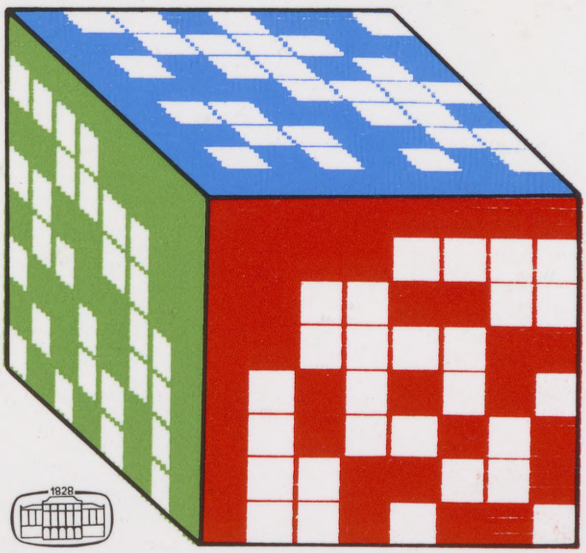
F. Schipp • W. R. Wade • P. Simon

with assistance from J. Pál

WALSH SERIES

an introduction to dyadic harmonic analysis



1828

WALSH SERIES

AN INTRODUCTION TO DYADIC HARMONIC ANALYSIS

WALSH SERIES

AN INTRODUCTION TO DYADIC HARMONIC ANALYSIS

F. Schipp, Eötvös L. University, Budapest
W. R. Wade, University of Tennessee, Knoxville
P. Simon, Eötvös L. University, Budapest

with assistance from
J. Pál, Eötvös L. University, Budapest

AKADÉMIAI KIADÓ, BUDAPEST 1990

ISBN 963 05 5880 7

Published in co-edition with Adam Hilger, Bristol and New York

Printed in Hungary
by Akadémiai Kiadó és Nyomda Vállalat, Budapest

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

Chapter 2

WALSH-FOURIER COEFFICIENTS

Chapter 3

DYADIC MARTINGALES AND HARDY SPACES

Chapter 4

CONVERGENCE IN NORM

Chapter 5

APPROXIMATION AND BASES

Chapter 6

ALMOST EVERYWHERE CONVERGENCE AND SUMMABILITY OF WALSH-FOURIER SERIES

Chapter 7

UNIQUENESS

Chapter 8

REPRESENTATION BY WALSH SERIES

Chapter 9

THE WALSH-FOURIER TRANSFORM

APPENDICES

PREFACE

The Walsh functions form an orthonormal system which can be applied in many situations. The Walsh system can perform all the usual applications of orthogonal systems (e.g., data transmission, multiplexing, filtering, image enhancement, and pattern recognition) and can perform them more efficiently. Also, they are easy to implement on high speed computers and can be used with very little storage space[1]. This is due in part to the fact that the Walsh functions take on only the values $+1$ and -1.

The Walsh system is also interesting from a theoretical point of view. First, it is the simplest non-trivial model for harmonic analysis but shares many properties with the trigonometric system. Secondly, it has been used to solve some fundamental problems in analysis (e.g., the basis problem). And third of all, it has played a role in the development of other areas of mathematics. For example, the three series theorem from probability was first discovered for the Rademacher system (a subset of the Walsh system) and the fundamental theorem of martingales was proved first by Paley for the Walsh system.

There are several books which describe applications of the Walsh system or give an introduction to certain specialized areas of Walsh analysis but there is no broadly based, theoretical monograph [2]. We attempt to fill this gap.

We give a thorough introduction to the foundations of Walsh-Fourier analysis including material on the dyadic group, the dyadic field, dyadic martingales, Walsh-Fourier coefficients and series, the Walsh-Fourier transform, the dyadic derivative, and various maximal operators. Our book does not cover all phases of the subject (for example lacunary series are barely mentioned). Instead we have tried to introduce the main techniques and fundamental problems of Walsh-Fourier analysis and make the literature accessible. In addition, we have attempted to show how the theory of Walsh-Fourier analysis relates to other aspects of harmonic analysis.

Concerning this last aim, we consider systems closely related to the Walsh system including the double Walsh system, product systems (including weakly and strongly multiplicative systems) and the systems of Haar, Faber-Schauder, Franklin, and Ciesielski. We do not give these systems a thorough treatment. Instead they are included to show how the techniques of Walsh-Fourier analysis can be used in a broader context to solve problems in related systems, to discuss equivalence of bases and the basis problem, and to obtain explicit isomorphisms between certain dyadic spaces and their classical counterparts.

While writing this book, we saw several concepts emerge which proved to be fundamental: quasi-measures, dyadic martingales, and non-linear sequence spaces (i.e., spaces of "sequences" of numbers or functions ordered by the tree-like collection of integer dyadic intervals rather than the linearly ordered collection of positive integers). These concepts both enriched and unified our subject.

First, there is a 1-1 correspondence between quasi-measures and the entire collection of Walsh series which allows certain problems to be recast as measure theoretic questions. In some cases this provided simple explanations of known results. For example, the fact that no Walsh series can diverge to $+\infty$ on a set of positive measure is a reflection of the fact that a quasi-measure is either a.e. differentiable or has upper derivate $+\infty$ and lower derivate $-\infty$ a.e. In other cases this provided new insight into the nature of the problem itself. For example, the existence of null series reduces to a question about certain measure preserving transformations.

Secondly, the non-linear sequence spaces and dyadic martingales unified the theory in a way not evident before. By introducing non-linear martingales and generalizing the Burkholder-Gundy theory of martingale transforms we see that the inequalities of Khintchin, Paley, and Sjölin as well as a.e. convergence of Walsh-Fourier series are all part of a general theory of non-linear martingale transforms. Moreover, by indexing the Haar and Franklin systems in a natural way to make them non-linear sequences, we see that the canonical isomorphism these bases induce an explicit isomorphism from the dyadic Hardy spaces and dyadic BMO to their classical (trigonometric!) counterparts. This gives a natural way to get classical theorems from dyadic ones and vice versa.

The book is organized as follows. The first two chapters are introductory. The reader interested in applications should read these together with Chapter 9 which deals with Walsh transforms. Chapter 3

[1] To save weight they were used in the Mariner spacecraft which explored the surface of Mars.

[2] Since this was written we have become aware of a book published in Russian by "Nauka". It is "Walsh Series and Transforms" by B.I. Golubov, A.V. Efimov, and V.A. Skvorcov.

introduces the probabalistic techniques, the dyadic Hardy spaces, and gives a self-contained proof of a.e. convergence of Walsh-Fourier series of functions in L^p for $p > 1$. Chapters 4 and 6 continue to investigate convergence of Walsh-Fourier series, both in norm and pointwise on the interval $[0, 1)$. Chapter 5 examines approximation by Walsh polynomials and various questions about bases. It contains the explicit isomorphism between classical and dyadic Hardy spaces mentioned above. Chapters 7 and 8 examine some problems concerning general Walsh series, namely uniqueness and representation.

These chapters are followed by Appendices which contain material of a classical nature and a brief introduction to Vilenkin systems. We also include historical notes for each chapter which give bibliographic references and discuss generalizations to Vilenkin and other closely related systems. The bibliography is long but not complete. More information of this type can be found in the papers cited at the beginning of the historical notes. The index is thorough and is meant to be used as a supplement to the Table of Contents to locate all results in the book which pertain to a given concept.

We would like to thank the Hungarian Academy of Sciences, the National Science Foundation, Eötvös Loránd University, and the University of Tennessee for the opportunity to collaborate on this book. We would like to thank Professors K. Tandori and F. Móricz for carefully reading the manuscript and serving as referees. We also acknowledge the help of S. Fridli in preparing the historical notes for the first four chapters and helping with the final proofreading, Rosi Templin for typing an earlier version of several chapters, and Cindi Blair, who typed, retyped, and in some cases, typed a third time various portions of this manuscript. Finally, we would like to thank Benjamin, who put the final version into a micro-computer and caught many typographical errors which still survived at that point.

Budapest and Knoxville, March 24, 1988

Chapter 1

INTRODUCTION

1.1 The Walsh Functions. We shall denote the set of *non-negative integers* by $\mathbf{N}$, the set of *positive integers* by $\mathbf{P}$, the set of *real numbers* by $\mathbf{R}$, the *complex plane* by $\mathbf{C}$, and the set of *dyadic rationals* in the unit interval $[0,1)$ by $\mathbf{Q}$. In particular, each element of $\mathbf{Q}$ has the form $p/2^n$ for some $p, n \in \mathbf{N}$, $0 \le p < 2^n$.

We shall use the notation $x = (x_\alpha, \alpha \in A)$ to represent a collection x indexed by a set A. Thus a sequence will be represented in the form $(x_n, n \in \mathbf{N})$.

Let r be the function defined on $[0,1)$ by

(1)
$$r(x) := \begin{cases} 1 & x \in [0, \frac{1}{2}) \\ -1 & x \in [\frac{1}{2}, 1) \end{cases}$$

extended to $\mathbf{R}$ by periodicity of period 1. The *Rademacher system* $\boldsymbol{r} := (r_n, n \in \mathbf{N})$ is defined by

(2)
$$r_n(x) := r(2^n x) \qquad (x \in \mathbf{R},\, n \in \mathbf{N}).$$

Given $n \in \mathbf{N}$ it is possible to write n uniquely as

(3)
$$n = \sum_{k=0}^{\infty} n_k 2^k$$

where $n_k = 0$ or 1 for $k \in \mathbf{N}$. This expression will be called the *binary expansion* of n and the numbers n_k will be called the *binary coefficients* of n.

In the literature, the term "Walsh functions" refers to one of three orthonormal systems: the *Walsh-Paley system* (which we shall call the *Walsh system*), the *original Walsh system*, or the *Walsh-Kaczmarz system.* These systems contain the same functions and differ only in enumeration. Each is a complete orthonormal system on $[0,1)$, contains the Rademacher system, and enjoys many properties analogous to the classical trigonometric, Sturm-Liouville, and Legendre systems.

The Walsh(-Paley) system $\boldsymbol{w} := (w_n, n \in \mathbf{N})$ was introduced by Paley [1] in 1932 as products of Rademacher functions in the following way. If $n \in \mathbf{N}$ has binary coefficients $(n_k, k \in \mathbf{N})$ then

(4)
$$w_n := \prod_{k=0}^{\infty} r_k^{n_k}.$$

Notice that this product is always finite because $n_k = 0$ for k sufficiently large. Notice also by definition that $w_0 = 1$ and $w_{2^n} = r_n$ for $n \in \mathbf{N}$. Moreover, it is clear that the

Walsh system is closed under finite products, each Walsh function is piecewise constant with finitely many jump discontinuities on [0,1), and takes only the values +1 or −1.

The original Walsh system $\boldsymbol{\phi} := (\phi_n, n \in \mathbf{N})$ was introduced by Walsh [1] in 1923. His definition was recursive and essentially the following one. Set

$$\phi_0 := 1 \quad \text{and} \quad \phi_1 := r_0.$$

For each integer $n \geq 2$ choose $m, k \in \mathbf{P}$ such that $1 \leq k \leq 2^{m-1}$ and $n = 2^{m-1} + k - 1$. Set

$$\phi_n := \phi_m^{(k)}$$

where the functions $\phi_m^{(k)}$ are periodic of period 1 and generated recursively by the following process. Define

$$\phi_1^{(1)} := \phi_1$$

and

$$\phi_2^{(k)}(x) := \begin{cases} \phi_1^{(1)}(2x) & x \in [0, \frac{1}{2}) \\ (-1)^k \phi_1^{(1)}(2x) & x \in [\frac{1}{2}, 1) \end{cases}$$

for $k = 1, 2$. If $m = 2, 3, \ldots$ and $1 \leq k \leq 2^{m-1}$ then define

$$\phi_{m+1}^{(2k-1)}(x) := \begin{cases} \phi_m^{(k)}(2x) & x \in [0, \frac{1}{2}) \\ (-1)^{k+1} \phi_m^{(k)}(2x) & x \in [\frac{1}{2}, 1) \end{cases}$$

and

$$\phi_{m+1}^{(2k)}(x) := \begin{cases} \phi_m^{(k)}(2x) & x \in [0, \frac{1}{2}) \\ (-1)^k \phi_m^{(k)}(2x) & x \in [\frac{1}{2}, 1). \end{cases}$$

The Walsh-Kaczmarz system $\boldsymbol{\kappa} := (\kappa_n, n \in \mathbf{N})$ was introduced by Šneider [1] in 1948. He also used products of Rademacher functions but differently from Paley. He began with

$$\kappa_0 := 1$$

but for each $n \in \mathbf{P}$ he defined

$$\kappa_n := r_m \prod_{k=0}^{m-1} r_k^{n_{m-k-1}}, \tag{5}$$

where $m \in \mathbf{N}$ satisfies $2^m \leq n < 2^{m+1}$ and the n_k's are the binary coefficients of n. Notice that $\kappa_{2^n} = r_n = w_{2^n}$ for $n \in \mathbf{N}$ but other than this the ordering of the system $\boldsymbol{\kappa}$ is quite different from that of $\boldsymbol{w}$. Nevertheless, it is obvious that the Walsh- Kaczmarz system is a rearrangement of the Walsh system. (It is not so easy to see that the original Walsh system is yet a third enumeration. We shall take this up in 1.4 below.)

Of the three enumerations, the original Walsh system $\boldsymbol{\phi}$ is used for most applications. This is so because of its recursive definition and because each ϕ_n has exactly n jumps on $[0,1)$. On the other hand, the Walsh system $\boldsymbol{w}$ is used more frequently by theoreticians. This is so for several reasons: definition (4) is extremely tractable, the paper by Paley was important and widely read, and the Walsh system has a stronger analogy to the trigonometric system. In regard to this last statement, we shall see that many trigonometric theorems have exact analogues for the Walsh system but not for the Walsh-Kaczmarz system. For example, the Dirichlet kernels for the Walsh system and the trigonometric system have the same order of growth, but the Dirichlet kernels in the Walsh-Kaczmarz system grow more rapidly (see 1.6 below).

The theory of Walsh series is extremely rich and the purpose of this chapter is to develop relationships between the Walsh system and several other areas of mathematics and to introduce some of the basic techniques. In 1.2 we show that Walsh analysis is a special case of the study of harmonic analysis on compact abelian groups. In 1.3 we show that the Walsh system is a complete orthonormal system on $[0,1)$. In 1.4 we show how to get from one enumeration to the other, and relate the Walsh system to the Haar system, one which plays a central role in the general study of orthogonal functions (see Olevskiĭ[1]). In 1.5 and 1.6 we define Walsh- Fourier coefficients and Walsh-Fourier series and examine the Walsh-Dirichlet kernels. In 1.7 we introduce the dyadic derivative. And in 1.8 we discuss summability and the Walsh- Fejér kernels.

1.2 The Dyadic Group. In 1949 Fine [1] made the fundamental observation that the Walsh functions can be viewed as characters of the dyadic group. This fact, in somewhat more general form, had been formulated by Vilenkin [1] in 1947.

In this section we shall define the dyadic group, identify its characters, Haar measure, and topology, and prove several theorems concerning its structure and that of its dual. (For definitions and elementary facts about topological groups see Appendices 0.3 and 0.4.)

Let $\mathbf{Z}_2$ be the discrete cyclic group of order 2, i.e., the set $\{0,1\}$ with the discrete topology and modulo 2 addition. Clearly, $\mathbf{Z}_2$ is a compact abelian group. The *dyadic group* $\mathbf{G}$ is defined to be the compact abelian group formed by taking the cartesian product of countably many copies of $\mathbf{Z}_2$, say

(6) $$\mathbf{G} := \mathbf{Z}_2 \times \mathbf{Z}_2 \times \ldots.$$

Thus $\mathbf{G}$ consists of sequences $x = (x_n, n \in \mathbf{N})$ where $x_n = 0$ or 1. The zero element of $\mathbf{G}$ is the sequence $0 := (x_n := 0, n \in \mathbf{N})$ and the group operation is given by

$$x + y := (|x_n - y_n|, n \in \mathbf{N})$$

for any $x = (x_n, n \in \mathbf{N})$ and $y = (y_n, n \in \mathbf{N})$ in $\mathbf{G}$. In particular, each $x \in \mathbf{G}$ is its own inverse, i.e., $x + x = 0$.

The subset

$$\mathbf{G}_0 := \{x \in \mathbf{G} : x_n \to 0 \quad \text{as } n \to \infty\}$$

is a countable subgroup of $\mathbf{G}$. Since each component $x_n = 0$ or 1, the subgroup $\mathbf{G}_0$ consists of those $x \in \mathbf{G}$ such that $x_n = 0$ for n sufficiently large.

Set $I_0(x) := \mathbf{G}$ for all $x \in \mathbf{G}$. For each $x \in \mathbf{G}$ and $n \in \mathbf{P}$ define

(7) $$I_n(x) := \{y \in \mathbf{G} : y_i = x_i \text{ for } 0 \le i < n\}.$$

We shall call these sets the ***dyadic intervals*** of **G**. Since the topology of **G** is the product topology (see (6) above), it has a countable base given by the family

$$\{I_n(x) : x \in \mathbf{G}_0, n \in \mathbf{P}\}.$$

Thus the dyadic intervals play an important role for the dyadic group. Since the dyadic intervals are evidently both open and closed, the dyadic group is totally disconnected.

The dyadic group is metrizable. Indeed, consider the map from **G** onto $[0,1]$ defined by

$$|x| := |(x_k, k \in \mathbf{N})| = \sum_{k=0}^{\infty} x_k 2^{-(k+1)}. \tag{8}$$

In view of the remarks above, $|x| = 0$ if and only if $x = 0$, and

$$\big|\, |x| - |y| \,\big| \le |x + y| \le |x| + |y|$$

for all $x, y \in \mathbf{G}$ i.e., the map $|\cdot|$ is a norm on **G**. Moreover,

$$\{y \in \mathbf{G} : |x + y| < 2^{-n}\} \subset I_n(x) \subset \{y \in \mathbf{G} : |x + y| \le 2^{-n}\} \tag{9}$$

for $n \in \mathbf{P}$ and $x \in \mathbf{G}$. Thus the dyadic group is metrizable with the distance between two points $x, y \in \mathbf{G}$ defined to be $|x + y|$.

It is useful to notice that the dyadic group is actually a vector space over the field $\mathbf{Z}_2$. Thus **G** is a normed linear space. It contains a closed system. Indeed, for each $i \in \mathbf{N}$ let e_i denote the element $(x_n, n \in \mathbf{N})$ which satisfies $x_i = 1$ and $x_n = 0$ for $n \neq i$. Notice that each element of $\mathbf{G}_0$ is a finite linear combination of the e_i's and that $\mathbf{G}_0$ is dense in **G**. Thus $\boldsymbol{e} := (e_i, i \in \mathbf{N})$ is a closed system in the normed linear space **G**. We shall call $\boldsymbol{e}$ the ***usual closed system.*** (See Exercise 1.14).

Define a measure on $\mathbf{Z}_2$ by assigning each singleton the measure $1/2$. Let μ represent the product measure on **G** inherited from $\mathbf{Z}_2$ (see Appendix 0.4). By definition

$$\mu(I_n(x)) = 2^{-n}$$

for $x \in \mathbf{G}$ and $n \in \mathbf{N}$. Thus

$$\mu(I_n(x) + y) = \mu(I_n(x))$$

for $x, y \in \mathbf{G}$ and $n \in \mathbf{N}$. Since the Borel sets of **G** are generated by the dyadic intervals in **G**, it is easy to see that μ is a translation invariant Borel measure on the dyadic group whose total variation is 1. In particular, μ is a Haar measure on **G**.

A ***character*** on **G** is a continuous complex-valued map which satisfies

$$f(x + y) = f(x)f(y) \quad \text{and} \quad |f(x)| = 1 \qquad (x, y \in \mathbf{G}). \tag{10}$$

The collection of characters on **G** is denoted by $\widehat{\mathbf{G}}$. Since $f \in \widehat{\mathbf{G}}$ implies

$$f(0) = 1 \quad \text{and} \quad f^2(x) = f(x + x) = f(0) \quad (x \in \mathbf{G}),$$

it follows that a character on the dyadic group can take only the values +1 or −1, i.e., each $f \in \widehat{\mathbf{G}}$ is real- valued. Since characters are continuous it is also clear that two characters coincide on $\mathbf{G}$ if and only if they coincide on the usual closed system $\mathbf{e} = (e_i, i \in \mathbf{N})$.

We shall describe the characters of $\mathbf{G}$ in two ways (see (12) and (17) below). For $n \in \mathbf{N}$ and $x = (x_n, n \in \mathbf{N}) \in \mathbf{G}$ set

$$\rho_n(x) := (-1)^{x_n}. \tag{11}$$

Since the sequence $(I_k(0), k \in \mathbf{N})$ forms a local base each ρ_n is continuous on $\mathbf{G}$. Moreover, it is clear that each ρ_n satisfies (10). Thus each ρ_n is a character on $\mathbf{G}$.

The collection $(\rho_n, n \in \mathbf{N})$ can be used to generate all characters of the dyadic group in the same way that the Rademacher functions generate the Walsh system (compare (12) below with (4) above).

THEOREM 1. *For each $n \in \mathbf{N}$ with binary expansion (3) let*

$$\psi_n := \prod_{k=0}^{\infty} \rho_k^{n_k}. \tag{12}$$

Then $\psi_n \in \widehat{\mathbf{G}}$.

Conversely, if $f \in \widehat{\mathbf{G}}$ then $f = \psi_n$ for some $n \in \mathbf{N}$.

PROOF. Since each ψ_n is a finite product of ρ_k's it is clear that $\psi_n \in \widehat{\mathbf{G}}$.

Conversely, let $f \in \widehat{\mathbf{G}}$. Since f is continuous on $\mathbf{G}$ and $e_i \to 0$ in $\mathbf{G}$ as $i \to \infty$, we have $f(e_i) \to f(0) = 1$ as $i \to \infty$. But f takes on only the values +1 or −1. Thus there exists an $M \in \mathbf{N}$, depending on f, such that $f(e_i) = 1$ for $i > M$. Consequently, there is a sequence $n_i = 0$ or 1 such that $n_i = 0$ for $i > M$ and $f(e_i) = (-1)^{n_i}$ for all $i \in \mathbf{N}$. Set

$$n := \sum_{i=0}^{\infty} n_i 2^i.$$

By definition we have

$$f(e_i) = (-1)^{n_i} = \psi(e_i).$$

Thus $f = \psi_n$ on the closed system $\mathbf{e}$ and it follows that $f = \psi_n$ on $\mathbf{G}$. ∎

Define the dyadic sum of a pair of integers $n, m \in \mathbf{N}$ by

$$n \oplus m := \sum_{k=0}^{\infty} |n_k - m_k| 2^k \tag{13}$$

where $(n_k, k \in \mathbf{N})$ and $(m_k, k \in \mathbf{N})$ are the binary coefficients of n and m. Notice that

$$\rho_k^{n_k + m_k} = \rho_k^{|n_k - m_k|} \qquad (k \in \mathbf{N}).$$

Thus we have by Theorem 1 that

$$\psi_n \psi_m = \psi_{n \oplus m}. \tag{14}$$

Hence $\widehat{\mathbf{G}}$ is an abelian group under pointwise multiplication. Moreover, the groups $\widehat{\mathbf{G}}$, $\mathbf{G}_0$ and $(\mathbf{N}, \oplus)$ are all isomorphic. An isomorphism from $\mathbf{N}$ to $\widehat{\mathbf{G}}$ is given by $n \to \psi_n$ and an isomorphism from $\mathbf{N}$ to $\mathbf{G}_0$ is given by

(15) $$n = \sum_{k=0}^{\infty} n_k 2^k \to (n_0, n_1, \dots).$$

In view of this last isomorphism, we shall use the same notation for an element $n = (n_k, k \in \mathbf{N})$ of $\mathbf{G}_0$ and an integer n of the form (3). This abuse of notation yields a concise formula for the characters of $\mathbf{G}$. Indeed, if

(16) $$\langle n, x \rangle := n_0 x_0 + n_1 x_1 + \dots \pmod 2$$

for $n \in \mathbf{G}_0$ and $x \in \mathbf{G}$, where multiplication on the right side of (16) is that of $\mathbf{Z}_2$, then (11) and (12) combine to produce

(17) $$\psi_n(x) = (-1)^{\langle n, x \rangle}$$

for $n \in \mathbf{N}$ and $x \in \mathbf{G}$. Notice that the map $\langle \cdot \rangle$ is a bilinear form from $\mathbf{G}_0 \times \mathbf{G}$ into $\mathbf{Z}_2$.

The collection $\widehat{\mathbf{G}}$ is an orthonormal system.

THEOREM 2. *If $n, m \in \mathbf{N}$ then*

$$\int_{\mathbf{G}} \psi_n \psi_m \, d\mu = \begin{cases} 1 & n = m \\ 0 & n \neq m. \end{cases}$$

PROOF. Since $\psi_n \psi_m = \psi_{n \oplus m}$ it suffices to show

$$\int_{\mathbf{G}} f \, d\mu = 0$$

for all $f \in \widehat{\mathbf{G}} \setminus \{\psi_0\}$.

Toward this fix such an f and choose $y \in \mathbf{G}$ such that $f(y) = -1$. By translation invariance

$$\begin{aligned} \int_{\mathbf{G}} f \, d\mu &= \int_{\mathbf{G}} f(x + y) \, d\mu(x) \\ &= f(y) \int_{\mathbf{G}} f(x) \, d\mu(x). \end{aligned}$$

Since $f(y) = -1$ this integral must vanish. ∎

An immediate corollary of Theorem 2 is the following important computation.

PALEY'S LEMMA. *If* $D_{2^n} := \sum_{k=0}^{2^n-1} \psi_k$ *then*

$$D_{2^n}(x) = \begin{cases} 2^n & x \in I_n(0) \\ 0 & x \in \mathbf{G} \setminus I_n(0) \end{cases}$$

for $n \in \mathbf{N}$.

PROOF. Since $\psi_k = 1$ on $I_n(0)$ for $0 \le k < 2^n$, it suffices to show that D_{2^n} vanishes off $I_n(0)$. But

$$\int_{I_n(0)} |D_{2^n}|^2 \, d\mu = 2^{2n} \mu(I_n(0)) = 2^n$$

and by orthogonality,

$$\int_{\mathbf{G}} |D_{2^n}|^2 \, d\mu = \sum_{k=0}^{2^n-1} \int_{\mathbf{G}} |\psi_k|^2 \, d\mu = 2^n.$$

Consequently, D_{2^n} must vanish off $I_n(0)$. ∎

The operation $\oplus$ satisfies the following partition property.

THEOREM 3. *For each* $k \in \mathbf{N}$ *let*

$$A_k := \{i \in \mathbf{N} : 2^k \le i < 2^{k+1}\}.$$

For each $n \in \mathbf{N}$ *with binary coefficients* n_k *let*

$$B_{n,k} := \begin{cases} \emptyset & n_k = 0 \\ n \oplus A_k & n_k = 1. \end{cases}$$

Then

$$\{j \in \mathbf{N} : 0 \le j < n\} = \bigcup_{k=0}^{\infty} B_{n,k}$$

for all $n \in \mathbf{N}$.

PROOF. Fix $n \in \mathbf{N}$. Since $B_{n,k} \cap B_{n,j} = \emptyset$ for $k \ne j$ and $\oplus$ enjoys the cancellation property, it is clear that the number of elements in

$$B_n := \bigcup_{k=0}^{\infty} B_{n,k}$$

is precisely

$$\sum_{k=0}^{\infty} n_k 2^k = n.$$

In particular, it suffices to show that $0 \le \ell < n$ for every $\ell \in B_n$.

Fix $\ell \in B_n$ and observe by construction that $\ell \in n \oplus A_k$ for some k which satisfies $n_k = 1$. Thus $\ell = n \oplus m$ where

$$m = 2^k + \sum_{j=0}^{k-1} m_j 2^j.$$

Consequently, the definition of dyadic addition implies

$$\ell = \sum_{j=0}^{k-1} |m_j - n_j| 2^j + |m_k - n_k| 2^k + \sum_{j=k+1}^{\infty} n_j 2^j$$

$$\leq 2^k - 1 + \sum_{j=k+1}^{\infty} n_j 2^j \leq n - 1 < n. \blacksquare$$

The abelian groups $\mathbf{G}$ and $(\mathbf{N}, \oplus)$ cannot be ordered in the usual algebraic sense (see Exercise 1.2). Nevertheless, given $x = (x_n, n \in \mathbf{N})$ and $y = (y_n, n \in \mathbf{N})$ in $\mathbf{G}$ we shall say that $x < y$ if there is an $n \in \mathbf{N}$ such that $x_n = 0, y_n = 1$ and $x_j = y_j$ for all $j < n$. By $x \leq y$ we shall mean $x < y$ or $x = y$. Notice that the relation $\leq$ is a linear ordering on $\mathbf{G}$.

It is evident that the map (8) is increasing, i.e., if $x \leq y$ in $\mathbf{G}$ then $|x| \leq |y|$. It is not strictly increasing. In fact, every nonzero element x in $\mathbf{G}_0$ has a conjugate element x^* in $\mathbf{G}$ such that $|x| = |x^*|$. Indeed, for each $x = (x_0, x_1, \ldots, x_m, 0, 0, \ldots)$ with $x_m = 1$ for some $m \in \mathbf{N}$ define

$$x^* := (x_0, x_1, \ldots, x_{m-1}, 0, 1, 1, \ldots).$$

Also define

$$0^* := (1, 1, \ldots).$$

It is clear that $|x| = |x^*|$ for all $x \in \mathbf{G}_0 \setminus \{0\}$. It is also easy to check that if $|x| = |y|$ then either $x = y$ or $x \in \mathbf{G}_0$ and $x = y^*$. (For a relationship between convergence, the ordering, and these conjugate elements see Exercise 1.5).

1.3 The Representation of the Dyadic Group on the Interval $[0,1]$. In this section we show that the dyadic group can be identified with the interval $[0,1]$ in such a way that the characters $\widehat{\mathbf{G}}$ correspond to the Walsh system $\boldsymbol{w}$, Haar measure μ corresponds to Lebesgue measure, the countable dense subgroup $\mathbf{G}_0$ corresponds to the dyadic rationals $\mathbf{Q}$, and the ordering introduced in the preceding section corresponds to the usual ordering on $[0,1]$. Thus investigation of the Walsh system may proceed in two directions:

a) use the system $\boldsymbol{w}$ and Lebesgue integration to study functions defined on $[0,1]$, or

b) use the characters $\widehat{\mathbf{G}}$ and Haar integration to study functions defined on the dyadic group.

A similar dichotomy prevails for classical Fourier series. One can investigate trigonometric series on $[0, 2\pi]$ or exponential series on the circle group $\mathbf{T}$. In Walsh analysis, the nature of the problem being considered sometimes dictates which direction to pursue. We shall use both points of view in subsequent chapters often with nothing more than expedience as our guide. In any event, this identification allows one to translate results from one system to the other and this task will frequently be left to the reader.

Any $x \in [0,1]$ can be written in the form

$$x = \sum_{k=0}^{\infty} x_k 2^{-(k+1)}, \tag{18}$$

where each $x_k = 0$ or 1. For each $x \in [0,1] \setminus \mathbf{Q}$ there is only one expression of this form. We shall call it the *dyadic expansion* of x. When $x \in \mathbf{Q}$ there are two expressions of this form, one which terminates in 0's and one which terminates in 1's. By the dyadic expansion of an $x \in \mathbf{Q}$ we shall mean the one which terminates in 0's. Notice that $1 \notin \mathbf{Q}$ so the dyadic expansion of $x = 1$ terminates in 1's.

Let $\mathbf{G}_0^* := \{x \in \mathbf{G} : x = y^* \text{ for some } y \in \mathbf{G}_0\}$. Define Fine's map $\varrho : [0,1] \to \mathbf{G}$ by

$$\varrho(x) := (x_0, x_1, \dots) \tag{19}$$

where x has dyadic expansion (18). Then ϱ is a strictly increasing, 1-1 map from $[0,1)$ onto $\mathbf{G} \setminus \mathbf{G}_0^*$. Moreover, it is easy to prove that ϱ satisfies

$$\begin{cases} \varrho(x+) = \varrho(x-) = \varrho(x) & x \in (0,1) \setminus \mathbf{Q} \\ \varrho(x+) = \varrho(x),\ \varrho(x-) = \varrho^*(x) & x \in \mathbf{Q} \\ \varrho(0+) = 0,\ \varrho(1-) = 0^*. \end{cases} \tag{20}$$

Here $\varrho(x+)$ represents the limit of ϱ at x from the right in the usual topology on $[0,1]$ and $\varrho(x-)$ that from the left.

Let $C(\mathbf{G})$ represent the collection of functions $f : \mathbf{G} \to \mathbf{R}$ which are continuous from the dyadic topology on $\mathbf{G}$ to the usual topology on $\mathbf{R}$. Let C_W represent the collection of functions $g : [0,1) \to \mathbf{R}$ which are continuous at every dyadic irrational, continuous from the right on $[0,1)$, and have a finite limit from the left on $(0,1]$, all this in the usual topology.

We shall call the map $f \to f \circ \varrho$ the canonical isomorphism. It is easy to see that this map is a vector space isomorphism from $C(\mathbf{G})$ onto C_W. First, it is clear by (20) that if $f \in C(\mathbf{G})$ and $g := f \circ \varrho$ then

$$\begin{cases} g(x+) = g(x-) = g(x) & x \in (0,1) \setminus \mathbf{Q} \\ g(x+) = g(x),\ g(x-) = f(\varrho^*(x)) & x \in \mathbf{Q} \\ g(0+) = f(0),\ g(1-) = f(0^*). \end{cases} \tag{21}$$

Thus the canonical isomorphism takes $C(\mathbf{G})$ into C_W. Next, notice by construction that the canonical isomorphism is a vector space homomorphism, i.e., preserves function addition and scalar multiplication. Finally, if $g \in C_W$ then the map f defined on $\mathbf{G}$ by

$$\begin{cases} f(y) := g(x) & y = \varrho(x),\ x \in [0,1) \setminus \mathbf{Q} \\ f(y^*) := g(x-) & y = \varrho(x),\ x \in \mathbf{Q} \\ f(0^*) := g(1-) & \end{cases} \tag{22}$$

is continuous on $\mathbf{G}$ (see Exercise 1.7).

It is also easy to see that the canonical isomorphism takes the character system $\widehat{\mathbf{G}}$ onto the Walsh system $\boldsymbol{w}$. Indeed if $x \in [0,1)$ has dyadic expansion (18) then the definition of the Rademacher functions given in 1.1 can be rewritten as

$$r_n(x) = (-1)^{x_n} \qquad (n \in \mathbf{N}).$$

Comparing this with (11) in 1.2, we see that $r_n = \rho_n \circ \boldsymbol{\varrho}$ on $[0,1)$, and $\rho_n(x) = r_n(|x|)$ for $x \in \mathbf{G} \setminus \mathbf{G}_0^*$ and every $n \in \mathbf{N}$. It follows from (4) and (12) that

$$w_n = \psi_n \circ \boldsymbol{\varrho}$$

and

$$\psi_n(x) = w_n(|x|) \qquad (x \in \mathbf{G} \setminus \mathbf{G}_0^*)$$

for every $n \in \mathbf{N}$.

Fine's map can be used to define a new addition and a new topology on $[0,1)$ which are closely related to those on $\mathbf{G}$. Indeed, define the *dyadic sum* of two numbers $x, y \in [0,1)$ and the *dyadic distance* between these numbers by

$$x \dot{+} y := |\boldsymbol{\varrho}(x) + \boldsymbol{\varrho}(y)|.$$

In terms of the dyadic expansions of x and y (see (18) above) we have

$$x \dot{+} y = \sum_{k=0}^{\infty} |x_k - y_k| \, 2^{-(k+1)}.$$

Hence $\dot{+}$ is evidently a metric and a commutative binary operation on $[0,1)$ which satisfies $x \dot{+} x = 0$. We shall call the topology generated by $\dot{+}$ on $[0,1)$ the dyadic topology. Note, $[0,1)$ is not a group under $\dot{+}$ (see Exercise 1.4).

The Walsh functions almost behave like characters with respect to dyadic addition, namely,

$$w_n(x \dot{+} y) = w_n(x) w_n(y) \quad (n \in \mathbf{N},\ x, y \in [0,1),\ x \dot{+} y \notin \mathbf{Q}). \tag{23}$$

(This identity will be used many times in subsequent chapters.) To prove (23) fix $m \in \mathbf{N}$ and $x, y \in [0,1)$. Notice that

$$r_m(x) r_m(y) = \rho_m \circ \boldsymbol{\varrho}(x)\, \rho_m \circ \boldsymbol{\varrho}(y) = \rho_m(\boldsymbol{\varrho}(x) + \boldsymbol{\varrho}(y)),$$

and that

$$r_m(x \dot{+} y) = \rho_m(\boldsymbol{\varrho}(|\boldsymbol{\varrho}(x) + \boldsymbol{\varrho}(y)|)).$$

Since $|\boldsymbol{\varrho}(x) + \boldsymbol{\varrho}(y)|$ is a dyadic irrational when $x \dot{+} y$ is, it is clear that

$$\boldsymbol{\varrho}(|\boldsymbol{\varrho}(x) + \boldsymbol{\varrho}(y)|) = \boldsymbol{\varrho}(x) + \boldsymbol{\varrho}(y)$$

for $x \dot{+} y \notin \mathbf{Q}$. Consequently, (23) holds for the Rademacher case, i.e., for $n = 2^m$. But the general case follows immediately since Walsh functions are finite products of Rademacher functions. Since for each fixed $y \in [0,1)$ the set of points x which satisfy $x \dot{+} y \in \mathbf{Q}$ is a countable set, we observe that (23) holds for a.e. $x, y \in [0,1)$.

By a dyadic interval in $[0,1)$ we shall always mean an interval of the form

$$(24) \qquad I(p,n) := [p2^{-n}, (p+1)2^{-n}) \qquad (0 \le p < 2^n,\ n, p \in \mathbf{N}).$$

Clearly, the dyadic topology is generated by the collection of dyadic intervals. Moreover, each dyadic interval is both open and closed in the dyadic topology. It follows that each Walsh function is continuous in the dyadic topology. Thus the dyadic topology differs from the usual topology in an essential way. (See also Exercise 1.11.)

For each $x \in [0,1)$ and $n \in \mathbf{N}$ we shall denote the dyadic interval of length 2^{-n} which contains x by $I_n(x)$. Thus

$$I_n(x) := I(p,n)$$

where $0 \le p < 2^n$ is uniquely determined by the relationship $x \in I(p,n)$. This is the same notation used for dyadic intervals on the group (compare with (7)) but will not cause problems because context will make it clear whether the dyadic interval is in the group or inside the unit interval.

A function $f : [0,1) \to \mathbf{R}$ which is continuous from the dyadic topology to the usual topology will be called *W-continuous*. Since

$$|x - y| \le x \dot{+} y \qquad (x, y \in [0,1)),$$

it is clear that every classically continuous function on [0,1) is W- continuous. In fact, every function in C_W is uniformly W-continuous on the unit interval. On the other hand, not every W-continuous function belongs to C_W (see Exercise 1.12).

Let L^0 represent the collection of a.e. finite, Lebesgue measurable functions from $[0,1)$ into $[-\infty, \infty]$. For $0 < p < \infty$ let L^p represent the collection of $f \in \mathrm{L}^0$ for which

$$\|f\|_p := \left(\int_0^1 |f|^p \right)^{1/p}$$

is finite. Let L^∞ represent the collection of $f \in \mathrm{L}^0$ for which

$$\|f\|_\infty := \inf\{y \in \mathbf{R} : |f(x)| \le y \text{ for a.e. } x \in [0,1)\}$$

is finite. It is well known that L^p is a Banach space for each $1 \le p \le \infty$.

For any set E denote the characteristic function of E by $\chi(E)$, i.e.,

$$\chi(E)(x) := \begin{cases} 1 & x \in E \\ 0 & x \notin E. \end{cases}$$

By a *dyadic step function* we shall mean a finite linear combination of characteristic functions of dyadic intervals in $[0,1)$. By a *Walsh polynomial* we shall mean a finite linear combination of Walsh functions. We shall denote the collection of Walsh polynomials by $\mathcal{P}$.

Since w_m is constant on $I(p,n)$ for each $0 \le p < 2^n$ and $0 \le m < 2^n$, it is clear that each Walsh polynomial is a dyadic step function. On the other hand, notice by the Paley lemma that

$$\chi(I_n(t))(x) = 2^{-n} \sum_{k=0}^{2^n-1} w_k(x \dot{+} t)$$

for each $x, t \in [0,1)$ and $n \in \mathbf{N}$. Thus each dyadic step function is a Walsh polynomial. It follows that the collection of dyadic step functions coincides with the collection of Walsh polynomials. Since Lebesgue measure is regular, it follows from Lusin's theorem that given $f \in \mathrm{L}^0$ there exist Walsh polynomials $P_1, P_2, \ldots$ such that $P_n \to f$ a.e. as $n \to \infty$. Moreover, any $f \in \mathrm{L}^1$ can be written in the form $f = g - h$ where the functions g, h are a.e. limits of increasing sequences of non-negative Walsh polynomials. In particular, $\mathcal{P}$ is dense in L^p for $1 \le p < \infty$.

On the group, let $\mathcal{P}$ represent the collection of finite linear combinations of characteristic functions of the dyadic intervals $\{I_n(x) : x \in \mathbf{G}_0, n \in \mathbf{N}\}$. Let $\mathrm{L}^0(\mathbf{G})$ represent the collection of functions which are a.e. $[\mu]$ limits of sequences in $\mathcal{P}$. For $0 < p < \infty$, let $\mathrm{L}^p(\mathbf{G})$ represent the collection of $f \in \mathrm{L}^0(\mathbf{G})$ such that

$$\|f\|_p := \left(\int_{\mathbf{G}} |f|^p \, d\mu \right)^{1/p}$$

is finite. Let $\mathrm{L}^\infty(\mathbf{G})$ represent the collection of functions $f \in \mathrm{L}^0(\mathbf{G})$ such that

$$\|f\|_\infty := \inf\{y \in \mathbf{R} : |f(x)| \le y \text{ for a.e. } [\mu]\ x \in \mathbf{G}\}$$

is finite. Notice that $\mathrm{L}^p(\mathbf{G})$ is a Banach space for each $1 \le p \le \infty$ and that $\mathcal{P}$ is dense in $\mathrm{C}(\mathbf{G})$ and in $\mathrm{L}^p(\mathbf{G})$ for $1 \le p < \infty$.

Let ℓ^0 represent the collection of sequences $\boldsymbol{a} = (a_n,\ n \in \mathbf{N})$ such that $a_n \in \mathbf{R}$. For each $0 < p < \infty$ let ℓ^p represent the collection of $\boldsymbol{a} \in \ell^0$ such that

$$\|\boldsymbol{a}\|_p := \left(\sum_{n=0}^{\infty} |a_n|^p \right)^{1/p}$$

is finite. Let ℓ^∞ represent the collection of $\boldsymbol{a} \in \ell^0$ such that

$$\|\boldsymbol{a}\|_\infty := \sup_{n \in \mathbf{N}} |a_n|$$

is finite. Recall that ℓ^p is a Banach space for each $1 \le p \le \infty$.

It is easy to see that

$$\sum_{k=0}^{n} a_k b_k = \sum_{k=0}^{n-1} \left(\sum_{j=0}^{k} a_j\right)(b_k - b_{k+1}) + b_n \sum_{j=0}^{n} a_j$$

holds for all $(a_k,\ k \in \mathbf{N}), (b_k,\ k \in \mathbf{N}) \in \ell^0$ and $n \in \mathbf{N}$. This algebraic identity will be referred to as ***Abel's transformation***.

Translation on the group and ***dyadic translation*** on the unit interval will both be denoted by $\boldsymbol{\tau}$. Thus if f is defined on $\mathbf{G}$ and $y \in \mathbf{G}$ then

$$(\boldsymbol{\tau}_y f)(x) := f(x+y) \quad (x \in \mathbf{G}),$$

and if f is defined on $[0,1)$ then

$$(\boldsymbol{\tau}_y f)(x) := f(x \dot{+} y) \quad (x \in [0,1)).$$

For any subset E of the unit interval, define the translation of E by y by

$$\boldsymbol{\tau}_y(E) := \{x \dot{+} y : x \in E\}.$$

When $E \subseteq [0,1)$ is measurable let $|E|$ represent its Lebesgue measure. The following result shows that Lebesgue measure is translation invariant on $[0,1)$ with respect to dyadic addition.

THEOREM 4. *Let $y \in [0,1)$. If $E \subseteq [0,1)$ is measurable then $\boldsymbol{\tau}_y(E)$ is measurable and*

$$|\boldsymbol{\tau}_y(E)| = |E|.$$

If $f \in \mathrm{L}^1$ then $\boldsymbol{\tau}_y f \in \mathrm{L}^1$ and

$$\int_0^1 \boldsymbol{\tau}_y f = \int_0^1 f. \tag{25}$$

PROOF. Let $I = [0, 2^{-n})$ for some $n \in \mathbf{N}$ and let $\sum_{k=0}^{\infty} y_k 2^{-(k+1)}$ represent the dyadic expansion of y. Choose $p \in \mathbf{N}$ such that $I(p,n) = I_n(y)$ and define

$$\bar{y}_k := \begin{cases} 0 & 0 \le k < n \\ |1 - y_k| & k \ge n. \end{cases}$$

Set $\bar{y} := \sum_{k=0}^{\infty} \bar{y}_k 2^{-(k+1)}$ and observe by the definition of dyadic addition that

$$\boldsymbol{\tau}_y(I) = \begin{cases} [p2^{-n}, (p+1)2^{-n}) & \bar{y} \notin I \\ [p2^{-n}, (p+1)2^{-n}] & \bar{y} \in I. \end{cases}$$

It follows that $|\boldsymbol{\tau}_y(I)| = |I|$ for all dyadic intervals I. Thus (25) holds for all $f \in \mathcal{P}$. By the monotone convergence theorem and remarks above, it follows that (25) holds for all $f \in \mathrm{L}^1$. Specializing to the case $f = \chi(E)$ we conclude that $|\boldsymbol{\tau}_y(E)| = |E|$. ∎

In view of the identification of $\mathbf{G}$ with $[0,1)$, it is natural to enquire about the relationship between Haar integration on the group and Lebesgue integration on the unit interval. The following result, whose proof is similar to Theorem 4, gives a complete answer.

THEOREM 5. *Let ϱ denote Fine's map.*

i) If $f \in L^0(G)$ then $f \circ \varrho \in L^0$. Conversely, if $g \in L^0$ and

$$f(x) := g(|x|) \qquad (x \in G) \tag{26}$$

then $f \in L^0(G)$.

ii) If $f \in L^1(G)$ then $f \circ \varrho \in L^1$, and

$$\int_G f\,d\mu = \int_0^1 f \circ \varrho.$$

Conversely, if $g \in L^1$ and f is defined by (26) then $f \in L^1(G)$ and

$$\int_0^1 g = \int_G f\,d\mu.$$

It follows that the canonical isomorphism $f \to f \circ \varrho$ is an isometry from $L^p(G)$ onto L^p for $1 \le p < \infty$.

The *modulus of continuity* of an $f \in C(G)$ is defined by

$$\omega(f,\delta) := \sup\{\|\tau_y f - f\|_\infty : y \in G, |y| < \delta\}, \tag{27}$$

for $\delta > 0$. The *L^p modulus of continuity* of an $f \in L^p(G)$ is defined by

$$\omega^{(p)}(f,\delta) := \sup\{\|\tau_y f - f\|_p : y \in G, |y| < \delta\}, \tag{28}$$

for $\delta > 0$ and $1 \le p < \infty$.

For $\alpha > 0$ denote by $\mathrm{Lip}(\alpha, G)$ the collection of functions $f \in C(G)$ which satisfy

$$|f(x+y) - f(x)| \le C|y|^\alpha \qquad (x, y \in G), \tag{29}$$

where C is a non-zero constant which depends only on f. Notice that $f \in \mathrm{Lip}(\alpha, G)$ implies $\omega(f,\delta) = O(\delta^\alpha)$ as $\delta \to 0$. Thus for each $\alpha > 0$ and $1 \le p < \infty$, we shall denote by $\mathrm{Lip}(\alpha, L^p(G))$ the collection of functions $f \in L^p(G)$ which satisfy

$$\omega^{(p)}(f,\delta) = O(\delta^\alpha) \quad \text{as} \quad \delta \to 0.$$

Analogously, for each $\delta > 0$ define the *dyadic modulus of continuity* of an $f \in C_W$ by

$$\omega(f,\delta) := \sup\{|f(x \dot{+} y) - f(x)| : x, y \in [0,1), 0 \le y < \delta\},$$

and the *dyadic L^p modulus of continuity* of an $f \in L^p$ by

$$\omega^{(p)}(f,\delta) := \sup\{\left(\int_0^1 |f(x \dot{+} y) - f(x)|^p\,dx\right)^{1/p} : y \in [0,1), 0 \le y < \delta\}.$$

For each $\alpha > 0$ let $\mathrm{Lip}(\alpha, W)$ denote image of $\mathrm{Lip}(\alpha, \mathbf{G})$ under the canonical isomorphism and $\mathrm{Lip}(\alpha, \mathrm{L}^p)$ denote the image of $\mathrm{Lip}(\alpha, \mathrm{L}^p(\mathbf{G}))$ under the canonical isomorphism. For example, a function f belongs to $\mathrm{Lip}(\alpha, W)$ if and only if $f = g \circ \boldsymbol{\varrho}$ for some $g \in \mathrm{Lip}(\alpha, \mathbf{G})$. In particular, if $f \in \mathrm{Lip}(\alpha, W)$ then there is a constant $C > 0$ such that

$$|f(x \dot{+} y) - f(x)| \leq C|y|^\alpha$$

for all $x, y \in [0,1)$ which satisfy $\boldsymbol{\varrho}(x \dot{+} y) = \boldsymbol{\varrho}(x) + \boldsymbol{\varrho}(y)$.

It should be pointed out that these definitions do not satisfy all the usual properties. For example, for each $n \in \mathbf{N}$ the condition $\delta < 2^{-n}$ implies $\omega(\rho_n, \delta) = 0$. Thus $\omega(f, 2\delta) \leq 2\omega(f, \delta)$ does not hold for all f. Moreover, $\mathrm{Lip}(\alpha, \mathbf{G})$ is non-trivial for $\alpha > 1$.

For functions defined on the unit interval the *local moduli of continuity* are defined as follows. For each dyadic interval I and each W-continuous function f let

$$\omega(f, I) := \sup\{|f(x \dot{+} y) - f(x)| : x \in I, 0 \leq y < |I|\}.$$

For each $1 \leq p < \infty$ and each $f \in \mathrm{L}^p$ let

$$\omega^{(p)}(f, I) := \sup_{0 \leq y < |I|} \left(\frac{1}{|I|} \int_I |f(x \dot{+} y) - f(x)|^p \, dx \right)^{1/p}.$$

The local modulus of continuity $\omega(f, I)$ is a measure of the oscillation of f on I. Thus we say that a function f defined on the unit interval is of *p-bounded fluctuation* for some $1 \leq p < \infty$ if

$$\sup_{n \in \mathbf{P}} \left(\sum_{k=0}^{2^n - 1} |\omega(f, I(k, n))|^p \right)^{1/p} < \infty.$$

A function is said to be of *bounded fluctuation* if it is of 1-bounded fluctuation. In this case its *total fluctuation* is defined by

$$\mathcal{F}\ell(f) := \sup_{n \in \mathbf{P}} \left(\sum_{k=0}^{2^n - 1} |\omega(f, I(k, n))| \right).$$

Clearly, every function of bounded variation on $[0,1)$ is also of bounded fluctuation but not conversely.

1.4 Transformations and Rearrangements of the Walsh System. For the first half of this section, let $(\Omega, \mathcal{G}, \nu)$ be a fixed probability space. By a measure preserving transformation on Ω we mean a map $T' : \Omega \to \mathbf{G}$ which satisfies

$$\nu(\{t \in \Omega : T'(t) \in E\}) = \mu(E) \tag{30}$$

for all measurable $E \in \mathbf{G}$. Given any sequence $\boldsymbol{f} = (f_n, n \in \mathbf{N})$ in $\mathrm{L}^1(\mathbf{G})$ we shall denote the sequence $(f_n \circ T', n \in \mathbf{N})$ by $\boldsymbol{f} \circ T'$. Notice that

$$\int_\Omega f \circ T' \, d\nu = \int_{\mathbf{G}} f \, d\mu$$

for all $f \in L^1(G)$ and all measure preserving transformations T'. Thus if $\boldsymbol{f}$ is an orthogonal system in $L^2(G)$ and T' is measure preserving then $\boldsymbol{f} \circ T'$ is orthogonal in $L^2_\nu(\Omega)$.

Let $\boldsymbol{\gamma} = (\gamma_n, n \in \mathbf{N})$ be a sequence of bounded, $\mathcal{G}$-measurable functions on Ω. The *product system* generated by $\boldsymbol{\gamma}$ is the system $\boldsymbol{g} = (g_n, n \in \mathbf{N})$ defined by

$$g_m := \prod_{n=0}^{\infty} \gamma_n^{m_n},$$

where the numbers m_n are the binary coefficients of m. The system $\boldsymbol{\gamma}$ is called *multiplicative* if $\int_\Omega g_m \, d\nu = 0$ for $m \in \mathbf{P}$ and called *strongly multiplicative* if the system $(g_m, m \in \mathbf{N})$ is orthogonal. Clearly, every strongly multiplicative system is multiplicative. On the other hand, there exist systems which are multiplicative but not strongly multiplicative. (See Exercise 1.28.)

The system $\boldsymbol{\gamma}$ is called *sign-like* if each γ_n takes on only the values $+1$ or -1. Notice that the Rademacher system $\boldsymbol{r} := (r_n, n \in \mathbf{N})$ is a sign-like, strongly multiplicative system. Also notice for sign-like systems that $\boldsymbol{\gamma}$ is strongly multiplicative if and only if it is multiplicative.

An important class of sign-like multiplicative systems is the collection of subsystems of the Walsh system generated by $\mathbf{Z}_2$-linearly independent indices. Indeed, let $\ell_0, \ell_1, \ldots$ be $\mathbf{Z}_2$-linearly independent. By definition if

$$\epsilon_0 \ell_0 \oplus \cdots \oplus \epsilon_n \ell_n = 0$$

for some $n \in \mathbf{N}$ and numbers $\epsilon_i \in \mathbf{Z}_2 \quad (i = 0, 1, \ldots, n)$ then $\epsilon_0 = \cdots = \epsilon_n = 0$. Consequently, the system $(w_{\ell_k}, k \in \mathbf{N})$ is sign- like and multiplicative.

Let $\boldsymbol{g} = (g_n, n \in \mathbf{N})$ be a system in $L^1_\nu(\Omega)$ and T be a map from $\mathbf{N}$ to $\mathbf{N}$. We shall denote the system $(g_{T(n)}, n \in \mathbf{N})$ by $T\boldsymbol{g}$. When T is 1-1 and onto we call $T\boldsymbol{g}$ a *rearrangement* of $\boldsymbol{g}$. By a *linear rearrangement* of $\boldsymbol{g}$ we mean a rearrangement $T\boldsymbol{g}$ where T is $\mathbf{Z}_2$-linear, i.e., satisfies

$$T(n \oplus m) = T(n) \oplus T(m)$$

for $n, m \in \mathbf{N}$.

It is important to notice that a linear rearrangement $T\boldsymbol{w}$ of the Walsh system is always a product system and is generated by the rearrangement $T\boldsymbol{r}$. Indeed, by definition

$$T\boldsymbol{r} = (w_{T(2^n)}, n \in \mathbf{N}).$$

Moreover, the sequence $(T(2^n), n \in \mathbf{N})$ is evidently $\mathbf{Z}_2$- linearly independent when T is both 1-1 and $\mathbf{Z}_2$-linear. Thus the system $T\boldsymbol{r}$ is multiplicative. But $\mathbf{Z}_2$-linearity of T implies

$$g_m = \prod_{n=0}^{\infty} w_{T(2^n)}^{m_n} = \prod_{n=0}^{\infty} w_{m_n T(2^n)} = w_{m_0 T(2^0) \oplus m_1 T(2^1) \oplus \ldots} = w_{T(m)}$$

for $m \in \mathbf{N}$. Therefore, $T\boldsymbol{w}$ is the product system generated by $T\boldsymbol{r}$.

The Rademacher system and measure preserving transformations generate all sign-like multiplicative systems:

THEOREM 6. *Suppose* $\boldsymbol{\gamma} = (\gamma_n, n \in \mathbf{N})$ *is a sign-like multiplicative system on some probability space* $(\Omega, \mathcal{G}, \nu)$. *Then there is a measure preserving transformation* $T' : \Omega \to \mathbf{G}$ *such that*

$$\gamma_n = \rho_n \circ T' \qquad (n \in \mathbf{N}).$$

Moreover, if $(g_n, n \in \mathbf{N})$ *is the product system generated by* $\boldsymbol{\gamma}$ *then*

$$g_n = \psi_n \circ T' \qquad (n \in \mathbf{N}).$$

PROOF. For every $x \in \mathbf{G}$ define a set $A_x \subseteq \Omega$ (which may be empty for some x's) by

$$A_x := \{t \in \Omega : \gamma_n(t) = \rho_n(x) \quad \text{for all } n \in \mathbf{N}\}.$$

Each A_x is measurable, $A_x \cap A_y = \emptyset$ for $x \neq y$, and

$$\Omega = \bigcup_{x \in \mathbf{G}} A_x.$$

In particular, given $t \in \Omega$ there is a unique $x \in \mathbf{G}$ such that $t \in A_x$. Denote this x by $T'(t)$.

We have defined a map T' from Ω into $\mathbf{G}$ such that $x = T'(t)$ if and only if $\rho_n(x) = \gamma_n(t)$ for all $n \in \mathbf{N}$. Thus

$$\gamma_n(t) = \rho_n \circ T'(t)$$

for all $n \in \mathbf{N}$ and $t \in \Omega$. It remains to see that T' is measure preserving, i.e., the set

$$\{t \in \Omega : T'(t) \in I\}$$

belongs to $\mathcal{G}$ for every interval $I = I_n(y)$ in $\mathbf{G}$ and

$$\nu(\{t \in \Omega : T'(t) \in I\}) = \mu(I) = 2^{-n}. \tag{31}$$

By definition

$$\begin{aligned} \{t \in \Omega : T'(t) \in I\} &= \bigcup_{x \in I} A_x \\ &= \{t \in \Omega : \gamma_j(t) = \rho_j(y) \quad \text{for } j = 0, 1, \ldots, n-1\} \\ &= \{t \in \Omega : 2^{-n} \prod_{j=0}^{n-1} (1 + \gamma_j(t)\rho_j(y)) = 1\}. \end{aligned} \tag{32}$$

Identity (31) follows easily from (32) and the fact that $\boldsymbol{\gamma}$ is a multiplicative system. Indeed, since the product in (32) takes only the values 0 and 1 we get

$$\begin{aligned} \nu(\{t \in \Omega : T'(t) \in I\}) &= 2^{-n} \int_\Omega \prod_{j=0}^{n-1} (1 + \rho_j(y)\gamma_j)\, d\nu \\ &= 2^{-n} \int_\Omega \left(\sum_{k=0}^{2^n - 1} \psi_k(y) g_k \right) d\nu = 2^{-n}. \end{aligned}$$

We conclude that T' is measurable and by (31) is measure preserving. ∎

In particular, given any $\mathbf{Z}_2$-linear map T of $\mathbf{N}$ into itself there is a measure preserving transformation $T' : \mathbf{G} \to \mathbf{G}$ such that $T\widehat{\mathbf{G}} = \widehat{\mathbf{G}} \circ T'$, i.e., $\psi_{T(m)}(x) = \psi_m(T'(x))$ for $m \in \mathbf{N}$ and $x \in \mathbf{G}$. Using the identification of $\mathbf{G}_0$ and $\mathbf{N}$ we have equivalently that

$$(33) \qquad \langle T(m), x\rangle = \langle m, T'(x)\rangle \qquad (m \in \mathbf{N}, x \in \mathbf{G})$$

(see (17) in 1.2). Since this identity uniquely determines T' we shall call T' the adjoint of T.

By using the identification of $\mathbf{G}_0$ with $\mathbf{N}$ and the usual closed system $\boldsymbol{e}$ we can represent each $\mathbf{Z}_2$-linear map T by a matrix $(t_{ij})_{i,j=0}^{\infty}$ where $t_{ij} = 0$ or 1 and

$$T(e_i) := (t_{ki}, k \in \mathbf{N}) \qquad (i \in \mathbf{N}).$$

Indeed, the binary coefficients of $T(m)$ are given by

$$T(m)_i = \sum_{j=0}^{\infty} t_{ij} m_j \pmod 2$$

for $i, m \in \mathbf{N}$. Notice, therefore, that the representing matrix of the adjoint T' of T is the one whose entries satisfy

$$t'_{ij} = t_{ji} \qquad (i, j \in \mathbf{N}).$$

In general, the adjoint of T may not be 1-1. However, if T is 1-1, onto then so is T' and its inverse is identical to the adjoint of T^{-1}. Indeed, T' is onto since T is 1-1 and onto. T' is 1-1 since T is onto. And by (33) we have

$$\begin{aligned}\left\langle T^{-1}(n), x\right\rangle &= \left\langle T^{-1}(n), T'\left((T')^{-1}(x)\right)\right\rangle \\ &= \left\langle n, (T')^{-1}(x)\right\rangle\end{aligned}$$

for $n \in \mathbf{N}$ and $x \in \mathbf{G}$. In particular,

$$(T^{-1})' = (T')^{-1}.$$

We summarize these observations as follows.

THEOREM 7. *If T is a $\mathbf{Z}_2$-linear map from $\mathbf{N}$ into $\mathbf{N}$ then there exists a unique measure preserving map T' from $\mathbf{G}$ into $\mathbf{G}$ such that*

$$\psi_{T(m)} = \psi_m \circ T' \qquad (m \in \mathbf{N})$$

and (33) holds. Moreover, if T is 1-1, onto then so is T', in which case

$$(T^{-1})' = (T')^{-1}.$$

Hence every linear rearrangement $T\widehat{\mathbf{G}}$ of the characters of $\mathbf{G}$ can be written in the form

$$T\widehat{\mathbf{G}} = \widehat{\mathbf{G}} \circ T'.$$

In the next two paragraphs we shall examine the original Walsh system $\boldsymbol{\phi}$. First we show that $\boldsymbol{\phi}$ is a linear rearrangement of the Walsh system $\boldsymbol{w}$. Consider the map $T_0 : \mathbf{G}_0 \to \mathbf{G}_0$ determined by the matrix $(t_{ij})_{i,j=0}^{\infty}$ where

$$t_{ij} := \begin{cases} 1 & j = i, i+1 \\ 0 & \text{otherwise.} \end{cases}$$

The map T_0 is obviously linear, 1-1, and onto, so we need only show

$$\phi_n = w_{T_0(n)} \qquad (n \in \mathbf{N}). \tag{34}$$

Toward this recall that $\phi_0 = 1, \phi_1 = r_0, \phi_2 = r_0 r_1$, and $\phi_3 = r_1$. Fix $n \in \mathbf{N}$ and write $n = 2^{m-1} + k - 1$ for $1 \le k \le 2^{m-1}, m \in \mathbf{P}$. By definition we have

$$\phi_{2n}(x) = \phi_{m+1}^{(2k-1)}(x) = r_0^{k+1}(x)\phi_n(2x)$$

and

$$\phi_{2n+1}(x) = \phi_{m+1}^{(2k)}(x) = r_0^{k}(x)\phi_n(2x)$$

for $x \in [0,1)$. Consequently, if $[n/2]$ represents the greatest integer in $n/2$ and $n \ge 4$ has binary expansion $\sum_{i=0}^{\infty} n_i 2^i$ then

$$\phi_n(x) = r_0^{n_0+n_1}(x)\phi_{[n/2]}(2x) \qquad (x \in \mathbf{R}).$$

Applying this identity repeatedly we arrive at

$$\phi_n(x) = \prod_{j=0}^{\infty} r_j^{n_j+n_{j+1}}(x) = (-1)^{\langle T_0(n), x\rangle}$$

for $x \in [0,1)$. We conclude by (17) that

$$\phi_n = w_{T_0(n)}$$

as required.

Secondly, we prove that each ϕ_n changes signs on $[0,1)$ exactly n times. Fix $n \in \mathbf{N}$ and write $n = 2^m + k$ for $0 \le k < 2^m$. Observe that ϕ_n changes signs at a point

$$x = \frac{x_0}{2} + \frac{x_1}{2^2} + \cdots + \frac{x_{j-1}}{2^j} + \frac{1}{2^{j+1}}$$

if and only if $j \le m$ and

$$\phi_n(x - \frac{1}{2^{m+1}}) = -\phi_n(x).$$

This condition is equivalent to

$$n_{j+1} \oplus n_{j+2} \oplus \cdots \oplus n_m = n_j \oplus n_{j+1} \oplus 1$$

since

$$x - \frac{1}{2^{m+1}} = \sum_{i=0}^{j-1} \frac{x_i}{2^{i+1}} + \sum_{i=j+1}^{m} \frac{1}{2^{i+1}}.$$

Hence ϕ_n changes signs at a point x if and only if $n_j = 1$. But for each $j \in \mathbf{N}$ the number of such points in $[0, 1)$ is $n_j 2^j$. Therefore, the number of sign changes of ϕ_n on $[0, 1)$ is precisely

$$\sum_{j=0}^{m} n_j 2^j = n.$$

Among W-continuous sign-like systems, this property characterizes the original Walsh system. Indeed, if $(g_n, n \in \mathbf{N})$ is an orthonormal system on $[0, 1)$ where each g_n continuous is from the right, takes on only the values $+1$ or -1, and changes signs exactly n times on $[0, 1)$ then $g_n = \phi_n$ for $n \in \mathbf{N}$ (see Byrnes and Swick [1] and Levizov [1]).

We have seen that the original Walsh system is a linear rearrangement of the system $\boldsymbol{w}$. The Walsh-Kaczmarz system is not. Instead, it is a *piecewise linear rearrangement.* To describe the situation we need additional notation.

For each $n \in \mathbf{N}$ set

$$\mathbf{G}_0^{(n)} := \{x \in \mathbf{G} : x_k = 0 \quad \text{for} \quad k \geq n\}.$$

Under the identification of $\mathbf{G}_0$ with $\mathbf{N}$ each $e_n + \mathbf{G}_0^{(n)}$ corresponds to the dyadic block

$$\{k \in \mathbf{N} : 2^n \leq k < 2^{k+1}\}.$$

Thus $\mathbf{G}_0$ can be written as a disjoint union in the form

$$\mathbf{G}_0 = \mathbf{G}_0^{(0)} \cup \left(\bigcup_{n=0}^{\infty} (e_n + \mathbf{G}_0^{(n)}) \right).$$

In particular, any map R defined on $\mathbf{G}_0$ is determined by its values on the sets $\mathbf{G}_0$ and $e_n + \mathbf{G}_0^{(n)} \quad (n \in \mathbf{N})$.

A map $R : \mathbf{G}_0 \to \mathbf{G}_0$ is called *piecewise linear* if there exist 1-1, $\mathbf{Z}_2$-linear maps R_n of $\mathbf{G}_0^{(n)}$ onto $\mathbf{G}_0^{(n)}$ such that

$$(35) \qquad R(m) := \begin{cases} m & m = 0 \quad \text{or} \quad e_0 \\ e_n + R_n(m + e_n) & m \in e_n + \mathbf{G}_0^{(n)} \quad (n \in \mathbf{P}). \end{cases}$$

By a piecewise linear rearrangement of a system $\boldsymbol{g}$ we shall mean a rearrangement of the form $R\boldsymbol{g}$ for some piecewise linear map R. Notice, then, that piecewise linear rearrangements are rearrangements within dyadic blocks. Indeed, if $(f_k, k \in \mathbf{N})$ is a piecewise linear rearrangement of $(g_k, k \in \mathbf{N})$ then

$$\{f_k : 2^n \leq k < 2^{n+1}\} = \{g_k : 2^n \leq k < 2^{n+1}\}$$

for $n \in \mathbf{N}$.

To show that the Walsh-Kaczmarz system is a piecewise linear rearrangement of the Walsh system, fix $n \in \mathbf{N}$ and let $R_n : \mathbf{G}_0^{(n)} \to \mathbf{G}_0^{(n)}$ be the $\mathbf{Z}_2$-linear map induced by the matrix $(t_{ij}^{(n)})_{i,j=0}^{\infty}$ where

$$t_{ij}^{(n)} := \begin{cases} 1 & i+j=n-1 \\ 0 & \text{otherwise.} \end{cases}$$

Let R be defined by (35). To show $\boldsymbol{\kappa} = R\boldsymbol{w}$ write $n = 2^m + \ell$ where $m, \ell \in \mathbf{N}$ and $0 \le \ell < 2^m$. Let $(\ell_k, k \in \mathbf{N})$ represent the binary coefficients of ℓ and observe by identity (5) in 1.1 that

$$\kappa_n(x) = w_{2^m}(x)(-1)^{\sum_{k=0}^{m-1} x_k \ell_{m-k-1}}$$

for $x \in [0,1)$ with dyadic expansion given by (18). It follows from the definition of R_n and (17) in 1.2 that

$$\kappa_n(x) = w_{2^m}(x)(-1)^{\langle R_m(\ell), x\rangle} = w_{2^m}(x) w_{R_m(\ell)}(x) = w_{2^m + R_m(\ell)}(x) = w_{R(n)}(x),$$

for $x \in [0,1)$. Thus $\boldsymbol{\kappa}$ is a piecewise linear rearrangement of $\boldsymbol{w}$.

For each $n \in \mathbf{N}$ set

$$a_{kj}^{(n)} := 2^{-n/2} w_k\left(\frac{j}{2^n}\right),$$

for $0 \le j, k < 2^n$, $j, k \in \mathbf{N}$. By a *Hadamard- Paley matrix* we shall mean a matrix of the form

$$A^{(n)} := (a_{kj}^{(n)})_{j,k=0}^{2^n-1} \qquad (n \in \mathbf{N}).$$

Notice by (4) in 1.1 that

$$w_j(k/2^n) = w_k(j/2^n) \tag{36}$$

for $0 \le j, k < 2^n$, $\quad j, k \in \mathbf{N}$. Thus each Hadamard-Paley matrix $A^{(n)}$ is a real, symmetric $2^n \times 2^n$ matrix. Moreover, since

$$\sum_{i=0}^{2^n-1} a_{ki}^{(n)} a_{ji}^{(n)} = \sum_{i=0}^{2^n-1} 2^{-n} w_k\left(\frac{i}{2^n}\right) w_j\left(\frac{i}{2^n}\right) = \sum_{i=0}^{2^n-1} \int_{I(i,n)} w_k w_j = \int_0^1 w_k w_j,$$

it is clear that each $A^{(n)}$ is orthogonal.

The *Hadamard transform* $\perp$ is defined on ℓ^0 as follows. Given $\boldsymbol{b} = (b_k, k \in \mathbf{N}) \in \ell^0$ define a sequence $\boldsymbol{b}^\perp := (b_k^\perp, k \in \mathbf{N}) \in \ell^0$ by the following process. Set $b_0^\perp := b_0$. For $0 \le k < 2^n, n \in \mathbf{N}$ set

$$b_{2^n+k}^\perp := \sum_{j=0}^{2^n-1} a_{kj}^{(n)} b_{2^n+j}.$$

Since the Hadamard-Paley matrices are symmetric and orthogonal, it is clear that $\perp$ takes ℓ^0 onto ℓ^0 and satisfies $(\boldsymbol{b}^\perp)^\perp = \boldsymbol{b}$. In particular, if $\boldsymbol{a}$ is the Hadamard transform of $\boldsymbol{b}$ then $\boldsymbol{b}$ is the Hadamard transform of $\boldsymbol{a}$.

Extend the Hadamard transform to function sequences in the obvious way. Notice by definition that

$$\begin{aligned}\sum_{k=0}^{2^n-1} b_{2^n+k} f^{\perp}_{2^n+k} &:= \sum_{k=0}^{2^n-1} b_{2^n+k} \Big(\sum_{j=0}^{2^n-1} a^{(n)}_{kj} f_{2^n+j} \Big) \\ &= \sum_{j=0}^{2^n-1} \Big(\sum_{k=0}^{2^n-1} a^{(n)}_{jk} b_{2^n+k} \Big) f_{2^n+j} \\ &= \sum_{j=0}^{2^n-1} b^{\perp}_{2^n+j} f_{2^n+j}\end{aligned}$$

for any function sequence $(f_k, k \in \mathbf{N})$ and $(b_k, k \in \mathbf{N}) \in \ell^0$. In particular,

$$\sum_{k=2^{n-1}}^{2^n-1} b_k f^{\perp}_k = \sum_{k=2^{n-1}}^{2^n-1} b^{\perp}_k f_k \tag{37}$$

for $n \in \mathbf{P}$. We shall refer to this property by calling the Hadamard transform self adjoint.

The *Haar system* $\boldsymbol{h} = (h_n, n \in \mathbf{N})$ is defined as follows. Set $h_0 := 1$. For $n, k \in \mathbf{N}$ with $0 \le k < 2^n$ define h_n on $[0,1)$ by

$$h_{2^n+k}(x) := \begin{cases} 2^{n/2} & x \in I(2k, n+1) \\ -2^{n/2} & x \in I(2k+1, n+1) \\ 0 & \text{otherwise.} \end{cases}$$

Extend each Haar function to $\mathbf{R}$ by periodicity of period 1. Thus each Haar function is continuous from the right, and the Haar system $\boldsymbol{h}$ is orthonormal on $[0,1)$. (Most authors define Haar functions at the points of discontinuity by averaging. This small change has little bearing on the results we prove here.)

The Walsh and Haar systems are Hadamard transforms of each other. Indeed, fix $0 \le k < 2^n, n, k \in \mathbf{N}$ and observe by definition that

$$h_{2^n+k} = 2^{n/2} r_n \chi(I(k,n)). \tag{38}$$

Also observe by the Paley lemma that

$$\sum_{j=0}^{2^n-1} w_j(x \dot{+} \frac{k}{2^n}) = 2^n \chi(I(k,n))(x) \tag{39}$$

for $x \in [0,1)$. Since

$$r_n(x) w_j(x \dot{+} \frac{k}{2^n}) = w_j(\frac{k}{2^n}) w_{2^n+j}(x),$$

it follows from (36), (38), and (39) that

$$h_{2^n+k} = \sum_{j=0}^{2^n-1} a_{kj}^{(n)} w_{2^n+j}. \tag{40}$$

Therefore, $\boldsymbol{h} = \boldsymbol{w}^{\perp}$ and by the self adjoint property we have

$$w_{2^n+k} = \sum_{j=0}^{2^n-1} a_{kj}^{(n)} h_{2^n+j} \quad (0 \le k < 2^n, n, k \in \mathbf{N}) \tag{41}$$

and

$$\sum_{k=2^{n-1}}^{2^n-1} b_k w_k = \sum_{k=2^{n-1}}^{2^n-1} b_k^{\perp} h_k \quad (n \in \mathbf{P}). \tag{42}$$

The Hadamard matrices $H^{(n)}$ are generated by the Kaczmarz ordering in the following way. Set

$$H^{(0)} := (1).$$

For $n \in \mathbf{P}$ set

$$H^{(n)} := \left(r_n \kappa_{2^n+j}\left(\frac{k}{2^n}\right) \right)_{j,k=0}^{2^n-1}.$$

For example,

$$H^{(1)} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

and

$$H^{(2)} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

An easy inductive argument establishes that the Hadamard matrices can be recursively generated by the Kronecker product, namely

$$H^{(n+1)} = H^{(1)} \otimes H^{(n)} := \begin{pmatrix} H^{(n)} & H^{(n)} \\ H^{(n)} & -H^{(n)} \end{pmatrix}$$

for $n \in \mathbf{N}$. We also notice by induction that the k-th row of $H^{(n)}$ is given by

$$(w_{\mathbf{n}(k)}(\frac{j}{2^n}), j = 0, 1, \dots 2^n - 1)$$

where for each $n \in \mathbf{N}$ and $0 \le k < 2^n$ with binary coefficients $(k_\ell, \ell \in \mathbf{N})$ the index $\mathbf{n}(k)$ is defined by

$$\mathbf{n}(k) := \sum_{\ell=0}^{n-1} k_{n-\ell-1} 2^\ell.$$

This remark has significance for certain applications of Walsh functions to spectroscopy. The map $k \to \mathbf{n}(k)$ is called the bit-reversal map and facilitates computation of the Fast Walsh Transform. These topics will be discussed in 9.7.

1.5 Walsh-Fourier Coefficients and Walsh-Fourier Series. The *Walsh-Fourier coefficients* of an $f \in L^1(G)$ (respectively, L^1) are defined by

$$\widehat{f}(n) := \int_G f\psi_n \, d\mu \qquad \text{(respectively, } := \int_0^1 fw_n)$$

for $n \in \mathbf{N}$. Thus the map $f \to \widehat{f}$ takes integrable functions to ℓ^∞ sequences and

$$\|\widehat{f}\|_\infty \le \|f\|_1. \tag{43}$$

A fundamental question of dyadic analysis (indeed of harmonic analysis in general) is what affect do conditions on f (e.g., continuity, differentiability, integrability) have on the sequence $\widehat{f}$. The Riesz-Fischer theorem (see Appendix 0.1) is one response to this question: an integrable function f belongs to L^2 if and only if $\widehat{f}$ belongs to ℓ^2.

The situation is more complicated for almost any other class of functions besides L^2. For example, the following result shows that the Walsh-Fourier coefficients of L^1 functions vanish at infinity. But we shall see (Theorem 3 in 8.1 and remarks following) that the rate of decay can be as slow as one wishes. Thus one cannot characterize L^1 functions by the growth of their Walsh-Fourier coefficients.

THE RIEMANN-LEBESGUE LEMMA. *If* $f \in L^1$ *then* $\widehat{f}(n) \to 0$ *as* $n \to \infty$.

PROOF. Since Walsh polynomials are dense in L^1, given $\varepsilon > 0$ there is a $g \in \mathcal{P}$ such that $\|f-g\|_1 < \varepsilon$. Since $\widehat{g}(n) = 0$ for large n we have

$$\int_0^1 fw_n = \int_0^1 (f-g)w_n$$

for large n. Therefore,

$$\limsup_{n\to\infty} |\widehat{f}(n)| \le \|f-g\|_1 < \varepsilon$$

as required. ∎

Let $*$ denote convolution on the group or dyadic convolution on the unit interval. Thus

$$(f*g)(x) := \int_G f(t)g(x+t)\, d\mu(t)$$

for $f, g \in L^1(G)$, $x \in G$, and

$$(f*g)(x) := \int_0^1 f(t)g(x \dot{+} t)\, dt$$

for $f, g \in L^1$, $x \in [0,1)$. Recall by the Fubini theorem and Hölder's inequality that

$$\|f*g\|_p \le \|f\|_1 \|g\|_p \tag{44}$$

for $1 \le p \le \infty$. Consequently, $L^1(G)$ and L^1 are Banach algebras under convolution (see Exercise 1.15).

Under the map $\widehat{\ }$, convolution of functions corresponds to pointwise multiplication of sequences, i.e.,

$$\widehat{f * g}(k) = \widehat{f}(k)\widehat{g}(k) \quad (k \in \mathbf{N}). \tag{45}$$

This is easy to verify on the group or on the unit interval. For example, on the unit interval we have by Fubini's theorem, (23) in 1.3, and translation invariance of Lebesgue measure that

$$\widehat{f \dot{*} g}(k) = \int_0^1 \int_0^1 f(t)g(x \dot{+} t)w_k(x \dot{+} t)w_k(t)\,dx\,dt = \widehat{f}(k)\widehat{g}(k).$$

Under the map $\widehat{\ }$ the shift operators $f \to \psi_n f$ on $L^1(G)$ and $f \to w_n f$ on L^1 correspond to dyadic translation in ℓ^0, e.g.,

$$\widehat{\psi_n f}(k) = \widehat{f}(n \oplus k) \tag{46}$$

for $n, k \in \mathbf{N}$. These identities follow immediately from the definitions and (14) in 1.2.

Conversely, translation in $L^1(G)$ or L^1 corresponds to a shift in ℓ^0. For example,

$$\widehat{\tau_y f}(k) = \psi_k(y)\widehat{f}(k) \quad (k \in \mathbf{N}) \tag{47}$$

holds for all $f \in L^1(G)$ and $y \in G$. This is evident by the translation invariance of μ and the fact that $\psi_n(x + y) = \psi_n(x)\psi_n(y)$ $(x, y \in G)$.

By a *Walsh series* we shall mean a series of the form $S := \sum_{k=0}^{\infty} a_k f_k$ where the coefficients a_k are real numbers and $f_k = \psi_k$ or w_k. The *n-th partial sums* of the Walsh series S will be denoted by $S_0 := 0$ and

$$S_n := \sum_{k=0}^{n-1} a_k f_k \qquad (n \in \mathbf{P}).$$

The *Walsh-Fourier series* of an $f \in L^1(G)$ is the Walsh series

$$Sf := \sum_{k=0}^{\infty} \widehat{f}(k)\psi_k.$$

Similarly, the Walsh-Fourier series of an $f \in L^1$ is the Walsh series

$$Sf := \sum_{k=0}^{\infty} \widehat{f}(k)w_k.$$

These are formal definitions; we make no prior assumptions about convergence. The n-th partial sums of these series are defined in the obvious manner. For example,

$$S_n f := \sum_{k=0}^{n-1} \widehat{f}(k)\psi_k \qquad (n \in \mathbf{P}).$$

For convenience we set $S_0 f := 0$.

Another fundamental question of dyadic analysis is the following one. Can a given function f be approximated by some Walsh series? Since the system $\widehat{\mathbf{G}}$ is orthonormal, it is easy to see that if a Walsh series $\sum_{k=0}^{\infty} a_k \psi_k$ converges uniformly on $\mathbf{G}$ to an integrable function f then $a_k = \widehat{f}(k)$ for $k = 0, 1, \ldots$, i.e., in this case the approximating Walsh series must be a Walsh-Fourier series. Thus this general question is often replaced by a more specific one: does the Walsh-Fourier series of a given integrable f converge to f in some sense?

We illustrate these ideas by passing to the unit interval and analyzing the indefinite integrals $\boldsymbol{J}_k$ of the Walsh functions w_k for $k \in \mathbf{N}$ extended to $[0, \infty)$ by periodicity of period 1. We will obtain their Walsh-Fourier series by representing them as uniformly convergent Walsh series.

First let $[x]$ represent the greatest integer in x and notice that

$$\boldsymbol{J}_0(x) = x - [x] \qquad (x \in [0, \infty)).$$

Since the binary coefficients of an $x \in [0, 1)$ satisfy

$$x_k = \frac{1 - r_k(x)}{2} \qquad (k \in \mathbf{N})$$

it is clear that

$$x - [x] = \frac{1}{2} - \frac{1}{4} \sum_{k=0}^{\infty} 2^{-k} r_k(x).$$

Consequently,

$$\boldsymbol{J}_0 = \frac{1}{2} - \frac{1}{4} \sum_{k=0}^{\infty} 2^{-k} w_{2^k}$$

uniformly on $[0, \infty)$.

Next observe since

$$\boldsymbol{J}_k(x) := \int_0^x w_k \qquad (k \in \mathbf{P}, x \in [0, \infty))$$

that

$$\boldsymbol{J}_1(x) = \frac{1}{2} + w_1(x) \left(x - [x] - \frac{1}{2} \right).$$

Using the expression for $\boldsymbol{J}_0$ derived above we see that

$$\boldsymbol{J}_1 = \frac{1}{4} - \frac{1}{4} \sum_{k=1}^{\infty} 2^{-k} w_{2^k + 1}$$

uniformly on $[0, \infty)$.

Finally, let $k = 2^n + \ell$ for $0 \le \ell < 2^n$ and recall that w_ℓ is constant on the intervals $I(p, n)$ for all $0 \le p < 2^n$. Since $w_k = r_n w_\ell$ and r_n changes signs from $I(2p, n+1)$ to

$I(2p+1, n+1)$ it follows that the integral of w_k over any dyadic interval $I(p,n)$ is zero. Hence $\boldsymbol{J}_k(p/2^n) = 0$ for $0 \le p < 2^n$ and

$$\boldsymbol{J}_k(x) = 2^{-n} w_\ell(x)\boldsymbol{J}_1(2^n x) = w_\ell(x)\boldsymbol{J}_{2^n}(x) \qquad (x \in [0,1)).$$

Using the representation for $\boldsymbol{J}_1$ derived above we see that

$$\begin{aligned}\boldsymbol{J}_{2^n}(x) &= 2^{-n}\boldsymbol{J}_1(2^n x)\\ &= 2^{-n-2}\left(1 - \sum_{j=1}^{\infty} 2^{-j} w_{2^j+1}(2^n x)\right)\\ &= 2^{-n-2}\left(1 - \sum_{j=1}^{\infty} 2^{-j} w_{2^{j+n}+2^n}(x)\right).\end{aligned}$$

We conclude that

$$\boldsymbol{J}_k = 2^{-n-2}\left(w_\ell - \sum_{j=1}^{\infty} 2^{-j} w_{2^{n+j}+k}\right) \tag{48}$$

uniformly on $[0,\infty)$ for $k = 2^n + \ell$, $0 \le \ell < 2^n$, $n \in \mathbf{N}$.

A useful corollary of these calculations is the following. Let $\operatorname{dist}(x,n)$ represent the distance from a point $x \in [0,1)$ to the nearest dyadic rational of the form $p/2^n$ for some $0 \le p < 2^n$. Then

$$\boldsymbol{J}_k(x) = w_\ell(x)\operatorname{dist}(x,n)$$

for $k = 2^n + \ell$, $0 \le \ell < 2^n$, $n \in \mathbf{N}$ and $x \in [0,1)$.

Let D_n denote the *Walsh-Dirichlet kernel* of order n, i.e., $D_0 := 0$ and

$$D_n := \sum_{k=0}^{n-1} \psi_k \qquad (n \in \mathbf{P}).$$

By interchanging the order of summation and integration, it is clear that

$$(S_n f)(x) = \int_{\mathbf{G}} f(t) \sum_{k=0}^{n-1} \psi_k(t)\psi_k(x)\, d\mu(t) = \int_{\mathbf{G}} f(t) D_n(x+t)\, d\mu(t),$$

for $n \in \mathbf{P}$ and $x \in \mathbf{G}$. In particular, we have the fundamental identity

$$S_n f = f * D_n \tag{49}$$

for $f \in \mathrm{L}^1(\mathbf{G})$ and $n \in \mathbf{N}$.

The general problem of determining when a Fourier series converges and in what sense (e.g., uniformly, a.e., in norm) is a delicate one. In view of (49) it comes as no surprise that the Walsh- Dirichlet kernels play a prominent role here. The following formulae will be useful in this regard.

THEOREM 8. *Let $n \in \mathbf{N}$ have binary coefficients $(n_k, k \in \mathbf{N})$. Then*

(i) $$D_{2^k} = \prod_{j=0}^{k-1}(1+\rho_j) \qquad (k \in \mathbf{N})$$

(ii) $$D_n = \psi_n \sum_{k=0}^{\infty} n_k (D_{2^{k+1}} - D_{2^k})$$

and

(iii) $$D_n = \psi_n \sum_{k=0}^{\infty} n_k \rho_k D_{2^k}.$$

Moreover, if $0 < n < 2^m$ then

(iv) $$D_n = \frac{1}{2}(D_{2^m} - \psi_n) - \frac{1}{2}\sum_{k=0}^{m-1}(\tau_{e_k}\psi_n)\rho_k D_{2^k}.$$

PROOF. For $0 \le k < 2^n$, $n \in \mathbf{N}$ we have by definition that

$$\psi_{2^n+k} = \rho_n \psi_k.$$

Consequently

$$D_{2^{n+1}} = \sum_{k=0}^{2^n-1}\psi_k + \sum_{k=0}^{2^n-1}\psi_{2^n+k} = D_{2^n} + \rho_n D_{2^n} = (1+\rho_n)D_{2^n}$$

for $n \in \mathbf{N}$ and (i) follows at once. We have also proved

(50) $$D_{2^{n+1}} - D_{2^n} = \rho_n D_{2^n} \qquad (n \in \mathbf{N}).$$

By Theorem 3,

$$D_n = \sum_{k=0}^{\infty} n_k \sum_{j \in n \oplus A_k} \psi_j$$

and by (14) in 1.2 we have

$$\sum_{j \in n \oplus A_k} \psi_j = \sum_{j=2^k}^{2^{k+1}-1} \psi_{n \oplus j} = \psi_n (D_{2^{k+1}} - D_{2^k}).$$

Thus (ii) holds.

Since (50) and (ii) imply (iii) it remains to verify (iv). Fix $0 < n < 2^m$. Recall that $\psi_n = 1$ on $I_m(0)$. Thus by the Paley lemma,

$$\psi_n D_{2^m} = D_{2^m}.$$

On the other hand, the obvious identity

$$n_k \psi_n = \frac{1}{2}(1 - (-1)^{n_k})\psi_n = \frac{1}{2}(\psi_n - \tau_{e_k}\psi_n)$$

and (iii) imply that

$$D_n = \frac{1}{2}\psi_n \sum_{k=0}^{m-1} \rho_k D_{2^k} - \frac{1}{2}\sum_{k=0}^{m-1}(\tau_{e_k}\psi_n)\rho_k D_{2^k}.$$

It follows from (50) that

$$D_n = \frac{1}{2}(D_{2^m} - \psi_n) - \frac{1}{2}\sum_{k=0}^{m-1}(\tau_{e_k}\psi_n)\rho_k D_{2^k}$$

for each $n \in \mathbf{N}$. ∎

All of these definitions and formulae have analogues on the unit interval for the system $\boldsymbol{w}$ in place of $\widehat{\mathbf{G}}$ and $\dot{+}$ in place of the group operation $+$. As the need arises we shall use these below.

Identity (49) will be used often. In conjunction with the Paley lemma it implies that 2^n-th partial sums of Walsh-Fourier series are averages, namely,

$$(51) \qquad (S_{2^n})f(x) = \frac{1}{|I_n(x)|}\int_{I_n(x)} f(t)\,dt \qquad (x \in [0,1),\ n \in \mathbf{N}).$$

This simplifies the study of convergence of 2^n-th partial sums of Walsh-Fourier series and shows it to be closely allied with classical differentiation of indefinite integrals.

For example, since the derivative of $\int_0^x f$ equals $f(x)$ a.e. when $f \in \mathrm{L}^1$ it follows from (51) that

$$\lim_{n\to\infty} S_{2^n} f = f \quad \text{a.e.} \quad \text{for} \quad f \in \mathrm{L}^1.$$

By Theorem 5 in 1.3 we also see that $\lim_{n\to\infty} S_{2^n} f = f$ a.e. $[\mu]$ for $f \in \mathrm{L}^1(\mathbf{G})$. Therefore, if $\widehat{f}(k) = 0$ for all $k \in \mathbf{N}$ then $f = 0$ a.e. on $[0,1)$ or $\mathbf{G}$. In particular, the systems $\boldsymbol{w}$ and $\widehat{\mathbf{G}}$ are both complete.

A similar argument shows that $S_{2^n} f \to f$ uniformly on $\mathbf{G}$ as $n \to \infty$ for all $f \in \mathrm{C}(\mathbf{G})$. By using the canonical isomorphism we also see that $S_{2^n} f \to f$ uniformly on $[0,1)$ as $n \to \infty$ for all $f \in \mathrm{C}_W$, in particular for each function classically continuous on $[0,1]$.

The *spectrum* of an integrable f is defined by

$$sp(f) := \{n \in \mathbf{N} : \widehat{f}(n) \neq 0\}.$$

Thus f belongs to $\mathcal{P}$ if and only if $sp(f)$ is a finite set. Given $P \in \mathcal{P}$ the number $\sup(sp(P))$ will be called the order of P. Clearly, if n is greater than the order of P then $S_n P = P$.

Let $\mathcal{A}$ denote the algebra of sets generated by the dyadic intervals in $[0,1)$. By a *quasi-measure* we shall mean a real-valued set function which is finitely additive on $\mathcal{A}$. Clearly, every finite Borel measure on $[0,1)$ is a quasi-measure but not conversely.

We shall denote the collection of quasi-measures by QM. For each $\nu \in$ QM define the ***Walsh-Fourier-Stieltjes coefficients*** of ν by

$$\widehat{\nu}(n) := \int_0^1 w_n \, d\nu \qquad (n \in \mathbf{N}).$$

Since each Walsh function is constant on sufficiently small dyadic intervals this definition makes sense. Also, if ν is a finite Borel measure and $\|\nu\|$ represents its total variation on $[0,1)$ then $\widehat{\nu}$ belongs to ℓ^∞ with

$$\|\widehat{\nu}\|_\infty \leq \|\nu\|.$$

We shall prove that the map $\nu \to \widehat{\nu}$ is a 1-1 function from QM onto ℓ^0. First, define the Walsh- Fourier-Stieltjes series $S\nu$ of an $\nu \in$ QM by

$$S\nu := \sum_{k=0}^{\infty} \widehat{\nu}(k) w_k.$$

For each $n \in \mathbf{N}$ set

$$(S_n \nu)(x) := \sum_{k=0}^{n-1} \widehat{\nu}(k) w_k$$

and observe for each $x \in [0,1)$ that

$$(S_n \nu)(x) = \int_0^1 D_n(x \dotplus t) \, d\nu(t).$$

Thus by the Paley lemma we have

$$(S_{2^n} \nu)(x) = 2^n \nu(I_n(x))$$

for $x \in [0,1)$ and $n \in \mathbf{N}$. In particular, the map $\widehat{\ }$ is 1-1 and linear on QM.

To show this map is onto, let $(a_n, n \in \mathbf{N}) \in \ell^0$ and consider the Walsh polynomials

$$S_{2^n} := \sum_{k=0}^{2^n-1} a_k w_k \qquad (n \in \mathbf{N}).$$

Define ν on $\mathcal{A}$ by

$$\nu(I) := \lim_{n \to \infty} \int_I S_{2^n}.$$

This limit exists because $\int_I w_j = 0$ for all $j \geq 2^{n_0}$ when $|I| = 2^{-n_0}$. Hence

$$\nu(I) = \int_I S_{2^{n_0}}$$

for $|I| = 2^{-n_0}$ and it follows that $\nu \in \mathrm{QM}$ and $\widehat{\nu}(n) = a_n$ for $n \in \mathbf{N}$.

Summarizing these observations, given any Walsh series

$$S := \sum_{k=0}^{\infty} a_k w_k$$

there is a unique quasi-measure ν on $\mathcal{A}$ such that

$$\widehat{\nu}(n) = a_n \quad \text{and} \quad S_{2^n}(x) = 2^n \nu(I_n(x)) \tag{52}$$

for all $n \in \mathbf{N}$ and $x \in [0,1)$. We shall call ν the quasi-measure associated with the series S.

The *upper binary derivate* of a quasi-measure ν at a point $x \in [0,1)$ is defined by

$$(\overline{\mathrm{D}}\nu)(x) := \limsup_{n\to\infty} \frac{\nu(I_n(x))}{|I_n(x)|},$$

the *lower binary derivate* by

$$(\underline{\mathrm{D}}\nu)(x) := \liminf_{n\to\infty} \frac{\nu(I_n(x))}{|I_n(x)|}.$$

A quasi-measure ν is said to have a *binary derivate* at x if $(\overline{\mathrm{D}}\nu)(x) = (\underline{\mathrm{D}}\nu)(x)$ in which case the common value will be denoted by $(\mathrm{D}\nu)(x)$. By (52), the 2^n-th partial sums of a Walsh series converge at a point x if and only if the binary derivate of its associated quasi-measure exists at x. This connects convergence properties of 2^n-th partial sums of Walsh series with measure theory.

For example, by the Radon-Nikodym theorem if ν is an absolutely continuous measure then its binary derivate exists a.e. and equals the Radon-Nikodym derivative of ν with respect to Lebesgue measure m. Thus

$$\lim_{n\to\infty} S_{2^n}\nu = \left[\frac{d\nu}{dm}\right]$$

a.e. $[m]$ on $[0,1)$ for all absolutely continuous ν. Again, if ν is a non-negative, singular measure then its binary derivate is 0 a.e. $[m]$ (see Theorem 6 in 6.2) and ∞ a.e. $[\nu]$ (see Lemma 3 in 7.1). Thus for such measures

$$\lim_{n\to\infty} S_{2^n}\nu = \begin{cases} 0 & \text{a.e.}\,[m] \\ +\infty & \text{a.e.}\,[\nu]. \end{cases}$$

And finally, for general quasi-measures (see Theorem 14 in Appendix 0.6), either the binary derivate exists a.e. or $(\overline{\mathrm{D}}\nu)(x) = +\infty$ and $(\underline{\mathrm{D}}\nu)(x) = -\infty$ for a.e. $x \in [0,1)$. Thus, the 2^n-th partial sums of a Walsh series either converges a.e. or satisfies

$$\limsup_{n\to\infty} S_{2^n} = +\infty, \quad \text{and} \quad \liminf_{n\to\infty} S_{2^n} = -\infty$$

a.e. on $[0,1)$.

It is well known that a function g on $[0,1]$ generates a quasi- measure by means of

$$\nu(\alpha,\beta] := g(\beta) - g(\alpha) \qquad (\alpha,\beta \in \mathbf{Q}).$$

It is natural to enquire about the relation between the Walsh- Fourier-Stieltjes coefficients of ν and the Walsh-Fourier coefficients of g. We shall close this section by answering this question for classically continuous g.

We begin by claiming that

$$g(\frac{k+1}{2^n}) - g(\frac{k}{2^n}) = \sum_{\ell=n}^{\infty} \sum_{s=2^\ell}^{2^{\ell+1}-1} \widehat{g}(s) \left(w_s(\frac{k+1}{2^n} - \frac{1}{2^{\ell+1}}) - w_s(\frac{k}{2^n}) \right)$$

holds for all integers $0 \le k < 2^n$.

To prove this claim fix such n, k and recall that $S_{2^\ell}\, g \to g$ uniformly on $[0,1)$ as $\ell \to \infty$. Hence by telescoping we have

$$g - S_{2^n} g = \sum_{\ell=n}^{\infty} (S_{2^{\ell+1}}\, g - S_{2^\ell}\, g)$$

uniformly on $[0,1)$. Since

$$S_{2^n}(\frac{k+1}{2^n}-) = S_{2^n}(\frac{k}{2^n})$$

it follows that

$$g(\frac{k+1}{2^n}-) - g(\frac{k}{2^n}) = \sum_{\ell=n}^{\infty} \sum_{s=2^\ell}^{2^{\ell+1}-1} \widehat{g}(s) \left(w_s(\frac{k+1}{2^n}-) - w_s(\frac{k}{2^n}) \right).$$

Since g is continuous and each w_s which appears in this sum is constant on dyadic intervals of the form $I(p, \ell+1)$, the claim follows at once.

Next, fix $j \in \mathbf{N}$ and choose an integer i such that $2^{i-1} \le j < 2^i$. Since w_j is constant on the intervals $I(p,i), 0 \le p < 2^i$, we have by definition that

$$\widehat{\nu}(j) = \sum_{k=0}^{2^i-1} w_j(\frac{k}{2^i}) \left(g(\frac{k+1}{2^i}) - g(\frac{k}{2^i}) \right).$$

In fact, since w_j is constant on the intervals $I(p,n), 0 \le p < 2^n$ for any $n \ge i$, we can expand this formula (by "untelescoping") to write

$$\widehat{\nu}(j) = \sum_{k=0}^{2^n-1} w_j(\frac{k}{2^n}) \left(g(\frac{k+1}{2^n}) - g(\frac{k}{2^n}) \right),$$

for any $n \ge i$. In particular, we have by the claim that

$$\widehat{\nu}(j) = \sum_{\ell=n}^{\infty} \sum_{s=2^\ell}^{2^{\ell+1}-1} \widehat{g}(s) \sum_{k=0}^{2^n-1} w_j(\frac{k}{2^n}) \left(w_s(\frac{k+1}{2^n} - \frac{1}{2^{\ell+1}}) - w_s(\frac{k}{2^n}) \right).$$

Fix $s = 2^\ell + t2^n + u$ where $t, u \in \mathbf{N}$ and $0 \le u < 2^n, 0 \le t < 2^{\ell-n}$. Then $w_s = w_{2^\ell+t2^n} w_u$ and since w_u is constant on the intervals $I(p,n)$ for $0 \le p < 2^n$ it is clear that

$$w_s(\frac{k+1}{2^n} - \frac{1}{2^{\ell+1}}) - w_s(\frac{k}{2^n}) = w_u(\frac{k}{2^n}) \left(w_q(\frac{k+1}{2^n} - \frac{1}{2^{\ell+1}}) - w_q(\frac{k}{2^n}) \right),$$

where $q = 2^\ell + t2^n$. But w_q does not depend on the first n binary coefficients of its argument and

$$\frac{k+1}{2^n} - \frac{1}{2^{\ell+1}} = \frac{k}{2^n} + \frac{1}{2^{n+1}} + \cdots + \frac{1}{2^{\ell+1}}.$$

Consequently,

$$w_s\left(\frac{k+1}{2^n} - \frac{1}{2^{\ell+1}}\right) - w_s\left(\frac{k}{2^n}\right) = w_u\left(\frac{k}{2^n}\right) \left(w_q(1 - \frac{1}{2^{\ell+1}}) - 1 \right).$$

Substituting this identity into the expression for $\widehat{\nu}$ above we obtain

$$\widehat{\nu}(j) = \sum_{\ell=n}^{\infty} \sum_{t=0}^{2^{\ell-n}-1} \sum_{u=0}^{2^n-1} \widehat{g}(2^\ell + t2^n + u) \sum_{k=0}^{2^n-1} w_{j\oplus u}(\frac{k}{2^n}) \left(w_{2^\ell+t2^n}(1 - \frac{1}{2^{\ell+1}}) - 1 \right).$$

Since the sum over k is zero for $j \ne u$ and 2^n for $j = u$ we conclude that

$$\widehat{\nu}(j) = 2^n \sum_{\ell=n}^{\infty} \sum_{t=0}^{2^{\ell-n}-1} \widehat{g}(2^\ell + t2^n + j) \left(w_{2^\ell+t2^n}(1 - \frac{1}{2^{\ell+1}}) - 1 \right). \quad (53)$$

We shall use this calculation in 2.3 to study the growth of Walsh- Fourier coefficients of continuous functions.

1.6 Dirichlet Kernels. Theorem 8 in 1.5 contains several exact representations for the Walsh-Dirichlet kernels D_k. It is convenient to have estimates for the growth of D_k as $k \to \infty$.

First we estimate the *Lebesgue constants*

$$L_k := \|D_k\|_1 \qquad (k \in \mathbf{N}).$$

In view of (49) in 1.5, these constants are important for the study of convergence of the full sequence of partial sums of Walsh-Fourier series.

If k is a positive integer written in the form

$$k = \sum_{i=1}^{j} 2^{n_i} \quad \text{for integers} \quad n_1 > n_2 > \cdots > n_j \geq 0$$

then it is not difficult to see that

$$L_k = j - \sum_{1 \leq p < r \leq j} 2^{n_r - n_p}. \tag{54}$$

Indeed, let $k = 2^n + \ell, 0 \leq \ell < 2^n$ where $n, \ell \in \mathbf{N}$ and observe by definition that

$$D_k = D_{2^n} + w_{2^n} D_\ell.$$

Since $w_{2^n} = +1$ on $[0, 2^{-n-1})$ and -1 on $[2^{-n-1}, 2^{-n})$, it follows from the Paley lemma and (4) that

$$D_k(x) = \begin{cases} 2^n + \ell & 0 \leq x < 2^{-n-1} \\ 2^n - \ell & 2^{-n-1} \leq x < 2^{-n} \\ w_{2^n} D_\ell(x) & 2^{-n} \leq x < 1. \end{cases} \tag{55}$$

Consequently,

$$\begin{aligned} L_k &= \int_0^{2^{-n-1}} (2^n + \ell) + \int_{2^{-n-1}}^{2^{-n}} (2^n - \ell) + \int_{2^{-n}}^{1} |D_\ell| \\ &= 1 + L_\ell - \int_0^{2^{-n}} |D_\ell| = 1 + L_\ell - \frac{\ell}{2^n}. \end{aligned}$$

Iteration of this identity leads to (54).

Equation (54) can be used to show that $L_k = O(\log k)$ as $k \to \infty$ (see Exercise 1.17). We shall obtain this estimate by a more direct method. First, define the ***variation*** of an $n \in \mathbf{N}$ with binary coefficients $(n_k, k \in \mathbf{N})$ by

$$V(n) := \sum_{k=1}^{\infty} |n_k - n_{k-1}| + n_0. \tag{56}$$

THEOREM 9. *If $n \in \mathbf{N}$ then*

$$\frac{V(n)}{8} \leq L_n \leq V(n). \tag{57}$$

PROOF. By Theorem 8 and Abel's transformation it is clear that

$$(58)\qquad \begin{aligned} |D_n| &= |\sum_{k=0}^{\infty} n_k(D_{2^{k+1}} - D_{2^k})| \\ &= |\sum_{k=1}^{\infty}(n_{k-1} - n_k)D_{2^k} - n_0|. \end{aligned}$$

Since the L^1 norm of D_{2^k} is 1, the right side of (57) follows immediately from (58). To verify the left side choose indices $0 \le \ell_1 \le m_1 < \ell_2 \le m_2 < \cdots < \ell_s \le m_s < \ell_{s+1} = \infty$ such that $m_j + 1 < \ell_{j+1}$ for $j = 1, 2, \ldots, s$, $n_k = 0$ for $0 \le k < \ell_1$, and

$$n_k = \begin{cases} 1 & \ell_j \le k \le m_j \\ 0 & m_j < k < \ell_{j+1}, \end{cases}$$

Use (58) to write

$$|D_n| = |\sum_{j=1}^{s}\sum_{k=\ell_j}^{m_j}(D_{2^{k+1}} - D_{2^k})| = |\sum_{j=1}^{s}(D_{2^{m_j+1}} - D_{2^{\ell_j}})|,$$

and set $A_j := [2^{-m_j-1}, 2^{-\ell_j})$ for $j = 1, 2, \ldots s$. Notice by the Paley lemma that

$$|D_n(x)| = 2^{\ell_j} - \sum_{i=1}^{j-1}(2^{m_i+1} - 2^{\ell_i}) \ge 2^{\ell_j} - 2^{m_{j-1}+1} \ge \frac{1}{2}2^{\ell_j}$$

for $x \in A_j$ and $j = 2, 3, \ldots s$. Since this inequality also holds for $j = 1$ it follows that

$$L_n \ge \sum_{j=1}^{s}\int_{A_j}|D_n| \ge \frac{1}{2}\sum_{j=1}^{s}2^{\ell_j}|A_j| > \frac{1}{4}s = \frac{1}{8}V(n)$$

for any $n \in \mathbf{N}$. ∎

Since $V(n) = O(\log n)$ we see that $L_n = O(\log n)$ as $n \to \infty$. This estimate is sharp. Indeed, a routine calculation verifies $V(N_n) = 2n$ for $N_n := \sum_{k=0}^{n-1} 2^{2k}, n \in \mathbf{P}$. Thus Theorem 9 implies that

$$L_{N_n} \ge \frac{n}{4} \ge \frac{\log(N_n)}{8}$$

for any $n \in \mathbf{P}$.

Next we turn our attention to pointwise estimates of the Walsh- Dirichlet kernels.

THEOREM 10. *If $x \in (0,1)$ and $n \in \mathbf{N}$ then*

$$|D_n(x)| \le \min\{n, \frac{2}{x}\}.$$

PROOF. Clearly, $\|D_n\|_\infty \le n$. Thus we need only show that $|D_n(x)| \le 2/x$. Fix $n \in \mathbf{N}, x \in (0,1)$ and choose $j \in \mathbf{P}$ such that

$$2^{-j} \le x < 2^{-j+1}. \tag{59}$$

Let $(n_k, k \in \mathbf{N})$ denote the binary coefficients of n and use Theorem 8 together with the Paley lemma to see

$$\begin{aligned}|D_n(x)| &= |\sum_{k=0}^{j-1} n_k r_k(x) D_{2^k}(x)| \\ &\le \sum_{k=0}^{j-1} n_k 2^k < 2^j.\end{aligned}$$

Since (59) implies $2^j < 2/x$, we conclude that $|D_n(x)| \le 2/x$. ∎

Lower pointwise estimates can also be obtained. For example, by Theorem 8 in 1.5 and (2) in 1.1 we can write

$$|D_n(x)| = |\sum_{k=0}^{j-2} n_k 2^k - n_{j-1}2^{j-1}| \tag{60}$$

for x and j satisfying (59). Hence when $n_{j-1} + n_{j-2} = 1$ it follows that

$$\begin{aligned}|D_n(x)| &\ge 2^{j-2} \\ &\ge \frac{1}{4x}.\end{aligned}$$

In particular, if $n = \sum_{k=0}^{m} 2^{2k}$, then

$$|D_n(x)| \ge \frac{1}{4x} \qquad (2^{-2m-1} \le x < 1). \tag{61}$$

Notice by Theorem 10 that $\varepsilon_n |D_n(x)| \to 0$ as $n \to \infty$ for every $x \in (0,1)$ and every $\varepsilon_n \to 0$ as $n \to \infty$. Dirichlet kernels in the Kaczmarz ordering do not satisfy this property (see Theorem 11 below). Consequently, it is harder to obtain pointwise convergence results for Walsh-Kaczmarz- Fourier series than for Walsh-Fourier series.

Let D_n^κ denote the Walsh-Kaczmarz-Dirichlet kernels. Thus $D_0^\kappa := 0$ and

$$D_n^\kappa := \sum_{j=0}^{n-1} \kappa_j \qquad (n \in \mathbf{P}).$$

There is a simple relationship between Walsh-Dirichlet kernels and Walsh-Kaczmarz-Dirichlet kernels. Indeed, temporarily define

$$T_m(x) := \sum_{i=0}^{m-1} x_{m-i-1} 2^{-i-1} + \sum_{i=m}^{\infty} x_i 2^{-i-1}$$

for $x \in [0,1)$ with dyadic expansion given by (18) and $m \in \mathbf{P}$. Then we have by (5) and (35) that

$$D_n^\kappa = D_{2^m} + (r_m)(D_k \circ T_m) \tag{62}$$

for $n = 2^m + k, 0 \le k < 2^m$, $n, m, k \in \mathbf{N}$. This formula will be used to show:

THEOREM 11. *There is a number $\alpha > 0$ such that*

$$\limsup_{n\to\infty} \frac{|D_n^\kappa(x)|}{\log n} > \alpha$$

for a.e. $x \in [0,1)$.

PROOF. Let $m \in \mathbf{N}$ and set

$$N_m := \sum_{2k<m} 2^{2k}.$$

By the Paley lemma and (62) it suffices to show

$$\limsup_{m\to\infty} \frac{|D_{N_m} \circ T_m|}{m} > \alpha$$

a.e. on $[0,1)$ for some $\alpha > 0$.

Let $q(m) := s$ where $s \in \mathbf{N}$ satisfies $2^s \le m < 2^{s+1}$, and set

$$E_m := T_m^{-1}([0,2^{-q(m)})) = T_m([0,2^{-q(m)})).$$

By (61) and the definition of N_m it is easy to see that

$$|D_{N_m}(x)| \ge \frac{2^{q(m)}}{4}$$

for any $x \in [0,2^{-q(m)})$. Since $2^{q(m)} \ge m/2$ it follows that $|D_{N_m}(T_m(y))| \ge m/8$ for all $y \in E_m$. Thus the proof of this theorem will be complete when we show

$$|\bigcap_{m=2^M}^{\infty} ([0,1) \setminus E_m)| = 0$$

for every $M \in \mathbf{N}$.

Toward this, let χ_m represent the characteristic function of the set E_m. Observe that

$$\begin{aligned}\chi_m &= 2^{-q(m)} \prod_{j=0}^{q(m)-1} (1 + r_j \circ T_m) \\ &= 2^{-q(m)} \prod_{j=0}^{q(m)-1} (1 + r_{m-1-j}).\end{aligned}$$

Set $\alpha_i := [2^i/i] - 2$ for $i \ge 4$ and let $(m_n, n \in \mathbf{N})$ represent the subsequence of integers defined by

$$\{2^i + ij : 0 \le j \le \alpha_i, i = 4,5,\dots\}.$$

Then $m_{n+1} - m_n \ge q(m_{n+1})$ for $n \in \mathbf{N}$ and the expression above shows us that the functions χ_{m_n} are stochastically independent. Consequently,

$$\int_0^1 \prod_{m=2^M}^{\infty}(1-\chi_m) \le \int_0^1 \prod_{m_n \ge 2^M}(1-\chi_{m_n})$$
$$= \prod_{m_n \ge 2^M}(1-2^{-q(m_n)}).$$

Therefore, it follows from the choice of m_n that

$$\left|\bigcap_{m=2^M}^{\infty}([0,1)\setminus E_m)\right| \le \prod_{i=M}^{\infty}\prod_{j=0}^{\alpha_i}(1-2^{-q(2^i+ij)}).$$

Since $1+x \le e^x$ for any real x we conclude that

$$\left|\bigcap_{m=2^M}^{\infty}([0,1)\setminus E_m)\right| \le \prod_{i=M}^{\infty}\prod_{j=0}^{\alpha_i}(1-2^{-i}) \le \prod_{i=M}^{\infty}\left(e^{-2^{-i}}\right)^{\alpha_i+1}$$
$$\le \prod_{i=M}^{\infty} e^{-\frac{1}{i}+2^{-i+1}} = 0$$

as required. ∎

1.7 The Dyadic Derivative. Although techniques may differ, the study of Walsh series frequently parallels that of trigonometric series. For example, most concepts introduced above have obvious classical counterparts. The dyadic version can be obtained by replacing the classical structure by the dyadic one, e.g., $+$ by $\dot{+}$ and the circle group $\mathbf{T}$ by $\mathbf{G}$. This approach will not work for differentiation. Indeed, since each Walsh function is locally constant the "derivative" defined by

$$\lim_{h\to 0}\frac{f(x\dot{+}h)-f(x)}{h}$$

leads to a differential operator which cannot distinguish one Walsh function from another.

For pedagogical reasons, we briefly shift our attention from Fourier series to Fourier transforms. In Chapter 9 we extend the Walsh functions from the index set $\mathbf{N}$ to the index set $[0,\infty)$ and dyadic addition from $[0,1)$ to $[0,\infty)$ in such a way that

$$w_{2^k}(t) = w_t(2^k) \tag{63}$$

(compare with (36) in 1.4) and

$$w_t(x\dot{+}y) = w_t(x)w_t(y) \tag{64}$$

for $k \in \mathbf{N}$ and a.e. $x, y, t \in [0, \infty)$. We shall introduce the Walsh transform

$$\widehat{f}(t) := \int_0^\infty f(x) w_t(x)\, dx$$

which is analogous to the classical Fourier transform

$$\mathcal{F}(f)(t) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-\imath xt}\, dx,$$

where $\imath := \sqrt{-1}$. We shall show under suitable conditions (e.g., see Theorem 11 in 9.4), that inversion holds, i.e.,

$$f(x) = \int_0^\infty \widehat{f}(t) w_t(x)\, dt \qquad (x \in [0, \infty)). \tag{65}$$

Recall that classical differentiation and the Fourier transform satisfy

$$\mathcal{F}(f')(t) = \imath t \mathcal{F}(f)(t)$$

for $t \in \mathbf{R}$ and smooth f. We would like to define a dyadic derivative which satisfies the analogous formula

$$\widehat{\mathbf{d}f}(t) = t\widehat{f}(t) \tag{66}$$

for $t \in [0, \infty)$. How should $\mathbf{d}$ be defined?

Proceeding formally, fix $n \in \mathbf{N}$ and use (48) in 1.5 to write the identity function on the interval $[0, 2^{n+1})$ in the form

$$t = 2^n - \sum_{k=0}^{\infty} 2^{n-k-1} w_{2^k}\left(\frac{t}{2^{n+1}}\right).$$

Observe by (66), (63), and (64) that

$$\begin{aligned}\int_0^{2^{n+1}} \widehat{\mathbf{d}f}(t) w_x(t)\, dt &= \int_0^{2^{n+1}} t\widehat{f}(t) w_x(t)\, dt \\ &= \int_0^{2^{n+1}} 2^n \widehat{f}(t) w_x(t)\, dt - \sum_{k=0}^{\infty} 2^{n-k-1} \int_0^{2^{n+1}} \widehat{f}(t) w_t(x \dot{+} 2^{k-n-1})\, dt.\end{aligned}$$

If $\widehat{f}$ is supported on $[0, 2^{n+1})$ it follows by inversion that

$$\begin{aligned}\int_0^{2^{n+1}} \widehat{\mathbf{d}f}(t) w_x(t)\, dt &= 2^n f(x) - \sum_{k=0}^{\infty} 2^{n-k-1} f(x \dot{+} 2^{k-n-1}) \\ &= \sum_{j=-\infty}^{n} 2^{j-1} \left(f(x) - f(x \dot{+} 2^{-j-1})\right).\end{aligned}$$

Hence

$$\int_0^\infty \widehat{\mathbf{d}f}(t)w_x(t)\,dt = \sum_{j=-\infty}^{\infty} 2^{j-1}\left(f(x) - f(x \dot{+} 2^{-j-1})\right).$$

We see then that the dyadic derivative of a function on $[0,\infty)$ should be $\mathbf{d}f(x) = \sum_{j=-\infty}^{\infty} 2^{j-1}\left(f(x) - f(x \dot{+} 2^{-j-1})\right)$ (see (12) in 9.1). If f is defined on $[0,1)$ and extended to $[0,\infty)$ by periodicity of period one then $f(x) = f(x \dot{+} 2^{-j-1})$ for all $j < 0$. Thus the dyadic derivative of a function on $[0,1)$ should be

$$\mathbf{d}f(x) = \sum_{j=0}^{\infty} 2^{j-1}\left(f(x) - f(x \dot{+} 2^{-j-1})\right).$$

Accordingly, for each f defined on $[0,1)$ and each $n \in \mathbf{P}$ set

$$\mathbf{d}_n f(x) := \sum_{j=0}^{n-1} 2^{j-1}\left(f(x) - f(x \dot{+} 2^{-j-1})\right)$$

for $x \in [0,1)$. We shall say that f is *dyadically differentiable* at x if

$$f^{[1]}(x) := \lim_{n\to\infty} \mathbf{d}_n f(x)$$

exists and is finite, and call $f^{[1]}$ the *dyadic derivative* of f at x. Higher order derivatives are defined recursively by

$$f^{[r]} := (f^{[r-1]})^{[1]}, \qquad r = 2, 3, \ldots.$$

We shall see that this derivative satisfies some of the usual properties but not all. For example, since $\mathbf{Q}$ is closed under the translation map $x \to x \dot{+} 2^{-j-1}$ for any $j \in \mathbf{N}$, it is clear that the function $\chi(\mathbf{Q})$ is everywhere dyadically differentiable but nowhere W-continuous.

The situation is worse if we look at classical continuity:

THEOREM 12. *If f is classically continuous on $[0,1)$ and dyadically differentiable at all but countably many points in $[0,1)$ then f is constant.*

PROOF. Let $x \in [0,1)$ be a dyadic irrational point at which $f^{[1]}$ exists. By definition, f satisfies

$$\lim_{j\to\infty} 2^j\left(f(x) - f(x \dot{+} 2^{-j})\right) = 0. \tag{67}$$

Let x have dyadic expansion given by (18) and choose integers $0 \le j_1 < j_2 < \ldots$ such that $x_{j_n} = 0$ for $n \in \mathbf{P}$. Then $x \dot{+} 2^{-j_n} = x + 2^{-j_n}$ and it follows from (67) that

$$\lim_{n\to\infty} \frac{f(x + 2^{-j_n}) - f(x)}{2^{-j_n}} = 0.$$

Hence by hypothesis the upper and lower right Dini derivates of f satisfy

$$D^+f(x) \geq 0 \geq D_-f(x)$$

for all but countably many $x \in [0,1)$. In particular, it follows from a classical theorem of Dini (see Theorem 12 in Appendix 0.6) that f is constant on $[0,1)$. ∎

In sharp contrast, each Walsh function is everywhere dyadically differentiable on $[0,1)$ and in fact,

$$w_k^{[r]} = k^r w_k \qquad (k \in \mathbf{N}, r \in \mathbf{P}). \tag{68}$$

To verify this, set $r = 1$ and fix $k \in \mathbf{N}$. Let $(k_j, j \in \mathbf{N})$ represent the binary coefficients of k and observe that

$$1 - (-1)^{k_j} = 2k_j \qquad (j \in \mathbf{N}).$$

Hence $w_k - w_k(2^{-j-1})w_k = 2k_j w_k$ for $j \in \mathbf{N}$ and it follows from definition that

$$\mathbf{d}_n w_k = \left(\sum_{j=0}^{n-1} k_j 2^j\right) w_k \qquad (n \in \mathbf{P}). \tag{69}$$

Since the sum on the right side of (69) is identically k for large n we have verified (68) for $r = 1$. Iteration establishes (68) for arbitrary r. (Notice that $w_k^{[1]} = kw_k$ is analogous to the classical formula $(e^{\imath kt})' = \imath k e^{\imath kt}$.)

Differentiation on the dyadic group is defined similarly except the usual closed system $\boldsymbol{e}$ is used in place of the sequence $(2^{-j-1}, j \in \mathbf{N})$. Specifically, set

$$\mathbf{d}_n f := \sum_{j=0}^{n-1} 2^{j-1}(f - \boldsymbol{\tau}_{e_j} f)$$

and call f differentiable at a point $x \in \mathbf{G}$ if

$$f^{[1]}(x) = \lim_{n\to\infty} \mathbf{d}_n f(x)$$

exists and is finite. As above,

$$\mathbf{d}_n \psi_k = \left(\sum_{j=0}^{n-1} k_j 2^j\right) \psi_k \qquad (n \in \mathbf{P}). \tag{70}$$

In particular, each character of $\mathbf{G}$ is infinitely differentiable everywhere on $\mathbf{G}$ and

$$\psi_k^{[r]} = k^r \psi_k \qquad (k \in \mathbf{N}, r \in \mathbf{P}).$$

For the remainder of this section let $\boldsymbol{X}$ represent a Banach space of functions in $L^1(G)$ such that $\tau_y f \in \boldsymbol{X}$, $f * g \in \boldsymbol{X}$,

$$\|\tau_y f\| = \|f\| \tag{71}$$

$$\|f\|_1 \leq \|f\|, \tag{72}$$

and

$$\|f * g\| \leq \|f\| \, \|g\|_1, \tag{73}$$

for each $f \in \boldsymbol{X}, g \in L^1(G)$, and $y \in G$. Notice that $\boldsymbol{X} = L^p(G), 1 \leq p < \infty$, and $\boldsymbol{X} = C(G)$ satisfy these conditions.

A function $f \in \boldsymbol{X}$ is called ***strongly differentiable*** in $\boldsymbol{X}$ if there is a $g \in \boldsymbol{X}$ such that $\mathbf{d}_n f \to g$ in the norm of $\boldsymbol{X}$ as $n \to \infty$. Notice by (72) that $\mathbf{d}_n f \to g$ in $L^1(G)$ norm so g is uniquely determined by f up to sets of μ measure zero. Thus we call g the ***strong derivative*** of f (in $\boldsymbol{X}$) and denote it by $\mathbf{d}f$. Higher order strong derivatives, when they exist, are defined recursively by

$$\mathbf{d}^{[1]} := \mathbf{d} \quad \text{and} \quad \mathbf{d}^{[r]} f := \mathbf{d}(\mathbf{d}^{[r-1]} f)$$

for $r = 2, 3, \ldots$. Since the sum in (70) is constant for large n it is clear that if ψ_k belongs to $\boldsymbol{X}$ then it is strongly differentiable in $\boldsymbol{X}$ and

$$\mathbf{d}^{[r]} \psi_k = k^r \psi_k \qquad (k \in \mathbf{N}, r \in \mathbf{P}).$$

Clearly, $\mathbf{d}$ is linear. Moreover, it is easy to see that $\mathbf{d}$ commutes with translations, i.e.,

$$\tau_y(\mathbf{d}f) = \mathbf{d}(\tau_y f) \qquad (y \in G)$$

holds for all strongly differentiable f. This serves as a limited replacement for the chain rule which does not hold in general for $\mathbf{d}$.

Walsh-Fourier coefficients of strongly differentiable functions behave as expected.

THEOREM 13. *If f is strongly differentiable in $\boldsymbol{X}$ then*

$$\widehat{\mathbf{d}f}(k) = \lim_{n \to \infty} \widehat{\mathbf{d}_n f}(k) \qquad (k \in \mathbf{N}). \tag{74}$$

If f is r times differentiable in $\boldsymbol{X}$ for some $r \in \mathbf{P}$ then

$$\widehat{\mathbf{d}^{[r]} f}(k) = k^r \widehat{f}(k) \qquad (k \in \mathbf{N}). \tag{75}$$

PROOF. Fix $n \in \mathbf{P}$ and $k \in \mathbf{N}$. Since $\widehat{\ }$ is linear, we have by (72) that

$$\begin{aligned} |\widehat{\mathbf{d}_n f}(k) - \widehat{\mathbf{d}f}(k)| &= |(\mathbf{d}_n f - \mathbf{d}f)\widehat{\ }(k)| \\ &\leq \|\mathbf{d}_n f - \mathbf{d}f\|_1 \\ &\leq \|\mathbf{d}_n f - \mathbf{d}f\|. \end{aligned}$$

Therefore, (74) follows at once from the strong differentiability of f.

To verify (75) we may suppose that $r = 1$. Observe by definition and (47) in 1.5 that

$$\begin{aligned}\widehat{\mathbf{d}_n f}(k) &= \int_{\mathbf{G}} \sum_{j=0}^{n-1} 2^{j-1}(f - \boldsymbol{\tau}_{e_j} f)\psi_k \, d\mu \\ &= \sum_{j=0}^{n-1} 2^{j-1}\left(\widehat{f}(k) - \psi_k(e_j)\widehat{f}(k)\right).\end{aligned}$$

Since $\widehat{f}(k) - \psi_k(e_j)\widehat{f}(k) = 2k_j\widehat{f}(k)$ it follows that

$$\widehat{\mathbf{d}_n f}(k) = k\widehat{f}(k)$$

for sufficiently large n. Thus (75) follows by letting $n \to \infty$ and applying (74). ∎

To identify the strong antiderivative, consider the operator

$$\mathbf{I}^{[r]} f := f * W_r \qquad (f \in \boldsymbol{X}),$$

where for each $r \in \mathbf{P}$ the function W_r is determined by

$$\widehat{W}_r(k) = \begin{cases} 1 & k = 0 \\ k^{-r} & k \in \mathbf{P}. \end{cases}$$

Notice by the Riesz-Fischer theorem (see Appendix 0.1) that W_r exists and belongs to $\mathrm{L}^2(\mathbf{G})$. In particular, it follows from (73) that $\mathbf{I}^{[r]} f \in \boldsymbol{X}$ for each $f \in \boldsymbol{X}$ and $r \in \mathbf{P}$. (The function W_r actually belongs to $\mathrm{Lip}(1, \mathrm{L}^1(\mathbf{G}))$ (see Exercise 3.25).)

The operator $\mathbf{I} := \mathbf{I}^{[1]}$ is sometimes referred to as the *dyadic integral*. Justification of this terminology is contained in the following result.

THEOREM 14. *If f satisfies $\widehat{f}(0) = 0$ and is r times strongly differentiable in $\boldsymbol{X}$ for some $r \in \mathbf{P}$ then*

(76) $$f = \mathbf{I}^{[r]}(\mathbf{d}^{[r]} f) \quad \text{a.e. } [\mu] \text{ on } \mathbf{G}.$$

PROOF. Fix $r, k \in \mathbf{P}$ and set $g := \mathbf{d}^{[r]} f$, $a_k := \widehat{\mathbf{I}^{[r]} g}(k)$. By the definition of W_r and (75) above, we have

$$\begin{aligned}\widehat{f}(k) &= k^r \widehat{f}(k)\widehat{W}_r(k) \\ &= \widehat{g}(k)\widehat{W}_r(k).\end{aligned}$$

On the other hand, since $\mathbf{I}^{[r]} g = W_r * g$ it follows from (45) in 1.5 that $a_k = \widehat{g}(k)\widehat{W}_r(k)$. Therefore,

$$\widehat{f}(k) = a_k$$

for all $k \in \mathbf{P}$. Since $a_0 = \widehat{f}(0) = 0$ and the system $\widehat{\mathbf{G}}$ is complete (76) follows immediately. ∎

We shall show (see Theorem 6 in 5.2 and Corollary 7 in 6.2) that (76) is valid when the order of the operators $\mathbf{d}$ and $\mathbf{I}$ is reversed and that the corresponding formulae hold for the pointwise derivative as well. These results will be referred to as the *Fundamental Theorem of Dyadic Calculus*.

The following result will be used to estimate the dyadic integral.

THEOREM 15. *Suppose x is a non-zero element of* $\mathbf{G}$. *Then*

$$\sum_{k=1}^{\infty} \frac{\psi_k(x)}{k} \quad \text{converges.}$$

Moreover,

$$| \sum_{k=1}^{2^N-1} \frac{\psi_k(x)}{k} | \le \sum_{k=0}^{\infty} 2^{-k+2} D_{2^k}(x), \tag{77}$$

for each $N \in \mathbf{P}$.

PROOF. Let $n \in \mathbf{N}$ have binary coefficients $(n_k, k \in \mathbf{N})$. Recall from Theorem 8 that

$$D_n = \psi_n \sum_{k=0}^{\infty} n_k \rho_k D_{2^k}. \tag{78}$$

Thus

$$|D_n(x)| \le \sum_{k=0}^{\infty} D_{2^k}(x) =: C(x)$$

holds for any $x \in \mathbf{G}$. If $x \neq 0$ then $C(x)$ is finite by the Paley lemma. Hence by Abel's transformation and the fact that $(1/k, k \in \mathbf{N})$ is monotone decreasing we have

$$| \sum_{k=M}^{N} \frac{\psi_k(x)}{k} | \le C(x) \frac{2}{M}$$

for any integers $0 < M \le N$. Hence the left side of (77) converges for each $x \in \mathbf{G} \setminus \{0\}$.

To obtain the sharper estimate (77) fix $N \in \mathbf{P}$, combine Abel's transformation with (78), and write

$$\begin{aligned} \sum_{n=1}^{2^N-1} \frac{\psi_n(x)}{n} &= \sum_{n=2}^{2^N-1} \frac{D_n}{n(n-1)} + \frac{D_{2^N}}{2^N-1} - 1 \\ &= \sum_{k=0}^{N-1} \rho_k D_{2^k} \left(\sum_{n=2}^{2^N-1} \frac{n_k}{n(n-1)} \psi_n \right) + \frac{D_{2^N}}{2^N-1} - 1. \end{aligned}$$

Since $n_k = 0$ for $n < 2^k$ and

$$\sum_{n=2^k}^{\infty} \frac{1}{n(n-1)} = \frac{1}{2^k - 1} \le \frac{2}{2^k}$$

for $k \in \mathbf{P}$ we conclude that

$$\sum_{k=1}^{2^N-1} \frac{\psi_k(x)}{k} \le 2 + \sum_{k=0}^{\infty} \frac{2}{2^k} D_{2^k} \le 4 \sum_{k=0}^{\infty} \frac{D_{2^k}}{2^k}$$

as required. ∎

1.8 Cesàro Summability. The *Cesàro* (or (C,1)) *means* of a series $\sum_{k=0}^{\infty} a_k$ are defined to be $\sigma_0 := 0$ and

$$\sigma_n := \frac{1}{n}\sum_{k=1}^{n} A_k \qquad (n \in \mathbf{P}),$$

where $A_k := \sum_{j=0}^{k-1} a_j$ for $k = 1, 2, \ldots$. The series $\sum_{k=0}^{\infty} a_k$ is said to be *Cesàro summable* to A if the Cesàro means σ_n converge to A as $n \to \infty$.

Since

$$|\sigma_n - A| \le \frac{1}{n}\sum_{k=1}^{n} |A_k - A|,$$

it is easy to see that if $A_k \to A$ as $k \to \infty$ then $\sigma_n \to A$ as $n \to \infty$. Thus every convergent series is Cesàro summable. On the other hand, $\sum_{k=0}^{\infty}(-1)^k$ is Cesàro summable but not convergent. Thus if a series fails to converge one may ask whether it is Cesàro summable.

For Walsh series the distinction between Cesàro summability and convergence is somewhat blurred (see remarks at the end of 5.1), especially in the case when the 2^n-th means are used. In fact, we shall see (Theorem 16 in 7.6) that if S is a Walsh series whose 2^n-th Cesàro means satisfy

$$\limsup_{n\to\infty} |\sigma_{2^n}(x)| < \infty$$

for $x \in [0, 1)$, then $S_{2^n}(x)$ converges, as $n \to \infty$, to a finite real number for a.e. $x \in [0, 1)$. Nevertheless, when using full partial sums of Walsh-Fourier series, some improvement may be obtained by passing to Cesàro means.

The *Walsh-Fejér kernels* on $\mathbf{G}$ are defined by

$$K_n := \sum_{k=0}^{n-1}\left(1 - \frac{k}{n}\right)\psi_k \qquad (n \in \mathbf{P}).$$

We shall use the same notation for the Walsh-Fejér kernels on the unit interval

$$K_n := \sum_{k=0}^{n-1}\left(1 - \frac{k}{n}\right)w_k \qquad (n \in \mathbf{P}).$$

In either case, it is clear that

$$K_n = \frac{1}{n}\sum_{i=1}^{n} D_i \qquad (n \in \mathbf{P}).$$

The Walsh-Fejér kernels play the same role for Cesàro summability that the Walsh-Dirichlet kernels do for convergence. Namely, if the Cesàro means of Sf are defined by $\sigma_0 f := 0$ and

$$\sigma_n f := \frac{1}{n}\sum_{i=1}^{n} S_i f,$$

then it is evident that

$$\sigma_n f = \frac{1}{n}\sum_{i=1}^{n} D_i * f = K_n * f \qquad (n \in \mathbf{P}).$$

Moreover, it is easy to see that

$$K_n = D_n - \frac{1}{n} D_n^{[1]} \qquad (n \in \mathbf{P}).$$

These identities will be used in Chapter VI to study a.e. Cesàro summability of Walsh-Fourier series and Walsh-Fourier-Stieltjes series.

It will be convenient to have several other identities at our disposal.

THEOREM 16.

i) If $n, k \in \mathbf{N}$ and $0 \le k \le 2^n$ then

$$(2^n + k)K_{2^n+k} = 2^n K_{2^n} + kD_{2^n} + \psi_{2^n} kK_k.$$

ii) If $k = \sum_{i=1}^{\ell} 2^{n_i}$ for some $\ell \in \mathbf{P}$, where $n_1 > n_2 > \cdots > n_\ell \ge 0$ are integers, and if $k^{(1)} := k$, $k^{(i)} := k^{(i-1)} - 2^{n_{i-1}}$ for $2 \le i \le \ell$, then

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

iii) If $n \in \mathbf{N}$ then

$$K_{2^n} = \frac{1}{2}\left(2^{-n} D_{2^n} + \sum_{j=0}^{n} 2^{j-n} \tau_{e_j} D_{2^n}\right).$$

In particular, $K_{2^n} \ge 0$ everywhere on $\mathbf{G}$ and $[0, 1)$.

iv) If $n, m \in \mathbf{P}$ and $2^{n-1} \le m < 2^n$ then

$$|K_m| \le \sum_{j=0}^{n-1} 2^{j-n} \sum_{i=j}^{n-1} (D_{2^i} + \tau_{e_j} D_{2^i}).$$

v) If $m \in \mathbf{N}$ then

$$\|K_m\|_1 \le 2.$$

PROOF. To establish i) observe since characters are homomorphisms on the dyadic group that

$$D_{2^n+i} = D_{2^n} + \psi_{2^n} D_i$$

holds for any $1 \le i \le 2^n$. Hence by definition,

$$\begin{aligned}(2^n + k)K_{2^n+k} &= \sum_{j=1}^{2^n} D_j + \sum_{i=1}^{k} D_{2^n+i} \\ &= 2^n K_{2^n} + kD_{2^n} + \psi_{2^n} kK_k.\end{aligned}$$

Part ii) follows from i) by iteration.
To prove iii) observe by (70) that

$$\begin{aligned} K_{2^n} &= D_{2^n} - 2^{-n}\mathbf{d}_n D_{2^n} \\ &= D_{2^n} - 2^{-n-1}(2^n-1)D_{2^n} + \sum_{j=0}^{n-1} 2^{j-n-1}\tau_{e_j} D_{2^n} \\ &= \frac{1}{2}\left((1+2^{-n})D_{2^n} + \sum_{j=0}^{n-1} 2^{j-n}\tau_{e_j} D_{2^n}\right) \\ &= \frac{1}{2}\left(2^{-n}D_{2^n} + \sum_{j=0}^{n} 2^{j-n}\tau_{e_j} D_{2^n}\right). \end{aligned}$$

To prove iv) use (70) to write

$$\begin{aligned} mK_m &= mD_m - \mathbf{d}_n(D_m) \\ &= mD_m - \sum_{j=0}^{n-1} 2^{j-1}(D_m - \tau_{e_j} D_m). \end{aligned}$$

Substitute $m = \sum_{j=0}^{n-1} m_j 2^j$ and $D_m = \psi_m \sum_{i=0}^{n-1} m_i \rho_i D_{2^i}$ into this identity. Thus obtain

$$mK_m = \sum_{j=0}^{n-1} 2^j \sum_{i=0}^{n-1} c_{ij}\psi_m,$$

where

$$c_{ij} := m_i m_j \rho_i D_{2^i} - \frac{m_i\rho_i}{2}\left(D_{2^i} - \psi_m(e_j)\rho_i(e_j)\tau_{e_j} D_{2^i}\right).$$

Notice that $c_{ij} = 0$ for $0 \le i < j < n$ since $\tau_{e_j} D_{2^i} = D_{2^i}$ for $i \le j$, $\rho_i(e_j) = 1$ for $i \ne j$, and

$$\frac{1-\psi_m(e_j)}{2} = m_j$$

for $j \in \mathbf{N}$. It follows that

$$\begin{aligned} |mK_m| &= |\sum_{j=0}^{n-1} 2^j \sum_{i=0}^{n-1} c_{ij}| \\ &\le \sum_{j=0}^{n-1} 2^{j-1} \sum_{i=j}^{n-1}(D_{2^i} + \tau_{e_j} D_{2^i}). \end{aligned}$$

Since $m \ge 2^{n-1}$ this inequality implies iv).

To prove v) fix $k \in \mathbf{N}$ and observe by ii) that

$$k\|K_k\|_1 \le \sum_{i=1}^{\ell} 2^{n_i}\|K_{2^{n_i}}\|_1 + \sum_{i=1}^{\ell-1} k^{(i+1)}\|D_{2^{n_i}}\|_1.$$

Recall that $\|D_{2^n}\|_1 = 1$ and use iii) to see that

$$\|K_{2^n}\|_1 = \int_0^1 K_{2^n} = 1$$

for $n \in \mathbf{N}$. Therefore,

$$\begin{aligned} k\|K_k\|_1 &\le \sum_{i=1}^{\ell} 2^{n_i} + \sum_{i=2}^{\ell} k^{(i)} \\ &\le 2\sum_{i=1}^{\ell} 2^{n_i} \\ &= 2k. \end{aligned}$$

as required. ∎

The following estimate will be used in Chapter 2 to study growth of Walsh-Fourier coefficients.

THEOREM 17. *If* $m \in \mathbf{P}$ *then*

$$\|\sum_{k=2^m}^{\infty} \frac{w_k}{k}\|_1 = O(2^{-m}) \quad \text{as} \quad m \to \infty.$$

PROOF. Apply Abel's transformation twice to obtain

$$\begin{aligned} \sum_{k=2^m}^{\infty} \frac{w_k}{k} &= \sum_{k=2^m+1}^{\infty} \frac{D_k}{k(k-1)} - \frac{D_{2^m}}{2^m} \\ &= \sum_{k=2^m+1}^{\infty} \frac{kK_k - (k-1)K_{k-1}}{k(k-1)} - \frac{D_{2^m}}{2^m} \\ &= \sum_{k=2^m+1}^{\infty} K_k\left(\frac{1}{k-1} - \frac{1}{k+1}\right) + \frac{K_{2^m-1}}{2^m+1} - \frac{D_{2^m}}{2^m}. \end{aligned}$$

Consequently, by the Paley lemma and Theorem 16 v) we have

$$\begin{aligned} \|\sum_{k=2^m}^{\infty} \frac{w_k}{k}\|_1 &\le 2\left(2^{-m} + \sum_{k=2^m+1}^{\infty} (\frac{1}{k-1} - \frac{1}{k+1})\right) + 2^{-m} \\ &\le 7 \cdot 2^{-m} \end{aligned}$$

as required. ∎

EXERCISES

1.1 For $\boldsymbol{x} = (x_n, n \in \mathbf{N}) \in \mathbf{G}$ and $n \in \mathbf{N}$ let

$$\boldsymbol{x}_n^\star := (x_0, x_1, \ldots, x_{n-1}, 1 - x_n, 1 - x_{n+1}, \ldots)$$

and

$$\boldsymbol{x}_n^\circ := (x_0, x_1, \ldots, x_{n-2}, 1 - x_{n-1}, x_n, x_{n+1}, \ldots).$$

Show that $|\boldsymbol{x} \dotplus \boldsymbol{x}_n^\star| = 2^{-n}$, $|\boldsymbol{x} \dotplus \boldsymbol{x}_n^\circ| = 2^{-n}$, $\boldsymbol{x}_n^\star \in I_n(\boldsymbol{x})$, $\boldsymbol{x}_n^\circ \notin I_n(\boldsymbol{x})$ and

$$I_n(\boldsymbol{x}) = \{\boldsymbol{y} \in \mathbf{G} : |\boldsymbol{x} \dotplus \boldsymbol{y}| < 2^{-n}\} \cup \{\boldsymbol{x}_n^\star\}$$

for all $\boldsymbol{x} \in \mathbf{G}$ and $n \in \mathbf{N}$.

1.2 An abelian group $(G, +)$ is algebraically ordered if there is a linear ordering $\leq$ on G such that $a \leq b$ implies $a + c \leq b + c$ for all $a, b, c \in G$. Show that $(\mathbf{N}, \oplus)$ and $\mathbf{G}$ are not algebraically ordered.

1.3 For every $x, y \in \mathbf{G}$ with $x < y$ set

$$(x, y) := \{z \in \mathbf{G} : x < z < y\},$$

$$[x, y) := \{z \in \mathbf{G} : x \leq z < y\},$$

$$(x, y] := \{z \in \mathbf{G} : x < z \leq y\},$$

$$[x, y] := \{z \in \mathbf{G} : x \leq z \leq y\}.$$

Show that $(x^*, x) = \emptyset$ for all $x \in \mathbf{G}_0 \setminus \{\mathbf{0}\}$. Show for

$$x = (x_0, x_1, \ldots x_{n-1}, 0, 0, \ldots), \quad y = (x_0, x_1, \ldots x_{n-1}, 1, 1, \ldots)$$

that

$$I_n(x) = I_n(y) + [x, y].$$

1.4 Show $[0, 1)$ is not closed under $\dotplus$ and that $\dotplus$ is not associative.

1.5 Suppose $x \in \mathbf{G}_0$, $x_n, y_n \in \mathbf{G}$ and $x_n \to x, y_n \to x^*$ as $n \to \infty$. Show $x \leq x_n$ and $y_n \leq x^*$ hold for all but finitely many $n \in \mathbf{N}$.

1.6 Prove (20).

1.7 Prove that the function f defined in (22) is continuous on $\mathbf{G}$.

1.8 Prove that the map

$$(x_n, n \in \mathbf{N}), (y_n, n \in \mathbf{N}) \to (x_0, y_0, x_1, y_1, \ldots)$$

is a homeomorphism from $\mathbf{G} \times \mathbf{G}$ onto $\mathbf{G}$.

1.9 Show that the characteristic function of the interval $[0, x)$ is W-continuous for some $x \in (0, 1)$ if and only if $x \in \mathbf{Q}$.

1.10 Let $f = \chi([0,1/2))$. Show $f \in C_W$ but $f(|x|)$ is not continuous on $\mathbf{G}$.

1.11 Prove that the map $x \to |x|$ is continuous from $\mathbf{G}$ to $[0,1]$ if the usual metric is used on $[0,1]$. Show this map is not continuous if the dyadic metric $\dot{+}$ is used instead.

1.12 Find an unbounded W-continuous function.

1.13 Show that the uniform closure of $\mathcal{P}$ is C_W.

1.14 Show that $\mathbf{Z}_2$ is a field. (It is called the second order Galois field.) Show that $\mathbf{G}$ is a normed linear space over $\mathbf{Z}_2$.

1.15 A Banach space $(X, \|.\|)$ is called a Banach algebra if there is an associative product $*$ on X such that

$$x*(y+z) = x*y + z*y$$

$$(x+y)*z = x*z + y*z$$

$$a(x*y) = (ax)*y = x*(ay)$$

and

$$\|x*y\| \le \|x\|\,\|y\|$$

for all $x, y, z \in X$ and scalars a. Show that ℓ^∞ and $L^1(\mathbf{G})$ are Banach algebras and that the map $\widehat{\ }$ is a 1-1 continuous Banach algebra homomorphism $L^1(\mathbf{G})$ into ℓ^∞.

1.16 Prove that

$$\operatorname{Lip}(\alpha, W) \subset \operatorname{Lip}(\alpha, L^q) \subset \operatorname{Lip}(\alpha, L^p) \subset \operatorname{Lip}(\alpha, L^1)$$

for $1 < q < p < \infty$ and $\alpha > 0$.

1.17 Use (55) in 1.6 to show that the Lebesgue constants satisfy

$$L_{2k} = L_k \qquad (k \in \mathbf{N})$$

and

$$L_{2k+1} = L_k + 1 - \frac{1}{2}\sum_{p=1}^{j} 2^{-n_p}$$

for $k = \sum_{p=1}^{j} 2^{n_p}$ and $n_1 > n_2 > \cdots > n_j \ge 0$. Prove that

$$L_{2k+1} = \frac{1 + L_k + L_{k+1}}{2} \qquad (k \in \mathbf{N})$$

and use this to show that $L_k = O(\log k)$ as $k \to \infty$. Also show there is an absolute constant $C > 0$ such that

$$\frac{1}{n}\sum_{k=1}^{n} L_k \ge C \log n \qquad (n \in \mathbf{P}).$$

[Fine [1]]

1.18 i) Show that $S_{2^n} f \to f$ in L^1 norm as $n \to \infty$ for all $f \in L^1$.

ii) Show that the set

$$\Upsilon := \{f \in L^1 : \widehat{f}(k) \to 0 \quad \text{as} \quad k \to \infty\}$$

is a closed subset of L^1.

iii) Prove that $\Upsilon = L^1$.

1.19 Suppose $f \in L^1$ and $S_n f$ converges a.e. on $[0,1)$ as $n \to \infty$. Prove that

$$f = \lim_{n\to\infty} S_n f \quad \text{a.e. on} \quad [0,1).$$

1.20 i) Given $f \in L^p(\mathbf{G}), 1 \le p < \infty$ and $\varepsilon > 0$ show there exists a Walsh polynomial P on $\mathbf{G}$ such that

$$\|P - f\|_p < \varepsilon.$$

ii) Given $f \in C(\mathbf{G})$ and $\varepsilon > 0$ show there exists a Walsh polynomial P on $\mathbf{G}$ such that

$$\|P - f\|_\infty < \varepsilon.$$

1.21 Let $1 \le p < \infty$.

i) Prove that

$$\|S_{2^n} f\|_p \le \|f\|_p$$

for all $f \in L^p$ and $n \in \mathbf{N}$.

ii) Prove that $S_{2^n} f \to f$ in L^p as $n \to \infty$ for all $f \in L^p$.

1.22 Let $(\varphi_n, n \in \mathbf{N})$ be any orthonormal system such that $\varphi_n \in L^\infty$ for $n \in \mathbf{N}$. Define a sequence of linear maps on L^1 by

$$\Lambda_n(f) := \int_0^1 f\varphi_n \qquad (n \in \mathbf{N}).$$

Prove that $\|\Lambda_n\| = \|\varphi_n\|_\infty$ and that $\Lambda_n(f) \to 0$ as $n \to \infty$ for every $f \in L^1$ if and only if

$$\sup_{n\in\mathbf{N}} \|\varphi_n\|_\infty < \infty.$$

1.23 Show that

$$\lim_{y\to 0} \omega^{(1)}(f, y) = 0$$

for every $f \in L^1$. Use this fact to prove the Riemann- Lebesgue lemma. (See also 1.18 and 1.22).

1.24 Suppose that

$$\sup_{n\in\mathbf{N}} \sum_{j=0}^{2^n-1} \omega^{(p)}(f, I(j,n)) < \infty$$

for some $1 \le p < \infty$. Show that

$$\widehat{f}(k) = O(\frac{1}{k}) \quad \text{as} \quad k \to \infty.$$

1.25 Recall that $\boldsymbol{J}_k$ represents the indefinite integral of w_k. Show that

$$f(x) := \sum_{k=0}^{\infty} \boldsymbol{J}_{2^k}(x) \qquad (x \in [0,1))$$

is continuous but nowhere differentiable on $[0,1)$.

[Van der Waerden, see also F. Riesz and B. Sz.-Nagy [1]]

1.26 For $f \in L^1$ represent the Haar-Fourier coefficients of f by

$$\overset{\bullet}{f}(k) := \int_0^1 f h_k \qquad (k \in \mathbf{N}).$$

i) If f is classically continuous on $[0,1)$ and $2^n \le k < 2^{n+1}$ show that

$$2^{n/2}|\overset{\bullet}{f}(k)| \to 0$$

as $n \to \infty$.

ii) Show that $(\overset{\bullet}{f}(k), k \in \mathbf{N})$ and $(\widehat{f}(k), k \in \mathbf{N})$ are Hadamard transforms of each other.

iii) Use i) and ii) to prove the Riemann-Lebesgue lemma for Walsh- Fourier coefficients.

1.27 Show that $\psi_0 + \psi_1 \in \mathrm{Lip}(\alpha, W)$ for all $\alpha > 0$.

1.28 i) Prove that the system

$$\gamma_n(x) := \cos(4^n \pi x) \qquad (n \in \mathbf{N}, x \in [0,1])$$

is strongly multiplicative.

ii) Construct a multiplicative system which is not strongly multiplicative.

Chapter 2

WALSH-FOURIER COEFFICIENTS

The first half of this chapter contains results which estimate the growth of Walsh-Fourier coefficients for various classes of functions, e.g., L^p functions, continuous functions, and absolutely continuous functions. The second half will identify conditions sufficient for pointwise convergence and absolute convergence of Walsh-Fourier series.

2.1 Estimates of Walsh-Fourier Coefficients. According to the Riemann-Lebesgue lemma, $\widehat{f}(k) \to 0$ as $k \to \infty$ for each $f \in L^1$. The rapidity with which $\widehat{f}$ tends to zero depends on properties of the function f.

Our first result relates the growth of $\widehat{f}$ to the dyadic L^1 modulus of continuity of f.

THEOREM 1. *If* $f \in L^1$ *then*

(1)
$$|\widehat{f}(k)| \le \frac{1}{2}\omega^{(1)}(f, \frac{1}{k}) \qquad (k \in \mathbf{P}).$$

PROOF. Fix $k \in \mathbf{P}$ and choose $n \in \mathbf{N}$ such that $2^n \le k < 2^{n+1}$. By (47) in 1.5 and linearity of $\widehat{\ }$ we see that

$$(f - \tau_{2^{-(n+1)}}f)\widehat{\ }(k) = \widehat{f}(k)(1 - w_k(2^{-(n+1)})).$$

Since the choice of n implies $w_k(2^{-(n+1)}) = -1$ it follows that

$$\widehat{f}(k) = \frac{1}{2}(f - \tau_{2^{-(n+1)}}f)\widehat{\ }(k).$$

Therefore,

$$\begin{aligned}|\widehat{f}(k)| &= \frac{1}{2}\left|\int_0^1 (f(x) - f(x \dotplus 2^{-(n+1)}))\, w_k(x)\, dx\right| \\ &\le \frac{1}{2}\omega^{(1)}(f, 2^{-(n+1)}) \\ &\le \frac{1}{2}\omega^{(1)}(f, \frac{1}{k}). \qquad \blacksquare\end{aligned}$$

If f belongs to $L^p, 1 < p < \infty$, or to C_W then (1) holds for the corresponding moduli of continuity, i.e.,

$$|\widehat{f}(k)| \le \frac{1}{2}\omega^{(p)}(f, \frac{1}{k}) \qquad (k \in \mathbf{P}, f \in L^p),$$

and

$$|\widehat{f}(k)| \le \frac{1}{2}\omega(f, \frac{1}{k}) \qquad (k \in \mathbf{P}, f \in C_W).$$

These inequalities follow immediately from Theorem 1 since

$$\omega^{(1)}(f,\delta) \le \omega^{(p)}(f,\delta) \le \omega(f,\delta)$$

for $f \in L^1$ and $\delta > 0$. In particular, we have

$$\widehat{f}(k) = O(k^{-\alpha}) \quad \text{as} \quad k \to \infty \tag{2}$$

for all $f \in \mathrm{Lip}(\alpha, W)$ and $\alpha > 0$.

The Walsh-Fourier coefficients of a strongly differentiable function f can be estimated by the L^1 modulus of continuity of $\mathbf{d}f$.

THEOREM 2. *If f is strongly differentiable in L^1 then*

$$|\widehat{f}(k)| = O\left(\frac{1}{k}\omega^{(1)}(\mathbf{d}f, \frac{2}{k})\right), \quad \text{as} \quad k \to \infty.$$

PROOF. Fix $k \in \mathbf{P}$ and choose $n \in \mathbf{N}$ such that $2^n \le k < 2^{n+1}$. Clearly, for any $g \in L^1$ we have

$$\frac{1}{2^n}\omega^{(1)}(g, \frac{1}{2^n}) \le \frac{2}{k}\omega^{(1)}(g, \frac{2}{k}).$$

Thus by (1) it suffices to show that

$$\omega^{(1)}(f, 2^{-n}) = O\left(2^{-n}\omega^{(1)}(\mathbf{d}f, 2^{-n})\right) \quad \text{as} \quad n \to \infty. \tag{3}$$

Set

$$W_1^{(n)} := \sum_{k=2^n}^{\infty} \frac{w_k}{k},$$

fix $0 \le y < 2^{-n}$ and consider the function F defined by

$$F := W_1^{(n)} * (\mathbf{d}f - \tau_y \mathbf{d}f).$$

We claim that $F = f - \tau_y f$ a.e. on $[0,1)$. Since $\boldsymbol{w}$ is complete, it is enough to show

$$\widehat{F}(j) = (f - \tau_y f)\widehat{\ }(j) \qquad (j \in \mathbf{N}). \tag{4}$$

The right side of (4) is easy to evaluate. By (47) in 1.5 we have

$$(f - \tau_y f)\widehat{\ }(j) = \begin{cases} 0 & 0 \le j < 2^n \\ (1 - w_j(y))\widehat{f}(j) & j \ge 2^n. \end{cases}$$

On the other hand, combining (45) and (47) in 1.5 with (75) in 1.7 we see that

$$\widehat{F}(j) = \widehat{W_1^{(n)}}(j)(1 - w_j(y))j\widehat{f}(j) \qquad (j \in \mathbf{N}).$$

Since $\widehat{W_1^{(n)}}(j) = 0$ for $0 \le j < 2^n$ and $\widehat{W_1^{(n)}}(j) = 1/j$ for $j \ge 2^n$ it follows that

$$\widehat{F}(j) = \begin{cases} 0 & 0 \le j < 2^n \\ (1 - w_j(y))\widehat{f}(j) & j \ge 2^n. \end{cases}$$

This verifies (4).

To obtain (3), observe by definition that

$$\omega^{(1)}(f, 2^{-n}) = \sup_{|y| \le 2^{-n}} \|f - \tau_y f\|_1.$$

Since $\|G * H\|_1 \le \|G\|_1 \, \|H\|_1$ holds for any $G, H \in \mathrm{L}^1$ we can use the claim to estimate

$$\omega^{(1)}(f, 2^{-n}) \le \|W_1^{(n)}\|_1 \omega^{(1)}(\mathbf{d}f, 2^{-n}).$$

But Theorem 17 in 1.8 shows that $\|W_1^{(n)}\|_1 = O(2^{-n})$ as $n \to \infty$. Thus (3) holds as promised. ∎

Estimate (2) holds for $\alpha = 1$ for any function of bounded fluctuation. In fact,

THEOREM 3. *If f is of bounded fluctuation then*

$$|\widehat{f}(k)| \le \frac{\mathcal{F}\ell(f)}{k} \qquad (k \in \mathbf{P}). \tag{5}$$

PROOF. Since Lebesgue measure is translation invariant with respect to $\dot{+}$ we can write

$$2\widehat{f}(k) = \int_0^1 f(x \dot{+} 2^{-(n+1)}) w_k(x \dot{+} 2^{-(n+1)})\, dx + \int_0^1 f(x) w_k(x)\, dx$$

for any integers $n, k \in \mathbf{N}$. Consequently, if $2^n \le k < 2^{n+1}$ then

$$\begin{aligned} 2\widehat{f}(k) &\le \int_0^1 |f(x \dot{+} 2^{-(n+1)}) - f(x)|\, dx \\ &= \sum_{p=0}^{2^n - 1} \int_{I(p,n)} |f(x \dot{+} 2^{-(n+1)}) - f(x)|\, dx. \end{aligned}$$

Since $|I(p,n)| = 2^{-n}$ for $0 \le p < 2^n$ it follows from the definition of local moduli of continuity that

$$2|\widehat{f}(k)| \le 2^{-n} \sum_{p=0}^{2^n - 1} |\omega(f, I(p,n))| \le 2^{-n} \mathcal{F}\ell(f).$$

Estimate (5) follows immediately from this inequality. ∎

Thus every function f of bounded variation satisfies $\widehat{f}(k) = O(1/k)$ as $k \to \infty$. In the next section (see Theorem 5) we show that "O" cannot be replaced by "o".

2.2 Walsh-Fourier Coefficients of Absolutely Continuous Functions. For the trigonometric system there is a direct relationship between smoothness of a function and how rapidly its Fourier coefficients tend to zero. For example, the Fourier coefficients of a function of bounded variation are $O(1/k)$ whereas those of absolutely continuous functions are $o(1/k)$, as $k \to \infty$.

In this section we prove that this is not the case for Walsh- Fourier coefficients. First we need a technical result concerning Walsh-Fourier coefficients of indefinite integrals.

THEOREM 4. *Suppose $f \in \mathrm{L}^1$ and $\widehat{f}(0) = 0$. If*

$$F(x) := \int_0^x f \qquad (x \in [0,1))$$

then

$$\widehat{F}(2^n + \ell) = -2^{-(n+2)}\widehat{f}(\ell) + o(2^{-n}) \tag{6}$$

holds uniformly in ℓ as $n \to \infty$.

PROOF. Fix $\ell \in \mathbf{N}$ and choose $n \in \mathbf{N}$ such that $2^n > \ell$. Set $k := 2^n + \ell$ and recall from (48) in 1.5 that the indefinite integral $\boldsymbol{J}_k$ satisfies

$$\boldsymbol{J}_k = 2^{-n-2}(w_\ell - \sum_{i=1}^{\infty} 2^{-i} w_{2^{n+i}+k}). \tag{7}$$

By hypothesis $F(1) = 0$. Since $\boldsymbol{J}_k(0) = 0$, integration by parts yields

$$\begin{aligned} \widehat{F}(2^n + \ell) &= \int_0^1 F(x) w_k(x)\, dx \\ &= -\int_0^1 \boldsymbol{J}_k(x) f(x)\, dx. \end{aligned}$$

Thus (7) implies

$$\begin{aligned} |\widehat{F}(2^n + \ell) + 2^{-(n+2)}\widehat{f}(\ell)| &= 2^{-(n+2)} \left| \sum_{i=1}^{\infty} 2^{-i} \widehat{f}(2^{n+i} + k) \right| \\ &\leq 2^{-(n+2)} \sup_{j > 2^n} |\widehat{f}(j)|. \end{aligned}$$

In particular, (6) follows from the Riemann-Lebesgue lemma. ∎

There exist functions of bounded variation whose Walsh-Fourier coefficients are not $o(1/k)$ as $k \to \infty$. In fact, THEOREM 5. *If F is absolutely continuous on $[0,1]$ and if*

$$\widehat{F}(k) = o(\frac{1}{k}) \quad \text{as} \quad k \to \infty \tag{8}$$

then F is constant on $[0,1]$.

PROOF. Let $f := F'$, fix $\ell \in \mathbf{N}$, and observe by Theorem 4 that

$$\lim_{n\to\infty} 2^n \widehat{F}(2^n + \ell) = -\frac{1}{4}\widehat{f}(\ell).$$

It follows from (8) that $\widehat{f}(\ell) = 0$ for all $\ell \in \mathbf{N}$. Therefore, $f = 0$ a.e. and F is constant on $[0,1]$. ∎

Thus the Walsh-Fourier coefficients of a smooth function cannot decay too rapidly. This is no surprise since the Walsh functions are not themselves continuous. For example, the absolutely convergent series

$$f := \sum_{k=0}^{\infty} \frac{w_{k+1}}{2^k}$$

cannot be continuous. Indeed, the initial jump of size 2 which w_1 contributes cannot be cancelled out by the rest of the series since $\sum_{k=1}^{\infty} 1/2^k = 1$. On the other hand, the Walsh-Fourier series of a function in C_W can decay as rapidly as one wishes.

The next result restores analogy with the trigonometric system by revealing that on the average, Walsh-Fourier coefficients of absolutely continuous functions are $o(1/k)$ as $k \to \infty$.

THEOREM 6. *Suppose $f \in L^1$ and $\widehat{f}(0) = 0$. If*

$$F(x) := \int_0^x f \qquad (x \in [0,1))$$

then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} j|\widehat{F}(j)| = 0. \tag{9}$$

PROOF. Set

$$\eta(k) := \frac{1}{k}\sum_{j=1}^{k} j|\widehat{F}(j)| \qquad (k \in \mathbf{P}).$$

Since $\eta(2^{m-1})/2 \leq \eta(k) \leq 2\eta(2^m)$ for $2^{m-1} \leq k < 2^m$ it suffices to show

$$\lim_{m\to\infty} \eta(2^m) = 0.$$

Let $\varepsilon > 0$. By Theorem 4 and the Riemann-Lebesgue lemma choose $N \in \mathbf{N}$ such that

$$(2^n + \ell)|\widehat{F}(2^n + \ell)| \leq |\widehat{f}(\ell)| + \varepsilon \leq 2\varepsilon$$

for $2^N \le \ell < 2^n$ and $n \ge N$. Thus for $m > N+1$ we have

$$\eta(2^m) = 2^{-m} \sum_{k=1}^{2^{N+1}-1} k|\widehat{F}(k)| + 2^{-m} \sum_{n=N+1}^{m-1} \left(\sum_{\ell=0}^{2^N-1} + \sum_{\ell=2^N}^{2^n-1} \right) (2^n+\ell)|\widehat{F}(2^n+\ell)|$$
$$\le 2^{-m}2^{2N+2}\|F\|_1 + 2^{-m}(m-N-1)2^N(\|f\|_1+\varepsilon) + 2\varepsilon.$$

We conclude that

$$\limsup_{m\to\infty} \eta(2^m) \le 2\varepsilon. \qquad \blacksquare$$

Theorems 5 and 6 contain the following information. If F is a non-constant absolutely continuous function on $[0,1]$ then

$$\limsup_{k\to\infty} k|\widehat{F}(k)| > 0 \tag{10}$$

but

$$\liminf_{k\to\infty} k|\widehat{F}(k)| = 0. \tag{11}$$

2.3 Walsh-Fourier Coefficients of Continuous Functions. At this point it is natural to ask how rapidly the Walsh-Fourier coefficients of a non-constant continuous function can decay. One answer to this question is contained in the following result.

THEOREM 7. *Let f be classically continuous on $[0,1]$, and let $(\varepsilon_k, k \in \mathbf{N})$ be a non-increasing sequence in ℓ^1. If*

$$\widehat{f}(k) = O(\varepsilon_k) \quad as \quad k \to \infty \tag{12}$$

then f is constant on $[0,1]$.

PROOF. Let ν be the quasi-measure defined by

$$\nu([\alpha,\beta)) := f(\beta) - f(\alpha) \qquad (\alpha,\beta \in \mathbf{Q}). \tag{13}$$

We will show that $\widehat{\nu}(j) = 0$ for all $j \in \mathbf{N}$.

Fix $j \in \mathbf{N}$. Recall from (53) in 1.5 that

$$\widehat{\nu}(j) = 2^n \sum_{\ell=n}^{\infty} \sum_{t=0}^{2^{\ell-n}-1} \widehat{f}(2^\ell + t2^n + j) \left(w_{2^\ell+t2^n}(1 - \frac{1}{2^{\ell+1}}) - 1 \right)$$

holds for any $n \ge i$ and $2^{i-1} \le j < 2^i$. Thus

$$|\widehat{\nu}(j)| \le 2^{n+1} \sum_{\ell=n}^{\infty} 2^{\ell-n} \max_{2^\ell \le s < 2^{\ell+1}} |\widehat{f}(s)|$$

and it follows from (12) that

$$|\widehat{\nu}(j)| \leq C \sum_{\ell=n}^{\infty} 2^{\ell} \varepsilon_{2^{\ell}}$$

for some constant C and n sufficiently large. Since $(\varepsilon_k, k \in \mathbf{N})$ is non-increasing we conclude that

$$|\widehat{\nu}(j)| \leq C \sum_{k=2^n}^{\infty} \varepsilon_k$$

for n sufficiently large.

We have proved that $\widehat{\nu}(j) = 0$ for $j \in \mathbf{N}$. But the map $\widehat{\ }$ is 1-1 from QM onto ℓ^0. Consequently, the quasi-measure ν must be identically zero. In particular, f is constant by (13). ∎

(We notice that the hypotheses of Theorem 7 imply that the Walsh- Fourier series of f converges absolutely. For the case of continuous functions whose Walsh-Fourier series do not converge absolutely see Corollary 3 in 8.2 and the example which follows its proof.)

An immediate consequence of this result is the following.

COROLLARY 1. *If f is continuous on $[0,1]$ and*

$$\widehat{f}(k) = O\left(\frac{1}{k(\log k)^{\alpha}}\right) \quad \text{as} \quad k \to \infty \tag{14}$$

for some $\alpha > 1$, then f is constant.

This result fails when $\alpha = 1$ (see Theorem 8 below). First we introduce the Cantor-Lebesgue functions.

Let $(I_n := [\alpha_n, \beta_n], n \in \mathbf{P})$ be a sequence of closed intervals in the interval $[0,1]$ which satisfies the following four properties:

(i) $$I_{2^{n+1}+2s+i} \subset I_{2^n+s} \qquad (i = 0, 1)$$

(ii) $$I_{2^{n+1}+2s} \cap I_{2^{n+1}+2s+1} = \emptyset$$

(iii) $$\alpha_{2^n+s} = \alpha_{2^{n+1}+2s}, \qquad \beta_{2^n+s} = \beta_{2^{n+1}+2s+1}$$

for $0 \leq s < 2^n$, $n \in \mathbf{N}$, and

(iv) $$\lim_{n\to\infty} \max_{0 \leq s < 2^n} |I_{2^n+s}| = 0.$$

The set

$$E := \bigcap_{k=1}^{\infty} \left(\bigcup_{0 \leq s < 2^k} I_{2^k+s} \right) \tag{15}$$

is called the Cantor set generated by the intervals $(I_n, n \in \mathbf{P})$. Of course, the "middle thirds" Cantor set eventuates when the intervals $I_1 = [0,1], I_2 = [0,1/3], I_3 = [2/3,1], \ldots$ are used.

Associate with the set E a monotone non-decreasing function as follows. For each closed interval I set

$$R_I(x) := \int_0^x \frac{\chi(I)}{|I|}. \tag{16}$$

For each integer $k \in \mathbf{N}$ set

$$F_k := 2^{-k} \sum_{s=0}^{2^k-1} R_{I_{2^k+s}}.$$

Observe by properties (i), (ii), and (iii) above that F_k is continuous and non-decreasing on $[0,1]$. In fact, each F_k is linear on the intervals I_{2^k+s} and constant on each component of the set

$$[0,1] \setminus \left(\bigcup_{0 \le s < 2^k} I_{2^k+s} \right).$$

In particular, it is not difficult to see that

$$F(x) := \lim_{k \to \infty} F_k(x) \qquad (x \in [0,1]) \tag{17}$$

exists, is continuous, monotone non-decreasing and satisfies $F(0) = 0, F(1) = 1$. This function will be called the Cantor- Lebesgue function associated with the set E.

THEOREM 8. *There is a non-constant continuous function F on $[0,1]$ whose Walsh-Fourier coefficients satisfy*

$$\widehat{F}(k) = O\left(\frac{1}{k(\log k)}\right) \quad \text{as} \quad k \to \infty. \tag{18}$$

PROOF. Let E be the Cantor set determined by intervals which satisfy $I_1 = [0,1]$ and $|I_n| = 2^{-n}$ for $n \ge 2$ (they will be specified below). Let F be the Cantor-Lebesgue function associated with E. Since F is monotone non-decreasing, it is easy to see that

$$|\widehat{F}(2^m + j)| \le |\widehat{F}(2^m)|$$

for $0 \le j < 2^m, m \in \mathbf{N}$. In particular, (18) will be established when we show

$$\widehat{F}(2^m) = O\left(\frac{1}{m2^m}\right) \quad \text{as} \quad m \to \infty. \tag{19}$$

This estimate is delicate and we have broken it into several steps. First, let

$$I = [\alpha, \alpha + 2^{-n}]$$

be a closed subinterval of $[0,1]$ and R_I be defined by (16) above. Since the Paley lemma implies

$$\chi(I)(t) = 2^{-n}D_{2^n}(t \dot{+} \alpha) \qquad (t \in [0,1)),$$

we can write R_I in the form

$$R_I(x) = \sum_{k=0}^{2^n-1} w_k(\alpha)\boldsymbol{J}_k(x) \qquad (x \in [0,1)).$$

Now the indefinite integrals $\boldsymbol{J}_k$ of w_k satisfy

$$\boldsymbol{J}_{2^s+\ell}(x) = w_\ell(x)\boldsymbol{J}_{2^s}(x)$$

for $0 \le \ell < 2^s$ and $x \in [0,1)$. It follows that

$$R_I(x) = x + \sum_{s=0}^{n-1}\sum_{\ell=0}^{2^s-1} w_{2^s+\ell}(\alpha)w_\ell(x)\boldsymbol{J}_{2^s}(x).$$

Therefore, for such an interval I we have that

$$(20) \qquad R_I(x) = x + \sum_{s=0}^{n-1} r_s(\alpha)D_{2^s}(x \dot{+} \alpha)\boldsymbol{J}_{2^s}(x) \qquad (x \in [0,1)).$$

Next, let $I' = [\alpha, \alpha + 2^{-2n}]$ and $I'' = [\alpha'', \alpha'' + 2^{-2n-1}]$ where $\alpha'' = \alpha + 2^{-n} - 2^{-2n-1}$. Consider the function

$$T_I := R_{I'} + R_{I''} - 2R_I.$$

Since $\alpha'' + 2^{-2n-1} = \alpha + 2^{-n}$ and I is a dyadic interval, it is evident that

$$r_j(\alpha) = r_j(\alpha'') \qquad (0 \le j < n),$$

$$r_j(\alpha) = -r_j(\alpha'') = 1 \qquad (n \le j \le 2n),$$

and

$$D_{2^j}(x \dot{+} \alpha) = D_{2^j}(x \dot{+} \alpha'') \qquad (0 \le j \le n).$$

Consequently, (20) implies that

$$\begin{aligned} T_I(x) &= \sum_{s=n+1}^{2n-1} \left(D_{2^s}(x \dot{+} \alpha) - D_{2^s}(x \dot{+} \alpha'')\right)\boldsymbol{J}_{2^s}(x) - D_{2^{2n}}(x \dot{+} \alpha'')\boldsymbol{J}_{2^{2n}}(x) \\ &= \sum_{s=n+1}^{2n-1} \boldsymbol{J}_{2^s}(x) \sum_{k=2^n}^{2^s-1}(w_k(\alpha) - w_k(\alpha''))w_k(x) - \boldsymbol{J}_{2^{2n}}(x) \sum_{k=0}^{2^{2n}-1} w_k(\alpha'')w_k(x) \end{aligned}$$

for $x \in [0,1)$. Since $\sum_{s=m+1}^{2n-1} 2^{-s-1} = 2^{-m-1} - 2^{-2n}$ it follows from (48) in 1.5 that

$$\widehat{T}_I(2^m) = \begin{cases} -r_m(\alpha)2^{-2n-2} & m < n \\ 2^{-m-1} - 3(2^{-2n-2}) & n \le m \le 2n-1 \\ 0 & m \ge 2n. \end{cases} \tag{21}$$

Now let $I_n := [\alpha_n, \alpha_n + 2^{-n}]$ for $n \ge 2$, where the numbers α_n are chosen so that the intervals $I = I_n, I' = I_{2n}$, and $I'' = I_{2n+1}$ enjoy the relationship used to derive (21). Define

$$T_{I_n} := R_{I_{2n}} + R_{I_{2n+1}} - 2R_{I_n}$$

for $n \in \mathbf{P}$ and notice by definition that

$$F = F_0 + \sum_{k=1}^{\infty} 2^{-k-1} \sum_{s=0}^{2^k-1} T_{I_{2^k+s}}.$$

Thus

$$F = F_0 + \sum_{n=2}^{\infty} \widetilde{T}_n \tag{22}$$

for

$$\widetilde{T}_n := 2^{-k-1} T_{I_n} \quad (2^k \le n < 2^{k+1}, n \in \mathbf{N}).$$

Moreover, we have by (21) that

$$\widehat{\widetilde{T}_n}(2^m) = \begin{cases} -r_m(\alpha_n)2^{-2n-k-3} & m < n \\ 2^{-m-k-2} - 3(2^{-2n-k-3}) & n \le m \le 2n-1 \\ 0 & m \ge 2n \end{cases} \tag{23}$$

where n and k are related by $2^k \le n < 2^{k+1}$.

To verify (19) fix $m \in \mathbf{P}$ and use (22) to split F into three pieces, $F = F_0 + G + H$, where

$$G := \sum_{n \in [\frac{m}{2} + \frac{1}{2}, m]} \widetilde{T}_n$$

and

$$H := \sum_{n \notin [\frac{m}{2} + \frac{1}{2}, m]} \widetilde{T}_n$$

To estimate $\widehat{F}_0$ use $F_0(x) := x$ and recall by (48) in 1.5 that

$$x = \frac{1}{2} - \frac{1}{4} \sum_{m=0}^{\infty} 2^{-m} w_{2^m}(x)$$

for $x \in [0,1)$. Consequently,

$$\widehat{F}_0(2^m) = -2^{-m-2}.$$

To estimate $\widehat{G}$ let $m := 2^k + 2u + v$ where $v = 0$ or 1 and $0 \leq u < 2^{k-1}$. Observe by (23) that

$$\begin{aligned}\widehat{G}(2^m) &= 2^{-k-1} \sum_{n=2^{k-1}+u+1}^{2^k-1} (2^{-m} - 3 \cdot 2^{-2n-1}) + 2^{-k-2} \sum_{n=2^k}^{2^k+2u+v} (2^{-m} - 3 \cdot 2^{-2n-1}) \\ &= \left((2^{k-1} - u - 1)2^{-k-1} + (2u + v + 1)2^{-k-2}\right) 2^{-m} + O(2^{-2^k-2u}2^{-k}) \\ &= 2^{-m-2} + O(\frac{1}{m2^m}) \quad \text{as} \quad m \to \infty.\end{aligned}$$

Similarly, since

$$H = \sum_{n=2}^{2^{k-1}+u+v} \widetilde{T}_n + \sum_{n=m+1}^{\infty} \widetilde{T}_n,$$

it also follows from (23) that

$$\widehat{H}(2^m) = O(2^{-k}2^{-2m}) = O(\frac{1}{m2^m}) \quad \text{as} \quad m \to \infty.$$

This completes the proof of (19). ∎

2.4 Absolutely Convergent Walsh-Fourier Series. We shall denote the collection of functions $f \in \mathrm{L}^1$ whose Walsh-Fourier series converge absolutely by A. Since $|w_k| = 1$ for $k \in \mathbf{N}$ it is clear that $f \in \mathrm{A}$ if and only if

$$\|f\|_{\mathrm{A}} := \sum_{k=0}^{\infty} |\widehat{f}(k)|$$

is finite. Moreover, each $f \in \mathrm{A}$ is a W-continuous function since it has a uniformly convergent Walsh-Fourier series.

It is evident that $\|.\|_{\mathrm{A}}$ is a norm. Since $\widehat{\ }$ is a 1-1 linear isometry from A onto ℓ^1 it is also evident that A is a Banach space. In fact, A is a Banach algebra. First, it is closed under pointwise multiplication. Indeed, by rearranging the Walsh-Fourier series of $f, g \in \mathrm{A}$ we can write

$$\begin{aligned} fg &= \sum_{m,n=0}^{\infty} \widehat{f}(m)\widehat{g}(n)w_{m\oplus n} \\ &= \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \widehat{f}(m)\widehat{g}(m \oplus k)w_k. \end{aligned}$$

Since this last series converges uniformly, it must be the Walsh- Fourier series of fg. Next, it follows that

$$\widehat{fg}(m) = \sum_{k=0}^{\infty} \widehat{f}(m)\widehat{g}(m \oplus k) \qquad (m \in \mathbf{N})$$

and

$$\|fg\|_{\mathrm{A}} \le \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} |\widehat{f}(m)\widehat{g}(m \oplus k)|$$
$$= \|f\|_{\mathrm{A}} \, \|g\|_{\mathrm{A}}.$$

Hence A is a Banach algebra.

We shall obtain several conditions sufficient for a function f to belong to A. The first one is a condition on the oscillation of f as measured by the local modulus of continuity.

THEOREM 9. *If* $f \in \mathrm{C}_W, 1 \le p \le 2$, *and*

$$\sum_{n=1}^{\infty} \left(\sum_{k=0}^{2^n-1} |\omega(f, I(k,n))|^p \right)^{1/p} < \infty,$$

then $f \in \mathrm{A}$.

PROOF. Fix $n \in \mathbf{N}$. Let q be the index conjugate to p and observe by Hölder's inequality that

$$\sum_{j=2^n}^{2^{n+1}-1} |\widehat{f}(j)| \le (2^n)^{1/p} \left(\sum_{j=2^n}^{2^{n+1}-1} |\widehat{f}(j)|^q \right)^{1/q}. \tag{24}$$

Fix $y \in I(1, n+1)$ and set

$$g := \tau_y f - f.$$

Then $w_j(y) = -1$ for $2^n \le j < 2^{n+1}$. Moreover, by (47) in 1.5 we have

$$\widehat{g}(j) = \widehat{f}(j)(w_j(y) - 1) \qquad (j \in \mathbf{N}).$$

Therefore,

$$\widehat{g}(j) = -2\widehat{f}(j) \qquad (2^n \le j < 2^{n+1}).$$

Hence it follows from the Hausdorff-Young inequality (see Appendix 0.1) that

$$\left(\sum_{j=2^n}^{2^{n+1}-1} |\widehat{f}(j)|^q \right)^{1/q} = \frac{1}{2} \left(\sum_{j=2^n}^{2^{n+1}-1} |\widehat{g}(j)|^q \right)^{1/q}$$
$$\le \left(\int_0^1 |g(x)|^p \, dx \right)^{1/p}$$
$$= \left(\int_0^1 |f(x \dotplus y) - f(x)|^p \, dx \right)^{1/p}.$$

Combine this inequality with (24). Thus verify

$$\sum_{j=2^n}^{2^{n+1}-1} |\widehat{f}(j)| \le \left(2^n \int_0^1 |f(x \dot{+} y) - f(x)|^p \, dx\right)^{1/p}$$
$$= \left(2^n \sum_{k=0}^{2^n-1} \int_{I(k,n)} |f(x \dot{+} y) - f(x)|^p \, dx\right)^{1/p}.$$

Since $|I(k,n)| = 2^{-n}$ for $0 \le k < 2^n$ it follows from the definition of the local modulus of continuity that

$$\sum_{j=2^n}^{2^{n+1}-1} |\widehat{f}(j)| \le \left(\sum_{k=0}^{2^n-1} |\omega(f, I(k,n))|^p\right)^{1/p}.$$

Summing this inequality over $n \in \mathbf{N}$ we conclude that

$$\sum_{\ell=1}^{\infty} |\widehat{f}(\ell)| < \infty. \qquad \blacksquare$$

COROLLARY 2. *If the global modulus of continuity of a function f satisfies*

$$\sum_{n=0}^{\infty} 2^{n/2} \omega(f, 2^{-n}) < \infty$$

then $f \in \mathrm{A}$.

PROOF. For each $n \in \mathbf{N}$ we have by definition that

$$\omega(f, I(k,n)) \le \omega(f, 2^{-n}) \qquad (0 \le k < 2^n).$$

Thus the hypothesis implies

$$\sum_{n=0}^{\infty} \left(\sum_{k=0}^{2^n-1} |\omega(f, I(k,n))|^2\right)^{1/2} \le \sum_{n=0}^{\infty} \left(2^n |\omega(f, 2^{-n})|^2\right)^{1/2} < \infty.$$

In particular, f belongs to A by Theorem 9. $\blacksquare$

The following result is an analogue of Bernstein's theorem for trigonometric Fourier series.

COROLLARY 3. *If $f \in \mathrm{Lip}(\alpha, W)$ for some $\alpha > 1/2$ then $f \in \mathrm{A}$.*

PROOF. By definition, such an f satisfies

$$\omega(f, 2^{-n}) \le C 2^{-n\alpha} \qquad (n \in \mathbf{N})$$

for some constant $C > 0$ which depends only on f. Therefore,

$$\sum_{n=0}^{\infty} 2^{n/2}\omega(f, 2^{-n}) \le C \sum_{n=0}^{\infty} 2^{n(\frac{1}{2}-\alpha)}.$$

Since $\alpha > 1/2$ this last series converges. Hence by Corollary 2, f belongs to A. ∎

Corollary 3 cannot be relaxed to $\alpha = 1/2$. Indeed, for each $n \in \mathbf{P}$ set

$$P_n(x) := 2^{-2n} r_{2n}(x) \sum_{k=0}^{2^n-1} w_{k2^n}(x) D_{2^n}(x \dot{+} \frac{k}{2^n}) \qquad (x \in [0,1)).$$

Observe since the functions $\tau_{k2^{-n}} D_{2^n}$, $k = 0, 1, \ldots 2^n - 1$ have pairwise non-overlapping supports that

(25) $$\|P_n\|_\infty \le 2^{-n}.$$

Hence the function

$$f(x) := \sum_{n=1}^{\infty} P_n(x) \qquad (x \in [0,1))$$

belongs to $\mathbf{C}_W$. Moreover, since the spectra of P_n satisfy

$$sp(P_n) \subseteq [2^{2n}, 2^{2n+1})$$

and

$$\sum_{k=0}^{\infty} |\widehat{P}_n(k)| = 1 \qquad (n \in \mathbf{P})$$

it is clear that $\widehat{f} \notin \ell^1$.

It remains to see that $f \in \mathrm{Lip}(\frac{1}{2}, W)$. Fix $y \in [0,1)$ and choose $n_0 \in \mathbf{P}$ such that

$$2^{-2n_0-1} \le y < 2^{-2n_0+1}.$$

By construction,

$$|P_n(x \dot{+} y) - P_n(x)| \le \begin{cases} 0 & n < n_0 \\ 2\|P_n\|_\infty & n \ge n_0. \end{cases}$$

Therefore, it follows from (25) that

$$|f(x \dot{+} y) - f(x)| \le 4(2^{-n_0}) \le 8\sqrt{y}.$$

In particular, f belongs to $\mathrm{Lip}(\frac{1}{2}, W)$.

The condition on α can be relaxed if f is of bounded variation. In fact, COROLLARY 4. *If f is of p-bounded fluctuation for some $1 \le p < 2$, and if $f \in \mathrm{Lip}(\alpha, W)$ for some $\alpha > 0$ then $f \in \mathrm{A}$.*

PROOF. By hypothesis, the number

$$\mathcal{F}\ell_p(f) := \sup_{n\in\mathbf{P}} \left(\sum_{k=0}^{2^n-1} |\omega(f, I(k,n))|^p \right)^{1/p}$$

is finite. To apply Theorem 9, notice that

$$\begin{aligned}\Delta &:= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{2^n-1} |\omega(f, I(k,n))|^2 \right)^{1/2} \\ &= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{2^n-1} |\omega(f, I(k,n))|^{2-p} |\omega(f, I(k,n))|^p \right)^{1/2} \\ &\le \sum_{n=0}^{\infty} |\omega(f, 2^{-n})|^{(2-p)/2} \left(\sum_{k=0}^{2^n-1} |\omega(f, I(k,n))|^p \right)^{1/2} .\end{aligned}$$

Since $f \in \mathrm{Lip}(\alpha, W)$ it follows that there is a constant $C > 0$ which depends only on f such that

$$\Delta \le \sum_{n=0}^{\infty} C(2^{-n})^{\alpha(2-p)/2} |\mathcal{F}\ell_p(f)|^{p/2}.$$

Since the exponent of 2^{-n} is positive we conclude that Δ is finite. Hence $f \in \mathrm{A}$ by Theorem 9. ∎

Since there exists $f \in \mathrm{Lip}(\alpha, W)$ such that $\widehat{f} \notin \ell^1$ it is natural to ask whether $\widehat{f} \in \ell^\beta$ for any $\beta > 1$. The following answer to this question is an analogue to a trigonometric theorem of Szász.

THEOREM 10. *If $f \in \mathrm{Lip}(\alpha, W)$ for some $\alpha > 0$ and if $\beta > 2/(2\alpha + 1)$ then*

$$\sum_{k=0}^{\infty} |\widehat{f}(k)|^\beta < \infty.$$

PROOF. Since $\mathrm{Lip}(\alpha, W) \subset \mathrm{L}^2$ we have $\widehat{f} \in \ell^2 \subset \ell^\beta$ for all $\beta \ge 2$. Hence we may suppose that $\beta < 2$.

Fix $n \in \mathbf{N}$. Apply Hölder's inequality to obtain

$$\begin{aligned}\Delta_n &:= \sum_{k=2^n}^{2^{n+1}-1} |\widehat{f}(k)|^\beta \\ &\le \left(\sum_{k=2^n}^{2^{n+1}-1} |\widehat{f}(k)|^2 \right)^{\beta/2} (2^n)^{1-\beta/2}.\end{aligned}$$

Retrace the proof of Theorem 9 to verify

$$\Delta_n \leq \left(\frac{1}{4}\sum_{k=0}^{2^n-1}\frac{1}{2^n}|\omega(f,I(k,n))|^2\right)^{\beta/2}(2^n)^{1-\beta/2}$$
$$\leq (\frac{1}{2})^\beta\left(2^n\frac{1}{2^n}|\omega(f,2^{-n})|^2\right)^{\beta/2}(2^n)^{1-\beta/2}.$$

Thus by hypothesis there is a constant $C > 0$ such that

$$\Delta_n \leq C(2^n)^{1-\alpha\beta-\beta/2}.$$

Summing this inequality over $n = 0, 1, \ldots$ we conclude that

$$\sum_{k=1}^{\infty}|\widehat{f}(k)|^\beta \leq C\sum_{n=0}^{\infty}(2^n)^{1-\alpha\beta-\beta/2} < \infty. \qquad \blacksquare$$

Theorem 10 is sharp. The method used to show Corollary 3 is sharp can be used to construct an $f \in \text{Lip}(\alpha, W)$ for which $\widehat{f} \notin \ell^\beta$ for $\beta = 2/(2\alpha+1)$.

THEOREM 11. *If* $f \in \text{Lip}(\alpha, W)$ *for some* $0 < \alpha \leq 1$ *then*

$$\sum_{k=1}^{\infty}k^{-\beta}|\widehat{f}(k)| < \infty$$

for any $\beta > 1/2 - \alpha$.

PROOF. Retrace the proof of Theorem 9 with $p = 2$ to see that

$$\sum_{k=2^n}^{2^{n+1}-1}|\widehat{f}(k)| \leq C(2^n)^{1/2-\alpha}$$

for some constant $C > 0$ and all $n \in \mathbf{N}$. On the other hand, it is clear that

$$\Delta_n := \sum_{k=2^n}^{2^{n+1}-1}k^{-\beta}|\widehat{f}(k)|$$
$$\leq 2^{|\beta|}(2^n)^{-\beta}\sum_{k=2^n}^{2^{n+1}-1}|\widehat{f}(k)|.$$

Consequently,

$$\Delta_n \leq C2^{|\beta|}(2^n)^{1/2-\alpha-\beta}$$

for all $n \in \mathbf{N}$. Adding these inequalities over $n = 0, 1, \ldots$ we obtain

$$\sum_{k=1}^{\infty}k^{-\beta}|\widehat{f}(k)| \leq C2^{|\beta|}\sum_{n=0}^{\infty}(2^n)^{1/2-\alpha-\beta}$$

for all β. Since by hypothesis $1/2 - \alpha - \beta < 0$, the proof of the theorem is complete. ∎

For each $f \in L^2$ and $n \in \mathbf{P}$ let

$$e_n(f) := \inf\{\|f - P\|_2\}$$

where the infimum is taken over all Walsh polynomials P whose spectra contain at most n points. Thus $e_n(f)$ represents a measure of how well a function f can be approximated in L^2 norm by a Walsh polynomial which has at most n non-zero terms.

As Stečkin did for the trigonometric case, it is possible to characterize functions $f \in A$ by the growth of $(e_n(f), n \in \mathbf{N})$.

THEOREM 12. *A function $f \in L^2$ belongs to A if and only if*

$$\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} e_n(f) < \infty.$$

PROOF. Let $(a_k, k \in \mathbf{N})$ be a non-increasing rearrangement of $(|\widehat{f}(k)|, k \in \mathbf{N})$. Thus

$$\sum_{k=0}^{\infty} |\widehat{f}(k)| = \sum_{k=0}^{\infty} a_k$$

and by the Riesz-Fischer theorem we have

$$e_n(f) = \left(\sum_{k=n}^{\infty} a_k^2\right)^{1/2}.$$

Since both $(a_k, k \in \mathbf{N})$ and $(e_k(f), k \in \mathbf{N})$ are non- increasing, it is easy to see that the series $\sum_{n=0}^{\infty} a_n$, $\sum_{n=0}^{\infty} 2^n a_{2^n}$ are equiconvergent, and the series

$$\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} e_n(f), \qquad \sum_{n=1}^{\infty} 2^{n/2} e_{2^n}(f)$$

are equiconvergent. Moreover, since

$$2^{n/2} a_{2^{n+1}} \le e_{2^n}(f) \le \sum_{k=n}^{\infty} 2^{k/2} a_{2^k} \qquad (n \in \mathbf{N})$$

we have that the series

$$\sum_{n=0}^{\infty} 2^n a_{2^n} \quad \text{and} \quad \sum_{n=1}^{\infty} 2^{n/2} e_{2^n}(f)$$

are equiconvergent. Consequently, the series

$$\sum_{k=0}^{\infty} a_k \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} e_n(f)$$

are equiconvergent. ∎

This theorem contains an analogue of a result of Bernstein.

COROLLARY 5. *If* $f \in \mathrm{L}^2$ *and*

$$\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} \|f - S_n f\|_2 < \infty$$

then $f \in \mathrm{A}$.

2.5 Pointwise Convergence and Summability. Results which obtain convergence of a Walsh-Fourier series at a specific point are based on the following observation.

LEMMA 1. *If* $f \in \mathrm{L}^1$ *and* $x \in [0,1)$ *then*

$$\lim_{n\to\infty} \int_{[0,1)\setminus I_N(x)} f(t) D_n(x \dotplus t)\, dt = 0 \tag{26}$$

for every $N \in \mathbf{N}$.

PROOF. Fix $N \in \mathbf{N}$ and recall from the Paley lemma that the function $\tau_x D_{2^k}$ vanishes off $I_N(x)$ for all $k \geq N$. Hence it follows from Theorem 8 in 1.5 that

$$\int_{[0,1)\setminus I_N(x)} f(t) D_n(x \dotplus t)\, dt = \sum_{k=0}^{N-1} n_k \widehat{g}_k(n)$$

where $(n_k, k \in \mathbf{N})$ are the binary coefficients of n and

$$g_k := \chi([0,1) \setminus I_N(x)) f\, \tau_x(r_k D_{2^k}) \qquad (k \in \mathbf{N}).$$

But each $g_k \in \mathrm{L}^1$ so the Riemann-Lebesgue lemma implies that $\widehat{g}_k(n) \to 0$ as $n \to \infty$. Consequently, (26) is established. ∎

One important consequence of this observation is the following.

