Interpolation of Operators

This is Volume 129 in
PURE AND APPLIED MATHEMATICS

A list of titles in this series appears at the end of this volume.

Interpolation of Operators

Colin Bennett
Robert Sharpley

Department of Mathematics
University of South Carolina
Columbia, South Carolina

ACADEMIC PRESS, INC.
Harcourt Brace Jovanovich, Publishers
Boston San Diego New York
Berkeley London Sydney
Tokyo Toronto

ACADEMIC PRESS, INC.
Orlando, Florida 32887

United Kingdom Edition published by
ACADEMIC PRESS INC. (LONDON) LTD.
24-28 Oval Road, London NW1 7DX

Library of Congress Cataloging-in-Publication Data

Bennett, Colin, 1946-
Interpolation of operators.

(Pure and applied mathematics; v. 129)
Bibliography: p. 443
Includes index.
1. Operator theory. 2. Interpolation spaces.
I. Sharpley, Robert. II. Title. III. Series: Pure and applied mathematics (Academic Press) ; 129.
QA3.P8 vol. 129 510 s 515.7'42 87-15324
[QA329]
ISBN 0-12-088730-4

88 89 90 91 9 8 7 6 5 4 3 2 1
Printed in the United States of America

To Margaret and Carla

Contents

Preface

Three classical interpolation theorems form the foundation of the modern theory of interpolation of operators. They are the M. Riesz convexity theorem (1926), G.O. Thorin's complex version of Riesz' theorem (1939), and the J. Marcinkiewicz interpolation theorem (1939). The ideas of Thorin and Marcinkiewicz were reworked some twenty years later into an abstract theory of interpolation of operators on Banach spaces and more general topological spaces. Thorin's technique has given rise to what is now known as the complex method of interpolation, and Marcinkiewicz' to the real method. Both have found widespread application, have extensive literatures attached to them, and remain very much alive as subjects of current research.

This is a book about the real method of interpolation. Our goal has been to motivate and develop the entire theory from its classical origins, that is, through the theory of spaces of measurable functions. Although the influence of Riesz, Thorin, and Marcinkiewicz is everywhere evident, the work of G. H. Hardy, J. E. Littlewood, and G. Pólya on rearrangements of functions also plays a seminal role. It is through the Hardy–Littlewood–Pólya relation that spaces of measurable functions and interpolation of operators come together, in a simple blend which has the capacity for great generalization. Interpolation between L^1 and L^∞ is thus the prototype for interpolation between more general pairs of Banach spaces. This theme airs constantly throughout the book.

The theory and applications of interpolation are as diverse as language itself. Our goal is not a dictionary, or an encyclopedia, but instead a brief biography of interpolation, with a beginning and an end, and (like interpolation itself) some substance in between.

The book should be accessible to anyone familiar with the fundamentals of real analysis, measure theory, and functional analysis. The standard advanced undergraduate or beginning graduate courses in these disciplines should suffice. The exposition is essentially self-contained.

We wish to thank Dr. Carl Riehm and McMaster University, without whose support this project would not have begun, and the University of South Carolina, which provides such a stimulating and exciting environment in which to work. We are particularly indebted to Dr. Sherman Riemenschneider and Mr. Natarajan Sivakumar of the University of Alberta, who suggested many improvements and caught several errors. We thank them for their generosity and their humor, which will someday surely be repaid. Our thanks are due also to Department of Mathematics staff members Jane Squires, Jessie Smith, Sue Darlington, and Dana Ward for their invaluable assistance in preparing the manuscript. It is a pleasure to acknowledge the efficiency and professionalism of our editors, Bill Sribney and Pascha Gerlinger of Harcourt Brace Jovanovich. Above all, we thank our families for their patience and support during the preparation of this work.

Columbia, South Carolina

C. Bennett
R. Sharpley

1 Banach Function Spaces

Although the Lebesgue spaces L^p $(1 \leq p \leq \infty)$ play a primary role in many areas of mathematical analysis, there are other classes of Banach spaces of measurable functions that are also of interest. The larger classes of Orlicz spaces and Lorentz spaces, for example, are of intrinsic importance. There is a considerable literature dealing with each of these classes. In this chapter, however, we shall concentrate not on the differences between such classes but instead on their similarities. This common ground provides the foundation for the abstract theory of *Banach function spaces.*

Banach function spaces are Banach spaces of measurable functions in which the norm is related to the underlying measure in an appropriate way. This allows for a fruitful interplay between functional-analytic and measure-theoretic techniques. The theory is further enriched by the presence of a natural order structure on the function elements themselves, and so may be subsumed in a more general treatment of *Banach lattices*, or *Riesz spaces*, as they are sometimes called. For our purposes, however, this more general point of view will be neither necessary nor desirable.

The Banach function space axioms are displayed in Section 1, where some elementary properties are derived from them. The concept of the *associate space* is introduced in Section 2 and this sets the scene for the discussion of duality, reflexivity, and separability in Sections 3, 4, and 5.

The program follows the same lines as any standard development of the L^p-spaces. In fact, the reader may find it instructive to keep the L^p-spaces in mind as a model for the entire theory of Banach function spaces.

The reader may also find it useful to reflect on the motivation for the particular choice of axioms. The literature shelters more than one axiomatic system under the general umbrella of Banach function spaces. Some use weaker versions of the Fatou property (property (P3) in Definition 1.1), while others rely on a different class of distinguished "bounded" sets in the underlying measure space (properties (P4), (P5)). For example, in a totally σ-finite measure space (R, μ), one could select once and for all an increasing sequence $(R_n)_{n=1}^{\infty}$ of measurable subsets of finite measure whose union is all of R. A measurable subset of R might then be declared "bounded" if it is contained in some set R_n. The approach we have adopted is simpler: the "bounded" sets are just the sets of finite measure. The resulting theory is less technical but also less general at this initial stage. There is, however, no real loss of generality when we specialize to the rearrangement-invariant spaces in the next chapter.

1. BANACH FUNCTION SPACES

Let (R, μ) be a measure space, in the sense described above. Let $\mathscr{M}^+$ be the cone of μ-measurable functions on R whose values lie in $[0, \infty]$. The characteristic function of a μ-measurable subset E of R will be denoted by χ_E.

Definition 1.1. A mapping $\rho : \mathscr{M}^+ \to [0, \infty]$ is called a *Banach function norm* (or simply a *function norm*) if, for all f, g, f_n, $(n = 1, 2, 3, \ldots)$, in $\mathscr{M}^+$, for all constants $a \geq 0$, and for all μ-measurable subsets E of R, the following properties hold:

(P1) $\rho(f) = 0 \;\Leftrightarrow\; f = 0$ μ-a.e.; $\quad \rho(af) = a\rho(f)$;
$\rho(f + g) \leq \rho(f) + \rho(g)$

(P2) $0 \leq g \leq f$ μ-a.e. $\;\Rightarrow\; \rho(g) \leq \rho(f)$

(P3) $0 \leq f_n \uparrow f$ μ-a.e. $\;\Rightarrow\; \rho(f_n) \uparrow \rho(f)$

(P4) $\mu(E) < \infty \;\Rightarrow\; \rho(\chi_E) < \infty$

(P5) $\mu(E) < \infty \;\Rightarrow\; \int_E f \, d\mu \leq C_E \rho(f)$

for some constant C_E, $0 < C_E < \infty$, depending on E and ρ but independent of f.

Among the simplest examples of Banach function norms are those

associated with the Lebesgue spaces L^p $(1 \leq p \leq \infty)$. Let

$$\rho_p(f) = \begin{cases} \left\{ \int_R f^p \, d\mu \right\}^{1/p}, & (1 \leq p < \infty), \\ \operatorname{ess\,sup}_R f, & (p = \infty), \end{cases} \qquad f \in \mathscr{M}^+. \tag{1.1}$$

Theorem 1.2. *The Lebesgue functionals ρ_p, $(1 \leq p \leq \infty)$, are function norms.*

Proof. The triangle inequality for ρ_p is the classical Minkowski inequality. The remaining parts of (P1) are obvious, as are (P2) and (P4). Property (P3) follows from the monotone convergence theorem, and (P5) from Hölder's inequality: if $1 < p < \infty$ and $1/p + 1/p' = 1$, then

$$\int_E f \, d\mu = \int_R f\chi_E \, d\mu \leq \left(\int_R f^p \, d\mu \right)^{1/p} \left(\int_R \chi_E^{p'} \, d\mu \right)^{1/p'} = C_E \rho(f),$$

with $C_E = \mu(E)^{1/p'}$. The cases $p = 1$ and $p = \infty$ are easier so their proofs are omitted. ∎

Let $\mathscr{M}$ denote the collection of all extended scalar-valued (real or complex) μ-measurable functions on R and $\mathscr{M}_0$ the class of functions in $\mathscr{M}$ that are finite μ-a.e. As usual, any two functions coinciding μ-a.e. will be identified. The natural vector space operations are well defined on $\mathscr{M}_0$ (although not on all of $\mathscr{M}$), and when $\mathscr{M}_0$ is given the topology of convergence in measure on sets of finite measure it becomes a metrizable topological vector space (cf. Exercise 1).

Definition 1.3. Let ρ be a function norm. The collection $X = X(\rho)$ of all functions f in $\mathscr{M}$ for which $\rho(|f|) < \infty$ is called a *Banach function space*. For each $f \in X$, define

$$\|f\|_X = \rho(|f|). \tag{1.2}$$

Theorem 1.4. *Let ρ be a function norm and let $X = X(\rho)$ and $\|\cdot\|_X$ be as in Definition 1.3. Then under the natural vector space operations, $(X, \|\cdot\|_X)$ is a normed linear space for which the inclusions*

$$S \subset X \hookrightarrow \mathscr{M}_0 \tag{1.3}$$

hold, where S is the set of μ-simple functions on R. In particular, if $f_n \to f$ in X, then $f_n \to f$ in measure on sets of finite measure, and hence some subsequence converges pointwise μ-a.e. to f.

Proof. It follows from Definition 1.3 and property (P5) of Definition 1.1 that every function in X is locally integrable and hence finite μ-a.e. (because μ is σ-finite). The set X therefore inherits the vector space operations from $\mathcal{M}_0$ and then there is no difficulty in using (P1) and (1.2) to verify that $(X, \|\cdot\|_X)$ is a normed linear space. Property (P4) shows that X contains the characteristic function of every set of finite measure and hence, by linearity, every μ-simple function. This establishes the set-theoretic inclusions in (1.3).

It remains to show that the inclusion map from X to $\mathcal{M}_0$ is continuous. Since both spaces are metrizable it will suffice to show that every sequence convergent in X is convergent also in $\mathcal{M}_0$ (to the same limit, of course). But if $f_n \to f$ in X, then (1.2) shows that $\rho(|f - f_n|) \to 0$ as $n \to \infty$. Let $\varepsilon > 0$ and let E be any subset of R having finite measure. By property (P5),

$$\mu\{x \in E : |f(x) - f_n(x)| > \varepsilon\} \leq \int_E \frac{1}{\varepsilon}|f - f_n|\, d\mu$$

$$\leq \frac{1}{\varepsilon} C_E \rho(|f - f_n|),$$

which converges to 0 as $n \to \infty$ since C_E is independent of n. This shows that $f_n \to f$ in measure on every set of finite measure or, what is the same thing, $f_n \to f$ in $\mathcal{M}_0$. A standard result in measure theory [Ro, p.92] now provides the desired pointwise a.e.-convergent subsequence. ■

The Banach function spaces arising from the functionals ρ_p in (1.1) are of course the familiar Lebesgue spaces $L^p = L^p(R, \mu)$:

$$\|f\|_{L^p} = \begin{cases} \left(\int_R |f|^p\, d\mu\right)^{1/p}, & (1 \leq p < \infty) \\ \operatorname*{ess\,sup}_R |f|, & (p = \infty). \end{cases} \tag{1.4}$$

The next result shows that one of the cornerstones of the L^p-theory, namely Fatou's lemma, has a natural analogue in every Banach function space. Note that the Fatou property (P3) plays a central role here.

Lemma 1.5. *Let $X = X(\rho)$ be a Banach function space and suppose $f_n \in X$, $(n = 1, 2, \ldots)$.*

(i) *If $0 \leq f_n \uparrow f$ μ-a.e., then either $f \notin X$ and $\|f_n\|_X \uparrow \infty$, or $f \in X$ and $\|f_n\|_X \uparrow \|f\|_X$.*

(ii) (*Fatou's lemma*) *If $f_n \to f$ μ-a.e., and if $\liminf_{n\to\infty} \|f_n\|_X < \infty$, then $f \in X$ and*

$$\|f\|_X \leq \liminf_{n\to\infty} \|f_n\|_X.$$

Proof. The first assertion is an immediate consequence of Definition 1.3 and the Fatou property (P3). For part (ii), let $h_n(x) = \inf_{m \geq n} |f_m(x)|$ so that $0 \leq h_n \uparrow |f|$ μ-a.e. By the lattice property (P2) and the Fatou property (P3),

$$\rho(|f|) = \lim_{n\to\infty} \rho(h_n) \leq \lim_{n\to\infty} \left(\inf_{m \geq n} \rho(|f_m|)\right)$$

$$= \liminf_{n\to\infty} \|f_n\|_X < \infty. \tag{1.5}$$

But f is certainly measurable (being the pointwise limit of a sequence of measurable functions) so (1.5) shows that f belongs to X and $\|f\|_X \leq \liminf_{n\to\infty} \|f_n\|_X$. ■

Fatou's lemma is the key to the completeness of the Banach function spaces, which is established in the next result.

Theorem 1.6. *Let X be a Banach function space. Suppose $f_n \in X$, $(n = 1, 2, \ldots)$ and*

$$\sum_{n=1}^{\infty} \|f_n\|_X < \infty. \tag{1.6}$$

Then $\sum_{n=1}^{\infty} f_n$ converges in X to a function f in X and

$$\|f\|_X \leq \sum_{n=1}^{\infty} \|f_n\|_X. \tag{1.7}$$

In particular, X is complete.

Proof. Let $t = \sum_{n=1}^{\infty} |f_n|$, $t_N = \sum_{n=1}^{N} |f_n|$, $(N = 1, 2, 3, \ldots)$, so $0 \leq t_N \uparrow t$. Since

$$\|t_N\|_X \leq \sum_{n=1}^{N} \|f_n\|_X \leq \sum_{n=1}^{\infty} \|f_n\|_X, \qquad (N = 1, 2, 3, \ldots),$$

it follows from (1.6) and Lemma 1.5(i) that t belongs to X. By (1.3), the series $\sum |f_n(x)|$ converges pointwise μ-a.e. and hence so does $\sum f_n(x)$. Thus if

$$f = \sum_{n=1}^{\infty} f_n, \qquad s_N = \sum_{n=1}^{N} f_n, \qquad (N = 1, 2, 3, \ldots),$$

then $s_N \to f$ μ-a.e. Hence, for any M, we have $s_N - s_M \to f - s_M$ μ-a.e. as $N \to \infty$. Furthermore,

$$\liminf_{N\to\infty} \|s_N - s_M\|_X \leq \liminf_{N\to\infty} \sum_{n=M+1}^{N} \|f_n\|_X = \sum_{n=M+1}^{\infty} \|f_n\|_X,$$

which tends to 0 as $M \to \infty$ because of (1.6). It follows therefore from Fatou's lemma (Lemma 1.5(ii)) that $f - s_M$ belongs to X (hence so does f) and $\|f - s_M\|_X \to 0$ as $M \to \infty$. But then, for every $M = 1, 2, 3, \ldots$,

$$\|f\|_X \leq \|f - s_M\|_X + \|s_M\|_X \leq \|f - s_M\|_X + \sum_{n=1}^{M} \|f_n\|_X.$$

The desired inequality (1.7) follows from this by letting M tend to infinity.

The property just established (that absolute convergence in the norm of X implies convergence in X) is often called the *Riesz-Fischer property*. It is routine to verify that the Riesz-Fischer property is equivalent to completeness in *any* normed linear space. Hence X is complete (cf. also Exercise 11). ■

The next theorem summarizes for future reference the basic properties of Banach function spaces established thus far.

Theorem 1.7. *Suppose ρ is a function norm and let*

$$X = \{f \in \mathscr{M} : \rho(|f|) < \infty\}.$$

For each $f \in X$, let $\|f\|_X = \rho(|f|)$. Then $(X, \|\cdot\|_X)$ is a Banach space and the following properties hold for all f, g, f_n, $(n = 1, 2, \ldots)$, in $\mathscr{M}$ and all measurable subsets E of R:

(i) (*the lattice property*) *If $|g| \leq |f|$ μ-a.e. and $f \in X$, then $g \in X$ and $\|g\|_X \leq \|f\|_X$; in particular, a measurable function f belongs to X if and only if $|f|$ belongs to X, and in that case f and $|f|$ have the same norm in X.*

(ii) (*the Fatou property*) *Suppose $f_n \in X$, $f_n \geq 0$, $(n = 1, 2, \ldots)$, and $f_n \uparrow f$ μ-a.e. If $f \in X$, then $\|f_n\|_X \uparrow \|f\|_X$ whereas if $f \notin X$, then $\|f_n\|_X \uparrow \infty$.*

(iii) (*Fatou's lemma*) *If $f_n \in X$, $(n = 1, 2, \ldots)$, $f_n \to f$ μ-a.e., and $\liminf_{n\to\infty} \|f_n\|_X < \infty$, then $f \in X$ and*

$$\|f\|_X \leq \liminf_{n\to\infty} \|f_n\|_X.$$

(iv) *Every simple function belongs to X.*

(v) *To each set E of finite measure there corresponds a constant C_E satisfying $0 < C_E < \infty$ such that*

$$\int_E |f|\, d\mu \leq C_E \|f\|_X$$

for all $f \in X$.

(vi) *If $f_n \to f$ in X, then $f_n \to f$ in measure on every set of finite measure; in particular, some subsequence of $\{f_n\}$ converges to f pointwise μ-a.e.*

We conclude this section with a simple, but useful, observation about Banach function space topologies.

Theorem 1.8. *Let X and Y be Banach function spaces over the same measure space. If $X \subset Y$, then in fact $X \hookrightarrow Y$; equivalently,*

$$\|f\|_Y \leq C\|f\|_X, \qquad (f \in X), \tag{1.8}$$

for some constant C independent of f.

Proof. Suppose $X \subset Y$ but (1.8) fails. Then there exist functions f_n in X for which

$$\|f_n\|_X \leq 1, \qquad \|f_n\|_Y > n^3, \qquad (n = 1, 2, \ldots).$$

Replacing each f_n with its absolute value, we may assume $f_n \geq 0$ for all n. It follows from the Riesz-Fischer property (Theorem 1.6) that $\sum n^{-2} f_n$ converges in X to some function f in X. The hypothesis that X is a subset of Y then shows that f belongs also to Y. But this is impossible because $0 \leq n^{-2} f_n \leq f$ and so $\|f\|_Y \geq n^{-2}\|f_n\|_Y > n$, for all n. Hence (1.8) must hold for some C independent of f. This shows that the inclusion map from X to Y is continuous, that is, $X \hookrightarrow Y$. ■

Corollary 1.9. *If two Banach function spaces consist of the same set of functions, then their norms are equivalent.*

2. THE ASSOCIATE SPACE

The classical Hölder inequality asserts that

$$\int_R |fg|\, d\mu \leq \|f\|_{L^p} \|g\|_{L^{p'}}, \tag{2.1}$$

for all $f \in L^p$ and $g \in L^{p'}$, where

$$1 \leq p \leq \infty, \qquad \frac{1}{p} + \frac{1}{p'} = 1. \tag{2.2}$$

The inequality is sharp in the sense that

$$\|g\|_{L^{p'}} = \sup\left\{\int_R |fg|\,d\mu : f \in L^p, \|f\|_{L^p} \le 1\right\}, \tag{2.3}$$

for all $g \in L^{p'}$ and for all p and p' satisfying (2.2).

Note that the space $L^{p'}$ associated with L^p in (2.1) is described explicitly in terms of L^p by (2.3). A similar situation exists for the Lorentz spaces and the Orlicz spaces. These results will be proved in Chapter IV. For the present, however, we can proceed with the abstract theory by using a suitable analogue of (2.3) to define a space X' corresponding to a given Banach function space X (Definition 2.1). An immediate consequence of this construction is that a sharp form of Hölder's inequality is valid for X and X' (Theorem 2.4). It is less clear, however, that the space X' defined in this way is itself a Banach function space. This is one of the main results of the present section (Theorem 2.2). The other is the Lorentz-Luxemburg theorem (Theorem 2.7) which shows that the second associate space $X'' = (X')'$ always coincides with the original space X.

Definition 2.1. If ρ is a function norm, its *associate norm* ρ' is defined on $\mathcal{M}^+$ by

$$\rho'(g) = \sup\left\{\int_R fg\,d\mu : f \in \mathcal{M}^+, \rho(f) \le 1\right\}, \qquad (g \in \mathcal{M}^+). \tag{2.4}$$

Theorem 2.2. *Let ρ be a function norm. Then the associate norm ρ' is itself a function norm.*

Proof. If $\rho(f) \le 1$, then f is finite μ-a.e. by (1.3). Hence, if $g = 0$ μ-a.e., then $\int fg\,d\mu = 0$ and so $\rho'(g) = 0$ by (2.4). Conversely, if $\rho'(g) = 0$, then $\int fg\,d\mu = 0$ for all $f \in \mathcal{M}^+$ with $\rho(f) \le 1$. If E is any measurable subset of R for which $0 < \mu(E) < \infty$, then properties (P1) and (P4) for ρ give $0 < \rho(\chi_E) < \infty$. Taking $f = \chi_E/\rho(\chi_E)$, so $\rho(f) = 1$, we obtain

$$0 = \int_R fg\,d\mu = \rho(\chi_E)^{-1}\int_E g\,d\mu.$$

This implies that $g = 0$ μ-a.e. on E. Since E is an arbitrary set of (nonzero) finite measure and μ is σ-finite, it follows that $g = 0$ μ-a.e. The remaining properties in (P1) and (P2) for ρ' are easily verified.

For the Fatou property (P3), suppose $g_n, g \in \mathcal{M}^+$, $(n = 1, 2, \ldots)$, and $0 \le g_n \uparrow g$ μ-a.e. By the lattice property (P2) for ρ' (which was established above), the sequence $\rho'(g_n)$ increases with n and $\rho'(g_n) \le \rho'(g)$ for all n. Hence, if $\rho'(g_n) = \infty$ for some n, there is nothing to prove. We therefore

assume that $\rho'(g_n) < \infty$ for all n. Let ξ be any number satisfying $\xi < \rho'(g)$. By (2.4), there is a function f in $\mathcal{M}^+$ with $\rho(f) \le 1$ such that $\int fg\,d\mu > \xi$. Now $0 \le fg_n \uparrow fg$ μ-a.e. so the monotone convergence theorem shows that $\int fg_n \uparrow \int fg$. Hence there is an integer N such that $\int fg_n > \xi$ for all $n \ge N$. But then it follows from (2.4) that $\rho'(g_n) > \xi$ for all $n \ge N$. This shows that $\rho'(g_n) \uparrow \rho'(g)$ and hence establishes the Fatou property (P3) for ρ'.

If $\mu(E) < \infty$, then property (P5) for ρ gives a constant $C_E < \infty$ for which

$$\int_R \chi_E f\,d\mu \le C_E\rho(f), \qquad (f \in \mathcal{M}^+).$$

Together with (2.4) this gives $\rho'(\chi_E) \le C_E < \infty$. Hence (P4) holds for ρ'.

For the remaining property (P5), fix E with $\mu(E) < \infty$. If $\mu(E) = 0$, there is nothing to prove so we may assume $\mu(E) > 0$. In this case, properties (P1) and (P4) for ρ show that the constant $C'_E \equiv \rho(\chi_E)$ satisfies $0 < C'_E < \infty$. The function $f = \chi_E/\rho(\chi_E)$ therefore has ρ-norm equal to 1. Hence, for any $g \in \mathcal{M}^+$, we obtain from (2.4) the estimate

$$\int_E g\,d\mu = C'_E\int_R fg\,d\mu \le C'_E\rho'(g),$$

which shows that (P5) holds for ρ'. ■

Definition 2.3. Let ρ be a function norm and let $X = X(\rho)$ be the Banach function space determined by ρ as in Definition 1.3. Let ρ' be the associate norm of ρ. The Banach function space $X(\rho')$ determined by ρ' is called the *associate space* of X and is denoted by X'.

It follows from (1.2) and (2.4) that the norm of a function g in the associate space X' is given by

$$\|g\|_{X'} = \sup\left\{\int_R |fg|\,d\mu : f \in X, \|f\|_X \le 1\right\}. \tag{2.5}$$

Theorem 2.4. (Hölder's inequality). *Let X be a Banach function space with associate space X'. If $f \in X$ and $g \in X'$, then fg is integrable and*

$$\int_R |fg|\,d\mu \le \|f\|_X\|g\|_{X'}. \tag{2.6}$$

Proof. If $\|f\|_X = 0$, then $f = 0$ μ-a.e. so both sides of (2.6) are zero. If $\|f\|_X > 0$, then the function $f/\|f\|_X$ has norm 1 so (2.5) gives

$$\int_R \left|\left(\frac{f}{\|f\|_X}\right)g\right| d\mu \le \|g\|_{X'}.$$

Multiplying through by $\|f\|_X$ we obtain (2.6). ■

Theorem 2.5. *If p and p' satisfy (2.2), then $L^{p'}$ is the associate space of L^p.*

Proof. This is an immediate consequence of (2.3) and Definitions 2.1 and 2.3 ■

The next result provides a useful converse to the integrability assertion of Theorem 2.4. It is sometimes referred to as *Landau's resonance theorem.*

Lemma 2.6. *In order that a measurable function g belong to the associate space X', it is necessary and sufficient that fg be integrable for every f in X.*

Proof. The necessity follows from Theorem 2.4. In the other direction, suppose that $\rho'(|g|) = \infty$ but that fg is integrable for every f in X. By (2.4), there exist nonnegative functions f_n satisfying

$$\|f_n\|_X \le 1, \qquad \int_R |f_n g|\,d\mu > n^3, \qquad (n = 1, 2, \ldots).$$

It follows from Theorem 1.6 that the function $f = \sum_{n=1}^{\infty} n^{-2} f_n$ belongs to X. However, the product fg cannot be integrable because

$$\int_R |fg|\,d\mu \ge n^{-2} \int_R |f_n g|\,d\mu > n, \qquad (n = 1, 2, \ldots).$$

This contradiction establishes the sufficiency. ■

Theorem 2.7 (G. G. Lorentz; W. A. J. Luxemburg). *Every Banach function space X coincides with its second associate space X''. In other words, a function f belongs to X if and only if it belongs to X'', and in that case*

$$\|f\|_X = \|f\|_{X''}. \tag{2.7}$$

Proof. If $f \in X$, then Hölder's inequality (2.6) shows that fg is integrable for every $g \in X'$. It follows therefore from Lemma 2.6 (applied to X' instead of X) that $f \in X''$. Hence, $X \subset X''$. We also obtain from (2.5) and (2.6),

$$\|f\|_{X''} = \sup\left\{\int_R |fg|\,d\mu : \|g\|_{X'} \le 1\right\} \le \|f\|_X.$$

Hence, in order to complete the proof we need only show $X'' \subset X$ and

$$\|f\|_X \le \|f\|_{X''}, \qquad (f \in X''). \tag{2.8}$$

First, choose an increasing sequence of sets R_N, $(N = 1, 2, \ldots)$, of finite measure whose union is R (possible because (R, μ) is totally σ-finite). For each

N and each $f \in X''$, let

$$f_N(x) = \min(|f(x)|, N)\chi_{R_N}(x). \tag{2.9}$$

Since $0 \le f_N \le N\chi_{R_N}$, that is, f_N is dominated by a simple function, it follows from Theorem 1.7 ((i), (iv)) that f_N belongs to X and to X''. Furthermore, since $0 \le f_N \uparrow |f|$, it follows from the Fatou properties for X and X'' (Theorem 1.7 (ii)) that in order to establish (2.8) it will suffice to show that

$$\|f_N\|_X \le \|f_N\|_{X''}, \qquad (N = 1, 2, \ldots). \tag{2.10}$$

For the remainder of the proof we suppose therefore that f and N are fixed. Clearly we may assume $\|f_N\|_X > 0$ since otherwise there is nothing to prove.

Let L_N^1 denote the space of μ-integrable functions on R having supports in R_N. With norm $g \to \int_{R_N} |g|\,d\mu$, it is clear that L_N^1 is a Banach space. If S denotes the closed unit ball of X, then the set $U = S \cap L_N^1$ is evidently a convex subset of L_N^1. It is also closed. For if $h_n \in U, (n = 1, 2, \ldots)$, and $h_n \to h$ in L_N^1, then some subsequence $h_{n(k)}$, $(k = 1, 2, \ldots)$, converges to h pointwise μ-a.e. on R. Since every $h_{n(k)}$ belongs to S, Fatou's lemma (Theorem 1.7(iii)) then shows that the same is true of h. Hence $h \in U$. It follows therefore that U is a closed convex subset of L^1.

If λ is any constant satisfying $\lambda > 1$, the function $g = \lambda f_N / \|f_N\|_X$ belongs to L_N^1 but not to U. By the Hahn-Banach theorem [Ru, Theorem 3.4], some closed hyperplane separates g and U, that is, there exists a nonzero $\phi \in L^\infty(R, \mu)$, which may be chosen with support in R_N, such that

$$\operatorname{Re}\left(\int_{R_N} \phi h\,d\mu\right) < \gamma < \operatorname{Re}\left(\int_{R_N} \phi g\,d\mu\right) \tag{2.11}$$

for some real number γ and all h in U. In fact, by writing $\phi = |\phi|\psi$ in polar form and observing that $\psi|h|$ belongs to U if and only if h does, we obtain

$$\sup_{h \in U} \int_{R_N} |\phi h|\,d\mu \le \gamma < \operatorname{Re}\left(\int_{R_N} \phi g\,d\mu\right) \le \int_{R_N} |\phi g|\,d\mu. \tag{2.12}$$

Now an arbitrary function h in S, when restricted to R_N, is the pointwise limit of the increasing sequence of truncations

$$h_n(x) = \min(h(x), n)\chi_{R_N}(x),$$

each of which is in L_N^1 and hence in U. It follows therefore from the monotone convergence theorem that the supremum over U in (2.12) may be replaced by the supremum over S. Hence, from (2.5) and (2.12),

$$\|\phi\|_{X'} = \sup_{h \in S} \int_{R_N} |\phi h|\,d\mu \le \gamma < \frac{\lambda}{\|f_N\|_X} \int_{R_N} |\phi f_N|\,d\mu;$$

equivalently,

$$\|f_N\|_X < \lambda \int_R \left| f_N \frac{\phi}{\|\phi\|_{X'}} \right| d\mu \leq \lambda \|f_N\|_{X''},$$

where the last inequality follows from Hölder's inequality (2.6) for X' and X''. Letting $\lambda \to 1$, we obtain the desired inequality (2.10). ■

With the next result we can begin to explore the relationship between the associate space and the Banach space dual.

Lemma 2.8. *The norm of a function g in the associate space X' is given by*

$$\|g\|_{X'} = \sup\left\{ \left| \int_R fg\, d\mu \right| : f \in X, \|f\|_X \leq 1 \right\}. \tag{2.13}$$

Proof. Since $|\int fg| \leq \int |fg|$, it is clear from (2.5) that the quantity on the right of (2.13) does not exceed $\|g\|_{X'}$. Hence, we need only establish the reverse inequality, which by (2.5) may written

$$\sup_{f \in S} \int_R |fg|\, d\mu \leq \sup_{f \in S} \left| \int_R fg\, d\mu \right|, \tag{2.14}$$

where both suprema extend over all f in the unit ball S of X.

On the set $E = \{x \in R : g(x) \neq 0\}$ we may write $g(x)$ in polar form $g(x) = |g(x)|\phi(x)$, where $|\phi(x)| = 1$. Hence, $|g(x)| = g(x)\bar{\phi}(x)$ on E. For any f in S, we thus have

$$\int_R |fg|\, d\mu = \int_E |fg|\, d\mu = \int_E |f|\bar{\phi} g\, d\mu. \tag{2.15}$$

If $h = |f|\bar{\phi}$ on E, and $h = 0$ off E, then $|h| \leq |f|$ on R and so $h \in S$. Hence, (2.15) gives

$$\int_R |fg|\, d\mu = \int hg\, d\mu \leq \left| \int_R hg\, d\mu \right| \leq \sup_{f \in S} \left| \int_R fg\, d\mu \right|.$$

Taking the supremum on the left over all f in S, we obtain (2.14). ■

Recall that a closed linear subspace B of the dual space X^* of a Banach space X is said to be *norm-fundamental* if

$$\|f\|_X = \sup\{|L(f)| : L \in B, \|L\|_{X^*} \leq 1\}$$

for every $f \in X$. Thus B is norm-fundamental if it contains sufficiently many functionals to reproduce the norm of every element of X.

Theorem 2.9. *The associate space X' of a Banach function space X is canonically isometrically isomorphic to a closed norm-fundamental subspace of the Banach space dual X^* of X.*

Proof. For each $g \in X'$, the linear functional L_g on X given by

$$L_g(f) = \int_R fg\, d\mu, \qquad (f \in X), \tag{2.16}$$

is bounded by virtue of Hölder's inequality (2.6). Furthermore, if $L_g(f) = 0$ for all $f \in X$, then (2.5) shows that $g = 0$. Hence, $g \to L_g$ is an isomorphism of X' onto a linear subspace of X^*. It is isometric because of (2.13):

$$\|L_g\|_{X^*} = \sup\{|L_g(f)| : \|f\|_X \leq 1\} = \|g\|_{X'}. \tag{2.17}$$

Since X' is complete, this implies that the isomorphism has closed range in X^*. Finally, if $f \in X$, then (2.7) and (2.13) give

$$\|f\|_X = \|f\|_{X''} = \sup\left\{ \left| \int_R fg\, d\mu \right| : g \in X', \|g\|_{X'} \leq 1 \right\}.$$

Using (2.16) and (2.17), we therefore obtain

$$\|f\|_X = \sup\{|L_g(f)| : g \in X', \|L_g\|_{X^*} \leq 1\},$$

which shows that the image of X' is norm-fundamental in X^*. ■

Proposition 2.10. *If X and Y are Banach function spaces and $X \hookrightarrow Y$, then $Y' \hookrightarrow X'$. In fact, if $\|f\|_Y \leq c\|f\|_X$ for all f in X, then $\|g\|_{X'} \leq c\|g\|_{Y'}$ for all g in Y'.*

Proof. It follows from (2.5) and the hypothesis on X and Y that

$$\begin{aligned} \|g\|_{X'} &\leq \sup\left\{ \int_R |fg|\, d\mu : \|f\|_Y \leq c \right\} \\ &= c \sup\left\{ \int_R |fg|\, d\mu : \|f\|_Y \leq 1 \right\} \\ &= c\|g\|_{Y'} \end{aligned}$$

for any g in Y'. ■

3. ABSOLUTE CONTINUITY OF THE NORM

In view of Theorem 2.9, it is natural to attempt to characterize the Banach function spaces X for which the associate space X' and the dual space X^* coincide. The present section lays some of the groundwork that will be needed for the solution of this and other duality problems in §4. The analysis hinges on an interplay between two distinguished subspaces X_a and X_b of X. These are respectively the subspace of functions of absolutely continuous norm and the subspace consisting of the closure of the set of simple functions.

Throughout this section we shall use $\{E_n\}_{n=1}^{\infty}$ to denote an arbitrary sequence of μ-measurable subsets of R. We shall write $E_n \to \varnothing$ μ-a.e. if the characteristic functions χ_{E_n} converge to 0 pointwise μ-a.e; if, in addition, the sequence $\{E_n\}$ is decreasing, we shall write $E_n \downarrow \varnothing$ μ-a.e. It is not hard to see that $E_n \to \varnothing$ μ-a.e if and only if the upper limit

$$\limsup_{n\to\infty} E_n = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} E_n$$

of the sequence $\{E_n\}$ is a set of μ-measure zero. Hence $E_n \to \varnothing$ μ-a.e. if and only if $\{E_n\}$ can be modified on a set of measure zero so as to converge to the empty set $\varnothing$. Notice that the sets E_n are not required to have finite measure.

