\beta_n) - G(\alpha_n)) \leqq 0.$$

Then $G(x)$ is monotone non-increasing in (a, b).

We shall first make some definitions. By I_n we shall mean an interval of the form $k \cdot 2^n \leqq x < (k+1)2^{-n}$, and by a small Greek letter with a subscript, such as ξ_n, we shall mean a number of the form $k \cdot 2^{-n}$. Let

$$\delta(I_n) = 2^n(G(\overline{k+1} \cdot 2^{-n}) - G(k \cdot 2^{-n})).$$

We shall say that I_m is *properly nested* in I_n, written $I_m \prec I_n$, $m > n$, if $I_m \subset I_n$ and if the right-hand end point of I_m is to the left of the right-hand end point of I_n. By a *properly nested sequence* of intervals we shall mean a nested sequence in which infinitely many of the intervals are properly nested in the preceding ones. Every properly nested sequence $\{I_{N+k}\}$, $k = 0, 1, 2, \cdots$, defines a unique number x such that $\alpha_{N+k}(x)$, $\beta_{N+k}(x)$ are the end points of I_{N+k}.

Now suppose that $G(w) < G(z)$ for some pair of points w, z, with $w < z$. By (i), we can find $\xi_n < \eta_n$ such that $G(\xi_n) < G(\eta_n)$. Hence we can find an $I_n = (\alpha_n, \beta_n)$ such that $\delta(I_n) = c_n > 0$. We shall now prove that given any v, $0 < v < 1$, we can find $I_{n+1}, I_{n+2}, \cdots, I_{n+p}$, satisfying

---

(7) Saks, *Théorie de l'integrale*, pp. 132–133.

$$(9.17)\qquad I_n \supset I_{n+1} \supset \cdots \supset I_{n+p-1} \succ I_{n+p},$$

$$(9.18)\qquad \delta(I_{n+r}) \geqq (1-v)c_n, \qquad r = 0, 1, 2, \cdots, p.$$

Draw a line $\lambda(x) = G(\alpha_n) + (1-v)c_n(x-\alpha_n)$, and consider the sequence of points $\xi_n = \alpha_n$, $\xi_{n+r} = (\xi_{n+r-1} + \beta_n)/2$, $r = 1, 2, \cdots$. By (i), $G(\xi_{n+r}) \to G(\beta_n)$, so there is a first element, say ξ_{n+p}, such that $G(\xi_{n+p}) > \lambda(\xi_{n+p})$; clearly $p \geqq 1$. The intervals

$$\begin{aligned} I_n &= (\xi_n, \beta_n), \\ I_{n+1} &= (\xi_{n+1}, \beta_n), \\ &\cdots\cdots\cdots \\ I_{n+p-1} &= (\xi_{n+p-1}, \beta_n), \\ I_{n+p} &= (\xi_{n+p-1}, \xi_{n+p}) \end{aligned}$$

then satisfy (9.17) and (9.18).

Now if $\delta(\xi_{n+p}, \beta_n) \geqq (1-v)c_n$, we can apply the method just used and obtain a sequence $I'_{n+p} = (\xi_{n+p}, \beta_n)$, $I'_{n+p+1}, \cdots, I'_{n+q}$, satisfying

$$(9.19)\qquad I'_{n+p} \supset \cdots \supset I'_{n+q-1} \succ I'_{n+q},$$

$$(9.20)\qquad \delta(I'_{n+r}) \geqq (1-v)^2 c_n, \qquad r = p, p+1, \cdots, q.$$