THEOREM 13. [THE LOCALIZATION PRINCIPLE] *If* $f, g \in \mathrm{L}^1$ *and* $f = g$ *on some open subset* V *of* $[0,1)$ *then* $(S_n f)(x)$ *and* $(S_n g)(x)$ *are equiconvergent for each* $x \in V$.

PROOF. Fix $x \in V$ and choose $N \in \mathbf{N}$ so large that $I_N(x) \subset V$. Then $f = g$ everywhere on $I_N(x)$ and we have

$$(S_n f)(x) - (S_n g)(x) = \int_{[0,1)\setminus I_N(x)} (f(t) - g(t)) D_n(x \dotplus t)\, dt \to 0$$

for all $n \in \mathbf{N}$. Hence the localization principle follows at once from Lemma 1. ∎

The next two results give sufficient conditions for convergence of a Walsh-Fourier at a given point.

THEOREM 14. [DINI'S TEST] *Let $x \in [0,1)$, $f \in L^1$, and define g for $t \in [0,1) \setminus \{x\}$ by*

$$g(t) := \frac{f(t) - f(x)}{x \dot{+} t}.$$

If g is integrable on $[0,1)$ then

$$(S_n f)(x) \to f(x) \quad \text{as} \quad n \to \infty.$$

PROOF. By Theorem 10 in 1.6,

$$|D_n(x \dot{+} t)| < \frac{2}{x \dot{+} t} \qquad (t \neq x, n \in \mathbf{N}).$$

Hence

$$|(S_n f)(x) - f(x)| \leq \left| \int_{[0,1)\setminus I_N(x)} (f(t) - f(x)) D_n(x \dot{+} t)\, dt \right| + 2 \int_{I_N(x)} |g(t)|\, dt$$

for $N \in \mathbf{N}$. Since $g \in L^1$, this last integral tends to zero as $N \to \infty$. We finish the proof by applying Lemma 1. ∎

It follows that if

$$|f(x \dot{+} t) - f(x)| = O\left((\log(\frac{1}{t}))^{-1-\varepsilon} \right) \quad \text{as} \quad t \to 0$$

for some $\varepsilon > 0$ then $(S_n f)(x) \to f(x)$ as $n \to \infty$.

THEOREM 15. [DIRICHLET'S TEST] *If f is finite- valued and of bounded variation on $[0,1)$ and if*

$$\lim_{t \to 0} |f(x \dot{+} t) - f(x)| = 0$$

for some $x \in [0,1)$ then $(S_n f)(x) \to f(x)$ as $n \to \infty$.

PROOF. We may suppose that f is monotone non-decreasing. Fix $x \in [0,1), N \in \mathbf{N}$, let α_N, β_N represent the endpoints of $I_N(x)$ and set

$$\Theta(\delta) := \sup_{0<t\leq\delta} |f(x \dot{+} t) - f(x)|.$$

By the mean value theorem for integrals, there is a point $\xi \in I_N(x)$ such that

$$\begin{aligned}(U_{n,N} f)(x) &:= \int_{I_N(x)} (f(t) - f(x)) D_n(x \dot{+} t)\, dt \\ &= (f(\alpha_N) - f(x)) \int_{\alpha_N}^{\xi} D_n(x \dot{+} t)\, dt + (f(\beta_N) - f(x)) \int_{\xi}^{\beta_N} D_n(x \dot{+} t)\, dt\end{aligned}$$

for $n \in \mathbf{N}$. Hence it follows that

$$|(U_{n,N}f)(x)| \le 4\Delta\Theta(2^{-N})$$

for N sufficiently large, where

$$\Delta := \sup_{x,y\in[0,1),n\in\mathbf{N}} |\int_0^y D_n(x \dot{+} t)\,dt|.$$

In view of the hypotheses and Lemma 1 it suffices to show that $\Delta \le 2$.

Toward this, fix $n \in \mathbf{N}, x, y \in [0,1)$, and choose integers $m \in \mathbf{N}$ and $0 \le \ell < 2^m$ such that $n = 2^m + \ell$. Then we can write

$$|\int_0^y D_n(x \dot{+} t)\,dt| \le |\int_0^y D_{2^m}(x \dot{+} t)\,dt| + |\int_0^y \left(\sum_{k=0}^{\ell-1} w_k(x)w_{2^m+k}(t)\right) dt|$$
$$=: I_1 + I_2.$$

Since $\|D_{2^m}\|_1 = 1$, it is clear that $I_1 \le 1$.

To estimate I_2, let $\boldsymbol{J}_j$ represent the indefinite integral of w_j and observe that

$$I_2 = |\sum_{k=0}^{\ell-1} w_k(x)\boldsymbol{J}_{2^m+k}(y)|.$$

Recall from (48) in 1.5 that

$$|\boldsymbol{J}_{2^m+k}(y)| \le 2^{-m} \qquad (0 \le k < \ell).$$

It follows that

$$I_2 \le \frac{\ell}{2^m} \le 1.$$

Therefore, $\Delta \le 2$ as promised. ∎

Thus the Walsh-Fourier series of a function f of bounded variation converges to f at every point of W-continuity. This result does not hold if W-continuity is relaxed (see Exercise 2.9).

The proof of Theorem 15 can be modified to show that if f belongs to C_W and is of bounded variation, then $S_n f$ converges to f uniformly on $[0,1)$ as $n \to \infty$ (see Exercise 2.8).

When dealing with Cesàro summability, the condition that f be of bounded variation can be dropped.

THEOREM 16. *Let $x \in [0,1)$ and suppose $f \in \mathrm{L}^1$ is W- continuous at x. Then*

$$\lim_{m\to\infty} (\sigma_m f)(x) = f(x).$$

PROOF. Recall from 1.8 that

$$(\sigma_m f)(x) = \int_0^1 f(t)K_m(x \dot{+} t)\,dt \qquad (m \in \mathbf{N})$$

and that

$$\|K_m\|_1 \le 2. \tag{27}$$

Thus

$$(\sigma_m f)(x) - f(x) = \int_{I_N(x)} (f(t) - f(x))K_m(x \dot{+} t)\,dt + \int_{[0,1)\setminus I_N(x)} (f(t) - f(x))K_m(x \dot{+} t)\,dt.$$

By Lemma 1 the second term tends to zero as $m \to \infty$ for any $N \in \mathbf{N}$. To estimate the first term, observe by (27) that

$$\begin{aligned} \Big| \int_{I_N(x)} (f(t) - f(x))K_m(x \dot{+} t)\,dt \Big| &\le \sup_{u \in I_N(x)} |f(u) - f(x)| \int_0^1 |K_m(x \dot{+} t)|\,dt \\ &\le 2 \sup_{u \in I_N(x)} |f(u) - f(x)|. \end{aligned}$$

Since f is W-continuous at x, this completes the proof of Theorem 16. ∎

The proof of Theorem 16 can be modified to show that the Walsh- Fourier series of a function in C_W is uniformly Cesàro summable on $[0, 1)$ (see Exercise 2.10).

EXERCISES

2.1 Suppose that $1 \le p < \infty$, that $f \in L^p$, and that

$$\sup_{n \in \mathbf{N}} \sum_{k=0}^{2^n - 1} \omega^{(p)}(f, I(k, n)) < \infty.$$

Prove that $\widehat{f}(k) = O(1/k)$ as $k \to \infty$.

2.2 Prove that (10) and (11) hold for any non-constant F which is absolutely continuous on [0,1].

2.3 Suppose g is continuous on [0,1] , $(a_n, n \in \mathbf{N}) \in \ell^1$, and that

$$|\widehat{g}(k)| \le \frac{a_n}{2^n}$$

for $0 \le k < 2^n, n \in \mathbf{N}$. Prove that g is constant on [0,1].

2.4 Let E be the Cantor set formed by removing successive middle halves from the interval [0,1]. Identify the corresponding intervals $(I_n, n \in \mathbf{N})$ determined by (15), and the Walsh- Fourier coefficients of the associated Cantor-Lebesgue function.

2.5 Let E be any Cantor set of the form (15) and let F be the associated Cantor-Lebesgue function. Prove that F is continuous, monotone increasing, $F' = 0$ a.e., and $F(0) = 0, F(1) = 1$.

2.6 Suppose that $f \in \mathrm{L}^p, 1 \le p < 2$, and that the L^p local modulus of continuity of f satisfies

$$\sum_{n=0}^{\infty} \left(2^{-n} \sum_{k=0}^{2^n-1} |\omega^{(p)}(f, I(k,n))|^p \right)^{1/p} < \infty.$$

Show that $f \in \mathrm{A}$.

2.7. i) Show that if $f \in \mathrm{L}^p, 1 \le p < 2$, and

$$\sum_{n=0}^{\infty} 2^{n/p} \omega^{(p)}(f, 2^{-n}) < \infty$$

then $f \in \mathrm{A}$.

ii) Show that if $f \in \mathrm{L}^p, 1 \le p < 2$, and $f \in \mathrm{Lip}(\alpha, \mathrm{L}^p)$ for some $\alpha > 1/p$ then $f \in \mathrm{A}$.

iii) Show that if $f \in \mathrm{L}^p, 1 \le p < 2$, and $f \in \mathrm{Lip}(\alpha, \mathrm{L}^p)$ for some $0 < \alpha \le 1$ then

$$\sum_{k=1}^{\infty} k^{-\beta} |\widehat{f}(k)| < \infty$$

for any $\beta > 1/p - \alpha$.

2.8 Prove that $S_n f$ converges to f uniformly as $n \to \infty$ for every function $f \in \mathrm{C}_W$ which is of bounded variation.

2.9 Let $f := \chi([0, 1/3))$. Show that $(S_n f)(1/3)$ does not converge as $n \to \infty$.

2.10 Prove that $\sigma_n f$ converges to f uniformly as $n \to \infty$ for every function $f \in \mathrm{C}_W$.

2.11 Use the Riesz-Fischer theorem to show that $f \in \mathrm{A}$ if and only if $f = g * h$ for some $g, h \in \mathrm{L}^2(\mathbf{G})$.

2.12 Suppose f is W-continuous at 0 and $\widehat{f}(k) \ge 0$ for $k \in \mathbf{N}$. Prove $f \in \mathrm{A}$.

2.13 Suppose f is twice continuously differentiable in the classical sense. Prove $f \in \mathrm{A}$.

Chapter 3

DYADIC MARTINGALES AND HARDY SPACES

3.1 Dyadic Martingales and the Dyadic Maximal Function. For each $n \in \mathbf{N}$ let $\mathcal{A}^n$ represent the σ-algebra generated by the collection of dyadic intervals $\{I(k,n) : k = 0, 1, \ldots, 2^n - 1\}$. Thus every element of $\mathcal{A}^n$ is a finite union of intervals of the form $[k/2^n, (k+1)/2^n)$.

Let $\mathrm{L}(\mathcal{A}^n)$ represent the collection of $\mathcal{A}^n$-measurable functions on $[0,1)$. By the Paley lemma, $\mathrm{L}(\mathcal{A}^n)$ coincides with the collection of Walsh polynomials of order less than 2^n. In particular, $\mathcal{P} = \bigcup_{n \in \mathbf{N}} \mathrm{L}(\mathcal{A}^n)$.

A sequence of functions $(f_n, n \in \mathbf{N})$ is called a *dyadic martingale* if each f_n belongs to $\mathrm{L}(\mathcal{A}^n)$ and

$$\int_E f_{n+1} = \int_E f_n \qquad (E \in \mathcal{A}^n, n \in \mathbf{N}).$$

Since $\int_E w_j = 0$ for all $E \in \mathcal{A}^n$, $j \geq 2^n$ and $n \in \mathbf{N}$, it is clear that the 2^n-th partial sums of any Walsh series is a dyadic martingale. Conversely, it is easy to see that every dyadic martingale can be obtained in this way. Thus investigation of 2^n-th partial sums of Walsh series leads to the study of dyadic martingales. Notice also that the map $\nu \rightarrow (S_{2^n}\nu, n \in \mathbf{N})$ is an isomorphism from QM onto the collection of dyadic martingales. Since we are only concerned with dyadic martingales, the adjective dyadic will be dropped from now on.

For each $f \in \mathrm{L}^1$ the *conditional expectation* of f given $\mathcal{A}^n$ is defined by

$$\mathcal{E}_n f := S_{2^n} f \qquad (n \in \mathbf{N}).$$

For convenience, we set $\mathcal{E}_{-1} f := 0$ and $\mathcal{A}^{-1} := \mathcal{A}^0$. It is clear that $(f_n, n \in \mathbf{N})$ is a dyadic martingale if and only if $f \in \mathrm{L}(\mathcal{A}_n)$ and

(1) $$\mathcal{E}_n(f_{n+1}) = f_n \qquad (n \in \mathbf{N}).$$

A martingale $(f_n, n \in \mathbf{N})$ will be called *regular* if there is a function $f \in \mathrm{L}^1$ such that $f_n = \mathcal{E}_n f$ for all $n \in \mathbf{N}$.

The operators $\mathcal{E}_n : \mathrm{L}^1 \rightarrow \mathrm{L}(\mathcal{A}^n)$ satisfy six important properties. Each is evidently linear. Since the Walsh system is orthogonal, it is easy to see that

(2) $$\mathcal{E}_n \circ \mathcal{E}_m = \mathcal{E}_{\min\{m,n\}} \qquad (m, n \in \mathbf{N}).$$

By Parseval's identity we have

(3) $$\int_0^1 (\mathcal{E}_n f) g = \int_0^1 f(\mathcal{E}_n g) = \int_0^1 (\mathcal{E}_n f)(\mathcal{E}_n g)$$

for all $f, g \in \mathrm{L}^1$ and $n \in \mathbf{N}$. And, since

$$(S_{2^n} f)(x) = \frac{1}{|I_n(x)|} \int_{I_n(x)} f \qquad (x \in [0,1)),$$

we also have

$$\|\mathcal{E}_n f\|_p \leq \|f\|_p \qquad (f \in \mathrm{L}^p, 1 \leq p \leq \infty, n \in \mathbf{N}), \tag{4}$$

$$|\mathcal{E}_n f| \leq \mathcal{E}_n |f| \qquad (f \in \mathrm{L}^1, n \in \mathbf{N}), \tag{5}$$

and

$$\mathcal{E}_n(\lambda f) = \lambda \mathcal{E}_n f \qquad (f \in \mathrm{L}^1, \lambda \in \mathrm{L}(\mathcal{A}^n), n \in \mathbf{N}). \tag{6}$$

Since the operators $\mathcal{E}_n$ are averages they inherit certain inequalities from the general theory of integrals. We shall have occasion to use the following two. The conditional Hölder inequality is

$$|\mathcal{E}_n(fg)| \leq (\mathcal{E}_n |f|^p)^{1/p} (\mathcal{E}_n |g|^{p'})^{1/p'}$$

which holds for all $f \in \mathrm{L}^p, g \in \mathrm{L}^{p'}, n \in \mathbf{N}$ and conjugate exponents p and p'. The conditional Jensen inequality is

$$\psi \circ \mathcal{E}_n |f| \leq \mathcal{E}_n(\psi \circ |f|)$$

which holds for all $f \in \mathrm{L}^0, n \in \mathbf{N}$ and ψ convex on $[0, \infty)$.

Given a real number y and a function g measurable on $[0,1)$, we shall usually denote the set $\{x \in [0,1) : g(x) > y\}$ by $\{g > y\}$. Notice by Fubini's theorem that

$$\begin{aligned} \|g\|_p^p &= p \int_0^1 \int_0^{|g(x)|} y^{p-1}\, dy\, dx \\ &= p \int_0^\infty y^{p-1} |\{|g| > y\}|\, dy \end{aligned}$$

holds for any measurable g and $p > 0$. We shall use this identity many times.

Let T be a map from L^1 into the collection of functions measurable on $[0,1)$ which is *sublinear*, i.e.,

$$|T(af)| = |a||Tf|, \qquad |T(f+g)| \leq |Tf| + |Tg|$$

hold a.e. on $[0,1)$ for $f, g \in \mathrm{L}^1$ and $a \in \mathbf{R}$. Let $1 \leq p \leq \infty$ and $1 \leq q \leq \infty$. The map T is said to be of *type* (p,q) if there is a constant $C > 0$ depending only on p and q such that

$$\|Tf\|_q \leq C \|f\|_p$$

for all $f \in L^p$. The map T is said to be of *weak type* (p,q) if $0 < q < \infty$ and there is a constant $C > 0$ depending only on p and q such that

$$|\{Tf > y\}| \le \left(\frac{C\|f\|_p}{y}\right)^q$$

for all $f \in L^p$ and $y > 0$. Since

$$\|Tf\|_p^p = p\int_0^\infty y^{p-1}|\{|Tf| > y\}|\,dy$$

it is clear that if T is of type (p,q) then T is of weak type (p,q). Notice also that if T is of weak type (p,q) for some p,q then Tf is a.e. finite for all $f \in L^p$.

Recall that the positive part of a function h is defined by $h^+ := \max\{h, 0\}$. We shall denote the positive part of the natural logarithm function by $\log^+ y$. The collection of $f \in L^0$ which satisfy

$$\int_0^1 |f|\log^+|f| < \infty$$

will be denoted by $L\log^+ L$.

Any sequence of sublinear maps which enjoys properties (4) and (6) above, gives rise to a maximal operator of weak type (1,1) and of type (p,p), $1 < p < \infty$ which also takes L^1 into $L\log^+ L$. In fact,

THEOREM 1. *Let $T_n : L^1 \to L(\mathcal{A}^n)$, $n \in \mathbf{N}$, be a sequence of sublinear maps which satisfy*

$$|T_n(\lambda f)| = |\lambda T_n f| \qquad (\lambda \in L(\mathcal{A}^n)),$$

and

$$\|T_n f\|_1 \le C\|f\|_1 \qquad (n \in \mathbf{N}, f \in L^1),$$

where C is an absolute constant independent of n and f. Define

$$T^*f := \sup_{n\in\mathbf{N}} |T_n f| \qquad (f \in L^1).$$

Then

$$y|\{T^*f > y\}| \le C\int_{\{T^*f>y\}} |f|$$

for all $f \in L^1$ and $y > 0$.

Suppose ψ is non-decreasing and $\psi(x)/x$ is locally integrable on $[0,\infty)$. For $y \ge 0$, define

$$\Phi(y) := \int_0^y \psi(t)\,dt \quad \text{and} \quad \Psi(y) := \int_0^y \frac{\psi(t)}{t}\,dt.$$

Then

$$\int_0^1 \Phi\circ(T^*f) \le C\int_0^1 |f|\Psi\circ(T^*f).$$

If $1 < p < \infty$ *then*

$$\|T^*f\|_p \le \frac{Cp}{p-1}\|f\|_p$$

holds for every $f \in L^p$.

If $C < e$ *then*

$$\|T^*f\|_1 \le \frac{Ce}{e-C}\int_0^1 |f|\log^+|f| + \frac{e}{e-C}$$

for all $f \in L\log^+ L$. *(Here, e denotes the base of the natural logarithm* log.*)*

PROOF. Fix $y > 0, f \in L^1$, set $T^*_{-1}f := 0$ and

$$T^*_n f := \sup_{0\le k\le n} |T_k f| \qquad (n \in \mathbf{N}).$$

For each $n \in \mathbf{N}$ set

$$E_n := \{T^*_n f > y \quad \text{and} \quad T^*_{n-1}f \le y\}.$$

Clearly, $E_n \in \mathcal{A}^n, E_n \subset \{|T_n f| > y\}, E_n \cap E_m = \emptyset$ for $n \ne m \quad (n, m \in \mathbf{N})$, and

$$\{T^*f > y\} = \bigcup_{n\in\mathbf{N}} E_n.$$

Consequently, we have

$$\begin{aligned} y|\{T^*f > y\}| &= y\sum_{n\in\mathbf{N}} |E_n| \\ &\le \sum_{n\in\mathbf{N}} \int_{E_n} |T_n f| \\ &= \sum_{n\in\mathbf{N}} \int_0^1 \chi(E_n)|T_n f|. \end{aligned}$$

Since $\chi(E_n) \in L(\mathcal{A}^n)$ it follows from hypothesis that

$$\begin{aligned} \int_0^1 \chi(E_n)|T_n f| &= \int_0^1 |T_n(\chi(E_n)f)| \\ &\le C\int_0^1 \chi(E_n)|f|. \end{aligned}$$

Consequently,

$$\begin{aligned} y|\{T^*f > y\}| &\le C\sum_{n\in\mathbf{N}} \int_0^1 \chi(E_n)|f| \\ &= C\int_{\{T^*f>y\}} |f|. \end{aligned}$$

To verify the estimate involving Ψ and Φ notice that we have proved

$$\int_0^1 \chi(\{T^*f > y\}) \leq \frac{C}{y} \int_0^1 \chi(\{T^*f > y\})|f|$$

for all $y > 0$. Multiply this inequality by $\psi(y)$ and integrate with respect to y over $[0, \infty)$. We have by Fubini's theorem that

$$\begin{aligned} \int_0^1 \Phi \circ (T^*f) &= \int_0^1 \left(\int_0^{(T^*f)(x)} \psi(y)\, dy \right) dx \\ &\leq C \int_0^1 |f(x)| \left(\int_0^{(T^*f)(x)} \frac{\psi(y)}{y}\, dy \right) dx \\ &= C \int_0^1 |f| \Psi \circ (T^*f). \end{aligned}$$

To verify the L^p inequalities, fix $1 < p < \infty$ and specialize to the case $\psi(y) = py^{p-1}$. In this case, $\Phi(y) = y^p$ and $\Psi(y) = p/(p-1)y^{p-1}$. Thus by the general inequality above together with Hölder's inequality, we have

$$\begin{aligned} \int_0^1 |T^*f|^p &\leq \frac{Cp}{p-1} \int_0^1 |f||T^*f|^{p-1} \\ &\leq \frac{Cp}{p-1} \|f\|_p \|T^*f\|_p^{p-1}. \end{aligned}$$

Finally, let

$$\psi(y) := \begin{cases} 0 & 0 \leq y \leq 1 \\ 1 & 1 < y. \end{cases}$$

Then $\Phi(y) = (y-1)^+$ and $\Psi(y) = \log^+ y$. Hence by the general inequality above we have

$$\int_0^1 (T^*f - 1)^+ \leq C \int_0^1 |f| \log^+(T^*f).$$

But $\log^+$ is concave on $[1, \infty)$. Thus

$$\log^+ u \leq u/e$$

for $u > 0$ and

$$a \log^+ b = a \log^+ (\frac{b}{a} a) \leq \frac{b}{e} + a \log^+ a$$

for all $a, b > 0$. Applying this to $a = |f|$ and $b = T^*f$, we conclude that

$$\int_0^1 T^*f \leq C \int_0^1 |f| \log^+ |f| + \frac{C}{e} \int_0^1 T^*f + 1. \qquad \blacksquare$$

We shall apply Theorem 1 to three maximal operators $\mathcal{E}^*, \widetilde{\mathcal{E}}$, and $\mathcal{E}^\natural$.

First, define the ***dyadic maximal operator*** by

$$\mathcal{E}^* f := \sup_{n \in \mathbf{N}} |\mathcal{E}_n f|$$

for $f \in \mathrm{L}^1$.

Next, define $\widetilde{\mathcal{E}}$ as follows. Set

$$\widetilde{\mathcal{E}}_0 f := |\mathcal{E}_0 f| \qquad (f \in \mathrm{L}^1).$$

For $n \in \mathbf{P}$ and $f \in \mathrm{L}^1$ set

$$\widetilde{\mathcal{E}}_n f := \sup_{0 \le m \le n} \left(|\mathcal{E}_m f| + |\mathcal{E}_m(r_m f)|\right).$$

Define

$$\widetilde{\mathcal{E}} f := \sup_{n \in \mathbf{N}} \widetilde{\mathcal{E}}_n f.$$

Notice that each $\widetilde{\mathcal{E}}_n$ is sublinear and that $\widetilde{\mathcal{E}}_n f$ is $\mathcal{A}^n$ measurable for all $n \in \mathbf{N}$ and $f \in \mathrm{L}^1$. Also notice that

$$\mathcal{E}^* f \le \widetilde{\mathcal{E}} f \le 3\mathcal{E}^* f \qquad (f \in \mathrm{L}^1). \tag{7}$$

Indeed, the left inequality of (7) is obvious. To prove the right inequality observe first that

$$D_{2^m} = D_{2^{m-1}} + r_{m-1} D_{2^{m-1}}.$$

Next, set

$$\mathcal{E}_n^* f := \sup_{0 \le k \le n} |\mathcal{E}_k f|$$

and verify that

$$\mathcal{E}_m f = \mathcal{E}_{m-1} f + r_{m-1} \mathcal{E}_{m-1}(r_{m-1} f)$$

and

$$|\mathcal{E}_{m-1}(r_{m-1} f)| \le |\mathcal{E}_m f| + |\mathcal{E}_{m-1} f| \le 2\mathcal{E}_m^* f$$

both hold for $m \in \mathbf{P}$ and $f \in \mathrm{L}^1$. This establishes (7).

Finally, define

$$\mathcal{E}^\natural f := \sup_{n \in \mathbf{N}} \mathcal{E}_n^\natural f, \tag{8}$$

where

$$\mathcal{E}_n^\natural f := \sup_{k \in \mathbf{N}} |\mathcal{E}_n(f w_k)| \qquad (n \in \mathbf{N}, f \in \mathrm{L}^1).$$

Notice once and for all that

$$|\mathcal{E}^\natural f| \le \mathcal{E}^*(|f|)$$

for $f \in \mathrm{L}^1$.

It is immediate from Theorem 1 that the following is true.

COROLLARY 1. *If $f \in L^1$ and $y > 0$ then*

$$y|\{\mathcal{E}^* f > y\}| \leq \int_{\{\mathcal{E}^* f > y\}} |f|,$$

$$y|\{\widetilde{\mathcal{E}} f > y\}| \leq 2 \int_{\{\widetilde{\mathcal{E}} f > y\}} |f|,$$

and

$$y|\{\mathcal{E}^\natural f > y\}| \leq \int_{\{\mathcal{E}^\natural f > y\}} |f|. \tag{9}$$

If $1 < p < \infty$ then

$$\|f\|_p \leq \|\mathcal{E}^* f\|_p \leq \frac{p}{p-1}\|f\|_p,$$

$$\|\widetilde{\mathcal{E}} f\|_p \leq \frac{2p}{p-1}\|f\|_p, \tag{10}$$

and

$$\|\mathcal{E}^\natural f\|_p \leq \frac{p}{p-1}\|f\|_p \tag{11}$$

for all $f \in L^p$.

If $f \in L \log^+ L$ then

$$\|\mathcal{E}^* f\|_1 \leq \frac{e}{e-1}\int_0^1 |f| \log^+ |f| + \frac{e}{e-1},$$

$$\|\widetilde{\mathcal{E}} f\|_1 \leq \frac{2e}{e-2}\int_0^1 |f| \log^+ |f| + \frac{e}{e-2},$$

and

$$\|\mathcal{E}^\natural f\|_1 \leq \frac{e}{e-1}\int_0^1 |f| \log^+ |f| + \frac{e}{e-1}.$$

Thus $\mathcal{E}^* f, \widetilde{\mathcal{E}} f$, and $\mathcal{E}^\natural f$ are a.e. finite for all $f \in L^1$.

There is a close connection between a.e. convergence of the sequences $T_n f$ and weak type inequalities for the corresponding maximal operators $T^* f$.

THEOREM 2. *Suppose $(\boldsymbol{X}, \|\cdot\|)$ is a normed linear space in L^0 and $\boldsymbol{X}_0$ is dense in $\boldsymbol{X}$. Let T, T_n $(n \in \mathbf{N})$, be linear maps from $\boldsymbol{X}$ to L^p for some $1 \leq p < \infty$ with T bounded and $T_n f \to T f$ a.e. on $[0,1)$, as $n \to \infty$, for each $f \in \boldsymbol{X}_0$. Set*

$$T^* f := \sup_{n \in \mathbf{N}} |T_n f| \qquad (f \in \boldsymbol{X}).$$

If there is a constant $C > 0$, independent of f and n, such that

$$y^p|\{T^*f > y\}| \leq C\|f\|^p$$

for all $y > 0$ and $f \in \boldsymbol{X}$ then

$$Tf = \lim_{n\to\infty} T_n f$$

a.e. on $[0,1)$ for every $f \in \boldsymbol{X}$.

PROOF. Fix $f \in \boldsymbol{X}$ and set

$$\boldsymbol{\Delta} := \limsup_{n\to\infty} |T_n f - Tf|.$$

It suffices to show that $\boldsymbol{\Delta} = 0$ a.e. on [0,1).

Choose $f_m \in \boldsymbol{X}_0$ ($m \in \mathbf{N}$) such that $\|f - f_m\| \to 0$ as $m \to \infty$. Observe by hypothesis that

$$\begin{aligned}\boldsymbol{\Delta} &\leq \limsup_{n\to\infty} |T_n(f - f_m)| + \limsup_{n\to\infty} |T_n(f_m) - T(f_m)| + |T(f - f_m)| \\ &\leq T^*(f - f_m) + |T(f - f_m)|\end{aligned}$$

a.e. for all $m \in \mathbf{N}$. In particular,

$$\begin{aligned}|\{\boldsymbol{\Delta} > 2y\}| &\leq |\{T^*(f - f_m) > y\}| + |\{|T(f - f_m)| > y\}| \\ &\leq \frac{C\|f - f_m\|^p}{y^p} + \|T\| \frac{\|f - f_m\|^p}{y^p}\end{aligned}$$

holds for any $y > 0$ and $m \in \mathbf{N}$. Since $f_m \to f$ in $\boldsymbol{X}$ as $m \to \infty$ it follows that

$$|\{\boldsymbol{\Delta} > 2y\}| = 0$$

for all $y > 0$. We conclude that $\boldsymbol{\Delta} = 0$ a.e. on [0,1). ∎

Applying this result to $\boldsymbol{X} = \mathrm{L}^1, T_n = \mathcal{E}_n, \boldsymbol{X}_0 = \mathcal{P}$, and $Tf = f$ we have another proof that $S_{2^n} f \to f$ a.e. as $n \to \infty$ for all $f \in \mathrm{L}^1$.

Let $1 \leq p < \infty$. Applying this result to $\boldsymbol{X} = \mathrm{L}^p, T_n = S_n, \boldsymbol{X}_0 = \mathcal{P}$, and $Tf = f$ we see that a sufficient condition for a.e. convergence of Walsh-Fourier series of L^p functions is that the maximal operator

$$S^*f := \sup_{n\in\mathbf{P}} |S_n f|$$

be of weak type (p,p). This condition is also necessary for the case $1 \leq p \leq 2$ (see Theorem 23 in 6.6). In fact, for $1 \leq p \leq 2$, the converse of Theorem 2 holds for any linear operators T_n which commute with the translation operators (see Stein [1]). Thus Sf converges a.e. on [0,1) for all $f \in \mathrm{L}^p$ for $1 \leq p \leq 2$ if and only if S^* is of weak type (p,p).

We shall prove that S^* is of weak type (p,p) for $1 < p < \infty$ and show it to be false for $p = 1$ (see Theorem 14 in 3.7 and Theorem 12 in 4.5). One step toward this is the following technical consequence of Corollary 1.

COROLLARY 2. *Let $(f_k, k \in \mathbf{N})$ be a sequence of non-negative integrable functions defined on $[0,1]$ and $(n_k, k \in \mathbf{N})$ be a sequence in $\mathbf{N}$. If $1 \le p < \infty$ then*

$$\int_0^1 \left(\sum_{k=0}^{\infty} \mathcal{E}_{n_k} f_k\right)^p \le p^p \int_0^1 (\sum_{k=0}^{\infty} f_k)^p$$

and

$$\int_0^1 (\sum_{k=0}^{\infty} f_k)^{1/p} \le 4 \int_0^1 \left(\sum_{k=0}^{\infty} \mathcal{E}_{n_k} f_k\right)^{1/p}.$$

PROOF. These inequalities are obvious for $p = 1$.

If $1 < p < \infty$ and p' is the index conjugate to p, then the usual inner product $\langle F, G\rangle := \int_0^1 FG$ and the Riesz representation theorem can be used to write

$$\|\sum_{k=0}^{\infty} \mathcal{E}_{n_k} f_k\|_p = \sup_{\|g\|_{p'} \le 1} \left|\left\langle \sum_{k=0}^{\infty} \mathcal{E}_{n_k} f_k, g\right\rangle\right|.$$

But (3), Hölder's inequality, and Corollary 1 imply

$$\begin{aligned}
\left|\left\langle \sum_{k=0}^{\infty} \mathcal{E}_{n_k} f_k, g\right\rangle\right| &\le \sum_{k=0}^{\infty} |\langle f_k, \mathcal{E}_{n_k} g\rangle| \\
&\le \left\langle \sum_{k=0}^{\infty} f_k, \mathcal{E}^* g\right\rangle \\
&\le \|\sum_{k=0}^{\infty} f_k\|_p \|\mathcal{E}^* g\|_{p'} \\
&\le p\|\sum_{k=0}^{\infty} f_k\|_p.
\end{aligned}$$

The first inequality is established.

To establish the second one, let r be the index conjugate to $2p$ and set $g_k := (f_k)^{1/(2p)}$ for $k \in \mathbf{N}$ and $\boldsymbol{g} := (g_k, k \in \mathbf{N})$. For any sequence $\boldsymbol{\lambda} = (\lambda_k, k \in \mathbf{N})$ of measurable functions set

$$\|\boldsymbol{\lambda}\|_{L^2(\ell^r)} := \|(\sum_{k=0}^{\infty} |\lambda_k|^r)^{1/r}\|_2.$$

With this notation, it is easy to see that

$$\begin{aligned}
\left(\int_0^1 (\sum_{k=0}^{\infty} f_k)^{1/p}\right)^{1/2} &= \|\boldsymbol{g}\|_{L^2(\ell^{2p})} \\
&= \sup_{\|\boldsymbol{\lambda}\|_{L^2(\ell^r)} \le 1} |\int_0^1 (\sum_{k=0}^{\infty} g_k \lambda_k)|.
\end{aligned}$$

Moreover, it follows from Hölder's inequality that

$$\begin{aligned}
|\int_0^1 (\sum_{k=0}^{\infty} g_k\lambda_k)| &\le \sum_{k=0}^{\infty}\int_0^1 \mathcal{E}_{n_k}(|g_k\lambda_k|) \\
&\le \sum_{k=0}^{\infty}\int_0^1 (\mathcal{E}_{n_k} g_k^{2p})^{1/(2p)}(\mathcal{E}_{n_k}|\lambda_k|^r)^{1/r} \\
&= \sum_{k=0}^{\infty}\int_0^1 (\mathcal{E}_{n_k} f_k)^{1/(2p)}(\mathcal{E}_{n_k}|\lambda_k|^r)^{1/r} \\
&\le \int_0^1 (\sum_{k=0}^{\infty}\mathcal{E}_{n_k} f_k)^{1/(2p)}(\sum_{k=0}^{\infty}\mathcal{E}_{n_k}|\lambda_k|^r)^{1/r} \\
&\le \left(\int_0^1 (\sum_{k=0}^{\infty}\mathcal{E}_{n_k} f_k)^{1/p}\int_0^1 (\sum_{k=0}^{\infty}\mathcal{E}_{n_k}|\lambda_k|^r)^{2/r}\right)^{1/2}.
\end{aligned}$$

Since $1 < 2/r = 2 - 1/p < 2$ it follows from the first inequality that

$$\begin{aligned}
\int_0^1 (\sum_{k=0}^{\infty}\mathcal{E}_{n_k}|\lambda_k|^r)^{2/r} &\le \left(\frac{2}{r}\right)^{2/r}\int_0^1 (\sum_{k=0}^{\infty}|\lambda_k|^r)^{2/r} \\
&\le \|\boldsymbol{\lambda}\|^2_{L^2(\ell^r)} \\
&\le 4.
\end{aligned}$$

This completes the proof of the second inequality. ∎

3.2 An Interpolation Theorem and the Canonical Decomposition. We begin this section with an interpolation result which will be used in 3.3 and 3.7.

LEMMA 1. *Let $F, G : [0,1] \to [0,\infty)$ be measurable functions which satisfy*

$$|\{G > \beta y, F \le \delta y\}| \le \varepsilon |\{G > y\}| \qquad (y > 0), \tag{12}$$

for some $\beta > 1, \delta > 0$, and $\varepsilon > 0$. Suppose further that Φ is a non-negative, continuous, non-decreasing function on $[0,\infty)$ which satisfies $\Phi(0) = 0$. If there exist real numbers $\gamma > 0, \eta > 0$ which satisfy

$$\Phi(\beta y) \le \gamma\Phi(y), \quad \Phi(\frac{y}{\delta}) \le \eta\Phi(y) \qquad (y > 0), \tag{13}$$

and if $\gamma\varepsilon < 1$, then

$$\int_0^1 \Phi\circ G \le \frac{\gamma\eta}{1-\gamma\varepsilon}\int_0^1 \Phi\circ F. \tag{14}$$

PROOF. Consider the Lebesgue-Stieltjes measure $d\Phi$ determined by

$$\int_{[a,b)} d\Phi := \Phi(b) - \Phi(a) \qquad (0 \le a < b < \infty).$$

Clearly, $d\Phi$ is a non-negative, σ-finite Borel measure on $[0,\infty)$. Moreover, if $f : [0,1] \to [0,\infty)$ is a measurable function then

$$\Phi(f(x)) = \int_0^{f(x)} d\Phi = \int_0^\infty \chi\{f > y\}(x)\, d\Phi(y)$$

holds for any $x \in [0,1]$. Hence Fubini's theorem implies

$$\int_0^1 \Phi \circ f = \int_0^\infty |\{f > y\}|\, d\Phi(y). \tag{15}$$

To verify (14), write $G = \beta G/\beta$ and use (13) to obtain

$$\int_0^1 \Phi \circ G \leq \gamma \int_0^1 \Phi \circ (\frac{G}{\beta}).$$

Thus (15) implies

$$\int_0^1 \Phi \circ G \leq \gamma \int_0^\infty |\{\frac{G}{\beta} > y\}|\, d\Phi(y). \tag{16}$$

But by (12) we have

$$\begin{aligned} |\{\frac{G}{\beta} > y\}| &\leq |\{G > \beta y, F \leq \delta y\}| + |\{G > \beta y, F > \delta y\}| \\ &\leq \varepsilon|\{G > y\}| + |\{F > \delta y\}|. \end{aligned}$$

Therefore, it follows from (16) and (15) that

$$\int_0^1 \Phi \circ G \leq \gamma\varepsilon \int_0^1 \Phi \circ G + \gamma \int_0^1 \Phi \circ (\frac{F}{\delta}).$$

In particular, (14) follows from (13) and the fact that $\gamma\varepsilon < 1$. ∎

Notice that hypothesis (13) is satisfied by all Φ which enjoy the inequality

$$\Phi(2y) \leq C_0\Phi(y) \qquad (y > 0) \tag{17}$$

for some constant C_0 independent of y. Therefore, Lemma 1 applies to the functions y^p for any $0 < p < \infty$ and to $y \log y$ for $y \geq 0$. (It does not apply to e^y for $y \geq 0$).

Conditions like hypothesis (12) arises naturally from certain types of decompositions.

LEMMA 2. *Suppose that*

$$G := \sum_{k\in\mathbf{Z}} \chi\{\lambda^* > 2^k\} g_{2^k}$$

for some non-negative functions $\lambda^*, g_{2^k} \in \mathrm{L}^0$, $k \in \mathbf{Z}$. *If* $0 < p < q$ *and there is a number* $B > 0$ *such that*

$$|\{g_{2^k} > t2^k\}| \leq Bt^{-q}2^{-pk}$$

holds for all $k \in \mathbf{Z}$ *and* $t > 0$, *then*

$$|\{G > y, \lambda^* \le y\}| \le BC(p,q)y^{-p},$$

where

$$C(p,q) := \frac{2^q}{(2^s - 1)^q(2^{qs} - 1)}$$

and $s := (q-p)/(2q)$.

PROOF. Fix $y > 0$ and choose $k_0 \in \mathbf{Z}$ such that

$$2^{k_0 - 1} < y \le 2^{k_0}.$$

Since $\chi\{\lambda^* \le y\}\chi\{\lambda^* > 2^k\} = 0$ for all $k \ge k_0$, it is clear that

$$\chi\{\lambda^* \le y\}G \le \sum_{k<0} g_{2^{k+k_0}}.$$

Moreover, since

$$(2^s - 1)\sum_{k<0} 2^{sk} = 1$$

for any $s > 0$, it is easy to see that

$$\begin{aligned}|\{G > y, \lambda^* \le y\}| &\le \sum_{k<0} |\{g_{2^{k+k_0}} > (2^s-1)2^{sk}2^{k_0-1}\}| \\ &= \sum_{k<0} |\{g_{2^{k+k_0}} > \left(\frac{2^s-1}{2}\right) 2^{k(s-1)}2^{k+k_0}\}|.\end{aligned}$$

It follows from hypothesis, therefore, that

$$|\{G > y, \lambda^* \le y\}| \le B2^q(2^s-1)^{-q}2^{-pk_0}\sum_{k<0} 2^{k(q-qs-p)}.$$

Since the choice of s implies $q - qs - p = qs$, we conclude from the choice of k_0 that

$$\begin{aligned}|\{G > y, \lambda^* \le y\}| &\le B(2^{qs}-1)C(p,q)y^{-p}\sum_{k<0} 2^{kqs} \\ &= BC(p,q)y^{-p}. \qquad \blacksquare\end{aligned}$$

For the next few pages, let $\overline{\mathbf{N}} := \mathbf{N} \cup \{\infty\}$. A function $\varsigma : [0,1) \to \overline{\mathbf{N}}$ is called a *dyadic stopping time* if

(18) $$\{\varsigma = n\} := \{x \in [0,1) : \varsigma(x) = n\} \in \mathcal{A}^n$$

for all $n \in \mathbf{N}$. We shall use dyadic stopping times to construct decompositions which satisfy the hypotheses of Lemma 2.

Notice that (18) is equivalent to the condition $\{\varsigma \le n\} \in \mathcal{A}^n$. Thus if ς and $\bar{\varsigma}$ are dyadic stopping times then so are

$$\varsigma \vee \bar{\varsigma} := \max\{\varsigma, \bar{\varsigma}\}$$

and

$$\varsigma \wedge \bar{\varsigma} := \min\{\varsigma, \bar{\varsigma}\}.$$

In particular, the collection of dyadic stopping times forms a net.

We shall call a sequence $(\lambda_n, n \in \mathbf{N})$ *adapted* if each λ_n is $\mathcal{A}^n$ measurable. Such a sequence generates a dyadic stopping time for each $y \in \mathbf{R}$ by means of

$$\tau(x) := \min\{n \in \mathbf{N} : \lambda_n(x) > y\} \qquad (x \in [0,1)). \tag{19}$$

Indeed, for each $n \in \mathbf{N}$ we have

$$\{\tau = n\} = \{\lambda_n > y\} \bigcap \left(\bigcup_{k=0}^{n-1} \{\lambda_k \le y\} \right). \tag{20}$$

(In this section, we have adopted the convention that min $\emptyset := \infty$.)

Let $(\lambda_n, n \in \mathbf{N})$ be adapted and suppose further that

$$\lambda_{-1} := 0 \le \lambda_0 \le \lambda_1 \le \dots.$$

Then it is easy to see that the dyadic stopping time τ defined by (19) satisfies

$$\begin{cases} \{\tau > n\} = \{\lambda_n \le y\}, \\ \{\tau \le n\} = \{\lambda_n > y\}, \\ \{\tau < \infty\} = \{\lambda^* > y\} \end{cases} \tag{21}$$

for $n \in \mathbf{N}, y \in \mathbf{R}$ where $\lambda^* := \lim_{n\to\infty} \lambda_n$.

We shall use martingales to generate our decompositions. Suppose for a moment that $(f_n, n \in \mathbf{N})$ is an a.e. convergent sequence of measurable functions defined on [0,1]. Set $f_{-1} := 0$,

$$f := f_\infty := \lim_{n\to\infty} f_n$$

and

$$d_n := f_{n+1} - f_n \qquad (n \ge -1).$$

For each dyadic stopping ς set

$$f_\varsigma := \sum_{n \in \overline{\mathbf{N}}} \chi\{\varsigma = n\} f_n.$$

Clearly, $f_\varsigma(x) = f_{\varsigma(x)}(x)$ everywhere on the set $\{\varsigma < \infty\}$ and a.e. on the set $\{\varsigma = \infty\}$. Moreover, it is easy to see that

$$\sum_{k=-1}^{n-1} \chi\{\varsigma > k\} d_k = f_{\varsigma \wedge n}$$

and

$$\sum_{k=-1}^{n-1} \chi\{\varsigma \le k\} d_k = f_n - f_{\varsigma \wedge n}$$

for $n \in \mathbf{N}$. In particular, the series

$$\sum_{-1 \le k < \varsigma} d_k = \sum_{k=-1}^{\infty} \chi\{\varsigma > k\} d_k$$

converges a.e. to f_ς, and

$$\sum_{\varsigma \le k} d_k = \sum_{k=-1}^{\infty} \chi\{\varsigma \le k\} d_k$$

converges a.e. to $f - f_\varsigma$.

Suppose in addition that ς and $\bar{\varsigma}$ are dyadic stopping times which satisfy $\varsigma \le \bar{\varsigma}$ on [0,1). Consider the series

$$\sum_{\varsigma \le k < \bar{\varsigma}} d_k = \sum_{k=-1}^{\infty} \chi\{\varsigma \le k < \bar{\varsigma}\} d_k. \tag{22}$$

Since

$$\sum_{k=-1}^{n-1} \chi\{\varsigma \le k < \bar{\varsigma}\} d_k = \sum_{k=-1}^{n-1} \chi\{\bar{\varsigma} > k\} d_k - \sum_{k=-1}^{n-1} \chi\{\varsigma > k\} d_k \tag{23}$$

it is evident that (22) converges a.e. to $f_{\bar{\varsigma}} - f_\varsigma$.

When the sequence $(f_n, n \in \mathbf{N})$ is a martingale we can say much more.

LEMMA 3. *Let $(f_n, n \in \mathbf{N})$ be a dyadic martingale which converges to a function f a.e. on $[0,1)$. Let ς and $\bar{\varsigma}$ be dyadic stopping times which satisfy $\varsigma \le \bar{\varsigma}$ and set*

$$g_n := \sum_{k=-1}^{n-1} \chi\{\varsigma \le k < \bar{\varsigma}\}(f_{k+1} - f_k),$$

$$g_n^* := \sup_{m \le n} |g_m| \qquad (n \in \mathbf{N}).$$

Then $(g_n, n \in \mathbf{N})$ is a dyadic martingale, $g := \lim_{n\to\infty} g_n$ exists a.e. and equals $f_{\bar{\varsigma}} - f_\varsigma$,

$$\{g_n^* \ne 0\} \subseteq \{\varsigma < \infty\}, \quad \{g \ne 0\} \subseteq \{\varsigma < \infty\}, \tag{24}$$

and

$$\chi\{\varsigma = m\} \mathcal{E}_m g_n = 0 \tag{25}$$

for all $n, m \in \mathbf{N}$.

Moreover, if $(f_n, n \in \mathbf{N})$ *is regular then* $g \in \mathrm{L}^1$, $g_n \to g$ *as* $n \to \infty$ *in the norm of* L^1, *and*

$$\chi\{\varsigma = m\}\mathcal{E}_m g = 0 \qquad (m \in \mathbf{N}). \tag{26}$$

PROOF. Since g and (22) coincide, we have already verified that $g = f_{\overline{\varsigma}} - f_\varsigma$ a.e.

Condition (18) implies both $\{\varsigma \le k\}$ and $\{\overline{\varsigma} > k\}$ belong to $\mathcal{A}^k$ for $k \in \mathbf{N}$. Consequently, $\{\varsigma \le k < \overline{\varsigma}\} \in \mathcal{A}^k$ and

$$\begin{aligned}
\mathcal{E}_k(g_{k+1}) - g_k &= \mathcal{E}_k(g_{k+1} - g_k) \\
&= \mathcal{E}_k\left(\chi\{\varsigma \le k < \overline{\varsigma}\}(f_{k+1} - f_k)\right) \\
&= \chi\{\varsigma \le k < \overline{\varsigma}\}\mathcal{E}_k(f_{k+1} - f_k) \\
&= 0.
\end{aligned}$$

In particular, $(g_k, k \in \mathbf{N})$ is a dyadic martingale.

Since g_n vanishes identically on the set $\{\varsigma = \infty\}$ for every $n \in \mathbf{N}$, it is clear that (24) holds.

To prove (25), observe by the definition of g_n that

$$\chi\{\varsigma = m\}g_n = 0$$

for $n \le m$. On the other hand, if $n > m$ then

$$\chi\{\varsigma = m\}g_n = \sum_{k=m}^{n-1} \chi\{\varsigma = m, \overline{\varsigma} > k\}(f_{k+1} - f_k).$$

It follows from (2) and (6) in 3.1 that

$$\chi\{\varsigma = m\}\mathcal{E}_m g_n = \sum_{k=m}^{n-1} \mathcal{E}_m\left(\mathcal{E}_k(\chi\{\varsigma = m, \overline{\varsigma} > k\}(f_{k+1} - f_k))\right) = 0.$$

Suppose now that $f_n = \mathcal{E}_n f$ for $n \in \mathbf{N}$ and some $f \in \mathrm{L}^1$. By (2) and (5) in 3.1 we have

$$\|f_\varsigma - f_{\varsigma \wedge n}\|_1 = \sum_{m \in \mathbf{N},\, m > n} \int_{\{\varsigma = m\}} |\mathcal{E}_m f - \mathcal{E}_n f|$$

and

$$\begin{aligned}
\int_{\{\varsigma = m\}} |\mathcal{E}_m f - \mathcal{E}_n f| &= \int_{\{\varsigma = m\}} |\mathcal{E}_m(f - \mathcal{E}_n f)| \\
&\le \int_{\{\varsigma = m\}} \mathcal{E}_m |f - \mathcal{E}_n f| \\
&= \int_{\{\varsigma = m\}} |f - \mathcal{E}_n f|
\end{aligned}$$

for $m > n$. It follows that

$$\|f_\varsigma - f_{\varsigma\wedge n}\|_1 \le \sum_{m\in\overline{\mathbf{N}},\, m>n} \int_{\{\varsigma=m\}} |f - \mathcal{E}_n f| \\ \le \|f - \mathcal{E}_n f\|_1.$$

In particular, $f_{\varsigma\wedge n} \to f_\varsigma$ in L^1 norm as $n \to \infty$.

It follows from (23) that $g_n \to g$ in L^1 norm as $n \to \infty$. Hence $g \in \mathrm{L}^1$ and $\mathcal{E}_m g_n \to \mathcal{E}_m g$ everywhere, as $n \to \infty$, for each $m \in \mathbf{N}$. Hence (25) implies (26) and the proof of the lemma is complete. ∎

Given a regular martingale $(\mathcal{E}_n f, n \in \mathbf{N})$ and an adapted sequence $(\lambda_n, n \in \mathbf{N})$, the stopping time (19) induces a decomposition of f by means of

$$f = f_\tau + (f - f_\tau).$$

For a suitable choice of λ_n's, this decomposition is the classical one of Calderon-Zygmund.

LEMMA 4. [THE CALDERON-ZYGMUND DECOMPOSITION] *Let $f \in \mathrm{L}^1$ and $y > \|f\|_1$. There exist non-overlapping dyadic intervals $I_0, I_1, \ldots$ and functions g, b such that*

$$f = g + b,$$

$$\|g\|_\infty \le 2y,$$

$$\{b \ne 0\} \subseteq \Omega,$$

where $\Omega := \cup_{k=0}^{\infty} I_k$,

$$|\Omega| \le \frac{1}{y}\|f\|_1,$$

$$\int_{I_k} b = 0,$$

and

$$y < \frac{1}{|I_k|}\int_{I_k} |f| \le 2y \qquad (k \in \mathbf{N}). \tag{27}$$

PROOF. Let $f_n := \mathcal{E}_n f$,

$$\lambda_n := \sup_{m\le n} \mathcal{E}_m|f| \qquad (n \ge -1),$$

and $\lambda^* := \sup_{n\in\mathbf{N}} \lambda_n$. Clearly, $f_n \to f$ a.e. as $n \to \infty$, $(\lambda_n, n \in \mathbf{N})$ is adapted, and $|f_n| \le \lambda_n$ for $n \in \mathbf{N}$.

Fix $m \in \mathbf{N}$ and let τ be defined by (19). It is easy to see that

$$\chi\{\tau = m\}|f_\tau| \le 2y. \tag{28}$$

In fact, since

$$|f_m| \leq \mathcal{E}_m|f| \leq 2\mathcal{E}_{m-1}|f|,$$

it is clear by (19) and (20) that

$$\begin{aligned}\chi\{\tau = m\}|f_\tau| &= \chi\{\tau = m\}|f_m| \\ &\leq 2\lambda_{m-1}\chi\{\tau = m\} \\ &\leq 2y,\end{aligned}$$

and

$$\chi\{\tau = \infty\}|f_\tau| = \chi\{\tau = \infty\}|f| \leq y$$

hold a.e. Hence the function $g := f_\tau$ satisfies $\|g\|_\infty \leq 2y$.

Set $b := \lim_{n\to\infty} \sum_{k=-1}^{n} \chi\{\tau \leq k\}(f_{k+1} - f_k)$ and $\Omega := \{\mathcal{E}^*|f| > y\}$. Apply Lemma 3 to $\overline{\varsigma} := \infty$ and $\varsigma := \tau$, noticing that b belongs to L^1 and $b = f - f_\tau$ a.e. Thus it follows from (24) and (21) that

$$\begin{aligned}\{b \neq 0\} &\subseteq \{\tau < \infty\} \\ &= \{\lambda^* > y\} \\ &= \Omega.\end{aligned}$$

Also, it follows from Corollary 1 in 3.1 that

$$|\Omega| \leq \frac{1}{y}\|f\|_1.$$

To see that Ω is a union of non-overlapping dyadic intervals, notice that each set in $\mathcal{A}^n$ can be split into a finite union of dyadic intervals of length 2^{-n}. But

$$\Omega = \{\tau < \infty\} = \bigcup_{n=0}^{\infty}\{\tau = n\},$$

and each $\{\tau = n\}$ belongs to $\mathcal{A}^n$. Thus Ω can be written as a countable union of non-overlapping dyadic intervals $I_0, I_1, \ldots$

Let $m \in \mathbf{N}$ and I be a dyadic interval belonging to $\mathcal{A}^m$. By (26) we have

$$\{\tau = m\}\mathcal{E}_m b = 0.$$

Consequently

$$\int_{I_k} b = 0 \qquad (k \in \mathbf{N}).$$

Moreover, the definition of τ and (28) imply

$$y < \mathcal{E}_k|f| \leq 2\mathcal{E}_{k-1}|f| \leq 2y$$

on the set $\{\tau = k\}$ for each $k \in \mathbf{N}$. Since by construction each I_j is contained in some $\{\tau = k\}$, (27) follows at once. ∎

A sequence $(f_n, n \in \mathbf{N})$ of measurable functions is said to be *predictable* by a sequence $\boldsymbol{\lambda} = (\lambda_n, n \in \mathbf{N})$ if $\boldsymbol{\lambda}$ is adapted, if $\lambda_n \le \lambda_{n+1}$ and $|f_{n+1}| \le \lambda_n$ hold for $n \in \mathbf{N}$, and if

$$\lambda^* := \lim_{n\to\infty} \lambda_n$$

is finite a.e. on [0,1).

It is important to notice that for each $f \in \mathrm{L}^1$ the regular martingale

$$f_n := \mathcal{E}_n f \qquad (n \in \mathbf{N})$$

is predictable by the adapted sequence

$$\lambda_n := \widetilde{\mathcal{E}}_n f \qquad (n \in \mathbf{N}).$$

Indeed, by definition

$$\begin{aligned} |\mathcal{E}_{n+1} f| &= |\mathcal{E}_n f + r_n \mathcal{E}_n(r_n f)| \\ &\le \sup_{0 \le m \le n} (|\mathcal{E}_m f| + |\mathcal{E}_m(r_m f)|) \\ &= \widetilde{\mathcal{E}}_n f, \end{aligned}$$

and by Theorem 2 in 3.1

$$\widetilde{\mathcal{E}} f = \lim_{n\to\infty} \widetilde{\mathcal{E}}_n f$$

is a.e. finite.

A function $f \in \mathrm{L}^1$ is said to be of *mean zero* if $\widehat{f}(0) = 0$. Predictable sequences induce decompositions of functions of mean zero in the following way.

THEOREM 3. *Let $f \in \mathrm{L}^1$ be of mean zero and suppose $f_n := \mathcal{E}_n f, n \in \mathbf{N}$, is predictable by $(\lambda_n, n \in \mathbf{N})$. For each $x \in [0,1)$ and each $k \in \mathbf{Z}$, define dyadic stopping times by*

$$\tau_k(x) := \min\{n \in \mathbf{N} : \lambda_n(x) > 2^k\}.$$

Then

$$\|f_{\tau_k}\|_\infty \le 2^k,$$

$$f_{\tau_{k+1}} - f_{\tau_k} = \sum_{m\in\mathbf{N}} \chi\{2^k < \lambda_m \le 2^{k+1}\}(f_{m+1} - f_m),$$

and

$$\mathcal{E}^*(f_{\tau_{k+1}} - f_{\tau_k}) \le 2^{k+2}\chi\{\lambda^* > 2^k\}$$

hold for each $k \in \mathbf{Z}$, and

$$f = \sum_{k\in\mathbf{Z}} (f_{\tau_{k+1}} - f_{\tau_k}) \tag{29}$$

holds a.e. on $[0,1)$.

Moreover, if λ^* *belongs to* L^1 *then (29) converges in* L^1 *norm and*

$$\sum_{k\in\mathbf{Z}} \|f_{\tau_{k+1}} - f_{\tau_k}\|_1 \le 8\|\lambda^*\|_1.$$

PROOF. Set $\lambda_{-1} := 0$, fix $k \in \mathbf{Z}$, let $m \le n$, $(m, n \in \mathbf{N})$, and observe by the definition of τ_k that on the set $\{\tau_k = m\}$ we have

$$\begin{aligned}|f_{\tau_k\wedge n}| &= |f_{\tau_k}|\\ &= |f_m|\\ &\le \lambda_{m-1}\\ &\le 2^k.\end{aligned}$$

Secondly, observe for $n < m < \infty$ that on the set $\{\tau_k = m\}$ we have

$$\begin{aligned}|f_{\tau_k\wedge n}| &= |f_n|\\ &\le \lambda_{m-1}\\ &\le 2^k.\end{aligned}$$

Third of all, observe that on the set $\{\tau_k = \infty\}$ we have

$$\begin{aligned}|f_{\tau_k\wedge n}| &= |f_n|\\ &\le \lim_{n\to\infty} \lambda_m\\ &= \lambda^*\\ &\le 2^k.\end{aligned}$$

It follows that

$$|f_{\tau_k\wedge n}| \le 2^k \quad \text{a.e..} \tag{30}$$

In particular, we obtain $\|f_{\tau_k}\|_\infty \le 2^k$ by letting $n \to \infty$.

Let

$$\Sigma_{MN} := \sum_{k=M}^{N} (f_{\tau_{k+1}} - f_{\tau_k})$$

for $-\infty < M < N < \infty$. Then $\Sigma_{MN} = f_{\tau_{N+1}} - f_{\tau_M}, \|f_{\tau_M}\|_\infty \le 2^M$, and $f_{\tau_N} = f$ on the set $\{\tau_N = \infty\} = \{\lambda^* \le 2^N\}$. Since λ^* is a.e. finite it follows that $\Sigma_{MN} \to f$ a.e. as $M \to -\infty$ and $N \to \infty$. Thus (29) converges a.e.

By Lemma 3 and (21), the series

$$\sum_{m=0}^{\infty} \chi\{\tau_k \le m < \tau_{k+1}\}(f_{m+1} - f_m) = \sum_{m=0}^{\infty} \chi\{2^k < \lambda_m \le 2^{k+1}\}(f_{m+1} - f_m)$$

converges to $f_{\tau_{k+1}} - f_{\tau_k}$ a.e. and in L^1 norm. Consequently,

$$\begin{aligned} g_n^{(k)} &:= \mathcal{E}_n(f_{\tau_{k+1}} - f_{\tau_k}) \\ &= \sum_{m=0}^{n-1} \chi\{\tau_k \le m < \tau_{k+1}\}(f_{m+1} - f_m) \\ &= f_{\tau_{k+1}\wedge n} - f_{\tau_k \wedge n}. \end{aligned}$$

Applying (24) we obtain

$$\{g_n^{(k)} \ne 0\} \subseteq \{\tau_k < \infty\} = \{\lambda^* > 2^k\}.$$

In particular, it follows from (30) that

$$\mathcal{E}^*(f_{\tau_{k+1}} - f_{\tau_k}) \le 2^{k+2}\chi\{\lambda^* > 2^k\}.$$

It remains to see that (29) converges in L^1 norm when $\lambda^* \in L^1$. But the inequality above implies

$$\sum_{k\in\mathbf{Z}} \|f_{\tau_{k+1}} - f_{\tau_k}\|_1 \le 4\|\sum_{k\in\mathbf{Z}} 2^k\chi\{\lambda^* > 2^k\}\|_1.$$

Hence we have by Abel's transformation that

$$\sum_{k\in\mathbf{Z}} \|f_{\tau_{k+1}} - f_{\tau_k}\|_1 \le 8\|\sum_{k\in\mathbf{Z}} 2^k\chi\{2^k < \lambda^* \le 2^{k+1}\}\|_1 \le 8\|\lambda^*\|_1.$$

Thus (29) converges in L^1 norm when λ^* belongs to L^1. ∎

COROLLARY 3. [THE CANONICAL DECOMPOSITION] *If $f \in L^1$ has mean zero and*

$$f^{(k)} := \sum_{n=0}^{\infty} \chi\{2^k < \widetilde{\mathcal{E}}_n f \le 2^{k+1}\}\Delta_n f,$$

where $\Delta_n f := \mathcal{E}_{n+1} f - \mathcal{E}_n f$, then

$$\mathcal{E}^*(f^{(k)}) \le 2^{k+2}\chi\{\widetilde{\mathcal{E}} f > 2^k\}$$

for all $k \in \mathbf{Z}$ and

$$f = \sum_{k\in\mathbf{Z}} f^{(k)}$$

a.e. on $[0,1)$.

Moreover, if $\mathcal{E}^* f \in L^1$ then this series converges to f in L^1 norm.

PROOF. Set $\lambda_n := \widetilde{\mathcal{E}}_n f, f_n := \mathcal{E}_n f$ for $n \in \mathbf{N}$, observe by (21) that

$$\chi\{\tau_k \le n < \tau_{k+1}\} = \chi\{2^k < \lambda_n \le 2^{k+1}\},$$

and apply Theorem 3. ∎

We shall use the canonical decomposition of f in the next section to prove certain classical inequalities, in 3.4 to study dyadic Hardy spaces, and in 3.7 to prove that the Walsh-Fourier series of an L^p function converges a.e. when $p > 1$.

3.3 Martingale Transforms. In this section we introduce martingale transforms and use them to obtain several classical inequalities.

Let $\boldsymbol{A}$ denote the collection of sequences $\boldsymbol{\alpha} = (\alpha_n, n \in \mathbf{N})$ which satisfy $\alpha_n \in \mathrm{L}(\mathcal{A}^n)$ for $n \in \mathbf{N}$ and

$$\|\boldsymbol{\alpha}\| := \sup_{n\in\mathbf{N}} \|\alpha_n\|_\infty < \infty.$$

Observe that $\boldsymbol{A}$ is an algebra with identity $(1, 1, \dots)$ under the operations

$$\boldsymbol{\alpha} + \boldsymbol{\beta} := (\alpha_n + \beta_n, n \in \mathbf{N}),$$

$$\boldsymbol{\alpha} \cdot \boldsymbol{\beta} := (\alpha_n \beta_n, n \in \mathbf{N}),$$

and

$$a\boldsymbol{\alpha} := (a\alpha_n, n \in \mathbf{N})$$

for $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \boldsymbol{A}$ and $a \in \mathbf{R}$. Also observe that $\|\cdot\| : \boldsymbol{A} \to [0, \infty)$ is a norm on $\boldsymbol{A}$.

Given $\boldsymbol{\alpha} \in \boldsymbol{A}$ and $f \in \mathrm{L}^1$, the *martingale transform* is defined by

$$\boldsymbol{T}(\boldsymbol{\alpha})f := \sum_{n=0}^{\infty} \alpha_n \Delta_n f,$$

where $\Delta_n f := \mathcal{E}_{n+1} f - \mathcal{E}_n f$ for $n \in \mathbf{N}$. The *maximal martingale transform* is defined by

$$\boldsymbol{T}^*(\boldsymbol{\alpha})f := \mathcal{E}^*(\boldsymbol{T}(\boldsymbol{\alpha})f) = \sup_{n\in\mathbf{N}} |\sum_{k=0}^{n} \alpha_k \Delta_k f|.$$

And the *square function* is defined by

$$\mathcal{Q}f := \left(\sum_{k=0}^{\infty} |\Delta_k f|^2\right)^{1/2}.$$

Notice that $\boldsymbol{T}(\boldsymbol{\alpha})f$ and $\mathcal{Q}f$ are finite sums for any $f \in \mathcal{P}$ since $\Delta_n f = 0$ for 2^n larger than the order of f. We shall use this and Theorem 2 in 3.1 to see that $\boldsymbol{T}(\boldsymbol{\alpha})f$ and $\mathcal{Q}f$ converge a.e. on [0,1) for all $f \in \mathrm{L}^1$ and $\boldsymbol{\alpha} \in \boldsymbol{A}$ (see Theorem 4 below).

$\boldsymbol{T}(\boldsymbol{\alpha})f$ is called a martingale transform because the sequence of partial sums

$$\boldsymbol{T}_N(\boldsymbol{\alpha})f := \sum_{n=0}^{N-1} \alpha_n \Delta_n f \qquad (N \in \mathbf{P})$$

forms a dyadic martingale. This fact follows immediately from the $\mathcal{A}^n$ measurability of each α_n.

Martingale transforms arise naturally in the study of partial sums of Walsh-Fourier series. Indeed, if $n \in \mathbf{N}$ has binary coefficients $(n_k, k \in \mathbf{N})$ then we have by Theorem 8 in 1.5 that

$$w_n S_n f = \sum_{k=0}^{\infty} n_k \Delta_k (f w_n).$$

Therefore,

$$|S_n f| = |\boldsymbol{T}(\boldsymbol{\alpha})(f w_n)|. \tag{31}$$

for $f \in \mathrm{L}^1$, and $\boldsymbol{\alpha} := (n_0, n_1, \ldots)$. (For a direct connection between L^p convergence of Walsh- Fourier series and martingale transforms, see Exercises 3.2 and 3.3.)

Let $\boldsymbol{L}$ denote the collection of linear maps from $\mathcal{P}$ into $\mathcal{P}$. It is well known that $\boldsymbol{L}$ is an algebra with identity. Since $\Delta_n f = 0$ for $f \in \mathcal{P}$ and n sufficiently large, it is clear that the map $\boldsymbol{\alpha} \to \boldsymbol{T}(\boldsymbol{\alpha})$ takes the algebra $\boldsymbol{A}$ into the algebra $\boldsymbol{L}$. This map is an algebra homomorphism. Indeed, for $\boldsymbol{\alpha}$, $\boldsymbol{\beta} \in \boldsymbol{A}$ and $a \in \mathbf{R}$ the identities

$$\boldsymbol{T}(\boldsymbol{\alpha} + \boldsymbol{\beta}) = \boldsymbol{T}(\boldsymbol{\alpha}) + \boldsymbol{T}(\boldsymbol{\beta})$$

and

$$\boldsymbol{T}(a\boldsymbol{\alpha}) = a\boldsymbol{T}(\boldsymbol{\alpha})$$

are trivial. And, the identity

$$\boldsymbol{T}(\boldsymbol{\alpha} \cdot \boldsymbol{\beta}) = \boldsymbol{T}(\boldsymbol{\alpha}) \circ \boldsymbol{T}(\boldsymbol{\beta})$$

follows from the fact that

$$\alpha_n \Delta_n \circ \alpha_m \Delta_m = \begin{cases} \alpha_n^2 \Delta_n & n = m \\ 0 & n \neq m. \end{cases} \tag{32}$$

(This fact follows easily from the orthogonality of the Walsh system and (6) in 3.1.)

The canonical decomposition introduced in the previous section is a sum of martingale transforms. To see this, let $f \in \mathrm{L}^1$ and define $\boldsymbol{\epsilon}^{(k)} \in \boldsymbol{A}$ for $k \in \mathbf{Z}$ by

$$\boldsymbol{\epsilon}^{(k)} := (\chi\{2^k < \widetilde{\mathcal{E}}_n f \leq 2^{k+1}\}, n \in \mathbf{N}).$$

In the notation of Corollary 3 in 3.2, we have

$$f^{(k)} = \boldsymbol{T}(\boldsymbol{\epsilon}^{(k)}) f \qquad (k \in \mathbf{Z}).$$

Therefore, the canonical decomposition of f has the form

$$f = \sum_{k \in \mathbf{Z}} \boldsymbol{T}(\boldsymbol{\epsilon}^{(k)}) f.$$

It is easy to see that each martingale transform is an operator of type (2,2). Indeed, fix $\boldsymbol{\alpha} \in \boldsymbol{A}$, $f \in \mathrm{L}^2$ and observe by (32) that

$$\begin{aligned} \|\sum_{n=N}^{M} \alpha_n \Delta_n f\|_2^2 &= \sum_{n=N}^{M} \|\alpha_n \Delta_n f\|_2^2 \\ &\leq \|\boldsymbol{\alpha}\|^2 \sum_{n=N}^{M} \|\Delta_n f\|_2^2 \\ &= \|\boldsymbol{\alpha}\|^2 \|\mathcal{E}_{M+1} f - \mathcal{E}_N f\|_2^2 \end{aligned}$$

for $M, N \in \mathbf{N}$ and $M \geq N$. Hence (2) and (4) in 3.1 imply

$$\|\sum_{n=N}^{M} \alpha_n \Delta_n f\|_2 \leq \|\boldsymbol{\alpha}\| \, \|f - \mathcal{E}_N f\|_2.$$

It follows that $\boldsymbol{T}(\boldsymbol{\alpha})f$ converges in L^2 norm and that

$$\|\boldsymbol{T}(\boldsymbol{\alpha})f\|_2 \leq \|\boldsymbol{\alpha}\| \, \|f\|_2 \qquad (\boldsymbol{\alpha} \in \boldsymbol{A},\ f \in \mathrm{L}^2). \tag{33}$$

This simple observation can be used to show that the martingale transform of an L^1 function converges a.e. In fact, we will use it to prove the following.

THEOREM 4. *There exists an absolute constant A such that*

$$y|\{\mathcal{Q}f > y\}| \leq A\|f\|_1$$

and

$$y|\{\boldsymbol{T}^*(\boldsymbol{\alpha})f > y\}| \leq A\|\boldsymbol{\alpha}\| \, \|f\|_1$$

for all $f \in \mathrm{L}^1, y > 0$, and $\boldsymbol{\alpha} \in \boldsymbol{A}$.

PROOF. Let Λ represent either $\mathcal{Q}$ or $\boldsymbol{T}^*(\boldsymbol{\alpha})$. We may suppose that $\|\boldsymbol{\alpha}\| \leq 1$. Since

$$\Lambda f = \lim_{N \to \infty} \Lambda(\mathcal{E}_N f)$$

and $\|\mathcal{E}_N f\|_1 \leq \|f\|_1$ for $N \in \mathbf{N}$, we may also suppose that f is a Walsh polynomial.

Fix $f \in \mathcal{P}, y > 0$ and set $\lambda^* := \tilde{\mathcal{E}} f$. Observe that

$$|\{\Lambda f > y\}| \leq |\{\Lambda f > y, \lambda^* \leq y\}| + |\{\lambda^* > y\}|.$$

Since by Corollary 1 in 3.1 we have

$$y|\{\lambda^* > y\}| \leq 2\|f\|_1$$

it suffices to show

$$y|\{\Lambda f > y, \lambda^* \leq y\}| \leq A\|f\|_1$$

for some absolute constant A.

Let $f = \sum_{k \in \mathbf{Z}} f^{(k)}$ represent the canonical decomposition of f and consider the functions

$$g_{2^k} := \Lambda(f^{(k)}) \qquad (k \in \mathbf{Z}).$$

Since Λ is sublinear it is clear that

$$\Lambda f \leq \sum_{k \in \mathbf{Z}} g_{2^k}.$$

Moreover, by Corollary 3 in 3.2 we have $g_{2^k} = \chi\{\lambda^* > 2^k\}g_{2^k}$. Consequently,

$$\Lambda f \leq G := \sum_{k\in\mathbf{Z}} \chi\{\lambda^* > 2^k\}g_{2^k}.$$

We shall apply Lemma 2 in 3.2. Fix $k \in \mathbf{Z}$ and $t > 0$ and notice that

$$|\{g_{2^k} > t2^k\}| \leq \left(\frac{1}{t2^k}\right)^2 \|g_{2^k}\|_2^2.$$

When $\Lambda = \mathcal{Q}$ we have

$$\|g_{2^k}\|_2 = \|\mathcal{Q}f^{(k)}\|_2 = \|f^{(k)}\|_2.$$

When $\Lambda = \boldsymbol{T}^*(\boldsymbol{\alpha})$ we have by Corollary 1 in 3.1 and (33) above that

$$\begin{aligned}\|g_{2^k}\|_2 &= \|\boldsymbol{T}^*(\boldsymbol{\alpha})(f^{(k)})\|_2\\ &= \|\mathcal{E}^*(\boldsymbol{T}(\boldsymbol{\alpha})f^{(k)})\|_2\\ &\leq 2\|\boldsymbol{T}(\boldsymbol{\alpha})f^{(k)}\|_2\\ &\leq 2\|f^{(k)}\|_2.\end{aligned}$$

Therefore, in either case we have

$$|\{g_{2^k} > t2^k\}| \leq 4\left(\frac{1}{t2^k}\right)^2 \|f^{(k)}\|_2^2. \tag{34}$$

Recall from Theorem 3 in 3.2 that

$$|f^{(k)}| \leq 2^{k+2}\chi\{\lambda^* > 2^k\}.$$

Hence it follows from (34) and Corollary 1 in 3.1 that

$$|\{g_{2^k} > t2^k\}| \leq 64t^{-2}|\{\lambda^* > 2^k\}| \leq 128\,\|f\|_1 t^{-2}2^{-k}$$

for $k \in \mathbf{Z}$ and $t > 0$. Therefore, by Lemma 2 in 3.2, with $p = 1$ and $q = 2$, we have

$$y|\{\Lambda f > y, \lambda^* \leq y\}| \leq 128\,C(1,2)\|f\|_1. \qquad \blacksquare$$

By Theorem 4 and the Marcinkiewicz interpolation theorem, we have that the operators $\mathcal{Q}$ and $\boldsymbol{T}^*(\boldsymbol{\alpha})$ are of type (p,p) for $1 < p < \infty$. However, much more is true.

THEOREM 5. *Let Φ be a continuous, non-decreasing function on $[0,\infty)$ which satisfies $\Phi(0) = 0$ and $\Phi(2y) \leq C_0\Phi(y)$ for all $y > 0$, where C_0 is a positive constant independent of y. There exist positive constants A, B, C which depend only on Φ such that*

$$A\int_0^1 \Phi\circ\mathcal{E}^*f \leq \int_0^1 \Phi\circ\mathcal{Q}f \leq B\int_0^1 \Phi\circ\mathcal{E}^*f$$

and

$$\text{(35)} \qquad \int_0^1 \Phi \circ \boldsymbol{T}^*(\boldsymbol{\alpha})f \le C \int_0^1 \Phi \circ \mathcal{E}^* f$$

for all $f \in \mathrm{L}^1$ *and* $\boldsymbol{\alpha} \in \boldsymbol{A}$ *with* $\|\boldsymbol{\alpha}\| \le 1$.

PROOF. Fix $f \in \mathrm{L}^1$, set $f_{-1} := f_0 := 0$ and let $f_n \in \mathrm{L}(\mathcal{A}^n)$, $n = 1, 2, \ldots$ be a sequence of Walsh polynomials which converges a.e. to f on [0,1). Let $\lambda_n \in \mathrm{L}(\mathcal{A}^n)$ be an increasing sequence of functions (to be specified below) which satisfy

$$\text{(36)} \qquad |f_{n+1} - f_n| \le \lambda_n \qquad (n \in \mathbf{N}).$$

Set $g_n := \max_{m \le n} |f_m|$, $G := \lim_{n\to\infty} g_n$, $\lambda_{-1} := 0$, and $\lambda := \lim_{n\to\infty} \lambda_n$. Fix $y > 0$, $\beta > 1$, $N > 1$ and consider the stopping times

$$\begin{aligned} \upsilon &:= \min\{n \in \mathbf{N} : g_n > y\}, \\ \varsigma &:= \min\{n \in \mathbf{N} : g_n > \beta y\}, \\ \tau &:= \min\{n \in \mathbf{N} : N\lambda_n > \beta y\}. \end{aligned}$$

Set

$$\begin{aligned} F &:= \sum_{k=0}^{\infty} \chi\{\upsilon \le k < \tau \wedge \varsigma\}(f_{k+1} - f_k) \\ &= f_{\tau\wedge\varsigma} - f_{\tau\wedge\upsilon}. \end{aligned}$$

Clearly, if $x \in \{G > \beta y, \lambda \le \beta y/N\} = \{\varsigma < \infty, \tau = \infty\}$, then

$$\begin{aligned} F(x) &= f_\varsigma(x) - f_\upsilon(x) \\ &= f_\varsigma(x) - f_{\upsilon-1}(x) - (f_\upsilon(x) - f_{\upsilon-1}(x)). \end{aligned}$$

Hence (36) implies $|F(x)| > \beta y - y - \lambda_{\upsilon-1}(x) \ge y(\beta - 1 - \beta/N)$. In particular, we have proved

$$\text{(37)} \qquad \{G > \beta y, \lambda \le \frac{\beta y}{N}\} \subseteq \{|F| > y(\beta - 1 - \frac{\beta}{N})\}.$$

We apply these observations first to the case $f_n := \mathcal{E}_n f$ and

$$\lambda_n := \left(\sum_{k=0}^{n} |\Delta_k f|^2\right)^{1/2} \qquad (n \in \mathbf{N}).$$

Thus $G = \mathcal{E}^* f$, $\lambda = \mathcal{Q} f$ and F is a martingale transform. Set

$$\Lambda := \sum_{k=0}^{\infty} \chi\{\upsilon \le k < \tau\}(\lambda_k^2 - \lambda_{k-1}^2)$$

and observe that

$$\begin{aligned}\|F\|_2^2 &= \int_0^1 \sum_{k=0}^{\infty} \chi\{v \le k < \tau \wedge \varsigma\}|\Delta_k f|^2 \\ &\le \int_0^1 \sum_{k=0}^{\infty} \chi\{v \le k < \tau\}|\Delta_k f|^2 \\ &= \int_0^1 \Lambda.\end{aligned}$$

But

$$\{\Lambda \neq 0\} \subseteq \{v < \infty\} = \{G > y\},$$

and by construction

$$\Lambda \le \lambda_{\tau-1}^2 \le \left(\frac{\beta y}{N}\right)^2 \chi\{G > y\}.$$

Therefore, it follows from (37) that

$$y^2(\beta - 1 - \frac{\beta}{N})^2|\{G > \beta y, \lambda \le \frac{\beta y}{N}\}| \le \|F\|_2^2 \le \|\Lambda\|_1 \le (\frac{\beta y}{N})^2|\{G > y\}|.$$

In particular,

$$|\{\mathcal{E}^* f > \beta y, \mathcal{Q} f \le \frac{\beta y}{N}\}| \le \delta_N |\{\mathcal{E}^* f > y\}|,$$

for

$$\delta_N := \frac{(\beta/N)^2}{(\beta - 1 - \beta/N)^2}.$$

This inequality is similar to (12) in 3.2. Since $\delta_N \to 0$ as $N \to \infty$ we can choose N so large that the hypotheses of Lemma 1 are satisfied. We conclude that there is an absolute constant $A > 0$ depending only on Φ such that

$$A \int_0^1 \Phi \circ \mathcal{E}^* f \le \int_0^1 \Phi \circ \mathcal{Q} f.$$

To obtain the inequality involving B, set $f_n := \sum_{k=0}^{n-1} |\Delta_k f|^2$ and $\lambda_n := |\widetilde{\mathcal{E}}_n f|^2$. In this case, $G = |\mathcal{Q} f|^2, \lambda = |\widetilde{\mathcal{E}} f|^2$ and the corresponding function F can be estimated by

$$\begin{aligned}\|F\|_1 &\le \int_0^1 \sum_{k=0}^{\infty} \chi\{v \le k < \tau\}|\Delta_k f|^2 \\ &= \int_0^1 |\sum_{k=0}^{\infty} \chi\{v \le k < \tau\}\Delta_k f|^2 \\ &\le \int_0^1 |\mathcal{E}_{\tau \wedge v} f - \mathcal{E}_\tau f|^2 \chi\{v < \infty\} \\ &\le 4 \int_0^1 |\widetilde{\mathcal{E}}_{\tau-1} f|^2 \chi\{v < \infty\} \\ &\le 4y\frac{\beta}{N}|\{G > y\}|.\end{aligned}$$

If we set $\widetilde{\beta} := \sqrt{\beta}, \widetilde{y} := \sqrt{y}, \widetilde{N} := \sqrt{N}$ and

$$\widetilde{\delta}_N := \frac{4(\beta/N)}{(\beta - 1 - \beta/N)}$$

then it follows from (37) that

$$\begin{aligned} |\{\mathcal{Q}f > \widetilde{\beta}\widetilde{y}, \widetilde{\mathcal{E}}f \le \frac{\widetilde{\beta}\widetilde{y}}{\widetilde{N}}\}| &= |\{G > \beta y, \lambda \le \frac{\beta y}{N}\}| \\ &\le \widetilde{\delta}_N |\{G > y\}| \\ &= \widetilde{\delta}_N |\{\mathcal{Q}f > \widetilde{y}\}|. \end{aligned}$$

Thus Lemma 1 applies as before.

Finally, to verify (35) set $f_n := \sum_{k=0}^{n-1} \alpha_k(\mathcal{E}_{k+1}f - \mathcal{E}_k f), \lambda_n := \widetilde{\mathcal{E}}_n f$ and apply Lemma 1 once more. We obtain

$$|\{\boldsymbol{T}^*(\boldsymbol{\alpha})f > \beta y, \widetilde{\mathcal{E}}f \le \frac{\beta y}{N}\}| \le \delta_N^* |\{\boldsymbol{T}^*(\boldsymbol{\alpha})f > y\}|,$$

for

$$\delta_N^* := \frac{4(\beta/N)^2}{(\beta - 1 - \beta/N)^2}. \qquad ∎$$

Specializing to the case $\Phi(y) := y^p$ we obtain the following inequalities.

COROLLARY 4. *For each* $0 < p < \infty$ *there are positive constants* A_p, B_p, C_p *such that*

$$A_p \|\mathcal{E}^* f\|_p \le \|\mathcal{Q}f\|_p \le B_p \|\mathcal{E}^* f\|_p, \tag{38}$$

and

$$\|\boldsymbol{T}^*(\boldsymbol{\alpha})f\|_p \le C_p \|\boldsymbol{\alpha}\| \, \|\mathcal{E}^* f\|_p$$

for all $f \in \mathrm{L}^1$ *and* $\boldsymbol{\alpha} \in \boldsymbol{A}$.

The proof of Theorem 5 and evaluation of the constant

$$\varpi := \frac{\gamma\eta}{(1 - \gamma\varepsilon)}$$

which appears on the right side of (14), yields estimates for these constants A_p, B_p, C_p. Indeed, fix $0 < p < \infty$ and set

$$\beta := 1 + \frac{1}{p}, \qquad N := \frac{\beta p}{c}$$

for some $0 < c < 1$. Then $\varepsilon := \delta_N = c^2/(1-c)^2, \gamma := \beta^p = (1 + 1/p)^p$ and $\eta := (N/\beta)^p$. If we choose c so small that $\gamma\varepsilon < 1$ then we have

$$\varpi = \frac{(1-c)^2\gamma}{(1-c)^2 - c^2\gamma} \left(\frac{p}{c}\right)^p.$$

Similarly, we obtain $\delta_N^* = 4c^2/(1-c)^2$ and

$$\varpi = \frac{(1-c)^2\gamma}{(1-c)^2 - 4c^2\gamma}\left(\frac{p}{c}\right)^p.$$

In particular,

$$A_p^{-p} \le C_p^p = A\left(\frac{p}{c}\right)^p$$

holds for certain absolute constants A and c. (For example, $A = 7$ and $c = 1/12$ will work.)

To estimate the constant B_p, set

$$\widetilde{\beta} := 1 + \frac{1}{p}, \qquad \widetilde{N} := \widetilde{\beta}\sqrt{\frac{p}{c}}$$

and verify that

$$\widetilde{\delta}_N = \frac{4c}{(2-c) + 1/p^2} \le \frac{4c}{2-c},$$

$\gamma = (1+1/p)^p$ and $\eta = (p/c)^{p/2}$. Consequently, in this case we have

$$\varpi \le \frac{(2-c)\gamma}{2-c-4c\gamma}\left(\frac{p}{c}\right)^{p/2}.$$

We conclude that

$$B_p^p \le A\left(\frac{p}{c}\right)^{p/2}$$

for some absolute constants A and c.

If we combine these estimates with Corollary 1 in 3.1 we obtain the following.

COROLLARY 5. *There is a positive constant* $\widetilde{A}$ *such that*

$$\widetilde{A}^{-1}\frac{1}{p}\|f\|_p \le \|\mathcal{Q}f\|_p \le \widetilde{A}\frac{p^{3/2}}{(p-1)}\|f\|_p \tag{39}$$

and

$$\|\boldsymbol{T}^*(\boldsymbol{\alpha})f\|_p \le \widetilde{A}\frac{p^2}{(p-1)}\|\boldsymbol{\alpha}\|\|f\|_p$$

hold for all $\boldsymbol{\alpha} \in \boldsymbol{A}$, $f \in \mathrm{L}^p$, *and* $1 < p < \infty$.

Inequality (39) is called Paley's inequality. An exponential version is given in Exercise 3.7. It fails to hold when $p = 1$ and $p = \infty$ (see Exercise 3.6).

In view of (31), Corollary 5 contains the following result.

COROLLARY 6. *For each $1 < p < \infty$ there is a positive constant C_p such that*

$$\sup_{n\in\mathbf{N}} \|S_n f\|_p \le C_p \|f\|_p$$

for all $f \in \mathrm{L}^p$.

This inequality and the Banach-Steinhaus theorem imply that the Walsh-Fourier series of any $f \in \mathrm{L}^p, 1 < p < \infty$, converges to f in L^p norm (see Exercise 3.3).

Since the square function of a Rademacher series $\sum_{k=0}^{\infty} c_k r_k$ is precisely $(\sum_{k=0}^{\infty} |c_k|^2)^{1/2}$, Corollary 4 contains the following result.

COROLLARY 7. [KHINTCHIN'S INEQUALITY] *For each $0 < p < \infty$ there exist positive constants A_p, B_p such that*

$$A_p \|\mathbf{c}\|_{\ell^2} \le \left(\int_0^1 |\sum_{k=0}^{\infty} c_k r_k|^p \right)^{1/p} \le B_p \|\mathbf{c}\|_{\ell^2}$$

for all $\mathbf{c} = (c_k, k \in \mathbf{N}) \in \ell^2$.

Fix $n \in \mathbf{N}, E \in \mathcal{A}^n, f \in \mathrm{L}^p$ and set $g_n := f - \mathcal{E}_n f$. Observe that

$$\chi(E)\boldsymbol{T}^*(\boldsymbol{\alpha})g_n = \boldsymbol{T}^*(\boldsymbol{\alpha})(\chi(E)g_n),$$

and

$$\chi(E)\mathcal{Q}g_n = \mathcal{Q}(\chi(E)g_n).$$

Specializing to $E := I(k,n)$ and applying Corollary 5 to $f := g_n$, we obtain:

COROLLARY 8. *There exists an absolute constant A such that*

$$\mathcal{E}_n(|\boldsymbol{T}^*(\boldsymbol{\alpha})(f - \mathcal{E}_n f)|^p) \le \left(\frac{Ap^2}{p-1}\right)^p \mathcal{E}_n(|f - \mathcal{E}_n f|^p)$$

and

$$\mathcal{E}_n(|\mathcal{Q}(f - \mathcal{E}_n f)|^p) \le \left(\frac{Ap^{3/2}}{p-1}\right)^p \mathcal{E}_n(|f - \mathcal{E}_n f|^p)$$

for $1 < p < \infty, n \in \mathbf{N}, f \in \mathrm{L}^p$, and $\boldsymbol{\alpha} \in \boldsymbol{A}$ with $\|\boldsymbol{\alpha}\| \le 1$.

We have seen that the map $\boldsymbol{\alpha} \to \boldsymbol{T}(\boldsymbol{\alpha})$ is an algebra homomorphism from $\boldsymbol{A}$ into $\boldsymbol{L}$. We close this section with a slightly more general point of view.

Fix $1 < p < \infty$ and let $\boldsymbol{\mathcal{L}}^p$ represent the collection of linear operators which are bounded from L^p into itself. Since the operators $\boldsymbol{T}(\boldsymbol{\alpha}) : \mathrm{L}^p \to \mathrm{L}^p$ are bounded on the subspace $\mathcal{P}$ and this subspace is dense in L^p, it is clear that $\boldsymbol{T}(\boldsymbol{\alpha}) \in \boldsymbol{\mathcal{L}}^p$ for every $\boldsymbol{\alpha} \in \boldsymbol{A}$. Moreover,

$$\boldsymbol{T}(\boldsymbol{\alpha} \cdot \boldsymbol{\beta})f = (\boldsymbol{T}(\boldsymbol{\alpha}) \circ \boldsymbol{T}(\boldsymbol{\beta}))(f)$$

for $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \boldsymbol{A}$ and $f \in \mathrm{L}^p$, and the map $\boldsymbol{\alpha} \to \boldsymbol{T}(\boldsymbol{\alpha})$ is an algebra homomorphism from $\boldsymbol{A}$ into $\boldsymbol{\mathcal{L}}^p$.

3.4 Dyadic Hardy Spaces and BMO. For each $f \in L^1$ and $0 < p < \infty$ set

$$\|f\|_{H^p} := \|\mathcal{E}^* f\|_p. \tag{40}$$

By Corollary 4 in 3.3, a function f satisfies $\|f\|_{H^p} < \infty$ if and only if $\mathcal{Q}f$ belongs to L^p.

On the collection of Walsh polynomials $\mathcal{P}$, the map $f \to \|f\|_{H^p}$ is a norm for $1 \le p < \infty$ and a quasi-norm for $0 < p < 1$. Define the dyadic Hardy space H^p to be the closure of $\mathcal{P}$ in the (quasi-) norm $\|\cdot\|_{H^p}$ for $0 < p < \infty$. Thus for $1 \le p < \infty$, the space H^p is precisely the collection of functions $f \in L^1$ such that $\|f\|_{H^p} < \infty$. Notice also by Corollary 5 in 3.3 that the spaces H^p and L^p are isomorphic for $1 < p < \infty$.

We shall denote H^1 by H. Since $\|f\|_1 \le \|f\|_{\mathrm{H}}$ and L^1 is complete, it can be shown that H is complete. Moreover, it is easy to see that

$$\lim_{n\to\infty} \|S_{2^n} f - f\|_{\mathrm{H}} = 0 \qquad (f \in \mathrm{H}) \tag{41}$$

(see also Theorem 7 in 4.4).

The space H lies between the spaces $L^p, p > 1$, and the space L^1. Indeed,

$$L^p = H^p \subset \mathrm{H} \subset L^1$$

for all $p > 1$ (see Exercise 3.14).

A function $\beta \in L^\infty$ is called a *(dyadic) atom* if either $\beta = 1$ or if there is a dyadic interval $I \subseteq [0, 1)$ such that

$$\{\beta \ne 0\} \subseteq I, \tag{42}$$

$$\|\beta\|_\infty \le \frac{1}{|I|}, \tag{43}$$

and

$$\int_0^1 \beta = 0. \tag{44}$$

If (42) and (43) hold, we shall say that β is *supported* on I. Thus if β is supported on I then

$$|\beta| \le \frac{\chi(I)}{|I|}. \tag{45}$$

Moreover, its dyadic maximal operator satisfies

$$\mathcal{E}^*\beta \le \frac{\chi(I)}{|I|}. \tag{46}$$

To see this, choose $N \in \mathbf{N}$ such that $|I| = 2^{-N}$ and observe that $\mathcal{E}_n\beta = 0$ for $0 \le n \le N$ by (44), and that

$$|\mathcal{E}_n\beta| \le \mathcal{E}_n\left(\frac{\chi(I)}{|I|}\right) = \frac{\chi(I)}{|I|}$$

for $n > N$ by (45).

Dyadic atoms characterize the dyadic Hardy space H.

THEOREM 6. *A function* $f \in \mathrm{L}^1$ *belongs to* H *if and only if there exist dyadic atoms* $\beta_0, \beta_1, \dots$ *and a sequence* $\boldsymbol{a} = (a_n, n \in \mathbf{N}) \in \ell^1$ *such that*

$$f = \sum_{n=0}^{\infty} a_n \beta_n. \tag{47}$$

Furthermore, if

$$\varrho(f) := \inf\{ \|\boldsymbol{a}\|_{\ell^1} \},$$

where the infimum is taken over all sequences $\boldsymbol{a} = (a_n, n \in \mathbf{N}) \in \ell^1$ *such that (47) holds for some atoms* $\beta_0, \beta_1, \dots$, *then*

$$\|f\|_{\mathrm{H}} \leq \varrho(f) \leq 25\|f\|_{\mathrm{H}} \tag{48}$$

for all $f \in \mathrm{H}$.

PROOF. Notice by (45) that (47) converges in L^1 norm. Consequently, if (47) holds then

$$\|\mathcal{E}^* f\|_1 \leq \sum_{n=0}^{\infty} |a_n| \, \|\mathcal{E}^* \beta_n\|_1.$$

In particular, such an f belongs to H and the left side of (48) is true.

Conversely, suppose $f \in \mathrm{H}$. Let $\sum_{k \in \mathbf{Z}} f^{(k)}$ represent the canonical decomposition of $f - \hat{f}(0)$. Thus $f^{(k)} := \boldsymbol{T}(\boldsymbol{\epsilon}^{(k)}) f$ where

$$\boldsymbol{\epsilon}^{(k)} := (\epsilon_n^{(k)}, n \in \mathbf{N}) := (\chi\{2^k < \widetilde{\mathcal{E}}_n f \leq 2^{k+1}\}, n \in \mathbf{N}) \qquad (k \in \mathbf{Z}).$$

By Corollary 3 in 3.2,

$$f - \hat{f}(0) = \sum_{k \in \mathbf{Z}} f^{(k)}$$

in L^1 norm. Moreover, since $\mathcal{E}^* f$ dominates $|f|$, it follows from Corollary 3 in 3.2 that

$$|f^{(k)}| \leq 2^{k+2} \chi\{\widetilde{\mathcal{E}} f > 2^k\} \qquad (k \in \mathbf{Z}). \tag{49}$$

Fix $k \in \mathbf{Z}$ and write $\chi\{\widetilde{\mathcal{E}} f > 2^k\}$ as a union of pairwise disjoint sets

$$E_n^{(k)} := \{\widetilde{\mathcal{E}}_n f > 2^k, \widetilde{\mathcal{E}}_{n-1} f \leq 2^k\} \qquad (n \in \mathbf{N}),$$

with the convention that $\widetilde{\mathcal{E}}_{-1} f := 0$. Since $E_n^{(k)} \in \mathcal{A}^n$ for each $n \in \mathbf{N}$, we can split this set into pairwise disjoint dyadic intervals, say

$$E_n^{(k)} = \bigcup_{I \in \mathcal{T}_n^{(k)}} I$$

for some collection $\mathcal{T}_n^{(k)} \subseteq \mathcal{A}^n$. Thus we can write $f^{(k)}$ in the form

$$f^{(k)} = \sum_{n=0}^{\infty} \sum_{I \in \mathcal{T}_n^{(k)}} \chi(I) f^{(k)}.$$

In particular, the canonical decomposition takes on the form

$$f = \widehat{f}(0) + \sum_{k \in \mathbf{Z}} \sum_{n \in \mathbf{N}} \sum_{I \in \mathcal{T}_n^{(k)}} a_I \beta_I \tag{50}$$

where for each $I \in \mathcal{T}_n^{(k)}$ we have defined

$$a_I := 2^{(k+2)}|I|, \qquad \beta_I := 2^{-(k+2)}|I|^{-1}\chi(I) f^{(k)}.$$

We shall show that (50) is an atomic decomposition of f, i.e., that each β_I is an atom and the sequence $\boldsymbol{a} := (a_I, I \in \mathcal{T}_n^{(k)}, k \in \mathbf{Z}, n \in \mathrm{N})$ belongs to ℓ^1.

Fix $I \in \mathcal{T}_n^{(k)}$. By construction, $\{\beta_I \neq 0\} \subseteq I$. Moreover, since $m < n$ implies

$$\chi(I)\epsilon_m^{(k)} = 0,$$

it is clear by (49) that $\|\beta_I\|_\infty \leq |I|^{-1}$ and

$$\chi(I) f^{(k)} = \sum_{m=n}^{\infty} \epsilon_m^{(k)} \Delta_m f \chi(I).$$

It follows that $\int_I f^{(k)} = 0$ and each β_I is a dyadic atom.

To compute $\|\boldsymbol{a}\|_{\ell^1}$, observe that

$$\begin{aligned}\|\boldsymbol{a}\|_{\ell^1} &= |\widehat{f}(0)| + \sum_{k \in \mathbf{Z}} \sum_{n \in \mathbf{N}} \sum_{I \in \mathcal{T}_n^{(k)}} |a_I| \\ &= |\widehat{f}(0)| + \sum_{k \in \mathbf{Z}} 2^{k+2} |\{\widetilde{\mathcal{E}} f > 2^k\}|.\end{aligned}$$

By Abel's transformation this second term can be written as

$$4 \sum_{k \in \mathbf{Z}} (2^{k+1} - 2^k) |\{\widetilde{\mathcal{E}} f > 2^k\}| = 4 \sum_{k \in \mathbf{Z}} 2^k |\{2^{k-1} < \widetilde{\mathcal{E}} f \leq 2^k\}| \leq 8 \|\widetilde{\mathcal{E}} f\|_1.$$

We conclude by (7) in 3.1 that

$$\|\boldsymbol{a}\|_{\ell^1} \leq 25 \|\mathcal{E}^* f\|_1. \qquad \blacksquare$$

We shall measure the *dyadic mean oscillation* of a $\psi \in \mathrm{L}^2$ by

$$\|\psi\|_{\mathrm{BMO}} := \sup_{n \in \mathbf{N}} \| \left(\mathcal{E}_n |\psi - \mathcal{E}_n \psi|^2\right)^{1/2} \|_\infty.$$

Observe that $\|\cdot\|_{\mathrm{BMO}}$ is a seminorm but $\|\psi\|_{\mathrm{BMO}} = 0$ for any constant ψ. Thus we shall say that a $\psi \in \mathrm{L}^2$ is of ***bounded dyadic mean oscillation*** if ψ satisfies $\|\psi\|_{\mathrm{BMO}} < \infty$ and ψ is of mean zero . The collection of all functions of bounded dyadic mean oscillation will be denoted by BMO. In particular, $(\mathrm{BMO}, \|\cdot\|_{\mathrm{BMO}})$ is a normed linear space.

In the next section we shall show that BMO is essentially the dual of H. It follows that BMO is a Banach space. Notice by definition that $\|\psi\|_{\mathrm{BMO}} \leq 2\|\psi\|_\infty$. Since H lies between $\mathrm{L}^p, p > 1$, and L^1, it also follows that

$$\mathrm{L}^\infty \subset \mathrm{BMO} \subset \mathrm{L}^p \qquad (p < \infty).$$

It is easy to see that these spaces are all distinct (see Exercise 3.15, for example).

We shall denote the collection of dyadic intervals in [0,1) by $\mathcal{I}_0$ and set

$$\mathcal{I} := \mathcal{I}_0 \cup \{\emptyset\}.$$

Since

$$(\mathcal{E}_n f)(x) = \frac{1}{|I_n(x)|}\int_{I_n(x)} f$$

for all $f \in \mathrm{L}^1, x \in [0,1)$, and $n \in \mathbf{N}$, it is evident that a function ψ belongs to BMO if and only if ψ is of mean zero and

$$\sup_{I\in\mathcal{I}_0}\left(\frac{1}{|I|}\int_I |\psi - \frac{1}{|I|}\int_I \psi|^2\right)^{1/2} < \infty.$$

Hence the term "bounded dyadic mean oscillation" is well chosen.

A function $\psi \in \mathrm{BMO}$ is said to be of ***vanishing dyadic mean oscillation*** if

$$\lim_{n\to\infty} \|(\mathcal{E}_n|\psi - \mathcal{E}_n\psi|^2)^{1/2}\|_\infty = 0.$$

The collection of functions of vanishing dyadic mean oscillation will be denoted by VMO. Clearly, VMO is a closed subspace of BMO.

The Walsh polynomials are dense in VMO. Indeed, let $\psi \in \mathrm{VMO}$ and notice by (2) in 3.1 that

$$\mathcal{E}_n|\psi - \mathcal{E}_n\psi|^2 = \mathcal{E}_n(\sum_{k=n}^\infty |\Delta_k\psi|^2) \qquad (n \in \mathbf{N}).$$

Hence we can write

$$\|\psi - \mathcal{E}_n\psi\|_{\mathrm{BMO}} = \sup_{m\in\mathbf{N}} \|(\mathcal{E}_m(\sum_{k=\max\{m,n\}}^\infty |\Delta_k\psi|^2))^{1/2}\|_\infty.$$

Now,

$$\begin{aligned}\|(\mathcal{E}_m(\sum_{k=\max\{m,n\}}^\infty |\Delta_k\psi|^2))^{1/2}\|_\infty &= \|(\mathcal{E}_m(\mathcal{E}_{\max\{m,n\}}(\sum_{k=\max\{m,n\}}^\infty |\Delta_k\psi|^2)))^{1/2}\|_\infty \\ &\leq \|\mathcal{E}_{\max\{m,n\}}(\sum_{k=\max\{m,n\}}^\infty |\Delta_k\psi|^2)\|_\infty^{1/2}.\end{aligned}$$

Consequently,

$$\|\psi - \mathcal{E}_n\psi\|_{\mathrm{BMO}} \le \sup_{m\ge n} \| \left(\mathcal{E}_m|\psi - \mathcal{E}_m\psi|^2\right)^{1/2} \|_\infty.$$

It follows that $S_{2^n}\psi \to \psi$ in BMO norm as $n \to \infty$ for every $\psi \in \mathrm{VMO}$.

THEOREM 7. [FEFFERMAN'S INEQUALITY] *If $f \in \mathrm{H}$ and $\psi \in \mathrm{BMO}$ then*

$$\left|\int_0^1 \mathcal{E}_m f \mathcal{E}_m \psi\right| \le 25\|f\|_{\mathrm{H}}\|\psi\|_{\mathrm{BMO}}$$

for all $m \in \mathbf{N}$.

PROOF. Fix $\psi \in \mathrm{BMO}$ and let β be a nontrivial dyadic atom supported on an interval I. Choose $N \in \mathbf{N}$ such that $|I| = 2^{-N}$. Since β is of mean zero we have $\mathcal{E}_n\beta = 0$ for $n \le N$. Hence

$$\int_0^1 \mathcal{E}_m\beta\mathcal{E}_N\psi = \int_0^1 \psi\mathcal{E}_{\min\{N,m\}}\beta = 0$$

for all $m \in \mathbf{N}$. It follows from (46) that

$$\begin{aligned}\left|\int_0^1 \mathcal{E}_m\beta\mathcal{E}_m\psi\right| &= \left|\int_0^1 (\mathcal{E}_m\beta)(\psi - \mathcal{E}_N\psi)\right| \\ &\le \frac{1}{|I|}\int_I |\psi - \mathcal{E}_N\psi| \qquad (m \in \mathbf{N}).\end{aligned}$$

Thus

$$\left|\int_0^1 \mathcal{E}_m\beta\mathcal{E}_m\psi\right| \le \|\psi\|_{\mathrm{BMO}} \tag{51}$$

for all $m \in \mathbf{N}$ and all dyadic atoms β.

Suppose $f \in \mathrm{H}$ has atomic decomposition

$$f = \sum_{n=0}^{\infty} a_n\beta_n.$$

We have by (51) that

$$\begin{aligned}\left|\int_0^1 \mathcal{E}_m f\mathcal{E}_m\psi\right| &\le \sum_{n=0}^{\infty} |a_n| \left|\int_0^1 \mathcal{E}_m\beta_n\mathcal{E}_m\psi\right| \\ &\le \sum_{n=0}^{\infty} |a_n|\,\|\psi\|_{\mathrm{BMO}} \qquad (m \in \mathbf{N}).\end{aligned}$$

Therefore Fefferman's inequality follows from Theorem 6. ∎

The product of an H function with a BMO function may not be integrable (see Exercise 3.16). Nevertheless, we can define an inner product between these spaces by

$$\langle f, \psi \rangle := \lim_{m\to\infty} \int_0^1 \mathcal{E}_m f \mathcal{E}_m \psi.$$

Notice by Hölder's inequality that

$$\begin{aligned} | \int_0^1 (\mathcal{E}_m f \mathcal{E}_m \psi - \mathcal{E}_n f \mathcal{E}_n \psi) | &= | \int_0^1 \mathcal{E}_m (f - \mathcal{E}_n f) \mathcal{E}_m \psi | \\ &\le \|\psi\|_{\mathrm{BMO}} \|f - \mathcal{E}_n f\|_{\mathrm{H}}, \end{aligned}$$

for $m \ge n \ge 0$. Since this last factor converges to zero when $f \in \mathrm{H}$, it is evident that the inner product is defined for all $f \in \mathrm{H}$ and $\psi \in \mathrm{BMO}$. Since Fefferman's inequality implies

$$| \langle f, \psi \rangle | \le 25 \|\psi\|_{\mathrm{BMO}} \|f\|_{\mathrm{H}},$$

the map $(f, \psi) \to \langle f, \psi \rangle$ is evidently a bounded, bilinear functional from $\mathrm{H} \times \mathrm{BMO}$ into $\mathbf{R}$. In particular, the maps $f \to \langle f, \psi \rangle$ for each $\psi \in \mathrm{BMO}$, and $\psi \to \langle f, \psi \rangle$ for each $f \in \mathrm{H}$ are bounded linear functionals on H and BMO, respectively. In the next section we shall take up the converse of this statement.

Concerning the growth of Walsh-Fourier coefficients of functions in H and BMO, the following is true.

THEOREM 8.

i) There is an absolute constant A such that

$$\sum_{k=1}^{\infty} \frac{|\hat{f}(k)|}{k} \le A \|f\|_{\mathrm{H}}$$

for all $f \in \mathrm{H}$.

ii) If $(c_k, k \in \mathbf{N})$ is a sequence of real numbers which satisfies

$$M := \sup_{n \in \mathbf{N}} \left(2^n \sum_{k=2^n}^{\infty} |c_k|^2 \right)^{1/2} < \infty$$

(in particular, if $c_k = O(1/k)$ as $k \to \infty$), then there is a function $\psi \in \mathrm{BMO}$ such that $\widehat{\psi}(k) = c_k$ for $k \in \mathbf{P}$ and $\|\psi\|_{\mathrm{BMO}} \le M$.

PROOF. Let $\beta \ne 1$ be a dyadic atom supported on some $I \in \mathcal{I}_0$. Choose $n \in \mathbf{N}$ such that $|I| = 2^{-n}$ and observe that $\widehat{\beta}(k) = 0$ for $0 \le k < 2^n$. Hence Bessel's inequality and (45) imply

$$\begin{aligned} \sum_{k=1}^{\infty} \frac{|\widehat{\beta}(k)|}{k} &\le \left(\sum_{k=2^n}^{\infty} \frac{1}{k^2} \right)^{1/2} \left(\sum_{k=2^n}^{\infty} |\widehat{\beta}(k)|^2 \right)^{1/2} \\ &\le \sqrt{\frac{2}{2^n}} \|\beta\|_2 \\ &\le \sqrt{2}. \end{aligned}$$

Suppose that $f \in \mathrm{H}$ with atomic decomposition

$$f = \sum_{n=0}^{\infty} a_n \beta_n.$$

The inequality above implies

$$\sum_{k=1}^{\infty} \frac{|\widehat{f}(k)|}{k} \le \sum_{n=0}^{\infty} |a_n| \sum_{k=1}^{\infty} \frac{|\widehat{\beta}(k)|}{k} \le \sqrt{2} \sum_{n=0}^{\infty} |a_n|.$$

Thus i) follows from Theorem 6 with $A = 25\sqrt{2}$.

To prove ii) use the Riesz-Fischer theorem to choose $\psi \in \mathrm{L}^2$ such that $\widehat{\psi}(k) = c_k$ for $k \in \mathbf{P}$ and $\widehat{\psi}(0) = 0$. Set

$$P_\ell^{(n)} := \sum_{k=0}^{2^n - 1} \widehat{\psi}(\ell 2^n + k) w_k \qquad (\ell \in \mathbf{P}, n \in \mathbf{N}),$$

and observe that

$$\psi - \mathcal{E}_n \psi = \sum_{\ell=1}^{\infty} P_\ell^{(n)} w_{\ell 2^n}$$

for $n \in \mathbf{N}$. Since each $P_\ell^{(n)}$ belongs to $\mathrm{L}(\mathcal{A}^n)$ and

$$\mathcal{E}_n(w_{\ell 2^n} w_{s 2^n}) = 0$$

for $\ell \neq s, \ell, s \in \mathbf{P}$, it follows that

$$\begin{aligned} \mathcal{E}_n |\psi - \mathcal{E}_n \psi|^2 &= \sum_{\ell=1}^{\infty} |P_\ell^{(n)}|^2 \\ &\le 2^n \sum_{k=2^n}^{\infty} |\widehat{\psi}(k)|^2. \end{aligned}$$

By hypothesis we conclude that $\psi \in \mathrm{BMO}$ and $\|\psi\|_{\mathrm{BMO}} \le M$. ∎

Let $\imath$ represent $\sqrt{-1}$. In the classical case, each Hardy space consists of functions F, analytic on the unit disc, which satisfy

$$\|F\|_{\mathcal{H}^p} := \sup_{0 \le r < 1} \left(\frac{1}{2\pi} \int_0^{2\pi} |F(re^{\imath\theta})|^p \, d\theta \right)^{1/p} < \infty.$$

This condition is equivalent to L^p integrability of the non- tangential maximal function

$$F_s^*(e^{\imath\theta}) := \sup_{z \in \Gamma_s(\theta)} |F(z)|,$$

where for each $0 < s < 1$ and $\theta \in [0, 2\pi)$, $\Gamma_s(\theta)$ represents the convex hull of the sets $\{e^{i\theta}\}$ and $\{z \in \mathbf{C} : |z| \leq s\}$ (see Garnet [1], for example). Thus the role that the non-tangential maximal function plays in the classical case is played by the dyadic maximal function $\mathcal{E}^* f$ in the dyadic case. Besides this analogy, it is natural to ask whether there is a more direct connection between the classical Hardy spaces and the dyadic ones.

By taking real parts of the boundary functions

$$\lim_{r \uparrow 1} F(re^{i\theta}) \qquad (0 \leq \theta < 2\pi)$$

and identifying the boundary of the unit disc with the interval [0,1), one generates the classical real Hardy spaces $\mathcal{H}^p$, for $0 < p < \infty$. As in the dyadic case, $\mathcal{H}^p$ and L^p are isomorphic for $1 < p < \infty$, and $\mathcal{H} := \mathcal{H}^1$ is a proper subspace of L^1. Moreover, $\mathcal{H}$ has an atomic characterization just like that given for H in Theorem 6 (see Coifman and Weiss [1] for details). The essential difference is that an atom for $\mathcal{H}$ can be supported on non-dyadic intervals.

We see that every dyadic atom is an atom for $\mathcal{H}$ but not conversely. Hence

$$\mathrm{H} \subset \mathcal{H}.$$

On the other hand, almost every translate of an $f \in \mathcal{H}$ belongs to the dyadic space H:

THEOREM 9. *Extend each $f \in \mathcal{H}$ from $[0,1)$ to $\mathbf{R}$ by periodicity of period 1. For each $\theta \in [0,1)$ set*

$$f_\theta(x) := f(x - \theta) \qquad (x \in [0,1)).$$

There is an absolute constant $C > 0$ such that

$$\int_0^1 \int_0^1 (\mathcal{E}^* f_\theta)(x)\, dx\, d\theta \leq C \|f\|_{\mathcal{H}}$$

for all $f \in \mathcal{H}$.

PROOF. By the classical analogue of Theorem 6 it suffices to show

$$\int_0^1 \int_0^1 (\mathcal{E}^* f_\theta)(x)\, dx\, d\theta \leq 8$$

for each atom f of $\mathcal{H}$.

Let f be an atom of $\mathcal{H}$ supported on an interval I. Choose $N \in \mathbf{N}$ such that

$$2^{-(N+1)} < |I| \leq 2^{-N}.$$

(Remember, I is not necessarily dyadic). Let $\widetilde{I} := \{e^{2\pi i t} : t \in I\}$. For each $\theta \in [0,1)$ let $\widetilde{I}_\theta := e^{-2\pi i\theta}\widetilde{I}$ and let $\widetilde{J}_\theta$ denote the arc in $\{e^{2\pi i t} : t \in [0,1)\}$ which has the same center as $\widetilde{I}_\theta$ but is three times larger. Set

$$I_\theta := \{t \in [0,1) : e^{2\pi i t} \in \widetilde{I}_\theta\}$$

and

$$J_\theta := \{t \in [0,1) : e^{2\pi i t} \in \widetilde{J}_\theta\}.$$

Suppose ω is a dyadic interval with $|\omega| = 2^{-n}$ for some $n \in \mathbf{N}$. If $n > N$ then the choice of N implies $|\omega| < |I|$. Hence, if $\omega \cap I_\theta \neq \emptyset$ then $\omega \subset J_\theta$. It follows from the definition of atoms that

$$\begin{aligned}|(\mathcal{E}_n f_\theta)(t)| &= \frac{1}{|\omega|}\Big|\int_\omega f_\theta(x)\,dx\Big|\\ &\leq \frac{1}{|I|}\chi(J_\theta)(t)\end{aligned}$$

for each $t \in \omega$. In particular,

$$\int_0^1 \int_0^1 \sup_{n>N} |(\mathcal{E}_n f_\theta)(t)|\,dt\,d\theta \leq \int_0^1 \frac{|J_\theta|}{|I|}\,d\theta = 3.$$

On the other hand, if $n \leq N$ then $\omega = I(k,n)$ for some $0 \leq k < 2^n$, and $|\omega| \geq |I|$. Since $\int_\omega f_\theta = 0$ when $\omega \cap I_\theta = \emptyset$ or $\omega \supset I_\theta$ we may suppose that $\omega \cap I_\theta \neq \emptyset$ and I_θ is not a subset of ω. The fact that ω is a dyadic interval implies

$$|\{\theta \in [0,1) : \omega \cap I_\theta \neq \emptyset \quad \text{and } I_\theta \text{ is not a subset of } \omega\}| \leq 2|I|.$$

Moreover, the fact that f_θ is supported on I_θ implies

$$\Big|\int_\omega f_\theta\Big| \leq \frac{|\omega \cap I_\theta|}{|I|} \leq 1.$$

It follows, therefore, that

$$\int_0^1 \int_0^1 \frac{1}{|\omega|}\left|\int_\omega f_\theta(t)\right| \chi(\omega)(t)\,dt\,d\theta \leq 2|I|.$$

But

$$|\mathcal{E}_n f_\theta| = \sum_{k=0}^{2^n-1} \frac{1}{|I(k,n)|}\left|\int_{I(k,n)} f_\theta(t)\,dt\right| \chi(I(k,n)).$$

Consequently,

$$\begin{aligned}\int_0^1 \int_0^1 \sup_{n\leq N} |(\mathcal{E}_n f_\theta)(t)|\,dt\,d\theta &\leq \sum_{n=0}^{N} \int_0^1 \int_0^1 |(\mathcal{E}_n f_\theta)(t)|\,dt\,d\theta\\ &\leq \sum_{n=0}^{N} 2|I|2^n\\ &< 4|I|2^N.\end{aligned}$$

Since $|I|2^N \leq 1$, the proof of the theorem is complete. ∎

In Chapter 5 we shall prove that as Banach spaces, $\mathcal{H}$ and H are isomorphic. To lay some foundations for this we introduce the sequence spaces $\mathbf{h}, \mathbf{h}_0$, **bmo**, and **vmo**.

Index the Haar system $\boldsymbol{h}$ by $\mathcal{I}$ in the following way. Set $h_\emptyset := 1$. For $I \in \mathcal{I}_0$ let h_I denote the Haar function h_n which satisfies $\{h_n \neq 0\} = I$. Denote the Haar-Fourier coefficients of an $f \in \mathrm{L}^1$ by

$$\dot{f}_I := \int_0^1 f h_I \qquad (I \in \mathcal{I})$$

and the partial sums of the Haar-Fourier series of f by

$$S_m^{\boldsymbol{h}} f := \sum_{k=0}^{m-1} \left(\int_0^1 f(t) h_k(t)\, dt \right) h_k$$

for $m \in \mathbf{P}$. By (42) in 1.4, $S_{2^n}^{\boldsymbol{h}} f = S_{2^n} f$ for all $n \in \mathbf{N}$. Hence it is obvious that

$$\chi(I)|\Delta_n f|^2 = |\dot{f}_I|^2 \frac{\chi(I)}{|I|} \qquad (I \in \mathcal{I}, |I| = 2^{-n}, n \in \mathrm{N}).$$

Therefore, the square function can be written in the form

$$(\mathcal{Q}f)(x) = \left(|\dot{f}_\emptyset|^2 + \sum_{I \in \mathcal{I}, I \supset \{x\}} \frac{|\dot{f}_I|^2}{|I|} \right)^{1/2}$$

and

$$\|f\|_{\mathrm{BMO}} = \sup_{I \in \mathcal{I}_0} \left(\frac{1}{|I|} \sum_{J \in \mathcal{I}, J \subseteq I} |\dot{f}_J|^2 \right)^{1/2}$$

for any $f \in \mathrm{L}^1$ of mean zero.

These observations lead naturally to the following definitions. Define the square function of a sequence $\boldsymbol{b} = (b_I, I \in \mathcal{I})$, indexed by $\mathcal{I}$, by

$$(\mathcal{Q}\boldsymbol{b})(x) := \left(|b_\emptyset|^2 + \sum_{I \in \mathcal{I}, I \supset \{x\}} \frac{|b_I|^2}{|I|} \right)^{1/2} \qquad (x \in [0,1)).$$

Let $\mathbf{h}$ represent the collection of sequences $\boldsymbol{b}$ which satisfy

$$\|\boldsymbol{b}\|_{\mathbf{h}} := \|\mathcal{Q}\boldsymbol{b}\|_1 < \infty,$$

let **bmo** denote the collection of sequences $\boldsymbol{b}$ which satisfy $b_\emptyset = 0$ and

$$\|\boldsymbol{b}\|_{\mathbf{bmo}} := \sup_{I \in \mathcal{I}_0} \left(\frac{1}{|I|} \sum_{J \in \mathcal{I}, J \subseteq I} |b_J|^2 \right)^{1/2} < \infty,$$

let $\mathbf{h}_0$ represent the collection of $\boldsymbol{b} \in \mathbf{h}$ which satisfy $b_\emptyset = 0$, and let **vmo** represent the collection of $\boldsymbol{b} \in$ **bmo** which satisfy

$$\lim_{|I|\to 0} \left(\frac{1}{|I|} \sum_{J\in\mathcal{I}, J\subseteq I} |b_J|^2 \right)^{1/2} = 0. \tag{52}$$

Let H_0 (respectively, $\mathcal{H}_0$) represent the collection of functions in H (respectively, $\mathcal{H}$) which are of mean zero. The definitions above imply that the map $f \to \dot{f}$ is a Banach space isomorphism from H onto **h**, H_0 onto $\mathbf{h}_0$, BMO onto **bmo**, and VMO onto **vmo**. In 5.4 we shall carry out a similar program for the Franklin-Fourier coefficient map and the classical spaces $\mathcal{H}, \mathcal{H}_0, \mathcal{BMO}$, and $\mathcal{VMO}$. In particular (see Corollary 4 in 5.5), we will show that these classical spaces and their dyadic counterparts are isomorphic as Banach spaces.

3.5 Duality in Dyadic Hardy Spaces. Given a Banach space $\boldsymbol{X}$, let $\boldsymbol{X}'$ represent its dual, i.e., $\boldsymbol{X}'$ represents the collection of bounded linear functionals on $\boldsymbol{X}$. Let $\|F\|$ represent the functional norm of an $F \in \boldsymbol{X}'$. In this section we show that H_0' is isometric and homeomorphic to BMO and VMO$'$ is isometric and homeomorphic to H_0. It will follow that $\mathbf{h}_0'$, **bmo** and $\mathbf{vmo}', \mathbf{h}_0$ are also isomorphic and homeomorphic pairs.

THEOREM 10.

i) Let $F : \mathrm{H}_0 \to \mathbf{R}$ *be a bounded linear functional. There exists a unique* $\psi \in$ BMO *such that*

$$F(f) = \langle f, \psi \rangle \qquad (f \in \mathrm{H}_0)$$

and

$$\frac{1}{4}\|\psi\|_{\mathrm{BMO}} \le \|F\| \le 25\|\psi\|_{\mathrm{BMO}}.$$

ii) Let $F :$ VMO $\to \mathbf{R}$ *be a bounded linear functional. There exists a unique* $f \in \mathrm{H}_0$ *such that*

$$F(\psi) = \langle f, \psi \rangle \qquad (\psi \in \mathrm{VMO})$$

and

$$\frac{1}{3}\|f\|_{\mathrm{H}} \le \|F\| \le 25\|f\|_{\mathrm{H}}.$$

PROOF. Fix $F \in \mathrm{H}_0'$. Consider the 1-1 linear map $\boldsymbol{\Delta} : \mathrm{H}_0 \to \mathrm{L}^1(\ell^2)$ defined by

$$\boldsymbol{\Delta} f := (\Delta_n f, n \in \mathrm{N}).$$

(See Appendix 0.2 for information on mixed norm spaces.) By Corollary 4 in 3.3, the map $F \circ \boldsymbol{\Delta}^{-1}$ is a bounded linear functional on the subspace $\boldsymbol{\Delta}(\mathrm{H}_0)$ of $\mathrm{L}^1(\ell^2)$).

Extend $F \circ \boldsymbol{\Delta}^{-1}$ to $\mathrm{L}^1(\ell^2)$ by the Hahn- Banach theorem and apply the Riesz representation theorem to this extension. Namely, choose a sequence $\boldsymbol{\lambda} = (\lambda_n, n \in \mathbf{N}) \in \mathrm{L}^\infty(\ell^2)$ such that

$$\|F\| = \|F \circ \boldsymbol{\Delta}^{-1}\| = \|\boldsymbol{\lambda}\|_{L^\infty(\ell^2)}, \tag{53}$$

and

$$(F \circ \boldsymbol{\Delta}^{-1})(\boldsymbol{g}) = \sum_{n=0}^{\infty} \int_0^1 g_n \lambda_n$$

for all $\boldsymbol{g} = (g_n, n \in \mathbf{N}) \in \mathrm{L}^1(\ell^2)$. Specializing to $\boldsymbol{g} := \boldsymbol{\Delta} f$ we obtain

$$F(f) = \sum_{n=0}^{\infty} \int_0^1 (\Delta_n f)\lambda_n.$$

In particular, by (3) in 3.1

$$F(f) = \sum_{n=0}^{\infty} \int_0^1 f \Delta_n \lambda_n.$$

for all $f \in \mathrm{H}$.

The series $\sum_{n=0}^{\infty} \Delta_n \lambda_n$ converges in L^2 norm to a function ψ which belongs to BMO. To see this, notice by the conditional Hölder inequality that

$$|\Delta_k \lambda_k|^2 \leq 2(\mathcal{E}_{k+1}|\lambda_k|^2 + \mathcal{E}_k|\lambda_k|^2)$$

holds for any $\lambda_k \in \mathrm{L}^2, k \in \mathbf{N}$. Consequently,

$$\mathcal{E}_s(|\sum_{k=n}^{m} \Delta_k \lambda_k|^2) = \mathcal{E}_s(\sum_{k=n}^{m} |\Delta_k \lambda_k|^2) \leq 4\mathcal{E}_s(\sum_{k=n}^{m} |\lambda_k|^2) \tag{54}$$

for all integers $0 \leq s \leq n \leq m$. Using (54) for $s = 0$ we see that the series in question converges in L^2 norm, say to ψ. Using (54) for $s = n$ and letting $m \to \infty$ we see that

$$\begin{aligned} \mathcal{E}_n(|\psi - \mathcal{E}_n\psi|^2) &\leq 4\mathcal{E}_n(\sum_{k=n}^{m} |\lambda_k|^2) \\ &\leq 4\|\boldsymbol{\lambda}\|^2_{L^\infty(\ell^2)}. \end{aligned}$$

Hence

$$\|\psi\|_{\mathrm{BMO}} \leq 4\|F\|, \tag{55}$$

by (53).

To see that ψ represents F, fix $f \in \mathrm{L}(\mathcal{A}^N)$ and $m \geq N$ in $\mathbf{N}$. Then (3) in 3.1 implies

$$\begin{aligned} F(f) &= \sum_{n=0}^{N} \int_0^1 f \Delta_n \lambda_n \\ &= \int_0^1 f(\sum_{n=0}^{m} \Delta_n \lambda_n) \\ &= \int_0^1 (\mathcal{E}_m f)(\mathcal{E}_m \psi). \end{aligned}$$

Let $m \to \infty$. We obtain $F(f) = \langle f, \psi \rangle$ for all $f \in \mathcal{P}$. Since $\mathcal{P}$ is dense in H, we conclude that ψ represents F. In particular, by Fefferman's inequality $\|F\| \le 25\|\psi\|_{\mathrm{BMO}}$ and the proof of i) is complete.

To prove ii), fix $n \in \mathbf{N}$ and set $Y_n := X_n := \mathrm{L}^2$. Define

$$\|\xi\|_{X_n} := \|(\mathcal{E}_n|\xi|^2)^{1/2}\|_\infty$$

for $\xi \in X_n$ and

$$\|\lambda\|_{Y_n} := \|(\mathcal{E}_n|\lambda|^2)^{1/2}\|_1$$

for $\lambda \in Y_n$. Notice for each pair $\xi \in X_n$, $\lambda \in Y_n$ that

$$\begin{aligned} \left|\int_0^1 \xi\lambda\right| &= \left|\int_0^1 \mathcal{E}_n(\xi\lambda)\right| \\ &\le \|\xi\|_{X_n}\|\lambda\|_{Y_n}. \end{aligned}$$

Thus it follows from the Riesz representation theorem on L^2 that each bounded linear functional on X_n looks like

$$\int_0^1 \xi\lambda \qquad (\xi \in X_n)$$

and has functional norm $\|\lambda\|_{Y_n}$, for some $\lambda \in Y_n$. (See also Exercise 3.20.)

Let $\boldsymbol{X} := X_0 \times X_1 \times \dots$, and set

$$\|\boldsymbol{\xi}\|_{\boldsymbol{X}} := \sup_{n\in\mathbf{N}} \|\xi_n\|_{X_n}$$

for $\boldsymbol{\xi} = (\xi_n, n \in \mathbf{N}) \in \boldsymbol{X}$. Let

$$\boldsymbol{X}_0^\infty := \{\boldsymbol{\xi} \in \boldsymbol{X} : \lim_{n\to\infty} \|\xi_n\|_{X_n} = 0\}.$$

Then (see Appendix 0.0), each bounded linear functional $\Lambda \in (\boldsymbol{X}_0^\infty)'$ has the following form

$$\Lambda(\xi) = \sum_{n=0}^\infty \int_0^1 \xi_n\lambda_n \qquad (\boldsymbol{\xi} = (\xi_n, n \in \mathbf{N}) \in \boldsymbol{X}_0^\infty),$$

and satisfies

$$\|\Lambda\| = \sum_{n=0}^\infty \|(\mathcal{E}_n|\lambda_n|^2)^{1/2}\|_1,$$

for some functions $\lambda_n \in \mathrm{L}^2$.

Fix $F \in \mathrm{VMO}'$ and consider the map

$$\boldsymbol{\Upsilon}\psi := (\psi - \mathcal{E}_n\psi, n \in \mathbf{N}).$$

Clearly, $\boldsymbol{\Upsilon} : \mathrm{VMO} \to \boldsymbol{X}_0^\infty$ and $\|\psi\|_{\mathrm{BMO}} = \|\boldsymbol{\Upsilon}\psi\|_{\boldsymbol{X}}$ for all $\psi \in \mathrm{BMO}$. Thus $F \circ \boldsymbol{\Upsilon}^{-1}$ is a bounded linear functional on the subspace $\boldsymbol{\Upsilon}(\mathrm{VMO})$ of $\boldsymbol{X}_0^\infty$.

Extend $F\circ\Upsilon^{-1}$ to $\boldsymbol{X}_0^\infty$ by the Hahn-Banach theorem. Apply the representation theorem cited above to choose functions $\lambda_n \in \mathrm{L}^2$ such that

$$\|F\| = \|F\circ\Upsilon^{-1}\| = \sum_{n=0}^{\infty} \|(\mathcal{E}_n|\lambda_n|^2)^{1/2}\|_1$$

and

$$F(\psi) = \sum_{n=0}^{\infty}\int_0^1 \lambda_n(\psi - \mathcal{E}_n\psi) \qquad (\psi \in \mathrm{VMO}).$$

Since

$$\int_0^1 \lambda_n(\psi - \mathcal{E}_n\psi) = \int_0^1 (\lambda_n - \mathcal{E}_n\lambda_n)\psi \qquad (n \in \mathbf{N}),$$

the representation of F above can be rewritten as

(56)
$$F(\psi) = \sum_{n=0}^{\infty}\int_0^1 (\lambda_n - \mathcal{E}_n\lambda_n)\psi \qquad (\psi \in \mathrm{VMO}).$$

We shall prove that the series $\sum_{n=0}^{\infty}(\lambda_n - \mathcal{E}_n\lambda_n)$ converges in L^1 norm to a function f which satisfies

(57)
$$\|f\|_{\mathrm{H}} \le 3\|F\|.$$

Indeed, the choice of the λ_n's implies

$$\begin{aligned}\sum_{n=0}^{\infty}\|\mathcal{E}_n\lambda_n\|_1 &\le \sum_{n=0}^{\infty}\|\lambda_n\|_1\\ &\le \sum_{n=0}^{\infty}\|(\mathcal{E}_n(|\lambda_n|^2))^{1/2}\|_1\\ &= \|F\|.\end{aligned}$$

Hence the series in question converges in L^1 norm, say to f. The identity

$$\mathcal{E}_m f = \sum_{n=0}^{m-1}\mathcal{E}_m\lambda_n - \sum_{n=0}^{m-1}\mathcal{E}_n\lambda_n \qquad (m \in \mathbf{P})$$

implies

(58)
$$\mathcal{E}^* f \le \sum_{n=0}^{\infty}\sup_{m\ge n}|\mathcal{E}_m\lambda_n| + \sum_{n=0}^{\infty}|\mathcal{E}_n\lambda_n|.$$

Now, since $\mathcal{E}_n\lambda_n$ is an average of λ_n, we have

$$\|\mathcal{E}_n\lambda_n\|_1 \le \|(\mathcal{E}_n(|\lambda_n|^2))^{1/2}\|_1.$$

Moreover, by Hölder's inequality and Corollary 1 in 3.1 (with $p = 2$), we have

$$\| \sup_{m\geq n} |\mathcal{E}_m \lambda_n| \|_1 = \| \mathcal{E}_n(\sup_{m\geq n} |\mathcal{E}_m \lambda_n|) \|_1$$
$$\leq 2\|(\mathcal{E}_n(|\lambda_n|^2))^{1/2}\|_1$$

for each $n \in \mathbf{N}$. Therefore, it follows from (58) that

$$\|\mathcal{E}^* f\|_1 \leq 3 \sum_{n=0}^{\infty} \|(\mathcal{E}_n(|\lambda_n|^2))^{1/2}\|_1 .$$

This verifies (57).

Finally, since $\mathcal{P}$ is dense in VMO it remains to show that

$$F(\psi) = \langle f, \psi \rangle \tag{59}$$

for all $\psi \in \mathcal{P}$. But (56) is finite when ψ is a Walsh polynomial. Consequently,

$$F(\psi) = \lim_{m\to\infty} \sum_{n=0}^{m} \int (\lambda_n - \mathcal{E}_n \lambda_n)\psi = \langle f, \psi \rangle . \quad \blacksquare$$

The proof of Theorem 10 contains the following information concerning the structure of H and BMO.

COROLLARY 9. *Let $f \in \mathrm{H}_0$. There exist $\lambda_n \in \mathrm{L}^2$ for $n \in \mathbf{N}$ such that*

$$f = \sum_{n=0}^{\infty} (\lambda_n - \mathcal{E}_n \lambda_n),$$

and

$$\frac{1}{25} \sum_{n=0}^{\infty} \|(\mathcal{E}_n(|\lambda_n|^2))^{1/2}\|_1 \leq \|f\|_{\mathrm{H}} \leq 3 \sum_{n=0}^{\infty} \|(\mathcal{E}_n(|\lambda_n|^2))^{1/2}\|_1 < \infty.$$

For the next several pages we shall use the notation $\Delta_{-1} f := \widehat{f}(0)$.

COROLLARY 10. *Let $\psi \in \mathrm{L}^0$ with*

$$\|\psi\|_{\mathrm{BMO}} < \infty.$$

There exists a sequence $\boldsymbol{\lambda} = (\lambda_n, n \in \mathbf{N}) \in \mathrm{L}^\infty(\ell^2)$ such that

$$\psi = \sum_{n=0}^{\infty} \Delta_{n-1} \lambda_n$$

and

$$\frac{1}{25} \|\boldsymbol{\lambda}\|_{L^\infty(\ell^2)} \leq \|\psi\|_{\mathrm{BMO}} \leq 4 \|\boldsymbol{\lambda}\|_{L^\infty(\ell^2)} .$$

(Note: If $\psi \in \mathrm{BMO}$ then since ψ is of mean zero the function λ_0 above can be taken to be identically zero.)

We close this section with a few remarks about H^p for $0 < p < 1$. Since these spaces are not used in subsequent chapters, the details are left to the reader.

A function $\beta \in \mathrm{L}^\infty$ is called a p-atom if $\beta = 1$ or if there is a dyadic interval I such that

$$\{\beta \neq 0\} \subseteq I,$$

$$\|\beta\|_\infty \leq |I|^{-1/p},$$

and

$$\int_0^1 \beta = 0.$$

Straightforward modifications of the proof of Theorem 6 lead to the following atomic characterization of H^p for $0 < p < 1$.

THEOREM 11. *Let $0 < p < 1$. A function f belongs to H^p if and only if*

$$f = \sum_{n=0}^{\infty} a_n \beta_n$$

for some p-atoms $\beta_0, \beta_1, \ldots$ and a sequence $\boldsymbol{a} = (a_n, n \in \mathbf{N}) \in \ell^p$, where this series converges in the L^p quasi-norm.

For each $\alpha \geq 0$ and each $g \in \mathrm{L}^1$, let

$$\|g\|_{Lip\alpha} := \sup_{I \in \mathcal{I}_0} \frac{1}{|I|^{1+\alpha}} \int_I |g - \frac{1}{|I|} \int_I g|,$$

and denote by $Lip\alpha$ the collection of functions g of mean zero which satisfy $\|g\|_{Lip\alpha} < \infty$. Notice that $Lip0$ is identical with BMO.

The argument of Theorem 7 can be used to show that there is an absolute constant C such that

$$|\int_0^1 \mathcal{E}_m f \mathcal{E}_m g| \leq C \|f\|_{H^p} \|g\|_{Lip\alpha}$$

for $f \in \mathrm{H}^p, g \in Lip\alpha,\ 0 < p < 1, m \in \mathbf{N}$, and $\alpha := (1/p) - 1$. This inequality can be used to extend Theorem 10. Namely, one can show that $(\mathrm{H}^p)'$ is essentially $Lip((1/p) - 1)$.

3.6 Interpolation in Dyadic Hardy Spaces. In this section we prove an interpolation theorem for dyadic Hardy spaces which will be used in Chapter 5 to study bases in L^1.

For each $2 \leq q \leq \infty$ let K^q denote the collection of functions $g \in \mathrm{L}^2$ which can be represented as

$$g = \sum_{n=0}^{\infty} \Delta_{n-1} \lambda_n \tag{60}$$

for some sequence $\boldsymbol{\lambda} = (\lambda_n, n \in \mathrm{N}) \in \mathrm{L}^q(\ell^2)$. Notice that such a representation will converge in L^2 norm and a.e. on [0,1) for each $\boldsymbol{\lambda} \in \mathrm{L}^q(\ell^2)$. Hence K^q is a linear subspace of L^2.

For each $g \in \mathrm{K}^q$, define

$$\|g\|_{K^q} := \inf\{\|\boldsymbol{\lambda}\|_{L^q(\ell^2)}\},$$

where this infimum is taken over all sequences $\boldsymbol{\lambda} = (\lambda_n, n \in \mathrm{N}) \in \mathrm{L}^q(\ell^2)$ which satisfy (60). It is easy to check that this is a norm on K^q for each $2 \leq q \leq \infty$. By Corollary 10 in 3.5, this norm is equivalent to the BMO norm when $q = \infty$. In particular, K^∞ and BMO are essentially isomorphic as Banach spaces.

When $2 \leq q < \infty$, the norms of K^q and H^q are equivalent. Indeed, if $g \in \mathrm{H}^q$ then the sequence $\boldsymbol{\lambda} = (\Delta_{n-1} g, n \in \mathrm{N})$ satisfies

$$\|\mathcal{Q}g\|_q = \|\boldsymbol{\lambda}\|_{L^q(\ell^2)}.$$

Thus $\mathrm{H}^q \subseteq \mathrm{K}^q$ and $\|g\|_{K^q} \leq \|\mathcal{Q}g\|_q$. On the other hand, if $g \in \mathrm{K}^q$ and $\boldsymbol{\lambda}$ satisfy (60) then by Corollary 2 in 3.1 we have

$$\begin{aligned}\|\mathcal{Q}g\|_q &\leq \left(\int_0^1 (2\sum_{n=0}^{\infty}(|\mathcal{E}_n\lambda_n|^2 + |\mathcal{E}_{n-1}\lambda_n|^2))^{q/2}\right)^{1/q} \\ &\leq 4q\left(\int_0^1 (\sum_{n=0}^{\infty}|\lambda_n|^2)^{q/2}\right)^{1/q} \\ &= 4q\|\boldsymbol{\lambda}\|_{L^q(\ell^2)}.\end{aligned}$$

Hence $\mathrm{K}^q \subseteq \mathrm{H}^q$ and $\|\mathcal{Q}g\|_q \leq 4q\|g\|_{K^q}$.

We shall prove that one can interpolate for operators between any two of the spaces $\mathrm{L}^p, \mathrm{H}^p$, and K^p. So that our theorem can be stated in compact form, we temporarily introduce the following notation.

If $X = \mathrm{L}$ and $1 \leq p \leq \infty$, then X^p will represent $\mathrm{L}^p(\Omega, \nu)$ for some σ-finite measure space (Ω, ν). If $X = \mathrm{H}$ and $1 \leq p < \infty$, then X^p will represent the dyadic Hardy space H^p. And, if $X = \mathrm{K}$ and $2 \leq p \leq \infty$, then X^p will represent the space K^p defined above. Thus given $X \in \{\mathrm{L}, \mathrm{H}, \mathrm{K}\}$ and certain parameters p, we have defined Banach spaces X^p. We use this notation and the convention that $1/\infty = 0$ in the following result.

THEOREM 12. *Let $X, Y \in \{\mathrm{L}, \mathrm{H}, \mathrm{K}\}$. Suppose $p_0 \neq p_1, q_0 \neq q_1$, $(1 \leq p_0, p_1, q_0, q_1 \leq \infty)$, and*

$$T : X^{p_j} \to Y^{q_j}$$

is a bounded linear operator with operator norm M_j for $j = 0, 1$. If

$$\frac{1}{p} := \frac{1-t}{p_0} + \frac{t}{p_1} \quad \text{and} \quad \frac{1}{q} := \frac{1-t}{q_0} + \frac{t}{q_1}$$

for some $0 < t < 1$ then T is bounded from X^p to Y^q with corresponding operator norm M which satisfies

$$M \leq CM_0^{1-t}M_1^t$$

for some absolute constant C.

PROOF. Let q' denote the index conjugate to q. Since both p and q' are finite, the collection of Walsh polynomials $\mathcal{P}$ is dense in X^p and $Y^{q'}$. Consequently, M is equivalent to

$$\widetilde{M} := \sup\{|\langle Tf, g\rangle| : f, g \in \mathcal{P}, \|f\|_{X^p} \le 1, \|g\|_{Y^{q'}} \le 1\},$$

where $\langle\, ,\, \rangle$ represents the usual inner product or the one on H $\times$ BMO defined in 3.4. Therefore, it suffices to show

$$\widetilde{M} \le \widetilde{C} M_0^{1-t} M_1^t$$

for some absolute constant $\widetilde{C}$. Our proof uses the complex method (see Appendix 0.2).

Denote the infinite vertical strip $\{z \in \mathbf{C} : 0 \le Re(z) \le 1\}$ by Ξ. For each $z \in \Xi$ define functions p_z, q'_z by

$$\frac{1}{p_z} := \frac{1-z}{p_0} + \frac{z}{p_1} \quad \text{and} \quad \frac{1}{q'_z} := \frac{1-z}{q'_0} + \frac{z}{q'_1},$$

where each q'_j is the index conjugate to q_j for $j = 0, 1$. Notice that $p_t = p$ and $q'_t = q'$.

Fix $j = 0$ or 1, $y \in \mathbf{R}$, and $f \in \mathcal{P} \cap X^p$ with $\|f\|_{X^p} \le 1$. We shall introduce a family of functions $(f_z, z \in \Xi)$ which satisfies

$$f_t = f \tag{61}$$

and

$$\|f_{j+iy}\|_{X^{p_j}} \le C_1 \qquad (y \in \mathbf{R}, j = 0, 1), \tag{62}$$

for some absolute constant C_1 which depends neither on f nor y. (Here the symbol i represents $\sqrt{-1}$.)

If $X = \text{L}$ set

$$f_z := |f|^{p/p_z} \operatorname{sgn} f \qquad (z \in \Xi).$$

Clearly, $f_t = f$. Moreover, for the case $p_j \ne \infty$,

$$\begin{aligned} \|f_{j+iy}\|_{X^{p_j}} &\le \left(\int_\Omega (|f|^{p/p_j})^{p_j}\, d\nu\right)^{1/p_j} \\ &= \|f\|_{X^p}^{p/p_j} \\ &\le 1 \end{aligned}$$

by the choice of f. The same inequality holds for the case $p_j = \infty$. Consequently, (62) is valid when $X = \text{L}$.

If $X = \text{H}$ set

$$f_z := \sum_{k \in \mathbf{Z}} 2^{kp/p_z} 2^{-k} f^{(k)} \qquad (z \in \Xi),$$

where $f = \sum_{k\in\mathbf{Z}} f^{(k)}$ is the canonical decomposition of f. Clearly, $f_t = f$, and

$$\mathcal{E}^*(f_z) \le \sum_{k\in\mathbf{Z}} |2^{kp/p_z}| 2^{-k}\mathcal{E}^*(f^{(k)}) \qquad (z \in \Xi).$$

But by Corollary 3 in 3.2,

$$\mathcal{E}^*(f^{(k)}) \le 2^{k+2}\chi\{\widetilde{\mathcal{E}}f > 2^k\} \qquad (k \in \mathbf{Z}).$$

Hence by Abel's transformation we have

$$\begin{aligned}
\mathcal{E}^*(f_{j+iy}) &\le 4\sum_{k\in\mathbf{Z}} 2^{kp/p_j}\chi\{\widetilde{\mathcal{E}}f > 2^k\} \\
&= \frac{4}{2^{p/p_j}-1}\sum_{k\in\mathbf{Z}}(2^{(k+1)p/p_j} - 2^{kp/p_j})\chi\{\widetilde{\mathcal{E}}f > 2^k\} \\
&= \frac{4\cdot 2^{p/p_j}}{2^{p/p_j}-1}\sum_{k\in\mathbf{Z}} 2^{kp/p_j}\chi\{2^k < \widetilde{\mathcal{E}}f \le 2^{k+1}\}.
\end{aligned}$$

Therefore, it follows from (7) in 3.1 that

$$\begin{aligned}
\|f_{j+iy}\|_{H^{p_j}} &= \|\mathcal{E}^*(f_{j+iy})\|_{p_j} \\
&\le \frac{4\cdot 2^{p/p_j}}{2^{p/p_j}-1}\|\widetilde{\mathcal{E}}f\|_p^{p/p_j} \\
&\le C_1\|f\|_{H^p}^{p/p_j}.
\end{aligned}$$

Thus (62) holds for $X = \mathrm{H}$.

If $X = \mathrm{K}$ set

$$f_z := \sum_{n=0}^{\infty} \Delta_{n-1}(\alpha_{\boldsymbol{\lambda}}^{p/p_z-1}\lambda_n)$$

where

$$\alpha_{\boldsymbol{\lambda}} := (\sum_{n=0}^{\infty} |\lambda_n|^2)^{1/2}$$

and $\boldsymbol{\lambda} = (\lambda_n, n \in \mathrm{N}) \in \mathrm{L}^p(\ell^2)$ satisfies (60) for $g := f$. Clearly, $f_t = f$. Moreover, for the case $p_j \ne \infty$, the construction of $\boldsymbol{\lambda}$ implies

$$\begin{aligned}
\|f_{j+iy}\|_{X^{p_j}} &\le \|\alpha_{\boldsymbol{\lambda}}^{p/p_z-1}\boldsymbol{\lambda}\|_{L^{p_j}(\ell^2)} \\
&= (\int_0^1 \alpha_{\boldsymbol{\lambda}}^p)^{1/p_j} \\
&= \|\boldsymbol{\lambda}\|_{L^p(\ell^2)}^{p/p_j} \\
&\le 1.
\end{aligned}$$

This inequality is nearly trivial when $p_j = \infty$. Consequently, (62) holds for $X = \mathrm{K}$.

By a similar argument we can show that given $g \in \mathcal{P} \cap Y^{q'}$ with $\|g\|_{Y^{q'}} \leq 1$, there is a family of functions $(g_z, z \in \Xi)$ and an absolute constant C_2 which depends neither on g nor y such that $g_t = g$ and

(63) $$\|g_{j+\imath y}\|_{(Y^{q_j})'} \leq C_2$$

for $j = 0, 1, y \in \mathbf{R}$, and $Y \in \{\mathrm{L}, \mathrm{H}, \mathrm{K}\}$. Here $(Y^q)'$ represents the dual of Y^q if $1 \leq q < \infty$, $\mathrm{L}^1(\Omega, \nu)$ if $q = \infty$ and $Y = \mathrm{L}$, but H if $q = \infty$ and $Y = \mathrm{K}$.

For each $z \in \Xi$ set

$$\zeta(z) := \langle Tf_z, g_z \rangle .$$

Observe by Hölder's inequality or Fefferman's inequality that

$$|\zeta(z)| \leq C\|Tf_z\|_{Y^{q_j}} \|g_z\|_{(Y^{q_j})'}$$

for $j = 0, 1$ and $z \in \Xi$. Hence it follows from (62), (63), and hypothesis that

$$|\zeta(j + \imath y)| \leq CC_1C_2M_j$$

for $j = 0, 1$ and $y \in \mathbf{R}$. But f, g are dyadic step functions so the function ζ is a finite sum of functions analytic on the interior of the strip Ξ and continuous on its closure. It follows from the three lines theorem (see Appendix 0.2) that

$$|\zeta(t)| \leq CC_1C_2M_0^{1-t}M_1^t.$$

Since $\zeta(t) = \langle Tf, g \rangle$, we conclude from (61) that $\widetilde{M} \leq CC_1C_2M_0^{1-t}M_1^t$ as required. ∎

We close this section with some remarks concerning the complex method of interpolation due to Calderon (see Stein and Weiss [1], p. 210).

Let $\boldsymbol{X}$ be a Banach space and Ξ denote the strip $\{z \in \mathbf{C} : 0 \leq Re(z) \leq 1\}$. A map $f : \Xi \to \boldsymbol{X}$ is called ***analytic*** on the interior Ξ^o of Ξ if $\Lambda \circ f : \Xi \to \mathbf{C}$ is analytic in the classical sense on Ξ^o for every functional $\Lambda \in \boldsymbol{X}'$.

Let $\boldsymbol{X}^0, \boldsymbol{X}^1$ be fixed Banach spaces continuously embedded in some topological vector space such that $\boldsymbol{X}^0 \cap \boldsymbol{X}^1 \neq \{0\}$. Let $\mathcal{F}(\boldsymbol{X}^0, \boldsymbol{X}^1)$ be the set of functions $f : \Xi \to \boldsymbol{X}^0 + \boldsymbol{X}^1$ which are analytic on Ξ^o, continuous on Ξ with $\|f\|_{\boldsymbol{X}^0 + \boldsymbol{X}^1}$ bounded on Ξ, and such that
i) the map $y \to f(j + \imath y)$ is continuous from $\mathbf{R}$ into $\boldsymbol{X}^j$, and
ii) the map $y \to \|f(j + \imath y)\|_{\boldsymbol{X}^j}$ is bounded on $\mathbf{R}$ for $j = 0, 1$.

The collection $\mathcal{F} := \mathcal{F}(\boldsymbol{X}^0, \boldsymbol{X}^1)$ is a Banach space under the norm

$$\|f\| := \max\{\sup_{y \in \mathbf{R}} \|f(\imath y)\|_{\boldsymbol{X}^0}, \sup_{y \in \mathbf{R}} \|f(1 + \imath y)\|_{\boldsymbol{X}^1}\}.$$

Let $\mathcal{N}_t := \{f \in \mathcal{F} : f(t) = 0\}$. Set $\boldsymbol{X}_t := \mathcal{F}/\mathcal{N}_t$. Then $\boldsymbol{X}_t$ can be identified with the collection of elements $f(t) \in \boldsymbol{X}^0 + \boldsymbol{X}^1$ for $f \in \mathcal{F}$. Moreover, it is a Banach space under the usual quotient norm

$$\|x\|_{\boldsymbol{X}_t} := \inf\{\|f\|_{\mathcal{F}} : f \in \mathcal{F} \quad \text{and} \quad f(t) = x\}.$$

The space $\boldsymbol{X}_t$ is called a space of interpolation between $\boldsymbol{X}^0$ and $\boldsymbol{X}^1$ for the following reason. Let $\boldsymbol{X}^0, \boldsymbol{X}^1, \boldsymbol{Y}^0$, and $\boldsymbol{Y}^1$ be Banach spaces and $\boldsymbol{X}_t, \boldsymbol{Y}_t$ be defined as above. Suppose for each $j = 0, 1$ that T is a bounded linear operator from $\boldsymbol{X}^j$ to $\boldsymbol{Y}^j$ which satisfies

$$\|Tx\|_{\boldsymbol{Y}^j} \leq M_j \|x\|_{\boldsymbol{X}^j} \qquad (x \in \boldsymbol{X}^j).$$

Then T is a bounded linear operator from $\boldsymbol{X}_t$ to $\boldsymbol{Y}_t$ for every $0 < t < 1$ and

$$\|Tx\|_{\boldsymbol{Y}_t} \leq M_0^{1-t} M_1^t \|x\|_{\boldsymbol{X}_t} \qquad (x \in \boldsymbol{X}_t).$$

Thus Theorem 12 is part of a general theory of interpolation.

The collection $\boldsymbol{X}_t, 0 \leq t \leq 1$, is called the collection of intermediate spaces between $\boldsymbol{X}^0$ and $\boldsymbol{X}^1$. It is known (see Echandia [1]) that $\mathrm{H}^p, 1 \leq p \leq 2$, is the collection of intermediate spaces between H and L^2. In the classical case, it is well known that $\mathcal{H}^p, 1 \leq p \leq 2$, is the collection of intermediate spaces between $\mathcal{H}$ and L^2. Since $\mathcal{H}^p = \mathrm{H}^p = \mathrm{L}^p$ for $1 < p < 2$, the general interpolation theorem cited above contains the following result.

THE MIXED INTERPOLATION THEOREM. *If T is a bounded linear operator from* H *to* $\mathcal{H}$ *and from* L^2 *to* L^2, *then T is a bounded linear operator from* L^p *to* L^p *for* $1 < p \leq 2$.

We shall use this result in 5.5 to show that the Haar and Franklin systems are equivalent bases in L^p for $1 < p < \infty$.

3.7 Martingale Trees and Almost Everywhere Convergence of Walsh-Fourier Series. In 3.1 we saw there is a close connection between a.e. convergence of a sequence of partial sums and weak inequalities for the corresponding maximal operator. For example, $S_n f$ converges a.e., as $n \to \infty$, for every $f \in \mathrm{L}^p$ if the maximal operator S^* is of weak type (p,p). In 3.3, using martingale techniques, we obtained weak type (p,p) inequalities for the martingale transform which included the case of the dyadic maximal operator $\mathcal{E}^*$. It is natural to ask whether there exists a martingale structure which carries out the same program in the more demanding case of the maximal operator S^*.

The purpose of this section is to define such a structure and obtain the necessary estimates. Specifically, we introduce a martingale structure, indexed by the tree-like set of integer intervals $\mathcal{J}$ instead of the linearly ordered set of integers, which we call a martingale tree. We form a martingale transform for martingale trees and obtain the corresponding martingale maximal inequality (compare Theorem 4 in 3.3 with Theorem 13 below). Our proof in this case is substantially the same. We use interpolation, replace $\widetilde{\mathcal{E}}$ by a predictor $\mathcal{E}^\sharp$ (compare the definition of $\widetilde{\mathcal{E}}$ in 3.1 with (85) below), and use it to generate a canonical decomposition (see (90) below).

The collection $\mathcal{J}$ consists of the empty set together with sets of the form

$$I(k,-n) := [k2^n, (k+1)2^n) \cap \mathrm{N} \qquad (k, n \in \mathrm{N}).$$

For each $I \in \mathcal{J}$, let $|I|$ represent the number of elements in I. For $m, n \in \mathrm{N}$, let $I_{-n}(m)$ denote the element $I \in \mathcal{J}$ which contains m and satisfies $|I| = 2^n$.

The collection $\mathcal{J}$ is partially ordered under set inclusion $\subseteq$. This partial order enjoys the following properties:

(64) $\qquad \{J \in \mathcal{J} : I \subseteq J\}$ is linearly ordered for each $\quad I \in \mathcal{J}, I \neq \emptyset$,

(65) $\{J \in \mathcal{J} : J \subseteq I\}$ is a finite collection for each $I \in \mathcal{J}$,

and, $\mathcal{J}$ is an upward directed set, i.e.,

(66) given $I, J \in \mathcal{J}$ there is a $K \in \mathcal{J}$ such that $I, J \subseteq K$.

We shall call the map $n \to I(0,-n)$ the *natural embedding* of $\mathbf{N}$ in $\mathcal{J}$. It is clear that the natural embedding is order preserving, i.e., $n \leq m$ implies $I(0,-n) \subseteq I(0,-m)$.

Define a σ-algebra $\mathcal{A}^I$ for each $I \in \mathcal{J}$ as follows. If $I = \emptyset$ set $\mathcal{A}^\emptyset := \{\emptyset, [0,1)\}$. If $|I| = 2^n$ for some $n \in \mathbf{N}$ set $\mathcal{A}^I := \mathcal{A}^n$. Notice that this collection is increasing, i.e., $I \subseteq J$ implies $\mathcal{A}^I \subseteq \mathcal{A}^J$.

Extend the operator sequence $(\mathcal{E}_n, n \in \mathbf{N})$ from the linearly ordered index set $\mathbf{N}$ to the non-linearly ordered index set $\mathcal{J}$ by

$$\mathcal{E}_I f := \sum_{k \in I} \hat{f}(k) w_k \tag{67}$$

for $I \in \mathcal{J}$ and $f \in \mathrm{L}^1$, where the empty sum is defined to be zero. Notice for $n, m \in \mathbf{N}$ and $I = I_{-n}(m)$ that

$$\sum_{k \in I} w_k = \sum_{s=0}^{2^n - 1} w_{m \oplus s} = w_m D_{2^n}.$$

Consequently,

$$\mathcal{E}_I f = f * (w_m D_{2^n}) = w_m \mathcal{E}_n(w_m f) \tag{68}$$

for $I = I_{-n}(m)$ and any $f \in \mathrm{L}^1$.

The operators $\mathcal{E}_I$ behave as the operators $\mathcal{E}_n$ did. Specifically, by (68) above and (2) through (6) in 3.1, it is clear that each $|\mathcal{E}_I f|$ is $\mathcal{A}^I$ measurable,

$$\mathcal{E}_I \circ \mathcal{E}_J = \mathcal{E}_{I \cap J} \qquad (I, J \in \mathcal{J}), \tag{69}$$

each $\mathcal{E}_I$ is self-adjoint in L^2,

$$\|\mathcal{E}_I f\|_p \leq \|f\|_p \qquad (1 \leq p \leq \infty), \tag{70}$$

and

$$\mathcal{E}_I(\lambda f) = \lambda \mathcal{E}_I f \tag{71}$$

for $f \in \mathrm{L}^1$, $I \in \mathcal{J}$, and each $\mathcal{A}^I$ measurable function λ. The operator

$$\mathcal{E}^\natural f := \sup_{I \in \mathcal{J}} |\mathcal{E}_I f| \qquad (f \in \mathrm{L}^1) \tag{72}$$

plays the role that $\mathcal{E}^*$ did in 3.3. (Compare with (8) in 3.1).

To define Δ_I, fix $I \in \mathcal{J}$, $f \in L^1$, and let I_+ denote the *double* of I, i.e., let I_+ be the unique element $J \in \mathcal{J}$ which satisfies $|J| = 2|I|$ and $I \subset J$. (Here as elsewhere, $I \subset J$ means $I \subseteq J$ but $I \neq J$.) Set

$$\Delta_I f := \mathcal{E}_{I_+} f - \mathcal{E}_I f. \tag{73}$$

Also define

$$\mathcal{Q}_I f := \left(\sum_{J \supseteq I} |\Delta_J f|^2 \right)^{1/2}$$

and

$$\mathcal{Q}^* f := \sup_{I \in \mathcal{J}} \mathcal{Q}_I f.$$

Let $\boldsymbol{A}$ denote the collection of elements of the form $\boldsymbol{\alpha} = (\alpha_I, I \in \mathcal{J})$, where each α_I is $\mathcal{A}^I$-measurable, real valued, and

$$\|\boldsymbol{\alpha}\| := \sup_{I \in \mathcal{J}} |\alpha_I| < \infty.$$

For each $\boldsymbol{\alpha} \in \boldsymbol{A}$, $f \in L^1$, and $I \in \mathcal{J}$, set

$$\boldsymbol{T}_I(\boldsymbol{\alpha}) f := \sum_{J \supseteq I} \alpha_J \Delta_J f, \tag{74}$$

$$\boldsymbol{T}_I^*(\boldsymbol{\alpha}) f := \sup_{K \supset I} \Big| \sum_{I \subseteq J \subset K} \alpha_J \Delta_J f \Big|, \tag{75}$$

and

$$\boldsymbol{T}^*(\boldsymbol{\alpha}) f := \sup_{I \in \mathcal{J}} \boldsymbol{T}_I^*(\boldsymbol{\alpha}) f.$$

It is important to realize that each $\boldsymbol{T}_I^*(\boldsymbol{\alpha}) f$ is a linear maximal martingale transform, i.e., one of the type considered in 3.3. Indeed, if $I = I_{-n}(m)$ for some $m, n \in \mathbf{N}$ then

$$\{J \in \mathcal{J} : J \supseteq I\} = \{I_{-k}(m) : k \geq n\}.$$

Consequently, it is obvious by (68) that

$$w_m \boldsymbol{T}_I(\boldsymbol{\alpha}) f = \sum_{k=n}^{\infty} \alpha_{I_{-k}(m)} \Delta_k (f w_m) \tag{76}$$

for $I = I_{-n}(m)$. In particular, this series converges a.e. and the index set is linear as promised. Similarly, each $\mathcal{Q}_I f$ is a linear square function with

$$\mathcal{Q}_I f = \left(\sum_{k=n}^{\infty} |\Delta_k (f w_m)|^2 \right)^{1/2}$$

for $I = I_{-n}(m)$. Thus the inequalities obtained in 3.3 hold for each $\boldsymbol{T}^*_I(\boldsymbol{\alpha})f$ and each $Q_I f$.

The non-linear maximal martingale transform $\boldsymbol{T}^*(\boldsymbol{\alpha})f$ can be used to estimate the maximal function S^*f. Indeed, let $m \in \mathrm{N}, f \in \mathrm{L}^1$, and let $(m_j, j \in \mathrm{N})$ represent the binary coefficients of m. Given $I \in \mathcal{J}$ with $I = I_{-n}(m)$ set $\alpha_I := m_n$, and consider the function sequence

$$\boldsymbol{\alpha} := (\alpha_I, I \in \mathcal{J}).$$

The element $\boldsymbol{\alpha}$ is well defined, for if $I_{-n}(m) = I_{-n}(k)$ for some $k \in \mathrm{N}$ with binary coefficients $(k_j, j \in \mathrm{N})$, then $m_n = k_n$. It is also clear that $\boldsymbol{\alpha} \in \boldsymbol{A}$. And, by Theorem 8 in 1.5 we have

$$\begin{aligned} S_m f &= \sum_{k=0}^{\infty} m_k w_m (\mathcal{E}_{k+1} - \mathcal{E}_k)(f w_m) \\ &= \sum_{I \supseteq \{m\}} \alpha_I \Delta_I f. \end{aligned}$$

Therefore, given $m \in \mathrm{N}$ there is an $\boldsymbol{\alpha} \in \boldsymbol{A}$ such that

$$|S_m f| \le \boldsymbol{T}^*_{\{m\}}(\boldsymbol{\alpha}) f. \tag{77}$$

Thus the maximal function S^*f is controlled by the non-linear maximal martingale transform $\boldsymbol{T}^*(\boldsymbol{\alpha})f$.

We shall prove that $\boldsymbol{T}^*(\boldsymbol{\alpha})f$ is of weak type (p,p) for $1 < p < \infty$ (see Theorem 13 below). To this end, recall that the Marcinkiewicz quasi-norm of an $h \in \mathrm{L}^0$ is defined by

$$\|h\|_{M^p} := \sup_{y>0} y|\{|h| > y\}|^{1/p} \qquad (0 < p < \infty).$$

We extend this quasi-norm to function sequences as follows. If $\boldsymbol{f} = (f_n, n \in \mathrm{N})$ is a sequence of measurable functions and $0 < p < \infty$ then

$$\|\boldsymbol{f}\|_{M^p} := \sup_{y>0} y \left(\sum_{n=0}^{\infty} |\{|f_n| > y, \ |f_k| \le y \ \text{ for all } k < n \}| \right)^{1/p}. \tag{78}$$

Notice that this extends the classical Marcinkiewicz quasi-norm if we identify a function h with the function sequence $(h_n := h, n \in \mathrm{N})$.

We shall denote the collection of function sequences $\boldsymbol{f} = (f_n, n \in \mathrm{N})$ which satisfy $\|\boldsymbol{f}\|_{M^p} < \infty$ and $f_n \in \mathrm{L}^0$, $n \in \mathrm{N}$, by M^p for each $0 < p < \infty$. Define a maximal function for such a sequence by

$$f^* := \sup_{n \in \mathrm{N}} |f_n|.$$

It is easy to see that $\boldsymbol{f}$ is of finite M^p quasi-norm (as a sequence) if and only if f^* is of finite M^p quasi-norm (as a function). Indeed, since

$$\{f^* > y\} = \bigcup_{n=0}^{\infty} \{|f_n| > y, \ |f_k| \le y \text{ for all } k < n \}$$

and the sets on the right side of this identity are pairwise disjoint, we have

$$\|\boldsymbol{f}\|_{M^p} = \sup_{y>0} y|\{f^* > y\}|^{1/p} = \|f^*\|_{M^p}.$$

The inequalities of Corollary 1 in 3.1 can be reformulated using the quasi-norm (78):

$$\|(\mathcal{E}_n f, n \in \mathbf{N})\|_{M^1} \leq \|f\|_1,$$

$$\|(\widetilde{\mathcal{E}}_n f, n \in \mathbf{N})\|_{M^1} \leq 2\|f\|_1, \tag{79}$$

and

$$\|(\mathcal{E}_n^\natural f, n \in \mathbf{N})\|_{M^1} \leq \|f\|_1.$$

It is these inequalities which we will generalize to martingale trees.

First, for $0 < p, q < \infty$ and $\boldsymbol{g} = (g_I, I \in \mathcal{J})$ with $g_I \in \mathrm{L}^0$, let

$$\|\boldsymbol{g}\|_{M^{pq}} := \sup_{y>0} y \left(\int_0^1 (\sum_{I \in \mathcal{J}} \chi\{|g_I| > y, |g_J| \leq y \quad \text{for all } J \subset I\})^{p/q} \right)^{1/p}.$$

Denote the collection of such non-linearly ordered sequences $\boldsymbol{g}$ which satisfy $\|\boldsymbol{g}\|_{M^{pq}} < \infty$ by M^{pq}. Abusing the notation slightly, set $\mathrm{M}^p := \mathrm{M}^{pp}$. Analogous to (78) we have

$$\|\boldsymbol{g}\|_{M^p} = \sup_{y>0} y \left(\sum_{I \in \mathcal{J}} |\{|g_I| > y, |g_J| \leq y \quad \text{for all } J \subset I\}| \right)^{1/p}, \tag{80}$$

but here the sets on the right side are not in general pairwise disjoint.

Notice for each fixed function sequence $\boldsymbol{g}$ that the numbers $\|\boldsymbol{g}\|_{M^{pq}}$ decrease as q increases. Hence we define

$$\|\boldsymbol{g}\|_{M^{p\infty}} := \lim_{q \to \infty} \|\boldsymbol{g}\|_{M^{pq}}.$$

Clearly, the maximal function

$$g^* := \sup_{I \in \mathcal{J}} |g_I|$$

satisfies

$$\|\boldsymbol{g}\|_{M^{p\infty}} = \|g^*\|_{M^p} \qquad (0 < p < \infty).$$

To describe the analogue of (79) for the non-linear case we introduce more notation. For each $I \in \mathcal{J}$ with $|I| \geq 2$ let I' denote the left half of I and I'' denote the right half of I. In the case $|I| = 1$ set $I' := I$ and $I'' := \emptyset$.

LEMMA 5. *Let $\boldsymbol{\Delta}$ be the interior in $\mathbf{R}^2$ of the triangle with vertices* $(0,0)$, $(1/2,1/2)$, *and* $(1,0)$. *Let $1 \le p, q < \infty$ satisfy $(1/p, 1/q) \in \boldsymbol{\Delta}$. Then there is a constant $A(p,q)$ depending only on p and q such that*

$$\|(|\mathcal{E}_{I'}f| + |\mathcal{E}_{I''}f|, I \in \mathcal{J})\|_{M^{pq}} \le A(p,q)\|f\|_p \tag{81}$$

for all $f \in \mathrm{L}^p$.

PROOF. Fix $(1/p, 1/q) \in \boldsymbol{\Delta}$ and $y > 0$. For each $I \in \mathcal{J}$ set

$$g_I := |\mathcal{E}_{I'}f| + |\mathcal{E}_{I''}f|$$

and

$$\beta_I := \chi\{g_I > y, g_J \le y \quad \text{for all } J \subset I\}.$$

We first show that $(\beta_I \mathcal{E}_I, I \in \mathcal{J})$ is a collection of pairwise orthogonal projections on L^2. Indeed, fix $h \in \mathrm{L}^2$ and $I, J \in \mathcal{J}$. Observe by construction that

$$\beta_I \beta_J = 0$$

when $I \subset J$ or $J \subset I$. Thus for the case $I \subset J$ we have by $\mathcal{A}^I$-measurability of β_I that

$$\beta_I \mathcal{E}_I(\beta_J \mathcal{E}_J h) = \mathcal{E}_I(\beta_I \beta_J \mathcal{E}_J h) = 0.$$

For the case $J \subseteq I$ we have

$$\beta_I \mathcal{E}_I(\beta_J \mathcal{E}_J h) = \beta_I \beta_J \mathcal{E}_I(\mathcal{E}_J h) = \begin{cases} \beta_I \mathcal{E}_I h & J = I \\ 0 & J \subset I. \end{cases}$$

Finally, for the case $I \cap J = \emptyset$ we have by (69) and (71) that

$$\beta_I \mathcal{E}_I(\beta_J \mathcal{E}_J h) = \beta_I \mathcal{E}_I(\mathcal{E}_J(\beta_J \mathcal{E}_J h)) = 0.$$

Thus the $\beta_I \mathcal{E}_I$'s are pairwise orthogonal projections.

Apply Bessel's inequality to these projections (see Theorem 1 in Appendix 0.1) to verify

$$\sum_{I \in \mathcal{J}} \|\beta_I \mathcal{E}_I h\|_2^2 \le \|h\|_2^2 \qquad (h \in \mathrm{L}^2).$$

Since each β_I is $\mathcal{A}^{I'}$-measurable it is clear that $\beta_I \mathcal{E}_{I'} h$ and $\beta_I \mathcal{E}_{I''} h$ are orthogonal in L^2. It follows that

$$\sum_{I \in \mathcal{J}} (\|\beta_I \mathcal{E}_{I'} h\|_2^2 + \|\beta_I \mathcal{E}_{I''} h\|_2^2) \le \|h\|_2^2 \tag{82}$$

for each $h \in \mathrm{L}^2$.

Consider the map

$$Th := (\beta_I \mathcal{E}_{I'} h, \beta_I \mathcal{E}_{I''} h, I \in \mathcal{J})$$

defined for each $h \in L^1$. By (82) the map T is of type $(L^2, L^2(\ell^2))$. By (72) and Corollary 1 in 3.1 T is of type $(L^s, L^s(\ell^\infty))$. Indeed,

$$\begin{aligned}\|\sup_{I\in\mathcal{J}}(\max\{\beta_I|\mathcal{E}_{I'}h|, \beta_I|\mathcal{E}_{I''}h|\})\|_s &\le \|\mathcal{E}^\natural h\|_s \\ &\le \frac{s}{s-1}\|h\|_s\end{aligned}$$

for $1 < s < \infty$ and $h \in L^s$. It follows by interpolation (see Corollary 2 in Appendix 0.2) that

$$\|Th\|_{L^p(\ell^q)} \le \frac{s}{s-1}\|h\|_s \tag{83}$$

for each $h \in L^s$ where

$$s := s(p,q) := \frac{p(2-q)}{(p-q)}.$$

By the definition of $\|\cdot\|_{M^{pq}}$ we have

$$\|(g_I, I\in\mathcal{J})\|_{M^{pq}} = \sup_{y>0} y\left(\int_0^1 (\sum_{I\in\mathcal{J}}\beta_I)^{p/q}\right)^{1/p}. \tag{84}$$

By construction

$$\begin{aligned}\beta_I(|\mathcal{E}_{I'}f|^q + |\mathcal{E}_{I''}f|^q) &\ge 2^{1-q}\beta_I(|\mathcal{E}_{I'}f| + |\mathcal{E}_{I''}f|)^q \\ &\ge 2^{1-q}y^q\beta_I \qquad (I\in\mathcal{J}).\end{aligned}$$

Consequently,

$$\begin{aligned}y\left(\int_0^1 (\sum_{I\in\mathcal{J}}\beta_I)^{p/q}\right)^{1/p} &\le 2^{1-1/q}\left(\int_0^1 (\sum_{I\in\mathcal{J}}\beta_I(|\mathcal{E}_{I'}f|^q + |\mathcal{E}_{I''}f|^q))^{p/q}\right)^{1/p} \\ &= 2^{1-1/q}\|Tf\|_{L^p(\ell^q)}.\end{aligned}$$

We conclude by (83) and (84) that

$$\|(g_I, I\in\mathcal{J})\|_{M^{pq}} \le 2^{1-1/q}\left(\frac{s}{s-1}\right)\|f\|_p$$

where $s = s(p,q)$. In particular,

$$A(p,q) := p2^{1-1/q}\left(\frac{2-q}{p+q-pq}\right). \qquad \blacksquare$$

For each $I \in \mathcal{J}$ with $|I| = 2^n$ set $\mathcal{A}^{I-} := \mathcal{A}^{n-1}$. A sequence $(f_I, I\in\mathcal{J})$ of measurable functions is said to be predictable by a sequence $\boldsymbol{\lambda} = (\lambda_I, I\in\mathcal{J})$ if $\boldsymbol{\lambda}$ increases in I, if each λ_I is $\mathcal{A}^{I-}$-measurable, and

$$|f_I| \le \lambda_I \qquad (I\in\mathcal{J}).$$

For each $I \in \mathcal{J}$ set

(85) $$\mathcal{E}_I^\sharp f := \sup_{J \subseteq I}(|\mathcal{E}_{J'} f| + |\mathcal{E}_{J''} f|).$$

Clearly for each $f \in \mathrm{L}^1$ the sequence $(\mathcal{E}_I f, I \in \mathcal{J})$ is predictable by $(\mathcal{E}_I^\sharp, I \in \mathcal{J})$. Also, if

(86) $$\mathcal{E}^\sharp f := \sup_{I \in \mathcal{J}} \mathcal{E}_I^\sharp f$$

then it is evident that

(87) $$\mathcal{E}^\natural f \leq \mathcal{E}^\sharp f \leq 2\mathcal{E}^\natural f$$

for all $f \in \mathrm{L}^1$.

Notice by definition that

$$\|(\mathcal{E}_I^\sharp f, I \in \mathcal{J})\|_{M^{pq}} = \|(|\mathcal{E}_{I'} f| + |\mathcal{E}_{I''} f|, I \in \mathcal{J})\|_{M^{pq}}$$

for all $0 < p < \infty$, $0 < q \leq \infty$, and $f \in \mathrm{L}^1$. Hence by Lemma 5 we have for each $(1/p, 1/q) \in \boldsymbol{\Delta}$ there is a constant $A(p,q)$ such that

(88) $$\|(\mathcal{E}_I^\sharp f, I \in \mathcal{J})\|_{M^{pq}} \leq A(p,q)\|f\|_p$$

for $f \in \mathrm{L}^p$.

Predictability can be used to estimate certain kinds of martingale trees.

LEMMA 6. *Let $\boldsymbol{\alpha} \in \boldsymbol{A}$ with $\|\boldsymbol{\alpha}\| \leq 1$. Suppose that $\boldsymbol{\lambda} = (\lambda_I, I \in \mathcal{J})$ is a prediction of $(\mathcal{E}_I f, I \in \mathcal{J})$ for some $f \in \mathrm{L}^1$. If $1 < p < \infty$ and $q > p$ then there is an absolute constant $B(p,q)$, depending only on p and q, such that*

$$\|\boldsymbol{T}^*(\boldsymbol{\alpha})f\|_{M^p} \leq B(p,q)\|\boldsymbol{\lambda}\|_{M^{pq}}$$

and

$$\|\mathcal{Q}^* f\|_{M^p} \leq B(p,q)\|\boldsymbol{\lambda}\|_{M^{pq}}.$$

PROOF. Using (69) and (75) it is easy to see that the sequence $(\boldsymbol{T}^*(\boldsymbol{\alpha})(\mathcal{E}_n f), n \in \mathrm{N})$ is increasing and everywhere convergent to $\boldsymbol{T}^*(\boldsymbol{\alpha})f$. A similar statement holds for $\mathcal{Q}^* f$. Therefore, it is enough to prove the lemma for a Walsh polynomial f. For simplicity we deal only with $\boldsymbol{T}^*(\boldsymbol{\alpha})$. A similar argument works for $\mathcal{Q}^*$. For this case we provide the facts but not the details.

The proof proceeds in three steps. First we generate a canonical decomposition for the non-linear case. Next we use it to show the non-linear maximal martingale transform and non-linear square function can be estimated by a supremum of their linear counterparts (see (94) and (97) below). Finally, we combine this estimate with linear martingale inequalities derived earlier to finish the proof.

For a canonical decomposition set

$$\epsilon_I^k := \chi\{2^k < \lambda_{I_+} \le 2^{k+1}\} \qquad (I \in \mathcal{J}) \tag{89}$$

and

$$\boldsymbol{\epsilon}^k := (\epsilon_I^k, I \in \mathcal{J}) \qquad (k \in \mathbf{Z}).$$

Observe by predictability that for each $I \in \mathcal{J}$ we have

$$\sum_{k\in\mathbf{Z}} \epsilon_I^k \Delta_I f = \chi\{\lambda_{I_+} > 0\}\Delta_I f = \Delta_I f.$$

Thus

$$\sum_{I\subseteq J\subset K} \alpha_J \Delta_J f = \sum_{k\in\mathbf{Z}} \sum_{I\subseteq J\subset K} \epsilon_J^k \alpha_J \Delta_J f$$

for all $I \subset K, I, K \in \mathcal{J}$. Consequently,

$$\boldsymbol{T}^*(\boldsymbol{\alpha})f \le \sum_{k\in\mathbf{Z}} \boldsymbol{T}^*(\boldsymbol{\alpha}\boldsymbol{\epsilon}^k)f. \tag{90}$$

For $\mathcal{Q}^*$ the facts are

$$\mathcal{Q}_I(\boldsymbol{\epsilon}^k)f = (\sum_{J\supseteq I} \epsilon_J^k |\Delta_J f|^2)^{1/2} \qquad (I \in \mathcal{J}),$$

$$\mathcal{Q}^*(\boldsymbol{\epsilon}^k)f = \sup_{I\in\mathcal{J}} \mathcal{Q}_I(\boldsymbol{\epsilon}^k f) \qquad (k \in \mathbf{Z}),$$

and

$$\mathcal{Q}^* f \le \sum_{k\in\mathbf{Z}} \mathcal{Q}^*(\boldsymbol{\epsilon}^k)f. \tag{91}$$

This completes step 1.

Next, fix $k \in \mathbf{Z}$ and for each $J \in \mathcal{J}$ set

$$\gamma_J^k := \chi\{\lambda_J > 2^k\} \tag{92}$$

and

$$\beta_J^k := \chi\{\lambda_J > 2^k, \lambda_K \le 2^k \quad \text{for all } K \subset J\}. \tag{93}$$

We claim that

$$\boldsymbol{T}^*(\boldsymbol{\alpha}\boldsymbol{\epsilon}^k)f \le 2 \sup_{I\in\mathcal{J}} \beta_I^k \boldsymbol{T}_I^*(\boldsymbol{\alpha}\boldsymbol{\epsilon}^k)f + 2^{k+2}\chi\{\lambda^* > 2^k\} \tag{94}$$

for $\lambda^* := \sup_{J\in\mathcal{J}} \lambda_J$. Fix $x \in [0,1)$ and a pair $I, K \in \mathcal{J}$ which satisfy $I \subset K$. Set

$$\Sigma_{IK}(x) := \sum_{I\subseteq J\subset K} (\alpha_J \epsilon_J^k \Delta_J f)(x)$$

and observe by (75) that (94) will follow from

$$|\Sigma_{IK}(x)| \leq 2 \sup_{J\in\mathcal{J}} (\beta_J^k \boldsymbol{T}_J^*(\boldsymbol{\alpha\epsilon}^k) f)(x) + 2^{k+2}\chi\{\lambda^* > 2^k\}(x). \tag{95}$$

To prove (95) notice by construction that

$$\epsilon_J^k = \gamma_{J_+}^k \epsilon_J^k \qquad (J \in \mathcal{J}).$$

Consequently, if the set

$$\{J \in \mathcal{J} : I \subseteq J \subset K \quad \text{and} \quad \gamma_{J_+}^k(x) = 1\}$$

is empty then $\Sigma_{IK}(x) = 0$ and (95) is trivial. If this set is non-empty, denote one of its intervals of minimal length by I^0. Let

$$\mathcal{J}_0 := \{M \in \mathcal{J} : M \subseteq I_+^0 \quad \text{and} \quad \gamma_M^k(x) = 1\}$$

and let I^1 be an interval of $\mathcal{J}_0$ of minimal length. Then $I^1 \subseteq I_+^0$ and $\gamma_{I^1}^k(x) = 1$ but $\gamma_L^k(x) = 0$ for every $L \subset I^1$. By comparing (92) and (93) we have $\beta_{I^1}(x) = 1$. Thus by minimality of I^0 and I^1,

$$\begin{aligned}
\Sigma_{IK}(x) &= \Sigma_{I^0K}(x) \\
&= \left(\alpha_{I^0}\epsilon_{I^0}^k \Delta_{I^0} f\right)(x) + \Sigma_{I_+^0 K}(x) \\
&= \left(\alpha_{I^0}\epsilon_{I^0}^k \Delta_{I^0} f\right)(x) + \beta_{I^1}^k(x)\left(\Sigma_{I^1K}(x) - \Sigma_{I^1 I_+^0}(x)\right) \\
&:= F_1(x) + F_2(x).
\end{aligned}$$

Now (89) implies

$$\chi\{\lambda^* > 2^k\}\epsilon_J^k = \epsilon_J^k \qquad (J \in \mathcal{J}). \tag{96}$$

Consequently by predictability we have

$$|F_1(x)| \leq 2 \cdot 2^{k+1}\chi\{\lambda^* > 2^k\}(x).$$

By definition we have

$$|F_2(x)| \leq 2|\beta_{I^1}^k(x)|(\boldsymbol{T}_{I^1}^*(\boldsymbol{\alpha\epsilon}^k)f)(x).$$

This establishes (95) and the proof of (94) is complete. For the case $\mathcal{Q}^*$ it turns out that

$$\mathcal{Q}^*(\boldsymbol{\epsilon}^k) \leq \sup_{I\in\mathcal{J}} \beta_I^k \mathcal{Q}_I(\boldsymbol{\epsilon}^k f) + 2^{k+2}\chi\{\lambda^* > 2^k\}. \tag{97}$$

This completes step 2.

For step 3 set

$$g_{2^k}^{(1)} := 2^k \chi\{\lambda^* > 2^k\},\ g_{2^k}^{(2)} := \sup_{I \in \mathcal{J}} \beta_I^k \boldsymbol{T}_I^*(\boldsymbol{\alpha\epsilon}^k) f$$

and suppose for a moment that there are constants $B^{(i)} := B^{(i)}(p,q)$ depending only on p and q such that

$$|\{g_{2^k}^{(i)} > t2^k\}| \le B^{(i)} t^{-q} 2^{-pk} \|\boldsymbol{\lambda}\|_{M^{pq}}^p \tag{98}$$

for all $k \in \mathbf{Z}$, $t > 0$, and $i = 1, 2$. Notice by (94) that

$$\boldsymbol{T}^*(\boldsymbol{\alpha\epsilon}^k) f \le 4(g_{2^k}^{(1)} + g_{2^k}^{(2)}) \qquad (k \in \mathbf{Z}).$$

Also notice by definition and (96) that

$$g_{2^k}^{(i)} \chi\{\lambda^* > 2^k\} = g_{2^k}^{(i)} \qquad (k \in \mathbf{Z}, i = 1, 2).$$

Hence by Lemma 2 in 3.2 and (98) we have

$$y^p |\{\boldsymbol{T}^*(\boldsymbol{\alpha}) f > y, \lambda^* \le y\}| \le 8^p (B^{(1)} + B^{(2)}) C(p,q) \|\boldsymbol{\lambda}\|_{M^{pq}}^p$$

for all $y > 0$. Since

$$|\{\boldsymbol{T}^*(\boldsymbol{\alpha}) f > y\}| \le |\{\boldsymbol{T}^*(\boldsymbol{\alpha}) f > y, \lambda^* \le y\}| + |\{\lambda^* > y\}|$$

and

$$y^p |\{\lambda^* > y\}| \le \|\lambda^*\|_{M^p}^p = \|\boldsymbol{\lambda}\|_{M^{p\infty}}^p \le \|\boldsymbol{\lambda}\|_{M^{pq}}^p$$

it follows that

$$\|\boldsymbol{T}^*(\boldsymbol{\alpha}) f\|_{M^p} \le B(p,q) \|\boldsymbol{\lambda}\|_{M^{pq}}$$

where $B(p,q)$ depends only on p and q. Thus it remains to prove (98).

For the case $i = 1$, (98) is easy. Indeed,

$$\begin{aligned}
|\{g_{2^k}^{(1)} > t2^k\}| &\le \frac{2^{kq} \int_0^1 \chi\{\lambda^* > 2^k\}}{2^{kp} t^q} \\
&\le \frac{\|\lambda^*\|_{M^p}}{2^{kp} t^q} \\
&= \frac{\|\boldsymbol{\lambda}\|_{M^{p\infty}}^p}{2^{kp} t^q} \\
&\le \frac{\|\boldsymbol{\lambda}\|_{M^{pq}}^p}{2^{kp} t^q} \qquad (k \in \mathbf{Z}, t > 0).
\end{aligned}$$

For the case $i = 2$ use $\boldsymbol{T}_I(\boldsymbol{\alpha\epsilon}^k) = \boldsymbol{T}_I(\boldsymbol{\alpha}) \boldsymbol{T}_I(\boldsymbol{\epsilon}^k)$ to verify

$$\boldsymbol{T}_I^*(\boldsymbol{\alpha\epsilon}^k) \le \boldsymbol{T}_I^*(\boldsymbol{\alpha})(\boldsymbol{T}_I(\boldsymbol{\epsilon}^k) f) \qquad (k \in \mathbf{Z}, I \in \mathcal{J}).$$

Moreover, use predictability and Theorem 3 in 3.2 to see that

$$T_I^*(\epsilon^k)f \le 2^{k+2}\chi\{\lambda^* > 2^k\} \qquad (k \in \mathbf{Z}, I \in \mathcal{J}).$$

Consequently, if $\mathcal{E}(h|\mathcal{A}^I)$ represents the conditional expectation of an $h \in \mathrm{L}^1$ with respect to $\mathcal{A}^I$, we have by the $\mathcal{A}^I$-measurability of β_I^k and Corollary 8 in 3.3 that

$$\begin{aligned}\beta_I^k \mathcal{E}(|T_I^*(\alpha\epsilon^k)f|^s \mid \mathcal{A}^I) &\le \left(\frac{As^2}{s-1}\right)^s \mathcal{E}(\beta_I^k |T_I(\epsilon^k)f|^s \mid \mathcal{A}^I) \\ &\le \left(\frac{As^2 2^{k+2}}{s-1}\right)^s \beta_I^k\end{aligned}$$

for every $k \in \mathbf{Z}, I \in \mathcal{J}, 1 < s < \infty$, and some absolute constant A. Combining this inequality for $s \ge q$ with the definition of $g_{2^k}^{(2)}$, it follows from Corollary 2 in 3.1 that

$$\begin{aligned}|\{g_{2^k}^{(2)} > t2^k\}| &\le t^{-q}2^{-kq} \int_0^1 (\sum_{I \in \mathcal{J}} \beta_I^k |T_I^*(\alpha\epsilon^k)f|^s)^{q/s} \\ &\le 4t^{-q}2^{-kq} \int_0^1 (\sum_{I \in \mathcal{J}} \beta_I^k \mathcal{E}(|T_I^*(\alpha\epsilon^k)f|^s \mid \mathcal{A}^I))^{q/s} \\ &\le 4t^{-q}\left(\frac{4As^2}{s-1}\right)^q \int_0^1 (\sum_{I \in \mathcal{J}} \beta_I^k)^{q/s}\end{aligned}$$

for $k \in \mathbf{Z}$ and $t > 0$. Finally, let $s := q^2/p$. By the definition of $\|\cdot\|_{M^{pq}}$ we conclude that

$$|\{g_{2^k}^{(2)} > t2^k\}| \le B^{(2)} t^{-q} 2^{-pk} \|\boldsymbol{\lambda}\|_{M^{pq}}^p$$

for some constant $B^{(2)}$ depending only on p and q. ∎

Since $(\mathcal{E}_I f, I \in \mathcal{J})$ is predictable by $(\mathcal{E}_I^\sharp f, I \in \mathcal{J})$, Lemma 6 and (88) combine to establish the following result.

THEOREM 13. *Let $1 < p < \infty$. There is a constant C_p depending only on p such that*

$$\|T^*(\alpha)f\|_{M^p} \le C_p \|f\|_p$$

and

$$\|\mathcal{Q}^* f\|_{M^p} \le C_p \|f\|_p$$

for all $f \in \mathrm{L}^p$ and all $\alpha \in \boldsymbol{A}$ with $\|\alpha\| \le 1$.

In particular, we have shown that the non-linear maximal operator $T^*(\alpha)$ is of weak type (p,p) for all $1 < p < \infty$, and $\alpha \in \boldsymbol{A}$ with $\|\alpha\| \le 1$. Thus by Theorem 2 in 3.1 and (77) above, we have proved the following convergence theorem.

THEOREM 14. *If $f \in \mathrm{L}^p$ for some $p > 1$ then $S_n f \to f$ a.e. on $[0,1)$ as $n \to \infty$.*

We have also proved that S^* is of weak type (p,p) for all $1 < p < \infty$. Thus it follows from the Marcinkiewicz interpolation theorem (see Appendix 0.2)

COROLLARY 11. *The maximal operator S^* is of type (p,p) for all $1 < p < \infty$.*

Exercises

3.1 Show the dyadic maximal operator $\mathcal{E}^*$ is of weak type (1,1).

3.2 Let $1 < p < \infty$ and suppose the maximal operator S^* is of type (p,p). Prove that $S_n f \to f$, as $n \to \infty$, in L^p norm and a.e. on [0,1).

3.3 Let $n = 2^{n_1} + 2^{n_2} + \cdots + 2^{n_\ell}$ where $n_1 > n_2 > \cdots > n_\ell \geq 0$ are integers and $\ell \in \mathbf{P}$.

a) Show

$$w_n D_n = \sum_{j=1}^{\ell} \sum_{k=2^{n_j}}^{2^{n_j+1}-1} w_k.$$

b) Let $g := f w_n$ for $f \in \mathrm{L}^1$ and prove

$$w_n S_n f = \sum_{j=1}^{\ell} \Delta_{n_j} g.$$

c) Use Paley's inequality to show there is a constant C_p for each $1 < p < \infty$ such that

$$\|S_n f\|_p \leq C_p \|f\|_p$$

for $n \in \mathbf{N}$ and $f \in \mathrm{L}^p$.

d) Show the inequality in c) holds if and only if $S_n f \to f$ in L^p norm as $n \to \infty$ for all $f \in \mathrm{L}^p$.

3.4 Prove there is a constant $c > 0$ such that

$$|\{S_n f > y\}| \leq \frac{c\|f\|_1}{y}$$

for all $f \in \mathrm{L}^1, n \in \mathbf{N}$, and $y > 0$.

3.5 Use 3.4 to show there exist constants $c_p > 0$ for $0 \leq p \leq 1$ such that

$$\|S_n f\|_1 \leq c_1 \int_0^1 |f| \log^+ |f| + c_0$$

and

$$\|S_n f\|_p \leq c_p \|f\|_1 \qquad (0 < p < 1)$$

both hold for all $f \in \mathrm{L}^1$ and $n \in \mathbf{N}$.

3.6 Let

$$f := \sum_{k=0}^{\infty} (-1)^k \frac{r_k D_{2^k}}{2^k}.$$

Show that $f \in \mathrm{L}^\infty$ but $\mathcal{Q} f \notin \mathrm{L}^\infty$.

Hint: Is $\mathcal{Q} f \geq \sqrt{n}$ on $[2^{-n}, 2^{-n+1})$ for some $n \in \mathbf{P}$?