Definition 3.1. A function f in a Banach function space X is said to have *absolutely continuous norm* in X if $\|f\chi_{E_n}\|_X \to 0$ for every sequence $\{E_n\}_{n=1}^{\infty}$ satisfying $E_n \to \varnothing$ μ-a.e. The set of all functions in X of absolutely continuous norm is denoted by X_a. If $X_a = X$, then the space X itself is said to have *absolutely continuous norm.*

The next result shows that in the preceding definition we could have restricted our attention to decreasing sequences $\{E_n\}$.

Proposition 3.2. *A function f in a Banach function space X has absolutely continuous norm if and only if $\|f\chi_{E_n}\|_X \downarrow 0$ for every sequence $\{E_n\}_{n=1}^{\infty}$ satisfying $E_n \downarrow \varnothing$ μ-a.e.*

Proof. The necessity is obvious. For the sufficiency, suppose f has the stated property for decreasing sequences and let $\{F_n\}_{n=1}^{\infty}$ be an arbitrary sequence for which $F_n \to \varnothing$ μ-a.e. Then the sequence $E_n = \bigcup_{m\geq n} F_m (n = 1, 2, \ldots)$ is decreasing and has the same upper limit as $\{F_n\}$, namely a set of measure zero. The hypothesis on f therefore implies $\|f\chi_{E_n}\| \downarrow 0$. Since $F_n \subset E_n$ for all n, it follows that $\|f\chi_{F_n}\| \to 0$ and hence that f has absolutely continuous norm. ■

Example 3.3. If $1 \leq p < \infty$, the Lebesgue space $L^p(R, \mu)$ has absolutely continuous norm. This is an immediate consequence of the preceding

proposition and the dominated convergence theorem. When $p = \infty$, the situation is somewhat different in that the absolute continuity depends on the structure of the underlying measure space (R, μ). For example, if (R, μ) is nonatomic, then the absolutely continuous part $(L^\infty)_a$ of L^∞ contains only the zero function. However, if (R, μ) is completely atomic, as in the case of the natural numbers, then the absolutely continuous part of l^∞ is the subspace c_0 consisting of the sequences which converge to zero at infinity (cf. Exercise 2).

We shall next give two useful characterizations of the functions of absolutely continuous norm (Propositions 3.5 and 3.6). The following lemma will be needed.

Lemma 3.4. *If f has absolutely continuous norm, then to each $\varepsilon > 0$ there corresponds $\delta > 0$ such that $\mu(E) < \delta$ implies $\|f\chi_E\|_X < \varepsilon$.*

Proof. Suppose not. Then, for some $\varepsilon > 0$, there are measurable sets E_n, $(n = 1, 2, \ldots)$, satisfying

$$\mu(E_n) < 2^{-n}, \qquad \|f\chi_{E_n}\|_X \geq \varepsilon.$$

But then the estimate

$$\mu\left(\bigcup_{n=m}^{\infty} E_n\right) \leq \sum_{n=m}^{\infty} \mu(E_n) < 2^{-m+1}, \qquad (m = 1, 2, \ldots)$$

shows that $E_n \to \varnothing$ μ-a.e., and so the norm-estimate above contradicts the fact that f has absolutely continuous norm. ■

Proposition 3.5. *A function f in a Banach function space X has absolutely continuous norm if and only if $\|f_n\|_X \downarrow 0$ for every sequence $\{f_n\}_{n=1}^{\infty}$ of μ-measurable functions satisfying $|f| \geq f_n \downarrow 0$ μ-a.e.*

Proof. The sufficiency is an immediate consequence of Proposition 3.2 (take $f_n = f\chi_{E_n}$). For the necessity, suppose f has absolutely continuous norm and $|f| \geq f_n \downarrow 0$ μ-a.e. Let $\varepsilon > 0$ and let $\{R_N\}_{N=1}^{\infty}$ be an increasing sequence of sets of nonzero finite measure whose union is R. Since the complements $Q_N = R - R_N \downarrow \varnothing$, the absolute continuity of the norm of f shows that $\|f\chi_{Q_N}\|_X < \varepsilon/2$, for some suitably large N. If $\alpha = \varepsilon/(4\|\chi_{R_N}\|)$, let $E_n = \{x \in R_N : f_n(x) > \alpha\}$, $(n = 1, 2, \ldots)$. Since $f_n \downarrow 0$ μ-a.e. and R_N has finite measure we see that $\mu(E_n) \downarrow 0$. Hence, Lemma 3.4 gives $\|f\chi_{E_n}\|_X < \varepsilon/4$ for all sufficiently large n. For these values of n we have

$$\begin{aligned}\|f_n\|_X &\le \|f_n\chi_{Q_N}\|_X + \|f_n\chi_{R_N}\|_X\\ &\le \|f_n\chi_{Q_N}\|_X + \|f_n\chi_{E_n}\|_X + \|f_n\chi_{R_N-E_n}\|_X\\ &\le \|f\chi_{Q_N}\|_X + \|f\chi_{E_n}\|_X + \alpha\|\chi_{R_N-E_n}\|_X\\ &< \frac{\varepsilon}{2}+\frac{\varepsilon}{4}+\frac{\varepsilon}{4}=\varepsilon.\end{aligned}$$

Hence $\|f_n\|_X \downarrow 0$ as required. ■

The monotone convergence theorem is at our disposal in every Banach function space X in the form of the Fatou property. The next result exhibits X_a as the largest subspace of X for which a suitable dominated convergence theorem holds.

Proposition 3.6. *A function f in a Banach function space X has absolutely continuous norm if and only if the following condition holds: whenever f_n, $(n = 1, 2, \ldots)$, and g are μ-measurable functions satisfying $|f_n| \le |f|$ for all n and $f_n \to g$ μ-a.e., then $\|f_n - g\|_X \to 0$.*

Proof. The sufficiency follows as in the previous result by letting $f_n = f\chi_{E_n}$ and $g = 0$. For the necessity, suppose f has absolutely continuous norm and let f_n, $(n = 1, 2, \ldots)$, and g be functions satisfying $|f_n| \le |f|$ μ-a.e. and $f_n \to g$ μ-a.e. If

$$h_n(x) = \sup_{m\ge n}|f_m(x) - g(x)|, \qquad (n = 1, 2, \ldots),$$

then clearly $2|f| \ge h_n \downarrow 0$ μ-a.e. By Proposition 3.5, we therefore have $\|h_n\|_X \downarrow 0$ and hence $\|f_n - g\|_X \le \|h_n\|_X \downarrow 0$. ■

Definition 3.7. A closed linear subspace Y of a Banach function space X is called an *order ideal* of X if it has the property:

$$f \in Y \text{ and } |g| \le |f| \ \mu\text{-a.e.} \quad \Rightarrow \quad g \in Y. \tag{3.1}$$

Clearly the zero subspace and the space X itself are order ideals of X.

Theorem 3.8. *The subspace X_a of functions of absolutely continuous norm is an order ideal of the Banach function space X. Furthermore, if $0 \le f_n \uparrow f$ μ-a.e. and $f \in X_a$, then $\|f - f_n\|_X \downarrow 0$.*

Proof. It follows immediately from Definition 3.1 that X_a is a subspace of X and satisfies (3.1). Hence, we will have shown that X_a is an order ideal of X if we show that X_a is closed. Suppose $f_n \in X_a$ $(n = 1, 2, \ldots)$ and $f_n \to f$ in X. Then, given $\varepsilon > 0$, we shall have $\|f - f_N\|_X < \varepsilon/2$ for some suitably large N. Suppose $\{E_m\}_{m=1}^{\infty}$ is a sequence satisfying $E_m \downarrow \varnothing$ μ-a.e. Since f_N has absolutely continuous norm, there exists M such that

$$\|f_N\chi_{E_m}\|_X < \frac{\varepsilon}{2}, \qquad (m \ge M).$$

Hence,

$$\begin{aligned}\|f\chi_{E_m}\|_X &\le \|(f - f_N)\chi_{E_m}\|_X + \|f_N\chi_{E_m}\|_X\\ &\le \|f - f_N\|_X + \|f_N\chi_{E_m}\|_X\\ &< \frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon, \qquad (m \ge M).\end{aligned}$$

This shows that $\|f\chi_{E_m}\|_X \downarrow 0$ and hence that $f \in X_a$. It follows that X_a is closed. The remaining assertion of the theorem, on the norm-convergence of increasing sequences in X_a, is a direct consequence of Proposition 3.6. ■

Recall (from Theorem 1.7) that the simple functions are contained in every Banach function space X. We now consider the closed subspace of X generated by the simple functions and explore its relationship to the subspace X_a of functions of absolutely continuous norm.

Definition 3.9. Let X be a Banach function space. The closure in X of the set of simple functions is denoted by X_b.

Proposition 3.10. *The subspace X_b is the closure in X of the set of bounded functions supported in sets of finite measure.*

Proof. We need only show that every bounded function supported in a set of finite measure belongs to X_b. Suppose therefore that f is bounded and $E = \{x : f(x) \neq 0\}$ has finite measure. If $f \ge 0$, it is not hard to construct a sequence of nonnegative simple functions f_n, $(n = 1, 2, \ldots)$, supported on E such that $f_n \to f$ uniformly. But then

$$\|f_n - f\|_X = \|(f_n - f)\chi_E\|_X \le \|f_n - f\|_{L^\infty}\|\chi_E\|_X,$$

which tends to 0 as $n \to \infty$, since E has finite measure and hence $\|\chi_E\|_X < \infty$.

The general case follows from this one by splitting f into its real and imaginary parts, and each of these into its positive and negative parts. ■

Theorem 3.11. *The subspace X_b is an order ideal of X and*

$$X_a \subset X_b \subset X. \tag{3.2}$$

Proof. Since X_b is evidently a closed subspace of X, to show that X_b is an order ideal we need only establish the lattice property (3.1). So suppose $f \in X_b$ and $|g| \leq |f|$ μ-a.e. Let $(f_n)_{n=1}^{\infty}$ be a sequence of simple functions such that $f_n \to f$ in X. Then each of the functions

$$g_n(x) = \operatorname{sgn}(g(x)) \cdot \min\{|f_n(x)|, |g(x)|\}, \qquad (n = 1, 2, \ldots)$$

is bounded and has support in a set of finite measure. Furthermore,

$$|g - g_n| = \max\{|g| - |f_n|, 0\} \leq ||f| - |f_n|| \leq |f - f_n|$$

so

$$\|g - g_n\|_X \leq \|f - f_n\|_X \to 0.$$

Using Proposition 3.10, we see that $g \in X_b$ and hence that X_b is an order ideal.

Suppose now that f has absolutely continuous norm. If $(R_n)_{n=1}^{\infty}$ is an increasing sequence of sets of finite measure whose union is R, then by Proposition 3.10 each of the functions

$$f_n(x) = \operatorname{sgn}(f(x)) \cdot \min\{|f(x)|, n\chi_{R_n}(x)\}, \qquad (n = 1, 2, \ldots)$$

belongs to X_b. Since $|f_n| \leq |f|$ and $f_n \to f$ μ-a.e., it follows from Proposition 3.6 that $f_n \to f$ in X. But X_b is closed so we have $f \in X_b$. Hence $X_a \subset X_b$. ■

The subspace X_b is always relatively large in the sense that it is a norm-fundamental subspace of $(X')^*$ (Theorem 3.12). By contrast, the subspace X_a may contain only the zero-function (Example 3.3) and so the inclusion $X_a \subset X_b$ may be proper (cf. also Exercise 3). We shall be interested in the opposite extreme when X_a and X_b coincide (cf. Theorem 4.1). Theorem 3.13 will show that this occurs if and only if the characteristic functions of the sets of finite measure all have absolutely continuous norms.

Theorem 3.12. *The subspace X_b of X is isometrically isomorphic to a norm-fundamental subspace of $(X')^*$.*

Proof. By Theorems 2.7 and 2.9 we may identify X, and hence X_b, with a closed subspace of $(X')^*$. We have to show that X_b is norm-fundamental. To this end, suppose $g \in X'$ and let $\varepsilon > 0$. By (2.5), there is a function f in the unit ball of X for which

$$\|g\|_{X'} \leq \int_R |fg| \, d\mu + \varepsilon. \tag{3.3}$$

Let $(R_N)_{N=1}^{\infty}$ be any increasing sequence of sets of finite measure whose union is R, and let

$$f_N = \min(|f|, N) \cdot \chi_{R_N}, \qquad (N = 1, 2, \ldots).$$

By Proposition 3.10, each f_N belongs to the unit ball $B = \{h \in X_b : \|h\|_X \leq 1\}$ of X_b. Since $0 \leq f_N \uparrow |f|$, the monotone convergence theorem and (3.3) give

$$\|g\|_{X'} \leq \sup_N \int_R |f_N g| \, d\mu + \varepsilon \leq \sup_{h \in B} \int_R |hg| \, d\mu + \varepsilon.$$

Letting $\varepsilon \to 0$, and observing that the reverse inequality is an immediate consequence of Hölder's inequality, we thus obtain

$$\|g\|_{X'} = \sup_{h \in B} \int_R |hg| \, d\mu. \tag{3.4}$$

Finally, note that (3.4) remains valid when the absolute value is placed outside the integral as in Lemma 2.8. Indeed, the only property needed in the proof of that lemma was the lattice property for X, and this remains true for X_b since it is an order ideal (Theorem 3.11). With this modification the identity (3.4) shows that X_b is norm-fundamental in $(X')^*$. ■

Theorem 3.13. *The subspaces X_a and X_b coincide if and only if the characteristic function χ_E has absolutely continuous norm for every set E of finite measure.*

Proof. The necessity is obvious since χ_E belongs to X_b for every set E of finite measure. Conversely, if every χ_E with $\mu(E) < \infty$ has absolutely continuous norm, then so does every simple function. Since X_a is closed it follows that $X_b \subset X_a$. The reverse inclusion is supplied by (3.2) so $X_a = X_b$. ■

4. DUALITY AND REFLEXIVITY

Let Y be a closed subspace of a Banach function space X and suppose Y contains the simple functions. Each function g in the associate space X' induces a bounded linear functional $L_g : f \to \int fg \, d\mu$ on X, and hence on Y.

The resulting map $g \to L_g$ is injective (because Y contains the characteristic function of every set of finite measure) and isometric (by Theorem 3.12 and the fact that Y contains X_b). Hence X' may be regarded as a closed subspace of Y^*. When Y is also an order ideal, there is a simple necessary and sufficient condition (cf. Theorem 4.1) for X' and Y^* to coincide. Through a series of corollaries this result provides answers to many of the duality questions raised in the preceding sections.

Theorem 4.1. *Let Y be an order ideal of a Banach function space X and suppose Y contains the simple functions. Then $Y^* = X'$ if and only if $Y \subset X_a$. In this case,*

$$Y = X_a = X_b. \tag{4.1}$$

Proof. The last assertion is clear. For if Y contains the simple functions, then it also contains their closure, namely the subspace X_b. Hence, if $Y \subset X_a$, then (3.2) gives $Y \subset X_a \subset X_b \subset Y$, from which (4.1) follows.

Turning to the main assertion of the theorem, we suppose first that $Y \subset X_a$ and hence that (4.1) holds. By the remarks preceding the statement of the theorem, we need only show $Y^* \subset X'$. So suppose L belongs to Y^*. We shall exhibit a function g in X' such that, for all $f \in Y$, we have

$$L(f) = \int_R fg \, d\mu. \tag{4.2}$$

Since (R, μ) is σ-finite, there is a sequence $\{S_N\}_{N=1}^{\infty}$ of disjoint subsets of R, each of which has finite measure, and whose union is all of R. For each $N = 1, 2, \ldots$, let $\mathscr{A}_N$ denote the σ-algebra of all μ-measurable subsets of S_N, and define a set-function λ_N on $\mathscr{A}_N$ by

$$\lambda_N(A) = L(\chi_A), \qquad (A \in \mathscr{A}_N).$$

Note that $\lambda_N(A)$ is well-defined for all $A \in \mathscr{A}_N$ because χ_A belongs to X_b and hence, by (4.1), to Y.

We claim that λ_N is countably additive on $\mathscr{A}_N$. To see this, let $(A_i)_{i=1}^{\infty}$ be a sequence of disjoint sets from $\mathscr{A}_N$ and let

$$B_n = \bigcup_{i=1}^{n} A_i, \qquad (n = 1, 2, \ldots), \qquad A = \bigcup_{i=1}^{\infty} A_i = \bigcup_{n=1}^{\infty} B_n.$$

Since $A \in \mathscr{A}_N$ we have $\chi_A \in X_b = X_a$, by (4.1). Furthermore, it is clear that $\chi_A \geq \chi_{A - B_n} \downarrow 0$, so Proposition 3.5 shows that $\|\chi_A - \chi_{B_n}\|_X \downarrow 0$ as $n \to \infty$. The

continuity and linearity of L on Y therefore give

$$\lambda_N(A) = L(\chi_A) = \lim_{n\to\infty} L(\chi_{B_n}) = \lim_{n\to\infty} \sum_{i=1}^{n} L(\chi_{A_i}) = \sum_{i=1}^{\infty} \lambda_N(A_i),$$

and this establishes the σ-additivity of λ_N.

It is easy to obtain a bound for the measure by estimating

$$|\lambda_n(A)| = |L(\chi_A)| \leq \|L\|_{Y^*} \|\chi_A\|_X \leq \|L\|_{Y^*} \|\chi_{S_N}\|_X, \qquad (A \in \mathscr{A}_N),$$

and observing that $\|\chi_{S_N}\|_X$ is finite because S_N has finite measure (Theorem 1.7). The estimate

$$|\lambda_N(A)| \leq \|L\|_{Y^*} \|\chi_A\|_X, \qquad (A \in \mathscr{A}_N)$$

shows also that λ_N is absolutely continuous with respect to μ (restricted to $\mathscr{A}_N$) because $\mu(A) = 0$ implies $\|\chi_A\|_X = 0$. Hence, by the Radon-Nikodym theorem, there is a unique (μ-a.e.) function g_N on S_N such that

$$L(\chi_A) = \lambda_N(A) = \int_R \chi_A g_N \, d\mu, \qquad (A \in \mathscr{A}_N).$$

Since the sets S_N are disjoint we may define a function g on all of R by setting $g = g_N$ on each S_N. Clearly,

$$L(\chi_A) = \int_R \chi_A g \, d\mu, \tag{4.3}$$

for all sets A in $\bigcup_{N=1}^{\infty} \mathscr{A}_N$.

This establishes (4.2) in a special case. Before passing to the general case we first show that g belongs to X'. Let h be a nonnegative simple function with support in some set $R_n \equiv \bigcup_{m=1}^{n} S_m$. If we suppose for the moment that g is real-valued, then clearly $h \cdot \operatorname{sgn}(g)$ is also a simple function with support in R_n. In particular, this function is a finite linear combination of characteristic functions of sets in $\bigcup_{N=1}^{\infty} \mathscr{A}_N$. Hence, we may apply (4.3) and use the linearity of L to obtain

$$\int |hg| \, d\mu = \int h \cdot \operatorname{sgn}(g) g \, d\mu = L(h \cdot \operatorname{sgn}(g)).$$

But L is bounded on Y so

$$\int |hg| \, d\mu \leq \|L\|_{Y^*} \|h \cdot \operatorname{sgn}(g)\|_Y \leq \|L\|_{Y^*} \|h\|_X. \tag{4.4}$$

If now f is an arbitrary function in X, then we may construct a sequence $(h_n)_{n=1}^{\infty}$ of simple functions, each h_n having support in R_n, such that $0 \leq h_n \uparrow |f|$ μ-a.e. Applying (4.4) to each h_n in turn, we see from the monotone convergence theorem that the left-hand sides converge to $\int |fg| \, d\mu$, and from the Fatou property that the right-hand sides converge to $\|L\|_{Y^*}\|f\|_X$. Hence,

$$\int_R |fg|\,d\mu \le \|L\|_{Y^*}\|f\|_X, \qquad (f \in X).$$

This, together with Lemma 2.6, implies that g belongs to X'. If g is complex-valued, then the same argument applied separately to the real and imaginary parts of g shows that each of these is in X' and hence that g again belongs to X'.

As in the remarks preceding the statement of the theorem, we may thus regard the linear functional L_g induced by g as a member of Y^*. In this way, we may interpret (4.3) as the assertion that the two functionals L and L_g in Y^* coincide on all simple functions with support in some R_n. It is now fairly easy to show that L and L_g coincide on all of Y. For if f is any real-valued function in Y, then, as above, there exists a sequence of simple functions h_n, $(n = 1, 2, \dots)$, with the support of h_n contained in R_n, such that $0 \le h_n \uparrow |f|$ μ-a.e. The functions $f_n = h_n \cdot \operatorname{sgn}(f)$ are also simple and have supports in the sets R_n, and $f_n \to f$ μ-a.e. But f has absolutely continuous norm (by (4.1)) so Proposition 3.6 shows that $f_n \to f$ in Y. Since $L(f_n) = L_g(f_n)$ and both functionals are continuous on Y, we conclude that $L(f) = L_g(f)$. Finally, if f is complex-valued, then both its real and imaginary parts belong to Y (because Y is an order ideal). The obvious argument therefore shows that L and L_g coincide on all of Y. As we remarked previously, this establishes that $X' = Y^*$.

Conversely, assuming $X' = Y^*$, we have to show that $Y \subset X_a$. Suppose to the contrary that some function f in Y does not have absolutely continuous norm. Since Y is an order ideal we may suppose $f \ge 0$. In that case, there is a sequence $E_n \downarrow \varnothing$ μ-a.e. and a positive number ε such that

$$\|f\chi_{E_n}\|_X \ge \varepsilon, \qquad (n = 1, 2, \dots). \tag{4.5}$$

It is an immediate consequence of the dominated convergence theorem that the sets

$$G_n = \left\{ g \in X' : \left|\int f\chi_{E_n} g\,d\mu\right| < \frac{\varepsilon}{2} \right\}, \qquad (n = 1, 2, \dots)$$

cover X'. But $X' = Y^*$ so we may also regard $(G_n)_{n=1}^{\infty}$ as a cover of Y^*. Now $f\chi_{E_n}$ belongs to Y (because Y is an order ideal) so in fact each G_n is a weak*-open subset of Y^*. Hence, by Alaoglu's theorem, there are finitely many indices $n(1), n(2), \dots, n(k)$ such that $(G_{n(i)})_{i=1}^{k}$ covers the unit ball of Y^*. Every g in X' with $\|g\|_{X'} \le 1$ therefore lies in some $G_{n(i)}$, so since $f \ge 0$ and $E_n \downarrow$ we have

$$\int |f\chi_{E_n} g|\,d\mu \le \int |f\chi_{E_{n(i)}} g|\,d\mu < \frac{\varepsilon}{2},$$

for all $n \ge N = \max\{n(1), n(2), \dots, n(k)\}$. This, together with Theorem 2.7 and Lemma 2.8, gives

$$\|f\chi_{E_n}\|_X = \sup_{\|g\|_{X'} \le 1} \left|\int f\chi_E g\,d\mu\right| \le \frac{\varepsilon}{2}, \qquad (n \ge N),$$

which contradicts (4.5). Hence $Y \subset X_a$. ■

Corollary 4.2. *Let X be a Banach function space. If X_a contains the simple functions, then $(X_a)^* = X'$.*

Proof. If $Y = X_a$, then Y is an order ideal (Theorem 3.8) containing the simple functions. The result follows by applying Theorem 4.1 to Y. ■

The space X has absolutely continuous norm if and only if $X = X_a$. Hence, applying Theorem 4.1 (to $Y = X$), we obtain the following result.

Corollary 4.3. *The Banach space dual X^* of a Banach function space X is canonically isometrically isomorphic to the associate space X' if and only if X has absolutely continuous norm.*

Corollary 4.4. *A Banach function space X is reflexive if and only if both X and its associate space X' have absolutely continuous norm.*

Proof. If both X and X' have absolutely continuous norm, then successive applications of Theorem 4.1 give

$$X^{**} = (X^*)^* = (X')^* = (X')' = X''.$$

However, Theorem 2.7 shows that $X'' = X$ so we have $X^{**} = X$. Since all of the identifications are the canonical ones, we conclude that X is reflexive.

Conversely, suppose X is reflexive. Recall (Theorem 2.9) that X' is a closed norm-fundamental subspace of X^*. If X' is a proper subspace of X^*, then the Hahn-Banach theorem provides a nonzero functional $F \in X^{**}$ that annihilates X'. The reflexivity of X allows us to represent F by evaluation at some f in X, in which case

$$\int fg\,d\mu = F(g) = 0,$$

for all $g \in X'$. Since X' is norm-fundamental in X^*, this implies $f = 0$ μ-a.e. and hence F is identically zero. From this contradiction we conclude that $X' = X^*$ and hence from Corollary 4.3 that X has absolutely continuous norm. This information, together with Theorem 2.7 and the fact that X is reflexive, now gives

$$(X')^* = (X^*)^* = X = X'' = (X')'.$$

Applying Corollary 4.3 once more, we deduce that X' also has absolutely continuous norm. ■

A closely related criterion which is sufficient for reflexivity is that the Banach function space have separable dual (Corollary 5.8). This will emerge from the general discussion of separability in the next section.

5. SEPARABILITY

The proofs of the main results on separability (Theorem 5.5 and its corollaries) will require some elementary facts about weak topologies. Let X be a Banach function space and suppose Z is an order ideal of X' containing the simple functions. Then Z contains $(X')_b$ and so it is a norm-fundamental subspace of X^* (Theorems 2.7 and 3.12). The collection of seminorms

$$f \to \left|\int fg\,d\mu\right|, \qquad (g \in Z) \tag{5.1}$$

on X is therefore a separating family which endows X with the structure of a Hausdorff locally convex topological vector space [Ru, pp. 60–66]. The topology in question is called the weak topology on X generated by Z and is denoted by $\sigma(X, Z)$. Note that $f_n \to f$ in the $\sigma(X, Z)$-topology if and only if $\int f_n g\,d\mu \to \int fg\,d\mu$ for all g in Z.

A subset A of X is $\sigma(X, Z)$-*bounded* if it is bounded in each of the seminorms (5.1), that is, if

$$\sup\left\{\left|\int fg\,d\mu\right| : f \in A\right\} < \infty \tag{5.2}$$

for each g in Z.

Lemma 5.1. *Let X be a Banach function space and suppose Z is an order ideal of X' containing the simple functions. Then a subset A of X is $\sigma(X, Z)$-bounded if and only if it is norm-bounded in X.*

Proof. Hölder's inequality (2.6) shows that norm-boundedness implies $\sigma(X, Z)$-boundedness. Conversely, suppose A is $\sigma(X, Z)$-bounded. Then (5.2) holds for each g in Z. Now each f in A defines a linear functional

$$F(g) = \int fg\,d\mu, \qquad (g \in Z)$$

in Z^* (Hölder's inequality), and $\|F\|_{Z^*} = \|f\|_X$ because Z is norm-fundamental

in X^* (Theorem 3.12). But (5.2) asserts that

$$\sup_{f \in A} |F(g)| \le C_g < \infty$$

for each g in Z. Hence, by the uniform boundedness principle,

$$\sup_{f \in A} \|f\|_X = \sup_{f \in A} \|F\|_{Z^*} < \infty,$$

which shows that A is norm-bounded. ■

Theorem 5.2. *Let X be a Banach function space and let Z be an order ideal of X' containing the simple functions. Then X is $\sigma(X, Z)$-complete.*

Proof. Let $(f_n)_{n=1}^\infty$ be a $\sigma(X, Z)$-Cauchy sequence in X. This means that $(\int f_n g\,d\mu)_{n=1}^\infty$ is a Cauchy sequence of scalars for each g in Z, and from this it is apparent that $(f_n)_{n=1}^\infty$ is $\sigma(X, Z)$-bounded. Hence, by Lemma 5.1, there is a constant $M > 0$ such that

$$\|f_n\|_X \le M, \qquad (n = 1, 2, \ldots).$$

For each $n = 1, 2, \ldots$, the measure ν_n defined by

$$\nu_n(E) = \int_E f_n\,d\mu$$

is absolutely continuous with respect to μ (that is, $\nu_n \ll \mu$). Let $(R_N)_{N=1}^\infty$ be an increasing sequence of sets of finite μ-measure whose union is all of R, and fix N. By hypothesis, $(\nu_n(E))_{n=1}^\infty$ is Cauchy and so

$$\nu(E) = \lim_{n\to\infty} \nu_n(E)$$

exists and is finite for every μ-measurable subset E of R_N. By the Hahn-Saks theorem [HS, p. 339], the sequence $(\nu_n)_{n=1}^\infty$, restricted to R_N, is uniformly absolutely continuous with respect to μ, and, on R_N, the set function ν is a measure which is absolutely continuous with respect to μ. It follows, by letting $N \to \infty$, that there is a locally integrable function f_0 (defined μ-a.e. on R) such that

$$\lim_{n\to\infty} \nu_n(E) = \nu(E) = \int_R f_0 \chi_E\,d\mu,$$

whenever $\mu(E) < \infty$.

If g is a simple function with support of finite measure, we therefore have

$$\lim_{n\to\infty} \int_R f_n g\,d\mu = \int_R f_0 g\,d\mu. \tag{5.3}$$

If g also satisfies $\|g\|_{X'} \le 1$, then

$$\left|\int_R f_0 g\,d\mu\right| \le \limsup_{n\to\infty} \|f_n\|_X \|g\|_{X'} \le M.$$

Hence, we see from Theorem 2.7 and the monotone convergence theorem that f_0 belongs to X and $\|f_0\|_X \le M$.

If now g is a bounded function with support of finite measure, then by uniformly approximating g by simple functions we find that (5.3) holds also for all such g.

We need to show that (5.3) holds for all g in Z. Fix such a function g and set

$$\omega_n(E) = \int_E f_n g\,d\mu, \qquad (n = 1, 2, \ldots).$$

The set functions ω_n are finite measures satisfying $\omega_n \ll \mu$. Since $g\chi_E$ belongs to Z for each measurable set E, the hypothesis implies that the sequence $(\omega_n(E))$ converges. Again by the Hahn-Saks theorem, the sequence (ω_n) is uniformly absolutely continuous with respect to μ, that is, $\omega_n(E) \to 0$ uniformly in n as $\mu(E) \to 0$. The measure ω_0 defined by $\omega_0(E) = \int_E f_0 g\,d\mu$ also satisfies $\omega_0 \ll \mu$. Thus, if we define a decreasing sequence of sets E_m by

$$E_m = \{|g| > m\} \cup R_m^c, \qquad (m = 1, 2, \ldots),$$

then, given $\varepsilon > 0$, we may choose N such that

$$|\omega_n(E_N)| < \frac{\varepsilon}{3}, \qquad (n = 0, 1, 2, \ldots).$$

If F_N denotes the complement of E_N, then $g\chi_{F_N}$ is bounded and has support of finite measure. Since (5.3) holds for such functions, we may choose n_0 such that

$$\left|\int_{F_N} f_n g\,d\mu - \int_{F_N} f_0 g\,d\mu\right| < \frac{\varepsilon}{3}, \qquad (n \ge n_0).$$

By splitting the range of integration R into $R = E_N \cup F_N$ we therefore obtain

$$\left|\int_R f_n g\,d\mu - \int_R f_0 g\,d\mu\right| < \frac{\varepsilon}{3} + |\omega_n(E_N)| + |\omega_0(E_N)| < \varepsilon,$$

whenever $n \ge n_0$. This establishes (5.3) for an arbitrary function g in Z. Hence, $f_n \to f_0$ in the $\sigma(X, Z)$ topology and so X is $\sigma(X, Z)$-complete. ∎

Corollary 5.3. *Every Banach function space X is $\sigma(X, X')$-complete.*

Let $\mathscr{A}$ denote the collection of all subsets of R of finite measure, where any

two such sets which differ by a set of μ-measure zero are identified. If

$$d(E, F) = \int |\chi_E - \chi_F|\,d\mu, \qquad (E, F \in \mathscr{A}), \tag{5.4}$$

then $(\mathscr{A}, d)$ is a complete metric space (Exercise 9).

Definition 5.4. A measure μ is said to be *separable* if the corresponding metric space $(\mathscr{A}, d)$ is separable [Ha, p. 168].

Theorem 5.5. *Let X be a Banach function space with underlying measure space (R, μ). Suppose Y is an order ideal of X containing the simple functions. Then Y is separable if and only if Y has absolutely continuous norm and μ is a separable measure.*

Proof. Suppose first that Y is separable and let $(f_n)_{n=1}^{\infty}$ be a dense subset of Y. If Y contains a function f_0 that does not have absolutely continuous norm, then by Proposition 3.2 there is a sequence $E_n \downarrow \varnothing$ μ-a.e. and a positive number ε such that

$$\|f_0 \chi_{E_n}\|_X \ge \varepsilon, \qquad (n = 1, 2, \ldots).$$

Consequently, since X' is norm-fundamental in X^* (Theorem 2.9), there are functions g_n, $(n = 1, 2, \ldots)$, in the unit ball of X' such that

$$\int |f_0 g_n \chi_{E_n}|\,d\mu \ge \frac{\varepsilon}{2}, \qquad (n = 1, 2, \ldots). \tag{5.5}$$

The functions

$$h_n = \operatorname{sgn}(\bar{f}_0)|g_n|\chi_{E_n}, \qquad (n = 1, 2, \ldots)$$

are also in the unit ball of X' so Hölder's inequality (2.6) gives

$$\left|\int f_m h_n\,d\mu\right| \le \|f_m\|_X, \qquad (m, n = 1, 2, \ldots).$$

In particular, the sequences $(\int f_m h_n\,d\mu)_{n=1}^{\infty}$, $(m = 1, 2, \ldots)$, are precompact. By diagonalization we thus obtain a subsequence $(h_{n(k)})_{k=1}^{\infty}$ with the property that $\int f_m h_{n(k)}\,d\mu$ converges as $k \to \infty$, for each $m = 1, 2, \ldots$. Since $(f_m)_{m=1}^{\infty}$ is dense in Y we may therefore use Hölder's inequality to deduce that $\lim_{k\to\infty} \int f h_{n(k)}\,d\mu$ exists for every f in Y. When Y is regarded as a subset of $X'' = X$, this merely asserts that the sequence $(h_{n(k)})_{k=1}^{\infty}$ is $\sigma(X', Y)$-Cauchy in X'. Hence, by Theorem 5.2 (with X' replacing X and Y replacing Z), there exists a function h in X' such that

$$\lim_{k\to\infty} \int f h_{n(k)}\,d\mu = \int f h\,d\mu, \qquad (f \in Y). \tag{5.6}$$

Recall that the functions $h_{n(k)}$ are supported in the decreasing sequence of sets $E_{n(k)}$. Thus, if E is a set of finite measure (so $f = \chi_E$ belongs to Y, by hypothesis) and if E is disjoint from some $E_{n(k)}$, then (5.6) shows that $\int_E h\,d\mu = 0$. Keeping this value of k fixed and allowing E to vary, we see that $h = 0$ μ-a.e. on the complement of $E_{n(k)}$, and then by allowing k to increase we conclude that h vanishes μ-a.e. because $E_{n(k)} \downarrow \varnothing$ μ-a.e. Hence (5.6), applied to f_0, gives

$$0 = \lim_{k\to\infty} \int f_0 h_{n(k)}\,d\mu = \lim_{k\to\infty} \int |f_0 g_{n(k)}| \chi_{E_{n(k)}}\,d\mu.$$

This contradicts (5.5) so we conclude that Y has absolutely continuous norm.