Since $I'_{n+p} \subset I_{n+p-1}$, we have obtained two sequences I_{n+r}, $r = 1, 2, \cdots, p$, and I'_{n+r}, $r = 1, 2, \cdots, q$, with the properties

$$(9.21)\qquad I_n \supset I_{n+1} \supset \cdots \supset I_{n+p-1} \succ I_{n+p},$$

$$(9.22)\qquad I_n \supset I'_{n+1} \supset \cdots \supset I'_{n+q-1} \succ I'_{n+q},$$

$$(9.23)\qquad \delta(I_{n+r}) \geqq (1-v)^2 c_n, \qquad r = 0, 1, \cdots, p,$$

$$(9.24)\qquad \delta(I'_{n+r}) \geqq (1-v)^2 c_n, \qquad r = 0, 1, \cdots, q,$$

$$(9.25)\qquad I_{n+p} \cdot I'_{n+q} = 0.$$

If (9.21)–(9.25) are satisfied for some p and q for an interval I_n with $\delta(I_n) = c_n > 0$ and for some v, $0 < v < 1$, we shall say that I_n is of type $D(c_n, v)$.

Now suppose that $\delta(\xi_{n+p}, \beta_n) < (1-v)c_n$; in this case it can be shown that $\delta(I_{n+p}) > (1+v)c_n$. Applying our method once again, this time to I_{n+p} with $\delta(I_{n+p}) = c_{n+p}$ and parameter v, we find that either I_{n+p} is of type $D(c_{n+p}, v)$ (from which follows that I_n is of type $D(c_n, v)$) or I_{n+p} contains a nested sequence

$$I_{n+p} \supset I_{n+p+1} \supset \cdots \supset I_{n+p_1-1} \succ I_{n+p_1}$$

such that

$$\delta(I_{n+r}) \geqq (1-v)c_{n+p} \geqq (1-v)(1+v)c_n \qquad (r = p, p+1, \cdots, p_1 - 1),$$

$$\delta(I_{n+p_1}) > (1+v)c_{n+p} \geqq (1+v)^2 c_n.$$

Continuing in this way, we obtain the sequence

$$I_n \supset \cdots \supset I_{n+p-1} \succ I_{n+p} \supset \cdots \supset I_{n+p_1-1}$$
$$\succ I_{n+p_1} \supset \cdots \supset \cdots \supset I_{n+p_k-1} \succ I_{n+p_k} \supset \cdots ,$$

with

$$\begin{aligned}
\delta(I_n) &= c_n, \\
\delta(I_{n+r}) &\geqq (1-v)c_n, && 0 < r < p, \\
\delta(I_{n+p}) &\geqq (1+v)c_n, \\
\delta(I_{n+r}) &\geqq (1-v)(1+v)c_n, && p < r < p_1, \\
\delta(I_{n+p_1}) &\geqq (1+v)^2 c_n, \\
&\cdots\cdots\cdots\cdots, \\
\delta(I_{n+r}) &\geqq (1-v)(1+v)^k c_n, && p_{k-1} < r < p_k, \\
\delta(I_{n+p_k}) &\geqq (1+v)^{k+1} c_n.
\end{aligned}$$

If this sequence does not break off, it defines a number x such that

$$2^N(G(\beta_N(x)) - G(\alpha_N(x))) = \delta(I_N), \qquad N \geqq n.$$

But $\lim_{N\to\infty} \delta(I_N) = +\infty$, contradicting (ii). Hence the sequence *does* end, and we find that I_{n+p_k} is of type $D(c_{n+p_k}, v)$ for some k; therefore I_n is of type $D(c_n, v)$.

We have just proved that *every* interval I_n with $\delta(I_n) = c_n > 0$ is of type $D(c_n, v)$ for *every* v such that $0 < v < 1$. Now order the elements of the set E, say $x^{(1)}, x^{(2)}, \cdots, x^{(k)}, \cdots$. Let $v_1, v_2, \cdots$ be chosen so that $0 < v_k < 1$ and

$$\prod_{k=1}^{\infty} (1 - v_k)^2 = \frac{1}{2} \cdot$$

Since I_n is of type $D(c_n, v_1)$, we can find two disjoint sub-intervals I_{n+p} and I'_{n+q} such that $c_{n+p} = \delta(I_{n+p}) \geqq (1-v_1)^2 c_n$ and $c'_{n+q} = \delta(I'_{n+q}) \geqq (1-v_1)^2 c_n$. Now choose that interval which does not contain $x^{(1)}$ and continue the process with v_2, avoiding the point $x^{(2)}$, and so on. In this way we obtain a properly nested sequence of intervals which define a point x that cannot belong to the set E. But for this point x,

$$\liminf_{N\to\infty} 2^N(G(\beta_N) - G(\alpha_N)) \geqq c_n \prod_{k=1}^{\infty} (1 - v_k)^2 = \frac{1}{2} c_n > 0,$$

which contradicts (iii). Our lemma is therefore established.