3.7 a) Using the fact that the constant C_p in Corollary 4 is $O(p)$, as $p \to \infty$, prove

$$\int_0^1 \chi\{\widetilde{\mathcal{E}} f \leq y\} \exp \frac{|\boldsymbol{T}(\boldsymbol{\alpha}) f|}{3y} \leq c$$

for $y > 0, \boldsymbol{\alpha} \in \boldsymbol{A}$, $\|\boldsymbol{\alpha}\| \leq 1$, and $f \in \mathrm{L}^1$. Here

$$c := \sum_{p=0}^{\infty} \frac{(2C_p)^p}{p!}.$$

b) Show that

$$\int_0^1 \exp \frac{|\boldsymbol{T}(\boldsymbol{\alpha})f|}{3\|f\|_\infty} \leq c$$

for $\boldsymbol{\alpha} \in \boldsymbol{A}$, $\|\boldsymbol{\alpha}\| \leq 1$, and $f \in \mathrm{L}^\infty$.

c) Prove there exists an absolute constant B such that

$$\int_0^1 \exp \frac{|\mathcal{Q}f|^2}{3\|f\|_\infty} \leq B$$

for all $f \in \mathrm{L}^\infty$.

3.8 For each $\boldsymbol{\alpha} \in \boldsymbol{A}$ and $0 < p \leq \infty$ define the p-norm of the martingale transform $\boldsymbol{T}(\boldsymbol{\alpha})$ by

$$\|\boldsymbol{T}(\boldsymbol{\alpha})\|_p := \sup_{f \in \mathcal{P}, \|f\|_p \leq 1} \|\boldsymbol{T}(\boldsymbol{\alpha})f\|_p.$$

Prove

$$\|\boldsymbol{T}(\boldsymbol{\alpha})\|_p = \|\boldsymbol{T}(\boldsymbol{\alpha})\|_q$$

for all $1 < p < \infty$, where $1/p + 1/q = 1$.

3.9 Let $f \in \mathrm{BMO}$ and $y > 0$. Show there exist absolute constants $A, B > 0$ such that

$$|\{|f| > y\}| \leq A \exp(-By/\|f\|_{\mathrm{BMO}}).$$

[John and Nirenberg [1]]

3.10 a) Show $\mathrm{L}\log^+ \mathrm{L} \subseteq \mathrm{H}$.

b) Show that if $f \in \mathrm{H}$ and $f \geq 0$ then $f \in \mathrm{L}\log^+ \mathrm{L}$.

c) Let

$$f(x) := (-1)^n \frac{2^n}{n^2}, \qquad (x \in (2^{-n}, 2^{-n+1}), n \in \mathbf{P}).$$

Prove $f \in \mathrm{H}$ but $f \notin \mathrm{L}\log^+ \mathrm{L}$.

Hint: Show $|\mathcal{Q}f|^2 \chi(2^{-n}, 2^{-n+1}) \leq 2^{2n} \sum_{k=n}^{\infty} k^{-4}$ for $n \in \mathbf{P}$.

3.11 Suppose $1 \leq p \leq 2$ and $f \in \mathrm{H}^p$. Prove there is an absolute constant C such that

$$\left(\sum_{k=1}^{\infty} \frac{|\widehat{f}(k)|^p}{k^{2-p}} \right)^{1/p} \leq C\|f\|_{H^p}.$$

Hint: Let $Tf := (n\widehat{f}(n), n \in \mathbf{P}), \boldsymbol{X} := \mathrm{L}^1(\mathbf{P}, \nu)$, where $\nu(\{n\}) := 1/n^2$ for $n \in \mathbf{P}$, and use interpolation (see Theorem 12 in 3.6).

3.12 Let $2 \le q \le \infty$ and let K^q denote the collection of $f \in \mathrm{L}^2$ for which there exists a $\gamma \in \mathrm{L}^q$ such that

$$\mathcal{E}_n|f - \mathcal{E}_n f|^2 \le \mathcal{E}_n(\gamma^2), \tag{97}$$

for $n \in \mathbf{N}$. Define a norm on K^q by

$$\|f\| := \inf \|\gamma\|_q,$$

where this infimum is taken over all $\gamma \in \mathrm{L}^q$ such that (97) holds. Prove that K^q and the space K^q introduced in 3.6 are equivalent Banach spaces for each $2 \le q \le \infty$.

3.13 a) Let $F : \mathcal{P} \to \mathbf{R}$ be a linear functional and set

$$\nu(I) := F(\chi(I)) \qquad (I \in \mathcal{I}).$$

Prove that ν is a quasi-measure and

$$F(f) = \int_0^1 f \, d\nu$$

for every $f \in \mathcal{P}$.

b) Show that to each linear map $F : \mathcal{P} \to \mathbf{R}$ there is a dyadic martingale $(g_n, n \in \mathbf{N})$ such that

$$F(f) = \lim_{n \to \infty} \int_0^1 g_n f$$

for every $f \in \mathcal{P}$.

c) Using b), prove that $(\mathrm{L}^p)' = \mathrm{L}^q$ for $1 \le p < \infty$ where p and q are conjugate indices.

3.14 Prove

$$f := \sum_{n=1}^{\infty} \frac{2^n}{n^2} \chi[2^{-n}, 2^{-n+1})$$

belongs to L^1 but does not belong to H.

3.15 Prove

$$\psi := \sum_{n=1}^{\infty} n \chi[2^{-n}, 2^{-n+1})$$

belongs to BMO but does not belong to L^∞.

3.16 Show there exist functions $f \in \mathrm{H}$ and $\psi \in \mathrm{BMO}$ such that $f\psi \notin \mathrm{L}^1$.

3.17 Prove BMO is closed under square roots, and

$$\|\sqrt{\psi}\|_{\mathrm{BMO}} \le 4\sqrt{\|\psi\|_{\mathrm{BMO}}}$$

for $\psi \ge 0, \psi \in \mathrm{BMO}$.

3.18 Prove

$$\|f * g\|_{\mathrm{BMO}} \le \|f\|_1 \, \|g\|_{\mathrm{BMO}}$$

for all $f \in L^1$ and $g \in \mathrm{BMO}$.

3.19 Use the martingale maximal theorem (Corollary 1 in 3.1) and Fubini's theorem to prove that a non-negative $f \in L^1$ belongs to H if and only if $f \in L\log^+ L$.

3.20 Let $(X_n, \|\cdot\|_n), n \in \mathbf{N}$ be a sequence of Banach spaces. Set

$$\boldsymbol{X} := X_0 \times X_1 \times \ldots,$$

$$\|\boldsymbol{\xi}\|_{\ell^\infty} := \sup_{n\in\mathbf{N}} \|\xi_n\|_n,$$

and

$$\|\boldsymbol{\xi}\|_{\ell^p} := \left(\sum_{n=0}^{\infty} \|\xi_n\|_n^p\right)^{1/p}$$

for $1 \le p < \infty$ and $\boldsymbol{\xi} := (\xi_0, \xi_1, \ldots) \in \boldsymbol{X}$.

a)Prove

$$\boldsymbol{X}^p := \{\boldsymbol{\xi} \in \boldsymbol{X} : \|\boldsymbol{\xi}\|_{\ell^p} < \infty\}$$

is a Banach space for $1 \le p \le \infty$.

b) Let $1 \le p < \infty$ and p' be the index conjugate to p. Denote the dual of X_n by Y_n, let

$$\boldsymbol{Y} := Y_0 \times Y_1 \times \ldots$$

and form $\boldsymbol{Y}^{p'}$ as $\boldsymbol{X}^p$ was formed above. Prove that $\Lambda \in (\boldsymbol{X}^p)'$ if and only if there is a $\Phi = (\Phi_n, n \in \mathbf{N}) \in \boldsymbol{Y}^{p'}$ such that

$$\|\Lambda\| = \|\Phi\|_{\boldsymbol{Y}^{p'}}$$

and

$$\Lambda(\boldsymbol{\xi}) = \sum_{n=0}^{\infty} \Phi_n(\xi_n)$$

for all $\boldsymbol{\xi} = (\xi_n, n \in \mathbf{N}) \in \boldsymbol{X}^p$.

3.21 Suppose $f_n \in L^\infty$ for $n \in \mathbf{N}$ and

$$\sum_{n=0}^{\infty} |f_n| \le 1.$$

a) Show

$$\mathcal{E}_m(|\sum_{n=m}^{\infty} \mathcal{E}_n f_n|^2) \le 2 \qquad (m \in \mathbf{N}).$$

b) Let

$$\psi := \sum_{n=0}^{\infty} \mathcal{E}_n f_n.$$

Prove $\psi \in \mathrm{BMO}$ and

$$\|\psi\|_{\mathrm{BMO}} \le \sqrt{2}.$$

3.22 Let $\boldsymbol{X} \subset \mathrm{L}^1(\ell^\infty)$ be the collection of sequences $f = (f_n, n \in \mathbf{N})$ such that $f_{n+1} = f_n$ for n sufficiently large. Set $f_\infty := \lim_{n\to\infty} f_n$. For $n, N \in \mathbf{N}$, let $\epsilon_n := (\delta_{nk}, k \in \mathbf{N})$ and $\epsilon^{(N)} := (0, 0, \dots, 0, 1, 1, \dots)$ where the first 1 in $\epsilon^{(N)}$ starts at the N-th coordinate.

a) Verify that $f \in \boldsymbol{X}$ implies there is an $N \in \mathbf{N}$ such that

$$f = \sum_{n=0}^{N-1} f_n \epsilon_n + f_\infty \epsilon^{(N)}.$$

b) Show that any $F \in \boldsymbol{X}'$ has the form

$$F(f) = \sum_{n=0}^{N-1} \langle \epsilon_n, f_n \rangle + \left\langle \eta^{(N)}, f_\infty \right\rangle \qquad (f \in \boldsymbol{X})$$

where $\sum_{n=0}^{N-1} |\epsilon_n| + |\eta^{(N)}| \le \|F\|, \eta^{(N)} = \epsilon_N + \eta^{(N+1)}$, and N depends on f and is given by a).

c) Prove that given $F \in \boldsymbol{X}'$ there exist $\eta, \epsilon_k \in \mathrm{L}^\infty$ for $k \in \mathbf{N}$ such that

$$F(f) = \sum_{k=0}^{\infty} \langle \epsilon_k, f_k \rangle + \langle \eta, f_\infty \rangle \qquad (f \in \boldsymbol{X})$$

and

$$\| \, |\eta| + \sum_{k=0}^{\infty} |\epsilon_k| \, \|_\infty \le \|F\|.$$

3.23 Let $\boldsymbol{X}$ represent the space introduced in Exercise 3.22. By considering the map from H into $\boldsymbol{X}$ defined by

$$f \to (\mathcal{E}_n f, n \in \mathbf{N}),$$

prove that every $\psi \in \mathrm{BMO}$ can be written in the form

$$\psi = \eta + \sum_{k=0}^{\infty} \mathcal{E}_k \epsilon_k$$

where $\eta, \epsilon_k \in \mathrm{L}^\infty$ for $k \in \mathbf{N}$, and

$$\| \, |\eta| + \sum_{k=0}^{\infty} |\epsilon_k| \, \|_\infty \le \|\psi\|_{\mathrm{BMO}}.$$

3.24 Let $(a_n, n \in \mathbf{P})$ be a sequence of positive numbers which satisfy

$$\sup_{n\in\mathrm{N}} (2^n \sum_{k=2^n}^{\infty} a_k^2) \le 1.$$

Prove there is a constant K such that

$$\sum_{k=1}^{\infty} a_k |\widehat{f}(k)| \le K \|f\|_{\mathrm{H}}$$

for all $f \in \mathrm{H}$.

3.25 a) Show

$$W_1(x) := 1 + \sum_{k=1}^{\infty} \frac{w_k(x)}{k}$$

converges for $x \in (0,1)$ and belongs to $\mathrm{Lip}(1, \mathrm{L}^1)$.

b) Show $\mathrm{Lip}(1, \mathrm{L}^1) \subset \mathrm{BMO}$ but $\mathrm{Lip}(1, \mathrm{L}^1)$ is not contained in L^∞.

[Ladhawala [1]]

3.26 Let $\widetilde{\mathcal{I}}$ be the collection of intervals which are the union of any pair of adjacent dyadic intervals of the same size. Call $\beta \in \mathrm{L}^\infty$ an $\widetilde{\mathcal{I}}$-atom if $\beta := 1$ or if β is of mean zero, and satisfies

$$\{\beta \ne 0\} \subseteq I,$$

$$\|\beta\|_\infty \le \frac{1}{|I|},$$

for some interval $I \in \widetilde{\mathcal{I}}$. Set

$$\|f\|_{\widetilde{\mathcal{BMO}}} := \sup_{I \in \widetilde{\mathcal{I}}} \left(\frac{1}{|I|} \int_I |f - \frac{1}{|I|} \int_I f|^2 \right)^{1/2}$$

and

$$\|f\|_{\widetilde{\mathcal{H}}} := \inf\{ \quad \|\boldsymbol{a}\|_{\ell^1} \},$$

where the infimum is taken over all sequences $\boldsymbol{a} := (a_n, n \in \mathbf{N}) \in \ell^1$ which satisfy $f = \sum_{n=0}^{\infty} a_n \beta_n$ for some $\widetilde{\mathcal{I}}$-atoms $\beta_0, \beta_1, \ldots$. Similarly define atomic $\mathcal{H}$ and $\mathcal{BMO}$ norms where $\widetilde{\mathcal{I}}$ is replaced by the entire collection of subintervals of [0,1).

a) Prove

$$\|f\|_{\mathcal{H}} \le \|f\|_{\widetilde{\mathcal{H}}} \le 4\|f\|_{\mathcal{H}}.$$

b) Show $f \in \widetilde{\mathcal{BMO}}$ if and only if there is a constant c such that

$$M := \sup_{I \in \widetilde{\mathcal{I}}} \left(\frac{1}{|I|} \int_I |f - c|^2 \right)^{1/2} < \infty,$$

in which case $\|f\|_{\widetilde{\mathcal{BMO}}} \le 2M$.

c) Prove

$$\|f\|_{\widetilde{\mathcal{BMO}}} \le \|f\|_{\mathcal{BMO}} \le 8\|f\|_{\widetilde{\mathcal{BMO}}}.$$

Chapter 4

CONVERGENCE IN NORM

4.1 L^p Convergence of Walsh-Fourier Series. We have seen that the Walsh-Fourier series of any $f \in L^p, p > 1$, converges to f a.e. It is natural to ask if Sf converges also in L^p norm.

THEOREM 1. *If $1 < p < \infty$ and $f \in L^p$ then $S_n f \to f$, as $n \to \infty$, in L^p norm.*

PROOF. Let $\varepsilon > 0$ and choose a Walsh polynomial P such that $\|f - P\|_p < \varepsilon$. Since $S_n P = P$ for n large, it is clear that

$$\|f - S_n f\|_p \le \|f - P\|_p + \|S_n P - S_n f\|_p.$$

Thus Corollary 6 in 3.3 implies

$$\limsup_{n\to\infty} \|f - S_n f\|_p \le 2C_p\varepsilon. \qquad \blacksquare$$

Since L^∞ is not separable, Theorem 1 fails for $p = \infty$. It also fails for $p = 1$ (see Theorem 2 below). However, the situation changes if we pass to 2^n-th partial sums. Indeed, since

$$\begin{aligned}\|S_{2^n} f\|_p &= \|D_{2^n} * f\|_p \\ &\le \|D_{2^n}\|_1 \|f\|_p \\ &= \|f\|_p\end{aligned}$$

for all $n \in \mathbf{N}$, the proof of Theorem 1 establishes that $S_{2^n} \to f$, as $n \to \infty$, in L^p norm for all $1 \le p < \infty$ and $f \in L^p$.

Much more is true. Let $\Phi : [0, \infty) \to [0, \infty)$ be an increasing, continuous, convex function such that $\Phi(0) = 0$ and $\Phi(2x) \le C\Phi(x)$ for $x \ge 0$ and some absolute constant $C > 0$. We shall show that

$$\lim_{n\to\infty} \int_0^1 \Phi(|S_{2^n} f - f|) = 0 \tag{1}$$

for any $f \in L^0$ which satisfies

$$\int_0^1 \Phi(|f|) < \infty.$$

First, apply the conditional Jensen inequality $\Phi(\mathcal{E}_n|f|) \le \mathcal{E}_n(\Phi(|f|))$ to verify that f is integrable and thus that $S_{2^n} f \to f$, as $n \to \infty$, a.e. and in L^1 norm.

Next, fix $k \in \mathbf{P}$ and choose by Egoroff's theorem a measurable set $E_k \subseteq [0,1)$ such that $|E_k| < 1/k$ and $S_{2^n}f \to f$ uniformly as $n \to \infty$ on $\overline{E}_k := [0,1) \setminus E_k$. Write

$$\int_0^1 \Phi(|S_{2^n}f - f|) = \int_{E_k} \Phi(|S_{2^n}f - f|) + \int_{\overline{E}_k} \Phi(|S_{2^n}f - f|)$$
$$:= I_1^n + I_2^n$$

for each $n \in \mathbf{N}$. Since $\Phi(0) = 0$ it is clear that $I_2^n \to 0$ as $n \to \infty$. On the other hand, the condition on Φ implies

$$I_1^n \leq C \int_{E_k} \Phi(S_{2^n}|f|) + C \int_{E_k} \Phi(|f|).$$

Moreover,

$$\int_{E_k} \Phi(S_{2^n}|f|) \leq \int_{E_k} S_{2^n}(\Phi(|f|))$$

by the conditional Jensen inequality, and

$$\int_{E_k} S_{2^n}(\Phi(|f|)) \to \int_{E_k} \Phi(|f|) \quad \text{as} \quad n \to \infty$$

since $\Phi(|f|) \in \mathrm{L}^1$. It follows that

$$\limsup_{n\to\infty} I_1^n \leq 2C \int_{E_k} \Phi(|f|)$$

which verifies (1).

The following result gives a sufficient condition for convergence in L^1 norm of the full sequence of partial sums $S_n f$.

THEOREM 2. *If* $f \in \mathrm{L}^1$ *and*

$$\omega^{(1)}(f,\delta) = o\left(\log\frac{1}{\delta}\right)^{-1} \quad \text{as} \quad \delta \to 0 \tag{2}$$

then $S_n f \to f$ *in* L^1 *norm as* $n \to \infty$. *However, there exists an* $F \in \mathrm{L}^1$ *with*

$$\omega^{(1)}(F,\delta) = O\left(\log\frac{1}{\delta}\right)^{-1} \quad \text{as} \quad \delta \to 0$$

such that SF *does not converge in* L^1 *norm.*

PROOF. Suppose $f \in \mathrm{L}^1$ satisfies (2). For each $n \in \mathbf{N}$ write

$$n = 2^k + \ell \qquad (0 \leq \ell < 2^k)$$

and recall that

$$S_n f = S_{2^k} f + f * (r_k D_\ell)$$

(see Theorem 8 in 1.5). Since $S_{2^k} f \to f$ in L^1 norm it suffices to show

$$\sup_{0 \le \ell < 2^k} \|f * (r_k D_\ell)\|_1 \to 0 \quad \text{as} \quad k \to \infty. \tag{3}$$

Toward this, fix $0 \le \ell < 2^k$ and use the fact that $r_k(2^{-k-1}) = -1$ to write

$$(r_k D_\ell)(t) = -(r_k D_\ell)(t \dotplus 2^{-k-1})$$

for $t \in [0, 1)$. By Fubini's theorem we have

$$\begin{aligned} \|f * (r_k D_\ell)\|_1 &\le \frac{1}{2} \int_0^1 \int_0^1 |f(t) - f(t \dotplus 2^{-k-1})| \, |D_\ell(x \dotplus t)| \, dt \, dx \\ &= \frac{1}{2} \|D_\ell\|_1 \int_0^1 \int_0^1 |f(t) - f(t \dotplus 2^{-k-1})| \, dt. \end{aligned}$$

Since, $\|D_\ell\|_1 = O(\log \ell)$ as $\ell \to \infty$ (see the remarks following Theorem 9 in 1.6), it follows that

$$\|f * (r_k D_\ell)\|_1 \le C k \omega^{(1)}(f, 2^{-k})$$

for $k \in \mathbf{N}$ and some absolute constant C. Consequently, (2) implies (3), and the proof that Sf converges in L^1 norm is complete.

To construct the function F, set $n_m := 2^m$,

$$M_m := \sum_{k < m/2} n_k^2 \qquad (m \in \mathbf{N}),$$

and

$$F := \sum_{k=0}^{\infty} 2^{-k} r_{n_k} D_{2^{n_k}}.$$

Since $\|D_{2^m}\|_1 = 1$, it is clear that the series defining F converges in L^1 norm, and consequently is the Walsh-Fourier series of F. Hence by Theorem 9 in 1.6,

$$\begin{aligned} \|S_{2^{n_k} + M_{n_k}} F - S_{2^{n_k}} F\|_1 &= 2^{-k} \|S_{M_{n_k}}(D_{2^{n_k}})\|_1 \\ &= 2^{-k} \|D_{M_{n_k}}\|_1 \\ &\ge C_1 > 0 \end{aligned}$$

for $k \in \mathbf{N}$. In particular, the Walsh-Fourier series of F does not converge in L^1 norm.

To estimate $\omega^{(1)}(F, \delta)$, fix $0 < \delta < 1$, choose $m \in \mathbf{N}$ so that

$$2^{-n_{m+1}} \le \delta < 2^{-n_m},$$

and observe that

$$|F(x \dot{+} y) - F(x)| \le \sum_{k=m}^{\infty} 2^{-k}(D_{2^{n_k}}(x) + D_{2^{n_k}}(x \dot{+} y))$$

for every $0 \le y < \delta$. Therefore,

$$\begin{aligned}\int_0^1 |F(x \dot{+} y) - F(x)|\, dx &\le 4 \cdot 2^{-m} \\ &= \frac{4}{n_m} \\ &= O\left(\log \frac{1}{\delta}\right)^{-1}\end{aligned}$$

as $\delta \to 0$. ∎

4.2 Uniform Convergence of Walsh-Fourier Series. By Corollary 4 in 2.4, functions of bounded fluctuation which belong to $\mathrm{Lip}(\alpha, \mathbf{G})$ for some $\alpha > 0$ have uniformly convergent Walsh-Fourier series. We shall see presently that the Lipschitz condition can be relaxed considerably.

Several times in this chapter we shall use the following notation. If j is an integer of the form

$$j = \sum_{s=0}^{k-1} j_s 2^s$$

where $k \in \mathbf{P}$ and $j_s = 0$ or 1, then elements $x_j^{(k)}$ of $\mathbf{G}$ are defined by

$$x_j^{(k)} := (j_{k-1}, j_{k-2}, \ldots, j_0, 0, 0, \ldots).$$

(This is simply the group analogue of bit-reversal introduced in 1.4.) For each f defined on $\mathbf{G}$ and $x \in \mathbf{G}$, the numbers

$$U_k(x) := \sum_{j=1}^{2^k} \frac{1}{j} |f(x + x_{j-1}^{(k)}) - f(x + x_{j-1}^{(k)} + e_k)| \qquad (k \in \mathbf{N})$$

give an indication of how much f wiggles near x. (Here, $(e_k, k \in \mathbf{N})$ is the usual closed system in $\mathbf{G}$ introduced in 1.2).

THEOREM 3. *If $f \in \mathrm{C}(\mathbf{G})$ and $\|U_k\|_\infty \to 0$ as $k \to \infty$, then Sf converges uniformly on* **G**.

PROOF. Fix $k \in \mathbf{N}$ and $0 \le n < 2^k$. Set $N := 2^k + n$ and recall that $S_N f = S_{2^k} f + f * \rho_k D_n$. Since $S_{2^k} f \to f$ uniformly on $\mathbf{G}$ as $k \to \infty$, it suffices to prove

$$\lim_{k \to \infty} \max_{0 \le n < 2^k} \|f * (\rho_k D_n)\|_\infty = 0.$$

To show this fix $x \in \mathbf{G}$, $k \in \mathbf{N}$, $0 \le n < 2^k$ and write

$$\begin{aligned}(f*(\rho_k D_n))(x) &= \int_{\mathbf{G}} f(x+t)\rho_k(t)D_n(t)\,d\mu(t)\\ &= \sum_{j=0}^{2^k-1}\sum_{i=0}^{1}\int_{x_j^{(k)}+ie_k+I_{k+1}(0)} f(x+t)\rho_k(t)D_n(t)\,d\mu(t)\\ &= D_n(0)\int_{I_{k+1}(0)}\sum_{i=0}^{1}(-1)^i f(x+ie_k+t)\,d\mu(t)\\ &\quad+\int_{I_{k+1}(0)}\sum_{j=1}^{2^k-1} D_n(x_j^{(k)})\sum_{i=0}^{1}(-1)^i f(x+x_j^{(k)}+ie_k+t)\,d\mu(t)\\ &=: J_1+J_2.\end{aligned}$$

Since $D_n(0) = n < 2^k$ and $\mu(I_{k+1}(0)) = 2^{-k-1}$, it is clear that

$$|J_1| \le \frac{1}{2}\omega(f, 2^{-k}).$$

Since f is continuous, it follows that

$$\lim_{k\to\infty}\max_{0\le n<2^k}\|J_1\|_\infty = 0.$$

To estimate J_2, observe by definition that given $0 \le j < 2^k$ there is an integer ℓ_0 such that $x_j^{(k)} \in I_{\ell_0} \setminus I_{\ell_0+1}$ and $j < 2^{k-\ell_0}$. Consequently, Theorem 8 of 1.5 implies

$$\begin{aligned}|D_n(x_j^{(k)})| &= |\sum_{\ell=0}^{k-1} n_\ell(\rho_\ell D_{2^\ell})(x_j^{(k)})|\\ &= |\sum_{\ell=0}^{\ell_0} n_\ell\rho_\ell(x_j^{(k)})2^\ell|\\ &\le \sum_{\ell=0}^{\ell_0} n_\ell 2^\ell\\ &< 2^{\ell_0+1}\\ &< \frac{2^{k+1}}{j}.\end{aligned}$$

It follows that

$$\begin{aligned}|J_2| &\le \int_{I_{k+1}(0)}\sum_{j=1}^{2^k-1}|D_n(x_j^{(k)})|\,|f(x+x_j^{(k)}+t)-f(x+x_j^{(k)}+e_k+t)|\,d\mu(t)\\ &\le 2^{k+1}\int_{I_{k+1}(0)}\sum_{j=1}^{2^k-1}\frac{1}{j}|f(x+x_j^{(k)}+t)-f(x+x_j^{(k)}+e_k+t)|\,d\mu(t).\end{aligned}$$

In particular, the hypothesis implies

$$\lim_{k\to\infty} \max_{0\le n<2^k} \|J_2\|_\infty \le \lim_{k\to\infty} \|U_k\|_\infty = 0. \qquad \blacksquare$$

COROLLARY 1. *If $f \in C(\mathbf{G})$ is of bounded fluctuation, then Sf converges uniformly on $\mathbf{G}$.*

PROOF. Since $\omega(f, 2^{-k}) \to 0$ as $k \to \infty$, we may choose integers $m_0, m_1, \ldots$ such that $m_k < 2^k - 1$, $m_k \to \infty$ as $k \to \infty$, and

$$\lim_{k\to\infty} \omega(f, 2^{-k}) \log m_k = 0.$$

Observe by definition that

$$\begin{aligned}|U_k(x)| &\le \sum_{j=1}^{m_k} \frac{1}{j}\,\omega(f, 2^{-k}) + \sum_{j=m_k+1}^{2^k-1} \frac{1}{j}\omega(f, x + E_{kj}) \\ &\le (1 + \log m_k)\omega(f, 2^{-k}) + \frac{1}{m_k+1} \sum_{j=0}^{2^k-1} \omega(f, E_{kj})\end{aligned}$$

for $x \in \mathbf{G}$, $k \in \mathbf{N}$, and $E_{kj} := x_j^{(k)} + I_k(0)$. In view of Theorem 3, it suffices to show that the sum

$$\Delta := \sum_{j=0}^{2^k-1} \omega(f, E_{kj})$$

is uniformly bounded in k. However, under the identification of $\mathbf{G}$ with $[0, 1)$, the elements $x_j^{(k)}, 0 \le j < 2^k$, correspond to the dyadic rationals $p/2^k, 0 \le p < 2^k$. Consequently,

$$\Delta \le \sum_{p=0}^{2^k-1} \omega(f, I(p, k)).$$

This sequence is bounded in k because f is of bounded fluctuation. $\blacksquare$

By itself, continuity of f is not sufficient to conclude that Sf converges uniformly on $\mathbf{G}$. Indeed, fix $x \in \mathbf{G}$ and consider the functionals

$$\Lambda_n f := (S_n f)(x) \qquad (n \in \mathbf{N}, f \in C(\mathbf{G})).$$

Each Λ_n is linear. Moreover, we have by (49) in 1.5 that

$$|\Lambda_n f| \le \|f\|_\infty \|D_n\|_1.$$

In particular, each Λ_n is a bounded linear functional on $C(\mathbf{G})$ with functional norm satisfying

$$\|\Lambda_n\| \le \|D_n\|_1.$$

This inequality is an equality. Indeed, D_n is a dyadic step function and so is $g := \operatorname{sgn}(\tau_x D_n)$. Thus $g \in C(\mathbf{G})$. Evaluation of Λ_n at g shows $\|\Lambda_n\| = \|D_n\|_1$ for all $n \in \mathbf{N}$. But $\sup_{n\in\mathbf{N}} \|D_n\|_1 = \infty$. Therefore, it follows from the Banach- Steinhaus theorem that

$$\sup_{n\in\mathbf{N}} |\Lambda_n f| = \infty$$

for all f belonging to a dense $\mathcal{G}_\delta$ set in $C(\mathbf{G})$. In particular, given $x \in \mathbf{G}$ there is an $f \in C(\mathbf{G})$ such that $(S_n f)(x)$ does not converge, as $n \to \infty$.

Such a function cannot be found if its modulus of continuity decays rapidly enough. (Compare with Theorem 2 in 4.1.)

THEOREM 4. *Let $f \in C(\mathbf{G})$ and suppose that*

$$\omega(f,\delta) = o\left(\log\frac{1}{\delta}\right)^{-1} \quad \text{as} \quad \delta \to 0.$$

Then $S_n f$ converges to f uniformly on $\mathbf{G}$, as $n \to \infty$.

PROOF. Let $n \in \mathbf{N}$ and choose integers $k \in \mathbf{N}$ and $0 \le \ell < 2^k$ such that $n = 2^k + \ell$. Then

$$S_n f = S_{2^k} f + S_{k,\ell} f,$$

where

$$S_{k,\ell} f := f * (\psi_{2^k} D_\ell)$$

for $0 \le \ell < 2^k, k \in \mathbf{N}$. Since $S_{2^k} f$ converges to f uniformly on $\mathbf{G}$ it suffices to show

$$\lim_{k\to\infty} \max_{0\le\ell<2^k} \|S_{k,\ell} f\|_\infty = 0. \tag{4}$$

Toward (4), fix $x \in \mathbf{G}$ and recall that $\psi_{2^k}(e_k) = -1$ but $\psi_j(e_k) = 1$ for $0 \le j < 2^k$. Thus it is evident that

$$\begin{aligned} S_{k,\ell} f(x) &= \int_{\mathbf{G}} f(t)\psi_{2^k}(x+t) D_\ell(x+t)\, d\mu(t) \\ &= -\int_{\mathbf{G}} f(t)\psi_{2^k}(x+t+e_k) D_\ell(x+t+e_k)\, d\mu(t). \end{aligned}$$

Since μ is translation invariant, this identity implies

$$2\|S_{k,\ell}\|_\infty \le \sup_{x\in\mathbf{G}} \int_{\mathbf{G}} |f - \tau_{e_k} f|\, |\tau_x D_\ell|\, d\mu \le \omega(f, 2^{-k}) \|D_\ell\|_1.$$

In particular, (4) follows from hypothesis and the fact that

$$\|D_\ell\|_1 = O(\log \ell) = O(\log 2^k) \quad \text{as} \quad k \to \infty. \qquad \blacksquare$$

This result is called the Dini-Lipschitz theorem. We shall show (Theorem 6 in 4.3) that "o" cannot be replaced by "O".

In view of the fact that the Walsh-Fourier series of a continuous f may diverge at a point, it is natural to ask whether uniform estimates for the partial sums S_nf can be obtained. Since

$$\|S_nf\|_\infty = \|D_n * f\|_\infty \leq \|D_n\|_1 \|f\|_\infty$$

and $\|D_n\|_1 = O(\log n)$ as $n \to \infty$, an immediate answer to this question is

$$\|S_nf\|_\infty = O(\log n) \quad \text{as} \quad n \to \infty.$$

Thus the operators

$$\frac{1}{\log n} S_n : \mathrm{C}(\mathbf{G}) \to \mathrm{C}(\mathbf{G}) \qquad (n \geq 2)$$

are uniformly bounded. Since the collection of Walsh polynomials is dense in $\mathrm{C}(\mathbf{G})$ and

$$\lim_{n\to\infty} \frac{\|S_nP\|_\infty}{\log n} = 0$$

for all $P \in \mathcal{P}$, it follows from the Banach-Steinhaus theorem that

$$S_nf = o(\log n) \quad \text{as} \quad n \to \infty$$

uniformly on $\mathbf{G}$ for each $f \in \mathrm{C}(\mathbf{G})$. We shall see (Theorem 6 in 4.3) that this estimate cannot be improved.

Although the Walsh-Fourier series of a continuous function need not converge at every point, it is always uniformly Cesàro summable on $\mathbf{G}$. Indeed, since the Walsh polynomials are dense in $\mathrm{C}(\mathbf{G})$, $\|K_n\|_1 \leq 2$ for $n \in \mathbf{N}$ (see Theorem 16 in 1.8), and $\sigma_nP \to P$ uniformly on $\mathbf{G}$, as $n \to \infty$, for any $P \in \mathcal{P}$, it is easy to see that

$$\lim_{n\to\infty} \|\sigma_nf - f\|_\infty = 0$$

for every $f \in \mathrm{C}(\mathbf{G})$. (See also Exercise 2.10.)

4.3 Walsh-Fejér Polynomials. These are the functions $Q_0, Q_1 \ldots$ defined by

$$Q_n := \psi_{N_n} \sum_{k=0}^{2^{2n}-1} c_k\psi_k \qquad (n \in \mathbf{N}),$$

where $N_n := \sum_{k=0}^{n-1} 2^{2k}$ for $n \in \mathbf{N}$, $c_0 := 0$, and

$$c_{2^k+\ell} := (-1)^k 2^{-k}$$

for $0 \leq \ell < 2^k, k \in \mathbf{N}$.

The Walsh-Fejér polynomials enjoy a number of useful properties.

THEOREM 5

i)
$$\|S_kQ_n\|_\infty \le 2n \qquad (k, n \in \mathbf{N}),$$

ii)
$$Q_n(0) = 0 \quad \text{and} \quad \|Q_n\|_\infty \le 2 \qquad (n \in \mathbf{N}),$$

iii)
$$|(S_kQ_n)(x)| \le 4m \qquad (x \notin I_m(0), k \in \mathbf{N}, m, n \ge 2),$$

iv)
$$(S_{N_n}Q_n)(0) = n \qquad (n \in \mathbf{N}).$$

PROOF. To prove i) observe by definition that

$$\begin{aligned}\|S_kQ_n\|_\infty &\le \sum_{j=0}^{2n-1} \sum_{k=2^j}^{2^{j+1}-1} |c_k| \\ &= \sum_{j=0}^{2n-1} 1 \\ &= 2n\end{aligned}$$

for $k, n \in \mathbf{N}$.

To prove ii), fix $n \in \mathbf{N}$, set $I_k := I_k(0)$ and $J_k := I_k \setminus I_{k+1}$ for $k \in \mathbf{N}$. Use the Paley lemma to see that

$$2^{-k}\rho_k(x)D_{2^k}(x) = \begin{cases} 1 & x \in I_{k+1} \\ -1 & x \in J_k \\ 0 & x \notin I_k \end{cases}$$

holds for $k \in \mathbf{N}$. Thus for $j < 2n$ and $x \in J_j$ we have

$$\begin{aligned}|Q_n(x)| &= \Big| \sum_{k=0}^{2^{2n}-1} c_k\psi_k(x) \Big| \\ &= \Big| \sum_{k=0}^{j} (-1)^k 2^{-k}\rho_k(x)D_{2^k}(x) \Big| \\ &= \Big| \sum_{k=0}^{j-1} (-1)^k + (-1)^{j+1} \Big| \\ &\le 2.\end{aligned}$$

On the other hand, $Q_n(x) = \sum_{k=0}^{2n-1}(-1)^k = 0$ for $x \in I_{2n}$.

To prove iii), fix $x \notin I_m$, $m,n \geq 2$, $k \in \mathbf{N}$, and choose $j_0 \in \mathbf{N}$ so that $x \in J_{j_0}$. Then $j_0 < m$ and by i) we may suppose that $m < 2n$. Decompose $Q_n = Q_n^1 + Q_n^2$ where

$$Q_n^1 := (1 - \chi(J_{j_0}))Q_n \quad \text{and} \quad Q_n^2 := \chi(J_{j_0})Q_n.$$

Recall (see Theorem 8 in 1.5) that

$$(S_k Q_n^1)(x) = \int_{\mathbf{G}} Q_n^1(t)\psi_k(x+t) \sum_{j=0}^{\infty} k_j(\rho_j D_{2^j})(x+t)\, d\mu(t).$$

Since $Q_n^1(t)D_{2^j}(x+t) = 0$ for $j > j_0, t \in \mathbf{G}$, it follows that

$$|(S_k Q_n^1)(x)| \leq \int_{\mathbf{G}} \sum_{j=0}^{j_0} |Q_n^1(t)| D_{2^j}(x+t)\, d\mu(t).$$

Consequently, ii) and $\|D_{2^j}\|_1 = 1$ imply

$$|(S_k Q_n^1)(x)| \leq 2(j_0 + 1).$$

On the other hand, since $Q_n^2 = a\psi_{N_n}\chi(J_{j_0})$ for $a = -1$ or 2, we have for any $t \in \mathbf{G}$ that

$$\begin{aligned} Q_n^2(t) &= a\, 2^{-(j_0+1)} \psi_{N_n}(t) D_{2^{j_0+1}}(t + e_{j_0}) \\ &= a\, 2^{-(j_0+1)} \psi_{N_n}(t) \sum_{j=0}^{2^{j_0+1}-1} \psi_j(t + e_{j_0}). \end{aligned}$$

It follows that

$$|(S_k Q_n^2)(x)| \leq 2 \cdot 2^{-(j_0+1)} \sum_{j=0}^{2^{j_0+1}-1} 1 = 2.$$

Combining this estimate with the earlier one, we conclude that

$$\begin{aligned} |(S_k Q_n)(x)| &\leq 2(j_0 + 1) + 2 \\ &< 2m + 4 \\ &\leq 4m. \end{aligned}$$

To prove iv) observe by Theorem 3 in 1.2 and the definition of N_n that

$$\{0, 1, 2, ..., N_n - 1\} = N_n \oplus \bigcup_{k=0}^{n-1} A_{2k}$$

where $A_{2k} := \{\ell \in \mathbf{N} : 2^{2k} \leq \ell < 2^{2k+1}\}$ for $k \in \mathbf{N}$. Thus by the definition of the c_ℓ's we have

$$\begin{aligned} (S_{N_n} Q_n)(0) &= \sum_{k=0}^{n-1} \sum_{\ell \in A_{2k}} c_\ell \\ &= \sum_{k=0}^{n-1} \sum_{\ell \in A_{2k}} \frac{1}{2^{2k}} \\ &= n. \quad \blacksquare \end{aligned}$$

These polynomials can be used to construct functions $f \in C(\mathbf{G})$ with divergent Walsh-Fourier series.

THEOREM 6

i) There is a function $f \in C(\mathbf{G})$ and a set $E \subset \mathbf{G}$ such that $(S_n f, n \in \mathbf{N})$ is uniformly bounded, $E \cap I_k(x)$ is uncountable for all $x \in \mathbf{G}$ and $k \in \mathbf{N}$, but Sf diverges on E.

ii) There is a function $f \in C(\mathbf{G})$ whose Walsh-Fourier series is everywhere convergent, but not uniformly convergent.

iii) Given any sequence $(\lambda_n, n \in \mathbf{N})$ which satisfies $\lambda_n = o(\log n)$ as $n \to \infty$, there is a function $f \in C(\mathbf{G})$ such that

$$\limsup_{n\to\infty} \frac{|(S_n f)(0)|}{\lambda_n} = \infty.$$

iv) There is a function $f \in C(\mathbf{G})$ whose Walsh-Fourier series diverges at some point in $\mathbf{G}$, but whose modulus of continuity satisfies

$$\omega(f,\delta) = O\left(\log \frac{1}{\delta}\right)^{-1} \qquad \text{as} \quad \delta \to 0.$$

PROOF. Let $(a_k, k \in \mathbf{P})$ belong to ℓ^1, $(n_k, k \in \mathbf{P})$ be a strictly increasing sequence of positive integers, and $(x^k, k \in \mathbf{P})$ be a sequence of points in $\mathbf{G}$. Consider the function

$$f(x) := \sum_{k=1}^{\infty} a_k \rho_{2n_k}(x) Q_{n_k}(x + x^k) \qquad (x \in \mathbf{G}). \tag{5}$$

By Theorem 5 ii), it is clear that $f \in C(\mathbf{G})$. Moreover, since (5) converges uniformly and the spectrum of each term satisfies

$$sp(\rho_{2n_k} Q_{n_k}) \subseteq [2^{2n_k}, 2^{2n_k+1}), \tag{6}$$

it is clear that (5) represents the Walsh-Fourier series of f. Hence, if for any $k \in \mathbf{P}$ and $x \in \mathbf{G}$ we denote

$$(S_{2^{2n_k}+N_k} f)(x) - (S_{2^{2n_k}} f)(x)$$

by $\delta_k f(x)$, then

$$|\delta_k f(x)| = |a_k| \, |(S_{N_{n_k}} Q_{n_k})(x + x^k)|. \tag{7}$$

To prove i), let $a_k := 2^{-k}, n_k := 2^k$ for $k \in \mathbf{P}$, and let $(x^k, k \in \mathbf{P})$ represent a sequence of points in $\mathbf{G}$ to be specified below. By (6) it is clear that given any $n \in \mathbf{N}$ there is an index $k \in \mathbf{P}$ such that

$$(S_n f)(x) = \sum_{j=1}^{k-1} \frac{1}{2^j} \rho_{2^{j+1}}(x) Q_{2^j}(x + x^j) + a_k \rho_{2^{k+1}}(x) R(x),$$

where R is a partial sum of the polynomial $\tau_{x^k} Q_{n_k}$. It follows from Theorem 5 i) and ii) that

$$\|S_n f\|_\infty \le 2 \sum_{j=1}^{\infty} \frac{1}{2^j} + \frac{1}{2^k} 2 \cdot 2^k \le 4.$$

Therefore, the partial sums of Sf are uniformly bounded. Furthermore, by (7) and Theorem 5 iv) it is clear that

$$|\delta_k f(x^k)| \geq 1 \qquad (k \in \mathbf{P}). \tag{8}$$

In particular, given any $k \in \mathbf{P}$ there is an integer m_k so large that

$$|\delta_k f(x)| \geq 1 \qquad (x \in I_{m_k}(x^k)).$$

Hence the Walsh-Fourier series of f diverges on the set

$$E := \bigcap_{j=1}^{\infty} \bigcup_{k=j}^{\infty} I_{m_k}(x^k).$$

Now, given integers m_k one can choose points $x^k \in \mathbf{G}$ so that the sets $I_{m_k}(x^k)$ "fill up" $\mathbf{G}$. In particular, the points $x^k \in \mathbf{G}$ can be chosen so that each $E \cap I_\ell(x)$ is uncountable for all $x \in \mathbf{G}$. (Compare this with the construction of Cantor sets in 2.3.) This verifies i).

To prove ii) set $n_k := 2^k, a_k := 2^{-k}$, and $x^k := e_k$ for $k \in \mathbf{P}$. By (8), the series Sf cannot be uniformly convergent. To see that Sf is everywhere convergent, fix $x \in \mathbf{G}$, set

$$\Theta_k := \sup_{m \in \mathbf{N}} |(S_m Q_{n_k})(x + x^k)|$$

for $k \in \mathbf{P}$, and observe that it is enough to show that $2^{-k}\Theta_k \to 0$ as $k \to \infty$. If $x = 0$ then Theorem 5 iii) and the fact that $x^k \notin I_{k+2}(0)$ imply

$$\Theta_k \leq 4(k+2) \qquad (k \in \mathbf{P}).$$

If $x \neq 0$ then since $x^k \to 0$ as $k \to \infty$, there are integers $\ell_0, k_0 \in \mathbf{P}$ such that $x + x^k \notin I_{\ell_0}(0)$ for $k > k_0$. Hence by Theorem 5 iii),

$$\Theta_k \leq 4\ell_0.$$

In either case, then, $2^{-k}\Theta_k \to 0$ as $k \to \infty$.

To prove iii) we may assume that $\lambda_n = \varepsilon_n \log n$ for $n \in \mathbf{P}$, where both $\varepsilon_n \downarrow 0$ and $\lambda_n \to \infty$ as $n \to \infty$. Choose positive integers $n_1, n_2, \ldots$ so that

$$\sum_{k=1}^{\infty} \varepsilon_{n_k} < \infty.$$

Choose $q_k \uparrow \infty$ so that

$$\varepsilon := \sum_{k=1}^{\infty} \varepsilon_{n_k} q_k < \infty$$

and set $a_k := \varepsilon_{n_k} q_k$, $x^k := 0$ for $k \in \mathbf{P}$. Use (7) and Theorem 5 iv) to verify

$$\begin{aligned}|(S_{2^{2n_k}+N_{n_k}} f)(0)| &\geq a_k |(S_{N_{n_k}} Q_{n_k})(0)| - |(S_{2^{2n_k}} f)(0)| \\ &\geq a_k n_k - 2\varepsilon.\end{aligned}$$

Since $\varepsilon_n/\lambda_n \to 0$ as $n \to \infty$, it follows that

$$\begin{aligned}\limsup_{n\to\infty} \frac{|(S_n f)(0)|}{\lambda_n} &\geq \limsup_{k\to\infty} \frac{\varepsilon_{n_k} q_k n_k}{\lambda_{2^{2n_k}+N_{n_k}}} \\ &\geq \frac{1}{2} \limsup_{k\to\infty} q_k = \infty.\end{aligned}$$

To prove iv) let $n_k := 2^k, a_k := 2^{-k}$, and $x^k := 0$ for $k \in \mathbf{P}$. The proof of i) above verifies that $(Sf)(0)$ diverges. To show that the modulus of continuity of f grows as it should, set $n_0 := 0$ and choose an integer $j(y)$, depending on $y \in \mathbf{G} \setminus \{0\}$, which satisfies

$$y \in I_{2n_{j(y)}} \setminus I_{2n_{j(y)+1}}.$$

Let

$$f_1(x) := \sum_{k=1}^{j(y)-1} a_k \rho_{2n_k}(x) Q_{n_k}(x + x^k),$$

$$f_2(x) := \sum_{k=j(y)}^{\infty} a_k \rho_{2n_k}(x) Q_{n_k}(x + x^k) \qquad (x \in \mathbf{G}),$$

and observe that $f_1(x+y) = f_1(x)$ for $x \in \mathbf{G}$, and that $f = f_1 + f_2$. But Theorem 5 ii) implies

$$|f_2(x+y) - f_2(x)| \leq 4 \sum_{k=j(y)}^{\infty} a_k = \frac{8}{n_{j(y)}}$$

for all $x, y \in \mathbf{G}$. Consequently,

$$\begin{aligned}\omega(f,\delta) &\leq \sup_{|y|<\delta} \frac{8}{n_{j(y)}} \\ &\leq 32 \left(\log \frac{1}{\delta}\right)^{-1}. \qquad \blacksquare\end{aligned}$$

4.4 Summability of Walsh-Fourier Series in Homogeneous Banach Spaces. We have seen that the Walsh-Fourier series of a function $f \in \mathrm{C}(\mathbf{G})$ is uniformly Cesàro summable on $\mathbf{G}$. In this section, we shall investigate the norm convergence of certain sequences of dyadic convolution operators (see Theorem 7 below).

Let $\boldsymbol{X}$ be a Banach space with norm $\|\cdot\|$. The space $\boldsymbol{X}$ is called a *homogeneous Banach space* if $\mathcal{P} \subseteq \boldsymbol{X} \subseteq \mathrm{L}^1(\mathbf{G})$ and if the following three properties hold:

(i) $$\|f\|_1 \le \|f\| \qquad (f \in \boldsymbol{X}),$$

(ii) $$\tau_x f \in \boldsymbol{X}, \ \|\tau_x f\| = \|f\| \qquad (x \in \mathbf{G}, f \in \boldsymbol{X}),$$

and, given $f \in \boldsymbol{X}$ there is a sequence of Walsh polynomials $(P_n, n \in \mathbf{N})$ such that

(iii) $$\lim_{n\to\infty} \|P_n - f\| = 0.$$

It is clear that $L^p(\mathbf{G})$, dyadic $H^p, 1 \le p < \infty$, and $C(\mathbf{G})$ are homogeneous Banach spaces.

The familiar inequality $\|f * g\|_p \le \|f\|_1 \|g\|_p$ holds for any homogeneous Banach space:

LEMMA 1. *If $\boldsymbol{X}$ is a homogeneous Banach space, if $f \in L^1(\mathbf{G})$ and $g, h \in \boldsymbol{X}$, then*

$$\|f * g - h \int_{\mathbf{G}} f \, d\mu\| \le \int_{\mathbf{G}} \|\tau_t g - h\| \, |f(t)| \, d\mu(t).$$

In particular,

$$\|f * g\| \le \|f\|_1 \|g\|$$

for all $f \in L^1(\mathbf{G})$ and $g \in \boldsymbol{X}$.

PROOF. Fix $f \in L^1(\mathbf{G}), h \in \boldsymbol{X}$, and suppose for a moment that $g = P$ is a Walsh polynomial of order 2^n. For each $0 \le k \le 2^n$ with binary coefficients $(k_j, j \in \mathbf{N})$ define a point $x^k \in \mathbf{G}$ by

$$x^k := (k_0, k_1, ..., k_{n-1}, 0, 0, \dots).$$

Since

$$\mathbf{G} = \bigcup_{k=0}^{2^n-1} I_n(x^k)$$

we can choose real numbers c_k such that

$$P = \sum_{k=0}^{2^n-1} c_k \chi(I_n(x^k)).$$

Fix $x \in \mathbf{G}$ and observe that

$$\begin{aligned}(f * P)(x) - h(x) \int_{\mathbf{G}} f \, d\mu &= \int_{\mathbf{G}} f(t)(P(x+t) - h(x)) \, d\mu(t) \\ &= \sum_{k=0}^{2^n-1} \int_{I_n(x^k)} f(t)(\tau_t P(x) - h(x)) \, d\mu(t).\end{aligned}$$

In particular, it follows from the choice of the c_k's that

$$\begin{aligned}\|f * P - h \int_{\mathbf{G}} f\,d\mu\| &= \| \sum_{k=0}^{2^n-1} (\tau_{x^k} P - h) \int_{I_n(x^k)} f\,d\mu\| \\ &\le \sum_{k=0}^{2^n-1} \|\tau_{x^k} P - h\| \int_{I_n(x^k)} |f|\,d\mu \\ &= \sum_{k=0}^{2^n-1} \int_{I_n(x^k)} \|\tau_t P - h\|\,|f|\,d\mu(t) \\ &= \int_{\mathbf{G}} \|\tau_t P - h\|\,|f|\,d\mu(t).\end{aligned}$$

Hence the lemma is true when $g = P$ is a Walsh polynomial.

Suppose now that $g \in \boldsymbol{X}$. By condition (iii) we can choose Walsh polynomials P_n, for $n \in \mathbf{N}$, such that $\|g - P_n\| \to 0$ as $n \to \infty$. It follows from the case already considered that for all $m, n \in \mathbf{N}$

$$\begin{aligned}\|f * P_n - f * P_m\| &\le \int_{\mathbf{G}} \|\tau_t(P_n - P_m)\|\,|f(t)|\,d\mu(t) \\ &= \|P_n - P_m\|\,\|f\|_1.\end{aligned}$$

Hence $(f * P_n, n \in \mathbf{N})$ is Cauchy in $\boldsymbol{X}$. Since $\boldsymbol{X}$ is a Banach space and satisfies (i), it follows that $f * P_n$ converges in $\mathrm{L}^1(\mathbf{G})$ norm. But

$$\begin{aligned}\|f * P_n - f * g\|_1 &\le \|f\|_1 \|P_n - g\|_1 \\ &\le \|f\|_1 \|P_n - g\|,\end{aligned}$$

for all $n \in \mathbf{N}$, so $f * P_n \to f * g$ in $\boldsymbol{X}$ as $n \to \infty$. Consequently,

$$\begin{aligned}\|f * g - h \int_{\mathbf{G}} f\,d\mu\| &= \lim_{n\to\infty} \|f * P_n - h \int_{\mathbf{G}} f\,d\mu\| \\ &\le \lim_{n\to\infty} \int_{\mathbf{G}} \|\tau_t P_n - h\|\,|f(t)|\,d\mu(t).\end{aligned}$$

In particular, an easy calculation verifies the inequalities. ▮

A sequence $(P_n, n \in \mathbf{N})$ of functions in $\mathrm{C}(\mathbf{G})$ is called an *approximate identity* (for dyadic convolution) if

(9) $$\|P_n\|_1 = 0(1) \quad \text{as} \quad n \to \infty,$$

(10) $$\int_{\mathbf{G}} P_n\,d\mu = 1 \qquad (n \in \mathbf{N}),$$

and

(11) $$\lim_{n\to\infty} \int_{\mathbf{G}\setminus I_k(0)} |P_n|\,d\mu = 0$$

for each $k \in \mathbf{N}$.

It is clear by the Paley lemma that $(D_{2^n}, n \in \mathbf{N})$ is an approximate identity. The Fejér kernels also form an approximate identity:

LEMMA 2. *The sequence*

$$K_n = \frac{1}{n}\sum_{i=1}^{n} D_i \qquad (n \in \mathbf{P})$$

is an approximate identity.

PROOF. By Theorem 16 v) in 1.8, $\|K_n\|_1 \leq 2$ for $n \in \mathbf{P}$. Hence (9) is satisfied.

Since $\int_{\mathbf{G}} D_n \, d\mu = 1$ for $n \in \mathbf{P}$, it is clear that (10) is satisfied.

To verify (11), recall that $D_{2^j} = 0$ on $\mathbf{G} \setminus I_k(0)$ for $j \geq k$. Hence by Theorem 16 iv) in 1.8, if $n \in \mathbf{P}$ and $2^{n-1} \leq m < 2^n$ then

$$\begin{aligned}\int_{\mathbf{G}\setminus I_k(0)} |K_m| \, d\mu &\leq \sum_{i=0}^{k-1} 2^{i-n} \sum_{j=i}^{k-1} \int_{\mathbf{G}\setminus I_k(0)} D_{2^j} \, d\mu \\ &\quad + \sum_{i=0}^{n-1} 2^{i-n} \sum_{j=i}^{n-1} \int_{\mathbf{G}\setminus I_k(0)} \tau_{e_i} D_{2^j} \\ &:= J_1 + J_2.\end{aligned}$$

Since k is fixed and $n \to \infty$ as $m \to \infty$, it is clear that

$$\lim_{m\to\infty} J_1 = 0.$$

To estimate J_2, observe that

$$(\mathbf{G} \setminus I_k(0)) \cap I_j(e_i) = \emptyset$$

when $\min\{i, j\} \geq k$. Consequently,

$$\begin{aligned}|J_2| &= |\sum_{i=0}^{k-1} 2^{i-n} \sum_{j=i}^{n-1} \int_{\mathbf{G}\setminus I_k(0)} \tau_{e_i} D_{2^j}| \\ &\leq 2^{-n} \sum_{i=0}^{k-1} 2^i (n-i).\end{aligned}$$

Since $n2^{-n} \to 0$ as $n \to \infty$, it follows that

$$\lim_{m\to\infty} J_2 = 0. \qquad \blacksquare$$

The following result shows that in any homogeneous Banach space, $S_{2^n} f$ and $\sigma_n f$ converge to f in norm as $n \to \infty$.

THEOREM 7. *Let $(P_n, n \in \mathbf{N})$ be an approximate identity, and $\boldsymbol{X}$ be a homogeneous Banach space. Then*

$$\lim_{n\to\infty} \|P_n * f - f\| = 0$$

for all $f \in \boldsymbol{X}$.

PROOF. For each $n \in \mathbf{N}$ set

$$\Lambda_n f := P_n * f.$$

By Lemma 1 the linear operators $\Lambda_n : \boldsymbol{X} \to \boldsymbol{X}$ are uniformly bounded for $n \in \mathbf{P}$. Moreover, the collection $\{\boldsymbol{\tau}_x D_{2^k} : x \in \mathbf{G} \text{ and } k \in \mathbf{N}\}$ is closed in $\boldsymbol{X}$ and by (ii) satisfies

$$\begin{aligned}\|P_n * \boldsymbol{\tau}_x D_{2^k} - \boldsymbol{\tau}_x D_{2^k}\| &= \|\boldsymbol{\tau}_x(P_n * D_{2^k} - D_{2^k})\| \\ &= \|P_n * D_{2^k} - D_{2^k}\| \qquad (n, k \in \mathbf{N}).\end{aligned}$$

Hence by the Banach-Steinhaus theorem we need only consider the special case $f := D_{2^k}$ for some fixed $k \in \mathbf{N}$.

Using the notation introduced in 4.2, let

$$\mathbf{G}_k := \{y \in \mathbf{G} : y = x_j^{(k)} \quad \text{for some} \quad 0 \le j < 2^k\}.$$

Fix $n \in \mathbf{N}$ and write

$$\boldsymbol{\Delta}(n) := \|P_n * D_{2^k} - D_{2^k}\| = \|\sum_{y\in\mathbf{G}_k} \boldsymbol{\tau}_y D_{2^k} \int_{I_k(y)} P_n \, d\mu - D_{2^k}\|.$$

By (10) we have

$$\boldsymbol{\Delta}(n) = \|\sum_{y\in\mathbf{G}_k} (\boldsymbol{\tau}_y D_{2^k} - D_{2^k}) \int_{I_k(y)} P_n \, d\mu\|.$$

Separating the term $y = 0$ from other y's in $\mathbf{G}_k$, and using translation invariance of the norm of $\boldsymbol{X}$, we obtain the following estimate:

$$\begin{aligned}\boldsymbol{\Delta}(n) &\le \sum_{y\in\mathbf{G}_k, y\neq 0} \|\boldsymbol{\tau}_y D_{2^k} - D_{2^k}\| \int_{I_k(y)} |P_n| \, d\mu \\ &\le 2\|D_{2^k}\| \int_{\mathbf{G}\setminus I_k(0)} |P_n| \, d\mu.\end{aligned}$$

Therefore, (11) implies that $\boldsymbol{\Delta}(n) \to 0$ as $n \to \infty$. ∎

We shall denote the Cesàro means of a Walsh series $S := \sum_{k=0}^{\infty} a_k \psi_k$ by

$$\sigma_n := \frac{1}{n} \sum_{k=1}^{n} S_k \qquad (n \in \mathbf{P}),$$

where for each $k \in \mathbf{P}$

$$S_k := \sum_{j=0}^{k-1} a_j \psi_j.$$

We close this section with a characterization, in terms of σ_n, of Walsh-Fourier series and Walsh-Fourier-Stieltjes series.

THEOREM 8. *Let $S = \sum_{k=0}^{\infty} a_k\psi_k$ be a Walsh series and $\boldsymbol{X}$ be a homogeneous Banach space. Then*

i) S is the Walsh-Fourier series of a function $f \in \boldsymbol{X}$ if and only if σ_n converges in $\boldsymbol{X}$ as $n \to \infty$.

ii) S is the Walsh-Fourier series of a function $f \in \mathrm{L}^\infty(\mathbf{G})$ if and only if there exists a constant M such that

$$\|\sigma_n\|_\infty \le M < \infty \qquad (n \in \mathbf{N}). \tag{12}$$

iii) S is the Walsh-Fourier-Stieltjes series of a finite Borel measure ν on $\mathbf{G}$ if and only if there exists a constant M such that

$$\|\sigma_n\|_1 \le M < \infty \qquad (n \in \mathbf{N}). \tag{13}$$

PROOF. To prove i), observe by Theorem 7 that if $S = Sf$ for some $f \in \boldsymbol{X}$ then $\lim_{n\to\infty} \sigma_n = f$ in the norm of $\boldsymbol{X}$. Conversely, if σ_n converges in $\boldsymbol{X}$, as $n \to \infty$, let f denote its limit. Then $f \in \boldsymbol{X}$ and $\sigma_n \to f$ in L^1 norm. In particular, for each $j \in \mathbf{N}$ we have by orthogonality that

$$\begin{aligned}
\widehat{f}(j) &= \lim_{n\to\infty} \int_{\mathbf{G}} \sigma_n \psi_j \, d\mu \\
&= \lim_{n\to\infty} \sum_{k=0}^{n-1} (1 - \frac{k}{n}) a_k \int_{\mathbf{G}} \psi_k \psi_j \, d\mu \\
&= \lim_{n\to\infty} (1 - \frac{j}{n}) a_j \\
&= a_j.
\end{aligned}$$

To prove ii), observe that if $S = Sf$ for some $f \in \mathrm{L}^\infty(\mathbf{G})$ then

$$\|\sigma_n f\|_\infty = \|K_n * f\|_\infty \le \|K_n\|_1 \|f\|_\infty \qquad (n \in \mathbf{P}).$$

In particular, (12) holds with $M = 2\|f\|_\infty$ by Theorem 16 in 1.8. Conversely, if (12) holds then for each integer $N < n$ we have

$$\begin{aligned}
\sum_{j=0}^{N} |a_j|^2 (1 - \frac{j}{n})^2 &\le \sum_{j=0}^{n-1} |a_j|^2 (1 - \frac{j}{n})^2 \\
&= \int_{\mathbf{G}} |\sigma_n|^2 \, d\mu \\
&\le M^2.
\end{aligned}$$

Let $n \to \infty$ and then $N \to \infty$. We obtain

$$\sum_{j=0}^{\infty} |a_j|^2 \le M^2.$$

In particular, by the Riesz-Fischer theorem (see Appendix 0.1), S is the Walsh-Fourier series of some function $f \in L^2(\mathbf{G})$. By Theorem 7, $\sigma_n \to f$ in $L^2(\mathbf{G})$, as $n \to \infty$. Thus there exists a subsequence of integers $n_1 < n_2 < \ldots$ such that $\sigma_{n_k} \to f$ a.e. as $k \to \infty$. Therefore, it follows from (12) that $f \in L^\infty(\mathbf{G})$.

To prove iii), observe that if S is the Walsh-Fourier-Stieltjes series of a finite Borel measure ν of total variation $\|\nu\|$, then by definition and Fubini's theorem we have for $n \in \mathbf{P}$ that

$$\begin{aligned}\|\sigma_n\|_1 &= \int_{\mathbf{G}} \left| \int_{\mathbf{G}} K_n(x+y)\, d\nu(y) \right| d\mu(x) \\ &\leq \int_{\mathbf{G}} \int_{\mathbf{G}} |K_n(x+y)|\, d\mu(x)\, d\nu(y).\end{aligned}$$

Since μ is translation invariant and $\|K_n\|_1 \leq 2$, it follows that (13) holds with $M = 2\|\nu\|$.

Conversely, suppose (13) holds. The functionals

$$\Lambda_n f := \int_{\mathbf{G}} f \sigma_n \, d\mu \qquad (n \in \mathbf{P}, f \in C(\mathbf{G}))$$

are bounded and linear on $C(\mathbf{G})$. In fact, by (13) they satisfy

$$|\Lambda_n f| \leq M \|f\|_\infty$$

for all $n \in \mathbf{P}$ and $f \in C(\mathbf{G})$. It follows from the Banach- Alaoglu theorem (Appendix 0.0) that there is a subsequence of integers $n_1, n_2, \ldots$ and a finite Borel measure ν on $\mathbf{G}$ such that

$$\int_{\mathbf{G}} f \, d\nu = \lim_{k \to \infty} \int_{\mathbf{G}} f \sigma_{n_k} \, d\mu$$

for all $f \in C(\mathbf{G})$. Applying this identity to $f = \psi_j$ yields $\widehat{\nu}(j) = a_j$ for all $j \in \mathbf{N}$. ∎

The important thing in Theorem 8 is that the Fejér kernels are uniformly L^1 bounded. Thus Theorem 8 also holds for 2^n-th partial sums of Walsh series (see Exercises 4.6 through 4.8).

4.5 Sets of Divergence. Let $\boldsymbol{X}$ be a homogeneous Banach space. A set $E \subseteq \mathbf{G}$ is called a *set of divergence* for $\boldsymbol{X}$ if there exists a function $f \in \boldsymbol{X}$ whose Walsh-Fourier series diverges on E.

Recall that the maximal function of an $f \in L^1(\mathbf{G})$ is defined by

$$S^* f := \sup_{n \in \mathbf{P}} |S_n f|.$$

It is clear that S^* is a sublinear map from $L^1(\mathbf{G})$ into the collection of measurable functions on [0,1), that $S^* f = 0$ if only if $f = 0$ a.e. on [0,1), and that the Walsh-Fourier series of f diverges at an $x \in \mathbf{G}$ whenever $(S^* f)(x) = \infty$.

Sets of divergence for homogeneous Banach spaces can be characterized by using the maximal operator S^*.

THEOREM 9. *If $\boldsymbol{X}$ is a homogeneous Banach space and if E is a set of divergence for $\boldsymbol{X}$ then there is a function $f \in \boldsymbol{X}$ such that $S^*f = \infty$ on E.*

PROOF. We claim that given any $g \in \boldsymbol{X}$, there is an unbounded monotone increasing sequence $\boldsymbol{\lambda} = (\lambda_j, j \in \mathbf{N})$ of positive real numbers and a function $f \in \boldsymbol{X}$ such that

$$\widehat{f}(j) = \lambda_j \widehat{g}(j) \qquad (j \in \mathbf{N}). \tag{14}$$

To prove this claim use Theorem 7 to choose a strictly increasing sequence of positive integers $n_1, n_2, \ldots$ such that

$$\|S_{2^{n_k}} g - g\| < 2^{-k} \qquad (k \in \mathbf{P}). \tag{15}$$

Consider the function f defined by

$$f := g + \sum_{k=1}^{\infty} (g - S_{2^{n_k}} g).$$

By (15), this series converges in the norm of $\boldsymbol{X}$, hence in $\mathrm{L}^1(\mathbf{G})$ norm as well. In particular, f belongs to $\boldsymbol{X}$ and

$$\widehat{f}(j) = \widehat{g}(j) + \sum_{k=1}^{\infty} \int_{\mathbf{G}} (g - S_{2^{n_k}} g)\psi_j \, d\mu$$

for $j \in \mathbf{N}$. Therefore, the claim follows from orthogonality if we set

$$\lambda_j := 1 + \sum_{2^{n_k} \le j} 1 \qquad (j \in \mathbf{N}).$$

To prove the theorem, suppose that $g \in \boldsymbol{X}$ is a function whose Walsh-Fourier series diverges on E. Use the claim to choose a monotone increasing, unbounded sequence $\boldsymbol{\lambda}$ which satisfies (14). By Abel's transformation,

$$\begin{aligned} S_n g - S_m g &= \sum_{j=m}^{n-1} (S_{j+1} f - S_j f) \frac{1}{\lambda_j} \\ &= \frac{S_n f}{\lambda_{n-1}} - \frac{S_m f}{\lambda_m} + \sum_{j=m+1}^{n-1} \left(\frac{1}{\lambda_{j-1}} - \frac{1}{\lambda_j} \right) (S_j f) \end{aligned}$$

for any integers $n, m \in \mathbf{N}$ with $n > m$. Since $\boldsymbol{\lambda}$ is increasing, it follows that

$$|S_n g - S_m g| \le \frac{2}{\lambda_m} S^* f \qquad (n, m \in \mathbf{N}, n > m).$$

Since $\boldsymbol{\lambda}$ is unbounded, it follows that Sg converges at x when $S^*f(x)$ is finite. In particular, $(S^*f)(x) = \infty$ for all $x \in E$. ∎

A homogeneous Banach space $\boldsymbol{X}$ is called *shift invariant* if

$$\psi_n f \in \boldsymbol{X}, \ \|\psi_n f\| = \|f\|$$

for all $f \in \boldsymbol{X}$ and $n \in \mathbf{N}$. Notice that the dyadic Hardy space H is not shift invariant. However, $\mathrm{C}(\mathbf{G})$ and $\mathrm{L}^p(\mathbf{G})$ are all shift invariant for $1 \le p < \infty$. Sets of divergence in shift invariant homogeneous Banach spaces have the following useful characterization.

Theorem 10. *Let $\mathbf{X}$ be a shift invariant, homogeneous Banach space. A set $E \subseteq \mathbf{G}$ is a set of divergence for $\mathbf{X}$ if and only if there exist Walsh polynomials $P_1, P_2, \ldots$ such that*

$$\sum_{j=1}^{\infty} \|P_j\| < \infty \tag{16}$$

and

$$\sup_{j \in \mathbf{P}} S^* P_j(x) = \infty \qquad (x \in E). \tag{17}$$

Proof. Suppose first that E is a set of divergence for $\mathbf{X}$. Let $g \in \mathbf{X}$ be a function whose Walsh-Fourier series diverges on E. By repeating the proof of Theorem 9, we can choose an unbounded, monotone increasing positive sequence $(\lambda_j, j \in \mathbf{N})$ and a function $f \in \mathbf{X}$ such that

$$S_n g - S_m g = \frac{S_n f}{\lambda_{n-1}} - \frac{S_m f}{\lambda_m} + \sum_{j=m+1}^{n-1} \left(\frac{1}{\lambda_{j-1}} - \frac{1}{\lambda_j} \right) (S_j f)$$

for all integers $n, m \in \mathbf{N}, m < n$. In particular, if $(\omega_j, j \in \mathbf{N})$ is any sequence of positive numbers which satisfies $\omega_j = o(\lambda_j)$ as $j \to \infty$, and

$$\sum_{j=1}^{\infty} (\frac{1}{\lambda_j} - \frac{1}{\lambda_{j+1}}) \omega_j < \infty,$$

then $|(S_j f)(x)| \neq O(\omega_j)$ as $j \to \infty$, for any $x \in E$. Consequently, we can choose an unbounded, monotone increasing sequence $(\omega_n, n \in \mathbf{N})$ such that for each $x \in E$, the inequality

$$|(S_n f)(x)| > \omega_n \tag{18}$$

holds for infinitely many integers $n \in \mathbf{N}$.

Use Theorem 7 to choose strictly increasing sequences of positive integers $(n_j, j \in \mathbf{N})$ and $(m_j, j \in \mathbf{N})$ which satisfy $n_j < m_j + 1$,

$$\|f - S_{2^{n_j}} f\| < 2^{-j}, \tag{19}$$

and

$$\|S^*(S_{2^{n_j}} f)\|_\infty < \frac{\omega_{m_j}}{2} \qquad (j \in \mathbf{N}). \tag{20}$$

Consider the functions defined by

$$P_j := S_{2^{m_{j+1}}} (f - S_{2^{n_j}} f) \qquad (j \in \mathbf{P}).$$

Clearly, these functions are Walsh polynomials. We will show that they satisfy (16) and (17). Since $\|S_{2^n}g\| \le \|g\|$ for $n \in \mathbf{N}$ and $g \in \boldsymbol{X}$, (16) is a direct consequence of (19). To verify (17), fix $x \in E$ and choose an $n \in \mathbf{N}$ satisfying (18) which is large enough that $m_j < n \le m_{j+1}$ for some $j \in \mathbf{P}$. Since the definition of P_j implies

$$S_nP_j = S_nf - S_n(S_{2^{n_j}}f),$$

we have by (18) and (20) that

$$|(S_nP_j)(x)| \ge |(S_nf)(x)| - \frac{\omega_{m_j}}{2} \ge \frac{1}{2}\omega_{m_j}.$$

Hence (17) follows from the fact that $\omega_m \to \infty$ as $m \to \infty$.

Conversely, suppose that

$$P_j := \sum_{k=0}^{2^{m_j}-1} c_k^{(j)}\psi_k \qquad (m_j \in \mathbf{N},\ j \in \mathbf{P})$$

is a sequence of Walsh polynomials which satisfies (16) and (17). Let $n_1 := m_1$ and for $j > 1$ set $n_j := 1 + \max\{n_{j-1}, m_j\}$. Then $(n_j, j \in \mathbf{N})$ is a strictly increasing sequence of integers and it is easy to see that

$$2^{n_{j+1}} \oplus k_1 > 2^{n_j} \oplus k_0 \tag{21}$$

for any choice of integers k_1 and k_0 which satisfy $0 \le k_0 \le 2^{m_j}, 0 \le k_1 \le 2^{m_{j+1}}$ and $j \in \mathbf{N}$.

Let

$$f := \sum_{j=1}^{\infty} \psi_{2^{n_j}} P_j$$

and observe by (16) that $f \in \boldsymbol{X}$. Since the norm of $\boldsymbol{X}$ dominates the L^1 norm, it is clear that the series defining f converges in $\mathrm{L}^1(\mathbf{G})$ norm. Consequently, this series is the Walsh-Fourier series of f. Moreover, (21) can be used to see that

$$S_{2^{n_j}+k}f - S_{2^{n_j}}f = \psi_{2^{n_j}} S_k P_j$$

for $0 \le k < 2^{n_{j+1}} - 2^{n_j}, j \in \mathbf{P}$. In particular, (17) implies the Walsh-Fourier series of f diverges at each $x \in E$. ∎

COROLLARY 2. *If $\boldsymbol{X}$ is a shift invariant, homogeneous Banach space, and $E_1, E_2, \ldots$ are sets of divergence for $\boldsymbol{X}$, then $E := \bigcup_{n=1}^{\infty} E_n$ is a set of divergence for $\boldsymbol{X}$.*

PROOF. Apply Theorem 10 to choose Walsh polynomials $P_1^{(n)}, P_2^{(n)}, \ldots$ such that

$$\sum_{j=1}^{\infty} \|P_j^{(n)}\| < \infty$$

and

$$\sup_{j\in\mathbf{P}}(S^*P_j^{(n)})(x) = \infty \qquad (x \in E_n,\ n \in \mathbf{P}). \tag{22}$$

Thus there exist integers $m_1 < m_2 < \ldots$ such that

$$\sum_{j=m_n}^{\infty} \|P_j^{(n)}\| < \frac{1}{2^n} \qquad (n \in \mathbf{P}).$$

Let $(Q_j, j \in \mathbf{P})$ be any enumeration of the polynomials $\{P_j^{(n)} : j \geq m_n, n = 1, 2, \ldots\}$, e.g., $Q_1 := P_{m_1}^{(1)}, Q_2 := P_{m_2}^{(2)}, Q_3 := P_{m_1+1}^{(1)}, \ldots$. Each Q_j is a Walsh polynomial and

$$\sum_{j=1}^{\infty} \|Q_j\| \leq \sum_{n=1}^{\infty} \frac{1}{2^n} < \infty.$$

In particular, by Theorem 10 it suffices to show that

$$\sup_{j\in\mathbf{P}}(S^*Q_j)(x) = \infty$$

for $x \in E$. But this follows from the construction and from (22) since every $x \in E$ necessarily belongs to some E_n. ▮

We turn our attention to specific homogeneous Banach spaces.

THEOREM 11. *If $1 \leq p < \infty$ and $E \subseteq \mathbf{G}$ is a set of Haar measure zero then E is a set of divergence for $L^p(\mathbf{G})$.*

PROOF. We begin with a general remark. If $A \subseteq \mathbf{G}$ is a finite union of dyadic intervals $I_1, I_2, \ldots, I_n$ for some $n \in \mathbf{P}$, and if N is any non-negative integer, then there exists a Walsh polynomial

$$P = \sum_{k=2^N}^{2^M-1} c_k \psi_k$$

for some $M \in \mathbf{N}$ such that

$$|P(x)| = 1 \qquad (x \in A),$$

and

$$\int_{\mathbf{G}} |P|^p \, d\mu = \mu(A).$$

Indeed, if $k_0 := \max\{2^N, 1/\mu(I_j) : 1 \leq j \leq n\}$, then

$$P := \chi(A)\psi_{k_0}$$

is such a polynomial.

To prove the theorem, suppose $E \subseteq \mathbf{G}$ satisfies $\mu(E) = 0$. Cover E with dyadic intervals $(I_k, k \in \mathbf{N})$ such that

$$\sum_{k=0}^{\infty} \mu(I_k) < 1$$

and each $x \in E$ belongs to infinitely many of the sets I_k. Set $n_0 := 0$ and choose integers $n_0 < n_1 < n_2 \ldots$ such that

$$\sum_{k=n_j}^{\infty} \mu(I_j) < 2^{-j} \qquad (j \in \mathbf{N}).$$

Apply the general remark above successively to the sets

$$A_j := \bigcup_{k=n_j}^{n_{j+1}-1} I_k \qquad (j \in \mathbf{N})$$

to generate integers $m_0 := 0 < m_1 < m_2 < \ldots$ and Walsh polynomials $P_0, P_1, \ldots$ such that

$$sp(P_j) \subset [2^{m_j}, 2^{m_{j+1}}), \tag{23}$$

$$\|P_j\|_p^p \leq 2^{-j}, \tag{24}$$

and

$$|P_j(x)| = 1 \qquad (x \in A_j) \tag{25}$$

for $j \in \mathbf{N}$.

Set

$$f := \sum_{j=1}^{\infty} P_j,$$

observing by (24) that this series converges in $L^p(\mathbf{G})$ norm. Hence $f \in L^p(\mathbf{G})$ and this series is the Walsh-Fourier series of f. Moreover, since the spectra of the polynomials P_j are pairwise disjoint, we have

$$S_{2^{m_{j+1}}} f - S_{2^{m_j}} f = P_j \qquad (j \in \mathbf{N}).$$

Since every $x \in E$ belongs to infinitely many of the sets A_j, it follows from (25) that the Walsh-Fourier series of f diverges at every point $x \in E$. ∎

This theorem cannot be improved for $1 < p < \infty$. Indeed, in this case the Walsh-Fourier series of an $f \in L^p$ converges a.e. (see Theorem 14 in 3.7).

It can be improved considerably for $p = 1$.

Theorem 12. *There is a function $f \in L^1(\mathbf{G})$ whose Walsh-Fourier series diverges everywhere.*

Proof. Fix $n \in \mathbf{P}$ and use the remarks preceding Theorem 10 in 1.6 to choose a number $C > 0$ and an integer $m_n \in [2^{n-1}, 2^n)$ such that

$$\|\sum_{k=0}^{m_n-1} \psi_k\|_1 > Cn.$$

The constant C does not depend on n.

Consider the function

$$g_n := \operatorname{sgn}\left(\sum_{k=0}^{m_n-1} \psi_k\right).$$

It is constant on any set of the form $I(k,n), 0 \le k < 2^n$. Hence g_n is a Walsh polynomial of order at most 2^n. Moreover, by the choice of m_n we have

$$(S_{m_n} g_n)(x) > Cn \qquad (x \in I_n(0)).$$

Set

$$Q_n := \prod_{k=0}^{2^n-1} (1 + \rho_{n+k} \tau_{x_k^{(n)}} g_n)$$

where the points $x_k^{(n)} \in \mathbf{G}, 0 \le k < 2^n$ were defined in 4.2. Clearly, Q_n is a Walsh polynomial of order at most 2^{2^n+n}, and $Q_n \ge 0$.

By expanding the product used to define Q_n, it is easy to see for $k = 0, 1, ..., 2^n - 1$ that

$$S_{2^{n+k}+m_n} Q_n - S_{2^{n+k}} Q_n = \rho_{n+k} S_{m_n}(\tau_{x_k^{(n)}} g_n).$$

The choice of the integers m_n therefore imply

$$|(S_{2^{n+k}+m_n} Q_n)(x) - (S_{2^{n+k}} Q_n)(x)| > Cn \qquad (x \in I_n(x_k^{(n)})). \tag{26}$$

Moreover, since the terms of the expanded product have pairwise disjoint spectra, it follows that

$$\|Q_n\|_1 = 1. \tag{27}$$

Let $n_1 < n_2 < ...$ be positive integers chosen so that

$$\sum_{j=1}^{\infty} \frac{1}{\sqrt{n_j}} < \infty,$$

and set

$$P_j := \frac{Q_{n_j}}{\sqrt{n_j}} \qquad (j \in \mathbf{P}).$$

By (27) it is evident that

$$\sum_{j=1}^{\infty} \|P_j\|_1 < \infty.$$

Moreover, for a fixed $x \in \mathbf{G}$ it is possible to choose integers $0 \le k(j) < 2^{n_j}$ such that

$$x \in I_{n_j}(x_{k(j)}^{(n_j)}) \qquad (j \in \mathbf{P}).$$

Hence (26) implies

$$(S^* P_j)(x) \ge \frac{C}{2}\sqrt{n_j}$$

for $j \in \mathbf{P}$, and $x \in \mathbf{G}$. Consequently, $\mathbf{G}$ is a set of divergence for $\mathrm{L}^1(\mathbf{G})$ by Theorem 10. ∎

What about sets of divergence for $\mathrm{C}(\mathbf{G})$? We saw in 4.3 that every singleton is a set of divergence for $\mathrm{C}(\mathbf{G})$. Hence it follows from Corollary 2 above that every countable subset of $\mathbf{G}$ is a set of divergence for $\mathrm{C}(\mathbf{G})$. On the other hand, no set of divergence for $\mathrm{C}(\mathbf{G})$ can be of positive μ measure (see Theorem 14 in 3.7). It is an open question whether every subset of $\mathbf{G}$ of μ zero is a set of divergence for $\mathrm{C}(\mathbf{G})$. A partial answer to this question is given by the following result.

THEOREM 13. *If $E \subseteq \mathbf{G}$ is compact and $\mu(E) = 0$, then E is a set of divergence for* $\mathrm{C}(\mathbf{G})$.

PROOF. By Theorem 10 we need to show there exist Walsh polynomials $P_1, P_2, \dots$ such that

$$\sum_{j=1}^{\infty} \|P_j\|_\infty < \infty$$

and

$$\sup_{j \in \mathbf{P}} (S^* P_j)(x) = \infty$$

for all $x \in E$. If we can construct Walsh polynomials $R_1, R_2, \dots$ such that $\|R_\ell\|_\infty \le 1$ and

$$(S^* R_\ell)(x) \ge \frac{\ell}{2}$$

for $x \in E$ and $\ell \in \mathbf{P}$ then the proof will be complete. Indeed, set

$$P_j := \frac{1}{j^2} R_{2^j} \qquad (j \in \mathbf{P})$$

and verify that these P_j's satisfy the required properties.

The construction of the polynomials R_ℓ will be presented in three steps.

For the first step, suppose for $i = 0$ and 1, that A_i is a finite union of pairwise disjoint dyadic intervals of length 2^{-k_i} for some $k_i \in \mathbf{N}$ with $k_0 < k_1$. For example,

$$A_0 := \bigcup_{j=1}^{M} I_j \qquad (M \in \mathbf{P})$$

with $\mu(I_j) = 2^{-k_0}$. Suppose further that $A_0 \supset A_1$ and that

$$\mu(I_j \cap A_1) \leq \mu(I_j \setminus A_1), \tag{28}$$

for $1 \leq j \leq M$. We claim that given any Walsh polynomial R of order less than 2^{k_0} which satisfies $\|R\|_\infty \leq 1$, $R = 0$ on A_0, and

$$|(S_m R)(x)| \geq \frac{\ell}{2} \tag{29}$$

for some integers $m, \ell \in \mathbf{N}$ and all $x \in A_0$, there exists a Walsh polynomial Q of order less than 2^{k_1} which satisfies $\|Q\|_\infty \leq 1$, $|Q| \leq 1/2$ on A_0, and

$$|(S_{m+q}Q)(x)| \geq \frac{\ell+1}{2}$$

for all $x \in A_1$, where $q := 2^{k_1} - 2^{k_0}$.

To prove this claim set

$$g(x) := \begin{cases} \frac{1}{2} & (S_m R)(x) \geq 0 \\ -\frac{1}{2} & (S_m R)(x) < 0, \end{cases}$$

$$\alpha_j := \frac{-1}{\mu(I_j \setminus A_1)} \int_{I_j \cap A_1} g\, d\mu \qquad (1 \leq j \leq M),$$

and define a function f on $\mathbf{G}$ by

$$f(x) := \begin{cases} g(x) & x \in A_1 \\ \alpha_j & x \in I_j \setminus A_1, 1 \leq j \leq M \\ 0 & x \in [0,1) \setminus A_0. \end{cases}$$

By the definition of the α_j's, it is clear that

$$\frac{1}{\mu(I_j)} \int_{I_j} f\, d\mu = 0$$

for $1 \leq j \leq M$. Since f vanishes off A_0, it follows that f is of mean zero on $I(p, k_0)$ for all $0 \leq p < 2^{k_0}$. Consequently, $\widehat{f}(j) = 0$ for $0 \leq j < 2^{k_0}$. On the other hand, since f is constant on all intervals of the form $I(p, k_1), 0 \leq p < 2^{k_1}$, it is clear that $\widehat{f}(j) = 0$ for $j \geq 2^{k_1}$. Hence f is a Walsh polynomial and

$$f = \sum_{j=2^{k_0}}^{2^{k_1}-1} \widehat{f}(j)\psi_j.$$

The choice of q implies $j \oplus q < q$ for $2^{k_0} \leq j < 2^{k_1}$. In particular, it follows from (46) in 1.5 that

$$S_q(\psi_q f) = \psi_q f.$$

By assumption, $\widehat{R}(j) = 0$ for $j \geq 2^{k_0}$. Since the choice of q implies $\psi_j \psi_q = \psi_{j+q}$ for $0 \leq j < 2^{k_0}$, we have

$$\psi_q R = \sum_{j=0}^{2^{k_0}-1} \widehat{R}(j)\psi_{j+q}.$$

Therefore,

$$S_{m+q}(\psi_q R) = \psi_q S_m(R).$$

Consequently, if $Q := (f + R)\psi_q$ then

$$S_{m+q}(Q) = (f + S_m(R))\psi_q.$$

To see that Q enjoys the promised properties, let $x \in A_1$ and suppose that $(S_m R)(x) \neq 0$. The identity above, the definition of f, and (29) imply

$$\begin{aligned} |(S_{m+q}Q)(x)| &= |f(x) + (S_m R)(x)| \\ &= |\frac{1}{2}\operatorname{sgn}(S_m R)(x) + (S_m R)(x)| \\ &= \frac{1}{2} + |(S_m R)(x)| \\ &\geq \frac{\ell + 1}{2}. \end{aligned}$$

On the other hand, if $(S_m R)(x) = 0$, then (29) implies that $\ell = 0$. Hence

$$\begin{aligned} |(S_{m+q}Q)(x)| &= |f(x) + (S_m R)(x)| \\ &= \frac{1}{2} + 0 \\ &= \frac{\ell + 1}{2}. \end{aligned}$$

Thus $|S_{m+q}Q| \geq (\ell + 1)/2$ on A_1.

To obtain bounds for Q, observe that the supports of f and R are disjoint by definition. Since both f and R are dominated by 1, it follows that $\|Q\|_\infty \leq 1$. Moreover, (28) implies that $|f| \leq 1/2$ on A_0. Since R vanishes on A_0, we see that $|Q| \leq 1/2$ on A_0. Finally, by construction Q is a Walsh polynomial of order less than 2^{k_1}. This completes the first step.

For the second step, use the fact that E is compact to choose sets

$$E_0 := \mathbf{G} \supset E_1 \supset \ldots$$

such that

$$E \subseteq \bigcap_{\ell=1}^{\infty} E_\ell,$$

and such that each E_ℓ is a finite union of dyadic intervals of length $2^{-k(\ell)}$, say

$$E_\ell := \bigcup_{j=1}^{M(\ell)} I_j^{(\ell)}.$$

Since $\mu(E) = 0$ we may suppose that $k(\ell) < k(\ell+1)$ and when constructing $E_{\ell+1}$ from E_ℓ that each interval $I_j^{(\ell)}$ has been diminished by at least half, i.e.,

$$\mu(I_j^{(\ell)} \cap E_{\ell+1}) \leq \mu(I_j^{(\ell)} \setminus E_{\ell+1}) \tag{30}$$

for $1 \leq j \leq M(\ell)$ and $\ell \in \mathbf{N}$. This completes the second step.

For the third step we show that to each $\ell \in \mathbf{N}$ there correspond a Walsh polynomial R_ℓ and an integer $m(\ell)$ such that $\|R_\ell\|_\infty \leq 1, R_\ell = 0$ on $E_{2\ell}$, and

$$|(S_{m(\ell)}R_\ell)(x)| \geq \frac{\ell}{2}$$

for all $x \in E_{2\ell}$.

Set $R_0 := 0$ and suppose that such R_ℓ and $m(\ell)$ have been chosen. Apply the claim to $A_0 := E_{2\ell}, A_1 := E_{2\ell+1}, M := M(2\ell), k_0 := k(2\ell), k_1 := k(2\ell+1), m := m(\ell)$ and $R := R_\ell$. Thus choose a Walsh polynomial Q such that $\|Q\|_\infty \leq 1$, $|Q| \leq 1/2$ on $E_{2\ell}$, and

$$|(S_{m+q}Q)(x)| \geq \frac{\ell+1}{2}$$

for all $x \in E_{2\ell+1}$ where

$$q := 2^{k(2\ell+1)} - 2^{k(2\ell)}.$$

Let

$$\beta_j := \frac{1}{\mu(I_j^{(2\ell+1)} \setminus E_{2\ell+2})} \int_{I_j^{(2\ell+1)} \cap E_{2\ell+2}} Q\, d\mu$$

for $1 \leq j \leq M(2\ell+1)$ and define h on $\mathbf{G}$ by

$$h(x) := \begin{cases} -Q(x) & x \in E_{2\ell+2} \\ \beta_j & x \in I_j^{(2\ell+1)} \setminus E_{2\ell+2},\ 1 \leq j \leq M(2\ell+1) \\ 0 & x \in [0,1) \setminus E_{2\ell+1}. \end{cases}$$

Finally, set $R_{\ell+1} := h + Q$ and $m(\ell+1) := m + q$.

Let $x \in E_{2\ell}$ and observe by (30) and the choice of Q that $|h(x)| \leq 1/2$. Hence

$$|R_{\ell+1}(x)| \leq |h(x)| + |Q(x)| \leq \frac{1}{2} + \frac{1}{2}.$$

On the other hand, if $x \in [0,1) \setminus E_{2\ell}$ then

$$|R_{\ell+1}(x)| = |Q(x)| \leq 1.$$

Consequently, $\|R_{\ell+1}\|_\infty \leq 1$.

Let $x \in E_{2\ell+2}$. By construction $h(x) = -Q(x)$. Hence it is clear that $R_{\ell+1}$ vanishes on $E_{2\ell+2}$.

Finally, observe that h is of mean zero on $I(p, k(2\ell+1))$ for each $0 \le p < 2^{k(2\ell+1)}$, and h is constant on $I(p, k(2\ell+2))$ for each $0 \le p < 2^{k(2\ell+2)}$. Therefore,

$$R_{\ell+1} = \sum_{j=0}^{2^{k(2\ell+1)}-1} \widehat{Q}(j)\psi_j + \sum_{j=2^{k(2\ell+1)}}^{2^{k(2\ell+2)}-1} \widehat{h}(j)\psi_j.$$

Since $m(\ell+1) < 2^{k(2\ell+1)}$, we conclude from the choice of Q that

$$|(S_{m(\ell+1)}R_{\ell+1})(x)| = |(S_{m+q}Q)(x)| \ge \frac{\ell+1}{2}$$

for all $x \in E_{2\ell+1}$. ∎

By combining this result with Corollary 2 we see that every $\mathcal{F}_\sigma$ set in **G** of Haar measure zero is a set of divergence for C(**G**).

4.6 Adjustment of Functions. We have seen that the Walsh-Fourier series of an $f \in \mathrm{L}^1$ may diverge everywhere on $[0,1)$. Nevertheless, it is always possible to adjust f, by changing its values on a set of small measure, to obtain a new function whose Walsh-Fourier series converges uniformly (see Theorem 14 below).

We begin with a technical result concerning polynomial approximations on dyadic intervals.

LEMMA 3. *Let $0 \le p < 2^n, m \ge n$, and $\ell > 0$ be integers. If h is the characteristic function of $I(p,n)$, then there exists a Walsh polynomial*

$$g = \sum_{i=2^m}^{2^{m+\ell}-1} c_i w_i$$

which vanishes off $I(p,n)$ such that

$$|\{g \ne h\}| \le 2^{-n-\ell}, \tag{31}$$

$$|g(x)| < 2^\ell,$$

and

$$|S_N g(x)| < 2^\ell$$

for $x \in [0,1)$ and $N \in \mathbf{N}$.

PROOF. Denote the dyadic expansion of $x \in [0,1)$ by $\sum_{k=0}^\infty x_k/2^{k+1}$. For each pair of integers $m, \ell \in \mathbf{P}$ let $E(m,\ell)$ denote the set of points $x \in [0,1)$ which satisfy $x_{m+j} = 0$ for $0 \le j < \ell$. Then $E(m,\ell)$ is a union of 2^m dyadic intervals of length $2^{-m-\ell}$. In particular,

$$|I(p,n) \cap E(m,\ell)| = 2^{-n-\ell}.$$

Let

$$g(x) := \begin{cases} 0 & x \in [0,1) \setminus I(p,n) \\ 1-2^\ell & x \in I(p,n) \cap E(m,\ell) \\ 1 & \text{otherwise.} \end{cases}$$

Then g satisfies (31), vanishes off $I(p,n)$, and is dominated by 2^ℓ. Also, if χ represents the characteristic function of $E(m,\ell)$ then

$$g = (1 - 2^\ell \chi)h. \tag{32}$$

Both h and χ can be represented in product form. Indeed, the definition of Rademacher functions implies that

$$1 + r_k(x)r_k(y) = \begin{cases} 2 & x_k = y_k \\ 0 & x_k \neq y_k, \end{cases}$$

for $k \in \mathbf{N}$ and $x,\ y \in [0,1)$. Hence for any fixed $y \in I(p,n)$ we have

$$h = 2^{-n} \prod_{k=0}^{n-1} (1 + r_k r_k(y)).$$

Since Walsh functions are products of Rademacher functions, this product when expanded yields

$$h = 2^{-n} \sum_{j=0}^{2^n-1} w_j w_j(y). \tag{33}$$

Similarly,

$$\chi = 2^{-\ell} \prod_{k=m}^{m+\ell-1} (1 + r_k) \tag{34}$$

which when expanded yields

$$\chi = 2^{-\ell}(1 + \Sigma),$$

where Σ represents the sum of all Walsh functions whose indices have the form

$$2^{k_1} + \cdots + 2^{k_i}$$

for $i \leq \ell$ and $m \leq k_1 < \cdots < k_i < m + \ell$. It follows from (32) and (34) that $g = -h \cdot \Sigma$. Thus by (33), g is a Walsh polynomial of the form

$$g = 2^{-n} \sum_{j=2^m}^{2^{m+\ell}-1} c_j w_j,$$

where $c_j = \pm 1$ for exactly $2^{n+\ell} - 2^n$ values of j, and $c_j = 0$ otherwise. Therefore, if $N \in \mathbf{N}$ and $x \in [0,1)$, then

$$\begin{aligned} |(S_N g)(x)| &\leq 2^{-n} \sum_{j=2^m}^{2^{m+\ell}-1} |c_j| \\ &\leq 2^{-n}(2^{n+\ell} - 2^n) \\ &= 2^\ell - 1. \quad \blacksquare \end{aligned}$$

THEOREM 14. *If $p_1 < q_1 < p_2 < q_2 < \ldots$ are positive integers with $(q_k/p_k, k \in \mathbf{P})$ unbounded, if f belongs to L^0, and if $\varepsilon > 0$, then there is a function g whose Walsh-Fourier series converges uniformly, whose Walsh-Fourier coefficients satisfy $\widehat{g}(j) \neq 0$ only when $p_k < j \leq q_k$, $k = 1, 2, \ldots$, such that*

$$|\{f \neq g\}| < \varepsilon.$$

PROOF. By Lusin's theorem, we may suppose that f is continuous and $\|f\|_\infty \leq 1$. Write f as a uniformly convergent sum of dyadic step functions. Specifically, let

$$f = \sum_{j=0}^{\infty} s_j$$

where each s_j is constant on dyadic intervals of the form $I(k, n_j)$ for $0 \leq k < 2^{n_j}$ and some $n_j \in \mathbf{N}$. Since f is continuous we may choose $(n_j, j \in \mathbf{N})$ increasing so rapidly that

$$|s_j| < 2^{-2j} \qquad (j \in \mathbf{N}). \tag{35}$$

Arrange the intervals $I(k, n_j)$, $0 \leq k < 2^{n_j}$, $j \geq 0$, into a sequence by the following process. Let $v_{-1} := 0$,

$$v_j := \sum_{\ell=0}^{j} 2^{n_\ell},$$

and set

$$J_i := I(k, n_j)$$

for $i = v_{j-1} + k < v_j$, $j \in \mathbf{N}$. Thus if f_i represents the characteristic function of the interval J_i, then there exist constants $a_0, a_1, \ldots$ such that

$$f = \sum_{i=0}^{\infty} a_i f_i.$$

By the choice of the functions s_j, this series converges uniformly. And in view of (35), we also have

$$|a_i| < 2^{-2j} \tag{36}$$

for $v_{j-1} \leq i < v_j$ and $j \in \mathbf{N}$.

We shall use Lemma 3 to adjust each f_i. First, choose j_0 so large that

$$\sum_{j=j_0}^{\infty} 2^{-j} < \varepsilon.$$

Next, suppose $v_{j-1} \leq i < v_j$ is fixed and use the fact that $(q_k/p_k, k \in \mathbf{P})$ is unbounded to choose an index k_i and an integer m_i such that

$$p_{k_i} \leq 2^{m_i} < 2^{m_i + j + j_0} < q_{k_i} \quad \text{and} \quad m_{i-1} + j + j_0 < m_i.$$

Apply Lemma 3 to $h := f_i$, $\ell := j + j_0$, $n := n_j$, and $m := m_i$. Thus choose a function g_i which is a linear combination of Walsh functions whose indices t satisfy $2^{m_i} \le t < 2^{m_i+j+j_0}$ such that $g_i = 0$ on $[0,1) \setminus J_i$, $|g_i| < 2^{j+j_0}$, $|S_N g_i| < 2^{j+j_0}$ for $N \in \mathbf{N}$, and $g_i = f_i$ except on a set E_i of measure $2^{-n_j-j-j_0}$.

Set

$$g := \sum_{i=0}^{\infty} a_i g_i. \tag{37}$$

This series converges uniformly. Indeed, if $x \in [0,1)$ and $j \in \mathbf{N}$ are fixed then there is only one index i which satisfies $g_i(x) \neq 0$ and $v_{j-1} \le i < v_j$. Hence it follows from (36) and the inequality $|g_i| < 2^{j+j_0}$ that

$$|a_i g_i(x)| < 2^{-2j}\, 2^{j+j_0} = 2^{j_0-j}.$$

In particular, (37) converges uniformly by the Weierstrass M-test.

Set

$$E := \bigcup_{i=0}^{\infty} E_i$$

and observe by construction that $f = g$ except on the set E. But the choice of j_0 and the choice of the sets E_i imply that

$$\begin{aligned} |E| &= \sum_{j=0}^{\infty} \sum_{i=v_{j-1}}^{v_j-1} |E_i| \\ &= \sum_{j=0}^{\infty} 2^{n_j}\, 2^{-n_j-j-j_0} \\ &= \sum_{j=j_0}^{\infty} 2^{-j} < \varepsilon. \end{aligned}$$

Consequently, g differs from f only on a set of measure less than ε .

It remains to verify that the Walsh-Fourier series of g converges uniformly on $[0,1)$. Let $N > 0$ be an integer and use uniform convergence of (37) to write

$$S_N g = \sum_{i=0}^{\infty} a_i S_N g_i.$$

By the choice of g_i,

$$S_N g_i = \begin{cases} 0 & N < 2^{m_i} \\ g_i & N \ge 2^{m_{i+1}}. \end{cases}$$

Since for each $N > 0$ it is possible to choose an integer ℓ such that $2^{m_\ell} \le N < 2^{m_{\ell+1}}$, it follows that

$$S_N g = \sum_{i=0}^{\ell-1} a_i g_i + a_\ell S_N g_\ell. \tag{38}$$

But $v_{j-1} \le \ell < v_j$ for some integer j. Hence by (36) and the fact that $|S_N g_\ell| < 2^{j+j_0}$. we have

$$|a_\ell(S_N g_\ell)| < 2^{-j+j_0}.$$

Since $j \to \infty$ as $\ell \to \infty$, it follows that $a_\ell(S_N g_\ell) \to 0$ uniformly as $\ell \to \infty$. Hence by (38) and the fact that $N \to \infty$ as $\ell \to \infty$ we conclude that

$$S_N g \to g$$

uniformly on $[0,1)$ as $N \to \infty$. ∎

Theorem 14 is not true if "converges uniformly" is replaced by "converges absolutely." To verify this we need several preliminary results.

For the next few pages let $\mathbf{G}_k$ represent the subgroups $\mathbf{G}_0^{(k)}$ introduced in 1.4, i.e.,

$$\mathbf{G}_k := \{x = (x_0, x_1, \dots) \in \mathbf{G} : x_n = 0 \quad \text{for} \quad n \ge k\}$$

for $k \in \mathbf{N}$. Notice each $\mathbf{G}_k$ is a subgroup of $\mathbf{G}$ of order 2^k. The discrete Walsh-Fourier transform of an $f \in \mathrm{L}^1(\mathbf{G})$ on $\mathbf{G}_k$ is defined by

$$\widehat{f}(m; \mathbf{G}_k) := 2^{-k} \sum_{y \in \mathbf{G}_k} f(y)\psi_m(y) \tag{39}$$

for $0 \le m < 2^k$ and $k \in \mathbf{N}$. Since

$$\sum_{m=0}^{2^k-1} \psi_m(y)\psi_m(t) = D_{2^k}(y+t) = \delta_{yt} 2^k \tag{40}$$

for $y, t \in \mathbf{G}_k$, $k \in \mathbf{N}$ (where δ_{yt} is the Kronecker delta), we have an inversion formula for this discrete Walsh- Fourier transform. Indeed, by definition and (40),

$$f(y) = \sum_{m=0}^{2^k-1} \widehat{f}(m; \mathbf{G}_k)\psi_m(y) \tag{41}$$

for $y \in \mathbf{G}_k$ and $k \in \mathbf{N}$.

The discrete Walsh-Fourier transform of the translate of a function is essentially a Walsh series (see (45) below). First notice that if f is a Walsh polynomial whose spectrum satisfies $sp(f) < 2^k$ then

$$\widehat{f}(m; \mathbf{G}_k) = \widehat{f}(m) \tag{42}$$

for $0 \le m < 2^k$. Next, verify the orthogonality relation

$$2^{-k} \sum_{y \in \mathbf{G}_k} \psi_s(y)\psi_{\ell 2^k + m}(y) = \delta_{ms} \tag{43}$$

for $\ell \in \mathbf{N}$ and $0 \le m, s < 2^k$. Use this to see that

(44) $$\widehat{\tau_x f}(m; \mathbf{G}_k) = \sum_{\ell=0}^{\infty} \widehat{f}(\ell 2^k + m)\psi_{\ell 2^k+m}(x)$$

for $f \in \mathrm{A}(\mathbf{G})$, $x \in \mathbf{G}$, $k \in \mathbf{N}$, and $0 \le m < 2^k$. Finally, observe that the Walsh series on the right side of (44) is the Walsh-Fourier series of the function

$$\widehat{\tau_x f}(m; \mathbf{G}_k) \qquad (x \in \mathbf{G})$$

(see Exercise 4.12). Consequently, the 2^n-th partial sums of this Walsh series converge a.e. $[\mu]$ when $f \in \mathrm{L}^1(\mathbf{G})$ and uniformly when $f \in \mathrm{C}(\mathbf{G})$. In particular,

(45) $$\widehat{\tau_x f}(m; \mathbf{G}_k) = \lim_{n\to\infty} \sum_{\ell=0}^{2^n-1} \widehat{f}(\ell 2^k + m)\psi_{\ell 2^k+m}(x)$$

for a.e. $[\mu]$ $x \in \mathbf{G}$ when $f \in \mathrm{L}^1(\mathbf{G})$ and for all $x \in \mathbf{G}$ when $f \in \mathrm{C}(\mathbf{G})$.

These observations will be used in the next two lemmas.

Lemma 4. *If $f \in \mathrm{A}(\mathbf{G})$, $k \in \mathbf{N}$, and g is a function defined on $\mathbf{G}_k$, then*

(46) $$2^{-k}\Big|\sum_{y\in\mathbf{G}_k} f(y)g(y)\Big| \le \|f\|_{\mathrm{A}} \max_{0\le m<2^k} |\widehat{g}(m; \mathbf{G}_k)|.$$

Proof. By inversion and orthogonality (i.e., (41) and (43)) we have

$$2^{-k}\sum_{y\in\mathbf{G}_k} f(y)g(y) = \sum_{m=0}^{2^k-1} \widehat{f}(m; \mathbf{G}_k)\widehat{g}(m; \mathbf{G}_k).$$

Consequently, it follows from (44) that

$$\begin{aligned} 2^{-k}\Big|\sum_{y\in\mathbf{G}_k} f(y)g(y)\Big| &\le \sum_{m=0}^{2^k-1}\sum_{\ell=0}^{\infty} |\widehat{f}(\ell 2^k + m)|\,|\widehat{g}(m; \mathbf{G}_k)| \\ &\le \|f\|_{\mathrm{A}} \max_{0\le m<2^k} |\widehat{g}(m; \mathbf{G}_k)|. \end{aligned}$$ ∎

Lemma 5. *Let $A \subseteq \mathbf{G}$ be measurable and $\varepsilon > 0$. Then for every $k \in \mathbf{N}$ there is an $x \in I_k(0)$ such that*

(47) $$2^{-k}\sum_{y\in\mathbf{G}_k} \chi(A)(x+y) > \mu(A) - \varepsilon.$$

Proof. Fix $x \in \mathbf{G}$, set $f := \chi(A)$ and observe by inversion and orthogonality that

$$\begin{aligned} 2^{-k}\sum_{y\in\mathbf{G}_k} \chi(A)(x+y) &= 2^{-k}\sum_{y\in\mathbf{G}_k} |(\tau_x f)(y)|^2 \\ &= \sum_{m=0}^{2^k-1} |\widehat{\tau_x f}(m; \mathbf{G}_k)|^2 \end{aligned}$$

for every $k \in \mathbf{N}$.

Suppose the lemma is false. Then there is a $k \in \mathbf{N}$ such that

$$\sum_{m=0}^{2^k-1} |\widehat{\tau_x f}(m; \mathbf{G}_k)|^2 \leq \mu(A) - \varepsilon$$

for all $x \in I_k(0)$. Since the left side of (47) is invariant under translation by elements of $\mathbf{G}_k$ it follows that this inequality holds for all $x \in \mathbf{G}$. In particular,

$$\mu(A) - \varepsilon \geq \int_{\mathbf{G}} \sum_{m=0}^{2^k-1} |\widehat{\tau_x f}(m; \mathbf{G}_k)|^2 \, d\mu(x).$$

Consequently, by Parseval's identity

$$\begin{aligned} \mu(A) - \varepsilon &\geq \sum_{m=0}^{2^k-1} \int_{\mathbf{G}} \left| \sum_{\ell=0}^{\infty} \widehat{f}(\ell 2^k + m) \psi_{\ell 2^k + m} \right|^2 \, d\mu(x) \\ &= \|\widehat{f}\|_{\ell^2}^2 \\ &= \int_{\mathbf{G}} |f|^2 \, d\mu \\ &= \mu(A) \end{aligned}$$

which is a contradiction. ∎

A local variant of the A-norm proves useful in this context. For each dyadic interval I and $f \in \mathrm{A}$ set

$$\|f\|_{\mathrm{A}}^I := \|\chi(I) \left(f - \frac{1}{\mu(I)} \int_I f \, d\mu \right) \|_{\mathrm{A}}.$$

We shall show that

$$\|f\|_{\mathrm{A}}^I \leq \sum_{m=1/\mu(I)}^{\infty} |\widehat{f}(m)| \tag{48}$$

and consequently that

$$\lim_{n \to \infty} \|f\|_{\mathrm{A}}^{I_n(x)} = 0 \tag{49}$$

for every $f \in \mathrm{A}$ and $x \in \mathbf{G}$.

To prove (48) fix a dyadic interval I, fix $x \in I$, and choose $n \in \mathbf{N}$ such that $\mu(I) = 2^{-n}$. Set

$$g := \chi(I) \left(f - \frac{1}{\mu(I)} \int_I f \, d\mu \right).$$

Since by Paley's lemma

$$\frac{\chi(I)}{\mu(I)} = \tau_x D_{2^n}$$

it is clear that $\widehat{g}(m) = \widehat{\chi(I)f}(m)$ for any $m \geq 2^n$. Moreover, by definition,

$$\widehat{\chi(I)f}(m) = \int_{\mathbf{G}} 2^{-n} D_{2^n}(x+t) f(t) \psi_m(t)\, d\mu(t)$$
$$= 2^{-n} \sum_{\ell=0}^{2^n-1} \psi_\ell(x) \int_{\mathbf{G}} f(t) \psi_{m\oplus\ell}(t)\, d\mu(t)$$
$$= 2^{-n} \sum_{\ell=0}^{2^n-1} \psi_\ell(x) \widehat{f}(m \oplus \ell).$$

Consequently,

$$|\widehat{g}(m)| \leq 2^{-n} \sum_{\ell=0}^{2^n-1} |\widehat{f}(m \oplus \ell)|$$

for every $m \geq 2^n$. On the other hand, it is clear since $\mu(I) = 2^{-n}$ that $\widehat{g}(m) = 0$ for $0 \leq m < 2^n$. We conclude that

$$\|g\|_{\mathrm{A}} \leq 2^{-n} \sum_{m=2^n}^{\infty} \sum_{\ell=0}^{2^n-1} |\widehat{f}(m \oplus \ell)|$$
$$= \sum_{m=2^n}^{\infty} |\widehat{f}(m)|.$$

This completes the proof of (48).

Finally we introduce a sequence of Walsh polynomials whose $C(\mathbf{G})$ norm is uniformly bounded, whose $A(\mathbf{G})$ norm is not, and such that if a function f is close to one of these polynomials on a certain discrete set then the $A(\mathbf{G})$ norm of f must be large (see Lemma 6 below).

To this end, use the isomorphism introduced by (15) in 1.2 to identify each

$$p = \sum_{k=0}^{\infty} p_k 2^k \in \mathbf{N}$$

with the element

$$\widetilde{p} := (p_0, p_1, \ldots) \in \mathbf{G}.$$

For each $n \in \mathbf{P}$ set

$$P_n := 2^{-n} \sum_{p=1}^{2^n-1} \psi_{p2^n} \tau_{\widetilde{p}} D_{2^n}.$$

Notice that the terms of this sum have pairwise disjoint spectra and non-overlapping supports. Consequently, we have

(50) $$\|P_n\|_\infty = 1$$

(51) $$\|P_n\|_{\mathrm{A}} = 2^n - 1$$

and

(52) $$\max_{m\in\mathbf{N}} |\widehat{P}_n(m)| = 2^{-n}.$$

Furthermore, it is clear by construction that

(53) $$sp(P_n) \subseteq [2^n, 2^{2n}).$$

(These polynomials will be used in 7.4 to construct null series with non-negative partial sums).

For the following, if A is a discrete set then $|A|$ represents the number of elements in A.

LEMMA 6. *Let $g \in \mathrm{A}(\mathbf{G})$, I be a dyadic interval with $\mu(I) = 2^{-k}$ and suppose A is a subset of $\mathbf{G}_{4k} \cap I$. If*

(54) $$|A| > \frac{3}{4}|\mathbf{G}_{4k} \cap I| = \frac{3}{4}2^{3k}$$

and

(55) $$|(g - P_{2k})\chi(A)| < \frac{1}{4}$$

then

(56) $$\|\chi(I)g\|_{\mathrm{A}} \geq \frac{1}{2}(k-3).$$

PROOF. Set
$$G(x) := \sin(g(x)) \qquad (x \in \mathbf{G}),$$

and suppose for a moment that $G \in \mathrm{A}(\mathbf{G})$. Notice by (55) that $|G(y)| > 1/2$ and

$$\mathrm{sgn}G(y) = \mathrm{sgn}P_{2k}(y) = P_{2k}(y)$$

for every $y \in A$. Hence the number of points in A can be estimated by

$$\frac{1}{2}|A| < \sum_{y\in A} G(y)P_{2k}(y) \leq \sum_{y\in\mathbf{G}_{4k}\cap I} G(y)P_{2k}(y) + |(\mathbf{G}_{4k} \cap I) \setminus A|.$$

Therefore, (54) and (46) imply

$$\frac{1}{8}2^{3k} < 2^{4k} \max_{0\leq m<2^{4k}} |\widehat{P}_{2k}(m; \mathbf{G}_{4k})| \, \|\chi(I)G\|_{\mathrm{A}}.$$

Consequently, we have by (42) that

(57) $$\frac{1}{8}2^k \leq \|\chi(I)G\|_{\mathrm{A}}.$$

On the other hand, since $G(x) = \sin(g(x))$ can be expanded in a Taylor series and since A(G) is a Banach algebra (see 2.4), it is obvious that

$$\begin{aligned} \|\chi(I)G\|_A &\le \sum_{n=1}^{\infty} \frac{\|\chi(I)g\|_A^{2n+1}}{(2n+1)!} \\ &\le \exp \|\chi(I)g\|_A. \end{aligned}$$

We conclude from (57) that

$$2^{k-3} \le \exp \|\chi(I)g\|_A. \qquad \blacksquare$$

We are now prepared to show that there exist continuous functions which cannot be adjusted on a set of positive measure to have absolutely convergent Walsh-Fourier series.

THEOREM 15. *There is an $F \in C(G)$ such that if $f \in A(G)$ then*

$$\mu(\{F = f\}) = 0.$$

PROOF. Set

$$F := \sum_{n=0}^{\infty} 2^{-4n} P_{2^{4n+1}}$$

and observe by (50) that this series converges uniformly on **G**. Thus $F \in C(G)$. Moreover, this series is the Walsh-Fourier series of F since by (53) its terms have pairwise disjoint spectra.

Suppose the theorem is false, i.e., that there exists an $f \in A(G)$ such that $E := \{F = f\}$ is of positive μ measure. Let $x^0 \in E$ be a point of density for E, i.e.,

$$\lim_{m\to\infty} \frac{(S_{2^m}\chi(E))(x^0)}{2^{-m}} = \lim_{m\to\infty} \frac{\mu(I_m(x^0)\cap E)}{2^{-m}} = 1. \tag{58}$$

For each $n \in \mathbf{P}$ set $k_n := 2^{4n-2}$, and

$$E_0 := I_{k_n}(x^0) \cap E.$$

Choose $N \in \mathbf{P}$ by (58) so that $n \ge N$ implies

$$\mu(E_0) > \frac{3}{4} 2^{-k_n}.$$

Choose by Lemma 5 an $x^n \in I_{4k_n}(0)$ such that

$$2^{-4k_n} |\mathbf{G}_{4k_n} \cap \tau_{x^n} E_0| > \frac{3}{4} 2^{-k_n}. \tag{59}$$

Finally set

$$g_n := 4k_n \tau_{x^n} (f - S_{2^{k_n}} F).$$

Fix an integer $n \geq N$. We will show that the hypotheses of Lemma 6 are satisfied by $g := g_n$, $k := k_n$, $I := I_{k_n}(x^0)$, and

$$A = \mathbf{G}_{4k_n} \cap \tau_{x^n} E_0.$$

First notice that $g \in \mathrm{A}(\mathbf{G})$ because $f \in \mathrm{A}(\mathbf{G})$. Next notice by (59) that hypothesis (54) is satisfied. To see that hypothesis (55) also holds break F into three pieces:

$$\begin{aligned} F &= \sum_{i=0}^{n-1} 2^{-4i} P_{2^{4i+1}} + 2^{-4n} P_{2^{4n+1}} + \sum_{i=n+1}^{\infty} 2^{-4i} P_{2^{4i+1}} \\ &=: F_n + \frac{1}{4k_n} P_{2k_n} + R_n. \end{aligned}$$

Clearly, $F_n = S_{2^{k_n}} F$. Moreover, since $x^n \in I_{4k_n}(0)$ we have $\tau_{x^n} P_{2k_n} = P_{2k_n}$. Consequently, it follows from the definition of A that

$$\begin{aligned} \chi(A)\tau_{x^n}(f - S_{2^{k_n}} F) &= \chi(A)\tau_{x^n}(F - S_{2^{k_n}} F) \\ &= \frac{1}{4k_n}\chi(A) P_{2k_n} + \chi(A)\tau_{x^n} R_n. \end{aligned}$$

Consequently, (50) implies

$$|(g_n - P_{2k_n})\chi(A)| \leq 4k_n \sum_{i=n+1}^{\infty} \frac{1}{k_i} = \frac{1}{15}.$$

In particular, we have by Lemma 6 that

$$\|\chi(I_{k_n}(x^0)) g_n\|_{\mathrm{A}} \geq \frac{1}{2}(k_n - 3)$$

for all $n \in \mathbf{N}$ and $n \geq N$. Since $x^n \in I_{4k_n}(0) \subset I_{k_n}(0)$ and the $\|\cdot\|_{\mathrm{A}}$-norm is translation invariant, we have by the definition of g_n that

$$\begin{aligned} \frac{1}{2}(k_n - 3) &\leq 4k_n \|\chi(I_{k_n}(x^0)\tau_{x^n}(f - S_{2^{k_n}} F)\|_{\mathrm{A}} \\ &= 4k_n \|\chi(I_{k_n}(x^0)((S_{2^{k_n}} F)(x^0) - f)\|_{\mathrm{A}}. \end{aligned} \tag{60}$$

Set

$$\varepsilon_n := \left| 2^{k_n} \int_{I_{k_n}(x^0)} f \, d\mu - (S_{2^{k_n}} F)(x^0) \right|.$$

By definition

$$\|f\|_{\mathrm{A}}^{I_{k_n}(x^0)} \geq \|((S_{2^{k_n}} F)(x^0) - f)\,\chi(I_{k_n}(x^0))\|_{\mathrm{A}} - \varepsilon_n \|\chi(I_{k_n}(x^0))\|_{\mathrm{A}}.$$

Moreover, we have by (60) that

$$\|((S_{2^{k_n}}F)(x^0) - f)\,\chi(I_{k_n}(x^0))\|_{\mathrm{A}} \geq \frac{1}{16}.$$

It follows, therefore, that

$$\|f\|_{\mathrm{A}}^{I_{k_n}(x^0)} \geq \frac{1}{8} - \varepsilon_n \tag{61}$$

for $n \geq N$. However by construction

$$\begin{aligned}\varepsilon_n &\leq 2^{k_n} \int_{I_{k_n}(x^0)} |F - f|\, d\mu \\ &= 2^{k_n} \int_{I_{k_n}(x^0)\setminus E} |F - f|\, d\mu \\ &\leq (\|F\|_\infty + \|f\|_\infty) 2^{k_n} \mu(I_{k_n}(x^0) \setminus E).\end{aligned}$$

This last sequence converges to zero as $n \to \infty$ by (58). In particular, (61) contradicts (49). ∎

Exercises

4.1 Let $1 \leq p < \infty$. Prove that $S_n f \to f$ in L^p norm as $n \to \infty$ for all $f \in \mathrm{L}^p$ if and only if there is a constant $M < \infty$ such that

$$\sup_{n\in\mathbf{N}} \|S_n f\|_p \leq M\|f\|_p$$

for $f \in \mathrm{L}^p$.

4.2 Suppose $a_0 \geq a_1 \geq a_2 \geq \cdots \geq a_k \to 0$ as $k \to \infty$, and let

$$S := \sum_{k=0}^{\infty} a_k w_k.$$

Show that S converges uniformly on compact subsets of (0,1), and that S converges on [0,1) if and only if S is absolutely convergent.

4.3 Let $\psi : [0,\infty) \to \mathbf{R}$ be a continuous, monotone increasing, unbounded function which satisfies $\psi(0) = 0$, and let Φ, Ψ be the corresponding Young's function pair, i.e.,

$$\Phi(x) := \int_0^x \psi, \quad \text{and} \quad \Psi(x) := \int_0^x \psi^{-1}$$

for $x \in [0,\infty)$. A function $f \in \mathrm{L}^0(\mathrm{G})$ is said to be of generalized bounded Φ-fluctuation if

$$\sup_{n\in\mathbf{P}} \left(\sum_{k=0}^{2^n-1} \Phi(\omega(f, I(k,n))) \right) < \infty.$$

Show that the Walsh-Fourier series of any continuous function of generalized bounded Φ-fluctuation converges uniformly on $\mathbf{G}$ when

$$\sum_{k=1}^{\infty} \Psi\left(\frac{1}{k}\right) < \infty.$$

[Onneweer and Waterman [2]]

4.4 Suppose $f \in C(\mathbf{G})$ and its modulus of continuity satisfies

$$\omega(f, 2^{-k}) = o\left(\frac{1}{k}\right), \quad \text{as} \quad k \to \infty.$$

Prove Sf converges uniformly on $\mathbf{G}$.

4.5 a) For each $0 < r < 1$ let

$$P_r := \sum_{k=0}^{\infty} r^k \psi_k.$$

Show $\|P_r\|_1 = O(1)$ as $r \uparrow 1$,

$$\int_{\mathbf{G}} P_r \, d\mu = 1 \qquad (0 < r < 1),$$

and

$$\lim_{r \uparrow 1} \int_{\mathbf{G} \setminus I_k(0)} |P_r| \, d\mu = 0$$

for $k \in \mathbf{N}$.

b) Let $\boldsymbol{X}$ be any homogeneous Banach space. Show that the Abel means

$$\sum_{k=0}^{\infty} r^k \widehat{f}(k) \psi_k$$

of the Walsh-Fourier series of f converge to f in the norm of $\boldsymbol{X}$ as $r \uparrow 1$, for every $f \in \boldsymbol{X}$.

4.6 Let $S = \sum_{k=0}^{\infty} a_k \psi_k$ be a Walsh series.

a) Show S is the Walsh-Fourier series of an $f \in L^{\infty}$ if and only if $\| \sup_{n \in \mathbf{N}} S_{2^n} \|_{\infty} < \infty$.

b) Let $\boldsymbol{X}$ be any homogeneous Banach space. Show S is the Walsh-Fourier series of an $f \in \boldsymbol{X}$ if and only if $(S_{2^n}, n \in \mathbf{N})$ converges in the $\boldsymbol{X}$ norm.

4.7 For each Walsh series $S = \sum_{k=0}^{\infty} a_k \psi_k$ let

$$\sigma_n := \sum_{k=0}^{n-1} \left(1 - \frac{k}{n}\right) a_k \psi_k \qquad (n \in \mathbf{P}).$$

Show S is the Walsh-Fourier-Stieltjes series of some finite Borel measure on $\mathbf{G}$ if and only if

$$\sup_{n \in \mathbf{N}} \|\sigma_n\|_1 < \infty.$$

What is the corresponding characterization of Walsh-Fourier- Stieltjes series on [0,1]?

Hint: If S is the Walsh-Fourier-Stieltjes series of some finite Borel measure on [0,1] does

$$\lim_{n\to\infty} \frac{1}{n} S_n(x-) = 0$$

for every $x \in \mathbf{Q}$?

[Fine[3]]

4.8 A Borel measure ν on $\mathbf{G}$ is called non-atomic if $\nu(\{x\}) = 0$ for all $x \in \mathbf{G}$. Prove that a Walsh series S is the Walsh-Fourier series of a non-atomic measure if and only if

$$\sup_{n\in\mathbf{N}} \|\sigma_n\|_1 < \infty$$

for $n \in \mathbf{N}$ in which case

$$\lim_{n\to\infty} \frac{1}{n} S_n = 0$$

uniformly on $\mathbf{G}$.

[Fine[3]]

4.9 Show that the function g in Theorem 14 can be chosen so that the following conditions are satisfied. Given any non-decreasing function ω on $[0,\infty)$ which satisfies $\omega(\delta) \to 0$ as $\delta \downarrow 0$, and given $\varepsilon > 0$ there is a set E (depending only on ε and ω such that $|E| < \varepsilon$, $g = f$ except on E, and Sg converges uniformly for any f whose modulus of continuity satisfies

$$\omega(f,\delta) \le \omega(\delta)$$

for $\delta > 0$.

[Price[4]]

4.10 Prove that if f is measurable, a.e. finite function on [0,1) then there is a Walsh series S which converges to f in measure on $[0,1)$.

4.11 Define π^n on $\mathbf{G}$ by

$$\pi^n(x_0, x_1, \dots) := (x_n, x_{n+1}, \dots)$$

i) Prove

$$\psi_m \circ \pi^n = \psi_{m2^n}$$

for $m, n \in \mathbf{N}$.

ii) Show π^n is measure preserving, i.e.,

$$\int_{\mathbf{G}} g \circ \pi^n \, d\mu = \int_{\mathbf{G}} g \, d\mu$$

for $n \in \mathbf{N}$ and $g \in \mathrm{L}^1(\mathbf{G})$.

iii) Prove

$$\widehat{f \circ \pi^n}(2^n k + s) = \begin{cases} \widehat{f}(k) & s = 0, \\ 0 & s \ne 0, \end{cases}$$

for $n, k, s \in \mathbf{N}$, $0 \le s < 2^n$, and $f \in \mathrm{L}^1(\mathbf{G})$.

4.12 Prove that the Walsh-Fourier series of the function

$$\widehat{\tau_x f}(m; \mathbf{G}_k) = 2^{-k} \sum_{y \in \mathbf{G}_k} f(x+y)\psi_m(y) \qquad (f \in \mathrm{L}^1(\mathbf{G}), x \in \mathbf{G})$$

is the Walsh series on the right side of (44).

Chapter 5

APPROXIMATION AND BASES

5.1 Approximation by Walsh Polynomials. For each $n \in \mathbf{P}$ let

$$\mathcal{P}_n := \{f \in \mathrm{C}(\mathbf{G}) : sp(f) \subseteq [0, n)\}.$$

Thus $\mathcal{P}_n$ is the collection of Walsh polynomials of order less than n and in the notation of 3.1, $\mathcal{P}_{2^m} = \mathrm{L}(\mathcal{A}^m)$ for $m \in \mathbf{N}$.

Throughout this section let $(\boldsymbol{X}, \|.\|_{\boldsymbol{X}})$ be a homogeneous Banach space over $\mathbf{G}$. Given an operator T on $\boldsymbol{X}$ we shall denote its operator norm by

$$\|T\|_{\boldsymbol{X}} := \sup_{f \in \boldsymbol{X}, \|f\|_{\boldsymbol{X}} \leq 1} \|Tf\|_{\boldsymbol{X}}.$$

In the case that $\boldsymbol{X} = \mathrm{L}^p(\mathbf{G})$ we shall abbreviate $\|T\|_{\boldsymbol{X}}$ by $\|T\|_p$.

Notice that each $\mathcal{P}_n$ is an n-dimensional subspace of $\boldsymbol{X}$ and that the partial sum operator S_n is a projection of $\boldsymbol{X}$ onto $\mathcal{P}_n$, i.e., $S_n P = P$ for all $P \in \mathcal{P}_n$, and $S_n f \in \mathcal{P}_n$ for all $f \in \boldsymbol{X}$. This projection is minimal in the following sense:

THEOREM 1. *Let $n \in \mathbf{P}$ and $T_n : \boldsymbol{X} \to \mathcal{P}_n$ be a projection. Then the operator norms of S_n and T_n are related by*

$$\|S_n\|_{\boldsymbol{X}} \leq \|T_n\|_{\boldsymbol{X}}.$$

PROOF. We shall establish the formula

$$S_n f = \int_{\mathbf{G}} \boldsymbol{\tau}_y T_n \boldsymbol{\tau}_y f \, d\mu(y)$$

for $f \in \boldsymbol{X}$ and $n \in \mathbf{N}$. If this formula holds then Lemma 1 of 4.4 implies

$$\begin{aligned} \|S_n f\|_{\boldsymbol{X}} &= \left\| \int_{\mathbf{G}} \boldsymbol{\tau}_y T_n \boldsymbol{\tau}_y f \, d\mu(y) \right\|_{\boldsymbol{X}} \\ &\leq \int_{\mathbf{G}} \|T_n\|_{\boldsymbol{X}} \|f\|_{\boldsymbol{X}} \, d\mu(y) \\ &= \|T_n\|_{\boldsymbol{X}} \|f\|_{\boldsymbol{X}} \end{aligned}$$

and the proof of the theorem will be complete.

The Walsh polynomials are dense in $\boldsymbol{X}$ (see (iii) at the beginning of 4.4). Hence we need only verify the formula for $f = \psi_k, k \in \mathbf{N}$ fixed. Suppose first that $n > k$. Then $S_n \psi_k = \psi_k$. Moreover, since T_n is a projection we have for every $y \in \mathbf{G}$ that

$$T_n(\boldsymbol{\tau}_y \psi_k) = T_n(\psi_k(y)\psi_k) = \psi_k(y)\psi_k.$$

Consequently,

$$\begin{aligned}\int_{\mathbf{G}} \tau_y T_n \tau_y \psi_k \, d\mu(y) &= \psi_k \int_{\mathbf{G}} \psi_k^2(y) \, d\mu(y) \\ &= \psi_k \\ &= S_n \psi_k\end{aligned}$$

for $n > k$. On the other hand, if $n \le k$ then

$$\begin{aligned}T_n \tau_y \psi_k &= \psi_k(y) T_n \psi_k \\ &= \psi_k(y) \sum_{\ell=0}^{n-1} c_\ell^{(n)} \psi_\ell\end{aligned}$$

for a suitable choice of the coefficients $c_\ell^{(n)} \in \mathbf{R}$. It follows from orthogonality that

$$\begin{aligned}\int_{\mathbf{G}} \tau_y T_n \tau_y \psi_k \, d\mu(y) &= \int_{\mathbf{G}} \psi_k(y) \sum_{\ell=0}^{n-1} c_\ell^{(n)} \psi_\ell(y) \psi_\ell \, d\mu(y) \\ &= 0 \\ &= S_n \psi_k\end{aligned}$$

for $n \le k$. We conclude the formula holds for $f = \psi_k$. ∎

To measure of the rate of approximation of an $f \in \boldsymbol{X}$ by polynomials in $\mathcal{P}_n$, define

$$\mathbf{E}_n(f, \boldsymbol{X}) := \inf_{P \in \mathcal{P}_n} \|f - P\|_{\boldsymbol{X}}.$$

Since each $\mathcal{P}_n$ is a finite dimensional subspace of $\boldsymbol{X}$, it is clear that for every $f \in \boldsymbol{X}$ there is at least one polynomial $P_n \in \mathcal{P}_n$ such that

$$\mathbf{E}_n(f, \boldsymbol{X}) = \|f - P_n\|_{\boldsymbol{X}},$$

i.e., the infimum above is attained.

Such a polynomial P_n will be called a best approximation of f in $\mathcal{P}_n$. It need not be unique. Indeed, consider the case $\boldsymbol{X} := \mathrm{C}(\mathbf{G})$. Fix $f \in \mathrm{C}(\mathbf{G})$, $n \in \mathbf{N}$, and let

$$P(x) := \frac{1}{2}\left(\max_{t \in I_n(x)} f(t) + \min_{t \in I_n(x)} f(t)\right) \qquad (x \in \mathbf{G}).$$

Then $P \in \mathcal{P}_{2^n}$ and it is easy to check that

$$\mathbf{E}_{2^n}(f, \boldsymbol{X}) = \|f - P\|_{\boldsymbol{X}}.$$

In particular, P is a best approximation to f in $\mathcal{P}_{2^n}$. However, let

$$\theta(k, n) := \frac{1}{2}\left(\max_{t \in I(k,n)} f(t) + \min_{t \in I(k,n)} f(t)\right),$$

$$\theta(n) := \max_{0 \le k < 2^n} \theta(k,n),$$

and suppose R is any Walsh polynomial in $\mathcal{P}_{2^n}$ which satisfies

$$|R(x)| \le \theta(n) - \theta(k,n)$$

for $x \in I(k,n)$ and $0 \le k < 2^n$. Then $P + R$ is also a best approximation of f in $\mathcal{P}_{2^n}$. It follows that if $\theta(k,n) \ne \theta(n)$ for some integer $k \in [0, 2^n)$ then f has infinitely many best approximations in $\mathcal{P}_{2^n}$ with respect to the norm of $\mathrm{C}(\mathbf{G})$.

It is obvious that

$$\mathbf{E}_n(f, \boldsymbol{X}) \le \|f - S_n f\|_{\boldsymbol{X}}$$

for $f \in \boldsymbol{X}$ and $n \in \mathbf{N}$. Conversely, for each $n \in \mathbf{N}$ and $f \in \boldsymbol{X}$ it is clear that

$$\|f - S_n f\|_{\boldsymbol{X}} \le \|f - P_n\|_{\boldsymbol{X}} + \|S_n(f - P_n)\|_{\boldsymbol{X}}$$

for any polynomial $P_n \in \mathcal{P}_n$. Therefore,

$$\|f - S_n f\|_{\boldsymbol{X}} \le (1 + \|S_n\|_{\boldsymbol{X}})\mathbf{E}_n(f, \boldsymbol{X}). \tag{1}$$

Notice that

$$\|S_n\|_{\boldsymbol{X}} \le \|D_n\|_1 = L_n \qquad (n \in \mathbf{N})$$

and that this inequality is sharp when $\boldsymbol{X} = \mathrm{C}(\mathbf{G})$ (see 4.2). Thus Theorem 9 in 1.6 and (1) above imply

$$\|f - S_n f\|_{\boldsymbol{X}} \le (1 + V(n))\mathbf{E}_n(f, \boldsymbol{X}) \tag{2}$$

for $f \in \boldsymbol{X}$, $n \in \mathbf{N}$, and $V(n)$ the variation of n.

A sharper estimate can be obtained for $\boldsymbol{X} := \mathrm{L}^p(\mathbf{G}), 1 < p < \infty$. Indeed, for each such p there exists a constant C_p (depending only on p) such that

$$\|S_n\|_p \le C_p \qquad (n \in \mathbf{N})$$

(see Corollary 6 in 3.3). Consequently, (1) implies

$$\|f - S_n f\|_p \le (1 + C_p)\mathbf{E}_n(f, \mathrm{L}^p(\mathbf{G}))$$

for $f \in \mathrm{L}^p(\mathbf{G}), n \in \mathbf{N}$, and $1 < p < \infty$. Thus for a given $f \in \mathrm{L}^p(\mathbf{G}), 1 < p < \infty$, the rate of approximation by Walsh polynomial of order n is no better than that of the Walsh-Fourier partial sum $S_n f$.

Define the modulus of continuity in $\boldsymbol{X}$ of an $f \in \boldsymbol{X}$ by

$$\omega^{(\boldsymbol{X})}(f, \delta) := \sup_{|y| < \delta} \|f - \tau_y f\|_{\boldsymbol{X}} \qquad (\delta > 0).$$

This modulus of continuity gives sharp estimates for the rate of approximation by Walsh polynomials of order 2^n :

THEOREM 2. *Let $f \in \boldsymbol{X}$ and $n \in \mathbf{N}$. Then*

$$\frac{1}{2}\omega^{(\boldsymbol{X})}(f,2^{-n}) \le \|f - S_{2^n}f\|_{\boldsymbol{X}} \le \omega^{(\boldsymbol{X})}(f,2^{-n})$$

and

$$\frac{1}{2}\omega^{(\boldsymbol{X})}(f,2^{-n}) \le \mathrm{E}_{2^n}(f,\boldsymbol{X}) \le \omega^{(\boldsymbol{X})}(f,2^{-n}).$$

PROOF. Fix $f \in \boldsymbol{X}, n \in \mathbf{N}$ and observe for every $t \in I_n(0)$ and $P \in \mathcal{P}_{2^n}$ that

$$\tau_t f - f = \tau_t(f - P) + (f - P).$$

Since $\boldsymbol{X}$ is a homogeneous Banach space, we have

$$\|\tau_t f - f\|_{\boldsymbol{X}} \le 2\|f - P\|_{\boldsymbol{X}}.$$

Consequently,

$$\omega^{(\boldsymbol{X})}(f,2^{-n}) \le 2\|f - P\|_{\boldsymbol{X}},$$

for any $P \in \mathcal{P}_{2^n}$. It follows that $\omega^{(\boldsymbol{X})}(f,2^{-n})$ is dominated by $2\|f - S_{2^n}f\|_{\boldsymbol{X}}$ and by $2E_{2^n}(f,\boldsymbol{X})$.

On the other hand, we know that

$$S_{2^n}f - f = \int_{\mathbf{G}} (\tau_t f - f) D_{2^n}(t)\, d\mu(t).$$

Consequently, Lemma 1 in 4.4 implies

$$\begin{aligned}\|S_{2^n}f - f\|_{\boldsymbol{X}} &\le \int_{\mathbf{G}} \|\tau_t f - f\|_{\boldsymbol{X}}\, D_{2^n}(t)\, d\mu(t)\\ &= 2^n \int_{I_n(0)} \|\tau_t f - f\|_{\boldsymbol{X}}\, d\mu(t)\\ &\le \omega^{(\boldsymbol{X})}(f,2^{-n}).\end{aligned}$$

Since $E_{2^n}(f,\boldsymbol{X}) \le \|S_{2^n}f - f\|_{\boldsymbol{X}}$ we also have

$$E_{2^n}(f,\boldsymbol{X}) \le \omega^{(\boldsymbol{X})}(f,2^{-n}). \qquad \blacksquare$$

This final inequality is the Walsh analogue of the classical Jackson inequality. The other inequalities in Theorem 2 have no trigonometric analogue.

For each $\alpha > 0$, Lipschitz classes in $\boldsymbol{X}$ can be defined by

$$\mathrm{Lip}(\alpha,\boldsymbol{X}) := \{f \in \boldsymbol{X} : \omega^{(\boldsymbol{X})}(f,\delta) = O(\delta^\alpha) \quad \text{as} \quad \delta \to 0\}.$$

Unlike the classical case, $\mathrm{Lip}(\alpha,\boldsymbol{X})$ is not trivial when $\alpha > 1$. For example, the function $f := \psi_0 + \psi_1$ belongs to $\mathrm{Lip}(\alpha,\boldsymbol{X})$ for all $\alpha > 0$ since

$$\omega^{(\boldsymbol{X})}(f,\delta) = 0 \tag{3}$$

when $0 < \delta < 1/2$.

The following result shows that these Lipschitz classes can be used to characterize functions by their rate of approximation by Walsh polynomials of order n and of order 2^n.

THEOREM 3. *Let $f \in \boldsymbol{X}$ and $\alpha > 0$. Then the following five conditions are equivalent :*

a) $f \in \text{Lip}(\alpha, \boldsymbol{X})$,

b) $\|f - S_{2^n} f\|_{\boldsymbol{X}} = O(2^{-\alpha n})$ as $n \to \infty$,

c) $\mathbf{E}_{2^n}(f, \boldsymbol{X}) = O(2^{-\alpha n})$ as $n \to \infty$,

d) $\mathbf{E}_n(f, \boldsymbol{X}) = O(n^{-\alpha})$ as $n \to \infty$,

e) $\omega^{(\boldsymbol{X})}(f, 2^{-n}) = O(2^{-\alpha n})$ as $n \to \infty$.

PROOF. By the Banach-Steinhaus theorem, a homogeneous Banach space satisfies

$$\lim_{|y| \to 0} \|f - \tau_y f\|_{\boldsymbol{X}} = 0.$$

Consequently,

$$\lim_{\delta \to 0} \omega^{(\boldsymbol{X})}(f, \delta) = 0 \tag{4}$$

for all $f \in \boldsymbol{X}$.

By Theorem 2, a) implies b).

Since

$$\mathbf{E}_{2^n}(f, \boldsymbol{X}) \le \|f - S_{2^n} f\|_{\boldsymbol{X}}$$

for $f \in \boldsymbol{X}, n \in \mathbf{N}$, it is clear that b) implies c).

By definition, $\mathbf{E}_k(f, \boldsymbol{X}) \le \mathbf{E}_j(f, \boldsymbol{X})$ for $f \in \boldsymbol{X}$ and $k \ge j$. Consequently,

$$\mathbf{E}_{2^{m+1}}(f, \boldsymbol{X}) \le \mathbf{E}_n(f, \boldsymbol{X}) \le \mathbf{E}_{2^m}(f, \boldsymbol{X})$$

for $f \in \boldsymbol{X}, 2^m \le n < 2^{m+1}$, and $m \in \mathbf{N}$. Hence it is evident that c) implies d).

By Theorem 2, d) implies e).

Finally, the fact that $\omega^{(\boldsymbol{X})}(f, \delta)$ decreases as δ decreases shows that e) implies a). ∎

The following result estimates approximation of $\text{Lip}(\alpha, \boldsymbol{X})$ functions by Cesàro means of Walsh-Fourier series.

THEOREM 4. *Suppose $f \in \text{Lip}(\alpha, \boldsymbol{X})$ and $\alpha > 0$. Then*

$$\|\sigma_n f - f\|_{\boldsymbol{X}} = \begin{cases} O(n^{-\alpha}) & 0 < \alpha < 1 \\ O(\log n / n) & \alpha = 1 \\ O(1/n) & \alpha > 1, \end{cases}$$

as $n \to \infty$.

PROOF. Let $n \in \mathbf{P}$ and choose $s \in \mathbf{N}$ such that $2^s \le n < 2^{s+1}$. Set

$$\boldsymbol{\Delta}(s) := 2^{-s\alpha} + 2^{-s} \sum_{k=0}^{s-1} 2^{k(1-\alpha)}.$$

We will show that

$$\|\sigma_{2^s} f - f\|_{\boldsymbol{X}} = O(\boldsymbol{\Delta}(s)), \qquad \text{as } s \to \infty . \tag{5}$$

Suppose for a moment that (5) holds. Set $P := S_{2^s}f$ and observe that

$$\sigma_n f - f = \sigma_n(f - P) + (P - f) + (\sigma_n P - P).$$

Since

$$\sigma_n P - P = \frac{1}{n}\sum_{\ell=1}^{2^s}(S_\ell f - S_{2^s}f) = \frac{2^s}{n}(\sigma_{2^s}P - P)$$

it follows from Theorem 2, Theorem 16 v) in 1.8, and Lemma 1 in 4.4 that

$$\begin{aligned}\|\sigma_n f - f\|_{\mathbf{X}} &\le \|(f - P) * K_n\|_{\mathbf{X}} + \|f - P\|_{\mathbf{X}} + \|\sigma_{2^s}P - P\|_{\mathbf{X}} \\ &\le 3\omega^{(\mathbf{X})}(f, 2^{-s}) + \|\sigma_{2^s}P - P\|_{\mathbf{X}}.\end{aligned}$$

Since

$$\|\sigma_{2^s}P - P\|_{\mathbf{X}} = \|(f * K_{2^s} - f) * D_{2^s}\|_{\mathbf{X}} \le \|\sigma_{2^s}f - f\|_{\mathbf{X}}$$

we see by (5) that $\|\sigma_n f - f\|_{\mathbf{X}}$ can be estimated by $\mathbf{\Delta}(s)$. But for $0 < \alpha < 1$ we have

$$\begin{aligned}\mathbf{\Delta}(s) &= O\left(2^{-s\alpha} + 2^{-s}2^{s(1-\alpha)}\right) \\ &= O(2^{-s\alpha}) \\ &= O(n^{-\alpha}) \quad \text{as} \quad n \to \infty.\end{aligned}$$

For $\alpha > 1$ we have

$$\begin{aligned}\mathbf{\Delta}(s) &= 2^{-s\alpha} + 2^{-s}\sum_{k=0}^{s-1}2^{k(1-\alpha)} \\ &= O\left(n^{-\alpha} + \frac{1}{n}\right) \\ &= O\left(\frac{1}{n}\right) \quad \text{as} \quad n \to \infty.\end{aligned}$$

And, for $\alpha = 1$ we have

$$\begin{aligned}\mathbf{\Delta}(s) &= O\left(2^{-s} + s2^{-s}\right) \\ &= O\left(\frac{s}{n}\right) \\ &= O\left(\frac{\log n}{n}\right) \quad \text{as} \quad n \to \infty.\end{aligned}$$

It remains to establish (5).

Recall that

(6) $$\sigma_{2^s}f - f = \int_{\mathbf{G}}(\boldsymbol{\tau}_t f - f)\,K_{2^s}(t)\,d\mu(t) \qquad (f \in \mathrm{L}^1(\mathbf{G})),$$

and that

$$K_{2^s} = \frac{1}{2}\left(2^{-s}D_{2^s} + \sum_{j=0}^{s} 2^{j-s}\tau_{e_j} D_{2^s}\right)$$

for $s \in \mathbf{N}$ (see Theorem 16 iii) in 1.8). Consequently, $x \in I_s(0)$ implies

$$|K_{2^s}(x)| \leq 2^s.$$

Since

$$\mathbf{G} = I_s(0) \bigcup \left(\bigcup_{k=0}^{s-1}(I_k(0) \setminus I_{k+1}(0))\right)$$

is a disjoint decomposition of $\mathbf{G}$, it follows from (6) that

$$\begin{aligned}\|\sigma_{2^s} f - f\|_{\boldsymbol{X}} \leq 2^s \int_{I_s(0)} &\|\boldsymbol{\tau}_t f - f\|_{\boldsymbol{X}}\, d\mu(t) \\ &+ \sum_{k=0}^{s-1} 2^{j-s} \int_{I_k(0)\setminus I_{k+1}(0)} \|\boldsymbol{\tau}_t f - f\|_{\boldsymbol{X}} D_{2^s}(t + e_k)\, d\mu(t) \\ \leq \omega^{(\boldsymbol{X})}(f, 2^{-s}) &+ \sum_{k=0}^{s-1} 2^{k-s}\omega^{(\boldsymbol{X})}(f, 2^{-k}),\end{aligned}$$

for $n \in \mathbf{P}$. Consequently, (5) holds for all $f \in \mathrm{Lip}(\alpha, \boldsymbol{X})$. ∎

The rate of approximation by Cesàro means is not necessarily improved as α increases beyond 1. Indeed, the function f which satisfied (3) also satisfies

$$\|f - \sigma_n f\|_{\boldsymbol{X}} = \frac{1}{n}\|\sum_{k=1}^{n}(f - S_k f)\|_{\boldsymbol{X}} = \frac{1}{n}\|\psi_1\|_{\boldsymbol{X}}$$

for $n \geq 2$.

In addition, the following is true.

THEOREM 5. *If*

$$\|\sigma_{2^n} f - f\|_{\boldsymbol{X}} = o(2^{-n}) \quad \text{as} \quad n \to \infty$$

for some $f \in \boldsymbol{X}$ then f is constant.

PROOF. Since $\mathbf{E}_{2^n}(f, \boldsymbol{X}) \leq \|\sigma_{2^n} f - f\|_{\boldsymbol{X}}$ we have by hypothesis and Theorem 2 that

$$\|S_{2^n} f - f\|_{\boldsymbol{X}} = o(2^{-n}) \quad \text{as} \quad n \to \infty.$$

Moreover,

$$\sigma_{2^n} f = \sum_{k=0}^{2^n-1}\left(1 - \frac{k}{2^n}\right)\widehat{f}(k)\psi_k = S_{2^n} f - 2^{-n}\sum_{k=0}^{2^n-1} k\widehat{f}(k)\psi_k.$$

Consequently,

$$\begin{aligned}\|\sum_{k=0}^{2^n-1} k\widehat{f}(k)\psi_k\|_{\boldsymbol{X}} &= 2^n\|\sigma_{2^n}f - S_{2^n}f\|_{\boldsymbol{X}} \\ &\leq 2^n\|\sigma_{2^n}f - f\|_{\boldsymbol{X}} + 2^n\|S_{2^n}f - f\|_{\boldsymbol{X}}.\end{aligned}$$

Since these last two terms tend to zero as $n \to \infty$ and $\|\cdot\|_1 \leq \|\cdot\|_{\boldsymbol{X}}$ it follows that

$$\begin{aligned}|j\widehat{f}(j)| &= \lim_{n\to\infty} |\int_0^1 \psi_j(\sum_{k=0}^{2^n-1} k\widehat{f}(k)\psi_k)| \\ &\leq \limsup_{n\to\infty} \|\sum_{k=0}^{2^n-1} k\widehat{f}(k)\psi_k\|_{\boldsymbol{X}} \\ &= 0\end{aligned}$$

for all $j \in \mathbf{P}$. We conclude that $\widehat{f}(j) = 0$ for $j \in \mathbf{P}$ and therefore f is constant. ∎

By Theorem 4, the rate of approximation by $\sigma_n f$ is as good as that by $S_{2^n}f$ for $\alpha < 1$. However, the σ_n's are not projections from $\boldsymbol{X}$ onto $\mathcal{P}_n$. These two important properties are reunited by the de la Valée Poussin means

$$V_n := 2\sigma_{2n} - \sigma_n \qquad (n \in \mathbf{N}).$$

Indeed, by definition $V_nP = P$ for all $P \in \mathcal{P}_n$, and consequently,

$$\|V_nf - f\|_{\boldsymbol{X}} \leq \|V_nf - P\|_{\boldsymbol{X}} + \|P - f\|_{\boldsymbol{X}} \leq (1 + \|V_n\|_{\boldsymbol{X}})\, \mathbf{E}_n(f,\boldsymbol{X})$$

holds for $f \in \boldsymbol{X}, n \in \mathbf{P}$ and P a best approximation of f in $\mathcal{P}_n$. Moreover,

$$\|V_n\|_{\boldsymbol{X}} \leq \|K_n\|_1 + 2\|K_{2n}\|_1 \qquad (n \in \mathbf{P}).$$

Since $\|K_m\|_1 \leq 2$ for $m \in \mathbf{P}$ it follows that

$$\|V_nf - f\|_{\boldsymbol{X}} \leq 7\mathbf{E}_n(f,\boldsymbol{X})$$

for $f \in \boldsymbol{X}, n \in \mathbf{P}$.

5.2 The Strong Derivative and Approximation. A homogeneous Banach space satisfies conditions (71), (72), and (73) in 1.7 (see Lemma 1 in 4.4). Thus the operators

$$\mathbf{d}_nf := \sum_{j=0}^{n-1} 2^{j-1}(f - \tau_{e_j}f) \qquad (n \in \mathbf{N}, f \in \boldsymbol{X})$$

can be used to define the strong dyadic derivative for any homogeneous Banach space $\boldsymbol{X}$.

We summarize the situation. A function $f \in \boldsymbol{X}$ is (strongly) differentiable in $\boldsymbol{X}$ if there is a function $\mathbf{d}^{[1]}f := \mathbf{d}f \in \boldsymbol{X}$ such that

$$\lim_{n\to\infty} \|\mathbf{d}_n f - \mathbf{d}f\|_{\boldsymbol{X}} = 0.$$

If f has a derivative in $\boldsymbol{X}$ of order r and if $\widehat{f}(0) = 0$, then

$$\widehat{\mathbf{d}^{[r]}f}(k) = k^r \widehat{f}(k)$$

for $r \in \mathbf{P}$ and $k \in \mathbf{N}$. The strong antiderivative operator (see Theorem 14 in 1.7) satisfies

$$\text{(7)} \qquad \mathbf{I}^{[r]}(\mathbf{d}^{[r]}f) = f \qquad (r \in \mathbf{P},\ \widehat{f}(0) = 0).$$

And, if f is differentiable in $\boldsymbol{X}$ then

$$\text{(8)} \qquad \lim_{n\to\infty} \widehat{\mathbf{d}_n f}(k) = \widehat{\mathbf{d}f}(k) \qquad (k \in \mathbf{N})$$

(see Theorem 13 in 1.7).

Clearly, (7) is half of a fundamental theorem of dyadic calculus. To establish the other half, fix $r \in \mathbf{P}$ and recall that $\mathbf{I}^{[r]}f = W_r * f$ for some $W_r \in \mathrm{L}^2$. Thus, by Lemma 1 in 4.4 $\mathbf{I}^{[r]}f$ is defined and belongs to $\boldsymbol{X}$ for every $f \in \boldsymbol{X}$.

THEOREM 6. *If $f \in \boldsymbol{X}$ and $\widehat{f}(0) = 0$ then*

$$\mathbf{d}^{[r]}(\mathbf{I}^{[r]}f) = f \qquad (r \in \mathbf{P}).$$

PROOF. We may suppose $r = 1$. The proof of Theorem 17 in 1.8 shows that

$$|\sum_{k=2^n}^{\infty} \frac{\psi_k}{k}| \le 7 \cdot 2^{-n} \qquad (n \in \mathbf{P}).$$

In particular, the definition of $\mathbf{d}_n$ implies

$$\begin{aligned} \|\mathbf{d}_n\left(\sum_{k=2^n}^{\infty} \frac{\psi_k}{k}\right)\|_1 &\le \sum_{j=0}^{n-1} 2^j \|\sum_{k=2^n}^{\infty} \frac{\psi_k}{k}\|_1 \\ &\le 7 \sum_{j=0}^{n-1} 2^{j-n} \\ &\le 7 \qquad (n \in \mathbf{P}). \end{aligned}$$

Recall (Theorem 7 in 4.4) that

$$\text{(9)} \qquad \lim_{n\to\infty} \|S_{2^n} f - f\|_{\boldsymbol{X}} = 0$$

for any $f \in \boldsymbol{X}$. Fix $n \in \mathbf{P}$ and $f \in \boldsymbol{X}$. Since $\mathbf{I}f := \mathbf{I}^{[1]}f \in \boldsymbol{X}$, we can write

$$\mathbf{d}_n(\mathbf{I}f) = \lim_{m\to\infty} \sum_{k=1}^{2^m-1} \frac{\widehat{f}(k)}{k} \mathbf{d}_n \psi_k,$$

where this series converges a.e. and in the norm of $\boldsymbol{X}$. Since $\mathbf{d}_n\psi_k = k\psi_k$ for $k < 2^n$ and

$$\mathbf{d}_n\psi_k = \sum_{j=0}^{n-1} k_j 2^j \psi_k$$

for $k \geq 2^n$ (see (70) in 1.7), we can also write

$$\sum_{k=1}^{2^m-1} \frac{\widehat{f}(k)}{k} \mathbf{d}_n \psi_k = S_{2^n} f + \sum_{k=2^n}^{2^m-1} \frac{\widehat{f}(k)}{k} \left(\sum_{j=0}^{n-1} k_j 2^j \psi_k \right) \qquad (m > n).$$

Hence it follows from (45) in 1.5 that

$$\mathbf{d}_n(\mathbf{I}f) = S_{2^n} f + \lim_{m\to\infty} (f - S_{2^n} f) * \left(\mathbf{d}_n \left(\sum_{k=2^n}^{2^m-1} \frac{\psi_k}{k} \right) \right).$$

Therefore, the estimate above implies

$$\|\mathbf{d}_n(\mathbf{I}f) - f\|_{\boldsymbol{X}} \leq 8\|S_{2^n} f - f\|_{\boldsymbol{X}}.$$

We conclude by (9) that $\mathbf{d}(\mathbf{I}f) = f$. ∎

The operator $\mathbf{I}$ plays the role of indefinite dyadic integration. Theorem 6 and identity (7) can be interpreted as a fundamental theorem of calculus. Unlike the classical case, $\mathbf{I}f$ is not always continuous (see Exercises 5.6, 5.8, and 5.9).

Theorem 6 contains two corollaries which will be used in the sequel.

COROLLARY 1. *Let $f \in \boldsymbol{X}$. Then f is differentiable in $\boldsymbol{X}$ if and only if there is a function $g \in \boldsymbol{X}$ which satisfies*

$$\widehat{g}(k) = k\widehat{f}(k) \qquad (k \in \mathbf{N}), \tag{10}$$

in which case $g = \mathbf{d}f$.

PROOF. Necessity was proved in Theorem 13 in 1.7.

To prove sufficiency suppose (10) holds for some $g \in \boldsymbol{X}$. Then (45) in 1.5 and the definition of $\mathbf{I}$ imply $\mathbf{I}g$ and $f - \widehat{f}(0)$ have the same Walsh-Fourier coefficients. Since the Walsh system is complete, it follows that

$$\mathbf{I}g = f - \widehat{f}(0).$$

Therefore, we have by Theorem 6 that f is differentiable in $\boldsymbol{X}$ and

$$\mathbf{d}f = \mathbf{d}(\mathbf{I}g + \widehat{f}(0)) = g. \quad \blacksquare$$

COROLLARY 2. *Let $\mathcal{D} \subset \boldsymbol{X}$ denote the collection of functions $f \in \boldsymbol{X}$ which are differentiable in $\boldsymbol{X}$. Then the differential operator $\mathbf{d} : \mathcal{D} \to \boldsymbol{X}$ is closed.*

PROOF. Let $f_n \in \mathcal{D}, f_n \to f$ in $\boldsymbol{X}$, and $\mathbf{d}f_n \to g$ in $\boldsymbol{X}$, as $n \to \infty$. By Corollary 1, it suffices to show (10) holds. Since

$$|\widehat{h}(k)| \leq \|h\|_{\boldsymbol{X}}$$

for each $k \in \mathbf{N}$ and $h \in \boldsymbol{X}$, the assumptions imply $\widehat{f}_n(k) \to \widehat{f}(k)$ and $\widehat{\mathbf{d}f_n}(k) \to \widehat{g}(k)$ as $n \to \infty$. We conclude by Corollary 1 that

$$\begin{aligned}\widehat{g}(k) &= \lim_{n\to\infty} \widehat{\mathbf{d}f_n}(k) \\ &= \lim_{n\to\infty} k\widehat{f}_n(k) \\ &= k\widehat{f}(k). \quad \blacksquare\end{aligned}$$

A simple relationship occurs between the moduli of continuity of a function and its strong derivative.

LEMMA 1. *Let $f \in \boldsymbol{X}$ be differentiable in $\boldsymbol{X}$ of order r for some $r \in \mathbf{P}$. Then*

$$\omega^{(\boldsymbol{X})}(f,\delta) = O(\delta^r \omega^{(\boldsymbol{X})}(\mathbf{d}^{[r]}f,\delta)) \quad \text{as} \quad \delta \to 0.$$

PROOF. We may suppose $r = 1$. Fix $n \in \mathbf{N}$, let

$$W_1^{(n)} := \sum_{k=2^n}^{\infty} \frac{\psi_k}{k},$$

and recall from Theorem 17 in 1.8 that

$$\|W_1^{(n)}\|_1 = O(2^{-n}) \quad \text{as} \quad n \to \infty. \tag{11}$$

Let $y \in I_n(0)$ and $f \in \boldsymbol{X}$. By comparing the Walsh- Fourier coefficients of both sides, observe that

$$f - \boldsymbol{\tau}_y f = W_1^{(n)} * \mathbf{d}f - \boldsymbol{\tau}_y(W_1^{(n)} * \mathbf{d}f).$$

Since

$$\boldsymbol{\tau}_y(W_1^{(n)} * \mathbf{d}f) = W_1^{(n)} * (\boldsymbol{\tau}_y \mathbf{d}f)$$

we obtain

$$f - \boldsymbol{\tau}_y f = W_1^{(n)} * (\mathbf{d}f - \boldsymbol{\tau}_y \mathbf{d}f).$$

Consequently,

$$\|f - \boldsymbol{\tau}_y f\|_{\boldsymbol{X}} \leq \|W_1^{(n)}\|_1 \, \|\mathbf{d}f - \boldsymbol{\tau}_y \mathbf{d}f\|_{\boldsymbol{X}}.$$

In particular, if we choose $2^{-n-1} < \delta \leq 2^{-n}$, then (11) implies

$$\begin{aligned}\omega^{(\boldsymbol{X})}(f,\delta) &= O(2^{-n}\omega^{(\boldsymbol{X})}(\mathbf{d}f,\delta)) \\ &= O(\delta\omega^{(\boldsymbol{X})}(\mathbf{d}f,\delta)) \quad \text{as} \quad \delta \to 0. \quad \blacksquare\end{aligned}$$

Differentiability plays a role in the rate of approximation by Walsh polynomials.

THEOREM 7. *Let r, $n \in \mathbf{P}$. If $f \in \boldsymbol{X}$ is differentiable in $\boldsymbol{X}$ of order r, then*

$$\mathrm{E}_n(f,\boldsymbol{X}) \le Cn^{-r}\|\mathbf{d}^{[r]}f\|_{\boldsymbol{X}}$$

where C is a constant which depends only on r.

PROOF. We may suppose that $r=1$. Let $2^s \le n < 2^{s+1}$. By Theorem 2 in 5.1,

$$\mathrm{E}_n(f,\boldsymbol{X}) \le \mathrm{E}_{2^s}(f,\boldsymbol{X}) \le \omega^{(\boldsymbol{X})}(f,2^{-s}).$$

Moreover, by Lemma 1 we have

$$\begin{aligned}\omega^{(\boldsymbol{X})}(f,2^{-s}) &= O\left(2^{-s}\omega^{(\boldsymbol{X})}(\mathbf{d}f,2^{-s})\right)\\ &= O(2^{-s}\|\mathbf{d}f\|_{\boldsymbol{X}}) \quad \text{as} \quad s\to\infty.\end{aligned}$$

Consequently,

$$\mathrm{E}_n(f,\boldsymbol{X}) = O\left(\frac{1}{n}\|\mathbf{d}f\|_{\boldsymbol{X}}\right) \quad \text{as} \quad n\to\infty. \qquad \blacksquare$$

We shall see (Theorem 8) that the rate of approximation by Walsh polynomials characterizes the collection of functions whose strong derivatives belong to $\mathrm{Lip}(\alpha,\boldsymbol{X})$ for some $\alpha>0$. First, we prove the following.

LEMMA 2. *For every $r,n \in \mathbf{P}$ and every polynomial $P \in \mathcal{P}_n$,*

$$\|\mathbf{d}^{[r]}P\|_{\boldsymbol{X}} \le 2^r n^r \|P\|_{\boldsymbol{X}}.$$

PROOF. We may suppose $r=1$. Fix $n \in \mathbf{N}$ with binary expansion

$$n = 2^s + \sum_{i=0}^{s-1} n_i 2^i.$$

Observe since $\mathcal{P}_n \subseteq \mathcal{P}_{2^{s+1}}$ that

$$\frac{1}{n}\|\mathbf{d}P\|_{\boldsymbol{X}} = \frac{1}{n}\|\sum_{j=0}^{s} 2^{j-1}(P-\boldsymbol{\tau}_{e_j}P)\|_{\boldsymbol{X}} \le \frac{1}{n}\sum_{j=0}^{s} 2^j\|P\|_{\boldsymbol{X}},$$

holds for any $P \in \mathcal{P}_n$. Since $2^{s+1}/n \le 2$ we conclude that

$$\|\mathbf{d}P\|_{\boldsymbol{X}} \le 2n\|P\|_{\boldsymbol{X}}. \qquad \blacksquare$$

THEOREM 8. *Let $f \in \boldsymbol{X}, \alpha > 0$, and $r \in \mathbf{P}$. Then the following four conditions are equivalent:*

a) $\mathbf{d}^{[r]}f$ *exists in $\boldsymbol{X}$ and belongs to* $\mathrm{Lip}(\alpha,\boldsymbol{X})$,
b) $\mathbf{d}^{[r]}f$ *exists in $\boldsymbol{X}$ and* $\omega^{(\boldsymbol{X})}(\mathbf{d}^{[r]}f,1/n) = O(n^{-\alpha})$ *as* $n\to\infty$,
c) $\mathrm{E}_n(f,\boldsymbol{X}) = O(n^{-r-\alpha})$ *as* $n\to\infty$,

d) $\mathbf{d}^{[j]}f$ *exists in* $\boldsymbol{X}$ *for* $j \le r$ *and*

$$\|\mathbf{d}^{[j]}f - \mathbf{d}^{[j]}P_n\|_{\boldsymbol{X}} = O(n^{-r-\alpha+j})$$

as $n \to \infty$*, for each best approximation* P_n *of* f *in* $\mathcal{P}_n$ *and each* $0 \le j \le r$*.*

PROOF. By definition, a) implies b).

For $n, s \in \mathbf{P}$ satisfying $2^s \le n < 2^{s+1}$, we have by Theorem 2 in 5.1 that

$$\mathbf{E}_n(f,\boldsymbol{X}) \le \mathbf{E}_{2^s}(f,\boldsymbol{X}) \le \omega^{(\boldsymbol{X})}(f, 2^{-s}).$$

Moreover, Lemma 1 and condition b) imply

$$\begin{aligned}\omega^{(\boldsymbol{X})}(f, 2^{-s}) &= O\left(2^{-rs}\omega^{(\boldsymbol{X})}(\mathbf{d}^{[r]}f, 2^{-s})\right)\\ &= O(2^{-rs}2^{-s\alpha})\\ &= O(n^{-r-\alpha}) \quad \text{as} \quad n \to \infty.\end{aligned}$$

Thus b) implies c).

To prove c) implies d), fix $0 \le j \le r$, set $U_1 := P_2$ and

$$U_k := P_{2^k} - P_{2^{k-1}} \qquad (k \in \mathbf{P}, k \ge 2). \tag{12}$$

It is evident that

$$f = \sum_{k=1}^{\infty} U_k,$$

and this series converges in the norm of $\boldsymbol{X}$. Moreover, condition c) together with Lemma 2 imply that

$$\|\mathbf{d}^{[j]}U_k\|_{\boldsymbol{X}} = O(2^{kj}\|U_k\|_{\boldsymbol{X}}) = O(2^{kj}2^{-k(r+\alpha)})$$

as $k \to \infty$. Since each operator $\mathbf{d}^{[j]}$ is closed (see Corollary 2), it follows that

$$\mathbf{d}^{[j]}f = \sum_{k=1}^{\infty} \mathbf{d}^{[j]}U_k,$$

and

$$\|\mathbf{d}^{[j]}f - \mathbf{d}^{[j]}(P_{2^m})\|_{\boldsymbol{X}} = \|\sum_{k=m+1}^{\infty} \mathbf{d}^{[j]}U_k\|_{\boldsymbol{X}} = O(2^{m(j-\alpha-r)})$$

as $m \to \infty$. In particular, condition d) is verified when $n = 2^m$.

For arbitrary $n \in \mathbf{N}$, let $n = 2^m + i$ where $0 \le i < 2^m$ and observe that

$$\|\mathbf{d}^{[j]}f - \mathbf{d}^{[j]}P_n\|_{\boldsymbol{X}} \le \|\mathbf{d}^{[j]}f - \mathbf{d}^{[j]}(P_{2^m}f)\|_{\boldsymbol{X}} + \|\mathbf{d}^{[j]}(P_nf - P_{2^m}f)\|_{\boldsymbol{X}}.$$

By Lemma 2 and the case already considered, we continue this estimate as follows:

$$\begin{aligned}\|\mathbf{d}^{[j]}f - \mathbf{d}^{[j]}P_n\|_{\boldsymbol{X}} &= O(2^{m(j-\alpha-r)}) + O(n^j\|P_nf - P_{2^m}f\|_{\boldsymbol{X}}) \\ &= O\left(2^{m(j-\alpha-r)} + n^j(n^{-r-\alpha} + 2^{-m(r+\alpha)})\right) \\ &= O(n^{j-\alpha-r})\end{aligned}$$

as $n \to \infty$. Consequently, c) implies d).

Finally, to show d) implies a) let $n := 2^m + i, 0 \le i < 2^m$ be integers and use Theorem 2 in 5.1 to obtain

$$\begin{aligned}\omega^{(\boldsymbol{X})}(\mathbf{d}^{[r]}f, \frac{1}{n}) &\le \omega^{(\boldsymbol{X})}(\mathbf{d}^{[r]}f, 2^{-m}) \\ &\le 2\|\mathbf{d}^{[r]}f - S_{2^m}(\mathbf{d}^{[r]}f)\|_{\boldsymbol{X}}.\end{aligned}$$

Since $S_{2^m}(\mathbf{d}^{[r]}f) = \mathbf{d}^{[r]}S_{2^m}f$ (see Corollary 1 above), it follows that

$$\begin{aligned}\omega^{(\boldsymbol{X})}(\mathbf{d}^{[r]}f, \frac{1}{n}) &\le 2\|\mathbf{d}^{[r]}f - \mathbf{d}^{[r]}(P_{2^m})\|_{\boldsymbol{X}} + 2\|\mathbf{d}^{[r]}(P_{2^m} - S_{2^m}f)\|_{\boldsymbol{X}} \\ &= 2\|\mathbf{d}^{[r]}f - \mathbf{d}^{[r]}(P_{2^m})\|_{\boldsymbol{X}} + 2\|S_{2^m}\left(\mathbf{d}^{[r]}(P_{2^m}) - \mathbf{d}^{[r]}f\right)\|_{\boldsymbol{X}}\end{aligned}$$

In particular, d) implies

$$\omega^{(\boldsymbol{X})}(\mathbf{d}^{[r]}f, \frac{1}{n}) = O(2^{m(r-r-\alpha)})$$

as $n \to \infty$. By definition, then, $\mathbf{d}^{[r]}f$ belongs to $\mathrm{Lip}(\alpha, \boldsymbol{X})$. ∎

This result shows there is a difference between approximation properties of the trigonometric and the Walsh systems. For example, in the trigonometric case the class of functions whose rate of approximation by polynomials of order n is $O(1/n)$ is characterized only by a higher order modulus of continuity. Hence the equivalence of b) and c) in the trigonometric case does not hold.

The following result estimates the approximation order of the partial sums of a Walsh-Fourier series when the function in question is strongly differentiable.

THEOREM 9. *Let $f \in \boldsymbol{X}$ be differentiable in $\boldsymbol{X}$ of order r for some $r \in \mathbf{P}$. Then*

$$\|f - S_nf\|_{\boldsymbol{X}} = O(n^{-r}\log n)\omega^{(\boldsymbol{X})}(\mathbf{d}^{[r]}f, \frac{1}{n})$$

as $n \to \infty$. In particular, if $\mathbf{d}^{[r]}f$ belongs to $\mathrm{Lip}(\alpha, \boldsymbol{X})$ for some $\alpha > 0$ then

$$\|f - S_nf\|_{\boldsymbol{X}} = O\left(\frac{\log n}{n^{r+\alpha}}\right) \quad \text{as} \quad n \to \infty.$$

PROOF. By (2) in 5.1, it is clear that

$$\|f - S_nf\|_{\boldsymbol{X}} = O(\log n)\mathbf{E}_n(f, \boldsymbol{X})$$

as $n \to \infty$. Moreover, if $n = 2^m + i$ where $0 \le i < 2^m$, then Theorem 2 implies

$$\mathbf{E}_n(f,\boldsymbol{X}) \le \mathbf{E}_{2^m}(f,\boldsymbol{X}) \le \omega^{(\boldsymbol{X})}(f,2^{-m}).$$

It follows, therefore, from Lemma 1 that

$$\begin{aligned}\|f - S_n f\|_{\boldsymbol{X}} &= O(\log n)\omega^{(\boldsymbol{X})}(f,2^{-m}) \\ &= O(n^{-r}\log n)\omega^{(\boldsymbol{X})}(\mathrm{d}^{[r]}f, \frac{1}{n})\end{aligned}$$

as $n \to \infty$. ∎

It is interesting to compare this result for $r = 0$ with Theorem 4. Apparently, the rate of approximation of an $f \in \mathrm{Lip}(\alpha,\boldsymbol{X})$ by the Cesàro means $\sigma_n f$, compared to the rate of approximation by the partial sums $S_n f$, is better for $0 < \alpha < 1$, the same for $\alpha = 1$, but worse for $\alpha > 1$.

This same phenomenon occurs when a derivative of f belongs to $\mathrm{Lip}(\alpha,\boldsymbol{X})$, but the index which marks the change from better to worse is shifted by the order of differentiablity. Specifically, the following estimates hold:

THEOREM 10. *Suppose* $\mathrm{d}^{[r]}f$ *exists in* $\boldsymbol{X}$ *and belongs to* $\mathrm{Lip}(\alpha,\boldsymbol{X})$ *for some* $r \in \mathbf{P}$ *and* $\alpha > 0$. *Then*

$$\|\sigma_n f - f\|_{\boldsymbol{X}} = \begin{cases} O(n^{-r-\alpha}) & \alpha + r < 1, \\ O(\log n/n) & \alpha + r = 1, \\ O(1/n) & \alpha + r > 1. \end{cases}$$

PROOF. Fix $r \in \mathbf{P}$ and $\alpha > 0$. If $\mathrm{d}^{[r]}f$ exists in $\boldsymbol{X}$ and belongs to $\mathrm{Lip}(\alpha,\boldsymbol{X})$ then by Theorem 3 in 5.1 and Lemma 1 above we have that f belongs to $\mathrm{Lip}(r+\alpha,\boldsymbol{X})$. Consequently, the order growth conditions follow immediately from Theorem 4 in 5.1. ∎

5.3 The Haar, Walsh, and Faber-Schauder Systems as Bases. Let $(\boldsymbol{X}, \|.\|_{\boldsymbol{X}})$ be a Banach space. A sequence $\boldsymbol{\epsilon} = (\epsilon_n, n \in \mathbf{N})$ in $\boldsymbol{X}$ is called a (Schauder) *basis* in $\boldsymbol{X}$ if to each $x \in \boldsymbol{X}$ there corresponds a unique sequence of numbers $(x_k, k \in \mathbf{N})$, called the coordinates of x (with respect to $\boldsymbol{\epsilon}$), such that

$$x = \sum_{k=0}^{\infty} x_k \epsilon_k \tag{13}$$

in $\boldsymbol{X}$, i.e., the sequence of partial sums

$$S_n^{\epsilon} x := \sum_{k=0}^{n-1} x_k \epsilon_k$$

converges to x in the norm of $\boldsymbol{X}$ as $n \to \infty$.

In the next three sections we shall identify, for each of the systems Haar, Walsh, Faber-Schauder, Franklin, and Ciesielski, subspaces of L^1 in which the given system is a basis. A summary of these results is located at the end of 5.5.

We begin with several elementary observations of a general nature.

A basis in $\boldsymbol{X}$ is always a closed system in $\boldsymbol{X}$.

Given a basis $\boldsymbol{\epsilon} = (\epsilon_n, n \in \mathrm{N})$ in $\boldsymbol{X}$, coordinate functionals $\epsilon'_n \in \boldsymbol{X}'$ can be defined by

$$\epsilon'_n(x) := x_n \qquad (x \in \boldsymbol{X}, n \in \mathrm{N}),$$

where the x_n's are the coordinates of x with respect to $\boldsymbol{\epsilon}$. Using the notation $\langle x, x' \rangle$ for $x'(x)$ when $x \in \boldsymbol{X}$ and $x' \in \boldsymbol{X}'$, observe by definition that

$$\text{(14)} \qquad \langle \epsilon_n, \epsilon'_m \rangle = \delta_{nm} \qquad (n, m \in \mathrm{N}),$$

where δ_{nm} is the Kronecker delta. Thus the systems $\boldsymbol{\epsilon}$ and $\boldsymbol{\epsilon}' := (\epsilon'_n, n \in \mathrm{N})$ are biorthogonal.

For every closed system $\boldsymbol{\epsilon} = (\epsilon_n, n \in \mathrm{N})$ in $\mathbf{X}$ there is at most one system $\boldsymbol{\epsilon}' = (\epsilon'_n, n \in \mathrm{N})$ in $\boldsymbol{X}'$ such that $\boldsymbol{\epsilon}$ and $\boldsymbol{\epsilon}'$ are biorthogonal, i.e., satisfy (14). If $\boldsymbol{\epsilon}$ and $\boldsymbol{\epsilon}'$ are biorthogonal, then the biorthogonal expansion of $x \in \boldsymbol{X}$ is given by

$$\text{(15)} \qquad \sum_{n=0}^{\infty} \langle x, \epsilon'_n \rangle \, \epsilon_n.$$

Clearly, if a series of the form (13) converges to x in the norm of $\boldsymbol{X}$, then biorthogonality implies $x_n = \langle x, \epsilon'_n \rangle$ for $n \in \mathrm{N}$. Consequently, the partial sums of (15) will also be denoted by $S^{\boldsymbol{\epsilon}}_n x$ for $n \in \mathrm{N}$.

The Banach-Steinhaus theorem (see Appendix 0.0) gives a simple characterization of bases. A closed system $\boldsymbol{\epsilon} = (\epsilon_n, n \in \mathrm{N})$ in $\boldsymbol{X}$ is a basis in $\boldsymbol{X}$ if and only if there is a system $\boldsymbol{\epsilon}' = (\epsilon'_n, n \in \mathrm{N})$ in $\boldsymbol{X}'$ biorthogonal to $\boldsymbol{\epsilon}$ such that

$$\text{(16)} \qquad \sup_{n \in \mathrm{N}} \| S^{\boldsymbol{\epsilon}}_n \|_{\boldsymbol{X}} < \infty.$$

For any system $\boldsymbol{\epsilon} = (\epsilon_n, n \in \mathrm{N})$ in a vector space $\boldsymbol{X}$, let $\mathrm{L}(\boldsymbol{\epsilon})$ represent the linear hull of $\boldsymbol{\epsilon}$. Thus given $x \in \mathrm{L}(\boldsymbol{\epsilon})$ there is an $N \in \mathrm{N}$ and numbers $b_0, b_1, \ldots, b_N$ such that

$$x = \sum_{n=0}^{N} b_n \epsilon_n.$$

When $\boldsymbol{X}$ is a normed linear space we shall denote the closure of $\mathrm{L}(\boldsymbol{\epsilon})$ in $\boldsymbol{X}$ by $\mathrm{CL}(\boldsymbol{\epsilon})$. Thus for $\boldsymbol{X} := \mathrm{L}^2$ and $\boldsymbol{\epsilon}$ the Walsh system we have $\mathrm{L}(\boldsymbol{\epsilon}) = \mathcal{P}$ and $\mathrm{CL}(\boldsymbol{\epsilon}) = \mathrm{L}^2$.

We shall show that the coordinate functional system $\boldsymbol{\epsilon}'$ generated by a basis $\boldsymbol{\epsilon}$ in a Banach space $\boldsymbol{X}$ is a basis in $\mathrm{CL}(\boldsymbol{\epsilon}')$. Indeed, recall that

$$x(x') := \langle x, x' \rangle \qquad (x \in \boldsymbol{X}, x' \in \boldsymbol{X}')$$

isometrically embeds $\boldsymbol{X}$ in $\boldsymbol{X}''$. In particular,

$$\|x\|_{\boldsymbol{X}} = \sup_{\|x'\|_{\boldsymbol{X}'} \le 1} |\langle x, x' \rangle| \qquad (x \in \boldsymbol{X}),$$

and the series

$$\sum_{n=0}^{\infty} \langle \epsilon_n, x' \rangle \epsilon'_n \tag{17}$$

can be viewed as the biorthogonal expansion of $x' \in \boldsymbol{X}'$ with respect to the system $\boldsymbol{\epsilon}'$. Denote the partial sums of (17) by $S_n^{\boldsymbol{\epsilon}'} x'$. By definition,

$$\left\langle x, S_n^{\boldsymbol{\epsilon}'} x' \right\rangle = \sum_{k=0}^{n-1} \langle \epsilon_k, x' \rangle \langle x, \epsilon'_k \rangle$$

for $n \in \mathrm{N}$ and $x \in \boldsymbol{X}$. Consequently,

$$\begin{aligned} \|S_n^{\boldsymbol{\epsilon}'} x'\|_{\boldsymbol{X}'} &= \sup_{\|x\|_{\boldsymbol{X}} \le 1} \left| \left\langle x, S_n^{\boldsymbol{\epsilon}'} x' \right\rangle \right| \\ &= \sup_{\|x\|_{\boldsymbol{X}} \le 1} \left| \langle S_n^{\boldsymbol{\epsilon}} x, x' \rangle \right| \\ &\le \sup_{\|x\|_{\boldsymbol{X}} \le 1} \|S_n^{\boldsymbol{\epsilon}} x\|_{\boldsymbol{X}} \|x'\|_{\boldsymbol{X}'} \\ &= \|S_n^{\boldsymbol{\epsilon}}\|_{\boldsymbol{X}} \|x'\|_{\boldsymbol{X}'} \end{aligned}$$

for $x' \in \boldsymbol{X}'$ and $n \in \mathrm{N}$. It follows that $\|S_n^{\boldsymbol{\epsilon}'}\|_{\boldsymbol{X}'} \le \|S_n^{\boldsymbol{\epsilon}}\|_{\boldsymbol{X}}$ for $n \in \mathrm{N}$. A similar argument establishes the reverse inequality. Therefore,

$$\|S_n^{\boldsymbol{\epsilon}'}\|_{\boldsymbol{X}'} = \|S_n^{\boldsymbol{\epsilon}}\|_{\boldsymbol{X}} \qquad (n \in \mathrm{N}). \tag{18}$$

In view of (16), we conclude that $\boldsymbol{\epsilon}'$ is a basis in $\mathrm{CL}(\boldsymbol{\epsilon}')$.

For specific situations, the following result will be used to estimate the operator norm of $S_n^{\boldsymbol{\epsilon}}$.

LEMMA 3. *Let* $[0,1)^2 := [0,1) \times [0,1)$. *Suppose* $K \in \mathrm{L}^{\infty}([0,1)^2)$ *and that the numbers*

$$\|K\|_{[1,\infty]} := \sup_{t \in [0,1]} \int_0^1 |K(x,t)|\, dx$$

and

$$\|K\|_{[\infty,1]} := \sup_{x \in [0,1]} \int_0^1 |K(x,t)|\, dt$$

are finite. If $1 \le p \le \infty$, p' *is the index conjugate to* p, *and*

$$(\boldsymbol{K}f)(x) := \int_0^1 K(x,t) f(t)\, dt \qquad (f \in \mathrm{L}^1, x \in [0,1)),$$

then

$$\sup_{\|f\|_p \le 1} \|\boldsymbol{K}f\|_p \le \frac{1}{p} \|K\|_{[1,\infty]} + \frac{1}{p'} \|K\|_{[\infty,1]}.$$

PROOF. For $p = 1$ or $p = \infty$ the inequality is trivial.
For $1 < p < \infty$ observe that

$$|uv| \le \frac{|u|^p}{p} + \frac{|v|^{p'}}{p'} \qquad (u, v \in \mathbf{R}).$$

Consequently, if $\|f\|_p \le 1$ and $\|\psi\|_{p'} \le 1$ then

$$\begin{aligned} |\int_0^1 (\boldsymbol{K}f)\psi| &= |\int_0^1 \int_0^1 K(x,t)f(t)\psi(x)\,dx\,dt| \\ &\le \int_0^1 \frac{|f(t)|^p}{p}\left(\int_0^1 |K(x,t)|\,dx\right)dt + \int_0^1 \frac{|\psi(x)|^{p'}}{p'}\left(\int_0^1 |K(x,t)|\,dt\right)dx \\ &\le \frac{1}{p}\|K\|_{[1,\infty]} + \frac{1}{p'}\|K\|_{[\infty,1]}. \end{aligned}$$

Therefore, the proof of this lemma is completed by taking the supremum of this inequality over all functions ψ which satisfy $\|\psi\|_{p'} \le 1$. ∎

All Banach spaces considered in the remainder of this chapter will be subspaces $\boldsymbol{X}$ of L^1 which satisfy

$$\|f\|_1 \le \|f\|_{\boldsymbol{X}} \qquad (f \in \boldsymbol{X}). \tag{19}$$

This allows L^∞ to be viewed as a subspace of $\boldsymbol{X}'$. Indeed, the L^2 inner product of $f \in \boldsymbol{X}$ with $\psi \in L^\infty$ satisfies

$$|\langle f, \psi\rangle| \le \|\psi\|_\infty \|f\|_1 \le \|\psi\|_\infty \|f\|_{\boldsymbol{X}}$$

and thus the map $f \to \langle f, \psi\rangle$ is a bounded linear functional on $\boldsymbol{X}$ for each $\psi \in L^\infty$.

Let $\boldsymbol{X}$ be a Banach space which satisfies (19). Any orthonormal system $\boldsymbol{g} = (g_n, n \in \mathrm{N})$ in L^2 whose elements g_n belong to $\boldsymbol{X} \cap L^\infty$ can be considered a biorthonormal system with $g_n' := g_n$ for $n \in \mathrm{N}$. Moreover, the biorthogonal expansion of an $f \in L^1$ is the Fourier series of f with respect to the system $\boldsymbol{g}$. In particular, if we define Dirichlet kernels with respect to $\boldsymbol{g}$ by

$$D_n^{\boldsymbol{g}}(x,t) := \sum_{k=0}^{n-1} g_k(x)g_k(t) \qquad (x, t \in [0,1), n \in \mathrm{N})$$

then the biorthogonal expansion of an integrable f satisfies

$$(S_n^{\boldsymbol{g}} f)(x) = \int_0^1 D_n^{\boldsymbol{g}}(x,t)f(t)dt$$

for $x \in [0,1)$ and $n \in \mathrm{N}$. Moreover, it is clear that $D_n^{\boldsymbol{g}}(x,t) = D_n^{\boldsymbol{g}}(t,x)$. Therefore, it follows from Lemma 3 that the operator norm of $S_n^{\boldsymbol{g}}(x,t) : L^p \to L^p$ satisfies

$$\|S_n^{\boldsymbol{g}}\|_{L^p} \le \sup_{x\in[0,1)} L_n^{\boldsymbol{g}}(x)$$

for $1 \le p \le \infty, n \in \mathbf{N}$, where $L_n^{\boldsymbol{g}}$ represents the Lebesgue functions associated with $\boldsymbol{g}$, i.e.,

$$L_n^{\boldsymbol{g}}(x) := \int_0^1 |D_n^{\boldsymbol{g}}(x,t)|\, dt \qquad (x \in [0,1), n \in \mathbf{N}).$$

In particular, if the Lebesgue functions associated with an orthonormal system $\boldsymbol{g}$ are uniformly bounded then for each $1 \le p \le \infty$ the system $\boldsymbol{g}$ is a basis in the closed subspace of L^p generated by $\boldsymbol{g}$.

We apply these observations first to the Haar system $\boldsymbol{h}$ and the Walsh system $\boldsymbol{w}$. By (42) in 1.4 we have that $S_{2^n}^{\boldsymbol{h}} = S_{2^n}^{\boldsymbol{w}}$ for $n \in \mathbf{N}$. Hence $\boldsymbol{h}$ is a closed system in C_W, and in L^p for $1 \le p \le \infty$, and

$$D_{2^n}^{\boldsymbol{h}}(x,t) = D_{2^n}^{\boldsymbol{w}}(x,t) = D_{2^n}(x \dot{+} t)$$

for $x, t \in [0,1), n \in \mathbf{N}$. Moreover, the definition of the Haar system implies

$$D_{2^n+k}^{\boldsymbol{h}}(x,t) = \begin{cases} D_{2^{n+1}}^{\boldsymbol{h}}(x,t) & x,t \in [0, k/2^n) \\ D_{2^n}^{\boldsymbol{h}}(x,t) & \text{otherwise,} \end{cases}$$

for $0 \le k < 2^n$. Hence the Lebesgue functions associated with $\boldsymbol{h}$ satisfy

$$L_m^{\boldsymbol{h}}(x) = 1 \qquad (m \in \mathbf{N}, x \in [0,1)).$$

Therefore, $\boldsymbol{h}$ is a basis in C_W and in L^p for $1 \le p < \infty$.

The Haar system is also a basis in the dyadic Hardy space H. Indeed, since

$$\mathcal{E}^*(S_n^{\boldsymbol{h}} f) \le \mathcal{E}^* f \qquad (f \in \mathrm{L}^1, n \in \mathbf{N}),$$

it follows by definition that $\|S_n^{\boldsymbol{h}} f\|_{\mathrm{H}} \le \|f\|_{\mathrm{H}}$ for $n \in \mathbf{N}$.

For the space VMO, notice first that

$$\int_0^1 f\, S_n^{\boldsymbol{h}} \psi = \int_0^1 \psi\, S_n^{\boldsymbol{h}} f \qquad (f, \psi \in \mathrm{L}^1, n \in \mathbf{N}).$$

Consequently, Theorem 10 in 3.5 implies

$$\|S_n^{\boldsymbol{h}}\|_{\mathrm{VMO}} \le 2C \|S_n^{\boldsymbol{h}}\|_{\mathrm{H}}$$

for some absolute constant $C > 0$. Next, notice that the Haar function $h_0 \notin \mathrm{BMO}$ since all functions in BMO are of mean zero. Finally, observe since $S_{2^n}^{\boldsymbol{h}} = S_{2^n}^{\boldsymbol{w}}$ that the Haar system is closed in VMO. It follows that the shifted Haar system $(h_n, n \in \mathbf{P})$ is a basis in VMO.

The situation is different for the Walsh system. We shall prove that $\boldsymbol{w}$ is a basis in L^p for $1 < p < \infty$ but not a basis in L^1, in C_W, in H, or in VMO.

For H and VMO, fix $m \in \mathbf{N}$, let $f := r_m D_{2^m}$ and set $n := 2^m + k$ for some $0 \le k < 2^m$. Then $\|f\|_{\mathrm{H}} = 1$ and

$$S_n^{\boldsymbol{w}} f = S_n f = r_m D_k.$$

Since $2^m \leq n < 2^{m+1}$ implies $\mathcal{E}_{m+1}(S_n f) = S_n f$, it follows that

$$\|\mathcal{E}^*(S_n f)\|_1 \geq \|D_k\|_1.$$

Since $\limsup_{k\to\infty} \|D_k\|_1 = \infty$ (see Theorem 9 in 1.6), we conclude that

$$\sup\{\|S_n f\|_{\mathrm{H}} : \|f\|_{\mathrm{H}} \leq 1, n \in \mathbf{N}\} = \infty.$$

Therefore, $\boldsymbol{w}$ is not a basis in H or in VMO.

For C_W recall from the paragraph preceding Theorem 4 in 4.2 that the operators $S_n : \mathcal{P} \to \mathrm{L}^\infty$ have operator norms which satisfy

$$\limsup_{n\to\infty} \|S_n\|_\infty = \limsup_{n\to\infty} \|D_n\|_1 = \infty.$$

Therefore, $\boldsymbol{w}$ is not a basis in C_W.

For $\mathrm{L}^p, 1 \leq p < \infty$, see Theorem 1 and Theorem 2 in 4.1.

It is interesting to note that the Walsh system remains a basis in L^p, $1 < p < \infty$, under any piecewise linear rearrangement; hence the Walsh-Kaczmarz system and the original Walsh system are both bases in L^p for $1 < p < \infty$. To see this, first suppose that $T : \mathbf{N} \to \mathbf{N}$ is linear and consider the linear rearrangement

$$T\boldsymbol{w} := (w_{T(n)}, n \in \mathbf{N}).$$

Choose by Theorem 7 in 1.4 a $1-1$, measure preserving transformation $T' : [0,1) \to [0,1)$ such that

$$w_{T(n)}(x) = w_n(T'(x)) \qquad (x \in [0,1), n \in \mathbf{N}).$$

Hence if T^* represents the inverse of T', the $T\boldsymbol{w}$- Fourier coefficients of an $f \in \mathrm{L}^1$ are given by

$$\begin{aligned}\int_0^1 f w_{T(n)} &= \int_0^1 (f)(w_n \circ T') \\ &= \int_0^1 (f \circ T^*)\, w_n \\ &= \widehat{(f \circ T^*)}(n) \qquad (n \in \mathbf{N}).\end{aligned}$$

Consequently, the partial sums of the $T\boldsymbol{w}$-Fourier series of f satisfy

$$\begin{aligned}(S_n^{T\boldsymbol{w}} f)(x) &= \sum_{k=0}^{n-1} \widehat{(f \circ T^*)}(k) w_k(T'(x)) \\ &= S_n(f \circ T^*)(T'(x))\end{aligned}$$

for $x \in [0,1), n \in \mathbf{N}$. Since T' and T^* are measure preserving, it follows that

$$\sup_{\|f\|_p \leq 1} \|S_n^{T\boldsymbol{w}} f\|_p = \sup_{\|g\|_p \leq 1} \|S_n^{\boldsymbol{w}} g\|_p$$

for $1 \le p \le \infty$. In particular, $T\boldsymbol{w}$ is a basis in L^p for some $1 \le p \le \infty$ if and only if $\boldsymbol{w}$ is.

Next, suppose $R : \mathbf{N} \to \mathbf{N}$ is piecewise linear. Consider the rearrangement

$$R\boldsymbol{w} := (w_{R(n)}, n \in \mathbf{N}).$$

By (35) in 1.4 it is easy to see that

$$S_{2^n}^{R\boldsymbol{w}} f = S_{2^n} f$$

and

$$(S_{2^n+k}^{R\boldsymbol{w}} f - S_{2^n}^{R\boldsymbol{w}} f)(x) = r_n(x) S_k(r_n \Delta_n f \circ R_n^*)(R_n'(x))$$

for $0 \le k < 2^n, n \in \mathbf{N}$, and $x \in [0,1)$. Consequently,

$$\|S_{2^n+k}^{R\boldsymbol{w}} f\|_p \le 2 \sup_{m \in \mathbf{N}} \|S_m f\|_p$$

for $1 \le p \le \infty$ and $n, k \in \mathbf{N}$. We conclude that $R\boldsymbol{w}$ is a basis in L^p for $1 < p < \infty$.