Next, we show that μ is separable. Let $(R_N)_{N=1}^{\infty}$ be an increasing sequence of sets of finite measure whose union is all of R, and let $\mathscr{A}_N$ denote the σ-algebra of all measurable subsets of R_N, $(N = 1, 2, \ldots)$. By hypothesis, the characteristic function of a set in any $\mathscr{A}_N$ belongs to Y. Now Y is separable so any subset of Y is also separable (Exercise 10). Hence, for each $N = 1, 2, \ldots$, there is a countable collection $(\chi_{E_{N,m}})_{m=1}^{\infty}$ of characteristic functions of sets $E_{N,m} \in \mathscr{A}_N$ which is norm-dense in the set of all characteristic functions of sets in $\mathscr{A}_N$. We shall show that the countable family $\mathscr{F}$ consisting of all sets $E_{N,m}$, $(N, m = 1, 2, \ldots)$, is dense in the metric space $(\mathscr{A}, d)$ (cf. (5.4)). This will establish the separability of μ.

Let F be any set in $\mathscr{A}$, that is, any set of finite measure. The dominated convergence theorem implies that $\mu(F \setminus R_N) \to 0$ as $N \to \infty$ since $\mu(F) < \infty$ and $F \cap R_N \uparrow F$. For arbitrary $\varepsilon > 0$, pick N so that

$$\mu(F \setminus R_N) < \frac{\varepsilon}{2}.$$

Now, by Theorem 1.7(v), there is a constant C_N such that

$$\int_{R_N} |f|\,d\mu \le C_N \|f\|_X, \qquad (f \in X). \tag{5.7}$$

By the definition of the family $\mathscr{F}$, we may pick $E \in \mathscr{F}$, $E \subset R_N$, so that

$$\|\chi_E - \chi_{F\cap R_N}\|_X < \frac{\varepsilon}{2C_N}. \tag{5.8}$$

Combining (5.7) and (5.8), we observe that

$$\begin{aligned} d(F, E) &= \int |\chi_F - \chi_E|\,d\mu \\ &= \mu(F \setminus R_N) + \int_{R_N} |\chi_E - \chi_{F\cap R_N}|\,d\mu \\ &< \frac{\varepsilon}{2} + C_N \|\chi_E - \chi_{F\cap R_N}\|_X < \varepsilon. \end{aligned}$$

Hence, $\mathscr{F}$ is dense in $(\mathscr{A}, d)$ and so μ is a separable measure.

Conversely, suppose μ is separable and Y has absolutely continuous norm. Then $Y = X_a = X_b$, by Theorem 4.1. Let $\mathscr{F}_1$ be a countable dense subset of $(\mathscr{A}, d)$, say $\mathscr{F}_1 = \{E_1, E_2, \ldots\}$. Let $(R_N)_{N=1}^{\infty}$ be an increasing sequence of sets of finite measure whose union is R, and let $\mathscr{F}$ denote the countable collection $\{E_j \cap R_N : j, N = 1, 2, \ldots\}$. Finally, let $\mathscr{D}$ denote the class of simple functions of the form

$$f = \sum_{k=1}^{K} r_k \chi_{F_k},$$

where the coefficients are rational (real and imaginary parts rational in the complex case) and the sets F_k belong to $\mathscr{F}$. Clearly $\mathscr{D}$ is countable. We shall show that $\mathscr{D}$ is dense in X_b and hence that $Y\,(= X_b)$ is separable.

Now X_b is the closure of the set of simple functions

$$g = \sum_{n=1}^{m} c_n \chi_{G_n},$$

where the coefficients c_n are arbitrary scalars and the sets G_n are arbitrary sets of finite measure. Since each c_n may be approximated arbitrarily closely by rationals, it is clear that in order to show that $\mathscr{D}$ is dense in X_b it will suffice to show that the characteristic function of an arbitrary set of finite measure can be approximated in the norm of X to any required degree of accuracy by characteristic functions of sets from $\mathscr{F}$.

Suppose then that G has finite measure and let $\varepsilon > 0$. Now χ_G belongs to $X_b = X_a$ and $\chi_{G\cap R_N} \uparrow \chi_G$, so by Proposition 3.6 we may choose N so that

$$\|\chi_G - \chi_{G\cap R_N}\|_X < \frac{\varepsilon}{2}.$$

Hence, it suffices to approximate $\chi_{G\cap R_N}$ by characteristic functions of sets from $\mathscr{F}$. For each n, we may choose $F_n \in \mathscr{F}$ for which $F_n \subset R_N$ and

$$\lim_{n\to\infty} d(G \cap R_N, F_n) = 0. \tag{5.9}$$

This merely asserts that $\chi_{F_n} \to \chi_{G\cap R_N}$ in $L^1(R)$ so we may select a subsequence $\chi_{F_{n(j)}} \to \chi_{G\cap R_N}$ μ-a.e. as $j \to \infty$. But all of these functions are dominated by χ_{R_N}, which belongs to X_a. Hence, the dominated convergence theorem (Proposition 3.6) shows that

$$\|\chi_{F_{n(j)}} - \chi_{G\cap R_N}\|_X \to 0.$$

As we remarked above, this suffices to establish the separability of Y. ■

Corollary 5.6. *A Banach function space X is separable if and only if it has absolutely continuous norm and its underlying measure μ is separable.*

Corollary 5.7. *Suppose $X_a = X_b$. Then X_a is separable if and only if μ is separable.*

Corollary 5.8. *If the dual space X^* of a Banach function space X is separable, then X is reflexive.*

Proof. Any Banach space with separable dual is itself separable [DS, p. 65]. Hence X is separable. Furthermore, any subset of X^* must also be separable. Hence the associate space X' is separable. But then both X and X' have absolutely continuous norm, by Theorem 5.5. In that case, Corollary 4.4 shows that X is reflexive. ∎

EXERCISES AND FURTHER RESULTS FOR CHAPTER 1

1. Let $(S_N)_{N=1}^{\infty}$ be a sequence of disjoint subsets in a measure space (R, μ), each of finite measure, and with union equal to R. For each f and g in $\mathcal{M}_0(R, \mu)$, let

$$d(f, g) = \sum_{N=1}^{\infty} 2^{-N} \frac{1}{\mu(S_N)} \int_{S_N} \frac{|f-g|}{1+|f-g|} \, d\mu.$$

Then $(\mathcal{M}_0, d)$ is a complete metric space and $f_n \to f$ in $(\mathcal{M}_0, d)$ if and only if $f_n \to f$ in measure on sets of finite measure.

2. (a) If (R, μ) is non-atomic, then a function f in $L^\infty(R, \mu)$ is of absolutely continuous norm if and only if $f = 0$ μ-a.e.

(b) The absolutely continuous part ℓ_a^∞ of the space ℓ^∞ of bounded sequences is the space c_0 of sequences that converge to zero at infinity.

(c) For which measure spaces is L_a^∞ equal to L^∞?

3. Let (R, μ) be the infinite interval $(0, \infty)$ with Lebesgue measure. For each f in $\mathcal{M}_0^+$, let

$$\rho(f) = \int_0^1 f(x)\,dx + \operatorname*{ess\,sup}_{1 \le x < \infty} f(x).$$

Then ρ is a function norm. If $X = X(\rho)$, determine X_a and X_b, and hence show that each of the inclusions $\{0\} \subset X_a \subset X_b \subset X$ is proper. What is the associate space X'?

4. Let (R, μ) be the real line with Lebesgue measure. For each f in $\mathcal{M}_0^+$, let

$$\rho(f) = \max\left\{ \int_{-\infty}^{\infty} f(x)\,dx,\ \operatorname*{ess\,sup}_{-\infty < x < \infty} f(x) \right\}.$$

Then ρ is a function norm and hence $L^1 \cap L^\infty$ is a Banach function space (cf. also Chapter II.6). If $X = X(\rho) = L^1 \cap L^\infty$, determine X_a and X_b.

5. Let (R, μ) be as in the preceding exercise and let

$$\rho(f) = \sup_{|E|=1} \int_E f(x)\,dx,$$

where the supremum is taken over all subsets E of R of measure 1. Then ρ is a function norm. If $X = X(\rho)$, determine X_a and X_b. The associate norm of ρ is the function norm described in the preceding exercise (G. G. Gould [1]; cf. also Chapter II.6).

6. Let (R, μ) be the interval $[0, 1]$ with Lebesgue measure. For each f in $\mathcal{M}_0^+$, let

$$\rho(f) = \sup_{0<h\le 1} \frac{1}{h} \int_0^h f(x)\,dx$$

(B. I. Korenblyum, S. G. Krein & B. Ya. Levin [1]).

(a) ρ is a function norm.

(b) If $f, g \in \mathcal{M}_0^+$ and g is decreasing, then

$$\int_0^1 f(x)g(x)\,dx \le \rho(f) \int_0^1 g(x)\,dx$$

(I. P. Natanson [1], [2]).

(c) For each g in $\mathcal{M}_0^+$, let

$$\tilde{g}(x) = \operatorname*{ess\,sup}_{x \le y \le 1} g(y).$$

Then $\tilde{g}$ is decreasing and $g(x) \le \tilde{g}(x)$ a.e. in $[0, 1]$.

(d) For each g in $\mathcal{M}_0^+$, let

$$\lambda(g) = \int_0^1 \tilde{g}(x)\,dx.$$

Then λ is a function norm whose associate norm is ρ. (W. A. J. Luxemburg & A. C. Zaanen [3]).

7. Prove Lemma 2.6 using the uniform boundedness principle.

8. Let K be a subset of the absolutely continuous part X_a of a Banach function space X. Then K is said to be of *uniformly absolutely continuous norm* if, whenever $E_n \to \varnothing$ μ-a.e. and $\varepsilon > 0$, there is an integer N such that

$$f \in K, \quad n \ge N \quad \Rightarrow \quad \|f\chi_{E_n}\|_X < \varepsilon.$$

(W. A. J. Luxemburg & A. C. Zaanen [1]).

Let K be an arbitrary subset of X_a. Then K is sequentially precompact in X if and only if K is sequentially precompact in $\mathcal{M}_0$ and is of uniformly absolutely continuous norm.

The proof of the necessity is straightforward. The sufficiency may be established as follows:

(i) Let $(f_n)_{n=1}^{\infty} \subset K$. Then some subsequence $(f_{n(j)})_{j=1}^{\infty}$ converges in $\mathcal{M}_0$ and pointwise μ-a.e. to a function f in $\mathcal{M}_0$.

(ii) Let $\varepsilon > 0$. Then there is a subset P of R with $0 < \mu(P) < \infty$ such that

$$\|g\chi_{R \setminus P}\|_X < \varepsilon$$

for all g in K. Furthermore $f\chi_{R\setminus P}$ belongs to X and $\|f\chi_{R\setminus P}\|_X < \varepsilon$.

(iii) For each $j = 1, 2, \ldots$, let

$$E_j = \left\{ x \in P : |f(x) - f_{n(j)}(x)| > \frac{\varepsilon}{\|\chi_P\|_X} \right\}.$$

Then there exists J such that $\|g\chi_{E_j}\|_X < \varepsilon$ for all g in K and all $j \geq J$. Furthermore, $f\chi_{E_j}$ belongs to X and $\|f\chi_{E_j}\|_X < \varepsilon$ for all $j \geq J$.

(iv) $f\chi_{P \setminus E_j}$ belongs to X, and $f_{n(j)} \to f$ in X.

9. With the notation of Section 5 (preceding Definition 5.4), show that $(\mathscr{A}, d)$ is a complete metric space.

10. (a) If (X, d) is a separable metric space, then any subspace of (X, d) is also separable.

(b) If X is a Banach space whose dual space X^* is separable, then X is separable ([DS, p. 65]).

(c) If X is a Banach function space whose second dual X^{**} is separable, then X is reflexive. This result fails for arbitrary Banach spaces (R. C. James [1]).

11. Let $\rho : \mathscr{M}_0^+(R, \mu) \to [0, \infty]$ be a functional which satisfies properties (P1), (P2), (P4), and (P5) of Definition I.1.1. Let L^ρ denote the subspace of $\mathscr{M}_0(R, \mu)$ consisting of those f for which the norm $\rho(|f|)$ is finite. The following conditions are equivalent:

(i) L^ρ is norm-complete;

(ii) $f_n \in \mathscr{M}_0^+, (n = 1, 2, \ldots) \quad \Rightarrow \quad \rho\left(\sum_{n=1}^\infty f_n\right) \leq \sum_{n=1}^\infty \rho(f_n)$;

(iii) $\left.\begin{array}{l} f_n \in \mathscr{M}_0^+, (n = 1, 2, \ldots), \\ \text{and } \sum_{n=1}^\infty \rho(f_n) < \infty \end{array}\right\} \Rightarrow \quad \rho\left(\sum_{n=1}^\infty f_n\right) < \infty.$

(iv) $\left.\begin{array}{l} f_n \in \mathscr{M}_0^+, (n = 1, 2, \ldots), \\ \text{and } \sum_{n=1}^\infty \rho(f_n) < \infty \end{array}\right\} \Rightarrow \left\{\begin{array}{l} \text{there exists } f \geq 0 \text{ in} \\ L^\rho \text{ with } f_n \leq f \text{ for all } n. \end{array}\right.$

Such norms are said to satisfy the *Riesz-Fischer property*. (I. Halperin & W. A. J. Luxemburg [1]; W. A. J. Luxemburg [3]; W. A. J. Luxemburg & A. C. Zaanen [2]).

12. Let $\rho : \mathscr{M}_0^+(R, \mu) \to [0, \infty]$ satisfy properties (P1), (P2), (P4), and (P5) of Definition 1.1. The functional ρ is said to have the *weak Fatou property* if

$$\left.\begin{array}{c} f_n \in \mathscr{M}_0^+, (n = 1, 2, \ldots), f_n \uparrow f \ \mu\text{-a.e.}, \\ \text{and } \sup_n \rho(f_n) < \infty \end{array}\right\} \Rightarrow \quad \rho(f) < \infty.$$

The Fatou property implies the weak-Fatou property, but not conversely (W. A. J. Luxemburg & A. C. Zaanen [2, p. 150]). The weak Fatou property implies the Riesz-Fischer property (I. Amemiya [1]), but not conversely (W. A. J. Luxemburg & A. C. Zaanen [2, p. 150]).

13. (I. Amemiya [1]) If ρ has the weak Fatou property, then there is a constant γ with $0 < \gamma \leq 1$ such that

$$f_n \in \mathscr{M}_0^+, (n = 1, 2, \ldots), f_n \uparrow f \ \mu\text{-a.e.} \quad \Rightarrow \quad \gamma\rho(f) \leq \lim_{n \to \infty} \rho(f_n).$$

The same conclusion holds for sequences $f_n \to f$ in measure on sets of finite measure.

14. (G. G. Lorentz, W. A. J. Luxemburg) Suppose ρ has the Riesz-Fischer property. Then ρ has the weak Fatou property if and only if there is a constant γ with $0 < \gamma \leq 1$ such that $\gamma\rho \leq \rho'' \leq \rho$ (in which case γ may be taken to be the constant in the preceding exercise). In particular, ρ has the Lorentz-Luxemburg property $\rho = \rho''$ if and only if ρ has the Fatou property (W. A. J. Luxemburg [1, 3], G. G. Lorentz [4, 6], I. Halperin & W. A. J. Luxemburg [1]).

NOTES FOR CHAPTER 1

Banach function spaces have a history almost as long as their ancestors, the Lebesgue L^p-spaces. That the L^p-spaces themselves are insufficient to describe the mapping properties of operators such as the conjugate-function operator was realized as early as 1925 by A. N. Kolmogorov [1] (weak-type estimates) and 1928 by A. Zygmund [1] and E. C. Titchmarsh [2], [3] ($L \log L$ estimates). Estimates of the latter type involve functionals of the form $\int \Phi(|f|)$ for convex increasing Φ, which had been investigated in some detail in the early twenties by W. H. Young. W. Orlicz [1], [3], in the early thirties, took the decisive step by furnishing the resulting function spaces with norms, thereby incorporating them in the general framework of the normed spaces introduced by S. Banach. Interestingly, G. H. Hardy & J. E. Littlewood [2] in 1930 had provided an explicit norm on the space $L \log L$, of a different type to that envisaged by Orlicz, which lay dormant for some twenty years and eventually resurfaced in the general theory of Lorentz spaces introduced in 1950 by G. G. Lorentz [1], [2].

Attempts to unify the Orlicz and Lorentz spaces in an axiomatic framework were made in the early fifties by I. Halperin [1], I. Halperin & H. W. Ellis [1], and G. G. Lorentz & D. G. Wertheim [1], the latter based on the Köthe-Toeplitz theory of sequence spaces which dates back also to the early thirties (cf. G. Köthe [1]). Such efforts culminated in the theory (as presented in the text) of Banach function norms and Banach function spaces, which were introduced and developed by W. A. J. Luxemburg [1] in 1955. Several important features of the theory (Theorem 2.7, for example) were obtained independently by G. G. Lorentz [6]. Thorough investigations of the properties of Banach function norms and semi-norms are made in the series of papers by W. A. J. Luxemburg & A. C. Zaanen [2] and in the book by A. C. Zaanen [1]. The theory of Banach lattices (or Riesz spaces) is treated in W. A. J. Luxemburg & A. C. Zaanen [4]; there is a massive literature on these spaces from the point of view of the geometry of Banach spaces, for which J. Lindenstrauss & L. Tzafriri [1] is an excellent source.

2 Rearrangement-Invariant Banach Function Spaces

For finite sequences $(a_i)_{i=1}^n$ and $(b_i)_{i=1}^n$ of nonnegative numbers, it is natural to say that (b_i) is a rearrangement of (a_i) if $b_i = a_{\sigma(i)}$, $(i = 1, 2, \ldots, n)$, for some permutation σ of the numbers $1, 2, \ldots, n$. In more general measure spaces, it is tempting to replace the notion of permutation with that of measure-preserving transformation and accordingly, for nonnegative measurable functions f and g, to say that g is a "rearrangement" of f if $g = f \circ \sigma$ for some measure-preserving transformation σ. This concept, while valid, is not broad enough for our purposes. Symmetry fails, for example. Thus, g may be a rearrangement of f in this sense without f being a rearrangement of g (cf. Example 7.7).

We adopt a broader definition: nonnegative functions f and g will be *rearrangements* of one another (or, in more precise terminology, will be deemed *equimeasurable*) if their distribution functions coincide (cf. Definition 1.2). This notion, which is clearly symmetric, also allows for equimeasurability of functions defined on different measure spaces. In particular, for each measurable function f, it enables the construction of a decreasing right-continuous function f^* on the interval $(0, \infty)$ that is equimeasurable with f. The function f^* is called the *decreasing rearrangement* of f (cf. Definition 1.5). Constructing f^* from a measurable function f is thus analogous to rearranging the terms of a finite sequence in decreasing order.

The mapping $f \to f^*$ may be viewed as a crude representation, in terms of which phenomena occurring on a general measure space may be viewed in terms of their counterparts on the specific measure space $(0, \infty)$ with Lebesgue measure. Much of the present chapter deals with properties of this representation.

Distribution functions and decreasing rearrangements are introduced in Section 1. Section 2 describes the kinds of measure spaces that admit a reasonable theory of rearrangements. The *maximal function* f^{**} and the *Hardy-Littlewood-Pólya relation* are defined in Section 3. These lead to the introduction of the *rearrangement-invariant spaces*, that is, the Banach function spaces in which functions and their equimeasurable rearrangements have the same norm. The *Luxemburg representation theorem* (Theorem 4.7) allows rearrangement-invariant spaces on more general measure spaces to be studied in the context of the interval $(0, \infty)$ with Lebesgue measure.

Two special families of rearrangement-invariant spaces—the *Lorentz* Λ *and* M *spaces*—are introduced in Section 5. Certain properties of rearrangement-invariant spaces depend only on the values their norms assign to characteristic functions of measurable sets E, that is, on the fundamental function $\varphi(t) = \|\chi_E\|$, where E is a set of measure t, $(0 < t < \infty)$. Modulo one or two minor technicalities, the smallest and the largest of all rearrangement-invariant spaces with given fundamental function φ are the Lorentz spaces $\Lambda(\varphi)$ and $M(\varphi)$, respectively (Corollary 5.14).

The Lorentz spaces $L^1 \cap L^\infty$ and $L^1 + L^\infty$, introduced in Section 6, are respectively the smallest and the largest of all rearrangement-invariant spaces, and play a special role in the theory of interpolation of operators.

Chapter II concludes with a brief discussion of measure-preserving transformations (Section 7). The main result is the theorem of G. G. Lorentz and J. V. Ryff (Theorem 7.5), to the effect that a nonnegative measurable function f on a finite measure space can be obtained from its decreasing rearrangement by composition with a suitable measure-preserving transformation. The converse is false (Example 7.7).

1. DISTRIBUTION FUNCTIONS AND DECREASING REARRANGEMENTS

As in the previous chapter, (R, μ) denotes a totally σ-finite measure space.

Definition 1.1. The *distribution function* μ_f of a function f in $\mathcal{M}_0 = \mathcal{M}_0(R, \mu)$ is given by

$$\mu_f(\lambda) = \mu\{x \in R : |f(x)| > \lambda\}, \qquad (\lambda \geq 0). \tag{1.1}$$

Observe that μ_f depends only on the absolute value $|f|$ of the function f, and μ_f may assume the value $+\infty$.

Definition 1.2. Two functions $f \in \mathcal{M}_0(R, \mu)$ and $g \in \mathcal{M}_0(S, \nu)$ are said to be *equimeasurable* if they have the same distribution function, that is, if $\mu_f(\lambda) = \nu_g(\lambda)$ for all $\lambda \geq 0$.

Proposition 1.3. *Suppose* f, g, f_n, $(n = 1, 2, \ldots)$, *belong to* $\mathcal{M}_0(R, \mu)$ *and let* a *be any nonzero scalar. The distribution function* μ_f *is nonnegative, decreasing, and right-continuous on* $[0, \infty)$. *Furthermore,*

$$|g| \leq |f| \ \mu\text{-a.e.} \quad \Rightarrow \quad \mu_g \leq \mu_f; \tag{1.2}$$

$$\mu_{af}(\lambda) = \mu_f(\lambda/|a|), \qquad (\lambda \geq 0); \tag{1.3}$$

$$\mu_{f+g}(\lambda_1 + \lambda_2) \leq \mu_f(\lambda_1) + \mu_g(\lambda_2), \qquad (\lambda_1, \lambda_2 \geq 0); \tag{1.4}$$

$$|f| \leq \liminf_{n \to \infty} |f_n| \ \mu\text{-a.e.} \quad \Rightarrow \quad \mu_f \leq \liminf_{n \to \infty} \mu_{f_n}; \tag{1.5}$$

in particular,

$$|f_n| \uparrow |f| \ \mu\text{-a.e.} \quad \Rightarrow \quad \mu_{f_n} \uparrow \mu_f.$$

Proof. It is clear that μ_f is nonnegative and decreasing (in the wide sense: $0 \leq \lambda_1 < \lambda_2 \ \Rightarrow \ \mu_f(\lambda_1) \geq \mu_f(\lambda_2)$). To establish right-continuity, let $E(\lambda) = \{x : |f(x)| > \lambda\}$, $(\lambda \geq 0)$, and fix $\lambda_0 \geq 0$. The sets $E(\lambda)$ increase as λ decreases, and

$$E(\lambda_0) = \bigcup_{\lambda > \lambda_0} E(\lambda) = \bigcup_{n=1}^{\infty} E\left(\lambda_0 + \frac{1}{n}\right).$$

Hence, by the monotone convergence theorem,

$$\mu_f\left(\lambda_0 + \frac{1}{n}\right) = \mu\left(E\left(\lambda_0 + \frac{1}{n}\right)\right) \uparrow \mu(E(\lambda_0)) = \mu_f(\lambda_0),$$

and this establishes the right-continuity.

Properties (1.2) and (1.3) are immediate consequences of Definition 1.1. Property (1.4) follows from the obvious fact that if $|f(x) + g(x)| > \lambda_1 + \lambda_2$, then either $|f(x)| > \lambda_1$ or $|g(x)| > \lambda_2$. To establish (1.5), fix $\lambda \geq 0$ and let

$$E = \{x : |f(x)| > \lambda\}, \qquad E_n = \{x : |f_n(x)| > \lambda\}, \qquad (n = 1, 2, \ldots).$$

Clearly, $E \subset \bigcup_{m=1}^{\infty} \bigcap_{n>m} E_n$. Hence,

$$\mu\left(\bigcap_{n>m} E_n\right) \leq \inf_{n>m} \mu(E_n) \leq \sup_m \inf_{n>m} \mu(E_n) = \liminf_{n \to \infty} \mu(E_n)$$

for each $m = 1, 2, \ldots$. But $\bigcap_{n>m} E_n$ increases with m, so an appeal to the monotone convergence theorem gives

$$\mu(E) \le \mu\left(\bigcup_{m=1}^{\infty} \bigcap_{n>m} E_n\right) = \lim_{m\to\infty} \mu\left(\bigcap_{n>m} E_n\right) \le \liminf_{n\to\infty} \mu(E_n).$$

This establishes the first of the assertions in (1.5). The second is an immediate consequence of the first. ■

Example 1.4. It will be worthwhile to formally compute the distribution function μ_f of a nonnegative simple function f. Suppose

$$f(x) = \sum_{j=1}^{n} a_j \chi_{E_j}(x), \tag{1.6}$$

where the sets E_j are pairwise disjoint subsets of R with finite μ-measure and $a_1 > a_2 > \cdots > a_n > 0$. If $\lambda \ge a_1$, then clearly $\mu_f(\lambda) = 0$. However, if $a_2 \le \lambda < a_1$, then $f(x)$ exceeds λ precisely on the set E_1, and so $\mu_f(\lambda) = \mu(E_1)$. Similarly, if $a_3 \le \lambda < a_2$, then $f(x)$ exceeds λ precisely on $E_1 \cup E_2$, and so $\mu_f(\lambda) = \mu(E_1 \cup E_2) = \mu(E_1) + \mu(E_2)$. In general, we have

$$\mu_f(\lambda) = \sum_{j=1}^{n} m_j \chi_{[a_{j+1}, a_j)}(\lambda), \qquad (\lambda \ge 0), \tag{1.7}$$

where

$$m_j = \sum_{i=1}^{j} \mu(E_i), \qquad (j = 1, 2, \ldots, n) \tag{1.8}$$

and a_{n+1} is defined to be zero (see Figure 1).

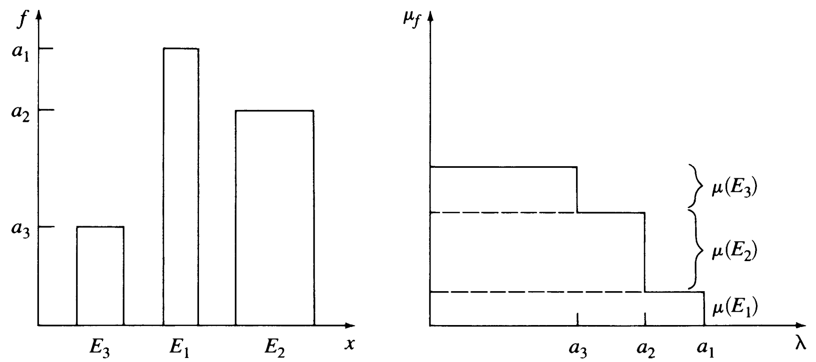


Figure 1. Graphs of $f(x)$ and $\mu_f(\lambda)$.

Definition 1.5. Suppose f belongs to $\mathscr{M}_0(R, \mu)$. The *decreasing rearrangement* of f is the function f^* defined on $[0, \infty)$ by

$$f^*(t) = \inf\{\lambda : \mu_f(\lambda) \le t\}, \qquad (t \ge 0). \tag{1.9}$$

We use here the convention that $\inf \varnothing = \infty$. Thus, if $\mu_f(\lambda) > t$ for all $\lambda \ge 0$, then $f^*(t) = \infty$. Also, if (R, μ) is a finite measure space, then the distribution function μ_f is bounded by $\mu(R)$ and so $f^*(t) = 0$ for all $t \ge \mu(R)$. In this case, we may regard f^* as a function defined on the interval $[0, \mu(R))$. Notice also that if μ_f happens to be continuous and strictly decreasing, then f^* is simply the inverse of μ_f on the appropriate interval. In fact, for general f, if we first form the distribution function μ_f and then form the distribution function m_{μ_f} of μ_f (with respect to Lebesgue measure m on $[0, \infty)$) we obtain precisely the decreasing rearrangement f^*. This is an immediate consequence of the identities

$$f^*(t) = \sup\{\lambda : \mu_f(\lambda) > t\} = m_{\mu_f}(t), \qquad (t \ge 0), \tag{1.10}$$

which follow from (1.9), the fact that μ_f is decreasing, and the definition of the distribution function.

Examples 1.6. (a) Now we compute the decreasing rearrangement of the simple function f given by (1.6). Referring to (1.9) and Figure 1, we see that $f^*(t) = 0$ if $t \ge m_3$. Also, if $m_3 > t \ge m_2$, then $f^*(t) = a_3$, and if $m_2 > t \ge m_1$, then $f^*(t) = a_2$, and so on. Hence,

$$f^*(t) = \sum_{j=1}^{n} a_j \chi_{[m_{j-1}, m_j)}(t), \qquad (t \ge 0), \tag{1.11}$$

where we have taken $m_0 = 0$.

Geometrically, we are merely rearranging the *vertical* blocks in the graph of f in decreasing order to obtain the decreasing rearrangement f^* (see Figure 2); the values of f^* at the jumps are determined by the right continuity (Proposition 1.7).

(b) It is sometimes more useful to section functions into *horizontal* blocks rather than vertical ones. Thus, the simple function f in (1.6) may be represented also as follows:

$$f(x) = \sum_{k=1}^{n} b_k \chi_{F_k}(x), \tag{1.12}$$

where the coefficients b_k are positive and the sets F_k each have finite measure and form an increasing sequence $F_1 \subset F_2 \subset \cdots \subset F_n$. Comparison with (1.6)

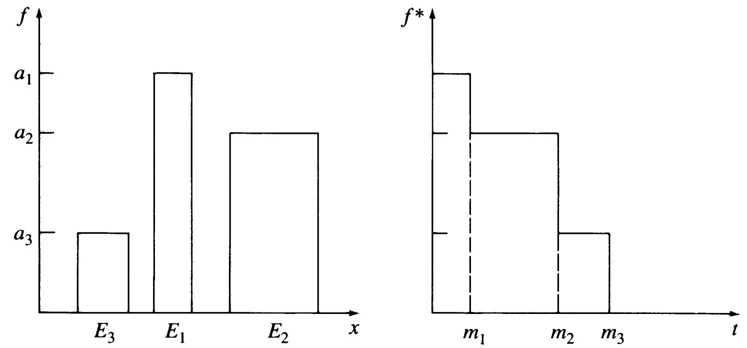


Figure 2. Graphs of f and f^*.

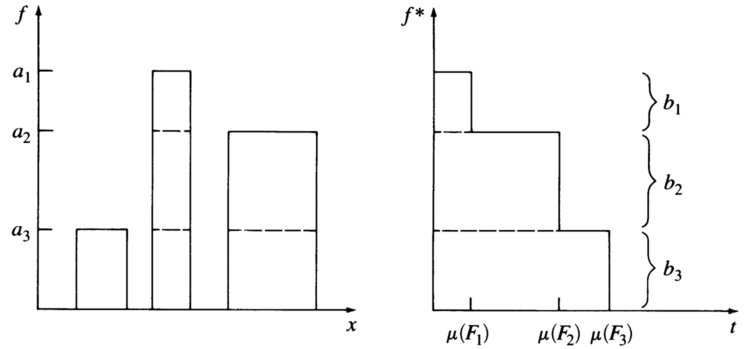


Figure 3. Graphs of f and f^*.

shows that

$$b_k = a_k - a_{k+1}, \qquad F_k = \bigcup_{j=1}^{k} E_j, \qquad (k = 1, 2, \ldots, n).$$

In this case, the decreasing rearrangement is viewed as being formed by sliding the blocks in each horizontal layer to form a single larger block positioned with its left-hand end against the vertical axis (see Figure 3). Thus

$$f^* = \sum_{k=1}^{n} b_k \chi_{[0, \mu(F_k))}. \tag{1.13}$$

(c) Let $f(x) = 1 - e^{-x}$, $(0 < x < \infty)$. The distribution function m_f (with respect to Lebesgue measure m on $(0, \infty)$) is infinite for $0 \le \lambda < 1$, and equal to

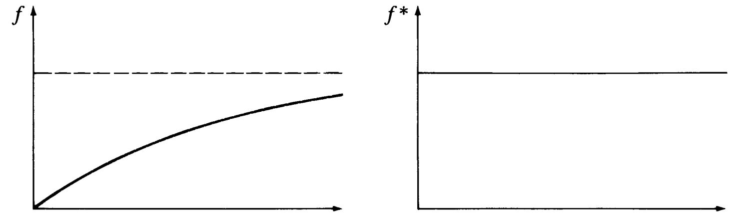


Figure 4. Graphs of $f(x) = 1 - e^{-x}$ and $f^*(t)$.

zero for all $\lambda \ge 1$. Hence $f^*(t) = 1$ for all $t \ge 0$ (cf. Figure 4). This example shows that a considerable amount of information may be lost in passing to the decreasing rearrangement. Such information, however, is irrelevant as far as L^p-norms (or any other rearrangement-invariant norms) are concerned. Thus, the L^p-norms of f and f^* are both infinite when $1 \le p < \infty$, and the L^∞-norms are both equal to 1.

Proposition 1.7. *Suppose f, g, and f_n, $(n = 1, 2, \ldots)$, belong to $\mathcal{M}_0(R, \mu)$ and let a be any scalar. The decreasing rearrangement f^* is a nonnegative, decreasing, right-continuous function on $[0, \infty)$. Furthermore,*

$$|g| \le |f| \ \mu\text{-a.e.} \quad \Rightarrow \quad g^* \le f^*; \tag{1.14}$$

$$(af)^* = |a| f^*; \tag{1.15}$$

$$(f + g)^*(t_1 + t_2) \le f^*(t_1) + g^*(t_2), \qquad (t_1, t_2 \ge 0); \tag{1.16}$$

$$|f| \le \liminf_{n\to\infty} |f_n| \ \mu\text{-a.e.} \quad \Rightarrow \quad f^* \le \liminf_{n\to\infty} f_n^*; \tag{1.17}$$

in particular,

$$|f_n| \uparrow |f| \ \mu\text{-a.e.} \quad \Rightarrow \quad f_n^* \uparrow f^*;$$

$$f^*(\mu_f(\lambda)) \le \lambda, \qquad (\mu_f(\lambda) < \infty); \qquad \mu_f(f^*(t)) \le t, \qquad (f^*(t) < \infty); \tag{1.18}$$

$$f \text{ and } f^* \text{ are equimeasurable}; \tag{1.19}$$

$$(|f|^p)^* = (f^*)^p, \qquad (0 < p < \infty). \tag{1.20}$$

Proof. That f^* is nonnegative, decreasing, and right-continuous follows from Proposition 1.3 and the fact that f^* is itself a distribution function (cf. (1.10)). The properties (1.14), (1.15), and (1.17) are immediate consequences of their counterparts in Proposition 1.3 and the definition of the decreasing rearrangement.

For property (1.18), fix $\lambda \geq 0$ and suppose $t = \mu_f(\lambda)$ is finite. Then (1.9) gives

$$f^*(\mu_f(\lambda)) = f^*(t) = \inf\{\lambda' : \mu_f(\lambda') \leq t = \mu_f(\lambda)\} \leq \lambda,$$

which establishes the first part of (1.18). For the second part, fix $t \geq 0$ and suppose $\lambda = f^*(t)$ is finite. By (1.9), there is a sequence $\lambda_n \downarrow \lambda$ with $\mu_f(\lambda_n) \leq t$, so the right-continuity of μ_f (Proposition 1.3) gives

$$\mu_f(f^*(t)) = \mu_f(\lambda) = \lim_{n\to\infty} \mu_f(\lambda_n) \leq t.$$

This establishes (1.18).