THEOREM XXVI. *Let $S(x)$ converge in $a < x < b$ except perhaps on a denumerable set E, where, however, $|s_{2^n}(x)| \leqq B(x) < \infty$; if the limit-function $f(x)$ is integrable, then*

$$L(x) = C + \int_a^x f(u)du$$

for some constant C and all x in (a, b).

Define $f(x)$ to be 0 on E, and apply Lemma 7 for an arbitrary $\epsilon > 0$. Let $G(x) = \phi(x) - L(x)$, $H(x) = L(x) - \psi(x)$. It is easy to see that part (iii) of Lemma 7 implies

$$\liminf_{n\to\infty} 2^n(\phi(\beta_n) - \phi(\alpha_n)) \leqq f(x) \leqq \limsup_{n\to\infty} 2^n(\psi(\beta_n) - \psi(\alpha_n)) \tag{9.26}$$

for all x in (a, b). By (9.13), for $x \notin E$,

$$\lim_{n\to\infty} 2^n(L(\beta_n) - L(\alpha_n)) = \lim_{n\to\infty} s_{2^n}(x) = f(x). \tag{9.27}$$

For all x in (a, b), by Theorem XXIV,

$$\limsup_{n\to\infty} \left| 2^n(L(\beta_n) - L(\alpha_n)) \right| \leqq B(x) < \infty. \tag{9.28}$$

From (9.26), (9.27), and (9.28), we find that $G(x)$ and $H(x)$ satisfy conditions (ii) and (iii) of Lemma 8.

Condition (i) follows from the continuity of ϕ and ψ and from the existence of $L(x)$ (Theorem XX, remark) together with Theorem XXV. Applying Lemma 8, therefore, we find that $G(x)$ and $H(x)$ are nonincreasing. Letting $\epsilon \to 0$, their limits, $\pm(L(x) - F(x))$ are also non-increasing, hence constant. Since $F(x)$ is an integral of $f(x)$, Theorem XXVI is established.

It is possible to modify Lemma 8 as follows:

LEMMA 9. *Let* $G(x)$ *be defined in* $a < x < b$ *and satisfy*:

(i) *Except perhaps on a denumerable set* E,

$$\limsup_{n\to\infty} 2^n(G(\beta_n) - G(\alpha_n)) \leqq 0.$$

(ii) *For all* x *in* (a, b), $G(\alpha_n') \to G(x)$, $G(\beta_n) \to G(x)$. *Then* $G(x)$ *is monotone nonincreasing in* (a, b).

The proof is not too different from that of Lemma 8 and will be omitted. Since condition (ii) of Lemma 9 is satisfied for $L(x)$ wherever it exists, the proof of Theorem XXVI carries over almost word for word to the following slightly stronger result.

THEOREM XXVII. *Let* $S(x)$ *converge in* $a < x < b$ *except perhaps on a denumerable set* E, *where however* $L(x)$ *exists; if the limit-function* $f(x)$ *is integrable, then*

$$L(x) = C + \int_a^x f(u)du$$

for some constant C and all x in (a, b).

If we take the interval (a, b) in Theorem XXVII to be $(0, 1)$, and apply Theorem XXIII, we obtain our next theorems.

THEOREM XXVIII. *Let $S(x)$ converge to an integrable function $f(x)$ except perhaps on a denumerable set E, where however $L(x)$ exists. Then $f(x)\sim S(x)$.*

COROLLARY. *If $S(x)$ converges to zero everywhere, then $a_k=0$, $k\geqq 0$.*

In order to weaken the "everywhere" condition in the corollary to Theorem XXVIII, let us assume that $S(x)$ converges to zero except perhaps on a denumerable set E. We shall examine the structure of the subset $D\subset E$ consisting of points x for which the partial sums $s_N(x)$ are unbounded. Here again we find the concept of the group G (§2) an extremely useful one.