Finally, we consider the space $\mathcal{C}[0,1]$ of classically continuous functions on $[0,1]$ and the *Faber-Schauder system* $\boldsymbol{\zeta} = (\zeta_n, n \in \mathbf{N})$. This system is defined by $\zeta_0 := 1$ and

$$\zeta_n(x) := \int_0^x h_{n-1} \qquad (x \in [0,1], n \in \mathbf{P}).$$

Thus ζ_0 is constant on $[0,1], \zeta_1$ is linear, and for $2^m < n \le 2^{m+1}, m \in \mathbf{N}$, the function ζ_n is continuous, piecewise linear, supported on $I := I(n - 2^m - 1, m)$, increasing on the left half of I, and decreasing on the right half of I.

The Faber-Schauder system is essentially biorthogonal to the Haar system. Indeed, by the Riesz representation theorem every functional Λ on $\mathcal{C}[0,1]$ can be written in the form

$$\Lambda g = \int_0^1 g \, dh \qquad (g \in \mathcal{C}[0,1])$$

for some function h of bounded variation which satisfies $\|\Lambda\| = Var(h)$. Thus each Haar function can be viewed as a functional on $\mathcal{C}[0,1]$. If we set

$$\widetilde{h}_0(x) := \begin{cases} -1 & x = 0 \\ 0 & 0 < x \le 1, \end{cases}$$

$$\widetilde{h}_1(x) := \begin{cases} -1 & 0 < x < 1 \\ 0 & x = 0 \text{ or } 1, \end{cases}$$

and

$$\widetilde{h}_n(x) := -h_{n-1}(x)$$

for $x \in [0,1]$ and $n \ge 2$, then it follows from definition that

$$\int_0^1 g \, d\widetilde{h}_0 = g(0)$$

and

$$\int_0^1 g\, d\widetilde{h}_1 = g(1) - g(0).$$

Moreover, if $n \geq 2$ and g is absolutely continuous with $g(0) = g(1) = 0$, then integration by parts yields

$$\int_0^1 g(x)\, d\widetilde{h}_n(x) = \int_0^1 \frac{d}{dx} g(x)\, h_{n-1}(x).$$

Since the Haar system is orthogonal, we conclude that

$$\int_0^1 \zeta_m\, d\widetilde{h}_n = \delta_{mn}$$

for $m, n \in \mathbf{N}$.

Partial sums of the biorthogonal expansion of an $f \in \mathcal{C}[0,1]$ with respect to the Faber-Schauder system are thus given by

(20) $$S^{\zeta}_{2^n+k+1} f = \sum_{m=0}^{2^n+k} c_m(f) \zeta_m$$

for $0 \leq k < 2^n$, where

(21) $$c_m(f) := \int_0^1 f\, d\widetilde{h}_m \qquad (m \in \mathbf{N}).$$

Notice that the partial sum (20) is piecewise linear with knots at the points

$$B^n_k := \{\frac{\ell}{2^{n+1}} : 0 \leq \ell < 2k\} \bigcup \{\frac{s}{2^n} : k \leq s \leq 2^n\}$$

for each $0 \leq k < 2^n, n \in \mathbf{N}$. Thus (23) below shows that $S^{\zeta}_{2^n+k+1}$ can be used to linearly interpolate over the points in B^n_k.

The classical derivative of a biorthogonal Faber-Schauder expansion is a Haar-Fourier series. Indeed, if

$$F(x) := \int_0^x f \qquad (x \in [0,1])$$

for some continuous f, then integration by parts yields

$$c_k(F) = -\int_0^1 \widetilde{h}_k\, dF = \int_0^1 f\, h_{k-1}$$

for $k \geq 1$. Hence by definition,

$$\begin{aligned}\frac{d}{dx}(S^{\zeta}_n F)(x) &= \sum_{k=1}^{n-1} \left(\int_0^1 f\, h_{k-1}\right) \left(\frac{d\zeta_k}{dx}(x)\right) \\ &= \sum_{k=1}^{n-1} \left(\int_0^1 f\, h_{k-1}\right) h_{k-1}(x)\end{aligned}$$

when this derivative exists. In particular,

$$\frac{d}{dx}(S_n^{\zeta}F)(x) = S_{n-1}^{h}f(x) \tag{22}$$

for $n \in \mathbf{P}$, and $x \in [0,1) \setminus \mathbf{Q}$.

For the next two sections let Ω denote the classical local modulus of continuity, i.e.,

$$\Omega(f,I) := \sup_{x,y\in I} |f(x) - f(y)|$$

for $f \in \mathcal{C}[0,1]$, and I any subinterval of $[0,1]$.

Recall that $\mathcal{I}_0$ represents the collection of dyadic intervals in $[0,1)$ and $\mathcal{I} := \mathcal{I}_0 \cup \{\emptyset\}$. The following result shows the Faber-Schauder system is a basis $\mathcal{C}[0,1]$.

THEOREM 11. *Let $f \in \mathcal{C}[0,1], n \in \mathbf{N}$, and $0 \le k < 2^n$. Then*

$$(S_{2^n+k+1}^{\zeta}f)(x) = f(x) \qquad (x \in B_k^n) \tag{23}$$

and

$$\|S_{2^n+k+1}^{\zeta}f - f\|_\infty \le 2\sup\{\Omega(f,I) : |I| = 2^{-n} \quad \text{and} \quad I \in \mathcal{I}_0\}.$$

PROOF. The interpolating property (23) is proved by induction on n.

Since (22) implies $c_0(f) = f(0)$ and $c_1(f) = f(1) - f(0)$, it is clear that (23) holds for $n = 0$.

Suppose (23) holds for $n-1$ and every $0 \le k < 2^{n-1}$. Fix $1 \le \ell \le 2^n$ and observe by (21) that

$$c_{2^n+\ell}(f) = 2^{n/2}\left(2f(\frac{\ell-1/2}{2^n}) - f(\frac{\ell-1}{2^n}) - f(\frac{\ell}{2^n})\right). \tag{24}$$

By definition $\zeta_{2^n+\ell}$ vanishes off $I := I(\ell-1,n)$, and at the midpoint of I satisfies

$$\zeta_{2^n+\ell}(\frac{\ell-1/2}{2^n}) = 2^{-n/2-1}.$$

It follows, therefore, from the inductive hypothesis that

$$\begin{aligned}(S_{2^n+k+1}^{\zeta}f)(\frac{\ell-1/2}{2^n}) &= \frac{1}{2}\left((S_{2^n+1}^{\zeta}f)(\frac{\ell-1}{2^n}) + (S_{2^n+1}^{\zeta}f)(\frac{\ell}{2^n})\right)\\ &\quad + 2^{-n/2-1}2^{n/2}\left(2f(\frac{\ell-1/2}{2^n}) - f(\frac{\ell-1}{2^n}) - f(\frac{\ell}{2^n})\right)\\ &= f(\frac{\ell-1/2}{2^n})\end{aligned}$$

for $\ell \le k < 2^n$. In particular, (23) holds for all $n \in \mathbf{N}$.

Fix $n \in \mathbf{N}, 0 \le k < 2^n$, and $I = [\alpha,\beta) \in \mathcal{I}$ such that $|I| = 2^{-n}$. Clearly, (23) implies

$$\begin{aligned}|(S_{2^n+k+1}^{\zeta}f)(x) - f(x)| &\le |(S_{2^n+k+1}^{\zeta}f)(x) - (S_{2^n+k+1}^{\zeta}f)(\alpha)| + |f(\alpha) - f(x)|\\ &\le |f(\frac{\alpha+\beta}{2}) - f(\alpha)| + |f(\alpha) - f(x)|\\ &\le 2\Omega(f,I),\end{aligned}$$

for all $x \in [\alpha,(\alpha+\beta)/2)$. Similarly,

$$|(S_{2^n+k+1}^{\zeta}f)(x) - f(x)| \le 2\Omega(f,I),$$

for all $x \in [(\alpha+\beta)/2,\beta)$. Consequently, the proof of the theorem is complete. ∎

5.4 The Franklin System. The Faber-Schauder system $\boldsymbol{\zeta}$ is linearly independent but not orthogonal on $[0,1]$. Let $\mathbf{f} = (\mathrm{f}_n, n \in \mathbf{N})$ be the system obtained from $\boldsymbol{\zeta}$ by the Gram-Schmidt orthogonalization process. Thus there is a uniquely determined triangular matrix $(\lambda_{\ell m})$ whose diagonal entries are positive such that

$$\mathrm{f}_m = \sum_{\ell=0}^{m} \lambda_{\ell m} \zeta_\ell \qquad (m \in \mathbf{N}).$$

In particular, each f_m is piecewise linear having m knots on $[0,1)$, each of which is located over a dyadic rational.

The system $\mathbf{f}$ is called the *Franklin system.* By construction it is complete and orthonormal on $[0,1]$.

It is convenient to index the Franklin system by dyadic intervals. Set $\mathrm{f}_\emptyset := \mathrm{f}_0$ and $\mathrm{f}_{[0,1)} := \mathrm{f}_1$. For each $I \in \mathcal{I}$ with $0 < |I| < 1$ set

$$\mathrm{f}_I := \mathrm{f}_{2^n+k}$$

where $I = I(k-1,n) =: [\alpha^I, \beta^I)$ for some $1 \le k \le 2^n$ and $n \in \mathbf{N}$. Denote the collection of knots of each f_I by $B^I := B_k^n$, where $I = I(k-1,n), 1 \le k \le 2^n, n \in \mathbf{P}$, and denote the points in B^I by

$$0 = t_0^I < t_1^I < \cdots < t_{2^n+k}^I = 1.$$

Our discussion of the Franklin system is not self contained. In addition to certain properties of classical Hardy spaces (see the material preceding the proof of Lemma 4 below), we have elected to use without proof the following facts: for every $I \in \mathcal{I}$ satisfying $0 < |I| < 1$

$$\text{(25)} \qquad \begin{cases} \|\mathrm{f}_I\|_1 \le 6\sqrt{3}|I|^{-1/2} \\ \|\sum_{\{J\in\mathcal{I}:|J|=|I|\}} |\mathrm{f}_J|\,\|_\infty \le 32\sqrt{3}|I|^{-1/2}, \end{cases}$$

$$\text{(26)} \qquad \frac{\sqrt{3}}{2}|I|^{-1/2} \le \mathrm{f}_I\left(\frac{\alpha^I+\beta^I}{2}\right) \le 2\sqrt{3}|I|^{-1/2},$$

$$\text{(27)} \qquad \frac{1}{3\sqrt{3}}|I|^{-1/2} \le -\mathrm{f}_I(\alpha^I), \mathrm{f}_I(\beta^I) \le \sqrt{3}|I|^{-1/2},$$

$$\text{(28)} \qquad \mathrm{f}_I(t_k^I) = (-1)^k \frac{\cosh(\frac{2\gamma}{|I|}t_k^I)}{\cosh(\frac{2\gamma}{|I|}\alpha^I)} \mathrm{f}_I(\alpha^I) \qquad (0 \le t_k^I \le \alpha^I),$$

and

$$\text{(29)} \qquad \mathrm{f}_I(t_k^I) = (-1)^k \frac{\cosh(\frac{\gamma}{|I|}|1-t_k^I|)}{\cosh(\frac{\gamma}{|I|}|1-\beta^I|)} \mathrm{f}_I(\beta^I) \qquad (\beta^I \le t_k^I \le 1),$$

where $\gamma := \log(2+\sqrt{3})$, i.e., $\cosh(\gamma) = 2$. Moreover, the Franklin-Dirichlet kernels satisfy

$$(30) \qquad |D_m^{\mathrm{f}}(t,x)| \le Cmq^{m|t-x|/2}$$

for $m \in \mathbf{P}$ and $t, x \in [0,1)$, where C is an absolute constant and $q := 2-\sqrt{3}$. Proofs of these facts can be found in Ciesielski [1] and [2].

Properties (25) through (29) are precise. For most purposes we shall use the less demanding properties (31) through (34) below. They show that the Franklin system is analogous to the Haar system in the following sense. As the Haar function h_I is supported on I and determined by its values at the end points of I for each $I \in \mathcal{I}_0$, even so, the same is nearly true for the Franklin function f_I. We shall develop this analogy further in the following section by showing that the Franklin and Haar systems are equivalent bases in many spaces.

For each pair $I, J \in \mathcal{I}_0$ and $t \in [0,1]$, set

$$r(t,I) := \frac{\mathrm{dist}(t,I)}{|I|}$$

and

$$r(J,I) := \frac{\mathrm{dist}(J,I)}{|I|}.$$

Properties (26) through (29), together with the well-known inequalities

$$\frac{1}{2}e^t < \cosh(t) < e^t \qquad (t > 0)$$

can be used to show there exist constants $C > 0$ and $q = 2-\sqrt{3}$ such that

$$(31) \qquad |\mathrm{f}_I(t)| \le C|I|^{-1/2}q^{r(t,I)},$$

$$(32) \qquad \sup_{t_1,t_2 \in J} |\mathrm{f}_I(t_1) - \mathrm{f}_I(t_2)| \le C|I|^{-3/2}|J|q^{r(J,I)},$$

$$(33) \qquad \left| \int_0^t \mathrm{f}_I(s)\,ds \right| \le C|I|^{1/2}q^{r(t,I)},$$

and

$$(34) \qquad Var(\mathrm{f}_I) \le C|I|^{-1/2}$$

for all $t \in [0,1]$ and $I, J \in \mathcal{I}_0$.

We shall establish certain fundamental inequalities (see (41) and (42) below) concerning Franklin-Fourier coefficients, and identify subspaces of L^1 in which the Franklin system is a basis. Even with the assumptions made above, our program is not an easy one.

At the end of 3.4, we showed that Haar-Fourier coefficients induce an isomorphism between H and **h**, and between BMO and **bmo**. To establish an analogous isomorphism for Franklin-Fourier coefficients between classical $\mathcal{H}$ and **h** and classical $\mathcal{BMO}$ and **bmo**, we need the following facts.

First, $\mathcal{H}$ can be characterized by atoms in the same way H is. The main difference is that the atoms for $\mathcal{H}$ can be supported on any interval, not just dyadic ones. This was mentioned in 3.4 and will be used in the proof of the next lemma.

Secondly, a function $f \in L^2$ of mean zero is said to belong to $\mathcal{BMO}$ if

$$\|f\|_* := \sup_{I\subseteq[0,1)} \left(\frac{1}{|I|}\int_I \left|f - \frac{1}{|I|}\int_I f\right|^2\right)^{1/2}$$

is finite. Here, the supremum is taken over *all* subintervals of $[0,1)$ not just the dyadic ones. Clearly, $\mathcal{BMO} \subset$ BMO.

Third of all, the dual of $\mathcal{H}$ is $\mathcal{BMO}$ and there exists an inner product, generalizing the L^2 inner product, which satisfies

$$|\langle f,\psi\rangle| \le C\|f\|_{\mathcal{H}}\|\psi\|_* \tag{35}$$

where C is an absolute constant independent of $f \in \mathcal{H}$ and $\psi \in \mathcal{BMO}$.

Finally, if $\mathcal{VMO}$ represents the collection of $f \in \mathcal{BMO}$ which satisfy

$$\lim_{|I|\to 0} \frac{1}{|I|}\int_I \left|f - \frac{1}{|I|}\int_I f\right|^2 = 0,$$

and $\mathcal{H}_0 := \{f \in \mathcal{H} : f \text{ is of mean zero }\}$, then the dual of $\mathcal{VMO}$ is isomorphic to $\mathcal{H}_0$, and the dual of $\mathcal{H}_0$ is isomorphic to $\mathcal{BMO}$. Moreover, $\mathcal{VMO}$ is the closure of $\mathcal{C}[0,1]$ in the $\mathcal{BMO}$ norm. Proofs of these facts can be found in Coifman and Weiss [1].

LEMMA 4. *There exist positive constants A, B, C, and an integer $n_0 \in \mathbf{N}$ such that*

$$2^{n/2}\|f_{2^n+k}\|_{\mathcal{H}} \le A \qquad (1 \le k \le 2^n, n \in \mathbf{N}), \tag{36}$$

$$2^{n/2}\|f_{2^n+2^{n-1}+1}\|_{\mathrm{H}} \ge Bn \qquad (n \ge n_0), \tag{37}$$

and

$$\left|\int_0^1 f_{2^n+2^{n-1}+1}h_{2^{n+1}+2^n}\right| \ge C \qquad (n \ge n_0). \tag{38}$$

PROOF. Fix $1 \le k \le 2^n, k \in \mathbf{N}$, set $I := I(k-1,n)$ and observe that $\mathrm{f}_I = \mathrm{f}_{2^n+k}$. By duality, (36) will follow if we show there is an absolute constant A_1 such that

$$\left|\int_0^1 \mathrm{f}_I\varphi\right| \le A_1|I|^{1/2}\|\varphi\|_*$$

for all $\varphi \in \mathcal{BMO}$.

Since f_I is orthogonal to $f_\emptyset$, we can write

$$\begin{aligned}\int_0^1 f_I\varphi &= \int_0^1 f_I\left(\varphi - \frac{1}{|I|}\int_I \varphi\right)\\ &= \sum_{J\in\mathcal{I},|J|=2^{-n}} \int_J f_I\left(\varphi - \frac{1}{|J|}\int_J\varphi\right) + \int_J f_I\left(\frac{1}{|J|}\int_J\varphi - \frac{1}{|I|}\int_I\varphi\right)\\ &=: U_1 + U_2.\end{aligned}$$

By (31), U_1 can be estimated by

$$\begin{aligned}|U_1| &\le C \sum_{J\in\mathcal{I},|J|=2^{-n}} |I|^{-1/2}q^{r(J,I)}|J|\,\|\varphi\|_*\\ &\le C|I|^{1/2}\|\varphi\|_* \sum_{|J|=2^{-n}} q^{r(J,I)}\\ &\le C_1|I|^{1/2}\|\varphi\|_*\end{aligned}$$

for some absolute constant $C_1 > 0$. On the other hand, if K is the smallest interval containing I and J then

$$\begin{aligned}\left|\frac{1}{|J|}\int_J\varphi - \frac{1}{|I|}\int_I\varphi\right| &\le \frac{2}{|I|}\int_K\left|\varphi - \frac{1}{|K|}\int_K\varphi\right|\\ &\le 4(r(J,I)+1)\|\varphi\|_*.\end{aligned}$$

Consequently, we have by (31) that

$$\begin{aligned}|U_2| &\le 4C|I|\,\|\varphi\|_* \sum_{J\in\mathcal{I},|J|=2^{-n}} |J|^{-1/2}r(J,I)q^{r(J,I)}\\ &\le C_2|I|^{1/2}\|\varphi\|_*\end{aligned}$$

for some absolute constant $C_2 > 0$. This completes the proof of (36).

To prove (37) fix an integer $n \ge 6$, set $J := [1/2, 1/2 + 2^{-n}), I^s = [1/2, 1/2 + 2^{-s})$ for $2 \le s \le n/2$,

$$g_n := 2^{n/2}f_{2^n+2^{n-1}+1} = |J|^{-1/2}f_J,$$

and observe that it suffices to show there is an absolute constant $B_1 > 0$ and an integer n_0 such that $n \ge n_0$ implies

$$\left|\int_{I^s} g_n\right| \ge B_1 \qquad (2 \le s \le n/2). \tag{39}$$

Indeed, if (39) were true then

$$(\mathcal{E}^*g_n)(x) \ge B_1 2^s \qquad (x \in I_s, 2 \le s \le n/2).$$

Consequently, we would have

$$\int_0^1 \mathcal{E}^* g_n \geq \sum_{s=2}^{n/2} \int_{1/2+2^{-s-1}}^{1/2+2^{-s}} \mathcal{E}^* g_n$$
$$\geq \frac{B_1}{2} \sum_{s=2}^{n/2} 1$$
$$\geq \frac{B_1}{8} n$$

for $n \geq n_0$.

To establish (39) we construct polygonals which approximate $\chi(I^s)$. Specifically, let

$$p(x) := p_s^n(x) := \begin{cases} 0 & x \notin [1/2 - 2^{-n-1}, 1/2 + 2^{-s} + 2^{-n}] \\ 1 & x \in I^s \\ \text{linear otherwise on } [0,1], \end{cases}$$

and fix $2 \leq s \leq n/2$. Notice by (21) and (24) that

$$\int_0^1 p_s^n \, d\widetilde{h}_m = 0$$

for $m > 2^n + 2^{n-1}$. Thus each function p_s^n is a linear combination of Franklin functions of order no greater than $2^n + 2^{n-1}$. Since g_n is a Franklin function of order $2^n + 2^{n-1} + 1$ we have

$$\int_0^1 p_s^n g_n = 0.$$

Consequently,

$$\int_{I_s} g_n = - \int_{1/2-2^{-n-1}}^{1/2} p g_n - \int_{1/2+2^{-s}}^{1/2+2^{-s}+2^{-n}} p g_n. \tag{40}$$

We consider terms on the right side of (40) separately. By (31), the second term is of order $q^{2^{n/2}} = o(1)$ as $n \to \infty$ since $s \leq n/2$. By linearity of g_n on $[1/2 - 2^{-(n+1)}, 1/2]$, the function pg_n is quadratic on this interval. Therefore, Simpson's rule implies

$$\left| \int_{1/2-2^{-n-1}}^{1/2} p g_n \right| = \frac{1}{3 \cdot 2^{n+2}} \left| g_n(1/2 - 2^{-n-1}) + 2 g_n(1/2) \right|$$
$$= \left(\frac{|\mathrm{f}_J(\alpha^J)|}{12 \cdot 2^{n/2}} \right) \left| 2 + \frac{\mathrm{f}_J(1/2 - 2^{-n-1})}{\mathrm{f}_J(\alpha^J)} \right|.$$

Moreover, (27) and (28) imply

$$\frac{|\mathrm{f}_J(\alpha^J)|}{2^{n/2}} \geq \frac{1}{3\sqrt{3}}$$

and

$$\frac{\mathrm{f}_J(1/2-2^{-n-1})}{\mathrm{f}_J(\alpha^J)} = -\lim_{n\to\infty}\frac{\cosh\gamma(2^n-1)}{\cosh\gamma 2^n}$$
$$= -e^{-\gamma}$$
$$= -q := \sqrt{3}-2.$$

It follows, therefore, that

$$\left|\int_{1/2-2^{-n-1}}^{1/2} g_n\right| \geq c > 0$$

for n sufficiently large where c is an absolute constant. These estimates verify (39). This completes the proof of (37).

To establish (38), observe for any function f, linear on a dyadic interval $I \in \mathcal{I}_0$, that

$$\int_0^1 f h_I = \frac{|I|^{1/2}}{4}(f(\alpha^I) - f(\beta^I)).$$

Hence (38) follows from (26) and (27) if we set

$$f := \mathrm{f}_{2^n+2^{n-1}+1}. \quad \blacksquare$$

The following observation is both useful and interesting. The facts for classical Hardy spaces do not change if we replace the collection of all subintervals of $[0,1)$ by a class only slightly larger than the dyadic ones. Indeed, let $\widetilde{\mathcal{I}}_0$ denote the collection of intervals formed by taking the union of any two adjacent dyadic intervals of the same size. For example, $[1/4, 3/4)$ belongs to $\widetilde{\mathcal{I}}_0$. Clearly, $\mathcal{I}_0 \subset \widetilde{\mathcal{I}}_0$ and given any subinterval J of $[0,1)$ there exists an $I \in \widetilde{\mathcal{I}}_0$ such that $J \subseteq I$ and $|J| \leq |I| \leq 4|J|$. From this it follows (see Exercise 3.26) that the atomic $\mathcal{H}$ norm generated by atoms supported on intervals in $\widetilde{\mathcal{I}}_0$ and the classical $\mathcal{H}$ norm are equivalent. Moreover, the norm defined by

$$\sup_{I\in\widetilde{\mathcal{I}}_0}\left(\frac{1}{|I|}\int_I\left|f-\frac{1}{|I|}\int_I f\right|^2\right)^{1/2}$$

is equivalent to the classical $\mathcal{BMO}$ norm. Therefore, the embeddings $\mathrm{H} \subset \mathcal{H}$ and $\mathcal{BMO} \subset$ BMO are continuous.

Denote the Franklin-Fourier coefficients of a function $g \in \mathrm{L}^1$ by

$$\overset{\diamond}{g}_I := \int_0^1 g\mathrm{f}_I \qquad (I \in \mathcal{I}),$$

Let $\mathcal{BMO}_0$ represent the collection of functions $g \in \mathcal{BMO}$ which satisfy $\overset{\diamond}{g}_{[0,1)} = 0$. Thus any function in $\mathcal{BMO}_0$ has both a zero constant coefficient and a zero linear coefficient.

Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be Banach spaces, let $\boldsymbol{X}_0 \subseteq \boldsymbol{X}$, and suppose $\Lambda : \boldsymbol{X} \to \boldsymbol{Y}$ is a function. If there exist absolute positive constants A and B such that

$$A\|x\|_{\boldsymbol{X}} \leq \|\Lambda x\|_{\boldsymbol{Y}} \leq B\|x\|_{\boldsymbol{X}}$$

for all $x \in \boldsymbol{X}_0$, we shall write $\|\Lambda x\|_{\boldsymbol{Y}} \sim \|x\|_{\boldsymbol{X}} \quad (x \in \boldsymbol{X}_0)$.

Clearly, if $\Lambda : \boldsymbol{X} \to \boldsymbol{Y}$ is a $1-1$, linear map which satisfies $\|\Lambda x\|_{\boldsymbol{Y}} \sim \|x\|_{\boldsymbol{X}}$ for all $x \in \boldsymbol{X}$, then Λ is a Banach space isomorphism of $\boldsymbol{X}$ onto $\Lambda(\boldsymbol{X})$. Therefore, the following result shows that $\mathcal{BMO}_0$ is isomorphic to **bmo**. Since by definition BMO is isomorphic to **bmo**, it follows that BMO and $\mathcal{BMO}_0$ are isomorphic Banach spaces.

THEOREM 12. *The map* $g \to \mathring{g}$ *is a* $1-1$, *linear map from* $\mathcal{BMO}_0$ *onto* **bmo** *which satisfies*

$$\|g\|_{\mathcal{BMO}} \sim \|\mathring{g}\|_{\mathbf{bmo}} \qquad (g \in \mathcal{BMO}_0).$$

PROOF. It is obvious that $g \to \mathring{g}$ is linear. Moreover, since **bmo** is a subspace of the sequence space $\ell^2(\mathcal{I})$, the Riesz-Fischer theorem implies that given $\boldsymbol{b} \in$ **bmo** there is a $g \in \mathrm{L}^2$ such that $\mathring{g} = \boldsymbol{b}$. Hence it suffices to show there exist absolute constants A and B such that

$$\|\mathring{g}\|_{\mathbf{bmo}} \leq A\|g\|_* \tag{41}$$

and

$$\|g\|_* \leq B\|\mathring{g}\|_{\mathbf{bmo}} \tag{42}$$

hold for all $g \in \mathcal{BMO}_0$.

The proof of these inequalities rests on three preliminary sets of estimates.

First we show there exist absolute constants A_1, A_3, A_4, A_5, and a positive function $A_2(t)$, which depends only on $t \in (0,\infty)$, such that for any nonempty interval $J \subseteq [0,1)$, dyadic or not, the following inequalities hold:

$$\sum_{|I|\geq|J|}{}' |I|^{-1} q^{r(J,I)} \leq A_1|J|^{-1}, \tag{43}$$

$$\sum_{|I|\leq|J|, I\cap J=\emptyset}{}' |I|^t q^{t\,r(J,I)} \leq A_2(t)|J|^t \qquad (t \in (0,\infty)), \tag{44}$$

$$\sum_{|I|\leq|J|, \operatorname{dist}(J,I)\geq|J|}{}' q^{r(J,I)} \leq A_3 \tag{45}$$

$$\int_J |\mathrm{f}_I|^2 \leq A_4 q^{r(J,I)} \qquad (I \in \mathcal{I}_0), \tag{46}$$

and

$$\int_{[0,1)\setminus J} |f_I| \le A_5 |I|^{1/2} (q^{r(\alpha^J, I)} + q^{r(\beta^J, I)}) \qquad (I \in \mathcal{I}_0, I \subseteq J). \tag{47}$$

Here $\sum'$ denotes summation over $I \in \mathcal{I}_0$.

To prove these inequalities fix J and choose $N \in \mathbf{N}$ such that $2^{-N} \le |J| < 2^{-N+1}$. Since

$$\begin{aligned}\sum_{|I|\ge|J|}{}' |I|^{-1} q^{r(J,I)} &\le \sum_{k=0}^{N} 2^k \sum_{|I|=2^{-k}} q^{r(J,I)} \\ &\le \frac{2}{1-q} \sum_{k=0}^{N} 2^k \\ &\le \frac{8}{1-q} |J|^{-1},\end{aligned}$$

(43) evidently holds. Similarly,

$$\begin{aligned}\sum_{|I|\le|J|, I\cap J=\emptyset}{}' |I|^t q^{t\,r(J,I)} &\le \sum_{k=N-1}^{\infty} 2^{-tk} \sum_{|I|=2^{-k}, I\cap J=\emptyset}{}' q^{r(J,I)} \\ &\le \frac{2}{1-q^t} \sum_{k=N-1}^{\infty} 2^{-tk} \\ &\le \frac{2^{1+t}}{(1-q^t)(1-2^{-t})} |J|^t,\end{aligned}$$

and

$$\begin{aligned}\sum_{|I|\le|J|, \mathrm{dist}(J,I)\ge|J|}{}' q^{r(J,I)} &\le 2A \sum_{k=N}^{\infty} \sum_{\ell=2^{k-N}}^{\infty} q^{\ell} \\ &\le \frac{2A}{1-q} \sum_{k=N}^{\infty} q^{2^{k-N}} \\ &\le \frac{2A}{(1-q)^2}\end{aligned}$$

for some constant A. Thus (44) and (45) hold.

Since (46) holds for $r(J,I) \le 1$ (with $A_4 = 1/q$), suppose that $r(J,I) \ge 1$. Then (31) implies

$$|f_I(t)|^2 \le C^2 |I|^{-1} q^{r(t,I)} q^{r(J,I)}$$

for all $t \in J$. In particular,

$$\int_J |f_I|^2 \le C^2 q^{r(J,I)} \int_1^{\infty} q^t \, dt.$$

Therefore, (46) holds for all $I \in \mathcal{I}_0$. Moreover, (31) also implies

$$|\mathrm{f}_I(t)| \le C|I|^{-1/2} q^{(t-\beta^I)/|I|}$$

for any $I \in \mathcal{I}_0, I \subseteq J$ and $\beta^J \le t \le 1$. Consequently,

$$\int_{\beta^J}^1 |\mathrm{f}_I| \le \frac{C}{\log(1/q)} |I|^{1/2} q^{r(\beta^J, I)}.$$

Since a similar estimate holds for the integral over $[0, \alpha^J)$, we conclude that (47) holds.

Next, we observe there is an absolute constant A_6 such that

$$|\mathring{g}_I| \le A_6 |I|^{1/2} \tag{48}$$

for all $I \in \mathcal{I}_0$ and all functions $g \in \mathcal{BMO}$ which satisfy either $\|g\|_* \le 1$ or $\|\mathring{g}\|_{\mathbf{bmo}} \le 1$. Indeed, (48) holds for $\|\mathring{g}\|_{\mathbf{bmo}} \le 1$ with $A_6 = 1$ by the definition of the **bmo** norm. On the other hand, if $\|g\|_* \le 1$ then Fefferman's inequality, (35), and (36) imply

$$\begin{aligned} |\mathring{g}_I| &= |\langle g, \mathrm{f}_I \rangle| \\ &\le C\|g\|_* \|\mathrm{f}_I\|_{\mathcal{H}} \\ &\le A_6 |I|^{1/2}. \end{aligned}$$

Our third set of preliminary estimates involves a decomposition for $\|g\|_{\mathcal{BMO}} \le 1$, namely,

$$\begin{aligned} g &= \sum_{|I|>|J|} \mathring{g}_I \mathrm{f}_I + \sum_{I \cap J = \emptyset, |I| \le |J|} \mathring{g}_I \mathrm{f}_I + \sum_{I \subseteq J, |I| \le |J|} \mathring{g}_I \mathrm{f}_I \\ &=: G_1 + G_2 + G_3 \end{aligned}$$

where J is a fixed interval in $\mathcal{I}_0$. To estimate the classical local modulus of continuity of G_1, use (32), (48), and (43) to obtain

$$\begin{aligned} \Omega(G_1, J) &\le \sum_{|I| \ge |J|} |\mathring{g}_I| \, \Omega(\mathrm{f}_I, J) \\ &\le C A_6 |J| \sum_{|I| \ge |J|} |I|^{-1} q^{r(J,I)} \\ &\le C A_1 A_6. \end{aligned}$$

Thus there is an absolute constant A_1' such that

$$\Omega(G_1, J) \le A_1'. \tag{49}$$

To estimate G_2, use the triangle inequality, (48), (46) and (44) with $t = 1/2$ to obtain

$$\begin{aligned}\left(\int_J |G_2|^2\right)^{1/2} &\leq \sum_{I\cap J=\emptyset,|I|\leq|J|} |\mathring{g}_I| \left(\int_J |\mathrm{f}_I|^2\right)^{1/2} \\ &\leq A_4^{1/2} A_6 \sum_{I\cap J=\emptyset,|I|\leq|J|} |I|^{1/2} q^{r(J,I)/2} \\ &\leq A_2(\frac{1}{2}) A_4^{1/2} A_6 |J|^{1/2}.\end{aligned}$$

Thus there is an absolute constant A_2' such that

$$\left(\frac{1}{|J|}\int_J |G_2|^2\right)^{1/2} \leq A_2'. \tag{50}$$

We obtain two estimates for G_3. First observe that

$$\int_0^1 \mathrm{f}_I = 0 \qquad (I \in \mathcal{I}_0)$$

and (47) imply for $I \subseteq J$ that

$$\left|\int_J \mathrm{f}_I\right| = \left|\int_{[0,1)\setminus J} \mathrm{f}_I\right| \leq A_5 |I|^{1/2} \left(q^{r(\alpha^J,I)} + q^{r(\beta^J,I)}\right).$$

Therefore, if $|J| = 2^{-N}$ it follows from (48) that

$$\begin{aligned}\left|\int_J G_3\right| &\leq \sum_{I\subseteq J} |\mathring{g}_I| \left|\int_J \mathrm{f}_I\right| \\ &\leq A_5 A_6 \sum_{I\subseteq J} |I| \left(q^{r(\alpha^J,I)} + q^{r(\beta^J,I)}\right) \\ &\leq A_5 A_6 \sum_{k=N}^{\infty} 2^{-k} \sum_{|I|=2^{-k}} \left(q^{r(\alpha^J,I)} + q^{r(\beta^J,I)}\right) \\ &\leq \frac{2A_5A_6}{1-q} \sum_{k=N}^{\infty} 2^{-k} \\ &= \frac{4A_5A_6}{1-q} |J|.\end{aligned}$$

Thus there is an absolute constant A_3' such that

$$\frac{1}{|J|}\left|\int_J G_3\right| \leq A_3'. \tag{51}$$

For the second estimate of G_3, use Parseval's identity and the triangle inequality to verify

$$\left(\sum_{I\subseteq J}|\mathring{g}_I|^2 - \int_J |G_3|^2\right)^{1/2} = \left(\int_{[0,1)\setminus J}|G_3|^2\right)^{1/2}$$
$$\le \sum_{I\subseteq J}|\mathring{g}_I|\left(\int_{[0,1)\setminus J}|f_I|^2\right)^{1/2}.$$

Moreover, observe by applying (46) to the components $[0,1)\setminus J =: J' \cup J''$ that

$$\left(\int_{[0,1)\setminus J}|f_I|^2\right)^{1/2} \le A_4^{1/2}\left(q^{r(J',I)/2} + q^{r(J'',I)/2}\right).$$

It follows, therefore, from (48) that

$$\sum_{I\subseteq J}|\mathring{g}_I|\left(\int_{[0,1)\setminus J}|f_I|^2\right)^{1/2} \le A_4^{1/2}A_6\sum_{I\subseteq J}|I|^{1/2}\left(q^{r(J',I)/2} + q^{r(J'',I)/2}\right)$$
$$\le A_4^{1/2}A_6\sum_{k=N}^{\infty}2^{-k/2}\sum_{I\subseteq J,|I|=2^{-k}}\left(q^{r(J',I)/2} + q^{r(J'',I)/2}\right)$$
$$\le \frac{2A_4^{1/2}A_6}{1-\sqrt{q}}\sum_{k=N}^{\infty}2^{-k/2}$$
$$< \frac{8A_4^{1/2}A_6}{1-\sqrt{q}}|J|^{1/2}.$$

Thus there is an absolute constant A_4' such that

(52) $$0 \le -\left(\frac{1}{|J|}\int_J |G_3|^2\right) + \sum_{I\subseteq J}|\mathring{g}_I|^2 \le A_4'.$$

This is the last of our third set of preliminary estimates.

By definition of the **bmo** norm, to establish (41) it suffices to show there is an absolute constant such that

(53) $$\sum_{I\subseteq J}|\mathring{g}_I|^2 \le A|J|$$

for every nonempty $J \in \mathcal{I}$ and every $\|g\|_* \le 1$. But (51) and (52) imply

$$\sum_{I\subseteq J}|\mathring{g}_I|^2 \le A_4'|J| + \int_J |G_3|^2$$
$$\le 2\int_J\left|G_3 - \frac{1}{|J|}\int_J G_3\right|^2 + A_5'|J|.$$

Moreover, since G_1 is continuous, (49) and (50) imply

$$\begin{aligned}\int_J\left|G_3-\frac{1}{|J|}\int_J G_3\right|^2 &\le 3\int_J\left|g-\frac{1}{|J|}\int_J g\right|^2+3\int_J\left|G_2-\frac{1}{|J|}\int_J G_2\right|^2\\ &\quad+3\int_J\left|G_1-\frac{1}{|J|}\int_J G_1\right|^2\\ &\le 3|J|+3\int_J|G_2|^2+3|J|(\Omega(G_1,J))^2\\ &\le A_6'|J|.\end{aligned}$$

Here A_5' and A_6' are absolute positive constants. Consequently, (53) holds and the proof of (41) is complete.

A similar argument can be used to establish (42). Indeed, it suffices to show the existence of an absolute constant $B>0$ such that

$$\int_J\left|g-\frac{1}{|J|}\int_J g\right|^2\le B|J| \tag{54}$$

for any interval $J\in\widetilde{\mathcal{I}}$ and any g satisfying $\|\overset{\diamond}{g}\|_{\mathbf{bmo}}\le 1$. Fix such an interval J and let J^* represent the interval twice as large as J which is centered where J is, or (in case that an endpoint of J is 0 or 1) let J^* represent the interval in $\widetilde{\mathcal{I}}$ twice as large as J which contains J. In either case, J^* is a union of at most four dyadic intervals. Also, if $J^*\cap I=\emptyset$ for some $I\in\mathcal{I}_0$, then

$$\operatorname{dist}(J^i,I)\ge|J^i|\qquad(i=1,2)$$

where $J=J^1\cup J^2$ and $J^i\in\mathcal{I}$ for $i=1,2$.

This time use the decomposition

$$\begin{aligned}g&=\sum_{|I|>|J^i|}\overset{\diamond}{g}_I\mathrm{f}_I+\sum_{I\cap J^*=\emptyset,|I|\le|J^i|}\overset{\diamond}{g}_I\mathrm{f}_I+\sum_{I\subseteq J^*,|I|\le|J^i|}\overset{\diamond}{g}_I\mathrm{f}_I\\ &=:H_1+H_2+H_3.\end{aligned}$$

The method used to establish (49) shows that $\Omega(H_1,J^i)\le B_1'$ for $i=1$ or 2 and some absolute constant $B_1'>0$. By (48) and (31) we have for $t\in J^i(i=1$ or $2)$ that

$$|H_2(t)|\le 2CA_6\sum_{|I|\le|J^i|,\operatorname{dist}(I,J^i)\ge|J^i|}q^{r(J,I)}\le B_2'$$

where B_2' is an absolute constant. And, Parseval's identity implies

$$\int_{J^i}|H_3|^2\le\int_0^1|H_3|^2\le\sum_{|I|\le|J^i|,I\subseteq J^*}|\overset{\diamond}{g}_I|^2.$$

Since J^* is a union of at most four dyadic intervals of length $|J^i|$, it follows from the assumption $\|\mathring{g}\|_{\mathbf{bmo}} \le 1$ that

$$\int_{J^i} |H_3|^2 \le 4|J^i| \qquad (i = 1 \quad \text{or} \quad 2).$$

Let c denote the value of H_1 at the midpoint of J. The estimates above together with (49) show

$$\begin{aligned}\int_{J^i} |g-c|^2 &\le 3\left(\int_{J^i} |H_1 - c|^2 + \int_{J^i} |H_2|^2 + \int_{J^i} |H_3|^2\right)\\ &\le B_3|J^i|\end{aligned}$$

for some absolute constant B_3. In particular,

$$\int_J \left| g - \frac{1}{|J|}\int_J g\right|^2 \le \int_J |g-c|^2 \le 2B_3|J|. \qquad \blacksquare$$

COROLLARY 3. *There exist absolute positive constants A, B such that*

$$A\|\sum_{k=1}^{N} c_k h_k\|_{\mathrm{H}} \le \|\sum_{k=1}^{N} c_k \mathrm{f}_{k+1}\|_{\mathcal{H}} \le B\|\sum_{k=1}^{N} c_k h_k\|_{\mathrm{H}}$$

hold for all $\mathbf{c} = (c_k, k \in \mathrm{N}) \in \ell^0$ *and* $N \in \mathrm{N}$.

PROOF. By definition, the map $f \to \dot{f}$ is an isomorphism of H onto **h**. By Theorem 12 the map

$$\Lambda(\boldsymbol{b}) := \sum_{k=1}^{\infty} b_k \mathrm{f}_{k+1}$$

is an isomorphism of **bmo** onto $\mathcal{BMO}_0$.

Let $\mathcal{H}_{00}$ represent those $g \in \mathcal{H}_0$ which satisfy $\mathring{g}_{[0,1)} = 0$. Since $\mathcal{H}_0{}' = \mathcal{BMO}$ it is easy to see that $\mathcal{H}_{00}{}' = \mathcal{BMO}_0$. It follows by duality that Λ is an isomorphism of **h** onto $\mathcal{H}_{00}$. Therefore, the map

$$g \mapsto \sum_{k=1}^{\infty} \langle g, h_k\rangle\, \mathrm{f}_{k+1}$$

is an isomorphism from H onto $\mathcal{H}_{00}$, i.e.,

$$\|g\|_{\mathrm{H}} \sim \|\sum_{k=1}^{\infty} \langle g, h_k\rangle \mathrm{f}_{k+1}\|_{\mathcal{H}}$$

for every $g \in \mathrm{H}_0$. $\blacksquare$

It is now easy to identify subspaces of L^1 in which the Franklin system is a basis.

By Theorem 11 the Franklin system is closed in $\mathcal{C}[0,1]$. Since $\|f\|_* \leq 2\|f\|_\infty$ for any bounded function f and $\mathcal{VMO}$ is the closure, in $\mathcal{BMO}$ norm, of $\{f \in \mathcal{C}[0,1] : \widehat{f}(0) = 0\}$ it follows that $\mathcal{VMO}$ is contained in the closed linear hull of $\mathbf{f}$. But Theorem 12 implies there are absolute constants A, B such that

$$\begin{aligned}\|S_n^{\mathrm{f}} g\|_* &\leq B\|(S_n^{\mathrm{f}} g)^\diamond\|_{\mathbf{bmo}} \\ &\leq B\|\overset{\diamond}{g}\|_{\mathbf{bmo}} \\ &\leq AB\|g\|_*\end{aligned}$$

for $n \in \mathbf{N}$ and $g \in \mathcal{BMO}_0$. Consequently, the operators $S_n^{\mathrm{f}} : \mathcal{BMO} \to \mathcal{BMO}$ are uniformly bounded in n. In particular, $(\mathrm{f}_n, n \in \mathbf{P})$ is a basis in $\mathcal{VMO}$.

By duality, $(\mathrm{f}_n, n \in \mathbf{N})$ is a basis in $\mathcal{H}$. Moreover, since (30) implies the Lebesgue functions associated with $\mathbf{f}$ are uniformly bounded, it is clear that the Franklin system is a basis in $\mathcal{C}[0,1]$ and in L^p for each $1 \leq p < \infty$.

Finally the Franklin system is not a basis in H. Indeed, by definition one verifies

$$2^{n/2}\|h_{2^n+k}\|_{\mathrm{H}} = 1$$

for $0 \leq k < 2^n, n \in \mathbf{N}$. Thus it follows from Lemma 4 that

$$\begin{aligned}\|(S_{m_n+2}^{\mathrm{f}} - S_{m_n+1}^{\mathrm{f}})(2^{(n+1)/2}h_{2m_n})\|_{\mathrm{H}} &= |\langle \mathrm{f}_{m_n+1}, h_{2m_n}\rangle|\,\|2^{(n+1)/2}\mathrm{f}_{m_n+1}\|_{\mathrm{H}} \\ &\geq \sqrt{2}BCn\end{aligned}$$

for $n \geq n_0$, where $m_n := 2^n + 2^{n-1}$. Consequently, the operators $S_m^{\mathrm{f}} : \mathrm{H} \to \mathrm{H}$ are not uniformly bounded in m.

We close this section with an application of Theorem 12 to *shift operators*. These operators will be used in the following section to study equivalence of bases.

For convenience extend indices of the Haar, Walsh, and Franklin systems by

$$h_{-1} := w_{-1} := \mathrm{f}_{-1} := 0.$$

For any $f \in \mathrm{L}^2$ define the Walsh shift operators $W^{(+)}$ and $W^{(-)}$ by

$$W^{(\pm)}f := \sum_{k=0}^{\infty} \widehat{f}(k) w_{k\pm 1}.$$

Similarly, define Haar shift operators by

$$H^{(\pm)}f := \sum_{k=0}^{\infty} \langle f, h_k\rangle\, h_{k\pm 1},$$

and Franklin shift operators by

$$F^{(\pm)}f := \sum_{k=0}^{\infty} \langle f, \mathrm{f}_k\rangle\, \mathrm{f}_{k\pm 1},$$

for $f \in \mathrm{L}^2$. Let $\boldsymbol{X} \subseteq \mathrm{L}^1$ be a Banach space which contains the Haar, Walsh, and Franklin polynomials, and suppose T is any one of these six shift operators. We shall say T is continuous from $\boldsymbol{X}$ to $\boldsymbol{X}$ if T has a continuous extension from the appropriate polynomials (which are themselves contained in L^2) to $\boldsymbol{X}$ such that $T(\boldsymbol{X})$ is a subset of $\boldsymbol{X}$.

THEOREM 13

i) For each $1 < p < \infty$ *the Walsh shift operators* $W^{(+)}$ *and* $W^{(-)}$ *are continuous from* L^p *to* L^p.

ii) The Haar shift operators $H^{(+)}$ *and* $H^{(-)}$ *are continuous from* H *to* H.

iii) The Franklin shift operators $F^{(+)}$ *and* $F^{(-)}$ *are continuous from* $\mathcal{H}$ *to* $\mathcal{H}$.

PROOF. By symmetry we consider only the positive shifts.

To prove i), fix $1 < p < \infty$. We shall prove

$$\|f\|_p \sim \|W^{(+)}f\|_p \tag{55}$$

for all Walsh polynomials f. Since the Walsh system is a basis for L^p, it will follow that $W^{(+)}$ has a continuous extension to L^p and that (55) holds for all $f \in L^p$.

Fix $n \in \mathbf{N}$ with binary expansion

$$n = \sum_{k=0}^{\infty} n_k 2^k$$

and choose $i \in \mathbf{N}$ such that $n_i = 0$ and $n_j = 1$ for $0 \leq j < i$. Thus

$$n + 1 = 2^i + \sum_{k=i+1}^{\infty} n_k 2^k$$

and the definition of the Walsh system implies

$$W^{(+)}w_n = r_i \prod_{k=i+1}^{\infty} r_k^{n_k} = w_n \prod_{j=0}^{i} r_j. \tag{56}$$

For each $N \in \mathbf{N}$ set $E_N := \{n \in \mathbf{N} : 0 \leq n < 2^{N+1}\}$, and for each $0 \leq k \leq N+1$ set

$$E_{Nk} := \{n \in E_N : n_0 = \cdots = n_{k-1} = 1, n_k = 0\}.$$

Then $\{E_{Nk} : k = 0, \ldots, N+1\}$ is a collection of pairwise disjoint sets which satisfy

$$E_N = \bigcup_{k=0}^{N+1} E_{Nk}.$$

Hence every Walsh polynomial

$$P = \sum_{i=0}^{2^N - 1} c_i w_i$$

can be written in the form

$$P = \sum_{k=0}^{N+1} \sum_{i \in E_{Nk}} c_i w_i =: \sum_{k=0}^{N+1} P_k.$$

Moreover,

$$W^{(+)}P = \sum_{k=0}^{N+1} W^{(+)}P_k. \tag{57}$$

For each $n \in E_N$ set

$$\grave{w}_n := \prod_{i=0}^{N} r_{N-i}^{n_i}.$$

Then $(\grave{w}_n, n \in \mathbf{N})$ is a linear rearrangement of the Walsh system. Notice, if $n \in E_{Nk}$ has binary coefficients $(n_j, j \in \mathbf{N})$ and

$$n^*(k) := 2^{N-k} + \sum_{j=0}^{N-k-1} n_{N-j}2^j,$$

then (56) implies

$$\begin{aligned} W^{(+)}w_n &= w_n \prod_{i=0}^{k} r_i \\ &= r_k \prod_{i=k+1}^{N} r_i^{n_i} \\ &= r_{N-(N-k)} \prod_{j=0}^{N-k-1} r_{N-j}^{n_{N-j}} \\ &= \grave{w}_{n^*(k)}. \end{aligned}$$

In particular, we have by (57) that

$$\begin{aligned} W^{(+)}P &= \sum_{k=0}^{N} W^{(+)}P_{N-k} \\ &= \sum_{k=0}^{N} \sum_{j=0}^{2^k-1} \grave{c}_{2^k+j}\grave{w}_{2^k+j} \end{aligned}$$

for a suitable choice of coefficients $\grave{c}_i$. Since by Theorem 7 in 1.4 there is a measure preserving transformation $T' : [0,1) \to [0,1)$ such that $\grave{w}_n = w_n \circ T'$ we obtain

$$\|W^{(+)}P\|_p \sim \|(\sum_{k=0}^{N} |W^{(+)}P_k|^2)^{1/2}\|_p$$

by applying Paley's inequality (Corollary 5 in 3.3) to the system $(\grave{w}_n, n \in \mathbf{N})$. This can be written as

$$\|W^{(+)}P\|_p \sim \|(\sum_{k=0}^{N} |P_k|^2)^{1/2}\|_p \tag{58}$$

because (56) implies $|W^{(+)}P_k| = |P_k|$.

Set

$$\grave{P}_k := P_k \prod_{i=0}^{N} r_i$$

for $k = 0, 1, \ldots, N$ and observe by (56) that each $\grave{P}_k$ belongs to the linear space generated by $\grave{w}_n$ for $2^{N-k} \le n < 2^{N-k+1}$. Thus (58) and Paley's inequality imply

$$\begin{aligned}
\|W^{(+)}P\|_p &\sim \|(\sum_{k=0}^{N} |P_k|^2)^{1/2}\|_p \\
&= \|(\sum_{k=0}^{N} |\grave{P}_k|^2)^{1/2}\|_p \\
&\sim \|\sum_{k=0}^{N} \grave{P}_k\|_p \\
&= \|\sum_{k=0}^{N} P_k\|_p \\
&= \|P\|_p.
\end{aligned}$$

This completes the proof of i).

To establish ii), identify $[0,1)$ with the circle group $\mathbf{T} := \{e^{2\pi i t} : t \in [0,1)\}$ and for any dyadic interval $I \subseteq [0,1)$, let I^* denote the "arc" centered where I is, but three times larger. It is clear for any dyadic atom $\beta \neq 1$ supported on I, that $\langle \beta, h_J \rangle = 0$ for all $J \in \mathcal{I}_0$ which satisfy either $|J| > |I|$ or $J \bigcap I = \emptyset$. Consequently, off I^* we have

$$\sup_{n \in \mathbf{N}} \left| \sum_{k=0}^{n} \langle \beta, h_k \rangle \, h_{k\pm1} \right| = 0.$$

Set

$$Tf := \sup_{n \in \mathbf{N}} \left| \sum_{k=0}^{n} \langle f, h_k \rangle \, h_{k\pm1} \right| \qquad (f \in \mathrm{L}^1).$$

Notice for each $n \in \mathbf{P}$ that the Haar functions $\{h_j : 2^n \le j < 2^{n+1}\}$ have non-overlapping supports. By repeating the argument of the proof of Corollary 1 in 3.1 we see that T is of type $(2,2)$. Hence Hölder's inequality implies

$$\begin{aligned}
\int_{I^*} |T\beta| &\le |I^*|^{1/2} \left(\int_{I^*} |T\beta|^2 \right)^{1/2} \\
&\le C|I|^{1/2} \left(\int_{I} |\beta|^2 \right)^{1/2} \\
&\le C \frac{|I|^{1/2}|I|^{1/2}}{|I|} \\
&= C
\end{aligned}$$

for any atom β supported on I. Let $f \in \mathrm{H}$ have atomic decomposition $\sum a_n\beta_n$ where each β_n is supported on a dyadic interval I_n. Then the identity above implies

$$\begin{aligned}\|Tf\|_1 &\leq \sum_{n=0}^{\infty} |a_n|\,\|T\beta_n\|_1 \\ &\leq \sum_{n=0}^{\infty} |a_n| \int_{I_n^*} |T\beta_n| \\ &\leq C\sum_{n=0}^{\infty} |a_n|.\end{aligned}$$

In particular, it follows from (42) in 1.4 and Theorem 6 in 3.4 that

$$\begin{aligned}\|H^{(\pm)}f\|_{\mathrm{H}} &= \|\mathcal{E}^*(H^{(\pm)}f)\|_1 \\ &\leq \|Tf\|_1 \\ &\leq 25C\|f\|_{\mathrm{H}}.\end{aligned}$$

This completes the proof of ii).

To establish iii) let f be any Franklin polynomial with neither constant nor linear term. Thus

$$f = \sum_{k=2}^{N} c_k \mathrm{f}_k$$

for some $N \in \mathrm{N}$ with $N \geq 2$. By Corollary 3 there exists a constant B such that

$$\begin{aligned}\|F^{(+)}f\|_{\mathcal{H}} &= \|\sum_{k=2}^{N} c_k \mathrm{f}_{k+1}\|_{\mathcal{H}} \\ &\leq B\|\sum_{k=2}^{N} c_k h_k\|_{\mathrm{H}} \\ &= B\|\sum_{k=1}^{N-1} c_{k+1} h_{k+1}\|_{\mathrm{H}} \\ &= B\|H^{(+)}\left(\sum_{k=1}^{N-1} c_{k+1} h_k\right)\|_{\mathrm{H}}.\end{aligned}$$

Hence by ii), there exists a constant B' such that

$$\|F^{(+)}f\|_{\mathcal{H}} \leq B'\|\sum_{k=1}^{N-1} c_{k+1} h_k\|_{\mathrm{H}}.$$

In particular, another application of Corollary 3 results in

$$\|F^{(+)}f\|_{\mathcal{H}} \leq B''\|f\|_{\mathcal{H}}$$

for some constant B'' and for all Franklin polynomials f with zero constant and linear coefficients. None of these constants depend on f. Since $\mathbf{f}$ is a basis in $\mathcal{H}$, we conclude that $F^{(+)}$ is continuous from $\mathcal{H}$ to $\mathcal{H}$. ∎

By duality, the Haar shift operators are continuous from BMO to BMO. Hence by Theorem 13 ii) and interpolation (see Theorem 12 in 3.6) we have that the Haar shift operators are continuous from L^p to L^p for $1 < p < \infty$.

In the classical case, one can interpolate between $\mathcal{H}$ and $\mathcal{BMO}$ (see Fefferman, Riviere, and Sagher [1]). Hence Theorem 13 iii) and duality establish that the Franklin shift operators are also continuous from L^p to L^p for $1 < p < \infty$. This statement is false for $p = 1$ (see Gevorkian [1]).

5.5 Equivalence of Bases. Let $\boldsymbol{X}$,$\boldsymbol{Y}$ be Banach spaces. Two linearly independent systems $\boldsymbol{\epsilon} = (\epsilon_n, n \in \mathrm{N})$ in $\boldsymbol{X}$ and $\boldsymbol{\tau} = (\tau_n, n \in \mathrm{N})$ in $\boldsymbol{Y}$ are said to be *equivalent* (in $(\boldsymbol{X},\boldsymbol{Y})$) if there is a constant $K > 0$ such that

$$K^{-1}\|\sum_{n=0}^{\infty} b_n\epsilon_n\|_{\boldsymbol{X}} \le \|\sum_{n=0}^{\infty} b_n\tau_n\|_{\boldsymbol{Y}} \le K\|\sum_{n=0}^{\infty} b_n\epsilon_n\|_{\boldsymbol{X}}$$

for all $\sum_{n=0}^{\infty} b_n\epsilon_n$ in the linear hull $\mathrm{L}(\boldsymbol{\epsilon})$. When $\boldsymbol{\epsilon}$ is equivalent to $\boldsymbol{\tau}$ in $(\boldsymbol{X},\boldsymbol{X})$, we shall say that $\boldsymbol{\epsilon}$ and $\boldsymbol{\tau}$ are equivalent in $\boldsymbol{X}$. When no ambiguity arises we shall use the notation $\boldsymbol{\epsilon} \sim \boldsymbol{\tau}$ for equivalence of systems.

By the Riesz-Fischer theorem, all complete orthonormal systems in L^2 are equivalent in L^2. In this section we prove that the Haar and Franklin systems, and Walsh and Ciesielski systems are equivalent in many other spaces, but that the Walsh and trigonometric systems are equivalent only in L^2.

Some authors call two bases $\boldsymbol{\epsilon}$ and $\boldsymbol{\tau}$ equivalent if the series

$$\sum_{n=0}^{\infty} b_n\epsilon_n, \quad \sum_{n=0}^{\infty} b_n\tau_n$$

are equiconvergent for every $(b_n, n \in \mathrm{N}) \in \ell^0$. By the Banach-Steinhaus theorem our definition is equivalent to this one when $\boldsymbol{\epsilon}$ and $\boldsymbol{\tau}$ are bases. In this last situation it is also clear that $\boldsymbol{\epsilon} \sim \boldsymbol{\tau}$ if and only if the sequence spaces

$$\{(\epsilon_n'(x), n \in \mathrm{N}) : x \in \boldsymbol{X}\}, \ \{(\tau_n'(y), n \in \mathrm{N}) : y \in \ Y\}$$

coincide.

Let $\boldsymbol{\epsilon} = (\epsilon_n, n \in \mathrm{N})$ be linearly independent in $\boldsymbol{X}$ and $\boldsymbol{\tau} = (\tau_n, n \in \mathrm{N})$ be linearly independent in $\boldsymbol{Y}$. The *canonical isomorphism* induced by the pair $\boldsymbol{\epsilon}$,$\boldsymbol{\tau}$ is the linear map $T : \mathrm{L}(\boldsymbol{\epsilon}) \to \mathrm{L}(\boldsymbol{\tau})$ defined by

$$T(\epsilon_n) := \tau_n \qquad (n \in \mathrm{N}).$$

By definition, $\boldsymbol{\epsilon} \sim \boldsymbol{\tau}$ if and only if T and its inverse are continuous on $\mathrm{L}(\boldsymbol{\epsilon})$ and $\mathrm{L}(\boldsymbol{\tau})$, respectively. In particular, if $\boldsymbol{\epsilon}$ and $\boldsymbol{\tau}$ are bases in $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively, and if $\boldsymbol{\epsilon} \sim \boldsymbol{\tau}$, then T extends to a continuous linear operator from $\boldsymbol{X}$ onto $\boldsymbol{Y}$.

A duality principle used several times in this section is the following one. If $\boldsymbol{\epsilon}$ and $\boldsymbol{\tau}$ are bases in $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively, and if $\boldsymbol{\epsilon} \sim \boldsymbol{\tau}$, then the coordinate functional systems $\boldsymbol{\epsilon}'$ and $\boldsymbol{\tau}'$ are equivalent in $(\boldsymbol{X}',\boldsymbol{Y}')$. To verify this, let T denote the canonical isomorphism induced by $\boldsymbol{\epsilon},\boldsymbol{\tau}$ and define its adjoint by

$$T'y' := \sum_{n\in\mathrm{N}} \langle \tau_n, y'\rangle\, \epsilon_n' \qquad (y' \in \mathrm{L}(\boldsymbol{\tau}')).$$

Then $T' : \mathrm{L}(\boldsymbol{\tau}') \to \mathrm{L}(\boldsymbol{\epsilon}')$ is linear and by definition,

$$\langle x, T'y'\rangle = \langle Tx, y'\rangle \qquad (x \in \mathrm{L}(\boldsymbol{\epsilon}), y' \in \mathrm{L}(\boldsymbol{\tau}')).$$

Consequently, $\|T\| = \|T'\|$ and T' has a continuous linear extension to $\mathrm{CL}(\boldsymbol{\tau}')$, the closure of $\mathrm{L}(\boldsymbol{\tau}')$ in $\boldsymbol{Y}'$. By symmetry, $(T^{-1})'$ has a continuous linear extension to $\mathrm{CL}(\boldsymbol{\epsilon}')$. Therefore, $\boldsymbol{\epsilon}'$ is equivalent to $\boldsymbol{\tau}'$ in $(\boldsymbol{X}',\boldsymbol{Y}')$.

Let $\boldsymbol{\epsilon} = (\epsilon_n, n \in \mathrm{N})$ be any system of non-zero elements in $\boldsymbol{X}$. We shall say that $\boldsymbol{\epsilon}$ has the B-property if there is a constant $B > 0$ depending only on $\boldsymbol{\epsilon}$ such that

$$\|\sum_{k=0}^{n} b_k\epsilon_k\|_{\boldsymbol{X}} \le B\|\sum_{k=0}^{\infty} b_k\epsilon_k\|_{\boldsymbol{X}}$$

for all $n \in \mathrm{N}$ and $\sum_{k=0}^{\infty} b_k\epsilon_k \in \mathrm{L}(\boldsymbol{\epsilon})$.

LEMMA 5. *Let $\boldsymbol{\epsilon} = (\epsilon_n, n \in \mathrm{N})$ be a system of non-zero elements in a Banach space $\boldsymbol{X}$. Then $\boldsymbol{\epsilon}$ has the B-property if and only if $\boldsymbol{\epsilon}$ is a basis in $\mathrm{CL}(\boldsymbol{\epsilon})$.*

PROOF. Every basis has the B-property. On the other hand, if $\boldsymbol{\epsilon}$ has the B-property then

$$\|\sum_{k=m}^{n} b_k\epsilon_k\|_{\boldsymbol{X}} \le 2B\|\sum_{k=0}^{\infty} b_k\epsilon_k\|_{\boldsymbol{X}}$$

for $0 \le m \le n$ and $\sum_{k=0}^{\infty} b_k\epsilon_k \in \mathrm{L}(\boldsymbol{\epsilon})$. Consequently, $\boldsymbol{\epsilon}$ is linearly independent and coordinate functionals

$$\boldsymbol{\epsilon}'(x) := b_n \qquad (n \in \mathrm{N})$$

can be defined for each $x = \sum_{k=0}^{\infty} b_k\epsilon_k \in \mathrm{L}(\boldsymbol{\epsilon})$. These functionals are bounded since

$$|\epsilon_n'(x)| = \frac{\|b_n\epsilon_n\|_{\boldsymbol{X}}}{\|\epsilon_n\|_{\boldsymbol{X}}} \le 2B\frac{\|x\|_{\boldsymbol{X}}}{\|\epsilon_n\|_{\boldsymbol{X}}}$$

for $n \in \mathrm{N}$ and $x = \sum_{k=0}^{\infty} b_k\epsilon_k \in \mathrm{L}(\boldsymbol{\epsilon})$. Consequently, partial sums of the biorthogonal expansion

$$S_n^{\boldsymbol{\epsilon}}x := \sum_{k=0}^{n-1} \epsilon_k'(x)\epsilon_k \qquad (n \in \mathrm{P})$$

are uniformly bounded on $\mathrm{L}(\boldsymbol{\epsilon})$. They can be extended to $\mathrm{CL}(\boldsymbol{\epsilon})$ with the same uniform bound. Hence by the Banach-Steinhaus theorem, $\|S_n^{\boldsymbol{\epsilon}}x - x\|_{\boldsymbol{X}} \to 0$ as $n \to \infty$ for all $x \in \mathrm{CL}(\boldsymbol{\epsilon})$. We conclude that $\boldsymbol{\epsilon}$ is a basis in $\mathrm{CL}(\boldsymbol{\epsilon})$. ∎

In particular, if two systems are equivalent, one of which is a basis, then both systems are bases (of the closed linear spaces they generate).

We apply these general observations first to the Haar and Franklin systems.

THEOREM 14. *The Haar and Franklin systems are equivalent in* $(\mathrm{H}, \mathcal{H})$, *in* $(\mathrm{VMO}, \mathcal{VMO})$, *and in* L^p *for* $1 < p < \infty$. *They are not equivalent in* L^1 *or in* L^∞.

PROOF. By definition, the system $\boldsymbol{\delta} := ((\delta_{IJ}, J \in \mathcal{I}_0), I \in \mathcal{I}_0)$ is a basis in **vmo** and is equivalent, in (**vmo**, VMO), to the shifted Haar system

$$\boldsymbol{h}(+) := (h_{n+1}, n \in \mathbf{N}).$$

By Theorem 12, $\boldsymbol{\delta}$ is equivalent to $\mathbf{f}(++) := (\mathrm{f}_{n+2}, n \in \mathbf{N})$ in $(\mathbf{vmo}, \mathcal{VMO}_0)$. Thus $\boldsymbol{h}(+)$ is equivalent to $\mathbf{f}(++)$ in $(\mathrm{VMO}, \mathcal{VMO}_0)$, and by the duality principle, in $(\mathrm{H}_0, \mathcal{H}_{00})$. On the other hand, we have by Theorem 13 in 5.4 that Haar shifts are continuous on H and Franklin shifts are continuous on $\mathcal{H}$. Consequently, the full systems $\boldsymbol{h}$ and $\mathbf{f}$ are equivalent in $(\mathrm{H}, \mathcal{H})$.

Let T be the canonical isomorphism induced by the pair $\boldsymbol{h}, \mathbf{f}$. Then T is continuous from H to $\mathcal{H}$. Moreover, by the Riesz-Fischer theorem T is continuous from L^2 to L^2. Hence by interpolation (see the remarks following Theorem 12 in 3.6), T is continuous from L^p to L^p for $1 < p \leq 2$. Similarly, the map T^{-1} is continuous also from L^p to L^p for $1 < p \leq 2$. Consequently, $\boldsymbol{h}$ and $\mathbf{f}$ are equivalent in L^p for $1 < p \leq 2$.

By the duality principle it remains to show that the Haar and Franklin systems are not equivalent in L^∞. Suppose to the contrary that $\boldsymbol{h}$ and $\mathbf{f}$ are equivalent in L^∞. Then the operators

$$T_n f := \sum_{k=0}^{2^n-1} \langle f, h_k \rangle \mathrm{f}_k \qquad (f \in \mathrm{L}^\infty, n \in \mathbf{N})$$

are uniformly bounded. However, if for a fixed $n \in \mathbf{P}$ we set

$$P := \sum_{k=0}^{2^n-1} \mathrm{f}_k(0) h_k$$

and $f := \operatorname{sgn} P$, then it turns out that

$$T_n f(0) \geq \frac{1}{3\sqrt{3}}(n-1).$$

Indeed, for each $k \in \mathbf{P}$ let $I_k := [0, 2^{-k}), I'_k := [2^{-k}, 2^{-k+1})$ and recall that

$$\|\mathring{g}_I h_I\|_1 \leq \int_I |g| \qquad (I \in \mathcal{I}_0, g \in \mathrm{L}^1).$$

Consequently,

$$\begin{aligned}
T_n f(0) &= \int_0^1 f(t) \left(\sum_{k=0}^{2^n-1} \mathrm{f}_k(0) h_k(t) \right) dt \\
&= \int_0^1 |P| \\
&\geq \sum_{k=1}^{n-1} \int_{I'_k} |P| \\
&\geq \sum_{k=1}^{n-1} \|\dot{P}_{I'_k} h_{I'_k}\|_1 \\
&\geq \sum_{k=1}^{n-1} 2^{-k/2} |\mathrm{f}_{I_k}(0)|.
\end{aligned}$$

In particular, the promised inequality follows from (27) and (28) in 5.4. ∎

We see, then, that the canonical isomorphism induced by the pair $\mathbf{f}, \boldsymbol{h}$ gives an explicit isomorphism between classical and dyadic structures.

COROLLARY 4. *Classical $\mathcal{H}$ is isomorphic to* H *and $\mathcal{VMO}$ is isomorphic to* VMO.

Next, we consider equivalence of the Walsh and Ciesielski systems.

The *Ciesielski system* $\mathbf{g} = (\mathrm{g}_n, n \in \mathbf{N})$ is defined by using Hadamard-Paley matrices and the Franklin system in the following way. Let $\mathrm{g}_0 := \mathrm{f}_0, \mathrm{g}_1 := \mathrm{f}_1$, and for $n \in \mathbf{N}, 0 \leq k < 2^n$, let

$$\mathrm{g}_{2^n+k+1} := \sum_{j=0}^{2^n-1} a_{kj}^{(n)} \mathrm{f}_{2^n+j+1}.$$

Hence the shifted Ciesielski system $\mathbf{g}(+) := (\mathrm{g}_{n+1}, n \in \mathbf{N})$ and the shifted Franklin system $\mathbf{f}(+) := (\mathrm{f}_{n+1}, n \in \mathbf{N})$ are Hadamard transforms of each other.

THEOREM 15. *The Ciesielski system is a basis in* L^p *which is equivalent to the Walsh system for* $1 < p < \infty$.

PROOF. Fix $1 < p < \infty$ and let L_0^p represent those $f \in \mathrm{L}^p$ of mean zero. It suffices to show the shifted Ciesielski system $\mathbf{g}(+)$ is a basis in L_0^p and is equivalent to $\boldsymbol{w}(+) := (w_{n+1}, n \in \mathbf{N})$ (see Theorem 13 above).

For each $m \in \mathbf{N}, x, t \in [0,1]$, and $f \in \mathrm{L}^1$, set

$$K_m(x,t) := \sum_{k=2^m}^{2^{m+1}-1} \mathrm{f}_{k+1}(x) h_k(t),$$

$$(\boldsymbol{K}_m f)(x) := \int_0^1 K_m(x,t) f(t) dt,$$

and

$$(\boldsymbol{K}'_m f)(t) := \int_0^1 K_m(x,t) f(x) dx.$$

Observe by orthogonality that

$$\boldsymbol{K}_m h_k = \mathrm{f}_{k+1}, \quad \boldsymbol{K}'_m \mathrm{f}_{k+1} = h_k \tag{59}$$

for $2^m \le k < 2^{m+1}, m \in \mathrm{N}$.

Since the supports of the Haar functions $h_k, 2^m \le k < 2^{m+1}$, are non-overlapping, we have by (25) in 5.4 that

$$\begin{aligned}\|K_m(\cdot,t)\|_1 &\le \sum_{k=2^m}^{2^{m+1}-1} \|\mathrm{f}_{k+1}\|_1 |h_k(t)| \\ &\le 6\sqrt{3},\end{aligned}$$

and, similarly,

$$\|K_m(x,\cdot)\|_1 \le C$$

for $m \in \mathrm{N}, x, t \in [0,1]$. It follows, therefore, from Lemma 3 in 5.3 that the operators $\boldsymbol{K}_m$ and $\boldsymbol{K}'_m$ are uniformly bounded on L^p, say

$$\|\boldsymbol{K}_m\|_p \le C', \qquad \|\boldsymbol{K}'_m\|_p \le C' \qquad (m \in \mathrm{N}).$$

Since $\mathrm{g}(+)$ is the Hadamard transform of $\mathbf{f}(+)$ and the Walsh system is the Hadamard transform of the Haar system, it is clear by (37) in 1.4 that

$$\sum_{k=2^m}^{2^{m+1}-1} b_k w_k = \sum_{k=2^m}^{2^{m+1}-1} b_k^{\perp} h_k,$$

$$\sum_{k=2^m}^{2^{m+1}-1} b_k \mathrm{g}_{k+1} = \sum_{k=2^m}^{2^{m+1}-1} b_k^{\perp} \mathrm{f}_{k+1}$$

for $(b_k, k \in \mathrm{N}) \in \ell^0$, and

$$\sum_{k=2^m}^{2^{m+1}-1} \mathrm{g}_{k+1}(x) w_k(t) = \sum_{k=2^m}^{2^{m+1}-1} \mathrm{f}_{k+1}(x) h_k(t)$$

for $x, t \in [0,1), m \in \mathrm{N}$. In particular, $\mathrm{g}(+)$ is a closed system in L^p_0. Moreover, in addition to (59) we have

$$\boldsymbol{K}_m w_k = \mathrm{g}_{k+1}, \qquad \boldsymbol{K}'_m \mathrm{g}_{k+1} = w_k \tag{60}$$

for $2^m \le k < 2^{m+1}$, and $m \in \mathrm{N}$.

To show $\mathrm{g}(+)$ is a basis in L_0^p fix $n \in \mathbf{N}$, choose $m \in \mathbf{N}$ satisfying $2^m \leq n < 2^{m+1}$, and let $\sum_{k=0}^{\infty} b_k \mathrm{g}_{k+1}$ be a Ciesielski polynomial. By (60),

$$\sum_{k=2^m}^{n} b_k \mathrm{g}_{k+1} = \boldsymbol{K}_m \left(\sum_{k=2^m}^{n} b_k w_k \right).$$

Since $\boldsymbol{w}$ is a basis in L^p we have by Lemma 5 and the estimates above that

$$\begin{aligned} \| \sum_{k=2^m}^{n} b_k \mathrm{g}_{k+1} \|_p &\leq C' B \| \sum_{k=2^m}^{2^{m+1}-1} b_k w_k \|_p \\ &= C' B \| \sum_{k=2^m}^{2^{m+1}-1} b_k^{\perp} h_k \|_p. \end{aligned}$$

Similarly, (59) and the fact that the Franklin system is a basis in L^p imply

$$\begin{aligned} \left\| \sum_{k=2^m}^{2^{m+1}-1} b_k^{\perp} h_k \right\|_p &= \left\| \boldsymbol{K}'_m \left(\sum_{k=2^m}^{2^{m+1}-1} b_k^{\perp} \mathrm{f}_{k+1} \right) \right\|_p \\ &\leq C' B \| \sum_{k=0}^{\infty} b_k^{\perp} \mathrm{f}_{k+1} \|_p \\ &= C' B' \| \sum_{k=0}^{\infty} b_k \mathrm{g}_{k+1} \|_p \end{aligned}$$

for some absolute constant B'. Writing

$$\sum_{k=0}^{n} b_k \mathrm{g}_{k+1} = \sum_{k=0}^{2^m-1} b_k \mathrm{g}_{k+1} + \sum_{k=2^m}^{n} b_k \mathrm{g}_{k+1},$$

it follows that there is an absolute constant B'' such that

$$\| \sum_{k=0}^{n} b_k \mathrm{g}_{k+1} \|_p \leq B'' \| \sum_{k=0}^{\infty} b_k \mathrm{g}_{k+1} \|_p$$

for all Ciesielski polynomials $\sum b_k \mathrm{g}_{k+1}$. Thus $\mathrm{g}(+)$ has the B-property and $\mathrm{g}(+)$ is a basis in L_0^p by Lemma 5.

Let $m \in \mathbf{N}$ and $(b_k, k \in \mathbf{N}) \in \ell^0$. Since

$$\| \sum_{k=0}^{2^m-1} b_k \mathrm{g}_{k+1} \|_p = \| \sum_{k=0}^{2^m-1} b_k^{\perp} \mathrm{f}_{k+1} \|_p$$

there exists by Theorem 14 an absolute constant $K > 0$ such that

$$K^{-1}\|\sum_{k=0}^{2^m-1} b_k g_{k+1}\|_p \le \|\sum_{k=0}^{2^m-1} b_k^{\perp} h_{k+1}\|_p \le K\|\sum_{k=0}^{2^m-1} b_k g_{k+1}\|_p.$$

Moreover, since the Haar and Walsh shift operators are continuous on L^p (see Theorem 13 in 5.4), it is clear that

$$\begin{aligned}\|\sum_{k=0}^{2^m-1} b_k^{\perp} h_{k+1}\|_p &\sim \|\sum_{k=0}^{2^m-1} b_k^{\perp} h_k\|_p \\ &= \|\sum_{k=0}^{2^m-1} b_k w_k\|_p \\ &\sim \|\sum_{k=0}^{2^m-1} b_k w_{k+1}\|_p.\end{aligned}$$

Consequently,

$$\|\sum_{k=0}^{2^m-1} b_k g_{k+1}\|_p \sim \|\sum_{k=0}^{2^m-1} b_k w_{k+1}\|_p$$

for all $(b_k, k \in \mathbf{N}) \in \ell^0$ and $m \in \mathbf{N}$. Since $\mathbf{g}(+)$ and $\boldsymbol{w}(+)$ are bases in L_0^p, it follows that they are equivalent in L^p. ∎

The argument used above works for certain W-systems (see Exercise 5.11) and a large class of product systems (see Theorem 16 below).

Let $\boldsymbol{\gamma} = (\gamma_k, k \in \mathbf{N})$ be a sequence of bounded, measurable functions defined on a probability space (Ω, ν). Recall that $\boldsymbol{\gamma}$ is strongly multiplicative if the product system $\boldsymbol{g} := (g_n, n \in \mathbf{N})$ defined by

$$g_n := \prod_{k=0}^{\infty} \gamma_k^{n_k} \tag{61}$$

(for $n \in \mathbf{N}$ with binary coefficients $(n_k, k \in \mathbf{N})$) is orthogonal in $L^2(\Omega)$. Notice that both the Walsh system and the original Walsh system are product systems of a strongly multiplicative system. Hence the following result shows that the Walsh system and the original Walsh system are equivalent in L^p for $1 < p < \infty$.

THEOREM 16. *Let $\boldsymbol{\gamma} = (\gamma_k, k \in \mathbf{N})$ be a strongly multiplicative system on a probability space (Ω, ν) with $\|\gamma_k\|_\infty \le 1$ for $k \in \mathbf{N}$. Let $\boldsymbol{g} := (g_n, n \in \mathbf{N})$ be the product system generated by $\boldsymbol{\gamma}$ and X^p denote the closure of the linear hull of $\boldsymbol{g}$ in $L^p(\Omega)$ norm. If*

$$\int_\Omega |g_n|^2 d\nu = \alpha \qquad (n \in \mathbf{P})$$

for some fixed $\alpha > 0$ *then* $\boldsymbol{g}$ *is a basis in* $\boldsymbol{X}^p$ *which is equivalent to the Walsh system in* $(\boldsymbol{X}^p, \mathrm{L}^p)$ *for all* $1 < p < \infty$.

PROOF. For each $m \in \mathbf{N}, x \in \Omega$, and $t \in [0,1)$, let

$$K_m(x,t) := \sum_{k=0}^{2^m-1} g_k(x) w_k(t),$$

$$(\boldsymbol{K}_m f)(x) := \int_0^1 K_m(x,t) f(t)\, dt \qquad (f \in \mathrm{L}^1),$$

and

$$(\boldsymbol{K}'_m g)(t) := \int_\Omega K_m(x,t) g(x)\, d\nu(x) \qquad (g \in \mathrm{L}^1(\Omega)).$$

By (61) and induction it is easy to verify

$$K_m(x,t) = \prod_{\ell=0}^{m-1} (1 + \gamma_\ell(x) r_\ell(t))$$

for $x \in \Omega, t \in [0,1)$ and $m \in \mathbf{N}$ (compare with Theorem 8 (i) in 1.5). Since $|\gamma_\ell| \leq 1$ a.e., it follows that $K_m \geq 0$ a.e. on $\Omega \times [0,1)$. Since $g_0 = 1$ we have by orthogonality that

$$\int_0^1 |K_m(x,t)|\, dt = 1$$

and

$$\int_\Omega |K_m(x,t)|\, d\nu(x) = 1.$$

Using the notation of Lemma 3 in 5.3, we have

$$\|K_m\|_{[\infty,1]} \leq 1, \qquad \|K_m\|_{[1,\infty]} \leq 1.$$

Consequently, the linear operators $\boldsymbol{K}_m$, respectively $\boldsymbol{K}'_m$, are uniformly bounded on L^p, respectively $\mathrm{L}^p(\Omega)$, for $1 \leq p \leq \infty$ and $m \in \mathbf{N}$, with $\|\boldsymbol{K}_m\|_p \leq 1$, and $\|\boldsymbol{K}'_m\|_p \leq 1$.

Fix $1 < p < \infty$. Clearly,

$$\boldsymbol{K}_m w_k = g_k \quad \text{and} \quad \boldsymbol{K}'_m g_k = \alpha w_k$$

for $0 \leq k < 2^m, m \in \mathbf{N}$. Hence for any polynomial $\sum_{k=0}^N b_k g_k$ and integers $n < N < 2^m$ we have

$$\left\| \sum_{k=0}^n b_k g_k \right\|_{L^p(\Omega)} = \left\| \boldsymbol{K}_m \left(\sum_{k=0}^n b_k w_k \right) \right\|_{L^p(\Omega)}$$
$$\leq \| \sum_{k=0}^n b_k w_k \|_p.$$

Since $\boldsymbol{w}$ is a basis in L^p we also have

$$\begin{aligned}\|\sum_{k=0}^{n} b_k w_k\|_p &\le B\|\sum_{k=0}^{N} b_k w_k\|_p \\ &= \frac{B}{\alpha}\left\|K'_m\left(\sum_{k=0}^{N} b_k g_k\right)\right\|_p \\ &\le \frac{B}{\alpha}\|\sum_{k=0}^{N} b_k g_k\|_{L^p(\Omega)}.\end{aligned}$$

Hence the system $(g_n, n \in \mathbf{N})$ has the B-property and is equivalent to the Walsh system in $(\boldsymbol{X}^p, \mathrm{L}^p)$. In particular, $(g_n, n \in \mathbf{N})$ is a basis in $\boldsymbol{X}^p$ by Lemma 5. ∎

The Walsh system is not equivalent to the Walsh-Kaczmarz system in L^p unless $p = 2$. To prove this we introduce the following notation.

Let T be a piecewise linear map on $\mathbf{N}$ induced by linear maps

$$T_n : \{0, 1, \ldots, 2^n - 1\} \to \{0, 1, \ldots, 2^n - 1\} \qquad (n \in \mathbf{N}).$$

For each $n \in \mathbf{N}$, let T'_n denote the adjoint of T_n and set $T^*_n := (T'_n)^{-1}$ (see Theorem 7 in 1.4). Then $T'_n : [0,1) \to [0,1)$ is a $1-1$, measure preserving transformation and

$$\begin{aligned} w_{T(2^n+k)}(x) &= w_{2^n+T_n(k)}(x), \\ w_k(T'_n(x)) &= w_{T_n(k)}(x) \end{aligned}$$

for $0 \le k < 2^n$ and $x \in [0,1)$.

Define $T^* : \mathcal{I} \to \mathcal{I}$ by $T^*(\emptyset) := \emptyset$ and $T^*(I) := T^*_n(I) := (T'_n)^{-1}(I)$ for $|I| = 2^{-n}$. Clearly,

$$|T^*(I)| = |I| \qquad (I \in \mathcal{I}).$$

(Such a map T^* will be called a measure preserving transformation on $\mathcal{I}$.) When $k < |I|^{-1}$, denote the value w_k takes on I by $w_k(I)$. Then the map $T^* : \mathcal{I} \to \mathcal{I}$ satisfies

$$w_{T(2^n+k)} = 2^{-n/2} \sum_{|I|=2^{-n}} w_k(I) h_{T^*(I)}. \tag{62}$$

Indeed, by (41) in 1.4,

$$w_{2^n+k} = 2^{-n/2} \sum_{|I|=2^{-n}} w_k(I) h_I. \tag{63}$$

Consequently,

$$\begin{aligned} w_{T(2^n+k)} &= w_{2^n+T_n(k)} \\ &= 2^{-n/2} \sum_{|I|=2^{-n}} w_{T_n(k)}(I) h_I \\ &= 2^{-n/2} \sum_{|I|=2^{-n}} w_k(T'_n(I)) h_I \\ &= 2^{-n/2} \sum_{|J|=2^{-n}} w_k(J) h_{T^*(J)} \end{aligned}$$

for $0 \le k < 2^n, n \in \mathbf{N}$.

For the following result, let

$$T^*\boldsymbol{h} := (h_{T^*(I)}, I \in \mathcal{I}).$$

LEMMA 6. *Let* $\boldsymbol{X}$ *and* $\boldsymbol{Y}$ *be Banach spaces. Suppose* $\boldsymbol{h}$ *and* $\boldsymbol{w}$ *are bases in* $\boldsymbol{X}$, *and* T^*h *and* $T\boldsymbol{w}$ *are bases in* $\boldsymbol{Y}$. *Then* $\boldsymbol{w}$ *and* $T\boldsymbol{w}$ *are equivalent in* $(\boldsymbol{X},\boldsymbol{Y})$ *if and only if* $\boldsymbol{h}$ *and* T^*h *are equivalent in* $(\boldsymbol{X},\boldsymbol{Y})$.

PROOF. Let $\boldsymbol{b} = (b_n, n \in \mathbf{N}) \in \ell^0$. Define a map $\boldsymbol{b} \to \boldsymbol{b}^\perp := (b_I^\perp, I \in \mathcal{I})$ by $b_\emptyset^\perp := b_\emptyset$, and

$$b_I^\perp := 2^{-n/2} \sum_{k=0}^{2^n-1} b_{2^n+k} w_k(I)$$

for $I \in \mathcal{I}, |I| = 2^{-n}, n \in \mathbf{N}$. Then $\perp$ is essentially the Hadamard transform and consequently is a bijection of ℓ^0 onto $\ell^0(\mathcal{I})$. Moreover, it follows directly from (63) and (62) that

$$\sum_{k=0}^{2^n-1} b_{2^n+k} w_{2^n+k} = \sum_{|I|=2^{-n}} b_I^\perp h_I$$

and

$$\sum_{k=0}^{2^n-1} b_{2^n+k} w_{T(2^n+k)} = \sum_{|I|=2^{-n}} b_I^\perp h_{T^*(I)}.$$

In particular, $\|S_{2^n}^{\boldsymbol{w}}\|_{\boldsymbol{X}} \sim \|S_{2^n}^{T\boldsymbol{w}}\|_{\boldsymbol{Y}}$ if and only if $\|S_{2^n}^{\boldsymbol{h}}\|_{\boldsymbol{X}} \sim \|S_{2^n}^{T^*\boldsymbol{h}}\|_{\boldsymbol{Y}}$. Since each of these systems is a basis, it follows that $\|S_n^{\boldsymbol{w}}\|_{\boldsymbol{X}} \sim \|S_n^{T\boldsymbol{w}}\|_{\boldsymbol{Y}}$ if and only if $\|S_n^{\boldsymbol{h}}\|_{\boldsymbol{X}} \sim \|S_n^{T^*\boldsymbol{h}}\|_{\boldsymbol{Y}}$. ∎

We see, then, that the problem of equivalence of linear rearrangements of the Walsh system reduces to equivalence of particular measure preserving rearrangements of the Haar system, i.e., rearrangements of the form $(h_{\varpi(I)}, I \in \mathcal{I})$, where ϖ is a $1-1$ map of $\mathcal{I}$ onto $\mathcal{I}$ which satisfies

$$|\varpi(I)| = |I| \qquad (I \in \mathcal{I}).$$

Let $\varpi : \mathcal{I} \to \mathcal{I}$ be measure preserving. Define the norm of ϖ by

$$\|\varpi\| := \sup_{I \in \mathcal{I}_0} \left(\frac{1}{|I|} |V_\varpi(I)| \right)^{1/2}$$

where

$$V_\varpi(I) := \bigcup \{\varpi(J) : J \in \mathcal{I} \quad \text{and} \quad J \subseteq I\}.$$

Define an operator R_ϖ on $\mathcal{P}$ by

$$R_\varpi f := \sum_{I \in \mathcal{I}} \dot{f}_I h_{\varpi(I)} \qquad (f \in \mathcal{P}),$$

where for each $I \in \mathcal{I}$, $\dot{f}_I := \langle f, h_I \rangle$ represents the corresponding Haar-Fourier coefficient of f.

The following result shows that various operator norms of R_ϖ are completely determined by the norm of ϖ.

THEOREM 17. *Let ϖ be a measure preserving bijection on $\mathcal{I}$. Then*

$$\|R_\varpi\|_{\mathrm{H}} = \|R_{\varpi^{-1}}\|_{\mathrm{BMO}} = \|\varpi\|.$$

Moreover, if $2 < p < \infty$ and q is the index conjugate to p, then there is a constant C_p depending only on p such that

$$\begin{aligned}\|\varpi\|^{(p-2)/p} &\leq \|R_\varpi\|_q \\ &= \|R_{\varpi^{-1}}\|_p \\ &\leq C_p\|\varpi\|^{(p-2)/p}.\end{aligned}$$

PROOF. We begin by showing

$$\|R_{\varpi^{-1}}\|_{\mathrm{BMO}} \leq \|\varpi\|. \tag{64}$$

Let $f \in \mathcal{P}$ with $\|f\|_{\mathrm{BMO}} \leq 1$ and recall that the map $f \to \overset{\bullet}{f}$ is an isometric isomorphism of BMO onto **bmo** (see 3.4). By definition, then, we have

$$\begin{aligned}\|R_{\varpi^{-1}}f\|_{\mathrm{BMO}} &= \|(R_{\varpi^{-1}}f)^\bullet\|_{\mathbf{bmo}} \\ &= \|(\overset{\bullet}{f}_{\varpi(J)}, J \in \mathcal{I})\|_{\mathbf{bmo}} \\ &= \sup_{I\in\mathcal{I}_0}\left(\frac{1}{|I|}\sum_{J\subseteq I}|\overset{\bullet}{f}_{\varpi(J)}|^2\right)^{1/2}.\end{aligned}$$

Write the set $V_\varpi(I)$ as a union of disjoint maximal intervals in the following way. For $x \in V_\varpi(I)$, let

$$n(x) := \min\{m : I_m(x) \subseteq V_\varpi(I) \quad \text{and} \quad \varpi^{-1}(I_m(x)) \subseteq I\}$$

and set

$$\mathcal{I}_I := \{I_{n(x)}(x) : x \in V_\varpi(I)\}.$$

Then $\mathcal{I}_I$ is the set of components of $V_\varpi(I)$ and

$$\{\varpi(J) : J \subseteq I\} \subseteq \{K \in \mathcal{I} : K \subseteq L, L \in \mathcal{I}_I\}.$$

In particular,

$$\begin{aligned}\frac{1}{|I|}\sum_{J\subseteq I}|\overset{\bullet}{f}_{\varpi(J)}|^2 &\leq \frac{1}{|I|}\sum_{L\in\mathcal{I}_I}\sum_{K\subseteq L}|\overset{\bullet}{f}_K|^2 \\ &\leq \frac{1}{|I|}\sum_{L\in\mathcal{I}_I}|L| \\ &= \frac{1}{|I|}|V_\varpi(I)|.\end{aligned}$$

for each $I \in \mathcal{I}_0$. This verifies (64).

To obtain the reverse inequality, let $\varepsilon > 0$ and choose an interval $I \in \mathcal{I}_0$ such that

$$\|\varpi\| - \varepsilon < \left(\frac{1}{|I|}|V_\varpi(I)|\right)^{1/2}.$$

Let

$$F_0 := \sum_{L \in \mathcal{I}_I} |L|^{1/2} h_L$$

and observe that $\|F_0\|_{\mathrm{BMO}} = 1$. Moreover, since $L \in \mathcal{I}_I$ implies $\varpi^{-1}(L) \subseteq I$ it is clear that

$$\begin{aligned}\frac{1}{|I|}\sum_{J \subseteq I} |(R_{\varpi^{-1}} F_0)^\bullet_J|^2 &= \frac{1}{|I|}\sum_{L \in \mathcal{I}_I} |L| \\ &= \frac{1}{|I|}|V_\varpi(I)| \\ &> (\|\varpi\| - \varepsilon)^2.\end{aligned}$$

It follows that

$$\|R_{\varpi^{-1}}\|_{\mathrm{BMO}} \geq \|R_{\varpi^{-1}} F_0\|_{\mathrm{BMO}} \geq \|\varpi\| - \varepsilon.$$

In particular

$$\|R_{\varpi^{-1}}\|_{\mathrm{BMO}} = \|\varpi\|.$$

By duality and the fact that $R^*_\varpi = R_{\varpi^{-1}}$, it remains to show

$$\|\varpi\|^{(p-2)/p} \leq \|R_\varpi\|_q = \|R_{\varpi^{-1}}\|_p \leq C_p \|\varpi\|^{(p-2)/p} \tag{65}$$

for $2 < p < \infty$.

To prove the left side of (65), fix $2 < p < \infty$ and observe that the function F_0 defined above satisfies $\|F_0\|_p = |V_\varpi(I)|^{1/p}$. By Hölder's inequality we have

$$\begin{aligned}|V_\varpi(I)| &= \|R_{\varpi^{-1}} F_0\|_2^2 \\ &= \int_0^1 \chi(I)|R_{\varpi^{-1}} F_0|^2 \\ &\leq \left(\int_0^1 \chi(I)\right)^{(p-2)/p} \|R_{\varpi^{-1}} F_0\|_p^2.\end{aligned}$$

Consequently,

$$\begin{aligned}\frac{\|R_{\varpi^{-1}} F_0\|_p}{\|F_0\|_p} &\geq |V_\varpi(I)|^{1/2}|I|^{-(p-2)/(2p)}|V_\varpi(I)|^{-1/p} \\ &= \left(\frac{1}{|I|}|V_\varpi(I)|\right)^{(p-2)/(2p)} \\ &> (\|\varpi\| - \varepsilon)^{(p-2)/p}.\end{aligned}$$

To prove the right side of (65) observe that $\|R_{\varpi}\|_2 = 1$. Hence by interpolation (see Theorem 12 in 3.6) with $1/q = (1-t)/2 + t$, i.e., $t = (2-q)/q = (p-2)/p$, it follows from (64) that

$$\|R_{\varpi}\|_{H^q} \leq C\|\varpi\|^{(p-2)/p}$$

for some absolute constant C. Since the H^p and L^p norms are equivalent, we conclude that (65) holds for some constant C_p depending only on p. ∎

COROLLARY 5. *Let $T : \mathcal{I} \to \mathcal{I}$ be a measure preserving bijection and let*

$$T\boldsymbol{h} := (h_{T(I)}, I \in \mathcal{I}).$$

i) The systems $\boldsymbol{h}$ and $T\boldsymbol{h}$ are equivalent in H *and in* BMO *if and only if*

$$\sup_{I\in\mathcal{I}_0} \frac{1}{|I|}\left|\bigcup_{J\subseteq I} T(J)\right| < \infty \quad \text{and} \quad \sup_{I\in\mathcal{I}_0} \frac{1}{|I|}\left|\bigcup_{J\subseteq I} T^{-1}(J)\right| < \infty.$$

ii) The systems $\boldsymbol{h}$ and $T\boldsymbol{h}$ are equivalent in L^p (for $1 < p < \infty$, $p \neq 2$) if and only if

$$\sup_{I\in\mathcal{I}_0} \frac{1}{|I|}\left|\bigcup_{J\subseteq I} T(J)\right| < \infty \quad \text{and} \quad \sup_{I\in\mathcal{I}_0} \frac{1}{|I|}\left|\bigcup_{J\subseteq I} T^{-1}(J)\right| < \infty.$$

PROOF. Both results are direct consequences of Theorem 17. For example, to prove i), set $\varpi := T$ and observe that

$$R_{\varpi}(h_I) = h_{T(I)} \qquad (I \in \mathcal{I}). \quad ∎$$

We apply these results to the Walsh-Kaczmarz rearrangement.

COROLLARY 6. *The Walsh (-Paley) and Walsh-Kaczmarz systems are not equivalent in* H, *in* BMO, *and in* L^p *for* $1 < p < \infty$ *and* $p \neq 2$.

PROOF. In 1.4 we proved that the piecewise linear transformation T which induces the Walsh-Kaczmarz rearrangement satisfies $T'_m = T^*_m$ for $m \in \mathbf{N}$.

For $m \in \mathbf{P}$ let $q(m) \in \mathbf{N}$ be defined by $2^{q(m)} \leq m < 2^{q(m)+1}$. Observe by the proof of Theorem 11 in 1.6 that the sets

$$E_m := T'_m([0, 2^{-q(m)})) = T^*_m([0, 2^{-q(m)}))$$

satisfy

$$\left|\bigcup_{m=M}^{\infty} E_m\right| = 1 \qquad (M \in \mathbf{P}).$$

Moreover, observe for $M \in \mathbf{P}$ and $I := [0, 2^{-q(M)})$ that

$$\bigcup_{J \subseteq I} T^*(J) \supseteq \bigcup_{m=q(M)}^{\infty} \left(\bigcup_{|J|=2^{-m}, J\subseteq I} T^*(J) \right)$$
$$= \bigcup_{m=q(M)}^{\infty} \left(\bigcup_{|J|=2^{-m}, J\subseteq I} T^*_m(J) \right)$$
$$\supseteq \bigcup_{m=q(M)}^{\infty} E_m.$$

It follows that

$$\sup_{I \in \mathcal{I}_0} \frac{1}{|I|} \left| \bigcup_{J \subseteq I} T^*(J) \right| = \infty.$$

By Corollary 5 and Lemma 6, we conclude that $\boldsymbol{w}$ and $T\boldsymbol{w}$ are not equivalent in H, in BMO, and in L^p for $1 < p < \infty$ and $p \neq 2$. ∎

Finally, we consider equivalence of the Walsh and trigonometric systems. In connection with this we introduce the following terminology.

For the next several pages, if Ω is a set then we shall denote the cartesian product $\Omega \times \Omega$ by Ω^2. Let (Ω, ν) be a complete measure space and $\boldsymbol{\epsilon} = (\epsilon_n, n \in \mathbf{N})$ be a system in $L^1(\Omega)$. The Kronecker product system generated by $\boldsymbol{\epsilon}$ is the system $\widetilde{\boldsymbol{\epsilon}} = (\widetilde{\epsilon}_{m,n}, m, n \in \mathbf{N})$ in $L^1(\Omega^2)$ defined by

$$\widetilde{\epsilon}_{m,n}(x, y) := \epsilon_m(x)\epsilon_n(y)$$

for $m, n \in \mathbf{N}, x, y \in \Omega$.

LEMMA 7. *Let (Ω, ν) be a complete measure space and suppose that $\boldsymbol{\epsilon} = (\epsilon_n, n \in \mathbf{N})$ and $\boldsymbol{\tau} = (\tau_n, n \in \mathbf{N})$ are equivalent systems in $L^p(\Omega)$ for some $1 \leq p \leq \infty$. Then the corresponding Kronecker product systems $\widetilde{\boldsymbol{\epsilon}}$ and $\widetilde{\boldsymbol{\tau}}$ are equivalent in $L^p(\Omega^2)$.*

PROOF. Fix a polynomial $F = \sum c_{m,n}\widetilde{\epsilon}_{m,n}$ in $L(\widetilde{\boldsymbol{\epsilon}})$ and set

$$G := \sum c_{m,n}\widetilde{\tau}_{m,n}.$$

For each $k \in \mathbf{N}$ set

$$\Phi_k := \sum_{\ell \in \mathbf{N}} c_{k,\ell}\epsilon_\ell$$

and

$$\Psi_k := \sum_{\ell \in \mathbf{N}} c_{\ell,k}\tau_\ell.$$

By the equivalence of ϵ and τ in $L^p(\Omega)$, and by Fubini's theorem, there exists a constant K such that

$$\begin{aligned}\|F\|_p &= \left(\int_\Omega\int_\Omega |\sum_{k\in\mathbb{N}} \Phi_k(y)\epsilon_k(x)|^p\, d\nu(x)\, d\nu(y)\right)^{1/p} \\ &\leq K\left(\int_\Omega\int_\Omega |\sum_{k\in\mathbb{N}} \Phi_k(y)\tau_k(x)|^p\, d\nu(x)\, d\nu(y)\right)^{1/p} \\ &= K\left(\int_\Omega\int_\Omega |\sum_{\ell\in\mathbb{N}} \Psi_\ell(x)\epsilon_\ell(y)|^p\, d\nu(x)\, d\nu(y)\right)^{1/p} \\ &\leq K^2\left(\int_\Omega\int_\Omega |\sum_{\ell\in\mathbb{N}} \Psi_\ell(x)\tau_\ell(y)|^p\, d\nu(x)\, d\nu(y)\right)^{1/p} \\ &= K^2\|G\|_p.\end{aligned}$$

By symmetry, we also have $\|G\|_p \leq K^2\|F\|_p$. It follows that the canonical isomorphism induced by the pair $\tilde{\epsilon},\tilde{\tau}$ is bicontinuous on $L^p(\Omega^2)$. Therefore the systems $\tilde{\epsilon}$ and $\tilde{\tau}$ are equivalent in $L^p(\Omega^2)$. ∎

THEOREM 18. *The trigonometric and Walsh (-Paley) systems are equivalent in* L^p *for* $1 < p < \infty$ *only when* $p = 2$.

PROOF. Fix $1 < p < \infty$, $p \neq 2$, and suppose to the contrary that the trigonometric and Walsh systems are equivalent in L^p.

Let $\mathbf{t} := (\mathrm{t}_k, k \in \mathbb{N})$ represent the trigonometric system normalized to $[0,1)$, i.e.,

$$\mathrm{t}_{2k}(x) := \cos 2\pi kx, \quad \mathrm{t}_{2k+1}(x) := \sin 2\pi kx$$

for $k \in \mathbb{N}, x \in [0,1)$. Let $\tilde{\mathbf{t}} := (\tilde{\mathrm{t}}_{k,\ell}, k,\ell \in \mathbb{N})$ be the Kronecker product system generated by $\mathbf{t}$, and $\tilde{\boldsymbol{w}} := (\tilde{w}_{m,n}, m,n \in \mathbb{N})$ be the Kronecker product system generated by the Walsh system. By Lemma 7, $\tilde{\mathbf{t}}$ and $\tilde{\boldsymbol{w}}$ are equivalent in $L^p([0,1)^2)$. Consequently, there is a constant $K > 0$ such that

$$\begin{aligned}K^{-1}\|\sum_{k,\ell\in\mathbb{N}} c_{k,\ell}\tilde{\mathrm{t}}_{k,\ell}\|_p &\leq \|\sum_{m,n\in\mathbb{N}} c_{m,n}\tilde{w}_{m,n}\|_p \\ &\leq K\|\sum_{k,\ell\in\mathbb{N}} c_{k,\ell}\tilde{\mathrm{t}}_{k,\ell}\|_p\end{aligned}$$

for every sequence $(c_{k,\ell}, k,\ell \in \mathbb{N})$ of real numbers which contains only finitely many non-zero entries.

Define multipliers U on $L(\tilde{\mathbf{t}})$ and V on $L(\tilde{\boldsymbol{w}})$ by

$$U(\sum_{k,\ell\in\mathbb{N}} c_{k,\ell}\tilde{\mathrm{t}}_{k,\ell}) := \sum_{\ell\leq k} c_{k,\ell}\tilde{\mathrm{t}}_{k,\ell}$$

and

$$V(\sum_{m,n\in\mathbf{N}} c_{m,n}\widetilde{w}_{m,n}) := \sum_{n\le m} c_{m,n}\widetilde{w}_{m,n}.$$

It is well known (see Zygmund [1], pp.253-266, or Exercise 5.12) that the operator U is continuous on $\mathrm{L}^p([0,1)^2)$. Since the choice of K above implies

$$K^{-1}\|Uf\|_p \le \|Vf\|_p \le K\|Uf\|_p,$$

it follows that V is continuous on $\mathrm{L}^p([0,1)^2)$. However, we shall prove (see Lemma 8 below) that V is continuous on $\mathrm{L}^p([0,1)^2)$ only when $p=2$. Therefore, $\mathbf{t}$ and $\boldsymbol{w}$ are equivalent in L^p only when $p=2$. ∎

LEMMA 8. *For each*

(66)
$$f = \sum_{(m,n)\in\mathbf{N}^2} c_{m,n}\widetilde{w}_{m,n} \in \mathrm{L}(\widetilde{\boldsymbol{w}})$$

define

(67)
$$Vf := \sum_{n\le m} c_{m,n}\widetilde{w}_{m,n}$$

If $1 \le p \le \infty$ *then* V *is bounded on* $\mathrm{L}^p([0,1)^2)$ *if and only if* $p=2$.

PROOF. By duality it suffices to show V is unbounded on $\mathrm{L}^p([0,1)^2)$ for each $1\le p<2$. Fix $1\le p<2, t\in\mathbf{P}$, and for each $0\le k<t$ set

(68)
$$R_k := \widetilde{w}_{2^k,0}\sum_{\ell=0}^{2^t-1}\widetilde{w}_{\ell,\ell}.$$

Consider the function

$$f := \sum_{k=0}^{t-1} R_k.$$

Notice that

$$R_k(x,y) = w_{2^k}(x)\,D_{2^t}(x\dot{+}y)$$

and consequently the R_k's are supported on the set

$$E := \{(x,y)\in[0,1)^2 : D_{2^t}(x\dot{+}y) = 2^t\}.$$

By translation invariance the measure of E can be calculated directly:

(69)
$$\begin{aligned}|E| &= 2^{-t}\int_0^1\int_0^1 D_{2^t}(x\dot{+}y)\,dx\,dy\\ &= 2^{-t}\int_0^1 D_{2^t}(x)\,dx\\ &= 2^{-t}.\end{aligned}$$

Moreover, this argument shows the R_k's are orthogonal with

$$\int_0^1 \int_0^1 R_k R_j = \delta_{kj} 2^t \tag{70}$$

for $0 \leq k, j < t$ where δ_{kj} is the Kronecker delta. Let $r = 2/p$ and r' be the index conjugate to r. By Hölder's inequality it follows from (68), (69), and (70) that

$$\begin{aligned} \|f\|_p &= \left(\iint_E |f|^p \right)^{1/p} \\ &\leq |E|^{1/(pr')} \left(\int_0^1 \int_0^1 |\sum_{k=0}^{t-1} R_k|^2 \right)^{1/2} \\ &= 2^{t(1/2 - 1/(pr'))} t^{1/2}. \end{aligned}$$

Therefore

$$\|f\|_p \leq 2^{t(1-1/p)} t^{1/2}. \tag{71}$$

To estimate Vf notice that

$$\widetilde{w}_{2^k,0} \widetilde{w}_{\ell,\ell} = \widetilde{w}_{2^k \oplus \ell, \ell}$$

and that $2^k \oplus \ell < \ell$ if and only if the k-th binary coefficient of ℓ satisfies $\ell_k = 1$. Thus if

$$A := \{\ell : 0 \leq \ell < 2^t, \ell_k = 0\}$$

and

$$B := \{\ell : 0 \leq \ell < 2^t, \ell_k = 1\}$$

then

$$Vf = \sum_{k=0}^{t-1} \widetilde{w}_{2^k,0} \sum_{\ell \in B} \widetilde{w}_{\ell,\ell}.$$

By Theorem 8 (i) in 1.5 we have

$$\begin{aligned} \sum_{\ell \in B} w_\ell(u) &= r_k(u) \prod_{j \neq k, j=0}^{t-1} (1 + r_j(u)) \\ &= \frac{r_k(u)}{2} \left(D_{2^t}(u) + D_{2^t}(u \dot{+} 2^{-k-1}) \right) \end{aligned}$$

for $u \in [0, 1)$ and $0 \leq k < t$. For such k the functions D_{2^t} and $\tau_{2^{-k-1}} D_{2^t}$ have pairwise non- overlapping supports. Hence applying this identity to $u = x \dot{+} y$ we have

$$|(Vf)(x,y)| \geq \frac{1}{2} \sum_{k=0}^{t-1} D_{2^t}(x \dot{+} y \dot{+} 2^{-k-1}).$$

Hence

$$\begin{aligned}\|Vf\|_p &\geq \frac{1}{2}\left(\sum_{k=0}^{t-1}\int_0^1\int_0^1 |D_{2^t}(x \dotplus y \dotplus 2^{-k-1})|^p\,dx\,dy\right)^{1/p}\\ &= \frac{1}{2}(t2^{t(p-1)})^{1/p}\\ &= \frac{1}{2}t^{1/p}2^{t(1-1/p)}.\end{aligned}$$

Therefore, it follows from (71) that

$$\frac{\|Vf\|_p}{\|f\|_p} \geq \frac{1}{2}t^{1/p-1/2}.$$

We have shown that for each $t \in \mathbf{P}$ there is a two-dimensional Walsh polynomial f_t such that

$$\frac{\|Vf_t\|_p}{\|f_t\|_p} \geq \frac{1}{2}t^{1/p-1/2}.$$

Let $t \to \infty$. Since $1 \leq p < 2$ we conclude that V is unbounded on $\mathrm{L}^p([0,1)^2)$. ∎

Let $\mathbf{N}^2$ be endowed with the group structure from the product of $(\mathbf{N}, \oplus)$ with itself (for product groups see Appendix 0.3). To show Lemma 8 holds for all the "lower triangular" multipliers

$$V_\beta f := \sum_{n \leq \beta m} c_{m,n}\widetilde{w}_{m,n} \qquad (\beta > 0),$$

we first show that there exist subgroups of $\mathbf{N}^2$ which lie close to the line $y = \beta x$.

LEMMA 9. *Let $\varepsilon > 0, \beta > 0$ and $t \in \mathbf{P}$. There exist $\mathbf{Z}_2$-linearly independent sets*

$$N_t := \{n_0, n_1, \ldots, n_{t-1}\} \subset \mathbf{N}$$

and

$$M_t := \{m_0, m_1, \ldots, m_{t-1}\} \subset \mathbf{N}$$

such that the group G_t in $\mathbf{N}^2$ generated by $\{(m_j, n_j) : 0 \leq j < t\}$ satisfies

$$|n - \beta m| < \varepsilon \qquad ((m,n) \in G_t). \tag{72}$$

PROOF. The proof is by induction on t.

Suppose first that $t = 1$. By a classical theorem of Kronecker (see Zygmund [1], Vol. II, p. 20) there are infinitely many fractions q/s which satisfy

$$|\beta - q/s| \leq s^{-2}.$$

Thus choose positive integers n_0, m_0 such that $m_0 > 1/\varepsilon$ and $|\beta - n_0/m_0| < 1/m_0^2$. Set

$$G_1 := \{(0,0), (m_0, n_0)\}$$

and observe that (72) holds for $t = 1$.

Next, suppose that (72) holds for some subgroup $G_t, t \geq 1$, with $\varepsilon/2$ in place of ε. Choose $r \in \mathbf{N}$ such that

$$2^r > \max\{m_j, n_j : 0 \leq j < t\}$$

and s and q such that

$$|2^r s\beta - 2^r q| < \varepsilon/2.$$

Set $m_t := 2^r s$, $n_t := 2^r q$ and verify that $m_t \oplus m = m_t + m$ and $n_t \oplus n = n_t + n$ for all $(m, n) \in G_t$. Set

$$N_{t+1} := N_t \cup \{n_t\}$$

and

$$M_{t+1} := M_t \cup \{m_t\}.$$

Thus by construction,

$$|n \oplus n_t - \beta(m \oplus m_t)| = |(n - \beta m) + (n_t - \beta m_t)| < \varepsilon.$$

We conclude by induction that the lemma is true for all $t \in \mathbf{P}$. ∎

THEOREM 19. *Let $\beta > 0$ and for each*

$$f = \sum_{(m,n)\in\mathbf{N}^2} c_{m,n}\widetilde{w}_{m,n} \in \mathrm{L}(\widetilde{\boldsymbol{w}})$$

define

$$V_\beta f := \sum_{n \leq \beta m} c_{m,n}\widetilde{w}_{m,n}.$$

If $1 \leq p \leq \infty$ then V_β is bounded on $\mathrm{L}^p([0,1)^2)$ if and only if $p = 2$.

PROOF. Fix $t \in \mathbf{N}, 1 \leq p < 2$ and $0 < \varepsilon < 1/2$. Apply Lemma 9 to choose sets M_t, N_t and the subgroup G_t in $\mathbf{N}^2$ such that (72) holds.

Since the sets M_t, N_t are $\mathbf{Z}_2$ linearly independent there are measure preserving transformations T_1, T_2 of $[0,1)$ such that

$$w_{m_j} = r_j \circ T_1$$

and

$$w_{n_j} = r_j \circ T_2$$

for $0 \leq j < t$. Define

$$T(x,y) := T_1(x)T_2(y)$$

for $(x,y) \in [0,1)^2$ and choose by the proof of Lemma 8 a two- dimensional Walsh polynomial f such that

$$\frac{\|V_1 f\|_p}{\|f\|_p} \geq \frac{1}{2} t^{1/2-1/p}.$$

Set

$$g := f \circ T$$

and notice by (72) and the stipulation $\varepsilon < 1/2$ that

$$V_\beta g = (V_1 f) \circ T.$$

In particular, we conclude that

$$\frac{\|V_\beta g\|_p}{\|g\|_p} = \frac{\|V_1 f\|_p}{\|f\|_p} \geq \frac{1}{2} t^{1/2-1/p}. \qquad \blacksquare$$

We close this section with the summary of results concerning bases which was promised in 5.3:

i) The Haar and Franklin systems are bases in L^1.

ii) The Franklin system is a basis in $\mathcal{H}$.

iii) The Haar system is a basis in H.

iv) The Haar, Walsh, Franklin, Ciesielski, original Walsh, and Walsh-Kaczmarz systems are bases in L^p for each $1 < p < \infty$. In fact, every piecewise linear rearrangement of the Walsh system is a basis in L^p for $1 < p < \infty$.

v) The shifted Haar system $(h_n, n \in \mathbf{P})$ is a basis in VMO.

vi) The shifted Franklin system $(\mathrm{f}_n, n \in \mathbf{P})$ is a basis in $\mathcal{VMO}$.

vii) The Haar system is a basis in $C_{\mathcal{W}}$.

viii) The Faber-Schauder and Franklin systems are bases in $\mathcal{C}[0,1]$.

5.6 The Basis Problem. This celebrated problem of Banach asks whether every separable Banach space has a basis. In this section we show the answer to this question is no by constructing separable Banach spaces, similar in spirit to the dyadic Hardy spaces, which fail to have a basis.

Using compact operators (see Appendix 0.0) we recast the problem. For the next several pages let $(\boldsymbol{X}, \|\cdot\|_{\boldsymbol{X}})$ be a Banach space, $\mathcal{L}(\boldsymbol{X})$ denote the collection of bounded linear operators from $\boldsymbol{X}$ to $\boldsymbol{X}$, and $\mathcal{K}(\boldsymbol{X})$ denote the collection of compact operators in $\mathcal{L}(\boldsymbol{X})$.

Call a set $\mathrm{K} \subseteq \boldsymbol{X}$ non-trivial if it contains at least one non-zero element. For each non-trivial compact subset K of $\boldsymbol{X}$ define a seminorm on $\mathcal{L}(\boldsymbol{X})$ by

$$\|\Lambda\|_{\mathrm{K}} := \sup_{x \in \mathrm{K}} \|\Lambda x\|_{\boldsymbol{X}} \qquad (\Lambda \in \mathcal{L}(\boldsymbol{X})).$$

Then $\boldsymbol{X}$ is said to have the *compact approximation property* (briefly, the CAP), if $\mathcal{K}(\boldsymbol{X})$ is dense in $\mathcal{L}(\boldsymbol{X})$ with respect to $\|\cdot\|_{\mathrm{K}}$ for every non-trivial compact set $\mathrm{K} \subseteq \boldsymbol{X}$.

It is easy to see that any Banach space with a basis has the CAP. Indeed, if $\boldsymbol{\epsilon}$ is a basis in $\boldsymbol{X}$ then the partial sums $S_n^{\boldsymbol{\epsilon}} x$ converge to x, as $n \to \infty$, uniformly on every compact subset of $\boldsymbol{X}$. Moreover, each $S_n^{\boldsymbol{\epsilon}}$ is a compact operator in $\mathcal{L}(\boldsymbol{X})$. Thus to solve the basis problem it suffices to construct a Banach space which fails to have the CAP.

Let $\langle\ ,\ \rangle$ denote the usual inner product in L^2. Suppose the Banach space $\boldsymbol{X}$ satisfies $\mathcal{P} \subseteq \boldsymbol{X} \subseteq L^1$. For each $\Lambda \in \mathcal{L}(\boldsymbol{X})$ and $n \in \mathbf{N}$ set

$$t_n(\Lambda) := \frac{1}{2^n} \sum_{k=0}^{2^n-1} \langle w_k, \Lambda w_k \rangle .$$

For convenience, set $t_{-1}(\Lambda) = 0$. Define the trace of Λ with respect to the Walsh system by

$$t(\Lambda) := \lim_{n\to\infty} t_n(\Lambda) \tag{73}$$

when this limit exists.

The trace function allows us to identify constructable conditions sufficient to conclude that $\boldsymbol{X}$ fails to have the CAP.

LEMMA 10. *Suppose* $\mathcal{P} \subseteq \boldsymbol{X} \subseteq \mathrm{L}^1$, $\|f\|_1 \le \|f\|_{\boldsymbol{X}}$ *for* $f \in \boldsymbol{X}$, *and*

$$\sup_{n\in\mathbf{N}} \|w_n\|_{\boldsymbol{X}} < \infty. \tag{74}$$

Suppose further that there exist non-trivial finite sets $K_0, K_1, \ldots$ *in* $\boldsymbol{X}$, *and a constant* $C > 0$, *such that*

$$\sum_{n=0}^{\infty} \sup_{f\in \mathrm{K}_n} \|f\|_{\boldsymbol{X}} < \infty \tag{75}$$

and

$$|t_n(\Lambda) - t_{n-1}(\Lambda)| \le C\|\Lambda\|_{\mathrm{K}_n} \tag{76}$$

holds for all $\Lambda \in \mathcal{L}(\boldsymbol{X})$ *and* $n \in \mathbf{N}$. *Then* $\boldsymbol{X}$ *fails to have the CAP.*

PROOF. Suppose for a moment that $t(\Lambda)$ exists and satisfies

$$|t(\Lambda)| \le C'\|\Lambda\|_{\mathrm{K}} \qquad (\Lambda \in \mathcal{L}(\boldsymbol{X})) \tag{77}$$

for some non-trivial compact set $\mathrm{K} \subseteq \boldsymbol{X}$ and some constant $C' > 0$.

Let $\varepsilon > 0$ and fix $\Lambda \in \mathcal{K}(\boldsymbol{X})$. Since Λ is compact, use (74) to choose $m \in \mathbf{N}$, a map $\varpi : \mathbf{N} \to 0, 1, \ldots, m$ and functions $f_0, f_1, \ldots, f_m \in \boldsymbol{X}$ such that

$$\|\Lambda w_n - f_{\varpi(n)}\| < \varepsilon.$$

The inequalities

$$|\langle f, w_k\rangle| \le \|f\|_1 \le \|f\|_{\boldsymbol{X}} \qquad (f \in \boldsymbol{X}, k \in \mathrm{N})$$

imply

$$\begin{aligned}
|t_n(\Lambda)| &= 2^{-n}\left|\sum_{k=0}^{2^n-1} \langle w_k, \Lambda w_k\rangle\right| \\
&\le 2^{-n}\sum_{k=0}^{2^n-1} \left|\left\langle w_k, \Lambda w_k - f_{\varpi(k)}\right\rangle\right| + 2^{-n}\sum_{k=0}^{2^n-1}\left|\left\langle w_k, f_{\varpi(k)}\right\rangle\right| \\
&\le \varepsilon + 2^{-n}\sum_{\ell=0}^{m}\sum_{k=0}^{2^n-1} |\langle w_k, f_\ell\rangle|
\end{aligned}$$

for $n \in \mathbf{N}$. Since $\langle w_k, f_\ell \rangle = \widehat{f_\ell}(k) \to 0$ as $k \to \infty$, we deduce that

$$\limsup_{n\to\infty} |t_n(\Lambda)| \le \varepsilon.$$

In particular, t vanishes identically on $\mathcal{K}(\boldsymbol{X})$. If $\boldsymbol{X}$ has the CAP, it follows from (77) that t vanishes identically on $\mathcal{L}(\boldsymbol{X})$. This contradicts the fact that if I is the identity in $\mathcal{L}(\boldsymbol{X})$ then $t(I) = 1$.

It suffices, therefore, to construct a non-trivial compact set $\mathrm{K} \subseteq \boldsymbol{X}$ such that (77) holds. Let

$$\alpha_n := \sup_{f \in \mathrm{K}_n} \|f\|_{\boldsymbol{X}} \qquad (n \in \mathbf{N})$$

and use (75) to choose $\omega_n > 0$ such that $\omega_n \to \infty$ as $n \to \infty$, and

$$\sum_{n=0}^{\infty} \alpha_n \omega_n < \infty.$$

Set

$$\mathrm{K} := \{0\} \bigcup \left(\bigcup_{n=0}^{\infty} \mathrm{K}_n/(\alpha_n\omega_n) \right)$$

where $\mathrm{K}_n/(\alpha_n\omega_n)$ represents the set of points $f \in \boldsymbol{X}$ which can be written as $f = g/(\alpha_n\omega_n)$ for some $g \in \mathrm{K}_n$. Since $\omega_n \to \infty$ as $n \to \infty$, the set K is compact. Moreover, by (76) we have

$$\begin{aligned} |t_n(\Lambda) - t_{n-1}(\Lambda)| &\le C\alpha_n\omega_n \|\Lambda\|_{\mathrm{K}_n/(\alpha_n\omega_n)} \\ &\le C\alpha_n\omega_n \|\Lambda\|_{\mathrm{K}} \end{aligned}$$

for $n \in \mathbf{N}$ and $\Lambda \in \mathcal{L}(\boldsymbol{X})$. Consequently, $t_n(\Lambda)$ converges as $n \to \infty$, and satisfies

$$\begin{aligned} |t(\Lambda)| &= |\sum_{n=0}^{\infty} t_n(\Lambda) - t_{n-1}(\Lambda)| \\ &\le C \left(\sum_{n=0}^{\infty} \alpha_n\omega_n \right) \|\Lambda\|_{\mathrm{K}} \end{aligned}$$

for every $\Lambda \in \mathcal{L}(\boldsymbol{X})$. ∎

For any σ-field $\mathcal{G}$ of measurable subsets of $[0,1)$ and each $f \in \mathrm{L}^1$, let $\mathcal{E}(f \mid \mathcal{G})$ denote the conditional expectation of f with respect to $\mathcal{G}$, i.e., $g := \mathcal{E}(f \mid \mathcal{G})$ is the $\mathcal{G}$ measurable function which satisfies

$$\int_E g = \int_E f \qquad (E \in \mathcal{G}).$$

(Such a g exists and is uniquely determined for each $f \in \mathrm{L}^1$ by the Radon-Nikodym theorem).

For each sequence $\mathcal{G} = (\mathcal{G}_n, n \in \mathbf{N})$ of σ- fields of measurable subsets of $[0,1)$ such that $\mathcal{G}_n \subset \mathcal{A}^{n+1}$ for $n \in \mathbf{N}$, set

$$\|f\|_{L^p(\mathcal{G})} := \sup_{n\in\mathbf{N}} \|\mathcal{E}(|f| \mid \mathcal{G}_n)\|_p \qquad (f \in \mathrm{L}^1).$$

Denote by $\mathrm{L}^p(\mathcal{G})$ the collection of $f \in \mathrm{L}^1$ which satisfy $\|f\|_{L^p(\mathcal{G})} < \infty$. It is obvious by definition that $\mathrm{L}^1(\mathcal{G})$ is isometrically isomorphic to L^1. Hence $\mathrm{L}^1(\mathcal{G})$ always has a basis. Moreover, when $\mathcal{G}$ is linearly ordered, $\mathrm{L}^p(\mathcal{G})$ is isometrically isomorphic to L^p for $1 < p < \infty$ (see Exercise 5.13) and thus also has a basis. Below we shall show that there exist non-linearly ordered $\mathcal{G}$ such that $\mathrm{L}^p(\mathcal{G})$ does not have a basis for any $1 < p < \infty$.

First notice for any sequence of σ-fields $\mathcal{G}$ and any $1 \le p \le \infty$, that $\mathrm{L}^p(\mathcal{G})$ is a Banach space $\boldsymbol{X}$ which satisfies $\mathcal{P} \subseteq \boldsymbol{X} \subseteq \mathrm{L}^1$, $\|f\|_1 \le \|f\|_{\boldsymbol{X}}$, and

$$\sup_{n\in\mathbf{N}} \|w_n\|_{\boldsymbol{X}} = 1.$$

In particular, if $\mathrm{L}_0^p(\mathcal{G})$ represents the closure of $\mathcal{P}$ in $\mathrm{L}^p(\mathcal{G})$, then $\mathrm{L}_0^p(\mathcal{G})$ is a separable Banach space which satisfies all the hypotheses of Lemma 10 except possibly (75) and (76). Therefore, we need to construct a sequence $\mathcal{G}$ of σ-fields in $[0,1)$ and a sequence $(\mathrm{K}_n, n \in \mathbf{N})$ of non-trivial finite sets in $\mathrm{L}_0^p(\mathcal{G}), 1 < p \le \infty$ such that (75) and (76) hold. It is here that the Walsh functions enter in an essential way.

Recall there is a $1-1$ map from the Walsh system $\boldsymbol{w}$ onto the $\mathbf{Z}_2$ linear space $\mathbf{G}_0$ (see 1.2). Thus the system $\boldsymbol{w}$ can be considered a $\mathbf{Z}_2$ linear space.

Let $\mathcal{S}$ denote the collection of all finite dimensional subspaces of $\boldsymbol{w}$. For each $S \in \mathcal{S}$ with basis R, the number of elements in S and R are related by

$$|S| = 2^{|R|}.$$

Moreover, S is the product system generated by the elements of R (see (61) in 5.5).

Fix $S \in \mathcal{S}$. Let $\mathcal{A}^S$ be the σ-algebra generated by S. Then $\mathcal{A}^S$ is atomic, i.e., there exist sets $A \in \mathcal{A}^S$ (called atoms) such that $B \subseteq A$ and $B \in \mathcal{A}^S$ imply $B = \emptyset$ or $B = A$. Moreover, the measure of each atom of $\mathcal{A}^S$ is precisely $|S|^{-1}$.

Denote the set of characteristic functions of the atoms of $\mathcal{A}^S$ by $\mathbf{At}(S)$ and the conditional expectation $\mathcal{E}(f \mid \mathcal{A}^S)$ by $\mathcal{E}^S(f)$. These definitions and orthogonality of the Walsh system imply

(78) $$w = |S| \sum_{a\in\mathbf{At}(S)} \langle w, a\rangle\, a \qquad (w \in S),$$

(79) $$a = \sum_{w\in S} \langle a, w\rangle\, w \qquad (a \in \mathbf{At}(S)),$$

and

(80) $$\mathcal{E}^S f = |S| \sum_{a\in\mathbf{At}(S)} \langle f, a\rangle\, a \qquad (f \in \mathrm{L}^1).$$

Combining these identities we see

$$\mathcal{E}^S f = \sum_{w \in S} \langle f, w \rangle \, w \qquad (f \in \mathrm{L}^1). \tag{81}$$

In particular,

$$\mathcal{E}^{S_1} \circ \mathcal{E}^{S_2} = \mathcal{E}^{S_1 \cap S_2} \qquad (S_1, S_2 \in \mathcal{S}). \tag{82}$$

For each pair of integers $n < m$ let $R_{n,m}$ denote the subspace of $\boldsymbol{w}$ generated by the Rademacher functions $\{r_k : n \leq k < m\}$. The connection between atoms and estimate (76) is revealed in the following identity:

$$t_n(\Lambda) - t_{n+m}(\Lambda) = 2^{-n} \sum_{a \in \mathbf{At}(R_{n,n+m})} \left(\sum_{w \in R_{0,n}} \langle aw, \Lambda((1-a)w) \rangle \right) \tag{83}$$

for all $\Lambda \in \mathcal{L}(\boldsymbol{X})$ and integers $n < m$. To verify this identity fix $\Lambda \in \mathcal{L}(\boldsymbol{X}), S \in \mathcal{S}, f \in \mathrm{L}^1$, and apply (78) and (79) to obtain

$$\begin{aligned} |S|^{-1} \sum_{w \in S} w \Lambda(fw) &= \sum_{w \in S} w \Lambda \left(\sum_{a \in \mathbf{At}(S)} \langle w, a \rangle \, af \right) \\ &= \sum_{a \in \mathbf{At}(S)} \left(\sum_{w \in S} \langle w, a \rangle \, w \right) \Lambda(af) \\ &= \sum_{a \in \mathbf{At}(S)} a \Lambda(af). \end{aligned}$$

Hence

$$\sum_{w \in S} w \Lambda(fw) = |S| \sum_{a \in \mathbf{At}(S)} a \Lambda(af). \tag{84}$$

We conclude from (81) and definition that

$$\begin{aligned} t_n(\Lambda) - t_{n+m}(\Lambda) &= 2^{-n} \sum_{w \in R_{0,n}} \left(\langle w, \Lambda w \rangle - 2^{-m} \sum_{w' \in R_{n,n+m}} \langle ww', \Lambda ww' \rangle \right) \\ &= 2^{-n} \sum_{w \in R_{0,n}} \left(\sum_{a \in \mathbf{At}(R_{n,n+m})} (\langle aw, \Lambda w \rangle - \langle aw, \Lambda aw \rangle) \right). \end{aligned}$$

One more definition is needed before we proceed with the construction of $\mathcal{G}$. Subspaces $S_1, S_2 \in \mathcal{S}$ are said to be independent if

$$S_1 \cap S_2 = \{w_0\}.$$

(Notice that w_0 is the "zero" of the linear space $\boldsymbol{w}$.) It is easy to see that two such subspaces S_1 and S_2 are independent if and only if the corresponding σ-fields $\mathcal{A}^{S_1}$ and $\mathcal{A}^{S_2}$ are independent in the probabilistic sense, i.e., if $E_j \in \mathcal{A}^{S_j} (j = 1, 2)$ then $|E_1 \cap E_2| = |E_1||E_2|$. Moreover, by construction, given $S_1 \subseteq S$ in $\mathcal{S}$, there is a third subspace $S_2 \subseteq S$ (called the complement of S_1 relative to S) such that S_1 and S_2 are independent and $S_1 + S_2 = S$.

THEOREM 20. *Let $1 < p < \infty$. There is a sequence of σ-fields $\mathcal{G}$ such that $L^p_0(\mathcal{G})$ fails to have a basis.*

PROOF. We have seen it suffices to construct a sequence $\mathcal{G}$ of σ-fields in $[0,1)$ and a sequence $(K_n, n \in N)$ of non-trivial finite sets in $\boldsymbol{X} := L^p(\mathcal{G})$ such that (75) and (76) hold.

We begin by showing there exist pairwise independent finite subspaces $S^{n,a}$ for each $a \in \mathbf{At}(R_{n,n+1})$ and $n \in N$ such that $S^{n,a} \subset R_{0,n}$ and

$$\text{(85)} \qquad \sum_{n=0}^{\infty} \left(\sum_{a \in \mathbf{At}(R_{n,n+1})} |S^{n,a}|^{-1/p'} \right) < \infty$$

where p' is the index conjugate to p.

First, take k copies of the group $\mathbf{Z}_2$ and form a product group $\mathbf{Z}_2^k$. Notice that the map $(n_0, n_1, \ldots, n_{k-1}) \to \sum_{j=0}^{k-1} n_j 2^j$ is a $1-1$ map from $\mathbf{Z}_2^k$ onto the set

$$\mathbf{F}_k := \{m \in N : 0 \le m < 2^k\}.$$

Since $\mathbf{F}_k$ is a field under dyadic addition and dyadic multiplication (see Appendix 0.3) it follows that there is a multiplication which makes $\mathbf{Z}_2^k$ a field. Denote this multiplication by $\odot$ and let

$$S_y := \{(x, x \odot y) \in \mathbf{Z}_2^k \times \mathbf{Z}_2^k : x \in \mathbf{Z}_2^k\}$$

for each $y \in \mathbf{Z}_2^k$. Clearly, each S_y is a linear subspace of $\mathbf{Z}_2^k \times \mathbf{Z}_2^k$ and $|S_y| = 2^k$. Moreover, since $S_{y_1} \cap S_{y_2} = \{0\}$ for $y_1 \ne y_2$, these subspaces are pairwise independent.

Next, fix $m \ge 5$ and set $k := 2^{m-2}$. Since the spaces

$$R_{2^{m-1},2^m} \quad \text{and} \quad \mathbf{Z}_2^k \times \mathbf{Z}_2^k$$

are isomorphic, these observations can be used to choose pairwise independent subspaces of $R_{2^{m-1},2^m}$, each containing 2^k elements. Since the condition $m \ge 5$ implies $2^k \ge 2 \cdot 2^m$, it follows that there exist pairwise independent subspaces

$$S^{n,a} \subset R_{2^{m-1},2^m} \subset R_{0,n}$$

for each integer $n \in [2^m, 2^{m+1})$ and each atom $a \in \mathbf{At}(R_{n,n+1})$ such that

$$|S^{n,a}| \ge 2^k.$$

Since $k > n/8$ for each $n \in [2^m, 2^{m+1})$, we have verified (85).

Let $\mathcal{G} = (\mathcal{G}_n, n \in N)$ be the sequence of σ- fields determined by

$$\mathbf{At}(\mathcal{G}_n) := \{ab : b \in \mathbf{At}(S^{n,a}), a \in \mathbf{At}(R_{n,n+1})\}.$$

Pairwise independence of the $S^{n,a}$'s together with (82) imply

$$\text{(86)} \qquad \mathcal{E}(f \mid \mathcal{G}_n) = 2 \sum_{a \in \mathbf{At}(R_{n,n+1})} a\mathcal{E}^{S^{n,a}}(fa)$$

for $f \in L^1$ and $n \in \mathbf{N}$.

For each $n \in \mathbf{N}$ let $\overline{S^{n,a}}$ represent the complement of $S^{n,a}$ relative to $R_{0,n}$. Set

$$\gamma_n := |S^{n,a}|^{1/p}$$

and define finite sets K_n by

$$K_n := \{\gamma_n w(1-a)b : w \in \overline{S^{n,a}}, b \in \mathbf{At}(S^{n,a}), a \in \mathbf{At}(R_{n,n+1})\}.$$

Then for any $\Lambda \in \mathcal{L}(L^p(\mathcal{G}))$ and $f \in K_n$ of the form $f = \gamma_n w(1-a)b$, we have

$$\begin{aligned} |\langle wab, \Lambda(w(1-a)b)\rangle| &\le \langle ab, \mathcal{E}(|\Lambda(w(1-a)b| \mid \mathcal{G}_n)\rangle \\ &\le \frac{1}{\gamma_n}\|ab\|_{p'} \sup_{f\in K_n} \|\Lambda f\|_{L^p(\mathcal{G})} \\ &\le \left(\frac{1}{\gamma_n |S^{n,a}|^{1/p'}}\right) \|\Lambda\|_{K_n}. \end{aligned}$$

Hence

$$|\langle wab, \Lambda(w(1-a)b)\rangle| \le |S^{n,a}|^{-1}\|\Lambda\|_{K_n} \tag{87}$$

for $\Lambda \in \mathcal{L}(L^p(\mathcal{G}))$ and $f = \gamma_n w(1-a)b \in K_n$.

To verify (76) combine (83), for $m = 1$, with (81) and (84) to obtain

$$\begin{aligned} t_n(\Lambda) - t_{n+1}(\Lambda) &= 2^{-n} \sum_{a\in \mathbf{At}(R_{n,n+1})} \sum_{w'\in S^{n,a}} \sum_{w''\in \overline{S^{n,a}}} \langle w'w''a, \Lambda(w'w''(1-a))\rangle \\ &= \sum_{a\in \mathbf{At}(R_{n,n+1})} |\overline{S^{n,a}}|^{-1} \sum_{w\in \overline{S^{n,a}}} \sum_{b\in \mathbf{At}(S^{n,a})} \langle wab, \Lambda(wb(1-a))\rangle \end{aligned}$$

for $\Lambda \in \mathcal{L}(L^p(\mathcal{G}))$ and $n \in \mathbf{N}$. In particular, inequality (87) implies

$$|t_n(\Lambda) - t_{n+1}(\Lambda)| \le 2\|\Lambda\|_{K_n}.$$

To verify (75) it suffices to show for any $n \in \mathbf{N}$ and $f_0 \in K_n$ that

$$\mathcal{E}(|f_0| \, |\mathcal{G}_m) \le 4 \sum_{a\in \mathbf{At}(R_{n,n+1})} |S^{n,a}|^{-1/p'} \tag{88}$$

is satisfied for all $m \in \mathbf{N}$. Indeed, by taking L^p norms, the supremum of these L^p norms, with respect to m, and then the supremum over $f_0 \in K_n$ we obtain from (88) that

$$\sup_{f\in K_n} \|f\|_{L^p(\mathcal{G})} \le 4 \sum_{a\in \mathbf{At}(R_{n,n+1})} |S^{n,a}|^{-1/p'}.$$

Hence summing this last inequality over all n and applying (85) we obtain (75).

To prove (88) let $f_0 = \gamma_n w(1-a_0)b_0 \in \mathrm{K}_n$ with $b_0 \in \mathbf{At}(S^{n,a_0})$ and $a_0 \in R_{n,n+1}$. Notice for $(m,a) \neq (n,a_0)$ that the σ-algebras $S^{m,a}$ and S^{n,a_0} are independent. Thus (82) implies

$$\begin{aligned}\mathcal{E}^{S^{m,a}}(a(1-a_0)b_0) &\le \mathcal{E}^{S^{m,a}} b_0 \\ &= \mathcal{E}^{S^{m,a}}\left(\mathcal{E}^{S^{n,a_0}} b_0\right) \\ &= \int_0^1 b_0 \\ &= |S^{n,a_0}|^{-1},\end{aligned}$$

for $a \neq a_0$. For $a = a_0$ we have $a(1-a_0)b_0 = 0$; consequently, this inequality holds for all $S^{m,a}$. But now (88) follows easily from (86) and the definition of γ_n. Indeed,

$$\begin{aligned}\mathcal{E}(|f_0|\ |\mathcal{G}_m) &\le 2\gamma_n \sum_{a\in\mathbf{At}(R_{m,m+1})} \mathcal{E}^{S^{m,a}}(a(1-a_0)b_0) \\ &\le 4\gamma_n |S^{n,a_0}|^{-1} \\ &= 4|S^{n,a_0}|^{-1/p'} \\ &\le 4 \sum_{a\in\mathbf{At}(R_{n,n+1})} |S^{n,a}|^{-1/p'}. \qquad \blacksquare\end{aligned}$$

Exercises

5.1 Let $\boldsymbol{X}$ be a homogeneous Banach space.

a) Prove

$$\mathbf{E}_{2^n}(f,\boldsymbol{X}) \le \|f - S_{2^n} f\|_{\boldsymbol{X}} \le 2\mathbf{E}_{2^n}(f,\boldsymbol{X})$$

for all $f \in \boldsymbol{X}, n \in \mathbf{N}$.

b) Show the constant 2 in a) is sharp.

5.2 Suppose f is dyadically differentiable in a homogeneous Banach space $\boldsymbol{X}$. Prove f is constant if and only if $\mathbf{d}f = 0$.

5.3 Let $\boldsymbol{X}$ be a homogeneous Banach space. For any $f \in \boldsymbol{X}$ and $n \in \mathbf{P}$, let $P_n f$ represent any best approximation to f in $\mathcal{P}_n$. Suppose $\alpha > 0, r \in \mathbf{P}$, and f is differentiable in $\boldsymbol{X}$ of order r. Consider the following conditions

(A) $$\mathbf{E}_n(f,\boldsymbol{X}) = O(n^{-r-\alpha}) \quad \text{as} \quad n \to \infty$$

(B) $$\|\mathbf{d}^{[k]}(P_n f)\|_{\boldsymbol{X}} = O(n^{k-r-\alpha}) \quad \text{as} \quad n \to \infty,$$

for $k > r + \alpha, k \in \mathbf{N}$.

a) Show (A) implies

$$\|\sum_{m=1}^{j} \mathbf{d}^{[k]}(P_{2^m} f - P_{2^{m-1}} f)\|_{\boldsymbol{X}} = O(n^{k-r-\alpha})$$

and

$$\|\mathbf{d}^{[k]}(P_{2^j}f - P_nf)\|_{\boldsymbol{X}} = O(n^{k-r-\alpha}),$$

as $n \to \infty$, for $k > r + \alpha$, where $j \in \mathbf{N}$ satisfies $2^j \le n < 2^{j+1}$.

b) Using a), show (A) implies (B).

c) Prove

$$\mathbf{E}_{2^j}(f,\boldsymbol{X}) \le \mathbf{E}_{2^{j+N}}(f,\boldsymbol{X}) + \sum_{m=1}^{N} \|P_{2^{j+m}}f - P_{2^{j+m-1}}f\|_{\boldsymbol{X}}$$

for $N \in \mathbf{P}$.

d) Using c) and Theorem 7 in 5.2, show (B) implies (A).

e) Show condition (B) is equivalent to $\mathbf{d}^{[r]}f \in \mathrm{Lip}(\alpha, \boldsymbol{X})$.

5.4 a) Let $\Delta_n^{(1)}f := 1/n\sum_{k=1}^{n}|S_kf - f|$ for $n \in \mathbf{P}, f \in \mathrm{L}^1$. Prove

$$\|\Delta_n^{(1)}f\|_\infty = O\left(\frac{\log n}{n}\right) \qquad \text{as} \quad k \to \infty,$$

for all $f \in \mathrm{Lip}(1, W)$.

b) Using Walsh-Fejér polynomials, prove there exists a function $F \in \mathrm{Lip}(1, W)$ and constant $c > 0$ such that

$$\|\Delta_n^{(1)}F\|_\infty \ge c\frac{\log n}{n} \qquad (n \in \mathbf{P}).$$

5.5 For $1 \le p < r < \infty$ set

$$\|f\|_{p,r} = \|\sup_{n\in\mathbf{N}} (S_{2^n}|f|^p)^{1/p}\|_r.$$

Prove $(\mathrm{L}^r, \|\cdot\|_{p,r})$ is a Banach space. If p, p' and r, r' are conjugate indices, show that the dual of $(\mathrm{L}^p, \|\cdot\|_{p,r})$ is isometrically isomorphic to $(\mathrm{L}^{p'}, \|\cdot\|_{p',r'})$.

5.6 a) Use Exercises 3.18 and 3.25 to prove that every function dyadically differentiable in L^1 belongs to BMO.

b) Show that every function dyadically differentiable in L^p, $1 < p < \infty$, belongs to $\mathrm{Lip}(1/q, W)$, where q is the index conjugate to p.

[Ladhawala [1]]

5.7 Show that every function f dyadically differentiable in H satisfies $\widehat{f} \in \ell^1$ and belongs to C_W.

5.8 Let

$$f := \sum_{n=2}^{\infty} \frac{w_n}{n\log n}.$$

Show f is dyadically differentiable in L^1 but is unbounded on $(0,1)$.

5.9 a) Prove that the indefinite dyadic integral of an H function belongs to C_W.

b) Prove that the indefinite dyadic integral of an L^1 function belongs to BMO but may be unbounded.

5.10 Let $\boldsymbol{\epsilon}$ be equivalent to $\boldsymbol{\tau}$ in $(\boldsymbol{X},\boldsymbol{Y})$, where $\boldsymbol{X}$ and $\boldsymbol{Y}$ are Banach spaces. Show that $\boldsymbol{\epsilon}$ is a basis in $\mathrm{CL}(\boldsymbol{\epsilon})$ if and only if $\boldsymbol{\tau}$ is a basis in $\mathrm{CL}(\boldsymbol{\tau})$.

5.11 Let $\boldsymbol{X} \subseteq \mathrm{L}^1$ be a Banach space. Let $(H_n, n \in \mathrm{N})$ and $(G_n, n \in \mathrm{N})$ be H-systems and let $(W_n, n \in \mathrm{N})$ and $(Z_n, n \in \mathrm{N})$ be the W-systems they generate (see Appendix 0.8). Suppose $(H_n, n \in \mathrm{N})$ is equivalent to $(G_n, n \in \mathrm{N})$ in $\boldsymbol{X}$ and that $(H_n, n \in \mathrm{N}), (G_n, n \in \mathrm{N})$, and $(W_n, n \in \mathrm{N})$ are bases in $\boldsymbol{X}$. Prove that $(Z_n, n \in \mathrm{N})$ is a basis in $\boldsymbol{X}$ and is equivalent to $(W_n, n \in \mathrm{N})$.

5.12 Let $\boldsymbol{\epsilon} = (\epsilon_k, k \in \mathbf{Z})$ where

$$\epsilon_k(t) := \exp(2\pi \imath k t)$$

for $k \in \mathbf{Z}$, $t \in [0,1)$, and $\imath := \sqrt{-1}$. Let $\widetilde{\boldsymbol{\epsilon}}$ be the Kronecker product system generated by $\boldsymbol{\epsilon}$. For each $A \subseteq \mathbf{R}^2$ and $f = \sum a_{m,n}\widetilde{\epsilon}_{m,n} \in \mathrm{L}(\widetilde{\boldsymbol{\epsilon}})$ define

$$T_A f := \sum_{(m,n)\in A} a_{m,n}\widetilde{\epsilon}_{m,n}.$$

a) Let $V, V' : \mathbf{R}^2 \to \mathbf{R}^2$ be defined by $V(x,y) := (x, x+y)$ and $V'(x,y) := (x+y, y)$. Set $V^* := (V')^{-1}$ and $Vf := f \circ V$ for $f \in \mathrm{L}(\widetilde{\boldsymbol{\epsilon}})$. Prove

$$V^{-1} \circ T_A \circ V = T_{V^*(A)}$$

for every $A \subseteq \mathbf{R}^2$.

b) Let$A := \{(k,\ell) \in \mathbf{Z}^2 : \ell \leq 0\}$ and prove

$$V^*(A) = \{(m,n) \in \mathbf{Z}^2 : n \leq m\}.$$

c) The trigonometric analogue of Corollary 6 in 3.3 is true. Thus

$$\| \sum_{k=n}^{N} a_k \epsilon_k \|_p \leq C_p \|f\|_p \tag{89}$$

for $1 < p < \infty$, $n, N \in \mathrm{N}$, and $f = \sum_{k=0}^{\infty} b_k \epsilon_k \in \mathrm{L}(\boldsymbol{\epsilon})$. Use (89) to show that

$$\|T_A f\|_p \leq C_p \|f\|_p$$

for any $f \in \mathrm{L}(\widetilde{\boldsymbol{\epsilon}})$, and $1 < p < \infty$.

d) Prove that the multiplier

$$Uf := \sum_{n \leq m} b_{m,n}\widetilde{\epsilon}_{m,n}$$

is bounded on $\mathrm{L}^p([0,1)^2)$ for any $1 < p < \infty$.

5.13 Let $\boldsymbol{\mathcal{G}} = (\mathcal{G}_n, n \in \mathrm{N})$ be any sequence of σ-fields of measurable subsets of $[0,1)$.

a) Prove that

$$\|f\|_{L^1(\mathcal{G})} = \|f\|_1 \qquad (f \in \mathrm{L}^1).$$

b) Suppose that $\mathcal{G}_0 \subset \mathcal{G}_1 \subset \ldots$. Prove that

$$\|f\|_{L^p(\mathcal{G})} \sim \|f\|_p \qquad (f \in \mathrm{L}^p)$$

for $1 < p \leq \infty$.

5.14 Let $(\boldsymbol{X}, \|\cdot\|_{\boldsymbol{X}})$ be a Banach space and $x_n \in \boldsymbol{X}$ for $n \in \mathrm{N}$. Denote by $\boldsymbol{X}_n$ the closure of the linear hull of $\{x_k : k \in \mathrm{N} \text{ and } k \neq n\}$. If $x_n \notin \boldsymbol{X}_n$ for every $n \in \mathrm{N}$ then $(x_n, n \in \mathrm{N})$ is called a minimal system.

a) Prove $(x_n, n \in \mathrm{N})$ has a biorthogonal system if and only if it is minimal.

b) Prove a system $\boldsymbol{\epsilon} = (\epsilon_n, n \in \mathrm{N})$ is a basis in $\boldsymbol{X}$ if and only if the following three conditions are satisfied:

i) $\boldsymbol{\epsilon}$ is a closed system,

ii) $\boldsymbol{\epsilon}$ is a minimal system,

iii) if $\boldsymbol{\epsilon}' = (\epsilon'_n, n \in \mathrm{N})$ is the biorthogonal system of $\boldsymbol{\epsilon}$ then there is a constant $M > 0$ such that for every $x \in \boldsymbol{X}$ and $N \in \mathrm{N}$,

$$\|\sum_{n=0}^{N} \langle x, \epsilon'_n \rangle \, \epsilon_n\|_{\boldsymbol{X}} \leq M\|x\|_{\boldsymbol{X}}.$$

[Kašin-Saakjan].

5.15 A basis $\boldsymbol{\epsilon} = (\epsilon_n, n \in \mathrm{N})$ in a Banach space $(\boldsymbol{X}, \|\cdot\|_{\boldsymbol{X}})$ is called unconditional if $(\epsilon_{\varpi(n)}, n \in \mathrm{N})$ is a basis in $\boldsymbol{X}$ for every bijection $\varpi : \mathrm{N} \to \mathrm{N}$.

i) Prove a closed, minimal system $\boldsymbol{\epsilon}$ (see Exercise 5.14) is an unconditional basis in $\boldsymbol{X}$ if and only if there is a number $M > 0$ such that

$$\|\sum_{n=0}^{N} r_n(t) \, \langle x, \epsilon'_n \rangle \, \epsilon_n\|_{\boldsymbol{X}} \leq M\|x\|_{\boldsymbol{X}}$$

for every $N \in \mathrm{N}, x \in \boldsymbol{X}$, and $t \in [0, 1)$, where $(\epsilon'_n, n \in \mathrm{N})$ is the biorthogonal system of $\boldsymbol{\epsilon}$.

[Kašin-Saakjan]

ii) Prove the Haar system is an unconditional basis in L^p for each $1 < p < \infty$.

[Marcinkiewicz [1]]

Chapter 6

ALMOST EVERYWHERE CONVERGENCE AND SUMMABILITY OF WALSH-FOURIER SERIES

6.1 Tests for Almost Everywhere Convergence. By Theorem 12 4.5, we know that the Walsh-Fourier series of an $f \in L^1$ may diverge everywhere. In this section we identify several conditions sufficient to conclude that Sf converges a.e. and obtain an estimate for the growth of the partial sums $S_n f$ valid for all $f \in L^1$.

By Theorem 2 in 4.1, if the L^1 modulus of continuity of f decays sufficiently rapidly, then Sf converges in L^1 norm. A similar phenomenon holds for a.e. convergence.

THEOREM 1. *Let $f \in L^1$. If any one of the three conditions*

$$\sum_{k=0}^{\infty} \int_0^1 \int_0^1 |f(x\dot{+}u) - f(x)| D_{2^k}(u)\, dx du < \infty, \tag{1}$$

$$\sum_{n=1}^{\infty} \frac{1}{n} \omega^{(1)}\left(f, \frac{1}{n}\right) < \infty, \tag{2}$$

or

$$\omega^{(1)}(f,\delta) = O\left(\log \frac{1}{\delta}\right)^{-1-\varepsilon}, \quad \text{as} \quad \delta \to 0, \quad \textit{for some} \quad \varepsilon > 0, \tag{3}$$

holds, then the Walsh-Fourier series of f converges a.e. on $[0,1)$.

PROOF. Suppose (1) holds. By Fubini's theorem we have

$$\sum_{k=0}^{\infty} \int_0^1 |f(x\dot{+}u) - f(x)| D_{2^k}(u)\, du < \infty \tag{4}$$

for a.e. $x \in [0,1)$. Given $m, n \in \mathbf{N}$ apply Theorem 8 in 1.5 to write

$$\begin{aligned} |S_n f(x) - f(x)| &= \left| \int_0^1 (f(x\dot{+}u) - f(x)) w_n(u) \left(\sum_{k=0}^{\infty} n_k r_k(u) D_{2^k}(u) \right) du \right| \\ &\le \sum_{k=0}^{m} \left| \int_0^1 (f(x\dot{+}u) - f(x)) r_k(u)\, D_{2^k}(u) w_n(u)\, du \right| \\ &\qquad + \sum_{k=m+1}^{\infty} \int_0^1 |\, f(x\dot{+}u) - f(x) \,|\, D_{2^k}(u)\, du \\ &=: \varepsilon(x) + R_m(x). \end{aligned}$$

Notice for each $x \in [0,1)$, that the terms of $\varepsilon(x)$ are Walsh-Fourier coefficients. In fact, if

$$G_x^{(k)}(u) := \big(f(x\dot{+}u) - f(x)\big)\, r_k(u) D_{2^k}(u)$$

for $k \in \mathbf{N}$ and $u \in [0,1)$ then

$$\varepsilon(x) \le \sum_{k=0}^{m} |\widehat{G}_x^{(k)}(n)|.$$

It follows, therefore, from the Riemann-Lebesgue lemma that

$$\limsup_{n\to\infty} |S_n f(x) - f(x)| \le R_m(x).$$

But (4) implies $R_m \to 0$ a.e., as $m \to \infty$. Consequently, $S_n f \to f$ a.e., as $n \to \infty$.

Next, suppose (2) holds. Clearly,

$$\begin{aligned}\int_0^1 \int_0^1 |f(x\dot{+}u) - f(x)| D_{2^k}(u)\, du dx &= \int_0^1 D_{2^k}(u) \left(\int_0^1 |f(x\dot{+}u) - f(x)|\, dx\right) du \\ &= 2^k \int_{I_k(0)} \|\tau_u f - f\|_1\, du \le \omega^{(1)}(f, 2^{-k})\end{aligned}$$

by definition. Moreover, since $\omega^{(1)}(f, 1/n)$ is monotone as $n \to \infty$, condition (2) is equivalent to

$$\sum_{n=0}^{\infty} \omega^{(1)}(f, 2^{-n}) < \infty.$$

Therefore, condition (2) implies condition (1). Since it is trivial that condition (3) implies condition (2), the proof of the theorem is complete. ∎

The next test is a Walsh analogue of a theorem of Marcinkiewicz. We need two preliminary results.

LEMMA 1. *Let E_0 be a measurable subset of $[0,1)$ and $\varepsilon > 0$ satisfy*

$$|E_0| + \varepsilon < 1.$$

There exists a collection $\mathcal{I}_*$ of pairwise disjoint dyadic intervals such that if $\overline{E}_0 := [0,1) \setminus E_0$ and $E := \bigcup_{I \in \mathcal{I}_*} I$, then $E_0 \subseteq E, |E \setminus E_0| < \varepsilon$, and

$$(5) \qquad I_+ \cap \overline{E}_0 \ne \emptyset \qquad \textit{for every} \qquad I \in \mathcal{I}_*.$$

PROOF. If $|E_0| = 0$ let $\mathcal{I}_1 = \emptyset$ and $\mathcal{I}_2$ be a collection of dyadic intervals which covers E_0 and satisfies

$$(6) \qquad \sum_{I \in \mathcal{I}_2} |I| < \varepsilon/2.$$

Set $\widetilde{\mathcal{I}} := \mathcal{I}_1 \cup \mathcal{I}_2, E := \bigcup_{I \in \widetilde{\mathcal{I}}} I$, and notice that $|E \setminus E_0| < \varepsilon/2$.

If $|E_0| \neq 0$ let

$$y := \left(1 + \frac{\varepsilon}{2|E_0|}\right)^{-1},$$

$E_1 := \{\mathcal{E}^*(\chi(E_0)) > y\}$, and $E_2 := E_0 \setminus E_1$. By Corollary 1 in 3.1

$$|E_1| \leq \|\chi(E_0)\|_1 \left(1 + \frac{\varepsilon}{2|E_0|}\right) = |E_0| + \frac{\varepsilon}{2},$$

hence $|E_1 \setminus E_0| \leq \varepsilon/2$. On the other hand, $\mathcal{E}_n(\chi(E_0)) \to \chi(E_0)$ a.e. as $n \to \infty$ implies $\chi(E_1) \geq \chi(E_0)$, hence $|E_2| = 0$. Evoke

$$E_1 = \bigcup_{n=0}^{\infty} \{\mathcal{E}_n^*(\chi(E_0)) > y, \mathcal{E}_{n-1}^*(\chi(E_0)) \leq y\}$$

to choose a collection $\mathcal{I}_1$ of dyadic intervals whose union is E_1. Since E_2 is of measure zero, we can cover E_2 with a collection $\mathcal{I}_2$ of dyadic intervals which satisfies (6). Again, set $\widetilde{\mathcal{I}} := \mathcal{I}_1 \cup \mathcal{I}_2$, $E := \cup_{I \in \widetilde{\mathcal{I}}} I$, and notice that $|E \setminus E_0| < \varepsilon$.

For each $x \in E$ let

$$\varpi(x) := \min\{n \in \mathbf{N} : I_n(x) \subseteq E\}$$

and set

$$\mathcal{I}_* := \{I_{\varpi(x)}(x) : x \in E\}.$$

Clearly, E is the union of $\mathcal{I}_*$ and $\mathcal{I}_*$ satisfies condition (5) by construction. Moreover, given $x, y \in E$ with $I_{\varpi(x)}(x) \cap I_{\varpi(y)}(y) \neq \emptyset$, it is also clear by construction that

$$I_{\varpi(x)}(x) = I_{\varpi(y)}(y).$$

Hence $\mathcal{I}_*$ is a collection of pairwise disjoint intervals and the proof of the lemma is complete. ∎

For the next several pages we shall write $\ell(I) := n + 1$ when I is a dyadic interval of length 2^{-n}.

The following result is a localization theorem (compare with Theorem 13 in 2.5).

LEMMA 2. *Let $\mathcal{I}_*$ be a collection of pairwise disjoint dyadic intervals and*

$$E := \bigcup_{I \in \mathcal{I}_*} I.$$

If $g \in \mathrm{L}^1$ satisfies $\{g \neq 0\} \subseteq E$ and

$$c' := \sup_{I \in \mathcal{I}_*} \frac{\ell(I)}{|I|} \int_I |g| < \infty,$$

then Sg converges a.e. to zero on $\overline{E} := [0,1) \setminus E$.

PROOF. For each $n \in \mathbf{N}$ set

$$g_n := \sum_{I \in \mathcal{I}_*} \frac{\chi(I)}{|I|} \sup_{m \geq n} \left| \int_I g w_m \right|$$

and

$$G_n := \sum_{I \in \mathcal{I}_*} \chi(I)\ell(I) g_n.$$

By the Riemann-Lebesgue lemma, G_n decreases monotonically to zero, as $n \to \infty$, everywhere on [0,1). Moreover, by hypothesis,

$$\begin{aligned} \|G_0\|_1 &\leq \sum_{I \in \mathcal{I}_*} \ell(I) \int_I |g(t)|\, dt \\ &\leq c' \sum_{I \in \mathcal{I}_*} |I| \\ &= c'|E| < \infty. \end{aligned}$$

Hence the Lebesgue dominated convergence theorem implies

$$\lim_{n \to \infty} \|G_n\|_1 = 0. \tag{7}$$

Set

$$T_n := \sum_{k=0}^{\infty} g_n * D_{2^k}$$

and observe that

$$\begin{aligned} \int_{\overline{E}} T_n &= \sum_{k=0}^{\infty} \int_{\overline{E}} g_n * D_{2^k} \\ &= \sum_{k=0}^{\infty} \sum_{I \in \mathcal{I}_*} \int_{\overline{E}} \left(\int_0^1 (\chi(I) g_n)(t) D_{2^k}(x \dot{+} t)\, dt \right) dx. \end{aligned}$$

Since $x \notin I$ and $k \geq \ell(I) - 1$ imply

$$\chi(I)(t) D_{2^k}(x \dot{+} t) = 0$$

for every $t \in [0,1)$, it follows that

$$\begin{aligned} \int_{\overline{E}} T_n &\leq \sum_{I \in \mathcal{I}_*} \sum_{k < \ell(I)} \int_0^1 (\chi(I) g_n)(t) \left(\int_0^1 D_{2^k}(x \dot{+} t)\, dx \right) dt \\ &= \sum_{I \in \mathcal{I}_*} \sum_{k < \ell(I)} \int_0^1 \chi(I) g_n = \int_0^1 G_n. \end{aligned}$$

In particular, (7) implies

$$\lim_{m\to\infty}\int_{\overline{E}} T_m = 0. \tag{8}$$

Fix $x \in \overline{E}$. Then x does not belong to any $I \in \mathcal{I}_*$, so $\boldsymbol{\tau}_x(r_k D_{2^k})$ is constant on each interval of $\mathcal{I}_*$ for $k = 0, 1, \ldots$. Hence

$$\sum_{I\in\mathcal{I}_*}\int_I \left(\left(\frac{1}{|I|}\int_I g(t)w_n(t)dt - gw_n\right)\boldsymbol{\tau}_x(r_k D_{2^k})\right) = 0$$

for every $n, k \in \mathbf{N}$. But Theorem 8 in 1.5 implies

$$\begin{aligned}|(S_n g)(x)| &= \left|\sum_{k=0}^{\infty}(n_k(gw_n) * (r_k D_{2^k}))(x)\right| \\ &= \left|\sum_{k=0}^{\infty} n_k \sum_{I\in\mathcal{I}_*}\frac{1}{|I|}\left(\int_I gw_n\right)(\chi(I) * r_k D_{2^k})(x)\right|.\end{aligned}$$

Consequently,

$$\sup_{n\geq m}|(S_n g)(x)| \leq T_m(x)$$

for every $x \in \overline{E}$ and $m \in \mathbf{N}$. We conclude by (8) that $S_n g \to 0$ a.e., as $n \to \infty$, on $\overline{E}$. ∎

For $f \in \mathrm{L}^1$ and $x \in [0,1)$, define

$$f^{\spadesuit}(x) := \sup_{n\in\mathbf{N}}\frac{n+1}{|I_n(x)|}\int_{I_n(x)}|f(t) - f(x)|\,dt.$$

The following result, which we shall call the Marcinkiewicz test, shows that if $f^{\spadesuit}$ belongs to L^0 then the Walsh-Fourier series of f converges a.e.

THEOREM 2. *If $f \in \mathrm{L}^1$ then Sf converges a.e. on the set $\{f^{\spadesuit} < \infty\}$.*

PROOF. Fix $N \in \mathbf{N}$ and set

$$E_0 := \{f^{\spadesuit} > N\}.$$

It suffices to show that Sf converges a.e. on the set $\overline{E}_0 := [0,1) \setminus E_0$.

We may suppose that $|E_0| < 1$. Choose $\varepsilon > 0$ such that $|E_0| + \varepsilon < 1$ and apply Lemma 1 to choose a collection $\mathcal{I}_*$ of pairwise disjoint dyadic intervals such that if $E = \bigcup_{I\in\mathcal{I}_*} I$ then $E_0 \subseteq E$, $|E \setminus E_0| < \varepsilon$, and (5) holds. In particular, for each $I \in \mathcal{I}_*$ choose a point

$$x_I \in I_+ \cap \overline{E}_0.$$

Let

$$g := \sum_{I\in\mathcal{I}_*}\chi(I)(f - f(x_I))$$

and

$$h := \sum_{I \in \mathcal{I}_*} f(x_I)\chi(I) + \chi(\overline{E})f$$

where $\overline{E} := [0,1) \setminus E$. Notice that $f = g + h$.

We shall prove that g satisfies the hypotheses of Lemma 2 and that h is bounded. It will follow from Lemma 2 and Theorem 14 in 3.7 that $Sg = 0$ a.e. on $\overline{E}$ and $Sh = h$ a.e. on [0,1). Since by Lemma 1, $|\overline{E}_0 \setminus \overline{E}| < \varepsilon$, the proof of the theorem will be complete.

Clearly $g \in \mathrm{L}^1$ and $\{g \neq 0\} \subseteq E$. Moreover, for each $I \in \mathcal{I}_*$, the condition $x_I \in I_+ \cap \overline{E}_0$ and the definition of E_0 imply

$$\begin{aligned}\frac{1}{|I|}\int_I |g| &= \frac{1}{|I|}\int_I |f(t) - f(x_I)|\,dt \\ &\leq \frac{2}{|I_+|}\int_{I_+} |f(t) - f(x_I)|\,dt \\ &\leq \frac{2N}{\ell(I_+)} \leq \frac{4N}{\ell(I)}.\end{aligned}$$

Consequently, the hypotheses of Lemma 2 are satisfied with $c' = 4N$.

To prove h is bounded, observe for any two points $x, y \in [0,1)$ that there exist $x', y' \in \overline{E}_0$ such that $h(x) = f(x')$ and $h(y) = f(y')$. Since $E_0 \neq [0,1)$ we have for all $x, y \in [0,1)$ that

$$\begin{aligned}|h(x) - h(y)| &\leq \int_0^1 |f(x') - f(u)|\,du + \int_0^1 |f(y') - f(u)|\,du \\ &\leq f^\spadesuit(x') + f^\spadesuit(y') \\ &\leq 2N. \quad \blacksquare\end{aligned}$$

It is interesting to note that one can show directly, without using Theorem 14 in 3.7, that Sh converges a.e. (see Exercise 6.18).

In 4.2 we saw that $\|S_n f\|_\infty / \log n \to 0$ as $n \to \infty$ for $f \in C_W$. Using the maximal function $\mathcal{E}^\natural$ (see (8) in 3.1), we can obtain a similar estimate for pointwise growth of partial sums of Walsh-Fourier series.

THEOREM 3. *If $f \in \mathrm{L}^1$ then*

$$\sup_{m \geq 2} \frac{|S_m f|}{\log_2 m} \leq 2\mathcal{E}^\natural f$$

and

$$\lim_{m \to \infty} \frac{S_m f}{\log_2 m} = 0 \qquad \text{a.e.} \tag{9}$$

PROOF. For $m, n \in \mathbf{P}, 2^{n-1} \leq m < 2^n$, recall from Theorem 8 in 1.5 that

$$S_m f = w_m \sum_{k=0}^{n-1} m_k r_k \mathcal{E}_k(r_k w_m f).$$

Hence it follows from the definition of $\mathcal{E}^\natural$ that

$$|S_m f| \le n\mathcal{E}^\natural f.$$

Let

$$T_m f := \frac{S_m f}{\log_2 m}$$

for $m \in \mathbf{N}$, $m \ge 2$, and $f \in \mathrm{L}^1$. Clearly, $T_m f \le 2\mathcal{E}^\natural f$. Therefore the maximal operator of the operators T_m is of weak type (1,1). Moreover, it is clear that $T_m f \to 0$ as $m \to \infty$, for every polynomial $f \in \mathcal{P}$. We conclude by Theorem 2 in 3.1 that $T_m f \to 0$ a.e. as $m \to \infty$ for every $f \in \mathrm{L}^1$. ∎

We shall show in 6.5 that the results of this section cannot be appreciably improved.

6.2 Almost Everywhere Summability of Walsh-Fourier Series and the Pointwise Dyadic Derivative. In this section we show the Walsh-Fourier series of every $f \in \mathrm{L}^1$ is a.e. Cesàro summable. We shall also prove a fundamental theorem of calculus for the pointwise dyadic derivative.

Let

$$\mathrm{L}_0^1 := \{f \in \mathrm{L}^1 : \int_0^1 f = 0\}.$$

An operator $T : \mathrm{L}^1 \to \mathrm{L}^0$ is said to be local if $f \in \mathrm{L}_0^1$ and $\{f \ne 0\} \subseteq I$ for some dyadic interval I imply

$$\{Tf \ne 0\} \subseteq I.$$

An operator $T : \mathrm{L}^1 \to \mathrm{L}^0$ will be called quasi- local if $f \in \mathrm{L}_0^1$ and $\{f \ne 0\} \subseteq I$ for some dyadic interval I imply that Tf is integrable off I with

$$\int_{[0,1)\setminus I} |Tf| \le C\|f\|_1 \tag{10}$$

for some constant C which does not depend on I or f.

Examples of local operators include the conditional expectations operators

$$\mathcal{E}_n f := S_{2^n} f \qquad (n \in \mathbf{N}, f \in \mathrm{L}^1).$$

It is clear that every local operator is quasi-local. Translation operators are quasi-local but not local. Another class of quasi- local operators is given by

$$\mathcal{E}_i^j f(x) := \mathcal{E}_i f(x \dot{+} 2^{-j-1}) \qquad (i, j \in \mathbf{N})$$

defined for each $f \in \mathrm{L}^1$ and $x \in [0,1)$.

In this section we shall denote the operator norm of a $T : \mathrm{L}^p \to \mathrm{L}^p$ by $\|T\|_p$, i.e., if $1 \le p \le \infty$ then

$$\|T\|_p := \sup_{\|f\|_p \le 1} \|Tf\|_p.$$

For certain kinds of quasi-local operators, the condition $\|T\|_\infty < \infty$ is sufficient to conclude that T is of weak type (1,1) and takes H to L^1.

THEOREM 4. *For each $n \in \mathbf{N}$ let $T_n : \mathrm{L}^1 \to \mathrm{L}^1$ be a bounded linear operator and for each $f \in \mathrm{L}^1$ let*

$$Tf := \sup_{n \in \mathbf{N}} |T_n f|.$$

If T is quasi-local and $A := \|T\|_\infty < \infty$, then there is a constant B (depending only on A and the constant C of quasi-locality) such that

$$y|\{Tf > y\}| \leq B\|f\|_1 \qquad (y > 0) \tag{11}$$

and

$$\|Tf\|_1 \leq B\|f\|_{\mathrm{H}}.$$

PROOF. The second inequality is almost obvious. For, if β is a dyadic atom supported in an interval I then by (10) and hypothesis we have

$$\begin{aligned} \left| \int_0^1 T\beta \right| &\leq \int_I |T\beta| + \int_{[0,1)\setminus I} |T\beta| \\ &\leq A\|\beta\|_\infty |I| + C\|\beta\|_1 \\ &\leq A + C. \end{aligned}$$

To prove (11), fix $f \in \mathrm{L}^1$, $y > \|f\|_1$, and choose by the Calderon-Zygmund decomposition non-overlapping dyadic intervals $I_0, I_1, \ldots$ and functions g, h such that $f = g + h$,

$$\|g\|_\infty \leq 2y \tag{12}$$

$$|\Omega| := \left| \bigcup_{k=0}^{\infty} I_k \right| \leq \frac{1}{y}\|f\|_1 \tag{13}$$

and if $h_k := \chi(I_k)h$ then $h_k \in \mathrm{L}_0^1$ and

$$\|h_k\|_1 \leq 4y|I_k| \qquad (k \in \mathbf{N}). \tag{14}$$

Since $Tf \leq Tg + Th$, and (12) implies $\|Tg\|_\infty \leq 2Ay$, we have by (13) that

$$\begin{aligned} |\{Tf > (1 + 2A)y\}| &\leq |\{Th > y\}| \\ &\leq |\Omega| + |\{x \notin \Omega : (Th)(x) > y\}| \\ &\leq \frac{\|f\|_1}{y} + \frac{1}{y}\int_{\overline{\Omega}} Th, \end{aligned}$$

where $\overline{\Omega} := [0, 1) \setminus \Omega$.

Since T is quasi-local and Ω is a union of non-overlapping dyadic intervals, we have

$$\int_{\overline{\Omega}} Th_k \leq C\|h_k\|_1$$

for $k \in \mathbf{N}$. Moreover, it is clear that

$$Th \le \sum_{k\in\mathbf{N}} Th_k.$$

It follows, therefore, from (14) and (13) that

$$\int_{\Omega} Th \le 4Cy \sum_{k\in\mathbf{N}} |I_k| \le 4C\|f\|_1.$$

Consequently,

$$y|\{Tf > (1+2A)y\}| \le (1+4C)\|f\|_1.$$

Since this inequality also holds for $y \le \|f\|_1$, we conclude by a change of variables that (11) holds. In fact, we may take

$$B := (1+4C)(1+2A). \quad \blacksquare$$

We shall apply this theorem to the operators

$$\mathcal{F}_n := 2^{-n} \sum_{j=0}^{n-1} 2^j \sum_{i=j}^{n-1} \mathcal{E}_i^j$$

and

$$\mathcal{R}_n := \sum_{j=0}^{n-1} 2^j \sum_{i=n}^{\infty} 2^{-i} \mathcal{E}_i^j \qquad (n \in \mathbf{P}).$$

Our interest in these operators stems from their relationship with summability and dyadic differentiation. For example, Theorem 16 in 1.8 implies

$$(15) \qquad |\sigma_m f| \le |\mathcal{F}_n f| + 2\sum_{i=0}^{n-1} 2^{i-n} |\mathcal{E}_i f|$$

for $2^{n-1} \le m < 2^n$. (See also (26) below.)

LEMMA 3. *The operators $\mathcal{F}_n$ and $\mathcal{R}_n$ are bounded on L^p for $1 \le p \le \infty$, $n \in \mathbf{P}$ with*

$$(16) \qquad \|\mathcal{F}_n\|_p \le 2, \|\mathcal{R}_n\|_p \le 2.$$

Moreover, if $\mathcal{F}^* f := \sup_{n\in\mathbf{P}} |\mathcal{F}_n f|$ and $\mathcal{R}^* f := \sup_{n\in\mathbf{P}} |\mathcal{R}_n f|$, then $\mathcal{F}^*$ and $\mathcal{R}^*$ are quasi-local.

PROOF. Clearly, $\|\mathcal{E}_i^j\|_p = 1$ for $i, j \in \mathbf{N}$ and $1 \le p \le \infty$. Consequently,

$$\begin{aligned}\|\mathcal{F}_n\|_p &\le \sum_{j=0}^{n-1} (n-j) 2^{-n+j} \\ &\le \sum_{k=1}^{\infty} k 2^{-k} \\ &\le 2.\end{aligned}$$

Fix a dyadic interval I, let $m \in \mathbf{N}$ be determined by $|I| = 2^{-m}$ and suppose $f \in L^1_0$ satisfies

$$\{f \neq 0\} \subseteq I.$$

Fix $x \notin I$ and observe that $j > m - 1$ implies $x \notin \tau_{2^{-j-1}}(I)$. Thus it is easy to see that

$$\mathcal{E}^j_i f(x) = 0$$

for all $x \notin I$ if $j \geq m, i \in \mathbf{N}$ or if $i \leq m$ and $j \in \mathbf{N}$. In particular,

$$|\mathcal{F}_n f(x)| \leq \sum_{j=0}^{m-1} 2^j \sum_{i=m+1}^{\infty} 2^{-i} |\mathcal{E}^j_i f(x)|$$

for all $x \notin I$ and $n \in \mathbf{N}$. Since the right side of this inequality is independent of n, it follows that

$$\int_{[0,1)\setminus I} \mathcal{F}^* f \leq \|f\|_1 \sum_{j=0}^{m-1} 2^j \sum_{i=m+1}^{\infty} 2^{-i} \leq \|f\|_1.$$

We conclude that $\mathcal{F}^*$ is quasi-local. A similar argument works for $\mathcal{R}^*$. Hence the proof of the lemma is complete. ∎

Since (16) implies $\|\mathcal{F}^*\|_\infty \leq 2$ and $\|\mathcal{R}^*\|_\infty \leq 2$, the following result follows immediately from Lemma 3 and Theorem 4.

Corollary 1. *The operators $\mathcal{F}^*$ and $\mathcal{R}^*$ are of weak type (1,1) and satisfy*

$$\|\mathcal{F}^* f\|_1 \leq C\|f\|_{\mathrm{H}}, \|\mathcal{R}^* f\|_1 \leq C\|f\|_{\mathrm{H}}$$

for all $f \in \mathrm{H}$, where $C > 0$ is an absolute constant.

It is now easy to prove that every Walsh-Fourier series is a.e. Cesàro summable.

Corollary 2. *Let*

$$\sigma^* f := \sup_{n>0} |\sigma_n f|$$

for $f \in L^1$. Then σ^* is of weak type (1,1), satisfies

$$\|\sigma^* f\|_1 \leq C\|f\|_{\mathrm{H}} \qquad (f \in \mathrm{H}) \tag{17}$$

for some absolute constant C. Moreover,

$$\lim_{n\to\infty} \sigma_n f = f \qquad \text{a.e.}$$

for all $f \in L^1$.

Proof. It is clear by (15) that

$$\sigma^* f \leq \mathcal{F}^* f + 2\mathcal{E}^* f.$$

Consequently, σ^* satisfies (17) and is of weak type (1,1). Moreover, it is clear that $\sigma_n f \to f$ as $n \to \infty$, for all Walsh polynomials f. Therefore, it follows from Theorem 2 in 3.1 that $\sigma_n f \to f$ a.e. as $n \to \infty$, for all $f \in L^1$. ∎

To generalize these results from Walsh-Fourier series to Walsh-Fourier-Stieltjes series, we introduce the following notation.

Let $\mathbb{M}$ denote the collection of finite Borel measures on [0,1). Given $\nu \in \mathbb{M}$ let $\|\nu\|$ denote its total variation. Let

$$\mathbb{M}_0 := \{\nu \in \mathbb{M} : \nu([0,1)) = 0\}$$

and call an operator $T : \mathbb{M} \to L^0$ quasi-local if given $\nu \in \mathbb{M}_0$ supported on some dyadic interval I, one has that $T\nu$ is integrable off I and

$$\int_{[0,1)\setminus I} T\nu \leq C\|\nu\|$$

for an absolute constant C which depends neither on ν nor I.

Notice that any operator $T : \mathbb{M} \to L^0$ can be viewed as an operator from L^1 to L^0 by restricting T to the absolutely continuous measures in $\mathbb{M}$. Indeed, recall that the map $f \to \nu_f$ defined by

$$\nu_f(E) := \int_E f \qquad (E \in \mathcal{A})$$

is a 1-1 map from L^1 onto the collection of absolutely continuous Borel measures on [0,1) and that

$$\|\nu_f\| = \|f\|_1.$$

In particular, if $T : \mathbb{M} \to L^0$ is quasi-local, then its restriction to L^1 is also quasi-local.

An operator $T : \mathbb{M} \to L^0$ is said to be of weak type if there is an absolute constant B such that

$$y|\{T\nu > y\}| \leq B\|\nu\| \tag{18}$$

for all $y > 0$ and $\nu \in \mathbb{M}$. Notice that the restriction to L^1 of an operator $T : \mathbb{M} \to L^0$ of weak type is necessarily of weak-type (1,1). Thus the following result generalizes (11) of Theorem 4.

THEOREM 5. *For each $n \in \mathbf{N}$ let $T_n : \mathbb{M} \to L^1$ be a bounded linear operator. For each $\nu \in \mathbb{M}$ let*

$$T\nu := \sup_{n\in\mathbf{N}} |T_n\nu|.$$

If T is quasi-local and $C_1 := \|T\|_\infty < \infty$, then there is a constant B (depending only on C_1 and the constant C of quasi-locality) such that (18) holds for all $y > 0$ and all $\nu \in \mathbb{M}$.

PROOF. By the proof of Theorem 4, it suffices to obtain a Calderon-Zygmund decomposition for measures. Specifically, we shall show that given $\nu \in \mathbb{M}$ and $y > \|\nu\|$ there exist non-overlapping dyadic intervals $I_0, I_1, \ldots$ satisfying

$$|\Omega| := \left| \bigcup_{k=0}^{\infty} I_k \right| \leq \frac{1}{y}\|\nu\|$$

and measures λ, ϑ satisfying $\nu = \lambda + \vartheta$ such that $\lambda = \nu_g$ for some $g \in L^\infty$ with

$$\|g\|_\infty \leq 2y,$$

and such that if $\vartheta_k(E) := \vartheta(I_k \cap E)$ for $E \in \mathcal{A}$ and $k \in \mathbf{N}$ then

$$\vartheta = \sum_{k=0}^{\infty} \vartheta_k$$

in the norm of $\mathbb{M}$, and

$$\|\vartheta_k\| \leq 4y|I_k|$$

for all $k \in \mathbf{N}$. By the Jordan decomposition we may suppose that ν is non-negative.

Fix $\nu \geq 0$ and $y > \|\nu\|$, set $f_n := S_{2^n}\nu$ for $n \in \mathbf{N}$ and consider the stopping time

$$\tau := \min\{n \in \mathbf{N} : f_n > y\}.$$

For each $n \in \mathbf{N}$ set $g_n := f_{n\wedge\tau}$, $h_n := (f_n - f_{n\wedge\tau})$ and observe that

$$f_n = g_n + h_n.$$

Fix a dyadic interval I and set

$$\lambda(I) := \lim_{n\to\infty} \int_I g_n.$$

This limit exists because $\int_I w_j = 0$ for j sufficiently large. Hence λ belongs to QM. But $f_n \geq 0$ and

$$f_n = \mathcal{E}_n f_{n+1} \leq 2\mathcal{E}_{n-1} f_{n+1} = 2f_{n-1}.$$

Hence it follows from the definition of τ that

$$0 \leq g_n \leq 2y \qquad (n \in \mathbf{N}).$$

We have shown that $\|g_n\|_\infty \leq 2y$ for $n \in \mathbf{N}$. Since each g_n is a 2^n-th partial sum of a Walsh-Fourier-Stieltjes series and the Dirichlet kernel D_{2^n} is non-negative for $n \in \mathbf{N}$, we can repeat the proof of Theorem 8 ii) in 4.4 to show that there is a function $g \in L^\infty$ and a subsequence $n_1, n_2, \ldots$ of integers such that $g_{n_k} \to g$ as $k \to \infty$, a.e. and in L^2 norm. Consequently,

$$\lambda(I) = \int_I g$$

and

$$0 \leq g \leq 2y \qquad \text{a.e.}$$

Let $f^* := \sup_{n\in\mathbf{N}} f_n$. By construction,

$$\{f^* > y\} = \bigcup_{n\in\mathbf{N}} \{\tau = n\}.$$

Moreover, each set $\{\tau = n\}$ is a finite union of dyadic intervals whose lengths decrease as n increases. Hence we can choose non-overlapping dyadic intervals $I_0, I_1, \dots$ such that each I_k is contained in some $\{\tau = n\}$ and

$$\{f^* > y\} = \bigcup_{k=0}^{\infty} I_k.$$

Since the proof of Theorem 1 in 3.1 establishes

$$|\{f^* > y\}| \le \frac{1}{y} \sup_{n\in\mathbf{N}} \|f_n\|_1,$$

it follows that

$$|\Omega\,| := |\bigcup_{k=0}^{\infty} I_k\,| \le \frac{1}{y} \sup_{n\in\mathbf{N}} \|f_n\|_1 = \frac{1}{y}\|\nu\|.$$

Finally, for each dyadic interval I set

$$\vartheta(I) := \lim_{n\to\infty} \int_I h_n,$$
$$\vartheta_k(I) := \vartheta(I \cap I_k) \qquad (k \in \mathbf{N}).$$

Notice by construction that $\nu = \lambda + \vartheta$ and that ϑ is supported on Ω. Fix $k \in \mathbf{N}$ and use the Jordan decomposition to obtain

$$\|\vartheta_k\| \le |\vartheta_k|(I_k).$$

Since each I_k is contained in some $\{\tau = n\}$, it follows that

$$\|\vartheta_k\| \le \int_{I_k} f_n + \int_{I_k} g_n \le 4y|I_k|.$$

Consequently, $\sum \vartheta_k$ is Cauchy in $\mathbb{M}$ and converges in $\mathbb{M}$. Since ϑ is supported on Ω and $\Omega = \bigcup I_k$, we conclude that

$$\vartheta = \sum_{k=0}^{\infty} \vartheta_k$$

as required. ∎

Extend the operators $\mathcal{E}_i^j$ from L^1 to $\mathbb{M}$ by

$$\mathcal{E}_i^j \nu(x) := S_{2^i}\nu(x \dot{+} 2^{-j-1}).$$

Since (52) in 1.5 implies

$$\mathcal{E}_i^j \nu(x) = 2^i \nu(I_i(x \dot{+} 2^{-j-1})),$$

we have

$$\|\mathcal{E}_i^j \nu\|_1 \le \|\nu\|$$

for $\nu \in \mathbb{M}$ and $i, j \in \mathbf{N}$. Thus the operators $\mathcal{F}_n$ and $\mathcal{R}_n$ can be extended from L^1 to $\mathbb{M}$ and they satisfy

$$\|\mathcal{F}_n\nu\|_1 \le 2\|\nu\|, \|\mathcal{R}_n\nu\|_1 \le 2\|\nu\|$$

for $\nu \in \mathbb{M}$ and $n \in \mathbf{P}$. Analogous to (15) we have

$$|\sigma_m \nu| \le |\mathcal{F}_n\nu| + 2\sum_{i=0}^{n-1} 2^{i-n}|\mathcal{E}_i\nu| \tag{19}$$

for $2^{n-1} \le m < 2^n$. In particular,

$$\sigma^*\nu \le \mathcal{F}^*\nu + 2\mathcal{E}^*\nu,$$

for $\mathcal{F}^*\nu := \sup_n|\mathcal{F}_n\nu|$, $\mathcal{E}^*\nu := \sup_n|\mathcal{E}_n\nu|$, $\sigma^*\nu := \sup_n|\sigma_n\nu|$ and $\nu \in \mathbb{M}$.

Set $\mathcal{R}^*\nu := \sup_n|\mathcal{R}_n\nu|$. The argument of Lemma 3 shows that $\mathcal{F}^*$ and $\mathcal{R}^*$ are quasi-local on $\mathbb{M}$. Hence Theorem 5 contains the following result.

COROLLARY 3. *There exist absolute constants B and C such that*

$$y|\{\mathcal{E}^*\nu > y\}| \le B\|\nu\|,$$
$$y|\{\mathcal{F}^*\nu > y\}| \le B\|\nu\|,$$
$$y|\{\mathcal{R}^*\nu > y\}| \le B\|\nu\|,$$

and

$$y|\{\sigma^*\nu > y\}| \le C\|\nu\|,$$

for all $\nu \in \mathbb{M}$ and $y > 0$.

We shall use this to prove a summability result concerning Walsh-Fourier-Stieltjes series. First we consider the following case, using the notation

$$\nu_E(I) := \nu(I \cap E)$$

to represent the restriction of $\nu \in \mathbb{M}$ to a set $E \in \mathcal{A}$.

THEOREM 6. *For each $n \in \mathbf{N}$ let $T_n : \mathbb{M} \to L^1$ be a bounded, positive linear operator, and for each $\nu \in \mathbb{M}$ let*

$$T\nu := \sup_{n \in \mathbf{N}} |T_n\nu|.$$

If T is of weak type on $\mathbb{M}$ and if $T_n(\nu_J) \to 0$ a.e. on $[0,1) \setminus J$ as $n \to \infty$ for every singular measure ν and every dyadic interval J, then

$$\lim_{n \to \infty} T_n\nu = 0 \qquad \text{a.e.}$$

for every singular measure $\nu \in \mathbb{M}$.

PROOF. We begin by showing that if ν is a singular, non-negative Borel measure in $\mathbb{M}$ then its binary derivate is 0 a.e. on [0,1), i.e., that

$$\overline{\mathrm{D}}\nu(x) := \limsup_{n\to\infty} \frac{\nu(I_n(x))}{|I_n(x)|} = 0$$

for a.e. $x \in [0,1)$.

Indeed, let $\varepsilon > 0$. Since ν is singular, choose a Borel set E such that $\nu(E) = 0$ and $|[0,1) \setminus E| = 0$. Fix $j \in \mathbf{P}$, set

$$E_j := \left\{ x \in E : \limsup_{n\to\infty} 2^n \nu(I_n(x)) > \frac{1}{j} \right\}$$

and choose by regularity of ν an open set $V \supset E$ such that $\nu(V) < \varepsilon$. By definition, given $x \in E_j$ there is an integer $n(x)$ so large that $I_{n(x)}(x) \subset V$ and

$$|I_{n(x)}(x)| < j\nu(I_{n(x)}(x)).$$

Since $\{I_{n(x)}(x) : x \in E_j\}$ is a collection of dyadic intervals which covers E_j we can choose non-overlapping dyadic intervals $(J_k, k \in \mathbf{P})$ such that

$$E_j \subseteq \bigcup_{k=1}^{\infty} J_k \subseteq V$$

and

$$|J_k| < j\nu(J_k) \qquad (k \in \mathbf{P}).$$

Therefore,

$$|E_j| \le \sum_{k=1}^{\infty} |J_k| \le j \sum_{k=1}^{\infty} \nu(J_k) \le j\nu(V) < j\varepsilon.$$

In particular, $|E_j| = 0$ for all $j \in \mathbf{P}$ and it follows that $\overline{\mathrm{D}}\nu = 0$ a.e. on [0,1).

To prove the theorem, we may suppose that $\nu \ge 0$. By hypothesis, then, $T_n\nu \ge 0$ for $n \in \mathbf{N}$. Recall from (52) in 1.5 that

$$(\mathcal{E}_n\nu)(x) = \frac{\nu(I_n(x))}{|I_n(x)|} \qquad (n \in \mathbf{N}, x \in [0,1)).$$

By what we just showed, given $\varepsilon > 0$ there exists an $N \in \mathbf{N}$ such that

$$|\{\sup_{n\ge N} \mathcal{E}_n\nu > \varepsilon^2\}| < \varepsilon.$$

Let

$$\Omega := \{\sup_{n\ge N} \mathcal{E}_n\nu > \varepsilon^2\}$$

and observe that Ω is a union of non-overlapping dyadic intervals belonging to $\bigcup_{n=N}^{\infty} \mathcal{A}^n$. Consequently, if $E := [0,1) \setminus \overline{E}$ where $\overline{E}$ is defined by

$$\overline{E} := \bigcup \{I(k,N) : 0 \le k < 2^N, \text{ and } I(k,N) \setminus \Omega \ne \emptyset\}$$

then $E \subset \Omega$ and if $I(k,N) \subseteq \overline{E}$ then $I(k,N) = I_N(x)$ for some $x \in [0,1) \setminus \Omega$. It follows that

$$\varepsilon^2 \ge (\mathcal{E}_N \nu)(x) = 2^N \nu(I(k,N)).$$

In particular,

$$\nu(\overline{E}) \le \sum_{I(k,N) \subseteq \overline{E}} \nu(I(k,N)) \le \varepsilon^2.$$

Write $\nu = \nu_E + \nu_{\overline{E}}$ and observe that

$$\limsup_{n\to\infty} T_n \nu \le \limsup_{n\to\infty} T_n \nu_E + T\nu_{\overline{E}}.$$

By hypothesis, $\lim_{n\to\infty} T_n \nu_E = 0$ a.e. on $\overline{E}$. Since T is of weak type, it follows that

$$\begin{aligned} |\{\limsup_{n\to\infty} T_n \nu > \varepsilon\}| &\le |E| + |\{x \in \overline{E} : T\nu_{\overline{E}}(x) > \varepsilon\}| \\ &\le \varepsilon + \frac{C\|\nu_{\overline{E}}\|}{\varepsilon} \\ &\le (1+C)\varepsilon. \end{aligned}$$

Therefore, $T_n \nu \to 0$ a.e. as $n \to \infty$. ∎

Corollary 4. *If ν is a singular measure in* $\mathbb{M}$ *then $\mathcal{F}_n \nu \to 0$ and $\mathcal{R}_n \nu \to 0$ a.e. as $n \to \infty$.*

Proof. Since $\mathcal{F}^*$ and $\mathcal{R}^*$ are of weak type (see Corollary 3), it suffices to show that $\mathcal{F}_n(\nu_J)$ and $\mathcal{R}_n(\nu_J)$ converge to zero a.e. on $[0,1) \setminus J$, as $n \to \infty$, for every singular ν and every dyadic interval J.

Fix such ν and J, and let $i, j \in \mathbf{N}$. Since $\mathcal{E}_i \nu \to 0$ a.e., as $i \to \infty$, we have for each j that $\mathcal{E}_i^j \nu(x) \to 0$ a.e. on [0,1) as $i \to \infty$.