Returning to (1.16), we may assume that $\lambda = f^*(t_1) + g^*(t_2)$ is finite since otherwise there is nothing to prove. Let $t = \mu_{f+g}(\lambda)$. Then by the triangle inequality and the second of the inequalities in (1.18) we have

$$\begin{aligned} t &= \mu\{x : |f(x) + g(x)| > f^*(t_1) + g^*(t_2)\} \\ &\leq \mu\{x : |f(x)| > f^*(t_1)\} + \mu\{x : |g(x)| > g^*(t_2)\} \\ &= \mu_f(f^*(t_1)) + \mu_g(g^*(t_2)) \\ &\leq t_1 + t_2. \end{aligned}$$

This shows in particular that t is finite. Hence, using the first of the inequalities in (1.18) and the fact that $(f+g)^*$ is decreasing, we obtain

$$\begin{aligned} (f+g)^*(t_1+t_2) \leq (f+g)^*(t) &= (f+g)^*(\mu_{f+g}(\lambda)) \\ &\leq \lambda = f^*(t_1) + g^*(t_2), \end{aligned}$$

and this establishes (1.16).

For an arbitrary function f in $\mathcal{M}_0$, we can find a sequence of nonnegative simple functions f_n, $(n = 1, 2, \ldots)$, such that $f_n \uparrow |f|$. It is clear (cf. Example 1.6(a)) that for each n the functions f_n and f_n^* are equimeasurable, that is,

$$\mu_{f_n}(\lambda) = m_{f_n^*}(\lambda), \qquad (\lambda \geq 0). \tag{1.21}$$

But $f_n \uparrow |f|$ and $f_n^* \uparrow f^*$ (by 1.17) so property (1.5), applied to each of the distribution functions in (1.21), shows that

$$\mu_f(\lambda) = m_{f^*}(\lambda), \qquad (\lambda \geq 0). \tag{1.22}$$

Hence, f and f^* are equimeasurable, as asserted by (1.19).

Finally, from (1.22) we have

$$\mu_{|f|^p}(\lambda) = \mu_f(\lambda^{1/p}) = m_{f^*}(\lambda^{1/p}) = m_{(f^*)^p}(\lambda), \qquad (\lambda \geq 0).$$

Passing to the decreasing rearrangements by means of (1.9), we obtain (1.20). ■

The next result gives alternative descriptions of the L^p-norm in terms of the distribution function and the decreasing rearrangement.

Proposition 1.8. *Let $f \in \mathcal{M}_0$. If $0 < p < \infty$, then*

$$\int_R |f|^p \, d\mu = p \int_0^\infty \lambda^{p-1} \mu_f(\lambda)\, d\lambda = \int_0^\infty f^*(t)^p \, dt. \tag{1.23}$$

Furthermore, in the case $p = \infty$,

$$\operatorname*{ess\,sup}_{x \in R} |f(x)| = \inf\{\lambda : \mu_f(\lambda) = 0\} = f^*(0). \tag{1.24}$$

Proof. In view of (1.5), (1.17), and the monotone convergence theorem, it will suffice to prove (1.23) for an arbitrary nonnegative simple function f. With f written in the form (1.6), we saw that its decreasing rearrangement f^* is given by (1.11). But then it is clear from (1.8) that

$$\int |f|^p \, d\mu = \sum_{j=1}^n a_j^p \mu(E_j) = \sum_{j=1}^n a_j^p m([m_{j-1}, m_j)) = \int_0^\infty (f^*)^p \, dm.$$

Similarly, using the expressions (1.6) and (1.7) for f and its distribution function μ_f, we have

$$\begin{aligned} p \int_0^\infty \lambda^{p-1} \mu_f(\lambda)\, d\lambda &= p \sum_{j=1}^n m_j \int_{a_{j+1}}^{a_j} \lambda^{p-1}\, d\lambda = \sum_{j=1}^n (a_j^p - a_{j+1}^p) m_j \\ &= \sum_{j=1}^n a_j^p \mu(E_j) = \int |f|^p \, d\mu, \end{aligned}$$

where the third equality follows from (1.8) and a summation by parts.

This establishes (1.23). The proof of (1.24) is straightforward and we omit it. ■

2. AN INEQUALITY OF HARDY AND LITTLEWOOD

While the decreasing rearrangement does not necessarily preserve sums or products of functions, there are nevertheless some basic inequalities that govern these processes. These will be derived in the present section (for products) and in the next section (for sums).

The starting point is an elementary inequality due to G. H. Hardy and J. E. Littlewood. The inequality involves finite sequences $(a_1, a_2, \ldots, a_n)$ and $(b_1, b_2, \ldots, b_n)$ of nonnegative real numbers, and asserts that

$$\sum_{j=1}^n a_j b_j \leq \sum_{j=1}^n a_j^* b_j^*, \tag{2.1}$$

where $(a_j^*)_{j=1}^n$ denotes the sequence of elements a_j arranged in decreasing order (and similarly for $(b_j^*)_{j=1}^n$). In other words, the sum attains its maximum value when the terms in each sequence are rearranged in decreasing order.

It will often be convenient to regard such a sequence $(a_j)_{j=1}^n$ as a simple function $f = \sum_{j=1}^n a_j\chi_{[j-1,j)}$ defined on the interval $[0, \infty)$. In that case, as Example 1.6(a) shows, the decreasing rearrangement f^* of f is just the simple function $f^* = \sum_{j=1}^n a_j^*\chi_{[j-1,j)}$ corresponding to the rearranged sequence $(a_j^*)_{j=1}^n$. Viewed in this way, the inequality (2.1) will be seen as a special case of the more general integral inequality (2.3) to be established in Theorem 2.2 below.

Lemma 2.1. *Let g be a nonnegative simple function on (R, μ) and let E be an arbitrary μ-measurable subset of R. Then*

$$\int_E g\, d\mu \le \int_0^{\mu(E)} g^*(s)\, ds. \tag{2.2}$$

Proof. We express g in the form of (1.12). Thus

$$g(x) = \sum_{j=1}^n b_j\chi_{F_j}(x),$$

where $F_1 \subset F_2 \subset \cdots \subset F_n$ and $b_j > 0$, $(j = 1, 2, \ldots, n)$. Using the formulation (1.13) for the decreasing rearrangement, we obtain

$$\begin{aligned}\int_E g\, d\mu &= \sum_{j=1}^n b_j\mu(E \cap F_j) \le \sum_{j=1}^n b_j \cdot \min(\mu(E), \mu(F_j)) \\ &= \sum_{j=1}^n b_j \int_0^{\mu(E)} \chi_{(0,\mu(F_j))}(s)\, ds = \int_0^{\mu(E)} g^*(s)\, ds. \quad \blacksquare\end{aligned}$$

Theorem 2.2. (G. H. Hardy & J. E. Littlewood). *If f and g belong to $\mathcal{M}_0 = \mathcal{M}_0(R, \mu)$, then*

$$\int_R |fg|\, d\mu \le \int_0^\infty f^*(s)g^*(s)\, ds. \tag{2.3}$$

Proof. Since f^* and g^* depend only on the absolute values of f and g, it is enough to establish (2.3) for nonnegative functions f and g. But then, in view of (1.17) and the monotone convergence theorem, there is no loss of generality in supposing f and g to be simple. In that case, we may write

$$f(x) = \sum_{j=1}^m a_j\chi_{E_j}(x),$$

where $E_1 \subset E_2 \subset \cdots \subset E_m$ and $a_j > 0$, $(j = 1, 2, \ldots, m)$ (cf. (1.12)). Then (cf. (1.13))

$$f^*(t) = \sum_{j=1}^m a_j\chi_{[0,\mu(E_j))}(t).$$

Hence, by Lemma 2.1,

$$\begin{aligned}\int |fg|\, d\mu &= \sum_{j=1}^m a_j \int_{E_j} g\, d\mu \le \sum_{j=1}^m a_j \int_0^{\mu(E_j)} g^*(s)\, ds \\ &= \int_0^\infty \sum_{j=1}^m a_j\chi_{[0,\mu(E_j))}(s)g^*(s)\, ds \\ &= \int_0^\infty f^*(s)g^*(s)\, ds,\end{aligned}$$

as desired. $\blacksquare$

An immediate consequence of the Hardy-Littlewood inequality (2.3) is that

$$\int_R |f\tilde{g}|\, d\mu \le \int_0^\infty f^*(t)g^*(t)\, dt \tag{2.4}$$

for every function $\tilde{g}$ on R equimeasurable with g. When f and g are finite sequences, as in (2.1), it is clear that equality is attained in (2.4) for a suitable $\tilde{g}$. Indeed, such a $\tilde{g}$ may evidently be obtained from g by means of a suitable permutation of the underlying point-set $\{1, 2, \ldots, n\}$. The situation is not as straightforward, however, for more general measure spaces and so it becomes necessary to isolate those classes of measure spaces with the desired properties.

Much of theory can be developed under a weaker requirement than the attainment of equality in (2.4). What is needed is that the *supremum* of the integrals on the left of (2.4) coincide with the value on the right; such measure spaces will be called *resonant*. If the supremum is in fact attained, then the measure space will be called *strongly resonant* (cf. Definition 2.3 below).

Naturally, we shall need to know which measure spaces possess these resonance properties. We shall see (Theorem 2.7) that a totally σ-finite measure space is resonant if and only if it is nonatomic or completely atomic with all atoms having equal measure. The strongly resonant measure spaces are precisely the resonant ones of finite measure (Theorem 2.6).

Definition 2.3. A totally σ-finite measure space (R, μ), is said to be *resonant* if, for each f and g in $\mathcal{M}_0(R, \mu)$, the identity

$$\int_0^\infty f^*(t)g^*(t)\, dt = \sup \int_R |f\tilde{g}|\, d\mu \tag{2.5}$$

holds, where the supremum is taken over all functions $\tilde{g}$ on R equimeasurable with g.

Similarly, (R, μ) is said to be *strongly resonant* if, for each pair of functions f and g in $\mathcal{M}_0(R, \mu)$, there exists a function $\tilde{g}$ on R equimeasurable with g such that

$$\int_0^\infty f^*(t) g^*(t)\, dt = \int_R |f\tilde{g}|\, d\mu. \tag{2.6}$$

Examples 2.4. (a) It is clear from (2.4) that every strongly resonant measure space is resonant. The converse, however, is false. Indeed, we shall see in Theorem 2.7 that the interval $[0, \infty)$ with Lebesgue measure is resonant, but the following example shows that it is not strongly resonant. Let f be the function described in Example 1.6(c). Thus, $f(x) = 1 - e^{-x}$, $(0 \le x < \infty)$, and $f^* \equiv 1$ (see Figure 4). Let $g = g^* = \chi_{[0,1)}$. If $\tilde{g}$ is equimeasurable with g, then $|\tilde{g}|$ is the characteristic function of some subset E of $(0, \infty)$ of measure 1. But then, as is evident from Figure 4,

$$\int_R |f\tilde{g}|\, d\mu = \int_E (1 - e^{-x})\, dx < 1 = \int_0^\infty f^*(t) g^*(t)\, dt,$$

so equality is never attained.

(b) Here is an example of a measure space that is not resonant. Let (R, μ) consist of two atoms a and b with $\mu(a) = 1$ and $\mu(b) = 2$. Let $f = \chi_b$ and $g = \chi_a$. Then $f^* = \chi_{[0,2)}$ and $g^* = \chi_{[0,1)}$ so $\int f^* g^* = 1$. But $\int |f\tilde{g}|\, d\mu = 0$ for every $\tilde{g}$ equimeasurable with g because f and $\tilde{g}$ must have disjoint supports. This argument shows in fact that no measure space containing two atoms with positive but unequal measures can be resonant.

Lemma 2.5. *Let (R, μ) be a finite nonatomic measure space. Suppose f belongs to $\mathcal{M}_0(R, \mu)$ and let t be any number satisfying $0 \le t \le \mu(R)$. Then there is a measurable set E_t, with $\mu(E_t) = t$, such that*

$$\int_{E_t} |f|\, d\mu = \int_0^t f^*(s)\, ds. \tag{2.7}$$

Moreover, the sets E_t can be constructed so as to increase with t:

$$0 \le s \le t \le \mu(R) \quad \Rightarrow \quad E_s \subset E_t. \tag{2.8}$$

Proof. We consider two separate cases, according to whether the number t does or does not lie in the range of the distribution function μ_f of f.

Suppose first that there exists $\alpha > 0$ for which $\mu_f(\alpha) = t$. In that case, it

follows from (1.9) that

$$f^*(t) = \inf\{\lambda : \mu_f(\lambda) = t\},$$

and then the right-continuity of μ_f gives $\mu_f(f^*(t)) = t$. Equivalently, the set $E_t = \{x : |f(x)| > f^*(t)\}$ has measure $\mu(E_t) = t$, and this shows that the distribution function of $f\chi_{E_t}$ is given by

$$\mu\{x \in E_t : |f(x)| > \lambda\} = \begin{cases} \mu_f(\lambda), & \lambda > f^*(t), \\ t, & 0 \le \lambda \le f^*(t). \end{cases}$$

Also, the distribution function of $f^*\chi_{[0,t)}$ is given by

$$m\{s \in [0, t] : f^*(s) > \lambda\} = \begin{cases} m_{f^*}(\lambda), & \lambda > f^*(t) \\ t, & 0 \le \lambda \le f^*(t). \end{cases}$$

But f and f^* are equimeasurable, so $\mu_f = m_{f^*}$. Hence, $f\chi_{E_t}$ and $f^*\chi_{[0,t)}$ are equimeasurable. In particular, by Proposition 1.8, their integrals are the same:

$$\int_{E_t} |f|\, d\mu = \int_R |f\chi_{E_t}|\, d\mu = \int_0^\infty f^*(s)\chi_{[0,t]}(s)\, ds = \int_0^t f^*(s)\, ds.$$

Hence, (2.7) holds. Notice that the sets E_t increase with t.

Next we consider the case where t is not in the range of μ_f (cf. Figure 5). Observe that since $\mu(R)$ is finite and $f \in \mathcal{M}_0$, the dominated convergence theorem gives

$$0 = \mu\{x : |f(x)| = \infty\} = \lim_{n\to\infty} \mu\{x : |f(x)| > n\} = \lim_{n\to\infty} \mu_f(n).$$

Hence, since μ_f is decreasing,

$$\lim_{\lambda\to\infty} \mu_f(\lambda) = 0. \tag{2.9}$$

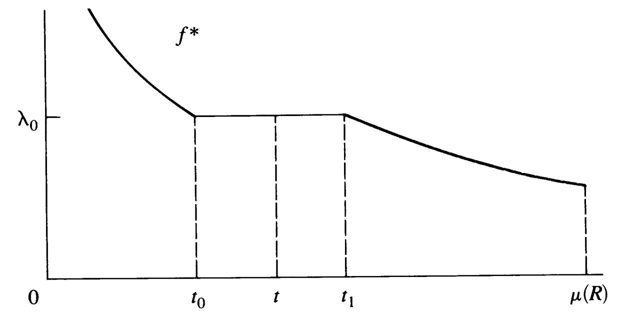


Figure 5

Let $\lambda_0 = f^*(t)$ and suppose first that $\lambda_0 > 0$. Since $t > 0$, it follows from (1.9) and (2.9) that λ_0 is finite. Also, since t is not in the range of μ_f, we see from (1.18) and (1.9) that $\mu_f(\lambda_0) < t < \mu_f(\lambda)$ for $0 < \lambda < \lambda_0$. Thus, if t_1 denotes the left-hand limit $\mu_f(\lambda_0 -)$ (which exists because μ_f is decreasing), then

$$t_0 \equiv \mu_f(\lambda_0) < t \leq \mu_f(\lambda_0 -) = t_1. \tag{2.10}$$

This, together with (1.9), shows in fact that

$$f^*(s) = \lambda_0, \qquad (t_0 \leq s < t_1) \tag{2.11}$$

(cf. Figure 5).

We claim that

$$t_1 = \mu\{x : |f(x)| \geq \lambda_0\}. \tag{2.12}$$

To see this, we need only express the set on the right as the intersection of the decreasing sequence of sets

$$F_n = \left\{x : |f(x)| > \lambda_0 - \frac{1}{n}\right\}, \qquad (n = 1, 2, \ldots),$$

observe that $\mu(R)$ is finite, and use the dominated convergence theorem to obtain

$$\mu\{x : |f(x)| \geq \lambda_0\} = \lim_{n\to\infty} \mu(F_n) = \lim_{n\to\infty} \mu_f\left(\lambda_0 - \frac{1}{n}\right) = \mu_f(\lambda_0 -).$$

It follows from (2.10) and (2.12) that the set $G = \{x : |f(x)| = \lambda_0\}$ has measure $t_1 - t_0$. Since (R, μ) is nonatomic, we may thus select a subset F_t of G with measure $t - t_0$ (notice that there is considerable freedom in this choice if $t < t_1$). The set E_t defined by the disjoint union

$$E_t = \{x : |f(x)| > \lambda_0\} \cup F_t \tag{2.13}$$

then has measure $\mu_f(\lambda_0) + (t - t_0) = t$, as desired. Moreover,

$$\int_{E_t} |f| \, d\mu = \int_{\{|f| > \lambda_0\}} |f| \, d\mu + \int_{F_t} |f| \, d\mu \tag{2.14}$$

Now, by (2.10), the value t_0 does lie in the range of μ_f. The first part of the proof shows therefore that the first of the integrals on the right of (2.14) has value $\int_0^{t_0} f^*(s) \, ds$. Also, since $|f| = \lambda_0$ on F_t, the second integral has value

$$\lambda_0 \mu(F_t) = \lambda_0(t - t_0) = \int_{t_0}^{t} f^*(s) \, ds,$$

by (2.11). Combining these results with (2.14), we obtain (2.7).

In the remaining case where $\lambda_0 = 0$, we obtain in place of (2.10) the inequality $\mu\{x : |f(x)| > 0\} \equiv t_0 < t$. In that case, we may choose F_t disjoint from the support of f, with measure $\mu(F_t) = t - t_0$. Defining E_t again by (2.13), and observing that $f^*(s)$ vanishes for $s \geq t_0$, we obtain (2.7) as before:

$$\int_{E_t} |f| \, d\mu = \int_0^{t_0} f^*(s) \, ds = \int_0^{t} f^*(s) \, ds.$$

We observed in the first part of the proof that the sets E_t satisfy (2.8) for all those values of t in the range of μ_f. For the other values of t, it is elementary (cf. Exercise 17) to show that the sets F_t in the second part of the proof may be chosen so as to increase with t (in each of the intervals $(t_0, t_1]$ in the complement of the range of μ_f). Given this construction, a moment's reflection will now show that (2.8) holds as desired. ∎

We are now in a position to characterize the strongly resonant measure spaces (Theorem 2.6) and the resonant measure spaces (Theorem 2.7).

Theorem 2.6. *A totally σ-finite measure space (R, μ) is strongly resonant if and only if it is a finite measure space of one of the following two types:*

(i) *nonatomic;*
(ii) *completely atomic, with all atoms having equal measure.*

Proof. First, we establish the sufficiency. Thus, given f and g in $\mathscr{M}_0(R, \mu)$, we need to find a function $\tilde{g}$ equimeasurable with g such that

$$\int_R |f\tilde{g}| \, d\mu = \int_0^\infty f^*(t) g^*(t) \, dt. \tag{2.15}$$

It is clear that we may assume f and g nonnegative.

If (R, μ) has finite measure and is of type (ii), then a suitable $\tilde{g}$ may be constructed from g merely by permuting the underlying (finitely many) atoms. Indeed, the desired permutation is just the composition of the inverse of the permutation that takes f to f^* with the permutation that takes g to g^*. Hence, we need only consider the case where (R, μ) is nonatomic and has finite measure.

Let $(g_n)_{n=1}^\infty$ be an increasing sequence of nonnegative simple functions such that $g_n \uparrow g$ μ-a.e. We shall construct an increasing sequence $(\tilde{g}_n)_{n=1}^\infty$ of nonnegative simple functions such that $\tilde{g}_n$ is equimeasurable with g_n and

$$\int_R |f\tilde{g}_n| \, d\mu = \int_0^\infty f^*(t) \tilde{g}_n^*(t) \, dt, \qquad (n = 1, 2, \ldots). \tag{2.16}$$

By (1.5), the function $\tilde{g}(x) = \lim_{n\to\infty} \tilde{g}_n(x)$ is then equimeasurable with g, and so (2.15) follows from (2.16), (1.17), and the monotone convergence theorem. It remains therefore only to construct the simple functions $\tilde{g}_n$ with the desired properties.

Fix $n = 1, 2, \ldots$, and, to save on notation, let $h = g_n$. As in Example 1.6(b), we may write

$$h = \sum_{j=1}^{m} b_j \chi_{F_j},$$

where $F_1 \subset F_2 \subset \cdots \subset F_m$ and $b_j > 0$, $(j = 1, 2, \ldots, m)$. Then, by Lemma 2.5, there exist sets $E_1 \subset E_2 \subset \cdots \subset E_m$ with $\mu(E_j) = \mu(F_j)$ and

$$\int_{E_j} f\,d\mu = \int_0^{\mu(F_j)} f^*(t)\,dt, \qquad (j = 1, 2, \ldots, m). \tag{2.17}$$

If we set

$$\tilde{h} = \sum_{j=1}^{m} b_j \chi_{E_j},$$

then, as in (1.13), the decreasing rearrangements of h and $\tilde{h}$ are given by

$$h^* = \sum_{j=1}^{m} b_j \chi_{[0,\mu(F_j))} = (\tilde{h})^*.$$

Hence, h and $\tilde{h}$ are equimeasurable. Moreover, from (2.17),

$$\begin{aligned}\int f\tilde{h}\,d\mu &= \sum_{j=1}^{m} b_j \int_{E_j} f\,d\mu = \sum_{j=1}^{m} b_j \int_0^{\mu(F_j)} f^*(t)\,dt \\ &= \int_0^\infty \sum_{j=1}^{m} b_j \chi_{[0,\mu(F_j))}(t) f^*(t)\,dt \\ &= \int_0^\infty f^*(t)h^*(t)\,dt,\end{aligned}$$

so if we set $\tilde{g}_n = \tilde{h}$ (recall that $h = g_n$), then $\tilde{g}_n$ is equimeasurable with g_n and, as the last identity shows, (2.16) holds. Finally, since the g_n increase with n, it is clear from the last assertion of Lemma 2.5 that the sets E_j in (2.17) may be constructed so that they, and hence the functions $\tilde{g}_n$, increase with n. As we remarked above, this shows that (R, μ) is strongly resonant.

Observe that if (R, μ) is resonant, then the argument in Example 2.4(b) shows that the atoms, if any exist, all have the same measure. Essentially the same argument shows also that (R, μ) cannot be a mixture of nontrivial atomic and nonatomic parts. Hence, a resonant measure space must be of types (i) or (ii) above. But if (R, μ) is strongly resonant, then $\mu(R)$ must be finite, for otherwise we could construct a counterexample along the lines of Example 2.4(a). We leave the details as an exercise. ■

Theorem 2.7. *A totally σ-finite measure space (R, μ) is resonant if and only if it is of one of the following two types:*

(i) *nonatomic;*
(ii) *completely atomic, with all atoms having equal measure.*

Proof. The necessity was established in the proof of the previous theorem. For the sufficiency, given f and g in $\mathcal{M}_0(R, \mu)$, we need to establish (2.5). Clearly, we may assume f and g nonnegative. We may also assume that the left-hand side of (2.5) is positive, since otherwise there is nothing to prove. Hence, it will suffice to show that, for each α satisfying

$$0 < \alpha < \int_0^\infty f^*(t)g^*(t)\,dt, \tag{2.18}$$

there exists a nonnegative function $\tilde{g}$ equimeasurable with g such that

$$\alpha < \int_R f\tilde{g}\,d\mu. \tag{2.19}$$

Since (R, μ) is totally σ-finite, there is a sequence $(R_n)_{n=1}^\infty$ of sets of finite measure whose union is all of R. Choose sequences $(f_n)_{n=1}^\infty$ and $(g_n)_{n=1}^\infty$ of simple functions, with f_n and g_n supported in R_n, $(n = 1, 2, \ldots)$, such that $0 \le f_n \uparrow f$ and $0 \le g_n \uparrow g$. By property (1.17), the monotone convergence theorem, and (2.18), there exists N such that

$$\alpha < \int_0^\infty f_N^*(t)g_N^*(t)\,dt. \tag{2.20}$$

Now the measure space (R_N, μ) is of the type considered in the previous theorem, and hence is strongly resonant. We can therefore find a nonnegative function h on R_N, equimeasurable with $g\chi_{R_N}$, such that

$$\int_{R_N} fh\,d\mu = \int_0^{\mu(R_N)} (f\chi_{R_N})^*(t)(g\chi_{R_N})^*(t)\,dt.$$

But $f\chi_{R_N} \ge f_N$ and $g\chi_{R_N} \ge g_N$ so the last identity and (2.20) combine to give

$$\alpha < \int_R fh\,d\mu, \tag{2.21}$$

where we have extended h to all of R by defining h to be zero outside R_N. Thus, if we set

$$\tilde{g} = h\chi_{R_N} + g\chi_{R\setminus R_N},$$

then $\tilde{g}$ is equimeasurable with g and $\tilde{g} \geq h$. Hence, by (2.21),

$$\alpha < \int_R fh\,d\mu \leq \int_R f\tilde{g}\,d\mu.$$

This establishes (2.19) and hence completes the proof. ■

The following result is an immediate consequence of Theorems 2.6 and 2.7.

Corollary 2.8. *A totally σ-finite measure space (R, μ) is strongly resonant if and only if it is resonant and $\mu(R)$ is finite.*

3. AN ELEMENTARY MAXIMAL FUNCTION

When g is the characteristic function of a set E of positive measure t, the Hardy-Littlewood inequality (2.3) reduces to

$$\frac{1}{\mu(E)}\int_E |f|\,d\mu \leq \frac{1}{t}\int_0^t f^*(s)\,ds, \qquad (f \in \mathcal{M}_0). \tag{3.1}$$

Hence, the average of $|f|$ over any set of measure t is dominated by the corresponding average of f^* over the interval $(0, t)$. Notice that the latter average is also maximal among all averages of f^* taken over sets of measure t (this follows directly from the fact that f^* is decreasing or from the special case of (3.1) in which (R, μ) is taken to be $[0, \infty)$ with Lebesgue measure and $f = f^*$). For this reason, the function on the right of (3.1) is called a *maximal function.* Since it will figure prominently in the subsequent analysis, we introduce a separate notation as follows.

Definition 3.1. Let f belong to $\mathcal{M}_0(R, \mu)$. Then f^{**} will denote the *maximal function* of f^* defined by

$$f^{**}(t) = \frac{1}{t}\int_0^t f^*(s)\,ds, \qquad (t > 0). \tag{3.2}$$

Some elementary properties of the *maximal operator* $f \to f^{**}$ are listed below.

Proposition 3.2. *Suppose f, g, and f_n, $(n = 1, 2, \ldots)$, belong to $\mathcal{M}_0$, and let a be any scalar. Then f^{**} is nonnegative, decreasing, and continuous on $(0, \infty)$. Furthermore, the following properties hold:*

$$f^{**} \equiv 0 \quad \Leftrightarrow \quad f = 0 \ \mu\text{-a.e;} \tag{3.3}$$

$$f^* \leq f^{**}; \tag{3.4}$$

$$|g| \leq |f| \ \mu\text{-a.e.} \quad \Rightarrow \quad g^{**} \leq f^{**}; \tag{3.5}$$

$$(af) = |a| f^{**}; \tag{3.6}$$

$$|f_n| \uparrow |f| \ \mu\text{-a.e.} \quad \Rightarrow \quad f_n^{**} \uparrow f^{**}. \tag{3.7}$$

Proof. Since f^* is decreasing, it follows from (3.2) that $f^{**}(t)$ is finite for any one value of t if and only if it is finite for all $t > 0$. In other words, the function f^{**} is either everywhere finite or everywhere infinite. In either case, it is nonnegative (if we include the value $+\infty$) and continuous.

The properties (3.3), (3.5), (3.6), and (3.7) are more-or-less immediate consequences of their counterparts for f^* in Proposition 1.7. The property (3.4) follows directly from the fact that f^* is decreasing. Thus,

$$f^{**}(t) = \frac{1}{t}\int_0^t f^*(s)\,ds \geq f^*(t)\frac{1}{t}\int_0^t ds = f^*(t).$$

Finally, f^* is decreasing so $f^*(v) \leq f^*(tv/s)$ if $0 < t \leq s$. Hence,

$$f^{**}(s) = \frac{1}{s}\int_0^s f^*(v)\,dv \leq \frac{1}{s}\int_0^s f^*\left(\frac{tv}{s}\right)dv = \frac{1}{t}\int_0^t f^*(u)\,du = f^{**}(t),$$

and so f^{**} is decreasing. ■

Now let us return to the Hardy-Littlewood inequality (3.1). When E is allowed to vary over all possible sets of measure t, we shall need to know whether the values of the integrals on the left come arbitrarily close to, or attain, the fixed value $f^{**}(t)$ on the right.

Proposition 3.3. *Let (R, μ) be a totally σ-finite measure space and let t be any positive number in the range of μ. Suppose f belongs to $\mathcal{M}_0(R, \mu)$.*
(a) *If (R, μ) is resonant, then*

$$f^{**}(t) = \frac{1}{t}\sup\left\{\int_E |f|\,d\mu : \mu(E) = t\right\}. \tag{3.8}$$

(b) *If (R, μ) is strongly resonant, then there is a subset E of R with $\mu(E) = t$ such that*

$$f^{**}(t) = \frac{1}{t}\int_E |f|\,d\mu. \tag{3.9}$$

Proof. Since t is in the range of μ, there is a measurable subset F of R with measure $\mu(F) = t$. Let $g = \chi_F$, so $g^* = \chi_{[0,t)}$. But a function $\tilde{g}$ can be equimeasurable with g if and only if $|\tilde{g}|$ is μ-a.e. equal to the characteristic function of some set E with measure $\mu(E) = \mu(F) = t$. With this observation, it is clear that (3.8) and (3.9) are immediate consequences of (2.5) and (2.6) respectively. ■

Observe that (3.8) implies a certain subadditivity property of the maximal operator $f \to f^{**}$. Thus,

$$(f+g)^{**}(t) \le f^{**}(t) + g^{**}(t), \tag{3.10}$$

provided t is in the range of μ and (R, μ) is resonant.

Now if (R, μ) is nonatomic, Theorem 2.7 shows that it is certainly resonant. If the space also has infinite measure, then the range of μ is the entire interval $[0, \infty)$ and so (3.10) holds for all $t > 0$. It is not hard to see that the same conclusion holds also for nonatomic measure spaces of finite measure. Indeed, if $t_0 = \mu(R)$ is finite, then (3.10) holds already for all t satisfying $0 < t \le t_0$. In this case, every f^* vanishes outside the interval $[0, t_0)$, so if $t > t_0$, then

$$f^{**}(t) = \frac{1}{t}\int_0^{t_0} f^*(s)\,ds = \frac{t_0}{t} f^{**}(t_0).$$

Hence, $f^{**}(t)$ satisfies (3.10) because $f^{**}(t_0)$ does. We have therefore shown that (3.10) holds for all $t > 0$ whenever (R, μ) is nonatomic.

The subadditivity of the maximal operator for arbitrary measure spaces can now be derived from the preceding result for nonatomic measure spaces by appealing to the so-called *method of retracts*. This procedure enables us to embed an arbitrary totally σ-finite measure space (R, μ) into a nonatomic measure space $(\bar{R}, \bar{\mu})$. As we shall see in Theorem 3.4, the subadditivity of the maximal operator on (R, μ) is then an easy consequence of the subadditivity of the maximal operator on $(\bar{R}, \bar{\mu})$.

The construction of $(\bar{R}, \bar{\mu})$ is as follows. If (R, μ) is an arbitrary totally σ-finite measure space, then R can be expressed as a disjoint union

$$R = R_0 \cup \left(\bigcup_j A_j\right),$$

where R_0 is atom-free and each A_j is an atom of finite positive measure (of which there are at most countably many because (R, μ) is σ-finite). Let

$$\bar{R} = R_0 \cup \left(\bigcup_j I_j\right),$$

where the sets I_j are disjoint subintervals of the real line satisfying $m(I_j) = \mu(A_j)$ for each j (should R_0 itself happen to be a subset of the real line, let us agree that the intervals I_j belong to a different copy of the line: the point is that the I_j's must be disjoint from R_0 as well as from one another).

A subset of $\bar{R}$ will be deemed measurable if its intersection with R_0 is μ-measurable and its intersection with each I_j is Lebesgue measurable. Clearly, the collection of all such measurable sets is a σ-algebra, so if $\bar{\mu}$ is defined on this σ-algebra in the obvious way, namely

$$\bar{\mu}(E) = \mu(E \cap R_0) + \sum_j m(E \cap I_j),$$

then $(\bar{R}, \bar{\mu})$ is evidently a nonatomic totally σ-finite measure space.

For each f in $\mathcal{M}_0(R, \mu)$, let $\mathscr{E}_1 f$ be the function in $\mathcal{M}_0(\bar{R}, \bar{\mu})$ which coincides with f on R_0 and which assumes the constant value on each I_j equal to the value of f on A_j. Then f and $\mathscr{E}_1 f$ are equimeasurable, and so

$$(\mathscr{E}_1 f)^* = f^*, \qquad (\mathscr{E}_1 f)^{**} = f^{**}. \tag{3.11}$$

Theorem 3.4. *Let (R, μ) be a totally σ-finite measure space and suppose f and g belong to $\mathcal{M}_0(R, \mu)$. Then*

$$(f+g)^{**}(t) \le f^{**}(t) + g^{**}(t), \qquad (0 < t < \infty). \tag{3.12}$$

Proof. The mapping $\mathscr{E}_1$ is linear so $\mathscr{E}_1(f+g) = \mathscr{E}_1 f + \mathscr{E}_1 g$. But $(\bar{R}, \bar{\mu})$ is nonatomic so

$$\begin{aligned}[\mathscr{E}_1(f+g)]^{**}(t) &= [\mathscr{E}_1 f + \mathscr{E}_1 g]^{**}(t) \\ &\le [\mathscr{E}_1 f]^{**}(t) + [\mathscr{E}_1 g]^{**}(t),\end{aligned}$$

because of the observation made above that (3.10) holds for all $t > 0$ for nonatomic measure spaces. The inequality (3.12) follows immediately from this one by applying (3.11). ■

Although the operation $f \to f^*$ satisfies the somewhat weaker relation (1.16), it is not subadditive (take $f = \chi_{[0,1)}$ and $g = \chi_{[1,2)}$, for example). Instead, the maximal function f^{**} will often play the decisive role. This is particularly true in connection with the order structure of the rearrangement-invariant spaces to be introduced in the next section. There, it is not the pointwise ordering $f_1^* \le f_2^*$ that is crucial so much as the weaker relation $f_1^{**} \le f_2^{**}$ between the two maximal functions.

Definition 3.5. If f_1 and f_2 belong to $\mathcal{M}_0(R, \mu)$, we write $f_1 \prec f_2$ if $f_1^{**} \le f_2^{**}$, that is, if

$$\int_0^t f_1^*(s)\,ds \le \int_0^t f_2^*(s)\,ds \tag{3.13}$$

for all $t > 0$.

The relation $\prec$, which is known as the *Hardy-Littlewood-Pólya relation*, will often be exploited in conjunction with the following result.