Let $\mu(x)$ be the function which maps the reals into G, and let $g(x)$ be an arbitrary real-valued function of period 1. We define the corresponding function on G by

$$\bar{g}(\bar{x}) = g(x) \qquad \text{if } \mu(x) = \bar{x} \text{ for some } x, \tag{9.29}$$

$$\bar{g}(\bar{x}) = \limsup_{\bar{y}\to\bar{x}} \bar{g}(\bar{y}) \qquad \text{if } \mu(x) \neq \bar{x} \text{ for any } x. \tag{9.30}$$

In (9.30) the lim sup is taken over those $\bar{y}$ which correspond to dyadic irrationals.

We recall that G is a complete metric space. The neighborhoods in G may be taken as the sets of points

$$\{t_1, t_2, \cdots, t_n, u_{n+1}, u_{n+2}, \cdots\}$$

in which $t_1, \cdots, t_n$ are fixed and $u_{n+1}, u_{n+2}, \cdots$ vary independently. These neighborhoods $\mathcal{N}(t_1, t_2, \cdots, t_n)$, $n=1, 2, \cdots$, form a complete system for G. The characteristic function of any neighborhood is continuous, since $\mathcal{N}$ is both open and closed. Any finite linear combination of such characteristic functions is also continuous. It is not too difficult to see that such linear combinations are precisely the images by (9.29) and (9.30) of all finite sums $\sum a_k\psi_k(x)$. Hence *the partial sums of an arbitrary Walsh series are continuous in the topology of G.*

Now let $\bar{D}$ be the set of $\bar{x}$ in G for which the sequence of continuous functions $\bar{s}_{2^n}(\bar{x})$ is unbounded. Then D is a G_δ, that is, a denumerable product of open sets[8]. By a theorem of W. H. Young[9], *in a complete space, every G_δ which contains a self-dense non-null subset has the power of the continuum.* It is clear that the set of $\bar{x}$ in G for which $\mu(x)=\bar{x}$ has no solution is denumerable, and that if this set is removed from $\bar{D}$ we have left precisely the image of D

(8) See F. Hausdorff, *Mengenlehre*, pp. 270–271.
(9) Ibid. pp. 136–137.

under μ. It will therefore be sufficient for our purpose to show that this image contains no isolated points. If $\bar{x}_0=\mu(x_0)$ is such a point, there must be a neighborhood $\mathcal{N}$ containing $\bar{x}_0$ but no other images of points of D. $\mathcal{N}$ is therefore the image under μ of a dyadic interval $k\cdot 2^{-p}\leqq x<(k+1)2^{-p}$ which contains x_0 but no other points of D.

Two cases now arise. If $x_0=k\cdot 2^{-p}$, we have the conditions of Theorem XXVI satisfied in $k\cdot 2^{-p}<x<(k+1)2^{-p}$, so that $L(x)$ is constant in that interval. Since x_0 is a dyadic rational, $L(x_0)$ exists; since $L(\beta_n(x_0))\rightarrow L(x_0)$, we find that $L(x)$ is constant in $k\cdot 2^{-p}\leqq x<(k+1)2^{-p}$. If $k\cdot 2^{-p}<x_0<(k+1)2^{-p}$ and x_0 is a dyadic rational, a similar argument shows that the same result is valid. If, however, x_0 is not a dyadic rational, we find that $L(x)$ is constant in two intervals abutting on x_0. Since $L(\beta_n(x_0))-L(\alpha_n(x_0))\rightarrow 0$, the two constants must be equal. In both cases, then, for $n>n_0$,

$$s_{2^n}(x_0) = 2^n(L(\beta_n(x_0)) - L(\alpha_n(x_0))) = 0. \tag{9.31}$$

This contradicts the assumption that x_0 belongs to D. Thus the image of D is self-dense and Young's theorem assures us that $\overline{D}$ has the power of the continuum unless D is vacuous. Since $\mu(D)$ and $\overline{D}$ differ by a denumerable set at most, and D is denumerable, D *is* the null-set. Another application of Theorem XXVI shows that $L(x)$ is constant, hence all the coefficients a_k of $S(x)$ vanish.