Let $|J| = 2^{m-1}$ and fix $x \in [0,1) \setminus J$. Notice that $\nu_J(I) = 0$ if $I \cap J = \emptyset$. Consequently,

$$\mathcal{E}_i^j \nu(x) = 2^{-i} \nu(I_i(x \dotplus 2^{-j-1})) = 0$$

for $i, j \ge m$. In particular, for $n \ge m$ we have

$$\mathcal{F}_n \nu_J(x) = \sum_{j=0}^{m-1} 2^{j-n} \sum_{i=j}^{n-1} \left(\mathcal{E}_i^j \nu\right)(x).$$

It follows, therefore, that $\mathcal{F}_n \nu_J \to 0$ a.e. on $[0,1) \setminus J$ as $n \to \infty$.

Similarly, we can show that

$$\mathcal{R}_n \nu_J(x) = \sum_{j=0}^{m-1} 2^j \sum_{i=n}^{\infty} 2^{-i} \left(\mathcal{E}_i^j \nu\right)(x)$$

for $x \in [0,1) \setminus J$. Since the inner sum tends to zero, as $n \to \infty$, for a.e. $x \in [0,1)$ and all $j \in \mathbf{N}$, it follows that $\mathcal{R}_n \nu_J \to 0$ a.e. on $[0,1) \setminus J$ as $n \to \infty$. ∎

COROLLARY 5. *Let $\nu \in \mathbb{M}$. If f is the Radon-Nikodym derivative of the absolutely continuous part of ν, then*

$$\lim_{n\to\infty} \sigma_n \nu = f \qquad \text{a.e.}$$

PROOF. There exists a singular measure λ and a function $f \in \mathrm{L}^1$ such that $\nu = \nu_f + \lambda$. By Corollary 2, $\sigma_n(\nu_f) \to f$ a.e., as $n \to \infty$. By Corollary 4 and (19), $\sigma_n(\lambda) \to 0$ a.e. Therefore,

$$\begin{aligned}\lim_{n\to\infty} \sigma_n \nu &= \lim_{n\to\infty} \sigma_n(\nu_f) + \lim_{n\to\infty} \sigma_n(\lambda)\\ &= f \qquad \text{a.e.} \quad \blacksquare\end{aligned}$$

We pass from summability to dyadic differentiation.

Recall that the dyadic integral $\mathbf{I}$ of an $f \in \mathrm{L}^1$ is defined by $\mathbf{I}f := f * W$ for $f \in \mathrm{L}^1$, where

$$W = 1 + \sum_{k=1}^{\infty} \frac{w_k}{k}.$$

For each $\nu \in \mathbb{M}$ define the dyadic integral $\mathbf{I}\nu$ by

$$\mathbf{I}\nu := \widehat{\nu}(0) + \lim_{n\to\infty} \sum_{k=1}^{2^n-1} \frac{\widehat{\nu}(k)}{k} w_k,$$

where this limit is taken in the L^1 norm. Notice that this series does indeed converge in L^1 norm, for by Theorem 17 in 1.8,

$$\Big\| \sum_{n=2^m}^{\infty} \frac{w_n}{n} \Big\|_1 = O(2^{-m}) \qquad \text{as} \qquad m \to \infty,$$

and if $(g * \nu)(x) := \int_0^1 g(x \dot{+} y)\, d\nu(y)$ then

$$\|g * \nu\|_1 \le \|g\|_1 \|\nu\|$$

for every $g \in \mathrm{L}^1$. Notice also that this definition extends $\mathbf{I}$ from L^1 to $\mathbb{M}$.

Recall that the pointwise dyadic derivative of a function f at $x \in [0,1)$ is defined by

$$f^{[1]}(x) := \lim_{n\to\infty} \mathbf{d}_n f(x)$$

where

$$\mathbf{d}_n f(x) := \sum_{j=0}^{n-1} 2^{j-1}(f(x) - f(x \dot{+} 2^{-j-1}))$$

for $x \in [0,1)$ and $n \in \mathbf{P}$. Also recall that

$$\mathbf{d}_n\,(w_{i2^n+k}) = k w_{i2^n+k} \tag{20}$$

$0 \le k < 2^n$ and $i, n \in \mathbf{N}$ (see (69) in 1.7).

Using the techniques developed above, we shall show that $\mathbf{I}\nu$ is a.e. dyadically differentiable and

$$\lim_{n\to\infty} \mathbf{d}_n(\mathbf{I}\nu) = f \qquad \text{a.e.}$$

for each $\nu \in \mathbb{M}$, where f is the Radon-Nikodym derivative of the absolutely continuous part of ν. We first consider the case when $\nu = \nu_f$ is absolutely continuous and notice that

$$\mathbf{d}_n(\mathbf{I}f) = f * (\mathbf{d}_n W) \qquad (n \in \mathbf{P}). \tag{21}$$

By (20) write

$$\begin{aligned}\mathbf{d}_n W &= D_{2^n} + \sum_{i=1}^{\infty}\sum_{k=1}^{2^n-1} \frac{\mathbf{d}_n(w_{i2^n+k})}{i2^n+k} \\ &= D_{2^n} + \sum_{i=1}^{\infty} \frac{w_{i2^n}}{i^2} \sum_{k=1}^{2^n-1} \frac{i^2 k w_k}{i2^n+k}.\end{aligned}$$

Thus with

$$b_k(i,n) := \frac{ik}{2^n} - \frac{i^2 k}{i2^n+k}$$

for $i, n \in \mathbf{P}, 0 \le k < 2^n$, one obtains

$$\mathbf{d}_n W = D_{2^n} + \sum_{i=1}^{\infty} \frac{w_{i2^n}}{i} 2^{-n} \sum_{k=1}^{2^n-1} k w_k - \sum_{i=1}^{\infty} \frac{w_{i2^n}}{i^2} \sum_{k=1}^{2^n-1} b_k(i,n) w_k.$$

In particular, if

$$U_n := 2^{-n} \left(\sum_{k=1}^{2^n-1} k w_k\right)\left(\sum_{i=1}^{\infty} \frac{w_{i2^n}}{i}\right)$$

and

$$V_n := \sup_{i\in\mathbf{P}} \left| \sum_{k=1}^{2^n-1} b_k(i,n) w_k \right|$$

then

$$|\mathbf{d}_n W| \le D_{2^n} + |U_n| + 2V_n \tag{22}$$

for $n \in \mathbf{P}$.

The operators $U_n * f$ and $V_n * f$ can be estimated using the operators $\mathcal{R}^*$ and σ^*. In fact, LEMMA 4. *If* $n \in \mathbf{P}$ *and* $x \in [0,1)$ *then*

$$|U_n(x)| \le 2\sum_{j=0}^{n-1} 2^{j+1} \sum_{i=n}^{\infty} 2^{-i} D_{2^i}(x \dotplus 2^{-j-1}) + 2\sum_{k=0}^{\infty} 2^{-k+1} D_{2^{n+k}}(x) \tag{23}$$

and

$$V_n \le 4 \cdot 2^{-n} \sum_{k=1}^{2^n} |K_k| + 2K_{2^n} + D_{2^n}. \tag{24}$$

PROOF. For (23) observe by definition that

$$\begin{aligned} \Big| \, 2^{-n} \sum_{k=0}^{2^n-1} k w_k(x) \, \Big| &= 2^{-n} |(\mathbf{d}_n D_{2^n})(x)| \\ &\le 2^{-n} \sum_{j=0}^{n-1} 2^{j-1} \left(D_{2^n}(x) + D_{2^n}(x \dotplus 2^{-j-1}) \right) \\ &\le 2^{-n-1} \sum_{j=0}^{n-1} 2^j D_{2^n}\left(x \dotplus 2^{-j-1}\right) + \frac{1}{2} D_{2^n}(x). \end{aligned}$$

Furthermore, by Theorem 15 in 1.7,

$$\begin{aligned} \Big| \sum_{i=1}^{\infty} \frac{w_{i2^n}(x)}{i} \Big| &= \Big| \sum_{i=1}^{\infty} \frac{w_i(2^n x)}{i} \Big| \\ &\le 4 \sum_{k=0}^{\infty} 2^{-k} D_{2^k}(2^n x). \end{aligned}$$

Since $j < n$ implies

$$\begin{aligned} D_{2^n}\left(x \dotplus 2^{-j-1}\right) D_{2^k}(2^n x) &= \prod_{\ell=0}^{n-1} \left(1 + r_\ell(x \dotplus 2^{-j-1})\right) \prod_{\ell=0}^{k-1} \left(1 + r_{\ell+n}(x \dotplus 2^{-j-1})\right) \\ &= D_{2^{n+k}}\left(x \dotplus 2^{-j-1}\right), \end{aligned}$$

(23) follows immediately from these two inequalities.

For (24), fix $i, n \in \mathbf{P}$, set $b_k := b_k(i,n)$, and use Abel's transformation twice to write

$$\sum_{k=0}^{2^n-1} b_k w_k = \left\{ \begin{array}{l} b_{2^n-1} D_{2^n} + (b_{2^n-2} - b_{2^n-1})(2^n-1) K_{2^n-1} \\ \qquad + \sum_{k=1}^{2^n-2} (b_{k-1} - 2b_k + b_{k+1})\, k K_k. \end{array} \right. \tag{25}$$

Let $g(t) := t^2/(1+t/i)$ for $t \ge 0$ and notice that

$$b_k = g\left(k 2^{-n}\right)$$

for $0 \le k < 2^n$. It is easy to show that

$$g(t) = it - i^2 + \frac{i^2}{1+t/i},$$
$$\frac{d}{dt}g(t) = \frac{2t+t^2/i}{(1+t/i)^2},$$

and

$$\frac{d^2}{dt^2}g(t) = 2\left(1+t/i\right)^{-3}.$$

Consequently, $0 \le g(t) \le 1$, $0 \le \frac{d}{dt}g(t) \le 2$, and $0 \le \frac{d^2}{dt^2}g(t) \le 2$ for all $t \in [0,1]$. In particular, it follows from the mean value theorem that

$$|b_{2^n-2} - b_{2^n-1}| \le 2 \cdot 2^{-n}$$

and

$$|b_{k-1} - 2b_k + b_{k+1}| \le 4 \cdot 2^{-2n}$$

for $0 \le k < 2^n$. These estimates together with (25) establish (24). ∎

For classical Fourier analysis, the Hardy-Littlewood maximal function is defined by

$$Mf(x) := \sup_{h \ne 0} \left| \frac{1}{h} \int_x^{x+h} f(t)\,dt \right|.$$

Its dyadic counterpart is given by

$$\mathbf{I}^* f := \sup_{n \in \mathbf{P}} |\mathbf{d}_n(\mathbf{I}f)|.$$

The following result will be called the dyadic Hardy-Littlewood maximal inequality.

COROLLARY 6. *The operator* $\mathbf{I}^*$ *is of weak type (1,1) and of weak type on* $\mathbb{M}$.

PROOF. By (21), (22), and Lemma 4 we have

$$\mathbf{I}^* f \le 7\mathcal{E}^*|f| + 12\sigma^*|f| + 2\mathcal{R}^*|f|. \tag{26}$$

Since the operators $\mathcal{E}^*$, σ^*, and $\mathcal{R}^*$ are of weak type (1,1) and of weak type on $\mathbb{M}$, the same is true for $\mathbf{I}^*$. ∎

Let P be a Walsh polynomial and observe that $\mathbf{d}_n(\mathbf{I}P) = P$ for n sufficiently large. Therefore, Corollary 6 above and Theorem 2 in 3.1 provide a fundamental theorem of pointwise dyadic calculus:

COROLLARY 7. *For every* $f \in \mathrm{L}^1$,

$$\lim_{n \to \infty} \mathbf{d}_n(\mathbf{I}f) = f \qquad \text{a.e..}$$

With only a little more work, we establish the following.

COROLLARY 8. *Let $\nu \in \mathbb{M}$. If the Radon-Nikodym derivative of the absolutely continuous part of ν is f, then*

$$\lim_{n\to\infty} \mathbf{d}_n(\mathbf{I}\nu) = f \qquad \text{a.e.}$$

PROOF. Since $\nu = \nu_f + \lambda$ for an $f \in L^1$ and a singular measure $\lambda \in M$, it suffices by Corollary 7 to show that

$$\lim_{n\to\infty} \mathbf{d}_n(\mathbf{I}\nu) = 0 \qquad \text{a.e.}$$

for any singular measure $\nu \in \mathbb{M}$.

Fix such a measure ν with $\nu \geq 0$. By the definition of $\mathbf{I}\nu$ we can choose a subsequence $N_1, N_2, \ldots$ of positive integers such that

$$\mathbf{d}_n(\mathbf{I}\nu) = \lim_{k\to\infty} \nu * \mathbf{d}_n W_{2^{N_k}} \qquad \text{a.e.,} \tag{27}$$

where

$$W_n := 1 + \sum_{j=1}^{n-1} \frac{w_j}{j}$$

for $n \in \mathbf{P}$. By repeating the steps which lead to (22) with $W_{2^{N_k}}$ in place of W we obtain an analogue of (22) for each $k \in \mathbf{P}$. Letting $k \to \infty$ and applying (27) we conclude that

$$|\mathbf{d}_n(\mathbf{I}\nu)| \leq \nu * D_{2^n} + \nu * |U_n| + 2\nu * V_n.$$

Consequently, it follows from Lemma 4 and Corollaries 4 and 5 that

$$\lim_{n\to\infty} \mathbf{d}_n(\mathbf{I}\nu) = 0$$

for every singular measure ν. ∎

6.3 Logarithm Spaces and Block Spaces. The fact that the maximal operator S^* is of type (p,p) for $1 < p < \infty$ has many consequences for pointwise convergence of Fourier series. We shall investigate some of these consequences here, in 6.4, and in 6.6.

If $f \in L^p$ for $p > 1$ then Sf converges a.e., but Sf may diverge everywhere when $f \in L^1$. In this section we shall identify some spaces $\boldsymbol{X}$ which satisfy

$$L^p \subset \boldsymbol{X} \subset L^1 \qquad (p > 1)$$

such that Sf converges a.e. when $f \in \boldsymbol{X}$.

We begin with an inequality which relates the maximal martingale transform for martingale trees to the predictor $\mathcal{E}^\sharp$ (for definitions see (75),(85), and (86) in 3.7).

LEMMA 5. *Let $E \subseteq [0,1)$ be a measurable set, let $f \in L^0$ with range in $[1/2,1)$ and set*

$$g := \chi(E)f.$$

There is an absolute constant $C > 0$, independent of E and f such that

$$|\{\boldsymbol{T}^*(\boldsymbol{\alpha})g > Cy\}| \leq C|\{\mathcal{E}^\sharp g > y/\log_2(1/y)\}| \tag{28}$$

for all $0 < y < 1/2$ *and* $\alpha \in \boldsymbol{A}$ *with* $\|\alpha\| \le 1$.

PROOF. Fix $\boldsymbol{\alpha} \in \boldsymbol{A}$ with $\|\boldsymbol{\alpha}\| \le 1$, and for each integer dyadic interval $I \in \mathcal{J}$ and each $k \in \mathbf{Z}$ set

$$R_I^k g := \beta_I^k \boldsymbol{T}_I^*(\boldsymbol{\alpha\epsilon}^k)g,$$

where β_I^k and $\boldsymbol{\epsilon}^k$ are defined by (89) and (93) in 3.7 except f has been replaced by g and λ_J has been replaced by $\mathcal{E}_J^\sharp g$. Set

$$R^k g := \sup_{I \in \mathcal{J}} R_I^k g + 2^{k+1}\chi\{\mathcal{E}^\sharp g > 2^k\}$$

and observe by (90) and (94) in 3.7 that

$$\boldsymbol{T}^*(\alpha)g \le 2\sum_{k\in\mathbf{Z}} R^k g. \tag{29}$$

Fix $I \in \mathcal{J}$ and recall that the operator $\boldsymbol{T}_I^*$ is a linear martingale transform. Hence by Corollary 4 in 3.3 and the estimate of C_p which follows it, there is an absolute constant $c_1 > 0$ such that

$$\begin{aligned}\|R_I^k g\|_q^q &= \|R_I^k(\beta_I^k g)\|_q^q \\ &\le c_1 q^q \|\beta_I^k \boldsymbol{T}_I^*(\epsilon^k)g\|_q^q \\ &\le c_1 2^{q(k+2)} q^q \|\beta_I^k\|_q^q\end{aligned}$$

for all $q > 0$ and $k \in \mathbf{Z}$. Using the exponential Taylor series, we find that

$$\int_0^1 \beta_I^k \exp\left(\frac{R_I^k g}{3\cdot 2^k}\right) \le c_1 \sum_{q=0}^{\infty} \frac{1}{q!}\left(\frac{q}{3}\right)^q \|\beta_I^k\|_1 \le c_2\|\beta_I^k\|_1$$

for some constant $c_2 > 0$. In particular,

$$|\{R_I^k g > 6|k|2^k\}| \le c_2 e^{-2|k|}\|\beta_I^k\|_1$$

for all $k \in \mathbf{Z}, I \in \mathcal{J}$, and $\alpha \in \boldsymbol{A}$ with $\|\alpha\| \le 1$.

Let

$$E_0 := \bigcup_{k \le s_0} \{R^k g > 8|k|2^k\}$$

where s_0 is some negative number to be specified below. The estimate above shows that

$$\begin{aligned}|E_0| &\le \sum_{k<0} |\{R^k g > 8|k|2^k\}| \\ &\le \sum_{k<0}\sum_{I\in\mathcal{J}} |\{R_I^k g > 6|k|2^k\}| \\ &\le c_2 \sum_{k<0} e^{-2|k|} \sum_{I\in\mathcal{J}} \|\beta_I^k\|_1.\end{aligned}$$

Consequently, (88) in 3.7 with $p = q = 2$ implies

$$|E_0| \leq 4c_2 \sum_{k\leq 0} e^{-2|k|}|E|2^{-2k}.$$

In particular, there is an absolute constant $c_3 > 0$ such that

$$|E_0| \leq c_3|E|. \tag{30}$$

Fix $0 < y \leq 1/2$ and choose $s \in \mathbf{Z}$ with $s \leq -1$ so that $2^{s-1} < y \leq 2^s$. Set $s_0 := s - \log_2 |s|$ and observe that

$$\frac{y}{\log_2(1/y)} \leq 2^{s_0}.$$

We claim that

$$\overline{E}_0 \cap \{\mathcal{E}^\sharp g \leq 2^{s_0}\} \subseteq \{\boldsymbol{T}^*(\alpha)g \leq 192y\}$$

where $\overline{E}_0 := [0,1) \setminus E_0$. To see this observe by construction that $\chi\{\mathcal{E}^\sharp g \leq 2^k\}R^k g = 0$ for all $k \in \mathbf{Z}$. Hence by (29) we have

$$\chi\{\mathcal{E}^\sharp g \leq 2^{s_0}\}\boldsymbol{T}^*(\alpha)g \leq 2 \sum_{k\leq s_0} R^k g.$$

It follows, therefore, from the definition of E_0 that

$$\begin{aligned}
\chi\{\mathcal{E}^\sharp g \leq 2^{s_0}\}\chi(\overline{E}_0)\boldsymbol{T}^*(\alpha)g &\leq 16 \sum_{k\leq s_0} |k|2^k \\
&\leq 64|s_0|2^{s_0} \\
&\leq 64\, 2^s(|s| + \log_2 |s|)/|s| \\
&\leq 128\, 2^s \\
&\leq 256y.
\end{aligned}$$

The claim and the choice of s_0 imply

$$|\{\boldsymbol{T}^*(\alpha)g > 256y\}| \leq |E_0| + |\{\mathcal{E}^\sharp g > y/\log_2(1/y)\}|.$$

Since $y \leq 1/2$ the second term on the right side of this inequality is not less than $|E|$. Therefore, (28) follows from (30) with $C := \max\{256, c_3 + 1\}$. ∎

Combining this result with Corollary 1 in 3.1 and inequalities (77) and (87) in 3.7 gives the following estimate for the maximal partial sum operator. For the case $y > 1/2$ we have used Theorem 13 in 3.7.

COROLLARY 9. *There is an absolute constant $\widetilde{c} > 0$ such that*

$$|\{S^*(\chi(E)g) > y\}| \leq \widetilde{c}\,\frac{1}{y}|\log(\frac{1}{y})|\,|E| \tag{31}$$

for all measurable sets $E \subseteq [0,1)$, *measurable functions* $g : [0,1) \to [1/2,1]$, *and* $y > 0$.

Recall from 3.1 that

$$\log^+ t := \begin{cases} \log t & t \geq 1 \\ 0 & \text{otherwise.} \end{cases}$$

For each $0 < p < \infty$ let $\mathrm{L}(\log^+ \mathrm{L})^p$ represent the collection of functions $f \in \mathrm{L}^0$ which satisfy

$$\int_0^1 |f|(\log^+ |f|)^p < \infty,$$

and let $\mathrm{L}\log^+ \mathrm{L}\log^+ \log^+ \mathrm{L}$ represent the collection of $f \in \mathrm{L}^0$ which satisfy

$$\int_0^1 |f| \log^+ |f| \log^+ \log^+ |f| < \infty.$$

Clearly, $\mathrm{L}^p \subset \mathrm{L}(\log^+ \mathrm{L})^p \subset \mathrm{L}\log^+ \mathrm{L}\log^+ \log^+ \mathrm{L} \subset \mathrm{L}\log^+ \mathrm{L}$ for every $p > 1$. Hence the following result improves Theorem 14 of 3.7.

THEOREM 7. *If* $f \in \mathrm{L}\log^+ \mathrm{L}\log^+ \log^+ \mathrm{L}$ *then* Sf *converges to* f *a.e. on* $[0,1)$.

PROOF. Let

$$\varrho(f) := \left(\int_0^1 |f| \, (\log^+ |f| \log^+ \log^+ |f| + 1) \right)^{1/5}.$$

It suffices to show there is a constant $c > 0$ such that

$$(32) \qquad |\{S^* f > c\varrho(f)\}| \leq c\varrho(f)$$

for all f satisfying $\varrho(f) \leq 1/2$. Indeed, set

$$\Delta := \limsup_{n \to \infty} |S_n f - f|$$

and observe for any Walsh polynomial f_m that

$$\begin{aligned} \Delta &\leq \limsup_{n \to \infty} |S_n(f - f_m)| + |f_m - f| \\ &\leq S^*(f - f_m) + |f - f_m|. \end{aligned}$$

Since $\|g\|_1 \leq \varrho^5(g)$ for any $g \in \mathrm{L}^1$, it would follow from (32) that

$$\begin{aligned} |\{\Delta \geq 2c\varrho(f - f_m)\}| &\leq c\varrho(f - f_m) + \frac{\|f - f_m\|_1}{c\,\varrho(f - f_m)} \\ &\leq c\varrho(f - f_m) + \frac{1}{c}\varrho^4(f - f_m) \end{aligned}$$

for $\varrho(f - f_m) \leq 1/2$. But $\varrho(f - f_m) \to 0$ as $m \to \infty$ for $f_m := S_{2^m} f$ (see (1) in 4.1). Consequently, $|\{\Delta > 0\}| = 0$ and $S_n f \to f$ a.e. as $n \to \infty$.

To prove (32) we may suppose $f \geq 0$ and $\varrho(f) \neq 0$. Write

$$f = f_0 + \sum_{k=4}^{\infty} f_k \tag{33}$$

where $f_0 := \chi\{f < 8\}f$ and $f_k := \chi\{2^{k-1} \leq f < 2^k\}f$ for $k \geq 4$.

Since $f_0 \in \mathrm{L}^2$ it follows from Corollary 11 in 3.7 that

$$\begin{aligned} |\{S^* f_0 > \varrho(f)\}| &\leq c_5 \|f_0\|_2^2/\varrho^2(f) \\ &\leq 8c_5 \|f_0\|_1/\varrho^2(f) \\ &\leq 8c_6 \varrho^3(f) \end{aligned}$$

for some absolute constants c_5 and c_6. Therefore

$$|\{S^* f_0 > \varrho(f)\}| \leq 8c_6 \varrho(f). \tag{34}$$

On the other hand, if $A_k := \{2^{k-1} \leq f < 2^k\}$ and $B_k := \{S^* f_k > \varrho^2(f)\}$ then applying Corollary 9 to $g := 2^{-k} f_k$ results in

$$|B_k| \leq \widetilde{c}|A_k| \frac{2^k}{\varrho^2(f)} \left(k + 2\log_2\left(\frac{1}{\varrho(f)}\right)\right) \leq c_7 k 2^k \varrho^{-3}(f)\,|A_k|$$

for some absolute constant c_7. Set

$$B := \bigcup_{k=4}^{\infty} B_k.$$

It follows that

$$|B| \leq c_7 \varrho^{-3}(f) \sum_{k=4}^{\infty} k 2^k |A_k| \leq c_8 \varrho^5(f) \varrho^{-3}(f)$$

for some absolute constant c_8. Consequently,

$$|B| \leq c_8 \varrho(f). \tag{35}$$

Let $\overline{B} := [0,1) \setminus B$ and $\overline{B}_k := [0,1) \setminus B_k$ for $k \geq 4$. Set

$$\lambda_k(y) := |\{S^* f_k > y\}| = |\{S^* g_k > 2^{-k} y\}|$$

and apply Corollary 9 to obtain

$$\lambda_k(y) \leq \widetilde{c}\,|A_k| \frac{2^k}{y} \log\left(\frac{2^k}{y}\right).$$

Consequently, for each $k \geq 4$ we have

$$\begin{aligned}\int_{\overline{B}_k} S^* f_k &= \int_0^{\varrho^2(f)} \lambda_k(y)\,dy \\ &\leq \int_{\varrho^2(f)/k^2}^{\varrho^2(f)} \lambda_k(y)\,dy + \int_0^{\varrho^2(f)/k^2} dy \\ &\leq \widetilde{c}|A_k|2^k \int_{\varrho^2(f)2^{-k}/k^2}^{\varrho^2(f)2^{-k}} \frac{1}{y}\log_2\frac{1}{y}\,dy + \varrho^2(f)/k^2 \\ &\leq c_9|A_k|2^k k\log k\log(1/\varrho(f)) + \varrho^2(f)/k^2\end{aligned}$$

for some absolute constant c_9. Set

$$g := \sum_{k=4}^{\infty} f_k.$$

It follows that

$$\int_{\overline{B}} S^* g \leq \sum_{k=4}^{\infty} \int_{\overline{B}_k} S^* f_k \leq c_{10}\left(\varrho^5(f)\log(\frac{1}{\varrho(f)}) + \varrho^2(f)\right).$$

Therefore, there exists an absolute constant $c_{11} > 0$ such that

$$|\overline{B} \cap \{S^* g > \varrho(f)\}| \leq c_{11}\varrho(f).$$

We conclude by (34) and (35) that

$$|\{S^* f > 2\varrho(f)\}| \leq c_{12}\varrho(f). \qquad \blacksquare$$

This proof can be used to obtain the following estimate for the L^1 norm of the maximal function $S^* f$. Set

$$\varrho_1(f) := \left(\int_0^1 |f(x)|((\log^+|f(x)|)^2 + 1)\,dx\right)^{1/2}.$$

There is an absolute constant $c > 0$ such that

$$\|S^* f\|_1 \leq c\varrho_1(f) \tag{36}$$

for all $f \in L^0$ satisfying $\varrho_1(f) \leq 1/2$. Indeed, use decomposition (33) and estimate $S^* f_0$ directly (using Corollary 11 in 3.7) by

$$\|S^* f_0\|_1 \leq \|S^* f_0\|_2 \leq C_2\|f_0\|_2 \leq C_2\varrho_1(f).$$

To estimate $S^* f_k$, write

$$\|S^* f_k\|_1 = \left(\int_0^{\varrho_1(f)2^{-k}} + \int_{\varrho_1(f)2^{-k}}^{\varrho_1(f)2^k} + \int_{\varrho_1(f)2^k}^{\infty}\right)\lambda_k(y)\,dy$$

and proceed as before. We shall use this estimate in 6.4 to study a.e. convergence of Walsh-Kaczmarz-Fourier series.

Another class of spaces between $L^p, p > 1$, and L^1 for which Walsh-Fourier series converge a.e. is the collection of ***block spaces***. These are complete quasi-normed linear spaces whose quasi-norm is defined by means of the function $\mathcal{N} : \ell^1 \to [0, \infty]$ given by

$$\mathcal{N}(\boldsymbol{c}) := \sum{}' |c_k| \left(1 + \log \frac{\sum_{j=1}^{\infty} |c_j|}{|c_k|}\right)$$

for $\boldsymbol{c} = (c_k, k \in \mathbf{P})$, where $\sum'$ indicates summation over all indices k which satisfy $c_k \neq 0$. Clearly, if $0 \le c_k \le \widetilde{c}_k$ for $k \in \mathbf{P}$ and $\boldsymbol{c} := (c_k, k \in \mathbf{P}), \widetilde{\boldsymbol{c}} := (\widetilde{c}_k, k \in \mathbf{P})$ then $\mathcal{N}(\boldsymbol{c}) \le \mathcal{N}(\widetilde{\boldsymbol{c}})$.

The function $\mathcal{N}$ enjoys two other important properties:

LEMMA 6. *If $\boldsymbol{c}, \widetilde{\boldsymbol{c}} \in \ell^1$ and a is a real number then*

$$\mathcal{N}(a\boldsymbol{c}) = |a| \mathcal{N}(\boldsymbol{c})$$

and

(37) $$\mathcal{N}(\boldsymbol{c} + \widetilde{\boldsymbol{c}}) \le (1 + \log 2)(\mathcal{N}(\boldsymbol{c}) + \mathcal{N}(\widetilde{\boldsymbol{c}})).$$

PROOF. That $\mathcal{N}$ is positive homogeneous is obvious. To establish (37) we may suppose both $\mathcal{N}(\boldsymbol{c})$ and $\mathcal{N}(\widetilde{\boldsymbol{c}})$ are finite.

Set

$$\psi(x) := 1 - x \log x - (1 - x) \log(1 - x)$$

for $x \in (0, 1)$ and $\psi(0) := \psi(1) := 1$. Notice that ψ is non-negative and continuous on $[0, 1]$, taking its maximum at $x = 1/2$. Hence

(38) $$\max_{x \in [0,1]} \psi(x) = 1 + \log 2.$$

Let $\boldsymbol{c} = (c_k, k \in \mathbf{P}), \widetilde{\boldsymbol{c}} = (\widetilde{c}_k, k \in \mathbf{P})$, set $b := \|\boldsymbol{c}\|_{\ell^1}, \widetilde{b} := \|\widetilde{\boldsymbol{c}}\|_{\ell^1}$, and $x := b/(b + \widetilde{b})$. Then

$$\begin{aligned} \mathcal{N}(\boldsymbol{c} + \widetilde{\boldsymbol{c}}) &\le \mathcal{N}(|\boldsymbol{c}| + |\widetilde{\boldsymbol{c}}|) \\ &\le b + \widetilde{b} + \sum{}' |c_k| \log \left(\frac{b + \widetilde{b}}{|c_k|}\right) + \sum{}' |\widetilde{c}_k| \log \left(\frac{b + \widetilde{b}}{|\widetilde{c}_k|}\right) \\ &= (b + \widetilde{b}) \psi(x) + \sum{}' |c_k| \log \frac{b}{|c_k|} + \sum{}' |\widetilde{c}_k| \log \frac{\widetilde{b}}{|\widetilde{c}_k|}. \end{aligned}$$

Consequently, it follows from (38) that

$$\mathcal{N}(\boldsymbol{c} + \widetilde{\boldsymbol{c}}) \le (1 + \log 2)(\mathcal{N}(\boldsymbol{c}) + \mathcal{N}(\widetilde{\boldsymbol{c}})). \quad \blacksquare$$

Let $1 < q \le \infty$. A *dyadic q-block* is a function $\beta \in L^q$ which is supported on some dyadic interval I such that

(39) $$\|\beta\|_q \le |I|^{1/q - 1}.$$

(Compare with dyadic atoms defined in 3.4.) Observe by Hölder's inequality that

(40) $$\|\beta\|_1 \le 1$$

for any q-block β.

The maximal operator S^* satisfies a weak type $(1, 1)$ inequality on the block spaces (see Theorem 8 below). We first verify this for q-blocks.

LEMMA 7. *Let $1 < q \le \infty$. There is a constant C, depending only on q, such that*

$$|\{S^*\beta > y\}| \le C/y$$

for all q-blocks β and all $y > 0$.

PROOF. For simplicity we suppose $q \neq \infty$.

Fix $y > 0$ and a q-block β supported on a dyadic interval I. Observe since S^* commutes with translations that we may suppose $I = I(0, r)$ for some $r \in \mathbf{N}$.

If $1 < y|I|$ then $|I|^{1-q}/y^q \le 1/y$. Hence by Corollary 11 in 3.7 and (39) above we have

$$|\{S^*\beta > y\}| \le C_q \left(\frac{\|\beta\|_q}{y}\right)^q \le C_q \frac{|I|^{1-q}}{y^q} \le C_q \frac{1}{y}.$$

If $1 \ge y|I|$ and $\overline{I} := [0,1) \setminus I$ then

$$\begin{aligned}|\{S^*\beta > y\}| &\le |I| + |\overline{I} \cap \{S^*\beta > y\}| \\ &\le 1/y + |\{x \ge 2^{-r} : S^*\beta(x) > y\}|.\end{aligned}$$

Hence it suffices to show that the second term on the right side of this inequality is dominated by $4/y$.

Toward this, fix $2^{-r} \le x < 1$ and observe that $|x \dot{+} t| > x/2$ for all $t \in I$. Consequently, by Theorem 10 in 1.6 we have

$$\sup_{t \in I} |D_n(x \dot{+} t)| \le 4/x. \tag{41}$$

But β is supported on I. It follows, therefore, from (41) and (40) that

$$|(S_n\beta)(x)| = |\int_I D_n(x \dot{+} t)\beta(t)\, dt| \le \frac{4}{x}\, \|\beta\|_1 \le \frac{4}{x}.$$

In particular,

$$|\{x \ge 2^{-r} : (S^*\beta)(x) > y\}| \le |\{x > 0 : \frac{4}{x} > y\}| = \frac{4}{y}. \qquad \blacksquare$$

Block spaces are defined using q-blocks as atoms. Specifically, for $1 < q \le \infty$ let $\mathcal{B}_q$ represent the collection of functions $f \in \mathrm{L}^0$ of the form

$$f = \sum_{k=1}^{\infty} c_k \beta_k \tag{42}$$

where $\boldsymbol{c} = (c_k, k \in \mathbf{P})$ satisfies $\mathcal{N}(\boldsymbol{c}) < \infty$, and $\beta_1, \beta_2, \ldots$ are q-blocks. Such a series always converges in L^1 norm since $\mathcal{N}(\boldsymbol{c}) < \infty$ implies $\boldsymbol{c} \in \ell^1$. Consequently, $\mathcal{B}_q \subseteq \mathrm{L}^1$.

On the other hand, if $f \in L^q$ is non-zero then $\beta_1 := \|f\|_q^{-1} f$ is a q-block supported on $[0, 1)$. Hence f can be written in the form (42) with $\beta_k := 0$ for $k > 1$. Therefore, $\mathcal{B}_q \supseteq L^q$ for each $q > 1$.

These set inequalities are proper. In fact, for each $q > 1$ there exist functions $f \in \mathcal{B}_q$ which belong to none of the spaces L^p for $p > 1$. To see this, let

$$\beta_k(x) := \begin{cases} 2^k & 2^{-k} \leq x < 2^{-k+1} \\ 0 & \text{otherwise,} \end{cases}$$

and observe that β_k is a q-block for each $k \in \mathbf{P}$. The function

$$f := \sum_{k=1}^{\infty} \frac{1}{k^2} \beta_k$$

belongs to $\mathcal{B}_q$. However,

$$\|f\|_p^p = \sum_{k=1}^{\infty} \frac{1}{k^{2p}} 2^{k(p-1)} = \infty$$

for all $p > 1$.

For each $f \in \mathcal{B}_q$ set

$$\|f\|_{\mathcal{B}_q} := \inf\{\mathcal{N}(\mathbf{c})\}$$

where the infimum is taken over all $\mathbf{c}$ which satisfy (42) for some q-blocks $\beta_1, \beta_2 \ldots$ By Lemma 7 we have

$$\|af\|_{\mathcal{B}_q} = |a| \|f\|_{\mathcal{B}_q}$$

and

$$\|f + g\|_{\mathcal{B}_q} \leq (1 + \log 2)(\|f\|_{\mathcal{B}_q} + \|g\|_{\mathcal{B}_q})$$

for scalars a, functions $f, g \in \mathcal{B}_q$, and $1 < q \leq \infty$. Hence each $\mathcal{B}_q$ is a quasi-normed linear space which lies between $L^p, p > 1$, and L^1.

It is useful to note that the decompositions of type (42) converge a.e. for $f \in \mathcal{B}_q$. In fact,

LEMMA 8. *Let* $\mathbf{c} = (c_k, k \in \mathbf{P}) \in \ell^1, 1 < q \leq \infty$, *and* $\beta_1, \beta_2, \ldots$ *be a sequence of functions which satisfy*

$$|\{|\beta_k| > y\}| \leq \frac{1}{y} \qquad (y > 0). \tag{43}$$

If f is defined by (42) then

$$|\{|f| > y\}| \leq 3\mathcal{N}(\mathbf{c})/y$$

for all $y > 0$.

PROOF. Since $\mathcal{N}$ is positive homogeneous we may suppose that $\|\mathbf{c}\|_{\ell^1} = 1$. Since

$$|\{|\sum c_k \beta_k| > y\}| \leq |\{\sum |c_k| |\beta_k| > y\}|$$

we may also suppose that $c_k > 0$ and $\beta_k \geq 0$ for all $k \in \mathbf{P}$.

Fix $y > 0$. For $x \in [0,1)$ and $k \in \mathbf{P}$ define

$$\epsilon_k(x) := \begin{cases} \beta_k(x) & \beta_k(x) > y/(2c_k) \\ 0 & \text{otherwise} \end{cases}$$

$$\lambda_k(x) := \begin{cases} \beta_k(x) & \beta_k(x) < y/2 \\ 0 & \text{otherwise} \end{cases}$$

and

$$\tau_k(x) := \beta_k(x) - \epsilon_k(x) - \lambda_k(x).$$

Set $\epsilon := \sum c_k \epsilon_k$, $\lambda := \sum c_k \lambda_k$, and $\tau := \sum c_k \tau_k$ and observe since these series converge in L^1 norm that $f = \epsilon + \lambda + \tau$. Hence

$$|\{f > 2y\}| \leq |\{\epsilon > y/2\}| + |\{\lambda > y/2\}| + |\{\tau > y\}|. \tag{44}$$

To estimate ϵ notice by definition that

$$\begin{aligned} |\{\epsilon > y/2\}| &\leq \sum_{k=1}^{\infty} |\{c_k \epsilon_k > y/2\}| \\ &= \sum_{k=1}^{\infty} |\{\beta_k > y/(2c_k)\}|. \end{aligned}$$

From (43) it follows that

$$|\{\epsilon > y/2\}| \leq \sum_{k=1}^{\infty} \frac{2c_k}{y} \leq 2\frac{\mathcal{N}(\mathbf{c})}{y}.$$

To estimate τ observe that

$$\int_0^1 \tau = \sum_{k=1}^{\infty} c_k \int_0^1 \tau_k = \sum_{k=1}^{\infty} c_k \int_0^{\infty} (-t)\, d|\{\tau_k > t\}|.$$

By the definition of τ_k, integration by parts leads to

$$\begin{aligned} \int_0^{\infty} (-t)\, d|\{\tau_k > t\}| &= \int_{y/2}^{y/(2c_k)} (-t)\, d|\{\beta_k > t\}| \\ &= \int_{y/2}^{y/(2c_k)} |\{\beta_k > t\}| - \frac{y}{2c_k}|\{\beta_k > y/(2c_k)\}| + \frac{y}{2}|\{\beta_k > y/2\}|. \end{aligned}$$

Consequently, it follows from (43) that

$$\begin{aligned} \int_0^1 \tau &\leq \sum_{k=1}^{\infty} c_k \left(\int_{y/2}^{y/(2c_k)} \frac{1}{t}\, dt + 1 \right) \\ &= \sum_{k=1}^{\infty} c_k \left(\log(y/(2c_k)) - \log(y/2) + 1 \right) \\ &= \mathcal{N}(\mathbf{c}) \end{aligned}$$

(recall that $\|\mathbf{c}\|_{\ell^1} = 1$). Hence

$$|\{\tau > y\}| \leq \frac{1}{y}\|\tau\|_1 \leq \frac{\mathcal{N}(\mathbf{c})}{y}.$$

Finally, we observe by construction that

$$|\{\lambda > y/2\}| = 0$$

Hence λ makes no contribution to (44) and the proof of the lemma is complete. ∎

By the definition of the $\mathcal{B}_q$ norm, Lemma 8 and (40) lead to the inequality

$$|\{|f| > y\}| \leq \frac{3\|f\|_{\mathcal{B}_q}}{y} \tag{45}$$

for all $y > 0, f \in \mathcal{B}_q$, and $1 < q \leq \infty$.

THEOREM 8. *Let $1 < q \leq \infty$. There is a constant C depending only on q such that*

$$|\{S^*f > y\}| \leq \frac{C\|f\|_{\mathcal{B}_q}}{y} \tag{46}$$

for all $y > 0$ and $f \in \mathcal{B}_q$. Moreover, if $f \in \mathcal{B}_q$ then Sf converges to f a.e. on $[0,1)$.

PROOF. Fix $f \in \mathcal{B}_q$. Choose $\mathbf{c} \in \ell^1$ with $\mathcal{N}(\mathbf{c}) < \infty$ and q-blocks $\beta_1, \beta_2, \ldots$ such that

$$f = \sum_{k=1}^{\infty} c_k \beta_k.$$

Since this series converges in L^1 norm, it is clear that

$$S_n f = \sum_{k=1}^{\infty} c_k S_n \beta_k \qquad (n \in \mathbf{N}).$$

Thus

$$S^* f \leq \sum_{k=1}^{\infty} |c_k| S^* \beta_k,$$

and (46) follows directly from Lemma 7 and Lemma 8.

To prove Sf converges a.e., let $\varepsilon > 0$ and choose N so large that

$$\sum_{k=N+1}^{\infty}{}' |c_k| \left(1 + \log\left(\frac{\sum_{j=N+1}^{\infty} |c_j|}{|c_k|}\right)\right) < \varepsilon. \tag{47}$$

Set $g := \sum_{k=1}^{N} c_k \beta_k$ and $h := f - g$. Since g is a finite sum of L^q functions, g belongs to L^q. Thus Sg converges to g a.e. on $[0,1)$.

To show Sh converges a.e. use (46) to verify

$$\begin{aligned}|\{\lim_{n\to\infty} \sup_{k,j \geq n} |S_k h - S_j h| > y/2\}| &\leq |\{S^* h > y/4\}| \\ &\leq \frac{4C}{y}\|h\|_{\mathcal{B}_q}.\end{aligned}$$

By construction, (47) implies $\|h\|_{\mathcal{B}_q} < \varepsilon$. We conclude that Sh converges a.e. ∎

6.4 Almost Everywhere Convergence of Rearrangements of Walsh-Fourier Series and Closely Related Systems. Let T be a 1-1 map from $\mathbf{N}$ onto $\mathbf{N}$. It induces a rearrangement of the Walsh system by

$$T\boldsymbol{w} := (w_{T(n)}, n \in \mathbf{N}).$$

We shall denote partial sums of Walsh-Fourier series in the system $T\boldsymbol{w}$ by

$$S_m^T f := \sum_{k=0}^{m-1} \widehat{f}(T(k)) w_{T(k)} \qquad (m \in \mathbf{P}),$$

and the corresponding maximal function by

$$S_*^T f := \sup_{m>0} |S_m^T f| \qquad (f \in \mathrm{L}^1).$$

We consider linear rearrangements first.

THEOREM 9. *Let $T\boldsymbol{w}$ be a linear rearrangement of the Walsh system. Then S_*^T is of type (p,p) for $1 < p < \infty$. Moreover, $S^T f$ converges a.e. to f for all $f \in \mathrm{L}\log^+ \mathrm{L}\log^+ \log^+ \mathrm{L}$.*

PROOF. Let $T : \mathbf{N} \to \mathbf{N}$ be a 1-1 linear bijection. Choose by Theorem 7 in 1.4 a 1-1 measure preserving transformation $T' : [0,1) \to [0,1)$ such that

$$w_{T(n)}(x) = w_n(T'(x))$$

for $x \in [0,1)$ and $n \in \mathbf{N}$. Set

$$T^* := (T^{-1})' = (T')^{-1}.$$

Notice for every $f \in \mathrm{L}^1$ and $n \in \mathbf{N}$, that

$$\begin{aligned}\widehat{f}(T^{-1}(n)) &= \int_0^1 f w_{T^{-1}(n)} \\ &= \int_0^1 (w_n \circ T^*) f \\ &= \int_0^1 (f \circ T') w_n \\ &= \widehat{f \circ T'}(n).\end{aligned}$$

Consequently,

$$\widehat{f}(T^{-1}(n)) = \widehat{f \circ T'}(n) \quad \text{and} \quad \widehat{f}(T(n)) = \widehat{f \circ T^*}(n).$$

We shall use these identities to show that $S_*^T f$ can be estimated using $S^* f$. We begin by establishing

$$S_m^T f = (S_m(f \circ T^*)) \circ T' = f * (D_m \circ T') \tag{48}$$

for every $m \in \mathbf{P}$ and $f \in \mathrm{L}^1$.

Fix $m \in \mathbf{P}, f \in \mathrm{L}^1$, and a subset $A \subseteq \mathbf{N}$. Define a projection

$$\Pi f := \sum_{n \in A} \widehat{f}(n) w_n$$

and observe that $\Pi f = f * D$ for

$$D := \sum_{n \in A} w_n.$$

Set

$$\Pi^T f := \sum_{n \in A} \widehat{f}(T(n)) w_{T(n)}$$

and verify that

$$\Pi^T f = (\Pi(f \circ T^*)) \circ T'. \tag{49}$$

Moreover, the definition of D implies

$$\begin{aligned} f * (D \circ T') &= \sum_{k \in \mathbf{N}} \widehat{f}(k) \widehat{D \circ T'}(k) w_k \\ &= \sum_{k \in \mathbf{N}} \widehat{f}(k) \widehat{D}(T^{-1}(k)) w_k \\ &= \sum_{n \in A} \widehat{f}(T(n)) w_{T(n)}. \end{aligned}$$

Therefore, we also have

$$\Pi^T f = f * (D \circ T'). \tag{50}$$

Specializing (49) and (50) to partial sums of the linear rearrangement $T\boldsymbol{w}$, (48) is obtained.

Since T' and T^* are measure preserving, it follows immediately from Corollary 11 in 3.7 that

$$\|S_*^T f\|_p = \|S^*(f \circ T^*)\|_p \leq C_p \|f \circ T^*\|_p = C_p \|f\|_p$$

for $1 < p < \infty$, and $f \in \mathrm{L}^p$, where C_p is an absolute constant depending only on p. Thus S_*^T is of type (p, p) for $1 < p < \infty$.

Similarly, we can transfer (32) in 6.3 from the Walsh system to any of its linear rearrangements. In particular, the argument of Theorem 7 in 6.3 shows that $S^T f$ converges to f a.e. for all $f \in \mathrm{L} \log^+ \mathrm{L} \log^+ \log^+ \mathrm{L}$. ∎

This result evidently includes the original Walsh system but not the Walsh-Kaczmarz system. Thus we turn our attention to piecewise linear rearrangements.

THEOREM 10. *Let $R\omega$ be a piecewise linear rearrangement of the Walsh system. Then S_*^R is of type (p,p) for all $2 \le p < \infty$ and $S^R f$ converges a.e. to f for all $f \in L^p, p \ge 2$.*

PROOF. Let $R : \mathbf{N} \to \mathbf{N}$ be a piecewise linear rearrangement generated by $\mathbf{Z}_2$-linear maps

$$R_n : \{0,1,\dots,2^n-1\} \to \{0,1,\dots,2^n-1\} \qquad (n \in \mathbf{N}).$$

Thus $R(2^n+m) = 2^n + R_n(m)$ and it follows from the proof of Theorem 9 that

$$S^R_{2^n+m} f = S_{2^n} f + r_n \left(S_m((r_n\Delta_n f)\circ R_n^*)\right)\circ R_n' \tag{51}$$

for $0 \le m < 2^n, n \in \mathbf{N}$, and $f \in L^1$, where R_n' and R_n^* are given by Theorem 7 in 1.4.

Fix $f \in L^p, 2 \le p < \infty$, and consider the maximal function

$$F^* f := \sup_{m,n} |\left(S_m((r_n\Delta_n f)\circ R_n^*)\right)\circ R_n'|.$$

Since the maps R_n' and R_n^* are measure preserving, it is clear by Corollary 11 in 3.7 that

$$\begin{aligned}\|F^* f\|_p^p &\le \sum_{n\in\mathbf{N}} \|S^*\left((r_n\Delta_n f)\circ R_n^*\right)\|_p^p \\ &\le C_p \sum_{n\in\mathbf{N}} \|(\Delta_n f)\circ R_n^*\|_p^p \\ &= C_p \sum_{n\in\mathbf{N}} \int_0^1 |\Delta_n f|^p\end{aligned}$$

where C_p is an absolute constant depending only on p. Using the fact that $p \ge 2$, we continue this inequality by

$$\sum_{n\in\mathbf{N}} \int_0^1 |\Delta_n f|^p \le \int_0^1 \left(\sum_{n=0}^{\infty} |\Delta_n f|^2\right)^{p/2} = \|\mathcal{Q}f\|_p^p.$$

It follows, therefore, from Corollary 5 in 3.3 that the map $f \to F^* f$ is of type (p,p). Since (51) implies

$$S_*^R f \le \mathcal{E}^* f + F^* f$$

we conclude that S_*^R is of type (p,p) for $2 \le p < \infty$. ∎

In particular, the Walsh-Kaczmarz-Fourier series of an $f \in L^p$ converges a.e. when $p \ge 2$. We will show this result holds also for $1 < p < 2$.

Let $R : \mathbf{N} \to \mathbf{N}$ be the piecewise linear map which generates the Kaczmarz rearrangement. For each $0 \le m < 2^n, n \in \mathbf{N}$ and $f \in L^1$ set

$$F_m^n f := S^R_{2^n+m} f - S^R_{2^n} f.$$

The maximal function

$$F_N^* f := \sup_{0\le n<N} \sup_{0\le m<2^n} |F_m^n f|$$

can be estimated using a martingale transform based on the Kaczmarz rearrangement (see (55) below).

Indeed, let R_n be the $\mathbf{Z}_2$-linear maps which generate R (see 1.4). Thus

$$R_n(m) = \sum_{j=0}^{n-1} m_j 2^{n-j-1}$$

for $0 \le m < 2^n$ and $n \in \mathbf{N}$. For each integer dyadic interval $I \in \mathcal{J}$ (see 3.7) let

$$\alpha_I := m_s \tag{52}$$

where $I := I(m, -s)$. Using the notation introduced in the proof of Theorem 9 we have

$$F_m^n f = \sum_{2^n+m \in I \subset [2^n, 2^{n+1})} \alpha_I \Delta_I^{R_n} f \tag{53}$$

(compare with (76) in 3.7).

Fix $N \in \mathbf{P}$ and let

$$\mathcal{J}_N := \{I \in \mathcal{J} : I \subset [0, 2^N) \quad \text{and} \quad I \neq [0, 2^n) \quad \text{for} \quad 0 \le n < N\}.$$

Given $I \in \mathcal{J}_N$, choose a unique $n \in \mathbf{N}$ such that $0 \le n < N$ and $I \subseteq [2^n, 2^{n+1})$. Consequently, I is of the form $I = I(2^n + m, -s)$ for some $0 \le m < 2^n$ and $0 \le s \le n$. Define a map $I \to \widetilde{I}$ from $\mathcal{J}_N$ to $\mathcal{J}$ by

$$\widetilde{I} := I(m2^{N-n}, -s - (N-n)).$$

It is easy to check that if $\boldsymbol{\alpha} = (\alpha_I, I \in \mathcal{J})$ is defined by (52), then $\boldsymbol{\alpha}$ is invariant under this map, i.e.,

$$\alpha_I = \alpha_{\widetilde{I}} \qquad (I \in \mathcal{J}_N).$$

Apply (51) to the operators

$$\mathcal{E}_I f := f * w_{2^n+m} D_{2^s}$$

for $I = I(2^n + m, -s)$ to obtain

$$\mathcal{E}_I^{R_n} f = f * (w_{2^n + R_n(m)}(D_{2^s} \circ R_n')) . \tag{54}$$

Use the definition of R_n to verify

$$\begin{aligned} D_{2^s} \circ R_n' &= \prod_{j=n-s}^{n-1} (1 + r_j) \\ &= D_{2^n} * \prod_{j=n-s}^{N-1} (1 + r_j) \\ &= D_{2^n} * (D_{2^{N-n+s}} \circ R_N') \end{aligned}$$

for all $0 \le s \le n < N$. Set

$$\widetilde{m} := R_N(2^n + R_n(m))$$

and observe that $\widetilde{m} = 2^{N-n-1} + m2^{N-n}$. Thus

$$w_{\widetilde{m}} D_{2^{N-n+s}} = w_{m2^{N-n}} D_{2^{N-n+s}} .$$

Apply (54) with R_N in place of R_n and $\widetilde{m}$ in place of m. It follows from the definition of $\widetilde{I}$ and (51) that

$$\begin{aligned}\Delta_I^{R_n} f &= D_{2^n} * f * \left(w_{R_N(\widetilde{m})} D_{2^{N-n+s}} \circ R'_N \right) \\ &= D_{2^N} * \Delta_{\widetilde{I}}^{R_N} f \\ &= \mathcal{E}_N(\Delta_{\widetilde{I}}^{R_N} f).\end{aligned}$$

Since $\alpha_I = \alpha_{\widetilde{I}}$ for $I \in \mathcal{J}_N$ we obtain

$$F_m^n f = \mathcal{E}_n \left(\sum_{I(m+2^n,0) \subseteq \widetilde{I} \subseteq I(m+2^n,-n)} \alpha_{\widetilde{I}} \Delta_{\widetilde{I}}^{R_N} f \right) .$$

Consequently,

$$F_N^* f \le \sup_{n \le N} \mathcal{E}_n(\boldsymbol{T}^{R_N}(\boldsymbol{\alpha}) f) \le \mathcal{E}^*(\boldsymbol{T}^{R_N}(\boldsymbol{\alpha}) f) \tag{55}$$

where

$$\boldsymbol{T}^{R_N}(\boldsymbol{\alpha}) f = \sup_{I \in \mathcal{J}} \boldsymbol{T}_I^{R_N}(\boldsymbol{\alpha}) f.$$

We are now prepared to prove the following.

THEOREM 11. *If $f \in \mathrm{L}(\log^+ \mathrm{L})^2$ (in particular if $f \in \mathrm{L}^p$ for any $p > 1$) then the Walsh-Kaczmarz-Fourier series of f converges a.e. on $[0, 1)$.*

PROOF. By (55) above and Corollary 1 in 3.1, it is clear that

$$\begin{aligned}|\{F_N^* f > y\}| &\le |\{\mathcal{E}^*(\boldsymbol{T}^{R_N}(\boldsymbol{\alpha}) f) > y\}| \\ &\le \frac{1}{y} \|\boldsymbol{T}^{R_N}(\boldsymbol{\alpha}) f\|_1\end{aligned}$$

for any $f \in \mathrm{L}^1$ and $n \in \mathbf{P}$. But

$$\boldsymbol{T}^{R_N}(\boldsymbol{\alpha}) f = (\boldsymbol{T}(\boldsymbol{\alpha})(f \circ R_N^*)) \circ R'_N .$$

Since both R'_N and R_N^* are measure preserving, it follows from (36) in 6.3 that

$$|\{F_N^* > y\}| \le \frac{c}{y} \varrho_1(f)$$

for any f satisfying

$$\varrho_1(f) := \int_0^1 \big(|f(x)|(\log^+|f(x)|)^2 + 1\big)\,dx < \infty.$$

Let $N \to \infty$ and use the density argument presented in the proof of Theorem 7 in 6.3. We conclude that $S^R f$ converges a.e. for all f satisfying $\varrho_1(f) < \infty$. ∎

These techniques apply to certain orthonormal systems which share some properties with the Walsh system. We shall consider product systems and W systems.

Let $\boldsymbol{g} = (g_n, n \in \mathbf{N})$ be a system of functions in some space $L^\infty(\Omega, \nu)$. Then $\boldsymbol{g}$ is called a *convergence system* if

$$\sum_{n=0}^{\infty} a_n g_n$$

converges a.e. for all $\boldsymbol{a} = (a_n, n \in \mathbf{N})$ belonging to ℓ^2. The density argument we have frequently used shows that $\boldsymbol{g}$ is a convergence system if there is an absolute constant B, depending only on $\boldsymbol{g}$, such that

$$\left\| \sup_{N \in \mathbf{N}} \Big| \sum_{n=0}^{N} a_n g_n \Big| \right\|_{L^2(\Omega)} \le B\|\boldsymbol{a}\|_{\ell^2} \tag{56}$$

for all $\boldsymbol{a} = (a_n, n \in \mathbf{N}) \in \ell^2$. Moreover, for orthogonal $\boldsymbol{g}$ the Riesz-Fischer theorem implies that $\boldsymbol{g}$ is a convergence system if and only if the $\boldsymbol{g}$-Fourier series of every $f \in L^2(\Omega, \nu)$ converges a.e. $[\nu]$.

Let (Ω, ν) by a probability space and let $\gamma_k \in L^\infty(\Omega)$ with $\|\gamma_k\|_\infty \le 1$ for $k \in \mathbf{N}$. Recall that the product system $\boldsymbol{g} = (g_n, n \in \mathbf{N})$ generated by the γ_k's is the system

$$g_n := \prod_{k=0}^{\infty} \gamma_k^{n_k} \qquad (n \in \mathbf{N}),$$

where $(n_k, k \in \mathbf{N})$ are the binary coefficients of n. The system $\boldsymbol{\gamma} = (\gamma_k, k \in \mathbf{N})$ is called weakly multiplicative if

$$A := \sum_{n=0}^{\infty} \left| \int_\Omega g_n \, d\nu \right| < \infty.$$

THEOREM 12. *The product system of a weakly multiplicative system is a convergence system.*

PROOF. Let $\boldsymbol{g} = (g_n, n \in \mathbf{N})$ be such a product system defined on a probability space (Ω, ν). It suffices to show (56) for some absolute constant B.

Fix $N \in \mathbf{N}$. For each $x \in \Omega$ and $t \in [0, 1)$ set

$$K_N(x, t) := \sum_{k=0}^{2^N - 1} g_k(x) w_k(t),$$

and

$$(\boldsymbol{K}_N f)(x) := \int_0^1 K_N(x,t) f(t)\,dt \qquad (f \in \mathrm{L}^1).$$

Recall from the proof of Theorem 16 in 5.5 that the kernel K_N is non-negative and

$$\boldsymbol{K}_N(w_n) = g_n \qquad (0 \le n < 2^N). \tag{57}$$

Non-negativity implies

$$\|K_N\|_{[\infty,1]} := \| \int_0^1 K_N(\cdot,t)\,dt\|_\infty = 1$$

and

$$\|K_N\|_{[1,\infty]} := \| \int_\Omega K_N(x,\cdot)\,d\nu(x)\|_\infty \le \sum_{k=0}^{2^N-1} \left| \int_\Omega g_k\,d\nu \right|.$$

Hence it follows from Lemma 3 in 5.3 and hypothesis that

$$\sup_{\|f\|_2 \le 1} \|\boldsymbol{K}_N f\|_{L^2(\Omega)} \le \frac{1}{2}(1+A).$$

By (57) we have

$$\sup_{0 \le m < 2^N} |\sum_{n=0}^m a_n g_n| \le \boldsymbol{K}_N \left(\sup_{0 \le m < 2^N} |\sum_{n=0}^m a_n w_n| \right).$$

Since the maximal operator S^* is of type $(2,2)$ we conclude that

$$\left\| \sup_{0 \le m < 2^N} |\sum_{n=0}^m a_n g_n| \right\|_{L^2(\Omega)} \le \frac{1}{2}(1+A) \left\| \sup_{0 \le m < 2^N} |\sum_{n=0}^m a_n w_n| \right\|_2 \le B\|\boldsymbol{a}\|_{\ell^2}$$

for some absolute constant $B > 0$. In particular, (56) is obtained by letting $N \to \infty$. ∎

Let (Ω, ν) be a probability space. A system $H = (H_n, n \in \mathbf{N})$ has the *H-property* if there is an absolute constant $M > 0$ such that

$$\left\| \sum_{k=2^n}^{2^{n+1}-1} |H_k| \right\|_{L^\infty(\Omega)} \le M 2^{n/2}$$

and

$$\|H_j\|_{L^1(\Omega)} \le M 2^{-n/2}$$

for $2^n \le j < 2^{n+1}$ and $n \in \mathbf{N}$. A system $W = (W_n, n \in \mathbf{N})$ is called a *W-system* if there is a system $H = (H_n, n \in \mathbf{N})$ which has the H-property such that W is the Hadamard transform of H. Since the Haar system $(h_n, n \in \mathbf{N})$ has the H- property, the Walsh system is a W-system.

Every W-system generated by a convergence system which has the H-property is itself a convergence system. In fact, THEOREM 13. *Let $H = (H_n, n \in \mathbf{N})$ have the H-property on some probability space (Ω, ν), and suppose*

$$\sum_{k=0}^{2^N-1} b_k H_k$$

converges a.e. $[\nu]$ as $N \to \infty$ for every $(b_k, k \in \mathbf{N}) \in \ell^2$. If $W = (W_n, n \in \mathbf{N})$ is the W-system generated by H then W is a convergence system.

PROOF. Fix $(a_k, k \in \mathbf{N}) \in \ell^2$. We must show that

$$\sum_{k=0}^{\infty} a_k W_k$$

converges a.e. $[\nu]$. Recall from (42) in 1.4 that

$$S_{2^N}^W := \sum_{k=0}^{2^N-1} a_k W_k = \sum_{k=0}^{2^N-1} a_k^{\perp} H_k$$

for $N \in \mathbf{N}$. Since the Hadamard transform is an isometry on ℓ^2, it follows from hypothesis that $S_{2^N}^W$ converges a.e. as $N \to \infty$. Therefore, it suffices to show

$$\eta_N := \max_{2^N \le m < 2^{N+1}} \left| \sum_{k=2^N}^{m} a_k W_k \right|$$

converges to zero a.e. $[\nu]$, as $N \to \infty$.

Fix $N \in \mathbf{N}$ and for each $x \in \Omega$, $t \in [0,1)$ set

$$K_N(x,t) := \sum_{k=2^N}^{2^{N+1}-1} H_k(x) h_k(t),$$

$$(\boldsymbol{K}_N f)(x) := \int_0^1 K_N(x,t) f(t)\, dt,$$

and

$$(\widetilde{\boldsymbol{K}}_N f)(x) := \int_0^1 |K_N(x,t)| f(t)\, dt \qquad (f \in \mathrm{L}^1).$$

Since the Haar system is the Hadamard transform of the Walsh system, it is easy to see that

$$W_k = \boldsymbol{K}_N(w_k)$$

for $2^N \le k < 2^{N+1}$. Thus

$$\eta_N \le \widetilde{\boldsymbol{K}}_N \left(\max_{2^N \le m < 2^{N+1}} \left| \sum_{k=2^N}^{m} a_k w_k \right| \right). \tag{58}$$

The H-property implies

$$\|K_N\|_{[\infty,1]} \le M, \qquad \|K_N\|_{[\infty,1]} \le M$$

for all $N \in \mathbf{N}$. Therefore the operators $\widetilde{\boldsymbol{K}}_N$ are uniformly bounded from L^2 to $L^2(\Omega)$. Since the maximal operator S^* is of type $(2,2)$, it follows from (58) that

$$\begin{aligned}\|\eta_N\|_{L^2(\Omega)} &\le M\| \max_{2^N \le m < 2^{N+1}} |\sum_{k=2^N}^{m} a_k w_k| \|_2 \\ &\le C \left(\sum_{j=2^N}^{2^{N+1}-1} |a_j|^2 \right)^{1/2}\end{aligned}$$

for $N \in \mathbf{N}$ and some absolute constant C. Consequently,

$$\nu(\{\sup_{n \ge N} \eta_n > y\}) \le \frac{C}{y^2} \sum_{j=2^N}^{\infty} |a_j|^2$$

for all $y > 0$ and $N \in \mathbf{N}$. We conclude that $\eta_N \to 0$ a.e. $[\nu]$ as $N \to \infty$. ∎

Since the Franklin system is a convergence system which has the H-property (see Appendix 0.7) we have:

COROLLARY 10. *The Ciesielski system is a convergence system.*

6.5 Divergent Walsh-Fourier Series. In this section we shall construct several divergent Walsh- Fourier series to show that many of the results in 6.1, 6.3, and 6.4 are best possible.

For every $f \in L^1$ let

$$\Delta^* f := \sup_{m,n \in \mathbf{N}} |S_n f - S_m f|.$$

Our constructions depend on the following simple observation.

LEMMA 9. *Let $(N_k, k \in \mathbf{N})$ and $(M_k, k \in \mathbf{N})$ be increasing sequences of natural numbers such that the intervals $[N_k, M_k)$ are pairwise disjoint for $k = 0, 1, \ldots$ Suppose $(P_k, k \in \mathbf{N})$ is a sequence of Walsh polynomials whose spectra satisfy*

$$sp(P_k) \subseteq [N_k, M_k)$$

such that

$$\sum_{k=0}^{\infty} \|P_k\|_1 < \infty. \tag{59}$$

If

$$f := \sum_{k=0}^{\infty} P_k \tag{60}$$

then Sf diverges everywhere on the set

$$\check{E} := \{\limsup_{k\to\infty} \Delta^* P_k > 0\}.$$

PROOF. By (59) the series (60) converges in L^1 norm. Thus f belongs to L^1 and is defined a.e. on [0,1). Moreover, for each $x \in \check{E}$ there exist indices $N_k \le n_k < m_k \le M_k$ and a number $\alpha > 0$ such that

$$|(S_{m_k} f)(x) - (S_{n_k} f)(x)| = |(S_{m_k} P_k)(x) - (S_{n_k} P_k)(x)| \ge \alpha$$

for infinitely many integers k. In particular, Sf diverges on $\check{E}$. ∎

Throughout this section $g_0, g_1, \ldots$ will represent a fixed sequence of functions which satisfies the following properties. For each $n \in \mathbf{N}, g_n : [0,1) \to \{-1,1\}$ is $\mathcal{A}^n$-measurable and there exists an integer $0 \le i_n < 2^n$ such that

$$|(S_{i_n} g_n)(x)| > cn \qquad (x \in I(0,n)), \tag{61}$$

where $c > 0$ is an absolute constant which does not depend on n. That such a sequence of functions exists is easily verified. For example, let $i_0 := i_1 := 1$, and for $n \ge 2$ let

$$i_n := \sum_{k=1}^{[n/2]} 2^{2(k-1)}.$$

Set

$$g_n := \operatorname{sgn} D_{i_n}$$

for $n \in \mathbf{N}$. By Theorem 9 in 1.6 there is an absolute constant $c > 0$ such that $\|D_{i_n}\|_1 > cn$. Since

$$(S_{i_n} g_n)(x) = (S_{i_n} g_n)(0) = \|D_{i_n}\|_1$$

for $x \in I(0,n)$ and $n \in \mathbf{N}$, it is thus clear that (61) holds for these g_n's.

Let

$$R_k^{(n)} := r_{n+k} r_{k2^{-n}} g_n$$

for $0 \le k < 2^n$. We shall refer to the sequence $(Q_n, n \in \mathbf{N})$ defined by

$$Q_n := \prod_{k=0}^{2^n-1} (1 + R_k^{(n)}) \tag{62}$$

as the *Walsh-Kolmogorov polynomials.* These Walsh polynomials enjoy several properties useful in constructing divergent Walsh- Fourier series.

First, they are essentially the Dirichlet kernels of the product system generated by the $R_k^{(n)}$'s. Indeed, if

$$W_m^{(n)} := \prod_{k=0}^{2^n-1} (R_k^{(n)})^{m_k}$$

for $0 \le m < 2^{2^n}$, where $(m_k, k \in \mathbf{N})$ represents the binary coefficients of $m \in \mathbf{N}$, then it is easy to see by induction that

$$Q_n = \sum_{m=0}^{2^{2^n}-1} W_m^{(n)} \qquad (n \in \mathbf{N}). \tag{63}$$

Thus

$$sp(W_m^{(n)}) \subseteq [m2^n, (m+1)2^n) \quad \text{and} \quad sp(Q_n) \subseteq [0, 2^{n+2^n}) \tag{64}$$

for $0 \le m < 2^{2^n}$ and $n \in \mathbf{N}$. In particular, the spectra of the $W_m^{(n)}$'s are pairwise disjoint for each fixed $n \in \mathbf{N}$.

Next, each Q_n takes only the values 0 and 2^{2^n}.

By (63) and (64) it is clear that

$$\|Q_n\|_1 = 1 \qquad (n \in \mathbf{N}). \tag{65}$$

By definition,

$$\begin{aligned} |(S_{2^{n+k}+i_n} Q_n - S_{2^{n+k}} Q_n)(x)| &= |(S_{2^{n+k}+i_n} W_{2^k}^{(n)})(x)| \\ &= |(S_{2^{n+k}+i_n} R_k^{(n)})(x)| \\ &= |(S_{i_n} g_n)(x \dot{+} k2^{-n})| \end{aligned}$$

for $x \in I(k,n)$ and $0 \le k < 2^n$. Thus by (61) we have

$$(\Delta^* Q_n)(y) > cn \qquad (y \in [0,1)). \tag{66}$$

In fact, we have

$$|(S_{2^{n+k}+i_n} Q_n - S_{2^{n+k}} Q_n)(x)| > cn \tag{67}$$

for $x \in I(k,n), 0 \le k < 2^n$, and $n \in \mathbf{N}$. Consequently, for every $x \in (0,1)$ there is an index

$$0 \le j_n(x) < 2^{n+2^n}$$

such that

$$|(S_{j_n(x)} Q_n)(x)| > \frac{cn}{2}. \tag{68}$$

To see how these properties can be used in conjunction with Lemma 9 to construct divergent Walsh-Fourier series we begin with the following

LEMMA 10. *If* $\boldsymbol{a} = (a_n, n \in \mathbf{N})$ *belongs to* ℓ^1 *and* $\boldsymbol{\omega} = (\omega_n, n \in \mathbf{N})$ *is an increasing sequence in* $\mathbf{N}$, *then*

$$P_n := a_n r_{2^{\omega_n+1}} Q_{\omega_n}$$

satisfies

$$sp(P_n) \subseteq [2^{2^{\omega_n+1}}, 2^{2^{\omega_{n+1}+1}})$$

for $n \in \mathbf{N}$, *and*

$$F := F_{\boldsymbol{a},\boldsymbol{\omega}} := \sum_{n=0}^{\infty} P_n$$

belongs to L^1. *Moreover, if*

$$\limsup_{n\to\infty} |a_n|\omega_n > 0 \tag{69}$$

then the Walsh-Fourier series of F *diverges everywhere* $[0,1)$.

PROOF. The construction implies

$$sp(P_n) \subseteq [2^{2^{\omega_n+1}}, 2^{2^{\omega_n+1}+2^{\omega_n}+\omega_n}).$$

Since $\boldsymbol{a} \in \ell^1$ it is clear that (59) holds. Thus $F \in \mathrm{L}^1$.

By Lemma 9, SF diverges everywhere on

$$\check{E} := \{\limsup_{n\to\infty} \Delta^* P_n > 0\}.$$

Since (66) and (69) imply $\check{E} = [0,1)$, the proof of the lemma is complete. ∎

Our first application of this result generalizes Theorem 12 in 4.5.

THEOREM 14. *Let* $\Phi, \Lambda : [0,\infty) \to [0,\infty)$ *be continuous, increasing functions which satisfy* $\Phi(u) = \Lambda(u) = 1$ *for* $0 \le u \le 4$ *and*

$$\Phi(u)\Lambda(u) = \log_2 \log_2 u$$

for $u \ge 4$. *Suppose* $(\lambda_n, n \in \mathbf{N})$ *is an increasing sequence of positive numbers such that*

$$\lambda_n = o(\Lambda(n)) \quad \text{as} \quad n \to \infty.$$

Then there is an $F \in \mathrm{L}^1$ *satisfying*

$$\int_0^1 F^* \Phi(F^*) < \infty \tag{70}$$

(where $F^* := \mathcal{E}^*(F)$*) such that* SF *diverges everywhere on* $[0,1)$ *and*

$$\limsup_{n\to\infty} \frac{|S_n F|}{\lambda_n} > 0 \tag{71}$$

a.e. *on* $[0,1)$.

PROOF. By hypothesis choose a decreasing sequence $(\epsilon_n, n \in \mathbf{N})$ of positive numbers such that $\epsilon_n \to 0$ as $n \to \infty$, and

$$\lambda_n \le \epsilon_n \Lambda(n) \qquad (n \in \mathbf{N}).$$

Observe for $n \ge 4$ that

$$\begin{aligned}\beta_n &:= \frac{n\epsilon_n}{\Phi(n)} \\ &= \frac{n\epsilon_n \Lambda(n)}{\log_2 \log_2 n} \\ &\ge \frac{n\lambda_n}{\log_2 \log_2 n}.\end{aligned}$$

Hence $\beta_n \to \infty$ as $n \to \infty$ and we can choose an increasing sequence of integers $(\omega_n, n \in \mathbf{N})$ with $\omega_0 \ge 4$ such that

$$\ell_n := 2^{2^{\omega_n}} \qquad (n \in \mathbf{N})$$

enjoys the following properties:

$$\beta_{\ell_n} \le \frac{1}{2}\beta_{\ell_{n+1}},$$

(72)
$$\frac{\epsilon_{\ell_n}}{\Phi(\ell_n)} < \frac{1}{2}$$

for $n \in \mathbf{N}$, and

$$\sum_{n=0}^{\infty} \epsilon_{\ell_n} < \infty.$$

Set $a_n := \epsilon_{\ell_n}/\Phi(\ell_n)$ for $n \in \mathbf{N}, \boldsymbol{\omega} := (\omega_n, n \in \mathbf{N})$ and $\boldsymbol{a} := (a_n, n \in \mathbf{N})$. Clearly $\boldsymbol{a}$ belongs to ℓ^1. Moreover, the construction implies

$$\begin{aligned}a_n\omega_n &:= \frac{\epsilon_{\ell_n}\omega_n}{\Phi(\ell_n)} \\ &= \frac{\epsilon_{\ell_n}\omega_n \Lambda(\ell_n)}{\log_2 \log_2 \ell_n} \\ &= \epsilon_{\ell_n}\Lambda(\ell_n) \\ &\ge \lambda_{\ell_n}\end{aligned}$$

for $n \in \mathbf{N}$. Hence choose $F := F_{\boldsymbol{a},\boldsymbol{\omega}}$ by Lemma 10 such that SF diverges everywhere.

To verify (70) use the facts

$$sp(Q_{\omega_n}) \subseteq [0, 2^{\omega_n + 2^{\omega_n}}), \quad 2^{\omega_n + 2^{\omega_n}} \le 2^{2^{\omega_n+1}}, \quad \text{and} \quad a_n Q_{\omega_n} \ge 0$$

to see that

$$F^* \leq \sum_{n=0}^{\infty} a_n Q_{\omega_n}.$$

Decompose

$$[0,1) = A^* \bigcup \left(\bigcup_{n=0}^{\infty} A_n \right),$$

where $A^* := \limsup_{n\to\infty}\{Q_{\omega_n} > 0\}$ and

$$A_n := \{Q_{\omega_n} > 0 \quad \text{and} \quad Q_{\omega_k} = 0 \quad \text{for all} \quad k > n\}$$

for $n \in \mathbf{N}$. Recall that Q_{ω_n} takes on only the values 0 and $2^{2^{\omega_n}}$. Consequently, the definition of A_n and (72) imply that

$$\begin{aligned}\chi(A_n)F^* &\leq \sum_{k=0}^{n} a_k 2^{2^{\omega_k}} \\ &\leq \sum_{k=0}^{n} \frac{\epsilon_{\ell_k}\ell_k}{\Phi(\ell_k)} \\ &= \sum_{k=0}^{n} \beta_{\ell_k}.\end{aligned}$$

Since $\beta_{\ell_k} \leq \beta_{\ell_{k+1}}/2$, we have

$$\chi(A_n)F^* \leq 2\beta_{\ell_n} = \frac{2\epsilon_{\ell_n}\ell_n}{\Phi(\ell_n)} \leq \ell_n$$

for each $n \in \mathbf{N}$. But (65) and the fact that Q_{ω_n} takes on the values 0 and $2^{2^{\omega_n}}$ imply

$$|\{Q_{\omega_n} \neq 0\}| = 2^{-2^{\omega_n}} = \frac{1}{\ell_n}.$$

Consequently, $|A_n| \leq 1/\ell_n$ and $|A^*| = 0$. It follows, therefore, that

$$\begin{aligned}\int_0^1 F^*\Phi(F^*) &= \sum_{n=0}^{\infty} \int_{A_n} F^*\Phi(F^*) \\ &\leq 2\sum_{n=0}^{\infty} \frac{\epsilon_{\ell_n}\ell_n}{\Phi(\ell_n)} \Phi(\ell_n)|A_n| \\ &= 2\sum_{n=0}^{\infty} \epsilon_{\ell_n} < \infty.\end{aligned}$$

To verify (71) fix $x \in [0,1) \setminus A^*$. Set $i_n := j_{\omega_n}(x)$ (see (68)) and

$$k_n := 2^{2^{\omega_n+1}} + i_n.$$

Observe that

$$(S_{k_n}F)(x) = (S_{2^{2^{\omega_n+1}}}F)(x) + a_n r_{2^{\omega_n+1}}(x)(S_{i_n}Q_{\omega_n})(x).$$

For any $x \notin A^*$, the first term of the right side of this identity is constant in n for n sufficiently large. Thus (68) implies

$$\limsup_{n\to\infty} \frac{|(S_{k_n}F)(x)|}{\lambda_{k_n}} \geq \limsup_{n\to\infty} c\frac{a_n\omega_n}{2\lambda_{k_n}}$$

for a.e. $x \in [0,1)$. But

$$\ell_n = 2^{2^{\omega_n}} \leq k_n = 2^{2^{\omega_n+1}} + j_{\omega_n}(x) \leq 2^{2^{\omega_n+2}}.$$

Consequently,

$$\begin{aligned}\log_2 \log_2 \ell_n &= \omega_n \\ &\geq \frac{1}{2}(\omega_n + 2) \\ &= \frac{1}{2}\log_2\log_2 2^{2^{\omega_n+2}} \\ &\geq \frac{1}{2}\log_2\log_2 k_n.\end{aligned}$$

We conclude that

$$\begin{aligned}\frac{a_n\omega_n}{\lambda_{k_n}} &= \frac{\epsilon_{\ell_n}\omega_n}{\Phi(\ell_n)\lambda_{k_n}} \\ &\geq \frac{1}{2}\frac{\epsilon_{k_n}\log_2\log_2 k_n}{\Phi(k_n)\lambda_{k_n}} \\ &= \frac{1}{2}\frac{\varepsilon_{k_n}\Lambda(k_n)}{\lambda_{k_n}} \geq \frac{1}{2}. \qquad \blacksquare\end{aligned}$$

Thus there exist $F \in \mathrm{H}$ and $F \in \mathrm{L}(\log^+\log^+\mathrm{L})^p$ for $0 < p < 1$ such that SF diverges everywhere on $[0,1)$. In fact, by separating the cases $\Lambda := 1$ from $\Phi := 1$, Theorem 13 contains the following results.

COROLLARY 11.

i) Let $\Phi : [0,\infty) \to [0,\infty)$ be an increasing continuous function which satisfies

$$\Phi(u) = o(\log\log u) \quad \text{as} \quad u \to \infty.$$

Then there is an $F \in \mathrm{L}^1$ such that (70) holds and SF diverges everywhere.

ii) Let $(\lambda_n, n \in \mathrm{N})$ be a sequence of real numbers which satisfies

$$\lambda_n = o(\log\log n) \quad \text{as} \quad n \to \infty.$$

Then there is a function $F \in \mathrm{H}$ such that (71) holds a.e.

The next result shows that condition (3) of Theorem 1 in 6.1 cannot be appreciably relaxed.

THEOREM 15. *There exists an $F \in L^1$ such that*

$$\omega^{(1)}(F,\delta) = O\left(\log\log\frac{1}{\delta}\right)^{-1} \quad \text{as} \quad \delta \to 0, \tag{73}$$

and SF diverges everywhere.

PROOF. Let $a_n := 2^{-n}, \omega_n := 2^n$ for $n \in \mathbf{N}$, and apply Lemma 10 to choose $F = F_{a\omega}$ such that SF diverges everywhere.

Fix $0 < y < 1/4$ and define $i \in \mathbf{N}$ by

$$2^{-2^{\omega_{i+1}+1}} \le y < 2^{-2^{\omega_i+1}}.$$

Notice by (64) that

$$(r_{2^{\omega_n+1}} Q_{\omega_n})(x \dot{+} y) = (r_{2^{\omega_n+1}} Q_{\omega_n})(x)$$

for $n < i$. Consequently,

$$\int_0^1 |F(x \dot{+} y) - F(x)|\, dx \le 2 \sum_{n=i}^{\infty} a_n \|Q_{\omega_n}\|_1$$

$$= 4 \cdot 2^i \le \frac{16}{\log_2 \log_2 1/y}.$$

In particular, F satisfies (73). ∎

The Marcinkiewicz test implies that Sf converges a.e. if $f \in L^1$ satisfies

$$\frac{1}{|I_n(x)|} \int_{I_n(x)} |f(t) - f(x)|\, dt \le \frac{M(x)}{n}$$

for $n \in \mathbf{N}, x \in [0,1)$, and some a.e. finite function M. The next result shows that on the right side of this inequality, n cannot be replaced by a sequence which satisfies $\lambda_n = o(n)$ as $n \to \infty$.

THEOREM 16. *Let $(\varepsilon_n, n \in \mathbf{N})$ be any sequence of positive numbers which decreases monotonically to zero. There is an $F \in L^1$ such that*

$$\limsup_{n\to\infty} \frac{n\varepsilon_n}{|I_n(x)|} \int_{I_n(x)} |F(t) - F(x)|\, dt < \infty \tag{74}$$

for a.e. $x \in [0,1)$ and SF diverges everywhere.

PROOF. We may suppose that $\varepsilon_0 \le 1/4$ and $(n\varepsilon_n, n \in \mathbf{N})$ is increasing. Choose a strictly increasing sequence $(\omega_n, n \in \mathbf{N})$ of positive integers such that

$$\sum_{n=0}^{\infty} \varepsilon_{\omega_n}^{1/3} < \infty$$

and

$$\sum_{n=0}^{\infty} \frac{1}{\omega_n} < \infty.$$

Set $a_n := 1/\omega_n, \mathbf{a} := (a_n, n \in \mathbf{N}), \boldsymbol{\omega} := (\omega_n, n \in \mathbf{N})$, and apply Lemma 10 to choose an integrable $F := F_{\mathbf{a},\boldsymbol{\omega}}$ such that SF diverges everywhere.

Observe by the definition of $R_k^{(m)}$ that

$$R_k^{(m)} = \pm r_{m+k} \qquad (0 \le k < 2^m)$$

on every $I(\ell, m)$ with $0 \le \ell < 2^m$. Hence the set

$$\begin{aligned} J_m^\ell &:= \{x \in I(\ell,m) : Q_m(x) \neq 0\} \\ &= \{x \in I(\ell,m) : Q_m(x) = 2^{2^m}\} \end{aligned}$$

is itself an interval of length 2^{-m-2^m} for each $0 \le \ell < 2^m$.

Set $q_m := [m + \log_2 \varepsilon_m^{-1/3}]$ and let $\widetilde{J}_m^\ell$ denote the dyadic interval of length 2^{-q_m} which contains J_m^ℓ. Set

$$A_m := \bigcup_{\ell=0}^{2^m-1} \widetilde{J}_m^\ell \qquad (m \in \mathbf{N}).$$

We shall show that

$$(\Theta_n Q_m)(x) := \frac{n\varepsilon_n}{|I_n(x)|} \int_{I_n(x)} Q_m \le m\varepsilon_m^{1/3} \tag{75}$$

for every $x \notin A_m$, and $n, m \in \mathbf{P}$.

Indeed, use the inequality

$$m + \log_2 t \le mt \qquad (t \ge 2)$$

to verify

$$q_m \varepsilon_{q_m} \le m\varepsilon_m^{2/3} \qquad (m \in \mathbf{P}). \tag{76}$$

Fix $n, m \in \mathbf{P}$ and $x \notin A_m$, and consider three cases separately: $n \ge q_m, m \le n < q_m$, and $n < m$.

If $n \ge q_m$ then $x \notin A_m$ implies Q_m is zero on the set $I_n(x)$. Hence (75) is trivial in this case.

If $m \le n < q_m$ then there exists at most one J_m^ℓ which is contained in $I_n(x)$. Since the value of Q_m is 2^{2^m} on J_m^ℓ and zero elsewhere on $I_n(x)$, we have

$$\int_{I_n(x)} Q_m = 2^{2^m} |J_m^\ell| = 2^{-m}.$$

Since $(n\varepsilon_n, n \in \mathbf{N})$ is increasing and $m \le n < q_m$, it follows from (76) that

$$\begin{aligned}(\Theta_n Q_m)(x) &= \frac{n\varepsilon_n}{|I_n(x)|} 2^{-m}\\ &\le n\varepsilon_n 2^{q_m - m}\\ &\le n\varepsilon_n 2^{\log_2 \varepsilon_m^{-1/3}}\\ &= \frac{n\varepsilon_n}{\varepsilon_m^{1/3}}\\ &\le \frac{q_m \varepsilon_{q_m}}{\varepsilon_m^{1/3}}\\ &\le m\varepsilon_m^{1/3}.\end{aligned}$$

Finally, if $n < m$ then

$$\begin{aligned}(\Theta_n Q_m)(x) &= n\varepsilon_n 2^n \sum_{I(\ell,m)\subseteq I_n(x)} \int_{I(\ell,n)} Q_m\\ &= n\varepsilon_n 2^n \, 2^{m-n} \, 2^{-m}\\ &= n\varepsilon_n\\ &\le m\varepsilon_m\\ &\le m\varepsilon_m^{1/3}.\end{aligned}$$

Thus (75) holds for all $x \notin A_m$ and $m, n \in \mathbf{P}$.

Set

$$A^* := \limsup_{k\to\infty} A_{\omega_k}.$$

Since

$$|A_{\omega_k}| \le 2^{\omega_k} \, 2^{-\omega_k - \log_2 \varepsilon_{\omega_k}^{-1/3} + 1} = 2\varepsilon_{\omega_k}^{1/3}$$

and $\sum \varepsilon_{\omega_k}^{1/3} < \infty$, it is clear that $|A^*| = 0$. Hence to establish (74) we need only consider $x \in [0,1) \setminus A^*$.

Fix such an x. There is an integer $k_0 = k_0(x)$ such that $k \ge k_0$ implies $x \notin A_{\omega_k}$ and $Q_{\omega_k}(x) = 0$. Decompose F by

$$F = \left(\sum_{k=0}^{k_0-1} + \sum_{k=k_0}^{\infty}\right) (a_k r_{2^{\omega_k+1}} Q_{\omega_k}) =: F_1 + F_2.$$

Since F_1 is constant on every interval $I(\ell, 2^{\omega_{k_0}+1})$ it is clear that

$$\frac{n\varepsilon_n}{|I_n(x)|} \int_{I_n(x)} |F_1(t) - F_1(x)| \, dt = 0$$

for $n \geq 2^{\omega_{k_0}+1}$. On the other hand, (75) implies

$$\frac{n\varepsilon_n}{|I_n(x)|}\int_{I_n(x)}|F_2(t)-F_2(x)|\,dt \leq \frac{n\varepsilon_n}{|I_n(x)|}\sum_{k=k_0}^{\infty}\frac{1}{\omega_k}\left(\int_{I_n(x)}Q_{\omega_k}\right)$$
$$\leq \sum_{k=k_0}^{\infty}\varepsilon_{\omega_k}^{1/3} < \infty$$

for all $n \in \mathbf{P}$. We conclude that (74) holds for $x \notin A^*$, hence for a.e. $x \in [0,1)$. ∎

A Walsh-Fourier series is said to diverge boundedly at $x \in [0,1)$ if Sf diverges at x but $S^*f(x) < \infty$.

To construct Walsh-Fourier series which diverge boundedly, we isolate another property common to Walsh-Kolmogorov polynomials.

LEMMA 11. *For every $n \in \mathbf{P}$,*

$$|\{S^*Q_n > 2n\}| \leq \frac{2}{n}.$$

PROOF. Fix $n \in \mathbf{P}$. The function system $\{W_k^{(n)} : 0 \leq k < 2^{2^n}\}$ is orthonormal in $[0,1)$. Thus by Theorem 6 in 1.4 there is a measure preserving transformation $\varpi : [0,1) \to [0,1)$ such that

$$W_k^{(n)} = w_k \circ \varpi$$

for $0 \leq k < 2^{2^n}$. Since

$$Q_n = \sum_{k=0}^{2^{2^n}-1} W_k^{(n)},$$

and the spectrum of each $W_k^{(n)}$ is contained in $[k2^n,(k+1)2^n)$, it follows that

$$S^*Q_n \leq \max_{0\leq m<2^{2^n}}|\sum_{k=0}^{m}W_k^{(n)}| + \max_{0\leq k<2^{2^n}} S^*W_k^{(n)}$$
$$= \max_{0\leq m<2^{2^n}}|D_{m+1}\circ\varpi| + \max_{0\leq k<2^{2^n}} S^*W_k^{(n)}.$$

Now, the definition of $W_k^{(n)}$ implies

$$W_k^{(n)} = w_{k2^n}\, g_k^{(n)}$$

for some $\mathcal{A}^n$ measurable function $g_k^{(n)}$ which takes only the values $+1$ and -1. Hence

$$S^*W_k^{(n)} = \max_{0\leq m<2^{2^n}}|S_m g_k^{(n)}|$$
$$\leq \max_{0\leq m<2^n} L_m$$
$$\leq n,$$

where the L_m's represent the Lebesgue constants introduced in 1.6. It follows that

$$|\{S^*Q_n > 2n\}| \le |\{D^* \circ \varpi > n\}| = |\{D^* > n\}|,$$

where

$$D^*(x) := \sup_{m \in \mathbf{P}} |D_m(x)|.$$

But

$$D^*(x) \le \frac{2}{x}$$

for all $x \neq 0$ (see Theorem 10 in 1.6). In particular,

$$|\{D^* > n\}| \le \frac{2}{n}$$

and the proof of Lemma 6 is complete. ∎

We now show that the example in Corollary 11 i) above can be chosen so that SF diverges boundedly a.e.

THEOREM 17. *Let $\Phi : [0, \infty) \to [0, \infty)$ be a continuous, monotone increasing function which satisfies*

$$\Phi(u) = o(\log \log u) \quad \text{as} \quad u \to \infty.$$

Then there is a function $F \in \mathrm{L}^1$ such that SF diverges boundedly a.e., diverges everywhere, and satisfies

$$\int_0^1 F^*\Phi(F^*) < \infty.$$

PROOF. Let $\varepsilon_n \to 0$ be a decreasing sequence of positive numbers such that

$$\Phi(n) = \varepsilon_n \log_2 \log_2 n \qquad (n \ge 4).$$

Let $(\omega_n, n \in \mathbf{N})$ be a strictly increasing sequence in $\mathbf{N}$ such that

$$\sum_{n=0}^{\infty} \frac{1}{\omega_n} < \infty$$

which also satisfies (72) for

$$\ell_n := 2^{2^{\omega_n}}.$$

Let

$$a_n := \frac{\varepsilon_{\ell_n}}{\Phi(\ell_n)} = \frac{1}{\omega_n} \qquad (n \in \mathbf{N})$$

and set $\boldsymbol{\omega} = (\omega_n, n \in \mathbf{N})$ and $\boldsymbol{a} = (a_n, n \in \mathbf{N})$. Since $a_n\omega_n = 1$, apply Lemma 10 to choose $F := F_{\boldsymbol{a},\boldsymbol{\omega}}$ such that SF diverges everywhere.

By repeating the proof of Theorem 14 one can show that

$$\int_0^1 F^*\Phi(F^*) < \infty.$$

Thus it remains to prove S^*F is a.e. finite.

Set

$$A_n := \{a_n S^* Q_{\omega_n} > 2\} = \{S^* Q_{\omega_n} > 2\omega_n\}$$

for $n \in \mathbf{N}$ and let

$$A^* := \limsup_{n\to\infty} A_n.$$

By Lemma 11 we have

$$\sum_{n=0}^{\infty} |A_n| \leq 2 \sum_{n=0}^{\infty} a_n < \infty.$$

Therefore, $|A^*| = 0$. But the construction of F implies

$$S^*F \leq F^* + \sup_{n\in\mathbf{N}} a_n S^* Q_{\omega_n}.$$

Since F^* is a.e. finite and since the second term is also finite on $[0,1) \setminus A^*$, we conclude that S^*F is a.e. finite on $[0,1)$. ∎

Thus the set on which an a.e. divergent Walsh-Fourier series diverges unboundedly can be small in a measure theoretic sense. The following result shows it must always be large in a topological sense.

THEOREM 18. *If $f \in \mathrm{L}^1$ and Sf diverges a.e. then Sf diverges unboundedly on a dense subset of* $[0,1)$.

PROOF. Suppose Sf does not diverge unboundedly on a dense subset of $[0,1)$. Then there is an interval $[a,b] \subseteq [0,1)$ such that $(S^*f)(x) < \infty$ for $x \in [a,b]$. Let

$$A_N := \{x \in [a,b] : (S^*f)(x) \leq N\} \qquad (N \in \mathbf{N}).$$

By the Baire category theorem there is an $N \in \mathbf{N}$ and a dyadic interval $I \subseteq [0,1)$ such that $B_N := A_N \cap I$ is dense in I.

Clearly,

$$|(S_n f)(x)| \leq N$$

for $x \in B_N$ and $n \in \mathbf{N}$. Since each $S_n f$ is continuous from the right, this inequality also holds for $x \in I$ and $n \in \mathbf{N}$. Therefore,

$$|f(x)| = \lim_{n\to\infty} |(S_{2^n} f(x)| \leq N$$

for a.e. $x \in I$.

Set $f_1 := \chi(I)f$ and observe by the localization theorem that Sf_1 and Sf are equiconvergent on I. Hence Sf_1 must diverge a.e. on I. However, since f_1 belongs to L^∞, we have by Theorem 14 in 3.7 that Sf_1 converges a.e. This contradiction establishes the theorem. ∎

We have seen that if $(a_n, n \in \mathbf{N})$ belongs to ℓ^2 then $\sum a_n w_n$ converges a.e. It is natural to ask whether this result holds for some weaker growth condition on the coefficients $(a_n, n \in \mathbf{N})$. The answer to this question is no. In fact, we will show that if $(a_n, n \in \mathbf{N})$ is any decreasing sequence of positive real numbers which does not belong to ℓ^2 there is an $F \in \mathrm{H}$ which satisfies $|\widehat{F}(n)| \leq a_n$ for $n \in \mathbf{N}$ such that SF diverges everywhere.

We need two technical lemmas.

LEMMA 12. *Let $\mathbf{a} = (a_n, n \in \mathbf{N})$ be a sequence of positive numbers which decreases to zero. If $\mathbf{a} \notin \ell^2$ and $q, r \in \mathbf{N}$ then there exist positive integers m and $p_1 \le p_2 \le \cdots \le p_m$ such that*

$$\sum_{k=1}^{m} 2^{-p_k} = 1 \tag{77}$$

and

$$a_{N_k} > 2^{-p_k} \tag{78}$$

for $N_k := q + 2^r(2^{p_1} + \cdots + 2^{p_k})$ and $1 \le k \le m$.

PROOF. Since $\mathbf{a}$ is monotone, the sequence

$$\widetilde{a}_n := a_{q+n2^r} \qquad (n \in \mathbf{N})$$

also does not belong to ℓ^2. Hence we need only consider the special case $q = r = 0$.

Let $p, N \in \mathbf{N}$. If

$$a_{N+2^s} \le 2^{-s}$$

for every $s \ge p$ then the fact that $\mathbf{a}$ is monotone implies

$$\begin{aligned} \infty &= \sum_{\ell=N+2^p}^{\infty} a_\ell^2 \\ &= \sum_{s=p}^{\infty} \sum_{j=0}^{2^s-1} a_{N+2^s+j}^2 \\ &\le \sum_{s=p}^{\infty} 2^s a_{N+2^s}^2 \\ &\le \sum_{s=p}^{\infty} 2^{-s} < \infty. \end{aligned}$$

Therefore, sets of the form

$$\{s \in \mathbf{P} : a_{N+2^s} > 2^{-s}, s \ge p\}$$

are non-empty for all $p, N \in \mathbf{N}$.

Choose $p_1 \le p_2 \le \ldots$ inductively as follows. Let

$$p_1 := \min\{s \in \mathbf{P} : a_{2^s} > 2^{-s}\}.$$

If $p_1, \ldots, p_k$ have been chosen, set $N_k := 2^{p_1} + \cdots + 2^{p_k}$ and

$$p_{k+1} := \min\{s \in \mathbf{P} : a_{N_k+2^s} > 2^{-s} \quad \text{for} \quad s \ge p_k\}.$$

Then (78) holds for $q = r = 0$ and all $k \in \mathbf{P}$, and it suffices to show (77) holds for some $m \in \mathbf{P}$.

Fix $k \in \mathbf{P}$ and notice by construction that

$$a_{N_k+2^{p_k+s}} \leq 2^{-p_k-s} \tag{79}$$

for $0 \leq s < p_{k+1} - p_k$.

Suppose $p_1 < p_{k+1}$, set

$$\ell := \min\{n \leq k+1 : p_n = p_{k+1}\}$$

and observe that $1 < \ell \leq k+1$ and $p_{\ell-1} < p_\ell$. Since $\boldsymbol{a}$ is monotone it follows from the definition of p_ℓ that

$$\begin{aligned} a_{N_{k+1}} &\leq a_{N_k+2^{p_{k+1}-1}} \\ &\leq a_{N_{\ell-1}+2^{p_\ell-1}} \\ &\leq 2^{-p_\ell+1} \\ &\leq 2^{-p_{k+1}+1}. \end{aligned}$$

Therefore, $p_1 < p_{k+1}$ implies $a_{N_{k+1}} \leq 2^{-p_{k+1}+1}$.

Combine this inequality with (79). We obtain

$$\begin{aligned} \sum_{\ell=N_k+2^{p_k}}^{N_{k+1}+2^{p_{k+1}}-1} a_\ell^2 &= \sum_{s=0}^{p_{k+1}-p_k-1} \sum_{\ell=N_k+2^{p_k+s}}^{N_k+2^{p_k+s+1}-1} a_\ell^2 + \sum_{\ell=N_{k+1}}^{N_{k+1}+2^{p_{k+1}}-1} a_\ell^2 \\ &\leq \sum_{s=0}^{p_{k+1}-p_k-1} 2^{p_k+s}2^{-2(p_k+s)} + 2^{p_{k+1}}\,2^{-2(p_{k+1}-1)} \\ &\leq 4(2^{-p_k} + 2^{-p_{k+1}}) \end{aligned}$$

for any $k \in \mathbf{P}$ which satisfies $p_{k+1} > p_1$. It follows that

$$\sum_{k=1}^{\infty} 2^{-p_k} = \infty.$$

Let

$$m := \max\{n \in \mathbf{N} : \sum_{j=1}^{n} 2^{-p_j} \leq 1\}$$

and

$$M := \sum_{k=1}^{m} 2^{-p_k}.$$

Suppose that $M < 1$. Since $M = p2^{-p_m}$ for some $p \in \mathbf{P}$, we have $M \leq 1 - 2^{-p_m}$. In particular,

$$\sum_{k=1}^{m+1} 2^{-p_k} \leq 1.$$

Since this inequality contradicts the choice of m, we conclude that $M = 1$, i.e., (77) holds as promised. ∎

This lemma will be used to construct a generalized Walsh-Kolmogorov polynomial which satisfies a certain coefficient condition.

LEMMA 13. *Suppose $\mathbf{a} = (a_n, n \in \mathbf{N})$ is a monotone decreasing sequence of positive numbers with $\mathbf{a} \notin \ell^2$ and let $n, q \in \mathbf{N}$. There is a Walsh polynomial*

$$\widetilde{Q} = Q_{n,q,\mathbf{a}}$$

such that $\widetilde{Q}$ takes only the values 0 and $\pm 2^{2^n}$, $\|\widetilde{Q}\|_1 = 1, \mathcal{E}^*(\widetilde{Q}) = |\widetilde{Q}|, \Delta^*\widetilde{Q} \geq cn$ for some absolute constant $c > 0$, $sp(\widetilde{Q}) \subset [q, \infty)$, and

$$(\widetilde{Q})\widehat{\ }(s) \leq a_s$$

for $s \geq q$.

PROOF. Let $r := 2^n + n + 1$. Apply Lemma 12 to q and r, choosing positive integers $p_1, p_2, \dots, p_m$ such that (77) and (78) hold.

Set $y_0 := 1$,

$$y_j := 2^{-p_1} + 2^{-p_2} + \cdots + 2^{-p_j}$$

and

$$U_j(x) := 2^{-p_j} D_{2^{p_j}}(x \dotplus y_{j-1}) Q_n(2^{p_j} x)$$

for $1 \leq j \leq m$, and $x \in [0,1)$ where Q_n is the Walsh- Kolmogorov polynomial given by (62) above. We will choose positive integers $\lambda_1, \lambda_2, \dots, \lambda_m$ such that if $B_0 := q$ and $B_j := 1 + \max\, sp(w_{\lambda_j} U_j)$ then

$$B_{j-1} < B_j,$$

(80) $$B_j \leq N_j := q + 2^r \sum_{i=1}^{j} 2^{p_i},$$

and

(81) $$sp(w_{\lambda_j} U_j) \subseteq [B_{j-1}, B_j)$$

for $1 \leq j \leq m$.

Let $s := 2^n + n = r - 1$,

$$\ell := \min\{k \in \mathbf{P} : q \leq k 2^{s+p_1}\}$$

and set $\lambda_1 := \ell 2^{s+p_1}$. Since $N < 2^{s+p}$ and $\ell \geq 1$ imply

(82) $$w_N w_{\ell 2^{s+p}} = w_{\ell 2^{s+p}+N},$$

it is clear that

$$\begin{aligned} B_1 &\leq (\ell+1)2^{s+p_1} \\ &= (\ell-1)2^{s+p_1} + 2 \cdot 2^{s+p_1} \\ &\leq q + 2^{r+p_1}. \end{aligned}$$

Thus (80) holds for $j = 1$. Since $sp(Q_n) \subseteq [0, 2^s)$ it is clear that

$$sp(U_j) \subseteq [0, 2^{s+p_j}]$$

for $1 \le j \le m$. In particular, (81) also holds for $j = 1$.

Suppose $\lambda_1, \lambda_2, \ldots, \lambda_j$ have been chosen so that (80) and (81) hold. Let

$$\ell := \min\{k \in \mathbf{P} : B_j \le k 2^{s+p_{j+1}}\}$$

and set $\lambda_{j+1} := \ell 2^{s+p_{j+1}}$. Then (82) implies (81) with $j+1$ in place of j. Moreover, by (80) we have

$$\begin{aligned} B_{j+1} &\le (\ell+1)2^{s+p_{j+1}} \\ &= (\ell-1)2^{s+p_{j+1}} + 2 \cdot 2^{s+p_{j+1}} \\ &\le B_j + 2 \cdot 2^{s+p_{j+1}} \\ &\le q + \sum_{i=1}^{j+1} 2^{r+p_i}. \end{aligned}$$

Consequently, (80) also holds with $j+1$ in place of j.

Define $\widetilde{Q}$ by

$$\widetilde{Q} := \sum_{j=1}^{m} w_{\lambda_j} U_j.$$

Since

$$2^{-p_k} D_{2^{p_k}}(x \dot{+} y_{k-1}) = \chi[y_{k-1}, y_k)(x),$$

it is clear that $\widetilde{Q}$ takes only the values 0 and $\pm 2^{2^n}$.

By definition,

$$\begin{aligned} \|U_j\|_1 &= \int_{y_{j-1}}^{y_j} Q_n(2^{p_j} x)\, dx \\ &= 2^{-p_j} \|Q_n\|_1 \\ &= 2^{-p_j} \end{aligned}$$

for $1 \le j \le m$. Therefore,

$$\begin{aligned} \|\widetilde{Q}\|_1 &= \sum_{j=1}^{m} \|U_j\|_1 \\ &= \sum_{j=1}^{m} 2^{-p_j} \\ &= 1. \end{aligned}$$

The choice of the λ_j's implies

$$sp(w_{\lambda_j} U_j) \subseteq [\ell 2^{s+p_j}, (\ell+1)2^{s+p_j})$$

for some $\ell \in \mathbf{P}$. Hence the partial sums

$$S_{2^i}(w_{\lambda_j} U_j)$$

are either 0 or $w_{\lambda_j} U_j$ for all $i \in \mathbf{N}$. Therefore, $\mathcal{E}^*(w_{\lambda_j} U_j) = U_j$ and

$$\mathcal{E}^*(\widetilde{Q}) = |\widetilde{Q}|.$$

To estimate $\Delta^* \widetilde{Q}$ observe that if Q and D are any Walsh polynomials, extended to $\mathbf{R}$ by periodicity of period 1, and if $sp(D) \subseteq [0, 2^p)$ for some $p \in \mathbf{N}$, then the function

$$V(x) := Q(2^p x) \qquad (x \in [0,1))$$

satisfies

$$S_{\ell 2^p}(VD)(x) = D(x)(S_\ell Q)(2^p x)$$

for $\ell, p \in \mathbf{P}$ and $x \in [0,1)$. Applying this observation to $Q := Q_n$ and $D := 2^{-p_j} D_{2^{p_j}}$ we have by (66) above that

$$\Delta^* \widetilde{Q} \geq cn.$$

Since (81) implies $\widehat{\widetilde{Q}}(s) = 0$ for $s < q$, it remains to estimate these Walsh-Fourier coefficients for $s \geq q$.

Fix $s \geq q$. If

$$\widehat{\widetilde{Q}}(s) \neq 0$$

then by (81) there is a $j \in \mathbf{P}$ such that

$$s \in [B_{j-1}, B_j).$$

Moreover, by (82) and a calculation above, we have

$$\begin{aligned} |\widehat{\widetilde{Q}}(s)| &= |\widehat{w_{\lambda_j} U_j}(s)| \\ &\leq \|U_j\|_1 \\ &= 2^{-p_j}. \end{aligned}$$

But $\boldsymbol{a}$ is monotone. Therefore, it follows from (80) and the choice of s that

$$2^{-p_j} \leq a_{N_j} \leq a_{B_j} \leq a_s.$$

We conclude that

$$|\widehat{\widetilde{Q}}(s)| \leq a_s$$

for all $s \geq q$. ∎

THEOREM 19. *Let $\Phi : [0, \infty) \to [0, \infty)$ be an increasing, continuous function which satisfies $\Phi(u) = o(\log\log u)$ as $u \to \infty$. Let $\mathbf{a} = (a_n, n \in \mathbf{N})$ be a monotone decreasing sequence of positive numbers with*

$$\sum_{n=0}^{\infty} a_n^2 = \infty.$$

Then there exists a function $F \in \mathrm{L}^1$ such that

$$|\widehat{F}(n)| \le a_n \qquad (n \in \mathbf{N}),$$

(83)
$$\int_0^1 F^*\Phi(F^*) < \infty,$$

and SF diverges everywhere.

PROOF. Choose a sequence $(\varepsilon_n, n \in \mathbf{N})$ of positive numbers converging to zero which satisfies

$$\Phi(n) = \varepsilon_n \log_2 \log_2 n \qquad (n \ge 4).$$

Choose a strictly increasing sequence $(\omega_n, n \in \mathbf{N})$ of positive integers such that $\omega_0 \ge 4$ and

(84)
$$\sum_{n=0}^{\infty} \varepsilon_{\ell_n} < \infty$$

for $\ell_n := 2^{2^{\omega_n}}, n \in \mathbf{N}$.

Use Lemma 13 to generate Walsh polynomials $\widetilde{Q}_k$ in the following way. Let $q_0 := 0$ and set $\widetilde{Q}_0 := Q_{\omega_0, q_0, \mathbf{a}}$. If $\widetilde{Q}_{k-1}$ has been chosen, let

$$q_k := 1 + \max\ sp(\widetilde{Q}_{k-1})$$

and set

$$\widetilde{Q}_k := Q_{\omega_k, q_k, \mathbf{a}}.$$

Thus $\widetilde{Q}_k$ takes on only the values 0 and $\pm\ell_k$, $\|\widetilde{Q}_k\|_1 = 1$, $\mathcal{E}^*(\widetilde{Q}_k) = |\widetilde{Q}_k|$, $\Delta^*\widetilde{Q}_k \ge c\omega_k$ for some absolute constant $c > 0$, $sp(\widetilde{Q}_k) \subset [q_k, \infty)$, and

$$|\widehat{\widetilde{Q}}_k(s)| \le a_s$$

for $s \ge q_k$.

Let

$$F := \sum_{k=0}^{\infty} \frac{\varepsilon_{\ell_k}}{\Phi(\ell_k)} \widetilde{Q}_k = \sum_{k=0}^{\infty} \omega_k^{-1} \widetilde{Q}_k.$$

By (84) this series converges in L^1 norm. Thus F belongs to L^1 and is defined up to sets of measure zero. Because the spectra of the $\widetilde{Q}_k$'s are contained in pairwise disjoint intervals, it is clear that

$$|\widehat{F}(s)| \le a_s$$

for $s \in \mathbf{N}$. Moreover, since $\omega_k \varepsilon_{\ell_k}/\Phi(\ell_k) = 1$ for $k \in \mathbf{N}$, the estimate of $\Delta^*\widetilde{Q}_k$ implies that SF diverges everywhere. It remains to verify (83).

Set

$$A_k := \{\widetilde{Q}_k \neq 0 \quad \text{and} \quad \widetilde{Q}_j = 0 \quad \text{for all} \quad j > k\}$$

and let

$$A^* := \limsup_{k\to\infty}\{\widetilde{Q}_k \neq 0\}.$$

Since $\|\widetilde{Q}_k\|_1 = 1$ and $\widetilde{Q}_k$ takes only the values 0 and $\pm\ell_k$, it is clear that

$$|A_k| \le |\{\widetilde{Q}_k \neq 0\}| = 2^{-2^{\omega_k}}$$

for $k \in \mathbf{N}$. Hence $|A^*| = 0$. Since the A_k's are pairwise disjoint it follows that

$$\int_0^1 F^*\Phi(F^*) = \sum_{k=0}^{\infty} \int_{A_k} F^*\Phi(F^*). \tag{85}$$

Fix $k \in \mathbf{N}$ and observe by the definition of A_k that

$$\begin{aligned}\chi(A_k)F^* &= \sum_{j=0}^{k} \omega_j^{-1}\widetilde{Q}_j\\ &\le \sum_{j=0}^{k} \omega_j^{-1}\mathcal{E}^*(\widetilde{Q}_j)\\ &= \sum_{j=0}^{k} \omega_j^{-1}|\widetilde{Q}_j|.\end{aligned}$$

Since $\widetilde{Q}_j \le \ell_j$ we have

$$\begin{aligned}\chi(A_k)F^* &\le \sum_{j=0}^{k} \omega_j^{-1}2^{2^{\omega_j}}\\ &\le \frac{2}{\omega_k}2^{2^{\omega_k}}\\ &\le \ell_k.\end{aligned}$$

Consequently,

$$\begin{aligned}\chi(A_k)F^*\Phi(F^*) &\le 2\chi(A_k)\omega_k^{-1}\ell_k\Phi(\ell_k)\\ &= 2\chi(A_k)\ell_k\,\varepsilon_{\ell_k}.\end{aligned}$$

Since $|A_k| \le 1/\ell_k$, we conclude from (85) that

$$\int_0^1 F^*\Phi(F^*) \le 2\sum_{k=0}^{\infty} \int_{A_k} \ell_k \varepsilon_{\ell_k} \le 2\sum_{k=0}^{\infty} \varepsilon_{\ell_k} < \infty. \quad \blacksquare$$

We close this section by constructing some a.e. divergent Walsh- Kaczmarz-Fourier series. These are easier to construct than divergent Walsh-Fourier series because the Walsh-Kaczmarz-Dirichlet kernels satisfy

$$\limsup_{n\to\infty} \frac{|D_n^\kappa(x)|}{\log n} > c \tag{86}$$

for a.e. $x \in [0,1)$ where $c > 0$ is an absolute constant (see Theorem 11 in 1.6).

This condition enters the picture in the following way.

LEMMA 14. *Let $(a_n, n \in \mathbf{N})$ be a decreasing sequence of positive numbers which converge to zero such that*

$$na_n > \delta > 0 \qquad (n \in \mathbf{N}). \tag{87}$$

If

$$F := \sum_{n=0}^{\infty} a_n r_n D_{2^n} \tag{88}$$

then F belongs to L^1 and its Walsh-Kaczmarz-Fourier series diverges a.e.

PROOF. By the Paley lemma, (88) converges pointwise on $(0,1)$.

Let

$$F_m := \sum_{n=0}^{m} a_n r_n D_{2^n} \qquad (m \in \mathbf{N})$$

and observe by Abel's transformation that

$$\begin{aligned} F_m &= \sum_{n=0}^{m} a_n (D_{2^{n+1}} - D_{2^n}) \\ &= \sum_{n=0}^{m-1} (a_n - a_{n+1}) D_{2^{n+1}} + a_m D_{2^{m+1}} - a_0 D_{2^0}. \end{aligned} \tag{89}$$

Since $\|D_{2^n}\|_1 = 1$ for $n \in \mathbf{N}$, it follows that

$$\|F_M - F_m\|_1 \le \sum_{n=m}^{M-1} (a_n - a_{n+1}) + a_m + a_M = 2a_m \tag{90}$$

for $M > m$. Therefore, (88) converges in L^1 norm and F belongs to L^1.

Let $S_m^\kappa F$ denote the m-th partial sum of the Walsh-Kaczmarz-Fourier series of F. Since $D_{2^n}^\kappa(x) = D_{2^n}(x) = 0$ for each $x \in (0,1)$ and n sufficiently large, we have by (87) that

$$\begin{aligned}|S_{2^n+m}^\kappa F - S_{2^n}^\kappa F| &= a_n|D_{2^n+m}^\kappa - D_{2^n}| \\ &> cna_n \\ &> c\delta > 0\end{aligned}$$

a.e. on $[0,1)$ for infinitely many $n \in \mathbf{N}$ and $m = m(n) < 2^n$. Hence the Walsh-Kaczmarz-Fourier series of F diverges a.e. ∎

As a first application of this technique we prove the following result (compare with Theorems 14 and 15 above).

THEOREM 20. *Suppose* $\Phi : [0,\infty) \to [0,\infty)$ *is a continuous, increasing function which satisfies*

$$\sum_{n=1}^{\infty} \frac{\Phi(2^n)}{n^2} < \infty.$$

There is a function $F \in L^1$ *whose Walsh-Kaczmarz-Fourier series diverges a.e. such that*

$$\omega^{(1)}(F,\delta) = O\left(\frac{1}{\log(1/\delta)}\right) \quad \text{as} \quad \delta \to 0 \tag{91}$$

and

$$\int_0^1 |F|\Phi(|F|) < \infty.$$

(In particular, F *belongs to* $L(\log^+ L)^p$ *for all* $0 < p < 1$.*)*

PROOF. Let $a_0 := 1$ and $a_n := 1/n$ for $n \in \mathbf{P}$, and define F by (88). By Lemma 14 the Walsh-Kaczmarz-Fourier series of F diverges a.e.

Use (89) to verify that

$$F = \sum_{n=0}^{\infty}(a_n - a_{n+1})D_{2^{n+1}} - 1.$$

Set $J_N := [2^{-N-1}, 2^{-N})$ and observe that F is constant on each J_N for $N \in \mathbf{N}$. Indeed, we have

$$F = \sum_{n=0}^{N-1}(a_n - a_{n+1})2^{n+1} - 1$$

on J_N. Therefore the choice of the a_n's implies

$$|F(x)| \leq \frac{8 \cdot 2^N}{N^2}$$

for $x \in J_N$ and $N \in \mathbf{P}$. Consequently,

$$\begin{aligned}\int_0^1 |F|\,\Phi(|F|) &= \int_{1/8}^1 |F|\,\Phi(|F|) + \sum_{N=3}^{\infty} \int_{J_N} |F|\,\Phi(|F|) \\ &\leq C + 4\sum_{N=3}^{\infty} \frac{\Phi(2^N)}{N^2} < \infty\end{aligned}$$

where C is finite and depends only on Φ. Thus it remains to verify (91).

Fix $0 < y < 1$ and choose $N \in \mathbf{N}$ such that $2^{-N-1} \leq y < 2^{-N}$. Surely

$$F = (F - F_{N-1}) + F_{N-1}$$

(see (89) above). Moreover, $F_{N-1}(x \dot{+} y) = F_{N-1}(x)$ for $x \in [0,1)$ and by (90) we have

$$\|F - F_{N-1}\|_1 \leq \frac{2}{N-1} < \frac{4}{N}.$$

Therefore

$$\begin{aligned}\|F - \boldsymbol{\tau}_y F\|_1 &= \|(F - F_{N-1}) - \boldsymbol{\tau}_y(F - F_{N-1})\|_1 \\ &\leq \frac{8}{N} \\ &\leq \frac{16}{\log_2(1/y)},\end{aligned}$$

and (91) follows at once. ∎

The same technique gives us the following Kaczmarz version of Corollary 11 ii).

THEOREM 21. *Let $(\varepsilon_n, n \in \mathbf{N})$ be a decreasing sequence of positive numbers which converges to zero. There exists an $F \in \mathrm{L}^1$ such that*

$$\limsup_{m\to\infty} \frac{|S_m^{\kappa} F|}{\varepsilon_m \log m} = \infty$$

a.e. on $[0,1)$.

PROOF. We may suppose that $\varepsilon_n \log n \to \infty$ as $n \to \infty$. Set $a_n := \sqrt{\varepsilon_{2^n}}$ and define F by (88). Let $2^n \leq m < 2^{n+1}$ and observe that

$$\begin{aligned}\frac{|S_m^{\kappa} F|}{\varepsilon_m \log m} &\geq \frac{\sqrt{\varepsilon_{2^n}}\,|D_m^{\kappa} - D_{2^n}|}{\varepsilon_m \log m} - \frac{|S_{2^n} F|}{\varepsilon_m \log m} \\ &\geq \frac{|D_m^{\kappa}|}{\sqrt{\varepsilon_m}\log m} - \frac{D_{2^n}}{\sqrt{\varepsilon_{2^n}}\log 2^n} - \frac{|S_{2^n} F|}{\varepsilon_m \log m}.\end{aligned}$$

Since the last two terms of this expression tend to zero a.e. on $[0,1)$, we appeal to (86) to finish the proof. ∎

6.6 Almost Everywhere Convergence of Double Walsh-Fourier Series. Let $\mathbf{N}^2 := \mathbf{N} \times \mathbf{N}$, $\mathbf{I}^2 := [0,1) \times [0,1)$, and let $\widetilde{\boldsymbol{w}} := (\widetilde{w}_{m,n}, (m,n) \in \mathbf{N}^2)$ represent the Kronecker product system generated by the Walsh system $\boldsymbol{w}$.

If $f \in \mathrm{L}^2(\mathbf{I}^2)$ then the double Walsh-Fourier coefficients of f are defined by

$$\widehat{f}(m,n) := \iint_{\mathbf{I}^2} f(x,y)\widetilde{w}_{m,n}(x,y)\,dxdy$$

for $(m,n) \in \mathbf{N}^2$. The double Walsh-Fourier series of f is the series

$$Sf := \sum_{(m,n)\in\mathbf{N}^2} \widehat{f}(m,n)\widetilde{w}_{m,n}.$$

As in the one-dimensional case, this is a formal definition. Convergence of Sf is neither implied nor presumed.

The rectangular partial sums of Sf are defined by

$$S_{M,N}f := \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \widehat{f}(m,n)\widetilde{w}_{m,n},$$

for $M, N \in \mathbf{P}$ and $f \in \mathrm{L}^1(\mathbf{I}^2)$. If

$$D_{M,N} := \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \widetilde{w}_{m,n},$$

then these definitions imply

$$(S_{M,N}f)(x,y) = \iint_{\mathbf{I}^2} D_{M,N}(x \dotplus u, y \dotplus v)f(u,v)\,dudv.$$

Hence,

$$(S_{2^M,2^N}f)(x,y) = 2^{M+N}\int_{I_M(x)}\int_{I_N(y)} f(u,v)\,dudv$$

for all $(x,y) \in \mathbf{I}^2$ and $M, N \in \mathbf{N}$. In particular,

$$\lim_{M,N\to\infty} S_{2^M,2^N}f$$

may not exist even if f is continuous.

Analogy with the one-dimensional case is restored if we pass from rectangular partial sums to square ones. For if $I_N(x,y) := I_N(x) \times I_N(y)$ then

$$(S_{2^N,2^N}f)(x,y) = \frac{1}{|I_N(x,y)|}\int_{I_N(x,y)} f(u,v)\,dudv.$$

It follows that $S_{2^N,2^N}f$ converges to f as $N \to \infty$ uniformly if f is continuous and a.e. if $f \in \mathrm{L}^1(\mathbf{I}^2)$.

What about the full sequence of square sums?

THEOREM 22. *If* $f \in L^2(\mathbf{I}^2)$ *then* $S_{N,N}f$ *converges to* f *a.e. on* $\mathbf{I}^2$ *as* $N \to \infty$.

PROOF. By Theorem 2 in 3.1 it suffices to show

$$\| \sup_{N\in\mathbf{N}} |S_{N,N}f| \|_{L^2(\mathbf{I}^2)} \le C\|f\|_{L^2(\mathbf{I}^2)} \tag{92}$$

for all $f \in L^2(\mathbf{I}^2)$, where C is an absolute constant which does not depend on f.

Fix $f \in L^2(\mathbf{I}^2)$ and suppose first that

$$\widehat{f}(m,n) = 0 \quad \text{for} \quad m < n. \tag{93}$$

Let $f_y(x) := f(x,y)$ for $x, y \in [0,1)$, and observe by Fubini's theorem that f_y belongs to L^2 for a.e $y \in [0,1)$. Hence by Corollary 11 in 3.7 we have

$$\|S^* f_y\|_2 \le C_2 \|f_y\|_2 \tag{94}$$

for a.e. $y \in [0,1)$ and some absolute constant C_2.

Set

$$g_m(y) := \int_0^1 f_y(x) w_m(x)\, dx \qquad (m \in \mathbf{N}, y \in [0,1))$$

and observe that

$$\begin{aligned} \|g_m\|_2 &:= \left(\int_0^1 \left| \int_0^1 f_y(x) w_m(x)\, dx \right|^2 dy \right)^{1/2} \\ &\le \left(\int_0^1 \int_0^1 |f_y(x)|^2 \, dy dx \right)^{1/2} \\ &= \|f\|_{L^2(\mathbf{I}^2)}. \end{aligned}$$

Hence $g_m \in L^2$ for all $m \in \mathbf{N}$. Moreover, since

$$\widehat{g}_m(n) = \int_0^1 g_m(y) w_n(y)\, dy = \widehat{f}(m,n),$$

it is clear by (93) that each g_m is a Walsh polynomial. Consequently,

$$\begin{aligned} (S_N f_y)(x) &= \sum_{m=0}^{N-1} \left(\sum_{n=0}^{m} \widehat{f}(m,n) w_n(y) \right) w_m(x) \\ &= (S_{N,N} f)(x,y) \end{aligned}$$

for $N \in \mathbf{N}$ and a.e. $x, y \in [0,1)$. In particular, (92) follows from (94) and the inequality is established when (93) is satisfied.

Reversing the roles of x and y we see that (92) also holds when $f \in L^2(\mathbf{I}^2)$ satisfies

$$\widehat{f}(m,n) = 0 \quad \text{for} \quad m \ge n.$$

Since every $f \in L^2(\mathbf{I}^2)$ can be written as $f = g + h$ where $\widehat{g}(m,n) = 0$ for $m < n$ and $\widehat{h}(m,n) = 0$ for $m \geq n$, we conclude that (92) holds for all $f \in L^2(\mathbf{I}^2)$. ∎

To consider partial sums over regions other than rectangles and squares, we introduce the following notation.

For the next several pages, let R be a bounded region in the first quadrant of $\mathbf{R}^2$ which contains the origin. Let

$$R(\gamma) := \{(m,n) \in \mathbf{N}^2 : (m/\gamma, n/\gamma) \in R\}$$

represent the dilation of R by $\gamma > 0$. For each $f \in L^1(\mathbf{I}^2)$ define R-sums of Sf by

$$S_{R(\gamma)}f := \sum_{(m,n)\in R(\gamma)} \widehat{f}(m,n)\widetilde{w}_{m,n}.$$

The double Walsh-Fourier series of an $f \in L^1(\mathbf{I}^2)$ is said to be R-summable to B at a point $(x,y) \in \mathbf{I}^2$ if

$$B = \lim_{\gamma\to\infty}(S_{R(\gamma)}f)(x,y).$$

Notice when R is a square with sides parallel to the axes that Sf is R-summable to B at (x,y) if and only if

$$B = \lim_{N\to\infty}(S_{N,N}f)(x,y).$$

Define a maximal function for R-summability by

$$\mathbf{S}^*f := \sup_{\gamma>0}|S_{R(\gamma)}f|.$$

By the usual density argument, one can show that a sufficient condition for a.e. R-summability of Sf for all $f \in L^p$, $1 \leq p < \infty$, is that $\mathbf{S}^*$ is of weak-type (p,p). We will show (Theorem 23 below) that this condition is also necessary. This result is a special case of a general theorem by Stein [1] valid for sequences of translation invariant operators. The proof here is an adaptation of the general case.

LEMMA 15. *Let $y > 0, f \in L^p(\mathbf{I}^2), 1 \leq p < \infty$, and suppose*

$$|\{\mathbf{S}^*f > y\}| > C$$

for some number $C > 0$. Then there is a double Walsh polynomial P such that

$$\|P\|_{L^p(\mathbf{I}^2)} \leq \|f\|_{L^p(\mathbf{I}^2)}$$

and

$$|\{\mathbf{S}^*P > y\}| > C.$$

PROOF. For each $\gamma > 0$ let $A_\gamma := \{|S_{R(\gamma)}f| > y\}$. Then $\{\mathbf{S}^*f > y\}$ is the union of the sets A_γ for $\gamma > 0$, and it follows from hypothesis that

$$|\bigcup_{0<\gamma\leq L} A_\gamma| > C$$

for some $L > 0$. Since R is bounded, we can choose an $M \in \mathbf{N}$ so large that

$$S_{R(\gamma)}P = S_{R(\gamma)}f$$

for $P := S_{2^M,2^M}f$. Since

$$\|S_{2^M,2^M}\|_{L^p(\mathbf{I}^2)} = 1,$$

P is the polynomial for which we search. ∎

LEMMA 16. *Suppose $(a_n, n \in \mathbf{N})$ is a sequence of positive numbers which converges to zero and $(\omega(n), n \in \mathbf{N})$ is a sequence of real numbers which tends to ∞. There is a non-decreasing sequence of non-negative integers $n_1, n_2, \ldots$ such that*

$$\sum_{k=1}^{\infty} a_{n_k} < \infty$$

but

$$\sum_{k=1}^{\infty} \omega(n_k) a_{n_k} = \infty.$$

PROOF. Set $\ell_0 := 0$. Given ℓ_{k-1} $k \in \mathbf{P}$, choose $\ell_k \geq \ell_{k-1}$ so that $b_k := a_{\ell_k} < 2^{-k}$ and $\omega(\ell_k) > 2^k$. Set $c_k := b_k \omega(\ell_k)$ and choose a positive integer m_k so that

$$(m_k - 1)b_k < 2^{-k} \leq m_k b_k.$$

Since

$$m_k c_k = m_k b_k \omega(\ell_k) > 2^{-k} 2^k = 1,$$

it is clear that $\sum_{k=0}^{\infty} m_k c_k = \infty$. On the other hand,

$$\begin{aligned}\sum_{k=0}^{\infty} m_k b_k &= \sum_{k=0}^{\infty} b_k + \sum_{k=0}^{\infty} (m_k - 1) b_k \\ &\leq 2 \sum_{k=0}^{\infty} (m_k - 1) b_k \\ &\leq 2 \sum_{k=0}^{\infty} 2^{-k} \\ &< \infty.\end{aligned}$$

Thus the sequence $(n_k, k \in \mathbf{P})$ can be taken to be $\ell_1, \ldots, \ell_1, \ell_2, \ldots \ell_2, \ldots$, where each ℓ_j is repeated m_j times. ∎

LEMMA 17. *If $A_1, A_2, \ldots$ are measurable sets in $\mathbf{I}^2$ which satisfy $\sum_{n=1}^{\infty} |A_n| = \infty$, then there exist points $x_1, x_2, \ldots$ in $\mathbf{I}^2$ such that a.e. $x \in \mathbf{I}^2$ belongs to infinitely many of the sets $x_n \dot{+} A_n$ $(n \in \mathbf{P})$.*

PROOF. Given $A \subseteq \mathbf{I}^2$ we shall denote $\mathbf{I}^2 \setminus A$ by $\overline{A}$.

The set of points belonging to finitely many of the sets $x_n \dot{+} A_n$ is

$$\bigcup_{p=1}^{\infty} \bigcap_{n=1}^{\infty} \overline{(x_{p+n} \dot{+} A_{p+n})}$$

and thus is a subset of

$$\bigcup_{\ell=0}^{\infty} \bigcap_{n=p_\ell+1}^{p_{\ell+1}} \overline{(x_n \dot{+} A_n)}$$

for any choice of $0 =: p_0 < p_1 < p_2 < \dots$

Let $\varepsilon > 0$, $n \in \mathbf{P}$, and $\chi_n := \chi(\overline{A}_n)$. Then the characteristic function of

$$\bigcap_{n=1}^{p_1} \overline{(x_n \dot{+} A_n)}$$

is $\chi_1(t \dot{+} x_1)\chi_2(t \dot{+} x_2)\dots\chi_{p_1}(t \dot{+} x_{p_1})$, $(t \in \mathbf{I}^2)$. Moreover,

$$\int_0^1 \cdots \int_0^1 \chi_1(t \dot{+} x_1)\dots\chi_{p_1}(t \dot{+} x_{p_1})\, dt dx_1 \dots dx_{p_1}$$
$$= \prod_{n=1}^{p_1}(1 - |A_n|).$$

Since the hypothesis $\sum_{n=1}^{\infty} |A_n| = \infty$ implies

$$\prod_{n=1}^{p_1}(1 - |A_n|) < \frac{\varepsilon}{2}$$

for p_1 large, it follows that there exist points $x_1, \dots x_{p_1}$ such that

$$\left| \bigcap_{n=1}^{p_1} \overline{(x_n \dot{+} A_n)} \right| < \frac{\varepsilon}{2}.$$

Continuing this process we choose positive integers $p_1 < p_2 < p_3 < \dots$ and points x_n for each $p_\ell < n \le p_{\ell+1}$ such that

$$| \bigcap_{n=p_\ell+1}^{p_{\ell+1}} \overline{(x_n \dot{+} A_n)}| < \left(\frac{1}{2}\right)^{\ell+1} \varepsilon$$

for $\ell = 0, 1, \dots$. Therefore,

$$\sum_{\ell=0}^{\infty} \left| \bigcap_{n=p_\ell+1}^{p_{\ell+1}} \overline{(x_n \dot{+} A_n)} \right| < \varepsilon.$$

We conclude that

$$\left| \bigcup_{p=1}^{\infty} \bigcap_{n=1}^{\infty} \overline{(x_{p+n} \dot{+} A_{p+n})} \right| = 0. \qquad \blacksquare$$

THEOREM 23. *Let $1 \le p \le 2$. If Sf is R- summable a.e. on $\mathbf{I}^2$ for every $f \in \mathrm{L}^p(\mathbf{I}^2)$ then $\mathbf{S}^*$ is of weak-type (p, p).*

PROOF. Suppose to the contrary that $\mathbf{S}^*$ is not of weak-type (p, p). Then for each $n \in \mathbf{P}$ there is a function f_n and positive numbers $\omega(n), y_n$ such that $\omega(n) \to \infty$,

$$\|f_n\|_{L^p(\mathbf{I}^2)} = 1,$$

and

$$|\{\mathbf{S}^* f_n > y_n\}| > \omega(n) y_n^{-p}.$$

We claim without loss of generality that each f_n is a polynomial with $\widehat{f}_n(0,0) = 0$, such that

(95) $$\sum_{n=0}^{\infty} y_n^{-p} < \infty,$$

(96) $$\sum_{n=0}^{\infty} \omega(n) y_n^{-p} = \infty,$$

(97) $$\sum_{n=0}^{\infty} |\{\mathbf{S}^* f_n > y_n\}| = \infty,$$

and

(98) $sp(f_n)$ are pairwise disjoint for $n \in \mathbf{N}$.

To see this notice first by Lemma 15 we may suppose each f_n is a polynomial. Also, since $\omega(n) \to \infty$ implies $y_n \to \infty$, we discard finitely many f_n's and suppose that $y_n > 1$. Hence by replacing f_n by

$$\frac{(f_n - \widehat{f}_n(0,0))}{\|f_n - \widehat{f}_n(0,0)\|_p}$$

we may suppose $\widehat{f}_n(0,0) = 0$ and $\|f_n\|_{L^p(\mathbf{I}^2)} = 1$.

Next choose by Lemma 16 a non-decreasing sequence $n_1, n_2, \ldots$ such that

$$\sum_{k=1}^{\infty} y_{n_k}^{-p} < \infty$$

and

$$\sum_{k=1}^{\infty} \omega(n_k) y_{n_k}^{-p} = \infty.$$

Since this last equation implies $\sum_{k=0}^{\infty} |\{\mathbf{S}^* f_{n_k} > y_{n_k}\}| = \infty$, we obtain (95), (96), and (97) by replacing $(f_n, n \in \mathbf{N})$ with $(f_{n_k}, k \in \mathbf{P})$.

Finally, extend each f_n to $\mathbf{R}^2$ by periodicity of period 1 and define

$$\Delta^k f_n(x,y) := f_n(2^k x, 2^k y) \qquad (k \in \mathbf{N}, (x,y) \in \mathbf{R}^2).$$

Since $\widehat{f}_n(0,0) = 0$ we choose, by induction, an increasing sequence $(k_n, n \in \mathbf{N})$ of positive integers such that the spectrum of each polynomial $\Delta^{k_n} f_n$ does not overlap with the spectra of $\Delta^{k_j} f_j$ for $j < n$, and also does not overlap with

$$\bigcup_{0<\gamma\leq \ell_n} R(\gamma),$$

where the number ℓ_n is determined by

$$\mathbf{S}^*(\Delta^{k_j} f_j) = \sup_{0<\gamma\leq \ell_n} |S_{R(\gamma)}(\Delta^{k_j} f_j)|$$

for $j < n$. Thus replacing f_n by $\Delta^{k_n} f_n$ results in (98). This replacement does not alter (97) because

$$S_{R(\gamma 2^k)}(\Delta^k f) = \Delta^k S_{R(\gamma)}(f) \qquad (\gamma > 0)$$

implies

$$|\{\mathbf{S}^* f > y_n\}| = |\{\mathbf{S}^*(\Delta^k f) > y_n\}|$$

for any $k \in \mathbf{P}$ and $f \in \mathrm{L}^1(\mathbf{R}^2)$. Therefore, the claim is established.

Choose by Lemma 17 a sequence $x_1, x_2 \ldots$ in $\mathbf{I}^2$ such that a.e. $x \in \mathbf{I}^2$ belongs to infinitely many of the sets $x_n \dot{+} \{\mathbf{S}^* f_n > y_n\}$. For $x \in \mathbf{I}^2$ let $P_n(x) := f_n(x \dot{+} x_n)/y_n$ and notice that $\|P_n\|_{L^p(\mathbf{I}^2))} = y_n^{-1}$. Consequently, it follows from a two-dimensional version of Paley's inequality that

$$\iint_{\mathbf{I}^2} \left(\sum_{n=1}^{\infty} \sum_{j=0}^{\infty} |\Delta_j P_n|^2 \right)^{p/2} \leq \sum_{n=1}^{\infty} \iint_{\mathbf{I}^2} \left(\sum_{j=0}^{\infty} |\Delta_j P_n|^2 \right)^{p/2}$$
$$\leq \sum_{n=1}^{\infty} A_p^p \|P_n\|_p^p < \infty,$$

where

$$\Delta_j f(x,y) := \frac{1}{|I_{j+1}(x,y)|} \iint_{I_{j+1}(x,y)} f - \frac{1}{|I_j(x,y)|} \iint_{I_j(x,y)} f$$

for $j \in \mathbf{N}, f \in \mathrm{L}^1(\mathbf{I}^2)$, and $(x,y) \in \mathbf{I}^2$. Since by construction

$$(\Delta_j P_{n_1})(\Delta_j P_{n_2}) = 0$$

for $n_1 \neq n_2$, it follows again from a two-dimensional version of Paley's inequality that

$$f := \sum_{n=1}^{\infty} P_n \in \mathrm{L}^p(\mathbf{I}^2).$$

However, if $x \dot{+} x_n \in \{\mathbf{S}^* f_n > y_n\}$ for some $x \in \mathbf{I}^2$ then there exist γ_1, γ_2 such that $(S_{R(\gamma_1)} f - S_{R(\gamma_2)} f)(x)$ exceeds $y_n\, y_n^{-1} = 1$. Thus $S_{R(\gamma)} f$ diverges as $\gamma \to \infty$, at any point belonging to infinitely many of the sets $x_n \dot{+} \{\mathbf{S}^* f_n > y_n\}$, i.e., diverges almost everywhere on $\mathbf{I}^2$. ∎

This result in conjunction with the Marcinkiewicz interpolation theorem proves the following.

COROLLARY 12. *Let $1 \le p < 2$. If Sf is R-summable a.e. on $\mathbf{I}^2$ for all $f \in L^p(\mathbf{I}^2)$ then $\mathbf{S}^*$ is of type (q,q) for all $p < q < 2$. In particular, the operators $S_{R(\gamma)}, \gamma > 0$, are uniformly bounded on $L^q(\mathbf{I}^2)$ for all $p < q < 2$.*

Thus we see that if the operators $S_{R(\gamma)}, \gamma > 0$, are not uniformly bounded on some $L^{p_0}(\mathbf{I}^2)$ for $1 < p_0 < 2$, then given $1 \le p < p_0$ there exists an $f \in L^p(\mathbf{I}^2)$ such that the R-partial sums of the double Walsh-Fourier series of f diverge on a set of positive measure.

We apply this observation first to triangular sums.

THEOREM 24. *Let $\beta > 0$ and let $R := \{(x,y) \in [0,\infty)^2 : y < 1 - \beta x\}$ be the triangular region bounded by the axes and the line $y = 1 - \beta x$. Then there exist $f \in L^p(\mathbf{I}^2)$ for $1 \le p < 2$ such that $S_{R(\gamma)} f$ diverges on a set of positive measure, as $\gamma \to \infty$.*

PROOF. It suffices to show that the operators $S_{R(\gamma)}, \gamma > 0$, are not uniformly bounded on $L^p(\mathbf{I}^2)$ for $1 \le p \ne 2$.

Fix $p \ne 2$ and recall from Theorem 19 in 5.5 that the operator

$$Vf := \sum_{n<\beta m} \widehat{f}(m,n)\widetilde{w}_{m,n}$$

is not uniformly bounded on $L^p(\mathbf{I}^2)$. Fix $M \in \mathbf{P}$. Let f be a double Walsh polynomial which satisfies $\widehat{f}(m,n) = 0$ for $m, n \ge 2^M$ and set $\gamma := \beta(2^M - 1)$. Then

$$\begin{aligned}
S_{R(\gamma)}(\widetilde{w}_{2^M-1,0} f) &= S_{R(\gamma)}\left(\sum_{k,n=0}^{2^M-1} \widehat{f}(2^M - k - 1, n)\widetilde{w}_{k,n}\right) \\
&= \sum_{\beta k + n < \beta(2^M-1)} \widehat{f}(2^M - k - 1, n)\widetilde{w}_{k,n} \\
&= \sum_{n<\beta m} \widehat{f}(m,n)\widetilde{w}_{2^M-m-1,n} \\
&= \widetilde{w}_{2^M-1,0} \sum_{n<\beta m} \widehat{f}(m,n)\widetilde{w}_{m,n} \\
&= \widetilde{w}_{2^M-1,0} Vf.
\end{aligned}$$

Hence $S_{R(\gamma)}, \gamma > 0$, are not uniformly bounded on $L^p(\mathbf{I}^2)$. ∎

This result can be extended to any region R provided part of its boundary is piecewise smooth and not parallel to the axes.

THEOREM 25. *Let h be a continuously differentiable non-negative function defined on some $[a,b] \subset [0,\infty)$. Suppose that there is a point $x_0 \in (a,b)$ such that*

$$\frac{d}{dx} h(x_0) < 0.$$

Let R denote the sector-like region bounded by the graph of h and the line segments from (0,0) through $(a, h(a))$ and from (0,0) through $(b, h(b))$. Then there exists an $f \in L^p(\mathbf{I}^2)$

for $1 \le p < 2$ such that $S_{R(\gamma)}f$ diverges, as $\gamma \to \infty$, on some subset of $\mathbf{I}^2$ of positive measure.

PROOF. Choose a rational number $\alpha > 0$ such that the line $y = \alpha x$ intersects the graph of h at $(c, h(c))$ for some $c \in (a, b)$. Clearly,

$$\frac{h(c)}{c} = \alpha \tag{99}$$

Moreover, if

$$\beta := -\frac{d}{dx}h(c), \tag{100}$$

we may assume that $\beta > 0$.

The idea behind our proof is this. Near $(c, h(c))$ the boundary of R looks like a line with slope $-\beta$. Thus for large γ the boundary of $R(\gamma)$ lies very close to a line with slope $-\beta$ for x in some huge interval. In particular, the partial sums $S_{R(\gamma)}f$ should eventually be like triangular partial sums, hence diverge by the previous theorem. Here are the details.

For each $\gamma > 0$ let $h_\gamma(x) := \gamma\, h(x/\gamma)$ for $x \in [\gamma a, \gamma b]$. Then the graph of h_γ forms part of the boundary of $R(\gamma)$.

Let $\gamma_0 := 1/(h(c) + \beta c)$ and for each $\gamma > 0$ let

$$\ell_\gamma(x) := -\beta x + \gamma/\gamma_0 \qquad (x \in \mathbf{R}).$$

We will show (see (102) and (103) below) that given $M \in \mathbf{N}$ the graph $y = h_\gamma(x)$ is near the line $y = \ell_\gamma(x)$ for all x in an interval of length 2^M.

Fix $M \in \mathbf{N}$ and set $\gamma_1 := \beta(2^M - 1)$. Let T be the triangular region in the first quadrant of $\mathbf{R}^2$ bounded by the axes and the line $y = 1 - \beta x$. Since $\mathbf{N}^2$ is discrete in $\mathbf{R}^2$, choose an $\varepsilon_0 > 0$ such that $\gamma_1 - \varepsilon_0 > 0$ and

$$S_{T(\gamma_1 - \varepsilon)}f = S_{T(\gamma_1)}f$$

for $0 < \varepsilon < \varepsilon_0$.

Let $\eta := 2^{-M-1}\varepsilon_0$ and choose $\delta > 0$ with $a < c - \delta < c + \delta < b$ such that

$$\left|\frac{d}{dx}h_\gamma(x) + \beta\right| < \eta \tag{101}$$

for all $\gamma > 0$ and $|x - \gamma c| \le \gamma\delta$.

Solve the equation defining γ_0 for $h(c)$ and verify that

$$h_{\gamma_0}(\gamma_0 c) = -\beta\gamma_0 c + 1.$$

This identity implies

$$h_\gamma(\gamma c) = \frac{\gamma}{\gamma_0}h_{\gamma_0}(\gamma_0 c) = \ell_\gamma(\gamma c).$$

Consequently,

$$\text{(102)} \qquad \begin{cases} \text{the line } \ell_\gamma \text{ passes through the point} \\ (\gamma c, h_\gamma(\gamma c)) \quad \text{for all} \quad \gamma > 0. \end{cases}$$

Moreover, by (101) and the mean value theorem, if γ is so large that $2^M < \gamma\delta$ then

$$\text{(103)} \qquad |h_\gamma(x) - \ell_\gamma(x)| < \eta\, 2^M = \frac{\varepsilon_0}{2}$$

for $\gamma > 0$ and $|x - \gamma c| \leq 2^M$.

Parametrize γ in terms of an $N \in \mathbf{N}$ and identify another point on ℓ_γ. Namely, fix $N \in \mathbf{N}$ write the rational α (see (99)) as $\alpha = p/q$ for $p, q \in \mathbf{P}$, and set

$$\text{(104)} \qquad \gamma := \gamma_0 \left(\gamma_1 + \frac{\varepsilon_0}{2} + \beta N q 2^M + N p 2^M \right).$$

The choice of γ_1 and (104) imply

$$Np2^M + \frac{\varepsilon_0}{2} = -\beta\left((Nq+1)2^M - 1\right) + \frac{\gamma}{\gamma_0}.$$

In particular,

$$\text{(105)} \qquad \begin{cases} \text{the line } \ell_\gamma \text{ passes through the point} \\ ((Nq+1)2^M - 1, Np2^M + \varepsilon_0/2). \end{cases}$$

Notice by (99) that

$$\frac{p}{q} = \alpha = \frac{h_\gamma(\gamma c)}{\gamma c}.$$

Hence the line $y = \alpha x$ passes through the points $(\gamma c, h_\gamma(\gamma c))$ and $(Nq2^M, Np2^M)$ for every $\gamma > 0$ and $N \in \mathbf{N}$. Define γ by (104) where N is chosen so large that $2^M < \gamma\delta$. Then (102), (103), and (105) all hold. It follows, therefore, that

$$Nq2^M \leq \gamma c \leq (Nq+1)2^M.$$

Consequently, (103) holds for all x in the interval

$$I_{M,N} := [Nq2^M, (Nq+1)2^M).$$

Suppose f is a double Walsh polynomial which satisfies $\widehat{f}(m,n) = 0$ for $m, n \geq 2^M$. Let

$$\mathcal{P}_{M,N} := \{g : g \quad \text{is a double Walsh polynomial with} \\ \widehat{g}(m,n) = 0 \quad \text{when } m,n \in \mathbf{N} \text{ and } m \notin I_{M,N} \text{ or } n < Np2^M\},$$

and set

$$\widetilde{w} := \widetilde{w}_{Nq2^M, Np2^M}.$$

Clearly,

$$\widehat{\widetilde{w}f}(m,n) = \widehat{f}(m \oplus Nq2^M, n \oplus Np2^M). \tag{106}$$

In the case $m \notin I_{M,N}$ we have $m \oplus Nq2^M \geq 2^M$. In the case $n < Np2^M$ we have $m \oplus Np2^M \geq 2^M$. Therefore $\widetilde{w}f$ belongs to $\mathcal{P}_{M,N}$.

By construction, part of the boundary of the dilation $T(\gamma/\gamma_0)$ is determined by the line ℓ_γ whereas part of the boundary of $R(\gamma)$ is determined by h_γ. Moreover, for $x \in I_{M,N}$ we have by (103) that $h_\gamma(x)$ and $\ell_\gamma(x)$ differ by at most $\varepsilon_0/2$. It follows, therefore, from (105) and the choice of ε_0 that $S_{T(\gamma/\gamma_0)}$ and $S_{R(\gamma)}$ are identical on $\mathcal{P}_{M,N}$. In particular,

$$S_{T(\gamma/\gamma_0)}(\widetilde{w}f) = S_{R(\gamma)}(\widetilde{w}f). \tag{107}$$

The choice of ε_0,and the identities (106), (104) imply

$$\begin{aligned}\widetilde{w}S_{T(\gamma_1)}f &= \widetilde{w}S_{T(\gamma_1-\varepsilon_0/2)}f \\ &= S_{T(\gamma_1-\varepsilon_0/2+\beta Nq2^M+Np2^M)}(\widetilde{w}f) \\ &= S_{T(\gamma/\gamma_0)}(\widetilde{w}f).\end{aligned}$$

We have therefore by (107) that

$$S_{R(\gamma)}(\widetilde{w}f) = \widetilde{w}S_{T(\gamma_1)}f$$

for any double Walsh polynomial f satisfying $\widehat{f}(m,n) = 0$ for $m,n \geq 2^M$. We conclude by the proof of Theorem 24 and Corollary 12 that as $\gamma \to \infty$, $S_{R(\gamma)}f$ must diverge on a set of positive measure for some $f \in L^p(I^2)$ for each $1 \leq p < 2$. ∎

Exercises

6.1 Prove there exist positive absolute constants c_1 and c_2 such that

$$c_1\|f\|_H \leq \|\sigma^* f\|_1 \leq c_2\|f\|_H$$

holds for all $f \in H$ with $f \geq 0$.

6.2 Let $f \in L^1$ and

$$\sigma^{**}f := \sup_{m \in \mathbf{P}} \frac{1}{m}\sum_{k=1}^{m}|S_k f|$$

i.e. $\sigma^{**}f$ is the maximal function associated with the strong means of the Walsh-Fourier series of f. Show that $\sigma^{**} : H \to L^1$ is not bounded.

(Hint: Use dyadic atoms and the fact that the arithmetic means of Lebesgue constants of the Walsh system are not bounded.)

6.3 Denote by $T(x)$ the one parameter group of martingale transform operators defined by

$$T(x)f = \sum_{n=0}^{\infty} \rho_n(x)\Delta_n f \qquad (x \in \mathbf{G}, f \in H).$$

Prove that

$$T(x) \circ T(y) = T(x+y) \qquad (x, y \in \mathbf{G})$$

and

$$T(0) = I.$$

Show there is a norm $\|\cdot\|$ on H equivalent to the original one such that

$$\|T(x)f\| = \|f\|$$

for all $x \in \mathbf{G}$ and $f \in \mathrm{H}$.

6.4 Let $T(x)$ be the operator defined in Exercise 6.3. Replace τ_{e_j} by $T(e_j)$ in the definition of dyadic differentiation and prove a "fundamental theorem" for this calculus.

6.5 Let

$$(Mf)(t) = \sup_I \frac{1}{|I|} \int_I |f(s)|\, ds$$

for $t \in [0,1), f \in \mathrm{L}^1$, where the supremum is taken over all subintervals I of $[0,1)$ containing t. A system of functions $(H_n, n \in \mathbf{N})$ is said to have the *H^*-property* if there is an absolute constant $C > 0$ such that

$$|\int_0^1 K_N(x,t) f(t)\, dt| \leq (Mf)(x)$$

for all $x \in [0,1)$, $N \in \mathbf{N}$, and all $f \in \mathrm{L}^1$, where

$$K_N(x,t) := \sum_{k=2^N}^{2^{N+1}-1} H_k(x) h_k(t)$$

for $x, t \in [0,1)$. Prove that the Haar system and the Franklin system both have the H^*-property.

6.6 Suppose $(H_n, n \in \mathbf{N})$ is an orthonormal system which has the H^*-property (see Exercise 6.5). Let $(W_n, n \in \mathbf{N})$ be its Hadamard transform and define an operator

$$W^* f = \sup_{n \in \mathbf{N}} \max_{0 \leq m < 2^n} |\sum_{k=2^n}^{2^n+m} \widehat{f}(k) W_k|$$

where $\widehat{f}(k)$ is the k-th Walsh-Fourier coefficient of $f \in \mathrm{L}^1$. Prove that W^* is of type (p,p) for every $1 < p < \infty$.

6.7 Use Exercise 6.6 for the Franklin and Ciesielski systems and the fact that Franklin-Fourier series of L^1 functions converge a.e. to prove that the Ciesielski-Fourier series of an $f \in \mathrm{L}^p, p > 1$, converges a.e. to f.

[Ciesielski, Simon, Sjölin [1]]

6.8 Let $1 \le p \le \infty$ and $\boldsymbol{g} = (g_n, n \in \mathbb{N})$ be the product system of $\boldsymbol{\gamma} = (\gamma_n, n \in \mathbb{N})$ defined on a probability space (Ω, ν). The system $\boldsymbol{\gamma}$ is called *p-weakly multiplicative* if

$$\sup_{n \in \mathbb{N}} \| \sum_{m=0}^{2^n-1} \left(\int_\Omega g_m \, d\nu \right) w_m \|_p < \infty.$$

Prove that if $\boldsymbol{\gamma}$ is weakly multiplicative then $\boldsymbol{\gamma}$ is p-weakly multiplicative for every p.

6.9 Let $\boldsymbol{\gamma}$ be a 2-weakly multiplicative system (see Exercise 6.8) of real-valued functions with $|\gamma_n| \le 1$ for $n \in \mathbb{N}$ and let $\boldsymbol{g}$ be its product system. Prove there is an absolute constant C such that

$$\| \sup_{m \in \mathbb{N}} | \sum_{k=0}^{m} a_k g_k | \|_{L^1(\Omega,\nu)} \le C \| \sup_{m \in \mathbb{N}} | \sum_{k=0}^{m} a_k w_k | \|_2$$

for every $(a_k, k \in \mathbb{N}) \in \ell^0$. Show $\boldsymbol{g}$ is a convergence system.

(Hint: Compare with Theorem 12 in 6.4.)

6.10 Prove that every uniformly bounded real orthonormal system has a weakly multiplicative subsystem. Show this result is false if "weakly" is removed from the conclusion.

6.11 Prove that every uniformly bounded orthonormal system has a subsystem which is a convergence system.

6.12 Let φ be a bounded function defined on $[0, \infty)$ of period 1 such that

$$\varphi(x + \frac{1}{2}) = -\varphi(x) \qquad (0 \le x \le \frac{1}{2}).$$

Set

$$\varphi_n(x) := \varphi(2^n x)$$

for $n \in \mathbb{N}$ and $x \in [0, \infty)$ and show that $(\varphi_n, n \in \mathbb{N})$ is a convergence system.

6.13 Show that the block space $\mathcal{B}_q$ is complete for every $q > 1$.

6.14 Suppose $q > 1$. Prove:

i) If $f \in \mathcal{B}_q$, and g is measurable with $|g| \le |f|$, then $\mathcal{N}_q(g) \le \mathcal{N}_q(f)$.

ii) If $f \in \mathcal{B}_q$ then $|f| \in \mathcal{B}_q$.

iii) If $f, g \in \mathcal{B}_q$ then $\sup(f, g), \inf(f, g) \in \mathcal{B}_q$.

6.15 Let $q > 1$. A *q-block* is a function β supported on an interval I (not necessarily dyadic) such that

$$\|\beta\|_q \le |I|^{\frac{1}{q}-1}.$$

Show any q-block is a sum of at most three dyadic q-blocks. Show that the dyadic block space $\mathcal{B}_q$ and the classical block space $\widetilde{\mathcal{B}}_q$ generated by q-blocks coincide.

6.16 Let $f \in \mathrm{L}^2$.

i) Prove

$$\max_{2^n \le m \le 2^{n+1}} |S_m f| \le \mathcal{E}^* f + n\mathcal{E}^* |S_{2^{n+1}} f - S_{2^n} f|$$

for all $n \in \mathbb{N}$.

ii) Prove there is an absolute constant $C > 0$ such that

$$\|S^*f\|_2 \leq C\left(|\widehat{f}(0)|^2 + \sum_{n=0}^{\infty} n \sum_{m=2^n}^{2^{n+1}-1} |\widehat{f}(m)|^2\right)^{1/2}$$

for every $f \in \mathrm{L}^2$.

[This is the Kolmogorov-Seliverstov-Plessner theorem for the Walsh system. For general orthonormal systems see Alexits [2], p. 175.]

6.17 Using Exercise 6.16 and Theorem 2 in 5.1 prove

$$\|S^*f\|_2 \leq C(\sum_{n=0}^{\infty} |\omega^{(2)}(f, 2^{-n})|^2)^{1/2}.$$

6.18 Using Exercise 6.17, prove that if $f \in \mathrm{C}_W$ and

$$\omega(f,\delta) = O\left(\log\frac{1}{\delta}\right)^{-\alpha} \quad \text{as} \quad \delta \to 0$$

for some $\alpha > \frac{1}{2}$ then Sf converges a.e.

[L. Pál]

Chapter 7

UNIQUENESS

7.1 Walsh Series and Quasi-measures. In previous chapters we have been primarily concerned with Walsh-Fourier series. In the next two chapters we shall prove a number of theorems about general Walsh series.

By a quasi-measure on the dyadic group we mean a finitely additive real-valued set function defined on the dyadic intervals of $\mathbf{G}$. The collection of quasi-measures on $\mathbf{G}$ will be denoted by $\mathrm{QM}(\mathbf{G})$. The collection of finite Borel measures on $\mathbf{G}$ will be denoted by $\mathbb{M}(\mathbf{G})$.

A quasi-measure ν on $\mathbf{G}$ is said to belong to $\mathbb{M}(\mathbf{G})$ if there is a measure $\lambda \in \mathbb{M}(\mathbf{G})$ such that the restriction of λ to dyadic intervals coincides with ν. There is a fundamental difference between quasi-measures on $[0,1)$ and quasi-measures on $\mathbf{G}$. For the interval $[0,1)$ there exists a non-negative quasi-measure which cannot be extended to a Borel measure (see Exercise 7.2). On the other hand, if ν is a non- negative quasi-measure on $\mathbf{G}$ then, since each dyadic interval in $\mathbf{G}$ is both open and closed, ν must be countably additive on the collection of dyadic intervals. Hence ν can be extended to a Borel measure on $\mathbf{G}$. In particular, if $\nu \in \mathrm{QM}(\mathbf{G})$ is non-negative then ν belongs to $\mathbb{M}(\mathbf{G})$.

The study of general Walsh series is equivalent to the study of Walsh-Fourier-Stieltjes series of quasi-measures. Indeed, for each $\nu \in \mathrm{QM}(\mathbf{G})$ define the Walsh-Fourier-Stieltjes coefficients of ν by

$$\widehat{\nu}(k) := \int_{\mathbf{G}} \psi_k \, d\nu \qquad (k \in \mathbf{N}) \tag{1}$$

and the Walsh-Fourier-Stieltjes series of ν by

$$S\nu := \sum_{k=0}^{\infty} \widehat{\nu}(k)\psi_k.$$

Denote the n-th partial sums of a Walsh series

$$S = \sum_{k=0}^{\infty} a_k \psi_k$$

by

$$S_n := \sum_{k=0}^{n-1} a_k \psi_k \qquad (n \in \mathbf{P}).$$

Repeating the arguments in 1.5 (see especially (52)) it is easy to see that the map $\nu \to S\nu$ is a 1-1 linear map from QM(G) onto the collection of all Walsh series. Moreover, if $S := S\nu$ then

$$2^{-n}S_{2^n}(x) = \nu(I_n(x)) \tag{2}$$

for $n \in \mathbf{N}$ and $x \in \mathbf{G}$.

A fundamental problem in the theory of general Walsh series is the problem of uniqueness. That is, when is a given Walsh series S the Walsh-Fourier series of some integrable function f? Since the Walsh system is orthogonal, an obvious solution to this problem is that uniqueness holds if S converges (for example) uniformly to f.

One might guess that if S converges a.e. to f then S is the Walsh-Fourier series of f. We shall see that this is not the case. In fact, the problem of uniqueness is a delicate one requiring distinctions among sets of measure zero not normally made in Lebesgue analysis.

Equation (2) suggests the problem of uniqueness is related to differentiation of quasi-measures, for we have

$$S_{2^n}(x) = \frac{\nu(I_n(x))}{\mu(I_n(x))} \tag{3}$$

for $n \in \mathbf{N}$ and $x \in \mathbf{G}$. In fact, this identity and a known result from the theory of binary derivates (see Theorem 13 in 0.6) lead to the following uniqueness theorem.

THEOREM 1. *If S is a Walsh series which satisfies*

$$\lim_{n\to\infty} 2^{-n}S_{2^n}(x) = 0 \tag{4}$$

for every $x \in \mathbf{G}$ and if

$$\lim_{n\to\infty} S_{2^n}(x) = 0 \tag{5}$$

for all but countably many $x \in \mathbf{G}$, then S is the zero series, i.e., the coefficients of S are identically zero.

Unless curiosity compels, there is no need to read Appendix 0.6 at this time. We have used the connection with binary derivates as a pedagogical tool. In particular, if ν is the quasi-measure associated with S then (4) can be viewed as a continuity condition on ν and (5) as an assumption that the binary derivate of ν is zero.

We shall say that a Walsh series satisfies the *$C - S$ condition* if (4) holds for all $x \in \mathbf{G}$. Notice that (4) cannot be relaxed at even one point in $\mathbf{G}$. Indeed, the Dirichlet series

$$S := \sum_{k=0}^{\infty} \psi_k$$

satisfies $S_{2^n}(x) \to 0$, as $k \to \infty$, for every $x \in \mathbf{G} \setminus \{0\}$ but

$$2^{-n}S_{2^n}(0) = 1 \qquad (n \in \mathbf{N}).$$

Notice that the $C - S$ condition is satisfied by any Walsh series whose full sequence of partial sums converges at some point of $\mathbf{G}$, in fact, by any Walsh series whose coefficients tend to zero. Indeed, if

$$S := \sum_{k=0}^{\infty} a_k \psi_k$$

then

$$2^{-n}|S_{2^n}| \leq 2^{-n} \sum_{k=0}^{2^n-1} |a_k|. \tag{6}$$

Hence if $a_k \to 0$ as $k \to \infty$ then $2^{-n} S_{2^n} \to 0$, as $n \to \infty$, uniformly on $\mathbf{G}$. In particular, the Walsh-Fourier series of an $f \in \mathrm{L}^1(\mathbf{G})$ satisfies the $C - S$ condition.

We shall prove that uniqueness holds for any Walsh series satisfying the $C - S$ condition under suitable convergence hypotheses. Our program is to show that the quasi-measure associated with such a series belongs to $\mathbb{M}(\mathbf{G})$ hence can be decomposed into an absolutely continuous part plus a singular part. By using (3) we shall prove the singular part is identically zero. It will follow that $S = S\nu$ for some ν satisfying

$$d\nu = f \, d\mu$$

and some $f \in \mathrm{L}^1(\mathbf{G})$. In particular, we have by definition

$$\widehat{\nu}(k) = \int_{\mathbf{G}} \psi_k \, d\nu = \int_{\mathbf{G}} f \psi_k \, d\mu = \widehat{f}(k)$$

for $k \in \mathbf{N}$, i.e., S is the Walsh-Fourier series of f.

We begin by identifying a condition sufficient to conclude a given quasi-measure belongs to $\mathbb{M}(\mathbf{G})$.

LEMMA 1. *If* $\nu \in \mathrm{QM}(\mathbf{G})$ *and*

$$\limsup_{n \to \infty} (S_{2^n} \nu)(x) \geq 0 \tag{7}$$

for all $x \in \mathbf{G}$ *then* ν *is non-negative and belongs to* $\mathbb{M}(\mathbf{G})$.

PROOF. By the remarks opening this section, it suffices to show $\nu \geq 0$. By considering $\nu + \alpha\mu$ for $\alpha > 0$ we may suppose that the inequality in (7) is strict. Thus for each $n \in \mathbf{N}$, $x \in \mathbf{G}$, and $y \in I_n(x)$ choose by (2) an integer $k(y) \geq n$ such that

$$\nu(I_{k(y)}(y)) > 0. \tag{8}$$

The compact set $I_n(x)$ is covered by the sets

$$\{I_{k(y)}(y) : y \in I_n(x)\}.$$

Thus we can choose finitely many points $y_1, y_2, \ldots, y_N$ in $I_n(x)$ such that

$$I_n(x) = \bigcup_{j=1}^{N} I_{k(y_j)}(y_j). \tag{9}$$

Since any two dyadic intervals are either disjoint or subsets of one another, we may suppose this union is disjoint. It follows from (9) that

$$\nu(I_n(x)) = \sum_{j=1}^{N} \nu(I_{k(y_j)}(y_j)) \geq 0.$$

We conclude by (8) that $\nu \geq 0$. ∎

If in Lemma 1 the series $S\nu$ satisfies the $C - S$ condition then hypothesis (7) can be relaxed at countably many points. In fact,

LEMMA 2. *Let* $\nu \in \mathrm{QM}(\mathbf{G})$, $S := S\nu$, *and* E *be a countable subset of* $\mathbf{G}$. *If*

$$\lim_{n\to\infty} 2^{-n} S_{2^n}(x) = 0 \qquad (x \in E) \tag{10}$$

and

$$\limsup_{n\to\infty} S_{2^n}(x) \geq 0 \qquad (x \notin E) \tag{11}$$

then ν *is non-negative and belongs to* $\mathbb{M}(\mathbf{G})$.

PROOF. Since E is countable there is a finite measure λ concentrated on E such that

$$\lambda(\{x\}) > 0 \qquad (x \in E). \tag{12}$$

Let $\alpha > 0$ and notice by (3) that

$$\limsup_{n\to\infty} (S_{2^n}(\alpha\lambda))(x) = \limsup_{n\to\infty} \alpha 2^n \lambda(I_n(x)) \geq 0$$

for all $x \in \mathbf{G}$. Hence we have by hypothesis (11) that

$$\limsup_{n\to\infty} 2^{-n}(S_{2^n}(\nu + \alpha\lambda))(x) \geq 0 \tag{13}$$

for $x \notin E$.

On the other hand, fix $x \in E$ and notice by (10) and (2) that

$$\begin{aligned} \lim_{n\to\infty} 2^{-n}(S_{2^n}(\nu + \alpha\lambda))(x) &= \lim_{n\to\infty} 2^{-n}(S_{2^n}\nu)(x) + \alpha \lim_{n\to\infty} 2^{-n}(S_{2^n}\lambda)(x) \\ &= \alpha \lim_{n\to\infty} \lambda(I_n(x)). \end{aligned}$$

Since $\{x\} = \cap_{n=0}^{\infty} I_n(x)$ implies $\lambda(\{x\}) = \lim_{n\to\infty} \lambda(I_n(x))$ we have by (12) that

$$\lim_{n\to\infty} 2^{-n}(S_{2^n}(\nu + \alpha\lambda))(x) > 0.$$

Therefore, (13) holds for all $x \in \mathbf{G}$.

Apply Lemma 1 to $\nu + \alpha\lambda$ and let $\alpha \to 0$. We conclude that ν is a non-negative, finite Borel measure on $\mathbf{G}$. ∎

In the classical situation, it is well known that the symmetric derivative of a singular measure λ is 0 a.e. with respect to Lebesgue measure on $[0, 1)$ and ∞ a.e. $[\lambda]$ on $[0, 1)$. This result is also true for binary derivates. We shall prove this and use it to obtain information about Walsh- Fourier-Stieltjes series of singular measures.

LEMMA 3. *Suppose $\lambda \in \mathbb{M}(\mathbf{G})$ is non- negative and singular with respect to Haar measure μ. Then*

$$\lim_{n\to\infty} S_{2^n}\lambda = 0 \quad \text{a.e.}\,[\mu] \tag{14}$$

and

$$\lim_{n\to\infty} S_{2^n}\lambda = \infty \quad \text{a.e.}\,[\lambda]. \tag{15}$$

PROOF. The proof of Theorem 6 in 6.2 shows that the binary derivate $\mathrm{D}\lambda$ exists and equals 0 a.e. $[\mu]$. By (3) this is equivalent to (14).

To prove (15), use the fact that λ is singular to choose a Borel set E in $\mathbf{G}$ such that $\lambda(\mathbf{G} \setminus E) = 0$ and $\mu(E) = 0$. We must show

$$\mathrm{D}\lambda(x) := \lim_{n\to\infty} \frac{\lambda(I_n(x))}{\mu(I_n(x))} = \infty$$

for a.e. $[\lambda]$ $x \in E$.

Fix $j \in \mathbf{P}$ and set

$$E_j := \{x \in E : \liminf_{n\to\infty} \frac{\lambda(I_n(x))}{\mu(I_n(x))} < j\}.$$

Let $\varepsilon > 0$ and choose an open set U in $\mathbf{G}$ such that $U \supset E_j$ and $\mu(U) < \varepsilon$. Let K be any compact subset of E_j and for each $x \in K$ choose a dyadic interval $J_x \subset U$ containing x such that

$$\lambda(J_x) < j\mu(J_x).$$

Since $\{J_x : x \in K\}$ covers K and consists of dyadic intervals we can choose a finite sequence $J_1, J_2, \dots, J_N$ of pairwise disjoint dyadic intervals such that

$$\lambda(J_\ell) < j\mu(J_\ell) \qquad (\ell = 1, 2, \dots, N)$$

and

$$K \subset \bigcup_{\ell=1}^{N} J_\ell \subset U.$$

It follows that

$$\begin{aligned}
\lambda(K) &\le \sum_{\ell=1}^{N} \lambda(J_\ell) \\
&< \sum_{\ell=1}^{N} j\mu(J_\ell) \\
&= j\mu\left(\bigcup_{\ell=1}^{N} J_\ell\right) \\
&\le j\mu(U) \\
&< j\varepsilon.
\end{aligned}$$

Thus $\lambda(K) = 0$.

We have shown that every compact subset of E_j is of λ- measure zero for all $j \in \mathbf{P}$. It follows that $\lambda(E_j) = 0$, i.e.,

$$\liminf_{n\to\infty} \frac{\lambda(I_n(x))}{\mu(I_n(x))} \geq j$$

for all $j \in \mathbf{P}$ and a.e. $[\lambda]$ $x \in E$. We conclude that $\lambda(I_n(x))/\mu(I_n(x)) \to \infty$ as $n \to \infty$ for a.e. $[\lambda]$ $x \in E$. ∎

These observations lead to a general solution to the uniqueness problem.

THEOREM 2. *Let S be a Walsh series, E be a countable subset of* $\mathbf{G}$, *and* $g \in \mathrm{L}^1(\mathbf{G})$. *If*

$$\lim_{n\to\infty} 2^{-n} S_{2^n}(x) = 0 \qquad (x \in E), \tag{16}$$

$$\limsup_{n\to\infty} |S_{2^n}(x)| < \infty \qquad (x \notin E) \tag{17}$$

and

$$\limsup_{n\to\infty} S_{2^n}(x) \geq g(x) \qquad (x \notin E) \tag{18}$$

then S is the Walsh-Fourier series of some $f \in \mathrm{L}^1(\mathbf{G})$.

PROOF. For each integer $k < 0$ choose V_k open in $\mathbf{G}$ such that

$$E_k := \{k \leq g < k+1\} \subset V_k$$

and

$$\mu(V_k \setminus E_k) \leq 2^k.$$

Set

$$\varphi := \sum_{k<0} k\chi(V_k).$$

Observe that $\varphi \leq g$ on $\mathbf{G}$, $\varphi \in \mathrm{L}^1(\mathbf{G})$, and if $\varphi(x) < \alpha$ for some $x \in \mathbf{G}$ and $\alpha \in \mathbf{R}$ then there is an open set V in $\mathbf{G}$ such that $x \in V$ and $\varphi(y) < \alpha$ for all $y \in V$. Thus φ is an upper semicontinuous function which satisfies

$$\varphi(x) \leq g(x)$$

for $x \in \mathbf{G}$.

Fix $x \in \mathbf{G}$, let $\alpha > \varphi(x)$ and choose $n_0 \in \mathbf{N}$ such that $y \in I_{n_0}(x)$ implies $\alpha > \varphi(y)$. Hence $n \geq n_0$ implies

$$(S_{2^n}\varphi)(x) = 2^n \int_{I_n(x)} \varphi \, d\mu < \alpha.$$

It follows that

$$\limsup_{n\to\infty} (S_{2^n}\varphi)(x) \leq \varphi(x). \tag{19}$$

Choose $\nu \in \mathrm{QM}(\mathbf{G})$ such that $S = S\nu$. Recall that any Walsh-Fourier series satisfies the $C - S$ condition. Thus if $x \in E$ we have by (16) that

$$\lim_{n\to\infty} 2^{-n}(S_{2^n}(\nu - \varphi))(x) = 0.$$

On the other hand, if $x \notin E$ we have by (18), (19), and the choice of φ that

$$\limsup_{n\to\infty}(S_{2^n}(\nu - \varphi))(x) \geq g(x) - \varphi(x) \geq 0.$$

Therefore, it follows from Lemma 2 that $\nu \in \mathbb{M}(\mathbf{G})$.

By the Lebesgue decomposition and the Radon-Nikodym theorem choose an $f \in \mathrm{L}^1(\mathbf{G})$ and a measure $\lambda \in \mathbb{M}(\mathbf{G})$ singular to μ such that

$$d\nu = f d\mu + d\lambda.$$

Thus

$$S = S\nu = Sf + S\lambda$$

and it suffices to show λ is the zero measure.

By considering the positive and negative parts of λ separately, we may suppose that $\lambda \geq 0$. Hence by Lemma 3 and (17), λ is concentrated on E. However, for each $x \in E$ we have by (2), the previous identity, and (16) that

$$\begin{aligned}\lambda(\{x\}) &= \lim_{n\to\infty} \lambda(I_n(x)) \\ &= \lim_{n\to\infty} 2^{-n}(S_{2^n}\lambda)(x) \\ &= 0.\end{aligned}$$

We conclude that $\lambda = 0$ as required. ∎

COROLLARY 1. *If S is a Walsh series which satisfies the $C-S$ condition, f is a finite-valued integrable function defined on all of* $\mathbf{G}$, *and*

$$\lim_{n\to\infty} S_{2^n}(x) = f(x)$$

for all but countably many $x \in \mathbf{G}$ then S is the Walsh-Fourier series of f.

COROLLARY 2. *If S is a Walsh series whose full partial sums converge, except perhaps at countably many points in* $\mathbf{G}$, *to a finite-valued, integrable function f defined on all of* $\mathbf{G}$, *then S is the Walsh-Fourier series of f.*

7.2 Uniqueness of Almost Everywhere Convergent Walsh Series. We have seen that uniqueness holds for Walsh series which converge to finite-valued, integrable functions, even if convergence is relaxed on a countable subset of $\mathbf{G}$. In this section we shall show that under suitable finiteness conditions, convergence can be relaxed on any set of measure zero.

It is important to recognize that a non-Fourier Walsh series which converges a.e. to an integrable function cannot have uniformly bounded 2^n-th partial sums.

LEMMA 4. *Let $f \in L^1(\mathbf{G})$, $J := I_{m_0}(y)$ for some $m_0 \in \mathbf{N}$, $y \in \mathbf{G}$, and M be any positive number. Let S be a Walsh series, set*

$$T_N := S_N - S_N f \qquad (N \in \mathbf{P}), \tag{20}$$

and suppose $T_{2^{m_0}} \neq 0$ on J. If $S_{2^n} \to f$ a.e. $[\mu]$, as $n \to \infty$, then there exist $n \in \mathbf{N}$ and $t \in \mathbf{G}$ such that $I := I_n(t) \subset J$, $T_{2^n} \neq 0$ on I and

$$|S_{2^n}(x)| > M \qquad (x \in I).$$

PROOF. This is immediate if there is at least one point x in some $I_n(t) \subset J$ such that

$$|S_{2^n}(x)| > M + |S_{2^n} f(x)|$$

holds, because both S_{2^n} and $S_{2^n} f$ are constant on $I_n(t)$ and $|T_{2^n}| \geq |S_{2^n}| - |S_{2^n} f|$.

Suppose to the contrary that

$$|S_{2^n}(x)| \leq M + |S_{2^n} f(x)|$$

for all $x \in J$ and $n \geq m_0$. Since $S_{2^n} f \to f$ in $L^1(\mathbf{G})$ norm it follows that the sequence T_{2^n} is bounded in the $L^1(\mathbf{G})$ norm. Consequently, the Banach-Alaoglu theorem can be used to choose a subsequence of integers $n_1, n_2, \ldots$ and a finite Borel measure ν on $\mathbf{G}$ such that

$$\lim_{j\to\infty} \int_J \varphi T_{2^{n_j}}\, d\mu = \int_J \varphi\, d\nu$$

for all $\varphi \in C(J)$. Specializing to $\varphi = 1$ and recalling from hypothesis that $T_{2^n} \to 0$ a.e. as $n \to \infty$, we find that

$$0 = \lim_{j\to\infty} \int_J T_{2^{n_j}}\, d\mu.$$

However, since $J = I_{m_0}(y)$ and $T_{2^{m_0}}$ is constant on J, we have that

$$\begin{aligned}\int_J T_{2^{n_j}}\, d\mu &= \int_J T_{2^{m_0}}\, d\mu \\ &= 2^{-m_0} T_{2^{m_0}}(y) \neq 0\end{aligned}$$

for $n_j \geq m_0$. This contradiction proves the lemma. ∎

LEMMA 5. *Let $x_1 \in \mathbf{G}$, $n_0 \in \mathbf{N}$, $I_0 := I_{n_0}(x_1)$, and T be Walsh series which satisfies the $C - S$ condition. If $T_{2^{n_0}}$ is non-zero on I_0, then there exists a set $J \subseteq I_0$ of the form $J = I_{m_0}(y)$ for some $y \in \mathbf{G}$, $m_0 \in \mathbf{N}$, such that $x_1 \notin J$ and $T_{2^{m_0}}$ is non-zero on J.*

PROOF. For each $k \in \mathbf{P}$, set $I_k := I_{n_0+k}(x_1)$ and $J_k := I_{k-1} \setminus I_k$. Observe for each $2^{n_0+k-1} \leq \ell < 2^{n_0+k}$ that ψ_ℓ is constant on I_k and on J_k, but changes signs from I_k to J_k. Thus

$$\psi_\ell(x) = -\psi_\ell(y) \tag{21}$$

for $x \in I_k$, $y \in J_k$, $2^{n_0+k-1} \leq \ell < 2^{n_0+k}$, and $k \in \mathbf{P}$.

Suppose the lemma is false. Since $x_1 \notin J_k$ for $k \in \mathbf{P}$, it follows that

$$T_{2^{n_0+k}} = 0 \tag{22}$$

on J_k for $k \in \mathbf{P}$. Since $T_{2^{n_0}}$ is constant and non-zero on I_0 it is clear that

$$T_{2^{n_0+k}}(x) = 2^k\alpha \qquad (x \in I_k) \tag{23}$$

for $k = 0$ and $\alpha := T_{2^{n_0}}(x_1)$. Suppose (23) holds for $k-1$ and some $k \in \mathbf{P}$. Then

$$T_{2^{n_0+k-1}} = 2^{k-1}\alpha \tag{24}$$

on I_k and J_k, and (22) implies

$$T_{2^{n_0+k}} - T_{2^{n_0+k-1}} = -2^{k-1}\alpha$$

on J_k. It follows from (21) that

$$T_{2^{n_0+k}} - T_{2^{n_0+k-1}} = 2^{k-1}\alpha$$

on I_k. Hence by (24),

$$T_{2^{n_0+k}} = 2^k\alpha$$

holds on I_k, i.e., (23) holds for all $k \in \mathbf{N}$. In particular,

$$2^{-m}T_{2^m}(x_1) = 2^{-n_0}\alpha \qquad (m \geq n_0)$$

and contrary to hypothesis T fails to satisfy the $C - S$ condition. ∎

It is now easy to prove a version of Corollary 1 in 7.1 valid for Walsh series which converge a.e.

THEOREM 3. *Suppose E is a countable subset of* $\mathbf{G}$ *and S is a Walsh series satisfying the $C - S$ condition such that*

$$\limsup_{n\to\infty} |S_{2^n}(x)| < \infty \tag{25}$$

for all $x \in \mathbf{G} \setminus E$. If

$$\lim_{n\to\infty} S_{2^n} = f$$

a.e. $[\mu]$ on $\mathbf{G}$ *for some function $f \in \mathrm{L}^1(\mathbf{G})$, then S is the Walsh-Fourier series of f.*

PROOF. Suppose the theorem is false. Then there exist $x \in \mathbf{G}$ and $n_0 \in \mathbf{N}$ such that

$$S_{2^{n_0}} \neq S_{2^{n_0}}(f)$$

on $I_{n_0}(x)$.

Set $E = \{x_1, x_2, \dots\}, I_0 := I_{n_0}(x)$, and $T := S - Sf$. Observe that T is a Walsh series satisfying the $C - S$ condition and that $T_{2^{n_0}}$ is non-zero on I_0.

It is easy to see there is a set $J \subseteq I_0$ of the form $I_{m_0}(y)$ such that $x_1 \notin J$ and $T_{2^{m_0}}$ is non-zero on J. Indeed, if $x_1 \in I_0$ use Lemma 5. If $x_1 \notin I_0$, set $J := I_0$ and $m_0 := n_0$.

Apply Lemma 4 to J, choosing a set $I := I_n(t)$ such that T_{2^n} is non-zero on I and

$$|S_{2^n}(x)| > 1 \qquad (x \in I). \tag{26}$$

Set $I_1 := I$, $n_1 := n$, and observe that $x_1 \notin I_1$, $I_1 \subset I_0$, $T_{2^{n_1}}$ is non-zero on I_1, and (26) holds.

Continued alternate applications of Lemma 5 and Lemma 4 generate nested compact sets $I_0 \supset I_1 \supset \dots$ such that $x_k \notin I_k$, and integers $n_1, n_2, \dots$, such that

$$|S_{2^{n_k}}(x)| > k \qquad (x \in I_k), \tag{27}$$

for $k = 1, 2, \dots$

Let

$$x_0 \in \bigcap_{k=1}^{\infty} I_k.$$

By construction, $x_0 \notin E$. Hence by hypothesis,

$$\limsup_{n\to\infty} |S_{2^n}(x_0)| < \infty.$$

On the other hand, (27) implies

$$\limsup_{n\to\infty} |S_{2^n}(x_0)| = \infty,$$

a contradiction. Therefore, S must be the Walsh-Fourier series of f. ∎

Hypothesis (25) can be relaxed slightly. Indeed, since the set of $x \in \mathbf{G}$ which satisfy (25) is a Borel set and since a Borel set is uncountable if and only if it contains no perfect subset, it is clear that Theorem 3 holds if "E is countable" is replaced by "E contains no perfect subsets". This hypothesis can be relaxed no further:

THEOREM 4. *Let E be an uncountable closed set in $\mathbf{G}$. There is a non-zero Walsh series S satisfying the $C - S$ condition such that*

$$\lim_{n\to\infty} S_{2^n}(x) = 0$$

for all $x \notin E$.

PROOF. We shall construct a measure ν on $\mathbf{G}$ such that $S := S\nu$ has the required properties.

If $\mu(E) > 0$ set

$$\nu(\omega) := \frac{\mu(\omega \cap E)}{\mu(E)}$$

for every Borel subset ω of $\mathbf{G}$. If $\mu(E) = 0$ consider $E_1 := \varrho^{-1}(E)$ where $\varrho : [0,1) \to \mathbf{G}$ is Fine's map. By hypothesis E_1 is an uncountable Borel set in $[0,1)$ hence contains a perfect set E_0, i.e., a closed set with no isolated points. Construct a non-negative, finite Borel measure ν_1 on E_0 using a Cantor-Lebesgue function, i.e., a function which increases on E_0 and has derivative zero off E_0. Set

$$\nu(\omega) := \nu_1(\varrho^{-1}(\omega))$$

for each Borel set ω in $\mathbf{G}$. In either case $\nu \in \mathbb{M}(\mathbf{G})$ is non-negative, non-zero, supported on E and satisfies

$$\nu(\{x\}) = 0 \qquad (x \in \mathbf{G}).$$

Let $S := S\nu$ and observe by (2) that

$$S_{2^n}(x) = 2^n \nu(I_n(x)) \tag{28}$$

for $n \in \mathbf{N}$ and $x \in \mathbf{G}$. Since $\nu(\{x\}) = 0$ for all $x \in \mathbf{G}$ it is clear that S satisfies the $C - S$ condition. Since

$$\widehat{\nu}(0) = \int_{\mathbf{G}} d\nu = \|\nu\|$$

it is also clear that S is not the zero series. Finally, if $x \notin E$ then the fact that E is closed implies

$$E \cap I_n(x) = \emptyset$$

for large n. It follows from (28) that $S_{2^n}(x) = 0$ for n sufficiently large. We conclude that $S_{2^n}(x) \to 0$ as $n \to \infty$ for every $x \notin E$. ∎

Theorem 3 can be generalized in another direction. Indeed, we verify without difficulty that the proof of Theorem 3 yields the following result.

THEOREM 5. *Suppose E is countable subset of $\mathbf{G}$, that S is a Walsh series satisfying the $C - S$ condition, and $n_1 < n_2 < \ldots$ is a sequence of positive integers such that*

$$\limsup_{j \to \infty} |S_{2^{n_j}}(x)| < \infty \qquad (x \notin E). \tag{29}$$

If $S_{2^{n_j}} \to f$ in measure (in particular, a.e. $[\mu]$) as $j \to \infty$ for some $f \in \mathrm{L}^1(\mathbf{G})$ then S is the Walsh- Fourier series of f.

7.3 Null Series and the Formal Product. In Theorem 4 we constructed non-zero Walsh series whose 2^n- th partial sums converge a.e. to zero. A *null series* is a non-zero Walsh series whose full sequence of partial sums converges a.e. to zero.

By Corollary 2 in 7.1, the set where a given null series does not converge to zero must be uncountable, i.e., thick in a set theoretic sense. Below we show that these sets can be thin in a measure theoretic sense. Our method for constructing these series involves the formal product of a Walsh series with a Walsh polynomial. This product is given in the following result.

LEMMA 6. *Suppose $S = \sum_{k=0}^{\infty} a_k w_k$ is a Walsh series, and*

$$P = \sum_{j=0}^{2^{n_0}-1} b_j w_j$$

is a Walsh polynomial. For each $k \in \mathbf{N}$ set

$$c_k := \sum_{j=0}^{2^{n_0}-1} b_j a_{k\oplus j}.$$

If $T := \sum_{k=0}^{\infty} c_k w_k$ then

$$T_{q2^{n_0}}(x) = P(x) S_{q2^{n_0}}(x) \tag{30}$$

for all $x \in [0,1)$, and $q \in \mathbf{P}$.

PROOF. Fix $x \in [0,1)$ and an integer $q \geq 1$. Since $w_{k\oplus j} = w_k w_j$ holds for all non-negative integers k and j, it is clear that

$$P(x) S_{q2^{n_0}}(x) = \sum_{j=0}^{2^{n_0}-1} \sum_{k=0}^{q2^{n_0}-1} b_j a_k w_{j\oplus k}(x). \tag{31}$$

Now, for fixed integers $0 \leq j < 2^{n_0}$ and $\ell \geq 0$, the dyadic sum $j \oplus k$ ranges over

$$\{\ell 2^{n_0}, \ell 2^{n_0} + 1, \ldots, (\ell+1)2^{n_0} - 1\}$$

as k does. Hence the identity $j \oplus k \oplus j = k$ allows (31) to be rewritten as

$$P(x) S_{q2^{n_0}}(x) = \sum_{j=0}^{2^{n_0}-1} \sum_{k=0}^{q2^{n_0}-1} b_j a_{j\oplus k} w_k(x).$$

The proof of the lemma is completed by interchanging the order of summation in this expression. ∎

LEMMA 7. *If S is a Walsh series whose coefficients satisfy $a_k \to 0$ as $k \to \infty$, and if $I := I(p, n_0)$ for some $0 \leq p < 2^{n_0}$, then there is a Walsh series T whose coefficients satisfy $c_k \to 0$ as $k \to \infty$, such that $\lim_{n\to\infty} T_n = 0$ uniformly on $[0,1) \setminus I$,*

$$T_{2^n}(x) = 0$$

for $n \geq n_0$ and $x \notin I$, and such that for each $x \in I, S(x)$ converges if and only if $T(x)$ converges, in which case $S(x) = T(x)$.

PROOF. Let $P := \chi(I)$. Then P is a Walsh polynomial of order $2^{n_0} - 1$. Apply Lemma 6 to choose a Walsh series T which satisfies

$$T_{q2^{n_0}} = P\, S_{q2^{n_0}} \qquad (q \in \mathbf{P}). \tag{32}$$

The coefficients of T are given by

$$c_k := \sum_{j=0}^{2^{n_0}-1} b_j a_{k\oplus j}.$$

Since $k \oplus j \to \infty$ as $k \to \infty$, it is clear by hypothesis that $c_k \to 0$ as $k \to \infty$. Thus given $\varepsilon > 0$, there is an integer q_0 so large that

$$\alpha_0 := \sup\{|a_k - c_k| : k \geq q_0 2^{n_0}\} < \varepsilon 2^{-n_0}.$$

If $x \notin I$ then $P(x) = 0$. Thus (32) implies that $T_{q2^{n_0}}(x) = 0$ for $q \in \mathbf{P}$. Hence $T_{2^n}(x) = 0$ for $n \geq n_0$. Moreover, since

$$|T_m(x)| \leq 2^{n_0} \sup\{|c_k| : k \geq q2^{n_0}\}$$

for $q \in \mathbf{P}$ and $q2^{n_0} \leq m < (q+1)2^{n_0}$ we see that $\lim_{n\to\infty} T_n = 0$ uniformly on $[0,1) \setminus I$.

If $x \in I$ then $P(x) = 1$. Hence (32) implies that

$$T_{q2^{n_0}}(x) = S_{q2^{n_0}}(x) \qquad (q \geq 1). \tag{33}$$

If $n \geq q_0 2^{n_0}$ then $q2^{n_0} \leq n < (q+1)2^{n_0}$ for some $q \geq q_0$. If follows, therefore, from the choice of q_0 and (33) that

$$|T_n(x) - S_n(x)| \leq \alpha_0\, 2^{n_0} < \varepsilon.$$

In particular, $S(x)$ and $T(x)$ are equiconvergent, and $S(x) = T(x)$ if either converges. ∎

THEOREM 6. *Suppose $n_1 < n_2 \ldots$ is a sequence of natural numbers, I is a dyadic interval, and S is a Walsh series whose coefficients tend to zero. If $S_{2^{n_j}} \to 0$ a.e. on I, as $j \to \infty$, and if*

$$\limsup_{j\to\infty} |S_{2^{n_j}}(x)| < \infty \tag{34}$$

for all but countably many $x \in I$, then S_n converges to zero, as $n \to \infty$, everywhere on I.

PROOF. Apply Lemma 7 to S and I, choosing a Walsh series T whose coefficients tend to zero, such that $T_{2^n} = 0$ on $[0,1) \setminus I$, for n large, and such that (33) holds for $x \in I$. Thus T satisfies the $C - S$ condition, $T_{2^{n_j}} \to 0$ a.e., as $j \to \infty$, and

$$\limsup_{j\to\infty} |T_{2^{n_j}}(x)| < \infty$$

for all but countably many $x \in [0,1)$. Hence by Theorem 5, T is the zero series. Since S and T are equiconvergent on I, it follows that $S_n(x) \to 0$, as $n \to \infty$, for all $x \in I$. ∎

Thus to construct null series which converge to zero off some closed set E in $\mathbf{G}$ it suffices to find Walsh series S whose coefficients tend to zero such that $S_{2^{n_j}} \to 0$ a.e. off E, as $j \to \infty$, for some subsequence of integers $n_1 < n_2 < \ldots$

We promised to show such sets E can be quite thin in a measure theoretic sense. To make precise what is meant by "thin", let Φ be a non-negative, increasing function on $[0,\infty)$ which satisfies $\Phi(t) \downarrow 0$ as $t \downarrow 0$. A set $E \subseteq [0,1)$ is said to be of Hausdorff measure zero with respect to Φ, if given $\varepsilon > 0$ there exist intervals $v_1, v_2, \ldots$ such that E is covered by the v_s's, and

$$\sum_{s=1}^{\infty} \Phi(|v_s|) < \varepsilon.$$

In this case we will write $\mathbf{\Phi}(E) = 0$.

Notice since $\Phi(t) \downarrow 0$ as $t \downarrow 0$, that the Hausdorff measure of any countable set is zero. On the other hand, if we take $\Phi(t) := t$, $t \geq 0$, then $\mathbf{\Phi}(E) = 0$ means that E is of Lebesgue measure zero. And, if $\Phi(t) := t^{1/p}$, $t \geq 0$, for some $1 < p < \infty$, then the larger p is, the fewer sets E of Lebesgue measure zero satisfy $\mathbf{\Phi}(E) = 0$.

THEOREM 7. *If Φ is any non-negative, increasing function on $[0,\infty)$ with $\Phi(0)=(+0)\Phi=0$, then there is a perfect set $E \subset [0,1]$ with $\mathbf{\Phi}(E) = 0$ and a non- zero Walsh series S such that $S(x)$ converges to zero for all $x \in [0,1] \setminus E$.*

PROOF. We begin by building a prototype for the dyadic blocks of S.