Proposition 3.6 (Hardy's lemma). *Let ξ_1 and ξ_2 be nonnegative measurable functions on $(0, \infty)$ and suppose*

$$\int_0^t \xi_1(s)\, ds \leq \int_0^t \xi_2(s)\, ds \tag{3.14}$$

for all $t > 0$. Let η be any nonnegative decreasing function on $(0, \infty)$. Then

$$\int_0^\infty \xi_1(s)\eta(s)\, ds \leq \int_0^\infty \xi_2(s)\eta(s)\, ds. \tag{3.15}$$

Proof. A simple argument using the monotone convergence theorem shows that it is enough to prove the result for η a nonnegative decreasing step-function. In that case, η may be expressed in the form

$$\eta(s) = \sum_{j=1}^n a_j \chi_{(0,t_j)}(s),$$

where the coefficients a_j are positive and $0 < t_1 < \cdots < t_n$. Using (3.14), we obtain

$$\int_0^\infty \xi_1 \eta\, ds = \sum_{j=1}^n a_j \int_0^{t_j} \xi_1\, ds \leq \sum_{j=1}^n a_j \int_0^{t_j} \xi_2\, ds = \int_0^\infty \xi_2 \eta\, ds,$$

which establishes (3.15). ■

Recall that in our discussion of the method of retracts we defined a mapping $\mathscr{E}_1$ from $\mathscr{M}_0(R, \mu)$ into $\mathscr{M}_0(\bar{R}, \bar{\mu})$. We now define an "inverse" operator from $\mathscr{M}_0(\bar{R}, \bar{\mu})$ into the measurable functions on (R, μ) by setting

$$\mathscr{E}_2(\bar{f}) = \bar{f}\chi_{R_0} + \sum_j \left[\frac{1}{m(I_j)} \int_{I_j} \bar{f}\, dm\right] \chi_{A_j} \tag{3.16}$$

for each $\bar{f}$ in $\mathscr{M}_0(\bar{R}, \bar{\mu})$. Note that $\mathscr{E}_2(\bar{f})$ does not belong to $\mathscr{M}_0(R, \mu)$ unless $\bar{f}$ is integrable on each of the intervals I_j. It is clear however that

$$\mathscr{E}_2\mathscr{E}_1 f = f, \tag{3.17}$$

for every f in $\mathscr{M}_0(R, \mu)$.

In the other direction, it is not true in general that $\mathscr{E}_1\mathscr{E}_2\bar{f}$ and $\bar{f}$ coincide since $\bar{f}$ need not be constant on the intervals I_j. For the same reason, there is no pointwise order relation between these functions or between their decreasing rearrangements. We can assert, however, that

$$\mathscr{E}_1\mathscr{E}_2\bar{f} \prec \bar{f}, \tag{3.18}$$

for every $\bar{f}$ in $\mathscr{M}_0(\bar{R}, \bar{\mu})$ (such that $\bar{f}$ is integrable on each I_j). This will follow from the next result because $\mathscr{E}_1\mathscr{E}_2$ is an averaging operator of the kind described below and $(\bar{R}, \bar{\mu})$ is nonatomic, hence resonant by Theorem 2.7.

Proposition 3.7. *Let (R, μ) be a resonant measure space and let $(E_j)_{j\in J}$ be a countable collection of pairwise disjoint subsets of R, each with finite positive measure. Let $E = R \setminus \bigcup_j E_j$. Suppose f belongs to $\mathscr{M}_0(R, \mu)$ and that f is integrable on each E_j. Let*

$$Af = f\chi_E + \sum_{j\in J} \left(\frac{1}{\mu(E_j)} \int_{E_j} f\, d\mu\right) \chi_{E_j}. \tag{3.19}$$

Then $Af \prec f$.

Proof. Suppose first that J has only one element so Af is of the form

$$Af = f\chi_E + \left(\frac{1}{\mu(E_1)} \int_{E_1} f\, d\mu\right) \chi_{E_1},$$

where $0 < \mu(E_1) < \infty$ and $E = R \setminus E_1$. To obtain the desired result $Af \prec f$, we need to show that

$$\int_0^t (Af)^*(s)\, ds \leq \int_0^t f^*(s)\, ds, \tag{3.20}$$

for all $t > 0$.

Suppose first that $0 < t < \infty$ and t lies in the range of μ. Let F be any subset of R with $\mu(F) = t$ and let $t_0 = \mu(F \cap E_1)$. Then

$$\int_F |Af|\, d\mu = \int_{F\cap E} |f|\, d\mu + t_0 \left|\frac{1}{\mu(E_1)} \int_{E_1} f\, d\mu\right|. \tag{3.21}$$

To estimate the last term, we write $f = f\chi_{E_1}$ on E_1 and use (3.1) and the fact that $(f\chi_{E_1})^{**}$ is decreasing to obtain

$$t_0 \left|\frac{1}{\mu(E_1)} \int_{E_1} f\, d\mu\right| \leq t_0 (f\chi_{E_1})^{**}(\mu(E_1)) \leq \int_0^{t_0} (f\chi_{E_1})^*(s)\, ds. \tag{3.22}$$

But (R, μ) is resonant so the measure space $(E_1, \mu | E_1)$ is strongly resonant (Corollary 2.8) because $\mu(E_1) < \infty$. The number t_0 is certainly in the range of $\mu | E_1$ because $\mu(F \cap E_1) = t_0$. Hence, by Proposition 3.3(b), there is a subset G of E_1 with $\mu(G) = t_0$ such that

$$\int_0^{t_0} (f\chi_{E_1})^*(s)\,ds = \int_G |f\chi_{E_1}|\,d\mu = \int_G |f|\,d\mu.$$

Combining this with (3.22) and (3.21), we obtain

$$\int_F |Af|\,d\mu \le \int_{F\cap E} |f|\,d\mu + \int_G |f|\,d\mu. \tag{3.23}$$

But $F \cap E$ and G are disjoint and

$$\mu((F\cap E)\cup G) = \mu(F\cap E) + \mu(G) = (t - t_0) + t = t,$$

so applying (3.1) to the right-hand side of (3.23) we obtain

$$\int_F |Af|\,d\mu \le \int_0^t f^*(s)\,ds.$$

Finally, taking the supremum over all sets F of measure t and applying Proposition 3.3(a), we obtain (3.20), at least for all t in the range of μ. Since (R, μ) is of one of the types described in Theorem 2.7, however, it is clear that (3.20) must then hold for all $t > 0$.

This establishes the proposition in the case where the index set J contains only a single element. The corresponding result for a finite index set J is derived from this one by successive iterations. For if $J = \{1, 2, \ldots, N\}$, say, then with

$$R_n = R \setminus \bigcup_{m=1}^n E_m$$

and

$$A_n f = f\chi_{R_n} + \sum_{m=1}^n \left(\frac{1}{\mu(E_m)} \int_{E_m} f\,d\mu\right)\chi_{E_m},$$

we have

$$A_n f \prec A_{n-1} f \prec \cdots \prec A_1 f \prec f.$$

Finally, suppose J is countably infinite, say $J = \{1, 2, \ldots\}$. Since $|Af| \le A(|f|)$, it is enough to establish $Af \prec f$ for f nonnegative. In that case, the functions

$$f_n = f\chi_E + \sum_{m=1}^n f\chi_{E_m}, \qquad (n = 1, 2, \ldots),$$

satisfy $0 \le f_n \uparrow f$ μ-a.e. and $0 \le Af_n \uparrow Af$ μ-a.e. But then

$$Af_n = A_n f_n \prec f_n \prec f, \qquad (n = 1, 2, \ldots),$$

so applying (3.7) to the left-hand side we obtain $Af \prec f$ as desired. ■

4. REARRANGEMENT-INVARIANT SPACES

The ℓ^p-norm $\left(\sum_{k=1}^n a_k^p\right)^{1/p}$ of a nonnegative vector $(a_1, a_2, \ldots a_n)$ evidently depends on the values $a_1, a_2, \ldots, a_n$ but not on the order in which they are arranged. In other words, the ℓ^p-norm provides some measure of the "magnitude" of a vector but without regard to how the entries are actually distributed over the underlying measure space. We shall want to isolate the Banach function spaces, over more general measure spaces, whose norms possess this kind of property. These spaces, the so-called *rearrangement-invariant spaces*, have a considerably richer structure than Banach function spaces in general.

Definition 4.1. Let ρ be a function norm over a totally σ-finite measure space (R, μ). Then ρ is said to be *rearrangement-invariant* if $\rho(f) = \rho(g)$ for every pair of equimeasurable functions f and g in $\mathcal{M}_0^+(R, \mu)$. In that case, the Banach function space $X = X(\rho)$ generated by ρ is said to be a *rearrangement-invariant space*.

Observe that a Banach function space X is rearrangement-invariant if and only if, whenever f belongs to X and g is equimeasurable with f, then g also belongs to X and $\|g\|_X = \|f\|_X$. It follows from Proposition 1.8 that the Lebesgue spaces $L^p(R, \mu)$ are rearrangement-invariant.

The decreasing rearrangement f^* of f is certainly equimeasurable with f, and in some sense f^* may be regarded as the canonical choice of a function equimeasurable with f. In this section we shall see that many of the Banach function space properties of rearrangement-invariant spaces have an interpretation in terms of f^* rather than f, at least when the underlying measure space is resonant.

Proposition 4.2. *Let ρ be a rearrangement-invariant function norm over a resonant measure space (R, μ). Then the associate norm ρ' is also rearrangement-invariant. Furthermore, the norms ρ and ρ' are given by*

$$\rho'(g) = \sup\left\{\int_0^\infty f^*(s)g^*(s)\,ds : \rho(f) \le 1\right\}, \qquad (g \in \mathcal{M}_0^+) \tag{4.1}$$

and

$$\rho(f) = \sup\left\{\int_0^\infty f^*(s)g^*(s)\,ds : \rho'(g) \le 1\right\}, \qquad (f \in \mathcal{M}_0^+). \tag{4.2}$$

Proof. The associate norm ρ' is defined by (cf. Definition I.2.1)

$$\rho'(g) = \sup\left\{\int_R fg\,d\mu : \rho(f) \le 1\right\}, \qquad (g \in \mathcal{M}_0^+). \tag{4.3}$$

Observe however that if $\rho(f) \le 1$, and if $\bar{f}$ is any nonnegative function on R equimeasurable with f, then $\rho(\bar{f}) \le 1$ because ρ is rearrangement-invariant. The supremum in (4.3) therefore extends over all f with $\rho(f) \le 1$ and over all $\bar{f}$ equimeasurable with such f. But then by (2.5) it must coincide with the supremum in (4.1) because (R, μ) is resonant. This establishes (4.1).

Since any two equimeasurable functions have the same decreasing rearrangement, it follows immediately from (4.1) that ρ' is rearrangement-invariant. In that case, we may apply the result to ρ' instead of ρ. The resulting identity analogous to (4.1) then coincides with (4.2) because $\rho'' = \rho$ by Theorem I.2.7. ■

Corollary 4.3. (Hölder's inequality). *Let ρ be a rearrangement-invariant norm over a resonant measure space (R, μ). If f and g belong to $\mathscr{M}_0^+(R, \mu)$, then*

$$\int_R fg \, d\mu \le \int_0^\infty f^*(s) g^*(s) \, ds \le \rho(f) \rho'(g). \tag{4.4}$$

Proof. The first inequality follows from Theorem 2.2, the second from the preceding result. ■

For convenience of reference we shall also formulate the last two results in terms of the Banach function spaces X and X' generated by ρ and ρ', respectively.

Corollary 4.4. *Let X be a Banach function space over a resonant measure space. Then X is rearrangement-invariant if and only if the associate space X' is, and in that case the norms are given by*

$$\|g\|_{X'} = \sup\left\{\int_0^\infty f^*(s) g^*(s) \, ds : \|f\|_X \le 1\right\}, \qquad (g \in X'), \tag{4.5}$$

and

$$\|f\|_X = \sup\left\{\int_0^\infty f^*(s) g^*(s) \, ds : \|g\|_{X'} \le 1\right\}, \qquad (f \in X). \tag{4.6}$$

Corollary 4.5. (Hölder's inequality). *Let X be a rearrangement-invariant space over a resonant measure space (R, μ). If f belongs to X and g to X', then*

$$\int_R |fg| \, d\mu \le \int_0^\infty f^*(s) g^*(s) \, ds \le \|f\|_X \|g\|_{X'}. \tag{4.7}$$

The significance of the Hardy-Littlewood-Pólya relation (Definition 3.5) to

the theory of rearrangement-invariant spaces stems from the following key result.

Theorem 4.6. *Let (R, μ) be a resonant measure space and suppose f_1 and f_2 belong to $\mathscr{M}_0^+(R, \mu)$. Let ρ be any rearrangement-invariant function norm over (R, μ). Then $f_1 \prec f_2$ implies $\rho(f_1) \le \rho(f_2)$.*

Proof. Because of (4.2), it need only be shown that

$$\int_0^\infty f_1^*(s) g^*(s) \, ds \le \int_0^\infty f_2^*(s) g^*(s) \, ds,$$

for every g satisfying $\rho'(g) \le 1$. But this is an immediate consequence of Hardy's lemma (Proposition 3.6) since $f_1 \prec f_2$ and g^* is nonnegative and decreasing. ■

Corollary 4.7. *Let X be a rearrangement-invariant space over a resonant measure space. Suppose f_1 belongs to $\mathscr{M}_0$ and f_2 belongs to X. If $f_1 \prec f_2$, then f_1 belongs to X and $\|f_1\|_X \le \|f_2\|_X$.*

An important application of Theorem 4.6 is that averaging operators A of the kind described by (3.19) are contractions in every rearrangement-invariant space X.

Theorem 4.8. *Let (R, μ) be a resonant measure space and let $(E_j)_{j \in J}$ be a countable collection of pairwise disjoint subsets of R, each with finite positive measure. Let $E = R \setminus \bigcup_j E_j$. For each f in $\mathscr{M}_0(R, \mu)$, let*

$$Af = f\chi_E + \sum_{j \in J} \left(\frac{1}{\mu(E_j)} \int_{E_j} f \, d\mu \right) \chi_{E_j}.$$

Then A is a contraction in every rearrangement-invariant space X over (R, μ), that is,

$$\|Af\|_X \le \|f\|_X, \qquad (f \in X).$$

Proof. If f belongs to X, then f is integrable over each E_j $(j \in J)$ by Theorem I.1.7(v). But then $Af \prec f$ by Proposition 3.7 and so the desired conclusion follows from Corollary 4.7. ■

Here is another application of Theorem 4.6.

Theorem 4.9. *Let (R, μ) be an arbitrary totally σ-finite measure space and suppose λ is a rearrangement-invariant function norm over $(\mathbf{R}^+, m)$. Then the* *functional $\underline{\lambda}$ defined by*

$$\underline{\lambda}(f) = \lambda(f^*), \qquad (f \in \mathscr{M}_0^+(R, \mu)) \tag{4.8}$$

is a rearrangement-invariant function norm over (R, μ).

Proof. The function-norm properties for $\underline{\lambda}$ (properties (P1),...,(P5) of Definition I.1.1) follow easily from their counterparts for λ. The only property requiring comment is the triangle inequality. Suppose f_1 and f_2 belong to $\mathscr{M}_0^+(R, \mu)$. By the subadditivity of the maximal operator (Theorem 3.4), we have

$$(f_1 + f_2)^{**} \le f_1^{**} + f_2^{**} = (f_1^* + f_2^*)^{**},$$

which shows that $(f_1 + f_2)^* \prec f_1^* + f_2^*$. Hence, since $(\mathbf{R}^+, m)$ is resonant, we may use Theorem 4.6 and the triangle inequality for λ to obtain

$$\lambda((f_1 + f_2)^*) \le \lambda(f_1^* + f_2^*) \le \lambda(f_1^*) + \lambda(f_2^*).$$

By virtue of (4.8), this establishes the triangle inequality for $\underline{\lambda}$. ■

The next result shows that *every* rearrangement-invariant function norm over (R, μ) arises in this way, provided the measure space (R, μ) is resonant.

Theorem 4.10 (Luxemburg representation theorem). *Let ρ be a rearrangement-invariant function norm over a resonant measure space (R, μ). Then there is a (not necessarily unique) rearrangement-invariant function norm $\bar{\rho}$ over $(\mathbf{R}^+, m)$ such that*

$$\rho(f) = \bar{\rho}(f^*), \qquad (f \in \mathscr{M}_0^+(R, \mu)). \tag{4.9}$$

Furthermore, if σ is any rearrangement-invariant function norm over $(\mathbf{R}^+, m)$ which represents ρ, in the sense that

$$\rho(f) = \sigma(f^*), \qquad (f \in \mathscr{M}_0^+(R, \mu)), \tag{4.10}$$

then the associate norm ρ' of ρ is represented in the same way by the associate norm σ' of σ, that is,

$$\rho'(g) = \sigma'(g^*), \qquad (g \in \mathscr{M}_0^+(R, \mu)). \tag{4.11}$$

Proof. If $\bar{\rho}$ is defined on $\mathscr{M}_0^+(\mathbf{R}^+, m)$ by

$$\bar{\rho}(h) = \sup\left\{\int_0^\infty g^* h^* \, ds : \rho'(g) \le 1\right\}, \tag{4.12}$$

then it is clear from (4.2) that (4.9) holds. In order to show that $\bar{\rho}$ is a function norm over $(\mathbf{R}^+, m)$, we need to verify the properties (P1),..., (P5) of Definition I.1.1.

For the triangle inequality, suppose h_1 and h_2 are arbitrary functions in $\mathscr{M}_0^+(\mathbf{R}^+, m)$. Then, as in the proof of the preceding theorem, $(h_1 + h_2)^* \prec h_1^* + h_2^*$. Hence, using Hardy's lemma (Proposition 3.6) and (4.12), we obtain

$$\bar{\rho}(h_1 + h_2) \le \bar{\rho}(h_1) + \bar{\rho}(h_2),$$

as desired. The remaining properties under (P1) are easy to verify. The lattice property (P2) follows at once from (1.14) and (4.12), and the Fatou property (P3) is a consequence of (1.17), (4.12), and the monotone convergence theorem.

The functional $\bar{\rho}$ is evidently rearrangement-invariant so in order to verify property (P4) it will be enough to show that $\bar{\rho}(\chi_{[0,t)})$ is finite for each $t > 0$. This is clear if t is in the range of μ, for then there is a subset F of R with measure $\mu(F) = t$. Since $\rho(\chi_F)$ is finite because of property (P4) for ρ, this implies $\bar{\rho}(\chi_{[0,t)})$ is finite because of (4.9). The triangle inequality for $\bar{\rho}$ now establishes the same result for any integral multiple of t, and the result for arbitrary t follows from this by appealing to the lattice property (P2). The proof of (P5) is similar. Hence, $\bar{\rho}$ is a rearrangement-invariant function norm.

Now let σ be any rearrangement-invariant function norm over $(\mathbf{R}^+, m)$ representing ρ in the sense that (4.10) holds. Since $(\mathbf{R}^+, m)$ is resonant (Theorem 2.7), the associate norm σ' of σ may be expressed as in (4.1) by

$$\sigma'(k) = \sup\left\{\int_0^\infty h^* k^* \, ds : \sigma(h) \le 1\right\}, \qquad (k \in \mathscr{M}_0^+(\mathbf{R}^+, m)).$$

In particular,

$$\sigma'(g^*) = \sup\left\{\int_0^\infty g^* h^* \, ds : \sigma(h) \le 1\right\}, \tag{4.13}$$

for any g in $\mathscr{M}_0^+(R, \mu)$. Hence, it follows from (4.10) and (4.1) that

$$\rho'(g) \le \sigma'(g^*), \qquad (g \in \mathscr{M}_0^+(R, \mu)). \tag{4.14}$$

For the reverse inequality, let $\mathscr{H}$ denote the class of nonnegative decreasing simple functions h that are supported on the interval $[0, \mu(R))$ and satisfy $\sigma(h) \le 1$. Since σ' has the Fatou property, it follows from (4.13) that

$$\sigma'(g^*) = \sup\left\{\int_0^\infty g^* h^* \, ds : h \in \mathscr{H}\right\}. \tag{4.15}$$

If (R, μ) is non-atomic, then every function h in $\mathscr{H}$ is the decreasing rearrangement of a simple function f on R (cf. Example 1.6(b)). Hence, it is clear from (4.1) and (4.15) that equality holds in (4.14). This establishes the desired result (4.11).

Since (R, μ) is resonant, the only remaining case is that of a completely atomic measure space (R, μ) with all atoms having the same positive measure, say α (cf. Theorem 2.7). In this case, the functions h in the class $\mathscr{H}$ considered above are not necessarily constant on the intervals $I_k = [(k-1)\alpha, k\alpha)$, $(k = 1, 2, \ldots)$, so they may not be representable as decreasing rearrangements of simple functions f on (R, μ). The remedy is simply to average each of the functions h in $\mathscr{H}$ over each of the intervals I_k. The integrals in (4.15) do not change because g^* is constant on each I_k, and the σ-norm of the averaged version of h cannot exceed $\sigma(h)$ because of Theorem 4.8. Hence, with this modification to $\mathscr{H}$, we may proceed exactly as before to conclude that (4.11) holds for ρ' and σ'. ■

The Luxemburg representation theorem shows in particular that the rearrangement-invariant spaces over a resonant measure space (R, μ) are completely determined by the rearrangement-invariant spaces over $(\mathbf{R}^+, m)$. The representation $\rho \to \bar{\rho}$ is in fact unique if (R, μ) is non-atomic and has infinite measure, although this is not the case in general. For example, the L^1-norm ρ over the finite interval $[0, 1)$ may be represented by either of the rearrangement-invariant norms

$$\sigma_1(h) = \int_0^1 h^*(s)\, ds, \qquad \sigma_2(h) = \int_0^\infty h^*(s)\, ds$$

over $(\mathbf{R}^+, m)$. The uniqueness can easily be restored however if $\mathbf{R}^+$ is replaced by $[0, \mu(R))$ (or, in the atomic case, by a suitable discrete analogue). The details are left as an exercise.

In measure spaces that are not resonant (hence contain a mixture of atomic and non-atomic parts, or a mixture of atoms of unequal measures), the concept of rearrangement-invariance has little practical significance. As is perhaps to be expected, many of the foregoing results break down in such a setting. For example, there may be rearrangement-invariant norms on such measure spaces that are not generated by any rearrangement-invariant norm on $\mathbf{R}^+$. For those that do arise from norms on $\mathbf{R}^+$ (as in Theorem 4.9), something of the theory can be salvaged. These are the so-called *universally rearrangement-invariant norms*; a sketch of the theory is provided in the exercises.

5. THE FUNDAMENTAL FUNCTION

Certain duality and separability properties for arbitrary Banach function spaces were established in Chapter I. We shall consider these properties again in the present section for rearrangement-invariant spaces and reformulate them in terms of a certain function φ_X, the *fundamental function* of X, associated with the rearrangement-invariant space X. In light of such results it is natural to attempt to classify rearrangement-invariant spaces according to fundamental function, and in particular to ask whether there are respectively smallest and largest rearrangement-invariant spaces with a given fundamental function. It turns out that concavity is quite crucial here but arbitrary fundamental functions do not necessarily possess this property. However, it is always possible to replace a given fundamental function φ_X with an "equivalent" function $\tilde{\varphi}_X$ which is concave, and with this modification it is possible to show that to any rearrangement-invariant space X there correspond rearrangement-invariant spaces $\Lambda(X)$ and $M(X)$ which are respectively the smallest and the largest of all rearrangement-invariant spaces whose fundamental functions coincide with $\tilde{\varphi}_X$. The spaces $\Lambda(X)$ and $M(X)$ are known as *Lorentz spaces*, and they will figure prominently in the weak-type interpolation theory to be developed in Chapters III and IV.

Definition 5.1. Let X be a rearrangement-invariant Banach function space over a resonant measure space (R, μ). For each finite value of t belonging to the range of μ, let E be a subset of R with $\mu(E) = t$ and let

$$\varphi_X(t) = \|\chi_E\|_X. \tag{5.1}$$

The function φ_X so defined is called the *fundamental function* of X.

Observe that the particular choice of set E with $\mu(E) = t$ is immaterial since if F is any other subset of R with $\mu(F) = t$, then χ_E and χ_F are equimeasurable and so $\|\chi_E\|_X = \|\chi_F\|_X$ because of the rearrangement-invariance of X. Hence, φ_X is well-defined by (5.1).

If (R, μ) is nonatomic, then the range of μ consists of the interval $[0, \mu(R)]$. In that case, a simple computation shows that the fundamental function of $L^p(R, \mu)$ is given by

$$\varphi_{L^p}(t) = t^{1/p}, \qquad (1 \le p < \infty), \qquad (0 \le t < \mu(R)), \tag{5.2}$$

and

$$\varphi_{L^\infty}(0) = 0; \qquad \varphi_{L^\infty}(t) = 1, \qquad (0 < t < \mu(R)). \tag{5.3}$$

By Theorem 2.7, the only other case of a resonant measure space is that in which (R, μ) is completely atomic with all atoms having the same measure. A typical example is that of the set $\mathbf{Z}$ of integers with counting measure, whose finite range is $\mathbf{Z}^+ = \{0, 1, 2, \ldots\}$. Writing ℓ^p for $L^p(\mathbf{Z})$, we thus have

$$\varphi_{\ell^p}(n) = n^{1/p}, \qquad (1 \le p < \infty), \qquad (n = 0, 1, 2, \ldots) \tag{5.4}$$

and

$$\varphi_{\ell^\infty}(0) = 0; \qquad \varphi_{\ell^\infty}(n) = 1, \qquad (n = 1, 2, \ldots). \tag{5.5}$$

One may observe from these identities that the product of the fundamental functions of $L^p(R, \mu)$ and its associate space $L^{p'}(R, \mu)$, $(1/p + 1/p' = 1)$, is identically equal to t. This is true generally, as the next result shows.

Theorem 5.2. *Let X be a rearrangement-invariant space over a resonant measure space (R, μ) and let X' be the associate space of X. Then*

$$\varphi_X(t)\varphi_{X'}(t) = t \tag{5.6}$$

for each finite value of t in the range of μ.

Proof. If $t = 0$, there is nothing to prove since every fundamental function vanishes at the origin. Hence, we may assume $0 < t < \infty$. Since t is in the range of μ, there is a subset E of R with $\mu(E) = t$, so by Hölder's inequality (4.7),

$$t = \int_E d\mu \le \|\chi_E\|_X \|\chi_E\|_{X'} = \varphi_X(t)\varphi_{X'}(t).$$

For the reverse inequality, write

$$\varphi_X(t) = \|\chi_E\|_X = \sup\left\{\int_E |g|\, d\mu : \|g\|_{X'} \le 1\right\}. \tag{5.7}$$

Now for any such g, Theorem 4.8 shows that the function

$$h = \left(\frac{1}{t}\int_E |g|\, d\mu\right)\chi_E$$

satisfies

$$\left(\frac{1}{t}\int_E |g|\, d\mu\right)\varphi_{X'}(t) = \|h\|_{X'} \le \|g\|_{X'} \le 1,$$

so taking the supremum over g and using (5.7) we obtain

$$\frac{\varphi_X(t)\varphi_{X'}(t)}{t} \le 1.$$

This completes the proof. ■

Corollary 5.3. *Let X be a rearrangement-invariant space over a resonant measure space (R, μ). Then the fundamental function φ_X of X satisfies:*

$$\varphi_X \text{ is increasing; } \varphi_X(t) = 0 \text{ iff } t = 0; \tag{5.8}$$

$$\varphi_X(t)/t \text{ is decreasing;} \tag{5.9}$$

$$\varphi_X \text{ is continuous, except perhaps at the origin.} \tag{5.10}$$

Proof. The fact that φ_X is increasing follows from the lattice property of Banach function spaces (Theorem I.1.7(i)). In that case, the preceding theorem shows that $\varphi_X(t)/t = 1/\varphi_{X'}(t)$ is decreasing. The continuity of φ_X requires proof only when (R, μ) is nonatomic since otherwise φ_X is defined on a discrete subset of $\mathbf{R}^+$ and hence is automatically continuous. In the nonatomic case, φ_X is an increasing function on the interval $(0, \mu(R))$ so at worst it can have jump discontinuities there. But clearly φ_X cannot have a jump discontinuity at any $t_0 > 0$ since then (5.9) would be violated to the right of t_0. This establishes (5.10). Note that (5.3) provides an example where φ_X does have a jump discontinuity at the origin. ■

Recall from §I.3 that X_a denotes the order ideal of functions in X with absolutely continuous norm whereas X_b signifies the closure in X of the simple functions. If (R, μ) is completely atomic with all of its (countably many) atoms having the same measure, then $X_a = X_b$ (by Theorem I.3.13) and $(X_b)^* = X'$ isometrically (by Theorem I.4.1). Furthermore, since such a measure space (R, μ) is evidently separable, it follows from Theorem I.5.5 that X_b is separable. We summarize these results as follows.

Theorem 5.4. *Let (R, μ) be completely atomic, consisting of countably many atoms of equal measure. If X is any rearrangement-invariant space over (R, μ), then*

(i) $X_a = X_b$;
(ii) $(X_b)^* = X'$;
(iii) *X_b is separable.*

The analogous result for nonatomic measure spaces requires an additional hypothesis on the fundamental function.

Theorem 5.5. *Let (R, μ) be a totally σ-finite nonatomic measure space and let X be an arbitrary rearrangement-invariant space over (R, μ).*

(a) *The following conditions on X are equivalent*:

(i) $\lim_{t\to 0^+} \varphi_X(t) = 0$;
(ii) $X_a = X_b$;
(iii) $(X_b)^* = X'$.

If, in addition, μ is separable, then each of properties (i), (ii), *and* (iii) *is equivalent to*

(iv) X_b is separable.

(b) *If* $\lim_{t\to 0^+} \varphi_X(t) > 0$, *then* $X_a = \{0\}$.

Proof. The equivalence of (ii) and (iii) follows from Theorem I.4.1, while the equivalence of (ii) and (iv), for μ separable, follows from Theorem I.5.5. Hence, the proof of part (a) will be complete once we establish the equivalence of (i) and (ii).

If (i) holds, then χ_E has absolutely continuous norm for any subset E of R with finite measure. For if $E_n \downarrow \varnothing$ μ-a.e., then $\mu(E \cap E_n) \downarrow 0$ by the dominated convergence theorem and so

$$\|\chi_E \chi_{E_n}\|_X = \|\chi_{E\cap E_n}\|_X \le \varphi_X(\mu(E \cap E_n)) \downarrow 0.$$

But then $X_a = X_b$ by Theorem I.3.13 and so (ii) holds.

Conversely, suppose $X_a = X_b$. Let E be any subset of R with positive measure. Since (R, μ) is nonatomic, we may select subsets E_n of E such that $\mu(E_n) = 2^{-n}\mu(E)$ and $E_{n+1} \subset E_n \subset E$, $(n = 1, 2, \ldots)$. The hypothesis $X_a = X_b$ shows in particular that χ_E has absolutely continuous norm so, since $E_n \downarrow \varnothing$ μ-a.e.,

$$\varphi_X(2^{-n}\mu(E)) = \|\chi_{E_n}\|_X = \|\chi_E \chi_{E_n}\|_X \downarrow 0.$$

This, together with the fact that φ_X is monotone, establishes (i).

To prove part (b), suppose to the contrary that X_a contains a function f which is nonzero on a set of positive measure. Then there exists $\varepsilon > 0$ and a subset E of R with finite measure such that $\varepsilon\chi_E \le |f|$. Hence, χ_E also belongs to X_a since X_a is an order ideal. But then, exactly as in the first part of the proof, we may conclude that $\varphi_X(t) \downarrow 0$ as $t \downarrow 0$. This establishes part (b). ■

We have seen in Corollary 5.3 that every fundamental function satisfies properties (5.8) and (5.9). Our next goal is to show that these properties in fact completely determine the functions φ that can arise as fundamental functions of rearrangement-invariant spaces. We introduce the following terminology.

Definition 5.6. Let φ be a nonnegative function defined on the interval $\mathbf{R}^+ = [0, \infty)$. If

$$\varphi(t) \text{ is increasing on } (0,\infty);\ \varphi(t) = 0 \iff t = 0; \tag{5.11}$$

$$\varphi(t)/t \text{ is decreasing on } (0,\infty), \tag{5.12}$$

then φ is said to be *quasiconcave*.

Observe that every nonnegative concave function on $[0,\infty)$ that vanishes only at the origin is quasiconcave. There are, however, quasiconcave functions that are not concave; take, for example, the function $\varphi(t) = \max(1, t)$, for $t > 0$, and $\varphi(0) = 0$.

The fundamental function of every rearrangement-invariant space over $(\mathbf{R}^+, m)$ is quasiconcave (Corollary 5.3). Conversely, we shall now show that to any quasiconcave function φ there corresponds a rearrangement-invariant space M over $(\mathbf{R}^+, m)$ whose fundamental function is φ.

Definition 5.7. Let φ be a quasiconcave function on $\mathbf{R}^+$. The *Lorentz space* $M_\varphi = M_\varphi(\mathbf{R}^+, m)$ consists of all functions f in $\mathcal{M}_0(\mathbf{R}^+, m)$ for which the functional

$$\|f\|_{M_\varphi} = \sup_{0<t<\infty} \{f^{**}(t)\varphi(t)\} \tag{5.13}$$

is finite.

Proposition 5.8. *If φ is quasiconcave, then the Lorentz space M_φ is a rearrangement-invariant Banach function space whose fundamental function coincides with φ.*

Proof. The Banach function norm properties (P1), (P2), and (P3) (cf. Definition I.1.1) for the functional

$$\rho(f) = \sup_{0<t<\infty} \{f^{**}(t)\varphi(t)\}, \qquad (f \in \mathcal{M}_0^+(\mathbf{R}^+, m))$$

follow easily from the elementary properties of f^{**} (the triangle inequality is a consequence of Theorem 3.4). Furthermore, M_φ is obviously rearrangement-invariant since its norm is defined in terms of f^*. To verify property (P4), let E be any subset of $\mathbf{R}^+$ with measure $m(E) = t$. Then $\chi_E^* = \chi_{[0,t)}$ and so

$$\begin{aligned}\|\chi_E\|_{M_\varphi} &= \sup_{0<s<\infty} \{\chi^{**}_{(0,t)}(s)\varphi(s)\} = \sup_{0<s<\infty} \left\{\min\left(1, \frac{t}{s}\right)\varphi(s)\right\} \\ &= \max\left\{\sup_{0<s<t} \varphi(s),\ t \cdot \sup_{t\le s\le\infty} \frac{\varphi(s)}{s}\right\} = \varphi(t),\end{aligned} \tag{5.14}$$

since φ increases and $\varphi(s)/s$ decreases. Because $\varphi(t) < \infty$, this establishes (P4). Finally, if f belongs to M_φ and E is an arbitrary subset of $\mathbf{R}^+$ of measure $t > 0$, then by (3.1),

$$\left|\int_E f(x)\,dx\right| \leq \int_0^t f^*(s)\,ds \leq \frac{t}{\varphi(t)} \cdot \sup_{0<s<\infty} \{f^{**}(s)\varphi(s)\} = C_t \|f\|_{M_\varphi}.$$

Since φ vanishes only at the origin, the constant $C_t = t/\varphi(t)$ is finite and so this establishes (P5). Hence M_φ is a rearrangement-invariant Banach function space and the identity (5.14) shows that its fundamental function coincides with φ. ■

The next result shows that M_φ is in fact the largest rearrangement-invariant space with fundamental function φ.

Proposition 5.9. *Let X be an arbitrary rearrangement-invariant space over $(\mathbf{R}^+, m)$. Then $X \hookrightarrow M_{\varphi_X}$ and the embedding has norm* 1:

$$\|f\|_{M_{\varphi_X}} \leq \|f\|_X, \qquad (f \in X). \tag{5.15}$$

Proof. Let $t > 0$. Then by Hölder's inequality and (5.6) we have, for any f in X,

$$\int_0^t f^*(s)\,ds \leq \|\chi_{(0,t)}\|_{X'} \|f\|_X = \frac{t}{\varphi(t)} \|f\|_X.$$

Hence,

$$f^{**}(t)\varphi(t) \leq \|f\|_X, \qquad (t > 0),$$

from which (5.15) follows. ■

It is natural to ask whether there is always a smallest rearrangement-invariant space with a given fundamental function. The simple answer is yes, but with some qualifications, and in order to provide the complete answer we shall need to delve deeper into the relationship between concave and quasiconcave functions.

Observe from (5.11) and (5.12) that every quasiconcave function satisfies $\varphi(t) \leq \varphi(1) \cdot \max(1, t)$, for all $t > 0$. In particular, φ is dominated by the concave function $\varphi(1)(1 + t)$. Hence, since the pointwise infimum of concave functions is itself concave, it follows that there is a smallest concave function, say $\tilde{\varphi}$, which dominates φ. The function $\tilde{\varphi}$ is called the *least concave majorant* of φ.