THEOREM XXIX. *If $S(x)$ converges to zero except perhaps on a denumerable set E, then $a_k=0$ for all k.*

We remark that the same method of proof is not effective in the case of an arbitrary integrable limit-function $f(x)$, although we feel quite sure that the theorem is true. It is possible to construct a series $S(x)$ whose partial sums of order 2^n are bounded at every point but one. To do this, consider the Fourier series of the function $f(x)$ defined as follows:

$$f(x) = 2^n/n^2 \quad \text{for} \quad 2^{-n} \leqq x < 2^{-n+1}, \qquad n = 1, 2, \cdots.$$

Clearly $f(x)$ is integrable, and

$$\int_0^1 f(x)dx = \sum_{n=1}^{\infty} \frac{1}{n^2}\cdot$$

It is easy to see that

$$s_{2^p}(x) = f(x) \quad \text{for} \quad 2^{-n} \leqq x < 2^{-n+1}, \qquad p \geqq n.$$

On the other hand,

$$s_{2^p}(0) = 2^p\int_0^{2^{-p}} f(x)dx = 2^p \sum_{k=p+1}^{\infty} \frac{1}{k^2} > \frac{2^p}{(p+1)^2} \rightarrow \infty.$$

If, however, we try to study the set C where $L(x)$ fails to exist, we can prove that $\mu(C)$ is self-dense. The difficulty here is that the set of points where a sequence of continuous functions fails to converge need not be a G_δ; it is known to be an $F_{\sigma\delta}$, but this is not sufficient to draw the desired conclusion.

We come now to questions of localization. If $a_k \to 0$, then $L(x)$ is completely determined by the partial sums of order 2^n. Conversely, if $L(x)$ is given in $a<x<b$, the formula

$$L(\beta_n(x)) - L(\alpha_n(x)) = 2^{-n}s_{2^n}(x) \tag{9.32}$$

shows that $s_{2^n}(x)$ is determined for $n \geqq n_0(x)$, $a<x<b$. Even more is true, however.

THEOREM XXX. *If $a_k \to 0$ and if $L(x)$ is constant in $a<x<b$, then $S(x)$ converges to zero uniformly in any sub-interval (α, β), $a<\alpha \leqq x<\beta<b$.*

It is sufficient to take (α, β) of the form I_p: $\rho\cdot 2^{-p} \leqq x < (\rho+1)2^{-p}$. By (9.32) the partial sums of order 2^{n+p}, $n\geqq 0$, are all zero in I_p. We shall prove that those of order $r\cdot 2^p$, $r \geqq 0$, also vanish. We have

$$\begin{aligned} s_{r\cdot 2^p}(x) &= \sum_{q=0}^{r-1} \sum_{m=0}^{2^p-1} a_{q\cdot 2^p+m}\psi_{q\cdot 2^p+m}(x) \\ &= \sum_{q=0}^{r-1} \psi_q(2^p x) \sum_{m=0}^{2^p-1} a_{q\cdot 2^p+m}\psi_m(x). \end{aligned} \tag{9.33}$$

Since $\psi_m(x)$ is constant on I_p for $m<2^p$, we are led to consider the series

$$\begin{aligned} S^*(y) &= \sum_{q=0}^{\infty} B_q\psi_q(y), \\ B_q &= \sum_{m=0}^{2^p-1} a_{q\cdot 2^p+m}\psi_m(\rho\cdot 2^{-p}). \end{aligned} \tag{9.34}$$