Let $A > 0$ and $\varepsilon > 0$. We claim there exists an integer N such that if I is a dyadic interval of length 2^{-n} for some $n \geq N$, if $m > n$ is an integer, $\delta > 0$, and $0 < B \leq A$ then there exist $\ell \in \mathbf{P}$ and a Walsh polynomial

$$G = \sum_{i=2^m}^{2^\ell - 1} a_i w_i$$

such that

$$|a_i| \leq \varepsilon \qquad (2^m \leq i < 2^\ell), \tag{35}$$

$$G(x) = 0 \qquad (x \in [0,1] \setminus I), \tag{36}$$

$$G(x) \geq -B \qquad (x \in [0,1]), \tag{37}$$

and

$$\begin{cases} \text{if } s \text{ is the number of dyadic intervals of length } 2^{-\ell} \\ \text{inside } I \text{ on which } G(x) \neq -B \text{ then } \quad s\Phi(2^{-\ell}) < \delta. \end{cases} \tag{38}$$

To verify this claim, choose N so large that $A2^{-N} \leq \varepsilon$. Let $m > n \geq N$ and I be a dyadic interval of length 2^{-n}. Divide I into 2^{m-n} dyadic subintervals of length 2^{-m}, say $J_1, J_2, \ldots, J_{2^{m-n}}$. Since $\Phi(t) \to 0$ as $t \to 0$ we may choose $\ell > m$ so large that

$$2^{m-n}\Phi(2^{-\ell}) < \delta. \tag{39}$$

For each integer $j \in [1, 2^{m-n}]$ let I_j be the left-most dyadic interval of length $2^{-\ell}$ which lies in J_j. Set

$$G(x) := \begin{cases} 0 & x \in [0,1] \setminus I \\ -B & x \in J_j \setminus I_j \quad \text{for some} \quad j \\ (2^{\ell-m}-1)B & x \in I_j \quad \text{for some} \quad j \end{cases}$$

and let $a_0, a_1, \ldots$ represent the Walsh-Fourier coefficients of G. Since $a_i = 0$ for $i \geq 2^\ell$ or $i < 2^m$, G is a Walsh polynomial. To verify (35), fix $2^m \leq i < 2^\ell$ and choose $m \leq q < \ell$ such that $2^q \leq i < 2^{q+1}$. Observe that

$$a_i = \sum_{p=0}^{2^q-1} a(p,i)$$

where

$$a(p,i) := \int_{I(p,q)} G(x) w_i(x)\, dx \qquad (p = 0, 1, \ldots, 2^q - 1).$$

If $I(p,q) \cap I_j = \emptyset$ for all $1 \leq j \leq 2^{m-n}$ then either G is constant on $I(p,q)$ or G is zero on $I(p,q)$. In any event, $a(p,i) = 0$.

If $I_j \subset I(p,q)$ for some j then the identity

$$a(p,i) = \pm\left(\int_{I(2p,q+1)} G - \int_{I(2p+1,q+1)} G\right)$$

can be used to verify $a(p,i) = \pm B\, 2^{-m}$. Since for each fixed q there are at most 2^{m-n} intervals $I(p,q)$ which contain an interval I_j, it follows that $|a_i| \leq B\, 2^{-n}$. Since $B \leq A$ and $n \geq N$, we have by the choice of N that $|a_i| \leq \varepsilon$. This verifies (35).

Conditions (36) and (37) follow immediately from the definition of G.

Condition (38) follows from (39). Indeed, if $x \in J_j$ and $G(x) \neq -B$ then $x \in I_j$. Hence the number s identified in (38) is precisely 2^{m-n}.

The claim is established.

Let $\varepsilon_k \downarrow 0$ as $k \to \infty$ with $\varepsilon_1 := \Phi(1)$ and set $n_0 := -\infty$. Suppose n_{k-1} has been chosen for some $k \in \mathbf{P}$. We shall show that there exist an integer $n_k > n_{k-1}$, and a Walsh polynomial Q_k whose coefficients satisfy (35) for $\varepsilon := \varepsilon_k$, such that

$$Q_k = \sum_{i=2^{n_{k-1}}}^{2^{n_k}-1} a_i w_i\,, \tag{40}$$

$$\sum_{i=1}^{k} Q_i \geq 0 \quad \text{on} \quad [0,1], \tag{41}$$

and

$$(42)\qquad \begin{cases} \text{if for each } \ell \in \mathbf{P} \text{ we set } E_\ell := \{\sum_{i=1}^{\ell} Q_j > 0\} \\ \text{and if } E_\ell = \bigcup_s v_s^{(\ell)} \text{ for some dyadic intervals of length } 2^{-n_\ell} \\ \text{then } E_{k+1} \subset E_k \text{ and } \sum_s \Phi(|v_s^{(k)}|) \le \varepsilon_k. \end{cases}$$

Indeed, let $n_1 := 0$, $Q_1 := \varepsilon_1$, $E_1 := [0,1]$, and suppose we have chosen Walsh polynomials $Q_1, Q_2, \dots, Q_{k-1}$ and integers $n_1, n_2, \dots, n_{k-1}$ so that (40), (41), and (42) hold for $k-1$ in place of k. In particular

$$A_{k-1} := \sup\{\sum_{i=1}^{k-1} Q_i(x) : x \in [0,1]\}$$

is a non-negative number, and we can apply the claim to $A := A_{k-1}$ and $\varepsilon := \varepsilon_k$ to choose the integer N_k.

Let $n \ge N_k$ be so large that the polynomials $Q_1, \dots, Q_{k-1}$ are all constant on dyadic intervals of the form $I(p,n)$ for $0 \le p < 2^n$. Let $I_1, I_2, \dots, I_{j_0}$ represent the dyadic intervals of length 2^{-n} which lie inside E_{k-1}, and observe that $1 \le j_0 \le 2^n$. Finally, for $1 \le j \le j_0$ define

$$B_j := \sum_{i=1}^{k-1} Q_i \quad \text{on} \quad I_j,\ 1 \le j \le j_0,$$

and observe by the choice of n that the B_j's are constant, and by definition that

$$0 < B_j \le A_{k-1}$$

for $1 \le j \le j_0$.

For each $1 \le j \le j_0$ apply the claim to $A := A_{k-1}, \varepsilon := \varepsilon_k, B := B_j, I := I_j, \delta := \varepsilon_k / j_0$, and an integer $m > n$ so large that $m \ge n_{k-1}$. Thus choose an integer ℓ, and polynomials $G_1, \dots, G_{j_0}$ whose coefficients satisfy (35) for $\varepsilon := \varepsilon_k$ such that (36), (37) and (38) hold for $G := G_j, I := I_j, B := B_j$, and $1 \le j \le j_0$.

Set $n_k := \ell$ and

$$Q_k := \sum_{j=1}^{j_0} G_j.$$

Clearly (36) and the definition of B_j imply that

$$\sum_{i=1}^{k} Q_i(x) = B_j + G_j(x) \qquad (x \in I_j,\ 1 \le j \le j_0).$$

Hence (41) follows from (37) and (42) follows from (38). Moreover, if we set $a_i := 0$ for $2^{n_{k-1}} \le i < 2^m$, then (40) also holds.

Let $S := \sum_{k=1}^{\infty} Q_k$, i.e.,

$$S_{2^{n_k}} = \sum_{i=1}^{k} Q_i$$

for $k = 1, 2, \ldots$, and let

$$E := \bigcap_{k=1}^{\infty} E_k.$$

By (42), $S_{2^{n_k}}(x) \to 0$, as $k \to \infty$, for all $x \in [0,1] \setminus E$, and $\Phi(E) = 0$. By (35), the coefficients a_i of S tend to zero as $i \to \infty$. Since the complement of E is a union of dyadic intervals, we conclude from Theorem 6 that $S(x)$ converges to zero for every $x \in [0,1) \setminus E$. ∎

The proof of Theorem 7 indicates that a null series can have large gaps. These gaps cannot be arbitrarily distributed. For example, there are no lacunary null series (see Exercise 7.14).

7.4 Null Series and Measure Preserving Transformations. In this section we introduce another method for constructing null series.

For each $k \in \mathbf{N}$ with binary coefficients $(k_j, j \in \mathbf{N})$ define the number

$$\ell(k) := \sum_{j=0}^{\infty} k_j.$$

For each non-zero real number $-1 \leq \alpha \leq 1$ consider the generalized Dirichlet series

$$D^{(\alpha)} := \sum_{k=0}^{\infty} \alpha^{\ell(k)} \psi_k.$$

Clearly $D^{(1)}$ is the usual Walsh-Dirichlet series, so by Paley's lemma, $D_{2^n}^{(1)}(x) \to 0$ as $n \to \infty$ for every $x \in \mathbf{G} \setminus \{0\}$, in particular, for a.e. $[\mu]$ $x \in \mathbf{G}$.

LEMMA 8. *If* $-1 \leq \alpha \leq 1$, $\alpha \neq 0$, *then*

$$\lim_{n\to\infty} D_{2^n}^{(\alpha)} = 0 \tag{43}$$

a.e. $[\mu]$ *on* $\mathbf{G}$.

PROOF. Let ν be the quasi-measure on $\mathbf{G}$ associated with $D^{(\alpha)}$. Notice that the partial sums of $D^{(\alpha)}$ satisfy

$$D_{2^n}^{(\alpha)} = \prod_{j=0}^{n-1} (1 + \alpha \rho_j) \qquad (n \in \mathbf{N}).$$

Since $|\alpha| \leq 1$ it follows that $D_{2^n}^{(\alpha)} \geq 0$ on $\mathbf{G}$ for all $n \in \mathbf{N}$. Hence by (3) in 7.1, ν is a nonnegative quasi-measure on $\mathbf{G}$, so, $\nu \in \mathbb{M}(\mathbf{G})$.

Choose non-negative $f \in L^1(G)$ and $\lambda \in \mathbb{M}(G)$ such that $d\nu = f d\mu + d\lambda$ where λ is singular to μ. Thus for each $n \in \mathbf{N}$,

$$D_{2^n}^{(\alpha)} = S_{2^n} f + S_{2^n} \lambda.$$

Moreover, by Lemma 3 in 7.1, $S_{2^n}\lambda \to 0$ a.e. $[\mu]$ as $n \to \infty$. In particular, the limit in (43) exists and is finite a.e. $[\mu]$ on $\mathbf{G}$.

Suppose x is a point in $\mathbf{G}$ where the limit in (43) is not zero. Then

$$1 = \lim_{n\to\infty} \frac{D_{2^{n+1}}^{(\alpha)}(x)}{D_{2^n}^{(\alpha)}(x)} = \lim_{n\to\infty} (1 + \alpha\rho_n(x))$$

which implies $\alpha = 0$. Since this case has already been handled, we conclude that (43) holds a.e. $[\mu]$ on $\mathbf{G}$. ∎

Since the terms of $D^{(\alpha)}$ do not converge to zero, $D^{(\alpha)}$ cannot be a null series. Nevertheless, certain measure preserving transformations of $D^{(\alpha)}$ are null series.

A map $T : \mathbf{G} \to \mathbf{G}$ is called simple if $T^{-1}(I)$ is a finite union of dyadic intervals for each dyadic interval I in $\mathbf{G}$. Let T be a simple, measure preserving transformation of $\mathbf{G}$. Define a system $\boldsymbol{\gamma} = (\gamma_n, n \in \mathbf{N})$ by

$$\gamma_n := \rho_n \circ T \qquad (n \in \mathbf{N}).$$

Each γ_n is a Walsh polynomial because T is simple. Moreover, the system $\boldsymbol{\gamma}$ is a strongly multiplicative, sign-like system because T is measure preserving. Conversely, every strongly multiplicative, sign-like system of Walsh polynomials can be generated by a simple, measure preserving transformation (see Theorem 6 in 1.4). In fact, if $\boldsymbol{g} = (g_m, m \in \mathbf{N})$ is the product system generated by $\boldsymbol{\gamma}$ then there is a simple, measure preserving transformation T of $\mathbf{G}$ such that

$$g_m = \psi_m \circ T \qquad (m \in \mathbf{N}). \tag{44}$$

These observations allow us to identify conditions sufficient to conclude that a measure preserving transformation of $D^{(\alpha)}$ is a null series.

LEMMA 9. *Let $-1 \le \alpha \le 1$, $\alpha \ne 0$. Suppose $(\gamma_n, n \in \mathbf{N})$ is a strongly multiplicative, sign-like system of Walsh polynomials, and $(g_m, m \in \mathbf{N})$ is its product system. Suppose further that*

$$q_n := \max_{2^{n-1} \le m < 2^n} sp(g_m) < \min_{2^n \le m < 2^{n+1}} sp(g_m) \qquad (n \in \mathbf{P}). \tag{45}$$

If the Walsh series

$$S^{(\alpha)} := 1 + \sum_{n=0}^{\infty} \Big(\sum_{m=2^n}^{2^{n+1}-1} \alpha^{\ell(m)} g_m \Big)$$

has partial sums which satisfy

$$\lim_{n\to\infty} S^*\left(S^{(\alpha)}_{q_{n+1}} - S^{(\alpha)}_{q_n}\right) = 0 \tag{46}$$

a.e. $[\mu]$ *on* $\mathbf{G}$, *then* $S^{(\alpha)}$ *is a null series.*

PROOF. Choose the simple, measure preserving transformation T which satisfies (44). By (44) the g_m's are Walsh polynomials and by (45) $S^{(\alpha)}$ is a Walsh series which satisfies

$$S^{(\alpha)}_{q_n} = D^{(\alpha)}_{2^n} \circ T \qquad (n \in \mathbf{N}). \tag{47}$$

Since T is measure preserving and $D^{(\alpha)}_{2^n} \to 0$ a.e. $[\mu]$, as $n \to \infty$, it follows from (47) that $S^{(\alpha)}_{q_n} \to 0$ a.e. $[\mu]$ as $n \to \infty$. But by (46) we see that $S^{(\alpha)}_{q_n} \to 0$ a.e. $[\mu]$ if and only if $S^{(\alpha)}_n \to 0$ a.e. $[\mu]$, as $n \to \infty$. We conclude that $S^{(\alpha)}$ is a null series for any $-1 \le \alpha \le 1, \alpha \neq 0$. ∎

By Theorem 2 in 7.1 a Walsh series S whose partial sums satisfy $S_n \ge 0$ for $n \in \mathbf{N}$ which converges to a finite limit at all but countably many points of $\mathbf{G}$ is a Walsh-Fourier series. The following result shows that this finiteness condition cannot be relaxed even if S converges to zero a.e.

THEOREM 8. *There exist null series S whose partial sums satisfy*

$$S_n \ge 0 \qquad (n \in \mathbf{N}).$$

PROOF. Fix $n \in \mathbf{N}$ and for each $m \in \mathbf{N}$ with binary expansion

$$m = \sum_{j=0}^{n-1} m_j 2^j$$

define a point $\overline{m} \in \mathbf{G}$ by

$$\overline{m} := (m_{n-1}, m_{n-2}, \ldots, m_0, 0, 0, \ldots).$$

(This is the group analogue of bit-reversal.) Set

$$Q_n(x) := 2^{-n} \sum_{m=0}^{2^n-1} \psi_{(m+1)2^n} D_{2^n}(x + \overline{m}) \qquad (x \in \mathbf{G}). \tag{48}$$

Recall from 4.6 (see especially (50) through (53)) that

$$|Q_n| = 1$$
$$|\widehat{Q}_n(k)| \le 2^{-n} \qquad (k \in \mathbf{N})$$
$$\widehat{Q}_n(0) = 0$$

and

$$\|S^*(Q_n)\|_\infty \le 2.$$

The shift operator on **G** is defined by

$$\boldsymbol{\pi}(x_0, x_1, \dots) = (x_1, x_2, \dots)$$

and observe that

$$\rho_m \circ \boldsymbol{\pi}^k = \rho_{m+k} \qquad (m, k \in \mathbf{N}). \tag{49}$$

Thus if $\ell_n := 2^{2n}$ and

$$\gamma_n := Q_{\ell_n} \circ \boldsymbol{\pi}^{\ell_n}$$

then γ_n is a Walsh polynomial which satisfies

$$|\gamma_n| = 1 \tag{50}$$

$$\|\widehat{\gamma}_n\|_{\ell^\infty} \le 2^{-\ell_n} \tag{51}$$

$$\widehat{\gamma}_n(0) = 0 \tag{52}$$

and

$$\|S^*(\gamma_n)\|_\infty \le 2. \tag{53}$$

Moreover, there is an increasing sequence of natural numbers $(p_k, k \in \mathbf{N})$ such that the index of each Rademacher function which appears in γ_n belongs to the interval $[p_n, p_{n+1})$. It follows that

$$\boldsymbol{\gamma} := (\gamma_k, k \in \mathbf{N})$$

is a strongly multiplicative, sign-like system of Walsh polynomials whose product system $(g_m, m \in \mathbf{N})$ satisfies (45).

Fix $0 < \alpha \le 1/4$ and let $S^{(\alpha)}$ be the Walsh series defined in Lemma 9. Each $m \in [2^{p_n}, 2^{p_{n+1}})$ can be written uniquely as

$$m = p2^{p_n} + q$$

for $p, q \in \mathbf{N}$ satisfying $p < 2^{p_{n+1}-p_n}$ and $q < 2^{p_n}$. Consequently,

$$S_m^{(\alpha)} - S_{2^{p_n}}^{(\alpha)} = \alpha\, S_p(\gamma_n)\, S_{2^{p_n}}^{(\alpha)} + \alpha\widehat{\gamma}_n(p2^{p_n})\psi_{p2^{p_n}}\, S_q^{(\alpha)}$$

for $2^{p_n} \le m < 2^{p_{n+1}}$. It follows from (51) and (53) that

$$|S_m^{(\alpha)} - S_{2^{p_n}}^{(\alpha)}| \le 2|\alpha|\, |S_{2^{p_n}}^{(\alpha)}| + |\alpha|\, 2^{-\ell_n} |S_q^{(\alpha)}|. \tag{54}$$

In particular, the condition $|\alpha| \le 1/4$ and the definition of the norm of A (see 2.4) imply

(55) $$\max_{2^{p_n} \le m < 2^{p_{n+1}}} |S_m^{(\alpha)} - S_{2^{p_n}}^{(\alpha)}| \le \frac{1}{2}|S_{2^{p_n}}^{(\alpha)}| + 2^{-\ell_n - 2}\|S_{2^{p_n}}^{(\alpha)}\|_{\mathrm{A}}$$

for all $n \in \mathbf{N}$.

By construction

$$\|S_{2^{p_n}}^{(\alpha)}\|_{\mathrm{A}} \le \prod_{j=0}^{n-1} (1 + \|\gamma_j\|_{\mathrm{A}})$$

and by (48) and (51) we have

$$\|\gamma_j\|_{\mathrm{A}} \le 2^{\ell_j} \qquad (j \in \mathbf{N}).$$

Hence the choice $\ell_j = 2^{2j}$ implies

$$\begin{aligned} 2^{-\ell_n}\|S_{2^{p_n}}^{(\alpha)}\|_{\mathrm{A}} &\le 2^{-2^{2n}} \prod_{j=0}^{n-1} \left(1 + 2^{2^{2j}}\right) \\ &\le 2^{-2^{2n} + 2^{2n-1} + n} \\ &\le 2^{-n} \end{aligned}$$

for $n \in \mathbf{N}$. Substituting this estimate into (55) we obtain

(56) $$\max_{2^{p_n} \le m < 2^{p_{n+1}}} |S_m^{(\alpha)} - S_{2^{p_n}}^{(\alpha)}| \le \frac{1}{2}|S_{2^{p_n}}^{(\alpha)}| + 2^{-n-2}$$

for all $n \in \mathbf{N}$.

It is now evident that hypothesis (46) of Lemma 9 is satisfied and thus $S^{(\alpha)}$ is a null series. It remains to verify $S_m^{(\alpha)} \ge 0$ for all $m \in \mathbf{N}$. Fix $m \in \mathbf{N}$ and choose $n \in \mathbf{N}$ such that $2^{p_n} \le m < 2^{p_{n+1}}$. Then (56) implies

$$|S_m^{(\alpha)}| \ge |S_{2^{p_n}}^{(\alpha)}| - |S_m^{(\alpha)} - S_{2^{p_n}}^{(\alpha)}| \ge \frac{1}{2}|S_{2^{p_n}}^{(\alpha)}| - 2^{-n-2}.$$

But by (47),

$$S_{2^{p_n}}^{(\alpha)} = D_{2^n}^{(\alpha)} \circ T = \prod_{j=0}^{n-1} (1 + \alpha\rho_j \circ T)$$

so the condition $0 < \alpha \le 1/4$ implies

$$|S_{2^{p_n}}^{(\alpha)}| \ge \prod_{j=0}^{n-1} \left(1 - \frac{1}{4}\right) > 2^{-n}.$$

We conclude that

$$|S_m^{(\alpha)}| > 2^{-n-1} - 2^{-n-2} > 0. \qquad \blacksquare$$

By the Riesz-Fischer theorem no null series has coefficients which belong to ℓ^2. The following result shows this observation is sharp.

THEOREM 9. *Let $M_1 < M_2 < \dots$ be a sequence of numbers which diverges to ∞. Then there is a null series whose coefficients satisfy*

$$\sum_{j=0}^{m-1} a_j^2 \le M_m \qquad (m \in \mathbf{P}). \tag{57}$$

PROOF. Choose an increasing sequence $(p_k, k \in \mathbf{N})$ in $\mathbf{P}$ such that

$$M_{2^{p_{k-1}}} \ge 2^k \qquad (k \in \mathbf{P}).$$

For each $n \in \mathbf{N}$ choose by (49) positive integers ℓ_n, m_n such that the index of every Rademacher function which appears in

$$\gamma_n := Q_{\ell_n} \circ \boldsymbol{\pi}^{m_n}$$

belongs to the interval $[p_n, p_{n+1})$. The system $\boldsymbol{\gamma} := (\gamma_n, n \in \mathbf{N})$, by the proof of Theorem 8, is strongly multiplicative, sign-like, consists of Walsh polynomials, and its product system $(g_m, m \in \mathbf{N})$ satisfies (45). Moreover, there exist a simple, measure preserving transformation T and a Walsh series $S^{(\alpha)}$, for each $-1 \le \alpha \le 1$, $\alpha \ne 0$, such that (47) and (54) hold.

Let $(a_j, j \in \mathbf{N})$ represent the coefficients of $S^{(1)}$. Fix $m \in \mathbf{N}$ and choose $k \in \mathbf{N}$ such that

$$2^{p_{k-1}} \le m < 2^{p_k}.$$

Since T is measure preserving we have by (47) that

$$\begin{aligned}\sum_{j=0}^{m-1} a_j^2 &\le \int_{\mathbf{G}} |S_{2^{p_k}}^{(1)}|^2 \, d\mu \\ &= \int_{\mathbf{G}} |D_{2^k} \circ T|^2 \, d\mu \\ &= \int_{\mathbf{G}} D_{2^k}^2 \, d\mu \\ &= 2^k \\ &\le M_{2^{p_{k-1}}} \\ &\le M_m.\end{aligned}$$

Thus it remains to see that $S^{(1)}$ is a null series. We shall apply Lemma 9.

Fix $n \in \mathbf{N}$ and set

$$D^* := \sum_{k=0}^{\infty} D_{2^k}.$$

Notice by (54) that

$$|S_m^{(1)}| \le 3|S_{2^{p_n}}^{(1)}| + |S_q^{(1)}|$$

for $2^{p_n} \le m < 2^{p_{n+1}}$ and some $q < 2^{p_n}$. Hence by induction one has

$$|S_m^{(1)}| \le 3 \sum_{k=0}^{n} |S_{2^{p_k}}^{(1)}|$$

for $2^{p_n} \le m < 2^{p_{n+1}}$, whence by (47),

$$|S_m^{(1)}| \le 3\,(D^* \circ T) \qquad (m \in \mathbf{N}).$$

Using the full power of (54), it follows that

$$\max_{2^{p_n} \le m < 2^{p_{n+1}}} |S_m^{(1)} - S_{2^{p_n}}^{(1)}| \le 3\,(2^{-\ell_n})(D^* \circ T) + 2\,(D_{2^n} \circ T). \tag{58}$$

But Paley's lemma implies $D^*(x)$ is finite and $D_{2^n} \to 0$ as $n \to \infty$ for all $x \in \mathbf{G} \setminus \{0\}$. Therefore, the right side of (58) converges to 0 a.e. $[\mu]$ as $n \to \infty$. We conclude that (46) holds for $\alpha = 1$. ▮

7.5 U-sets and M-sets. A set E is called a *set of uniqueness*, or U-set, if the only Walsh series converging to zero outside E is the zero series, i.e., the series all of whose coefficients are identically zero.

Every measurable U-set is of Lebesgue measure zero. Indeed, suppose $E \subseteq [0,1)$ satisfies $|E| > 0$ and let K be a perfect subset of E with $|K| > 0$. Set

$$f := \chi(\mathrm{K}).$$

Since $f \in \mathrm{L}^2, S_n(f) \to f$ a.e. as $n \to \infty$, and the coefficients of $S(f)$ tend to zero. Thus given any open dyadic interval $I \subset [0,1] \setminus K$, it follows from Theorem 6 in 7.3 that $S(f)$ converges to zero everywhere on I. Since the complement of K is a union of such open dyadic intervals, we conclude that $S(f)$ converges to zero outside K. Since $\widehat{f}(0) := |K| > 0$, $S(f)$ is not the zero series. Consequently, neither K nor E are U-sets.

Not every set of Lebesgue measure zero is a U-set. Indeed, Theorem 7 shows that a set can be quite thin and still fail to be a U-set. However, if E is so thin that it contains no perfect subsets, then E is a U-set. For if S is a Walsh series which converges to zero outside E then the set

$$B := \{x \in E : S(x) \quad \text{does not converge to zero}\}$$

is a Borel set which contains no perfect subsets. It follows that B is countable, whence by Corollary 2 in 7.1, S is the zero series.

Because of this, the union of two U-sets may not be a U-set. Indeed, let E be the perfect set corresponding to $\Phi(x) := x$, $x \ge 0$, in Theorem 7. Then E is not a U-set. However, E can be broken into two subsets E_1, E_2 which have no perfect subsets. Of course, neither E_1 nor E_2 are Borel sets.

It is not known whether the union of two Borel U-sets is a U-set, unless one of them is closed.

THEOREM 10. *The union of countably many closed* U*-sets is itself a* U*-set.*

PROOF. Let $E_1, E_2, \ldots$ be closed U-sets and set

$$E := \bigcup_{n=1}^{\infty} E_n.$$

Let $S := \sum_{k=0}^{\infty} a_k w_k$ be a Walsh series which converges to zero off E. It suffices to show that the set

$$B := \{x \in [0,1) : \limsup_{n\to\infty} |S_n(x)| = \infty\}$$

is countable. Indeed, then $\varrho(B)$ is countable. Since S converges a.e. we also see that $a_k \to 0$ as $k \to \infty$. It follows that the character series

$$S^{\odot} := \sum_{k=0}^{\infty} a_k \psi_k$$

satisfies the $C-S$ condition and the finiteness condition of Theorem 3 in 7.2. Hence, $S^{\odot}$ is identically zero. In particular, S is also identically zero and the set E is a U-set as required.

Notice that B can be written as the intersection of the sets

$$\{x \in [0,1] : |S_n(x)| > k, \quad \text{for some integer} \quad n\},$$

for $k = 1, 2, \ldots$ Since each S_n is continuous except at finitely many points in $[0,1)$, it follows from the Baire category theorem that B is empty, countable, or of the second category on itself.

Suppose that B is of the second category on itself. Since $B = \cup_{n=1}^{\infty}(B \cap E_n)$, there exist an open dyadic interval J and an integer n such that $B \cap E_n \cap J$ is dense in $B \cap J \neq \emptyset$. We shall obtain a contradiction by showing that $B \cap J = \emptyset$.

We may suppose that the endpoints of J do not lie in E_n. Thus the set $K := E_n \cap J$ is a closed U-set contained in J. Since $x \in J \setminus K$ implies $x \notin B$, it is clear that

$$\limsup_{n\to\infty} |S_n(x)| < \infty \qquad (x \in J \setminus K).$$

Let $I \subset J \setminus K$ be a dyadic interval and apply Lemma 7 to S to choose a Walsh series T satisfying the $C-S$ condition such that $T_{2^n}(x) = 0$ for $x \notin I$ and n sufficiently large, such that $T_n \to 0$ a.e. on I, as $n \to \infty$, and such that (see (33))

$$\limsup_{n\to\infty} |T_{2^n}(x)| < \infty$$

for $x \in I$. By Theorem 3, T is the zero series. It follows from Lemma 7 that $S(x)$ converges to zero for every $x \in I$. Since I was any dyadic interval contained in the open set $J \setminus K$, we have established that $S(x)$ converges to zero for every $x \in J \setminus K$.

Apply Lemma 7 again, this time to the interval J, choosing a Walsh series $\widetilde{S}$ which converges to zero outside J and which is equiconvergent with S on J. We have just seen that S converges to zero on $J \setminus K$. It follows that $\widetilde{S}$ converges to zero off K. Since K is a U-set, $\widetilde{S}$ is identically zero. It follows that $S(x)$ converges to zero for every $x \in J$. Hence by the definition of B, we have $J \cap B = \emptyset$. ∎

The following result gives a useful method for constructing U-sets.

THEOREM 11. *A set $E \subset [0,1)$ is a U- set if there exist Walsh polynomials*

$$P_\ell(x) = \sum_{j=1}^{N(\ell)} b_j(\ell) w_j(x) \qquad (\ell \in \mathbf{P}),$$

which satisfy the following four properties:

$$P_\ell(x) = 0 \qquad (x \in E,\, \ell \in \mathbf{P}), \tag{59}$$

$$\sum_{j=0}^{N(\ell)} |b_j(\ell)| \le A < \infty \qquad (\ell \in \mathbf{P}), \tag{60}$$

$$|b_0(\ell)| \ge C > 0 \qquad (\ell \in \mathbf{P}), \tag{61}$$

and

$$\lim_{\ell \to \infty} b_j(\ell) = 0 \qquad (j \in \mathbf{P}). \tag{62}$$

PROOF. Suppose $S = \sum_{k=0}^{\infty} a_k w_k$ is a Walsh series which converges to zero outside E. We first show that

$$a_k = -\frac{1}{b_0(\ell)} \sum_{j=1}^{N(\ell)} b_j(\ell) a_{k \oplus j} \tag{63}$$

for $k \in \mathbf{N}$ and $\ell \in \mathbf{P}$. Indeed, fix $\ell \in \mathbf{P}$ and apply Lemma 6 to choose a Walsh series T whose coefficients satisfy

$$c_k = \sum_{j=0}^{N(\ell)} b_j(\ell) a_{k \oplus j} \qquad (k \in \mathbf{N}),$$

such that

$$T_{2^n} := P\, S_{2^n}$$

for $n > \log_2 N(\ell)$. Hypothesis (59) and the choice of S imply that $T_{2^n} \to 0$ everywhere on $[0,1)$ as $n \to \infty$. By Theorem 3, therefore, T is the zero series, i.e., $c_k = 0$ for $k = 0, 1, \ldots$ Since $k \oplus 0 = k$, it follows that (63) holds for $k \in \mathbf{N}$ and $\ell \in \mathbf{P}$.

To show $a_k = 0$ fix $k \in \mathbf{N}$ and let $\varepsilon > 0$. Since $k \oplus j \to \infty$ as $j \to \infty$, choose j_0 so large that $j \ge j_0$ implies $|a_{k\oplus j}| \le \varepsilon$. It follows from (61) and (63) that

$$|a_k| \le \frac{1}{C} \left(\sum_{j=1}^{j_0 - 1} |b_j(\ell) a_{k \oplus j}| + \varepsilon \sum_{j=j_0}^{N(\ell)} |b_j(\ell)| \right)$$

for $\ell \in \mathbf{P}$. Let $\ell \to \infty$. By (60) and (62) we conclude that

$$|a_k| \le (A/C)\varepsilon.$$

Therefore, $a_k = 0$ as promised. ∎

A set which is not a U-set is called a *set of multiplicity*, or an M-set. Clearly, every M-set is uncountable and every set of positive Lebesgue measure is an M-set.

An M-set can be thin since by Theorem 7 in 7.3 there exist M-sets of Hausdorff measure zero. We shall presently show that no M-set can be so thin that certain of its dilations consistently miss a fixed interval. To define what is meant by dilations we need the dyadic product.

In 9.1 we shall introduce the dyadic field $(\mathbf{F}, +, \cdot)$ which contains the dyadic group $\mathbf{G}$ as an additive subgroup. We will show that the characters of $\mathbf{G}$ can be extended to $\mathbf{F}$ (see (11) in 9.1), and that $\mathbf{F}$ can be identified with $[0, \infty)$ in the same way that $\mathbf{G}$ was identified with $[0, 1)$ (see (5) in 9.1 and the remarks following Theorem 6 in 9.2). It follows that there is a dyadic product on $[0, \infty)$ which is distributive over dyadic addition $\dot{+}$. We shall denote the dyadic product of two elements $x, y \in [0, \infty)$ by $x \cdot y$.

The dyadic product and Walsh functions interact in the expected manner.

LEMMA 10. *If* $m, k \in \mathbf{N}$ *and* $x \in [0, 1)$ *then*

$$w_k(m \cdot x) = w_{m \cdot k}(x).$$

PROOF. By (9) and (11) in 9.1 and the identification of $\mathbf{F}$ with $[0, \infty)$ there is a map $\wp : [0, \infty) \to \{0, 1\}$ such that

$$w_k(x) = (-1)^{\wp(k \cdot x)}$$

for all $k \in \mathbf{N}$ and $x \in [0, 1)$. Therefore, the fact that the dyadic product is associative implies

$$w_k(m \cdot x) = (-1)^{\wp(k \cdot m \cdot x)} = w_{k \cdot m}(x)$$

for all $k, m \in \mathbf{N}$ and $x \in [0, 1)$. ∎

A subset E of $[0, 1)$ is called an H-*set* if there exist non-negative integers $n_1 < n_2 < \dots$ and an interval $I_0 \subset [0, 1)$ such that $n_\ell \cdot x$ never belongs to I_0 for any $x \in E$ and any $\ell \in \mathbf{P}$. Thus every H-set is a nowhere dense set which is in a certain sense, sparsely distributed in the interval $[0, 1)$.

THEOREM 12. *Every* H-*set is a* U-*set*.

PROOF. Let E be an H-set, let $0 \le n_1 < n_2 < \dots$ be integers, and let I_0 be a dyadic interval such that

$$(64) \qquad n_\ell \cdot x \notin I_0 \qquad (x \in E, \ell \in \mathbf{P}).$$

The function $f := \chi(I_0)$ is a Walsh polynomial and can be written as

$$f(x) = \sum_{j=0}^{N} a_j w_j(x) \qquad (x \in [0, 1)).$$

Set $P_\ell(x) := f(n_\ell \cdot x)$ for $\ell \in \mathbf{P}$, and $x \in [0,1)$. Observe by Lemma 10 that

$$(65) \qquad P_\ell(x) = \sum_{j=0}^{N} a_j w_{n_\ell \cdot j}(x) \qquad (\ell \in \mathbf{P}).$$

Hence $P_1, P_2, \ldots$ are Walsh polynomials.

We shall prove that these polynomials and the set E satisfy the hypotheses of Theorem 11. It will follow that E is a U-set.

Hypothesis (59) is satisfied because f vanishes off I_0 and $n_\ell \cdot x$ never belongs to I_0 when $x \in E$.

Hypothesis (60) is satisfied by

$$A := \sum_{j=0}^{N} |a_j|.$$

Since $m \cdot k = 0$ if and only if $m = 0$ or $k = 0$, the constant term of each polynomial P_ℓ is a_0. Hence hypothesis (61) is satisfied by $C := |a_0| = |I_0|$.

Finally, since for each fixed $j \in \mathbf{P}$, $n_\ell \cdot j \to \infty$ as $\ell \to \infty$, and since there are only finitely many non-zero coefficients in (65), it follows (using the notation of Theorem 11) that $b_j(\ell) = 0$ for ℓ sufficiently large. Therefore, hypothesis (62) is satisfied. ∎

A set $E \subset [0,1)$ is called a *Dirichlet set* if there exist non-negative integers $m_1, m_2, \ldots$ with

$$\limsup_{n\to\infty} m_n = \infty$$

such that $w_{m_n}(x) = 1$ for $x \in E$ and $n \in \mathbf{P}$.

THEOREM 13. *Every Dirichlet set is a* U*-set.*

PROOF. Let E be a Dirichlet set, let $m_1, m_2, \ldots$ be an unbounded sequence of positive integers and suppose that $w_{m_n}(x) = 1$ for $x \in E$ and $n \in \mathbf{P}$.

Since $(m_n, n \in \mathbf{P})$ is unbounded, it contains a subsequence $n_1 < n_2 < \ldots$ with $n_\ell \to \infty$ as $\ell \to \infty$. Moreover, by assumption and Lemma 10, $w_1(n_\ell \cdot x) = 1$ for $x \in E$ and $\ell \in \mathbf{P}$. In particular, $n_\ell \cdot x \notin I_0 := (1/2, 1)$ for $x \in E$ and $\ell \in \mathbf{P}$. It follows that E is an H-set. Therefore E is a U- set by Theorem 12. ∎

In particular, the Cantor set formed by removing middle halves from $[0,1)$ is a U-set. In fact, any Cantor set formed by removing middle intervals from $[0,1)$ with ratio of dissection equal 2^{-p} for some integer $p \in \mathbf{P}$ is a U-set. However, in the Walsh case no other symmetric Cantor set has been classified as a U-set or as an M-set. Thus it is not known whether the Cantor middle thirds set is a U- set for Walsh series.

Also unsolved is the problem of characterizing U-sets in the dyadic group, except for the class of closed subgroups.

THEOREM 14. *Let H be a closed subgroup of the dyadic group* $\mathbf{G}$. *Then H is a* U*-set if and only if H is of Haar measure zero.*

PROOF. Every U-set in $\mathbf{G}$ is of Haar measure zero.

Conversely, suppose $\mu(H) = 0$. Consider the factor group

$$\Gamma := \mathbf{G}/H$$

and suppose for a moment that its dual group $\widehat{\Gamma}$ contains infinitely many elements. It will follow from Theorem 8 in Appendix 0.3 that there exist non-negative integers $m_1, m_2, \ldots$ such that $\psi_{m_n} = 1$ on H for $n = 1, 2, \ldots$, and

$$\limsup_{n\to\infty} m_n = \infty.$$

Hence H is a Dirichlet set in $\mathbf{G}$, whence by the group analogue of Theorem 13, H is a U-set.

An elementary proof that $\widehat{\Gamma}$ is infinite appears below. The reader familiar with Pontryagin duality need read no further. Indeed, Γ is necessarily uncountable because $\mathbf{G}$ can be written as a union of cosets of H, and every coset of H is of Haar measure zero.

The elementary proof rests on showing that $\widehat{\Gamma}$ and $\widehat{\mathbf{G}}$ are the same groups. Indeed, we shall construct a homeomorphic isomorphism F from $\mathbf{G}$ onto Γ.

Let $e_0, e_1, \ldots$ denote the usual closed system in $\mathbf{G}$. Set $n(0) := 0$ and suppose $n(i-1)$ has been chosen for some $i \in \mathbf{P}$. Choose $n(i) > n(i-1)$ to be the smallest index k for which e_k does not belong to the $\mathbf{Z}_2$-linear span of

$$E := \{e_j + y : 0 \le j < k, y \in H\}.$$

This choice is possible, for if e_k belonged to the $\mathbf{Z}_2$- linear span of E for all $k > n(i-1)$ then the $\mathbf{Z}_2$-linear span $\mathbf{G}_0$ of the system $(e_j, j \in \mathbf{N})$ would be contained in a finite union of cosets of H. Since $\mathbf{G}_0$ is dense in $\mathbf{G}$ and the finite union of cosets of H is a closed set of Haar measure zero, it would follow that $\mu(\mathbf{G}) = 0$, a contradiction.

Therefore, there exists an infinite $\mathbf{Z}_2$-linearly independent system

$$(H + e_{n(i)}, i \in \mathbf{N})$$

whose $\mathbf{Z}_2$-linear span Γ_0 is dense in Γ. Set

$$F(e_i) := H + e_{n(i)} \qquad (i \in \mathbf{N})$$

and extend F to $\mathbf{G}_0$ by $\mathbf{Z}_2$-linearity. Use the dyadic metric on $\mathbf{G}$ to define a metric on Γ by

$$|H + x| := \inf_{y \in H} |x + y|.$$

Observe that $|x| \le |H + x|$ for all $x \in \mathbf{G}$. Since $n(i) \ge i$ for $i \in \mathbf{N}$, it follows that $|F(x)| \le |x|$ for all $x \in \mathbf{G}_0$. Hence F has a unique, continuous $\mathbf{Z}_2$-linear extension from the dense set $\mathbf{G}_0$ to all of $\mathbf{G}$. This extension takes $\mathbf{G}$ onto Γ because Γ_0 is dense in Γ. Moreover, it is 1-1 because $F(x) = 0$ implies $x = 0$.

It remains to verify that F is open. Since F is linear and since the compact open subgroups $I_0(0), I_1(0), \ldots$ form a neighborhood base at $0 \in \mathbf{G}$ we need only show that $F(\mathrm{K})$ is open in $\mathbf{G}/H$ for every compact, open subgroup K in $\mathbf{G}$. However, $\mathbf{G}$ is contained in a finite union of translates of K, so the compact group $\mathbf{G}/H$ is contained in a finite union of translates of $F(\mathrm{K})$. It follows from the Baire category theorem that one of those translates has interior, i.e., $F(\mathrm{K})$ has interior V in $\mathbf{G}/H$. Since $F(\mathrm{K})$ is a subgroup in $\mathbf{G}/H$, it is clear that $x + V \subset F(\mathrm{K})$ for all $x \in F(\mathrm{K})$. Consequently, $F(\mathrm{K})$ is open in $\mathbf{G}/H$ as required. ∎

7.6 Uniqueness and Cesàro Summability. The connection between quasi-measures and Walsh series can be used to show that 2^n-th partial sums of Walsh series are either very bad or very good.

THEOREM 15. *Let E be a measurable subset of $[0,1)$ and S be a Walsh series. Then except on a subset of E of Lebesgue measure zero, either*

$$\limsup_{n\to\infty} S_{2^n}(x) = +\infty \quad \textit{and} \quad \liminf_{n\to\infty} S_{2^n}(x) = -\infty$$

or

$$\lim_{n\to\infty} S_{2^n}(x) \quad \textit{exists and is finite for } x \in E.$$

In particular, if

$$\liminf_{n\to\infty} S_{2^n}(x) > -\infty \qquad (x \in E)$$

then S_{2^n} converges a.e. on E, as $n \to \infty$, to some finite limit.

PROOF. Let $\nu \in \mathrm{QM}$ such that $S = S\nu$. A classical theorem on binary derivates (see Theorem 14 in Appendix 0.6) says either

$$\limsup_{n\to\infty} \frac{\nu(I_n(x))}{|I_n(x)|} = +\infty, \text{ and } \liminf_{n\to\infty} \frac{\nu(I_n(x))}{|I_n(x)|} = -\infty$$

or

$$\lim_{n\to\infty} \frac{\nu(I_n(x))}{|I_n(x)|} \quad \text{exists and is finite for a.e. } x \in E.$$

Since by (3) in 7.1 we have

$$S_{2^n}(x) = \frac{\nu(I_n(x))}{|I_n(x)|}$$

for $n \in \mathbf{N}$ and $x \in [0,1)$, the proof of the theorem is complete. ∎

Let S be a Walsh series and let

$$\sigma_n(x) := \frac{1}{n}\sum_{j=1}^{n} S_j(x) \qquad (n \in \mathbf{P})$$

represent its Cesàro means. We shall show that Theorem 3 in 7.1 holds with σ in place of S.

It is a peculiarity of the Walsh system that S_{2^n} converges a.e. if σ_{2^n} is bounded. In fact,

THEOREM 16. *If $E \subset [0,1)$ is measurable and*

$$\limsup_{n\to\infty} |\sigma_{2^n}(x)| < \infty \qquad (x \in E), \tag{66}$$

then there is a finite-valued, measurable function f such that

$$\lim_{n\to\infty} S_{2^n}(x) = f(x)$$

and

(67) $$\liminf_{n\to\infty} \sigma_{2^n}(x) \le f(x) \le \limsup_{n\to\infty} \sigma_{2^n}(x)$$

for a.e. $x \in E$.

PROOF. Suppose first that $\mathrm{K} \subset [0,1)$ is a set of positive measure and let $x \in \mathrm{K}$ be a dyadic irrational point of density for K. Choose integers $0 \le p_m < 2^m$ so that

$$\{x\} = \bigcap_{m=1}^{\infty} I(p_m, m).$$

We claim that there exist a number M_0 and a sequence $t_m := t_m(x) \to 0$, as $m \to \infty$, such that if $m > M_0$ and $2^m \le n \le 2^{m+1}$, then

(68) $$x + t_m \in \mathrm{K},$$

(69) $$x, x + t_m \quad \text{belong to different halves of} \quad I(p_m, m)$$

and

(70) $$\frac{1}{2}(\sigma_n(x + t_m) + \sigma_n(x)) = \frac{2^m}{n}\sigma_{2^m}(x) + \frac{(n - 2^m)}{n} S_{2^m}(x).$$

To prove the claim, fix $m > 0$, let b_m be the midpoint of $I(p_m, m)$, and set $a_m := b_m - x$. If $a_m > 0$ (which means that x belongs to the left half of $I(p_m, m)$) then set $J_m^{(1)} := (x, b_m)$, $J_m^{(2)} := (x, a_m + b_m)$ and $J_m^{(3)} := (b_m, a_m + b_m)$. On the other hand, if $a_m < 0$ then set $J_m^{(1)} := (b_m, x)$, $J_m^{(2)} := (a_m + b_m, x)$, and $J_m^{(3)} := (a_m + b_m, b_m)$. In particular,

(71) $$J_m^{(3)} \text{ and } x \text{ belong to different halves of } I(p_m, m)$$

and

(72) $$|\mathrm{K} \cap J_m^{(3)}| = |\mathrm{K} \cap J_m^{(2)}| - |\mathrm{K} \cap J_m^{(1)}|.$$

Since x is a point of density and $|J_m^{(1)}| = |J_m^{(3)}| = |a_m| = |J_m^{(2)}|/2$, both

$$\frac{|\mathrm{K} \cap J_m^{(2)}|}{2|a_m|} \quad \text{and} \quad \frac{|\mathrm{K} \cap J_m^{(1)}|}{|a_m|}$$

converge to 1 as $m \to \infty$. Hence by (72) we have

$$\lim_{m\to\infty} \frac{|\mathrm{K} \cap J_m^{(3)}|}{|J_m^{(3)}|} = 1.$$

In particular, there is an M_0 so large that $\mathrm{K} \cap J_m^{(3)} \neq \emptyset$ for $m > M_0$.

Let $m > M_0$ and choose a number t_m such that $x_m := x + t_m \in \mathrm{K} \cap J_m^{(3)}$. Since $|J_m^{(3)}| \to 0$ it is clear that $x_m \to x$ and thus $t_m \to 0$, as $m \to \infty$. Moreover, (68) holds by the choice of t_m and (69) follows immediately from (71).

To verify (70), fix $2^m \le n \le 2^{m+1}$ and set

$$P_k := \sum_{j=2^m}^{k-1} a_j w_j \qquad (k > 2^m).$$

Thus $S_k = S_{2^m} + P_k$ for $2^m < k \le 2^{m+1}$, and by definition,

$$\sigma_n = \frac{1}{n}\sum_{k=1}^{2^m} S_k + \frac{1}{n}\sum_{k=2^m+1}^{n} S_k.$$

It follows, therefore, that

$$\sigma_n = \frac{2^m}{n}\sigma_{2^m} + \frac{1}{n}\sum_{k=2^m+1}^{n}(S_{2^m} + P_k). \tag{73}$$

Recall that w_j is constant on $I(p_m, m)$ for $0 \le j < 2^m$, and that w_j changes signs on $I(p_m, m)$ for every integer j satisfying $2^m \le j < 2^{m+1}$. Hence it follows from (69) that

$$\sigma_{2^m}(x) = \sigma_{2^m}(x + t_m),$$
$$S_{2^m}(x) = S_{2^m}(x + t_m),$$

and

$$P_k(x + t_m) = -P_k(x) \qquad (2^m < k \le n).$$

Evaluate (73) at x and $x + t_m$, and add the resulting equations. Since the terms involving P_k drop out, (70) results at once.

The claim is established.

To prove the theorem we may suppose that $|E| > 0$. Let $\varepsilon > 0$, $y \in E$, and set

$$F(y) := \liminf_{n\to\infty} \sigma_{2^n}(y).$$

By hypothesis, F is finite-valued on E. Moreover, if $F_m := \min\{F, \sigma_{2^{(m+1)}}\}$ then $F_m \to F$ a.e. on E, as $m \to \infty$. Hence by Egoroff's theorem and Lusin's theorem there is a perfect set $\mathrm{K} \subset E$ with $|\mathrm{K}| > |E| - \varepsilon$ such that F is continuous on K and $F_m \to F$ uniformly on K, as $m \to \infty$. Thus for any fixed $x \in \mathrm{K}$ and $x + t_m \in \mathrm{K}$ with $t_m \to 0$, we can choose m so large that

$$\begin{aligned}\sigma_{2^{(m+1)}}(x + t_m) &\ge F_m(x + t_m)\\ &\ge F(x + t_m) - \varepsilon/2\\ &\ge F(x) - \varepsilon.\end{aligned}$$

In particular,

(74) $$\sigma_{2(m+1)}(x+t_m) \geq \liminf_{n\to\infty} \sigma_{2^n}(x) - \varepsilon$$

for large m.

If we let x be a point of density for K it follows from (70) and (74) that

(75) $$\liminf_{n\to\infty} \sigma_{2^n}(x) - \varepsilon + \sigma_{2(m+1)}(x) \leq \sigma_{2^m}(x) + S_{2^m}(x)$$

for large m. Since this inequality implies

$$2\liminf_{n\to\infty} \sigma_{2^n}(x) \leq \limsup_{n\to\infty} \sigma_{2^n}(x) + S_{2^m}(x) + 3\varepsilon,$$

and since $|K| > |E| - \varepsilon$, we have by hypothesis that

(76) $$\liminf_{m\to\infty} S_{2^m}(x) > -\infty$$

for a.e. $x \in E$. Hence by Theorem 15 there is a finite-valued function f such that $S_{2^m} \to f$ a.e. on E, as $m \to \infty$.

Let E_1 represent the set of points $x \in E$ which are points of density for E and which satisfy $S_{2^m}(x) \to f(x)$ as $m \to \infty$. Then $|E_1| > |E| - \varepsilon$ and by (75) we have

$$\liminf_{n\to\infty} \sigma_{2^n}(x) - \varepsilon + \sigma_{2^m}(x) \leq \limsup_{n\to\infty} \sigma_{2^n}(x) + f(x) + 2\varepsilon$$

for $x \in E_1$ and m large. Take the limit supremum of this inequality, as $m \to \infty$. The following inequality eventuates:

$$\liminf_{n\to\infty} \sigma_{2^n}(x) \leq f(x) + 3\varepsilon \qquad (x \in E_1).$$

Thus the left inequality in (67) is obtained by letting $\varepsilon \to 0$.

A similar argument establishes the right inequality in (67). ∎

THEOREM 17. *Let S be a Walsh series which satisfies the $C - S$ condition and suppose $f \in L^1(G)$. If*

(77) $$\limsup_{n\to\infty} |\sigma_{2^n}(x)| < \infty$$

for all but countably many $x \in G$ and

$$\lim_{n\to\infty} \sigma_{2^n} = f$$

a.e. $[\mu]$ on G then S is the Walsh-Fourier series of f.

PROOF. It suffices to show

(78) $$S_{2^n}(x) = (S_{2^n}f)(x) \qquad (n \in \mathbf{N})$$

for all but countably many $x \in \mathbf{G}$. Since Fine's map takes [0,1] onto $\mathbf{G} \setminus \mathbf{G}_0^*$ and $\mathbf{G}_0^*$ is countable, we shift our attention to the interval $[0,1)$. Thus we suppose that S is a Walsh series on $[0,1)$ which satisfies the $C-S$ condition (see Exercise 7.1), that there is a countable subset E of $[0,1)$ such that (77) holds for $x \in [0,1) \setminus E$, and that $\sigma_{2^n} \to f$ a.e. as $n \to \infty$ for $f \in \mathrm{L}^1$. In particular, it follows from Theorem 16 that $S_{2^n} \to f$ a.e. as $n \to \infty$.

For each $n \in \mathbf{N}$, (78) is equivalent to

$$S_{2^n}(x) = 2^n \int_{I(p,n)} f, \tag{79}$$

where $I(p,n) = I_n(x)$. Thus if we let V denote the largest open subset of [0,1) such that (79) holds for all $I(p,n) \subset V$, then it suffices to show V is the open interval $(0,1)$.

Suppose to the contrary that $V \neq (0,1)$. Thus $K := [0,1] \setminus V$ is a non-empty, closed set. Apply Lemma 4 in 7.2 to the series $S - Sf$ to verify that K has no isolated points. In particular, K is a non-empty, perfect set. It follows from the Baire category theorem and hypothesis that

$$|\sigma_{2^n}(x)| < C \qquad (x \in K_1 \setminus E,\ n \in \mathbf{N}), \tag{80}$$

where $C < \infty$ is an absolute constant and $K_1 := I_0 \cap K$ for some dyadic interval I_0.

Suppose without loss of generality that $\mathbf{Q} \subset E$ and let $x \in K_1 \setminus E$. Thus to each $n \in \mathbf{P}$ there correspond non-negative integers $0 \le p_n < 2^{m_n}$ such that $m_n < m_{n+1}$, $x \in I(p_n, m_n)$ and each of the halves $I_1 := I(2p_n, m_n+1)$ and $I_2 := I(2p_n+1, m_n+1)$ intersect $K_1 \setminus E$, say $y_i \in I_i \cap (K_1 \setminus E)$ for $i = 1,2$. In particular, it follows from (80) that

$$|\sigma_{2^{m_n+1}}(y_i)| < C \qquad (i = 1,2). \tag{81}$$

However, $y_1 \in I_1$ and $y_2 \in I_2$ imply $w_j(y_1) = w_j(y_2)$ for $0 \le j < 2^{m_n}$ and $w_j(y_1) = -w_j(y_2)$ for $2^{m_n} \le j < 2^{m_n+1}$. Consequently,

$$\sigma_{2^{m_n+1}}(y_1) + \sigma_{2^{m_n+1}}(y_2) = \sigma_{2^{m_n}}(y_3) + S_{2^{m_n}}(y_3)$$

holds for any $y_3 \in I(p_n, m_n)$. It follows, therefore, from (81), (80), and the condition $x \in I(p_n, m_n)$, that

$$|S_{2^{m_n}}(x)| \le 3C \qquad (x \in K_1 \setminus E, n \in \mathbf{N}). \tag{82}$$

Let $\widetilde{I}$ be the open interval whose endpoints coincide with those of the portion K_1. Let J_0 be any dyadic interval satisfying $J_0 \subset \widetilde{I}$ and let $m \in \mathbf{N}$ be determined by $|J_0| = 2^{-m}$.

Write $[0,1) = E_1 \cup E_2 \cup E_3$ where $E_1 := K_1 \cap J_0, E_2 := V \cap J_0$, and $E_3 := [0,1) \setminus J_0$. For each $x \in E_1$ set $m(x) := +\infty$. For each $x \in E_2$ let $m(x)$ represent the smallest integer for which there exists a dyadic interval of the form $I = I(q(x), m(x))$ such that x belongs to the interior of I which in turn is a subset of E_2. The choice of V therefore implies

$$S_{2^{m(x)}}(x) = 2^{m(x)} \int_{I(q(x),m(x))} f \qquad (x \in E_2). \tag{83}$$

Let $\boldsymbol{a} = (a_k, k \in \mathbf{N})$ represent the coefficients of S, let $\boldsymbol{b} = (b_k, k \in \mathbf{N})$ be the Hadamard transform of $\boldsymbol{a}$, and consider the Haar series

$$H := \sum_{j=0}^{\infty} b_j h_j.$$

By (42) in 1.4,

$$S_{2^n} = H_{2^n} \tag{84}$$

for $n \in \mathbf{N}$.

Let $n_1 < n_2 \ldots$ represent the collection of all integers n which satisfy the conditions $0 \le n < 2^{m(x)}$ and $h_n(x) \ne 0$ for some $x \in E_1 \cup E_2$ and consider the subseries

$$\widetilde{H} := \sum_{j=1}^{\infty} b_{n_j} h_{n_j}.$$

By choice of the indices n_j, it is clear that

$$\widetilde{H}_{2^n}(x) = H_{2^n}(x) \qquad (x \in E_1, n \in \mathbf{N})$$

and

$$\widetilde{H}_{2^n}(x) = H_{2^{m(x)}}(x) \qquad (x \in E_2, n \ge m(x)).$$

Let $\widetilde{\boldsymbol{b}} = (\widetilde{b}_k, k \in \mathbf{N})$ where $\widetilde{b}_k := b_{n_j}$ if $k = n_j$ for some j, and $\widetilde{b}_k := 0$ otherwise. Let $\boldsymbol{c}$ be the Hadamard transform of $\widetilde{\boldsymbol{b}}$ and set

$$\widetilde{S} := \sum_{k=0}^{\infty} c_k w_k.$$

By (42) in 1.4, $\widetilde{S}_{2^n} = \widetilde{H}_{2^n}$ for $n \in \mathbf{N}$. It follows, therefore, from (84) and the construction of $\widetilde{H}$ that

$$\widetilde{S}_{2^n}(x) = S_{2^n}(x) \qquad (x \in E_1, n \in \mathbf{N}), \tag{85}$$

and

$$\widetilde{S}_{2^n}(x) = S_{2^{m(x)}}(x) \qquad (x \in E_2, n \ge m(x)). \tag{86}$$

Apply Lemma 6 in 7.3 to the Walsh series $\widetilde{S}$ and the Walsh polynomial $P := \chi(J_0)$. Hence choose a Walsh series T which satisfies

$$T_{2^n} = P\,\widetilde{S}_{2^n} \qquad (n \ge m).$$

Clearly, identities (85) and (86) imply

(87) $$T_{2^n}(x) = S_{2^n}(x) \qquad (x \in E_1,\, n \geq m),$$

(88) $$T_{2^n}(x) = S_{2^{m(x)}}(x) \qquad (x \in E_2,\, n \geq \max\{m, m(x)\}),$$

and

(89) $$T_{2^n}(x) = 0 \qquad (x \in E_3, n \geq m).$$

In particular, T satisfies the $C - S$ condition. Moreover, (82), (87), (88), and (89) imply

$$\limsup_{n\to\infty} |T_{2^{m_n}}(x)| < \infty \qquad (x \in [0,1] \setminus E).$$

Finally, by (87), (88), (89) it is clear that $T_{2^n} \to \varphi$ a.e., as $n \to \infty$, where $\varphi(x) := f(x)$ for $x \in E_1$, $\varphi(x) = S_{2^{m(x)}}(x)$ for $x \in E_2$, and $\varphi(x) = 0$ for $x \in E_3$. It follows from Theorem 5 in 7.2 that T is the Walsh-Fourier series of φ. In particular,

$$T_{2^m}(x) = 2^m \int_{J_0} \varphi$$

for $x \in J_0$ and $m \in \mathbf{N}$. Since (83) and the definition of φ imply

$$\int_{J_0} \varphi = \int_{J_0} f,$$

and since S_{2^m} is constant on J_0, we conclude from (87) that

$$S_{2^m}(x) = 2^m \int_{J_0} f$$

for $x \in J_0$. But J_0 was any dyadic interval contained in $\widetilde{I}$. Hence (79) holds for all dyadic intervals $I(p,n)$ whose interior is contained in $\widetilde{I}$. Since V is the largest open set enjoying this property, it follows that $\widetilde{I} \subset V$. This contradicts the fact that $\widetilde{I} \cap K_1 \neq \emptyset$. ∎

The techniques of 7.1 can be modified to prove the following Cesàro analogue of Theorem 2.

THEOREM 18. *Suppose* $f \in \mathrm{L}^1(\mathbf{G})$ *and* S *is a Walsh series which satisfies the* $C - S$ *condition. Suppose further that*

(90) $$\limsup_{n\to\infty} (S_{2^n}(x) - \sigma_{2^{n+1}}(x)) \geq 0$$

and

$$\limsup_{n\to\infty} |\sigma_{2^n}(x)| < \infty$$

for all but countably many $x \in \mathbf{G}$. *If*

$$\liminf_{n\to\infty} \sigma_{2^n}(x) \geq f(x) \qquad (x \in \mathbf{G}),$$

then S *is the Walsh-Fourier series of* f.

It is not known whether hypothesis (90) is superfluous, even in the case that f is bounded.

Exercises

7.1 Let $S^{\odot} = \sum a_k \psi_k$ and $S = \sum a_k w_k$. Show $S^{\odot}$ satisfies the $C - S$ condition if and only if

$$(91) \qquad \begin{cases} \lim_{n\to\infty} 2^{-n} S_{2^n}(x) = 0 & (x \in [0,1)) \\ \lim_{n\to\infty} 2^{-n} S_{2^n}(x-) = 0 & (x \in (0,1]). \end{cases}$$

7.2 Let $g : [0,1) \to \mathbf{R}$ be an increasing function and define $\nu \in \mathrm{QM}$ by

$$\nu(I) := g(\beta) - g(\alpha)$$

for $I = [\alpha, \beta)$ a dyadic interval. Prove ν is σ-additive on the collection of dyadic intervals if and only if g is left continuous on $\mathbf{Q}$.

7.3 A Borel measure ν on $[0,1]$ is called non-atomic if $\nu(\{x\}) = 0$ for all $x \in [0,1]$. Show that the Walsh-Fourier- Stieltjes series of a non-atomic measure satisfies

$$2^{-n} \sum_{k=0}^{2^n-1} a_k^2 \to 0$$

as $n \to \infty$.

[Fine [3]]

7.4 A Haar series S is said to satisfy the $C-S$ condition if (91) holds. Use the Hadamard transform to prove Theorems 3 and 5 for Haar series which satisfy the $C - S$ condition.

7.5 Use Exercise 7.4 to establish the following uniqueness result for Haar series. If S is a Haar series satisfying the $C - S$ condition, if $n_1 < n_2 < \dots$ is a sequence of integers such that

$$\limsup_{j\to\infty} |S_{n_j}(x)| < \infty$$

for all but countably many $x \in [0,1]$, and if $S_{n_j} \to f$ a.e. as $j \to \infty$, for some $f \in \mathrm{L}^1$, then S is the Haar-Fourier series of f.

Hint: Prove for any fixed $x \in [0,1]$ that S converges at x if and only if $S_{2^n}(x)$ converges, as $n \to \infty$.

7.6 A Haar series $\sum_{k=0}^{\infty} c_k h_k$ is said to satisfy the $A-T$ condition if given any $x_0 \in [0,1]$,

$$\lim_{j\to\infty} \frac{c_{k_j}}{h_{k_j}(x_0)} = 0$$

where $k_1 < k_2 < \dots$ are those indices ℓ which satisfy $h_\ell(x_0) \neq 0$. Use Exercise 7.5 to show that uniqueness holds for Haar series which satisfy the $A - T$ condition.

[Arutunjan and Talaljan [1]]

7.7 If S is a Walsh series which satisfies

$$\limsup_{n\to\infty} |S_{2^n}(x)| < \infty$$

for all $x \in [0,1]$ and $S_{2^n} \to f$ a.e., as $n \to \infty$, for some $f \in L^1$ then show S is the Walsh-Fourier series of f.

7.8 Use the proof of Lemma 4 directly to show that if

$$\limsup_{j\to\infty} |S_{2^{n_j}}(x\pm)| < \infty$$

for all $x \in [0,1)$ and $S_{2^{n_j}} \to f$ a.e., as $j \to \infty$, for some $f \in L^1$ and integers $n_1 < n_2 < \dots$, then S is the Walsh-Fourier series of f.

7.9 Prove that given any point $x_0 \in [0,1)$ there is a non-zero Walsh series S which satisfies $S_{2^n}(x) \to 0$, as $n \to \infty$, and

$$2^{-n} S_{2^n}(x) \to 0,$$

as $n \to \infty$, for all $x \neq x_0$.

7.10 If $f \in L^1$ and S is a Walsh series which satisfies

$$\lim_{n\to\infty} S_n(x) \geq f(x)$$

and

$$\limsup_{n\to\infty} |S_{2^n}(x)| < \infty$$

for all but countably many $x \in [0,1]$, then prove there is a $g \in L^1$ with $g \geq f$ such that S is the Walsh-Fourier series of g.

7.11 Let A_0 be the collection of Walsh series whose coefficients tend to zero and let A denote the set of all functions whose Walsh-Fourier series are absolutely convergent. Prove that given $f \in A$ and $S \in A_0$ there is a Walsh series $T \in A_0$ such that T_n and fS_n are equiconvergent, as $n \to \infty$, and $T = fS$.

7.12 Let A_0 and A be defined as in Exercise 7.11, and B represent those Walsh series whose coefficients are bounded. Show that A_0, A and B are Banach spaces whose duals are related by

$$A_0' = A, \; A' = B.$$

7.13 Let A be defined as in Exercise 7.11. A set $E \subseteq [0,1]$ is called an elementary U-set if there is a sequence of functions $f_n \in A$, for $n = 1, 2, \dots$, and a constant C such that $f_n(x) = 0$ for $x \in E$,

$$\sum_{k=0}^{\infty} |\widehat{f}_n(k)| \leq C$$

for $n = 1, 2, \dots$, $\widehat{f}_n(0) \to 1$, as $n \to \infty$, and $\widehat{f}_n(k) \to 0$, as $n \to \infty$, for $k \neq 0$.

a) Prove that every elementary U-set is a U- set (see Theorem 11 in 7.5).

b) Prove that E is an elementary U-set if and only if there exists a sequence $f_n \in A$, for $n = 1, 2, \dots$, such that $f_n = 0$ on E, $n \geq 1$, and such that if $T \in A'$ then $T(f_n - 1) \to 0$ as $n \to \infty$.

7.14 a) Suppose that $E \subset [0,1]$ satisfies $|E| > 1/2$ and $S = \sum_{k=0}^{\infty} a_k r_k$ is a Rademacher series which converges to zero off E. Show that $|E \cap (E + 2^{-N})| > 0$ for all integers $N \geq 1$, and $r_k(x + 2^{-N}) = r_k(x)$ unless $N = k$.

b) Show that the zero series is the only Rademacher series which converges to zero on a set E of Lebesgue measure greater than 1/2.

c) Suppose S is a Rademacher series which converges to a constant on a set of positive measure. Show that S is a Rademacher polynomial.

[Coury [1]]

7.15 a) Let $k_1 < m_1 < k_2 < m_2 < \ldots$ be integers which satisfy $m_j - k_j \to \infty$ as $j \to \infty$. Let V_1 denote the set of points $x \in [0,1]$ whose binary expansions satisfy $x_{k_1} = x_{m_1} = 1$. For $j > 1$ let V_j denote the set of points

$$x \in [0,1] \setminus \bigcup_{\ell<j} V_\ell$$

whose binary expansions satisfy $x_{k_j} = x_{m_j} = 1$. Set

$$E := [0,1] \setminus \bigcup_{\ell=1}^{\infty} V_\ell$$

and

$$P_\ell := (1 - r_{k_\ell})(1 - r_{m_\ell}).$$

Use Theorem 11 in 7.5 to prove that E is a U-set.

b) Show there exist open intervals $I_i^{(m)}$ and $J_i^{(m)}$, $1 \leq i \leq 2^m$, $m \in \mathbf{P}$ such that

$$E = \bigcap_{m=1}^{\infty} \bigcup_{i=1}^{2^m} I_i^{(m)}$$

where for each m, the intervals $I_1^{(m)}, \ldots, I_{2^m}^{(m)}$ are formed by discarding $J_j^{(m)}$ from $I_j^{(m-1)}$, for $j = 1, \ldots, 2^{m-1}$.

c) Prove that E is "thick" in the sense that

$$\sup_{1\leq j\leq 2^{m-1}} \frac{|J_j^{(m)}|}{|I_j^{(m-1)}|} \to 0, \quad \text{as} \quad m \to \infty.$$

[Skvorcov [18]]

7.16 Show that if n_0 is any non-negative integer and E is the Cantor set formed with constant ratio of dissection 2^{-n_0}, then E is a U-set.

[Šneider [1]]

7.17 Fix $q \in \mathbf{P}$. Let $A = (a_1, a_2, \ldots, a_q)$ and $B = (b_1, b_2, \ldots, b_q)$ be vectors with integer components. Define

$$A \circ B := (a_1 \cdot b_1) \oplus (a_2 \cdot b_2) \oplus \cdots \oplus (a_q \cdot b_q)$$

and for any $x \in [0,1]$ define

$$x \cdot A := (a_1 \cdot x, a_2 \cdot x, \dots a_q \cdot x),$$

where $\oplus$ represents dyadic addition on $\mathbf{N}$ and $\cdot$ represents dyadic multiplication on $[0,\infty)$. A subset E of $[0,1]$ is said to be an $H^{(q)}$-set if there is an open, connected set $\boldsymbol{I}_0 \subset \mathbf{R}^q$ and sequence of vectors A_ℓ with integer components such that $A_\ell \circ B \to \infty$ as $\ell \to \infty$, for any vector B, and such that $x \cdot A_\ell \notin \boldsymbol{I}_0$ for any $x \in E$ and any $\ell \geq 1$. Prove that every $H^{(q)}$-set is a U- set.

(Hint: If $I_1 \times \cdots \times I_q \subset \boldsymbol{I}_0$ for open intervals I_j, and $f_j := 1$ on I_j but $f_j := 0$ off I_j, then consider the Walsh polynomials

$$P_\ell(x) := \sum_{j=1}^{q} f_j(a_j^{(\ell)} \cdot x),$$

where $A_\ell := (a_1^{(\ell)}, \dots, a_q^{(\ell)})$, and apply Theorem 11 in 7.5.)

For the next three exercises let $S = \sum_{k=0}^{\infty} a_k w_k$ be a Walsh series and let

$$L(x) := \lim_{m\to\infty} \sum_{k=0}^{2^m-1} a_k \int_0^x w_k(t)\,dt$$

be its first integral.

7.18 a) Show $L(x)$ exists for all $x \in \mathbf{Q}$.

b) Let $[\alpha_n, \beta_n) := I_n(x)$ for $x \in [0,1)$ and prove

$$L(\beta_n) - L(\alpha_n) = 2^{-n} S_{2^n}(x).$$

c) Show for $x \in [0,1)$ that $L(x)$ is finite when

$$\limsup_{n\to\infty} |S_{2^n}(x)| < \infty.$$

7.19 Suppose $L(x)$ exists for all $x \in [0,1]$, that $\limsup_{n\to\infty} S_{2^n}(x) \geq f(x)$ for all but countably many $x \in [0,1]$ and some $f \in \mathrm{L}^1$. Suppose further that

$$\limsup_{n\to\infty} 2^{-n} S_{2^n}(x) \leq 0$$

for all $x \in [0,1)$, and that

$$\limsup_{n\to\infty} 2^{-n} S_{2^n}(x-) \leq 0$$

for all $x \in (0,1]$. Prove S is a Walsh-Fourier series.

7.20 Suppose that E is a countable subset of the dyadic group $\mathbf{G}$ and let Q^* represent all elements of $\mathbf{G}$ which terminate in 0's or terminate in 1's. Suppose f is integrable and S is a Walsh series which satisfies

$$\liminf_{n\to\infty} S_{2^n}(x) \leq f(x) \leq \limsup_{n\to\infty} S_{2^n}(x) \qquad (x \in \mathbf{G} \setminus E)$$

and

$$\liminf_{n\to\infty} 2^{-n} S_{2^n}(x) \le 0 \le \limsup_{n\to\infty} S_{2^n}(x) \qquad (x \in E \cup Q^*).$$

Prove that S is the Walsh-Fourier series of f.

(Hint: Prove $L(x) = \int_0^x f(\varrho(u))\,du$, where $\varrho : [0,1] \to \mathbf{G}$ is Fine's map. Then show by a direct calculation that the coefficients a_k of S satisfy

$$a_k = \int_{\mathbf{G}} f\psi_k \, d\mu.)$$

7.21 Suppose $f \in \mathrm{L}^1$, S is a Walsh series satisfying the $C - S$ condition and that either

$$\limsup_{n\to\infty} |S_{2^n}(x)| < \infty,\ x \notin E, \quad \text{and} \quad \limsup_{n\to\infty} |S_{2^n}(x)| \le |f(x)| \quad \text{a.e.}$$

or

$$\limsup_{n\to\infty} |\sigma_{2^n}(x)| < \infty,\ x \notin E, \quad \text{and} \quad \limsup_{n\to\infty} |\sigma_{2^n}(x)| \le |f(x)| \quad \text{a.e.}$$

where E is at most countable. Prove that S is the Walsh- Fourier series of $g := \lim_{n\to\infty} S_{2^n}$.

7.22 Let S be a Haar series. Show that S_n converges a.e. on a set E if and only if σ_n converges a.e. on E as $n \to \infty$, in which case $\lim_{n\to\infty} S_n = \lim_{n\to\infty} \sigma_n$ a.e. on E.

[Šaginjan [1]]

Chapter 8

REPRESENTATION BY WALSH SERIES

8.1 Walsh Series with Monotone Coefficients. Let $\boldsymbol{a} = (a_k, k \in \mathbf{N})$ be a sequence of real numbers which decays monotonically to zero, i.e., which satisfies $a_{k-1} \geq a_k \to 0$ as $k \to \infty$. Then the function

$$f(x) := \sum_{k=0}^{\infty} a_k w_k(x)$$

is defined for all $x \in (0,1)$. In fact, it follows immediately from Abel's transformation and Theorem 10 in 1.6 that the series representing f converges uniformly on $[\delta, 1)$ for all $\delta > 0$.

In this section we identify conditions on $\boldsymbol{a}$ sufficient to conclude that f is integrable or non-negative.

Integrability is considered first.

THEOREM 1. *If $\boldsymbol{a}$ decays monotonically to zero then*

$$f(x) := \sum_{k=0}^{\infty} a_k w_k(x)$$

is continuous on $[0,1) \setminus \mathbf{Q}$, and

$$a_k = \lim_{\delta \downarrow 0} \int_{\delta}^{1} f(x) w_k(x)\, dx \qquad (k \in \mathbf{N}).$$

In particular, f has an improper Riemann integral on $[0,1)$ and thus is Perron integrable on $[0,1]$.

(For the definition of Perron integration see Zygmund [1], Vol. II, p. 84.)

PROOF. Since each Walsh function is continuous on $[0,1)\setminus\mathbf{Q}$ and the series representing f converges uniformly on compact subsets of $[0,1)$ it is clear that f is continuous on $[0,1)\setminus\mathbf{Q}$. Moreover, integrating term by term we have that

$$\int_{\delta}^{1} f(x) w_k(x)\, dx = a_k - \sum_{j=0}^{\infty} a_j \int_{0}^{\delta} w_j(x) w_k(x)\, dx$$

for $\delta > 0$ and $k \in \mathbf{N}$. Since for each fixed $k \in \mathbf{N}$ the Walsh function $w_k(x)$ equals 1 for x near zero, it is clear that the proof of this theorem will be complete when we show

$$\lim_{\delta \downarrow 0} \sum_{j=0}^{\infty} a_j \int_{0}^{\delta} w_j(x)\, dx = 0. \tag{1}$$

For each $0 < \delta < 1$ choose an integer $n := n(\delta)$ such that $2^{-n} \leq \delta < 2^{-n+1}$, and break the series (1) into two pieces :

$$I_1(\delta) := \sum_{j=0}^{2^n-1} a_j \int_0^\delta w_j(x)\,dx$$

and

$$I_2(\delta) := \sum_{j=2^n}^{\infty} a_j \int_0^\delta w_j(x)\,dx.$$

Since $\delta < 2^{-n+1}$ it is clear that

$$|I_1(\delta)| \leq 2^{-n+1} \sum_{j=0}^{2^n-1} a_j.$$

Since $\mathbf{a}$ decays monotonically to zero and $n \to \infty$ as $\delta \to 0$ it follows that $I_1(\delta) \to 0$ as $\delta \to 0$.

To estimate $I_2(\delta)$ fix $\delta > 0$ (hence $n \in \mathbf{N}$) and recall that

$$\int_0^{2^{-n}} w_j(x)\,dx = 0 \qquad (j \geq 2^n).$$

Hence $I_2(\delta)$ can be written as $A + B$ where

$$A := \sum_{m=1}^{\infty} \left(\sum_{j=m2^n}^{(2m+1)2^{n-1}-1} a_j \int_{2^{-n}}^{\delta} w_j(x)\,dx \right),$$

and

$$B := \sum_{m=1}^{\infty} \left(\sum_{j=(2m+1)2^{n-1}}^{(m+1)2^n-1} a_j \int_{2^{-n}}^{\delta} w_j(x)\,dx \right).$$

Let $m2^n \leq j < (2m+1)2^{n-1}$. Then $j = m2^n + s$ where $0 \leq s < 2^{n-1}$. Since $w_s = 1$ on $[2^{-n}, 2^{-n+1}) \supset [2^{-n}, \delta]$ it is evident that $w_j = w_{m\,2^n}$ on $[2^{-n}, \delta]$. Thus

$$A = \sum_{m=1}^{\infty} \left(\sum_{j=m2^n}^{(2m+1)2^{n-1}-1} a_j \right) \int_{2^{-n}}^{\delta} w_{m2^n}(x)\,dx.$$

A similar argument establishes

$$B = \sum_{m=1}^{\infty} \left(\sum_{j=(2m+1)2^{n-1}}^{(m+1)2^n-1} a_j \right) \int_{2^{-n}}^{\delta} (-w_{m2^n}(x))\,dx.$$

(The minus sign eventuates because $w_{2^{n-1}} = r_{n-1} = -1$ on $[2^{-n}, \delta]$.) Consequently,

$$|I_2(\delta)| = |A+B| \le \delta \sum_{m=1}^{\infty} \left| \sum_{j=m2^n}^{(2m+1)2^{n-1}-1} a_j - \sum_{j=(2m+1)2^{n-1}}^{(m+1)2^n-1} a_j \right| .$$

Since $\boldsymbol{a}$ decays monotonically to zero and $\delta 2^{n-1} \le 1$, it follows that

$$I_2(\delta) \le a_{2^n}.$$

In particular, $I_2(\delta) \to 0$ as $\delta \to 0$. ∎

The function f in Theorem 1 may not be Lebesgue integrable. Namely, there exists a sequence $\boldsymbol{a}$ which decays monotonically to zero such that

$$f(x) := \sum_{k=0}^{\infty} a_k w_k(x) \qquad (x \in (0,1))$$

is not Lebesgue integrable on $[0,1)$. To verify this, set

$$n_k := \sum_{j=0}^{(k!)^2-1} 2^{2j}$$

for $k = 1, 2, \ldots$ Let a_{n_k} be determined by

$$a_{n_k} := \sum_{j=k+1}^{\infty} \frac{1}{j!}$$

and define $(a_j, j \in \mathbf{N})$ by $a_0 := a_1 := 1$ and $a_j := a_{n_k}$ for $n_k \le j < n_{k+1}$, $k \in \mathbf{P}$. Since $(a_j, j \in \mathbf{N})$ decays monotonically to zero, the series representing f converges uniformly on $[\delta, 1)$ for $\delta > 0$. Consequently, if

$$I(\delta) := \int_\delta^1 \left| \sum_{j=0}^{\infty} a_j w_j(x) \right| dx,$$

then

$$\int_\delta^1 |f(x)|\, dx = I(\delta).$$

In particular, it suffices to show that $I(\delta) \to \infty$, as $\delta \to 0$.

Use Abel's transformation to write

$$\begin{aligned} I(\delta) &= \int_\delta^1 \left| \sum_{j=0}^{\infty} (a_j - a_{j+1}) D_{j+1}(x) \right| dx \\ &= \int_\delta^1 \left| \sum_{k=1}^{\infty} \frac{1}{k!} D_{n_k}(x) \right| dx. \end{aligned}$$

Hence for any $p \in \mathbf{P}$ we have

$$I(\delta) \geq \frac{1}{p!}\int_{\delta}^{1} |D_{n_p}| - \sum_{k=0}^{p-1} \frac{1}{k!}\int_{0}^{1} |D_{n_k}| - \sum_{k=p+1}^{\infty} \frac{1}{k!}\int_{\delta}^{1} |D_{n_k}|.$$

Set $\delta_p := 2^{-2(p!)^2} < 1/n_p$. By Theorem 9 in 1.6 we have

$$\int_{\delta_p}^{1} |D_{n_p}| \geq \frac{(p!)^2}{16}, \quad \int_{0}^{1} |D_{n_k}| \leq 2(k!)^2$$

for all p and all $k \in \mathbf{N}$. Therefore

$$\sum_{k=0}^{p-1} \frac{1}{k!}\int_{0}^{1} |D_{n_k}| \leq 2\sum_{k=0}^{p-1} k! < 4(p-1)!$$

for $p \in \mathbf{P}$. Moreover, by Theorem 10 in 1.6

$$\begin{aligned}\sum_{k=p+1}^{\infty} \frac{1}{k!}\int_{\delta_p}^{1} |D_{n_k}| &\leq \sum_{k=p+1}^{\infty} \frac{2}{k!}\int_{\delta_p}^{1} \frac{dx}{x} \\ &\leq (4\log 2)p!\left(\frac{1}{p+1} + \frac{1}{(p+1)(p+2)} + \dots\right) \\ &\leq \frac{8\log 2}{(p+1)}p!\end{aligned}$$

Hence

$$I(\delta_p) \geq p!\left(\frac{1}{16} - \frac{4}{p} - \frac{8\log 2}{(p+1)}\right).$$

In particular, $I(\delta_p) \to \infty$ as $p \to \infty$ and f is not Lebesgue integrable on $[0,1)$.

To obtain conditions on a Walsh series sufficient to conclude its limit is Lebesgue integrable we prove the following.

THEOREM 2. *Let $2 \leq p < \infty$ and $\boldsymbol{a} = (a_n, n \in \mathbf{N}) \in \ell^0$. If q is the index conjugate to p then*

$$\|\sum_{m=0}^{2^n-1} a_m D_m\|_1 \leq 2p2^{n/p}\left(\sum_{m=0}^{2^n-1} |a_m|^q\right)^{1/q}$$

for every $n \in \mathbf{N}$. In particular,

$$\|\sum_{m=0}^{2^n-1} a_m D_m\|_1 \leq 2^{n+2}\|a\|_{\ell^\infty} \qquad (n \in \mathbf{N}).$$

PROOF. Fix $n \in \mathbf{N}$ and set

$$g_n := \operatorname{sgn}\left(\sum_{m=0}^{2^n-1} a_m D_m\right).$$

Use formula iii) of Theorem 8 in 1.5 to verify

$$\|\sum_{m=0}^{2^n-1} a_m D_m\|_1 = \sum_{m=0}^{2^n-1} a_m \int_0^1 (\sum_{j=0}^{n-1} m_j r_j D_{2^j}) w_m g_n$$
$$\leq \sum_{j=0}^{n-1} \sum_{m=0}^{2^n-1} |a_m| \left| \int_0^1 r_j D_{2^j} g_n w_m \right|.$$

These integrals are Walsh-Fourier coefficients of $F_j := r_j D_{2^j} g_n$. Consequently by the Hölder and Hausdorff-Young inequalities we get

$$\|\sum_{m=0}^{2^n-1} a_m D_m\|_1 \leq \left(\sum_{m=0}^{2^n-1} |a_m|^q\right)^{1/q} \sum_{j=0}^{n-1} \left(\sum_{m=0}^{2^n-1} |\widehat{F}_j(m)|^p\right)^{1/p}$$
$$\leq \left(\sum_{m=0}^{2^n-1} |a_m|^q\right)^{1/q} \sum_{j=0}^{n-1} \|g_n D_{2^j}\|_q.$$

Since Paley's lemma implies

$$\|g_n D_{2^j}\|_q \leq \|D_{2^j}\|_q = 2^{j(q-1)/q} = 2^{j/p}$$

we conclude that

$$\|\sum_{m=0}^{2^n-1} a_m D_m\|_1 \leq \left(\sum_{m=0}^{2^n-1} |a_m|^q\right)^{1/q} \frac{2^{n/p}}{2^{1/p}-1}.$$

Thus the proof is complete since

$$\frac{1}{2^{1/p}-1} \leq 2p$$

for $p \geq 2$. ∎

Define the first difference of a sequence $\boldsymbol{a} = (a_k, k \in \mathbf{N})$ by

$$\Delta \boldsymbol{a} := (a_k - a_{k+1}, k \in \mathbf{N}).$$

Set $\Delta^0 \boldsymbol{a} := \boldsymbol{a}$ and

$$\Delta^n \boldsymbol{a} := \Delta(\Delta^{n-1} \boldsymbol{a}) \qquad (n \in \mathbf{P}, \boldsymbol{a} \in \ell^0).$$

It is easy to see for every $n, k \in \mathbf{N}$ that

$$(\Delta^n \boldsymbol{a})_k = \sum_{j=0}^{n} (-1)^j \binom{n}{j} a_{k+j}.$$

We shall frequently denote $(\Delta \boldsymbol{a})_k$ by Δa_k for $k \in \mathbf{N}$.

A sequence $\boldsymbol{a}$ is said to be *convex* if for every $k \in \mathbf{N}$

$$(\Delta^2 \boldsymbol{a})_k \geq 0.$$

It is said to be *quasi-convex* if

$$\sum_{k=1}^{\infty}(k+1)|(\Delta^2 \boldsymbol{a})_k| < \infty.$$

Quasi-convexity sometimes replaces the monotone condition for integrability of Walsh series.

It is instructive to notice that every bounded, convex sequence is both monotone non-increasing and quasi-convex. Indeed, let $\boldsymbol{a} = (a_k, k \in \mathbf{N})$ be bounded and convex, and suppose that $\boldsymbol{a}$ is not monotone non-increasing. Then $\Delta a_m < 0$ for some $m \in \mathbf{N}$. Since $\boldsymbol{a}$ is convex we have $|\Delta a_k| \geq |\Delta a_m|$ for all $k \geq m$. Hence

$$a_n - a_m = -\sum_{k=m}^{n-1} \Delta a_k \geq (n-m)|\Delta a_m|$$

for $n > m$. It follows that $a_n \to \infty$ as $n \to \infty$, a contradiction since $\boldsymbol{a}$ is bounded. Consequently, $\boldsymbol{a}$ is monotone non-increasing. To show it is also quasi-convex, let $a := \lim_{k\to\infty} a_k$. Then a is finite because $\boldsymbol{a}$ is bounded. Also, it is clear that

$$a_0 - a = \sum_{k=0}^{\infty} \Delta a_k. \tag{2}$$

Thus $\Delta \boldsymbol{a}$ decays monotonically to zero, and since

$$\sum_{k=n}^{\infty} \Delta a_k \to 0 \quad \text{as} \quad n \to \infty,$$

we also have $n\Delta a_n \to 0$ as $n \to \infty$. Consequently, Abel's transformation and (2) imply

$$a_0 - a = \sum_{k=0}^{\infty}(k+1)(\Delta^2 \boldsymbol{a})_k,$$

and $\boldsymbol{a}$ is quasi-convex.

In connection with Theorem 2 we introduce the norms

$$\begin{aligned} \|\boldsymbol{a}\|_{\ell_*^r} &:= |a_0| + \sum_{n=0}^{\infty} 2^n \left(2^{-n} \sum_{k=2^n}^{2^{n+1}-1} |a_k|^r\right)^{1/r} \\ &= |a_0| + \sum_{n=0}^{\infty} 2^{n/r'} \left(\sum_{k=2^n}^{2^{n+1}-1} |a_k|^r\right)^{1/r} \end{aligned}$$

for $\boldsymbol{a} \in \ell^0$ and $1 < r < \infty$, where r' the index conjugate to r. It is clear that the ℓ_*^r-norms increase in r. Thus we define

$$\|\boldsymbol{a}\|_{\ell_*^1} := \|\boldsymbol{a}\|_{\ell^1},$$

and

$$\|a\|_{\ell_*^\infty} := \lim_{r\to\infty} \|a\|_{\ell_*^r} = |a_0| + \sum_{n=0}^{\infty} 2^n \max_{2^n \le k < 2^{n+1}} |a_k|.$$

Consequently, if $\ell_*^r := \{a \in \ell^0 : \|a\|_{\ell_*^r} < \infty\}$ then

$$\ell_*^\infty \subseteq \ell_*^r \subseteq \ell_*^1$$

for all $1 \le r \le \infty$.

It is easy to see that if a is quasi-convex then $\Delta a \in \ell_*^\infty$. Indeed, let $2^n \le m < 2^{n+1}$ and notice by definition that

$$|\Delta a_m| \le \sum_{k=m}^{\infty} |(\Delta^2 a)_k| \le \sum_{k=2^n}^{\infty} |(\Delta^2 a)_k|.$$

Consequently by Abel's transformation we have

$$\begin{aligned}
\|\Delta a\|_{\ell_*^\infty} &\le |\Delta a_0| + \sum_{n=0}^{\infty} 2^n \sum_{k=2^n}^{\infty} |(\Delta^2 a)_k| \\
&\le |\Delta a_0| + \sum_{s=0}^{\infty} \sum_{k=2^s}^{2^{s+1}-1} |(\Delta^2 a)_k| \sum_{n=0}^{s} 2^n \\
&\le |\Delta a_0| + 2\sum_{s=0}^{\infty} \sum_{k=2^s}^{2^{s+1}} k\,|(\Delta^2 a)_k| \\
&= |\Delta a_0| + 2\sum_{k=0}^{\infty} k\,|(\Delta^2 a)_k| < \infty.
\end{aligned}$$

THEOREM 3. *Let $a \in \ell^0$ and suppose that $a_n \to 0$ as $n \to \infty$. If*

$$\Delta a \in \ell_*^2$$

then the series

$$\sum_{k=0}^{\infty} a_k w_k$$

converges pointwise on $(0,1)$, its sum f belongs to L^1 and satisfies

$$\|f\|_1 \le 8\|\Delta a\|_{\ell_*^2}.$$

In particular, if a is quasi-convex then $f \in L^1$.

PROOF. Since

$$\|\Delta a\|_{\ell^1} = \|\Delta a\|_{\ell_*^1} \le \|\Delta a\|_{\ell_*^2} < \infty$$

and $|D_n(x)| \le 2/x$ for all $x \in (0,1)$ and $n \in \mathbf{N}$, it is easy to see by Abel's transformation that the series converges pointwise on $(0,1)$.

For each $n, m \in \mathbf{N}$ combine Abel's transformation and Theorem 2 to obtain

$$\begin{aligned}\| \sum_{k=2^n}^{2^m-1} a_k w_k \|_1 &= \| \sum_{k=2^n}^{2^m-1} \Delta a_k D_{k+1} + a_{2^m} D_{2^m} - a_{2^n} D_{2^n} \|_1 \\ &\le \sum_{s=n}^{\infty} \| \sum_{k=2^s}^{2^{s+1}-1} \Delta a_k D_{k+1} \|_1 + |a_{2^m}| + |a_{2^n}| \\ &\le 8 \sum_{s=n}^{\infty} 2^{s/2} \left(\sum_{k=2^s}^{2^{s+1}-1} |\Delta a_k|^2 \right)^{1/2} + |a_{2^n}| + |a_{2^m}|.\end{aligned}$$

It follows that 2^n-th partial sums of the series in question converge in L^1-norm and consequently the limit function f belongs to L^1. Moreover, substituting $n = 0$ and letting $m \to \infty$ in the inequality above we obtain $\|f\|_1 \le 8\|\boldsymbol{a}\|_{\ell^2_*}$ as required. ∎

Not every Lebesgue integrable Walsh series has coefficients which are quasi-convex. Indeed, let $\boldsymbol{a}$ decay monotonically to zero and suppose that

$$a_k = a_{2^n} \qquad (2^n \le k < 2^{n+1}, n \in \mathbf{N}).$$

By Abel's transformation

$$f(x) := \sum_{k=0}^{\infty} a_k w_k(x) = \sum_{n=0}^{\infty} (a_{2^n-1} - a_{2^n}) D_{2^n}(x).$$

Since $D_{2^n} \ge 0$ for $n \in \mathbf{N}$, it follows from the monotone convergence theorem that

$$\|f\|_1 = \sum_{n=0}^{\infty} (a_{2^n-1} - a_{2^n}) \, \|D_{2^n}\|_1 = a_0.$$

In particular, if such a sequence $\boldsymbol{a}$ decays monotonically to zero then f is Lebesgue integrable no matter how slowly $a_k \to 0$.

If the coefficients of a Walsh series decay monotonically to zero sufficiently rapidly, then the limit function is always Lebesgue integrable. In fact,

THEOREM 4. *If $(a_k, k \in \mathbf{N})$ decays monotonically to zero and if*

$$\sum_{k=1}^{\infty} \frac{a_k}{k} < \infty,$$

then $f := \sum_{k=0}^{\infty} a_k w_k$ converges in L^1 norm.

PROOF. Notice for monotone non-increasing sequences $(a_k, k \in \mathbf{N})$ that $\sum a_k/k < \infty$ is equivalent to

$$\sum_{k=1}^{\infty} \Delta a_k \log(k+1) < \infty.$$

By Abel's transformation,

$$\sum_{k=n}^{\infty} a_k w_k = \sum_{k=n}^{\infty} \Delta a_k D_{k+1} - a_n D_n$$

for any $n \in \mathbf{N}$. Since

$$\|D_k\|_1 = O(\log(k+1))$$

as $k \to \infty$ we see that

$$\begin{aligned} \|\sum_{k=0}^{n-1} a_k w_k - f\|_1 &= \|\sum_{k=n}^{\infty} a_k w_k\|_1 \\ &\leq a_n \|D_n\|_1 + \sum_{k=n}^{\infty} \Delta a_k \|D_{k+1}\|_1 \\ &\leq C_1 a_n \log(n+1) + C_1 \sum_{k=n}^{\infty} \Delta a_k \log(k+1) \\ &\leq C_2 \sum_{k=n}^{\infty} \Delta a_k \log(k+1) \end{aligned}$$

where C_1 and C_2 are absolute constants independent of n. It follows, therefore, from hypothesis that $f \in \mathrm{L}^1$ and the series representing f converges in L^1 norm. ∎

Theorem 4 holds for Walsh-Kaczmarz series since the Walsh-Kaczmarz- Dirichlet kernel also satisfies

$$\|D_k\|_1 = O(\log(k+1))$$

as $k \to \infty$ (see Exercise 8.6).

For the trigonometric case, convexity of the coefficients is enough to ensure both integrability and non-negativity of the limit function. This is not the case for Walsh series (see Theorems 5 and 6 below).

LEMMA 1. *For $x \in [0,1)$ and $0 \leq t < 1$ let*

$$P(x,t) := \sum_{k=0}^{\infty} t^k w_k(x).$$

Then $P(x,t)$ is positive and can be written in the form

$$P(x,t) = \prod_{k=0}^{\infty} \left(1 + t^{2^k} r_k(x)\right).$$

(Note: This is the Poisson kernel for the Walsh system.)

PROOF. For each integer $n \geq 0$ set

$$P_{2^n}(x,t) := \sum_{k=0}^{2^n-1} t^k w_k(x).$$

Then $P_{2^n}(x,t) \to P(x,t)$ as $n \to \infty$. Moreover, since $r_n w_k = w_{2^n+k}$ for $0 \leq k < 2^n$ and $n \in \mathbf{N}$, an easy inductive argument establishes that

$$P_{2^n}(x,t) = \prod_{k=0}^{n-1} \left(1 + t^{2^k} r_k(x)\right).$$

In particular, $P_{2^n}(x,t) \geq 0$ for $n \in \mathbf{N}$. Since $P(x,t)$ cannot be zero, the lemma is established by letting $n \to \infty$. ∎

This computation and a solution to the moment problem can be used to obtain sufficient conditions for Walsh series to have non- negative limits.

A sequence $\boldsymbol{a}$ is called completely monotone if $\Delta^n a_k \geq 0$ for all non-negative integers n and k. Thus every completely monotone sequence is convex.

Completely monotone sequences have a non-decreasing solution to the moment problem (see Theorem 11 in Appendix 0.5). Thus given a completely monotone sequence $\boldsymbol{a}$ there is a non-decreasing function g of bounded variation such that

$$a_k = \int_0^1 t^k \, dg(t) \qquad (k \in \mathbf{N}). \tag{3}$$

THEOREM 5. *If $\boldsymbol{a}$ is completely monotone then*

$$f := \sum_{k=0}^{\infty} a_k w_k$$

is non-negative and Lebesgue integrable on $[0,1)$.

PROOF. By Theorem 3, $f \in \mathbf{L}^1$.

To show that f is non-negative, choose a non-decreasing function g which satisfies (3). Using the notation introduced in the proof of Lemma 1, we have that

$$\sum_{k=0}^{2^n-1} a_k w_k(x) = \int_0^1 P_{2^n}(x,t) \, dg(t) \tag{4}$$

for all $x \in [0,1)$ and $n \in \mathbf{N}$.

We shall presently show that (4) implies

$$f(x) = \int_0^1 P(x,t) \, dg(t). \tag{5}$$

Since g is non-decreasing, and by Lemma 1 $P(x,t)$ is positive, it will follow that f is non-negative as required.

To verify (5) consider first the case when x is a dyadic rational. The factors $(1+t^{2^k}r_k(x))$ of $P(x,t)$ are thus equal to $(1+t^{2^k})$ for large k. Hence $P_{2^n}(x,t)$ is eventually monotone increasing in n for all $t \in [0,1)$. Consequently, (5) follows from (4) by means of the monotone convergence theorem.

Now suppose that x is a dyadic irrational. Let p be the least positive integer for which the p-th binary coefficient of x is 1 and let u be the left endpoint of $I_p(x)$. Then $P_{2^n}(x,t)$ is dominated by

$$P(u,t) = \prod_{k=0}^{p-1}\left(1+t^{2^k}r_k(x)\right)\prod_{k=p}^{\infty}(1+t^{2^k})$$

for all $0 \le t < 1$ and all $n > p$. Moreover, $P(u,t)$ is integrable with respect to $dg(t)$, since (5) holds for the dyadic rational $x = u$. Therefore, the Lebesgue dominated convergence theorem implies

$$\lim_{n\to\infty}\int_0^1 P_{2^n}(x,t)\,dg(t) = \int_0^1 P(x,t)\,dg(t).$$

In particular, (5) follows from (4) and the proof is complete. ∎

We close this section by showing that not all Walsh series with convex coefficients have non-negative limits.

THEOREM 6. *There is a convex sequence* $\boldsymbol{a}$ *which decays monotonically to zero for which the function*

$$f := \sum_{k=0}^{\infty} a_k w_k$$

assumes negative values.

PROOF. Let

$$g(t) := -\int_0^t r_0(u)\,du \qquad (0 \le t \le 1)$$

and observe that g is of bounded variation. Set

$$a_k := \int_0^1 t^{k+2}\,dg(t) = -\int_0^1 r_0(t)t^{k+2}\,dt \qquad (k \in \mathbf{N}),$$

and verify by direct calculation that

$$\Delta^2 a_k = -\int_0^1 r_0(t)t^2(1-t)^2t^k\,dt > 0 \qquad (k \in \mathbf{N}).$$

In particular, $\boldsymbol{a}$ is convex and decays monotonically to zero.

Next, observe that

$$\sum_{k=0}^{2^n-1} a_k w_k(x) = \int_0^1 t^2 P_{2^n}(x,t)\,dg(t)$$

for $x \in [0,1)$ and $n \in \mathbf{N}$. By repeating the argument used to establish Theorem 5, one can show that

$$f(x) = \int_0^1 t^2 P(x,t)\,dg(t)$$

for $x \in [0,1)$. Consequently, the choice of g implies

$$f(x) = \int_{1/2}^1 t^2 P(x,t)\,dt - \int_0^{1/2} t^2 P(x,t)\,dt.$$

In particular, $f(1-)$ will be negative if we can show that

$$\int_0^{1/2} t^2 P(1-,t)\,dt > \int_{1/2}^1 t^2 P(1-,t)\,dt,$$

or equivalently, that

$$(6) \qquad 2\int_0^{1/2} t^2 P(1-,t)\,dt > \int_0^1 t^2 P(1-,t)\,dt.$$

To estimate the left side of (6) observe for $t \le 1/2$ that

$$\begin{aligned} P(1-,t) &= 1 - t - t^2 + t^3 - \dots \\ &\ge 1 - t - t^2 + t^3 - (t^4 + t^5 + \dots) \\ &\ge 1 - t - t^2. \end{aligned}$$

Thus, Lemma 1 implies

$$\begin{aligned} P(1-,t) &= (1-t)P(1-,t^2) \\ &\ge (1-t)(1-t^2-t^4) \\ &= 1 - t - t^2 + t^3 - t^4 + t^5 \end{aligned}$$

for $t \le 1/2$. Hence

$$2\int_0^{1/2} t^2 P(1-,t)\,dt \ge 2\int_0^{\frac12} (t^2 - t^3 - t^4 + t^5 - t^6 + t^7)\,dt > .0435.$$

On the other hand,

$$\begin{aligned} P(1-,t) &\le (1-t)(1-t^2)(1-t^4) \\ &= 1 - t - t^2 + t^3 - t^4 + t^5 + t^6 - t^7. \end{aligned}$$

Therefore,

$$\int_0^1 t^2 P(1-,t)\,dt \le \frac13 - \frac14 - \frac15 + \frac16 - \frac17 + \frac18 + \frac19 - \frac1{10} < .0433. \qquad \blacksquare$$

8.2 Term by Term Dyadic Differentiation. Let x be a point in [0,1). A Walsh series $\sum_{k=0}^{\infty} a_k w_k$ is said to be *term by term dyadically differentiable* at x if

$$f(t) := \sum_{k=0}^{\infty} a_k w_k(t)$$

exists and is finite for $t = x, x \dot{+} 2^{-\ell}$ for $\ell \in \mathbf{P}$, and if the dyadic derivative of f exists at x and satisfies

$$f^{[1]}(x) = \sum_{k=1}^{\infty} k a_k w_k(x)$$

(see 1.7).

We shall identify several conditions sufficient for term by term dyadic differentiability. The first one uses a result from the preceding section.

THEOREM 7. *If* $(a_k, k \in \mathbf{N})$ *and* $(ka_k, k \in \mathbf{N})$ *are quasi-convex and* $ka_k \to 0$ *as* $k \to \infty$, *then*

$$f := \sum_{k=0}^{\infty} a_k w_k$$

is a.e. term by term dyadically differentiable.

PROOF. By Theorem 3 in 8.1 both

$$S := \sum_{k=0}^{\infty} a_k w_k \quad \text{and} \quad S^{[1]} := \sum_{k=0}^{\infty} k a_k w_k$$

are Walsh-Fourier series which converge to integrable functions f and g everywhere on $(0,1)$. It follows, therefore, from Corollary 1 in 5.2 that f is strongly dyadically differentiable with

$$\mathbf{d}^{[1]} f = g \quad \text{a.e.}$$

But $f = \mathbf{I}(\mathbf{d}^{[1]} f)$ a.e. Therefore, the fundamental theorem of dyadic calculus (Corollary 7 in 6.2) implies that f is a.e. dyadically differentiable with $f^{[1]} = \mathbf{d}^{[1]} f$ a.e. ∎

When $(ka_k, k \in \mathbf{N})$ decays monotonically to zero, the Walsh series $\sum_{k=0}^{\infty} a_k w_k(x)$ is term by term dyadically differentiable for all $x \in (0,1)$, $x \neq 2^{-j}, j = 1, 2, \ldots$. In fact,

THEOREM 8. *Let* $x \in (0,1)$, $x \neq 2^{-j}$ *for* $j = 1, 2, \ldots$, *let* $a_k \in \mathbf{R}$, $k \in \mathbf{N}$, *and define a sequence of numbers* $(R_n, n \in \mathbf{N})$ *by* $R_0 := 0$ *and*

$$R_n := \sum_{k=2^n}^{2^{n+1}-2} |a_k - a_{k+1}| \qquad (n \in \mathbf{P}).$$

If $a_k \to 0$ *as* $k \to \infty$, *and* $2^n R_n \to 0$ *as* $n \to \infty$, *then a necessary and sufficient condition that* $f := \sum_{k=0}^{\infty} a_k w_k$ *be dyadically differentiable at* x *is that*

$$S_{2^n}^{[1]}(x) := \sum_{k=0}^{2^n-1} k a_k w_k(x)$$

converge to a finite real number $g(x)$, as $n \to \infty$, in which case

$$f^{[1]}(x) = g(x).$$

PROOF. Use Abel's transformation to write

$$\sum_{k=2^m}^{2^n-1} a_k w_k = \sum_{k=2^m}^{2^n-2} (a_k - a_{k+1}) D_{k+1} + a_{2^n-1} D_{2^n} - a_{2^m} D_{2^m}$$

for $m, n \in \mathbf{N}$. Since $2^{-v} \le x < 1$ implies that $D_{2^v}(x) = 0$ for any integer v, it follows immediately from hypothesis that the 2^n-th partial sums of the series representing f converge on $(0,1)$ as $n \to \infty$.

In fact, the n-th partial sums converge. To see this use Abel's transformation for $2^m \le n < 2^{m+1}$ to write

$$\sum_{k=2^m}^{n} a_k w_k = \sum_{k=2^m}^{n-1} (a_k - a_{k+1}) D_{k+1} + a_n D_{n+1} - a_{2^m} D_{2^m}.$$

Since $|D_{k+1}(x)| \le 2/x$ for $k \in \mathbf{N}$ and $x \in (0,1)$ it follows that

$$\left|\sum_{k=2^m}^{n} a_k w_k(x)\right| \le \frac{2}{x}\left(R_m + 2 \max_{2^m \le k < 2^{m+1}} |a_k|\right),$$

for $x \in (0,1)$. Since this last expression converges to zero as $m \to \infty$, the series representing f must converge on $(0,1)$.

Decompose f on (0,1) by

$$f = \sum_{k=0}^{2^n-1} a_k w_k + g_n$$

where

(7) $$g_n = \sum_{k=2^n}^{\infty} a_k w_k \qquad (n \in \mathbf{N}).$$

Recall that

$$\mathbf{d}_n f(x) = \sum_{\ell=0}^{n-1} 2^{\ell-1}(f(x) - f(x \dotplus 2^{-\ell-1}))$$

and $\mathbf{d}_n(w_k) = k w_k$ for $0 \le k < 2^n$ (see (69) in 1.7). Thus

$$\mathbf{d}_n f(x) = S_{2^n}^{[1]}(x) + (\mathbf{d}_n g_n)(x)$$

for $x \in (0,1)$ and $n \in \mathbf{N}$. In particular, the proof of the theorem will be complete if we show that $(\mathbf{d}_n g_n)(x) \to 0$, as $n \to \infty$, for any $x \in (0,1)$ which is not the reciprocal of an integer power of 2.

Fix $x \in (0,1)$ with $x \neq 2^{-j}$, for $j = 1, 2, \ldots$ Choose $p = p(x) \in \mathbf{N}$ so large that neither x nor $x \dot{+} 2^{-\ell}$ belong to the interval $[0, 2^{-p})$ for $\ell = 1, 2, \ldots$ Since $m \geq p$ implies $D_{2^m}(t) = 0$ for $2^{-p} \leq t < 1$, it follows from the definition of $\mathbf{d}_n$ that

$$(\mathbf{d}_n D_{2^m})(x) = 0$$

for $m \geq p$ and $n \in \mathbf{P}$. Hence applying $\mathbf{d}_n$ to (7) results in

$$(8) \qquad (\mathbf{d}_n g_n)(x) = \sum_{m=n}^{\infty} \sum_{k=2^m}^{2^{m+1}-2} (a_k - a_{k+1})(\mathbf{d}_n D_{k+1})(x)$$

provided $n \geq p$.

To estimate

$$(9) \qquad (\mathbf{d}_n D_k)(x) = \sum_{\ell=0}^{n-1} 2^{\ell-1}(D_k(x) - D_k(x \dot{+} 2^{-\ell-1})),$$

recall that $|D_k(t)| \leq 2/t$ for any $t \neq 0$. Since x, $x \dot{+} 2^{-\ell} \geq 2^{-p}$ for $\ell \in \mathbf{P}$, it follows that $|D_k(x)|$ and $|D_k(x \dot{+} 2^{-\ell-1})|$ are dominated by the constant $C(x) := 2^{p(x)+1} < \infty$. This constant is evidently independent of both k and ℓ. Hence by (9) we have

$$\sup_{k \in \mathbf{P}} |(\mathbf{d}_n D_k)(x)| \leq 2C(x) \sum_{\ell=0}^{n-1} 2^{\ell-1}.$$

This estimate and (8) imply

$$|(\mathbf{d}_n g_n)(x)| \leq 2^n C(x) \sum_{m=n}^{\infty} R_m$$

for n sufficiently large. Since by hypothesis

$$2^n \sum_{m=n}^{\infty} R_m \to 0, \quad \text{as} \quad n \to \infty,$$

we conclude that $(\mathbf{d}_n g_n)(x) \to 0$ as $n \to \infty$. ∎

COROLLARY 1. *A Walsh series $\sum_{k=0}^{\infty} a_k w_k$ is term by term dyadically differentiable at a point $x \in (0,1)$, $x \neq 2^{-j}$ for $j = 1, 2, \ldots$, if any one of the following three conditions is satisfied:*

i) $2^n \sum_{k=2^n}^{\infty} |a_k - a_{k+1}| \to 0$, *as* $n \to \infty$, *and* $\sum_{k=1}^{\infty} k a_k w_k(x)$ *converges;*

ii) $a_{2^n} \geq a_{2^n+1} \geq \cdots \geq a_{2^{n+1}-1}$ *for n sufficiently large,*

$$\sum_{k=1}^{\infty} k a_k w_k(x) \qquad \text{converges}$$

and

$$2^n \sum_{k=n}^{\infty}(a_{2^k} - a_{2^{k+1}-1}) \to 0,$$

as $n \to \infty$;

iii) $(ka_k, k \in \mathbf{N})$ *decays monotonically to zero.*

Theorem 9 below shows that if the derived series $\sum ka_k w_k$ converges fast enough, then $\sum a_k w_k$ is term by term dyadically differentiable on all of $[0,1)$. Our proof rests on interchanging a limit and an infinite sum by using the following method.

LEMMA 2. *Suppose* $(\omega_m, m \in \mathbf{N})$ *is a non-decreasing sequence of positive real numbers and* $\sum_{k=0}^{\infty} x_k$ *is a convergent series of real numbers such that*

$$\sum_{k=m}^{\infty} x_k = o\left(\frac{1}{\omega_m}\right), \quad \text{as} \quad m \to \infty. \tag{10}$$

Suppose further that $(b_k, k \in \mathbf{P})$, $(b_k^{(n)}, k \in \mathbf{P})$ *are sequences of real numbers such that*

$$\lim_{n\to\infty} b_k^{(n)} = b_k \qquad (k \in \mathbf{P}). \tag{11}$$

If there is an absolute constant $M > 0$ *such that*

$$\sum_{k=1}^{\infty}\left(\frac{1}{\omega_k}\right)\left|b_k^{(n)} - b_{k+1}^{(n)}\right| \le M \qquad (n \in \mathbf{P}), \tag{12}$$

then the series $\sum_{k=0}^{\infty} b_k x_k$ *and* $\sum_{k=0}^{\infty} b_k^{(n)} x_k$ *converge to finite real numbers for* $n \in \mathbf{P}$, *and*

$$\lim_{n\to\infty}\sum_{k=1}^{\infty} b_k^{(n)} x_k = \sum_{k=1}^{\infty} b_k x_k. \tag{13}$$

PROOF. Fix $m \in \mathbf{P}$. Notice by (12) that

$$\sum_{k=1}^{m}\left(\frac{1}{\omega_k}\right)|b_k - b_{k+1}| \le \left(\frac{|b_1 - b_1^{(n)}|}{\omega_1}\right) + 2\sum_{k=1}^{m}\left(\frac{|b_{k+1} - b_{k+1}^{(n)}|}{\omega_k}\right) + M.$$

It follows from (11) that

$$\sum_{k=1}^{m}\frac{|b_k - b_{k+1}|}{\omega_k} \le M \qquad (m \in \mathbf{P}). \tag{14}$$

Let $k \ge 2$. Write

$$b_1 - b_k = \sum_{j=1}^{k-1}(b_j - b_{j+1}).$$

Since $(\omega_k, k \in \mathbf{N})$ is non-decreasing we have by (14) that

$$|b_1 - b_k| \le \omega_k \sum_{j=1}^{k-1} \frac{|b_j - b_{j+1}|}{\omega_j} \le \omega_k M.$$

Hence there is an absolute constant $L \ge 1$ such that

$$|b_k| \le \omega_k L < \infty \qquad (k \in \mathbf{P}). \tag{15}$$

Let $\varepsilon > 0$. By (10) choose N so large that

$$\omega_k L |\sum_{j=k}^{\infty} x_j| < \varepsilon$$

for $k \ge N$. Let $m > n \ge N$ be positive integers. Write

$$x_k = \sum_{j=k}^{\infty} x_j - \sum_{j=k+1}^{\infty} x_j.$$

By Abel's transformation we have

$$\sum_{k=n}^{m} b_k x_k = b_n \sum_{j=n}^{\infty} x_j + \sum_{k=n+1}^{m} (b_k - b_{k-1}) \sum_{j=k}^{\infty} x_j - b_m \sum_{j=m+1}^{\infty} x_j.$$

Hence it follows from (15) that

$$\left|\sum_{k=n}^{m} b_k x_k\right| \le \omega_n L \left|\sum_{j=n}^{\infty} x_j\right| + \sum_{k=n+1}^{\infty} \left(\frac{|b_k - b_{k-1}|}{\omega_k}\right) \omega_k \left|\sum_{j=k}^{\infty} x_j\right|$$
$$+ \omega_{m+1} L \left|\sum_{j=m+1}^{\infty} x_j\right|.$$

It follows from (14) and the choice of N that

$$|\sum_{k=n}^{m} b_k x_k| \le (2 + M)\varepsilon \qquad (n \ge N).$$

The partial sums of $\sum_{k=0}^{\infty} b_k x_k$ are evidently Cauchy so the series $\sum_{k=0}^{\infty} b_k x_k$ converges. A similar argument establishes that $\sum_{k=0}^{\infty} b_k^{(n)} x_k$ also converges for each $n \in \mathbf{P}$.

Let $\varepsilon > 0$. Since (11) and (10) hold and both $\sum_{k=0}^{\infty} x_k$, $\sum_{k=0}^{\infty} b_k x_k$ converge, choose an integer N so large that

$$|b_N^{(n)}| \le 2|b_N|, \tag{16}$$

(17)
$$\left|\sum_{j=N}^{\infty} b_j x_j\right| < \varepsilon,$$

(18)
$$\left|\sum_{j=N}^{m} x_j\right| < \varepsilon,$$

and

(19)
$$\omega_m \left|\sum_{j=m}^{\infty} x_j\right| < \varepsilon$$

for $m, n \geq N$.

Fix $n > N$. Since $\sum_{k=1}^{\infty} b_k^{(n)} x_k$ converges choose $N(n) > N$ such that

(20)
$$\left|\sum_{k=N(n)}^{\infty} b_k^{(n)} x_k\right| < \varepsilon.$$

As above, use Abel's transformation to write

$$\sum_{k=N}^{N(n)-1} b_k^{(n)} x_k = b_N^{(n)} \sum_{j=N}^{\infty} x_j + \sum_{k=N+1}^{N(n)-1} \left(b_k^{(n)} - b_{k-1}^{(n)}\right) \sum_{j=k}^{\infty} x_j$$
$$- b_{N(n)-1}^{(n)} \sum_{j=N(n)}^{\infty} x_j.$$

Observe by (16) that

$$\left|\sum_{k=N}^{N(n)-1} b_k^{(n)} x_k\right| \leq 2|b_N| \left|\sum_{j=N}^{\infty} x_j\right| + \sum_{k=N+1}^{N(n)-1} \left(\frac{\left|b_k^{(n)} - b_{k-1}^{(n)}\right|}{\omega_k}\right) \omega_k \left|\sum_{j=k}^{\infty} x_j\right|$$
$$+ 2|b_{N(n)-1}| \left|\sum_{j=N(n)}^{\infty} x_j\right|.$$

Thus it follows from (12), (18), (14), and (19) that there is a constant $C > 0$ such that

(21)
$$\left|\sum_{k=N}^{N(n)-1} b_k^{(n)} x_k\right| \leq C\varepsilon.$$

To verify (13) notice that

$$\left|\sum_{k=1}^{\infty} b_k^{(n)} x_k - \sum_{k=1}^{\infty} b_k x_k\right| \le \sum_{k=1}^{N-1} \left|\left(b_k^{(n)} - b_k\right) x_k\right| + \left|\sum_{k=N}^{\infty} b_k x_k\right| + \left|\sum_{k=N}^{N(n)-1} b_k^{(n)} x_k\right| + \left|\sum_{k=N(n)}^{\infty} b_k^{(n)} x_k\right|.$$

Consequently we have by (17), (21), and (20) that

$$\left|\sum_{k=1}^{\infty} b_k^{(n)} x_k - \sum_{k=1}^{\infty} b_k x_k\right| \le \sum_{k=1}^{N-1} \left|\left(b_k^{(n)} - b_k\right) x_k\right| + (2+C)\varepsilon$$

for any $n > N$. Let $n \to \infty$. Since N and C are fixed and $\varepsilon > 0$ was arbitrary we conclude that (13) holds. ▮

THEOREM 9. *Suppose $\gamma_0 \ge \gamma_1 \cdots \ge \gamma_k > 0$ for $k \in \mathbf{N}$.*

a) Let $x \in [0,1)$ and suppose that the series

$$f(t) := \sum_{k=0}^{\infty} a_k w_k(t)$$

converges at $t := x \dot{+} 2^{-j-1}$ for $j \in \mathbf{N}$. If

$$\sum_{k=0}^{\infty} \gamma_k < \infty \quad \text{and} \quad \sum_{k=m}^{\infty} a_k w_k(x) = o(\gamma_m)$$

as $m \to \infty$, or if

$$\sum_{k=1}^{\infty} \gamma_k / k < \infty \quad \text{and} \quad \sum_{k=m}^{\infty} k a_k w_k(x) = o(\gamma_m) \tag{22}$$

as $m \to \infty$, then f is term by term dyadically differentiable at x.

b) If

$$\sum_{k=0}^{\infty} \gamma_k < \infty \quad \text{and} \quad \sum_{k=0}^{\infty} \left(\frac{a_k}{\gamma_k}\right) w_k(x_0) < \infty$$

for some $x_0 \in [0,1)$ then $f := \sum_{k=0}^{\infty} a_k w_k$ converges absolutely on $[0,1)$ and is term by term dyadically differentiable at x_0.

PROOF. We begin with an opening remark. Let $x_k \in \mathbf{R}$ for $k \in \mathbf{N}$. If ω_k is a non-decreasing sequence of positive numbers and

$$\sum_{k=0}^{\infty} \omega_k x_k$$

converges to a real number, then

$$\sum_{k=m}^{\infty} x_k = o(1/\omega_m) \tag{23}$$

as $m \to \infty$ and $\sum_{k=0}^{\infty} x_k$ also converges. Indeed, set

$$y_m := \sum_{k=m}^{\infty} \omega_k x_k, \qquad (m \in \mathbf{N})$$

and observe by Abel's transformation that

$$\begin{aligned} \left|\sum_{k=m}^{\infty} x_k\right| &= \left|\sum_{k=m}^{\infty} \frac{y_k - y_{k+1}}{\omega_k}\right| \\ &\leq \frac{|y_m|}{\omega_m} + \sum_{k=m+1}^{\infty} |y_k| \left(\frac{1}{\omega_{k-1}} - \frac{1}{\omega_k}\right) \\ &\leq \frac{2}{\omega_m} \sup_{k \geq m} |y_k|. \end{aligned}$$

To prove a), suppose first that (22) holds. Apply the remark to $\omega_k := k$ and $x_k := a_k w_k(x)$. Thus by hypothesis f converges at x as well as at $x \dot{+} 2^{-j-1}$ for $j \in \mathbf{N}$. Hence $\mathbf{d}_n f(x)$ exists. For each $j \in \mathbf{N}$ and $n \in \mathbf{P}$ let

$$\langle k \rangle_n := \sum_{j=0}^{n-1} k_j 2^j,$$

where $(k_j, j \in \mathbf{N})$ represents the binary coefficients of k. By (69) in 1.7 we have

$$\begin{aligned} \mathbf{d}_n f(x) &= \sum_{k=0}^{\infty} \langle k \rangle_n \, a_k w_k(x) \\ &= \sum_{k=1}^{\infty} \frac{\langle k \rangle_n}{k} k a_k w_k(x). \end{aligned}$$

Moreover, it is clear that $\langle k \rangle_n / k \to 1$ as $n \to \infty$. We shall apply Lemma 2 to $\omega_k := 1/\gamma_k$ and

$$b_k^{(n)} := \frac{\langle k \rangle_n}{k}.$$

Thus it suffices to show that there exists a constant M, independent of n, such that

$$A_n := \sum_{k=1}^{\infty} \gamma_k \left| b_k^{(n)} - b_{k+1}^{(n)} \right| \leq 2M$$

for $n \in \mathbf{P}$.

Fix $n \in \mathbf{P}$. Since $\langle k \rangle_n = k$ for $0 \le k < 2^n$ we have

$$A_n = \sum_{k=2^n}^{\infty} \gamma_k \left| b_k^{(n)} - b_{k+1}^{(n)} \right| .$$

Thus A_n can be broken into two pieces, B_n which is the sum over indices k which belong to the set

$$\mathbf{N}_1 := \{k \ge 2^n : k = p2^n - 1 \quad \text{for some} \quad p \ge 2\}$$

and C_n which is the sum over indices k which belong to the set

$$\mathbf{N}_2 := \{k \ge 2^n : k \notin \mathbf{N}_1\}.$$

To estimate B_n notice that $\langle p2^n \rangle_n = 0$ and $\langle p2^n - 1\rangle_n = 2^n - 1$. Since $(\gamma_k, k \in \mathbf{N})$ is non-increasing we have

$$\begin{aligned} B_n &= \sum_{p=2}^{\infty} \gamma_{p2^n-1} \left| \frac{2^n - 1}{p2^n - 1} \right| \\ &\le \sum_{p=2}^{\infty} \left(\sum_{k=(p-1)2^n}^{p2^n-1} \gamma_k \frac{1}{(p2^n - 1)} \right) \\ &\le \sum_{k=1}^{\infty} \frac{\gamma_k}{k}. \end{aligned}$$

By hypothesis $M := \sum_{k=1}^{\infty} \gamma_k / k < \infty$. Hence

$$B_n \le M \qquad (n \in \mathbf{P}).$$

To estimate C_n notice that

$$\langle k+1 \rangle_n = \langle k \rangle_n + 1$$

for all $k \in \mathbf{N}_2$. Hence

$$\left| \frac{\langle k \rangle_n}{k} - \frac{\langle k+1 \rangle_n}{k+1} \right| = \left| \frac{k - \langle k \rangle_n}{k(k+1)} \right| \le \frac{k}{k(k+1)} \le \frac{1}{k}.$$

Consequently, $C_n \le \sum_{k=1}^{\infty} \gamma_k / k$ and again we obtain

$$C_n \le M \qquad (n \in \mathbf{P}).$$

We conclude that

$$A_n = B_n + C_n \le 2M$$

for $n \in \mathbf{P}$.

A similar argument establishes the rest of a).

To prove b), apply the opening remark to $\omega_k := 1/\gamma_k$ and $x_k := a_k w_k(x_0)$. Thus

$$\sum_{k=m}^{\infty} a_k w_k(x_0) = o(\gamma_m)$$

as $m \to \infty$. Moreover, since $a_k/\gamma_k \to 0$ as $k \to \infty$, it is clear by hypothesis that f converges absolutely on $[0, 1)$. Hence by a), f is term by term dyadically differentiable at x_0. ∎

COROLLARY 2. *A Walsh series* $\sum_{k=0}^{\infty} a_k w_k$ *is term by term dyadically differentiable on* $(0,1)$ *if any one of the following conditions is satisfied:*

i) for every $x \in (0,1)$, *there is an* $\varepsilon = \varepsilon(x) > 0$ *such that*

$$\sum_{k=m}^{\infty} a_k w_k(x) = o(m^{-1-\varepsilon})$$

as $m \to \infty$;

ii) for every $x \in (0,1)$, *there is an* $\varepsilon = \varepsilon(x) > 0$ *such that*

$$\sum_{k=1}^{\infty} k^{1+\varepsilon} a_k w_k(x)$$

converges;

iii) for some $\varepsilon > 0$, *the sequence* $(k^{1+\varepsilon} a_k, k \in \mathbf{N})$ *decays monotonically to zero.*

By Theorem 12 in 1.7 any classically continuous function whose Walsh-Fourier coefficients satisfy the hypotheses of Theorem 8 is necessarily constant. In the case of Theorem 9 a stronger statement can be made.

THEOREM 10. *For every pair of integers* $0 \le p < m$ *let*

$$A(p,m) := \{k \in \mathbf{N} : 2^m + (2\ell - 1)2^p \le k < 2^m + 2\ell 2^p, 1 \le \ell \le 2^{m-p-1}\},$$

and set $A(p,p) := \{k \in \mathbf{N} : 2^p \le k < 2^{p+1}\}$. *If* f *is classically continuous on some interval* $[a,b] \subseteq [0,1)$ *and*

$$\lim_{p\to\infty} 2^p \sum_{m=p}^{\infty} \left(\sum_{k \in A(p,m)} \widehat{f}(k) w_k(x) \right) \tag{24}$$

exists and is finite for all but countably many $x \in [a,b]$, *then* f *is constant on* $[a,b]$.

PROOF. Since f is continuous on $[a,b]$,

$$f(x) = \lim_{n\to\infty} \sum_{k=0}^{2^n-1} \widehat{f}(k) w_k(x)$$

uniformly for $x \in [a,b]$. In view of the identity

$$w_k(x \dot{+} 2^{-p-1}) = w_k(x) w_k(2^{-p-1}),$$

it follows that

$$f(x \dot{+} 2^{-p-1}) - f(x) = \lim_{n\to\infty} \sum_{k=0}^{2^n-1} \widehat{f}(k) w_k(x) \left(w_k(2^{-p-1}) - 1\right),$$

for $p = 0, 1, \ldots$. Since the factor $w_k(2^{-p-1}) - 1$ equals 0 or -2, and the latter case eventuates only when $k \in A(p,m)$ for some $m \geq p$, we deduce that

$$f(x \dot{+} 2^{-p-1}) - f(x) = -2 \sum_{m=p}^{\infty} \left(\sum_{k \in A(p,m)} \widehat{f}(k) w_k(x) \right),$$

for $p = 0, 1, \ldots$. In particular, if $g(x)$ represents the limit (24), then

$$\lim_{p \to \infty} \frac{f(x \dot{+} 2^{-p-1}) - f(x)}{2^{-p-1}} = -4g(x) \tag{25}$$

for all but countably many $x \in [a, b]$.

Let x be a dyadic irrational which satisfies (25). Since its binary expansion contains infinitely many zeros and ones, we can choose subsequences $(\ell_j, j \in \mathbf{N})$ and $(k_j, j \in \mathbf{N})$ of $(2^{-p-1}, p \in \mathbf{N})$ such that $x \dot{+} \ell_j = x + \ell_j$ and $x \dot{+} k_j = x - k_j$, for $j = 1, 2, \ldots$. It follows from (25), therefore, that

$$\lim_{j \to \infty} \frac{f(x + \ell_j) - f(x)}{\ell_j} = -\lim_{j \to \infty} \frac{f(x - k_j) - f(x)}{-(k_j)}.$$

In particular, the upper and lower Dini derivates of f satisfy

$$\overline{D} f(x) \geq 0 \geq \underline{D} f(x)$$

for all but countably many $x \in [a, b]$. Hence by Theorem 12 in Appendix 0.6 the function f is constant on $[a, b]$. ∎

COROLLARY 3. *Let $A(p,m)$ be defined as in Theorem 10. Set*

$$B(p,p) := A(p,p) \setminus \{2^{p+1} - 1\},$$

and

$$B(p,m) := A(p,m) \setminus \{2^m + \ell 2^{p+1} - 1 : 1 \leq \ell \leq 2^{m-p-1}\}.$$

If f is classically continuous on $[a, b] \subseteq [0, 1)$, and if its Walsh-Fourier coefficients satisfy

$$2^p \sum_{m=p}^{\infty} \left(\sum_{k \in B(m,p)} |\widehat{f}(k) - \widehat{f}(k+1)| \right) \to 0 \quad \text{as} \quad p \to \infty$$

(in particular, if $2^n \sum_{k=2^n}^{2^{n+1}-2} |\widehat{f}(k) - \widehat{f}(k+1)| \to 0$ as $n \to \infty$), then f is constant on $[a, b]$.

PROOF. Fix an integer p, and let $N := 2^{\ell_1} + 2^{\ell_2} + \cdots + 2^{\ell_{j_0}}$ where $\ell_1 > \ell_2 > \cdots > \ell_{j_0} \geq p$. Using the fact that each $D_{2^{\ell_j}}$ vanishes on $[2^{-p}, 1)$ it is not difficult to see that $D_N(x) = 0$

for $2^{-p} \le x \le 1$. Hence by Abel's transformation, given $x \in (0,1)$ one can choose an integer p so large that

$$\sum_{m=p}^{\infty}\left(\sum_{k\in A(p,m)} \widehat{f}(k)w_k(x)\right) = \sum_{m=p}^{\infty}\left(\sum_{k\in B(p,m)} (\widehat{f}(k)-\widehat{f}(k+1))D_{k+1}(x)\right).$$

Since $|D_{k+1}(x)| \le 2/x$, it follows from hypothesis that

$$2^p\sum_{m=p}^{\infty}\left(\sum_{k\in A(p,m)} \widehat{f}(k)w_k(x)\right) \to 0 \quad \text{as} \quad p\to\infty$$

for all $x \in (0,1)$. Hence by Theorem 10, the function f is constant on $[a,b]$. ∎

Suppose $a_0 := 0$ and

$$a_k := \frac{1}{n2^n} \qquad (2^n \le k < 2^{n+1}),$$

for $n \in \mathbf{P}$. Then

$$2^n \sum_{k=2^n}^{2^{n+1}-2} |a_k - a_{k+1}| \to 0$$

as $n \to \infty$, but $\sum_{k=0}^{\infty} |a_k| = \sum_{k=1}^{\infty} 1/k = \infty$. Thus Corollary 3 applies in situation where the Walsh series in question does not converge absolutely.

8.3 Representation of Measurable Functions. In 4.6 we showed that an a.e. finite-valued, measurable function g can be adjusted on a set of small measure so that its Walsh-Fourier series converges uniformly. If $g \in L^p, p > 1$, then g can be adjusted on a set E_0 so that its Walsh-Fourier coefficients are small, and its Walsh-Fourier series is L^p convergent off the set E_0. In fact,

THEOREM 11. *Let $B \subset [0,1)$ be measurable and suppose that $g \in L^p, p > 1$. If g vanishes off B and ε is any positive number, then there exist a function $G \in L^p$ and a measurable set $E_0 \subset B$ such that $|E_0| < \varepsilon$, $G(x) = g(x)$ for $x \notin E_0$, $|\widehat{G}(k)| \le \varepsilon$ for $k \in \mathbf{N}$, and*

$$\|S_nG\|_{L^p(E)} \le \varepsilon + \|g\|_{L^p(E)}$$

for all measurable sets $E \subset [0,1) \setminus E_0$ and all integers $n \ge 1$.

We postpone the proof of this result to the next section.

The purpose of this section is to use Theorem 11 to show that given any function f, measurable on $[0,1)$, there is a Walsh series which converges to f in measure, whether f is a.e. finite or not. Since no Walsh series can diverge to ∞ on a set of positive measure (see Theorem 15 in 7.6), this representation theorem is not true if "converges in measure" is replaced by "converges a.e.". To represent measurable f by a.e. convergent Walsh series, one must assume that f is a.e. finite- valued (see Exercise 8.12).

We begin with a corollary of Theorem 11. Throughout this section, given a subset E of $[0,1)$ the symbol $\overline{E}$ denotes the complement $[0,1) \setminus E$.

LEMMA 3. *Let $f_1 \in L^0$. If $\varepsilon > 0$ and $n_0 \in \mathbf{P}$ then there exist a measurable set $E_1 \subset [0,1)$, an integer $n_1 > n_0$, and real numbers a_k such that $|E_1| < \varepsilon$, $|a_k| < \varepsilon$ for $n_0 \le k < n_1$,*

$$\left\| \sum_{k=n_0}^{n_1-1} a_k w_k - f_1 \right\|_{L^p(\overline{E}_1)} < \varepsilon, \tag{26}$$

and

$$\left\| \sum_{k=n_0}^{m-1} a_k w_k \right\|_{L^p(E)} < \varepsilon + \|f_1\|_{L^p(E)} \tag{27}$$

for any measurable $E \subset \overline{E}_1$, and any integer $n_0 \le m < n_1$.

PROOF. Since f_1 is a.e. finite, we can choose a measurable set $B \subset [0,1)$ with $|\overline{B}| < \varepsilon/2$ and a bounded function g such that

$$g(x) = \begin{cases} f_1(x) & x \in B \\ 0 & x \in \overline{B}. \end{cases}$$

Apply Theorem 11 to choose $G \in L^p$ and the set E_0 such that $|E_0| < \varepsilon/2$, $G(x) = g(x)$ for $x \notin E_0$,

$$|\widehat{G}(k)| < \frac{\varepsilon}{2n_0} \qquad (k \in \mathbf{N}), \tag{28}$$

and

$$\|S_n G\|_{L^p(E)} \le \frac{\varepsilon}{2} + \|g\|_{L^p(E)} \tag{29}$$

for any measurable $E \subset \overline{E}_0$, and any integer $n \ge 1$.

Set $E_1 := \overline{B} \cup E_0$ and $a_k = \widehat{G}(k)$ for $k \in \mathbf{N}$. Then $|E_1| < \varepsilon$ and by (28), $|a_k| < \varepsilon$ for $k \in \mathbf{P}$.

To verify (26) observe that $S_n G \to G$ in L^p as $n \to \infty$. Since $f_1(x) = G(x)$ for $x \notin E_1$, it follows that $S_n G \to f_1$ in $L^p(\overline{E}_1)$ as $n \to \infty$. Thus we can choose $n_1 > n_0$ so large that

$$\|S_{n_1} G - f_1\|_{L^p(\overline{E}_1)} < \frac{\varepsilon}{2}.$$

Since (28) implies

$$\|S_{n_0} G\|_p \le \sum_{k=0}^{n_0-1} |a_k| \, \|w_k\|_p \le \frac{\varepsilon}{2},$$

it follows from Minkowski's inequality that

$$\|S_{n_1} G - S_{n_0} G - f_1\|_{L^p(\overline{E}_1)} < \varepsilon.$$

Thus (26) holds.

To verify (27) fix $E \subset \overline{E}_1$. Since E is a subset of the complement of E_0, we have by (29) that

$$\|S_m G - S_{n_0} G\|_{L^p(E)} \leq \frac{\varepsilon}{2} + \|g\|_{L^p(E)} + \|S_{n_0} G\|_p.$$

Since $\|S_{n_0} G\|_p \leq \varepsilon/2$, (27) follows immediately. ∎

A series $\sum_{k=1}^{\infty} f_k$ of functions is said to converge asymptotically to f in $L^p(A)$ for some measurable set A if given $\varepsilon > 0$ there is a measurable set $B \subset A$ with $|B| > |A| - \varepsilon$ such that

$$\lim_{N \to \infty} \int_B |\sum_{k=1}^{N} f_k - f|^p = 0.$$

Clearly, a sequence which converges asymptotically in $L^p(A)$ for some $1 \leq p < \infty$ also converges in measure on A.

THEOREM 12. *Let f be any function (finite-valued or not) which is measurable on $[0,1)$. If*

$$A := \{x \in [0,1) : |f(x)| < \infty\}$$

then there exists a Walsh series S which converges asymptotically to f in $L^p(A)$, for all $1 < p < \infty$, such that S converges to f in measure on $[0,1) \setminus A$.

PROOF. Let $A_+ := \{x : f(x) = +\infty\}$ and $A_- := [0,1) \setminus (A \cup A_+)$. Choose $\varepsilon_1 > \varepsilon_2 > \dots$ such that $\sum_{k=1}^{\infty} \varepsilon_k < \infty$ and set

$$f_1(x) := \begin{cases} f(x) & x \in A \\ 1 & x \in A_+ \\ -1 & x \in A_-. \end{cases}$$

Apply Lemma 3 to $\varepsilon := \varepsilon_1$ to choose an integer $n_1 > n_0$, a measurable set E_1, and coefficients a_k which satisfy $|a_k| < \varepsilon_1$ for $n_0 \leq k < n_1$, $|E_1| < \varepsilon$, (26), and (27).

Suppose $a_1, a_2, \dots, a_{n_{i-1}}$ have been chosen. Let

$$f_i(x) := \begin{cases} f(x) - \sum_{j=1}^{n_{i-1}} a_j w_j(x) & x \in A \\ 1 & x \in A_+ \\ -1 & x \in A_- \end{cases}$$

and apply Lemma 3 with f_i in place of f_1, n_{i-1} in place of n_0, and $\varepsilon = \varepsilon_i$. Thus choose a measurable set E_i, an integer $n_i > n_{i-1}$, and coefficients a_k such that

$$|E_i| < \varepsilon_i, \tag{30}$$

$$|a_k| < \varepsilon_i, \qquad (n_{i-1} \leq k < n_i), \tag{31}$$

(32)
$$\|\sum_{k=n_{i-1}}^{n_i-1} a_k w_k - f_i\|_{L^p(\overline{E}_i)} < \varepsilon_i,$$

and

(33)
$$\|\sum_{k=n_{i-1}}^{m-1} a_k w_k\|_{L^p(E)} < \varepsilon_i + \|f_i\|_{L^p(E)}$$

for any measurable $E \subset \overline{E}_i$ and any integer $n_{i-1} \le m < n_i$. In particular, there exist measurable sets $E_1, E_2, \ldots$, integers $n_0 < n_1 < \ldots$, and real numbers $a_1, a_2, \ldots$ such that conditions (30) through (33) hold for $i \in \mathbf{P}$.

Let $S := \sum_{k=1}^{\infty} a_k w_k$. To show that S converges to f asymptotically in $L^p(A)$, let $\varepsilon > 0$ and suppose without loss of generality that $A \ne \emptyset$. Choose i_0 so large that

$$\sum_{i=i_0}^{\infty} \varepsilon_i < \varepsilon,$$

and set

$$B := A \setminus \bigcup_{i=i_0}^{\infty} E_i.$$

Then $|B| > |A| - \varepsilon$.

Let $n \ge n_{i_0+1}$. Choose $i \ge i_0 + 1$ such that $n_i \le n < n_{i+1}$. Since $B \subset A$, the definition of f_i implies

$$\begin{aligned}\|S_n f - f\|_{L^p(B)} &\le \|\sum_{k=1}^{n_i-1} a_k w_k - f\|_{L^p(B)} + \|\sum_{k=n_i}^{n-1} a_k w_k\|_{L^p(B)} \\ &= \|\sum_{k=n_{i-1}}^{n_i-1} a_k w_k - f_i\|_{L^p(B)} + \|\sum_{k=n_i}^{n-1} a_k w_k\|_{L^p(B)}.\end{aligned}$$

However, $B \subset \overline{E}_i$ so (32) and (33) imply

(34)
$$\|S_n f - f\|_{L^p(B)} \le \varepsilon_i + \varepsilon_{i+1} + \|f_{i+1}\|_{L^p(B)}.$$

Since by construction

$$f_{i+1}(x) = f_i(x) - \sum_{k=n_{i-1}}^{n_i-1} a_k w_k(x)$$

for $x \in A$, it follows from (32) that the $L^p(B)$ norm of f_{i+1} is dominated by ε_i. Substituting this estimate into (34) we arrive at

$$\|S_n f - f\|_{L^p(B)} \le 2\varepsilon_i + \varepsilon_{i+1}.$$

Since $i \to \infty$ as $n \to \infty$, we conclude that $S_n f \to f$ in $L^p(B)$ norm, as $n \to \infty$, and thus that S converges to f asymptotically in $L^p(A)$.

By symmetry, it remains to see that S converges to $+\infty$ in measure on A_+. We may suppose that $|A_+| > 0$. Hence given $\varepsilon > 0$ and $M > 0$ we need to verify that

$$|\{x \in A_+ : S_n(x) \geq M\}| \geq |A_+| - \varepsilon$$

for n sufficiently large.

Choose i_0 so large that

$$\sum_{i=i_0}^{\infty} \varepsilon_i < \frac{\varepsilon}{4(2^p + 1)}.$$

Without loss of generality, suppose both ε and ε_i are less than 1 for $i \geq i_0$. Notice immediately that the construction of f_i and (32) imply

$$\int_{A_+ \cap \overline{E}_i} |\sum_{k=n_{i-1}}^{n_i - 1} a_k w_k - 1|^p < \varepsilon_i^p.$$

In particular, if

$$A_i := \{x \in A_+ : \sum_{k=n_{i-1}}^{n_i - 1} a_k w_k < \frac{1}{2}\},$$

then $|A_i \cap \overline{E}_i| \leq \varepsilon_i^p 2^p$. Thus by (30) we have that

$$|A_i| \leq (1 + 2^p)\varepsilon_i \qquad (i \in \mathbf{P}). \tag{35}$$

Since the Walsh functions are a.e. finite (a considerable understatement, but this remark pertains to generalizations discussed in the next section), we can choose an integer M_0 so large that

$$|\{x \in A_+ : \sum_{k=1}^{n_{i_0} - 1} a_k w_k(x) > -M_0\}| > |A_+| - \frac{\varepsilon}{4}.$$

Moreover, for any $j > i_0$ it is clear that

$$\{x \in A_+ : \sum_{k=n_{i_0}}^{n_j - 1} a_k w_k(x) < (j - i_0)\frac{1}{2}\} \subseteq \bigcup_{i=i_0}^{j-1} A_i.$$

It follows, therefore, from (35) and the choice of i_0 that

$$|\{x \in A_+ : \sum_{k=n_{i_0}}^{n_j - 1} a_k w_k(x) \geq (j - i_0)\frac{1}{2}\}| > |A_+| - \frac{\varepsilon}{4}.$$

Hence the choice of M_0 implies that

$$(36) \qquad |\{x \in A_+ : \sum_{k=n_{i_0}}^{n_j-1} a_k w_k(x) \geq (j-i_0)\frac{1}{2} - M_0\}| > |A_+| - \frac{\varepsilon}{2}$$

for any integer $j > i_0$.

Fix $j_0 > i_0$ so that

$$(j_0 - i_0)\frac{1}{2} - M_0 \geq M + 2\left(\frac{4}{\varepsilon}\right)^{1/p}$$

and set $N := n_{j_0}$. We shall complete the proof of this theorem by showing that

$$|\{x \in A_+ : S_n(x) < M\}| \leq \varepsilon$$

for any $n \geq N$.

Toward this, fix $n \geq N$ and choose $j \geq j_0$ so that $n_j \leq n < n_{j+1}$. Observe that

$$\begin{aligned} |\{x \in A_+ : S_n(x) < M\}| &\leq |\{x \in A_+ : S_n(x) - S_{n_j}(x) < -2(\frac{4}{\varepsilon})^{1/p}\}| \\ &\quad + |\{x \in A_+ : S_{n_j}(x) < M + 2(\frac{4}{\varepsilon})^{1/p}\}| \\ &=: I_1 + I_2. \end{aligned}$$

To estimate I_1, use (30) and the choice of i_0 to verify

$$|A_+ \cap \overline{E_{j+1}}| > |A_+| - \frac{\varepsilon}{4}$$

since $j \geq j_0 > i_0$. Moreover, by (33) for $i = j+1$ we have

$$\int_{\overline{E_{j+1}}} |S_n - S_{n_j}|^p \leq (1 + \varepsilon_{j+1})^p < 2^p.$$

In particular, it follows that $|I_1| \leq \varepsilon/2$.

To estimate I_2, use the choice of j_0 to see that if $j \geq j_0$ then

$$(j - i_0)\frac{1}{2} - M_0 \geq M + 2\left(\frac{4}{\varepsilon}\right)^{1/p}.$$

Hence by (36), (31) and the choice of i_0, we conclude that $|I_2| \leq \varepsilon/2$. ∎

COROLLARY 4. *If f is measurable on $[0,1)$ then there is a Walsh series with coefficients tending to zero such that $S_n \to f$ in measure as $n \to \infty$.*

If f belongs to L^0 we can also represent f asymptotically.

COROLLARY 5. *If f is measurable and a.e. finite on $[0,1)$, then there exists a Walsh series S with coefficients tending to zero such that S converges asymptotically to f in L^p for all $1 < p < \infty$.*

8.4 Normalized Convergence Systems. Throughout this section let $\mathbf{u} = (u_n, n \in \mathbf{N})$ represent a *normalized convergence system* in L^p for some $1 < p < \infty$, i.e., $\mathbf{u}$ is orthogonal in L^2, the Fourier series of any $f \in L^p$ with respect to the system $\mathbf{u}$ converges to f in the L^p norm, and $\|u_k\|_p = 1$, $\|u_k\|_q \le M < \infty$ for $k = 1, 2, \ldots$, where p and q are conjugate indices. Thus the Walsh, Ciesielski, and trigonometric systems are normalized convergence systems in L^p for all $1 < p < \infty$.

For each $f \in L^1$ and each integer $n \in \mathbf{P}$ let

$$a_n(f) := \int_0^1 f(t) u_n(t)\, dt$$

and

$$S_n(f) = \sum_{k=0}^{n-1} a_k(f) u_k.$$

The functions u_k are necessarily a.e. finite. Moreover, by the Banach-Steinhaus theorem there is a constant C_p, depending only on p, such that

$$\|S_n(f)\|_p \le C_p \|f\|_p \qquad (f \in L^p, n \in \mathbf{P}). \tag{37}$$

In particular, it is a routine exercise to check that the results of the previous section hold for any normalized convergence system, not just the Walsh system.

The purpose of this section is to show that Theorem 11 in 8.3 holds for any normalized convergence system (see Theorem 13 below). We begin with a technical result which localizes the construction of G in Theorem 13.

LEMMA 4. *Let I be an interval in $[0,1]$ and $B \subset I$ be measurable. Suppose $1 < p < \infty$ and that $g \in L^p$ vanishes off the set B. If $n > 0$ is an integer, if $0 < \varepsilon_0 < 1$ and $\varepsilon > 0$, then there exist a bounded function g_1 and a measurable set $E_1 \subset B$ such that g_1 vanishes off E_1, $|E_1| \le \varepsilon_0 |I|$,*

$$\int_I |g_1|^p \le \varepsilon_0^{1-p} \int_I |g|^p \tag{38}$$

and

$$\left| \int_I (g_1 - g) u_k \right| < \varepsilon \qquad (1 \le k \le n). \tag{39}$$

PROOF. We may suppose that $I = [\alpha, \beta)$ for some $0 \le \alpha < \beta < 1$. For each N partition I into even intervals I_j^N, $j = 1, 2, \ldots, N$ and choose subintervals J_j^N of I_j^N such that $|J_j^N| = \varepsilon_0 |I_j^N|$ for $1 \le j \le N$. Define operators $\boldsymbol{F}_N$ and $\boldsymbol{G}_N$ from $L^1(I)$ into $L^\infty(I)$ by

$$\boldsymbol{F}_N f := \sum_{j=1}^N \chi(I_j^N) \frac{1}{|I_j^N|} \int_{I_j^N} f$$

and

$$G_N f := \sum_{j=1}^{N} \chi(J_j^N) \frac{1}{|J_j^N|} \int_{I_j^N} f$$

for $f \in L^1(I)$ and $N \in \mathbf{N}$.

Suppose for a moment that

$$\lim_{N\to\infty} \int_I h G_N f = \int_I hf \tag{40}$$

for all $f \in L^p(I)$ and all $h \in L^q(I)$, where q is the index conjugate to p. Apply (40) to $h := u_k$, $1 \le k \le n$, and $f := g$ to choose N so large that

$$|\int_I (G_N g - g) u_k| < \varepsilon \tag{41}$$

for $1 \le k \le n$. Set

$$g_1 := \chi(B) G_N g,$$

and

$$E_1 := \bigcup_{j=1}^{N} B \cap J_j^N.$$

Clearly (39) follows from (41). By construction g_1 is a bounded function which vanishes off E_1, E_1 is a subset of B, and

$$|E_1| \le \sum_{j=1}^{N} |J_j^N| = \varepsilon_0 |I|.$$

Finally, by Hölder's inequality

$$\|G_N f\|_{L^p(I)} \le \varepsilon_0^{1/p-1} \|f\|_{L^p(I)} \tag{42}$$

for every $f \in L^p(I)$. Specializing to $f := g$ we see that (38) also holds. Consequently, it remains to prove (40).

Fix $1 \le r < \infty$. Notice for any interval $J \subseteq I$ and $h := \chi(J)$ that

$$\lim_{N\to\infty} \|F_N h - h\|_r = 0. \tag{43}$$

Since the collection of characteristic functions of subintervals of I forms a closed system in $L^r(I)$ it is evident that (43) holds for all $h \in L^r(I)$.

Fix $f \in L^p(I)$ and $h \in L^q(I)$. Notice by construction that

$$\int_{I_j^N} F_N f = \int_{I_j^N} f = \int_{I_j^N} G_N f \tag{44}$$

for $1 \le j \le N$ and $N \in \mathbf{N}$. Write

$$\int_I h(\boldsymbol{G}_N f - f) = \int_I (h - \boldsymbol{F}_N h)(\boldsymbol{G}_N f - f) + \int_I (\boldsymbol{F}_N h)(\boldsymbol{G}_N f - f).$$

By (44) the second term of this identity is zero. Therefore, it follows from (42) and Hölder's inequality that

$$\left| \int_I h(\boldsymbol{G}_N f - f) \right| \le \int_I |h - \boldsymbol{F}_N h| |\boldsymbol{G}_N f - f|$$
$$\le \left(1 + \frac{1}{\varepsilon_0}\right) \|f\|_p \|h - \boldsymbol{F}_N h\|_q.$$

Since $r := q \ne \infty$ it is now clear that (40) follows directly from (43). In particular, the function g_1 satisfies (39) and the proof of the lemma is complete. ∎

THEOREM 13. *Let $B \subset [0,1]$ be measurable and suppose that $\boldsymbol{u} = (u_k, k \in \mathrm{N})$ is a normalized convergence system in L^p for some $p > 1$. If $g \in \mathrm{L}^p$ vanishes off B and $\varepsilon > 0$ then there exist a function $G \in \mathrm{L}^p$ and a measurable set $E_0 \subset B$ such that $|E_0| < \varepsilon$, $G(x) = g(x)$ for $x \notin E_0$, $|a_k(G)| \le \varepsilon$ for $k \in \mathbf{N}$, and*

$$\|S_n(G)\|_{L^p(E)} \le \varepsilon + \|g\|_{L^p(E)}$$

for $n \in \mathbf{P}$ and all measurable $E \subset [0,1] \setminus E_0$.

PROOF. Since $\boldsymbol{u}$ is a normalized convergence system in L^p, it follows from (37) that

$$\sup_{n \in \mathbf{P}} \|S_n(h)\|_p \le C_p \|h\|_p$$

for some absolute constant C_p, and

$$|a_n(h)| \le \|u_n\|_q \|h\|_p \le M \|h\|_p$$

for $h \in \mathrm{L}^p$ and $n \in \mathbf{P}$. Consequently, there is a $0 < \delta < 1$ such that $\delta \le \varepsilon$ and

$$(45) \qquad \|h\|_p < \delta \quad \text{implies} \quad |a_n(h)|,\ \|S_n(h)\|_p < \frac{\varepsilon}{4} \qquad (n \in \mathbf{P}).$$

Divide $[0,1]$ into even, non-overlapping intervals $I_1, I_2, \dots, I_m$ such that

$$\int_{I_k} |g|^p < \frac{\delta^{2p}}{2^{p+1}} \qquad (1 \le k \le m).$$

Fix $n_1 \ge 1$ and apply Lemma 4 to g, $n := n_1$, $I := I_1$, $\varepsilon_0 := \delta$, with $\varepsilon/(4n_1)$ in place of ε, and $B \cap I_1$ in place of B. Thus choose a bounded function g_1 and a measurable set $E_1 \subset B \cap I_1$ such that g_1 vanishes off E_1, $|E_1| \le \varepsilon |I_1|$,

$$\int_{I_1} |g_1|^p \le \delta^{1-p} \int_{I_1} |g|^p,$$

and, if $a_n^{(1)} := \int_{I_1}(g - g_1)u_n$ and $S_n^{(1)} := \sum_{k=1}^n a_k^{(1)}u_k$, then

$$|a_n^{(1)}| < \frac{\varepsilon}{4n_1}$$

and

$$\|S_n^{(1)}\|_p \leq \frac{\varepsilon}{4} \qquad (1 \leq n < n_1).$$

(Recall that $\|u_k\|_p = 1$.)

Set

$$v_1(x) := \begin{cases} g(x) & x \in I_1 \setminus E_1 \\ 0 & \text{otherwise} \end{cases}$$

and

$$\widetilde{v}_1(x) := \begin{cases} g(x) - g_1(x) & x \in I_1 \\ 0 & \text{otherwise.} \end{cases}$$

Since g_1 vanishes off E_1, it is clear that

$$v_1(x) = \widetilde{v}_1(x) \qquad (x \notin E_1).$$

Also, our notation guarantees that

$$a_k^{(1)} = a_k(\widetilde{v}_1) \qquad (k \in \mathbf{P}).$$

Thus $S_n^{(1)}$ represents the partial sums of the Fourier series of $\widetilde{v}_1$ in the system $\boldsymbol{u}$. Since this system is a normalized convergence system in L^p, and since $g \in L^p$, it follows from construction that $S_n^{(1)} \to v_1$ in $L^p(\overline{E}_1)$ as $n \to \infty$. Thus choose n_2 so large that

$$|a_n^{(1)}| < \frac{\varepsilon}{4} \quad \text{and} \quad \|S_n^{(1)} - v_1\|_{L^p(\overline{E}_1)} \leq \frac{\varepsilon}{4}$$

for $n \geq n_2$.

Apply Lemma 4 to the interval I_2, choosing a bounded function g_2 and a measurable set $E_2 \subset I_2 \cap B$ such that g_2 vanishes off E_2, $|E_2| < \varepsilon |I_2|$,

$$\int_{I_2} |g_2|^p < \delta^{1-p} \int_{I_2} |g|^p$$

and if $a_n^{(2)} := \int_{I_2}(g - g_2)u_n$ and $S_n^{(2)} := \sum_{k=1}^n a_n^{(2)}u_k$, then $|a_n^{(2)}| < \varepsilon/8$ and

$$\|S_n^{(2)}\|_p \leq \frac{\varepsilon}{8} \qquad (1 \leq n < n_2).$$

As above, since $\boldsymbol{u}$ is a normalized convergence system for L^p we can choose n_3 so large that

$$|a_n^{(1)} + a_n^{(2)}| < \frac{\varepsilon}{8}$$

and construct a function v_2 (see the general case) which satisfies

$$\|S_n^{(1)} + S_n^{(2)} - v_2\|_{L^p(\overline{E_1 \cup E_2})} < \frac{\varepsilon}{8} \qquad (n \geq n_3).$$

Continuing in this manner, for each $1 \leq k \leq m$ we choose a bounded function g_k and a measurable set $E_k \subset I_k \cap B$ such that

(46) $$g_k \quad \text{vanishes off} \quad E_k,$$

(47) $$|E_k| < \varepsilon |I_k|,$$

(48) $$\int_{I_k} |g_k|^p < \delta^{1-p} \int_{I_k} |g|^p,$$

and if $a_n^{(k)} := \int_{I_k} (g - g_k) u_n$, $S_n^{(k)} := \sum_{j=1}^n a_j^{(k)} u_j$ then

(49) $$|a_n^{(k)}| < \frac{\varepsilon}{2^{k+1}} \qquad (1 \leq n < n_k),$$

(50) $$\|S_n^{(k)}\|_p \leq \frac{\varepsilon}{2^{k+1}} \qquad (n \geq n_k),$$

and

(51) $$|\sum_{j=1}^{k-1} a_n^{(j)}| < \frac{\varepsilon}{2^k} \qquad (n \geq n_k).$$

Moreover, if

$$v_k(x) := \begin{cases} g(x) & x \in \bigcup_{\ell=1}^k I_\ell \setminus \left(\bigcup_{\ell=1}^k E_\ell \right) \\ 0 & \text{otherwise,} \end{cases}$$

then

(52) $$\| \sum_{j=1}^k S_n^{(j)} - v_k \|_{L^p(\overline{E_1 \cup \cdots \cup E_k})} < \frac{\varepsilon}{2^{k+1}} \qquad (n \geq n_{k+1}).$$

Set $E_0 := E_1 \cup \cdots \cup E_m$ and

$$G(x) := g(x) - \sum_{k=1}^m g_k(x) \qquad (x \in [0,1]).$$

Since each g_k is bounded, it is clear that $G \in \mathrm{L}^p$. By (47),

$$|E_0| \leq \varepsilon \sum_{k=1}^{m} |I_k| = \varepsilon.$$

And, since each g_k is supported on E_k, it is evident that $G(x) = g(x)$ for $x \notin E_0$.

It remains to estimate $a_n(G)$ and the $\mathrm{L}^p(E)$ norm of $S_n(G)$. Notice once and for all that

$$a_n(G) = \sum_{j=1}^{m} a_n^{(j)} \qquad (n \in \mathbf{P}) \tag{53}$$

since

$$a_n(G) = \sum_{j=1}^{m} \int_{I_j} G\, u_n = \sum_{j=1}^{m} a_n^{(j)}.$$

Notice also that

$$\|G\|_{L^p(I_k)} < \delta \qquad (1 \leq k \leq m). \tag{54}$$

Indeed,

$$\begin{aligned} \|G\|_{L^p(I_k)}^p &= \int_{I_k} |g - g_k|^p \\ &\leq 2^p \int_{I_k} (|g|^p + |g_k|^p) \\ &\leq \frac{\delta^{2p}}{2} + 2^p \int_{I_k} |g_k|^p \end{aligned}$$

by the choice of I_k. Consequently, (54) follows immediately from (48).

Observe by definition that

$$a_n^{(j)} = \int_{I_k} G\, u_n, \qquad (1 \leq j, k \leq m,\ n \in \mathbf{N}).$$

Apply (45) to

$$h(x) := \begin{cases} G(x) & x \in I_j \\ 0 & x \notin I_j. \end{cases}$$

We see by (54) that

$$|a_n^{(j)}| \leq \frac{\varepsilon}{4} \quad \text{and} \quad \|S_n^{(j)}\|_p \leq \frac{\varepsilon}{4}, \tag{55}$$

for $1 \leq j \leq m$ and $n \in \mathbf{N}$.

Suppose $n < n_m$. Set $n_0 := 0$ and choose $2 \leq k \leq m$ such that $n_{k-1} \leq n < n_k$. It follows from (53) that

$$|a_n(G)| \leq |\sum_{j=1}^{k-2} a_n^{(j)}| + |a_n^{(k-1)}| + \sum_{j=k}^{m} |a_n^{(j)}|.$$

Thus by (51), (55), and (49), we have

$$|a_n(G)| \leq \frac{\varepsilon}{2^{k-1}} + \frac{\varepsilon}{4} + \frac{\varepsilon}{2^k} \leq \varepsilon.$$

On the other hand, if $n \geq n_m$ then (53) and (55) imply that

$$|a_n(G)| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2^m} \leq \varepsilon.$$

Hence this inequality holds for all $n \in \mathbf{N}$.

To estimate the $\mathrm{L}^p(E)$ norm of $S_n(G)$, fix $E \subset [0,1] \setminus E_0$ and observe by (53) that

$$S_n(G) = \sum_{j=1}^{m} S_n^{(j)} \qquad (n \in \mathbf{P}). \tag{56}$$

Again, we separate the cases $n < n_m$ and $n \geq n_m$. In the former case, choose $k \leq m$ so that $n_{k-1} \leq n < n_k$ and use (56) to write

$$\begin{aligned} \|S_n(G)\|_{L^p(E)} \leq \|\sum_{j=1}^{k-2} S_n^{(j)} - v_{k-2}\|_{L^p(E)} + \|v_{k-2}\|_{L^p(E)} \\ + \|S_n^{(k-1)}\|_{L^p(E)} + \sum_{j=k}^{m} \|S_n^{(j)}\|_{L^p(E)}. \end{aligned}$$

Hence by (52), the definition of v_{k-2}, the right side of (55), and (50), it follows that

$$\begin{aligned} \|S_n(G)\|_{L^p(E)} &\leq \frac{\varepsilon}{2^{k-1}} + \|g\|_{L^p(E)} + \frac{\varepsilon}{4} + \frac{\varepsilon}{2^k} \\ &\leq \varepsilon + \|g\|_{L^p(E)}, \end{aligned}$$

as required. On the other hand, if $n \geq n_m$ then by (56) we can write

$$\|S_n(G)\|_{L^p(E)} \leq \|\sum_{j=1}^{m-1} S_n^{(j)} - v_{m-1}\|_{L^p(E)} + \|v_{m-1}\|_{L^p(E)} + \|S_n^{(m)}\|_{L^p(E)}.$$

Using (52), the definition of v_{m-1}, and (55), we conclude that

$$\begin{aligned} \|S_n(G)\|_{L^p(E)} &\leq \frac{\varepsilon}{2^m} + \|g\|_{L^p(E)} + \frac{\varepsilon}{4} \\ &\leq \varepsilon + \|g\|_{L^p(E)}. \qquad \blacksquare \end{aligned}$$

Exercises

8.1. Prove that $\| \sup_n |D_n| \|_p < \infty$ for $0 < p < 1$ and show that if S is a Walsh series whose coefficients satisfy $a_k \downarrow 0$, as $k \to \infty$, then the limit of S belongs to L^p for $0 < p < 1$.

8.2. A function f is said to be A-integrable on $[0,1]$ if

$$(A)\int_0^1 f := \lim_{N\to\infty} \int_0^1 f_N(x)dx$$

exists and is finite, and if

$$\lim_{N\to\infty} N|\{x \in [0,1] : |f(x)| > N\}| = 0,$$

where

$$f_N(x) := \begin{cases} f(x) & |f(x)| \le N \\ 0 & |f(x)| > N. \end{cases}$$

Show that if S is a Walsh series whose coefficients satisfy $a_k \downarrow 0$, as $k \to \infty$, then its limit f is A- integrable and

$$a_k = (A)\int_0^1 f w_k \qquad (k \in \mathbf{N}).$$

[Rubinštein [1]]

8.3. Suppose $(a_k, k \in \mathbf{N})$ is quasi-convex and converges monotonically to 0. Prove $\sum_{k=0}^{\infty} a_k w_k$ converges in L^1 norm if and only if $a_n \log n \to 0$ as $n \to \infty$.

(Hint: Use the proof of Theorem 4.)

[Fomin]

8.4. Show that $(1/(n+1), n \in \mathbf{N})$ and $(t^n, n \in \mathbf{N})$, are completely monotone for each $0 < t < 1$.

8.5. Show that if $f(x) = \sum_{k=0}^{\infty} a_k w_k(x)$ is classically differentiable on $(0,1)$ and if $(a_k, k \in \mathbf{N})$ is completely monotone then f is constant on $(0,1)$.

[Coury[2]]

8.6. Prove that $\| \sum_{k=0}^{n} \kappa_k \|_1 = 0(\log(n+1))$, as $n \to \infty$, where $\kappa_0, \kappa_1, \ldots$ represents the Walsh-Kaczmarz system.

8.7. a) Let $f(x) := \sum_{k=2}^{\infty} w_k(x)/(k \log k)$ and show that

$$f^{[1]}(x) = \sum_{k=2}^{\infty} w_k(x)/\log k$$

for all $x \in (0,1), x \neq 2^{-j}, j = 1,2,\ldots$.

b) Let $f(x) := \sum_{k=1}^{\infty} k^{-\alpha} w_k(x)$ for $\alpha > 1$ and show that

$$f^{[1]}(x) = \sum_{k=1}^{\infty} k^{-\alpha+1} w_k(x)$$

for all $x \in (0,1), x \neq 2^{-j}, j = 1,2,\ldots$.

[Butzer and Wagner [2]]

8.8. a) Suppose that $g(x) = \sum_{k=1}^{\infty} ka_k w_k(x)$ converges absolutely for some $x \in [0,1)$. Prove that $f := \sum_{k=1}^{\infty} a_k w_k$ is everywhere dyadically differentiable with $f^{[1]}(x) = g(x)$ for $[0,1)$.

Hint: Use

$$(\mathbf{d}_n f)(x) = \sum_{k=0}^{\infty} \langle k \rangle_n \, a_k w_k(x)$$

directly.

b) Show the formula in Exercise 8.7 b) for $f^{[1]}(x)$ holds for all $x \in [0,1)$ if $\alpha > 2$.

[Butzer and Wagner [2]]

8.9. Suppose that $(ka_k, k \in \mathbf{N})$ is of bounded variation. Prove that $f := \sum_{k=0}^{\infty} a_k w_k$ converges on $(0,1)$ and is dyadically differentiable with

$$f^{[1]}(x) = \sum_{k=1}^{\infty} ka_k w_k(x)$$

for all $x \in (0,1), x \neq 2^{-j}, j = 1, 2, \ldots$.

[Schipp [12]]

8.10. Let $x \in (0,1), x \neq 2^{-j}, j = 1, 2, \ldots$ and let $a_0, a_1, \ldots$ be real numbers. Consider

$$f(t) := \sum_{k=0}^{\infty} a_k w_k(t)$$

and

$$g(t) := \lim_{n \to \infty} \sum_{k=0}^{2^n - 1} ka_k w_k(t)$$

whenever these limits exist.

a) If $ka_k \to 0$ and

$$k \sum_{j=k}^{2k} |a_j - a_{j+1}| \to 0 \quad \text{as} \quad k \to \infty,$$

then $f^{[1]}(x)$ exists if and only if $g(x)$ exists in which case

$$f^{[1]}(x) = g(x).$$

b) If $(a_k, k \in \mathbf{N})$ decays monotonically to zero and f is absolutely convergent, then $f^{[1]}(x)$ and $g(x)$ exist, in which case

$$f^{[1]}(x) = g(x).$$

8.11. If $S = \sum_{k=1}^{\infty} a_{n_k} w_{n_k}$ where $n_{k+1}/n_k \geq q > 1$ for $k = 1, 2, \ldots$, and if

$$g(x) = \sum_{k=1}^{\infty} n_k a_{n_k} w_{n_k}(x)$$

converges for some $x \in (0,1), x \neq 2^{-j}, j = 1,2,\ldots$, then S converges on $(0,1)$, is dyadically differentiable at x and satisfies $f^{[1]}(x) = g(x)$.

8.12. a) Let I be a dyadic interval and f be its characteristic function. Use the definition of the Haar system to show that given integers $n, m > 0$ there is a Haar polynomial

$$P = \sum_{j=n}^{N} a_j h_j$$

such that $h_j = 0$ off I for $n \leq j \leq N$,

$$|I \cap \{x : P(x) = f(x)\}| > |I|(1 - 2^m)$$

and

$$\sum_{j=n}^{N} |a_j h_j(x)| \leq 2^{m+1} \qquad (x \in I).$$

b) Let f be continuous on $[0,1]$, let ε and δ be positive numbers, and $n \in \mathbf{P}$. Then there exist a Haar polynomial

$$P = \sum_{j=n}^{N} a_j h_j$$

and a measurable subset E of $[0,1]$ with $|E| > 1 - \varepsilon$ such that

$$|\sum_{j=n}^{N} a_j h_j(x) - f(x)| < \delta$$

for $x \in E$ and

$$\sum_{j=n}^{N} |a_j h_j(x)| \leq \frac{8}{\varepsilon}(|f(x)| + \delta)$$

for $x \in [0,1]$.

c) Show that given any function f which is measurable and a.e. finite-valued on $[0,1]$ there is a Haar series which converges absolutely a.e. to f.

d) Use the Hadamard transform and c) above to prove that given any function f which is a.e. finite-valued and measurable on $[0,1]$, there is a Walsh series S such that $S_{2^n} \to f$ a.e., as $n \to \infty$.

[Arutunjan [1]]

Chapter 9

THE WALSH-FOURIER TRANSFORM

9.1 The Dyadic Field. Throughout this chapter let $\mathbf{F}$ denote the set of doubly infinite sequences

$$x = (x_n, n \in \mathbf{Z})$$

where $x_n = 0$ or 1 and $x_n \to 0$ as $n \to -\infty$. Denote the doubly infinite sequence whose entries are identically zero by 0. Thus to each $x \in \mathbf{F}$ with $x \neq 0$ there corresponds an integer $\boldsymbol{S}(x) \in \mathbf{Z}$ such that

(1) $$x_{\boldsymbol{S}(x)} = 1 \quad \text{but} \quad x_n = 0 \quad \text{for} \quad n < \boldsymbol{S}(x).$$

Let $x = (x_n, n \in \mathbf{Z})$ and $y = (y_n, n \in \mathbf{Z})$ be elements of $\mathbf{F}$. Define the sum of x and y by

(2) $$x + y := (|x_n - y_n|, n \in \mathbf{Z}).$$

Define the product of x and y by

(3) $$x \cdot y := (\xi_n, n \in \mathbf{Z})$$

where for each $n \in \mathbf{Z}$,

$$\xi_n := \sum_{i+j=n} x_i y_j \pmod 2.$$

Notice that $(\mathbf{F}, +)$ is an abelian group, $(\mathbf{F}, \cdot)$ is an abelian semigroup, and in fact, that $(\mathbf{F}, +, \cdot)$ is a commutative algebra over the finite field $\mathbf{Z}_2 := \{0, 1\}$. This algebra has an identity

(4) $$e_0 := (\delta_{0,n}, n \in \mathbf{Z})$$

and contains a subgroup

$$\mathbf{F}_0 := \{x \in \mathbf{F} : x_n = 0 \quad \text{for} \quad n < 0\}$$

which is isomorphic to the dyadic group $\mathbf{G}$.

The algebra $\mathbf{F}$ is normed. Indeed, for $x = (x_n, n \in \mathbf{Z}) \in \mathbf{F}$ define

(5) $$|x| := \sum_{n \in \mathbf{Z}} x_n 2^{-n-1}.$$

It is easy to see that $|x| \geq 0$,

$$|x + y| \leq |x| + |y|,$$

and

$$|x \cdot y| \le |x|\ |y|$$

for all $x, y \in \mathbf{F}$.

The map $x \to |x|$ takes $\mathbf{F}$ onto $\mathbf{R}^+ := [0,\infty)$ and is $1-1$ off a countable subset of $\mathbf{F}$. For, if $\mathbf{Q}^+ := \{p2^q : p \in \mathbf{N}, q \in \mathbf{Z}\}$ represents the dyadic rationals in $\mathbf{R}^+$ then every point in $\mathbf{R}^+ \setminus \mathbf{Q}^+$ has exactly one preimage in $\mathbf{F}$. On the other hand, each point in $\mathbf{Q}^+$ has two preimages in $\mathbf{F}$ (one which satisfies $x_n \to 0$ as $n \to \infty$ and one which satisfies $x_n \to 1$ as $n \to \infty$). At any rate, the map $x \to |x|$ allows us to identify $\mathbf{F}$ with $\mathbf{R}^+$ in essentially the same way that Fine's map allowed us to identify the dyadic group with $[0,1)$.

There is another norm, a non-Archimedean one, which can be defined on $\mathbf{F}$. Set $\|0\| := 0$ and for each $x \in \mathbf{F}$ with $x \neq 0$ set

$$\|x\| := 2^{-\mathcal{S}(x)} \tag{6}$$

where $\mathcal{S}(x)$ is defined in (1). Notice that

$$\|x+y\| \le \max(\|x\|, \|y\|)$$

and

$$\|x \cdot y\| = \|x\|\ \|y\| \tag{7}$$

for $x, y \in \mathbf{F}$. Also, by definition we have

$$\frac{1}{2}\|x\| \le |x| \le \|x\| \qquad (x \in \mathbf{F}).$$

Hence (6) is a norm on the algebra $\mathbf{F}$ which is equivalent to (5).

Let

$$\mathbf{B} := \{x \in \mathbf{F} : \|x\| = 1\}$$

denote the unit ball in $\mathbf{F}$. It is easy to see that $\mathbf{B}$ is a multiplicative subgroup of $\mathbf{F}$. Indeed, if $x \in \mathbf{B}$ then the equation $x \cdot y = e_0$ is equivalent to the system of equations

$$\|y\| = 1,$$

$$x_0\, y_0 = 1,$$

and

$$\sum_{i+j=n} x_i y_j \pmod 2 = 0 \qquad (n \in \mathbf{P}).$$

This system uniquely determines $y := x^{-1}$. Consequently, each $x \in \mathbf{B}$ has an inverse in $\mathbf{B}$ and $\mathbf{B}$ is a multiplicative group.

Define the usual closed system in $\mathbf{F}$ by

$$e_n := (\delta_{n,j}, j \in \mathbf{Z})$$

and observe that

$$e_n \cdot x = (x_{j-n}, j \in \mathbf{Z})$$

for each $n \in \mathbf{Z}$ and $x = (x_j, j \in \mathbf{Z}) \in \mathbf{F}$. Thus multiplication by e_n is a shift operator on $\mathbf{F}$. Clearly,

$$e_n \cdot e_m = e_{n+m}$$

and

$$\|e_n\| = 2^{-n} \qquad (n, m \in \mathbf{Z}).$$

Thus $\{e_n : n \in \mathbf{Z}\}$ forms a 1-parameter subgroup of $\mathbf{F}$ which is algebraically isomorphic to $\mathbf{Z}$.

Let $x \in \mathbf{F}$, $x \neq 0$. Choose $n \in \mathbf{Z}$ such that $\|x\| = 2^{-n}$. By (7) we have

$$\|e_{-n} \cdot x\| = 1.$$

Hence $e_{-n} \cdot x$ is invertible and it follows that x is invertible with

$$x^{-1} = (e_n \cdot e_{-n} \cdot x)^{-1} = (e_{-n} \cdot x)^{-1} \cdot e_{-n}.$$

Therefore $\mathbf{F}$ is a field. It is called the dyadic field.

Let $\mathbf{F}^* := \mathbf{F} \setminus \{0\}$. Then $\mathbf{F}^*$ is an abelian group under the multiplication defined in (3). We have just seen that $\{e_n : n \in \mathbf{Z}\}$ is a subgroup of $\mathbf{F}^*$ isomorphic to $\mathbf{Z}$, and that given $x \in \mathbf{F}^*$ there is an $n \in \mathbf{Z}$ such that $e_{-n} \cdot x \in \mathbf{B}$. It follows that

(8) $$\mathbf{F}^* \cong \mathbf{Z} \times \mathbf{B}$$

where $\cong$ represents an algebraic isomorphism whose existence is guaranteed by the fundamental theorem of homomorphisms from group theory. We shall use this observation in 9.6 to identify the characters of the multiplicative group $\mathbf{F}^*$.

It is easy to see that addition and multiplication are continuous maps from $\mathbf{F} \times \mathbf{F}$ into $\mathbf{F}$. Moreover, the inequality

$$\|(e_0 + x)^{-1} - e_0\| \leq \frac{\|x\|}{1 - \|x\|}$$

holds for all $x \in \mathbf{F}$ satisfying $\|x\| < 1$. Thus the map $x \to x^{-1}$ is continuous from $\mathbf{F}^*$ into $\mathbf{F}$. Therefore the dyadic field is a topological field.

Define projections $\pi_n : \mathbf{F} \to \{0, 1\}$ by

$$\pi_n(x) := \pi_n((x_j, j \in \mathbf{Z})) := x_n \qquad (n \in \mathbf{Z}).$$

Define the integer part of an $x \in \mathbf{F}$ by

$$[x] := (\ldots, x_{-2}, x_{-1}, 0, 0 \ldots),$$

i.e., $[x]$ is that element of $\mathbf{F}$ defined by

$$\pi_n([x]) = \begin{cases} 0 & n \geq 0 \\ x_n & n < 0. \end{cases}$$

Thus $|[x]|$ is an integer in $\mathbf{R}^+$ for each $x \in \mathbf{F}$ and we shall also denote this integer by $[x]$ when no confusion arises.

Characters of the additive group $(\mathbf{F}, +)$ can be generated in the following way. For each $x, y \in \mathbf{F}$ define

$$\psi_y(x) := (-1)^{\pi_{-1}(x \cdot y)}. \tag{9}$$

Since π_{-1} is linear, it is clear that each ψ_y is a character on $(\mathbf{F}, +)$, i.e., ψ_y is continuous on $\mathbf{F}$ and satisfies

$$\psi_y(x + x') = \psi_y(x)\psi_y(x') \qquad (x, x' \in \mathbf{F}).$$

It is also clear that

$$\psi_y(x) = \psi_x(y)$$

and

$$\psi_y(x) = \psi_{[y]}(x)\,\psi_{[x]}(y). \tag{10}$$

In particular, the group of characters of $\mathbf{F}$ is isomorphic to $(\mathbf{F}, +)$.

Concerning identity (10) and possible confusion with the characters of the dyadic group $\mathbf{G}$, we note that

$$\pi_{-1}([x] \cdot y) = \sum_{j=1}^{\infty} x_{-j} y_{j-1} \quad (\text{mod } 2)$$

for $x \in \mathbf{F}$ and $y \in \mathbf{F}_0$. Therefore,

$$\psi_{[x]}(y) = \psi_{|[x]|}(y) \tag{11}$$

for $x \in \mathbf{F}$, $y \in \mathbf{F}_0$ (compare with (17) in 1.2 and (36) in 1.4). In particular, the set $\{\psi_y : y \in \mathbf{F}\}$ can be viewed as an extension of the Walsh functions from the index set $\mathbf{N}$ to the index set $\mathbf{F}$.

The functions ψ_y $(y \in \mathbf{F})$ exhaust the characters of the additive group $(\mathbf{F}, +)$. Indeed, given any character ψ of $(\mathbf{F}, +)$ the condition

$$\psi^2(x) = \psi(x + x) = \psi(0) = 1$$

implies that $\psi(x) = \pm 1$ for each $x \in \mathbf{F}$. Hence given $n \in \mathbf{Z}$ there is a number $y_{-n-1} = 0$ or 1 such that

$$\psi(e_n) = (-1)^{y_{-n-1}}.$$

Since ψ is continuous and $e_n \to 0$ as $n \to \infty$, it follows that $\lim_{n \to -\infty} y_n = 0$. Hence the sequence $y := (y_n, n \in \mathbf{Z})$ belongs to $\mathbf{F}$ and by construction

$$\psi(e_n) = \psi_y(e_n) \qquad (n \in \mathbf{Z}).$$

The continuous functions ψ and ψ_y therefore agree on the usual closed system $(e_n, n \in \mathbf{Z})$. We conclude that $\psi = \psi_y$ everywhere on $\mathbf{F}$.

Differentiation of functions on $\mathbf{F}$ can be defined as follows. For each $n \in \mathbf{P}$ and each function f on $\mathbf{F}$ set

$$\text{(12)} \qquad \mathbf{d}_n f := \sum_{j=-n+1}^{n-1} 2^{j-1}(f - \boldsymbol{\tau}_{e_j} f)$$

where

$$\text{(13)} \qquad (\boldsymbol{\tau}_t f)(x) := f(x+t) \qquad (x, t \in \mathbf{F})$$

represents translation of f by an element $t \in \mathbf{F}$. If

$$f^{[1]}(x) := \lim_{n \to \infty} (\mathbf{d}_n f)(x)$$

exists at some point $x \in \mathbf{F}$, we shall say that f is differentiable at x and call $f^{[1]}(x)$ the pointwise derivative of f at x. Higher order derivatives are defined recursively.

Similarly, if $\boldsymbol{X}$ is some Banach space of functions on $\mathbf{F}$ which contains the additive characters ψ_y $(y \in \mathbf{F})$, and if the limit

$$\mathbf{d}f := \mathbf{d}^{[1]} f := \lim_{n \to \infty} \mathbf{d}_n f$$

exists in the norm of $\boldsymbol{X}$, then we shall say that f is strongly differentiable (in $\boldsymbol{X}$) and call $\mathbf{d}f$ the strong derivative of f.

The additive characters are always strongly differentiable and everywhere differentiable with

$$\text{(14)} \qquad \mathbf{d}\psi_y = \psi_y^{[1]} = |y| \psi_y \qquad (y \in \mathbf{F}).$$

Indeed, by the definitions of $\mathbf{d}_n$ and ψ_y, we have for each $n \in \mathbf{N}$ and $y \in \mathbf{F}$ that

$$\begin{aligned}
(\mathbf{d}_{n+1}\psi_y)(x) &= \sum_{j=-n}^{n} 2^{j-1}(\psi_y(x) - \psi_y(x + e_j)) \\
&= \psi_y(x) \sum_{j=-n}^{n} 2^{j-1}(1 - \psi_y(e_j)) \\
&= \psi_y(x) \sum_{j=-n}^{n} 2^{j-1}(1 - (-1)^{\pi_{-1}(y \cdot e_j)}) \\
&= \psi_y(x) \sum_{j=-n}^{n} 2^{j-1}(1 - (-1)^{y_{-j-1}}) \\
&= \psi_y(x) \sum_{j=-n}^{n} y_{-j-1} 2^j.
\end{aligned}$$

Therefore,

$$|y|\psi_y(x) - (\mathbf{d}_n\psi_y)(x) = \left(\sum_{|j|\geq n} y_{-j-1}2^j\right)\psi_y(x)$$

and (14) follows at once.

Notice that if f depends only on $x_0, x_1, \ldots$ then f can be considered as a function defined on $\mathbf{F}_0$. In this case $\boldsymbol{\tau}_{e_j} f = f$ for $j < 0$ and consequently,

$$\mathbf{d}_n f = \sum_{j=0}^{n-1} 2^{j-1}(f - \boldsymbol{\tau}_{e_j} f).$$

In particular, the difference operator defined before on $\mathbf{G}$ agrees with the one defined above on $\mathbf{F}$.

9.2 The Walsh-Fourier Transform. The additive field $(\mathbf{F}, +)$ is a locally compact abelian group and its unit ball

$$\mathbf{B} := \{x \in \mathbf{F} : \|x\| = 1\}$$

is compact. Hence there is a unique Haar measure μ on $\mathbf{F}$ which satisfies

$$\mu(\mathbf{B}) = 1.$$

The spaces $\mathrm{L}^p_\mu(\mathbf{F})$ will be denoted by $\mathrm{L}^p(\mathbf{F})$ for $1 \leq p \leq \infty$ and the corresponding norm by $\|\cdot\|_p$.

Given $f \in \mathrm{L}^1(\mathbf{F})$ the Walsh-Fourier transform of f is the function on $\mathbf{F}$ defined by

(15) $$\widehat{f}(y) := \int_{\mathbf{F}} f(x)\psi_y(x)\,d\mu(x) \qquad (y \in \mathbf{F}).$$

Thus the map $f \to \widehat{f}$ is a linear map from $\mathrm{L}^1(\mathbf{F})$ into $\mathrm{L}^\infty(\mathbf{F})$. Since

(16) $$\|\widehat{f}\|_\infty \leq \|f\|_1$$

it is also clear that

(17) $$\lim_{n\to\infty} \widehat{f}_n(y) = \widehat{f}(y) \qquad (y \in \mathbf{F})$$

for any sequence $(f_n, n \in \mathbf{N})$ which converges to f in $\mathrm{L}^1(\mathbf{F})$ norm.

The Walsh-Fourier transform of an integrable function is continuous and bounded on $\mathbf{F}$. In fact, THEOREM 1. *If* $f \in L^1(\mathbf{F})$ *then* $\widehat{f}$ *is uniformly continuous on* $\mathbf{F}$.

PROOF. Let $y, t \in \mathbf{F}$. Since

$$\widehat{f}(y+t) - \widehat{f}(y) = \int_{\mathbf{F}} f(x)\psi_y(x)(\psi_t(x) - 1)\, d\mu(x),$$

we can write

$$\begin{aligned} |\widehat{f}(y+t) - \widehat{f}(y)| &\le \int_{\mathbf{F}} |f(x)||\psi_t(x) - 1|\, d\mu(x) \\ &\le \int_{|x|\le R} |f(x)||\psi_t(x) - 1|\, d\mu(x) + 2\int_{|x|>R} |f(x)|\, d\mu(x) \\ &=: I_R + J_R \end{aligned}$$

for any $R > 0$. The Borel measure μ is locally finite on the locally compact space $\mathbf{F}$. Hence given $\varepsilon > 0$ we can choose R so large that

$$J_R < \varepsilon.$$

With R fixed, use the fact that $\{|x| \le R\}$ is compact together with the Lebesgue dominated convergence theorem to see that

$$\lim_{t\to 0} I_R = 0.$$

Thus choose $\delta > 0$ so that $|t| < \delta$ implies

$$I_R < \varepsilon.$$

It follows that

$$|\widehat{f}(y+t) - \widehat{f}(y)| \le 2\varepsilon$$

for all $y \in \mathbf{F}$ and $|t| < \delta$. In particular, f is uniformly continuous on $\mathbf{F}$. ∎

Thus $\widehat{f}$ is a continuous, bounded function on $\mathbf{F}$ which vanishes at ∞ (see (28) below).

Translation and character multiplication exhibit the same kind of duality under the Walsh-Fourier transform as they did under the Walsh-Fourier coefficient map (compare the following result with (46) and (47) in 1.5).

THEOREM 2. *Let* $f \in L^1(\mathbf{F})$ *and* $t \in \mathbf{F}$. *Then*

$$\widehat{\tau_t f} = \psi_t \widehat{f} \tag{18}$$

and

$$\widehat{\psi_t f} = \tau_t \widehat{f}. \tag{19}$$

PROOF. Let $y, t \in \mathbf{F}$. Since μ is translation invariant and ψ_y is an additive character, we have

$$\begin{aligned}\widehat{\tau_t f}(y) &= \int_{\mathbf{F}} f(x+t)\psi_y(x)\,d\mu(x)\\ &= \int_{\mathbf{F}} f(x)\psi_y(x+t)\,d\mu(x)\\ &= \int_{\mathbf{F}} f(x)\psi_y(x)\psi_y(t)\,d\mu(x).\end{aligned}$$

In view of (9), this verifies (18). Similarly,

$$\begin{aligned}\widehat{\psi_t f}(y) &= \int_{\mathbf{F}} \psi_t(x)f(x)\psi_y(x)\,d\mu(x)\\ &= \int_{\mathbf{F}} f(x)\psi_x(y+t)\,d\mu(x)\\ &= \widehat{f}(y+t). \qquad \blacksquare\end{aligned}$$

The convolution of two functions f and g in $L^1(\mathbf{F})$ is defined by

$$(f*g)(x) := \int_{\mathbf{F}} f(x+t)g(t)\,d\mu(t) \qquad (x \in \mathbf{F}). \tag{20}$$

By Fubini's theorem $f*g \in L^1(\mathbf{F})$ and

$$\|f*g\|_1 \le \|f\|_1\,\|g\|_1. \tag{21}$$

Thus $L^1(\mathbf{F})$ is a Banach algebra under function addition and convolution.

The Walsh-Fourier transform takes convolution to pointwise multiplication.

THEOREM 3. *If $f, g \in L^1(\mathbf{F})$, then*

$$\widehat{f*g} = \widehat{f}\,\widehat{g}. \tag{22}$$

PROOF. Let $y \in \mathbf{F}$ and use Fubini's theorem to write

$$\begin{aligned}\widehat{f*g}(y) &= \int_{\mathbf{F}} \left(\int_{\mathbf{F}} f(x+t)g(t)\,d\mu(t)\right)\psi_y(x)\,d\mu(x)\\ &= \int_{\mathbf{F}} \left(\int_{\mathbf{F}} f(x+t)\psi_y(x)\,d\mu(x)\right) g(t)\,d\mu(t).\end{aligned}$$

Since $\psi_y(x) = \psi_y(x+t)\psi_y(t)$ and μ is translation invariant, we conclude that

$$\begin{aligned}\widehat{f*g}(y) &= \int_{\mathbf{F}} \left(\int_{\mathbf{F}} f(\xi)\psi_y(\xi)\,d\mu(\xi)\right)\psi_y(t)g(t)\,d\mu(t)\\ &= \widehat{f}(y)\widehat{g}(y). \qquad \blacksquare\end{aligned}$$

Since $L^\infty(\mathbf{F})$ is a Banach algebra under pointwise addition and multiplication it follows that the Walsh-Fourier transform is a bounded Banach algebra homomorphism from $L^1(\mathbf{F})$ into $L^\infty(\mathbf{F})$. We shall see in 9.4 that this homomorphism is also 1-1.

The following multiplication formula proves useful in a variety of situations.

THEOREM 4. *If* $f, g \in \mathrm{L}^1(\mathbf{F})$ *then*

$$\int_{\mathbf{F}} \widehat{f}(y)g(y)\,d\mu(y) = \int_{\mathbf{F}} f(y)\widehat{g}(y)\,d\mu(y). \tag{23}$$

PROOF. This follows directly from definition, (9), and Fubini's theorem:

$$\begin{aligned}\int_{\mathbf{F}} \widehat{f}(y)g(y)\,d\mu(y) &= \int_{\mathbf{F}} \left(\int_{\mathbf{F}} f(x)\psi_y(x)\,d\mu(x)\right) g(y)\,d\mu(y)\\ &= \int_{\mathbf{F}} \left(\int_{\mathbf{F}} g(y)\psi_x(y)\,d\mu(y)\right) f(x)\,d\mu(x)\\ &= \int_{\mathbf{F}} f(x)\widehat{g}(x)\,d\mu(x). \qquad \blacksquare\end{aligned}$$

These results do not use the dyadic structure in an essential way. For any locally compact group G whose group of characters is $\widehat{G}$ and whose Haar measure is ν, the Fourier transform of an $f \in \mathrm{L}^1_\nu(G)$ can be defined by

$$\widehat{f}(\gamma) = \int_G f(x)\overline{\gamma}(x)\,d\nu(x) \qquad (\gamma \in \widehat{G})$$

where $\overline{\gamma}$ represents the complex conjugate of γ. It is not difficult to verify (see Rudin [1] or Hewitt and Ross [1], for example) that (16), (17), and Theorems 1 through 4 are satisfied for any such Fourier transform. We will prove other such results in 9.3. However, most of our theory (in particular, the rest of this section which concerns itself with dyadic differentiation) relies heavily on the dyadic structure and frequently has no general analogue for locally compact groups.

Before continuing, recall that $\mathbf{F}$ contains two other important groups, the unit ball $\mathbf{B}$ and the multiplicative group $\mathbf{F}^*$. We shall identify their characters in 9.6 and study the Fourier transform induced by $\mathbf{F}^*$, the so-called Mellin transform.

The following result shows what the Walsh-Fourier transform does to the strong L^1 derivative. This property (see (66) in 1.7) was used to motivate the definition and is essentially equivalent to it.

THEOREM 5. *Suppose* f *is strongly differentiable in* $\mathrm{L}^1(\mathbf{F})$. *Then*

$$\widehat{\mathbf{d}f}(y) = |y|\widehat{f}(y) \qquad (y \in \mathbf{F}). \tag{24}$$

PROOF. By hypothesis

$$\lim_{n\to\infty} \|\mathbf{d}_n f - \mathbf{d}f\|_1 = 0.$$

Hence by (17)

$$\widehat{\mathbf{d}f}(y) = \lim_{n\to\infty} \widehat{\mathbf{d}_n f}(y).$$

But (12) and (18) imply

$$\widehat{\mathbf{d}_{n+1}f}(y) = \sum_{j=-n}^{n} 2^{j-1}(\widehat{f}(y) - \psi_{e_j}(y)\widehat{f}(y))$$
$$= \widehat{f}(y)\left(\sum_{j=-n}^{n} 2^{j-1}(1 - \psi_{e_j}(y))\right)$$

for $n \in \mathbf{N}$. Consequently, (24) is obtained by letting $n \to \infty$. ∎

Thus the Walsh-Fourier transform takes a derivative of f to a polynomial multiple of $\widehat{f}$. The following result shows that the Walsh-Fourier transform takes a polynomial multiple of f to a derivative of $\widehat{f}$.

THEOREM 6. *Let $f \in \mathrm{L}^1(\mathbf{F})$ and for each $x \in \mathbf{F}$ set $g(x) := |x| f(x)$. If $g \in \mathrm{L}^1(\mathbf{F})$ then $\widehat{f}$ is pointwise differentiable on $\mathbf{F}$ and*

(25) $$(\widehat{f}\,)^{[1]}(y) = \widehat{g}(y) \qquad (y \in \mathbf{F}).$$

PROOF. Set

(26) $$\alpha_n(x) := \sum_{j=-n}^{n} 2^{j-1}(1 - \psi_{e_j}(x)) = \sum_{j=-n}^{n} \frac{x_{j-1}}{2^j}$$

for $x \in \mathbf{F}$, $n \in \mathbf{N}$. By definition, $\alpha_n(x) \to |x|$ as $n \to \infty$.