If $x > 0$ is fixed, it follows from (5.11) and (5.12) that $\varphi(t) \leq (1 + t/x)\varphi(x)$ for all $t \geq 0$. Thus $\varphi(t)$ is dominated by the concave function $\psi(t) = (1 + t/x)\varphi(x)$ and so the least concave majorant $\tilde{\varphi}$ must satisfy $\tilde{\varphi}(t) \leq \psi(t)$ for all t. In particular, when $t = x$, we have $\tilde{\varphi}(x) \leq \psi(x) \leq 2\varphi(x)$ and so we obtain the following result.

Proposition 5.10. *If φ is quasiconcave, then the least concave majorant $\tilde{\varphi}$ of φ satisfies*

$$\frac{1}{2}\tilde{\varphi} \leq \varphi \leq \tilde{\varphi}. \tag{5.16}$$

Recall that the quasiconcave functions are the ones that occur naturally as fundamental functions of rearrangement-invariant spaces. Such fundamental functions need not be concave, but with the aid of the last result we shall see that every rearrangement-invariant space can at least be equivalently renormed so that its fundamental function is concave.

Proposition 5.11. *Let X be a rearrangement-invariant space over $(\mathbf{R}^+, m)$. Then X can be equivalently renormed with a rearrangement-invariant norm in such a way that the resulting fundamental function is concave.*

Proof. The fundamental function $\varphi = \varphi_X$ of X is quasiconcave by Corollary 5.3, and so by Proposition 5.10, the least concave majorant $\tilde{\varphi}$ of φ satisfies (5.16). Let $M_{\tilde{\varphi}}$ be the Lorentz space associated with $\tilde{\varphi}$ (cf. Definition 5.7) and let

$$\nu(f) = \max(\|f\|_X, \|f\|_{M_{\tilde{\varphi}}}), \qquad (f \in \mathscr{M}_0^+(\mathbf{R}^+, m)).$$

Since both X and $M_{\tilde{\varphi}}$ are rearrangement-invariant spaces, it is a simple matter to check that ν is a rearrangement-invariant function norm. Furthermore, it follows from (5.15) and (5.16) that

$$\|f\|_X \leq \nu(f) \leq \max(\|f\|_X, 2\|f\|_{M_\varphi}) \leq 2\|f\|_X,$$

so ν is equivalent to the norm of X. Finally, since

$$\nu(\chi_{(0,t)}) = \max(\varphi(t), \tilde{\varphi}(t)) = \tilde{\varphi}(t),$$

we see that $X = X(\nu)$ has concave fundamental function $\tilde{\varphi}$. ■

Definition 5.12. Let X be a rearrangement-invariant space over $(\mathbf{R}^+, m)$ and suppose X has been renormed as in Proposition 5.11 so that its fundamental function φ_X is concave. The *Lorentz spaces* $\Lambda(X)$ and $M(X)$ are defined as follows. The space $M(X)$ is simply the space M_{φ_X} (cf. Definition 5.7) whose norm is given by

$$\|f\|_{M(X)} = \sup_{0<t<\infty} \{f^{**}(t)\varphi_X(t)\}. \tag{5.17}$$

The space $\Lambda(X)$ consists of all f in $\mathscr{M}_0^+(\mathbf{R}^+, m)$ for which

$$\|f\|_{\Lambda(X)} = \int_0^\infty f^*(s)\, d\varphi_X(s) \tag{5.18}$$

is finite.

We remark that the integral in (5.18) is well-defined because φ_X is increasing. Furthermore, since φ_X is nonnegative and concave, then φ_X may be represented as the integral of a nonnegative, decreasing function, say ϕ_X, on $(0, \infty)$. Hence the Riemann-Stieltjes integral in (5.18) may be rewritten in the form

$$\|f\|_{\Lambda(X)} = \|f\|_{L^\infty}\varphi_X(0+) + \int_0^\infty f^*(s)\phi_X(s)\, ds. \tag{5.19}$$

Theorem 5.13. *Let X be a rearrangement-invariant Banach function space over $(\mathbf{R}^+, m)$ and suppose X has been renormed to have concave fundamental function φ_X. Then the Lorentz spaces $\Lambda(X)$ and $M(X)$ are rearrangement-invariant Banach function spaces and each has fundamental function equal to φ_X. Furthermore,*

$$\Lambda(X) \hookrightarrow X \hookrightarrow M(X) \tag{5.20}$$

and each of the embeddings has norm 1.

Proof. Since $M(X) = M_{\varphi_X}$, the assertions involving $M(X)$ have already been established in Propositions 5.8 and 5.9.

To establish the triangle inequality in $\Lambda(X)$, recall that $(f+g)^* \prec f^* + g^*$ and, since φ_X is concave, its derivative ϕ_X is nonnegative and decreasing. Hence, by Hardy's lemma (Proposition 3.6),

$$\int_0^\infty (f+g)^*\phi_X\, ds \le \int_0^\infty f^*\phi_X\, ds + \int_0^\infty g^*\phi_X\, ds,$$

and this, together with the identity (5.19), establishes the triangle inequality for the norm in $\Lambda(X)$. The remaining function norm properties (P1) – (P3) follow easily from the corresponding properties of f^*. To verify properties (P4) and

(P5), let E be any set of measure $m(E) = t > 0$. Then

$$\|\chi_E\|_{\Lambda(X)} = \int_0^\infty \chi_{(0,t)}(s)\, d\varphi_X(s) = \varphi_X(t). \tag{5.21}$$

This shows that (P4) holds and also that $\Lambda(X)$ has fundamental function equal to φ_X. The property (P5) for $\Lambda(X)$ will follow directly from the corresponding property for X once we have established the norm-one embedding of $\Lambda(X)$ into X, that is,

$$\|f\|_X \le \|f\|_{\Lambda(X)}, \qquad (f \in \Lambda(X)). \tag{5.22}$$

Since both norms are rearrangement-invariant and have the Fatou property it will suffice to establish (5.22) for decreasing step functions $f = f^*$. In that case, we may write

$$f^* = \sum_{k=1}^n c_k \chi_{(0,t_k)},$$

where $c_k > 0$ and $0 < t_1 < t_2 < \cdots < t_n$ (cf. Example I.1.6(c)). But then

$$\begin{aligned}\|f\|_X &\le \sum_{k=1}^n c_k \|\chi_{(0,t_k)}\|_X = \sum_{k=1}^n c_k \varphi_X(t_k)\\ &= \int_0^\infty f^*(s)\, d\varphi_X(s) = \|f\|_{\Lambda(X)},\end{aligned}$$

as desired. ■

Corollary 5.14. *If X is a rearrangement-invariant space over $(\mathbf{R}^+, m)$ with concave fundamental function φ_X, then the Lorentz spaces $\Lambda(X)$ and $M(X)$ are respectively the smallest and the largest of all rearrangement-invariant spaces with fundamental function φ_X.*

In the latter part of this section we have worked exclusively with the measure space $(\mathbf{R}^+, m)$. The definitions and results can easily be carried over to arbitrary resonant measure spaces (R, μ), however, by appealing to the Luxemburg representation theorem.

6. THE SPACES $L^1 + L^\infty$ AND $L^1 \cap L^\infty$

Further examples of rearrangement-invariant spaces can be constructed from the Lebesgue spaces simply by taking sums and intersections. The procedure is illustrated in this section with the construction of $L^1 + L^\infty$ and

$L^1 \cap L^\infty$. These spaces play a special role in the theory in that they are respectively the largest and the smallest of all rearrangement-invariant spaces (Theorem 6.6).

Definition 6.1. Let (R, μ) be a totally σ-finite measure space.

(a) The space $L^1 + L^\infty = (L^1 + L^\infty)(R, \mu)$ consists of all functions f in $\mathcal{M}_0(R, \mu)$ that are representable as a sum $f = g + h$ of functions g in L^1 and h in L^∞. For each f in $L^1 + L^\infty$, let

$$\|f\|_{L^1+L^\infty} = \inf\{\|g\|_{L^1} + \|h\|_{L^\infty}\}, \tag{6.1}$$

where the infimum is taken over all representations $f = g + h$ of the kind described above.

(b) For each f in the intersection $L^1 \cap L^\infty$ of L^1 and L^∞, let

$$\|f\|_{L^1 \cap L^\infty} = \max\{\|f\|_{L^1}, \|f\|_{L^\infty}\}. \tag{6.2}$$

The expression on the right of (6.2) is the maximum of the quantities $\sup_{1 \le t < \infty} \int_0^t f^*(s)\,ds$ and $\sup_{0<t<1} f^{**}(t)$, so the norm in $L^1 \cap L^\infty$ has the following description in terms of the maximal function f^{**}:

$$\|f\|_{L^1 \cap L^\infty} = \sup_{0<t<\infty} \frac{\int_0^t f^*(s)\,ds}{\min(1,t)} = \sup_{0<t<\infty} (f^{**}(t) \cdot \max(1,t)). \tag{6.3}$$

The next result provides an analogous description of the norm in $L^1 + L^\infty$.

Theorem 6.2. *Let (R, μ) be a totally σ-finite measure space and suppose f belongs to $\mathcal{M}_0(R, \mu)$. Then*

$$\inf_{f=g+h} \{\|g\|_{L^1} + t\|h\|_{L^\infty}\} = \int_0^t f^*(s)\,ds = tf^{**}(t), \tag{6.4}$$

for all $t > 0$.

Proof. The second of the identities in (6.4) results from the definition of f^{**} so we need only establish the first. Fix f in $\mathcal{M}_0(R, \mu)$ and $t > 0$ and let α_t denote the infimum on the right of (6.4).

Let us show first that

$$\int_0^t f^*(s)\,ds \le \alpha_t. \tag{6.5}$$

We may assume that f belongs to $L^1 + L^\infty$ since otherwise the infimum α_t is infinite and there is nothing to prove. In that case f may be expressed as a sum $f = g + h$ with g in L^1 and h in L^∞. The subadditivity of f^{**} (Theorem 3.4) gives

$$\int_0^t f^*(s)\,ds \le \int_0^t g^*(s)\,ds + \int_0^t h^*(s)\,ds$$

and hence, by (1.23) and (1.24),

$$\int_0^t f^*(s)\,ds \le \|g\|_{L^1} + t\|h\|_{L^\infty}.$$

Taking the infimum over all possible representations $f = g + h$, we obtain (6.5).

For the reverse inequality

$$\alpha_t \le \int_0^t f^*(s)\,ds,$$

it will suffice to construct functions g in L^1 and h in L^∞ such that $f = g + h$ and

$$\|g\|_{L^1} + t\|h\|_{L^\infty} \le \int_0^t f^*(s)\,ds. \tag{6.6}$$

Clearly, the right-hand side may be assumed to be finite. The Hardy-Littlewood inequality (3.1) then guarantees the integrability of f over any subset of R of measure at most t. Thus, if we let $E = \{x : |f(x)| > f^*(t)\}$ and set $t_0 = \mu(E)$, then (1.18) gives $t_0 \le t$ and so f is integrable over E. In particular, the function

$$g(x) = \max\{|f(x)| - f^*(t), 0\} \cdot \operatorname{sgn} f(x)$$

belongs to $L^1(R, \mu)$, whereas the function

$$h(x) = \min\{|f(x)|, f^*(t)\} \cdot \operatorname{sgn} f(x)$$

lies in $L^\infty(R, \mu)$ with L^∞-norm at most $f^*(t)$. Hence, by (3.1),

$$\|g\|_{L^1} = \int_E |f|\,d\mu - \mu(E)f^*(t) \le \int_0^{t_0} f^*(s)\,ds - t_0 f^*(t),$$

so

$$\|g\|_{L^1} + t\|h\|_{L^\infty} \le \int_0^{t_0} f^*(s)\,ds + (t - t_0)f^*(t).$$

But f^* is constant and equal to $f^*(t)$ whenever $t_0 \le s \le t$, so the last estimate in fact coincides with (6.6). Since $f = g + h$, the proof is complete. ■

The preceding theorem shows in particular that the norm of a function f in $L^1 + L^\infty$ is simply $f^{**}(1)$. This characterization will enable us to show fairly easily that $L^1 + L^\infty$ is a rearrangement-invariant Banach function space. It also provides a useful link between the norms in $L^1 + L^\infty$ and $L^1 \cap L^\infty$, and this will lead to the identification in Theorem 6.4 of $L^1 + L^\infty$ as the associate space of $L^1 \cap L^\infty$. The requisite Hölder inequality is essentially contained in the following lemma.

Lemma 6.3. *Let ξ and η be nonnegative decreasing functions on $(0, \infty)$. Then*

$$\int_0^\infty \xi\eta\, ds \le \left(\int_0^1 \eta\, ds\right) \cdot \max\left\{\int_0^\infty \xi\, ds,\ \sup_{0<s<\infty} \xi(s)\right\}. \tag{6.7}$$

Proof. Since ξ is nonnegative and decreasing, we may use (6.3) to express the maximum, say A, in (6.7) in the form

$$A = \sup_{0<t<\infty} \frac{\int_0^t \xi(s)\, ds}{\min(1,t)}.$$

Then, for any $t > 0$,

$$\int_0^t \xi(s)\, ds \le A \cdot \min(1,t) \le \int_0^t A\chi_{(0,1)}(s)\, ds.$$

Since η is also nonnegative and decreasing, we may therefore apply Hardy's lemma (Proposition 3.6) to obtain

$$\int_0^\infty \xi(s)\eta(s)\, ds \le \int_0^\infty A\chi_{(0,1)}(s)\eta(s)\, ds = A\int_0^1 \eta(s)\, ds,$$

and this establishes (6.7). ■

Theorem 6.4. *The spaces $L^1 + L^\infty$ and $L^1 \cap L^\infty$ over a resonant measure space are rearrangement-invariant Banach function spaces, and their norms are given by*

$$\|f\|_{L^1+L^\infty} = \int_0^1 f^*(s)\, ds \tag{6.8}$$

and

$$\|f\|_{L^1\cap L^\infty} = \sup_t \frac{\int_0^t f^*(s)\, ds}{\min(1,t)}. \tag{6.9}$$

Furthermore, the spaces $L^1 + L^\infty$ and $L^1 \cap L^\infty$ are mutually associate, that is,

$$(L^1 + L^\infty)' = L^1 \cap L^\infty; \qquad (L^1 \cap L^\infty)' = L^1 + L^\infty. \tag{6.10}$$

Proof. The expressions (6.8) and (6.9), which follow directly from (6.4) and (6.3), respectively, may be written in the form

$$\rho(f) = \|f\|_{L^1+L^\infty} = \sup_{0<t<\infty} \{f^{**}(t)\varphi(t)\},$$

where $\varphi(t) = \min(1,t)$, and

$$\nu(f) = \|f\|_{L^1\cap L^\infty} = \sup_{0<t<\infty} \{f^{**}(t)\psi(t)\}, \tag{6.11}$$

where $\psi(t) = \max(1,t)$. Hence, it is clear that $L^1 + L^\infty$ and $L^1 \cap L^\infty$ are precisely the Lorentz spaces M_φ and M_ψ, respectively. In particular, by Proposition 5.8, both are rearrangement-invariant Banach function spaces.

In order to establish (6.10), it will suffice to show that $\rho' = \nu$, for then the remaining identity $\rho = \nu'$ will follow from Theorem I.2.7. In one direction, the Hölder inequality (6.7) gives

$$\int fg\, d\mu \le \int_0^\infty f^*(s)g^*(s)\, ds \le \rho(f)\nu(g),$$

so it follows at once from the definition of the associate norm that $\rho' \le \nu$. In the opposite direction, the Hölder inequality (4.4), applied to ρ and ρ' gives

$$\int_0^t g^*(s)\, ds \le \rho(\chi_{(0,t)})\rho'(g) = \min(1,t)\rho'(g).$$

It follows from this and (6.11) that $\nu \le \rho'$, and hence the proof is complete. ■

Corollary 6.5. *The fundamental functions of $L^1 + L^\infty$ and $L^1 \cap L^\infty$ are given by*

$$\varphi_{L^1+L^\infty}(t) = \min(1,t); \qquad \varphi_{L^1\cap L^\infty}(t) = \max(1,t). \tag{6.12}$$

Theorem 6.6. *Let X be an arbitrary rearrangement-invariant Banach function space over a resonant measure space. Then*

$$L^1 \cap L^\infty \hookrightarrow X \hookrightarrow L^1 + L^\infty. \tag{6.13}$$

Furthermore, the norm in X may be replaced by a constant multiple of itself in such a way that each of the embeddings in (6.13) has norm 1.

Proof. In view of the Luxemburg representation theorem (Theorem 4.10) and the respective representations (6.8) and (6.9) of the norms in $L^1 + L^\infty$ and $L^1 \cap L^\infty$, it will suffice to prove the result only for the measure space $(\mathbf{R}^+, m)$. In that case, using Hölder's inequality (4.7) we obtain

$$\int_0^1 f^*(s)\, ds \le \|\chi_{(0,1)}\|_{X'}\|f\|_X = \varphi_{X'}(1)\|f\|_X. \tag{6.14}$$

In view of (6.8), this establishes the embedding $X \hookrightarrow L^1 + L^\infty$ and shows that the inclusion map is of norm at most $\varphi_{X'}(1)$. In fact, by taking f in (6.14) to be $\chi_{(0,1)}$ and appealing to Theorem 5.2, we see that the inclusion has norm equal to $\varphi_{X'}(1)$.

Similarly, by interchanging the roles of X and X', we find that $X' \hookrightarrow L^1 + L^\infty$ with the inclusion map of norm $\varphi_X(1)$. Hence, it follows from Proposition I.2.10 and Theorem 6.4 that $L^1 \cap L^\infty \hookrightarrow X$ and

$$\|f\|_X \le \varphi_X(1)\|f\|_{L^1 \cap L^\infty}, \qquad (f \in L^1 \cap L^\infty). \tag{6.15}$$

The inequalities (6.14) and (6.15) together establish the assertions in (6.13). Furthermore, since by Theorem 5.2 the constants $\varphi_X(1)$ and $\varphi_{X'}(1)$ are mutually reciprocal, it is clear that if $\|\cdot\|_X$ is replaced by $\varphi_{X'}(1)\|\cdot\|_X$, then each of the embeddings in (6.13) will have norm 1. ■

If (R, μ) is finite, then $\|f\|_{L^1} \le \mu(R)\|f\|_{L^\infty}$. It follows from (6.2) that the spaces $L^1 \cap L^\infty$ and L^∞ coincide and have equivalent norms (identical if $\mu(R) = 1$). Of course, by Proposition I.2.10, the spaces $L^1 + L^\infty$ and L^1 also coincide and have equivalent norms (identical if $\mu(R) = 1$). Hence, we have the following result.

Corollary 6.7. *Let X be a rearrangement-invariant space over a finite measure space (R, μ). Then*

$$L^\infty \hookrightarrow X \hookrightarrow L^1. \tag{6.16}$$

Furthermore, if $\mu(R) = 1$ and $\|1\|_X = 1$, then both of the embeddings in (6.16) *have norm* 1.

In the discrete case, it is customary to denote the spaces $L^p(R, \mu)$ by ℓ^p, in which case we have the following analogue of Corollary 6.7. The proof is left as an exercise.

Corollary 6.8. *Suppose (R, μ) is completely atomic, consisting of countably many atoms each with positive measure α. If X is a rearrangement-invariant space over (R, μ), then*

$$\ell^1 \hookrightarrow X \hookrightarrow \ell^\infty. \tag{6.17}$$

Furthermore, if $\alpha = 1$ and the norm in X of the characteristic function of an atom is equal to 1, *then both of the embeddings in* (6.17) *have norm* 1.

We have seen in Corollary 6.5 that the fundamental function of $L^1 + L^\infty$ is equal to $\min(1, t)$. Since this function is concave we can immediately construct the Lorentz Λ- and M- spaces for $L^1 + L^\infty$:

$$\Lambda(L^1 + L^\infty) = L^1 + L^\infty = M(L^1 + L^\infty). \tag{6.18}$$

Hence, in view of Corollary 5.14, $L^1 + L^\infty$ is the only rearrangement-invariant space with fundamental function equal to $\min(1, t)$.

For $L^1 \cap L^\infty$, the situation is slightly different in that the fundamental function, which by Corollary 6.5 is equal to $\max(1, t)$, is not concave. According to Proposition 5.11, however, there is an equivalent norm on $L^1 \cap L^\infty$ which has a concave fundamental function. This is easy to see directly. For if the standard norm

$$\|f\|_{L^1 \cap L^\infty} = \max(\|f\|_{L^1}, \|f\|_{L^\infty}) \tag{6.19}$$

is replaced by the equivalent norm

$$\|f\| = \|f\|_{L^1} + \|f\|_{L^\infty}, \tag{6.20}$$

then the resulting fundamental function is evidently equal to $t + 1$, and this is nothing but the least concave majorant of the original fundamental function $\max(1, t)$. Computing the Λ- and M-spaces for $L^1 \cap L^\infty$, we see once again that both coincide with the original space:

$$\Lambda(L^1 \cap L^\infty) = L^1 \cap L^\infty = M(L^1 \cap L^\infty). \tag{6.21}$$

In particular, $L^1 \cap L^\infty$, with the norm in (6.20), is the only rearrangement-invariant space with fundamental function equal to $1 + t$.

As a final remark, let us note that the associate norm of the $(L^1 + L^\infty)$-norm is given by (6.19) rather than (6.20). Hence, concavity of the fundamental function in X does not guarantee concavity of the fundamental function in the associate space X'.

7. MEASURE-PRESERVING TRANSFORMATIONS

In the finite-dimensional case, where measurable functions may be regarded as finite sequences $(a_1, a_2, \ldots, a_n)$, say, equimeasurability arises from permutation of the underlying point-set $\{1, 2, \ldots, n\}$. Thus, two non-negative sequences $(a_1, a_2, \ldots, a_n)$ and $(b_1, b_2, \ldots, b_n)$ are equimeasurable if and only if there is a permutation σ of $\{1, 2, \ldots, n\}$ such that $b_i = a_{\sigma(i)}$ for $i = 1, 2, \ldots, n$. There are partial results of this nature for more general measure spaces and we shall examine some of them in this section. The notion of permutation is no longer available in this context and is replaced by that of "measure-preserving transformation".

Definition 7.1. Let (R_1, μ_1) and (R_2, μ_2) be totally σ-finite measure spaces. A mapping σ from R_1 into R_2 is said to be a *measure-preserving transformation* if, whenever E is a μ_2-measurable subset of R_2, the set $\sigma^{-1}E = \{x \in R_1 : \sigma(x) \in E\}$ is a μ_1-measurable subset of R_1 and

$$\mu_1(\sigma^{-1}E) = \mu_2(E). \tag{7.1}$$

That measure-preserving transformations induce equimeasurability is shown by the following result.

Proposition 7.2. *Let $\sigma : R_1 \to R_2$ be a measure-preserving transformation. If f_2 is a nonnegative μ_2-measurable function on R_2, then the function $f_1 = f_2 \circ \sigma$ is a nonnegative μ_1-measurable function on R_1 and f_1, f_2 are equimeasurable.*

Proof. For any real λ, the sets $E_i = \{f_i > \lambda\}$, $(i = 1, 2)$, satisfy $E_1 = \sigma^{-1}E_2$. Hence, since σ is measure-preserving, the μ_2-measurability of f_2 implies the μ_1-measurability of f_1. Moreover, it follows from (7.1) that $\mu_1(E_1) = \mu_2(E_2)$ and hence that f_1 and f_2 are equimeasurable. ■

The converse is false, as Example 1.6(c) plainly shows. However, there are some results of a positive nature along these lines. We begin with the following fundamental lemma.

Lemma 7.3 (G. G. Lorentz). *Let (R, μ) be a finite nonatomic measure space. Let f be a μ-measurable function on R, let g be a Lebesgue-measurable function on $[0, \mu(R)]$, and suppose f and g are nonnegative and equimeasurable. If C is any set of real numbers for which $g^{-1}(C)$ is Lebesgue measurable, then $f^{-1}(C)$ is μ-measurable and*

$$\mu(f^{-1}(C)) = m(g^{-1}(C)). \tag{7.2}$$

Proof. That (7.2) holds for all closed sets C follows easily from the fact that f and g are equimeasurable. Suppose now that C is any set for which $A = g^{-1}(C)$ is measurable. If $\varepsilon > 0$, then by Lusin's theorem, there exists a closed subset A_0 of A, on which g is continuous, such that $m(A) < m(A_0) + \varepsilon$. The set $C_0 = g(A_0)$ is closed so, by the observation made above, the set $B_0 = f^{-1}(C_0)$ is μ-measurable and $\mu(B_0) = m(g^{-1}(C_0)) \geq m(A_0)$. Now $B_0 \subset B = f^{-1}(C)$, so

$$m(A) < m(A_0) + \varepsilon \leq \mu(B_0) + \varepsilon \leq \mu_*(B) + \varepsilon,$$

where $\mu_*(B)$ denotes the inner μ-measure of B [Ro, pp. 274–280]. Letting

$\varepsilon \to 0$, we obtain $m(A) \leq \mu_*(B)$. A similar argument shows that $m(A^c) \leq \mu_*(B^c)$. Hence,

$$\mu_*(B) + \mu_*(B^c) \geq m(A) + m(A^c) = m([0, \mu(R)]) = \mu(R),$$

and from this it follows that B is μ-measurable and $\mu(B) = m(A)$. ■

Proposition 7.4. *Let (R, μ) be a finite nonatomic measure space with $\mu(R) = b$. Then for each real number a, there exists a measure-preserving transformation σ from R onto the interval $[a, a + b]$.*

Proof. There is an increasing family of μ-measurable subsets E_t of R such that $\mu(E_t) = t$ for $0 \leq t \leq b$ (cf. Exercise 17). Define a mapping $\sigma : R \to [0, b]$ by setting

$$\sigma(x) = \inf\{t : x \in E_t\}, \qquad (x \in R). \tag{7.3}$$

We claim that σ is measure-preserving. To see this, let $g(t) = t$, $(0 \leq t \leq b)$, and let $\lambda \geq 0$. Then

$$\{x : \sigma(x) > \lambda\} \subset \{x : x \notin E_\lambda\} \subset \{x : \sigma(x) > \lambda\} \cup \{x : \sigma(x) = \lambda\}.$$

But

$$\{x : \sigma(x) = \lambda\} = \bigcap_{n=1}^{n} (E_{\lambda + 1/n} \setminus E_{\lambda - 1/n})$$

so

$$\begin{aligned} \mu(\{x : \sigma(x) = \lambda\}) &= \lim_{n\to\infty} \mu(E_{\lambda + 1/n} \setminus E_{\lambda - 1/n}) \\ &= \lim_{n\to\infty} \left[\left(\lambda + \frac{1}{n}\right) - \left(\lambda - \frac{1}{n}\right)\right] = 0. \end{aligned}$$

Hence,

$$\mu\{x : \sigma(x) > \lambda\} = \mu\{x : x \notin E_\lambda\} = b - \lambda = m\{t : g(t) > \lambda\},$$

and this shows that σ and g are equimeasurable. If C is any Lebesgue-measurable subset of $[0, b]$, then $g^{-1}(C) = C$ is Lebesgue measurable. Hence, by Lemma 7.3, the set $\sigma^{-1}(C)$ is μ-measurable and

$$\mu(\sigma^{-1}(C)) = m(g^{-1}(C)) = m(C).$$

This shows that σ is a measure-preserving transformation of R onto $[0, b]$. Composing this with the translation $x \to x + a$ gives the desired measure-preserving transformation from R onto $[a, a + b]$. ■

Theorem 7.5 (J. V. Ryff). *Let (R, μ) be a finite nonatomic measure space and let f be a nonnegative μ-measurable function on R. Then there is a measure-preserving transformation $\sigma: R \to (0, \mu(R))$ such that $f = f^* \circ \sigma$ μ-a.e.*

Proof. By redefining f, if necessary, on a set of μ-measure zero, we may assume that f and f^* have the same range, say Ω. The equimeasurability of f and f^*, together with the fact that $\mu(R)$ is finite, guarantees that the sets

$$E_\lambda = \{x \in R : f(x) = \lambda\}, \qquad I_\lambda = \{t \in (0, \mu(R)) : f^*(t) = \lambda\}$$

satisfy

$$\mu(E_\lambda) = m(I_\lambda), \qquad (\lambda \in \Omega).$$

Since f^* is decreasing, each I_λ is either a singleton or an interval. In the former case, let $\sigma_\lambda : E_\lambda \to I_\lambda$ be the mapping which carries every point of E_λ into the unique value of t, say t_λ, for which $f^*(t) = \lambda$. The latter case, where I_λ is an interval, can occur for at most countably many values of λ. The set of such λ we denote by Γ. For each $\lambda \in \Gamma$, let σ_λ be a measure-preserving transformation from E_λ onto I_λ, the existence of which is guaranteed by Proposition 7.4. Now define a map $\sigma: R \to (0, \mu(R))$ by

$$\sigma(x) = \sigma_\lambda(x), \qquad (x \in E_\lambda;\ \lambda \in \Omega).$$

It is clear that $f = f^* \circ \sigma$, so it remains only to show that σ is measure-preserving. Let F be any measurable subset of $(0, \mu(R))$. We may write F as the disjoint union

$$F = \bigcup_{\lambda \in \Gamma} (F \cap I_\lambda) \cup A, \qquad A = \bigcup_{\lambda \in \Omega \setminus \Gamma} (F \cap I_\lambda).$$

Then $\sigma^{-1} F$ is the disjoint union

$$\sigma^{-1} F = \bigcup_{\lambda \in \Gamma} \sigma^{-1}(F \cap I_\lambda) \cup \sigma^{-1} A. \tag{7.4}$$

Now, for each $\lambda \in \Gamma$, the map $\sigma_\lambda : E_\lambda \to I_\lambda$ is measure-preserving. Hence, $\sigma^{-1}(F \cap I_\lambda) = \sigma_\lambda^{-1}(F \cap I_\lambda)$ is μ-measurable and

$$\mu(\sigma^{-1}(F \cap I_\lambda)) = m(F \cap I_\lambda), \qquad (\lambda \in \Gamma). \tag{7.5}$$

On the other hand, the function f^* is one-to-one on A. Thus, if we set $C = f^*(A)$ and $B = f^{-1}(C)$, then $A = (f^*)^{-1}(C)$ and $B = \sigma^{-1} A$. Since A is evidently measurable, we see from Lemma 7.3 that B is μ-measurable and $\mu(B) = m(A)$. In other words,

$$\mu(\sigma^{-1} A) = m(A). \tag{7.6}$$

Combining (7.4), (7.5), and (7.6), we see that $\sigma^{-1} F$ is μ-measurable and

$$\mu(\sigma^{-1} F) = \sum_{\lambda \in \Gamma} \mu(\sigma^{-1}(F \cap I_\lambda)) + \mu(\sigma^{-1} A)$$
$$= \sum_{\lambda \in \Gamma} m(F \cap I_\lambda) + m(A) = m(F).$$

Hence, σ is measure-preserving and the proof is complete. ∎

Example 1.6(c) shows that the preceding result fails if $\mu(R)$ is infinite. This difficulty is not a serious one, however, and can be overcome by adding the hypothesis that $f^*(t) \to 0$ as $t \to \infty$. The exact statement is as follows.

Corollary 7.6. *Let (R, μ) be a resonant measure space and let f be a nonnegative μ-measurable function on R for which $\lim_{t \to \infty} f^*(t) = 0$. Then there is a measure-preserving transformation σ from the support of f onto the support of f^* such that $f = f^* \circ \sigma$ μ-a.e. on the support of f.*

Proof. Suppose first that (R, μ) is nonatomic. The result has already been established in Theorem 7.5 for finite measure spaces so we may assume $\mu(R) = \infty$. In the same vein, if $f^*(t) = 0$ for some $t > 0$, then f is supported on a set of finite μ-measure and the result again reduces to the previous one. Hence, we may assume that $f^*(t) > 0$ for all $t > 0$. Our hypotheses on f^* are thus

$$\lim_{t \to \infty} f^*(t) = 0, \qquad 0 < f^*(t) \le \infty, \qquad (0 < t < \infty). \tag{7.7}$$

We define an increasing sequence $(t_n)_{n=0}^\infty$ of nonnegative numbers inductively as follows. Let $t_0 = 0$. It follows from (7.7) that there is a number s satisfying $0 < s < \infty$ for which $f^*(s) < f^*(0+)$. Let $t_1 = \inf\{t : f^*(t) \le f^*(s)\}$. Then $t_1 < \infty$ and the right-continuity of f^* implies

$$f^*(t_1) \le f^*(s) < f^*(0+).$$

Hence, $0 = t_0 < t_1 < \infty$. Suppose now that $t_0 < t_1 < \cdots < t_n$ have been defined and set

$$t_{n+1} = \inf\left\{t : f^*(t) \le \frac{1}{2} f^*(t_n)\right\}.$$

Again, it follows from (7.7) that t_{n+1} is finite, and from the right-continuity of f^* that $t_{n+1} > t_n$. Hence, we have $t_0 < t_1 < \cdots < t_n < t_{n+1} < \infty$ and the induction is complete.

Let $I_0 = (0, t_1)$ and $I_n = [t_n, t_{n+1})$, $(n = 1, 2, \ldots)$. The infinite interval

$(0,\infty)$ may thus be expressed as the countable union $\bigcup_{n=0}^{\infty} I_n$ of the disjoint finite intervals I_n. Observe that t belongs to I_n if and only if

$$\frac{1}{2} f^*(t_n) < f^*(t) \leq f^*(t_n), \qquad (n = 1, 2, \ldots),$$

and t belongs to I_0 if and only if $f^*(t) > f^*(t_1)$.

Let

$$E_0 = \{x \in R : f(x) > f^*(t_1)\}$$

and

$$E_n = \left\{x \in R : \frac{1}{2} f^*(t_n) < f(x) \leq f^*(t_n)\right\}, \qquad (n = 1, 2, \ldots).$$

It follows from the equimeasurability of f and f^* that each E_n has finite measure $\mu(E_n) = m(I_n) = t_{n+1} - t_n$, $(n = 0, 1, 2, \ldots)$. Furthermore, for $n = 0, 1, 2, \ldots$,

$$(f\chi_{E_n})^*(t) = (f^*\chi_{I_n})(t + t_n) = ((f^*\chi_{I_n}) \circ \tau_{t_n})(t) \tag{7.8}$$

where τ_{t_n} is the translation $\tau_{t_n}(t) = t + t_n$. By Theorem 7.5, there is a measure-preserving transformation $\sigma_n : E_n \to [0, m(I_n))$ such that $f\chi_{E_n} = (f\chi_{E_n})^* \circ \sigma_n$. Hence, by (7.8),

$$f\chi_{E_n} = (f^*\chi_{I_n}) \circ (\tau_{t_n} \circ \sigma_n). \tag{7.9}$$

Define a mapping σ from the support of f, say R', onto $(0, \infty)$ by

$$\sigma(x) = (\tau_{t_n} \circ \sigma_n)(x), \qquad (x \in E_n, \quad n = 0, 1, 2, \ldots). \tag{7.10}$$

If $x \in R'$, then x belongs to one and only one E_n, in which case

$$\sigma_n(x) \in [0, m(I_n)) = [0, t_{n+1} - t_n)$$

and

$$\sigma(x) = (\tau_{t_n} \circ \sigma_n)(x) = \sigma_n(x) + t_n$$

belongs to $[t_n, t_{n+1}) = I_n$. Hence, from (7.9),

$$f(x) = (f\chi_{E_n})(x) = (f^*\chi_{I_n})(\sigma_n(x) + t_n) = f^*(\sigma(x)).$$

Since each σ_n is measure-preserving, it is clear from the definition (7.10) that σ is measure-preserving, and with this the proof is complete in the nonatomic case.