For $\rho \leqq y < \rho+1$, we may write $y=2^p x$, $x \in I_p$. Hence in that interval of y,

$$\overset{*}{s}_{2^n}(y) = s_{2^{n+p}}(x) = 0, \qquad x = 2^{-p}y \in I_p, \qquad n \geqq 0, \tag{9.35}$$

where, of course, $s_N^*(y)$ is the Nth partial sum of series (9.34). If $L^*(y)$ is the formal integral of $S^*(y)$, it follows from (9.35) that $L^*(y)=0$ in $\rho \leqq y<\rho+1$. Since $B_0 = s_1^*(y) = 0$, $L^*(y)$ vanishes identically. But $B_q \to 0$, since $a_k \to 0$. Applying Theorem XXIII, we find that $B_q = 0$, $q \geqq 0$. By (9.35), for $x \in I_p$,

$$\begin{aligned} 0 = \psi_{q\cdot 2^p}(x)B_q &= \sum_{m=0}^{2^p-1} a_{q\cdot 2^p+m}\psi_{q\cdot 2^p+m}(x) \\ &= s_{(q+1)2^p}(x) - s_{q\cdot 2^p}(x) \qquad (q \geqq 0). \end{aligned}$$

Hence $s_{q\cdot 2^p}(x) = 0$ for $q \geqq 0$, $x \in I_p$. If $N = q\cdot 2^p + M$, $0 \leqq M < 2^p$,

$$s_N(x) = s_{q\cdot 2^p}(x) + o(M) = o(1)$$

uniformly for $x \in I_p$.

COROLLARY A. *If* $S(x) = \sum_{k=0}^{\infty} a_k\psi_k(x)$, $S^*(x) = \sum_{k=0}^{\infty} a_k^*\psi_k(x)$, $a_k - a_k^* \to 0$, *and if* $L(x) = L^*(x) + C$ *in* $a < x < b$, *then* $S(x)$ *and* $S^*(x)$ *are uniformly equiconvergent in any subinterval* (α, β), $a < \alpha \leqq x < \beta < b$.

COROLLARY B. *If* $a_k \to 0$, *if* $L(x) = C + \int_a^x f(u)du$ *in* $a < x < b$, *and if* $f^*(x)$ *is defined as* $f(x)$ *in* $a < x < b$ (mod 1), *zero otherwise, then* $S(x)$ *and the* WFS *of* $f^*(x)$ *are uniformly equiconvergent in any sub-interval* (α, β), $a < \alpha \leqq x < \beta < b$.

It is quite likely that Theorem XXX could be proved by means of a theory of formal multiplication analogous to that of Rajchman in trigonometric series(10). The task would be greatly facilitated by the fact that the characteristic function of an interval I_p is a finite linear combination of Walsh functions.

BIBLIOGRAPHY

1. M. Kac, *Sur les fonctions* $2^n t - [2^n t] - 1/2$, J. London Math. Soc. vol. 13 (1938) pp. 131–134.

2. ———, *On the distribution of values of sums of the type* $\sum f(2^k t)$, Ann. of Math. vol. 47 (1946) pp. 33–49.

3. S. Kaczmarz, *Über ein Orthogonalsystem*, Comptes Rendus du premier congrès des mathématiciens des pays slaves, Warsaw, 1929, pp. 189–192.

4. S. Kaczmarz and H. Steinhaus, *Le systeme orthogonal de M. Rademacher*, Studia Math. vol. 2 (1930) pp. 231–247.

5. R. E. A. C. Paley, *A remarkable series of orthogonal functions*, Proc. London Math. Soc. vol. 34 (1932) pp. 241–279.

6. H. A. Rademacher, *Einige Sätze über Reihen von allgemeinen Orthogonalenfunktionen*, Math. Ann. vol. 87 (1922) pp. 112–138.

7. J. L. Walsh, *A closed set of normal orthogonal functions*, Amer. J. Math. vol. 55 (1923) pp. 5–24.

8. A. Zygmund, *Trigonometrical series*, Warsaw, 1935.

UNIVERSITY OF PENNSYLVANIA,
PHILADELPHIA, PA.

(10) See Zygmund [8, p. 279 et seq.].