Fix $y \in \mathbf{F}$, $n \in \mathbf{N}$ and observe by definition that

$$(\mathbf{d}_{n+1}\widehat{f})(y) = \sum_{j=-n}^{n} 2^{j-1}\left(\int_{\mathbf{F}} f(x)\psi_y(x)\,d\mu(x) - \int_{\mathbf{F}} f(x)\psi_{y+e_j}(x)\,d\mu(x)\right)$$
$$= \int_{\mathbf{F}} f(x)\psi_y(x)\alpha_n(x)\,d\mu(x).$$

Thus we can write

$$|(\mathbf{d}_{n+1}\widehat{f})(y) - \widehat{g}(y)| = \left|\int_{\mathbf{F}} f(x)\psi_y(x)(\alpha_n(x) - |x|)\,d\mu(x)\right|$$
$$= \left|\int_{\mathbf{F}} |x| f(x)\psi_y(x)\left(\frac{\alpha_n(x)}{|x|} - 1\right) d\mu(x)\right|.$$

The integrand is dominated by $2|g(x)|$ and converges to zero as $n \to \infty$. We conclude by the Lebesgue dominated convergence theorem that

$$(\widehat{f}\,)^{[1]}(y) := \lim_{n\to\infty} (\mathbf{d}_n\widehat{f})(y) = \widehat{g}(y). \quad ∎$$

It is sometimes convenient to work on $\mathbf{R}^+$ instead of $\mathbf{F}$. The identification $x \to |x|$ takes Haar measure μ to Lebesgue measure on $\mathbf{R}^+$, the characters of $(\mathbf{F}, +)$ to generalized

Walsh functions on $\mathbf{R}^+$, differentiation to a dyadic differentiation, and induces a dyadic sum and product on $\mathbf{R}^+$ in exactly the same way that Fine's map carried the dyadic group structure to the unit interval $[0, 1)$. Details are similar to those in 1.3 and are left to the reader.

We shall denote dyadic addition by $\dot{+}$, but leave all other notation the same. In particular, we shall denote the generalized Walsh functions by ψ_x $(x \in \mathbf{R}^+)$. Thus by (11) in 9.1 and the identification above it is evident that

$$\psi_k(x) = w_k(x)$$

for $k \in \mathbf{N}$ and $x \in [0, 1)$. Also, it is clear that

$$\psi_y(x \dot{+} t) = \psi_y(x)\psi_y(t)$$

for $x, t \in \mathbf{R}^+$, and $x \dot{+} t$ dyadic irrational, that

$$\psi_x(y) = \psi_y(x),$$

$$\psi_y^{[1]}(t) = y\psi_y(t),$$

and

$$\psi_y(x) = \psi_{[y]}(x)\psi_{[x]}(y),$$

for $x, y, t \in \mathbf{R}^+$,where for $u \in \mathbf{R}^+$, $[u]$ represents the greatest integer in u.

The functions $(\psi_j, j \in \mathbf{N})$ form a complete orthonormal system in each interval of the form $[k, k+1)$, $k = 0, 1, \ldots$. Indeed, since $x = [x] \dot{+} (x - [x])$ it is clear that

$$\psi_j(x) = \psi_j([x])\, w_j(x - [x])$$

for $j \in \mathbf{N}$ and $x \in \mathbf{R}^+$. For $x \in [k, k+1)$ it follows that $\psi_j(x) = w_j(x - [x])$. In particular, we see that ψ_j is a periodic extension of w_j from $[0, 1)$ to $\mathbf{R}^+$.

Fix $k \in \mathbf{N}$. It is easy to see that the integral over $[k, k+1)$ of any generalized Walsh function must be 0, 1 or -1. Indeed

$$\int_k^{k+1} \psi_y(x)\, dx = \int_k^{k+1} \psi_{[y]}(x)\psi_{[x]}(y)\, dx = \psi_k(y) \int_k^{k+1} \psi_{[y]}(x)\, dx.$$

Therefore

$$\int_k^{k+1} \psi_y(x)\, dx = \begin{cases} w_k(y) & 0 \le y < 1 \\ 0 & y \ge 1. \end{cases} \tag{27}$$

The Walsh-Fourier transform of an $f \in \mathrm{L}^1(\mathbf{R}^+)$ is defined by

$$\widehat{f}(y) := \int_0^\infty f(x)\psi_y(x)\, dx \qquad (y \in \mathbf{R}^+).$$

It is easy to see that the results derived above have dyadic versions valid for this Walsh-Fourier transform as well. In particular, $\widehat{f}$ is a W-continuous L^∞ function on $\mathbf{R}^+$.

It is now also easy to see that a version of the Riemann-Lebesgue lemma holds for the Walsh-Fourier transform, namely,

$$\lim_{y\to\infty} \widehat{f}(y) = 0 \tag{28}$$

for all $f \in L^1(\mathbf{R}^+)$. Indeed, fix $y \in \mathbf{R}^+$, let $\varepsilon > 0$ and write

$$\widehat{f}(y) = \int_0^n f(x)\psi_y(x)\,dx + \int_n^\infty f(x)\psi_y(x)\,dx.$$

Choose $n \in \mathbf{N}$ so large that

$$\int_n^\infty |f(x)|\,dx < \varepsilon.$$

Notice that

$$\begin{aligned}\int_0^n f(x)\psi_y(x)\,dx &= \sum_{k=0}^{n-1} \int_k^{k+1} f(x)\psi_y(x)\,dx \\ &= \sum_{k=0}^{n-1} \psi_k(y) \int_k^{k+1} \psi_{[y]}(x)f(x)\,dx.\end{aligned}$$

In particular, $\widehat{f}(y)$ is dominated by ε plus a fixed sum of Walsh-Fourier coefficients of order $[y]$. Since each of these Walsh-Fourier coefficients tend to zero as $y \to \infty$, we conclude that (28) holds as promised.

9.3 The Walsh-Fourier-Plancherel Transform. In this section we extend the Walsh-Fourier transform from $L^1(\mathbf{R}^+) \cap L^p(\mathbf{R}^+)$ to $L^p(\mathbf{R}^+)$ for $1 < p \le 2$.

For each $k \in \mathbf{N}$ define a function Ω_k on $\mathbf{R}^+$ by

$$\Omega_k(x) := \begin{cases} \psi_k(x) & x \in [k, k+1) \\ 0 & \text{otherwise.} \end{cases}$$

The Walsh-Fourier transform of Ω_k is easy to compute:

$$\begin{aligned}\widehat{\Omega}_k(y) &= \int_{\mathbf{R}^+} \Omega_k(x)\psi_y(x)\,dx \\ &= \int_k^{k+1} \psi_k(x)\psi_{[y]}(x)\psi_{[x]}(y)\,dx \\ &= \psi_k(y) \int_k^{k+1} \psi_k(x)\psi_{[y]}(x)\,dx \\ &= \begin{cases} \psi_k(y) & y \in [k, k+1) \\ 0 & \text{otherwise.} \end{cases}\end{aligned}$$

Consequently,

$$\widehat{\Omega}_k = \Omega_k \qquad (k \in \mathbf{N}).$$

In particular, the functions Ω_k $(k \in \mathbf{N})$ are eigenfunctions of the Walsh-Fourier transform.

For each $k \in \mathbf{N}$ set

$$\Omega_{k,0,1} := \Omega_k$$

and

$$\Omega_{k,0,-1} := 0.$$

For $n \in \mathbf{P}$, $k \in \mathbf{N}$, and $j = \pm 1$ define a function $\Omega_{k,n,j}$ on $\mathbf{R}^+$ by

(29)
$$\Omega_{k,n,j}(x) := \begin{cases} \frac{1}{\sqrt{2}}\psi_{k+n}(x) & x \in [k, k+1) \\ \frac{j}{\sqrt{2}}\psi_k(x) & x \in [k+n, k+n+1) \\ 0 & \text{otherwise.} \end{cases}$$

Notice that

$$\begin{aligned} \widehat{\Omega}_{k,n,j}(y) &= \int_{\mathbf{R}^+} \Omega_{k,n,j}(x)\psi_y(x)\,dx \\ &= \int_k^{k+1} \frac{1}{\sqrt{2}}\psi_{k+n}(x)\psi_{[y]}(x)\psi_{[x]}(y)\,dx \\ &\qquad + j\int_{k+n}^{k+n+1} \frac{1}{\sqrt{2}}\psi_k(x)\psi_{[y]}(x)\psi_{[x]}(y)\,dx \\ &= \frac{1}{\sqrt{2}}\psi_k(y)\int_k^{k+1} \psi_{k+n}(x)\psi_{[y]}(x)\,dx \\ &\qquad + \frac{j}{\sqrt{2}}\psi_{k+n}(y)\int_{k+n}^{k+n+1} \psi_k(x)\psi_{[y]}(x)\,dx \\ &= \begin{cases} \frac{j}{\sqrt{2}}\psi_{k+n}(x) & y \in [k, k+1) \\ \frac{1}{\sqrt{2}}\psi_k(x) & y \in [k+n, k+n+1) \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Hence

(30)
$$\widehat{\Omega}_{k,n,j} = j\Omega_{k,n,j}$$

for $k, n \in \mathbf{N}$, and $j \in \{+1, -1\}$. Therefore the system

(31)
$$\mathbf{\Omega} := \{\Omega_{k,n,j} : n, k \in \mathbf{N}, j \in \{+1, -1\}\}$$

is orthonormal in $\mathrm{L}^2(\mathbf{R}^+)$ and consists entirely of eigenfunctions of the Walsh-Fourier transform.

This system is complete. For if $g \in L^2(\mathbf{R}^+)$ and

$$\int_{\mathbf{R}^+} g(x)\Omega_{k,n,j}(x)\,dx = 0 \qquad (k, n \in \mathbf{N}, j = \pm 1)$$

then

$$\int_k^{k+1} g(x)\psi_\ell(x)\,dx = 0 \qquad (k, \ell \in \mathbf{N}).$$

Since $(\psi_\ell, \ell \in \mathbf{N})$ is complete on each interval $[k, k+1)$ it follows that $g = 0$ a.e. on $[k, k+1)$ for $k = 0, 1, \dots$ Therefore, $\mathbf{\Omega}$ is a complete, orthonormal system in $L^2(\mathbf{R}^+)$.

For each $f \in L^2(\mathbf{R}^+)$ let

$$c_{k,n,j}(f) := \int_{\mathbf{R}^+} f(x)\Omega_{k,n,j}(x)\,dx \tag{32}$$

represent the $\mathbf{\Omega}$-Fourier coefficients of f. Define the Walsh-Fourier-Plancherel transform of f to be the formal Walsh- Fourier transform of the $\mathbf{\Omega}$-Fourier series of f, that is,

$$\mathcal{F}f := \sum_{k,n \in \mathbf{N}} \left(c_{k,n,1}(f)\,\widehat{\Omega}_{k,n,1} + c_{k,n,-1}(f)\,\widehat{\Omega}_{k,n,-1}\right). \tag{33}$$

This defines a function $\mathcal{F}f$ for each $f \in L^2(\mathbf{R}^+)$. In fact,

THEOREM 7. *If $f \in L^2(\mathbf{R}^+)$ then (33) converges in L^2 norm. Moreover,*

$$\|\mathcal{F}f\|_2 = \|f\|_2 \tag{34}$$

and

$$\mathcal{F}(\mathcal{F}f) = f. \tag{35}$$

PROOF. Since $\mathbf{\Omega}$ is a complete orthonormal system in $L^2(\mathbf{R}^+)$ it is clear that

$$f = \sum_{k,n \in \mathbf{N}} \left(c_{k,n,1}(f)\,\Omega_{k,n,1} + c_{k,n,-1}(f)\,\Omega_{k,n,-1}\right) \tag{36}$$

where this series converges in the $L^2(\mathbf{R}^+)$ norm. Moreover, by (30) we have

$$\sum_{k,n \in \mathbf{N}} \left(c_{k,n,1}(f)\,\widehat{\Omega}_{k,n,1} + c_{k,n,-1}(f)\,\widehat{\Omega}_{k,n,-1}\right)$$
$$= \sum_{k,n \in \mathbf{N}} \left(c_{k,n,1}(f)\,\Omega_{k,n,1} - c_{k,n,-1}(f)\,\Omega_{k,n,-1}\right).$$

Consequently, it follows from the Riesz-Fischer theorem that (33) converges in $L^2(\mathbf{R}^+)$ norm and

$$\|f\|_2^2 = \sum_{k,n \in \mathbf{N}} \left(|c_{k,n,1}(f)|^2 + |c_{k,n,-1}(f)|^2\right) = \|\mathcal{F}f\|_2^2.$$

This verifies (34) and shows that

$$\mathcal{F}f = \sum_{k,n\in\mathbf{N}} (c_{k,n,1}(f)\,\Omega_{k,n,1} - c_{k,n,-1}(f)\,\Omega_{k,n,-1})$$

for all $f \in \mathrm{L}^2(\mathbf{R}^+)$. In particular, (36) implies $\mathcal{F}(\mathcal{F}f) = f$ for all $f \in \mathrm{L}^2(\mathbf{R}^+)$. ∎

Thus we see that $\mathcal{F}$ is an isometric Banach space isomorphism of $\mathrm{L}^2(\mathbf{R}^+)$ onto itself.

To obtain a closed form for the Walsh-Fourier-Plancherel transform, define the generalized Dirichlet kernel by

$$D_t(y) := \int_0^t \psi_y(x)\,dx \tag{37}$$

for $t, y \in \mathbf{R}^+$. For integer values of t, the generalized Dirichlet kernel is a zero extension of the Walsh-Dirichlet kernels defined in 1.5. Indeed, if $n \in \mathbf{N}$ then (27) implies

$$\begin{aligned} D_n(y) &= \int_0^n \psi_y(x)\,dx \\ &= \sum_{k=0}^{n-1} \int_k^{k+1} \psi_y(x)\,dx \\ &= \begin{cases} \sum_{k=0}^{n-1} w_k(y) & 0 \le y < 1 \\ 0 & y \ge 1. \end{cases} \end{aligned}$$

THEOREM 8. *If* $f \in \mathrm{L}^2(\mathbf{R}^+)$ *then*

$$(\mathcal{F}f)(t) = \frac{d}{dt}\left(\int_{\mathbf{R}^+} f(y)D_t(y)\,dy\right) \tag{38}$$

for a.e. $t \in \mathbf{R}^+$.

PROOF. Fix $t \in \mathbf{R}^+$ and observe by definition that

$$\begin{aligned} &\int_0^t (\mathcal{F}f)(x)\,dx \\ &\qquad = \sum_{k,n\in\mathbf{N}} \left(c_{k,n,1}(f) \int_0^t \widehat{\Omega}_{k,n,1}(x)\,dx + c_{k,n,-1}(f) \int_0^t \widehat{\Omega}_{k,n,-1}(x)\,dx \right). \end{aligned}$$

Since by Fubini's theorem

$$\int_0^t \widehat{\Omega}_{k,n,j}(x)\,dx = \int_{\mathbf{R}^+} \Omega_{k,n,j}(y)D_t(y)\,dy$$

we can express the indefinite integral of $\mathcal{F}f$ as an $\mathbf{\Omega}$-Fourier series of D_t. Thus it follows from Bessel's equality that

$$\int_0^t (\mathcal{F}f)(x)\,dx = \int_{\mathbf{R}^+} f(y)D_t(y)\,dy.$$

We conclude by differentiating with respect to t that (38) holds for a.e. $t \in \mathbf{R}^+$. ∎

It is now easy to see that the Walsh-Fourier-Plancherel transform is an extension of the Walsh-Fourier transform:

COROLLARY 1. *If* $f \in L^1(\mathbf{R}^+) \cap L^2(\mathbf{R}^+)$ *then*

$$\mathcal{F}f = \widehat{f}. \tag{39}$$

PROOF. By the Lebesgue dominated convergence theorem, interchange the derivative and integral which appear on the right side of (38). We have by (37) that

$$\begin{aligned}(\mathcal{F}f)(t) &= \int_{\mathbf{R}^+} f(y)\frac{d}{dt}D_t(y)\,dy \\ &= \int_{\mathbf{R}^+} f(y)\psi_y(t)\,dy \\ &= \widehat{f}(t)\end{aligned}$$

for a.e. $t \in \mathbf{R}^+$. ∎

We can also obtain the Walsh-Fourier-Plancherel transform as a limit of truncated Walsh-Fourier transforms:

COROLLARY 2. *For* $f \in L^2(\mathbf{R}^+)$ *and* $t \in \mathbf{R}^+$ *define a function* f_t *on* $\mathbf{R}^+$ *by*

$$f_t(x) := \begin{cases} f(x) & x \in [0,t) \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\mathcal{F}f = \lim_{t\to\infty} \widehat{f_t} \tag{40}$$

in $L^2(\mathbf{R}^+)$ *norm.*

PROOF. Since $f_t \to f$ in $L^2(\mathbf{R}^+)$ it follows from (34) that $\mathcal{F}(f_t) \to \mathcal{F}(f)$ in $L^2(\mathbf{R}^+)$ as $t \to \infty$. Since $f_t \in L^1(\mathbf{R}^+) \cap L^2(\mathbf{R}^+)$ for each $t \in \mathbf{R}^+$, we conclude by Corollary 1 that $\widehat{f_t} \to \mathcal{F}f$ in $L^2(\mathbf{R}^+)$ as $t \to \infty$. ∎

Thus we see that

$$(\mathcal{F}f)(y) = \underset{t\to\infty}{l.i.m.} \int_0^t f(x)\psi_y(x)\,dx \tag{41}$$

for $f \in L^2(\mathbf{R}^+)$, where $l.i.m.$ stands for "limit in the mean" and represents convergence in the $L^2(\mathbf{R}^+)$ norm.

To extend the Walsh-Fourier transform to $L^p(\mathbf{R}^+)$ for $1 < p < 2$, set

$$L^1(\mathbf{R}^+) + L^2(\mathbf{R}^+) := \{f_1 + f_2 : f_i \in L^i(\mathbf{R}^+),\ i = 1 \text{ and } 2\}. \tag{42}$$

For any $f = f_1 + f_2 \in L^1(\mathbf{R}^+) + L^2(\mathbf{R}^+)$ define

$$\widehat{f} := \widehat{f_1} + \mathcal{F}f_2. \tag{43}$$

This definition does not depend on the representation of f. Indeed, if $f = f_1 + f_2 = g_1 + g_2$ for some f_i, $g_i \in \mathrm{L}^i(\mathbf{R}^+)$, $i = 1$ and 2, then $f_1 - g_1 \in \mathrm{L}^1(\mathbf{R}^+)$, $g_2 - f_2 \in \mathrm{L}^2(\mathbf{R}^+)$, and

$$f_1 - g_1 = g_2 - f_2.$$

Hence by Corollary 1, $\widehat{f_1} - \widehat{g_1} = \mathcal{F}g_2 - \mathcal{F}f_2$ so

$$\widehat{f_1} + \mathcal{F}f_2 = \widehat{g_1} + \mathcal{F}g_2.$$

Recall that $\mathrm{L}^p(\mathbf{R}^+) \subset \mathrm{L}^1(\mathbf{R}^+) + \mathrm{L}^2(\mathbf{R}^+)$ for any $1 < p < 2$. Indeed, given $f \in \mathrm{L}^p(\mathbf{R}^+)$ write $f = f_1 + f_2$ where

$$f_2(x) := \begin{cases} f(x) & |f(x)| \le 1 \\ 0 & |f(x)| > 1. \end{cases}$$

Since $u \le u^p$ when $u \ge 1$ and $t^2 \le t^p$ when $0 < t \le 1$, it is clear that

$$\int_{\mathbf{R}^+} |f_1| \le \int_{\mathbf{R}^+} |f|^p < \infty$$

and

$$\int_{\mathbf{R}^+} |f_2|^2 \le \int_{\mathbf{R}^+} |f|^p < \infty.$$

Therefore $\mathrm{L}^p(\mathbf{R}^+) \subset \mathrm{L}^1(\mathbf{R}^+) + \mathrm{L}^2(\mathbf{R}^+)$ and (43) defines the Walsh-Fourier transform of any $f \in \mathrm{L}^p(\mathbf{R}^+)$ for $1 < p < 2$.

The following result answers the question of integrability of $\widehat{f}$ so defined. Inequality (44) is called Hausdorff-Young inequality.

THEOREM 9. *Let $1 \le p \le 2$ and q be the index conjugate to p. If $f \in \mathrm{L}^p(\mathbf{R}^+)$ then $\widehat{f} \in \mathrm{L}^q(\mathbf{R}^+)$ and*

$$\|\widehat{f}\|_q \le \|f\|_p. \tag{44}$$

PROOF. By (16) in 9.2 the Walsh-Fourier transform is of type $(1, \infty)$. By (34) above the Walsh-Fourier transform is of type $(2, 2)$. The constant in both cases is 1. Therefore, the Riesz- Thorin theorem implies that the Walsh-Fourier transform is of type (p, q) with constant 1, i.e., that (44) holds for all $f \in \mathrm{L}^p(\mathbf{R}^+)$. ∎

9.4 Inversion of the Walsh-Fourier Transform. For each $f \in \mathrm{L}^1(\mathbf{R}^+)$ and each $t > 0$ define the Walsh-Dirichlet integral by

$$(S_t f)(x) := \int_0^t \widehat{f}(y)\psi_y(x)\,dy \qquad (x \in \mathbf{R}^+). \tag{45}$$

By Fubini's theorem it is clear that

$$(S_t f)(x) = \int_{\mathbf{R}^+} f(u) \left(\int_0^t \psi_y(x \dotplus u)\,dy \right) du.$$

In view of (37) we see that

(46) $$S_t f = f * D_t$$

for $t \in \mathbf{R}^+$ and $f \in \mathrm{L}^1(\mathbf{R}^+)$. Thus the generalized Dirichlet kernel D_t plays the same role for the Walsh-Dirichlet integral that the Walsh-Dirichlet kernels do for partial sums of Walsh-Fourier series.

Recall from 9.3 that for integer values of t, the generalized Dirichlet kernel is a zero extension of the Walsh-Dirichlet kernel outside $[0,1)$, namely

$$D_t(y) = \begin{cases} \sum_{k=0}^{n-1} w_k(y) & y \in [0,1) \\ 0 & y \in [1,\infty) \end{cases}$$

for $t = n \in \mathbf{N}$. To see what happens when $t \notin \mathbf{N}$ set

$$\boldsymbol{J}_0^{(1)}(t) := t - [t]$$

and

$$\boldsymbol{J}_\ell^{(1)}(t) := \int_0^t \psi_\ell(x)\,dx \qquad (\ell \in \mathbf{P}, t \in \mathbf{R}^+)$$

(compare with (48) in 1.5). Notice by (10) in 9.1 that

$$\int_{[t]}^t \psi_y(x)\,dx = \psi_{[t]}(y)\boldsymbol{J}_{[y]}^{(1)}(t).$$

Consequently, by definition

(47) $$D_t(y) = D_{[t]}(y) + \psi_{[t]}(y)\boldsymbol{J}_{[y]}^{(1)}(t)$$

for $y, t \in \mathbf{R}^+$.

The generalized Dirichlet kernel satisfies many of the same properties as the Walsh-Dirichlet kernel. For example, since

$$|\boldsymbol{J}_{[y]}^{(1)}(t)| \le 2^{-n-1}$$

for $2^n \le y < 2^{n+1}$, $n \in \mathbf{N}$, and $t \in \mathbf{R}^+$, it follows from (47) and Theorem 10 in 1.6 that

(48) $$|D_t(y)| \le \frac{3}{y} \qquad (y, t \in \mathbf{R}^+).$$

Also, we have by Paley's lemma that

$$D_{2^n}(y) = \begin{cases} 2^n & y \in [0, 2^{-n}) \\ 0 & y \in [2^{-n}, \infty), \end{cases}$$

and by (47) that

$$D_{2^{-n}}(y) = \boldsymbol{J}_{[y]}^{(1)}(2^{-n}) = \begin{cases} 2^{-n} & y \in [0, 2^n) \\ 0 & y \in [2^n, \infty), \end{cases}$$

for $n \in \mathbf{N}$. Therefore,

(49) $$D_{2^n} = 2^n \chi[0, 2^{-n}) \qquad (n \in \mathbf{Z}).$$

Combining (49) with (46), it is now evident that

$$(S_{2^n} f)(x) = 2^n \int_0^{2^{-n}} f(x \dotplus u)\,du$$

for $n \in \mathbf{Z}$ and $x \in \mathbf{R}^+$. Hence the following is true.

COROLLARY 3. *If* $f \in L^1(\mathbf{R}^+)$ *then*

i)
$$\lim_{n\to\infty} \|S_{2^n} f - f\|_1 = 0$$

ii)
$$\lim_{n\to\infty} S_{2^n} f = f \qquad \text{a.e. on} \quad \mathbf{R}^+$$

and

iii)
$$\lim_{n\to\infty} (S_{2^{-n}} f)(x) = 0 \qquad \text{for all} \quad x \in \mathbf{R}^+$$

In particular, if $f, g \in L^1(\mathbf{R}^+)$ *and* $\widehat{f} = \widehat{g}$ *a.e. then* $f = g$ *a.e. on* $\mathbf{R}^+$.

In sharp contrast to iii), the L^1 norm of the Walsh-Dirichlet integrals of order 2^{-n} converge to $|\widehat{f}(0)|$:

THEOREM 10. *If* $f \in L^1(\mathbf{R}^+)$ *then*

$$\lim_{n\to\infty} \|S_{2^{-n}} f\|_1 = \left| \int_{\mathbf{R}^+} f(x)\, dx \right|.$$

In particular,

$$\lim_{n\to\infty} \|S_{2^{-n}} f\|_1 = 0$$

if and only if $\widehat{f}(0) = 0$.

PROOF. For each $n \in \mathbf{N}$ let χ_n represent the characteristic function of the interval $[0, 2^n)$.

Consider the function

$$T_n f := S_{2^{-n}} f - 2^{-n} \chi_n \int_0^{2^n} f(x)\, dx.$$

By (46) we have

$$T_n f = 2^{-n} \left(f * \chi_n - \chi_n \int_0^{2^n} f(x)\, dx \right).$$

Since $\|f * g\|_1 \le \|f\|_1 \|g\|_1$ for $f, g \in L^1(\mathbf{R}^+)$, it follows that

(50)
$$\|T_n f\|_1 \le \|f\|_1 + \left| \int_0^{2^n} f(x)\, dx \right| \le 2\|f\|_1$$

for $n \in \mathbf{N}$.

Let g be any dyadic step function on $\mathbf{R}^+$. Choose $N \in \mathbf{N}$ such that

$$\{g \ne 0\} \subset [0, 2^N).$$

Then $n \geq N$ implies

$$S_{2^{-n}}g = 0 \qquad \text{on} \quad [2^n, \infty)$$

and

$$S_{2^{-n}}g = 2^{-n}\int_0^{2^n} g \qquad \text{on} \quad [0, 2^n).$$

Thus $T_n g = 0$ for large n. Since the dyadic step functions with compact support are dense in $L^1(\mathbf{R}^+)$ it follows that

$$\lim_{n\to\infty} \|T_n f\|_1 = 0$$

for all $f \in L^1(\mathbf{R}^+)$. Consequently the proof of the theorem is complete. ∎

Theorem 10 is a special case of a general theorem from martingale theory. Indeed, since $S_{2^{-n}}f$ is the conditional expectation of f with respect to the σ-algebra generated by the sets $[k2^n, (k+1)2^n)$, $k, n \in \mathbf{N}$, the sequence $(S_{2^{-n}}f, n \in \mathbf{N})$ is a reverse martingale.

We close this section with some inversion results for the Walsh- Fourier transform.

THEOREM 11. *Let $f \in L^1(\mathbf{R}^+)$ be W- continuous on $\mathbf{R}^+$. If $\widehat{f} \in L^1(\mathbf{R}^+)$ then*

$$f(t) = \int_{\mathbf{R}^+} \widehat{f}(y)\psi_t(y)\,dy \tag{51}$$

for all $t \in \mathbf{R}^+$.

PROOF. Let $\varepsilon > 0$, fix $t \in \mathbf{R}^+$, and choose by hypothesis an integer $n > 0$ such that

$$|f(t \dot{+} u) - f(t)| < \varepsilon$$

for all $u \in \mathbf{R}^+$ which satisfy $|u| < 2^{-n}$. Clearly,

$$\begin{aligned}\left|\int_0^{2^n} \widehat{f}(y)\psi_t(y)\,dy - f(t)\right| &= |(S_{2^n}f)(t) - f(t)| \\ &\leq 2^n \int_0^{2^{-n}} |f(t \dot{+} u) - f(t)|\,du \\ &< \varepsilon.\end{aligned}$$

Therefore

$$\lim_{n\to\infty} \int_0^{2^n} \widehat{f}(y)\psi_t(y)\,dy = f(t).$$

But

$$\left|\int_{2^n}^{\infty} \widehat{f}(y)\psi_t(y)\,dy\right| \leq \int_{2^n}^{\infty} |\widehat{f}(y)|\,dy \to 0$$

as $n \to \infty$ since $\widehat{f}$ is integrable. Therefore, (51) holds for all $t \in \mathbf{R}^+$. ∎

For a dyadic field analogue of Cesàro summability we introduce the Walsh-Fejér integrals

$$(\sigma_u f)(x) := \frac{1}{u}\int_0^u (S_t f)(x)\,dt$$

for $x, u \in \mathbf{R}^+$ and $f \in \mathrm{L}^1(\mathbf{R}^+)$. Integrating by parts, we have by the definition of S_t that

$$(52) \qquad (\sigma_u f)(x) = \int_0^u \left(1 - \frac{t}{u}\right) \widehat{f}(t)\psi_t(x)\, dt$$

for $x, u \in \mathbf{R}^+$ and $f \in \mathrm{L}^1(\mathbf{R}^+)$.

Define the generalized Fejér kernel by

$$\mathcal{K}_u(y) := \frac{1}{u}\int_0^u D_t(y)\, dt \qquad (y, u \in \mathbf{R}^+).$$

(We have used a different notation for the generalized Fejér kernel because it is not simply a zero extension of the Walsh- Fejér kernels introduced in 1.8 (see (56) below).) By definition, (48), and Fubini's theorem we have

$$\begin{aligned}(\sigma_u f)(x) &= \frac{1}{u}\int_0^u (f * D_t)(x)\, dt \\ &= \int_0^\infty \mathcal{K}_u(y) f(x \dot{+} y)\, dy.\end{aligned}$$

Thus

$$\sigma_u f = f * \mathcal{K}_u \qquad (u \in \mathbf{R}^+, f \in \mathrm{L}^1(\mathbf{R}^+)).$$

We will estimate the $\mathcal{K}_u$'s and show $\sigma_u f \to f$ in $\mathrm{L}^1(\mathbf{R}^+)$ norm as $u \to \infty$.

First introduce the second integrals

$$\boldsymbol{J}_\ell^{(2)}(t) := \int_0^t \boldsymbol{J}_\ell^{(1)}(x)\, dx \qquad (\ell \in \mathbf{N}, t \in [0, 1))$$

and extend the $\boldsymbol{J}_\ell^{(2)}$'s to $\mathbf{R}^+$ by periodicity of period 1. Notice that

$$\boldsymbol{J}_{2^n+k}^{(1)} = w_k \boldsymbol{J}_{2^n}^{(1)} \qquad (0 \le k < 2^n, n \in \mathbf{N}).$$

It follows that

$$(53) \qquad \boldsymbol{J}_{2^n+2^m+k}^{(2)}(1) = 0$$

and

$$(54) \qquad \sup_{t \in \mathbf{R}^+} |\boldsymbol{J}_{2^n+2^m+k}^{(2)}(t)| = 2^{-n-m-3}$$

hold for all $0 \le k < 2^m, 0 \le m < n$, and

$$(55) \qquad \boldsymbol{J}_{2^n}^{(2)}(1) = \sup_{t \in \mathbf{R}^+} |\boldsymbol{J}_{2^n}^{(2)}(t)| = 2^{-n-2}$$

for $n \in \mathbf{N}$. Let

$$f_0(y) := \sum_{j=0}^{\infty} 2^{-j} \sum_{i=0}^{j} D_{2^{-i}}(y \dotplus 2^{j-1}) \qquad (y \in \mathbf{R}^+).$$

Then $f_0(y) \geq 1$ for $y \in [0,1)$, $f_0(y) \geq 2^{-n-1}$ for $y \in [2^n, 2^n+1)$, and $f_0(y) \geq 2^{-n-m-2}$ for $y \in [2^n + 2^m, 2^n + 2^{m+1})$ and all $0 \leq m < n,\ m, n \in \mathbf{N}$. Thus by (54) and (55) we obtain

$$\|\boldsymbol{J}^{(2)}_{[y]}\|_\infty \leq \frac{f_0(y)}{2} \qquad (y \in \mathbf{R}^+). \tag{56}$$

Next extend the Walsh-Fejér kernels $(K_\ell, \ell \in \mathbf{P})$ to $\mathbf{R}^+$ by

$$K_\ell := \frac{1}{\ell}(D_1 + D_2 + \ldots D_\ell)$$

(i.e., $K_\ell := 0$ on $[1, \infty)$ for $\ell \in \mathbf{P}$). By definition and (47),

$$\begin{aligned} u\mathcal{K}_u(y) &= \sum_{k=0}^{[u]-1} \int_k^{k+1} D_t(y)\,dt + \int_{[u]}^{u} D_t(y)\,dt \\ &= \sum_{k=0}^{[u]-1} \left(D_k(y) + \psi_k(y)\boldsymbol{J}^{(2)}_{[y]}(1) \right) \\ &\qquad + (u - [u])D_{[u]}(y) + \psi_{[u]}(y)\boldsymbol{J}^{(2)}_{[y]}(u) \end{aligned}$$

for $u, y \in \mathbf{R}^+$. In particular, it follows from (56) that

$$|\mathcal{K}_u(y)| \leq |K_{[u]}(y)| + 2f_0(y) \qquad (y \in \mathbf{R}^+) \tag{57}$$

for $u \geq 1$.

THEOREM 12. *For every* $f \in \mathrm{L}^1(\mathbf{R}^+)$,

$$\lim_{u \to \infty} \|\sigma_u f - f\|_1 = 0. \tag{58}$$

PROOF. By Theorem 16 in 1.8

$$\|K_{[u]}\|_1 \leq 2 \qquad (u \geq 1)$$

and by Paley's lemma

$$\|f_0\|_1 \leq 4.$$

Thus it follows from (57) that

$$\|\mathcal{K}_u\|_1 \leq 10 \qquad (u \geq 1). \tag{59}$$

Fix $n \in \mathbf{N}$ and set $f := \chi[0,1)w_n$. By (10) in 9.1 we have

$$\widehat{f}(y) = \int_0^1 w_n(x)\psi_y(x)\,dx = \int_0^1 w_{n\oplus[y]}(x)\,dx$$

for all $y \in \mathbf{R}^+$. Consequently, $\widehat{f} = \chi[n, n+1)$. Therefore, it follows from (52) and inversion that

$$\begin{aligned} f(x) - (\sigma_u f)(x) &= \int_n^{n+1} \psi_x(t)\,dt - \int_n^{n+1} \left(1 - \frac{t}{u}\right)\psi_x(t)\,dt \\ &= \frac{1}{u}\int_n^{n+1} t\psi_x(t)\,dt \\ &= \frac{\psi_n(x)}{u}\int_0^1 (n+t)\psi_{[x]}(t)\,dt \end{aligned}$$

for $x \in \mathbf{R}^+$ and $u \geq 1$. In particular, if

$$g(t) := n + t \qquad (t \in [0,1))$$

then the Walsh-Fourier coefficients of g satisfy

$$\|\sigma_u f - f\|_1 = \frac{1}{u}\|\widehat{g}\|_{\ell^1}.$$

We conclude that (58) holds for $f = \chi[0,1)w_n$.

By translation invariance, (58) holds for every f in the closed system

$$\{\chi[m, m+1)w_n : m, n \in \mathbf{N}\}.$$

But the operators σ_u $(u \geq 1)$ are uniformly bounded on $L^1(\mathbf{R}^+)$ since by (59)

$$\|\sigma_u f\|_1 = \|f * \mathcal{K}_u\|_1 \leq \|f\|_1 \, \|\mathcal{K}_u\|_1 \leq 10\|f\|_1.$$

We conclude that (58) holds for all $f \in L^1(\mathbf{R}^+)$. ∎

9.5 The Inverse Dyadic Derivative. For each $n \in \mathbf{Z}$ define a function W_n by specifying that its Walsh-Plancherel transform satisfies

(60) $$(\mathcal{F}W_n)(y) := \widetilde{W}_n(y) := \begin{cases} 0 & y \in [0, 2^{-n}) \\ 1/y & y \in [2^{-n}, \infty). \end{cases}$$

Notice since $\mathcal{F}W_n \in L^2(\mathbf{R}^+)$ that (60) uniquely defines $W_n \in L^2(\mathbf{R}^+)$ (see Theorem 7 in 9.3).

The W_n's are also integrable.

THEOREM 13. *For each $n \in \mathbf{Z}$,*

$$W_n \in \mathrm{L}^1(\mathbf{R}^+) \cap \mathrm{L}^2(\mathbf{R}^+).$$

Moreover,

$$W_n(x) = \lim_{k\to\infty} \int_{2^{-n}}^{2^k} \frac{1}{y}\psi_x(y)\,dy$$

for $n \in \mathbf{Z}$ and a.e. $x \in \mathbf{R}^+$, where this limit exists both pointwise and in the L^1 norm.

PROOF. Fix $n \in \mathbf{Z}$ and set

$$W_{n,m}(x) := \int_{2^{-n}}^{2^m} \frac{1}{y}\psi_x(y)\,dy$$

for $x \in \mathbf{R}^+$, $m \in \mathbf{Z}$. Let $k \in \mathbf{N}$ with $k > m$ and integrate by parts twice to see

$$\begin{aligned} W_{n,k}(x) - W_{n,m}(x) &= \int_{2^m}^{2^k} \frac{1}{y}\psi_x(y)\,dy \\ &= \frac{D_{2^k}(x)}{2^k} - \frac{D_{2^m}(x)}{2^m} + \frac{\mathcal{K}_{2^k}(x)}{2^k} \\ &\quad - \frac{\mathcal{K}_{2^m}(x)}{2^m} + 2\int_{2^m}^{2^k} \frac{1}{y^2}\mathcal{K}_y(x)\,dy. \end{aligned}$$

Since $\|D_{2^j}\|_1 = 1$ for $j \in \mathbf{Z}$ and $\|\mathcal{K}_u\|_1 \le 14$ for $u \ge 1$ (see (49) and (59) in 9.4) it follows that $W_{n,k} - W_{n,m}$ and $W_{n,k}$ are integrable with

$$\|W_{n,k} - W_{n,m}\|_1 = O(2^{-m})$$

uniformly in n, as $k, m \to \infty$. Therefore, for each $n \in \mathbf{Z}$ there is an L^1 function U_n such that

$$U_n = \lim_{m\to\infty} W_{n,m} \tag{61}$$

in L^1 norm. Since by (48) in 9.4

$$|D_v(x)| \le \frac{3}{x}$$

and consequently

$$|\mathcal{K}_v(x)| \le \frac{3}{x} \qquad (x, v > 0)$$

it also follows that this limit exists pointwise on $\mathbf{R}^+$. In particular, (61) holds for a.e. $x \in \mathbf{R}$ and it remains to show

$$U_n = W_n \qquad (n \in \mathbf{Z}).$$

But by Corollary 2 in 9.3 and (35) we have

$$W_n(x) = (\mathcal{F}\widetilde{W_n})(x) = \underset{m\to\infty}{\text{l.i.m.}} \int_{2^{-n}}^{2^m} \frac{1}{y}\psi_x(y)\,dy = \underset{m\to\infty}{\text{l.i.m.}} W_{n,m}(x).$$

Since $W_{n,m} \to U_n$ a.e. as $m \to \infty$, we conclude that $U_n = W_n$ a.e. ∎

The functions $(W_n, n \in \mathbf{N})$ provide a kernel for dyadic integration. Specifically, if given $f \in \mathrm{L}^1(\mathbf{R}^+)$ there exists a function $g \in \mathrm{L}^1(\mathbf{R}^+)$ such that

$$\lim_{n\to\infty} \|W_n * f - g\|_1 = 0$$

then g is called the (strong) dyadic integral of f and will be denoted by $\mathbf{I}f$.

The following characterization of the dyadic integral is a transform analogue of Corollary 1 in 5.2.

THEOREM 14. *Suppose* $f, g \in \mathrm{L}^1(\mathbf{R}^+)$. *Then*

$$g = \mathbf{I}f$$

if and only if

$$\widehat{g}(y) = \begin{cases} 0 & y = 0 \\ \widehat{f}(y)/y & y > 0. \end{cases}$$

PROOF. Suppose $g = \mathbf{I}f$, i.e.,

$$\lim_{n\to\infty} \|W_n * f - g\|_1 = 0.$$

Then by (17) and Theorem 3 in 9.2 we have

$$\widehat{g}(y) = \lim_{n\to\infty} \widehat{W_n}(y)\widehat{f}(y) = \begin{cases} 0 & y = 0 \\ \widehat{f}(y)/y & y > 0. \end{cases}$$

Conversely, suppose $\widehat{g}(y) = (1/y)\widehat{f}(y)$ for $y > 0$ and $\widehat{g}(0) = 0$. Notice that $\widehat{\chi}[0,1) = \chi[0,1)$ and by a change of variables that

(62) $$\widehat{D}_{2^{-n}} = \chi[0, 2^{-n}) \qquad (n \in \mathbf{Z}).$$

Thus by hypothesis the Walsh-Fourier transforms of

$$W_n * f - W_m * f \quad \text{and} \quad S_{2^{-n}}g - S_{2^{-m}}g$$

are identical for all $n, m \in \mathbf{Z}$. Hence

$$\|W_n * f - W_m * f\|_1 = \|S_{2^{-n}}g - S_{2^{-m}}g\|_1.$$

But $\widehat{g}(0) = 0$ so it follows from Theorem 10 in 9.4 that $(W_n * f, n \in \mathbf{N})$ is a Cauchy sequence in L^1. Therefore there is a function $h \in \mathrm{L}^1(\mathbf{R}^+)$ such that $W_n * f \to h$ in L^1 norm. By definition

$$h = \mathbf{I}f.$$

Hence by what we have already proved

$$\widehat{h}(y) = \begin{cases} 0 & y = 0 \\ \widehat{f}(y)/y & y > 0. \end{cases}$$

Thus $\widehat{h}(y) = \widehat{g}(y)$ for $y \in \mathbf{R}^+$ and it follows from Corollary 3 in 9.4 that $g = h = \mathbf{I}f$. ∎

This leads easily to a partial fundamental theorem of dyadic calculus (see also Theorem 16 and Corollary 6 below).

Corollary 4. *If $f \in L^1(\mathbf{R}^+)$ is strongly dyadically differentiable and $\widehat{f}(0) = 0$ then*

$$f = \mathbf{I}(\mathbf{d}f).$$

Proof. By hypothesis and Theorem 5 in 9.2 we have

$$\widehat{f}(y) = \begin{cases} 0 & y = 0 \\ \widehat{\mathbf{d}f}(y)/y & y > 0. \end{cases}$$

Therefore the conclusion follows at once from Theorem 14. ∎

It turns out that classical integrals of dyadically integrable functions behave in a special way.

Corollary 5. *If f is dyadically integrable then*

$$\int_0^{2^n} f = o(2^{-n})$$

as $n \to \infty$.

Proof. Let $n \in \mathbf{Z}$ and apply Theorem 4 in 9.2 to f and $g = \chi[0, 2^{-n})$. In view of (62) we have

$$\int_0^{2^n} f = 2^n \int_0^{2^{-n}} \widehat{f}.$$

On the other hand, if $h = \mathbf{I}f$ then by Theorem 14 we have

$$\widehat{f}(y) = y\widehat{h}(y) \qquad (y > 0).$$

It follows that

$$2^n \int_0^{2^n} f = 2^{2n} \int_0^{2^{-n}} y\widehat{h}(y)\, dy$$

for all $n \in \mathbf{Z}$. Since $\widehat{h}(0) = 0$ and $\widehat{h}$ is W-continuous, we conclude that

$$2^n \int_0^{2^n} f \to 0$$

as $n \to \infty$. ∎

It is clear that not every $f \in L^1(\mathbf{R}^+)$ is dyadically integrable. A concrete example is given by $f := \chi[0, 1)$. Indeed, $\widehat{f} = f$ and

$$(W_n * f)\widehat{\ }(y) = \begin{cases} 0 & y \in [0, 2^{-n}) \\ 1/y & y \in [2^{-n}, 1) \\ 0 & y \in [1, \infty) \end{cases}$$

for $n \in \mathbf{N}$. Hence by inversion we have for a.e. $x \in [0,1)$ that

$$\begin{aligned}(W_n * f)(x) &= \int_{2^{-n}}^{1} \frac{1}{y}\psi_x(y)\,dy \\ &= \int_{2^{-n}}^{1} \frac{1}{y}\,dy \\ &= n\log 2.\end{aligned}$$

Therefore $W_n * f$ fails to converge, as $n \to \infty$, a.e. on $\mathbf{R}^+$ and f is not dyadically integrable.

In sharp contrast, the function $\chi[0,1)w_m$ is always dyadically integrable for each $m \in \mathbf{P}$. In fact,

THEOREM 15. *Let*

$$\boldsymbol{Y} := \{\chi[0,2^n)w_{m2^{-n}} : n \in \mathbf{N},\, m \in \mathbf{P}\}.$$

Then the linear hull of $\boldsymbol{Y}$ *is* L^1 *dense in the set of dyadically integrable functions.*

PROOF. Fix $m \in \mathbf{P}$. By repeating the argument above with $\chi[0,1)w_m$ in place of $\chi[0,1)$ we see that $\chi[0,1)w_m$ is dyadically integrable. Thus $\boldsymbol{Y}$ consists entirely of dyadically integrable functions.

Let f be dyadically integrable and $\varepsilon > 0$. Choose an $N \in \mathbf{N}$ such that

$$\int_{2^N}^{\infty} |f| < \varepsilon.$$

Set

$$f = f_1 + f_2 := \chi[0,2^N)f + \chi[2^N,\infty)f$$

and observe that $\|f_2\|_1 < \varepsilon$. Also notice that

$$P_n := S_{2^n}f_1 - 2^{-N}\chi[0,2^N)\int_0^{2^N} f$$

belongs to the linear hull of $\boldsymbol{Y}$ for all $n \geq -N$ since $\widehat{P}_n(0) = 0$.

Clearly,

$$\begin{aligned}\|f - P_n\|_1 &\leq \|f_1 - P_n\|_1 + \|f_2\|_1 \\ &\leq \|f_1 - S_{2^n}f_1\|_1 + \|2^{-N}\chi[0,2^N)\int_0^{2^N} f\|_1 + \|f_2\|_1 \\ &< \|f_1 - S_{2^n}f_1\|_1 + 2\varepsilon.\end{aligned}$$

Moreover, $S_{2^n}f_1 \to f_1$ in L^1 norm as $n \to \infty$ (see Corollary 3 in 9.4). We conclude that $\|P_n - f\|_1 < 3\varepsilon$ for all $n \in \mathbf{N}$ large enough. ∎

For the remainder of this section we shall investigate the strong and pointwise dyadic differentiability of the dyadic integral $\mathbf{I}f$. We begin by describing the difference operator $\mathbf{d}_n(\mathbf{I}f)$ (see (12) in 9.1) in terms of the function α_n introduced by (26) in 9.2.

LEMMA 1. *Suppose $f \in L^1(\mathbf{R}^+)$ is dyadically integrable. Then*

$$\mathbf{d}_n(\mathbf{I}f) = \mathbf{d}_n W_n * f$$

and

$$(\mathbf{d}_n W_n * f)(x) = \lim_{k\to\infty} \int_0^{2^k} \frac{\alpha_{n-1}(y)}{y} \widehat{f}(y)\psi_x(y)\, dy \tag{63}$$

for $n \in \mathbf{N}$, where the limit exists both in the L^1 norm and for a.e. $x \in \mathbf{R}^+$.

PROOF. We have by definition that

$$\begin{aligned}(\mathbf{d}_n W_{m,k})(x) &= \frac{1}{2} \sum_{j=-n+1}^{n-1} \int_{2^{-m}}^{2^k} \frac{2^j}{y} \left(\psi_x(y) - \psi_{x\dot{+}2^{-(j+1)}}(y)\right) dy \\ &= \int_{2^{-m}}^{2^k} \frac{\alpha_{n-1}(y)}{y} \psi_x(y)\, dy\end{aligned}$$

for $x \in \mathbf{R}^+$, $m \geq n$, and $k \in \mathbf{N}$. Since α_n vanishes on $[0, 2^{-n})$ it follows that

$$\mathbf{d}_n W_{m,k} = \mathbf{d}_n W_{n,k} = \widehat{V}_{n,k} \tag{64}$$

for any $m \geq n$ and $k \in \mathbf{N}$, where

$$V_{n,k}(y) := \frac{\alpha_n(y)}{y} \chi[2^{-n}, 2^k)(y) \qquad (n, k \in \mathbf{N},\ y \in \mathbf{R}^+).$$

By the proof of Theorem 13 above,

$$\mathbf{d}_n W_m = \lim_{k\to\infty} \mathbf{d}_n W_{m,k}$$

in L^1 norm and a.e. Since translation is an isometry in the L^1 norm, it follows that

$$\begin{aligned}\mathbf{d}_n(W_m * f) &= \mathbf{d}_n W_m * f \\ &= (\lim_{k\to\infty} \mathbf{d}_n W_{m,k}) * f \\ &= \lim_{k\to\infty} (\mathbf{d}_n W_{m,k} * f) = \mathbf{d}_n W_n * f\end{aligned}$$

for $m \geq n$. Consequently,

$$\mathbf{d}_n(\mathbf{I}f) = \lim_{n\to\infty} \mathbf{d}_n(W_m * f) = \mathbf{d}_n W_n * f.$$

It remains to verify (63). By construction and (21) in 9.2, it suffices to show

$$(\mathbf{d}_n W_{n,k} * f)(x) = \int_0^\infty V_{n,k}\psi_x \widehat{f}$$

for $k \in \mathbf{P}$ and a.e. $x \in \mathbf{R}^+$. But by Theorem 4 in 9.2 and by (19),(20) and (64),

$$\begin{aligned}\int_0^\infty V_{n,k}\psi_x\widehat{f} &= \int_0^\infty (V_{n,k}\psi_x)\widehat{\ } f \\ &= \int_0^\infty \tau_x \widehat{V}_{n,k} f \\ &= \int_0^\infty \tau_x \mathbf{d}_n W_{n,k} f = (\mathbf{d}_n W_{n,k} * f)(x)\end{aligned}$$

as promised. ∎

To estimate $\mathbf{d}_n W_n$ we introduce functions β_n on $\mathbf{R}^+$ by

$$\beta_n(y) := \sum_{j=-n}^{0} y_{j-1} 2^{-j}$$

for $n \in \mathbf{N}$ and $y \in \mathbf{R}^+$. Notice once and for all that

$$2^n \alpha_n(2^{-n} y) = \beta_{2n}(y) \tag{65}$$

for $n \in \mathbf{N}$ and $y \in \mathbf{R}^+$. We also introduce function f_n, g_n defined by

$$f_n(x) := 2^{-n} \sum_{j=-\infty}^{n} 2^j \sum_{i=j}^{n} D_{2^i}(x \dot{+} 2^{-j-1}) \tag{66}$$

and

$$g_n(x) := \sum_{j=-\infty}^{n} 2^j \sum_{i=n}^{\infty} 2^{-i} D_{2^i}(x \dot{+} 2^{-j-1})$$

for $n \in \mathbf{Z}$ and $x \in \mathbf{R}^+$. (The function f_0 was introduced above (56) in 9.4.) Clearly,

$$\|f_n\|_1 \le \sum_{j=-\infty}^{n} (n-j+1) 2^{-n+j} = 4$$

and

$$\|g_n\|_1 \le \sum_{j=-\infty}^{n} 2^{-n+j+1} = 4$$

for all $n \in \mathbf{Z}$.

LEMMA 2. *For each $n \in \mathbf{N}$ and $x \in \mathbf{R}^+$ let*

$$V_n(x) := \lim_{k \to \infty} \int_0^{2^k} \frac{\beta_n(y)}{y} \psi_x(y)\, dy.$$

Then V_n exists everywhere and

$$(\mathbf{d}_{n+1}W_{n+1})(x) = 2^{-n}V_{2n}(2^{-n}x)$$

for $n \in \mathbf{N}$ and a.e. $x \in \mathbf{R}^+$. Moreover,

$$|V_n| \le 10f_0 + g_0 + \chi[0,1)|\mathbf{d}_n W|$$

and

$$\|V_n\|_1 = O(1)$$

as $n \to \infty$.

PROOF. Fix $n \in \mathbf{N}$. By repeating the proof of Theorem 13 it is easy to see V_n exists everywhere on $\mathbf{R}^+$. Moreover, by using the change of variables $y = u2^n$ we have by (65) that

$$\begin{aligned}(\mathbf{d}_{n+1}W_{n+1})(x) &= \lim_{k\to\infty}\int_0^{2^k}\frac{\alpha_n(u)}{u}\psi_x(u)\,du\\ &= \lim_{k\to\infty} 2^{-n}\int_0^{2^{k+n}}\frac{\alpha_n(2^{-n}y)}{2^{-n}y}\psi_x(2^{-n}y)\,dy\\ &= 2^{-n}\lim_{k\to\infty}\int_0^{2^k}\frac{\beta_{2n}(y)}{y}\psi_{2^{-n}x}(y)\,dy\\ &= 2^{-n}V_{2n}(2^{-n}x)\end{aligned}$$

for $x \in \mathbf{R}^+$.

To estimate V_n, notice first that

$$\beta_n(y) = k$$

for $y \in [i2^n + k, i2^n + k + 1)$, $i, n \in \mathbf{N}$, $0 \le k < 2^n$. Next define $0 \le \langle j\rangle_n < 2^n$ for each $j \in \mathbf{N}$ by

$$j \equiv \langle j\rangle_n \pmod{2^n}.$$

Thus (compare with (21) in 6.2)

$$\begin{aligned}V_n(x) &= \sum_{j=1}^{\infty}\langle j\rangle_n\, w_j(x)\int_j^{j+1}\frac{\psi_{[x]}(y)}{y}\,dy \qquad (67)\\ &= \sum_{j=1}^{\infty}\langle j\rangle_n\, w_j(x)A_j(x) + \left(\chi[0,1)\sum_{j=1}^{\infty}\frac{\langle j\rangle_n\, w_j}{j}\right)(x)\\ &=: V_n^{(1)}(x) + (\chi[0,1)\mathbf{d}_n W)(x)\end{aligned}$$

where

$$A_j(x) := \int_j^{j+1}\left(\frac{1}{y}-\frac{1}{j}\right)\psi_{[x]}(y)\,dy$$

for $j \in \mathbf{P}$ and $x \in \mathbf{R}^+$, and

$$W := 1 + \sum_{k=1}^{\infty} \frac{w_k}{k}.$$

For each $j \in \mathbf{P}$ we have from integration by parts that

$$\begin{aligned} A_j(x) &= \int_j^{j+1} \frac{1}{y^2} \left(\boldsymbol{J}_{[x]}^{(1)}(y) - \boldsymbol{J}_{[x]}^{(1)}(1) \right) dy \\ &= \int_j^{j+1} \left(\frac{1}{y^2} - \frac{1}{j^2} \right) \left(\boldsymbol{J}_{[x]}^{(1)}(y) - \boldsymbol{J}_{[x]}^{(1)}(1) \right) dy + \frac{1}{j^2} \left(\boldsymbol{J}_{[x]}^{(2)}(1) - \boldsymbol{J}_{[x]}^{(1)}(1) \right) \\ &=: A_j^1(x) + A_j^2(x). \end{aligned}$$

Using (56) and the mean value theorem for integrals we see that for a suitable choice of $\xi \in (j, j+1)$ one has

$$\begin{aligned} |A_j^1(x)| &= \left(\frac{1}{j^2} - \frac{1}{(j+1)^2} \right) \left| \int_\xi^{j+1} (\boldsymbol{J}_{[x]}^{(1)}(y) - \boldsymbol{J}_{[x]}^{(1)}(1))\, dy \right| \\ &\le \frac{4}{j^3} f_0(x). \end{aligned}$$

Similarly, using (56) and the fact that

$$|\boldsymbol{J}_{[x]}^{(1)}(1)| \le \frac{f_0(x)}{2}$$

we can show

$$|A_j^2(x)| \le \frac{1}{j^2} f_0(x).$$

Therefore,

$$|V_n^1(x)| \le 4 f_0(x) \sum_{j=1}^{\infty} \frac{\langle j \rangle_n}{j^3} + |\sum_{j=1}^{2^n-1} \frac{w_j(x)}{j}|\, |\boldsymbol{J}_{[x]}^{(2)}(1) - \boldsymbol{J}_{[x]}^{(1)}(1)| + f_0(x) \sum_{j=2^n}^{\infty} \frac{\langle j \rangle_n}{j^2}$$

for $x \in \mathbf{R}^+$. Since

$$\sum_{j=1}^{\infty} \frac{\langle j \rangle_n}{j^3} < \sum_{j=1}^{\infty} \frac{1}{j^2} < 2$$

and

$$\sum_{j=2^n}^{\infty} \frac{\langle j \rangle_n}{j^2} = \sum_{i=1}^{\infty} \sum_{k=0}^{2^n-1} \frac{k}{(i2^n + k)^2} \le \sum_{i=1}^{\infty} \frac{2^{2n}}{i^2 2^{2n}} < 2$$

we have

$$|V_n^1(x)| \le 10 f_0(x) + |\sum_{j=1}^{2^n-1} \frac{w_j(x)}{j}|\, |\boldsymbol{J}_{[x]}^{(2)}(1) - \boldsymbol{J}_{[x]}^{(1)}(1)|.$$

But by (55) and the fact that $\boldsymbol{J}_\ell^{(1)}(1) = 0$ for $\ell \in \mathbf{P}$ it is clear that

$$|\boldsymbol{J}_{[x]}^{(2)}(1) - \boldsymbol{J}_{[x]}^{(1)}(1)| \le \frac{1}{2}D_1(x) + \frac{1}{4}\sum_{k=0}^{\infty} 2^{-k} D_1(x \dot{+} 2^k).$$

Moreover, Theorem 15 in 1.7 implies

$$|\sum_{j=1}^{2^n-1} \frac{w_j}{j}| \le \sum_{k=0}^{\infty} 2^{-k+2}\overline{D}_{2^k}$$

for $n \in \mathbf{N}$, where $\overline{D}_{2^k}$ is the periodic extension of D_{2^k} from $[0,1)$ to $\mathbf{R}^+$. In particular, it follows from the definition of g_0 that

$$|\sum_{j=1}^{2^n-1} \frac{w_j(x)}{j}|\,|\boldsymbol{J}_{[x]}^{(2)}(1) - \boldsymbol{J}_{[x]}^{(1)}(1)| \le g_0(x)$$

for $n \in \mathbf{N}$ and $x \in \mathbf{R}^+$.

We have proved that $|V_n^1| \le 10 f_0 + g_0$. By (67) this implies

$$|V_n| \le 10 f_0 + g_0 + \chi[0,1)|\mathbf{d}_n W|.$$

Furthermore, it is clear by (22), (23) and (24) in 6.2 that

$$\int_0^1 |\mathbf{d}_n W| = O(1)$$

as $n \to \infty$. We conclude by (66) that $\|V_n\|_1 = O(1)$ as $n \to \infty$. ∎

These estimates imply that every dyadic integral is strongly dyadically differentiable.

THEOREM 16. *If $f \in \mathrm{L}^1(\mathbf{R}^+)$ is dyadically integrable then $\mathbf{I}f$ is strongly dyadically differentiable in L^1 and*

$$\mathbf{d}(\mathbf{I}f) = f.$$

PROOF. By Lemma 1 we need to show $d_n W_n * f \to f$ as $n \to \infty$ for all dyadically integrable f. Define linear operators T_n on $\mathrm{L}^1(\mathbf{R}^+)$ by

$$T_n f := \mathbf{d}_{n+1} W_{n+1} * f \qquad (n \in \mathbf{N}, f \in \mathrm{L}^1(\mathbf{R}^+)). \tag{68}$$

By Lemma 2 these operators are uniformly bounded:

$$\begin{aligned} \|T_n\| &\le \|\mathbf{d}_{n+1} W_{n+1}\|_1 \\ &= \int_0^\infty 2^{-n} |V_{2n}(2^{-n}x)|\, dx \\ &= \int_0^\infty |V_{2n}(u)|\, du \\ &= O(1) \end{aligned}$$

as $n \to \infty$. Hence by Theorem 15 it suffices to show

$$\lim_{n\to\infty} \|T_n f - f\|_1 = 0$$

for

(69) $$f := \chi[0, 2^s)w_{m2^{-s}} \qquad (s \in \mathbf{N},\, m \in \mathbf{P}).$$

Fix $n, s \in \mathbf{N}$ and $m \in \mathbf{P}$. Set

$$\gamma(u) := \frac{1}{m+u},$$

$$\gamma_n(u) := \alpha_n(u2^{-s}) - u2^{-s} \qquad (0 \le u < 1),$$

and

$$F_n(x) := \left(\int_0^1 \gamma_n(u)\gamma(u)\psi_{[x]}(u)\,du\right)\psi_m(x) \qquad (x \in \mathbf{R}^+).$$

Notice that

$$\begin{aligned}\|F_n\|_{L^1(\mathbf{R}^+)} &= \sum_{k=0}^{\infty} |\widehat{\gamma\gamma_n}(k)| \\ &= \|\gamma_n\gamma\|_{\mathrm{A}} \\ &\le \|\gamma_n\|_{\mathrm{A}}\|\gamma\|_{\mathrm{A}}\end{aligned}$$

(see (45) in 1.5 and the opening remarks of 2.4). Since

$$\alpha_n(y) - y = -\sum_{k=n+1}^{\infty} \frac{y_{k-1}}{2^k} = \sum_{k=n+1}^{\infty} \frac{r_{k-1}(y)}{2^{k+1}} - \frac{1}{2^{n+1}}$$

for $y \le 2^s$ and $n > s$, we have by definition that

$$\gamma_n(u) = \sum_{k=n+1}^{\infty} \frac{r_{k-s-1}(u)}{2^{k+1}} - \frac{1}{2^{n+1}}$$

for $0 \le u \le 1$. Consequently,

$$\|\gamma_n\|_{\mathrm{A}} = \frac{1}{2^n}.$$

On the other hand, since γ is twice continuously differentiable in the classical sense (see Exercise 2.13) we have

$$\|\gamma\|_{\mathrm{A}} < \infty.$$

Therefore

(70) $$\lim_{n\to\infty} \|F_n\|_{L^1(\mathbf{R}^+)} = 0.$$

To finish the proof notice that (69) implies

$$\widehat{f} = 2^s \chi[\frac{m}{2^s}, \frac{m+1}{2^s}).$$

Hence we have by (63) that

$$(T_n f)(x) - f(x) = 2^s \int_{m/2^s}^{(m+1)/2^s} \frac{\alpha_n(y) - y}{y} \psi_x(y)\,dy$$

for $n \in \mathbf{N}$, $x \in \mathbf{R}^+$. Make the substitution $u = y2^s$ and apply (10) in 9.1. It follows that

$$(T_n f)(x) - f(x) = 2^s \int_{m}^{(m+1)} \frac{\alpha_n(u2^{-s}) - u2^{-s}}{u} \psi_{[x]}(u)\psi_{[u]}(x)\,du = 2^s F_n(x).$$

We conclude by (70) that $T_n f \to f$ in L^1 norm as $n \to \infty$. ∎

To investigate the pointwise dyadic differentiability of the dyadic integral we introduce the maximal function

$$T^* f := \sup_{n \in \mathbf{N}} |T_n f| \qquad (f \in \mathrm{L}^1(\mathbf{R}^+))$$

where the operators T_n were defined in (68) above.

Notice by Lemma 2 that the operator $f \to T^* f$ is of type (∞, ∞). The following result shows T^* is of weak type $(1,1)$.

THEOREM 17. *There is a constant $A > 0$ such that*

$$|\{T^* f > y\}| \le A \frac{\|f\|_1}{y}$$

for all $f \in \mathrm{L}^1(\mathbf{R}^+)$ and all $y > 0$.

PROOF. For $n \in \mathbf{Z}$ and any function h defined on $\mathbf{R}^+$ let us use the notation

$$h^{\langle n \rangle}(x) := 2^{-n} h(x 2^{-n}) \qquad (x \in \mathbf{R}^+).$$

Let f_n, g_n be defined by (66) and recall that

$$\|f_n\|_1, \quad \|g_n\|_1 \le 4 \qquad (n \in \mathbf{Z}).$$

Notice for $i, j \in \mathbf{Z}$ that

$$2^{-n} D_{2^i}(2^{-n}x \dotplus 2^{-(j+1)}) = D_{2^{i-n}}(x \dotplus 2^{-(j-n)-1})$$

for all $x \in \mathbf{R}^+$. Consequently,

$$g_m^{\langle n \rangle} = g_{m-n}, \quad f_m^{\langle n \rangle} = f_{m-n} \qquad (m, n \in \mathbf{Z}).$$

Moreover, for $2^{s-1} \le \ell < 2^s$, $s, \ell \in \mathbf{P}$, it follows from Theorem 16 in 1.8 that

$$|K_\ell| \le \sum_{j=0}^{s-1} 2^{j-s} \sum_{i=j}^{s-1} (D_{2^i} + \tau_{2^{-j-1}} D_{2^i})$$
$$\le \sum_{i=0}^{s} 2^{i-s} D_{2^i} + f_s.$$

By (22),(23), and (24) in 6.2, therefore,

$$\chi[0,1)|\mathbf{d}_n W| \le 4(D_{2^n} + \sum_{s=0}^{n} 2^{s-n} D_{2^s}) + 8(\sum_{s=0}^{n} 2^{s-n} f_s + f_n).$$

Since $D_{2^s}^{\langle n\rangle} = D_{2^{s-n}}$, we conclude by Lemma 2 that

$$|V_{2n}^{\langle n\rangle}| \le 10 f_{-n} + g_{-n} + 4(D_{2^n} + \sum_{s=0}^{2n} 2^{s-2n} D_{2^{s-n}}) + 8(f_n + \sum_{s=0}^{2n} 2^{s-2n} f_{s-n}).$$

For each $h \in \mathrm{L}^1(\mathbf{R}^+)$ let

$$E^* h := \sup_{i \in \mathbf{Z}} |D_{2^i} * h|,$$
$$G^* h := \sup_{i \in \mathbf{Z}} |g_i * h|,$$

and

$$F^* h := \sup_{i \in \mathbf{Z}} |f_i * h|$$

We have proved that

$$T^* f \le 12 E^*|f| + G^*|f| + 34 F^*|f|.$$

Now, the concept of quasi-locality for operators introduced in 6.2 can be extended from [0,1) to $\mathbf{R}^+$. It is not difficult to verify that E^*, F^*, and G^* are all quasi-local. Moreover, Theorem 4 in 6.2 has an analogue for $\mathbf{R}^+$. In particular, it follows that E^*, F^* and G^* are weak type $(1,1)$. Thus so is T^*. ∎

COROLLARY 6. *Suppose $f \in \mathrm{L}^1(\mathbf{R}^+)$ is dyadically integrable. Then $\mathbf{I}f$ is a.e. dyadically differentiable on $\mathbf{R}^+$ and*

$$(\mathbf{I}f)^{[1]} = f.$$

PROOF. Let $\boldsymbol{Z}$ represent the linear hull of the space $\boldsymbol{Y}$ defined in Theorem 15. Let $\boldsymbol{X}$ denote the collection of dyadically integrable functions in $\mathrm{L}^1(\mathbf{R}^+)$. By Theorem 15, $\boldsymbol{Z}$ is dense in $\boldsymbol{X}$. By the proof of Theorem 16, $T_n f \to f$ a.e. as $n \to \infty$ for every $f \in \boldsymbol{Z}$. And, by Theorem 17 the operator $f \to T^* f$ is of weak type $(1,1)$. We conclude by Theorem 2 in 3.1 that $T_n f \to f$ a.e. for all $f \in \boldsymbol{X}$. ∎

We also notice by the Marcinkiewicz interpolation theorem that T^* is of type (p,p) for $1 < p \le \infty$. Hence there is a constant $A_p > 0$ such that

$$\| \sup_{n \in \mathbf{N}} |\mathbf{d}_n W_n * f| \|_p \le A_p \|f\|_p$$

for $f \in \mathrm{L}^p(\mathbf{R}^+)$ and $1 < p \le \infty$.

9.6 The Mellin Transform. The Fourier transform induced by the locally compact abelian (multiplicative) group $\mathbf{F}^* := \mathbf{F} \setminus \{0\}$ is called the Mellin transform. Let ν be the multiplicatively invariant Borel measure on $\mathbf{F}^*$ which satisfies $\nu(\mathbf{B}) = 1$, i.e., let ν be a Haar measure on $\mathbf{F}^*$. Then the Mellin transform of an $f \in \mathrm{L}^1_\nu := \mathrm{L}^1_\nu(\mathbf{F}^*)$ is given by

$$\mathcal{M}f(\gamma) := \int_{\mathbf{F}^*} f(x)\overline{\gamma}(x)\, d\nu(x)$$

for each character γ on $\mathbf{F}^*$. Being a Fourier transform, $\mathcal{M}$ satisfies multiplicative analogues of (16),(17), and Theorems 1 through 4 above (see the remarks following Theorem 4 in 9.2). The purpose of this section is to describe a method for generating the characters of $\mathbf{F}^*$, obtain an explicit formula for $\mathcal{M}f$ using the identification of $\mathbf{F}^*$ with $\mathbf{Z} \times \mathbf{B}$, and show inversion holds for square integrable functions.

Since $\mathbf{F}^*$ is algebraically isomorphic to $\mathbf{Z} \times \mathbf{B}$ (see (8) in 9.1), the character group $\widehat{\mathbf{F}^*}$ must be isomorphic to $\widehat{\mathbf{Z}} \times \widehat{\mathbf{B}}$. Thus we begin by identifying the multiplicative characters of $\mathbf{B}$. For this it is important to realize that every element of $\mathbf{B}$ has a product representation.

LEMMA 3. *To each $x \in \mathbf{B}$ there corresponds a unique sequence of numbers $y_j = 0$ or 1 for $j \in \mathbf{P}$ (called multiplicative digits of x) such that*

$$x = \prod_{j=1}^{\infty} (e_0 + e_j)^{y_j}. \tag{71}$$

PROOF. Let $y \in \mathbf{B}$. According to (1) and (6) in 9.1 we can write y uniquely as

$$y = e_0 + \sum_{j=1}^{\infty} y_j e_j. \tag{72}$$

Consequently, we can define a formal product on $\mathbf{B}$ by

$$\Phi(y) := \prod_{j=1}^{\infty} (e_0 + y_j e_j).$$

Notice for each $y_j \in \{0, 1\}$ that $(e_0 + y_j e_j) = (e_0 + e_j)^{y_j}$. Consequently, the proof of the lemma will be complete if we show Φ is defined on all of $\mathbf{B}$ and is a 1-1 function from $\mathbf{B}$ onto $\mathbf{B}$.

To show Φ is defined on $\mathbf{B}$, fix $y \in \mathbf{B}$, set $P_0(y) := e_0$ and

$$P_n(y) := \prod_{j=1}^{n} (e_0 + y_j e_j) \qquad (n \in \mathbf{P}).$$

For $m \in \mathbf{N}$ with binary coefficients

$$m = \sum_{j=1}^{\infty} m_j 2^{j-1} \tag{73}$$

set

$$G_m(y) := \prod_{j=1}^{\infty} (y_j e_j)^{m_j}.$$

(Note: In this section we find it convenient to shift the indices of the binary coefficients of elements of **N** by 1 (compare (73) above with (3) in 1.1)). Clearly for each $n \in \mathbf{P}$, the 2^n-th partial sum of the series

$$\sum_{m=0}^{\infty} G_m(y) \tag{74}$$

coincides with $P_n(y)$. Since $\|e_j\| = 2^{-j}$ and

$$\begin{aligned}\sum_{m=0}^{2^n-1} \|G_m(y)\| &\le \sum_{m=0}^{2^n-1} \| \prod_{j=1}^{\infty} e_j^{m_j} \| \\ &\le \sum_{m=0}^{2^n-1} \prod_{j=1}^{\infty} \left(\frac{1}{2^j}\right)^{m_j} \\ &\le \prod_{\ell=1}^{n} (1+2^{-\ell}) \\ &\le \prod_{\ell=1}^{\infty} (1+2^{-\ell}) < \infty,\end{aligned}$$

it follows that the series (74) converges. Consequently, the partial products $P_n(y)$ also converge as $n \to \infty$ and $\Phi(y)$ is well defined.

To show Φ is 1-1 from **B** onto **B**, fix $x \in \mathbf{B}$. Since

$$e_j = e_1^j \qquad (j \in \mathbf{P})$$

we have for any choice of $n_1, n_2, \ldots, n_m \in \{0,1\}$ with $m = n_1 + 2n_2 + 3n_3 + \cdots + m\, n_m$ that

$$\prod_{j=1}^{m} e_j^{n_j} = \prod_{j=1}^{m} (e_1^j)^{n_j} = e_m.$$

Consequently,

$$\begin{aligned}\prod_{j=1}^{\infty} (e_0 + y_j e_j) &= \sum_{m=0}^{\infty} G_m(y) \\ &= e_0 + \sum_{m=1}^{\infty} e_m g_m(y)\end{aligned}$$

where

$$g_m(y) := \sum_{n_1+2n_2+3n_3+\cdots+m\,n_m=m} y_1^{n_1} y_2^{n_2} \ldots y_m^{n_m} \pmod 2$$

for $y \in \mathbf{B}$ and $m \in \mathbf{P}$. Since the g_m's can be written as

$$g_m(y) = y_m + \sum_{n_1+2n_2+3n_3+\cdots+(m-1)n_{m-1}=m} y_1^{n_1} y_2^{n_2} \ldots y_{m-1}^{n_{m-1}} \pmod 2$$

and x can be written as

$$x = e_0 + \sum_{j=1}^{\infty} x_j e_j,$$

it is evident that

$$x = \prod_{j=1}^{\infty} (e_0 + y_j e_j)$$

if and only if

$$x_1 = y_1$$

and

$$x_m = y_m + \sum_{n_1+2n_2+3n_3+\cdots+(m-1)n_{m-1}=m} y_1^{n_1} y_2^{n_2} \cdots y_{m-1}^{n_{m-1}} \pmod 2 \tag{75}$$

for $m = 2, 3, \ldots$, where the sum in (75) is by convention zero if the index set is empty. In particular, the map Φ is onto. In fact, by (75) the dyadic digits $y_1, y_2, \ldots$ can be defined recursively by

$$y_1 := x_1$$

and

$$y_m := x_m + A_m(x_1, \ldots, x_{m-1}) \pmod 2$$

for certain functions

$$A_m : \mathbf{Z}_2^{m-1} \to \mathbf{Z}_2$$

for $m = 2, 3, \ldots$. Consequently, Φ is also 1-1. ∎

Notice that every rearrangement of (74) converges to the same sum. Thus the product (71) actually converges unconditionally.

COROLLARY 7. *Let $x \in \mathbf{B}$ and $(y_j, j \in \mathbf{P})$ represent the multiplicative digits of x. Then*

$$x = \prod_{m=0}^{\infty} \prod_{k=0}^{\infty} (e_0 + e_{2m+1})^{2^k y_{2^k(2m+1)}}. \tag{76}$$

PROOF. For $j \in \mathbf{P}$ choose $k, m \in \mathbf{N}$ uniquely so that

$$j = 2^k(2m+1).$$

Then by unconditional convergence we can write (71) as

$$x = \prod_{m=0}^{\infty} \prod_{k=0}^{\infty} (e_0 + e_{2^k(2m+1)})^{y_{2^k(2m+1)}}. \tag{77}$$

However, an easy induction argument establishes

$$e_0 + e_{2^k(2m+1)} = (e_0 + e_{2m+1})^{2^k} \tag{78}$$

for $k, m \in \mathbf{N}$. Combining this with (77) we obtain (76). ∎

We are prepared to identify the multiplicative characters of the compact abelian group $\mathbf{B}$. Denote the circle group $\{e^{\imath\theta} : \theta \in [0, 2\pi)\}$ by $\mathbf{T}$. For each $\ell = 2^k(2m+1) \in \mathbf{P}$ define

$$\phi_\ell(x) := \exp \pi\imath(y_\ell + \sum_{s=1}^{k} y_{\ell 2^{-s}} 2^{-s}) \tag{79}$$

for $x \in \mathbf{B}$ with multiplicative digits $(y_j, j \in \mathbf{P})$, where $\imath = \sqrt{-1}$. For $n \in \mathbf{N}$ with binary expansion

$$n = \sum_{\ell=1}^{\infty} n_\ell 2^{\ell-1}$$

define

$$v_n := \prod_{\ell=1}^{\infty} (\phi_\ell)^{n_\ell}.$$

Notice each v_n is a finite product, hence continuous from $\mathbf{B}$ to $\mathbf{T}$ and

$$v_{2^{\ell-1}} = \phi_\ell$$

for each $\ell \in \mathbf{P}$.

THEOREM 18. *The full set of multiplicative characters of* $\mathbf{B}$ *is given by the system*

$$(v_n, n \in \mathbf{N}).$$

PROOF. Fix $k, m \in \mathbf{N}$ and set $\ell := 2^k(2m+1)$. To show each v_n is a character of $\mathbf{B}$ it suffices to show

$$\phi_\ell(x \cdot \tilde{x}) = \phi_\ell(x)\phi_\ell(\tilde{x}) \tag{80}$$

for $x, \tilde{x} \in \mathbf{B}$.

Fix such $x, \tilde{x}$ and represent their respective multiplicative digits by y_j, $\tilde{y}_j$ for $j \in \mathbf{P}$. For each integer $0 \le s \le k$ set

$$\alpha_s := y_{2^{k-s}(2m+1)} + \tilde{y}_{2^{k-s}(2m+1)}$$

and let β_s represent the multiplicative digit of $x \cdot \tilde{x}$ of index $2^{k-s}(2m+1)$. Notice by Lemma 3 and its corollary that

$$x \cdot \tilde{x} = \prod_{j=1}^{\infty}(e_0 + e_j)^{y_j + \tilde{y}_j} = \prod_{u=1}^{\infty}\prod_{v=1}^{\infty}(e_0 + e_{2^v(2u+1)})^{y_{2^v(2u+1)} + \tilde{y}_{2^v(2u+1)}}.$$

This is not a canonical product so it is not necessarily the case that $\alpha_s = \beta_s$. Nevertheless we see by (78) that they are related by

$$\alpha_s + t_{s+1} = 2t_s + \beta_s$$

for $s = k, k-1, \ldots, 1, 0$, $t_{k+1} = 0$, and $t_s \in \{0, 1\}$. Therefore,

$$\begin{aligned}
\exp(\pi\imath \sum_{s=0}^{k}(\alpha_s - \beta_s)2^{-s}) &= \exp(\pi\imath \sum_{s=0}^{k}(2t_s - t_{s+1})2^{-s}) \\
&= \exp\left(\pi\imath \sum_{s=0}^{k}\left(\frac{t_s}{2^{s-1}} - \frac{t_{s+1}}{2^s}\right)\right) \\
&= \exp \pi\imath \left(2t_0 - \frac{t_{k+1}}{2^k}\right) \\
&= \exp 2\pi\imath t_0 = 1.
\end{aligned}$$

In particular,

$$\exp(\pi\imath \sum_{s=0}^{k}\beta_s 2^{-s}) = \exp(\pi\imath \sum_{s=0}^{k}\alpha_s 2^{-s})$$

and (80) follows at once from (79).

Conversely, let f be any multiplicative character of $\mathbf{B}$. By Corollary 7 we have

$$f(x) = \prod_{m=0}^{\infty}\prod_{k=0}^{\infty} f(e_0 + e_{2m+1})^{2^k y_{2^k(2m+1)}} \tag{81}$$

for $x \in \mathbf{B}$ with multiplicative digits $(y_j, j \in \mathbf{P})$. Thus f is completely determined by its values

$$f(e_0 + e_{2m+1}) \qquad (m \in \mathbf{N}).$$

Fix $x \in \mathbf{B}$ with multiplicative digits $(y_j, j \in \mathbf{P})$, and observe that $f(e_0 + e_{2^k(2m+1)}) \to 1$ as $k \to \infty$. Thus for each $m \in \mathbf{N}$ there is a $q_m \in \mathbf{N}$ such that

$$f(e_0 + e_{2^{q_m}(2m+1)}) = (f(e_0 + e_{2m+1}))^{2^{q_m}} = 1.$$

Choose a natural number $0 \le p_m < 2^{q_m}$ such that

$$f(e_0 + e_{2m+1}) = \exp\left(\frac{2\pi\imath p_m}{2^{q_m}}\right) \tag{82}$$

for $m \in \mathbf{N}$. Suppose for a moment that for all $m \in \mathbf{N}$ there is an $s_m \in \mathbf{N}$ such that

$$f(e_0 + e_{2^{s_m}(2m+1)}) = -1.$$

Since $e_0 + e_{2^{s_m}(2m+1)} \to e_0$ as $m \to \infty$ it would follow from (82) and the continuity of f that $f(e_0) = -1$, a contradiction. Hence such an s_m can be found only for finitely many $m \in \mathbf{N}$, i.e., there is an $M \in \mathbf{N}$ such that

$$p_m = 0$$

for $m \geq M$. In view of (81) and (82), we have verified that

$$f(x) = \exp 2\pi\imath \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \frac{p_m}{2^{q_m}} 2^k y_{2^k(2m+1)} \tag{83}$$

and that these sums are finite.

It remains to show the right sides of (83) and (79) coincide for a suitable choice of $n \in \mathbf{N}$. Define $n \in \mathbf{N}$ by

$$n := \sum_{m=0}^{\infty} \sum_{v=0}^{\infty} n_v^m 2^{2^v(2m+1)-1}$$

where for each $m \in \mathbf{N}$ coefficients $n_v^m = 0$ or 1 are determined by

$$\frac{p_m}{2^{q_m}} = \sum_{v=0}^{\infty} n_v^m \, 2^{-(v+1)}$$

and $n_v^m := 0$ for each $v \geq q_m$. Thus the binary coefficients $(n_\ell, \ell \in \mathbf{P})$ of n satisfy

$$\frac{p_m}{2^{q_m}} = \sum_{j=0}^{\infty} n_{2^j(2m+1)} 2^{-(j+1)}. \tag{84}$$

By definition and (79)

$$\begin{aligned} v_n(x) &= \prod_{m=0}^{\infty} \prod_{j=0}^{\infty} \left(\phi_{2^j(2m+1)}(x)\right)^{n_{2^j(2m+1)}} \\ &= \prod_{m=0}^{\infty} \exp \pi\imath \sum_{j=0}^{\infty} \sum_{s=0}^{j} 2^{-s} n_{2^j(2m+1)} y_{2^{j-s}(2m+1)}. \end{aligned}$$

Moreover, it is easy to see by substituting k for $j - s$ that

$$\begin{aligned} &\exp \pi\imath \sum_{j=0}^{\infty} \sum_{s=0}^{j} 2^{-s} n_{2^j(2m+1)} \, y_{2^{j-s}(2m+1)} \\ &\qquad = \exp \pi\imath \sum_{j=0}^{\infty} \sum_{k=0}^{j} 2^{k-j} n_{2^j(2m+1)} y_{2^k(2m+1)} \\ &\qquad = \exp 2\pi\imath \sum_{j=0}^{\infty} 2^{-(j+1)} n_{2^j(2m+1)} \sum_{k=0}^{j} 2^k y_{2^k(2m+1)} \\ &\qquad = \exp 2\pi\imath \sum_{j=0}^{\infty} 2^{-(j+1)} n_{2^j(2m+1)} \sum_{k=0}^{\infty} 2^k y_{2^k(2m+1)}. \end{aligned}$$

Consequently, it follows from (84) and the calculation above that

$$v_n(x) = \exp 2\pi\imath \sum_{m=0}^{\infty}\sum_{k=0}^{\infty} \frac{p_m}{2^{q_m}} 2^k y_{2^k(2m+1)}$$

We conclude by (83) that $v_n(x) = f(x)$ as required. ∎

The set of characters of any compact abelian group is itself a group. Thus for each pair $n, \ell \in \mathbf{N}$ there is an integer $n \circ \ell \in \mathbf{N}$ which satisfies

$$v_{n\circ\ell} = v_n v_\ell.$$

An explicit description of the operation $\circ$ can be given. Indeed, fix $n, \ell \in \mathbf{N}$ and notice by the proof of Theorem 18 for $j = 2m+1$ that

$$v_n(e_0 + e_j) = \exp \pi\imath \sum_{s=0}^{\infty} n_{j2^s} 2^{-s}$$

for n with binary coefficients $(n_v, v \in \mathbf{P})$. Hence if $(\ell_v, v \in \mathbf{P})$ represents the binary coefficients of ℓ then for each $j \in \mathbf{P}$ we have

$$\begin{aligned} v_n(e_0 + e_j)v_\ell(e_0 + e_j) &= \exp \pi\imath \sum_{s=0}^{\infty} (n_{j2^s} + \ell_{j2^s}) 2^{-s} \\ &= v_{n\circ\ell}(e_0 + e_j) \\ &= \exp \pi\imath \sum_{s=0}^{\infty} (n \circ \ell)_{j2^s} 2^{-s}. \end{aligned}$$

In particular, the operation $\circ$ on $\mathbf{N}$ satisfies

$$(85) \quad \sum_{s=0}^{\infty} (n \circ \ell)_{(2m+1)2^s} 2^{-(s+1)} = \sum_{s=0}^{\infty} n_{(2m+1)2^s} 2^{-(s+1)} + \sum_{s=0}^{\infty} \ell_{(2m+1)2^s} 2^{-(s+1)} \pmod 1$$

for every $m \in \mathbf{N}$.

Let $\mathbf{Q}^\infty$ denote the product group formed by taking the cartesian product of countably many copies of the group of dyadic rationals $\mathbf{Q}$ (with addition modulo 1 as the group operation). For each $n \in \mathbf{N}$ with binary expansion

$$n = \sum_{j=1}^{\infty} n_j 2^{j-1}$$

define an element $x = (x_0, x_1, \dots) \in \mathbf{Q}^\infty$ by

$$x_m := \sum_{s=0}^{\infty} n_{(2m+1)2^s} 2^{-(s+1)} \qquad (m \in \mathbf{N}).$$

Then (85) implies that the map $n \to x$ is a group isomorphism of $(N, \circ)$ onto $\mathbf{Q}^\infty$. In particular, the character group $\widehat{\mathbf{B}}$ is isomorphic to $\mathbf{Q}^\infty$.

It is easy to see that the translation invariant measure μ when restricted to $\mathbf{B}$ is also multiplicatively invariant. Hence it follows from Theorem 18 that the system $(v_n, n \in \mathbf{N})$ is orthonormal on $\mathrm{L}^2_\mu(\mathbf{B})$. Now the definition of ϕ_ℓ shows that for each $\ell \in \mathbf{P}$ there is a function B_ℓ depending only on the digits $x_1, \dots, x_{\ell-1}$ of x (see (72)) such that

$$\phi_\ell(x) = \exp \pi\imath(x_\ell + B_\ell(x_1, \dots, x_{\ell-1}))$$

for $x \in \mathbf{B}$. Thus each ϕ_ℓ is $\mathcal{A}^\ell$- measurable. Moreover, since

$$\mathcal{E}(\phi_\ell | \mathcal{A}^{\ell-1}) = \exp \pi\imath B_\ell(x_1, \dots, x_{\ell-1}) \mathcal{E}(\exp \pi\imath x_\ell | \mathcal{A}^{\ell-1})$$

we also have

$$\mathcal{E}(\phi_\ell | \mathcal{A}^{\ell-1}) = 0 \qquad (\ell \in \mathbf{P}).$$

In particular (see Schipp [14]) the system $(v_n, n \in \mathbf{N})$ is a convergence system. That is, if $F \in \mathrm{L}^2_\mu(\mathbf{B})$ and

$$a_n(F) := \int_{\mathbf{B}} F\overline{v}_n \, d\mu \qquad (n \in \mathbf{N})$$

then

$$\lim_{N\to\infty} \sum_{n=0}^{N} a_n(F)\overline{v}_n = F \tag{86}$$

a.e. $[\mu]$ on $\mathbf{B}$.

We are now prepared to describe $\widehat{\mathbf{F}^*}$ and show inversion holds for the Mellin transform. As we noted near the beginning of this section, $\widehat{\mathbf{F}^*}$ can be identified with $\widehat{\mathbf{Z}} \times \widehat{\mathbf{B}}$ and $\mathbf{F}^*$ can be identified with $\mathbf{Z} \times \mathbf{B}$, where the group operation on $\mathbf{Z}$ is classical addition of integers and that on $\mathbf{B}$ is multiplication from $\mathbf{F}^*$. Thus by classical Fourier analysis and Pontryagin duality, $\widehat{\mathbf{Z}}$ is isomorphic to the circle group $\mathbf{T}$. By Theorem 18 above, $\widehat{\mathbf{B}}$ is isomorphic to $(\mathbf{N}, \circ)$. Consequently, $\widehat{\mathbf{F}^*}$ can be identified with the product group $[0,1) \times \mathbf{N}$ where the group operation on $[0,1)$ is addition modulo 1 and that on $\mathbf{N}$ is defined by (85) above.

With these isomorphisms in mind, we can identify Haar integrals on $\mathbf{F}^*$ and $\widehat{\mathbf{F}^*}$ as follows. If $f \in \mathrm{L}^1_\nu := \mathrm{L}^1_\nu(\mathbf{F}^*)$ then

$$\int_{\mathbf{F}^*} f \, d\nu = \sum_{k\in\mathbf{Z}} \int_{\mathbf{B}} f(k, x) \, d\mu(x). \tag{87}$$

If $g \in \mathrm{L}^1(\widehat{\nu}) := \mathrm{L}^1_{\widehat{\nu}}(\widehat{\mathbf{F}^*})$, and $\widehat{\nu}$ represents Haar measure on the group $\widehat{\mathbf{F}^*}$, then

$$\int_{\widehat{\mathbf{F}^*}} g \, d\widehat{\nu} = \sum_{n\in\mathbf{N}} \int_0^1 g(\alpha, n) \, d\alpha. \tag{88}$$

Moreover, using the fact that the characters of a product group are products of the characters of the factor groups, we see that the characters of $\mathbf{F}^*$ are given by

$$\gamma_{\alpha,n}(k,x) := e_\alpha(k)v_n(x)$$

for $(k,x) \in \mathbf{Z} \times \mathbf{B}$ and $(\alpha,n) \in [0,1) \times \mathbf{N}$, where

$$e_\alpha(k) := e_k(\alpha) := \exp(2\pi\imath k\alpha)$$

for $k \in \mathbf{Z}$ and $\alpha \in [0,1)$. In particular, the Mellin transform defined at the beginning of this section takes on the explicit form

$$(\mathcal{M}f)(\alpha,n) = \sum_{k\in\mathbf{Z}} \overline{e}_k(\alpha) \left(\int_{\mathbf{B}} f(k,x)\overline{v}_n(x)\, d\mu(x) \right) \tag{89}$$

for $(\alpha,n) \in [0,1) \times \mathbf{N}$ and $f \in \mathrm{L}^1_\nu$.

This definition makes sense for any f in $\mathrm{L}^1_\nu \cup \mathrm{L}^2_\nu$. In fact,

THEOREM 19. *i) If $f \in \mathrm{L}^1_\nu$ then the right side of (89) converges absolutely and uniformly on $[0,1) \times \mathbf{N}$, hence $\mathcal{M}f$ is continuous.*

ii) If $f \in \mathrm{L}^2_\nu$ then the right side of (89) converges in L^2 norm and a.e. on $[0,1) \times \mathbf{N}$. Moreover,

$$\|\mathcal{M}f\|_{L^2(\widehat{\nu})} = \|f\|_{L^2_\nu}. \tag{90}$$

PROOF. If $f \in \mathrm{L}^1_\nu$ then we have by (87) that

$$\sum_{k\in\mathbf{Z}} \int_{\mathbf{B}} |f(k,x)|\, d\mu(x) < \infty.$$

Hence i) follows easily from the Weierstrass M-test.

If $f \in \mathrm{L}^2_\nu$ and $n \in \mathbf{N}$ then the sequence $(c_k(n), k \in \mathbf{Z})$ defined by

$$c_k(n) := \int_{\mathbf{B}} f(k,x)\overline{v}_n(x)\, d\mu(x) \qquad (k \in \mathbf{Z})$$

belongs to $\ell^2(\mathbf{Z})$. Hence the right side of (89) is a trigonometric Fourier series, $T^{(n)}$, of some $\mathrm{L}^2[0,1)$ function. By Carleson's theorem (see Carleson [1]; compare with Theorem 14 in 3.7), $T^{(n)}$ converges a.e. on $[0,1)$ for each $n \in \mathbf{N}$. By the Riesz-Fischer theorem, $T^{(n)}$ converges in $\mathrm{L}^2[0,1)$ norm for each $n \in \mathbf{N}$. Moreover, by (87), Parseval's identity and (88), we also have

$$\begin{aligned}\|\mathcal{M}f\|^2_{L^2(\widehat{\nu})} &= \sum_{n\in\mathbf{N}} \int_0^1 |\sum_{k\in\mathbf{Z}} c_k(n)\overline{e}_k(\alpha)|^2\, d\alpha \\ &= \sum_{n\in\mathbf{N}} \left(\sum_{k\in\mathbf{Z}} |c_k(n)|^2 \right) \\ &= \|f\|^2_{L^2_\nu}.\end{aligned}$$

This verifies (90) and the proof of ii) is complete. ∎

By (88) the inverse Mellin transform of a $g \in L^1(\widehat{\nu})$ is defined by

$$(91) \qquad (\mathcal{M}'g)(k,x) := \sum_{n \in \mathbf{N}} v_n(x) \left(\int_0^1 g(\alpha, n) e_k(\alpha)\, d\alpha \right)$$

for $(k,x) \in \mathbf{Z} \times \mathbf{B}$.

THEOREM 20. *i) If $g \in L^1(\widehat{\nu})$ then the right side of (91) converges absolutely and uniformly on $\mathbf{Z} \times \mathbf{B}$.*

ii) If $g \in L^2(\widehat{\nu})$ then the right side of (91) converges in L^2 norm and a.e. on $\mathbf{Z} \times \mathbf{B}$. Moreover,

$$\|\mathcal{M}'g\|_{L^2_\nu} = \|g\|_{L^2(\widehat{\nu})}.$$

PROOF. Notice for each fixed $k \in \mathbf{Z}$ and $g \in L^2(\widehat{\nu})$ that the right side of (91) is a Fourier series of an $L^2_\nu(\mathbf{B})$ function with respect to the system $(v_n, n \in \mathbf{N})$. Moreover, recall by (86) that this system is a convergence system. Consequently, we can verify this theorem by repeating the proof of Theorem 19. ∎

It is now easy to see that inversion holds for the Mellin transform.

THEOREM 21. *i) If $f \in L^2_\nu$ then*

$$\mathcal{M}'(\mathcal{M}f) = f$$

a.e. and in L^2_ν norm.

ii) If $g \in L^2(\widehat{\nu})$ then

$$\mathcal{M}(\mathcal{M}'g) = g$$

a.e. and in $L^2(\widehat{\nu})$ norm.

PROOF. Fix $m \in \mathbf{N}$, $j \in \mathbf{Z}$ and define functions $f_{m,j} := f$ on $\mathbf{Z} \times \mathbf{B}$ and $g_{m,j} := g$ on $[0,1) \times \mathbf{N}$ by

$$f(k,x) := \delta_{jk} v_m(x)$$

and

$$g(\alpha, n) := \delta_{nm} \overline{e}_j(\alpha).$$

By definition and orthogonality we have

$$\begin{aligned}(\mathcal{M}f)(\alpha,n) &= \sum_{k \in \mathbf{Z}} \overline{e}_k(\alpha) \delta_{jk} \int_{\mathbf{B}} v_m \overline{v}_n \, d\mu \\ &= \delta_{nm} \overline{e}_j(\alpha) \\ &= g(\alpha, n)\end{aligned}$$

for $(\alpha, n) \in [0,1) \times \mathbf{N}$. Similarly

$$\begin{aligned}(\mathcal{M}'g)(\alpha,n) &= \sum_{n \in \mathbf{N}} v_n(x) \delta_{nm} \int_0^1 e_k \overline{e}_j \\ &= \delta_{jk} v_m(x) \\ &= f(k,x)\end{aligned}$$

for $(k,x) \in \mathbf{Z} \times \mathbf{B}$. Thus $\mathcal{M}f = g$, $\mathcal{M}'g = f$,

$$\mathcal{M}'(\mathcal{M}f_{m,j}) = f_{m,j},$$

and

$$\mathcal{M}(\mathcal{M}'g_{m,j}) = g_{m,j}$$

for each $m \in \mathbf{N}$ and $j \in \mathbf{Z}$. However, it is easy to see that

$$\{f_{m,j} : m \in \mathbf{N},\ j \in \mathbf{Z}\}$$

is a closed system in L^2_ν and

$$\{g_{m,j} : m \in \mathbf{N},\ j \in \mathbf{Z}\}$$

is a closed system in $L^2(\widehat{\nu})$. Hence inversion holds and the proof of Theorem 20 is complete. ▮

9.7 The Fast Walsh Transform. Let g be a discrete function of order N, i.e., a function defined on the set $\{0,1,\ldots,N-1\}$. The Walsh Transform of g is the discrete function of order N defined by

$$\hat{g}(j) := \frac{1}{N}\sum_{k=0}^{N-1} g(k)w_j(\frac{k}{N}) \tag{92}$$

for $0 \leq j < N$, where each w_j is a Walsh function as defined in 1.1.

The Walsh Transform plays the same role for discrete functions as the Walsh-Fourier transform does for functions in $L^1(\mathbf{F})$ and the Walsh-Fourier coefficient map does for functions in $L^1(\mathbf{G})$. It is easy to see that the Walsh Transform is linear on the collection of discrete functions of order N. Moreover, when N is a power of two, say $N = 2^n$ for some $n \in \mathbf{N}$, then the collection $\{0,1,\ldots,N-1\}$ forms a group under dyadic addition $\oplus$ whose characters are given by the maps $k \to w_j(k/N)$ for $j = 0,1,\ldots,N-1$. In particular, the Walsh Transform is a Fourier transform on a compact group, hence satisfies the kind of duality exhibited by (18) and (19) in 9.2.

There is a close connection between the Walsh Transform of order $N = 2^n$ and the Walsh-Fourier coefficients of $\mathcal{A}^n$-measurable functions. Indeed, if g is discrete of order 2^n and

$$G(x) := g(k) \tag{93}$$

for $x \in [k/2^n, (k+1)/2^n)$, $k = 0,1,\ldots,2^n-1$, then G is $\mathcal{A}^n$-measurable on $[0,1)$ and the map $g \to G$ is a linear isomorphism between the collection of discrete functions of order 2^n and the collection of $\mathcal{A}^n$-measurable functions on $[0,1)$. Moreover, the Walsh-Fourier coefficients of G obviously satisfy

$$\begin{aligned}\widehat{G}(j) &= \int_0^1 G(x)w_j(x)\,dx \\ &= \frac{1}{2^n}\sum_{k=0}^{2^n-1} G(\frac{k}{2^n})w_j(\frac{k}{2^n}) \\ &= \hat{g}(j)\end{aligned}$$

for $j = 0, 1, \ldots, 2^n - 1$.

Since an $\mathcal{A}^n$-measurable G satisfies

$$G(x) = (S_{2^n}G)(x) = \sum_{k=0}^{2^n-1} \widehat{G}(k)w_k(x)$$

for $x \in [0, 1)$, it is now clear that inversion holds for the Walsh Transform of order 2^n:

THEOREM 22. *If g is discrete of order $N = 2^n$ and $\hat{g}$ is its Walsh Transform then*

$$g(k) = \sum_{j=0}^{N-1} \hat{g}(j)w_k(\frac{j}{N})$$

for $k = 0, 1, \ldots, N - 1$.

The Walsh Transform can be used to approximate Walsh-Fourier coefficients.

THEOREM 23. *Let $n \in \mathbf{N}$, f be Riemann integrable on $[0, 1)$, and*

$$g_n(k) := f(\frac{k}{2^n})$$

for $k = 0, 1, \ldots, 2^n - 1$. Then the Walsh Transforms of g_n satisfy

$$\widehat{f}(j) = \lim_{n\to\infty} \hat{g}_n(j).$$

PROOF. Let $N = 2^n$ and suppose $N > j$. By hypothesis

$$\hat{g}_n(j) = \frac{1}{N} \sum_{k=0}^{N-1} f(\frac{k}{N})w_j(\frac{k}{N}). \tag{94}$$

Since $\{k/N : k = 0, 1, \ldots, N - 1\}$ generates an even partition of $[0, 1)$, it is clear that the right side of (94) is a Riemann sum. Consequently,

$$\lim_{n\to\infty} \hat{g}_n(j) = \int_0^1 f(t)w_j(t)\,dt = \widehat{f}(j). \qquad \blacksquare$$

The chief advantage of this observation lies in the fact that a Walsh Transform of order $N = 2^n$ can be computed very efficiently (see Theorem 24 below).

Let $\mathcal{E}_k$ represent the conditional expectation operator with respect to the σ-algebra $\mathcal{A}^k$ for $k \in \mathbf{N}$. Recall from 3.1 that for each $f \in \mathrm{L}^1$,

$$\mathcal{E}_0 f = \int_0^1 f \tag{95}$$

$$\mathcal{E}_k(\lambda f) = \lambda \mathcal{E}_k f \tag{96}$$

for each $\mathcal{A}^k$-measurable function λ, and

$$\mathcal{E}_k(\mathcal{E}_\ell f) = \mathcal{E}_\ell(\mathcal{E}_k f) = \mathcal{E}_\ell f \tag{97}$$

for $\ell \leq k$, $\ell, k \in \mathbf{N}$. Recall also that

$$(\mathcal{E}_k f)(x) = (S_{2^k} f)(x) = \frac{1}{|I_k(x)|}\int_{I_k(x)} f$$

for $x \in [0,1)$. Since each $\mathcal{A}^{k+1}$-measurable function λ is constant on any dyadic interval of length 2^{-k-1}, it follows that

$$(\mathcal{E}_k \lambda)(x) = \frac{1}{2}(\lambda(x_0) + \lambda(x_1)) \tag{98}$$

where x_0 is any point from the left half of $I_k(x)$ and x_1 is any point from the right half of $I_k(x)$.

Let g be discrete of order 2^n and fix an integer $0 \leq j < 2^n$. Let G be the $\mathcal{A}^n$-measurable function corresponding to g (see (93) above). Thus

$$\hat{g}(j) = \int_0^1 w_j G.$$

Write the binary expansion of j as

$$j = \sum_{v=0}^{n-1} j_v 2^v \tag{99}$$

and recall that $w_j = r_0^{j_0} \dots r_{n-1}^{j_{n-1}}$. Hence

$$\hat{g}(j) = \int_0^1 r_0^{j_0} \dots r_{n-1}^{j_{n-1}} G. \tag{100}$$

This suggests a recursive method for computing the Walsh Transform of g.

THEOREM 24. *Let $n \in \mathbf{N}$, $N = 2^n$, and suppose g is discrete of order N. Set*

$$F_n(j) = g(j) \qquad (j = 0, 1, \dots, N-1) \tag{101}$$

and recursively define discrete functions F_k of order n for $k = n-1, n-2, \dots, 0$ by

$$F_k(p2^{n-k} + q) = \frac{1}{2}\left(F_{k+1}(p2^{n-k} + q) + F_{k+1}(p2^{n-k} + 2^{n-k-1} + q)\right) \tag{102}$$

and

$$F_k(p2^{n-k} + 2^{n-k-1} + q) = \frac{1}{2}\left(F_{k+1}(p2^{n-k} + q) - F_{k+1}(p2^{n-k} + 2^{n-k-1} + q)\right) \tag{103}$$

for $0 \le p < 2^k$ *and* $0 \le q < 2^{n-k-1}$*. Then the Walsh Transform of* g *is given by*

$$\hat{g}(j) = F_0(\mathbf{n}(j)) \qquad (j = 0, 1, \ldots, N-1)$$

where $\mathbf{n}(j)$ *is the bit-reversal of* j *defined in at the end of 1.4.*

PROOF. Let G be the $\mathcal{A}^n$-measurable function corresponding to g. Set $G_n^0 := G_n$ and for each $0 \le k \le n$ and $0 \le s < 2^{n-k}$ let G_k^s be the $\mathcal{A}^k$-measurable function defined by

$$G_k^s(\frac{p}{2^k}) = F_k(p2^{n-k} + s) \qquad (0 \le p < 2^k).$$

By definition and by (102) and (103) we have

$$G_k^{i2^{n-k-1}+q}(\frac{p}{2^k}) = \frac{1}{2}\left(G_{k+1}^q(\frac{p}{2^k}) + (-1)^i G_{k+1}^q(\frac{p}{2^k} + \frac{1}{2^{k+1}})\right)$$

for $i \in \{0,1\}$. Consequently, (98) implies

$$G_k^{i2^{n-k-1}+q} = \mathcal{E}_k(r_k^i G_{k+1}^q) \qquad (0 \le q < 2^{n-k-1},\ i = 0, 1)$$

for $k = n-1, n-2, \ldots, 0$. Applying this identity successively, we arrive at

$$G_0^{\mathbf{n}(q)} = \mathcal{E}_0(r_0^{q_0} r_1^{q_1} \ldots r_{n-1}^{q_{n-1}} G)$$

for $q = q_0 + q_1 2 + \cdots + q_{n-1}2^{n-1}$ and $\mathbf{n}(q) = q_0 2^{n-1} + \cdots + q_{n-1}$. Specializing to $q_v := j_v$ where the j_v's are given by (99), we conclude from (95) and (100) that

$$F_0(\mathbf{n}(j)) = G_0^{\mathbf{n}(j)} = \mathcal{E}_0(r_0^{q_0} r_1^{q_1} \ldots r_{n-1}^{q_{n-1}} G) = \hat{g}(j)$$

for $j = 0, 1, \ldots 2^n - 1$. ∎

The algorithm suggested by (101), (102) and (103) is called the Fast Walsh Transform (FWT). It is presently the most efficient algorithm for computing the Walsh Transform. Notice for each fixed k that (102) requires N additions and (103) requires N subtractions. Hence apart from normalization by $1/2$, the n step algorithm FWT takes

$$nN = N \log_2 N$$

arithmetic operations to compute $\hat{g}$. In sharp contrast, definition (92) takes N^2 arithmetic operations. Thus FWT results in considerable savings. For example, if $N = 1024$ then $N \log_2 N = 10,240$ but $N^2 = 1,048,576$.

There is another "Walsh" transform which appears in the literature. Let $(H^{(n)}, n \in \mathbf{N})$ represent the Hadamard matrices. Thus the k-th row of $H^{(n)}$ is given by

(104) $$\{w_{\mathbf{n}(k)}(\frac{j}{2^n}) : j = 0, 1, \ldots, 2^n - 1\}$$

for each $n \in \mathbf{N}$, $0 \le k < 2^n$, where $k \to \mathbf{n}(k)$ represents the bit-reversal map (see 1.4). Let g be discrete of order $N = 2^n$. The Walsh-Hadamard Transform $\tilde{G}$ of g is defined by the matrix equation

$$\begin{pmatrix} \tilde{G}(0) \\ \tilde{G}(1) \\ \vdots \\ \tilde{G}(N-1) \end{pmatrix} = H^{(n)} \cdot \begin{pmatrix} g(0) \\ g(1) \\ \vdots \\ g(n) \end{pmatrix}.$$

Though not widely circulated, there is a simple relationship between the Walsh Transform and the Walsh-Hadamard Transform. Indeed by definition, if g is discrete of order $N = 2^n$ and $\tilde{G}$ is its Walsh-Hadamard Transform then

$$\tilde{G}(k) = \sum_{j=0}^{N-1} g(j) w_{\mathbf{n}(k)}(\frac{j}{2^n})$$

for $0 \le k < 2^n$. Since bit-reversal is its own inverse it follows that

$$\tilde{G}(k) = \hat{g}(\mathbf{n}(k))$$

for $k = 0, 1, \dots, 2^n - 1$, where $\hat{g}$ is the Walsh Transform of g. In particular, the Walsh-Hadamard transform has a two- sided inverse and can be computed efficiently by a trivial modification to any existing program which computes FWT.

The Walsh Transform has many applications.

Most of these are based on the fact (see Theorem 23 above) that FWT approximates the Walsh-Fourier coefficients of a sampled function.

For example, consider the problem of pattern recognition. A known quantity (such as human speech, cursive letters, or handwritten codes) is analyzed by determining the relative strengths and distribution of its Walsh-Fourier coefficients. A standard is formed based on this characterization. An unknown pattern can be recognized by sampling it, applying FWT, and comparing the results with the standard. The unknown pattern is identified with the known quantity when its Walsh-Fourier coefficients fit the standard.

Another example is image enhancement. A picture is reduced to square gray levels (called pixels) to produce a discrete function f. The Walsh-Fourier coefficients of f are approximated by FWT. These coefficients are altered to produce subtle changes in the original picture. Since higher order coefficients tend to be generated by edges and sharp transitions, diminishing higher order coefficients tends to blur or reduce contrast in the original picture. This technique is helpful for images which contain interference (so-called noisy transmissions). On the other hand, diminishing lower order coefficients tends to sharpen the original picture. This technique is useful for images where lighting was too intense, too weak, or where the image was blurred by atmospheric effects such as fog or a light cloud cover.

Other applications stem from the fact that FWT is incredibly fast. Take, for example, matrix multiplication and matrix inversion. Since the digits of a number can be treated as a discrete function which can be expanded in a Walsh-Fourier series, multiplication can be done in transform space more efficiently. Even for something as simple as the multiplication of two numbers, this method can be as much as 40 times faster for 1000-digit numbers.

All these applications work as well with any Fourier transform, not just the Walsh one. There are other applications which use the Walsh Transform in an essential way. We close this section with such an application. Define a function ϵ_k on $[0,1)$ by

$$\epsilon_k := \frac{1-w_k}{2} \qquad (k \in \mathbf{P}),$$

and consider the associated Fourier coefficients

$$\gamma(k) := \int_0^1 G\epsilon_k \tag{105}$$

for $k \in \mathbf{P}$ and $G \in \mathrm{L}^1$. Clearly each ϵ_k takes only values 0 and 1, and by definition the $\gamma(k)$'s can be obtained from the Walsh-Fourier coefficients of G by

$$\gamma(k) = \frac{1}{2}(\widehat{G}(0) - \widehat{G}(k)) \qquad (k \in \mathbf{P}).$$

In particular, if G is an $\mathcal{A}^n$-measurable function and if $\widehat{G}(0)$ and $\gamma(k)$ are known for $k = 0, 1, \dots, 2^n - 1$ then G can be reconstructed using FWT. This observation suggests a physical method for collection of data which simulates the Walsh Transform.

Suppose data is collectable in $2^n - 1$ separate channels and that the signal in each channel is constant over small time intervals. (Spectroscopy and Radio Astronomy offer good examples for this application.) For each $i = 1, 2, \dots, 2^n - 1$, let $g(i)$ represent the datum coming through the i-th channel. Let G represent the $\mathcal{A}^n$-measurable function corresponding to g. Under normal circumstances the data g would be collected by taking $2^n - 1$ readings, one per channel.

Consider a different scheme for collecting the data g. Read several channels at once, adding the results and dividing this sum by 2^{n-1}. Repeat this process $2^n - 1$ times to generate new data $\gamma(1), \gamma(2), \dots, \gamma(2^n - 1)$. At each reading, the decision regarding which channels to collect is mandated by the coefficients $\varepsilon_{ij} := \epsilon_i(j/2^n)$. If $\varepsilon_{ij} = 1$ then collect the j-th channel on the i- th reading. This process yields γ as defined in (105). In particular, the original data g can be recovered from the collected data γ by a single application of the Walsh Transform. This method is especially useful when background noise or interference is independent of the intensity of the signal and not multiplied across several channels. Indeed, the intensity of γ is 2^{n-1} times that of g. Thus for large n the background noise tends to affect γ very little.

Exercises

9.1. Suppose f and its Walsh-Fourier transform $\widehat{f}$ belong to $\mathrm{L}^1(\mathbf{R}^+)$. Suppose that f is W-continuous on $\mathbf{R}^+$ and sequency-limited, i.e., $\widehat{f}(y) = 0$ for $y \geq 2^n$ and some $n \in \mathbf{Z}$. Prove that f has the sampling representation

$$f(t) = \sum_{k=0}^{\infty} f(\frac{k}{2^n}) D_0(2^n t \dot{+} k) \qquad (t \in \mathbf{R}^+).$$

(Hint: Show

$$D_0(2^n t \dot{+} k) = \chi[k/2^n, (k+1)/2^n)(t)$$

for $t \in \mathbf{R}^+$, $k \in \mathbf{N}$, $n \in \mathbf{Z}$, and

$$f(t) = 2^n \int_{k/2^n}^{(k+1)/2^n} f$$

for $t \in [k/2^n, (k+1)/2^n)$, $k \in \mathbf{N}$, $n \in \mathbf{Z}$.)

9.2. Suppose f and $\widehat{f}$ belong to $\mathrm{L}^1(\mathbf{R}^+)$, that f is W-continuous on $\mathbf{R}^+$ and duration-limited, i.e., $f(t) = 0$ for $t \geq T$ for some $T \in \mathbf{R}^+$. Prove that f has the sampling representation

$$f(t) = \lim_{n \to \infty} \sum_{k=0}^{\infty} f(\frac{k}{2^n}) D_0(2^n t \dot{+} k) \qquad (t \in \mathbf{R}^+).$$

Moreover, show the truncation error

$$R_n(t) := f(t) - \sum_{k=0}^{\infty} f(\frac{k}{2^n}) D_0(2^n t \dot{+} k)$$

satisfies

$$\|R_n\|_\infty := \sup_{t \in \mathbf{R}^+} |R_n(t)| \leq 2 \int_{2^n}^{\infty} |\widehat{f}|.$$

9.3. Show, unlike the trigonometric case, that there exist functions which are simultaneously sequency-limited and duration- limited.

9.4. Show the Walsh-Fourier transform of any $f \in \mathrm{L}^1(\mathbf{R}^+)$ can be estimated by

$$|\widehat{f}(y)| \leq \frac{1}{2} \omega^{(1)}(f, \frac{2}{y}) \qquad (y > 0).$$

9.5. Suppose f and $\mathbf{d}^{[r]} f \in \mathrm{L}^1(\mathbf{R}^+)$ for some $r \in \mathbf{N}$. Show

$$\omega^{(1)}(f, \delta) = O(\delta^r \omega^{(1)}(\mathbf{d}^{[r]} f, \delta))$$

as $\delta \downarrow 0$.

9.6. Suppose f satisfies the assumptions in Exercise 9.2. Suppose further that either

i) $\mathbf{d}^{[r]} f \in \mathrm{L}^1(\mathbf{R}^+)$ exists and belongs to $\mathrm{Lip}\alpha$,

or

ii) $f \in \mathrm{Lip}(\alpha + r)$ for some $\alpha > 0$ and $r \in \mathbf{N}$.

Show there is a constant $L > 0$ such that

$$\|R_n\|_\infty \leq L \frac{2^{r+\alpha-1}}{r+\alpha-1} 2^{-n(r \dot{+} \alpha - 1)}.$$

For the remaining exercises we introduce the following notation. Let $(\Omega, \mathcal{A}, \nu)$ be a probability space, p and n be fixed integers greater than 1 and for $k = 0, 1, \ldots, n-1$ let

$$\Phi_k := \{\varphi_k^j : j = 0, 1, \ldots p-1\}$$

be a fixed sequence of $L^2(\Omega, \mathcal{A}, \nu)$ functions. Denote the product system of the Φ_k's by Ψ, i.e., for $m = (m_0, m_1, \ldots, m_{n-1})$ with $0 \le m_j < p$ set

$$\psi_m := \varphi_0^{m_0} \varphi_1^{m_1} \cdots \varphi_{n-1}^{m_{n-1}}.$$

Suppose further there are σ-algebras

$$\mathcal{A}_0 := \{\Omega, \emptyset\} \subseteq \mathcal{A}_1 \subseteq \cdots \subseteq \mathcal{A}_n \subseteq \mathcal{A}$$

such that Φ_k consists of $\mathcal{A}_{k+1}$-measurable functions for $k = 0, 1, \ldots, n-1$.

9.7. (Generalized Fast Fourier Transform: GFFT.) For a given $G \in L^2(\Omega, \mathcal{A}, \nu)$ set $F_n^{\emptyset} := \mathcal{E}_n G$ and

$$F_k^m := \mathcal{E}_k(\varphi_k^{m_k} \varphi_{k+1}^{m_{k+1}} \cdots \varphi_{n-1}^{m_{n-1}} G)$$

for $m = (m_k, m_{k+1}, \ldots, m_{n-1})$, $0 \le k \le n-1$, where $\mathcal{E}_k$ represents conditional expectation with respect to the σ-algebra $\mathcal{A}_k$.

i) Prove the recursion formula

$$F_k^m = \mathcal{E}_k(\varphi_k^{m_k} F_{k+1}^{m'}) \qquad (k = 0, 1, \ldots, n-1)$$

where $m' = (m_{k+1} \ldots, m_{n-1})$ for $k < n-1$ and $m' = \emptyset$ for $k = n-1$.

ii) Prove

$$F_0^m = \int_\Omega G\psi_m.$$

9.8. Adapt GFFT for the original Walsh system and the bit-reversed Walsh system and obtain an analogue of Theorem 24 in each case.

9.9. Repeat 9.8 for the two-dimensional Walsh system and for the multidimensional Walsh system.

9.10. Repeat 9.8 for the discrete trigonometric system and the multidimensional discrete trigonometric system.

9.11. Repeat 9.8 for the Vilenkin systems.

9.12. Suppose

$$\mathcal{E}_k(\varphi_k^i \varphi_k^j) = \delta_{ij}$$

for $0 \le i, j < p$. Prove the product system Ψ is orthogonal.

9.13. Let T be a $\mathbf{Z}_2$-linear transformation of $\{0, 1, \ldots, 2^n - 1\}$. Compute the Walsh-Fourier coefficients of $\mathcal{A}^n$-measurable functions with respect to the system

$$w_{T(k)} \quad (0 \le k < 2^n).$$

(Hint: In the notation of Theorem 7 in 1.4 prove

$$\int_0^1 f w_{T(k)} = \int_0^1 (f \circ T^*) w_k = (\widehat{f \circ T^*})(k).)$$

9.14. For $x = (x_0, x_1, \dots)$ and $y = (y_0, y_1 \dots)$ in $\mathbf{G}$ define

$$x \bullet y := (x_0, y_0, x_1, y_1, \dots) \in \mathbf{G}.$$

Prove that the map $(x, y) \to x \bullet y$ is a measure preserving bijection from $\mathbf{G} \times \mathbf{G}$ onto $\mathbf{G}$.

9.15 For $n = n_0 + n_1 2 + \dots$ and $m = m_0 + m_1 2 + \dots$ in $\mathbf{N}$ define

$$n \diamond m := n_0 + m_0 2 + n_1 2^2 + m_1 2^3 + \dots \in \mathbf{N}.$$

i) Prove that the map $(n, m) \to n \diamond m$ is a bijection of $\mathbf{N} \times \mathbf{N}$ onto $\mathbf{N}$.

ii) Let $\widetilde{\boldsymbol{\psi}}$ denote the Kronecker product system generated by the Walsh system $\boldsymbol{\psi} := \widehat{\mathbf{G}}$. Prove

$$\widetilde{\psi}_{n,m}(x, y) = \psi_{n \diamond m}(x \bullet y)$$

for $n, m \in \mathbf{N}$ and $x, y \in \mathbf{G}$.

iii) Let $\Theta : \mathbf{G} \to \mathbf{G}^2$ represent the inverse of the map $(x, y) \to x \bullet y$. Prove that

$$\begin{aligned} \widehat{f}(n, m) &:= \int_{\mathbf{G}^2} f \widetilde{\psi}_{n,m} \, d(\mu \times \mu) \\ &= \int_{\mathbf{G}} (f \circ \Theta) \psi_{n \diamond m} \\ &= (\widehat{f \circ \Theta})(n \diamond m). \end{aligned}$$

iv) Use (iii) to describe FWT for the two-dimensional Walsh system.

APPENDICES

0.0 Banach Spaces. Let F be a field. A *linear space* over F is a set X together with operations

$$+ : X \times X \to X, \quad \cdot : F \times X \to X$$

such that $(X, +)$ is a group (see 0.3) and

$$\begin{aligned} \alpha \cdot (x + y) &= \alpha \cdot x + \alpha \cdot y \\ (\alpha + \beta) \cdot x &= \alpha \cdot x + \beta \cdot x \\ 1 \cdot x &= x \\ \alpha \cdot (\beta \cdot x) &= (\alpha\beta) \cdot x \end{aligned}$$

for all $\alpha, \beta \in F$, $x, y \in X$, where 1 is the unit element of F. If $F := \mathbf{R} := (-\infty, \infty)$ then X is called a real linear space.

A *normed linear space* is (for us) a real linear space X together with a function

$$\|.\|_X := \|.\| : X \to [0, \infty)$$

such that

$$\|x\| = 0 \quad \text{if and only if} \quad x = 0,$$

$$\|\alpha \cdot x\| = |\alpha| \, \|x\|,$$

and

$$\|x + y\| \le \|x\| + \|y\|$$

for all $\alpha \in \mathbf{R}$ and $x, y \in X$. A function $\|.\| : X \to [0, \infty)$ is called a *seminorm* if it enjoys the last two of these three conditions. It is called a *quasi-norm* if it enjoys the first two conditions and, in place of the triangle inequality, there is a constant $C > 0$ such that

$$\|x + y\| \le C(\|x\| + \|y\|)$$

for all $x, y \in X$.

Let X be a normed linear space. A sequence $(x_n, n \in \mathbf{N})$ of elements in X is said to *converge* to y in X if $\|x_n - y\| \to 0$ as $n \to \infty$, and said to be *Cauchy* if $\|x_m - x_n\| \to 0$ as $n, m \to \infty$. Clearly every convergent sequence is Cauchy.

A *Banach space* is a complete normed linear space, i.e., a normed linear space in which all Cauchy sequences are convergent. For any complete measure space (Ω, ν) the spaces $L^p(\Omega)$ are Banach spaces for $1 \le p \le \infty$. The norms are given by

$$\|f\|_p := \left(\int_\Omega |f|^p \, d\nu \right)^{1/p}.$$

For $0 < p < 1$ the map $f \to \|f\|_p^p$ is a quasi-norm, but $L^p(\Omega)$ is still complete.

Let X, Y be normed linear spaces. A function $\Lambda : X \to Y$ is called a *linear operator* if

$$\Lambda(\alpha \cdot x + \beta \cdot y) = \alpha \cdot \Lambda(x) + \beta \cdot \Lambda(y)$$

for all $\alpha, \beta \in \mathbf{R}$, $x, y \in X$. A linear operator Λ from X to Y is said to be *bounded* if its *operator norm*

$$\|\Lambda\| := \sup_{\|x\|_X = 1} \|\Lambda(x)\|_Y = \sup_{x \neq 0} \frac{\|\Lambda(x)\|_Y}{\|x\|_X}$$

is finite. The collection of bounded linear operators from X to Y will be denoted by $\mathcal{L}(X;Y)$. This space is a normed linear space under the operator norm and a Banach space when Y is. Moreover, each $\Lambda \in \mathcal{L}(X;Y)$ is continuous from X to Y .

A *subspace* of a linear space X is a subset Z of X which is closed under the addition and scalar multiplication of X. Thus every subspace of a normed linear space is itself a normed linear space.

A *closed system* in a normed linear space X is a collection $(x_\gamma, \gamma \in \Gamma)$ of elements in X whose *linear hull* is dense in X, i.e., given $x \in X$ there is a sequence of finite linear combinations of the e_γ's which converges to x in the norm of X.

Denote the collection of natural numbers $\{0, 1, \dots\}$ by $\mathbf{N}$. Let $\Lambda_n \in \mathcal{L}(X;Y)$ for $n \in \mathbf{N}$. The Banach-Steinhaus theorem states either

$$\sup_{n \in \mathbf{N}} \|\Lambda_n\| < \infty$$

or there is a dense $\mathcal{G}_\delta$ set E in X such that

$$\sup_{n \in \mathbf{N}} \|\Lambda_n(x)\| = \infty \qquad (x \in E).$$

Thus $\Lambda_n(x)$ converges for every $x \in X$ if and only if there is a closed system Z of X and a constant $M > 0$ such that $\Lambda_n(z)$ converges for each $z \in Z$ and $\|\Lambda_n\| \leq M < \infty$ for all $n \in \mathbf{N}$.

An operator $\Lambda \in \mathcal{L}(X;Y)$ is called *compact* if any sequence $(x_n, n \in \mathbf{N})$ in X with $\|x_n\| \leq 1$ $(n \in \mathbf{N})$ has a subsequence such that $\Lambda(x_{n_k})$ converges in Y as $k \to \infty$. The collection $\mathcal{K}(X;Y)$ of compact operators from X to Y is a closed ideal in $\mathcal{L}(X;Y)$.

The Banach space $X' := \mathcal{L}(X;\mathbf{R})$ is called the *dual* of X, and its elements are called bounded linear functionals. There is natural embedding of X in $X'' := (X')'$, namely evaluation

$$x(x') := x'(x)$$

for $x \in X$ and $x' \in X'$. If $X = X''$ then X is said to be reflexive.

Let X be a normed linear space and Z a subspace of X. The Hahn-Banach theorem states that any $\Lambda \in Z'$ has a norm preserving extension to all of X, i.e., there is a $\Psi \in X'$ such that $\Lambda = \Psi$ on Z and $\|\Lambda\| = \|\Psi\|$. Thus given $x \in X$, x fails to lie in the closure of Z if and only if there is a functional $\Psi \in X'$ such that $\Psi(x) = 1$ but $\Psi = 0$ on Z.

Numbers $1 \le p, p' \le \infty$ are called ***conjugate exponents*** (or ***conjugate indices***) if

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

Let $1 \le p < \infty$ and p' be its conjugate exponent. Let (Ω, ν) be any σ-finite measure space. The Riesz representation theorem states that a functional Λ belongs to $(\mathrm{L}^p(\Omega))'$ if and only if there is a function $\varphi \in \mathrm{L}^{p'}(\Omega)$ such that

$$\|\Lambda\| = \|\varphi\|_{p'}$$

and

$$\Lambda(f) = \int_\Omega f\varphi \, d\nu \qquad (f \in \mathrm{L}^p(\Omega)).$$

Thus the dual of L^p is isometrically isomorphic to $\mathrm{L}^{p'}(\Omega)$ for $1 \le p < \infty$. For the case $\Omega := [0,1)$ and ν is Lebesgue measure we have from the isometry condition that

$$\|\varphi\|_p = \sup_{\|f\|_{p'}=1} \int_0^1 f\varphi$$

for $1 < p \le \infty$ and $\varphi \in \mathrm{L}^p[0,1)$. We shall occasionally use this identity to compute an L^p norm.

The following result is also called the Riesz representation theorem. If Ω is compact Hausdorff space and if $\mathcal{C}(\Omega)$ represents the continuous, real-valued functions on Ω then a functional Λ belongs to $(\mathcal{C}(\Omega))'$ if and only if there is a finite (real) Borel measure ν on Ω such that

$$\Lambda f = \int_\Omega f \, d\nu \qquad (f \in \mathcal{C}(\Omega)).$$

In particular, functionals on $\mathrm{L}^\infty(\Omega)$ can be represented as finite Borel measures through this identification.

The Banach-Alaoglu theorem states that the unit ball in any Banach space is always weak*-compact. The version necessary for us is that if $(\nu_n, n \in \mathbf{N})$ is a sequence of Borel measures on some compact Hausdorff space Ω whose total variations satisfy

$$\sup_{n \in \mathbf{N}} \|\nu_n\| \le M < \infty$$

then there is a subsequence of integers $(n_k, k \in \mathbf{P})$ and a finite Borel measure ν on Ω such that $\|\nu\| \le M$ and

$$\int_\Omega f \, d\nu = \lim_{k \to \infty} \int_\Omega f \, d\nu_{n_k}$$

for all $f \in \mathcal{C}(\Omega)$. This form is easy to verify by diagonalization since $\mathcal{C}(\Omega)$ is separable when Ω is compact.

Given Banach spaces X and Y define

$$(x,y) + (z,w) := (x+z, y+w)$$

$$\alpha \cdot (x, y) := (\alpha \cdot x, \alpha \cdot y)$$

and

$$\|(x, y)\| := \max\{\|x\|_X, \|y\|_Y\}$$

for $(x, y), (z, w) \in X \times Y$ and α real. It is easy to check that these definitions make $X \times Y$ into a Banach space. Moreover, any functional Λ belongs to $(X \times Y)'$ if and only if there exist functionals $\Phi \in X'$ and $\Psi \in Y'$ such that

$$\|\Lambda\| = \|\Phi\| + \|\Psi\|$$

and

$$\Lambda(x, y) = \Phi(x) + \Psi(y)$$

for $(x, y) \in X \times Y$. This observation can be generalized to countable cartesian products in the following way.

Let $(X_n, \|\cdot\|_n), n \in \mathbf{N}$ be a collection of Banach spaces, let $X := X_0 \times X_1 \ldots$ (see 0.4) and define addition and scalar multiplication on X coordinatewise. For $\xi = (\xi_n, n \in \mathbf{N}) \in X$ set

$$\|\xi\|_\infty := \sup_{n \in \mathbf{N}} \|\xi_n\|_n$$

and for $1 \leq p < \infty$ define

$$\|\xi\|_p := \left(\sum_{n=0}^{\infty} \|\xi_n\|_n^p \right)^{1/p} .$$

It is a routine exercise to verify that

$$X^p := \{\xi \in X : \|\xi\|_p < \infty\}$$

is a Banach space for each $1 \leq p \leq \infty$, and that

$$X_0^\infty := \{\xi \in X : \|\xi_n\|_n \to 0 \quad \text{as} \quad n \to \infty\}$$

is a closed subspace of X^∞.

For each $n \in \mathbf{N}$ denote the dual of X_n by $(Y_n, |\cdot|_n)$. Let X_0^∞ and Y^1 be defined as above. Then the dual of X_0^∞ is isometrically isomorphic to Y^1. Indeed, if $\mathbf{\Phi} = (\Phi_n, n \in \mathbf{N})$ belongs to Y^1 then the map

$$\Lambda(\xi) := \sum_{n=0}^{\infty} \Phi_n(\xi_n) \qquad (\xi = (\xi_n, n \in \mathbf{N}) \in X^\infty)$$

is clearly a bounded linear functional on X^∞ with

$$\|\Lambda\| \leq \|\mathbf{\Phi}\|_{Y^1}.$$

On the other hand, if Λ belongs to the dual of X_0^∞ and for each $n \in \mathbf{N}$ we define

$$\Lambda_n(\xi_0, \xi_1, \ldots, \xi_n) := \Lambda(\xi_0, \xi_1, \ldots, \xi_n, 0, 0 \ldots)$$

then Λ_n belongs to the dual of $X_0 \times X_1 \times \cdots \times X_n$. We have seen how to represent the dual of the cartesian product of two Banach spaces. Iterating this representation we find linear functionals $\Phi_j \in Y_j$ $(j = 0, 1, \ldots, n)$ such that

$$\|\Lambda_n\| = \|\Phi_1\| + \cdots + \|\Phi_n\|$$

and

$$\Lambda_n(\xi_0, \xi_1, \ldots, \xi_n) = \sum_{j=0}^{n} \Phi_j(\xi_j)$$

for all $(\xi_0, \xi_1, \ldots, \xi_n) \in X_0 \times X_1 \times \cdots \times X_n$. Since

$$(\xi_0, \xi_1, \ldots, \xi_n, 0, 0 \ldots) \to (\xi_0, \xi_1, \ldots)$$

in X^∞, as $n \to \infty$, for all $\xi = (\xi_n, n \in X) \in X_0^\infty$, and since Λ is continuous on X_0^∞, it follows that Λ has the desired representation. Moreover, since $\|\Lambda\| \geq \|\mathbf{\Phi}\|_{Y^1}$ for

$$\mathbf{\Phi} := (\Phi_n, n \in \mathbf{N}),$$

we have also verified the isometry condition. (A similar representation holds for the duals of X^p when $1 \leq p < \infty$ (see Exercise 3.20)).

0.1 Orthonormal Systems. A *Hilbert space* is a (real) Banach space H whose norm $\| \cdot \|$ arises from an inner product, i.e., there is a function

$$\langle \ , \ \rangle : H \times H \to \mathbf{R}$$

such that

$$\langle \alpha f + \beta g, h \rangle = \alpha \langle f, h \rangle + \beta \langle g, h \rangle,$$

$$(1) \qquad \langle f, g \rangle = \langle g, f \rangle$$

and

$$(2) \qquad \|f\|^2 = \langle f, f \rangle,$$

for all $f, g, h \in H$ and $\alpha, \beta \in \mathbf{R}$.

A *projection* on a Hilbert space H is a linear map $P : H \to H$ which is idempotent and self adjoint, i.e.,

$$P(\alpha f + \beta g) = \alpha P(f) + \beta P(g)$$

$$(3) \qquad P^2(f) := (P \circ P)(f) = P(f),$$

and

$$(4) \qquad \langle Pf, g \rangle = \langle f, Pg \rangle$$

for all $f, g \in H$ and $\alpha, \beta \in \mathbf{R}$. A collection of projections $P_1, P_2, \ldots$ is called orthogonal if

$$(5) \qquad P_i \circ P_j = 0 \qquad (i \neq j).$$

The following result and its corollary will be referred to as Bessel's inequality.

THEOREM 1. *If $P_1, P_2, \ldots$ is an orthogonal collection of projections on a Hilbert space H then*

$$\sum_{i=1}^{\infty} \|P_i f\|^2 \leq \|f\|^2,$$

for all $f \in H$.

PROOF. Fix an integer $n \geq 1$ and observe by (2),(1) and (5) that

$$\begin{aligned} 0 &\leq \|f - \sum_{i=1}^{n} P_i f\|^2 \\ &= \|f\|^2 - \sum_{i=1}^{n} \langle P_i f, f\rangle . \end{aligned}$$

Hence it follows from (3),(4) and (2) that

$$0 \leq \|f\|^2 - \sum_{i=1}^{n} \|P_i f\|^2$$

for any integer $n \geq 1$. The desired inequality follows by letting $n \to \infty$. ∎

A collection $(\psi_n, n \in \mathbf{N})$ of elements of a Hilbert space H is called an orthonormal system if

$$\langle \psi_i, \psi_j \rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases} \tag{6}$$

COROLLARY 1. *If $(\psi_n, n \in \mathbf{N})$ is an orthonormal system in a Hilbert space H, then*

$$\sum_{i=0}^{\infty} |\langle f, \psi_i\rangle|^2 \leq \|f\|^2$$

for all $f \in H$.

PROOF. For each integer $i \geq 0$ set

$$P_i f := \langle f, \psi_i \rangle \psi_i \qquad (f \in H),$$

and verify that $(P_n, n \in \mathbf{N})$ is an orthogonal collection of projections. Since

$$\|P_i f\| = |\langle f, \psi_i \rangle|$$

for $i \in \mathbf{N}$, the desired inequality follows from Theorem 1. ∎

An orthonormal system $(\psi_i, i \in \mathbf{N})$ is called complete if $\langle f, \psi_i \rangle = 0$ for $i = 0, 1, \ldots$ implies $f = 0$ in H. Bessel's inequality becomes an identity for complete orthonormal systems. In fact, THEOREM 2. *Let H be a Hilbert space and $(\psi_n, n \in \mathbf{N})$ be an orthonormal system in H. Then the following four conditions are equivalent:*

$$(\psi_n, n \in \mathbf{N}) \quad \textit{is complete;} \tag{7}$$

$$\begin{cases} \textit{if} \quad S_n f := \sum_{i=0}^{n-1} \langle f, \psi_i \rangle \psi_i & (f \in H, n \in \mathbf{N}), \\ \textit{then} \quad S_n f \to f \quad \textit{in } H \textit{ as } n \to \infty & \textit{for all } f \in H; \end{cases} \tag{8}$$

$$\langle f, g \rangle = \sum_{i=0}^{\infty} \langle f, \psi_i \rangle \langle g, \psi_i \rangle \qquad (f, g \in H); \tag{9}$$

and

$$\|f\|^2 = \sum_{i=0}^{\infty} |\langle f, \psi_i \rangle|^2 \qquad (f \in H). \tag{10}$$

PROOF. Fix f and g in H.

Suppose (7) holds. By orthogonality it is clear that

$$\|S_m f - S_n f\|^2 = \sum_{i=n}^{m-1} |\langle f, \psi_i \rangle|^2$$

for any integers $m > n \geq 0$. It follows from Corollary 1 that the sequence $(S_n f, n \in \mathbf{P})$ is Cauchy in H. Since H is a Banach space, there exists a $g \in H$ such that $S_n f \to g$ as $n \to \infty$. Thus for each integer $i \geq 0$ we have

$$\begin{aligned} 0 &= \lim_{n\to\infty} \langle S_n f - g, \psi_i \rangle \\ &= \lim_{n\to\infty} \sum_{j=0}^{n-1} \langle f, \psi_j \rangle \langle \psi_i, \psi_j \rangle - \langle g, \psi_i \rangle . \end{aligned}$$

In particular, (6) implies $\langle f - g, \psi_i \rangle = 0$. We conclude by (7) that $f - g = 0$, i.e., $S_n f \to f$ in H as $n \to \infty$.

If (8) holds then it is clear that

$$\langle f, g \rangle = \lim_{n\to\infty} \langle S_n f, S_n g \rangle .$$

However,

$$\begin{aligned} \langle S_n f, S_n g \rangle &= \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \langle f, \psi_j \rangle \langle \psi_i, \psi_j \rangle \langle g, \psi_i \rangle \\ &= \sum_{i=0}^{n-1} \langle f, \psi_i \rangle \langle g, \psi_i \rangle \end{aligned}$$

by orthogonality. Therefore, (9) follows by letting $n \to \infty$.

Since (9) implies (10) (set $g = f$), it remains to show that (10) implies (7). But if $\langle f, \psi_i \rangle = 0$ for $i = 0, 1, \ldots$, then (10) implies that $\|f\| = 0$. Hence $f = 0$ and (7) follows by definition. ∎

Equation (9) is called Parseval's identity.

Identity (10) shows that the map $f \to (\langle f, \psi_i \rangle, i \in \mathbf{N})$ takes H into the collection of square summable real sequences. The following result, called the Riesz-Fischer theorem, shows that this map is onto.

THEOREM 3. *Let H be a Hilbert space which contains a complete orthonormal system $(\psi_n, n \in \mathbf{N})$. If $(a_n, n \in \mathbf{N})$ is a sequence of numbers which satisfies*

$$\sum_{i=0}^{\infty} |a_i|^2 < \infty$$

then there is an $f \in H$ such that

$$a_i = \langle f, \psi_i \rangle \qquad (i \in \mathbf{N}), \tag{11}$$

and

$$\|f\| = \left(\sum_{i=0}^{\infty} |a_i|^2 \right)^{1/2} .$$

PROOF. By Theorem 2 it suffices to show (11). For each integer $n \geq 1$ set

$$S_n := \sum_{i=0}^{n-1} a_i \psi_i .$$

Observe by orthogonality that

$$\|S_m - S_n\|_2^2 = \sum_{i=n}^{m-1} |a_i|^2$$

for any integers $m > n \geq 0$. It follows by hypothesis that $(S_n, n \in \mathbf{N})$ is Cauchy in H. In particular, there is an $f \in H$ such that $S_n \to f$ as $n \to \infty$.

Let $n > i$ be integers. By orthogonality it is clear that

$$\langle S_n, \psi_i \rangle = a_i .$$

Hence let $n \to \infty$ to conclude that

$$\langle f, \psi_i \rangle = a_i$$

for $i \in \mathbf{N}$. ∎

The space L^2 of square integrable functions defined on the interval $[0, 1)$ is a Hilbert space with inner product

$$\langle f, g \rangle := \int_0^1 fg .$$

Given a function f integrable on $[0,1]$, its Fourier coefficients with respect to an orthonormal system $(\psi_n, n \in \mathbf{N})$ in L^2 are given by

$$\widehat{f}(i) := \int_0^1 f\psi_i \qquad (i \in \mathbf{N}).$$

Thus by the Riesz-Fischer theorem, an integrable function f belongs to L^2 if and only if its Fourier coefficients with respect to any orthonormal system in L^2 are square summable.

An orthonormal system $(\psi_n, n \in \mathbf{N})$ in L^2 is called uniformly bounded if there is a constant $M > 0$ such that

$$\|\psi_i\|_\infty \le M \qquad (i \in \mathbf{N}). \tag{12}$$

The following result is called the Hausdorff-Young theorem. It is an extension of Bessel's inequality from $p = 2$ to $1 \le p < 2$.

THEOREM 4. *Suppose $\psi_0, \psi_1, \dots$ is a uniformly bounded orthonormal system in L^2. If $f \in \mathrm{L}^p$ for some $1 \le p \le 2$ and q is the index conjugate to p, then*

$$\left(\sum_{i=0}^{\infty} |\widehat{f}(i)|^q\right)^{1/q} \le M^{(2/p-1)} \|f\|_p$$

where M is given by (12).

PROOF. For every $1 \le p < \infty$ let

$$\|\widehat{f}\|_{\ell^p} := \left(\sum_{i=0}^{\infty} |\widehat{f}(i)|^p\right)^{1/p}$$

and let

$$\|\widehat{f}\|_{\ell^\infty} := \sup_{i \in \mathbf{N}} |\widehat{f}(i)|.$$

By Bessel's inequality we have

$$\|\widehat{f}\|_{\ell^2} \le \|f\|_2. \tag{13}$$

By (12), it is clear that

$$\|\widehat{f}\|_{\ell^\infty} \le M\|f\|_1. \tag{14}$$

Fix $1 \le p \le 2$ and choose $0 \le t \le 1$ such that $p = 2/(1+t)$. If q is the exponent conjugate to p then $q = 2/(1-t)$. Thus

$$\frac{1}{p} = \frac{1-t}{2} + t \tag{15}$$

and

$$(16) \qquad \frac{1}{q} = \frac{1-t}{2} + 0.$$

In the next section (see the interpolation result Corollary 3 in 0.2) we shall prove that if $T : \mathrm{L}^1 \to \ell^\infty$ is a linear map such that

$$\|Tf\|_{\ell^2} \leq A\|f\|_2,$$

$$\|Tf\|_{\ell^\infty} \leq B\|f\|_1,$$

and p and q satisfy (15), (16) for some $0 \leq t \leq 1$, then

$$(17) \qquad \|Tf\|_{\ell^q} \leq A^{1-t}B^t\|f\|_p.$$

Consider the map $T : \mathrm{L}^1 \to \ell^\infty$ defined by

$$Tf := (\widehat{f}(i), i \in \mathbf{N}).$$

It is clear that T is linear. Hence by inequalities (13) and (14) and the interpolation result cited above, we have

$$\|\widehat{f}\|_{\ell^q} \leq M^{2/p-1}\|f\|_p. \qquad \blacksquare$$

The Hausdorff-Young theorem cannot be extended to the case $p > 2$. Indeed, for the trigonometric case there exist continuous f such that

$$\|\widehat{f}\|_{\ell^q} = \infty$$

for all $1 \leq q < 2$ (see Zygmund [1], p. 199). This fact also holds for the Walsh case (see the remark following Theorem 10 in 2.4).

0.2 Interpolation Theorems. Let $\mathbf{C}$ represent the set of complex numbers. The three lines theorem from classical complex variables states that if f is holomorphic on the open strip

$$\{z \in \mathbf{C} : Re\,(z) \in (0,1)\},$$

continuous and bounded on the closed strip $\{z \in \mathbf{C} : Re\,(z) \in [0,1]\}$, and

$$M(x) := \sup_{y \in \mathbf{R}} |f(x + \imath y)| \qquad (x \in (0,1))$$

(where $\imath := \sqrt{-1}$), then

$$(18) \qquad M(t) \leq M(0)^{1-t}M(1)^t \qquad (t \in (0,1)).$$

Let $\mathcal{I}$ be a countable index set and let $\mathbf{L}^0$ denote the collection of elements of the form $\boldsymbol{G} = (G_I, I \in \mathcal{I})$, where each G_I is a function measurable on $[0,1]$. For each pair of numbers $1 \le p, q \le \infty$ and each $\boldsymbol{G} \in \mathbf{L}^0$ let

$$\|\boldsymbol{G}\|_{L^p(\ell^q)} := \| \left(\sum_{I\in\mathcal{I}} |G_I|^q \right)^{1/q} \|_{L^p}.$$

It is easy to check that if p' (respectively, q') is the index conjugate to p (respectively, q) then

$$\int_0^1 |\sum_{I\in\mathcal{I}} F_I G_I| \le \|\boldsymbol{F}\|_{L^p(\ell^q)} \|\boldsymbol{G}\|_{L^{p'}(\ell^{q'})}$$

for all $\boldsymbol{F}, \boldsymbol{G} \in \mathbf{L}^0$. This inequality will be called Hölder's inequality for mixed norm spaces.

A linear map $T : \mathrm{L}^1 \to \mathbf{L}^0$ is said to be of type $(\mathrm{L}^r, \mathrm{L}^p(\ell^q))$ for some $1 \le r, p, q \le \infty$ if there exists a constant M, depending only on r, p, q, such that

$$\|Tf\|_{L^p(\ell^q)} \le M \|f\|_{L^r}$$

for all $f \in \mathrm{L}^r$.

Let $1 \le r_i, p_i, q_i \le \infty$ for $i = 0, 1$. The following result shows that if T is of type $(\mathrm{L}^{r_i}, \mathrm{L}^{p_i}(\ell^{q_i}))$ then T is of type $(\mathrm{L}^r, \mathrm{L}^p(\ell^q))$ for all points

$$(\frac{1}{p}, \frac{1}{q}, \frac{1}{r})$$

in $\mathbf{R}^3$ which lie on the line segment between

$$(\frac{1}{p_0}, \frac{1}{q_0}, \frac{1}{r_0}) \quad \text{and} \quad (\frac{1}{p_1}, \frac{1}{q_1}, \frac{1}{r_1}).$$

THEOREM 5. *Let $T : \mathrm{L}^1 \to \mathbf{L}^0$ be a linear map. Suppose $1 \le r_j, p_j, q_j \le \infty$ for $j = 0, 1$, and that*

$$\begin{cases} \dfrac{1}{p} = \dfrac{1-t}{p_0} + \dfrac{t}{p_1}, \\ \dfrac{1}{q} = \dfrac{1-t}{q_0} + \dfrac{t}{q_1}, \\ \dfrac{1}{r} = \dfrac{1-t}{r_0} + \dfrac{t}{r_1} \end{cases} \tag{19}$$

for some $0 < t < 1$. If there exist finite real numbers M_j such that

$$\|Tf\|_{L^{p_j}(\ell^{q_j})} \le M_j \|f\|_{L^{r_j}} \tag{20}$$

for $f \in \mathrm{L}^{r_j}$ and $j = 0, 1$, then

$$\|Tf\|_{L^p(\ell^q)} \le M_0^{1-t} M_1^t \|f\|_{L^r}$$

for all $f \in \mathrm{L}^r$.

PROOF. Fix $f \in \mathrm{L}^r$ and assume without loss of generality that $\|f\|_{L^r} = 1$. Suppose for simplicity that $1 < p, q, r < \infty$.

For any $1 < s < \infty$ let s' be the exponent conjugate to s. Observe that

$$\begin{aligned}\|Tf\|_{L^p(\ell^q)} &= \sup_{\|g\|_{L^{p'}}=1} \int_0^1 g\left(\sum_{I\in\mathcal{I}} |(Tf)_I|^q\right)^{1/q} \\ &= \sup_{\|g\|_{L^{p'}}=1}\left(\sup_{\|\gamma\|_{\ell^{q'}}=1}\int_0^1 g\left(\sum_{I\in\mathcal{I}}\gamma_I(Tf)_I\right)\right).\end{aligned}$$

In particular, it suffices to show that

$$\int_0^1 g\left(\sum_{I\in\mathcal{I}}\gamma_I(Tf)_I\right) \le M_0^{1-t}M_1^t \tag{21}$$

for each $f \in \mathrm{L}^r$, $\boldsymbol{\gamma} := (\gamma_I, I \in \mathcal{I}) \in \ell^{q'}$ and $g \in \mathrm{L}^{p'}$ with

$$\|f\|_{L^r} = \|\boldsymbol{\gamma}\|_{\ell^{q'}} = \|g\|_{L^{p'}} = 1.$$

Moreover, we may assume that both f and g are simple functions, and $\gamma_I = 0$ for all but finitely many $I \in \mathcal{I}$.

To establish (21) fix such $f, \boldsymbol{\gamma}$ and g, and for each complex number z set

$$f_z := |f|^{r((1-z)/r_0+z/r_1)}\operatorname{sgn} f,$$

$$g_z := |f|^{p'((1-z)/p'_0+z/p'_1)}\operatorname{sgn} g,$$

and

$$\boldsymbol{\gamma}_z := (|\gamma_I|^{q'((1-z)/q'_0+z/q'_1)}\operatorname{sgn}\gamma_I,\ I \in \mathcal{I}),$$

where for any real a,

$$\operatorname{sgn} a := \begin{cases} 1 & a > 0 \\ 0 & a = 0 \\ -1 & a < 0. \end{cases}$$

Hypothesis (19) implies $f_t = f$, $g_t = g$, and $\boldsymbol{\gamma}_t = \boldsymbol{\gamma}$. Moreover, the simplifying assumptions imply that

$$\|f_{j+\imath y}\|_{L^{r_j}} = \|g_{j+\imath y}\boldsymbol{\gamma}_{j+\imath y}\|_{L^{p'_j}(\ell^{q'_j})} = 1 \tag{22}$$

for $j = 0, 1$, and all $y \in \mathbf{R}$. For example, by construction

$$|f_{\imath y}| = |f|^{r Re\,((1-\imath y)/r_0+(\imath y)/r_1)} = |f|^{r/r_0},$$

thus

$$\|f_{\imath y}\|_{L^{r_0}} = (\|f\|_r)^{r/r_0} = 1.$$

Consider the function

$$F(z) := \int_0^1 \sum_{I\in\mathcal{I}} (Tf_z)_I(\boldsymbol{\gamma}_z)_I g_z \qquad (z \in \mathbf{C}).$$

Our simplifying assumptions force F to be a finite sum of holomorphic functions; hence F is holomorphic on the infinite open strip $\{z \in \mathbf{C} : Re\, z \in (0,1)\}$ and continuous and bounded on its closure. And, since $f_t = f$, $g_t = g$, $\gamma_t = \boldsymbol{\gamma}$, it is clear that the left side of (21) is precisely $F(t)$. Hence by the three lines theorem it suffices to show

$$\sup_{y\in\mathbf{R}} |F(j+\imath y)| \le M_j$$

for $j = 0, 1$. This inequality follows immediately from hypothesis. Indeed, by (20), (22), and the mixed norm Hölder inequality, we have for $j = 0, 1$ that

$$\begin{aligned}\sup_{y\in\mathbf{R}} |F(j+\imath y)| &= \sup_{y\in\mathbf{R}} \left| \int_0^1 g_{j+\imath y} \sum_{I\in\mathcal{I}} (\boldsymbol{\gamma}_{j+\imath y})_I (Tf_{j+\imath y})_I \right| \\ &\le M_j \sup_{y\in\mathbf{R}} \|f_{j+\imath y}\|_{L^{r_j}} \|g_{j+\imath y}\boldsymbol{\gamma}_{j+\imath y}\|_{L^{p_j'}(\ell^{q_j'})} \\ &\le M_j. \quad \blacksquare\end{aligned}$$

COROLLARY 2. *Let* $\boldsymbol{\Delta}$ *represent the interior in* $\mathbf{R}^2$ *of the triangle whose vertices are* $(0,0), (1/2,1/2)$, *and* $(1,0)$. *Let* T *be a linear map from* L^1 *into* $\mathbf{L}^0$ *which is of type* $(\mathrm{L}^2, \mathrm{L}^2(\ell^2))$ *and of type* $(\mathrm{L}^s, \mathrm{L}^s(\ell^\infty))$ *for each* $1 < s < \infty$. *If*

$$(\frac{1}{p}, \frac{1}{q}) \in \boldsymbol{\Delta}$$

then T *is of type* $(\mathrm{L}^p, \mathrm{L}^p(\ell^q))$. *In fact, if*

$$\|Tf\|_{L^2(\ell^2)} \le A_0 \|f\|_{L^2}$$

and

$$\|Tf\|_{L^s(\ell^\infty)} \le A_1 \|f\|_{L^s}$$

for all $1 < s < \infty$, *then*

$$\|Tf\|_{L^p(\ell^q)} \le A_0^{2/q} A_1^{(q-2)/q} \|f\|_{L^p}.$$

PROOF. Let

$$(\frac{1}{p}, \frac{1}{q}) \in \boldsymbol{\Delta}$$

and set

$$s := \frac{p(2-q)}{p-q}.$$

Notice that $(1/s, 0)$ is the X-intercept of the straight line through $(1/2, 1/2)$ and $(1/p, 1/q)$. In fact, if

$$t = \frac{q-2}{q}$$

then

$$\frac{1}{p} = \frac{1-t}{2} + \frac{t}{s}, \quad \text{and} \quad \frac{1}{q} = \frac{1-t}{2} + 0.$$

In particular, the corollary follows directly from Theorem 5 applied to $p_0 = q_0 = r_0 = 2$ and $p_1 = r_1 = s$, $q_1 = \infty$. ∎

Let L^0 represent the collection of functions measurable and a.e. finite on $[0, 1]$. The following result is called the Riesz-Thorin theorem.

COROLLARY 3. *Let T be a linear map which takes L^1 into L^0 and $1 \le p_j, r_j \le \infty$. Suppose there exist finite numbers M_j such that*

$$\|Tf\|_{L^{p_j}} \le M_j \|f\|_{L^{r_j}}$$

for $f \in L^{r_j}$ and $j = 0, 1$. If

$$\frac{1}{p} = \frac{1-t}{p_0} + \frac{t}{p_1}$$

and

$$\frac{1}{r} = \frac{1-t}{r_0} + \frac{t}{r_1}$$

for some $0 < t < 1$ then

$$\|Tf\|_{L^p} \le M_0^{1-t} M_1^t \|f\|_{L^r}$$

for all $f \in L^r$.

PROOF. Let $\mathcal{I} = \{1\}$, identify $\mathbf{L}^0$ with L^0, $L^p(\ell^1)$ with L^p, and apply Theorem 5 to $L^p(\ell^1)$ and L^r. ∎

The Riesz-Thorin theorem can be generalized. For each measurable set E of real numbers, let $|E|$ denote the Lebesgue measure of E and

$$\chi(E) := \begin{cases} 1 & x \in E \\ 0 & x \notin E \end{cases}$$

represent the characteristic function of E. For each measurable g define the distribution function λ_g on $(0, \infty)$ by

$$\lambda_g(y) := |\{x \in [0, 1) : |g(x)| > y\}|.$$

Notice for $0 < p < \infty$ that

$$\int_0^1 |g|^p = \int_0^1 \int_0^{|g(x)|} py^{p-1}\,dydx$$
$$= \int_0^\infty \int_0^1 py^{p-1}\chi([0,\ |g(x)|\,])(y)\,dxdy.$$

Consequently,

$$\|g\|_p^p = \int_0^\infty py^{p-1}\lambda_g(y)\,dy \tag{23}$$

for all measurable g and all $0 < p < \infty$.

A map T from L^p into L^0 is said to be of weak type (p, q) for some $1 \le p \le \infty$ and $1 \le q < \infty$ if there is a constant $A > 0$ such that

$$\lambda_{Tf}(y) \le \left(\frac{A}{y}\|f\|_p\right)^q$$

for all $f \in L^p$ and all $y > 0$. Notice by (23) that

$$\lambda_g(y) \le \left(\frac{\|g\|_p}{y}\right)^p$$

for all $y > 0, 0 < p < \infty$, and measurable g. Hence any map of type (L^p, L^q) is necessarily of weak type (p, q).

An operator T which maps a linear space of measurable functions on [0,1) in the collection of measurable functions on [0,1) is called ***sublinear*** if

$$|T(f+g)| \le |Tf| + |Tg|$$

a.e. on [0,1) and

$$|T(\alpha f)| = |\alpha|\,|Tf|$$

for all scalars α and all f in the domain of T. The following generalization of the Riesz-Thorin theorem is called the Marcinkiewicz interpolation theorem.

THEOREM 6. *Let T be sublinear and of weak type $(p_0, q_0), (p_1, q_1)$ for some parameters $1 \le p_i < \infty$, $1 \le q_i < \infty$ with $p_i \le q_i$ $(i = 0, 1)$. If p and q are defined by (19) for some $0 < t < 1$ then T is of type (L^p, L^q), i.e., there is a constant $B > 0$ such that*

$$\|Tf\|_q \le B\|f\|_p$$

for all $f \in L^p$.

PROOF. By symmetry we may suppose that $p_0 > p_1$. For simplicity we also suppose that $p_0 = q_0$ and $p_1 = q_1$, hence $p = q$. (These are the only cases used in our book.)

Fix f and suppose for a moment that $f = f_0 + f_1$. Since T is sublinear we have

$$\lambda_{Tf}(2y) \leq \lambda_{Tf_0}(y) + \lambda_{Tf_1}(y) \qquad (0 < y < \infty).$$

Hence (23) and hypothesis imply

$$\begin{aligned}\|Tf\|_p^p &= \int_0^\infty p y^{p-1} \lambda_{Tf}(y)\, dy \\ &= p2^p \int_0^\infty y^{p-1} \lambda_{Tf}(2y)\, dy \\ &\leq p2^p (A_0^{p_0} I_0 + A_1^p I_1)\end{aligned}$$

where for each $i = 0$ or 1,

$$I_i := \int_0^\infty \|f_i\|_{p_i}^{p_i}\, d\nu_i(y)$$

and

$$d\nu_i(y) := y^{p-p_i-1}\, dy.$$

To estimate I_0 use (23) to write

$$I_0 = p_0 \int_0^\infty \int_0^\infty t^{p_0-1} \lambda_{f_0}(t)\, dt\, d\nu_0(y).$$

Specify f_0 (hence f_1) by

$$f_0(x) := \begin{cases} f(x) & |f(x)| \leq y \\ y & \text{otherwise.} \end{cases}$$

Since $\lambda_{f_0}(t) = 0$ for $t > y$ and $\lambda_{f_0}(t) = \lambda_f(t)$ for $t \leq y$ we have by Fubini's theorem that

$$\begin{aligned}I_0 &= p_0 \int_0^\infty \int_0^y t^{p_0-1} \lambda_{f_0}(t)\, dt\, d\nu_0(y) \\ &= p_0 \int_0^\infty \int_t^\infty t^{p_0-1} \lambda_f(t)\, d\nu_0(y)\, dt.\end{aligned}$$

It follows from the definition of ν_0 and (23) that

$$\begin{aligned}|I_0| &\leq \frac{p_0}{p_0 - p} \int_0^\infty t^{p_0-1+(p-p_0)} \lambda_f(t)\, dt \\ &= \frac{p_0}{p(p_0 - p)} \|f\|_p^p.\end{aligned}$$

Similarly,

$$|I_1| \leq \frac{p_1}{p(p - p_1)} \|f\|_p^p.$$

Consequently,

$$\|Tf\|_p^p \le 2^p \left(\frac{p_0 A_0^{p_0}}{p_0 - p} + \frac{p_1 A_1^{p_1}}{p - p_1} \right) \|f\|_p^p.$$

We conclude that the theorem holds in this case with

$$B = 2 \left(\frac{p_0 A_0^{p_0}}{p_0 - p} + \frac{p_1 A_1^{p_1}}{p - p_1} \right)^{1/p}. \quad \blacksquare$$

We notice without difficulty that these arguments are valid for any measure space in which the Riesz representation theorem holds. In particular, Theorems 5 and 6 hold if [0,1) is replaced by any σ-finite measure space.

0.3 Locally Compact Abelian Groups. Let G be an *abelian group*, i.e., a set together with a binary operation $+$ defined on G such that

$$x + y = y + x \qquad (x, y \in G),$$
$$x + (y + z) = (x + y) + z \qquad (x, y, z \in G),$$

there is an element $0 \in G$ satisfying

$$x + 0 = x \qquad (x \in G),$$

and, given $x \in G$ there is an element $-x \in G$ such that

$$x - x := x + (-x) = 0.$$

Let H be a subgroup of G, i.e., a subset of G which satisfies

$$x + y \in H \qquad (x, y \in H).$$

The sets

$$H + x := \{x + y : y \in H\}$$

are called *cosets* of H.

Denote the collection of cosets of H by

$$G/H := \{H + x : x \in G\}.$$

Define equality and addition in G/H as follows:

$$H + x = H + y \quad \text{if and only if} \quad x - y \in H$$

and

$$(H + x) + (H + y) := H + (x + y) \qquad (x, y \in G).$$

It is easy to see that G/H is an abelian group, the so-called *quotient group* of G modulo H. The map $x \to H + x$ is called the ***natural homomorphism*** of G onto G/H.

A group G is said to be ***locally compact*** if there is a locally compact Hausdorff topology on G for which the map $(x, y) \to x - y$ is continuous from the product space $G \times G$ onto G. It follows for such a group that the ***translation map*** $x \to x + y$ is continuous for each $y \in G$, and the ***reflection*** $x \to -x$ is continuous on G. Hence if W is open in G and contains 0 then there is a compact neighborhood V of 0 such that

$$V + V \subset W.$$

Also, if W is open in G and E is any subset of G, it follows that

$$W + E := \{x + y : x \in W, y \in E\}$$

is open, for it is a union of open sets:

$$W + E = \bigcup_{y \in E} (W + y).$$

And, if K_1, K_2 are compact subsets of G then $K_1 + K_2$ is compact, since it is the image under the map $(x, y) \to x + y$ of the compact set $K_1 \times K_2$.

A set $E \subseteq G$ is called *symmetric* if $E = -E$. Since $E \cap (-E)$ is always symmetric, it is clear that every locally compact abelian group has a neighborhood base at 0 which consists of compact, symmetric sets. Also notice since $+$ is continuous, that given any neighborhood W of $0 \in G$ there is a neighborhood V of $0 \in G$ such that $V + V \subset W$.

Let G be a locally compact abelian group, H be a closed subgroup of G, and

$$h : G \to G/H$$

be the natural homomorphism. A set $U \subseteq G/H$ is said to be open if $U = h(V)$ for some V open in G. It is obvious that the collection of open subsets of G/H form a topology. Since by definition the natural homomorphism is both continuous and open, it follows that G/H is a locally compact topological space. Moreover, since

$$h(x - y) = (H + x) - (H + y)$$

it is easy to see that the group operations on G/H are continuous. Hence G/H is itself a locally compact abelian group if we can show that the topology on G/H is Hausdorff. To this end, suppose $H + x \neq H + y$ in G/H. Then $x - y \notin H$. Since $y + H$ is closed in G it follows that there exists a neighborhood W of $0 \in G$ such that

$$(x + W) \bigcap (y + H) = \emptyset.$$

Choose a symmetric, compact neighborhood V of $0 \in G$ such that $V + V \subset W$. In particular, we have

$$(x + H + V) \bigcap (y + H + V) = \emptyset.$$

Since $h(x+H+V)$ and $h(y+H+V)$ are open in G/H, we conclude that $H+x$ and $H+y$ have disjoint neighborhoods in G/H, i.e., G/H is Hausdorff.

A *Haar measure* on a topological abelian group G is a non- trivial non-negative, regular Borel measure μ on G which is *translation invariant*, i.e., such that

$$\mu(E+y) = \mu(E)$$

for all $y \in G$ and all Borel sets $E \subseteq G$.

It is a fact that every locally compact group (abelian or not) has a Haar measure (see Pontryagin [1], for example). For a slick construction of such a measure see Bredon [1]. In the abelian case, the construction can be carried out efficiently using Kakutani's fixed point theorem (see Rudin [2]).

We shall not construct Haar measure here because for the groups we consider, the construction is trivial. By regularity, Haar measure μ is finite when G is compact.

It is interesting to note that Haar measure is unique up to constant multiples. Indeed, suppose that μ and $\tilde{\mu}$ are both Haar measures on a locally compact abelian group G. Let f and g be functions continuous on G with compact support, and choose g so that

$$\int_G g\, d\tilde{\mu} = 1.$$

By translation invariance of μ we have

$$\int_G f\, d\mu = \int_G g(y) \int_G f(x+y)\, d\mu(x)\, d\tilde{\mu}(y).$$

It follows from Fubini's theorem and translation invariance of $\tilde{\mu}$ that

$$\int_G f\, d\mu = \int_G \int_G f(y) g(y-x)\, d\tilde{\mu}(y)\, d\mu(x) = \int_G f(y)\, d\tilde{\mu}(y) \int_G g(-x)\, d\mu(x).$$

Hence $\mu = c\tilde{\mu}$ for

$$c := \int_G g(-x)\, d\mu(x).$$

When G is compact we normalize its Haar measure μ by requiring that $\mu(G) = 1$. For this reason μ is called the Haar measure of a compact G, and we shall write

$$\mathrm{L}^1(G) := \mathrm{L}^1_\mu(G)$$

for the collection of real-valued integrable functions on G.

A *character* of G is a continuous function $\psi : G \to \mathbf{C}$ which satisfies

$$\psi(x+y) = \psi(x)\psi(y) \qquad (x, y \in G),$$

and

$$|\psi(x)| = 1 \qquad (x \in G).$$

Thus every character of G is a continuous group homomorphism from G into the *circle group*

$$\mathbf{T} := \{z \in \mathbf{C} : |z| = 1\}.$$

Let $\widehat{G}$ represent the set of characters of G. It is clear that $\widehat{G}$ contains the identity function

$$\psi_0(x) := 1 \qquad (x \in G),$$

and it is easy to see that $\widehat{G}$ is an abelian group under the operation

$$(\psi + \phi)(x) := \psi(x)\phi(x)$$

for $x \in G$, and $\psi, \phi \in \widehat{G}$. Thus, if $\overline{z}$ represents the complex conjugate of an element $z \in \mathbf{C}$, we have that

$$-\psi(x) = \overline{\psi(x)}.$$

The group $\widehat{G}$ satisfies the following orthogonality condition.

LEMMA 1. *If $\psi, \phi \in \widehat{G}$ and $\psi \neq \phi$ then*

$$\int_G \psi\overline{\phi}\, d\mu = 0.$$

PROOF. Fix $y \in G$ such that $\psi(y)\overline{\phi(y)} \neq 1$. Since μ is translation invariant we have

$$\int_G \psi\overline{\phi}\, d\mu = \int_G \psi(x+y)\overline{\phi(x+y)}\, d\mu(x) = \psi(y)\overline{\phi(y)} \int_G \psi\overline{\phi}\, d\mu.$$

The choice of y implies that this integral must be zero. ∎

For each compact set $K \subseteq G$ and each real number $r > 0$ let

$$N(K, r) := \{\psi \in \widehat{G} : |1 - \psi(x)| < r \quad \text{for all} \quad x \in K\}.$$

Then the identity ψ_0 of $\widehat{G}$ belongs to $N(K, r)$. Moreover, any finite intersection of sets of the form $N(K, r)$ contains a set of the form $N(K, r)$. Consequently, these sets can be used as a base for a topology of $\widehat{G}$ at ψ_0.

Accordingly, a subset $V \subseteq \widehat{G}$ is called *open* if given $\psi \in V$ there is a compact set $K \subseteq G$ and a real number $r > 0$ such that

$$\psi + N(K, r) \subseteq V.$$

It is clear that the map $(\psi, \phi) \to \psi - \phi$ is continuous from $\widehat{G} \times \widehat{G}$ into $\widehat{G}$, since

$$(\psi + N(K, \frac{r}{2})) - (\phi + N(K, \frac{r}{2})) \subseteq (\psi - \phi) + N(K, r).$$

Therefore $\widehat{G}$ is a topological abelian group.

One can show that $\widehat{G}$ is actually locally compact. In particular, it makes sense to talk about $(\widehat{G})\widehat{\ }$. As locally compact abelian groups, it turns out that G and $(\widehat{G})\widehat{\ }$ are isomorphic (see, for example, Hewitt and Ross [1] or Rudin [1]). This result is usually referred to as Pontryagin duality (see Pontryagin [1]).

We do not need such heavy machinery since in our situation the group $\widehat{G}$ is discrete. Indeed, THEOREM 7. *If G is a metrizable, compact, abelian group then $\widehat{G}$ is a countable discrete group.*

PROOF. For each $f \in \mathrm{L}^1(G)$ let

$$\widehat{f}(\psi) := \int_G f\overline{\psi}\,d\mu$$

where μ is normalized Haar measure on G. Then $\widehat{f} : \widehat{G} \to \mathbf{C}$ and it is easy to see that $\widehat{f}$ is continuous on $\widehat{G}$. Indeed, let $\varepsilon > 0$, set $K := G$, and define a real number r by

$$\frac{1}{r} := \int_G |f|\,d\mu.$$

Then for any $\psi \in N(K, r\varepsilon)$ it is clear that

$$\begin{aligned}|\widehat{f}(\psi) - \widehat{f}(\psi_0)| &= |\int_G f(\overline{\psi} - 1)\,d\mu| \\ &\leq \|f\|_{L^1(G)}\|\overline{\psi} - 1\|_{L^\infty(G)} \\ &< \varepsilon.\end{aligned}$$

Since $\widehat{f}$ is continuous on $\widehat{G}$, the set

$$V_f := \{\phi \in \widehat{G} : \widehat{f}(\phi) \in (\frac{1}{2}, \frac{3}{2})\}$$

is open in $\widehat{G}$ for all $f \in \mathrm{L}^1(G)$. But $\psi_0 := 1$ belongs to $\mathrm{L}^1(G)$ for compact G. And by orthogonality (see Lemma 1 above), we have

$$\widehat{\psi_0}(\phi) = \int_G \psi_0\overline{\phi}\,d\mu = \begin{cases} 0 & \psi_0 \neq \phi \\ 1 & \psi_0 = \phi. \end{cases}$$

In particular, for the case $f = \psi_0$ we have

$$V_f = \{\psi_0\}.$$

It follows that $\{\psi_0\}$ is open. Thus $\widehat{G}$ is discrete.

To see that $\widehat{G}$ is countable let $\mathcal{C}(G)$ denote the collection of continuous, complex-valued functions on G. Observe that $\widehat{G} \subset \mathcal{C}(G)$, and that $\mathcal{C}(G)$ is separable since G is metrizable. Consequently, $\widehat{G}$ will be countable if we show the distance between any two distinct characters ψ, ϕ in $\widehat{G}$ is bounded away from zero. However, this follows immediately from orthogonality. Indeed,

$$\|\psi - \phi\|^2_{L^\infty(G)} \geq \int_G |\psi - \phi|^2\,d\mu = 2. \quad \blacksquare$$

COROLLARY 4. *If G is a compact, metrizable abelian group then $\widehat{G}$ is at most countable, a locally compact abelian group, and an orthonormal system in $\mathrm{L}^2(G)$.*

Given a subgroup H of a topological group G, the *annihilator* of H is defined by

$$\mathcal{N}(H) := \{\psi \in \widehat{G} : \psi(x) = 1 \quad \text{for all} \quad x \in H\}.$$

There is an interesting relationship between quotient groups and annihilators.

THEOREM 8. *Let H be a closed subgroup of a locally compact abelian group G. Then there is a 1-1 group homomorphism Λ which takes $\mathcal{N}(H)$ onto $\widehat{G/H}$.*

PROOF. Let $\psi \in \mathcal{N}(H)$ and define

$$\Lambda(\psi)(H+x) := \psi(x) \qquad (x \in G).$$

Since ψ belongs to the annihilator of H, $\Lambda(\psi)$ is well defined. Indeed, if $H + x = H + y$ then $\psi(x - y) = 1$.

Since the natural homomorphism $x \to H + x$ is a homeomorphism, it is evident that $\Lambda(\psi)$ is a character on G/H. Thus Λ takes $\mathcal{N}(H)$ into $\widehat{G/H}$.

To show Λ is 1-1, suppose $\Lambda(\psi) = \Lambda(\phi)$ for some $\psi, \phi \in \mathcal{N}(H)$. Let $x \in G$ and choose $x_0 \in G, y_0 \in H$ such that $x = x_0 + y_0$. Since $\psi, \phi \in \mathcal{N}(H)$, we have that

$$\begin{aligned}\psi(x) &= \psi(x_0)\psi(y_0)\\ &= \Lambda(\psi)(H+x_0)\\ &= \Lambda(\phi)(H+x_0)\\ &= \phi(x_0)\phi(y_0)\\ &= \phi(x).\end{aligned}$$

Hence $\psi = \phi$ on G.

The definition of character addition shows that Λ is a group homomorphism. Hence it remains to see that Λ is onto.

Let $\gamma \in \widehat{G/H}$. To find the preimage of γ under Λ, fix $x \in G$, choose $x_0 \in G, y_0 \in H$ such that $x = x_0 + y_0$, and set

$$\psi(x) := \gamma(H + x_0).$$

The function ψ does not depend on the decomposition $x_0 + y_0$. Indeed, if $x = x_1 + y_1$ for some $x_1 \in G$ and $y_1 \in H$, then $x_0 - x_1 \in H$, whence the coset

$$H + x_0 - x_1$$

is the identity of G/H. In particular,

$$\gamma(H + x_0 - x_1) = 1,$$

and it follows that

$$\gamma(H + x_0) = \gamma(H + x_1).$$

Thus ψ is well-defined.

Since the natural homomorphism $x \to H + x$ is a homeomorphism, it is clear that ψ belongs to $\widehat{G}$. Also, by definition, $\Lambda(\psi) = \gamma$. Hence it remains to see that ψ belongs to the annihilator of H. But if $x \in H$ then $H + x$ is the zero of G/H. Hence

$$\psi(x) = \gamma(H + x) = 1,$$

so $\psi \in \mathcal{N}(H)$. ∎

The map Λ in Theorem 8 turns out to be a homeomorphism. This is easy to see when G is compact and metrizable since both G/H and $\mathcal{N}(H)$ are discrete.

Let

$$m = \sum_{j=0}^{\infty} m_j 2^j, \quad n = \sum_{j=0}^{\infty} n_j 2^j$$

represent the binary expansions of m and n. Define the dyadic sum of the pair $m, n \in \mathbf{N}$ by

$$m \oplus n := \sum_{j=0}^{\infty} |m_j - n_j| 2^j.$$

Define the dyadic product of the pair $m, n \in \mathbf{N}$ by

$$m \odot n := \sum_{k=0}^{\infty} \ell_k 2^k$$

where

$$\ell_k := \sum_{j=0}^{k} m_j n_{k-j} \pmod 2$$

for $k \in \mathbf{N}$.

Given $m \in \mathbf{N}$ we shall denote by $\vartheta(m)$ the unique $M \in \mathbf{N}$ which satisfies $2^M \leq m < 2^{M+1}$. Define a norm on $\mathbf{N}$ by

$$\|m\| := 2^{\vartheta(m)}$$

for $m \in \mathbf{P}$ and

$$\|0\| = 0.$$

It is easy to see that

$$1 \odot n = n \odot 1 = n$$

$$\|m \oplus n\| \leq \max\{\|m\|, \|n\|\}$$

and

$$\|m \odot n\| = \|m\| \, \|n\|$$

for all $m, n \in \mathbf{N}$. Hence $(\mathbf{N}, \oplus, \odot)$ is a normed algebra with identity.

This algebra enjoys a division algorithm. Namely, if $m, n \in \mathbf{N}$ and $n \neq 0$ there exist unique $q, r \in \mathbf{N}$ such that

$$m = q \odot n \oplus r$$

with $0 \leq \|r\| < \|n\|$.

To prove existence let $2^M = \|m\|$, $2^N = \|n\|$ and $2^Q = \|q\|$. Notice by definition that $\|q \odot n\| = \|q\| + \|n\|$. Thus we must find binary coefficients of q and r which satisfy

$$m_{M-i} = \sum_{j=0}^{i} n_{N-i+j} q_{Q-j} \pmod 2 \qquad (0 \leq i \leq M - N)$$

and

$$m_s = \sum_{j=0}^{s} n_{s-j} q_j + r_s \pmod 2 \qquad (0 \le s \le N-1)$$

where $n_t := 0$ for $t < 0$. By definition $n_N = 1$. Thus the matrix of the first system is triangular and every entry on the main diagonal is 1. It follows that the first system can be solved recursively to obtain the binary coefficients of q, hence q itself. Putting these values into the second system, we can solve for r. Moreover, the proof shows the solutions are unique. Hence the division algorithm is verified.

Let $m, n \in \mathbf{N}$. We shall call n a *dyadic divisor* of m if there is a $q \in \mathbf{N}$ such that

$$m = n \odot q.$$

We shall abbreviate "n is a dyadic divisor of m" by $n \| m$. Notice that every integer $m \in \mathbf{N}$ has at least two dyadic divisors, 1 and m. If it has no other dyadic divisors then it is called a *dyadic prime*.

An integer $r \in \mathbf{N}$ is called a *greatest common dyadic divisor* of a pair $m, n \in \mathbf{N}$ if $r \| m$, $r \| n$ and if $s \| r$ for any divisor s of m and n. Since each pair $m, n \in \mathbf{N}$ can have at most one greatest common dyadic divisor, we shall denote it by (m, n) when it exists.

Using the division algorithm it is easy to construct a *dyadic Euclid's algorithm* and thus show that (m, n) exists for every pair $m, n \in \mathbf{N}$. Indeed, suppose $\|m\| \ge \|n\| > 0$. Set $r_{-2} := m$, $r_{-1} := n$. Define $r_j \in \mathbf{N}$ recursively by

$$r_{j-2} := q_j \odot r_{j-1} \oplus r_j, \qquad \|r_j\| < \|r_{j-1}\|.$$

Since $\|r_j\|$ is strictly decreasing as j increases, there is a $k \in \mathbf{N}$ such that $\|r_k\| \ne 0$ but $\|r_{k+1}\| = 0$. The number $r := r_k$ is the greatest common dyadic divisor of m and n.

If $m, n \in \mathbf{N}$ and p is a dyadic prime then

$$p \| (m \odot n) \quad \text{implies} \quad p \| m \quad \text{or} \quad p \| n.$$

Indeed, if p is not a dyadic divisor of m then $(p, m) = 1$. Consequently, by Euclid's algorithm there exist $p', m' \in \mathbf{N}$ such that

$$p \odot p' \oplus m \odot m' = 1.$$

Multiplying by n we obtain,

$$p \odot (p' \odot n) \oplus (m \odot m') \odot n = n.$$

Hence by definition, $p \| n$.

Therefore, every positive integer has a unique dyadic prime factorization. Specifically, let $\mathbf{P}_r$ denote the collection of dyadic primes in $\mathbf{N}$. Given $m \in \mathbf{P}$ there are integers $\alpha_p \in \mathbf{N}$ for every $p \in \mathbf{P}_r$ such that

$$m = \dot{\prod_{p \in \mathbf{P}_r}} p^{\alpha_p}$$

where $\dot{\prod}$ represents the dyadic product, and p^{α_p} represents the dyadic product of p with itself α_p times (the empty product is 1). Using the multiplicativity of the dyadic norm, we notice that

$$\vartheta(m) = \sum_{p \in \mathbf{P}_r} \alpha_p \vartheta(p).$$

This observation gives us some information about the distribution of dyadic primes. For each $k \in \mathbf{N}$ let M_k represent the number of dyadic primes p which satisfy

$$\|p\| = 2^k.$$

We shall presently prove that $M_1 = 2,\ M_2 = 1,\ M_3 = 2$ and

$$M_k > \frac{2^k}{k} - 2^{k/2}$$

for $k \geq 4$. In particular, for each $k \in \mathbf{N}$ there exists a dyadic prime p such that $\|p\| = 2^k$.

To see this fix $x \in \mathbf{R}$ with $|x| < 1/4$ and observe by the prime factorization of m and the formula for $\vartheta(m)$ above that

$$\sum_{m=1}^{\infty} x^{\vartheta(m)} = \prod_{p \in \mathbf{P}_r} \left(\sum_{k=0}^{\infty} x^{k\vartheta(p)} \right).$$

By definition of ϑ this identity is equivalent to

$$\begin{aligned}
\frac{1}{1-2x} &= \sum_{n=0}^{\infty} 2^n x^n \\
&= \prod_{p \in \mathbf{P}_r} \frac{1}{1 - x^{\vartheta(p)}} \\
&= \prod_{k \in \mathbf{N}} \left(\prod_{p \in \mathbf{P}_r,\, \|p\| = 2^k} \frac{1}{1-x^k} \right) \\
&= \prod_{k \in \mathbf{N}} \frac{1}{1-x^k} M_k.
\end{aligned}$$

Consequently,

$$-\log(1-2x) = -\sum_{k=0}^{\infty} M_k \log(1-x^k).$$

Since $M_k \leq 2^k$ and $|\log(1-u)| \leq 2u$ for $|u| \leq 1/2$, this series converges absolutely and uniformly when $|x| \leq 1/4$ and the steps taken before are justified. Moreover, term by term differentiation yields

$$\frac{2}{1-2x} = \sum_{k=0}^{\infty} \frac{kx^{k-1} M_k}{1-x^k}.$$

Therefore,

$$\begin{aligned}\sum_{k=1}^{\infty} 2^k x^k &= \frac{2x}{1-2x} \\ &= \sum_{k=0}^{\infty} \frac{kx^k M_k}{1-x^k} \\ &= \sum_{s=0}^{\infty} sM_s \left(\sum_{\ell=1}^{\infty} x^{s\ell} \right).\end{aligned}$$

By comparing coefficients we conclude that

$$2^k = \sum_{s|k} sM_s$$

where $s|k$ means s is a divisor of k in the usual sense.

It is now clear that $M_1 = M_3 = 2$ and $M_2 = 1$. Moreover, for $k \geq 4$ we have

$$\begin{aligned}2^k &\leq kM_k + \sum_{1 \leq s \leq k/2} s2^s \\ &< k(M_k + 2^{k/2}).\end{aligned}$$

In particular,

$$M_k > \frac{2^k}{k} - 2^{k/2}$$

for $k \geq 4$.

Let $k \in \mathbf{N}$ and let p be a dyadic prime which satisfies $\|p\| = 2^k$. It is now easy to see that the finite group

$$\mathbf{F}_k := \{m \in \mathbf{N} : 0 \leq m < 2^k\}$$

is a field under dyadic addition and dyadic multiplication modulo p. Indeed, by the division algorithm for every $i, j \in \mathbf{F}_k$ there is a unique $q, r \in \mathbf{N}$ with $r \in \mathbf{F}_k$ such that

$$i \odot j = q \odot p \oplus r.$$

Define the product of i and j by

$$i \circ j = r.$$

Then $\circ$ is commutative, associative, and distributes over $\oplus$. Moreover,

$$1 \circ i = i$$

for all $i \in \mathbf{F}_k$ so 1 is the identity of $\mathbf{F}_k$. Thus it remains to see that every non-zero element of $\mathbf{F}_k$ has an inverse. But if $i \in \mathbf{F}_k$ and $i \neq 0$ then by Euclid's algorithm there exist $i' \in \mathbf{F}_k$ and $q \in \mathbf{N}$ such that

$$i \odot i' \oplus p \odot q = 1.$$

Hence by definition,

$$i \circ i' = 1$$

so i' is the multiplicative inverse of i.

0.4 Product Spaces. Let $X_1, X_2, \ldots$ be a sequence of sets. The infinite cartesian product

$$X := X_1 \times X_2 \times \ldots$$

is the collection of elements $(x_1, x_2, \ldots)$ such that $x_j \in X_j$ for $j = 1, 2, \ldots$.

If each X_j is a topological space, the ***product topology*** on X is defined in the following way. A subset V of X is called a basic open set if

$$V = \{(x_1, x_2, \ldots) : x_{j_k} \in V_k,\ 1 \leq k \leq n\}, \tag{24}$$

for some integer $n \geq 1$, integers $j_1 < j_2 < \cdots < j_n$ and open sets $V_k \subseteq X_{j_k}$. A subset of X is called open if it is a union of basic open sets.

It is easy to see that the product topology on X is Hausdorff when each X_j is a Hausdorff space. Much more is true:

TYCHONOFF'S THEOREM. *If $X_1, X_2, \ldots$ is a sequence of compact Hausdorff spaces, then the product space X is a compact Hausdorff space.*

Suppose that $(X_j, \mathcal{M}_j, \mu_j)$ is a non-negative measure space for $j = 1, 2, \ldots$ Recall that a subset

$$R \subset X_1 \times \cdots \times X_n$$

is called an ***n-dimensional rectangle*** if R can be written as

$$R = B_1 \times \cdots \times B_n$$

where $B_j \in \mathcal{M}_j$ for $1 \leq j \leq n$. Denote the smallest σ-algebra containing the n-dimensional rectangles by $\mathcal{M}_1 \times \cdots \times \mathcal{M}_n$. It is well known that there is a non-negative measure $\mu_1 \times \cdots \times \mu_n$ defined on $\mathcal{M}_1 \times \cdots \times \mathcal{M}_n$ such that

$$(\mu_1 \times \cdots \times \mu_n)(R) = \mu_1(B_1) \ldots \mu_n(B_n)$$

for every n-dimensional rectangle $R = B_1 \times \cdots \times B_n$.

A subset E of $X := X_1 \times X_2 \times \ldots$ is called a ***measurable cylinder*** with n-dimensional base B if

$$E = \{(x_1, x_2, \ldots) \in X : (x_1, x_2, \ldots, x_n) \in B\}$$

for some $B \in \mathcal{M}_1 \times \cdots \times \mathcal{M}_n$. The smallest σ-algebra containing the measurable cylinders of X will denoted by $\mathcal{M}$.

Product measure μ on X is constructed in the proof of the following result.

THEOREM 9. *For each $j \in \mathbf{P}$ let $(X_j, \mathcal{M}_j, \mu_j)$ be a non-negative measure space with $\|\mu_j\| = 1$. There exists a unique non-negative measure μ defined on $\mathcal{M}$ which satisfies $\|\mu\| = 1$ such that*

$$\mu(E) = (\mu_1 \times \cdots \times \mu_n)(B) \tag{25}$$

for every measurable cylinder E of X with n-dimensional base B.

PROOF. Define μ on measurable cylinders by (25). It is easy to see that μ does not depend on the representation of E. Indeed, suppose that E can also be represented with m-dimensional base B'. Assuming without loss of generality that $m < n$, it is clear that

$$B = B' \times X_{m+1} \times \cdots \times X_n.$$

Hence hypothesis $\|\mu_j\| = 1$ implies

$$(\mu_1 \times \cdots \times \mu_m)(B) = (\mu_1 \times \cdots \times \mu_m)(B'),$$

and μ is well defined.

Let $\mathcal{A}$ represent the algebra of measurable cylinders in X, and observe that μ is finitely additive on $\mathcal{A}$. Hence by the Caratheodory extension theorem, it suffices to show that μ is countably additive on $\mathcal{A}$.

We claim that if $C_1 \supset C_2 \supset \ldots$ belong to $\mathcal{A}$ and

$$\bigcap_{n=1}^{\infty} C_n = \emptyset, \tag{26}$$

then $\mu(C_n) \to 0$, as $n \to \infty$.

Suppose for a moment this claim is true. Let $A_1, A_2, \ldots$, and E belong to $\mathcal{A}$, and suppose

$$E = \bigcup_{n=1}^{\infty} A_n,$$

where the A_n's are pairwise disjoint. Write

$$E_n := \bigcup_{i=1}^{n} A_i$$

for $n = 1, 2, \ldots$ and observe that the hypotheses of the claim are met by $C_n := E \setminus E_n$ for $n = 1, 2, \ldots$ Consequently, it follows the identity

$$\mu(E) = \mu(E_n) + \mu(C_n)$$

that $\mu(E_n) \to \mu(E)$ as $n \to \infty$. Since μ is finitely additive, we conclude that

$$\sum_{k=1}^{n} \mu(A_k) \to \mu(E)$$

as $n \to \infty$. Hence μ is countable additive on $\mathcal{A}$.

To verify the claim we may assume that each cylinder C_n has n- dimensional base B_n. Since the condition $C_{n+1} \subset C_n$ implies $B_{n+1} \subset B_n \times X_{n+1}$, it is clear that the characteristic functions of B_{n+1} and B_n satisfy

$$\chi(B_{n+1})(x_1, x_2, \ldots, x_{n+1}) \leq \chi(B_n)(x_1, x_2, \ldots, x_n), \tag{27}$$

for $x_j \in X_j$, $1 \leq j \leq n+1$, and $n = 1, 2, \ldots$.

Let

$$g_n^{(1)}(x) := \int_{X_2} \cdots \int_{X_n} \chi(B_n)(x, x_2, \ldots, x_n)\, d\mu_n(x_n) \ldots d\mu_2(x_2),$$

for $x \in X_1$, and observe by definition that

$$\mu(C_n) = \int_{X_1} g_n^{(1)}\, d\mu_1.$$

By (27), $g_n^{(1)}(x_1)$ decreases, as $n \to \infty$, for each fixed $x_1 \in X_1$. It follows from the Lebesgue dominated convergence theorem that

$$\lim_{n\to\infty} \mu(C_n) = \int_{X_1} h_1\, d\mu_1$$

where

$$h_1 := \lim_{n\to\infty} g_n^{(1)}.$$

If the claim is false, then this integral is positive, i.e., there exists an $x_1' \in X_1$ such that $h_1(x_1') > 0$. This condition implies that $x_1' \in B_1$. Indeed, if $x_1' \notin B_1$ then

$$\chi(B_n)(x_1', x_2, \ldots, x_n) = 0,$$

whence $g_n^{(1)}(x_1') = 0$ for $n = 1, 2, \ldots$. This contradicts the choice that $h_1(x_1') > 0$.

For $n > 2$ observe that

$$g_n^{(1)}(x_1') = \int_{X_2} g_n^{(2)}\, d\mu_2$$

where

$$g_n^{(2)}(x) := \int_{X_3} \cdots \int_{X_n} \chi(B_n)(x_1', x, x_3, \ldots, x_n)\, d\mu_n(x_n) \ldots d\mu_3(x_3)$$

for $x \in X_2$. By repeating the argument above, we can show that

$$\lim_{n\to\infty} g_n^{(1)}(x_1') = \int_{X_2} h_2\, d\mu_2$$

where $g_n^{(2)} \downarrow h_2$ as $n \to \infty$. Since $h_1(x_1') > 0$, it follows that $h_2(x_2') > 0$ for some $x_2' \in X_2$. As above, this condition implies that $(x_1', x_2') \in B_2$.

Continuing in this manner, we generate points $x_j' \in X_j$ such that

$$(x_1', x_2', x_3', \ldots, x_n') \in B_n,$$

for $n = 1, 2, \ldots$ We conclude that

$$(x_1', x_2', \ldots) \in \bigcap_{n=1}^{\infty} C_n.$$

This contradicts (26). Therefore the claim is true. ∎

THEOREM 10. *If $G_1, G_2, \ldots$ is a sequence of compact abelian groups then*

$$G := G_1 \times G_2 \times \ldots$$

is a compact abelian group. Moreover, if μ represents the product measure on G inherited from the normalized Haar measures μ_j on G_j, then μ is Haar measure on G.

PROOF. By Theorem 9 and Tychonoff's theorem, G is a compact topological space and μ is a non-negative measure on G of total variation 1. It remains to verify that G is a group, that its group operations are continuous in the product topology, and that μ is translation invariant.

Define addition on G as follows. If $(x_1, x_2, \ldots)$ and $(y_1, y_2, \ldots)$ belong to G set

$$(x_1, x_2, \ldots) + (y_1, y_2, \ldots) := (x_1 + y_1, x_2 + y_2, \ldots).$$

It is easy to check that G is an abelian group with additive inverse given by

$$-(x_1, x_2, \ldots) = (-x_1, -x_2, \ldots).$$

Moreover, since each μ_j is translation invariant on G_j, it is evident that μ is translation invariant on all measurable cylinders in G. Consequently, it remains to check that the map $(z, y) \to z - y$ is continuous from $G \times G$ into G.