The atomic case is easier. If each of the (countably many) atoms has positive measure k, say, then these can be put into one-to-one correspondence with the intervals $[nk, (n+1)k)$, $(n = 0, 1, 2, \ldots)$, in the following way. Let $a_1 > a_2 > \cdots > a_m > \cdots$ be the distinct values in the range of f^*, arranged in decreasing order. The hypothesis that $f^*(t) \to 0$ as $t \to \infty$ ensures that

each of the values a_m is assumed by f^* on at most a finite number, say j_m, of the intervals $[nk, (n+1)k)$. Since f and f^* are equimeasurable, the function f assumes the value a_m on exactly j_m atoms. Let σ be any one-to-one correspondence of these j_m atoms onto the corresponding j_m intervals, for $m = 1, 2, \ldots$. The result is a measure-preserving transformation of the support of f onto the support of f^* such that $f = f^* \circ \sigma$ on the support of f. ■

The preceding results show, under suitably general conditions, that f may be reproduced from its decreasing rearrangement f^* by means of a measure-preserving transformation σ (so that $f = f^* \circ \sigma$). The reverse procedure is, unfortunately, not always possible. That is, there are functions f for which it is impossible to expresss f^* in the form $f^* = f \circ \tau$ with τ measure-preserving. This is demonstrated by the following example.

***Example** 7.7.* Let (R, μ) be the measure space $(0, 1)$ with Lebesgue measure. Let $f(t) = 1 - 2t \pmod 1$ so $f^*(t) = 1 - t$, $(0 < t < 1)$ (cf. Figure 6).

Let σ be the measure-preserving transformation given by $\sigma(t) = 2t \pmod 1$ (cf. Figure 7).

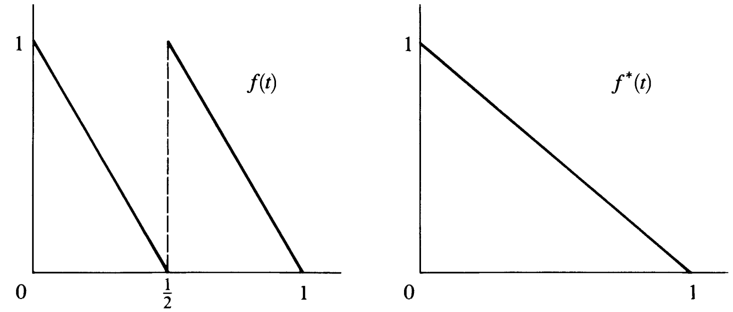


Figure 6. Graphs of f and f^*.

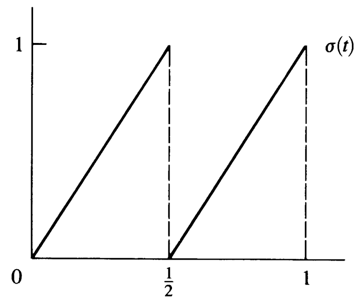


Figure 7. Graph of σ.

It is easy to check that $f = f^* \circ \sigma$ a.e. on $(0,1)$. Thus, if we suppose that there is a measure-preserving transformation τ with $f^* = f \circ \tau$ a.e. on $(0,1)$, then $f^* = (f^* \circ \sigma) \circ \tau$ a.e. on $(0,1)$. Since f^* is one-to-one, it follows that $(\sigma \circ \tau)(t) = t$ a.e. on $(0,1)$. By redefining τ to be $\tau(t) = 1/2$ on the exceptional set of measure zero, we have $\sigma(\tau(t)) = \sigma(t/2) = t$ for these values of t and hence we may assume that $(\sigma \circ \tau)(t) = t$ for all $t \in (0,1)$. In that case, the mapping τ is one-to-one from $(0,1)$ onto its range $A = \tau(0,1)$. In particular, $\tau^{-1}: A \to (0,1)$ exists. Note that τ^{-1} is simply the restriction of σ to A.

We claim that A is measurable. To see this, let

$$J = \tau^{-1}\left(\left(0, \frac{1}{2}\right)\right), \qquad K = \tau^{-1}\left(\left[\frac{1}{2}, 1\right)\right).$$

Since τ is measure-preserving, both J and K are measurable and $m(J) = m(K) = 1/2$. Now if $x \in J$, then $\tau(x) \in (0, 1/2)$ so $x = \sigma(\tau(x)) = 2\tau(x)$. Thus $\tau|_J$ is just the linear map $x \to x/2$ restricted to J. Hence, since $m(J) = 1/2$, it follows that $\tau(J)$ is measurable and $m(\tau(J)) = 1/4$. Similarly, if $x \in K$, then $\tau(x) \in [1/2, 1)$ so $x = \sigma(\tau(x)) = 2\tau(x) - 1$. The map $\tau|_K$ is therefore $x \to (1+x)/2$ restricted to K, so again we deduce that $\tau(K)$ is measurable and $m(\tau(K)) = 1/4$. Now A is the disjoint union of $\tau(J)$ and $\tau(K)$ so A is measurable and has measure $m(A) = 1/4 + 1/4 = 1/2$. Hence, since τ is measure-preserving, we obtain the contradiction that

$$m(0,1) = m(\tau^{-1}(A)) = m(A) = \frac{1}{2}.$$

Thus, it is impossible to have $f^* = f \circ \tau$ with τ measure-preserving.

As a further consequence of Ryff's theorem, we have the following result.

Corollary 7.8. *Let (R, μ) be a nonatomic totally σ-finite measure space. Then every nonnegative, decreasing, right-continuous function h on $(0, \infty)$ is the decreasing rearrangement of some μ-measurable function f on R.*

Proof. It is clear that there exists a μ-measurable function g on R with the property that $g(x)$ is never zero and $g^*(t) \to 0$ as $t \to \infty$. By Corollary 7.6, there is a measure-preserving transformation σ from R onto $(0, \mu(R))$ such that $g = g^* \circ \sigma$. Let $f = h \circ \sigma$. Proposition 7.2 shows that f and h are equimeasurable so $f^* = h^*$. But the hypotheses on h ensure that $h = h^*$. Hence $f^* = h$ and the proof is complete. ∎

EXERCISES AND FURTHER RESULTS FOR CHAPTER 2

1. Let (R, μ) be a totally σ-finite measure space and suppose $f \in L^1(R, \mu)$. Then

$$\int_R [|f| - \lambda]^+ \, d\mu = \int_0^\infty [f^* - \lambda]^+ \, dt,$$

for all $\lambda > 0$, where $[x]^+ = \max(x, 0)$. Furthermore,

$$\mu_f(\lambda) = -\frac{d}{d\lambda} \int_R [|f| - \lambda]^+ \, d\mu, \qquad (\lambda > 0)$$

and

$$f^{**}(t) = \inf_{\lambda \geq 0} \left\{ \lambda + \frac{1}{t} \int_R [|f| - \lambda]^+ \, d\mu \right\}, \qquad (t > 0).$$

2. If $f, g \in L^1(R, \mu)$, then for all $t > 0$,

$$t|f^{**}(t) - g^{**}(t)| \leq \int_R \big||f| - |g|\big| \, d\mu \leq \int_R |f - g| \, d\mu.$$

In particular, if $|f_n| \to |f|$ in L^1-norm, then $f_n^*(t) \to f^*(t)$ at every point of continuity of $f^*(t)$. (HINT: Fix $t, \varepsilon > 0$ and, for each $n = 1, 2, \ldots$, use Exercise 1 to show that

$$\frac{1}{\varepsilon} \int_t^{t+\varepsilon} f^* - e_n \leq f_n^*(t) \leq \frac{1}{\varepsilon} \int_{t-\varepsilon}^t f^* + e_n,$$

where $e_n = (2/\varepsilon) \int_R \big||f| - |f_n|\big|$; cf. A. M. Garsia & E. Rodemich [1]).

3. Let Φ be a nonnegative, increasing, left-continuous function on $[0, \infty)$ with $\Phi(0+) = 0$ and $\Phi(\infty) = \infty$. Let (R, μ) be a totally σ-finite measure space and let f belong to $\mathcal{M}_0(R, \mu)$. Then

$$\int_R \Phi(|f|) \, d\mu = \int_0^\infty \Phi(f^*) \, dt$$

(HINT: Approximate $|f|$ by an increasing sequence of simple functions).

4. If Φ is a convex function satisfying the conditions of Exercise 3, then Φ may be represented as the integral of some nonnegative increasing function ϕ:

$$\Phi(t) = \int_0^t \phi(\lambda) \, d\lambda, \qquad (t > 0);$$

equivalently, there is a positive measure ν on $\mathbf{R}^+$ such that

$$\Phi(t) = \int_0^\infty [t - \lambda]^+ \, d\nu(\lambda), \qquad (t > 0).$$

Hence

$$\int_R \Phi(|f|) \, d\mu = \int_0^\infty \phi(\lambda) \mu_f(\lambda) \, d\lambda = \int_R \int_0^\infty [|f| - \lambda]^+ \, d\nu \, d\mu.$$

5. The following conditions are equivalent:

(i) $f \prec g$;
(ii) for every convex Φ (as in the preceding exercise),

$$\int_R \Phi(|f|)\,d\mu \le \int_R \Phi(|g|)\,d\mu;$$

(iii) for every $\lambda \ge 0$,

$$\int_R [|f| - \lambda]^+\,d\mu \le \int_R [|g| - \lambda]^+\,d\mu$$

(HINT: Exercises 1, 4; cf. G. H. Hardy, J. E. Littlewood & G. Pólya [1, §249], A. M. Garsia & E. Rodemich [1], A. W. Marshall & I. Olkin [1]).

6. The Hardy-Littlewood inequality (2.1) asserts that the sum $\sum_{i=1}^n a_i b_i$ is greatest when both sequences (a_i) and (b_i) are arranged in decreasing order. The sum is smallest when one sequence is arranged in increasing order and the other in decreasing order (G. H. Hardy, J. E. Littlewood & G. Pólya [1, p. 261]).

7. The inequality $|f_1| \le |f_2|$ a.e. implies $f_1^* \le f_2^*$, which in turn implies $f_1^{**} \le f_2^{**}$. Show by example that these implications can not be reversed.

8. If (R, μ) is totally σ-finite, then

$$\int_0^\infty f^*(t)g^*(t)\,dt = \sup\left\{\int_R |f\tilde{g}|\,d\mu : \tilde{g} \prec g\right\}$$

and the supremum is attained if $\mu(R)$ is finite. (HINT: Treat the nonatomic case first and use the method of retracts to establish the result in general).

9. (C. Herz [1]) (a) If $f \in \mathcal{M}_0^+(R,\mu)$, let $\mathscr{A}(f)$ denote the set of positive functions ϕ defined on $(0,\infty)$ with the properties:

(i) $t\phi(t)$ is concave,
(ii) for each subset E of R with $0 < \mu(E) < \infty$, the inequality $(1/\mu(E))\int_E f\,d\mu \le \phi(\mu(E))$ holds.

Then f^{**} is the minimal element of $\mathscr{A}(f)$.
(b) Use this result to give an alternative proof of the subaddivity of the maximal operator $f \to f^{**}$ (Theorem 3.4). (HINT: Show that $\phi = f^{**} + g^{**}$ belongs to $\mathscr{A}(f+g)$).

10. The operator $f \to f^*$ is not in general subadditive. Neither is it submultiplicative. However, the partial integrals satisfy

$$\int_0^t (fg)^*(s)\,ds \le \int_0^t f^*(s)g^*(s)\,ds, \qquad (t \ge 0).$$

11. (H. D. Ruderman [1]) (a) Let (a_{ij}), $(i = 1,\ldots,m,\ j = 1,\ldots,n)$, be an $m \times n$ matrix with nonnegative entries a_{ij}. For each i, let a_{ij}^* denote the entries a_{ij} of the i-th row arranged in decreasing order. Then

$$\sum_{j=1}^n \prod_{i=1}^m a_{ij} \le \sum_{j=1}^n \prod_{i=1}^m a_{ij}^*$$

and

$$\prod_{j=1}^n \sum_{i=1}^m a_{ij} \ge \prod_{j=1}^n \sum_{i=1}^m a_{ij}^*.$$

(b) Use either of these inequalities to derive the inequality

$$nb_1b_2\cdots b_n \le b_1^n + b_2^n + \cdots + b_n^n$$

relating the geometric and arithmetic means for arbitrary nonnegative numbers $b_1, b_2, \ldots, b_n$. (HINT: It may be assumed without loss of generality that $b_1 \ge b_2 \ge \cdots \ge b_n \ge 0$. Let (a_{ij}) be the $n \times n$ matrix

$$\begin{bmatrix} b_1 & b_2 & \cdots & b_{n-1} & b_n \\ b_2 & b_3 & \cdots & b_n & b_1 \\ \vdots & \vdots & & \vdots & \vdots \\ b_n & b_1 & \cdots & b_{n-2} & b_{n-1} \end{bmatrix}.$$

12. (G. G. Lorentz [3]) (a) Let $\Phi = \Phi(x, u_1, u_2, \ldots, u_n)$ be a continuous function of the variables $0 < x < 1$, $u_i \ge 0$, and suppose that Φ has continuous second order partial derivatives with respect to all variables. Then in order that the inequality

$$\int_0^1 \Phi(x, f_1(x), \ldots, f_n(x))\,dx \le \int_0^1 \Phi(t, f_1^*(t), \ldots, f_n^*(t))\,dt$$

hold for all nonnegative bounded functions $f_1, f_2, \ldots, f_n$ on $[0,1]$, it is necessary and sufficient that

$$\frac{\partial^2\Phi}{\partial u_i \partial u_j} \ge 0 \qquad \text{and} \qquad \frac{\partial^2\Phi}{\partial x \partial u_i} \le 0$$

for all i and j (cf. also Ky Fan & G. G. Lorentz [1] for a related result).
(b) The above conditions hold if $\Phi(x, u_1, \ldots, u_n) = u_1 u_2 \cdots u_n$, so

$$\int_0^1 f_1 f_2 \cdots f_n\,dx \le \int_0^1 f_1^* f_2^* \cdots f_n^*\,dt.$$

(c) If $\Phi(x, u_1, u_2, \ldots, u_n) = F(u_1 + u_2 + \cdots + u_n)$, then the conditions above hold if and only if F is convex. This is true of the function $F(u) = -\log u$, in which case the inequality in part (a), interpreted for sums instead of integrals, reduces to the second of Ruderman's inequalities in the preceding exercise.

13. (G. G. Lorentz & T. Shimogaki [1, 3]; cf. also Chapter III.7) (a) If f and g are nonnegative integrable functions on $(0,1)$, then $f^* - g^* \prec f - g$ (use this result to give a direct proof of the first inequality in Exercise 2).
(b) Let f be a nonnegative decreasing function on $(0,1)$, and suppose $f = f_1 + f_2$ with $f_1, f_2 \ge 0$. Then there are nonnegative decreasing functions f_i' with the properties $f_i' \prec f_i$, $(i = 1, 2)$, and $f = f_1' + f_2'$.
(c) Let f be a nonnegative integrable function on $(0,1)$, and let g_1 and g_2 be nonnegative simple functions on $(0,1)$. If $g_1 + g_2 \prec f$, then there exist nonnegative functions f_1 and f_2 on $(0,1)$, with disjoint supports, such that $f = f_1 + f_2$ and $g_i \prec f_i$, $(i = 1, 2)$.

14. If (R, μ) is nonatomic and has infinite measure, then the representation $\rho \to \bar{\rho}$ in the Luxemburg representation theorem (Theorem 4.10) is unique.

15. If (R, μ) is resonant, construct a corresponding measure space $(\tilde{R}, \tilde{\mu})$ as follows:

(i) if (R, μ) is nonatomic, let $(\tilde{R}, \tilde{\mu})$ consist of the interval $[0, \mu(R))$ with Lebesgue measure;

(ii) if (R, μ) is completely atomic, with each of the countably many atoms having the same positive measure, say α, let $\tilde{R}$ denote the collection of intervals $[(k-1)\alpha, k\alpha)$, $(1 \le k \le \mu(R)/\alpha)$, and declare each of these intervals to be a $\tilde{\mu}$-atom of measure α. Now establish the following version of the Luxemburg representation theorem.

(a) If λ is a rearrangement-invariant function norm over $(\tilde{R}, \tilde{\mu})$, then the functional $\underset{\sim}{\lambda}$ defined by

$$\underset{\sim}{\lambda}(f) = \lambda(f^*), \qquad (f \in \mathscr{M}_0^+(R, \mu))$$

is a rearrangement-invariant function norm over (R, μ).

(b) If ρ is a rearrangement-invariant function norm over (R, μ), then there exists a *unique* rearrangement-invariant function norm $\tilde{\rho}$ over $(\tilde{R}, \tilde{\mu})$ such that

$$\rho(f) = \tilde{\rho}(f^*), \qquad (f \in M_0^+(R, \mu)).$$

(c) The following identities hold:

$$(\tilde{\rho})_\sim = \rho, \qquad (\underset{\sim}{\lambda})^\sim = \lambda,$$

$$(\tilde{\rho})' = (\rho')^\sim, \qquad (\underset{\sim}{\lambda})' = (\lambda')_\sim.$$

(d) There is a one-to-one correspondence between the rearrangement-invariant function norms over (R, μ) and those over $(\tilde{R}, \tilde{\mu})$.

16. (W. A. J. Luxemburg [3]) Let (R, μ) be an arbitrary totally σ-finite measure space. A function norm ρ over (R, μ) is said to be *universally rearrangement-invariant* if

$$f_1 \prec f_2 \quad \Rightarrow \quad \rho(f_1) \le \rho(f_2), \qquad (f_1, f_2 \in \mathscr{M}_0^+(R, \mu)),$$

and, in that case, the corresponding Banach function space $X = X(\rho)$ is said to be *universally rearrangement-invariant*.

(a) Every universally rearrangement-invariant space over (R, μ) is rearrangement-invariant, and the converse holds if (R, μ) is resonant.

(b) If λ is a rearrangement-invariant function norm over $(\mathbf{R}^+, m)$, then the functional $\underset{\sim}{\lambda}$ defined in Exercise 15(a) is a universally rearrangement-invariant function norm over (R, μ).

(c) If ρ is a function norm over (R, μ) whose associate norm ρ' is universally rearrangement-invariant, then there exists a rearrangement-invariant function norm λ over $(\mathbf{R}^+, m)$ such that $\rho = \underset{\sim}{\lambda}$.

(d) A function norm ρ over (R, μ) is universally rearrangement-invariant if and only if its associate norm ρ' is.

(e) A function norm ρ over (R, μ) is universally rearrangement-invariant if and only if $\rho = \underset{\sim}{\lambda}$ for some universally rearrangement-invariant function norm λ over $(\mathbf{R}^+, m)$.

17. Let (R, μ) be a finite nonatomic measure space. Then there is an increasing family $\{E_t : 0 \le t \le \mu(R)\}$ of μ-measurable subsets of R such that $\mu(E_t) = t$, $(0 \le t \le \mu(R))$. (HINT: Assume, without loss of generality, that $\mu(R) = 1$. Subdivide R into two subsets of measure $1/2$, then each of these into two subsets of measure $1/4$, and so on. Define E_t to be the union of certain of these sets (indexed in an appropriate way) according to the appearance of a 0 or a 1 in the corresponding place in the dyadic expansion of the number t).

18. Suppose $f \in L^2(\mathbf{R})$ and let $f_0(x) = xf(x)$. Then

$$\|f\|_{L^1} \le \{8\|f\|_{L^2}\|f_0\|_{L^2}\}^{1/2}$$

(HINT: If $E = (-\alpha, \alpha)$, then

$$\int_{\mathbf{R}} |f| = \int_E |f| + \int_{E^c} |f_0(t)| \left|\frac{1}{t}\right| dt \le (2\alpha)^{1/2}\|f\|_2 + \left(\frac{2}{\alpha}\right)^{1/2} \|f_0\|_2;$$

now minimize over α).

19. If a nonnegative measurable function f with the property that $f^*(+\infty) = 0$ assumes a constant value $\beta > 0$ on a set E of finite measure, then there is an interval I with $|I| = |E|$ such that $f^* = \beta$ on I. The result may fail if $f^*(+\infty) > 0$.

20. (a) (G. F. D. Duff [1], [2], [3]; K. M. Chong [1]) Given $a_1, \ldots, a_n \in \mathbf{R}$, let $A_k = a_{k+1} - a_k$ and $B_k = a_{k+1}^* - a_k^*$. Then $B \prec A$. If f is absolutely continuous on a finite interval, then $(f^*)' \prec f'$ (cf. also J. V. Ryff [2]).

(b) (J. Steiner [1, Vol. II, p. 265]) The *symmetric decreasing rearrangement* f^s of a measurable function f on the interval $(-a, a)$, $(0 < a \le \infty)$, is defined on $(-a, a)$ by $f^s(x) = f^*(2|x|)$, where f^* is the decreasing rearrangement. Thus, f, f^s and f^* are equimeasurable. Suppose f is a nonnegative continuous function on $[-1, 1]$, vanishing at both endpoints, with a continuous derivative on $(-1, 1)$ which vanishes at only finitely many points. Then the arc length of $y = f(x)$ exceeds the arc length of $y = f^s(x)$, with equality iff $f = f^s$. (HINT: Express the arc length integrals in terms of the endpoints of the intervals constituting the sets $\{f > y\}$ and $\{f^s > y\}$ and analyze the resulting inequality; cf. G. H. Hardy, J. E. Littlewood & G. Pólya [1, §404]).

(c) (H. A. Schwarz [1]) Define a two-dimensional "radial" decreasing rearrangement f^r by $f^r(x, y) = f^*(\pi(x^2 + y^2))$. For suitably regular nonnegative f, the area of the surface $z = f(x, y)$ exceeds the area of $z = f^r(x, y)$ (cf. G. H. Hardy, J. E. Littlewood & G. Pólya [1, §405]; for further results of this type, cf. also A. M. Garsia & E. Rodemich [1]).

(d) If f, g, h are nonnegative measurable functions on $\mathbf{R}$, then

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x)g(y)h(-x-y)\,dx\,dy \le \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f^s(x)g^s(y)h^s(-x-y)\,dx\,dy$$

(cf. G. H. Hardy, J. E. Littlewood & G. Pólya [1, Ch. X]).

21. It is often useful to have a rearrangement which respects the sign of the function. For an integrable function f on $[-\pi, \pi]$, let $\tilde{\mu}_f$ denote the distribution function $\tilde{\mu}_f(\lambda) = |\{f > \lambda\}|$ (no absolute values on f), and define the *signed symmetric decreasing rearrangement* $\tilde{f}^s$ by $\tilde{f}^s(x) = \inf\{\lambda : \tilde{\mu}_f(\lambda) \le 2x\}$. The *Baernstein* *function* f^B is defined by

$$f^B(x) = \int_{-x}^{x} \tilde{f}^s(y)\,dy = \sup_{|E|=2x} \int_E f, \qquad (0 \le x \le \pi).$$

A. Baernstein [1] has shown that if $u(re^{i\theta}) = u_r(\theta)$ is harmonic in the annulus $r_1 < r < r_2$, then u^B (where $u^B(re^{i\theta}) = (u_r)^B(\theta)$ for each r) is subharmonic in the semi-annulus $\{re^{i\theta} : r_1 < r < r_2, 0 < \theta < \pi\}$; among the consequences are deep extremal properties for univalent functions f in the disk, which arise from estimates such as $(\log|f|)^B \le (\log|k|)^B$, where k is the Koebe function $k(z) = z/(1-z)^2$.

22. (a) (Steffensen's inequality) Let $f, g \in L^1(0,a)$ with $f = f^*$ and $0 \le g \le 1$. Then

$$\int_{a-c}^{a} f \le \int_0^a fg \le \int_0^c f, \qquad \text{where } c = \int_0^a g$$

(HINT: For the second of the inequalities, let E be any measurable subset of $[0,a]$ with $|E| = c$ and show $g \prec \chi_E$; apply Hardy's lemma (Proposition 3.6) and the Hardy-Littlewood inequality (Theorem 2.2)).

(b) (Brunk-Olkin inequality) Suppose $1 \ge w_1 \ge \cdots \ge w_n \ge 0$ and $a_1 \ge \cdots \ge a_n \ge a_{n+1} = 0$. If f is convex on $[0, a_1]$, then

$$\left\{1 - \sum_{j=1}^{n} (-1)^{j-1} w_j\right\} f(0) + \sum_{j=1}^{n} (-1)^{j-1} w_j f(a_j) \ge f\left(\sum_{j=1}^{n} (-1)^{j-1} w_j a_j\right)$$

(HINT: Apply Steffensen's inequality to f' and g, where g assumes the value $w_1 - w_2 + \cdots + (-1)^{j-1} w_j$ on $[a_{j+1}, a_j)$ for $j = 1, 2, \ldots, n$). For these and further inequalities of this type, see K. M. Chong & N. M. Rice [1, §11], E. F. Beckenbach & R. Bellman [1], and A. W. Marshall & I. Olkin [1]).

NOTES FOR CHAPTER 2

The decreasing rearrangement received its first systematic treatment in the 1934 book of G. H. Hardy, J. E. Littlewood, and G. Pólya [1], which includes references to its use as far back as the early 1880's in the work of J. Steiner [1] and H. A. Schwarz [1] (cf. Exercise 20). The decreasing rearrangement plays a fundamental role in the 1930 paper by G. H. Hardy and J. E. Littlewood [2] on the maximal function. Detailed accounts have also been given by A. Grothendieck [1], W. A. J. Luxemburg [3], G. G. Lorentz [4], [8], J. V. Ryff [1], [2], P. W. Day [1], [2] and K. M. Chong & N. M. Rice [1]. A Baernstein has made extensive use of the symmetric decreasing rearrangement in function-theoretic applications; cf. A. Baernstein [1], [2] and the references cited there (cf. also Exercise 21).

The material in §§2-4 follows closely the accounts by W. A. J. Luxemburg [3], P. W. Day [1], and K. M. Chong & N. M. Rice [1]. The Hardy-Littlewood-Pólya relation (Definition 3.5) is something of a misnomer, having been

introduced by R. F. Muirhead [1] but generally attributed in the literature to the former authors because of their detailed treatment in [1]. It has been studied by numerous authors in connection with diverse fields; an extensive survey is provided in the book by A. W. Marshall & I. Olkin [1]. For the history of the rearrangement-invariant spaces of §4 and the Lorentz spaces of §5, we refer to the references of the preceding paragraph and to the Notes for Chapter 1 and the references cited there. Theorem 5.13 and properties of the fundamental function are due to E. M. Semenov [1]. The spaces $L^1 + L^\infty$ and $L^1 \cap L^\infty$ were studied by G. G. Gould [1] and W. A. J. Luxemburg & A. C. Zaanen [3], whose results constitute most of §6. The main results of §7, namely, Lemma 7.3 and Theorem 7.5, are due to G. G. Lorentz [4, p. 60] and J. V. Ryff [1], [2].

3 Interpolation of Operators on Rearrangement-Invariant Spaces

The preceding chapters have provided us with the rudiments of the theory of rearrangement-invariant Banach function spaces. Now we are in a position to consider the interpolation of operators on these spaces. The basic concepts of interpolation theory are purely functional-analytic in nature. Thus, we begin in Section 1 in the context of general Banach spaces. This will allow us to fix notation and to define the theory in a format that will be useful also for subsequent chapters.

The fundamental interpolation theorem for rearrangement-invariant spaces is established in Section 2. It asserts on the one hand that every rearrangement-invariant Banach function space is an interpolation space between L^1 and L^∞, and on the other that every Banach function space which is an interpolation space between L^1 and L^∞ is rearrangement-invariant.

Section 3 contains an analysis of the Hardy-Littlewood maximal operator M. The main result is Theorem 3.8, which establishes the equivalence of the functions $(Mf)^*$ and f^{**}. The direct part of this assertion is given by the inequality

$$(Mf)^*(t) \le c\frac{1}{t}\int_0^t f^*(s)\,ds, \qquad (0 < t < \infty).$$

A closely related inequality is established in Section 4 for the Hilbert transform H:

$$(Hf)^*(t) \le c\left\{\frac{1}{t}\int_0^t f^*(s)\,ds + \int_t^\infty f^*(s)\frac{ds}{s}\right\}, \qquad (0 < t < \infty).$$

The presence of the integrals over $(0,t)$ reflects the fact that both M and H are unbounded on L^1. The additional integral over (t,∞) for the Hilbert transform arises from the unboundedness of H on L^∞; that term is absent in the expression for the maximal operator M, which is bounded on L^∞. These kinds of estimates lie at the heart of *weak-type interpolation* and the Marcinkiewicz interpolation theorem. We shall refer to them here simply as *weak-type inequalities*.

In section 5, such weak-type inequalities are instituted in a formal definition of a concept of *weak type* (specifically, joint weak type $(p,p;q,q)$). The fundamental weak-type interpolation theorems of A. P. Calderón (Theorem 5.7) and D. W. Boyd (Theorem 5.16), characterize the joint weak-type $(p,p;q,q)$ interpolation spaces. The latter theorem is formulated in terms of a pair of parameters (the Boyd indices) which are associated with each rearrangement-invariant space. The boundedness of the Hardy-Littlewood maximal operator and of the Hilbert transform on rearrangement-invariant spaces can be fully described in terms of the Boyd indices of the spaces involved (cf. Theorems 5.17 and 5.18).

The corresponding results for the conjugate-function operator, which is the periodic analog of the Hilbert transform, are derived in Section 6. There are close connections with Fourier series. In particular, the norm-convergence of Fourier series in rearrangement-invariant spaces is completely described in terms of the Boyd indices (Corollary 6.11).

In Section 7, we consider two fundamental results of G. G. Lorentz and T. Shimogaki involving the Hardy-Littlewood-Pólya relation. The first (Theorem 7.4) deals with decreasing rearrangements and differences of functions. The second (Thereom 7.7) involves a certain "splitting" of measurable functions with respect to the Hardy-Littlewood-Pólya relation. The latter result, of interest in its own right, will play a crucial role in the characterization of monotone interpolation spaces in Chapter V.

1. INTERPOLATION SPACES.

We have seen in the last chapter that the sum $L^1 + L^\infty$ and the intersection $L^1 \cap L^\infty$ of L^1 and L^∞ are respectively the largest and the smallest of the rearrangement-invariant Banach function spaces. Because of this, they will play a special role in the interpolation theory to be developed in this chapter. Since much of the framework we shall need does not depend on the special properties of L^1 and L^∞, however, we shall develop it in the more abstract setting of Banach spaces.

Definition 1.1. A pair (X_0, X_1) of Banach spaces X_0 and X_1 is called a *compatible couple* if there is some Hausdorff topological vector space, say $\mathscr{X}$, in which each of X_0 and X_1 is continuously embedded.

Note that (L^1, L^∞) is a compatible couple because both L^1 and L^∞ are continuously embedded in the Hausdorff space $\mathscr{M}_0$ of measurable functions that are finite a.e. (Theorem I.1.4). Moreover, any pair (X, Y) of Banach spaces for which X is continuously embedded in Y (or vice versa) is a compatible couple because then we may choose for the Hausdorff space $\mathscr{X}$ the space Y itself.

Definition 1.2. Let (X_0, X_1) be a compatible couple, with corresponding Hausdorff space $\mathscr{X}$. Let $X_0 + X_1$ denote the *sum* of X_0 and X_1, that is, the set of elements x in $\mathscr{X}$ that are representable in the form $x = x_0 + x_1$ for some x_0 in X_0 and x_1 in X_1. For each x in $X_0 + X_1$, set

$$\|x\|_{X_0+X_1} = \inf\{\|x_0\|_{X_0} + \|x_1\|_{X_1} : x = x_0 + x_1\}, \tag{1.1}$$

where the infimum extends over all representations $x = x_0 + x_1$ of x with x_0 in X_0 and x_1 in X_1. For each element x in the intersection $X_0 \cap X_1$ of X_0 and X_1, set

$$\|x\|_{X_0 \cap X_1} = \max\{\|x\|_{X_0}, \|x\|_{X_1}\}. \tag{1.2}$$

Observe that when (X_0, X_1) is taken to be the compatible couple (L^1, L^∞), these definitions of the sum $L^1 + L^\infty$ and the intersection $L^1 \cap L^\infty$ agree with those given previously (Definition II.6.1). Notice also that $X_0 + X_1$ and $X_0 \cap X_1$ (and their norms) do not depend on the particular choice of Hausdorff space $\mathscr{X}$ associated with the couple (X_0, X_1).

Theorem 1.3. *If (X_0, X_1) is a compatible couple, then $X_0 + X_1$ and $X_0 \cap X_1$ are Banach spaces under the norms* (1.1) *and* (1.2), *respectively.*

Proof. It is evident that $X_0 \cap X_1$ is a normed linear space when furnished with the functional (1.2). In order to see that the same is true for $X_0 + X_1$ we need only show that the functional in (1.1) vanishes only on the zero element, since the other properties of the norm are clear. But if $\|x\|_{X_0+X_1} = 0$, then for each natural number n, there exist elements x_0^n in X_0 and x_1^n in X_1 such that

$$x = x_0^n + x_1^n \tag{1.3}$$

and

$$\|x_0^n\|_{X_0} + \|x_1^n\|_{X_1} \leq \|x\|_{X_0+X_1} + \frac{1}{n} = \frac{1}{n}.$$

The last estimate implies that $(x_0^n)_{n=1}^\infty$ converges to 0 in X_0, hence also in $\mathscr{X}$, since X_0 is continuously embedded in $\mathscr{X}$. Similarly, $(x_1^n)_{n=1}^\infty$ converges to 0 in $\mathscr{X}$. But then $(x_0^n + x_1^n)_{n=1}^\infty$ converges to 0 in $\mathscr{X}$, so we conclude from (1.3) that $x = 0$.

It remains only to establish completeness for both spaces. For $X_0 \cap X_1$, it follows immediately from (1.2) that if $(x_n)_{n=1}^\infty$ is a Cauchy sequence in $X_0 \cap X_1$, then it is Cauchy also in X_0 and in X_1. Hence, by completeness of these spaces, there are elements y_0 in X_0 and y_1 in X_1 such that

$$\|x_n - y_0\|_{X_0} \to 0, \qquad \|x_n - y_1\|_{X_1} \to 0. \tag{1.4}$$

Using again the fact that X_0 and X_1 are continuously embedded in $\mathscr{X}$, we deduce that $x_n \to y_0$ in $\mathscr{X}$ and $x_n \to y_1$ in $\mathscr{X}$, and hence, since $\mathscr{X}$ is Hausdorff, that $y_0 = y_1$. In particular, the element $y = y_0 = y_1$ belongs to $X_0 \cap X_1$, and it follows from (1.2) and (1.4) that $x_n \to y$ in $X_0 \cap X_1$. Hence, $X_0 \cap X_1$ is complete.

To establish the completeness of $X_0 + X_1$, it will suffice to show that every absolutely convergent series in $X_0 + X_1$ is convergent in $X_0 + X_1$. To this end, let $(x^n)_{n=1}^\infty$ be any sequence of elements of $X_0 + X_1$ satisfying $\sum_{n=1}^\infty \|x^n\|_{X_0+X_1} < \infty$. Each x^n is representable as a sum $x^n = x_0^n + x_1^n$ of elements x_i^n in X_i, $(i = 0, 1)$, satisfying

$$\|x_0^n\|_{X_0} + \|x_1^n\|_{X_1} < \|x^n\|_{X_0+X_1} + 2^{-n}.$$

Then, for $i = 0, 1$, the series $\sum_n \|x_i^n\|_{X_i}$ also converges and hence, by the completeness of X_i, there is an element y_i in X_i to which $\sum_n x_i^n$ converges in X_i. Thus, as $N \to \infty$,

$$\left\| \sum_{n=1}^N x_i^n - y_i \right\|_{X_i} \to 0, \qquad (i = 0, 1). \tag{1.5}$$

The element $y = y_0 + y_1$ evidently belongs to $X_0 + X_1$ and we have from (1.1),

$$\left\| \sum_{n=1}^N x^n - y \right\|_{X_0+X_1} \leq \left\| \sum_{n=1}^N x_0^n - y_0 \right\|_{X_0} + \left\| \sum_{n=1}^N x_1^n - y_1 \right\|_{X_1}.$$

Letting $N \to \infty$, we see therefore from (1.5) that $\sum x^n$ converges in $X_0 + X_1$ to y. Hence $X_0 + X_1$ is complete. ■

Definition 1.4. If (X_0, X_1) is a compatible couple, then a Banach space X is said to be an *intermediate space* between X_0 and X_1 if X is continuously embedded between $X_0 \cap X_1$ and $X_0 + X_1$:

$$X_0 \cap X_1 \hookrightarrow X \hookrightarrow X_0 + X_1. \tag{1.6}$$

Clearly, X_0 and X_1 are always intermediate spaces for the couple (X_0, X_1). Also, Theorem II.6.7 shows that every rearrangement-invariant Banach function space (over a resonant measure space) is an intermediate space for the couple (L^1, L^∞).