Toward this, let V be open in G and contain $z - y$. We may suppose V satisfies (24). In particular,

$$z_{j_k} - y_{j_k} \in V_k$$

for $1 \leq k \leq n$. Since V_k is open in G_{j_k}, there exist open sets U_k and W_k in G_{j_k} such that $z_{j_k} \in U_k$, $y_{j_k} \in W_k$, and $U_k - W_k \subset V_k$ for $1 \leq k \leq n$. Set

$$U := \{(x_1, x_2, \ldots) : x_{j_k} \in U_k,\ 1 \leq k \leq n\},$$

$$W := \{(x_1, x_2, \ldots) : x_{j_k} \in W_k,\ 1 \leq k \leq n\},$$

and observe that U and W are open in G. Since $z \in U$, $y \in W$, and $U - W \subset V$, it follows that the map $(z, y) \to z - y$ is continuous from $G \times G$ into G. ∎

0.5 The Moment Problem. A sequence of real numbers $(a_n, n \in \mathbf{N})$ is said to have a solution to the moment problem if there exists a function g of bounded variation on $[0, 1]$ such that

$$a_k = \int_0^1 t^k \, dg(t)$$

for $k \in \mathbf{N}$.

For every sequence of real numbers $(a_k, k \in \mathbf{N})$ and every pair of non-negative integers n, k, set

$$\Delta^n a_k := \sum_{j=0}^{n} (-1)^j \binom{n}{j} a_{k+j}.$$

The sequence $(a_k, k \in \mathbf{N})$ is called *completely monotone* if $\Delta^n a_k \geq 0$ for all integers $n, k \geq 0$.

THEOREM 11. *Completely monotone sequences have a non- decreasing solution to the moment problem.*

PROOF. Let $(a_k, k \in \mathbf{N})$ be completely monotone, and for each positive integer n define a step function g_n with jumps at m/n for $m = 0, 1, \ldots, n-1$ by the following process. Let

$$a(j,n) := \binom{n}{j}(-1)^{n-j}\Delta^{n-j}a_j$$

for $j = 0, 1, \ldots, n-1$, set $g_n(0) := 0$, $g_n(1) := a_0$, and

$$g_n(x) := \sum_{j=0}^{m-1} a(j,n) \qquad (\frac{m-1}{n} < x < \frac{m}{n}).$$

Extend g_n to $[0,1]$ by averaging g_n at all jumps.

For each polynomial $P(x) = \sum_{j=0}^{n} c_j x^j$, let

$$\Lambda(P) := \sum_{j=0}^{n} c_j a_j.$$

Consider the Bernstein polynomials

$$B(k,n)(x) := \sum_{j=0}^{n} \binom{n}{j} \left(\frac{j}{n}\right)^k x^j (1-x)^{n-j}$$

and observe that

$$\Lambda(B(k,n)) = \int_0^1 t^k \, dg_n(t) \tag{28}$$

for $n, k \in \mathbf{N}$.

Since $(a_k, k \in \mathbf{N})$ is completely monotone, it is clear that

$$\sum_{j=0}^{n} |a(j,n)| = \sum_{j=0}^{n} a(j,n) = a_0.$$

Hence the functions $g_0, g_1, \ldots$ are uniformly of bounded variation on $[0,1]$ with variation a_0. It follows from the Banach-Alaoglu theorem (see 0.0) that there is a function g of bounded variation on $[0,1]$, and a sequence of integers $n_1, n_2, \ldots$, such that

$$\lim_{j\to\infty} \int_0^1 t^k \, dg_{n_j}(t) = \int_0^1 t^k \, dg(t)$$

for $k \in \mathbf{N}$. Moreover, since each g_n is non-decreasing we may adjust g at points of discontinuity so that g is also non- decreasing. Therefore, by (28) it suffices to show

$$\lim_{n\to\infty} \Lambda(B(k,n)) = a_k \tag{29}$$

for $k \in \mathbf{N}$.

Since $a_0 := \Lambda(B(0,n))$, we may suppose that $k > 0$. A direct calculation verifies

$$a_k = \sum_{j=k}^{n} \frac{j(j-1)\dots(j-k+1)}{n(n-1)\dots(n-k+1)} a(j,n).$$

Indeed, $a_k = \Lambda(x^k)$ and by the binomial theorem we can write

$$x^k = x^k((1-x)+x)^{n-k} = \sum_{j=k}^{n} \frac{j(j-1)\dots(j-k+1)}{n(n-1)\dots(n-k+1)} \binom{n}{j} x^j (1-x)^{n-j}.$$

Consequently, the definition of Λ implies

$$\text{(30)} \qquad a_k - \Lambda(B(k,n)) = \sum_{j=k}^{n} \left(\frac{ny(ny-1)\dots(ny-k+1)}{n(n-1)\dots(n-k+1)} - y^k \right) a(j,n) - \sum_{j=0}^{k-1} y^k a(j,n)$$

for $y = j/n$. Since $(nx-i)/(n-i)$ converges uniformly to x on $[0,1]$, as $n \to \infty$, it is clear that

$$\lim_{n\to\infty} \prod_{i=0}^{k-1} \frac{nx-i}{n-i} = x^k$$

uniformly on $[0,1]$. Hence given $\varepsilon > 0$, there is an $n_0 > 0$ such that

$$\left| \frac{ny(ny-1)\dots(ny-k+1)}{n(n-1)\dots(n-k+1)} - y^k \right| < \varepsilon$$

for $n > n_0$, $y = j/n$, and $k \le j \le n$. Moreover, we can choose n_0 so large that

$$|\sum_{j=0}^{k-1} y^k a(j,n)| < \left(\frac{k}{n}\right)^k a_0 < \varepsilon$$

for $n > n_0$. Therefore, it follows from (30) that

$$|a_k - \Lambda(B(k,n))| < \varepsilon(a_0 + 1)$$

for $n > n_0$. ∎

0.6 Binary Derivates and Quasi-measures. Given an integer $n \ge 0$ and a point $x \in [0,1)$ we shall denote the dyadic interval of length 2^{-n} which contains x by $I_n(x)$. Thus

$$\text{(31)} \qquad I_n(x) := \left[\frac{p}{2^n}, \frac{p+1}{2^n}\right)$$

where the integer p is determined by $p \leq 2^n x < p+1$.

A *quasi-measure* is a finite-valued, finitely additive set function defined on the dyadic intervals. The *lower binary derivate* of a quasi-measure ν is defined by

$$\underline{D}\nu(x) := \liminf_{n\to\infty} 2^n \nu(I_n(x)) \qquad (x \in [0,1)). \tag{32}$$

The *upper binary derivate* of ν is defined similarly but "lim inf" is replaced by "lim sup".

The theory of binary derivates closely parallels that of classical *Dini derivates*. For example, the following classical result is well known.

THEOREM 12. *If F is continuous on an interval $[a,b]$ and if the lower Dini derivate of F satisfies*

$$D_+F(x) := \liminf_{h\to 0+} \frac{F(x+h)-F(x)}{h} \leq 0 \tag{33}$$

for all but countably many $x \in (a,b)$, then F is monotone non- increasing on $[a,b]$.

PROOF. We need to show $F(a) \geq F(b)$. Suppose to the contrary that $F(a) < F(b)$. By considering $F(x) - x/n$ we may assume that $D_+F(x)$ is strictly negative for all but countably many $x \in (a,b)$. Thus if

$$E := \{x \in [a,b] : D_+F(x) \geq 0\},$$

then the set $F(E)$ has no interior. In particular, there is a point y_0 in the interval $(F(a), F(b))$ which does not belong to $F(E)$.

Let

$$x_0 := \sup\{x \in [a,b] : F(x) \leq y_0\}.$$

Then $F(x_0) \leq y_0$. If $F(x_0) < y_0$ then $x_0 \neq b$ and by continuity we can choose an $x_1 \in (x_0, b)$ such that $F(x_1) < y_0$ contrary to the choice of x_0. It follows that $F(x_0) = y_0$. Hence for h positive but small,

$$\frac{F(x_0+h)-F(x_0)}{h} \geq \frac{y_0 - F(x_0)}{h} = 0.$$

Therefore $x_0 \in E$, $y_0 \in F(E)$ and we have reached a contradiction. ∎

Let $\mathbf{Q}$ denote the dyadic rationals in $[0,1)$. Set $I'_n(0) := \emptyset$ for $n \in \mathbf{N}$. For each $x \in \mathbf{Q}$ with $x \neq 0$ denote the left end point of $I_n(x)$ by $\alpha_n(x)$ and set

$$I'_n(x) := \left[\alpha_n(x) - 2^{-n}, \alpha_n(x)\right) \qquad (n \in \mathbf{N}). \tag{34}$$

Theorem 12 has a binary analogue.

THEOREM 13. *Let ν be a quasi-measure and E be a countable subset of* $[0,1)$. *If*

$$\liminf_{n\to\infty} \nu(I'_n(x)) \leq 0 \qquad (x \in \mathbf{Q}), \tag{35}$$

$$\liminf_{n\to\infty} \nu(I_n(x)) \le 0 \qquad (x \in [0,1)), \tag{36}$$

and

$$\underline{D}\nu(x) \le 0 \qquad (x \in [0,1) \setminus E), \tag{37}$$

then $\nu(I) \le 0$ *for every dyadic interval* I.

PROOF. By considering $\nu - \varepsilon m$ for $\varepsilon > 0$ (where m denotes Lebesgue measure) we may suppose the inequality in (37) is strict.

Suppose for the purposes of contradiction that there is dyadic interval I_0 with $\nu(I_0) > 0$. Choose $p \in \mathbf{N}$ such that $|I_0| = 2^{-p}$. We shall construct a nested sequence $I_0 \supset I_1 \supset \dots$ of closed, dyadic intervals such that $|I_n| = 2^{-p-n}$, $\nu(I_n) > 0$ for $n \in \mathbf{N}$, which enjoys the following property:

$$\begin{cases} \text{if } \mathbf{Q} \bigcap E \text{ contains a point which belongs to } \bigcap_{n=0}^{\infty} I_n \\ \text{then } \nu(I_n) \le \nu(I_{n+1}) \text{ for } n \text{ sufficiently large.} \end{cases} \tag{38}$$

Let $\mathbf{Q} \bigcap E = \{x_1, x_2, \dots\}$ and suppose that I_n has been chosen. Break I_n into two even, non-overlapping, closed subintervals J_1 and J_2. Then $|J_i| = 2^{-p-n-1}$ $(i = 1, 2)$, and $\nu(I_n) = \nu(J_1) + \nu(J_2)$. Thus $\nu(I_n) > 0$ implies $\nu(J_i) > 0$ for $i = 1$ or 2.

If $\mathbf{Q} \bigcap E$ has no points in I_n, set $I_{n+1} := J_i$ where i is the smallest index for which $\nu(J_i) > 0$.

Otherwise, let $k \ge 1$ be the smallest integer which satisfies $x_k \in I_n$. If x_k is the midpoint of I_n, again set $I_{n+1} := J_i$ where i is smallest and $\nu(J_i) > 0$. If x_k is not the midpoint of I_n, then choose $i = 1$ or 2 so that $x_k \in J_i \setminus J_{i \oplus 1}$. Set $I_{n+1} := J_{i\oplus 1}$ unless $\nu(J_{i \oplus 1}) \le 0$, in which case set $I_{n+1} := J_i$. Here $\oplus$ represents addition modulo 2.

It is clear that $|I_{n+1}| = 2^{-p-n-1}$ and $\nu(I_{n+1}) > 0$. Moreover, if k is the smallest index for which $x_k \in I_n$ and if $x_k \in I_{n+1}$ then by construction $\nu(I_n \setminus I_{n+1}) \le 0$. In particular, $\nu(I_n) \le \nu(I_{n+1})$ and it follows that (38) is satisfied by the intervals $I_0, I_1, I_2, \dots$

Let

$$x_0 \in \bigcap_{n=0}^{\infty} I_n.$$

We shall show that either (35) or (36) fails to hold for $x = x_0$.

First, suppose $x_0 \notin \mathbf{Q}$. By construction,

$$I_n = I_{n+p}(x_0) \qquad (n \in \mathbf{N}). \tag{39}$$

Since $\nu(I_n) \ge 0$ and $\underline{D}\nu(x) < 0$ for $x \notin E$, it follows that $x_0 \in E$. Hence by (38) there exists an integer $N > 0$ such that $\nu(I_n) \ge \nu(I_N) > 0$ for $n \ge N$. In particular,

$$\liminf_{n\to\infty} \nu(I_n(x_0)) \ge \nu(I_N) > 0$$

and (36) fails to hold for $x = x_0$.

Next, suppose that $x_0 \in \mathbf{Q}$ and that x_0 is eventually a left end point of the intervals I_n. Then (39) holds for large n, and again (36) fails to hold for $x = x_0$.

Finally, suppose that $x_0 \in \mathbf{Q}$ and that x_0 is eventually a right end point of the intervals I_n. Then

$$I_n = I'_{n+p}(x_0)$$

for large n. Hence (38) implies that

$$\liminf_{n\to\infty} \nu(I'_n(x_0)) \geq \nu(I_N) > 0$$

for N sufficiently large. Thus (35) fails to hold for $x = x_0$. ∎

The following result also has a classical analogue.

THEOREM 14. *Let ν be a quasi-measure. Then either*

$$\underline{\mathrm{D}}\nu(x) = -\infty \quad \text{and} \quad \overline{\mathrm{D}}\nu(x) = +\infty$$

for a.e. $x \in [0,1)$, or

$$\underline{\mathrm{D}}\nu(x) = \overline{\mathrm{D}}\nu(x) \tag{40}$$

is finite for a.e. $x \in [0,1)$.

PROOF. We begin with the following claim. If for some $\varepsilon > 0$, measurable set $E_0 \subseteq [0,1)$, real number r, and dyadic interval J_0 the following three conditions hold

$$|E_0 \cap J_0| > (1-\varepsilon)|J_0|, \tag{41}$$

$$\nu(I) > 0 \quad \text{for all dyadic intervals satisfying} \quad I \subset J_0,\ I \cap E_0 \neq \emptyset, \tag{42}$$

and

$$\overline{\mathrm{D}}\nu(x) > r \qquad (x \in E_0), \tag{43}$$

then

$$\nu(J_0) > r(1-2\varepsilon)|J_0|.$$

To establish this claim, let $\varepsilon > 0$, $x \in E_0$ and suppose without loss of generality that E_0 lies in the interior of J_0. By (43) there is a dyadic interval $J_x \subset J_0$ containing x such that $\nu(J_x) > r|J_x|$. Hence by (41) and Vitali's covering lemma there exist non-overlapping dyadic intervals $J_1, J_2, \ldots, J_n$ such that

$$\nu(J_i) > r|J_i|, \quad J_i \bigcap E_0 \neq \emptyset, \tag{45}$$

and

$$\sum_{i=1}^{n} |J_i| > (1-\varepsilon)|J_0|. \tag{46}$$

Let $\mathcal{I}_1 := \{J_1, J_2, \ldots, J_n\}$ and let $\mathcal{I}_2$ denote the collection of dyadic intervals $I \subset J_0$ such that $I \bigcap E_0 \neq \emptyset$ and at least one of the halves of I does not contain any interval from $\mathcal{I}_1$. Choose an integer n_0 so large that

$$(47) \qquad |J_i| > 2^{-n_0} \qquad (1 \leq i \leq n),$$

and consider the collection of dyadic intervals

$$\tilde{\mathcal{I}} := \{I \in \mathcal{I}_1 \cup \mathcal{I}_2 : |I| \geq 2^{-n_0}\}.$$

We claim that $\tilde{\mathcal{I}}$ covers J_0. Suppose to the contrary that some point of J_0 is not covered by $\tilde{\mathcal{I}}$. Then there is a dyadic interval $I_0 \subset J_0$ of length 2^{-n_0} which is not contained in any interval $I \in \tilde{\mathcal{I}}$. By (47) I_0 cannot contain any interval from $\mathcal{I}_1$, and by choice I_0 is not contained in any interval from $\tilde{\mathcal{I}}$. Thus we can choose dyadic intervals $I_0 \subset I_1 \subset \ldots$ with $|I_j| = 2|I_{j-1}|$ such that $I_j \bigcap E_0 = \emptyset$ for $j \geq 1$. However, $I_0 \subset J_0$ so for j sufficiently chosen $I_j = J_0$. Thus $J_0 \bigcap E_0 = \emptyset$ contradicting the fact that $E_0 \subset J_0$. In particular, $\tilde{\mathcal{I}}$ covers J_0.

Let

$$A := \bigcup_{i=1}^{n} J_i.$$

Since these intervals are dyadic, any two of them which overlap are comparable. Thus we may suppose that these J_i's are pairwise disjoint. Hence (45) implies that $\nu(A) \geq r|A|$. Moreover, since $\tilde{\mathcal{I}}$ covers J_0, the set $J_0 \setminus A$ consists of a finite number of non-overlapping intervals from $\mathcal{I}_2$. By the definition of $\mathcal{I}_2$ and hypothesis (42) it follows that $\nu(J_0 \setminus A) \geq 0$. Consequently,

$$\begin{aligned} \nu(J_0) &= \nu(A) + \nu(J_0 \setminus A) \\ &\geq \nu(A) \\ &\geq r|A|. \end{aligned}$$

However, (46) implies $|J_0 \setminus A| < \varepsilon |J_0|$ and it follows that $|A| > (1 - 2\varepsilon)|J_0|$. Therefore,

$$\nu(J_0) > r(1 - 2\varepsilon)|J_0|$$

and (44) is established.

By symmetry, we shall complete the proof of this theorem by showing that (40) holds for a.e. $x \in (0,1)$ satisfying $\underline{\mathrm{D}}\nu(x) > -\infty$.

Fix $\varepsilon > 0$ and consider the set

$$E := \{x \in (0,1) : \overline{\mathrm{D}}\nu(x) > \underline{\mathrm{D}}\nu(x) > -\infty\}.$$

By (52) in 1.5 the set E is measurable. Hence it remains to verify that E is of Lebesgue measure zero.

Suppose that $|E| > 0$. Then for some rational $r > 0$ and for some integer k, the Lebesgue measure of the set

$$\{x \in (0,1) : \overline{\mathrm{D}}\nu(x) - \underline{\mathrm{D}}\nu(x) > r, \quad k\varepsilon < \underline{\mathrm{D}}\nu(x) < (k+1)\varepsilon\}$$

is also positive. By replacing $\nu(I)$ by $\nu(I) - k\varepsilon|I|$ we may suppose that there is a number $\eta > 0$ and a measurable set E_0 with $|E_0| > 0$ such that

$$\overline{\mathrm{D}}\nu(x) > r, \quad 0 < \underline{\mathrm{D}}\nu(x) < 2\varepsilon \qquad (x \in E_0), \tag{48}$$

and

$$\begin{cases} \nu(I) > 0 & \text{for all dyadic intervals which satisfy} \\ & E_0 \bigcap I \neq \emptyset, \quad |I| < \eta. \end{cases} \tag{49}$$

Let $x \in E_0$ be a point of density of E_0. Then there exists a dyadic interval J_0 with $|J_0| < \eta$ such that (41) holds. Since (48) and (49) imply (42) and (43), it follows from (44) that $\nu(J_0) > r(1-2\varepsilon)|J_0|$. Moreover, in view of (48) one can shrink J_0 if necessary so that $\nu(J_0) < 2\varepsilon|J_0|$. Consequently,

$$2\varepsilon|J_0| > r(1-2\varepsilon)|J_0|$$

holds for all $\varepsilon > 0$. In particular, if we let $\varepsilon \to 0$ we conclude that $r \leq 0$. This contradicts the fact that r is a positive rational number.

The upshot of all this is that $\overline{\mathrm{D}}\nu(x) = \underline{\mathrm{D}}\nu(x)$ for a.e. x which satisfies $\underline{\mathrm{D}}\nu(x) > -\infty$. It remains to establish that $\overline{\mathrm{D}}\nu(x)$ is finite. However, if $\overline{\mathrm{D}}\nu(x) = +\infty$ on a set E_0 of positive Lebesgue measure, then (43) holds for every real number r. Hence by (44) we can choose a sufficiently small, fixed, dyadic interval J_0 so that

$$\nu(J_0) > r(1-2\varepsilon)|J_0|$$

holds for all $r > 0$. In particular, if we let $r \to +\infty$ we see that $\nu(J_0) = +\infty$. This contradicts the fact that ν is finite-valued. ∎

0.7 Vilenkin Systems. The Vilenkin systems were introduced in 1947 by Vilenkin [1]. They include as a special case the Walsh system and many of the proofs presented for the Walsh system in this book generalize readily to the Vilenkin case. Some indication of this has been given in the Exercises and in the Historical Notes.

Below we give a brief introduction to the Vilenkin systems and relate them to other types of orthonormal systems which have come up in the course of our narrative. For a detailed study of Vilenkin systems see the book by Agaev, Vilenkin, Džafarli, and Rubinštein [1].

Let $m = (m_k, k \in \mathbf{N})$ be a sequence of natural numbers such that $m_k \geq 2$ $(k \in \mathbf{N})$. We shall construct a compact abelian group G_m for each such sequence m. Let Z_{m_k} $(k \in \mathbf{N})$ be the m_k-th discrete cyclic group, i.e., Z_{m_k} can be represented by the set $\{0, 1, \dots, m_k - 1\}$, where the group operation is the mod m_k addition, and every subset is open. Haar measure on Z_{m_k} can be generated by insisting that the measure of a singleton is $1/m_k$. The group G_m is defined as the complete discrete product of the compact groups Z_{m_k} $(k \in \mathbf{N})$. Thus on G_m we use coordinate-wise addition as the group operation, the product topology and the product measure. Thus G_m is a compact abelian group.

The group G_m is metrizable. Let $M_0 := 1$ and $M_k := m_{k-1}M_{k-1}$ $(k \in \mathbf{P})$. Define the distance between the elements $(x_k, k \in \mathbf{N}) \in G_m$ and $(y_k, k \in \mathbf{N}) \in G_m$ by

$$\varrho(x, y) := \sum_{k=0}^{\infty} \frac{|x_k - y_k|}{M_{k+1}}.$$

It is obvious that the topology induced by this metric coincides with that of G_m. It is easy to give a base for the neighborhoods of G_m:

$$I_0(x) := G_m$$
$$I_n(x) := \{y = (y_i, i \in \mathbf{N}) \in G_m : y_i = x_i \quad \text{for} \quad i = 0, 1, \ldots, n-1\}$$

for $x \in G_m$, $n \in \mathbf{P}$. Moreover, if 0 denotes the null element of G_m, then the topology induced by the sets $I_n(0)$ $(n \in \mathbf{N})$ (the base of neighborhoods of 0) coincides with the topology of G_m. Furthermore, it can be shown that the product measure introduced on G_m is the normalized Haar measure of G_m.

Clearly, if $m_k := 2$ $(k \in \mathbf{N})$ then G_m coincides with the dyadic group. Thus no confusion will arise if we use some of the same notation introduced in the study of the dyadic group for these Vilenkin groups G_m.

Accordingly, for each $n \in \mathbf{N}$ set

$$\rho_n(x) := \exp \frac{2\pi \imath x_n}{m_n}$$

for $x = (x_k, k \in \mathbf{N}) \in G_m$, $n \in \mathbf{N}$. Clearly, each ρ_n is a character of G_m. In fact, every complex-valued function defined on G_m whose values depend only on one coordinate of the elements x is continuous on G_m in the topology of G_m.

To enumerate finite products of the functions ρ_n $(n \in \mathbf{N})$, write each $n \in \mathbf{N}$ uniquely with the help of the sequence $(M_n, n \in \mathbf{N})$ in the form

$$n = \sum_{k=0}^{\infty} n_k M_k \qquad (0 \le n_k < m_k,\ n_k \in \mathbf{N}). \tag{50}$$

As in the dyadic case, it can be shown that the set $\widehat{G}_m$ consists of the functions

$$\psi_n := \prod_{k=0}^{\infty} \rho_k^{n_k} \qquad (n \in \mathbf{N}),$$

where the sequence $(n_k, k \in \mathbf{N})$ has been defined in (50). The system $\widehat{G}_m := (\psi_n, n \in \mathbf{N})$ is called a Vilenkin system.

A group operation $\oplus$ corresponding to that introduced earlier in the set $\mathbf{N}$ can be defined as follows:

$$n \oplus j := \sum_{k=0}^{\infty} (n_k + j_k \pmod{m_k}) M_k$$

where

$$n = \sum_{k=0}^{\infty} n_k M_k \quad \text{and} \quad j = \sum_{k=0}^{\infty} j_k M_k.$$

It is clear, then, that $\psi_{n \oplus j} = \psi_n \psi_j$ $(n, j \in \mathbf{N})$ and $\widehat{G}_m$ isomorphic to $\mathbf{N}$.

As in the dyadic case, $\widehat{G}_m$ is a complete orthonormal system. First, the analogue of Theorem 3 in 1.2 can be realized as follows:

LEMMA. *Let $n = \sum_{k=0}^{\infty} n_k M_k \in \mathbf{N}$ and for each pair $1 \leq j < m_\ell$, ℓ of natural numbers, define*

$$A_\ell^j := \{t \in \mathbf{P} : jM_\ell \leq t < (j+1)M_\ell\}.$$

Then

$$\{0, 1, \ldots, n-1\} = \bigcup_{\ell=0}^{\infty} \bigcup_{k=0}^{n_\ell - 1} (n \oplus A_\ell^{m_\ell - n_\ell + k}).$$

Next, it can be easily verified with the help of this equality, that the kernels $D_n := \sum_{k=0}^{n-1} \psi_k$ $(n \in \mathbf{N})$ can be written in the form

$$D_n = \psi_n \sum_{k=0}^{\infty} \left(\sum_{j=m_k - n_k}^{m_k - 1} \rho_k^j D_{M_k} \right),$$

and

$$D_{M_k} = \prod_{j=0}^{k-1} (1 + \cdots + \rho_j^{m_j - 1})$$

where n and n_k are related by (50). Hence it is obvious that

$$D_{M_k}(x) = \begin{cases} M_k & x \in I_k(0) \\ 0 & \text{otherwise.} \end{cases}$$

In particular, the partial sums of the Fourier series of an $f \in \mathrm{L}^1(G_m)$ satisfy

$$(S_{M_n} f)(x) = M_n \int_{I_n(x)} f \qquad (x \in G_m,\ n \in \mathbf{N}).$$

Using the theorem on differentiability of indefinite integrals, it follows that $S_{M_n} f \to f$ a.e. $[\mu]$, as $n \to \infty$, for any $f \in \mathrm{L}^1(G_m)$. In particular, $\widehat{G}_m$ is complete.

Let $k = \sum_{s=0}^{\infty} k_s M_s \in \mathbf{N}$, $n \in \mathbf{P}$ and consider the sets $I_n(x_k)$ where

$$x_k := (k_0, k_1, \ldots, k_{n-1}, 0, 0, \ldots) \in G_m.$$

Clearly,

$$G_m = \bigcup_{k=1}^{M_n} I_n(x_k)$$

is a disjoint decomposition of G_m. Denote by $\mathcal{A}^n$ the σ-field generated by the sets $I_n(x_k)$ $(k = 0, \ldots, M_n - 1)$ and let $\mathcal{A}^0 := \{\emptyset, G_m\}$. One can easily prove that the conditional expectation operator $\mathcal{E}_n$ with respect to $\mathcal{A}^n$ has the form

$$\mathcal{E}_n f = S_{M_n} f \qquad (f \in \mathrm{L}^1(G_m), n \in \mathbf{N}).$$

From this one can see that the M_n-th partial sums form a martingale. This observation makes it possible to use the results from martingale theory to investigate Vilenkin systems.

The dyadic derivative can be defined on a Vilenkin group as follows. Let $n \in \mathbf{P}$ and set

$$d_n f(x) := \sum_{j=0}^{n-1} M_j \sum_{k=0}^{m_j-1} \frac{k}{m_j} \sum_{s=0}^{m_j-1} \overline{\rho_j(se_j)}^k f(x+se_j)$$

where $e_j := (0,0,\dots,0,1,0,\dots) \in G_m$ has its only non- zero component at the j-th position, and se_j represents the sum in G_m of s identical copies of e_j. The function f is called (strongly) differentiable if there exists a function $df \in L^1(G_m)$ such that

$$\lim_{n\to\infty} \|df - d_n f\|_1 = 0.$$

Higher order derivatives, pointwise differentiability and a corresponding anti-derivative can be defined analogously to the dyadic case.

A precise analogue of the representation on $[0,1)$ mentioned in 1.3 can also be constructed with the help of the mapping from G_m to [0,1) defined by

$$(x_k, k \in \mathbf{N}) \to \sum_{k=0}^{\infty} \frac{x_k}{M_{k+1}}.$$

As in the dyadic case, the order of growth of the Lebesgue functions for the system $\widehat{G}_m$ can be obtained. Boundedness (or lack thereof) of the sequence m plays a very important role in the investigations of Fourier series with respect to the system $\widehat{G}_m$. In the case when m is bounded, the theorems on the dyadic group can easily be extended to the group G_m. One reason for this is that the sums

$$\sum_{j=m_k-n_k}^{m_k-1} \rho_k^j \qquad (k \in \mathbf{N})$$

which appear in the representation of D_n $(n \in \mathbf{N})$ are uniformly bounded in this case. This makes it possible to adapt many of the dyadic methods to G_m when m is bounded. If the sequence m is not bounded, then the situation changes very much. Some theorems, which are true for the trigonometric system, for the Walsh system, and for G_m when m is bounded, fail to hold in the general case.

The system $\widehat{G}_m$, on the basis of what has been said, is often called a generalized Walsh system or Walsh type system.

We shall speak briefly about another type of system which is also connected with the name of Vilenkin, but starts from the Haar system and is therefore called a Haar type system. Taking the possibility of the representation on $[0,1)$ into consideration, the following construction will be given on the real line. Let $k \in \mathbf{P}$ and write k in the form

$$k = M_n + r(m_n - 1) + s - 1 \tag{51}$$

where $n \in \mathbf{N}$, $r = 0,\dots,M_n - 1$ and $s = 0,\dots,m_n - 1$. It is obvious that this expression is unique for each $k \in \mathbf{P}$. Let us write an arbitrary element $t \in [0,1)$ in the form

$$t = \sum_{k=0}^{\infty} \frac{t_k}{M_{k+1}} \qquad (0 \le t_k < m_k). \tag{52}$$

(There exist two such expressions for the so-called m-adic rational numbers. For such numbers we use the expression which contains only a finite number of terms different from zero.) Define the function system $(h_n, n \in \mathbf{N})$ by $h_0 := 1$ and

$$h_k(t) := \begin{cases} \sqrt{M_n}\exp\dfrac{2\pi\imath s t_n}{m_n} & \dfrac{r}{M_n} \le t < \dfrac{r+1}{M_n} \\ 0 & \text{otherwise} \end{cases} \tag{53}$$

where $\imath := \sqrt{-1}$. Extend h_n to all of $\mathbf{R}$ by periodicity of period 1. (The numbers t_n, s, r have been defined in (51) and (52).) One can easily see that the system $(h_n, n \in \mathbf{N})$ is a complete orthonormal system in $L^2[0,1)$. Concerning the Lebesgue functions

$$\alpha_n(x) := \int_0^1 |\sum_{k=0}^{n-1} \overline{h_k(x) h_k(t)}|\, dt \qquad (x \in [0,1), n \in \mathbf{P}),$$

it can be shown that

(i) $$\alpha_{M_n + r m_n}(t) = 1 \qquad (0 \le r < M_n,\ n \in \mathbf{N},\ t \in [0,1)),$$

(ii) $$\alpha_n(t) \le c_1 \log m_k \qquad (M_k < n \le M_{k+1},\ n \in \mathbf{N},\ t \in [0,1)),$$

(iii) $$\limsup_{n\to\infty} \frac{\alpha_n(t)}{\log m_k} \ge c_2 \qquad (M_k < n \le M_{k+1},\ n \in \mathbf{N},\ t \in [0,1)),$$

and

(iv) $$\alpha_n(t) \le c_3 \log n \qquad (t \in [0,1)),\ n = 2,3,\dots)$$

where c_1, c_2 and c_3 are positive constants.

In case $m_n := 2$ $(n \in \mathbf{N})$, the system $(h_n, n \in \mathbf{N})$ is the Haar system as defined in 1.4.

An orthonormal set $\boldsymbol{X}$ of L^2-functions is called multiplicative in Vilenkin sense if for any $f, g \in \boldsymbol{X}$ the functions $1/f$, fg belong to $\boldsymbol{X}$. It can be easily shown that the elements of $\boldsymbol{X}$ have the absolute value 1 and $\boldsymbol{X}$ is an abelian group with respect to the multiplication of functions. We give another characterization of these systems below. The reader should notice that the Vilenkin systems G_m are all multiplicative in the Vilenkin sense.

A great number of systems used in Fourier analysis can be expressed as a product system of certain systems. This property plays a very important role in questions related to convergence. Therefore, we shall introduce a generalization of the concept of product system in this section.

Let $\Phi_n := \{\phi_n^k : k \in J_n\}$ $(n \in \mathbf{P})$ be a sequence of function systems consisting of complex-valued functions, where $J_n \subset \mathbf{Z}$, $0 \in J_n$, $\phi_n^0 = 1 (n \in \mathbf{P})$. In order to define the product system generated by the sequence $(\Phi_n, n \in \mathbf{P})$ we introduce the set

$$J := \bigcup_{n=1}^{\infty} (J_1 \times \dots \times J_n \times \{0\} \times \{0\} \times \dots).$$

Define an order relation in the set J as follows. Given $\mathbf{p} = (p_1, p_2, \dots), \mathbf{q} = (q_1, q_2, \dots)$ in J we say $\mathbf{p} < \mathbf{q}$ if there exists an index $n \in \mathbf{P}$ such that $p_n < q_n$ and $p_m = q_m$ if $m > n$, and write $\mathbf{p} \le \mathbf{q}$ if $\mathbf{p} < \mathbf{q}$ or $\mathbf{p} = \mathbf{q}$. It is clear that $\le$ is a linear order relation on J. Let

$$\phi_{\mathbf{p}} := \prod_{n=1}^{\infty} \phi_n^{p_n} \tag{54}$$

for any index $\mathbf{p} = (p_1, p_2, \dots) \in J$. The system $\Phi := (\phi_{\mathbf{p}}, \mathbf{p} \in J)$ is called the product system of the function systems Φ_n $(n \in \mathbf{P})$. Since for $\mathbf{p} \in J$ we have $p_n = 0$ for n sufficiently large, the product in (54) is finite.

The original definition of a product system can be obtained by taking the special case $\Phi_n = \{1, \gamma_n\}$, $J_n = \{0, 1\}$ $(n \in \mathbf{P})$. In this case Φ is usually called the product system generated by the system $\boldsymbol{\gamma} = (\gamma_n, n \in \mathbf{N})$. Thus Φ consists of the finite products formed from the elements of $\boldsymbol{\gamma}$. In this case the mapping

$$\mathbf{p} = (p_1, p_2, \dots) \to \sum_{n=1}^{\infty} p_n 2^{n-1}$$

is a bijection from J to $\mathbf{N}$ which preserves the ordering. According to this, the index set J is often replaced by the set of the natural numbers. Product systems of a finite number of systems can be defined in a similar way.

It can be easily shown that Vilenkin system G_m is a product system in the above sense. Namely, let $J_n := \{0, 1, \dots, m_n - 1\}$, and $\phi_n^k(x) := \exp(kx2\pi\imath)$ for $x \in X_n := \{k/m_n : k \in J_n\}$ $(n \in \mathbf{P})$. Let X be the cartesian product of the sets X_n $(n \in \mathbf{P})$ and $\Pi_n : X \to X_n$ $(n \in \mathbf{P})$ be the n-th projection of the space X. Then the product system of the systems $\Phi_n := \{\phi_n^k \circ \Pi_n : k \in J_n\}$ $(n \in \mathbf{P})$ can be clearly identified with $\widehat{G}_m$. The order-preserving bijection

$$(n_1, n_2, \dots) \to \sum_{k=0}^{\infty} n_{k+1} M_k$$

from J to $\mathbf{N}$ makes it possible to identify J and $\mathbf{N}$.

It turns out that any countable system $F = (f_n, n \in \mathbf{N})$, multiplicative in Vilenkin sense, may be regarded as a product system. In fact, without loss of generality set $f_0 := 1$ and define the function sets $\Phi_n = \{\phi_n^k : k \in J_n\}$ $(n \in \mathbf{N})$ in the following way. Set $\Phi_0 := \{f_0\}$ and assume that the sets $\Phi_1, \Phi_2, \dots \Phi_n$ has already been defined. Denote by F_n the subgroup of F generated by the set $\cup_{k=0}^{n} \Phi_k$ and let $s \in \mathbf{N}$ be the minimal index such that $f_s \notin F_n$. Let us introduce the following notation:

$$s_{n+1} := \min\{k : k \in \mathbf{N}, f_s^k \in F_n\} \qquad (\min \emptyset := \infty)$$

$$J_{n+1} := \begin{cases} \{k \in \mathbf{N} : k < s_{n+1}\} & s_{n+1} < \infty \\ \mathbf{Z} & s_{n+1} = \infty, \end{cases}$$

$$\phi_{n+1}^k := f_s^k \qquad (k \in J_{n+1}),$$

$$\Phi_{n+1} := \{\phi_{n+1}^k : k \in J_{n+1}\}.$$

Then it is obvious that $F_1 \subset F_2 \subset \cdots \subset F_n \subset \cdots \subset F$, $\cup_{n=0}^{\infty} F_n = F$ and every $f \in F$ can be expressed in the form

$$f = \prod_{k=1}^{\infty} \phi_k^{n_k}.$$

We remark that when F is periodic i.e., if there exists a $k \in \mathbf{P}$ such that $f^k = 1$ $(f \in F)$, then F can be expressed as a product system of systems Φ_n $(n \in \mathbf{N})$, in which the number of the elements of J_n is finite for every $n \in \mathbf{N}$.

Let $(\Omega, \mathcal{G}, \nu)$ be a probability space and $\boldsymbol{\gamma} := (\gamma_n, n \in \mathbf{N})$ a (real-valued) function system in $\mathrm{L}^\infty(\Omega, \mathcal{G}, \nu)$. If the system $\boldsymbol{\gamma}$ is orthogonal in the space $\mathrm{L}^2(\Omega, \mathcal{G}, \nu)$ then the integral of the product of any two different elements equals zero. A number of systems possess the stronger property that the integral of any finite product of its element is also equal to zero. These systems play an important role not only in the theory of orthogonal series but also in probability theory.

Denote by g_n $(n \in \mathbf{N})$ the elements of the product system of $\boldsymbol{\gamma}$, that is, let

$$g_n := \prod_{k=0}^{\infty} \gamma_k^{n_k} \tag{55}$$

for $n = \sum_{i=0}^{\infty} n_i 2^i$ $(n_i \in \{0,1\})$.

We say that the system $\boldsymbol{\gamma}$ is

i) strongly multiplicative, if the system $(g_n, n \in \mathbf{N})$ is orthogonal;

ii) multiplicative, if $\int_\Omega g_n \, d\nu = 0 \qquad (n \in \mathbf{P})$;

iii) weakly multiplicative, if $\sum_{i=0}^{\infty} | \int_\Omega g_i \, d\nu| < \infty$.

It is clear that every strongly multiplicative system is multiplicative and every multiplicative system is weakly multiplicative.

It can be shown that every orthogonal system contains a weakly multiplicative subsystem. Additional properties of multiplicative systems and generalizations of the definition will be mentioned below. The Rademacher system is evidently strongly multiplicative. On the other hand, let $(\gamma_n, n \in \mathbf{N})$ $(\gamma_n \in \mathrm{L}^\infty[0,1])$ be any strongly multiplicative system, such that $|\gamma_n(x)| = 1$ for almost all $x \in [0,1]$ $(n \in \mathbf{N})$. Suppose further that the product system $(g_n, n \in \mathbf{N})$ is equinormed (that is,

$$\int_0^1 |g_n|^2 = c$$

is constant for n sufficiently large). Then the following two conditions are equivalent:

i) $(g_n, n \in \mathbf{N})$ is complete,

ii) there exists a measure preserving bijection $\varpi : [0,1] \to [0,1]$ such that

$$g_n(x) = w_n(\varpi(x))$$

for a.e. $x \in [0,1]$, $n \in \mathbf{N}$. In particular, if a function $f \in \mathrm{L}^1[0,1]$ has the expansion $\sum c_n g_n$ with respect to the system $(g_n, n \in \mathbf{N})$, then the function $f \circ \varpi^{-1}$ has the Walsh expansion $\sum c_n w_n$.

Generalizing a property of the Haar system, we introduce the H- property. Let $(\Omega, \mathcal{G}, \nu)$ be a measure space and $H_n \in L^\infty(\Omega, \mathcal{G}, \nu)$ $(n \in \mathbf{N})$. The system $(H_n, n \in \mathbf{N})$ is said to have the H-property if there exists a number $K > 0$ such that

$$\begin{cases} \| \sum_{i=0}^{2^n-1} |H_{2^n+i}| \|_\infty \leq K2^{n/2} \\ \|H_{2^n+j}\|_1 \leq K2^{-n/2} \end{cases} \tag{56}$$

for $j = 0, 1, \ldots, 2^n - 1$, $n \in \mathbf{N}$. A system which has the H-property is said to be of H-type. By definition (see 1.4), the Haar system has the H-property. The Franklin system $(\mathrm{f}_n, n \in \mathbf{N})$ introduced in 5.4 satisfies

$$\| \sum_{i=0}^{2^n-1} |\mathrm{f}_{2^n+i+1}| \|_\infty \leq 2^5 \sqrt{3}\, 2^{n/2}$$

and

$$\|\mathrm{f}_{2^n+i+1}\|_1 \leq 6\sqrt{3}\, 2^{-n/2}$$

for $i = 0, 1, \ldots, 2^n - 1$, $n \in \mathbf{N}$. (See Ciesielski [1] and [2] for details.) In particular, the system $(\mathrm{f}_{n+1}, n \in \mathbf{N})$ also has the H-property.

Let $\boldsymbol{\gamma} = (\gamma_n, n \in \mathbf{N})$ be an arbitrary function system, and $(g_n, n \in \mathbf{N})$ be the product system of $\boldsymbol{\gamma}$. Then the H-system generated by $\boldsymbol{\gamma}$ is the system $(H_n, n \in \mathbf{N})$ which is the Hadamard transform of $(g_n, n \in \mathbf{N})$ (see 1.4). It is clear that the H-system generated by the Rademacher system is the Haar system. We shall show that if $\gamma_n \in L^\infty(\Omega, \mathcal{G}, \nu)$ $(n \in \mathbf{N})$ is weakly multiplicative and $\|\gamma_n\|_\infty \leq 1$ $(n \in \mathbf{N})$, then the H-system generated by $\boldsymbol{\gamma}$ has the H-property.

Indeed, by definition, (56), and (55) we find that

$$\begin{aligned} H_{2^n+k} &= 2^{-n/2}\gamma_n \sum_{j=0}^{2^n-1} \prod_{i=0}^{n-1} (r_i(k2^{-n})\,\gamma_i)^{j_i} \\ &= 2^{-n/2}\gamma_n \prod_{i=0}^{n-1} (1 + r_i(k2^{-n})\,\gamma_i) \end{aligned}$$

for $j = \sum_{i=0}^{n-1} j_i 2^i$. It follows, therefore, that

$$\begin{aligned} \|H_{2^n+k}\|_1 &\leq 2^{-n/2} | \int_\Omega \prod_{i=0}^{n-1} (1 + r_i(k2^{-n})\,\gamma_i)\, d\nu | \\ &= 2^{-n/2} | \int_\Omega \sum_{j=0}^{2^n-1} w_j(k2^{-n}) g_j\, d\nu | \\ &\leq 2^{-n/2} \sum_{j=0}^{\infty} | \int_\Omega g_j\, d\nu |. \end{aligned}$$

Similarly

$$\begin{aligned}\sum_{k=0}^{2^n-1}|H_{2^n+k}| &\le 2^{-n/2}\left(\sum_{k=0}^{2^n-1}\sum_{j=0}^{2^n-1} w_j(k2^{-n})g_j\right)\\ &= 2^{-n/2}\left(\sum_{j=0}^{2^n-1} g_j\sum_{k=0}^{2^n-1} w_j(k2^{-n})\right)\\ &= 2^{-n/2}\left(\sum_{j=0}^{2^n-1} g_j 2^n\int_0^1 w_j\right)\\ &= 2^{n/2} \qquad (n\in\mathbf{N}).\end{aligned}$$

The Lebesgue-functions $F_n(x) := \|\sum_{k=0}^{n-1} H_k(x)H_k\|_1$ $(n \in \mathbf{N})$ of an H-type system generated by a weakly multiplicative system $\boldsymbol{\gamma} = (\gamma_n, n \in \mathbf{N})$ are uniformly bounded if $\|\gamma_n\|_\infty \le 1$ $(n \in \mathbf{N})$. Indeed, since the Hadamard-Paley matrices $A^{(n)}$ $(n \in \mathbf{N})$ are orthogonal, it follows from (56) that

$$\sum_{k=2^n}^{2^{n+1}-1} H_k(x)H_k = \sum_{k=2^n}^{2^{n+1}-1} g_k(x)g_k \qquad (x\in\Omega,\ n\in\mathbf{P}).$$

In particular, the Lebesgue-functions F_{2^n} are uniformly bounded:

$$\begin{aligned}\|\sum_{k=0}^{2^n-1} H_k(x)H_k\|_1 &= \|H_0(x)H_0 + \sum_{j=0}^{n-1}\sum_{k=2^j}^{2^{j+1}-1} H_k(x)H_k\|_1\\ &= \|\sum_{k=0}^{2^n-1} g_k(x)g_k\|_1\\ &= \|\prod_{j=0}^{n-1}(1+\gamma_j(x)\gamma_j)\|_1\\ &\le \sum_{k=0}^{\infty}|\int_\Omega g_k\,d\nu| < \infty.\end{aligned}$$

Therefore, to show that the F_n's are uniformly bounded it suffices to verify that

$$\boldsymbol{\Delta} := \sup\{\int_\Omega |\sum_{k=2^n}^{s} H_k(x)H_k(t)|\,d\nu(t) : 2^n \le s < 2^{n+1},\ n\in\mathbf{N}\} < \infty$$

for $x \in \Omega$. This fact follows immediately from the H- property since

$$\boldsymbol{\Delta} \le \sup\{\sum_{k=2^n}^{2^{n+1}-1}|H_k(x)| \sup_{2^n\le k<2^{n+1}}\|H_k\|_1 : n\in\mathbf{N}\} < \infty$$

for each $x \in X$.

A closely related concept is given by a W-system or W-type system. Let

$$(H_n, n \in \mathbf{N})$$

be a function system with $H_n \in \mathrm{L}^\infty(\Omega, \mathcal{G}, \nu)$. The system $(W_n, n \in \mathbf{N})$ given by the Hadamard transform of $(H_n, N \in \mathbf{N})$ is called the W-system generated by $(H_n, n \in \mathbf{N})$. Since the Hadamard-Paley matrices are orthogonal, every W-system generated by an orthonormal system is an orthonormal system. Furthermore if the system $(H_n, n \in \mathbf{N})$ has the H-property then the W- system generated by it is obviously uniformly bounded. Clearly, Walsh system is a W-system generated by the Haar system, and the Ciesielski system $(\mathrm{g}_{n+1}, n \in \mathbf{N})$ (see 5.5) is a W- system generated by the Franklin system $(\mathrm{f}_{n+1}, n \in \mathbf{N})$. Since the Franklin system is an H-type orthonormal system, the Ciesielski system is also an orthonormal system whose elements being continuous, piecewise linear functions are uniformly bounded. Every H-system is a W-system generated by the product system of its generator system.

HISTORICAL NOTES

Here we give additional comments and references about the material presented in this book. The bibliography following these notes is not exhaustive. More information of this type can be found in Balašov and Rubinštein [1], Olevskiĭ[4], Stanković [2], Talaljan [4], Ul'janov [5], and Wade [8].

CHAPTER 1

1.1. Walsh functions were used for the transposition of conductors in open wire lines as early as the late 1800's, and the complete system of Walsh functions seems to have been found around 1900 by J.A. Barrett (see Harmuth [1], p. 3).

The Rademacher functions were introduced to the mathematical world in 1922 by Rademacher [1]. They were probably unknown to Walsh [1] who introduced his system in 1923 and showed that the Walsh and Haar systems are Hadamard transforms of each other. Walsh's construction of the Walsh system, presented here, is somewhat cumbersome but can be generated recursively. This is one of the main reasons this enumeration is preferred for applications.

What we call the Walsh system is often referred to in the literature as the Walsh-Paley system. It was introduced by R.E.A.C. Paley [1] in 1932. He was first to recognize that Walsh functions are products of Rademacher functions. Exploiting this, he obtained fundamental inequalities (see 3.3 below) which bear his name and used them to prove that lacunary partial sums of Walsh- Fourier series of $f \in L^p$, $1 < p < \infty$, converge a.e. and that the full sequence of partial sums converges in L^p norm.

The Walsh-Kaczmarz enumeration was introduced in 1948 by Šneider [1]. This enumeration has not been studied as thoroughly as the other two, but it appears more frequently in the literature these days.

1.2. The fact that the Walsh system is the group of characters of a compact abelian group connects Walsh analysis with abstract harmonic analysis. It was discovered independently by Fine [1] and Vilenkin [1]. The latter actually introduced a large class of compact groups (now called Vilenkin groups) and corresponding characters which includes the dyadic group and the Walsh system as a special case. (For general references to harmonic analysis on groups see Pontryagin [1], Rudin [1], and Hewitt and Ross [1]. See also Appendix 0.3.)

Paley's lemma had its roots in the dissertations of Walsh [1] and Haar [1]. We have attached Paley's name to it since he used it both frequently and skillfully, and was first to write it explicitly as presented here.

Paley's lemma shows that the 2^n-th partial sums of the Walsh- Dirichlet kernel are non-negative. It is interesting to note that among the uniformly bounded, complete orthonormal systems whose functions alternate signs on finer and finer partitions of $[0,1]$, the Walsh system is the only one whose Dirichlet kernels have non- negative 2^n-th partial sums (Price [2]).

1.3. Theorems 4 and 5 are due to Morgenthaler [1]. The concepts of W-continuity and the dyadic modulii of continuity are due to Fine [1] and Morgenthaler [1]. Functions of bounded fluctuation were introduced by Onneweer and Waterman [1], [2] to investigate uniform convergence of Walsh-Fourier series.

1.4. Multiplicative systems were introduced by Alexits [2] who studied strongly and weakly multiplicative kinds.

Theorem 6 is due to Waterman [1], [2].

Schipp [9] proved Theorem 7 and introduced linear and piecewise linear rearrangements to study a.e. convergence of Walsh-Fourier series in the original and Kaczmarz ordering. Bahšečjan [2] has generalized these concepts.

The Haar system was introduced by Haar [1] in 1909. It came about in the following way. For most of the nineteenth century many analysts believed that trigonometric Fourier series of continuous functions converged everywhere if not uniformly on $[0,2\pi]$. After du- Bois Reymond showed this was not the case, work progressed by adding hypotheses sufficient to guarantee uniform convergence or by proving summability results directly. When Haar began working on his dissertation, Hilbert posed the question: does there exist an orthonormal system such that the Fourier series of every continuous function converges uniformly? Haar answered this question in the affirmative, showing Haar-Fourier series of continuous functions converge uniformly on $[0,1]$. (It is not an accident that the Haar system is unbounded. Olevskiĭ [1] has shown that no bounded orthonormal system has uniformly convergent Fourier series for all continuous functions.)

Walsh [1] first observed that the Haar and Walsh systems are Hadamard transforms of each other. The Hadamard-Paley matrices were exploited by Kaczmarz [1] and Alexits [2] to deduce convergence results for Walsh-Fourier series from Haar-Fourier series results. Hadamard matrices are important for certain applications (see 9.6).

1.5. For trigonometric Fourier series, the Riemann- Lebesgue lemma was proved by Riemann for Riemann integrable functions and Lebesgue for Lebesgue integrable functions. The Walsh version is due to Walsh [1]. It does not hold for improperly Riemann integrable functions which are not Lebesgue integrable, i.e., for functions which are not absolutely integrable.

The Walsh-Fourier coefficients of the indefinite integrals J_k were first computed by Fine [1]. He also proved (44) through (47).

Walsh-Fourier series and coefficients are special cases of the theory of general orthogonal series. In this regard see Kaczmarz and Steinhaus [1], Alexits [2] and Kašin and Saakjan [1]. For connections with functional analysis see Rudin [2] and Riesz and Sz.-Nagy [1].

The various formulae given in Theorem 8 for the Dirichlet kernels were used and generalized by many authors, e.g., Paley [1], Fine [1], Vilenkin [1], Bljumin [1], and Simon [1].

The fact that the 2^n-th partial sums of Walsh-Fourier series of continuous functions converge uniformly was recognized by Walsh [1]. He also obtained several refinements of this result not mentioned here.

The identification of Walsh series and quasi-measures has roots in Fine's work on uniqueness [1]. Yoneda was first to exploit the measure theoretic arguments this identification allows (see Wade and Yoneda [1]). He also introduced the concept of a Walsh- Fourier-Stieltjes series of a quasi-measure ν. Fine [4] had studied such series earlier in the case when ν is a finite Borel measure on $[0,1)$ or the dyadic group $\mathbf{G}$.

The fact that the 2^n-th partial sums of a Walsh series must either converge a.e. or have limit supremum $+\infty$ and limit infimum $-\infty$ was first discovered by Talaljan and Arutunjan [1]. A martingale proof was given by Gundy [2]. Arutunjan [1] has also shown that the 2^n-th partial sums of a Walsh series S_{2^n} converges a.e., as $n \to \infty$, if and only if

$$\sum_{k=0}^{\infty} |S_{2^{k+1}}(x) - S_{2^k}(x)| < \infty$$

for a.e. $x \in [0,1)$. (For a martingale proof of this characterization, see Gundy [1].)

The computations leading to identity (53) are due to Bočkarev [1]. He used them (as we do) to estimate the growth of Walsh- Fourier coefficients of non-constant, continuous functions (see 2.3).

1.6. Many of the estimates and results concerning Lebesgue constants for the Walsh system can be found in Fine [1], Šneider [3], and Vilenkin [1]. Theorem 9 is new. Theorem 10 is due to Fine [1] and has been generalized by Fridli and Simon [1]. Theorem 11 was proved by Šneider [1].

1.7. The Walsh-Fourier transform was introduced by Fine [2]. The idea of using it to motivate the definition of dyadic differentiation is due to Splettstößer.

The dyadic derivative has its roots in the work of Gibbs and Millard [1] who introduced it for discrete functions. The definitions presented in this section, together with most of the results (including Theorems 13, 14, 15), are due to Butzer and Wagner [1], [2].

Theorem 12 was discovered by Skvorcov (see Skvorcov and Wade [1]). The situation does not change much if the function f is allowed some jump discontinuities. Indeed, Engels [1] has shown that if f is defined and bounded, and possesses only countably many discontinuities (all exclusively of the first kind), and if the set of discontinuities of f has only finitely many cluster points, then f is dyadically differentiable at all but countably many points in $[0,1)$ if and only if f is piecewise constant. Butzer, Engels, and Wipperfürth [1], [2] have tried to remedy this narrowness by defining an extended dyadic derivative which includes classical polynomials among the "dyadically differentiable" functions. Although this extended dyadic derivative E does not agree with the dyadic derivative (for example, $Ew_3 = -w_3$ but $dw_3 = 3w_3$), it is the case that every dyadically differentiable function is differentiable in this extended sense. Moreover, this extended dyadic derivative enjoys a product rule which reduces to the classical one of Leibniz when the factors are

Rademacher functions. It remains to see whether this new derivative will play an important role in Walsh analysis.

The dyadic derivative has been generalized to Vilenkin groups by Onneweer [7] and to the dyadic field by Butzer and Wagner [3] (see also J. Pál [1]).

1.8. The first three formulae of Theorem 16 were obtained by Fine [1]. The remaining two are due to Schipp [14] and Yano [4], respectively. For generalizations to Vilenkin systems see Onneweer [7], [9], and Pál and Simon [1].

Theorem 17 appeared explicitly in Butzer and Wagner [1]. Penney [1] has other facts about the Walsh series W.

CHAPTER 2

2.1. Theorem 1 is due to Fine [1]. Estimates of the Walsh-Fourier coefficients of functions belonging to various other classes of functions can be found in Efimov [1], Horoško [1], [2], and Siddiqi [1]. For the Vilenkin case see Vilenkin [1] and Ohkuma [1].

Theorem 2 is due to Butzer and Wagner [1]. Theorem 3 was proved by Onneweer [5] who also considered the Vilenkin case. Trigonometric analogues of these results can be found in Bary [1] and Zygmund [1].

2.2. Theorems 4 and 5 were proved by Fine [1] and Theorem 6 by Morgenthaler [1]. Theorem 6 has a classical analogue but Theorems 4 and 5 do not.

For each $k \in \mathbf{N}$ let $\ell(k)$ represent the number of non-zero dyadic digits in the binary expansion of k. Thus if k has binary coefficients $(k_j, j \in \mathbf{N})$ then

$$\ell(k) := \sum_{j=0}^{\infty} k_j.$$

This number is called the Vielfalt (roughly, the diversity) of k by the Austrian school. It crops up in uniqueness as well (see 7.2 and 7.4 below).

A mild lacunarity condition is given by the assumption that $\hat{f}(k) = 0$ for all integers k satisfying $\ell(k) > n$ for some fixed $n \in \mathbf{N}$. P. Weiss [1] showed that if f is n- times classically continuously differentiable and if $\hat{f}(k) = 0$ for all $\ell(k) > n$ then f is a polynomial of order n (see also Liedl [1]). Roider [1] showed that if $f \in \mathrm{L}^1$ and $\hat{f}(k) = 0$ for $\ell(k) > n$ and if f assumes only finitely many values, or only integer values, then f must be a Walsh polynomial. In the integer valued case the coefficients $\hat{f}(k)$ must be integer multiples of 2^{-n}. He also investigates this result for Walsh series lacunary in the Hadamard sense.

2.3. The results of this section were obtained by Bočkarev [1]. Theorem 7 carries the tacit assumption that the Walsh- Fourier series of f converges absolutely. Coury [3] obtained conditions on the Walsh-Fourier coefficients of a continuous function f sufficient to conclude that f is constant which do not force Sf to converge absolutely (see comments at the end of 8.2). His results also show that a continuous f which satisfies $\sum_{k=0}^{\infty} k|\hat{f}(k)| < \infty$ must be constant. Other results of this type can be found in Powell and Wade [1], and Wade [8]. In particular, if f is continuously differentiable and satisfies

$$\sup_{n \in \mathbf{P}} \left| \sum_{k=0}^{n} \left(\sum_{j=2^n}^{2^{n+1}-1} \hat{f}(j) w_j(x_0) \right) \right| < \infty$$

for every dyadic rational x_0 then f is constant.

2.4. Theorem 9 and Corollaries 2 and 4 are due to Onneweer [3], [5], and [6]. His results hold for any bounded Vilenkin group. Corollary 3 was proved Fine [1]. The condition "$f \in \mathrm{Lip}\ \alpha,\ \alpha > 1/2$" has been weakened to "the series

$$\sum_{n=1}^{\infty} \omega^{(2)}(f, \frac{1}{n})/\sqrt{n}$$

converges "by Yoneda [1]. For any bounded Vilenkin group, Vilenkin and Rubinštein [1] have shown that if $f \in L^2$ and

$$\sum_{k=0}^{\infty} \sqrt{m_k}\omega^{(2)}(f, 1/m_k) < \infty$$

then the Vilenkin-Fourier series of f converges absolutely.

The Vilenkin-Fourier series of a function in $\mathrm{Lip}(\alpha, L^p) * \mathrm{Lip}(\beta, L^q)$ is always absolutely convergent for $1 \le p \le 2$, $q > 1$, $0 < \alpha, \beta \le 1$. This was proved for bounded Vilenkin groups by Onneweer [6] and later for unbounded ones by Quek and Yap [1].

The problem of determining conditions sufficient to conclude that

$$\sum_{k=1}^{\infty} k^{-\gamma} |\widehat{f}(k)|^{\beta}$$

converges is an old one. For a complete discussion see McLaughlin [1]. We have presented special cases in Theorems 10 and 11 (originally obtained for Vilenkin groups by Onneweer [5]). These theorems are analogues of trigonometric results of Bernstein- Szàsz and Hardy, respectively. They are sharp in the following sense. Bočkarev [2] has shown that there is no complete orthonormal system for which Theorems 10 and 11 hold if $\beta > 2/(2\alpha + 1)$ and $\beta > 1/2 - \alpha$ are replaced by $\beta \ge 2/(2\alpha + 1)$ and $\beta \ge 1/2 - \alpha$.

Theorem 12 is an analogue of a trigonometric result due to Stečkin. Yoneda [1] has shown that if f is continuous, of bounded variation, and satisfies

$$\sum_{k=1}^{\infty} \sqrt{\omega(f, 1/k)}/k < \infty$$

then the Walsh-Fourier series of f converges absolutely. On the other hand, Bočkarev [2] has shown that if ω is a modulus of continuity which satisfies

$$\sum_{k=1}^{\infty} \sqrt{\omega(1/k)}/k = \infty$$

then for any bounded, orthonormal system in L^2 there exists an absolutely continuous function whose Fourier coefficients in this system satisfy $\sum_{k=1}^{\infty} |a_k(f)| = \infty$ but whose classical modulus of continuity satisfies $\Omega(f, \delta) = O(\omega(\delta))$ as $\delta \to 0$.

2.5. Theorems 13, 14, and 16 are due to Fine [1]. The trigonometric analogue of Theorem 13 is called the Riemann localization principle. It was obtained for Vilenkin systems by Vilenkin [1], and for 2^n-th partial sums of Walsh-Fourier series by Walsh [1]. The trigonometric analogue of Theorem 14 is called Dini's test. It was obtained for bounded Vilenkin systems by Vilenkin [1].

Skvorcov [17] has shown that localization does not hold for Cesàro means of Walsh-Kaczmarz-Fourier series. Namely, he constructed an $f \in L^1$ which equals 0 on [0,1/2] but whose Walsh-Kaczmarz-Fourier series is not Cesàro summable at the point $x = 0$.

Theorem 15 was proved by Walsh [1]. Its trigonometric analogue is known as the Dirichlet-Jordan test.

Yano [3] obtained the following analogue of the Hardy-Littlewood test. If $f \in L^1$, if $|\widehat{f}(k)| = O(k^{-\delta})$ as $k \to \infty$ and if

$$2^n \int_{I_n(x)} |f(x \dot{+} t) - f(x)|\, dt = o(\frac{1}{n})$$

as $n \to \infty$, then the Walsh-Fourier series of f at x converges to $f(x)$. Vilenkin obtained this same result for bounded Vilenkin systems.

CHAPTER 3

3.1. Gundy ([1] and [2]) was first to exploit the fact that the 2^n-th partial sums of any Walsh series forms a martingale. This allows one to use results from the general theory of probability to obtain results about Walsh series.

Dyadic martingales are part of the broader study of general martingales. These are sequences $(f_n, n \in \mathbf{N})$ of functions on a probability space $(\Omega, \mathcal{G}, \nu)$ endowed with an increasing sequence of sub-σ-fields $\mathcal{B}_n$, $n \in \mathbf{N}$, such that each f_n is $\mathcal{B}_n$ measurable and the conditional expectation $\mathrm{E}(f_{n+1} \mid \mathcal{B}_n)$ is precisely f_n for $n \in \mathbf{N}$. Conditions (2) through (6) hold for any martingales if $\mathcal{E}_n$ is replaced by $\mathrm{E}(\cdot \mid \mathcal{B}_n)$ and $\mathcal{A}^n$ is replaced by $\mathcal{B}_n$. General references can be found in Burkholder [2], Garsia [1], and Neveu [1]. Martingales with non-discrete index sets are discussed in Lipčev and Siriaev [1].

Theorem 1 in this form seems to be new. It is easy to check that it holds for any increasing sequence $(\mathcal{B}_n, n \in \mathbf{N})$ of sub-σ-fields in any probability space $(\Omega, \mathcal{G}, \nu)$. Therefore, it contains the martingale maximal theorem and Doob's inequality: if $(f_n, n \in \mathbf{N})$ is a martingale then

$$y\nu\{f_n^* > y\} \le \int_{\{f_n^* > y\}} |f_n|\, d\nu \le \sup_{n\in\mathbf{N}} \int_\Omega |f_n|\, d\nu$$

for all $y > 0$, and

$$\sup_{n\in\mathbf{N}} \|f_n\|_p \le \sup_{n\in\mathbf{N}} \|f_n^*\|_p \le q \sup_{n\in\mathbf{N}} \|f_n\|_p$$

where p and q are conjugate indices, $1 < p < \infty$, and $f_n^* := \sup_{k\le n} |f_k|$. In particular, if the martingale is regular (i.e., $f_n = \mathrm{E}(f \mid \mathcal{B}_n)$ for some $f \in \mathrm{L}(\Omega)$) then

$$\|f\|_p \le \|\sup_{n\in\mathbf{N}} f_n^*\|_p \le q\|f\|_p.$$

Theorem 2 is a Banach-Steinhaus type theorem for a.e. convergence. The earliest result of this type we know of was obtained by Kolmogorov [1] who used a special case to show certain lacunary partial sums of Fourier series are a.e. convergent. For necessity of weak type conditions with regard to a.e. convergence in the general case, see Stein [1] and Nikišin [1].

Corollary 2 is a special case of the convexity lemmas of Burkholder, Davis, and Gundy [1]. Namely, if $(g_n, n \in \mathbf{N})$ is a sequence of non-negative, measurable functions on the probability space $(\Omega, \mathcal{G}, \nu)$ and if Φ is a non- decreasing, continuous, convex function on $[0,\infty)$ with $\Phi(0) = 0$ and $\Phi(2y) \le C\Phi(y)$ for all $y > 0$ then

$$\int_\Omega \Phi\left(\sum_{n=1}^{\infty} \mathrm{E}(g_n \mid \mathcal{B}_n)\right) d\nu \le C \int_\Omega \Phi\left(\sum_{n=1}^{\infty} g_n\right) d\nu$$

with the reverse inequality eventuating when "convex" is replaced by "concave". For a simple proof of these inequalities see Garsia [2]. For a connection between these inequalities and the martingale maximal theorem see Mogyoròdi [1].

3.2. Lemma 1 is due to Burkholder [2].

The definition and basic properties of stopping times in the general case are similar to the dyadic case presented here. In particular, Lemma 3 holds for general martingales (see Neveu [1]).

A decomposition for general martingales corresponding to Lemma 4 was obtained by Gundy [3]. A simple proof can be found in Burkholder [2].

Theorem 3 also holds for general martingales (see Garsia [1]), and again predictability plays an important role. For the trigonometric case, a decomposition similar to Theorem 3 can be proved if the dyadic maximal function is replaced by the harmonic maximal function. In this regard see Kašin and Saakjan [1].

3.3. Martingale transforms were first systematically studied by Burkholder. The inequalities of Theorem 4 for general martingales were proved by Burkholder [1]. An earlier version for the Walsh case dates back to Yano [5]. The proof of Theorem 4 presented here is due to Schipp [15] (see also Burkholder [3]).

Theorem 5 first appeared in Burkholder, Gundy, and Davis [1]. They obtained this result for predictable martingales and, in the case where Φ is convex, without assuming predictability.

For $1 < p < \infty$, Corollary 4 is a simple consequence of the inequalities of Paley and Doob. The case $p = 1$ was obtained by Davis [1] in 1970.

The first inequality of Corollary 5 is called Paley's inequality. It was obtained by Paley [1] in 1932. In fact, he proved the following. Let $(\lambda_n, n \in \mathrm{N})$ be a lacunary sequence, that is, a sequence of natural numbers tending to ∞ which satisfies $\lambda_{n+1}/\lambda_n \geq q > 1$ for $n \in \mathrm{N}$. If

$$\delta_n := S_{\lambda_n} f - S_{\lambda_{n-1}} f$$

and

$$q(f) = \left(\sum_{n=1}^{\infty} \delta_n^2 \right)^{1/2}$$

then for every $1 < p < \infty$ there are absolute constants A_p and B_p such that

$$A_p \|q(f)\|_p \leq \|f\|_p \leq B_p \|q(f)\|_p$$

for all $f \in \mathrm{L}^p$. Notice that $q(f)$ reduces to the square function Qf when $\lambda_n := 2^n$ for $n \in \mathrm{N}$. Bonami [1] has shown that Paley's result is sharp in the following sense. There exists a sequence $(\lambda_n, n \in \mathrm{N})$ tending monotonically to ∞ such that $\lambda_{n+1}/\lambda_n \to 1$ and $\lambda_{n+1} - \lambda_n \to \infty$ as $n \to \infty$ for which the inequality above fails to hold .

Paley's inequality holds for general martingales if $\Delta_n f$ is replaced by the martingale differences $d_n := f_{n+1} - f_n$ (see Burkholder [2]). The special case when the d_n's are independent and of mean zero was obtained by Marcinkiewicz and Zygmund [1].

Watari [1] proved the following version of Paley's inequality for Vilenkin systems of bounded type. If

$$\delta_{n,j}(f) := S_{(j+1)M_n} f - S_{jM_n} f$$

and

$$Q(f) := \left(|\widehat{f}(0)|^2 + \sum_{n=1}^{\infty} \sum_{j=1}^{M_n - 1} |\delta_{n,j}(f)|^2 \right)^{1/2}$$

then there exist absolute constants A_p, B_p such that

$$A_p \|Q(f)\|_p \leq \|f\|_p \leq B_p \|Q(f)\|_p$$

for all $f \in \mathrm{L}^p$, $1 < p < \infty$. He also considered these inequalities for the weights x^α, building on earlier pioneering work done for the Walsh system by Hirschman [1].

Paley's inequality fails to hold for Vilenkin systems of unbounded type. Watari [1] proved that given any such system and any $p < 2$ there is an $f \in \mathrm{L}^p$ such that $Q(f) = \infty$ a.e. on $[0, 1)$.

Paley's inequality also fails for $p = 1$, even in the Walsh case. However, Bočkarev [3] has the following result. Let $f \in \mathrm{L}^1$. There is a monotone increasing sequence of natural numbers $(n_k, k \in \mathrm{N})$ and a sequence of real numbers y_k such that

$$\| \left(\sum_{n=n_k+1}^{n_{k+1}} d_n^2 \right)^{1/2} \|_\infty \leq y_k$$

and if there exists a $C_1 > 0$ for which

$$\sum_{k=m}^{\infty} y_k < C_1 y_m \qquad (m \in \mathrm{N}),$$

then

$$\| \sum_{k=0}^{\infty} \frac{1}{y_k} \sum_{n=n_k+1}^{n_{k+1}} d_n^2 \|_1 \leq 4C_1 \|f\|_1,$$

where d_n represents the martingale difference introduced above.

Sjölin [1] has another proof of Paley's inequality, and Gosselin [4] has a vector-valued version.

The second inequality of Corollary 5 was proved by Burkholder [1].

Corollary 6 was obtained by Paley [1]. It is equivalent to convergence of Sf in L^p norm for each $1 < p < \infty$ (see Exercise 3.3). It holds for arbitrary Vilenkin systems, as was shown independently by Schipp [13], Simon [2], and Young [3]: Schipp proved that Fourier series of $f \in L^p$ for $1 < p < \infty$ converge in L^p norm for a large class of product systems which includes the Vilenkin systems; Simon introduced a conjugate function for Vilenkin systems, proved analogues of the inequalities of Riesz and Kolmogorov, and deduced Corollary 6 from them; and Wo- Sang Young worked with a modified Dirichlet kernel in the Vilenkin system, showing directly that the associated partial sum operators were uniformly of weak-type $(1,1)$.

3.4. Dyadic Hardy spaces and BMO are special cases of a general theory valid for any probability space with a fixed increasing sequence of sub-σ-fields $\mathcal{B}_n$. In place of $\mathcal{E}_n$ one uses $E(\cdot \mid \mathcal{B}_n)$, and in place of $\Delta_n f$ one uses $f_n - f_{n-1}$. Details can be found in Garsia [1].

In the classical case, Hardy spaces can be characterized in three ways: by the maximal function, by the atomic decomposition, and by the conjugate function. For Vilenkin groups of unbounded type these concepts may give rise to different Hardy spaces.

Simon [2] introduced a conjugate function for Vilenkin groups in the following way. Let

$$\widetilde{D}_N := -\imath \sum_{k=0}^{N} D_{M_k} \sum_{-[(m_k-1)/2]}^{[m_k/2]} \operatorname{sgn} j \, r_k^j$$

and set

$$\widetilde{f}_N := f * \widetilde{D}_N \qquad (N \in \mathrm{N}).$$

He showed that $\widetilde{f}_N$ converges a.e., as $N \to \infty$, for any $f \in L^1$. Moreover, if

$$\widetilde{f} := \lim_{N \to \infty} \widetilde{f}_N$$

then the map $f \to \widetilde{f}$ is of type (p,p) for $1 < p < \infty$ and of weak-type $(1,1)$. Simon [3] also introduced a Hardy-Littlewood maximal function Mf to Vilenkin groups and showed that the map $f \to Mf$ is of type (p,p) for $1 < p < \infty$ and of weak-type $(1,1)$. These operators are related by the Hunt type inequality:

$$|\{Mf \le y, \ |\widetilde{f}| > \lambda y\}| \le Ce^{\lambda y}$$

for $f \in L^1$, $\lambda, y > 0$ and C an absolute constant (see Simon [4]).

Fridli and Simon [1] have shown that martingale Hardy spaces for Vilenkin systems of unbounded type have no atomic structure.

Theorem 6 appeared for the first time in Coifman and Weiss [1]. For Vilenkin systems of bounded type, it was proved by Chao [3]. As indicated above, it is not true for Vilenkin systems of unbounded type.

For the classical case, functions of bounded mean oscillation were introduced by John and Nirenberg [1]. For general martingales there are two corresponding spaces: BMO^+ defined by

$$\sup_{n \in \mathbf{N}} \left(E\left(|f - E(f \mid \mathcal{B}_n)|^2 \mid \mathcal{B}_n\right)\right)^{1/2} < \infty$$

and BMO defined by

$$\sup_{n \in \mathbf{N}} \left(E\left(|f - E(f \mid \mathcal{B}_{n-1})|^2 \mid \mathcal{B}_n\right)\right)^{1/2} < \infty.$$

In the dyadic case (because of predictability) these two spaces are equivalent. We have used BMO^+ for convenience.

Theorem 7 was obtained first in the classical case by Fefferman [1]. Its martingale version (for BMO) can be found in Garsia. For BMO^+, it was proved by Herz [1]. For certain Orlicz spaces it has been investigated by Mogyoròdi [1].

Theorem 8 i) is due to Ladhawala [1]. It is an analogue of a classical result of Hardy (see Zygmund [1]). Theorem 8 ii) for special case $b_k = O(1/k)$ is also due to Ladhawala. These results hold for bounded Vilenkin systems (see Chao [4]) but fail for unbounded ones if the martingale Hardy space is used (see Fridli and Simon [1]).

Notice by Theorem 8 ii) and Theorem 1 in 2.1 that $\mathrm{Lip}(1, \mathrm{L}^1) \subseteq \mathrm{BMO}$. Also notice that $W := 1 + \sum_{k=1}^{\infty} w_k/k$ belongs to $\mathrm{Lip}(1, \mathrm{L}^1)$ but not to L^{∞} (see Exercise 3.25). Hence the dyadic case is unlike the classical case (see Zygmund [1], p. 180).

Theorem 9 was discovered by Davis [2]. It can be used to obtain information about classical Hardy spaces by looking first at the more transparent dyadic case (see Garnet and Jones [1]). The simple proof of Theorem 9 presented here is due to Chao.

3.5. Theorem 10 i) was proved for the classical case by Fefferman and Stein [1]. It holds for the general martingale case (see Garsia [1]) if BMO is used in place of BMO^+.

The classical version of Theorem 10 ii) can be found in Coifman and Weiss [1]. The Walsh case is due to Schipp [16], but it is not known whether it holds for the general martingale case. For the case when the sub-σ-fields are not monotone increasing, the spaces VMO have a very complicated structure. In fact, Schipp [18] has shown that they are separable but need not have a Schauder basis (see also 5.6).

For duality of martingale Hardy spaces with non-linearly ordered stochastic basis see F. Weisz [1]. He also obtained an analogue of Theorem 8 for two-dimensional dyadic Hardy spaces.

There are many characterizations of H_0 and BMO, even in the general case (see, e.g., Garsia [1]). We have included two here, Corollaries 9 and 10.

Theorem 11 gives an atomic characterization of H^p for $0 < p < 1$. In the classical case, it is due to Coifman (see Coifman and Weiss [1]). For the general martingale case, see Herz [2].

3.6. The usual definition of the K^q spaces is slightly different but equivalent to the one given here (see Garsia [1]). These spaces can be defined for general martingales and generalize the martingale version of Fefferman's inequality in the following way. If p, q are conjugate exponents with $1 \le p \le 2$, if $f \in \mathrm{L}^p$ and $\phi \in \mathrm{K}^q$, then the expectation operator can be extended to $f\phi$ in such a way that

$$|\mathrm{E}(f\phi)| \le \sqrt{\frac{2}{p}} \|F\|_{H^p} \|\phi\|_{K^q}.$$

Theorem 12 is new in this form. It can be proved quite easily using the concept of intermediate spaces. We have presented an elementary proof here which uses the tools developed in this chapter. See also the paper of Echandia [1].

Intermediate spaces (see Bergh and Löfström [1]) were used to verify interpolation between classical and dyadic Hardy spaces. We did this because it was simple and did not delve deeply into the trigonometric structure. A proof along the lines of Theorem 12 is possible for this kind of interpolation, but would require and exposition of a canonical decomposition for trigonometric Fourier series. (See Kašin and Saakjan [1].)

3.7. After 1873 when du Bois-Reymond gave an example of a continuous function whose trigonometric Fourier series diverges somewhere on $[0, 2\pi)$ it was not clear what should be sought in the way of convergence theorems for Fourier series. In 1915, motivated by connection between Fourier series and complex function theory, Lusin pointed in the right direction by making the conjecture that the Fourier series of any continuous function converges a.e. on $[0, 2\pi)$. In 1966 Carleson [1] verified Lusin's conjecture. In fact, he proved that the trigonometric Fourier series of any L^2 function converges a.e.

Hunt [1] showed this remains true if L^2 is replaced by L^p for any $p > 1$. Both Carleson's and Hunt's proof reduced convergence to showing the maximal partial sum operator S^* is of weak- type. This method was transferred to the Walsh system by Billard [1], Hunt [1] and Sjölin [1], and to bounded Vilenkin systems by Gosselin [1]. The proofs are difficult and at times hard to follow.

In this section we show that for the Walsh system, the necessary weak-type (p, p) estimates of S^* can be obtained by martingale techniques. The martingales are not indexed by a linearly ordered set as in the classical case, but by the tree-like collection of dyadic intervals (see (67)). This approach is due to Schipp [14], who showed that any product system of complex valued martingale differences with absolute value 1 is a convergence system. Thus if $f \in \mathrm{L}^2$ then the Walsh-Fourier series of f converges a.e. The case $p > 1$ can be found in Schipp [15]. For inequality (89), see Schipp [15]. For inequality (90), see Schipp [17]. For $p > 2$ and the non-predictable case, see Fridli and Schipp [1].

CHAPTER 4

4.1. Theorem 1 is due to Paley [1].

Theorem 2 is due to Onneweer [11]. A Walsh-Kaczmarz version of it has also been obtained by Skvorcov [18]. Similar results hold for Vilenkin systems of bounded type (see Onneweer [2], [11], Onneweer and Waterman [1], and Simon [5]). For Vilenkin systems of unbounded type, Simon [7] has proved the following. If $f \in L^1(G)$ and $\omega^{(1)}(f, M_k^{-1}) = o(\log M_{k+1})^{-1}$, as $k \to \infty$, then the Vilenkin-Fourier series f converges in $L^1(G)$ norm. Moreover, the condition $o(\log M_{k+1})^{-1}$ can be replaced by $o(k^{-1})$ if and only if $\limsup_{k\to\infty}(\log M_{k+1}/k) < \infty$. (His techniques also establish the corresponding statement for the martingale Hardy space $H^1(G)$.)

Concerning the rate of growth of the L^1 norm of partial sums of a Walsh-Fourier series, Simon [9] has obtained an analogue of a theorem of Smith [1]. In fact, he proved that for any Vilenkin group G of bounded type the sequence

$$\frac{1}{\log n}\sum_{k=1}^{n}\frac{1}{k}\|S_k f\|_1$$

converges to $\|f\|_1$, as $n \to \infty$, for all $f \in H^1(G)$. This result also holds for some functions $f \in L^1(G)\setminus H^1(G)$, but not all.

4.2. Theorem 3 is due to Onneweer [1]. It also holds for the Walsh-Kaczmarz system (see Skvorcov [19]) and for Vilenkin systems of bounded type (see Onneweer and Waterman [2]).

Some conditions on the coefficients are strong enough by themselves to imply a Walsh series converges uniformly. In fact, Watari [4] has obtained the following analogue of the Salem-Zygmund theorem. If $(a_k, k \in \mathbb{N})$ is a sequence of real numbers such that

$$\sum_{k=1}^{\infty} a_k^2(\log k)^{1+\varepsilon} < \infty$$

for some $\varepsilon > 0$ then the Walsh series $\sum_{k=0}^{\infty} a_k r_k(x) w_k$ converges uniformly for almost every $x \in [0,1)$.

Corollary 1 is due to Onneweer. It shows that any continuous function of bounded variation has a uniformly convergent Walsh- Fourier series. This special case of Corollary 1 goes back to Vilenkin [1], who proved it for Vilenkin systems of bounded type. For the unbounded case, Simon [7] has shown it holds if and only if $\limsup_{k\to\infty}(\log M_k/k) < \infty$.

Theorem 4 is due to Onneweer and Waterman [1], who obtained it for Vilenkin systems of bounded type. In [2], they strengthened Theorem 4 to include functions of harmonic bounded fluctuation. Theorem 4 also holds for the Walsh-Kaczmarz-Fourier series (see Skvorcov [17]).

Li Luan Tan [1] has obtained the following analogue of the Steinhaus theorem. If $g \in \operatorname{Lip} 1$ on $[a,b] \subset [0,1]$ and $f \in L^1$ then $S_n(fg) - gS_n f$ converges uniformly to 0 on compact subsets of $[a,b]$. This result has been generalized to Vilenkin systems of bounded type by Zubakin [1]. Further results along this line can be found in Tevzadze [1] and Efimov [1].

The estimate $S_n f = o(\log n)$ for $f \in C(G)$ was obtained first by Fine [1].

Guličev [2] has estimated the rate of uniform convergence of a Walsh-Fourier series using Lebesgue constants L_k and the classical modulus of continuity $\Omega(f, 1/k)$. He has shown that

$$\liminf_{k\to\infty}\frac{\|S_k f - f\|_\infty}{\Omega(f,1/l)L_k} = 0$$

for $f \in \mathcal{C}[0,1]$ and

$$\liminf_{k\to\infty}\frac{\|S_k f - f\|_\infty}{L_k} = 0$$

for $f \in L^\infty$. This is in sharp contrast to the trigonometric case where Oskolkov and Busko proved that the corresponding limit infima are strictly positive.

This problem has been considered for Vilenkin groups G by Fridli [1]. Let X be a Banach space of integrable functions defined on G and let

$$V_k(f,X) := \frac{\|S_k f - f\|_X}{\omega^X(f,1/k)L_k}$$

for $k \in \mathbf{P}$. If $X = C(G)$ or $L^p(G)$ for some $1 < p < \infty$, then $\liminf_{k\to\infty} V_k(f, X) = 0$ for $f \in X$. On the other hand, for $X = L^1(G)$ or $H^1(G)$ (the martingale Hardy space generated by the underlying Vilenkin system), analogy with the trigonometric system is restored. Indeed, $\liminf_{k\to\infty} V_k(f, X) > 0$ in both cases. When the Vilenkin group G is of bounded type, it is also the case that $\limsup_{k\to\infty} V_k(f, X) < \infty$ but if G is unbounded then there is an $f \in X$, such that $\limsup_{k\to\infty} V_k(f, X) = \infty$.

Fine [1] proved that the Walsh-Fourier series of $f \in C(G)$ is uniformly Cesàro summable. This result is true for any Vilenkin system of bounded type (see Vilenkin [1]), but Price [5] has shown that for every Vilenkin system of unbounded type there is a continuous function whose Cesàro means do not converge uniformly.

For uniform strong summability of Walsh-Fourier series of $f \in C(G)$ see Schipp [3].

4.3. Although du Bois-Reymond was the first to show that the trigonometric Fourier series of a continuous function may diverge at certain points, the first simple construction of such series was obtained by Fejér. His method centered on constructing certain polynomials with properties like those in Theorem 5.

The Walsh-Fejér polynomials were introduced by Schipp [1], and Theorems 5 and 6 are due to him. These results hold for Vilenkin systems, bounded or unbounded, as was proved by Simon [1], [2].

4.4. The definition of homogeneous Banach spaces on the circle group $\mathbf{T}$ is well-known (see Katznelson [1]). It can be formulated for subspaces of $L^1(G)$ for any compact group G (see Butzer and Nessel [1]). In this setting Lemma 1 and Theorem 7 are still true.

The definition given here is slightly different but (in our setting) equivalent to the usual one. We have substituted (iii) for

$$\text{(iii)'} \qquad \lim_{x\to x_0} \|\tau_x f - \tau_{x_0} f\|_X = 0 \qquad (f \in X, x_0 \in G).$$

Lemma 2 (especially that the Walsh-Fejér kernels K_n satisfy (9)) is due to Yano [4].

Theorem 8 i), for the case $X = L^p(G)$ where $1 < p < \infty$, follows immediately from the fact that the Walsh system is a basis for $L^p(G)$. The case $X = L^1(G)$ is due to Morgenthaler [1]. For the case $X = C(G)$, it shows that a Walsh series S is the Walsh-Fourier series of some function $f \in C_W$ if and only if the series S is uniformly Cesàro summable on $[0, 1)$. Fine [1] was first to recognize that this characterization fails to hold if C_W is replaced by $\mathcal{C}[0, 1]$, the collection of functions classically continuous on $[0, 1]$.

Theorem 8 ii) is due to Morgenthaler [1] and Theorem 8 iii) is due to Fine [1].

4.5. The proofs of Theorems 9, 10, 11 and Corollary 2 are routine adaptations of the corresponding trigonometric techniques developed by Kahane and Katznelson (see Katznelson [1]).

For the Walsh case they were first written down by Harris and Wade [1]. (For Vilenkin systems of bounded type see Heladze [2].)

Notice Theorem 11 is best possible for $1 < p < \infty$. Indeed, in view of Theorem 14 in 3.7, a measurable subset of G is a set of divergence for $L^p(G)$, $1 < p < \infty$, if and only if it is of Haar measure zero. It is not known whether the same holds for $L^\infty(G)$ and for $C(G)$.

The proof of Theorem 12 is an adaptation of the corresponding trigonometric proof due to Kahane and Katznelson. Of course such integrable functions had been constructed earlier (see Stein [1], and Schipp [5], for example). For multidimensional Walsh-Fourier series, Theorem 12 is due to Sanadze and Heladze [1]. Simon has proved Theorem 12 for Vilenkin systems (see Schipp and Simon [3]).

Sets of divergence for $C(G)$ are more difficult to construct. Schipp [1] showed every singleton is one, and in [3] constructed an uncountable one. Harris and Wade [1] also constructed uncountable ones by a Cantor-like process. And, Onneweer [10] has shown that every subset of G of logarithmic Hausdorff measure zero is one. It is thought that every subset of Haar measure zero is one, but this problem remains open. Theorem 13 is a partial solution to this problem. It is due to Harris [1], who proved it for arbitrary Vilenkin systems.

Sets of divergence for other classes of functions have been studied by Lukašenko [1], Harper [1], and Kobayashi [1]. Harper's result is the following one. For each $0 \le \alpha < 1$ let S_α represent the collection of functions f which satisfy

$$\sum_{k=0}^{\infty} 2^{k(1-\alpha)} \sum_{n=2^k}^{2^{k+1}-1} |\hat{f}(k)|^2 < \infty.$$

Then a closed set $E \subset G$ is a set of divergence for $\mathcal{S}_\alpha$ if and only if E is of α-capacity zero. This is a refinement of Theorem 11 for $p = 2$, because $\mathcal{S}_\alpha \subset L^2(G)$ for $0 \le \alpha < 1$ and $\mathcal{S}_\alpha = L^2(G)$ for $\alpha = 1$. Moreover, every countable set is a set of α-capacity zero, and every set of α-capacity is of Haar measure zero.

4.6. Theorem 14 was obtained for trigonometric Fourier series by Men'šov in 1940 (see Zygmund [1]). For the Walsh case, it was first verified by Kotljar [1]. The proof given here is due to Price [4]. His techniques show that the adjustment from f to g can be made to depend only on ε and the modulus of continuity of f (see Exercise 4.9). This result also holds for Vilenkin systems of bounded type (see Onneweer [4]).

Olevskiĭ [5] showed that there is a classically continuous function f on $[0,1)$ such that if g is any integrable function which coincides with f on a set of positive measure, then

$$\sum_{k=0}^{\infty} |\widehat{g}(k)|^p = \infty$$

for all $0 \le p < 2$. Thus Theorem 15 is a corollary of Olevskiĭ's result. The proof given here is essentially due to Katznelson [2]. It was adapted to the Walsh case by Guličev [1].

CHAPTER 5

5.1. Theorem 1 was first proved for the trigonometric case by Haršiladze and Ložinski (see Cheney [1] for example).

It holds for any system of characters.

We have shown a best approximation in $C(G)$ is not necessarily unique. The same is true of $L^1(G)$. However, since $L^p(G)$ is strictly convex for $1 < p < \infty$, the best approximation in one of these spaces is always unique (see Cheney [1]).

Theorem 2 shows a close connection between modulii of continuity, best approximations, and certain tails of Walsh-Fourier series. It is not difficult to see that Theorem 2 holds for any Vilenkin system. For the Walsh case and $X = C(G)$ or $L^p(G)$, Theorems 2 and 4 are due to Watari [2]. For the Vilenkin case they are due to Efimov [1].

Ul'janov has shown that the L^q norm of a function f can be estimated by the rate of best trigonometric approximations in L^p spaces for $p < q$. Golubov [1] has obtained a Walsh version of Ul'janov's results. Specifically, if $1 \le p, q < \infty$ and $f \in L^q$ then

$$\|f\|_q \le C\left(\|f\|_p + \Big(\sum_{k=0}^{\infty} k^{q/p-2}(\mathbf{E}_k(f, L^p))^q\Big)^{1/q}\right)$$

and

$$\mathbf{E}_k(f, L^q) \le C\left(\mathbf{E}_k(f, L^p)n^{1/p-1/q} + \Big(\sum_{k=0}^{\infty} k^{q/p-2}(\mathbf{E}_k(f, L^p))^q\Big)^{1/q}\right).$$

Here C is an absolute constant.

Theorem 3 was proved for the cases $X = C(G)$ and $L^p(G)$, $1 \le p < \infty$, by Watari [2]. It is easy to see that Theorem 3 holds for any bounded Vilenkin system. However, in the unbounded case only b), c) and e) are equivalent; conditions a) and d) do not follow from them.

By Theorem 7 in 4.4, $\sigma_n f$ converges to f in X, as $n \to \infty$ for all $f \in X$. Theorem 4 estimates the rate of convergence when $f \in \mathrm{Lip}(\alpha, X)$ for $\alpha > 0$. The cases $X = L^p(G)$, $1 \le p < \infty$ were originally discovered by Watari [2]. They now are known to follow from a more general result of Bljumin [1],[2] which also contains an analogue of a theorem of Stečkin and Fomin.

The rate of convergence of (C, β) sums has also been investigated for the special cases $X = L^p(G)$ and surprisingly, no change occurs. First, Yano [1] proved that if $f \in \mathrm{Lip}(\alpha, C[0,1])$ for some $0 < \alpha < 1$ and if $\beta > \alpha$, then $\|\sigma_n^\beta f - f\|_\infty = O(n^{-\alpha})$, as $n \to \infty$. And, if $f \in \mathrm{Lip}(\alpha, L^p)$ for some $0 < \alpha < 1$ and if $\beta > \alpha$ then $\|\sigma_n^\beta f - f\|_p = O(n^{-\alpha})$ as $n \to \infty$. More recently, Skvorcov [16] has shown that these estimates hold for $0 < \beta \le \alpha$ as well. Skvorcov also proves that if $f \in \mathrm{Lip}(1, L^p)$ for some $1 \le p < \infty$ and if $\beta > 0$ then $\|\sigma_n^\beta f - f\|_p = O(\log n/n)$ as $n \to \infty$. Thus the second estimate in Theorem 4 cannot be improved by passing to (C,β) sums. Skvorcov's techniques are quite general and hold for Walsh-Kaczmarz series as well. Skvorcov [19] has also obtained these estimates for Vilenkin systems of bounded type.

Siddiqi [1] has used (C,β) means to characterize $\mathrm{Lip}(\alpha, \mathrm{L}^p)$ for $1 < p < \infty$ and $0 < \alpha < 1$. He proved that $f \in \mathrm{Lip}(\alpha, \mathrm{L}^p(\mathrm{G}))$ if and only if $\|\sigma_n^\beta f - \sigma_m^\beta f\|_p = O(n^{-\alpha})$ as $n \to \infty$, for all $0 < \beta \leq 1$, and $m > n$. He also considered the case $p = \infty$.

Concerning (C,β) sums for $-1 < \beta < 0$, Survillo [2] has proved analogues of trigonometric theorems due to Men'šov [1], [2], and Tevzadze [2] has considered uniform (C,β) summability.

Kokilašvili [1] has obtained the following estimate for (C,β) sums when $\beta \geq 1$. If $f \in \mathcal{C}[0,1]$ and $\beta \geq 1$ then $\|\sigma_{n-1}^\beta f - f\|_\infty \leq (1/n)\sum_{k=1}^n \mathbf{E}_k(f, \mathcal{C}[0,1])$. The trigonometric analogue of this result is due to Stečkin [1].

Other means have been examined. Tateoka [1] studied Abel means (see also Fine [1]). De la Vallee Poussin means have been studied by Bljumin [2]. (His results hold for Vilenkin systems of bounded type.) Triangular summability has been studied by Li Luan Tan [1] and Survillo [1]. And the summability method of Rogosinski has been studied by Haršiladze [1] and Morgenthaler [1].

For strong summability see Schipp [3]. His results hold for a large class of positive permanent summability methods which was investigated in the trigonometric case by Leindler [1]. Special cases of his results include the work of Sunouchi [2] on strong (C,p) sums, and that of Yano [2] on (C,p) means. All these results have analogues for the trigonometric and other polynomial systems (see Alexits [1], [3], Alexits and Kràlik [1], [2], Stečkin [1], and Bljumin [1]). The trigonometric analogue of the first inequality of Theorem 5 goes all the way back to Marcinkiewicz (see Zygmund [1]).

Investigating Walsh analogues of results due to Marcinkiewicz and Zygmund, Sunouchi [1], [2] introduced the operators

$$U_p(f) := \left(\sum_{n=1}^{\infty} \frac{1}{n}|S_n f - \sigma_n f|^p\right)^{1/p} \qquad (p \geq 2)$$

and

$$T(f) = \left(\sum_{k=0}^{\infty} |S_{2^k} f - \sigma_{2^k} f|^2\right)^{1/2}$$

for $f \in \mathrm{L}^1$. He showed that the L^r norms of $T(f)$, $U_p(f)$, and f are equivalent for each $1 < r < \infty$. Since this result fails if $r = 1$, it is of interest to determine whether $T(f)$, $U_p(f)$, and f are equivalent in H norm. A partial answer to this problem was obtained by Simon [8]. He has proved that T is bounded from H to L^1 but U_p is not. His proof is valid for any Vilenkin system of bounded type.

For other results concerning summability and approximation by Walsh series, we refer the reader to Siddiqi [4].

5.2. The strong derivative on the dyadic group was introduced by Butzer and Wagner [1]. For the cases $X = \mathrm{L}^p(\mathrm{G})$ $1 \leq p < \infty$ and $X = \mathrm{C}(\mathrm{G})$, Theorem 6 and its two corollaries are due to them. The case $X = \mathrm{H}$ was investigated by Ladhawala [1]. He also showed that any function strongly differentiable in H is continuous on the group, but there exist BMO functions (not continuous on the group) which are strongly differentiable in $\mathrm{L}^1(\mathrm{G})$. (See also Penney [1].)

Onneweer [7] defined a "dyadic"derivative on Vilenkin groups. Pál and Simon [1] have proved a fundamental theorem of calculus for this derivative.

There are many properties satisfied by the strong derivative which fail to hold for the classical derivative. Lemma 1 shows one of them.

For the classes $X = \mathrm{L}^p(\mathrm{G})$ $1 \leq p < \infty$ and $X = \mathrm{C}(\mathrm{G})$, Lemmas 1, 2, and Theorems 7, 8 are due to Butzer and Wagner [1]. Their results have been generalized to Vilenkin groups of bounded type by Fridli [2].

Theorem 7 is the Walsh analogue of an inequality of Jackson. For a more general version see Butzer and Scherer [1].

Lemma 2 is a Walsh analogue of the Bernstein inequality. For Vilenkin groups of unbounded type, Fridli [3] has shown the following. Lemma 2 holds for $X = \mathrm{L}^p(G)$ $1 < p < \infty$ but fails to hold for $X = \mathrm{L}^1(G)$, $\mathrm{H}(G)$, or $\mathrm{C}(G)$. For these three exceptional cases, he obtains replacement inequalities which are valid for Vilenkin groups of unbounded type, and shows that these inequalities cannot be improved.

In connection with Theorem 8, we remark that Butzer and Wagner [1] have shown the following for $X = L^p(G)$ $1 \le p < \infty$ and $X = C(G)$. If $f \in \mathrm{Lip}(\alpha, X)$ for $\alpha > 1$ then f is strongly differentiable in X. On the other hand, if f is strongly differentiable in X then $f \in \mathrm{Lip}(\alpha, X)$ for all $0 < \alpha \le 1$.

Theorems 9 and 10 are consequences of known results.

5.3. In this section only particular bases are investigated. For general problems connected with bases see Banach [1], Kašin and Saakjan [1], Lindenstrauss and Tzarfriri [1], Ljusternik and Sobolev [1], and Olevskiĭ [4].

Haar [1] proved that the Haar-Fourier series of every continuous function converges uniformly on $[0,1]$. Schauder [2] showed that the Haar system is a basis in L^p for every $1 \le p < \infty$ and Orlicz [1] generalized this result to a class of separable Banach spaces which today bear his name.

Paley [1] proved that the Walsh system is a basis in L^p for $1 < p < \infty$. The fact that the Lebesgue functions (constants) of the Walsh system are not uniformly bounded follows from a more general theorem of Olevskiĭ [4] (see also Bočkarev [4]). Namely, for every uniformly bounded orthonormal system the Lebesgue functions are unbounded on a set of positive measure.

Faber [1] introduced the system ζ and proved it is a basis in $\mathcal{C}[0,1]$. Schauder [1] rediscovered this system seventeen years later and proved his basis property. We call this system the Faber-Schauder system.

The biorthogonal expansion with respect to the Faber-Schauder system was used by Ciesielski [5] to prove that the estimate $c_n(f) = O(n^{-\alpha})$ is equivalent to the condition that f belongs to the classical space $\mathcal{LIP}(\alpha)$. For approximation properties of the Faber-Schauder system see Matveev [1]. The Haar analogue of Theorem 11 is due to Sz.-Nagy [1].

5.4. In 1928 Franklin [1] introduced the system which bears his name. It was the first orthonormal basis in $\mathcal{C}[0,1]$. A systematic investigation of the Franklin system was begun by Ciesielski in [1] and [2]. Among others, the fundamental properties (25) through (34) are proved there. He also showed that the Franklin system is a basis in L^p for $1 \le p < \infty$.

The upper $\mathcal{H}$ estimate in Lemma 4 was given by Wojtaszczyk [1]. The proof presented here is taken from Chang and Ciesielski [1]. The lower H estimate is due to Schipp and Simon [2].

Theorem 12 is due to Wojtaszczyk [1]. The simple proof here is adapted from Chang and Ciesielski [1].

The Walsh shift operators were investigated by Ciesielski and Kwapien [1]. Part i) of Theorem 13 is taken from here. The rest of Theorem 13 is due to Simon [6].

5.5. The first non-trivial result on equivalence of concrete bases was the theorem of Ciesielski, Simon, and Sjölin [1] that the Haar and Franklin system are equivalent in L^p for $1 < p < \infty$. The equivalence of these systems in $(\mathrm{H}, \mathcal{H})$ and $(\mathrm{VMO}, \mathcal{VMO})$ was established by Wojtaszczyk [1]. (The proof of L^p equivalence presented here combines the Wojtaszczyk result with interpolation. This is quite different from the original Ciesielski-Simon-Sjölin proof.)

Corollary 4 was first proved by Maurey [1]. Later Carleson [2] gave a different proof using an explicit isomorphism. In showing that the Haar and Franklin systems are equivalent in $(\mathrm{H}, \mathcal{H})$, Wojtaszczyk applied Carleson's proof to the Franklin system.

The Ciesielski system was introduced by Ciesielski [1]. Ropela [1] proved this system is a basis in L^p for $1 < p < \infty$. Ciesielski, Simon, and Sjölin [1] proved the Ciesielski and Walsh systems are equivalent in L^p for $1 < p < \infty$. The proof of Theorems 15 and 16 presented here use the method Schipp introduced in [8]. Theorem 16 also follows directly from a result of Waterman [1].

The fact that the Walsh and Walsh-Kaczmarz systems are not equivalent in L^p unless $p = 2$ was proved by Bahšečjan [2]. This was based on results of Semjonov, formulated here in the second part of Theorem 17 and in the L^p part of Corollary 5.

The first part of Theorem 17, i.e., the exact form for the H and BMO norms of R and the proof presented here are due to Schipp.

Theorem 18 was proved by Wo-Sang Young [4]. The proof presented here is different from hers and is based on the multiplier result (Theorem 19) of Harris [2].

According to the celebrated result of M. Riesz that the Hilbert transform is bounded, $\mathcal{H}^p$ is Banach space isomorphic to L^p for $1 < p < \infty$. Consequently, the existence of a basis for such p is obvious. The first basis for $\mathcal{H}^1$ was constructed by Billard [2] by means of the Haar system. Using an analytic extension of the Franklin system, Bočkarev [4] constructed an orthogonal basis for the space of functions analytic on the interior of the unit disc and continuous on its boundary.

A basis $(\epsilon_n, n \in \mathbf{N})$ in a Banach space $\boldsymbol{X}$ is called unconditional if for every bijection $\varpi : \mathbf{N} \to \mathbf{N}$ the sequence $(\epsilon_{\varpi(n)}, n \in \mathbf{N})$ is a basis in $\boldsymbol{X}$. Using Paley's inequality, Marcinkiewicz [1] proved for $1 < p < \infty$ that the Haar system is an unconditional basis in L^p. The Franklin analogue of this result is due to Bočkarev [4]. It is easy to see that the Haar system is an unconditional basis in H and in VMO. Consequently, by Theorem 14 the Franklin system is an unconditional basis in $\mathcal{H}$ and in $\mathcal{VMO}$.

It is known that the space of functions analytic on the interior of the unit disc and continuous on its boundary has no unconditional basis (see Pelczynski [1]). Furthermore, neither do L^1 and $\mathcal{C}[0,1]$. (For details see Kašin and Saakjan [1] or Olevskiĭ[4].)

5.6. The basis problem was posed by Banach [1] in 1932. It was solved by Enflo [1] in 1973. Other authors have since constructed other separable Banach spaces which have no basis (see Lindenstrauss and Tzarfriri [1] and Szankowski [1], [2]). The fundamental step of Lemma 10 is due to Enflo.

The space BMO($\mathcal{G}$) defined by

$$\|f\|_{\mathrm{BMO}(\mathcal{G})} := \sup_{\mathcal{G}\in\mathcal{G}} \|\mathcal{E}(|f - \mathcal{E}(f|\mathcal{G})| \mid \mathcal{G})\|_\infty < \infty.$$

The space VMO($\mathcal{G}$) is defined to be the closure of the set of dyadic step functions in BMO($\mathcal{G}$). Schipp [18] has shown that VMO($\mathcal{G}$) does not have basis.

CHAPTER 6

6.1. Theorem 1 is due to Schipp [6]. This result has been generalized to arbitrary Vilenkin systems (Simon [8]) and to double Walsh-Fourier series (Sidorov [1]).

Yano [1] proved that if $1 < p < \infty$ and

$$\int_0^1 \int_0^1 \frac{|f(x+t) - f(x-t)|^p}{t}\, dt\, dx < \infty$$

then the Walsh-Fourier series of f converges a.e. Lemmas 1 and 2 and Theorem 2 were proved by Schipp [5]. A similar result where dyadic addition is replaced by the usual addition was obtained by Yano [1]. The trigonometric version is of course much older, and was discovered by Marcinkiewicz (see Bary [1], p. 374).

A trigonometric analogue of Theorem 3 can be found in Zygmund [1], p. 66. Theorem 3 also remains valid if $\log_2 n$ is replaced by $V(n)$, the variation of n (see Šneider [3]).

6.2. Fine [3] proved every Walsh-Fourier series is a.e. (C,α) summable for $\alpha > 0$. His argument is an adaptation of the older trigonometric analogue due to Marcinkiewicz. Schipp [12] gave a simpler proof for the case $\alpha = 1$ using the operators $\mathcal{F}_n$. The operators $\mathcal{R}_n$ were introduced by Schipp [11] to study dyadic differentiability of Stieltjes measures. These techniques have been expanded here to obtain the new results: Theorems 4, 5, and 6. The weak estimate of σ^* in Corollary 2 is due to Schipp [12]. That σ^* is $(\mathrm{L}^1, \mathrm{H})$ bounded was discovered by Fujii [1]. Corollary 5 is a result of Fine [4]. For generalizations to Vilenkin systems see Pál and Simon [1]. Corollary 5 also holds if S_{2^n} replaces σ_n.

Other types of generalizations replace Cesàro summability with strong summability. For example. Schipp [2] proved that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} |S_k f - f|^2 = 0$$

a.e. for every $f \in \mathrm{L}^1$. This result was generalized to strong (C,α) summability for $\alpha > 0$ by Cybertowicz [1] and to multiple Walsh-Fourier series by Šarašenidze [1].

Corollaries 7 and 8 and the dyadic analogue of the Hardy-Littlewood maximal inequality were proved by Schipp in [7] and [11]. These results have been generalized to bounded Vilenkin systems by Pál and Simon [2].

6.3. Corollary 9 and Theorem 7 are due to Sjölin [1]. The idea of using martingale trees to obtain Corollary 9 (in particular, Lemma 5) is new.

The block spaces were introduced by Taibleson and Weiss [1]. They proved Lemmas 6, 7, 8, and Theorem 8 for trigonometric Fourier series. Their proofs were adapted here to the dyadic case.

There is quite a bit known about these spaces. Every function of finite entropy belongs to the block space $\mathcal{B}_\infty$, hence to every block space (Taibleson and Weiss [1]). Thus functions of finite entropy have a.e. convergent Walsh-Fourier series. Moreover, Soria [1] proved that a function f belongs to $\mathcal{B}_\infty$ if and only if

$$\int_0^\infty \lambda_f(t)\left(1+\log^+\left(\frac{1}{t\lambda_f(t)}\right)\right) dt < \infty,$$

where $\lambda_f(t) := |\{|f| > t\}|$ is the distribution function of f.

The block space $\mathcal{B}_1$ is equivalent to L^1. On the other hand, even the block space $\mathcal{B}_\infty$ is not comparable to $\mathrm{L}\log^+\mathrm{L}$. In fact, $\mathcal{B}_\infty$ is not a subset of any Orlicz space except L^1. Nevertheless, $\mathcal{B}_\infty$ does contain the Dini class, i.e., those integrable functions which satisfy

$$\int_0^1\int_0^1 \frac{|f(x)-f(y)|}{|x-y|}\, dx\, dy < \infty.$$

Block spaces have been used in higher dimensions to study a.e. Bochner-Riesz summability at the critical index (see Shan-Zhen, Taibleson and Weiss [1], and Shan-Zhen [1]). Connections between block spaces and a.e. convergence of multiple Walsh-Fourier series remain unexplored.

6.4. Theorems 9 and 10 originally appeared in Schipp [11]. Bahšečjan [2] generalized these results to piecewise isomorphic rearrangements.

Theorem 11 was proved by Wo-Sang Young [2]. Theorem 12 was proved by Schipp [6] and is a generalization of a result of Alexits. It was Alexits who introduced the notion of weakly multiplicative systems. For other results concerning such systems see Alexits [2] and Mòricz [2].

Theorem 12 can be generalized. The best result along these lines is due to Tandori [5] who introduced the following notion. A system $\boldsymbol{\gamma} = (\gamma_n, n \in \mathrm{N})$ in $\mathrm{L}^\infty(\Omega,\nu)$ is called locally weakly 2-multiplicative if for every measurable subset $E \subseteq \Omega$ and for every $\varepsilon > 0$ there exists a set $E_0 \subset E$ with $\nu(E_0) > \nu(E) - \varepsilon$ such that

$$\sum_{k=0}^\infty \left|\int_{E_0} g_k\, d\nu\right|^2 < \infty$$

where $\boldsymbol{g} := (g_k, k \in \mathrm{N})$ is the product system generated by $\boldsymbol{\gamma}$. He proved under the condition $\|\gamma_n\|_\infty \le 1$, $n \in \mathrm{N}$, that the system $\boldsymbol{g}$ is a convergence system if and only if $\boldsymbol{\gamma}$ is locally weakly 2-multiplicative.

Theorem 13 and Corollary 10 were proved by Schipp [8]. Corollary 10 has been generalized by Ciesielski, Simon, and Sjölin [1]. Among other things, they showed that if $f \in \mathrm{L}(\log^+\mathrm{L})^3$ then the Ciesielski-Fourier series of f converges a.e.

6.5. Kolmogorov gave the first example of an integrable function with an everywhere divergent Fourier series (see Zygmund [1], p.310). For recent results connected with Kolmogorov's example see Ul'janov [8].

The existence of divergent Walsh-Fourier series was first proved by Stein [1]. Bočkarev [3] has shown these results to be part of the general theory of bounded orthonormal systems. Indeed, he proved that given any uniformly bounded orthonormal system there exists an integrable function whose Fourier series diverges on a set of positive measure. For bounded Vilenkin systems, see Heladze [2], [3].

Theorem 14 was proved in a slightly different formulation by Schipp [5]. This is an analogue of an earlier trigonometric result by Tandori [4]. From Theorem 14 in the special case $\Phi = 1$ we obtain a result of Ladhawala and Pankratz [1], namely that there exist functions in H with a.e. divergent Walsh-Fourier series. In the same paper they proved that if $(n_k, k \in \mathrm{N})$ is lacunary, i.e., $n_{k+1}/n_k \ge q > 1$ for $k \in \mathrm{N}$, then $S_{n_k} f \to f$ a.e., as $k \to \infty$, for every $f \in \mathrm{H}$. Another corollary of Theorem 14 is the result of Moon [1]. There exist functions $f \in \mathrm{L}(\log^+\log^+\mathrm{L})^{1-\varepsilon}$ for $\varepsilon > 0$ whose Walsh-Fourier series diverge everywhere.

Theorem 15 is due to Schipp [5]. It is an analogue of an earlier result of Prohorenko for the trigonometric system. Riesz products were also used by Schipp and Simon [3] to construct a.e. divergent trigonometric Fourier series, and by Simon [10] to construct a.e. divergent Vilenkin-Fourier series.

Theorems 16 and 17 are due to Schipp [5]. They are analogues of trigonometric results of Marcinkiewicz (see, for example, Bary [1]). For other counterexamples concerning the trigonometric system and for Theorem 18 see Körner [1].

Theorem 19 is a special case of result of Fridli and Schipp [2]. Its trigonometric analogue is due to Galstian [1]. Lemma 12 is a modification of a result of Stečkin (see, for example, Bary [1]).

Lemma 14 and Theorems 20 and 21 are due to Balašov [2].

Some rearrangements of the Walsh system are not convergence systems. This follows from a general theory obtained independently by Olevskiĭ [1] and Ul'janov [2]. They showed that every complete orthonormal system has a rearrangement which is not a convergence system.

An increasing sequence of real numbers $(\omega_n, n \in \mathbf{P})$ is called a Weyl multiplier for a.e. unconditional convergence for an orthogonal system $(\varphi_n, n \in \mathbf{P})$ if

$$\sum_{n=1}^{\infty} a_n^2 \omega_n < \infty$$

implies that every rearrangement of

$$\sum_{n=1}^{\infty} a_n \varphi_n$$

is a.e. convergent. The first general result in this direction is due to Tandori [2], who proved that sequences $\omega_n = o(\log\log n)$ are not Weyl multipliers for a.e. unconditional convergence for the Walsh system. A similar result for the trigonometric system was proved by Mòricz [1]. The best known result was obtained independently by Bočkarev [2] and Nakata [1]. They proved that if

$$\sum_{n=1}^{\infty} \frac{1}{n\omega_n} = \infty$$

then there is a sequence of real numbers $(a_n, n \in \mathbf{N})$ such that $\sum_{n=1}^{\infty} a_n^2 \omega_n < \infty$ and the Walsh series $\sum_{n=1}^{\infty} a_n w_n$ has an a.e. divergent rearrangement. Moreover, Bočkarev has shown that if ω is a modulus of continuity satisfying

$$\sum_{n=2}^{\infty} \frac{\omega(1/n)}{n\sqrt{\log n}} = \infty$$

then there is a continuous function f with classical modulus of continuity satisfying $\Omega(f, \delta) = O(\omega(\delta))$ as $\delta \to 0$ such that the Walsh-Fourier series of f has an a.e. divergent rearrangement.

Positive results for Weyl multipliers for the Walsh system are no better than those for the general orthonormal systems. For a general discussion of these results see Olevskiĭ [4] or Bočkarev [3].

6.6. Theorem 22 was proved by Mòricz [3]. Theorems 23, 24, and 25 can be found in Harris [2]. In regard to a.e. convergence of rectangular partial sums of Walsh-Fourier series and trigonometric Fourier series see Mòricz [4].

CHAPTER 7

7.1. Uniqueness of everywhere convergent Walsh series was considered by Walsh [1] and Vilenkin [1].

Fine [1] was first to obtain uniqueness for Walsh series which converge at all but countably many points in $[0, 1)$. In fact, he obtained Corollary 1 in the special case when $f = 0$ and the coefficients of S tend to zero (stronger than the C-S condition).

He conjectured this result would hold for any $f \in \mathrm{L}^1$. His conjecture was solved independently by Arutunjan and Talaljan [1] (see 7.2) and Crittenden and Shapiro [1].

Crittenden and Shapiro [1] obtained Theorems 1 and 2 by using the first integral (a dyadic adaptation of Riemann's approach to uniqueness of trigonometric series) and the Baire category theorem (see Exercises 7.18, 7.19, and 7.20). Their argument was simplified somewhat by Lindahl [1].

Grubb [1] proved Theorem 2 using quasi-measures. We present his simple approach here. Since it uses only that certain partial sums of Walsh-Fourier series are averages and every Walsh series is already a Walsh-Fourier-Stieltjes series, Theorem 2 is valid for any Vilenkin system, bounded or not. In fact, Grubb [1] has shown that uniqueness holds in certain zero-dimensional compact spaces with or without a group structure.

7.2. Arutunjan and Talaljan [1] solved Fine's conjecture (see 7.1 above) by obtaining uniqueness first for Haar series (see Exercise 7.6) and deducing the Walsh result as a corollary by means of the Hadamard transform. They obtained Lemmas 4 and 5 and Theorems 3 and 5 for Walsh series whose coefficients tend to zero and whose 2^{n_j}-th partial sums converge off a countable set. Wade [1] noticed their arguments applied to Walsh series whose 2^{n_j}-th partial sums converge a.e. if a finiteness condition like (29) is assumed. The fact that these results also hold for Walsh series satisfying the C-S condition is a recent discovery (see Wade [11]).

Generalizations of Theorems 3 and 5 have been intensively pursued by Skvorcov. Under the assumption that the coefficients of S tend to zero he has the following results. Theorem 5 holds if f is Denjoy integrable in the narrow sense [2], but need not hold if f is Denjoy integrable in the wide sense [15]. Theorem 5 holds if 2^{n_j} is replaced by a sequence of integers $(m_j, j \in \mathbf{N})$ provided the m_j's satisfy some additional conditions [6], [10]. Examples of these additional conditions are

$$2^j \leq m_j < 2^{j+1} \qquad (j \in \mathbf{N})$$

or

$$\ell(m_j) \leq M < \infty \qquad (j \in \mathbf{N}),$$

where $\ell(m_j)$ represents the number of non-zero dyadic digits in the expansion of m_j (see 2.2 and 7.4). Under the weaker assumption that S satisfies the C-S condition, Skvorcov [14] has shown if $2^j \leq m_j < 2^{j+1}$ for $j \in \mathbf{N}$, if f is a finite-valued Perron integrable function (i.e., Denjoy integrable in the narrow sense) such that

$$\lim_{j\to\infty} S_{m_j}(x) = f(x)$$

for all but countable many $x \in [0,1)$ then S is the Perron- Walsh-Fourier series of f.

Theorem 5 will not hold for arbitrary sequences. In fact, Skvorcov [7] has constructed a non-zero Walsh series S and a sequence of integers $(m_j, j \in \mathbf{N})$ such that $S_{m_j} \to 0$ everywhere on $[0,1)$ as $j \to \infty$.

For other results along these lines see Skvorcov [4], [13], Movsisjan [1], and Wade [5], [7].

Theorem 4 is due to Crittenden and Shapiro [1].

7.3. Lemmas 6 and 7 are due to Šneider [2]. He used them to show the Cantor middle halves set is a set of uniqueness for Walsh series (see 7.5). In the case when S_{2^n} converges to zero off a countable set, Theorem 6 goes back to Šneider [2]. The version here first appeared in Wade [2].

Theorem 7 is due to Skvorcov [11]. Skvorcov [8] has also shown that a null series can diverge properly to ∞ on a perfect set.

7.4. Lemmas 8 and 9 are new here. Theorem 8 is due to Schipp [4]. Theorem 9 was discovered by Skvorcov [12].

7.5. With the exception of Theorem 14, the results of this section are analogues of earlier results about the trigonometric system (see Zygmund [1], pp. 344-352).

The opening remarks concerning U-sets and M-sets are due to Šneider [2]. Coury [1], [5] has shown there exist M-sets of measure zero which are dense in $[0,1]$. Theorem 10, a Walsh analogue of a theorem of Bary, was verified by Wade [2]. Šneider [2] had proved it earlier in the special case when the E_n's were contained in intervals I_n which were pairwise disjoint. Šneider also proved Theorem 11. A related result was obtained by Lippman and Wade [1]. They showed every closed U-set is a countable union of elementary U-sets (see Exercise 7.13).

Theorem 12 can be found in Wade [6]. It has been generalized by Yoneda [5].

Theorem 13 is due to Yoneda [2]. He has other results on Dirichlet sets in [4] and [7].

Theorem 14 is due to Yoneda [2] and has no known trigonometric analogue. The proof here is new and quite simple.

In the trigonometric case there is a fascinating connection between sets of uniqueness and number theory. Indeed, Salem and Zygmund (see Zygmund [1], p. 152) completely determined which Cantor sets of constant dissection ratio were sets of uniqueness and which were not. It followed from their work (Zygmund [1], p. 349) that the property of being a set of uniqueness depends not so much on "thinness" as on the number theoretic properties that determine the set itself. In regard to the connection between algebraic numbers and harmonic analysis in general, see the book by Meyer [1]. For a dyadic group version of the theorem of Salem and Zygmund, see Aubertin [1].

If a growth condition is placed on the coefficients, then there exist U-sets of positive measure. In fact, Gevorkjan [1] proved given any sequence $(\varepsilon_n, n \in \mathbb{N})$ which decreases monotonically to zero there exists a set E in $[0,1)$ of measure 1 such that if $S := \sum_{k=0}^{\infty} a_k w_k$, $\lim_{j\to\infty} S_{2^{n_j}} = 0$ off E, for some subsequence of integers $(n_j, j \in \mathbb{N})$ and if $|a_k| \le \varepsilon_k$, $k \in \mathbb{N}$, then $a_k = 0$ for all $k \in \mathbb{N}$. A special case of this result had been obtained earlier by Shapiro [1]. For other results on sets of uniqueness for special classes of Walsh series see Crittenden and Shapiro [1], Wade [4], and Yoneda [3], [6], [8].

7.6. Theorem 15 is due to Arutunjan [2]. The idea of using derivates for its proof goes to Skvorcov [3].

Theorem 16 and the proof presented here are Šaginjan's [2].

Theorems 17 and 18 and the proofs presented here are due to Skvorcov [9]. By publishing them he solved a long-standing problem concerning uniqueness of Cesàro summable Walsh series. The problem was first posed by V.L. Shapiro, who suggested it to Crittenden for a dissertation topic. Handicapped by the cumbersome first integral technique (see comments in 7.1), Crittenden [1] obtained a partial solution to the problem, but was unable even to show that uniqueness holds for everywhere Cesàro summable Walsh series. Wade [3] made no additional progress but did solve the problem for Haar series. In so doing, he brought the original problem to the attention of Skvorcov, who solved it in [9].

CHAPTER 8

8.1. The observation that if $\mathbf{a} := (a_k, k \in \mathbb{N})$ decays monotonically to zero then the Walsh series $\sum_{k=0}^{\infty} a_k w_k$ converges uniformly on compact subsets of $[0,1)$ was made first by Šneider [1]. In sharp contrast, he showed that the corresponding Walsh-Kaczmarz series diverges a.e. unless $a_k = o(1/\log k)$ as $k \to \infty$.

Theorem 1 is due to Rubinštein [1], who also examined it for certain multiplicative systems [3]. It is an analogue of trigonometric result due to Ul'janov.

Theorems 2 and 3 appear in Mòricz and Schipp [1]. For the special case that $\mathbf{a}$ is quasi-convex, Theorem 3 is due to Yano [1]. The example which follows Theorem 3 is also Yano's. Fomin (see Balašov and Rubinštein [1], p. 774) has shown that if $\mathbf{a}$ is decreasing, quasi-convex and converges to zero then $\sum_{k=0}^{\infty} a_k w_k$ converges in L^1 if and only if $a_k = o(1/\log k)$ as $k \to \infty$. Mòricz [5], [6] has studied L^p convergence for $0 < p \le 1$ for Walsh-Fourier series with coefficients of generalized bounded variation.

Theorem 4 is due to Balašov [2]. The hypothesis cannot be weakened because there exist non-Fourier Walsh series whose coefficients satisfy $\sum_{k=1}^{\infty} a_k/k = \infty$ and $a_k \downarrow 0$ as $k \to \infty$.

Lemma 1 and Theorems 5 and 6 are due to Coury [3].

8.2. Butzer and Wagner [1], [2] introduced the dyadic derivative and began to study term by term dyadic differentiation of Walsh series. Theorem 7 can be found in Butzer and Wagner [2]. They also showed a Walsh series whose coefficients satisfy

$$\sum_{k=0}^{\infty} k|a_k| < \infty$$

is everywhere term by term dyadically differentiable. They conjectured that Corollary 1 iii) holds, and this conjecture was verified by Schipp [12] a year later.

Theorem 8 can be found in Skvorcov and Wade [1]. The condition on R_n had been used earlier by Coury [3] to obtain sufficient conditions on the Walsh-Fourier coefficients of a given f sufficient to conclude f is constant.

Lemma 2 and Theorem 9 are due to Powell (see Powell and Wade [1]). Theorem 10 and Corollary 3 can be found in Wade [9].

The problem of term by term dyadic differentiation of Rademacher series has been completely solved by Onneweer [7]. He showed that a Rademacher series is dyadically differentiable if and only if its term by term dyadic derivative converges. He also obtained this result for arbitrary Vilenkin systems. Thus a Rademacher series on a Vilenkin group is either differentiable a.e. or almost nowhere. It follows that there exist functions continuous on the group which are nowhere differentiable.

8.3, 8.4. The problem of representing measurable functions by trigonometric series goes all the way back to Lusin. Men'šov [3] finally solved the problem in 1939, showing that given any function f a.e. finite and measurable on $[0, 2\pi)$ there is a trigonometric series which converges a.e. to f. Later, he showed [4] that there exist universal trigonometric series, i.e., series S such that given an a.e. finite measurable f there is a subsequence $(n_k, k \in \mathbf{N})$ of integers such that $S_{n_k} \to f$ a.e. as $k \to \infty$. Men'šov [5] also showed that if f is measurable on $[0, 2\pi)$ (whether infinite on a set of positive measure or not) there is a trigonometric series S which converges in measure to f.

Talaljan has investigated these theorems for general orthonormal systems. In [1] he showed that there exist universal series in any complete orthonormal system. In [2], [3] and [4] he obtained analogues of Men'šov's representation theorems for normed bases of L^p and for normalized convergence systems. In particular, a measurable function f can be represented in measure by Walsh series, and a.e. when f is a.e. finite. A specific construction of this last result for 2^n-th partial sums was given by Arutunjan [1] (see Exercise 8.12). Talaljan and Arutunjan [1] also showed this result cannot be extended to functions which assume an infinite value on a set of positive measure. Indeed, they proved that the 2^n-th partial sums of a Walsh series cannot diverge to $+\infty$ on a set of positive measure. A proof of this result using martingale techniques appears in Gundy [2]. Its trigonometric analogue was proved only recently by Konjagin [1].

The results and proofs of 8.3 and 8.4 are due to Talaljan. An excellent expository article on the problem of representation can be found in Talaljan [5] and in Ul'janov [7].

CHAPTER 9

9.1. The dyadic field was introduced by Fine [2] in 1950. He also constructed the additive characters of the dyadic field (the so-called generalized Walsh functions). Pichler [1] introduced a version of the generalized Walsh functions analogous to the original ordering of the Walsh functions. He showed that on the average, the number of sign changes of a generalized Walsh function is given by its index.

9.2. The Walsh-Fourier transform was introduced by Fine [2]. For other properties of the Walsh-Fourier transform see Crittenden [2]. For generalizations see Selfridge [1] and the book by Taibleson [1].

9.3. The eigenfunctions $(\Omega_k, k \in \mathbf{N})$ of the Walsh-Fourier transform were introduced by Pál [5]. In [6] he used them (as we do) to define the Walsh-Fourier transform of square-integrable functions. A similar program was carried out earlier by Pichler [1]. Of course, the Plancherel theorem (Theorem 7) and the Hausdorff-Young inequality (Theorem 9) hold in much greater generality (see Taibleson [1] or Rudin [1]).

9.4. Corollary 3 is due to Crittenden [2]. Theorem 10 is due to Wagner [1]. The rest of the results in this section are dyadic analogues of well-known trigonometric results.

9.5. The dyadic derivative on $\mathbf{R}^+$ was introduced by Butzer and Wagner [3] and the dyadic integral was defined by Wagner [1] (see also Pál [2]). The strong dyadic integral was defined by Wagner [1]. The theorem for pointwise dyadic differentiability of the dyadic integral is new. A weaker version of it appears in Pál [3], [4].

9.6. The multiplicative group of the dyadic field was investigated by Hagmark [1]. He constructed the characters of $\mathbf{F}^*$ and defined the Mellin transform on the dyadic field. In this connection see also the book by Taibleson [1].

9.7. The Fast Fourier Transform originated in Cooley and Tukey [1]. The Fast Walsh Transform was discovered by Green [1] and generalized by Welch [1]. The idea of using iterated conditional expectations to develop the algorithm is due to Schipp [19]. A good computer program to evaluate FWT can be found in Gonzalez and Wintz [1] (p. 65). The Fast Walsh-Hadamard Transform was introduced by Whelchel and Guinn [1], and Pratt, Kane, and Andrews [1]. A thorough discussion of the sampling simulation of the Walsh- Hadamard Transform which appears here can be found in Decker and Harwit [1].

Several books have been written about applications of Walsh series. For further reading on this subject we suggest Harmuth [1], [2], Beauchamp [1], Maqusi [1], and Bass [1].

Some recent applications of Walsh series do not rely on the Fast Walsh Transform. Examples include genetic algorithms (see Bethke [1]), which optimize non-differentiable functions by letting solutions evolve over several iterations, and ophthalmology (see Optican and Richmond [1], Richmond and Optican [1], and Richmond, Optican, Podell, and Spitzer [1]), where there is some evidence that two dimensional Walsh functions may be used to decipher the neural code which transfers information from our eyes to our brains.

APPENDICES

0.0. For more information about Banach spaces see Banach [1], Riesz and Sz.-Nagy [1], or Rudin [2].

0.1. For more information about orthonormal systems see Alexits [2], Kašin and Saakjan [1], and Olevskiĭ [4].

0.2. For more information on interpolation see Bergh and Löfström [1].

0.3. The general material presented here can be found in Hewitt and Ross [1] or Rudin [1]. Material on the dyadic multiplicative structure of $\mathbf{N}$ can be found in Berlekamp [1].

0.4. A good general reference here is Ash [1].

0.5. The proof presented here comes from Widder [1].

0.6. Theorem 13 is due to Lindahl [1]. Theorem 14 is due to Ward (see Saks [1], p. 141).

0.7. For detailed information on Vilenkin systems see Agaev, Vilenkin, Dzafarli, and Rubinštein [1].

REFERENCES

Note: For the articles in Russian, the titles have been translated into English.

G.N. Agaev, N.Ja. Vilenkin, G.M. Dzafarli, and A.I. Rubinštein. 1. "Multiplicative systems and harmonic analysis on zero-dimensional groups," Baku "ELM", 1981.

G. Alexits. 1. *Une contribution à la théories constructive des fonctions*, Acta Sci. Math. **19** (1958), 149–157.

2. "Convergence problems of orthogonal functions," Pergamon Press, New York, 1961.
3. *Sur les bornes de la théorie de l' approximation des fonctions continues par polynomes*, Publ. Math. Inst. Hung. Acad. Sci. **8** (1963), 329–340.

G. Alexits and D. Kràlik. 1. *Über der Annäherungsgrad den Approximation im starken Sinne von stetigen Funktionen*, Publ. Math. Inst. Hung. Acad. Sci. **8** (1963), 317–327.

2. *Über die Approximation im starken Sinne*, Acta Math. Acad. Sci. Hungar. **16** (1965), 27–32.

F.G. Arutunjan. 1. *On the series in the Haar system*, Dokl. Akad. Nauk Armjan. SSR **17** (1966), 134–139.

2. *On the series in the Haar system*, Dokl. Akad. Nauk Armjan. SSR **42** (1968), 134–140.

F.G. Arutunjan and A.A. Talaljan. 1. *On uniqueness of Haar and Walsh series*, Izv. Akad. Nauk SSSR **28** (1964), 1391–1408.

R.B. Ash. 1. "Real Analysis and Probability," Academic Press, New York, N.Y., 1972.

B. Aubertin. 1. *Algebraic elements and sets of uniqueness in the group integers of a p-series field*, Canad. Math. Bull. **29** (1986), 177–184.

A.V. Bahšečjan. 1. *On some generalized systems of Rademacher and Walsh*, Izv. Akad. Nauk Armjan. SSR **10** (1975), 85–91.

2. *On permutations of the Walsh system*, Dokl. Akad. Nauk Armjan. SSR **69** (1979), 98–102.

A.S. Baiarstanova. 1. *Summation of subsequences of partial sums of Walsh-Fourier series*, Izv. Akad. Nauk Kazan. SSR Fiz.- Mat. **3** (1979), 59–62.

L.A. Balašov. 1. *Series with gaps*, Izv. Akad. Nauk SSSR **29** (1965), 631–644.

2. *Series with respect to the Walsh system with monotone coefficients*, Sibirsk. Mat. Ž. **12** (1971), 25–39.

L.A. Balašov and A.I. Rubinštein. 1. *Series with respect to the Walsh system and their generalizations*, J. Soviet Math. **1** (1973), 727–763.

S. Banach. 1. "Theorie des Operations Linears," Monog. Mat., Warszawa, 1932.

N.K. Bary. 1. "A treatise on trigonometric series," Pergamon Press, London, 1964.

C.A. Bass. 1. "Applications of Walsh Functions," Symposium and Workshop, Naval Research Laboratory Washington, D.C., 1970.

K.G. Beauchamp. 1. "Walsh Functions and Their Applications," Academic Press, New York, 1975.

J. Bergh and J. Löfström. 1. "Interpolation spaces," Springer-Verlag, Berlin, 1976.

E.R. Berlekamp. 1. "Algebraic coding theory," McGraw-Hill Book Company, New York, N.Y., 1968.

A.D. Bethke. 1. "Genetic Algorithms as Function Optimizers," Dissertation, University of Michigan, 1980.

M.G. Bicadze. 1. *Simple and multiple trigonometric series and series with respect to the Walsh system with monotone coefficients*, Soobšč. Akad. Nauk Gruzin. SSR **96** (1979), 545–448.

P. Billard. 1. *Sur la convergence presque partout des séries de Fourier-Walsh des fonctions de l'espace $L^2[0,1]$*, Studia Math. **28** (1967), 363–388.

2. *Bases dans H et bases de sous espaces de dimension finie dans A*, in "Linear Operators and Approximation," Birkhäuser Verlag, Basel, Editors: Butzer, Kahane, and Sz.-Nagy, 1972.

S.L. Bljumin. 1. *Linear methods of summation of Fourier series with respect to multiplicative systems*, Sibirsk. Mat. Ž. **9** (1968), 449–455.

2. *Certain properties of a class of multiplicative systems and problems of approximation of functions by polynomials with respect to those systems*, Izv. Vuzov Mat., No. 4 (1968), 13–22.

S.V. Bočkarev. 1. *The Walsh Fourier coefficients*, Izv. Akad. SSSR, Ser. Mat. **34** (1970), 203–208.

2. *A majorant of the partial sums for a rearranged Walsh system*, Dokl. Akad. Nauk SSSR **239** (1978), 509–510.

3. "A method of averaging in the theory of orthogonal series and some problems in the theory of bases," Trudy Mat. Inst. Steklov., No. 146, 1978.

4. *Existence of bases in the space of analytic functions and some properties of the Franklin system*, Mat. Sbornik **95(137)** (1974), 3–18.

N.A. Bokaev. 1. *On sums of closed U-sets for multiplicative orthonormal systems*, Vestnik Mosk. Univ. Ser. Mat. Meh. (1985), 93–96.

A. Bonami. 1. *A propos d'un theoreme de Paley sur les fonctions de Walsh*, Compt. Rendus Acad. Sci. Paris **266** (1968), A655–A657.

G. Bredon. 1. *A new treatment of the Haar integral*, Mich. Math. J. **10** (1963), 365–373.

D.L. Burkholder. 1. *Martingale transforms*, Ann. Math. Stat. **37** (1966), 1494–1504.

2. *Distribution function inequalities for martingales*, Ann. Prob. **1** (1973), 19–42.

3. *A sharp inequality for martingale transforms*, Ann. Prob. **7** (1979), 858–863.

4. *A sharp inequality for martingales and stochastic integrals*, Colloq. Paul Lévy, Palaiseau, 1987 Astérisque **157-158** (1988), 75–94.

D.L. Burkholder, B.J. Davis, and R.F. Gundy. 1. *Integral inequalities for convex functions of operators on martingales*, in "Proc. Sixth Berkeley Symp. Math. Stat. and Prob.," 1972, pp. 223–240.

E. Buško. 1. *Fonctions continues et fonctions borneés non adhérentes dans $L^\infty(T)$ a la suite de leurs sommes partielles de Fourier*, Studia Math. **34** (1970), 319–337.

P.L. Butzer, E. Engles, and U. Wipperfürth. 1. *An extension of the dyadic calculus with fractional order derivatives: general theory*, Comp. Maths. with Appls. **12 B** (1986), 1073 –1090.

2. *An extension of the dyadic calculus with fractional order derivatives. Further theory and applications*, Comp. Maths. with Appls. **12 A** (1986), 921–943.

P.L. Butzer and R.J. Nessel. 1. "Fourier Analysis and Approximations I," Birkhäuser (Basel) and Academic Press, New York, London, 1971.

P.L. Butzer and K. Scherer. 1. *Jackson and Bernstein type inequalities for families of commutative operators in Banach spaces*, J. Approx. Theory **5** (1972), 308–342.

P.L. Butzer and W. Splettstößer. 1. *Sampling principle for duration-limited signals and dyadic Walsh analysis*, Inform. Sci. **14** (1978), 93–106.

2. *On quantization, truncation, and jitter errors in the sampling theorem and its generalizations*, Signal Processing **2** (1980), 101–112.

P.L. Butzer and H.J. Wagner. 1. *Walsh series and the concept of a derivative*, Applic. Anal. **3** (1973), 29–46.

2. *On dyadic analysis based on the pointwise dyadic derivative*, Analysis Math. **1** (1975), 171–196.

3. *A calculus for Walsh functions defined on* $\mathbf{R}^+$, in "Proc. Symp. of Appl. of Walsh Functions," Washington, D.C., 1973, pp. 75–81.

J.S. Byrnes and D.A. Swick. 1. *Instant Walsh functions*, SIAM Review **12** (1970), p. 131.

L. Carleson. 1. *On convergence and growth of partial sums of Fourier series*, Acta Math. **116** (1966), 135–157.

2. *An explicit unconditional basis in* H^1, Bull. Sci. Math. **104** (1980), 405 -416.

S.-Y. Alice Chang and Z. Ciesielski. 1. *Spline characterizations of* H^1, Studia Math. **75** (1983), 183–192.

J.-A. Chao. 1. *Lusin area functions on local fields*, Pacif. J. Math. **59** (1975), 383–390.

2. *Conjugate characterizations of* H^1 *dyadic martingales*, Math. Annalen **240** (1979), 63–67.
3. H^p *and BMO regular martingales*, in "Harmonic Analysis," Lecture Notes in Mathematics, Vol. 908, Springer-Verlag, Berlin, 1982, pp. 274–284.
4. *Hardy spaces on regular martingales*, in "Martingale Theory in Harmonic Analysis and Banach Spaces," Lecture Notes in Mathematics, Vol. 939, Springer-Verlag, Berlin, 1982, pp. 18–28.

E.W. Cheney. 1. "Introduction to approximation theory," McGraw-Hill, New York, N.Y., 1966.

Z. Ciesielski. 1. *Properties of the orthonormal Franklin system*, Studia Math. **23** (1963), 141–157.

2. *Properties of the orthonormal Franklin system, II*, Studia Math. **27** (1966), 289–323.
3. *Bases and approximation by splines*, in "Proc. Intern. Congress Math.," Vancouver, B.C., 1974, pp. 47–51.
4. *Constructive function theory and spline systems*, Studia Math. **523** (1975), 277–302.
5. *Some properties of the Schauder basis of the space* $C[0,1]$, Bull. Acad. Polon. Sci. **8** (1960), 141–144.

Z. Ciesielski and S. Kwapien. 1. *Some properties of the Haar, Walsh-Paley, Franklin, and the bounded polygonal orthonormal bases in* L^p *spaces*, Comment. Math. **2** (1977), 37–42, Special Issue (dedicated to W. Orlicz on the occasion of his seventy-fifth birthday).

Z. Ciesielski, P. Simon, and P. Sjölin. 1. *Equivalence of Haar and Franklin bases in* L^p *spaces*, Studia Math. **60** (1977), 195–210.

R.R. Coifman. 1. *A real variable characterization of* H^p, Studia Math. **51** (1974), 269–274.

R.R. Coifman and G. Weiss. 1. *Extensions of Hardy spaces and their use in analysis*, Bull. Amer. Math. Soc. **83** (1977), 569–645.

J. Cooley and J. Tukey. 1. *An algorithm for the machine calculation complex Fourier series*, Math. Comp. **19** (1965), 297–301.

J.E. Coury. 1. *A class of Walsh M-sets of measure zero*, J. Math. Anal. Appl. **31** (1970), 318–320.

2. *Some results on lacunary Walsh series*, Pacif. J. Math. **45** (1973), 419– 425.
3. *Walsh series with coefficients tending monotonically to zero*, Pacif. J. Math. **54** (1974), 1–16.
4. *On the differentiation of Rademacher series*, Mich. Math. J. **21** (1974), 321–336.
5. *A class of Walsh M-sets of measure zero (Corrigendum)*, J. Math. Anal. Appl. **52** (1975), 669–671.

R.B. Crittenden. 1. "Theorem on the uniqueness of the (C,1) summability of Walsh series," Doctoral Dissertation, University of Oregon, 1964.

2. *Walsh-Fourier transform*, in "Proc. Symp. and Workshop of Applications of Walsh Functions," Washington, D.C., 1970, pp. 170–174.

R.B. Crittenden and V.L. Shapiro. 1. *Sets of uniqueness on the group* 2^ω, Annals of Math. **81** (1965), 550–564.

Z. Cybertowicz. 1. *On strong (C,r) summability of Walsh- Fourier series*, Funct. Approx. Comment Math. **3** (1970), 37–46.

B.J. Davis. 1. *On the integrability of the martingale square function*, Israel J. Math. **8** (1970), 187–190.

2. *Hardy spaces and rearrangements*, Trans. Amer. Math. Soc. **261** (1980), 211–233.

J. Decker and M. Harwit. 1. *Sequential encoding and multislit spectrometers*, Applied Optics **7** (1968), 2205–2209.

J.L. Doob. 1. "Stochastic Processes," J. Wiley, New York, N.Y., 1953.

A.V. Efimov. 1. *On some approximation properties of periodic multiplicative orthonormal systems*, Mat. Sbornik **69** (1965), 354–370.

P. Enflo. 1. *A counterexample to the approximation property in Banach spaces*, Acta Math. **139** (1973), 309–317.

V. Enchandia. 1. *Interpolation between dyadic Hardy spaces H^p: the real method*, Annales Univ. Sci. Budapest, Sect. Math. **31**, 261–266.

W. Engels. 1. *On the characterization of the dyadic derivative*, Acta Math. Acad. Sci. Hungar. **46** (1985), 47–56.

G. Faber. 1. *Über die Orthogonal Functionen des Herrn Haar*, Jahresber. Deutsch. Math. Verein. **19** (1910), 104–117.

C. Fefferman. 1. *Characterization of bounded mean oscillation*, Bull. Amer. Math. Soc. **77** (1971), 587–588.
2. *Pointwise convergence of Fourier series*, Annals of Math. **98** (1973), 551–571.

C. Fefferman, N.M. Rivier, and Y. Sagher. 1. *Interpolation between H^p spaces: The real method*, Trans. Amer. Math. Soc. **191** (1974), 75–81.

C. Fefferman and E.M. Stein. 1. *H^p spaces of several variables*, Acta Math. **129** (1972), 137–194.

N.J. Fine. 1. *On the Walsh functions*, Trans. Amer. Math. Soc. **65** (1949), 372–414.
2. *The generalized Walsh functions*, Trans. Amer. Math. Soc. **69** (1950), 66–77.
3. *Cesàro summability of Walsh-Fourier series*, Proc. Nat. Acad. Sci. USA **41** (1955), 588–591.
4. *Fourier-Stieltjes series of Walsh functions*, Trans. Amer. Math. Soc. **86** (1957), 246 –255.

G.A.O. Fomin. 1. *On linear methods of summation of Fourier series*, Mat. Sbornik **65** (1964), 144–152.

P. Franklin. 1. *A set of continuous orthogonal functions*, Math. Annalen **100** (1928), 522– 529.

S. Fridli. 1. *Approximation by Vilenkin-Fourier sums*, Acta Math. Acad. Sci. Hungar. **47** (1986), 33–44.
2. *Approximation by Vilenkin polynomials*, Annales Univ. Sci. Budapest Eötvös, Sect. Math. **28** (1986), 133–141.
3. *Bernstein type inequalities with respect to the Vilenkin systems*, in "Constructive Theory of Functions '84," Proc of Conf. in Sofia, 1984, pp. 352–355.

S. Fridli and F. Schipp. 1. *Tree-Martingales*, in "Proc. 5-th Pannonian Symp. Math. Stat.," (Visegràd) (1985), pp. 53–63.
2. *On the everywhere divergence of Vilenkin-Fourier series*, Acta Sci. Math. (Szeged) **48** (1985), 155–162.

S. Fridli and P. Simon. 1. *On the Dirichlet kernels and a Hardy space with respect to the Vilenkin system*, Acta Math. Acad. Sci. Hungar. **45** (1985), 223–234.

N. Fujii. 1. *A maximal inequality for H^1 functions on the generalized Walsh-Paley group*, Proc. Amer. Math. Soc. **77** (1979), 111–116.

S.S. Galstian. 1. *On everywhere divergent trigonometric Fourier series with quickly decreasing coefficients*, Mat. Sbornik **122(164)** (1983), 157–167.

J. B. Garnett and P.W. Jones. 1. *BMO and dyadic BMO*, Pacif. J. Math. **99** (1982), 351–371.

A.M. Garsia. 1. "Martingale inequalities," W.A. Benjamin, Inc., Reading, Mass., 1973.
2. *On a convex function inequality for martingales*, Ann. Prob. **1** (1973), 171–174.

G.G. Gevorkjan. 1. *Uniqueness sets for Haar and Walsh systems*, Dokl. Akad. Nauk Armjan. SSR **73** (1981), 91–96.
2. *On the unboundedness of the shift operator for the Franklin system in L^1*, Mat. Zametki **38** (1985), 523–533.

J.E. Gibbs and M.S. Millard. 1. *Walsh functions as solutions of a logical differential equation*, National Physical Lab. (1969), Middlesex, England, DES Report No. 1.

B. Golubov. 1. *Approximation of functions in the L^p metric by Haar and Walsh polynomials*, Mat. Sbornik **87** (1972), 254–274.

R.C. Gonzalez and P. Wintz. 1. "Digital Image Processing," Addison- Wesley Publ. Co., Reading, Mass., 1977.

J. Gosselin. 1. *Almost everywhere convergence of Vilenkin- Fourier series*, Trans. Amer. Math. Soc. **185** (1973), 345–370.
2. *A weighted norm inequality for Vilenkin-Fourier series*, Proc. Amer. Math. Soc. **49** (1975), 349–353.
3. *On the convergence of Walsh-Fourier series for $L^2[0,1]$*, in "Studies in Analysis, Advances in Math. (Suppl. Studies 4)," Academic Press, New York, N.Y., 1979, pp. 223–232.
4. *On the Littlewood-Paley theory for mixed norm spaces*, Trans. Amer. Math. Soc. **256** (1979), 113–124.

J. Gosselin and Wo-Sang Young. 1. *On rearrangements of Vilenkin-Fourier series which preserve a.e. convergence*, Trans. Amer. Math. Soc. **209** (1975), 157–174.

R.R. Green. 1. *A serial orthogonal decoder*, Space Programs Summary, Jet Propulsion Laboratory, Pasadena, Calif. **4** (1966), 247–251.

D.J. Grubb. 1. *Sets of uniqueness in compact, 0- dimensional metric groups*, Trans. Amer. Math. Soc. **301** (1987), 239–249.

N.V. Guličev. 1. *Correction of functions up to absolutely convergent Fourier series*, Vestnik Mosk. Univ. Ser. Mat. Meh. (1979), 12–15.
2. *Approximation to continuous functions by Walsh-Fourier series*, Analysis Math. **6** (1980), 269–280.

R.F. Gundy. 1. *Martingale theory and pointwise convergence of certain orthogonal series*, Trans. Amer. Math. Soc. **124** (1966), 228–248.
2. *On a class of martingale series*, in "Orthogonal Expansions and Their Continuous Analogues," Proc. Conf., Edwardsdale, Ill., 1967, pp. 99–102.
3. *A decomposition for L^1-bounded martingales*, Ann. Math. Stat. **39** (1968), 134–138.
4. *On the class $L \log L$ martingales and singular integrals*, Studia Math. **23** (1969), 109–118.

R.F. Gundy and R.L. Wheeden. 1. *Weighted integral inequalities for the non- tangential maximal function, Lusin area integral, and Walsh-Paley series*, Studia Math. **49** (1974), 107 –124.

A. Haar. 1. *Zur Theorie der orthogonalen Funktionen systeme*, Mat. Ann. **69** (1910), 331–371.

P.-E. Hagmark. 1. "On the multiplicative group of the dyadic field," Report No. A178, Helsinki Univ. of Technology, 1981.

H.F. Harmuth. 1. "Transmission of information by orthogonal functions," Springer-Verlag, Berlin, 1972.
2. "Sequency theory," Academic Press, New York, N.Y., 1977.

L. Harper. 1. *Capacities of sets and harmonic analysis on the group 2^ω*, Trans. Amer. Math. Soc. **126** (1967), 303–315.

D. Harris. 1. *Compact sets of divergence for continuous functions on a Vilenkin group*, Proc. Amer. Math. Soc. **98** (1986), 436–440.
2. *A.e. divergence of multiple Walsh-Fourier series*, Proc. Amer. Math. Soc. **101** (1987), 637–643.
3. *Sets of uniqueness and closed subgroups*, Analysis Math. (to appear).

D. Harris and W.R. Wade. 1. *Sets of divergence on the group 2^ω*, Trans. Amer. Math. Soc. **240** (1978), 385–392.

F.I. Haršiladze. 1. *On Walsh-Fourier series*, Tr. Gruz. Polytehn. Inst. (1962), 43–53.

Š.V. Heladze. 1. *The convergence of Fourier series almost everywhere and in L^1*, Dokl. Akad. Nauk SSSR **232** (1977), 535–537.
2. *Divergence everywhere of Fourier series in bounded Vilenkin systems*, Sakharth. SSR Mecn. Akad. Math. Inst. Šrom **58** (1978), 224–242.
3. *Divergence everywhere of Walsh-Fourier series*, Soobšč. Akad. Nauk Gruz. SSR **77** (1975), 305– 307.

C. Herz. 1. *Bounded mean oscillation and regulated martingales*, Trans. Amer. Math. Soc. **193** (1974), 199–215.

2. *H_p-spaces of martingales, $0 < p \leq 1$*, Z. Wahrsch. Verw. Geb. **28** (1974), 189–205.

E. Hewitt and K. Ross. 1. "Abstract Harmonic Analysis I," Springer-Verlag, Heidelberg, 1963.

I.I. Hirschman. 1. *The decomposition of Walsh and Fourier series*, Mem. Amer. Math. Soc. **65** (1955).

N.P. Horoško. 1. *An estimate of the upper bound of Fourier coefficients on classes of continuous functions in certain orthonormal systems*, Mat. Mežvuzov Konf. Molod. Učen.-Mat., Har'kov (1966), 116–124.

2. *An estimate of the upper bound of Fourier coefficients on certain classes of functions in the systems of Haar, Rademacher and Walsh*, Teor. Funkciĭ Funkcional Anal. i Prilozen Vyp. **15** (1972), 3–12.

R.A. Hunt. 1. *On the convergence of Fourier series*, in "Orthogonal expansions and their continuous analogues (Proc. Conf. Edwardsville, Ill., 1967)," S. Illinois Univ. Press, Carbondale, Ill., 1968, pp. 235- 255.

2. *Almost everywhere convergence of Walsh-Fourier series of L^2 functions*, Actes Congres. Intern. Math. **2** (1970), 655–661.

R.A. Hunt and Wo-Sang Young. 1. *A weighted norm inequality for Fourier series*, Bull. Amer. Math. Soc. **80**, 274–277.

M.A. Jastrebova. 1. *On approximation of functions, satisfying the Lipschitz condition, by arithmetic means of their Walsh-Fourier series*, Mat. Sbornik **71** (1966), 214–226.

F. John and L. Nirenberg. 1. *On functions of bounded mean oscillation*, Comm. Pure Appl. Math. **14** (1961), 415–426.

S. Kaczmarz. 1. *Über ein Orthogonal System*, Comp. Rend. Congres Math. (Warsaw, 1929).

S. Kaczmarz and G. Steinhaus. 1. "Theorie der Orthogonalreihen," Monogr. Mat. Vol. 6, Warsaw, 1935.

B.S. Kašin and A.A. Saakjan. 1. "Orthogonal series," Izdat. "Nauka", Moscow, 1984.

Y. Katznelson. 1. "Harmonic analysis," Wiley and Sons, New York, N.Y., 1968.

2. *On a theorem of Menchoff*, Proc. Amer. Soc. **53** (1975), 396–398.

M. Kobayashi. 1. *Capacities of sets and harmonic analysis on the group 2^ω*, Tôhoku Math. J. **21** (1969), 419–433.

V.M. Kokilašvili. 1. *Approximation of functions by averages of Walsh-Fourier series*, Soobšč. Akad. Nauk Gruz. SSR **45** (1967), 305–310.

A.M. Kolmogorov. 1. *Une contribution à l'étude de la convergence des series de Fourier*, Fund. Math. **5** (1924), 96–97.

2. *Une série de Fourier-Lebesgue divergente partout*, Comp. Rend. **183** (1926), 1327–1328.

S.V. Konjagin. 1. *Limits of indeterminancy of trigonometric series*, Math. Notes **44** (1988), 910–919.

A.N. Kõrner. 1. *Sets of divergence for Fourier series*, Bull. Lond. Math. Soc. **3** (1971), 152- -154.

B.D. Kotljar. 1. *Walsh series and the theorem of Men'šov on adjustment of functions*, Izv. Akad. Nauk SSSR **30** (1966), 1193–1200.

N.R. Ladhawala. 1. *Absolute summability of Walsh-Fourier series*, Pacific J. Math. **65** (1976), 103–108.

N.R. Ladhawala and D.C. Pankratz. 1. *Almost everywhere convergence of Walsh- Fourier series of H^1 functions*, Studia Math. **59** (1976), 85–92.

L. Leindler. 1. *Über die Approximation im starken Sinne*, Acta Math. Acad. Sci. Hungar. **16** (1965), 255–262.

S.V. Levizov. 1. *Some properties of the Walsh system*, Mat. Zametki **27** (1980), 715–720.

R. Liedl. 1. *Über gewisse Functionale in Raum $C^{(v)}[0,1]$ und Walsh-Fourierkoeffizienten*, Monat. Math. **72** (1968), 38–44.

J. Lindenstrauss and L. Tzarfriri. 1. "Classical Banach spaces I, II," Springer-Verlag, New York, N.Y., 1973, 1979.

Li Luan Tan. 1. *The system of orthogonal functions of Walsh*, Acta Sci. Mat. Univ. Pekin. **3** (1957), 403-422.

R. Lindahl. 1. *A differentiation theorem for functions defined on the dyadic rationals*, Proc. Amer. Math. Soc. **30** (1971), 349–352.

R.S. Lipčer and A.N. Siriaev. 1. "Statistic of stochastic processes," 3rd ed. Nauka, 1974.

G.E. Lippman and W.R. Wade. 1. *Pseudofunctions and uniqueness on the group of integers of a p-series field*, Acta Math. Acad. Sci. Hungar. **35** (1980), 1–12.

L.A. Ljusternik and W.I. Sobolev. 1. "Elemente der Funktionalanalysis," Akademie-Verlag, Berlin, 1955.

S. Ju. Lukašenko. 1. *The structure of sets of divergence of trigonometric and Walsh series*, Dokl. Akad. Nauk SSSR **253** (1980), 528–530.
2. *On the structure of divergence sets for Walsh-Fourier series*, Anal. Math. **10** (1984), 23–41.

Lu Shan-Zhen. 1. *A note on the almost everywhere convergence of Bochner-Riesz means of multiple conjugate Fourier series*, in "Lecture Notes," Vol. 908, Springer-Verlag, 1982, pp. 311–318.

Lu Shan-Zhen, M.H Taibleson, and G. Weiss. 1. *On the almost everywhere convergence of Bochner-Riesz means of multiple Fourier series*, in "Lecture Notes," Vol. 908, Springer-Verlag, 1982, pp. 319–325.

M. Maqusi. 1. "Applied Walsh Analysis," Heyden, London, 1981.

J. Marcinkiewicz. 1. *Quelques théorèmes sur les séries orthogonales*, Ann. Soc. Polon. Math. **16** (1937), 85–96.

J. Marcinkiewicz and A. Zygmund. 1. *Quelques théorèmes sure les fonctions independentes*, Studia Math. **7** (1938), 104–120.

V.A. Matveev. 1. *On series with respect to the Schauder system*, Mat. Zametki **2** (1967), 267–278.

B. Maurey. 1. *Isomorphisms entre espaces H_1*, Acta Math. **145** (1980), 79–120.

J.R. McLaughlin. 1. *Absolute convergence of series of Fourier coefficients*, Trans. Amer. Math. Soc. **184** (1973), 291–316.

D.E. Men'šov. 1. *Properties of Cesàro means of negative order and some other T-means for Fourier series of continuous functions*, Mat. Sbornik **86** (1971), 419–445.
2. *An application of the method of Cesàro summation of negative order to Fourier-Lebesgue series*, Dokl. Akad. Nauk SSSR **210** (1973), 271– 273.
3. *Sur la representation des fonctions measurables par des séries trigonometriques*, Mat. Sbornik **9** (1941), 667–692.
4. *On limits of sequences of partial sums of trigonometric series*, Dokl. Akad. Nauk **106** (1956), 777–780.
5. *Convergence in measure of trigonometric series*, Trudy Mat. Inst. Akad. Nauk SSSR **32** (1950), 1–99.

Y. Meyer. 1. "Algebraic Numbers and Harmonic Analysis," North-Holland, Amsterdam, 1972.

J. Mogyoròdi. 1. *Maximal inequalities, convexity inequality and their duality I, II*, Analysis Math. **7** (1981), 131-140, 185–197.

K.H. Moon. 1. *An everywhere divergent Fourier-Walsh series of the class $L(\log^+ \log^+ L)^{1-\epsilon}$*, Proc. Amer. Math. Soc. **50** (1975), 309–314.

C. Moraga. 1. *On some applications of the Chrestenson functions in logic design and data processing*, Mathematics and Computers in Simulation **27** (1985), 431–439.
2. *Spectral characterization of co-symmetries*, in "Theory and Applications of Spectral Techniques," Dortmund, 1988, pp. 127–139.

G.W. Morgenthaler. 1. *Walsh-Fourier series*, Trans. Amer. Math. Soc. **84** (1957), 472–507.

F. Mòricz. 1. *On the order of magnitude of the partial sums of rearrangements of Fourier series of square integrable functions*, Acta Sci. Math. **28** (1967), 155–167.
2. *On the convergence properties of weakly multiplicative systems*, Acta Sci. Math. **38** (1967), 127–144.
3. *On the convergence of double orthogonal series*, Analysis Math. **2** (1976), 287–304.
4. *Lebesgue functions and multiple function series, I*, Acta Math. Acad. Sci. Hungar. **37** (1981), 481–496.
5. *On L^1- convergence of Walsh-Fourier series*, Rend. Circ. Mat. Palermo **38** (1989), 411–418.
6. *Walsh-Fourier series with coefficients of generalized bounded variation*, J. Austral. Math. Soc. **47** (1989), 458–465.

F. Mòricz and F. Schipp. 1. *On the integrability and L^1 convergence of Walsh series*, J. Math. Anal. Appl. **146** (1990), 99–109.

H. Movsisjan. 1. *Uniqueness of double series in the Haar and Walsh systems*, Izv. Akad. Armjan. SSR Ser. Mat. **9** (1974), 40–61.

C.J. Mozzochi. 1. "On the pointwise convergence of Fourier series," Lecture Notes, Vol. 199, Springer-Verlag, New York, N.Y., 1971.

S. Nakata. 1. *On the unconditional convergence of Walsh series*, Analysis Math. **5** (1979).

I. Neveu. 1. "Discrete-Parameter martingales," North Holland Publ., Amsterdam, 1975.

E.M. Nikišin. 1. *Resonance theorems and linear operators*, Uspehi Mat. Nauk **25** (1970), 129– 191.

T. Ohkuma. 1. *On a certain system of orthogonal step functions*, Tôhoku Math. J. **5** (1953), 166–177.

A.M. Olevskiĭ. 1. *Orthogonal series with respect to complete systems*, Mat. Sbornik **58** (1962), 707–748.
2. *Order growth of the Lebesgue functions of bounded orthonormal systems*, Dokl. Akad. Nauk SSSR **176** (1967), 1247–1250.
3. *On singularities of the Carleman type*, Sibirsk. Mat. Ž. **8** (1967), 807–826.
4. "Fourier series with respect to general orthonormal systems," Springer-Verlag, Berlin, 1975.
5. *The existence of functions with unremovable Carleman singularities*, Dokl. Akad. Nauk SSSR **238** (1978).

C.W. Onneweer. 1. *On uniform convergence of Walsh-Fourier series*, Pacif. J. Math. **34** (1970), 117–122.
2. *On moduli of continuity and divergence of Fourier series on groups*, Proc. Amer. Math. Soc. **29** (1971), 109–112.
3. *Uniform convergence of Fourier series on groups, II*, Rocky Mt. J. Math. **1** (1971), 623–631.
4. "Adjustment on small sets of functions on certain groups," Technical Report # 225 (October, 1971), University of New Mexico.
5. *Absolute convergence of Fourier series*, Duke Math. J. **39** (1974), 599–609.
6. *Absolute convergence of Fourier series on certain groups, II*, Duke Math. J. **41** (1974), 679–688.
7. *Differentiability for Rademacher series on groups*, Acta Sci. Math. (Szeged) **39** (1977), 121–128.
8. "Differentiation of fractional order on p-groups," Technical Report # 330 (May, 1977), University of New Mexico.
9. *Differentiation on a p-adic or p-series fields*, in "Linear Spaces and Approximation (Proc. Conf. Oberwolfach, 1977)," pp. 187–198.
10. *Sets of divergence in the group of integers of a p-adic or p-series fields*, Annales Univ. Sci. Budapest Eötvös, Sect. Math. **21** (1978), 13–20.
11. *On L^1 convergence of Walsh-Fourier series*, Internat. J. Math. Sci. **1** (1978), 47–56.
12. *On the definition of dyadic differentiation*, Appl. Anal. **9** (1979), 267–278.

C.W. Onneweer and D. Waterman. 1. *Uniform convergence of Fourier series on groups, I*, Mich. Math. J. **18** (1971), 265–273.

2. *Fourier series of functions of harmonic bounded fluctuation*, J. d'Anal. Math. **27** (1974), 79–93.

L.M. Optican and B.J. Richmond. 1. *Temporal encoding of two dimensional patterns by single units in primate inferior temporal cortex, III*, J. Neurophysiology **57** (1987), 162–178.

W. Orlicz. 1. *Über eine gewisse Klasse von Räumen von Typus B.*, Bull. Internat. Acad. Polon. Sci., Ser. A **3** (1932), 207–220.

K.I. Oskolkov. 1. *Sequences of norms of Fourier sums of bounded functions*, Trudy Mat. Inst. Akad. Nauk SSSR **143** (1977), 129–142.

J. Pál. 1. *On the connection between the concept of a derivative defined on the dyadic field and the Walsh-Fourier transform*, Annales Univ. Sci. Budapest Eötvös, Sect. Math. **18** (1975), 49–54.

2. *On the concept of a derivative among functions on the dyadic field*, SIAM J. Math. Anal. **8** (1977), 375–391.
3. *On almost everywhere differentiability of dyadic integral functions on R^+*, Coll. Math. Soc. J. Bolyai, Fourier Analysis and Approximation Theory, Vol. II, Budapest, 1975 **19** (1978), 591–601.
4. *The almost everywhere differentiability of Wagner's dyadic integral function on R^+*, Coll. Math. Soc. J. Bolyai, A. Haar Memorial Conference, Vol. II, Budapest, 1985 **49** (1987), 695–701.
5. *The eigenfunctions of the Walsh-Fourier transform*, Proc. Int. Conf. Approximation and Function Spaces, Gdansk, 1979 (1981), 553–557.
6. *On the Walsh-Fourier transform of functions in $L^2(R^+)$*, Proc. Int. Conf. on Constructive Theory of Functions, Varna, 1981 (1983), 475–478.

J. Pál and P. Simon. 1. *On a generalization of the concept of derivative*, Acta Math. Acad. Sci. Hungar. **29** (1977), 155–164.

2. *On the generalized Butzer-Wagner type a.e. differentiability of integral function*, Annales Univ. Sci. Budapest Eötvös, Sect. Math. **20** (1977), 156–165.

L. Pál. 1. *On the role of stuctural properties of functions in the theory of their Walsh-Fourier expansions*, Annales Univ. Sci. Budapest Eötvös, Sect. Math. **10** (1967), 155–166.

L. Pál and F. Schipp. 1. *On Haar and Schauder series*, Acta Sci. Math. **31** (1970), 53–58.

R.E.A.C. Paley. 1. *A remarkable system of orthogonal functions*, Proc. Lond. Math. Soc. **34** (1932), 241–279.

A. Pelczynckі. 1. *Banach spaces of analytic functions and absolute summing operators*, in "Conf. of Math. Sci. Regional Conference series in Math.," Amer. Math. Soc., Providence, R.I., 1977.

R. Penney. 1. *On the rate of growth of the Walsh antidifferentiation operator*, Proc. Amer. Math. Soc. **55** (1976), 57–61.

F. Pichler. 1. "Das System der sal- und cal-Funktionen als Erweiterung des Systems der Walsh-Funktionen und die Theorie der sal- und cal-Fouriertransformation," Dissertation, University of Innsbruck, Innsbruck, Austria, 1967.

2. *Walsh-Fourier Synthese optimaler Filter*, Arch. Elektr. Übertr. **24** (1970), 350–360.
3. *Walsh functions, signals and systems*, in "Proc. Symp. Applic. Walsh Functions," Washington, D.C., 1974, pp. 177–188.
4. *Some historical remarks on the theory of Walsh functions and their applications in information engineering*, in "Theory and Applications of Gibbs Derivatives," Institute of Mathematics, Belgrade, 1989.

L.S. Pontryagin. 1. "Topological groups," Princeton Univ. Press, Princeton, N.J., 1946.

C. Powell and W.R. Wade. 1. *Term by term dyadic differentiation*, Canad. J. Math. **33** (1981), 247–256.

W.K. Pratt, J. Kane, and H.C. Andrews. 1. *Hadamard transform image coding*, Proc. IEEE **57** (1969), 58–68.

J.J. Price. 1. *Orthonormal sets with non-negative Dirichlet kernels, I*, Trans. Amer. Math. Soc. **95** (1960), 256–262.

2. *Orthonormal sets with non-negative Dirichlet kernels, II*, Trans. Amer. Math. Soc. **100** (1961), 153–161.
3. *A density theorem for Walsh functions*, Proc. Amer. Math. Soc. **18** (1967), 209–211.
4. *Walsh series and adjustment of functions on small sets*, Ill. J. Math. **13** (1969), 131–136.
5. *Sparse subsets of orthonormal systems*, Proc. Amer. Math. Soc. **35** (1972), 161–164.

V. Prohorenko. 1. *Divergent series with respect to the system of Haar*, Izv. Vysš Učebn. Zaved. Mat. No. **1(104)** (1971), 62–68.

T.S. Quek and L.Y.H. Yap. 1. *Absolute convergence of Vilenkin-Fourier series*, J. Math. Anal. Appl. **74** (1980), 1–14.

H.A. Rademacher. 1. *Einige Sätze über Reihen von allgemeinen Orthogonalenfunktionen*, Math. Annalen **87** (1922), 112–138.

R.R. Riad. 1. *On the conjugate of the Walsh series*, Annales Univ. Sci. Budapest Eötvös, Sect. Math. **30** (1987), 69–76.

B.J. Richmond and L.M. Optican. 1. *Temporal encoding of two dimensional patterns by single units in primate inferior temporal cortex, II*, J. Neurophysiology **57** (1987), 147–161.

B.J. Richmond, L.M. Optican, M. Podell, and H. Spitzer. 1. *Temporal encoding of two dimensional patterns by single units in primate inferior temporal cortex, I*, J. Neurophysiology **57** (1987), 132–146.

R. Riesz and B. Sz.-Nagy. 1. "Functional analysis," Frederick Ungar Publish. Co., New York, N.Y., 1955.

B. Roider. 1. "Über Stufenfunktionen Endlicher Vielfalt," Dissertation, University of Innsbruck, Innsbruck, Austria, 1969.

S. Ropela. 1. *Properties of bounded orthonormal spline bases*, Banach Cent. Publications, Approximation Theory **3** (1975).

A.I. Rubinštein. 1. *The A integral and series with respect to the Walsh system*, Usp. Mat. Nauk **18** (1963), 191–197.

2. *On series with gaps*, Izv. Vuzov Mat. **3** (1963), 137–148.
3. *Certain properties of a class of multiplicative periodic complete orthonormal systems*, Mat. Zametki **2** (1967), 279–288.

W. Rudin. 1. "Fourier Analysis on Groups," Interscience Pub., Wiley and Sons, New York-London, 1967.

2. "Functional Analysis," McGraw Hill Book Company, New York, N.Y., 1973.

L.A. Šaginjan. 1. *Summability of series in the Haar system by the $(C,1)$ method*, Mat. Zametki **15** (1974), 394–404.

2. *On indefinite limits and sets of limit functions of Walsh series*, Mat. Sbornik (N.S.) **95** (1974), 263–271.
3. *Multiple Walsh-Fourier series*, Soobšč. Akad. Nauk Gruz. SSR **81** (1976), 565–567.
4. *Summability to infinity of trigonometric series and series in the Walsh system*, Mat. Sbornik (N.S.) (150) **108** (1979), 457–470.
5. *On a theorem of Marcinkiewicz-Zygmund*, Izv. Akad. Nauk Armjan. SSR **16** (1981), 54–64.

S. Saks. 1. "Theory of the integral," Dover Publications, New York, N.Y., 1964.

D.K. Sanadze and Š.V. Heladze. 1. *Convergence and divergence of multiple Walsh- Fourier series*, Sakharth. SSR Mech. Akad. Math. Inst. Šrom. **55** (1977), 93–106.

T.V. Šarašenidze. 1. *Multiple Walsh-Fourier series*, Soobšč. Akad. Gruz. SSR **81** (1976), 565–567.

J. Schauder. 1. *Zur Theorie stetiger Abbildungen in Funktionalräumen*, Math. Z. **26** (1927), 47–65.

2. *Eine Eigenschaft des Haarschen Orthogonalsystems*, Math. Z. **28** (1928), 317–320.

F. Schipp. 1. *Bemerkung zur Divergenz der Walsh-Fourierreihen*, Annales Univ. Sci. Budapest Eötvös, Sect. Math. **11** (1968), 53–58.
2. *Über die starke Summation von Walsh- Fourierreihen*, Acta Sci. Math. (Szeged) **30** (1969), 77–87.
3. *On the strong approximation of Walsh-Fourier series*, MTA III. Oszt. Közl. (in Hungarian) **19** (1969), 101–111.
4. *Über Walsh-Fourierreihen mit nichtnegativen Partialsummen*, Annales Univ. Sci. Budapest Eötvös, Sect. Math. **12** (1969), 43–48.
5. *Über die Divergenz der Walsh-Fourierreihen*, Annales Univ. Sci. Budapest Eötvös, Sect. Math. **12** (1969), 49–62.
6. *Über die Konvergenz von Reihen nach Produktsystemen*, Acta Sci. Math. (Szeged) **35** (1973), 13–16.
7. *Über einen Ableitungsbegriff von P.L. Butzer und H.J. Wagner*, Math. Balkanica **4** (1974), 541–546.
8. *On a.e. convergence of expansion with respect to a bounded orthonormal system of polygonals*, Studia Math. **58** (1976), 287–290.
9. *Certain rearrangements of series in the Walsh series*, Mat. Zametki **18** (1975), 193–201.
10. *On term by term dyadic differentiability of Walsh series*, Analysis Math. **2** (1976), 149–154.
11. *On the dyadic derivative*, Acta Math. Acad. Sci. Hungar. **28** (1976), 145–152.
12. *Über gewiessen Maximaloperatoren*, Annales Univ. Sci. Budapest Eötvös, Sect. Math. **18** (1975), 189–195.
13. *On L^p-norm convergence of series with respect to product systems*, Analysis Math. **2** (1976), 49–64.
14. *Pointwise convergence of expansions with respect to certain product systems*, Analysis Math. **2** (1976), 63–75.
15. *Fourier series and martingale transforms*, in "Linear spaces and approximation," (Ed. P.L. Butzer and B.Sz.-Nagy) ISNM 40 Birkhäuser Verlag, Basel, 1978.
16. *The dual space of martingale VMO spaces*, Proc. 3rd Pannonian Symp. Math. Stat. (Visegràd) (1982), 305–311.
17. *On a Paley-type inequality*, Acta Sci. Math. (Szeged) (1983), 357–364.
18. *VMO spaces not having Schauder basis*, Analysis Math. **9** (1983), 313–322.
19. *Fast Fourier transform and conditional expectation*, Coll. Math. Soc. J. Bolyai, Numerical Methods, Keszthely, 1977, **22** (1980), 565–576.

F. Schipp and P. Simon. 1. *On some (L, H)-type maximal inequalities with respect to the Walsh-Paley system*, Proc. Inter. Conf. Function Series and Operators (1980), 22–29.
2. *Investigation of Haar and Franklin series in Hardy spaces*, Analysis Math. **8** (1982), 47–56.
3. *On a theorem of Kolmogorov*, Mat. Lapok (in Hungarian) **31** (1983), 117–123.

R.S. Selfridge. 1. *Generalized Walsh functions*, Pacific Math. J. **5** (1985), 451–480.

E.M. Semjonov. 1. *On the equivalence of rearrangements of the Haar system in L_p*, Dokl. Akad. Nauk SSSR **242** (1978), 1258–1260.

E.M. Semjonov and B. Stöckert. 1. *The rearrangements of the Haar system in the spaces L_p*, Analysis Math. **7** (1981), 227–295.

V.L. Shapiro. 1. *$U(\epsilon)$-sets for Walsh series*, Proc. Amer. Math. Soc. **16** (1965), 867–870.

A.H. Siddiqi. 1. *On the Walsh-Fourier coefficients of certain classes of functions*, Riv. Mat. Univ. Parma (2) **10** (1969), 157–163.
2. *On lacunary partial sums of Walsh-Fourier series*, Rend. Circ. Mat. Palermo (2) **18** (1969), 313–318.
3. *Walsh-Fourier series with monotone coefficients*, Aligarh Bull. Math. **3/4** (1973/74), 25–29.
4. "Walsh functions," Aligarh Muslim University Press, Aligarh, INDIA, 1977.

A.M. Sidorov. 1. *Remarks on double Walsh-Fourier series*, Sbornik Asp. Rabot. Kazan. Univ. Kazan (1973), 65 –71.

P. Simon. 1. *Verallgemeinerte Walsh-Fourierreihen I*, Annales Univ. Sci. Budapest Eötvös, Sect. Math. **16** (1973), 103–113.
2. *Verallgemeinerte Walsh-Fourierreihen II*, Acta Math. Acad. Sci. Hungar. **27** (1976), 329–341.

3. *On a maximal function*, Annales Univ. Sci. Budapest Eötvös, Sect. Math. **21** (1978), 41– 44.
4. *On the concept of a conjugate function*, Colloq. Math. Soc. J. Bolyai, Fourier Analysis and Approximation Theory, 1975 **19** (1978), 747–755.
5. *On convergence of Vilenkin-Fourier series*, Acta Math. Acad. Sci. Hungar. **33** (1979), 189–196.
6. *Remarks on the shift operators with respect to the Haar and Franklin systems*, Acta Math. Acad. Sci. Hungar. **39** (1982), 251–254.
7. *On the Parseval equality and the Dini-Lipschitz condition with respect to Vilenkin systems*, Analysis Math. **10** (1984), 151–161.
8. *(L^1, H)-type estimations for some operators with respect to the Walsh-Paley system*, Acta Math. Acad. Sci. Hungar. **46** (1985), 307–310.
9. *Strong convergence of certain means with respect to the Walsh-Fourier series*, Acta Math. Acad. Sci. Hungar. **49** (1987), 425–431.
10. *On the divergence of Vilenkin-Fourier series*, Acta Math. Acad. Sci. Hungar. **41** (1983), 359–370.

P. Sjölin. 1. *An inequality of Paley and convergence a.e. of Walsh-Fourier series*, Arkiv för Math. **8** (1969), 551–570.
2. *The Haar and Franklin systems are not equivalent in L^1*, Bull. Acad. Polon. Sci. **25** (1977), 1099–1100.

V.A. Skvorcov. 1. *A theorem of Cantor type for the Haar system*, Vestnik Mosk. Univ. Ser. Mat. Meh. **5** (1964), 3–6.
2. *Haar series with convergent subsequences of partial sums*, Dokl. Akad. Nauk SSSR **183** (1968), 784–786.
3. *Differentiation with respect to nets and the Haar series*, Mat. Zametki **4** (1968), 33–40.
4. *Some generalizations of uniqueness theorem for Walsh series*, Mat. Zametki **13** (1973), 367–372.
5. *The coefficients of multiple Haar and Walsh series*, Vestnik Mosk. Univ. Ser. Mat. Meh. **28** (1973), 77–79.
6. *The uniqueness of Walsh series with convergent subsequences of partial sums*, Mat. Zametki **16** (1974), 27–32.
7. *An example of a Walsh series with a subsequence of partial sums that converges to zero everywhere*, Mat. Sbornik (N.S.) **97** (1975), 517–539.
8. *An example of a zero-series in the Walsh system*, Mat. Zametki **19** (1976), 179–186.
9. *Uniqueness theorems for Walsh series that are summable by the $(C, 1)$ method*, Vestnik Mosk. Univ. Ser. Mat. Meh. **31** (1976), 73–80.
10. *Uniqueness conditions for the representation of functions by Walsh series*, Mat. Zametki **21** (1977), 187–197.
11. *The h-measure of M-sets for the Walsh systems*, Mat. Zametki **21** (1977), 335–340.
12. *The rate of convergence to zero of coefficients of zero series in the Haar and Walsh systems*, Izv. Akad. Nauk SSSR, Ser. Mat. **41** (1977), 703–716.
13. *Null series with respect to a multiplicative system*, Vestnik Mosk. Univ. Ser. Mat. Meh. **6** (1979), 63–67.
14. *Walsh series that are convergent with respect to partial sums*, Mat. Zametki **28** (1980), 45– 52.
15. *On an example of a series in the Walsh system*, Mat. Zametki **28** (1980), 737–748.
16. *Certain estimates of approximation of functions by Cesàro means of Walsh-Fourier series*, Mat. Zametki **29** (1981), 529–547.
17. *On Fourier series with respect to the Walsh- Kaczmarz system*, Analysis Math. **7** (1981), 141–150.
18. *Convergence in L_1 of Fourier series with respect to the Walsh-Kaczmarz system*, Vestnik Mosk. Univ. Ser. Mat. Meh. **6** (1981), 3–6.
19. *Cesàro means of Fourier systems with respect to multiplicative systems*, Vestnik Mosk. Univ. Ser. Mat. Meh. (1982), 7–11.
20. *An example of a U-set for the Walsh system*, Vestnik Mosk. Univ. Ser. Mat. Meh. (1982), 53–55.

V.A. Skvorcov and W.R. Wade. 1. *Generalizations of some results concerning Walsh series and the dyadic derivative*, Analysis Math. **5** (1979), 249–255.

B. Smith. 1. *A strong convergence theorem for $H^1(T)$*, in "Lecture Notes in Math.," Vol. 955, Springer-Verlag, New York, N.Y., 1983, pp. 169–173.

A.A. Šneider. 1. *On series with respect to Walsh functions with monotone coefficients*, Izv. Akad. Nauk SSSR, Ser. Mat. **12** (1948), 179–192.

2. *Uniqueness of expansions with respect to the Walsh system of functions*, Mat. Sbornik **24** (1949), 279–300.
3. *Convergence of Fourier series with respect to the Walsh functions*, Mat. Sbornik **34** (1954), 441–472.

F. Soria. 1. *Integral characterization of a space generated by blocks*, Proceedings of the Cortora Meeting in Harmonic Analysis, Cortora, Italy (1982).

W. Splettstößer. 1. *Error estimates for sampling approximation of non- band limited functions*, Math. Meth. in Appl. Sci. **1** (1979), 127–137.

R.S. Stanković. 1. *A note on differential operators on finite non- abelian groups*, Applicable Analysis **21** (1986), 31–41.

2. *Why IWGD-89? A look at the bibliography of Gibbs derivatives*, in "Theory and Applications of Gibbs Derivatives," Institute of Mathematics, Belgrade, 1989.

S.B. Stečkin. 1. *On approximation of periodic functions by Fejér sums*, Trudy Mat. Inst. Akad. Nauk SSSR **62** (1961), 48–60.

E.M. Stein. 1. *On the limits of sequences of operators*, Annals Math. (2) **74** (1961), 140–170.

E.M. Stein and G. Weiss. 1. "Introduction to Fourier Analysis on Euclidean Spaces," Princeton Univ. Press, Princeton, N.J., 1971.

G. Sunouchi. 1. *On the Walsh-Kaczmarz series*, Proc. Amer. Soc. **2** (1951), 5–11.

2. *Strong summability of Walsh-Fourier series*, Tôhoku Math. J. **16** (1964), 228–237.

G.S. Survillo. 1. *On summability by triangular matrices of Fourier series in multiplicative periodic systems of Vilenkin-Dzafarli type*, Trudy Novosibirsk. Gos. Pedag. Inst. **94** (1974), 34–35.

2. *An application of the Cesàro methods of summability of negative orders to Fourier series in multiplicative orthonormal systems*, Trudy Mosk. Obl. Pedag. Inst. **4** (1974), 241–253.

A. Szankowski. 1. *A Banach lattice without the approximation property*, Israel J. Math. **29** (1976), 329–337.

2. *Subspaces without the approximation property*, Israel J. Math. **30** (1978), 123–129.

B. Szőkefalvi-Nagy. 1. *Approximation properties of orthogonal expansions*, Acta Sci. Math. **15** (1953/54), 31–37.

M.H. Taibleson. 1. "Fourier Analysis on Local Fields," Princeton Univ. Press, Princeton, N.J., 1975.

M.H. Taibleson and G. Weiss. 1. *Certain function spaces connected with almost everywhere convergence of Fourier series*, in "Proc. Conf. A. Zygmund's 80'th birthday (Chicago, 1980)," Wadsworth Publ. Co., 1982.

A.A. Talaljan. 1. *On the convergence a.e. of subsequence of partial sums of general orthogonal series*, Akad. Nauk Armjan. SSR. **10** (1957), 17–34.

2. *The representation of measurable functions by series*, Uspehi Mat. Nauk **15** (1960), 77–141.
3. *On systems, the series with respect to which represents any measurable functions*, Mat. Sbornik **76** (1968), 39–51.
4. *On systems, the series with respect to which represent functions in the class $L_p[0,1]$, for $p \geq 1$*, Mat. Sbornik **78** (1969), 425–445.
5. *Representation and uniqueness problems in the theory of orthogonal series*, J. Soviet Math. **1** (1973), 635–676.

A.A. Talaljan and F.G. Arutunjan. 1. *Convergence of series with respect to the Haar system to $+\infty$*, Mat. Sbornik **66** (1965), 240–247.

K. Tandori. 1. *Bemerkung zur Divergenz der Fourierreihen stetiger Functionen*, Publ. Math. Debrecen **2** (1952), 191–193.

2. *Über die Divergenz der Walschen Reihen*, Acta Sci. Math. **27** (1966), 261–263.
3. *Bemerkung zur starken Summation des Fourierreihen*, Acta Math. Acad. Sci. Hungar. **19** (1968), 271–285.
4. *Ein Divergenzsatz für Fourierreihen*, Acta Sci. Math. (Szeged) **30** (1969), 43–48.
5. *Bemerkung zur Konvergenz der Reihen nach multiplicativen Funktionensystemen*, Acta Math. Acad. Sci. Hungar. **33** (1979), 211–215.

J. Tateoka. 1. *On the localization property of the double Walsh- Fourier series*, Acta Math. Acad. Sci. Hungar. 11 (1976), 407–411.

V.I. Tevzadze. 1. *Uniform convergence of Walsh-Fourier series*, Soobšč. Akad. Nauk Gruz. SSR **92** (1978), 301–304.

2. *Uniform convergence of Cesàro means of negative order of Walsh-Fourier series*, Soobšč. Akad. Nauk Gruz. SSR **102** (1981), 33–36.

P. Turán. 1. *On the strong summability of Fourier series*, J. Ind. Math. Soc. **12** (1948), 8–12.

P.L. Ul'janov. 1. *Divergent series with respect to the Haar system and to bases*, Dokl. Akad. Nauk SSSR **138** (1961), 556–559.

2. *On exact Weyl multipliers for unconditional convergence*, Dokl. Akad. Nauk SSSR **141** (1961), 1040–1049.
3. *On series with respect to the Haar system*, Dokl. Akad. Nauk SSSR **149** (1963), 532–534.
4. *On series with respect to the Haar system*, Mat. Sbornik **63** (1964), 356–391.
5. *Solved and unsolved problems in the theory of trigonometric and orthogonal series*, Uspehi Mat. Nauk **19** (1964), 3–69.
6. *On the approximation of functions*, Sibirsk. Mat. Ž. **5** (1964), 418–437.
7. *Representations of functions of series and classes $\phi(L)$*, Russian Math. Surveys **27** (1972), 1–54.
8. *A.N. Kolmogorov and divergence of Fourier series*, Uspehi Mat. Nauk **38** (1983), 51–90.

N. Ya. Vilenkin. 1. *A class of complete orthonormal series*, Izv. Akad. Nauk SSSR, Ser. Mat. **11** (1947), 363–400.

2. *On the theory of orthogonal systems with gaps*, Izv. Akad. Nauk SSSR, Ser. Mat. **13** (1949), 245–252.
3. *On the theory of Fourier integrals on topological groups*, Mat. Sbornik **30** (1952), 233–244.

N. Ya. Vilenkin and A. Rubinštein. 1. *A theorem of S.B. Stečkin on absolute convergence and character series for a zero dimensional, abelian group*, Izv. Vysš. Učeb. Zaved. (160) No. 9 (1975), 3–9.

W.R. Wade. 1. *A uniqueness theorem for Haar and Walsh series*, Trans. Amer. Math. Soc. **141** (1969), 187– 194.

2. *Summing closed U-sets for Walsh series*, Proc. Amer. Math. Soc. **29** (1971), 123–125.
3. *Uniqueness of Haar series which are $(C,1)$ summable to Denjoy integrable functions*, Trans. Amer. Math. Soc. **176** (1973), 489–498.
4. *Sets of uniqueness and α-capacity on the group 2^ω*, Trans. Amer. Math. Soc. **208** (1975), 309–315.
5. *Growth conditions and uniqueness of Walsh series*, Mich. Math. J. **24** (1977), 153–155.
6. *Sets of uniqueness for the group of integers of a p-series field*, Canad. J. Math. **31** (1979), 858–866.
7. *Uniqueness of Walsh series which satisfy an averaged growth condition*, SIAM J. Math. Anal. **11** (1980), 933–937.
8. *Recent developments in the theory of Walsh series*, International J. Math. **5** (1982), 625–673.
9. *Walsh-Fourier coefficients and locally constant functions*, Proc. Amer. Math. Soc. **87** (1983), 434–438.
10. *Decay of Walsh series and dyadic differentiation*, Trans. Amer. Math. Soc. **277** (1983), 413– 420.
11. *A unified approach to uniqueness of Walsh series and Haar series*, Proc. Amer. Math. Soc. **99** (1987), 61–65.

W.R. Wade and K. Yoneda. 1. *Uniqueness and quasi-measures on the group of integers of a p-series field*, Proc. Amer. Math. Soc. **84** (1982), 202–206.

J.L. Walsh. 1. *A closed set of normal orthogonal functions*, Amer. J. Math. **45** (1923), 5–24.

C. Watari. 1. *On generalized Walsh-Fourier series*, Tôhoku Math. J. **10** (1958), 211–241.
2. *Best approximation by Walsh polynomials*, Tôhoku Math. J. **15** (1963), 1–5.
3. *Mean convergence of Walsh-Fourier series*, Tôhoku Math. J. **16** (1964), 183–188.
4. *Multipliers for Walsh-Fourier series*, Tôhoku Math. J. **16** (1964), 239–251.

D. Waterman. 1. *W-systems are the Walsh functions*, Bull. Amer. Math. Soc. **75** (1969), 139–142.
2. *On systems of functions resembling the Walsh system*, Mich. Math. J. **29** (1982), 83–87.

P. Weiss. 1. *Zusammenhang von Walsh-Fourierreihen mit Polynomen*, Monat. Math. **71** (1967), 165–179.

F. Weisz. 1. *Martingale Hardy spaces for* $0 < p \leq 1$, Prob. Th. Rel. Fields **84** (1990), 361–376.

L.R. Welch. 1. *Computation of finite Fourier series*, Space Programs Summary, Jet Propulsion Laboratory, Pasadena, Calif. **4** (1966), 295–297.

J.E. Whelchel and D.F. Guinn. 1. *Fast Fourier-Hadamard transform and its use in signal representation and classification*, EASON' (1968), 561–573.

D.V. Widder. 1. "The Laplace Transform," Princeton Univ. Press, Princeton, N.J., 1941.

P. Wojtaszczyk. 1. *The Franklin system is an unconditional basis in* H_1, Arkiv för Mat. **20** (1982), 293–300.

Sh. Yano. 1. *On Walsh series*, Tôhoku Math. J. **3** (1951), 223–242.
2. *On approximation by Walsh functions*, Proc. Amer. Math. Soc. **2** (1951), 962–967.
3. *A convergence test for Walsh-Fourier series*, Tôhoku Math. J. **6** (1954), 226–230.
4. *Cesàro summability of Walsh-Fourier series*, Tôhoku Math. J. **9** (1957), 267–272.

K. Yoneda. 1. *On absolute convergence of Walsh-Fourier series*, Math. Japonica **18** (1973), 71–78.
2. *Perfect sets of uniqueness on the group* 2^ω, Canad. J. Math. **34** (1982), 759–764.
3. *Sets of uniqueness for a certain class* $\mathcal{M}_\varepsilon$ *on the dyadic group*, Proc. Amer. Math. Soc. **88** (1983), 279–284.
4. *A sufficient condition for a set to be Dirichlet set on the dyadic group*, Math. Japonica **29** (1984), 45–50.
5. *Summing generalized closed U-sets for Walsh series*, Proc. Amer. Math. Soc. **94** (1985), 110–114.
6. *Uniqueness theorems for a certain small class of Walsh series*, Math. Japonica **30** (1985), 321–326.
7. *Dirichlet sets and some uniqueness theorems for Walsh series*, Tôhoku Math. J. **38** (1986), 1–14.
8. *A generalized capacity and a uniqueness theorem on the dyadic group*, Proc. Amer. Math. Soc. **102** (1988), 52–56.

Young, Wo-Sang. 1. *On rearrangements of Walsh-Fourier series and Hardy-Littlewood type maximal inequalities*, Bull. Amer. Math. Soc. **80** (1974), 490–494.
2. *On a.e. convergence of Walsh-Kaczmarz-Fourier series*, Proc. Amer. Math. Soc. **44** (1974), 353–358.
3. *Mean convergence of generalized Walsh-Fourier series*, Trans. Amer. Soc. **218** (1976), 311–320.
4. *A note on Walsh-Fourier series*, Proc. Amer. Math. Soc. **59** (1976), 305–310.

A.M. Zubakin. 1. *On an analogue of a theorem of Steinhaus*, Sibirsk. Mat. Ž. **9** (1968), 206–210.

A. Zygmund. 1. "Trigonometric Series," Cambridge University Press, New York, N.Y., 1959.

AUTHOR INDEX

SUBJECT INDEX

NOTATION INDEX

This book is suitable for use as an introduction to harmonic analysis. Except for material usually covered in a first course on Lebesgue integration, concepts are developed as needed and in this sense, the book is nearly self-contained. In particular, it is accessible to beginning graduate students and the doctoral candidates in various specialities in engineering. This book includes very recent results and covers topics very broadly, and thus is also suitable as a reference for specialists.

A large part of this book is devoted to the study of Walsh functions. Topics covered include convergence and summability of Walsh–Fourier series, growth of Walsh–Fourier coefficients, uniqueness and representation by Walsh series, approximation by Walsh polynomials, inversion of the Walsh transform and the Fast Walsh Transform. We also examine the role that Walsh functions play in the development of new spaces (the dyadic Hardy spaces and BMO), resolution of the basis problem, and the study of other systems (e.g. the Franklin system and certain multiplicative systems).

Another important theme is that the Walsh functions provide a vehicle to link harmonic analysis, probability theory, tree-like nets, and the theory of finitely additive set functions (quasi-measures). This allows problems to be recast in several different ways and so provide an alternative approach to classical theorems. To illustrate this point, we offer a fairly simple proof of the most general theorem on a.e. convergence of Walsh–Fourier series, and construct an explicit isomorphism from classical Hardy space (defined by means of analytic functions) to dyadic Hardy space (defined by means of martingale maximal functions).

ISBN 963 05 5880 7

WALSH SERIES

F. Schipp

W.R.Wade

P. Simon

with assistance from

J. Pál

1828