We turn our attention now to the operators on these spaces. We shall denote by $\mathscr{B}(X, Y)$ (or $\mathscr{B}(X)$, if $X = Y$) the space of bounded linear operators from a Banach space X into a Banach space Y. The space $\mathscr{B}(X, Y)$ is itself a Banach space under the operator norm

$$\|T\|_{\mathscr{B}(X,Y)} = \sup\{\|Tx\|_Y : \|x\|_X \leq 1\}.$$

It will be convenient to use the same symbol T to denote an operator in $\mathscr{B}(X, Y)$ as well as the restriction of the operator to a linear subspace of X. We shall also use the terminology that "T is a bounded operator on X" to mean that T is a bounded operator from X into itself, that is, belongs to $\mathscr{B}(X)$.

In order to describe the interpolation property, we shall need to consider operators that are simultaneously bounded on each component X_0 and X_1 of a compatible couple (X_0, X_1). By this, we shall mean that the operators are *a priori* defined on the sum $X_0 + X_1$ and their restrictions to X_0 and to X_1 are bounded in an appropriate sense. The precise definition is as follows.

Definition 1.5. Let (X_0, X_1) and (Y_0, Y_1) be two compatible couples and let T be a linear operator defined on $X_0 + X_1$ and taking values in $Y_0 + Y_1$. Then T is said to be *admissible* with respect to the couples (X_0, X_1) and (Y_0, Y_1) if, for each $i = 0, 1$, the restriction of T to X_i maps X_i into Y_i and, furthermore, is a bounded operator from X_i into Y_i:

$$\|Tx\|_{Y_i} \leq \|T\|_{\mathscr{B}(X_i,Y_i)} \|x\|_{X_i}, \qquad (x \in X_i). \tag{1.7}$$

The class of admissible operators is denoted by

$$\mathscr{A} = \mathscr{A}(X_0, X_1; Y_0, Y_1).$$

The norm of an admissible operator T is given by

$$\|T\|_{\mathscr{A}} = \max_{i=0,1} \|T\|_{\mathscr{B}(X_i,Y_i)}. \tag{1.8}$$

Although we did not assume in Definition 1.5 that the operators in question are *bounded* operators from $X_0 + X_1$ into $Y_0 + Y_1$, this property is in fact an immediate consequence of the hypotheses embodied in (1.7). More precisely, we have the following result.

Proposition 1.6. *Every admissible operator T is a bounded operator from $X_0 + X_1$ into $Y_0 + Y_1$, and*

$$\|T\|_{\mathscr{B}(X_0+X_1, Y_0+Y_1)} \leq \|T\|_{\mathscr{A}}. \tag{1.9}$$

Proof. Suppose T is admissible and let x be an arbitrary element of $X_0 + X_1$. Let $x = x_0 + x_1$ be any representation of x as a sum of elements x_i in X_i, $(i = 0, 1)$. Then by (1.1), (1.7), and (1.8), we have

$$\begin{aligned}\|Tx\|_{Y_0+Y_1} &= \|Tx_0 + Tx_1\|_{Y_0+Y_1} \leq \|Tx_0\|_{Y_0} + \|Tx_1\|_{Y_1} \\ &\leq \|T\|_{\mathscr{B}(X_0,Y_0)}\|x_0\|_{X_0} + \|T\|_{\mathscr{B}(X_1,Y_1)}\|x_1\|_{X_1} \\ &\leq \|T\|_{\mathscr{A}}(\|x_0\|_{X_0} + \|x_1\|_{X_1}).\end{aligned}$$

Taking the infimum over all representations $x = x_0 + x_1$ of x, we see from (1.1) that

$$\|Tx\|_{Y_0+Y_1} \leq \|T\|_{\mathscr{A}}\|x\|_{X_0+X_1}.$$

This shows that T is a bounded operator from $X_0 + X_1$ into $Y_0 + Y_1$ and that (1.9) holds. ■

Proposition 1.7. *The class $\mathscr{A} = \mathscr{A}(X_0, X_1; Y_0, Y_1)$ of admissible operators is a Banach space when equipped with the norm* (1.8). *Furthermore, $\mathscr{A}$ is continuously embedded in $\mathscr{B}(X_0 + X_1; Y_0 + Y_1)$.*

Proof. Consider the subset

$$\mathscr{D} = \{(U, V) : U \in \mathscr{B}(X_0, Y_0),\ V \in \mathscr{B}(X_1, Y_1),\ U = V \text{ on } X_0 \cap X_1\}$$

of the product space $\mathscr{B}(X_0, Y_0) \times \mathscr{B}(X_1, Y_1)$, equipped with the product norm

$$\|(U, V)\|_{\mathscr{B}(X_0,Y_0)\times\mathscr{B}(X_1,Y_1)} = \max\{\|U\|_{\mathscr{B}(X_0,Y_0)}, \|V\|_{\mathscr{B}(X_1,Y_1)}\}.$$

We claim that $\mathscr{D}$ is a closed subspace of the product space and hence is a Banach space in its own right. To see this, suppose $(U_n, V_n)_{n=1}^{\infty}$ is a sequence in $\mathscr{D}$ that converges in the product norm to an element (U, V) in $\mathscr{B}(X_0, Y_0) \times \mathscr{B}(X_1, Y_1)$. Then

$$\|U_n - U\|_{\mathscr{B}(X_0,Y_0)} \to 0, \qquad \|V_n - V\|_{\mathscr{B}(X_1,Y_1)} \to 0 \tag{1.10}$$

as $n \to \infty$, and to prove our assertion we have only to show that (U, V) belongs to $\mathscr{D}$, that is, $U = V$ on $X_0 \cap X_1$. If x is an arbitrary element of $X_0 \cap X_1$, we may write

$$Ux - Vx = (Ux - U_n x) + (V_n x - Vx) \tag{1.11}$$

because $U_n = V_n$ on $X_0 \cap X_1$. But (1.10) shows that $Ux - U_n x \to 0$ in Y_0, and hence in $Y_0 + Y_1$. Similarly, $V_n x - Vx \to 0$ in $Y_0 + Y_1$, and so we deduce from (1.11) that $Ux = Vx$. Hence, $U = V$ on $X_0 \cap X_1$ and so (U, V) belongs to $\mathscr{D}$. As noted previously, this establishes that $\mathscr{D}$ is a Banach space.

The proof will thus be complete if $\mathscr{A}$ is shown to be isometrically isomorphic to $\mathscr{D}$. It is clear that the mapping $T \to (U, V)$, where U and V are the restrictions of T to X_0 and to X_1, respectively, is a linear mapping of $\mathscr{A}$ into $\mathscr{D}$. The mapping is obviously isometric, and it is injective because if the restrictions U and V both vanish identically, then T evidently vanishes on all of $X_0 + X_1$. Thus, it remains only to show that the mapping is surjective. Suppose then that (U, V) is an arbitrary element of $\mathscr{D}$. Define a corresponding operator T from $X_0 + X_1$ into $Y_0 + Y_1$ as follows. If x is an arbitrary element of $X_0 + X_1$, let $x = x_0 + x_1$ be any representation of x as a sum of elements x_0 in X_0 and x_1 in X_1. Set

$$Tx = Ux_0 + Vx_1. \tag{1.12}$$

Notice that if $x = x_0' + x_1'$ is any other representation of x, then $x_0 - x_0' = x_1' - x_1$ belongs to $X_0 \cap X_1$. Accordingly, since U and V coincide on the intersection, we have

$$U(x_0 - x_0') = V(x_0 - x_0') = V(x_1' - x_1),$$

and so $Ux_0 + Vx_1 = Ux_0' + Vx_1'$. This shows that the right-hand side of (1.12) is independent of the particular choice of representation of x, and hence that the operator T is well-defined on $X_0 + X_1$.

It is clear that T is a linear mapping of $X_0 + X_1$ into $Y_0 + Y_1$ and that its restrictions to X_0 and to X_1 coincide with U and V, respectively. Hence, T is admissible, and is transformed into the pair (U, V) in $\mathscr{D}$ by the mapping under consideration. This establishes surjectivity and hence completes the proof. ■

In the preamble to the definition of the admissible operators, we remarked that in order to formulate the interpolation property, we should need to consider operators that are *simultaneously* bounded on X_0 and on X_1. One approach was that adopted in Definition 1.5, namely, to consider operators that are *a priori* defined on the larger space $X_0 + X_1$ and then to require that their restrictions on X_0 and X_1 be bounded. Another possibility might be to begin with *two* operators U (in $\mathscr{B}(X_0, Y_0)$) and V (in $\mathscr{B}(X_1, Y_1)$) satisfying the consistency requirement that $U = V$ on $X_0 \cap X_1$. The preceding proof shows that the outcome is the same: every such pair (U, V) arises in one and only one way as the pair of restrictions to X_0 and X_1 of some admissible operator T on $X_0 + X_1$.

Now we can formulate the interpolation property.

Definition 1.8. Let (X_0, X_1) and (Y_0, Y_1) be two compatible couples. Let X and Y be intermediate spaces of the couples (X_0, X_1) and (Y_0, Y_1), respectively. The pair (X, Y) is said to have the *interpolation property* (or to be an *interpolation pair*) relative to (X_0, X_1) and (Y_0, Y_1) if every admissible operator maps X into Y.

It is clear that the pairs (X_0, Y_0) and (X_1, Y_1) are interpolation pairs, as are the pairs $(X_0 \cap X_1, Y_0 \cap Y_1)$ and $(X_0 + X_1, Y_0 + Y_1)$. The basic problem, of course, will be to construct more interesting examples and this will be one of our main tasks in subsequent chapters. For the remainder of this section, however, we shall content ourselves with exploring some general properties of interpolation pairs. The following lemma will prove useful.

Lemma 1.9. *Let A, B, C, and D be Banach spaces with A and B continuously embedded in C and D, respectively. Let L be a bounded linear operator from C into D. If L also maps A into B, then L is in fact a bounded operator from A into B.*

$$\begin{tikzcd} C \arrow[r, "L"] & D \\ A \arrow[u] \arrow[r, dashed, "L"] & B \arrow[u] \end{tikzcd}$$

Proof. We use the closed graph theorem. Suppose $a_n \to a$ in A and $La_n \to b$ in B. The continuous embeddings of A into C and B into D guarantee that $a_n \to a$ in C and $La_n \to b$ in D. But then the continuity of L from C into D implies that $La_n \to La$ in D, and so we conclude that $La = b$. Hence, by the closed graph theorem, L is a bounded operator from A into B. ■

Our first application of the lemma will be to establish the *boundedness* of admissible operators on interpolation pairs.

Proposition 1.10. *Let (X, Y) be an interpolation pair relative to the compatible couples (X_0, X_1) and (Y_0, Y_1). Then every admissible operator is a bounded operator from X into Y.*

Proof. Let T be an arbitrary admissible operator, that is, T belongs to $\mathscr{A}(X_0, X_1; Y_0, Y_1)$. Then T is a bounded operator from $X_0 + X_1$ into $Y_0 + Y_1$ (Proposition 1.6) and maps X into Y (Definition 1.8). Furthermore, X and Y are intermediate spaces (Definition 1.8) for the couples (X_0, X_1) and (Y_0, Y_1), respectively, and so, in particular, are continuously embedded in $X_0 + X_1$ and $Y_0 + Y_1$, respectively (Definition 1.4).

$$\begin{tikzcd} X_0 + X_1 \arrow[r, "T"] & Y_0 + Y_1 \\ X \arrow[u] \arrow[r, dashed, "T"] & Y \arrow[u] \end{tikzcd}$$

It therefore follows directly from Lemma 1.9 that T is a bounded operator from X into Y. ■

The operation of restricting an admissible operator to the component X of an interpolation pair (X, Y) may thus be regarded as a mapping from the Banach space $\mathscr{A}$ of admissible operators into the Banach space $\mathscr{B}(X, Y)$. The next result, which also is a consequence of Lemma 1.9, shows that this mapping is continuous.

Proposition 1.11. *Let (X, Y) be an interpolation pair for the couples (X_0, X_1) and (Y_0, Y_1). Then there is a constant C, depending only on the spaces involved, such that*

$$\|T\|_{\mathscr{B}(X,Y)} \leq C\|T\|_{\mathscr{A}}, \tag{1.13}$$

for every admissible operator T.

Proof. Let $R(T)$ denote the restriction of an operator T in $\mathscr{B}(X_0 + X_1, Y_0 + Y_1)$ to the interpolation space X. Since X is continuously embedded in $X_0 + X_1$, it is clear that R is a bounded linear operator from $\mathscr{B}(X_0 + X_1, Y_0 + Y_1)$ into $\mathscr{B}(X, Y_0 + Y_1)$, and, as we remarked above, R maps $\mathscr{A}$ into $\mathscr{B}(X, Y)$. Furthermore, the space $\mathscr{A}$ is continuously embedded in $\mathscr{B}(X_0 + X_1, Y_0 + Y_1)$ (Proposition 1.7), and the space $\mathscr{B}(X, Y)$ is continuously embedded in $\mathscr{B}(X, Y_0 + Y_1)$ (because $Y \hookrightarrow Y_0 + Y_1$).

$$\begin{tikzcd} \mathscr{B}(X_0 + X_1, Y_0 + Y_1) \arrow[r, "R"] & \mathscr{B}(X, Y_0 + Y_1) \\ A \arrow[u] \arrow[r, dashed, "R"] & \mathscr{B}(X, Y) \arrow[u] \end{tikzcd}$$

Hence, an application of Lemma 1.9 shows that R is a bounded operator from $\mathscr{A}$ into $\mathscr{B}(X, Y)$. ■

The interpolation pairs (X, Y) for which the constant C in (1.13) is equal to 1 will be of special interest. We make the following definition.

Definition 1.12. Let (X, Y) be an interpolation pair for the couples (X_0, X_1) and (Y_0, Y_1). If

$$\|T\|_{\mathscr{B}(X,Y)} \leq \|T\|_{\mathscr{A}}, \tag{1.14}$$

for every admissible operator T, then (X, Y) is said to be an *exact* interpolation pair.

Proposition 1.13. *In any interpolation pair (X, Y), the component X can be equivalently renormed so that the resulting interpolation pair is exact.*

Proof. For each x in X, let

$$|x|_X = \max(\|x\|_X, \sup\{\|Tx\|_Y : \|T\|_{\mathscr{A}} \leq 1\}). \tag{1.15}$$

It is clear that $|\cdot|_X$ is a norm on X, and that

$$\|x\|_X \leq |x|_X \leq \max(C, 1)\|x\|_X, \qquad (x \in X),$$

where C is the constant in (1.13). Hence, $|\cdot|_X$ and $\|\cdot\|_X$ are equivalent norms on X. Since every admissible operator T with $\|T\|_{\mathscr{A}} \leq 1$ evidently satisfies $\|Tx\|_Y \leq |x|_X$, for all x in X, it follows that $(X, |\cdot|_X)$ and $(Y, \|\cdot\|_Y)$ constitute an exact interpolation pair. ■

We shall often be interested in the case where the couple (Y_0, Y_1) coincides with (X_0, X_1). It is therefore natural to consider intermediate spaces X for which (X, X) is an interpolation pair. Such spaces will be referred to simply as *interpolation spaces.*

Definition 1.14. An intermediate space X of a compatible couple (X_0, X_1) is said to be an *interpolation space* for (X_0, X_1) (or an interpolation space between X_0 and X_1) if every admissible operator T (i.e., $T \in \mathscr{A}(X_0, X_1; X_0, X_1)$) maps X into itself. An interpolation space X is said to be *exact* if the pair (X, X) is exact in the sense of Definition 1.12, that is, if

$$\|T\|_{\mathscr{B}(X)} \leq \|T\|_{\mathscr{A}} \tag{1.16}$$

for every admissible operator T.

Proposition 1.13 shows that there is an equivalent norm $|\cdot|_X$ on any interpolation space X such that $(X, |\cdot|_X, X, \|\cdot\|_X)$ is an exact interpolation pair. Since two different norms are involved, this does not guarantee that X itself is an exact interpolation space. We conclude this section by showing that $(X, |\cdot|_X, X, |\cdot|_X)$ is also exact and hence that $(X, |\cdot|_X)$ is an exact interpolation space.

Proposition 1.15. *Every interpolation space X can be equivalently renormed so as to become exact.*

Proof. Let $(X, \|\cdot\|_X)$ be the given interpolation space and define $|\cdot|_X$ as in (1.15). Since the identity operator on $X_0 + X_1$ is evidently admissible and has $\mathscr{A}$-norm at most 1, it is clear that (1.15) reduces in this case to

$$|x|_X = \sup\{\|Tx\|_X : \|T\|_{\mathscr{A}} \leq 1\} \tag{1.17}$$

Thus, if S is any admissible operator with $\|S\|_{\mathscr{A}} \leq 1$, we have

$$|Sx|_X = \sup\{\|TSx\|_X : \|T\|_{\mathscr{A}} \leq 1\}.$$

But TS is again an admissible operator of norm at most 1, so we see from (1.17) that $|Sx|_X \leq |x|_X$. Hence S, as a bounded operator from $(X, |\cdot|_X)$ into itself, has norm at most 1. This establishes (1.16) and completes the proof. ■

2. INTERPOLATION BETWEEN L^1 AND L^∞

We turn now to the problem of describing the interpolation spaces for the couple (L^1, L^∞). We begin by showing that every rearrangement-invariant Banach function space is an exact interpolation space for (L^1, L^∞). The crux of the argument is the following characterization of the admissible operators in terms of the Hardy-Littlewood-Pólya relation (Definition II.3.5). As usual, we shall work throughout with a σ-finite measure space.

Proposition 2.1. *Let T be a linear operator defined on $L^1 + L^\infty$ and taking values in $\mathscr{M}_0$. Then T is admissible if and only if there is a constant C such that*

$$Tf \prec Cf \tag{2.1}$$

for all f in $L^1 + L^\infty$. Furthermore, the least constant C for which (2.1) *holds is $C = \|T\|_{\mathscr{A}}$.*

Proof. Suppose first that T is admissible. If f is an arbitrary function in $L^1 + L^\infty$, let $f = g + h$ be any representation of f as a sum of functions g in L^1 and h in L^∞. Then $Tf = Tg + Th$ belongs to $L^1 + L^\infty$ and, by Theorem II.6.2, for any $t > 0$,

$$\begin{aligned}\int_0^t (Tf)^*(s)\,ds &\le \|Tg\|_{L^1} + t\|Th\|_{L^\infty}\\ &\le \|T\|_{\mathscr{A}}(\|g\|_{L^1} + t\|h\|_{L^\infty}).\end{aligned}$$

Taking the infimum over all representations $f = g + h$ of f and appealing once again to Theorem II.6.2, we find that

$$\int_0^t (Tf)^*(s)\,ds \le \|T\|_{\mathscr{A}} \int_0^t f^*(s)\,ds,$$

and this establishes (2.1) with $C = \|T\|_{\mathscr{A}}$.

Suppose conversely that (2.1) holds for some constant C, that is,

$$\int_0^t (Tf)^*(s)\,ds \le C\int_0^t f^*(s)\,ds, \qquad (t > 0). \tag{2.2}$$

Taking $t = 1$, we see from Theorem II.6.5 that Tf belongs to $L^1 + L^\infty$ whenever f does, so T is a linear operator on $L^1 + L^\infty$. Letting $t \to \infty$ in (2.2), we see that T is bounded on L^1 with operator norm at most C, and dividing by t and letting $t \to 0$ in (2.2), we see that T is bounded on L^∞ with operator norm at most C. Hence, T is admissible and $\|T\|_{\mathscr{A}} \le C$. ∎

Theorem 2.2. *Let X be a rearrangement-invariant Banach function space over a resonant measure space. Then X is an exact interpolation space between L^1 and L^∞.*

Proof. It follows from Theorem II.6.7 that X is an intermediate space for the couple (L^1, L^∞). Furthermore, Proposition 2.1 shows that any admissible operator T satisfies $Tf \prec f\|T\|_{\mathscr{A}}$ for every f in $L^1 + L^\infty$. Hence, by Corollary II.4.7, the function Tf belongs to X whenever f does, and

$$\|Tf\|_X \le \|T\|_{\mathscr{A}}\|f\|_X, \qquad (f \in X).$$

In particular,

$$\|T\|_{\mathscr{B}(X)} \le \|T\|_{\mathscr{A}}$$

for every admissible operator T and so X is an exact interpolation space for (L^1, L^∞). ∎

The converse, to the effect that if a Banach function space is an exact interpolation space, then it must be rearrangement-invariant, lies deeper and its proof will occupy the remainder of this section.

Recall that an operator whose norm is at most 1 is called a *contraction*, and an operator with the property that $Tf \ge 0$ whenever $f \ge 0$ is said to be *positive*.

Definition 2.3. Let T be an admissible operator (for (L^1, L^∞)). If T is a positive contraction on L^1 and on L^∞, then T is said to be a *substochastic operator*.

The following description of substochastic operators results immediately from Proposition 2.1.

Proposition 2.4. *Let T be a positive linear operator defined on $L^1 + L^\infty$ and taking values in $\mathscr{M}_0$. Then T is substochastic if and only if $Tf \prec f$ for all f in $L^1 + L^\infty$.*

We see at once that if $g = Tf$, for some substochastic operator T, then $g \prec f$. We shall need the converse: if $g \prec f$ (for nonnegative f and g), then $g = Tf$ for some substochastic operator T. The proof will be broken into several stages corresponding to increasing levels of complexity in the underlying measure space.

We begin with the finite-dimensional case. Thus our functions are n-dimensional vectors $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ with real or complex entries. The corresponding Lebesgue spaces L^1 and L^∞ are then denoted by ℓ_n^1 and ℓ_n^∞, respectively, with respective norms given by

$$\|\mathbf{x}\|_{\ell_n^1} = \sum_{i=1}^n |x_i|, \qquad \|\mathbf{x}\|_{\ell_n^\infty} = \max_{1 \le i \le n} |x_i|.$$

We shall use the same symbol A, say, to denote an $n \times n$ matrix $A = (a_{ij})_{i,j=1}^n$ and the linear operator $\mathbf{x} \to A\mathbf{x}$ corresponding to the matrix in the usual way:

$$(A\mathbf{x})_i = \sum_{j=1}^n a_{ij}x_j, \qquad (i = 1, 2, \ldots, n).$$

Definition 2.5. An $n \times n$ matrix $A = (a_{ij})_{i,j=1}^{n}$ of nonnegative entries a_{ij} is said to be *substochastic* if

(a) $$\sum_{i=1}^{n} a_{ij} \leq 1, \qquad (j = 1, 2, \ldots, n)$$

(b) $$\sum_{j=1}^{n} a_{ij} \leq 1, \qquad (i = 1, 2, \ldots, n).$$

It is not hard to check that condition (a) is equivalent to the requirement that A be a contraction on ℓ_n^1. Similarly, condition (b) holds if and only if A is a contraction on ℓ_n^∞. Hence, in view of Definitions 2.3 and 2.5, we have the following result.

Proposition 2.6. *An $n \times n$ matrix A is substochastic if and only if the corresponding matrix operator A is substochastic.*

Recall that if $\mathbf{a} = (a_1, a_2, \ldots, a_n)$, then we denote by $\mathbf{a}^* = (a_1^*, a_2^*, \ldots, a_n^*)$ the vector whose entries are the values $|a_i|, (i = 1, 2, \ldots, n)$, arranged in decreasing order. In this case, two vectors $\mathbf{a}$ and $\mathbf{b}$ satisfy $\mathbf{b} \prec \mathbf{a}$ (cf. Definition II.3.5) if and only if

$$\sum_{i=1}^{j} b_i^* \leq \sum_{i=1}^{j} a_i^* \tag{2.3}$$

for each $j = 1, 2, \ldots, n$.

Theorem 2.7 (G. H. Hardy, J. E. Littlewood & G. Pólya). *Let* $\mathbf{a}$ *and* $\mathbf{b}$ *be any two n-dimensional vectors with nonnegative entries. Then* $\mathbf{b} \prec \mathbf{a}$ *if and only if there is an $n \times n$ substochastic matrix A such that* $\mathbf{b} = A\mathbf{a}$.

Proof. If A is a substochastic matrix and $\mathbf{b} = A\mathbf{a}$, then it follows directly from Propositions 2.6 and 2.4 that $\mathbf{b} \prec \mathbf{a}$.

Suppose conversely that $\mathbf{b} \prec \mathbf{a}$. We shall construct a substochastic matrix A such that $\mathbf{b} = A\mathbf{a}$. The proof proceeds by induction on the dimension n. The case $n = 1$ is obvious so we suppose $n \geq 2$ and that such matrices can always be constructed in dimensions less than n.

Since permutations of the point-set $\{1, 2, \ldots, n\}$ are evidently substochastic, and products of substochastic matrices are again substochastic, we may assume without loss of generality that the entries of $\mathbf{a}$ and $\mathbf{b}$ are arranged in decreasing order:

$$a_1 \geq a_2 \geq \cdots \geq a_n \geq 0, \qquad b_1 \geq b_2 \geq \cdots \geq b_n \geq 0.$$

The idea of the proof is to replace $\mathbf{a}$ with a substochastic image $\mathbf{a}'$ (i.e., $\mathbf{a}' = A'\mathbf{a}$ for some substochastic matrix A') such that $\mathbf{b} \prec \mathbf{a}'$. The vector $\mathbf{a}'$, however, is to have the additional property that $a_1' = b_1$. The $(n-1)$-dimensional vectors $\boldsymbol{\alpha} = (a_2', a_3', \ldots, a_n')$ and $\boldsymbol{\beta} = (b_2, b_3, \ldots, b_n)$ then also satisfy $\boldsymbol{\beta} \prec \boldsymbol{\alpha}$ so, by the induction hypothesis, there is an $(n-1) \times (n-1)$ substochastic matrix B, say, such that $\boldsymbol{\beta} = B\boldsymbol{\alpha}$. But then the $n \times n$ matrix

$$B' = \begin{bmatrix} 1 & 0 \\ 0 & B \end{bmatrix}$$

is also substochastic and satisfies $\mathbf{b} = B'\mathbf{a}'$. Hence, the matrix $A = B'A'$ is substochastic and we have $A\mathbf{a} = B'(A'\mathbf{a}) = B'\mathbf{a}' = \mathbf{b}$, as desired.

Thus, it remains only to construct the substochastic matrix A' such that $\mathbf{b} \prec \mathbf{a}' = A'\mathbf{a}$ and $b_1 = a_1'$. If $b_1 = a_1$, then we merely take $A' = I$, the identity matrix. Also, if $b_i \leq a_i$ for all i, then we may take for A' the diagonal matrix which transforms $\mathbf{a}$ into $\mathbf{b}$. If neither of these cases holds, then $b_1 < a_1$ (by (2.3)) and there is a smallest index k (necessarily $k > 1$) such that $b_k > a_k$. Hence,

$$\begin{aligned} &b_1 < a_1, \qquad b_j \leq a_j, \qquad (j = 2, 3, \ldots, k-1), \\ &b_k > a_k, \qquad \mathbf{b} \prec \mathbf{a}. \end{aligned} \tag{2.4}$$

Since the entries of $\mathbf{b}$ are in decreasing order, it follows that $a_1 > b_1 \geq b_k > a_k$, so there exists λ with $0 < \lambda < 1$ such that

$$b_1 = \lambda a_1 + (1 - \lambda) a_k. \tag{2.5}$$

Let π denote the $n \times n$ permutation matrix which interchanges the first and the k-th entries of any n-dimensional vector, leaving the others fixed. Now any permutation matrix (including the identity I) is substochastic, hence so is the convex combination

$$A' = \lambda I + (1 - \lambda)\pi. \tag{2.6}$$

If $\mathbf{a}' = A'\mathbf{a}$, it follows from (2.5) and (2.6) that

$$a_1' = \lambda a_1 + (1 - \lambda)(\pi \mathbf{a})_1 = \lambda a_1 + (1 - \lambda) a_k = b_1.$$

Hence, it remains only to show that $\mathbf{b} \prec \mathbf{a}'$. But this follows easily from the facts

$$\begin{aligned} a_1' &= b_1, \qquad a_j' = a_j, \qquad (j \neq 1, k), \\ a_1' + a_k' &= a_1 + a_k, \qquad \mathbf{b} \prec \mathbf{a}, \end{aligned}$$

which themselves are direct consequences of (2.4), (2.5), (2.6), and the definition of π. ■

The next step is to pass, by means of a limiting argument, from the finite-dimensional setting of Theorem 2.7 to the more general setting of an arbitrary nonatomic measure space. For this we shall need the notion of a Banach limit, that is, a positive norm-one linear functional λ on ℓ^∞ with the additional property that $\lambda(\mathbf{a}) = \lim_{n\to\infty} a_n$, whenever $\mathbf{a} = (a_n)_{n=1}^\infty$ is convergent.

Lemma 2.8. *There is a positive linear functional λ on ℓ^∞ such that, for all sequences $\mathbf{a} = (a_n)_{n=1}^\infty$ and $\mathbf{b} = (b_n)_{n=1}^\infty$, the following properties hold:*

$$|\lambda(\mathbf{a})| \le \lambda(|\mathbf{a}|) \le \|\mathbf{a}\|_{\ell^\infty}; \tag{2.7}$$

$$0 \le a_n \le b_n, \quad (n = 1, 2, \ldots) \quad \Rightarrow \quad \lambda(\mathbf{a}) \le \lambda(\mathbf{b}); \tag{2.8}$$

$$\lambda(\mathbf{a}) = \lim_{n\to\infty} a_n, \text{ if } \mathbf{a} \text{ is convergent.} \tag{2.9}$$

Proof. [Ru, p. 82]. ■

Proposition 2.9. *Let (R, μ) be a nonatomic totally σ-finite measure space, and let f and g be nonnegative μ-measurable functions on R. Suppose f belongs to $L^1 + L^\infty$ and $g \prec f$. Then there exists a substochastic operator T such that $Tf = g$ μ-a.e. Moreover, the operator T may be chosen so that*

$$|Th| \le T(|h|), \qquad \mu\text{-a.e.}, \tag{2.10}$$

for every h in $L^1 + L^\infty$.

Proof. Suppose first that g has the form

$$g = \sum_{k=1}^n b_k \chi_{F_k}, \tag{2.11}$$

where $b_k > 0$, $(k = 1, 2, \ldots, n)$, and $\{F_k\}_{k=1}^n$ is a pairwise-disjoint collection of μ-measurable subsets of R of common measure $\varepsilon > 0$. We assume also that

$$\int_0^{k\varepsilon} g^*(s)\,ds < \int_0^{k\varepsilon} f^*(s)\,ds, \qquad (k = 1, 2, \ldots, n) \tag{2.12}$$

(that is, we exclude the possibility of equality which might have arisen had we merely assumed $g \prec f$).

Now (R, μ) is totally σ-finite so by II.(3.7) the function f^{**} can be approximated arbitrarily closely pointwise, hence on any finite set of points, by functions of the form $(f\chi_Q)^{**}$, where Q is a subset of R of finite measure. Hence, there exists a set Q of finite measure at least $n\varepsilon$ such that (2.12) holds with f replaced by $f\chi_Q$, that is,

$$\int_0^{k\varepsilon} g^*(s)\,ds < \int_0^{k\varepsilon} (f\chi_Q)^*(s)\,ds, \qquad (k = 1, 2, \ldots, n). \tag{2.13}$$

Since (R, μ) is nonatomic, the measure space (Q, μ) satisfies the hypotheses of Lemma II.2.5. Hence, there exist sets $G_1 \subset G_2 \subset \cdots \subset G_n \subset Q$ such that $\mu(G_k) = k\varepsilon$ and

$$\int_{G_k} f\,d\mu = \int_{G_k} f\chi_Q\,d\mu = \int_0^{k\varepsilon} (f\chi_Q)^*(s)\,ds, \qquad (k = 1, 2, \ldots, n). \tag{2.14}$$

Observe that each of these numbers is finite because f, hence $f\chi_Q$, belongs to $L^1 + L^\infty$. Now let $E_1 = G_1$ and $E_k = G_k \setminus G_{k-1}$, $(k = 2, 3, \ldots, n)$, so $\mu(E_k) = \varepsilon$ for $k = 1, 2, \ldots, n$. But then it is clear from the construction and (2.14) that the numbers

$$a_k = \frac{1}{\mu(E_k)} \int_{E_k} f\,d\mu, \qquad (k = 1, 2, \ldots, n) \tag{2.15}$$

are finite and are arranged in decreasing order, that is, $\mathbf{a} = \mathbf{a}^* = (a_1^*, a_2^*, \ldots, a_n^*)$. If $\mathbf{b}$ denotes the vector $(b_1, b_2, \ldots, b_n)$ whose components are the coefficients of g (cf. (2.11)), then we see from (2.11), (2.13), (2.14), and (2.15) that

$$\begin{aligned}\sum_{j=1}^k b_j^* &= \frac{1}{\varepsilon}\int_0^{k\varepsilon} g^*\,ds \le \frac{1}{\varepsilon}\int_0^{k\varepsilon} (f\chi_Q)^*\,ds = \frac{1}{\varepsilon}\int_{G_k} f\,d\mu \\ &= \frac{1}{\varepsilon}\sum_{j=1}^k \int_{E_j} f\,d\mu = \sum_{j=1}^k a_j = \sum_{j=1}^k a_j^*,\end{aligned}$$

for each $k = 1, 2, \ldots, n$. Hence, $\mathbf{b} \prec \mathbf{a}$ and so by Theorem 2.7 there is an $n \times n$ substochastic matrix A such that

$$\mathbf{b} = A\mathbf{a}. \tag{2.16}$$

For each function h in $(L^1 + L^\infty)(R, \mu)$, let

$$\alpha_k(h) = \frac{1}{\mu(E_k)} \int_{E_k} h\,d\mu, \qquad (k = 1, 2, \ldots, n). \tag{2.17}$$

Setting $\boldsymbol{\alpha}(h) = (\alpha_1(h), \alpha_2(h), \ldots, \alpha_n(h))$, we define a vector

$$\boldsymbol{\beta}(h) = (\beta_1(h), \beta_2(h), \ldots, \beta_n(h))$$

by means of the identity

$$\boldsymbol{\beta}(h) = A(\boldsymbol{\alpha}(h)). \tag{2.18}$$

We may then define a linear operator T on $L^1 + L^\infty$ by setting

$$Th = \sum_{k=1}^{n} \beta_k(h)\chi_{F_k}, \qquad (h \in L^1 + L^\infty). \tag{2.19}$$

Observe that $\boldsymbol{\alpha}(f) = \mathbf{a}$ by (2.15) and (2.17), and so $\boldsymbol{\beta}(f) = \mathbf{b}$ by (2.16) and (2.18). Hence, a comparison of (2.11) and (2.19) shows immediately that $Tf = g$. Furthermore, the estimates

$$\|Th\|_{L^1(R,\mu)} = \varepsilon\|\boldsymbol{\beta}(h)\|_{\ell_n^1} \leq \varepsilon\|\boldsymbol{\alpha}(h)\|_{\ell_n^1} \leq \|h\|_{L^1(R,\mu)}$$

and

$$\|Th\|_{L^\infty(R,\mu)} = \|\boldsymbol{\beta}(h)\|_{\ell_n^\infty} \leq \|\boldsymbol{\alpha}(h)\|_{\ell_n^\infty} = \|h\|_{L^\infty(R,\mu)}$$

follow easily from the fact that A is substochastic. Hence, T is a substochastic operator and $Tf = g$. Observe, however, from (2.17) that $|\alpha_k(h)| \leq \alpha_k(|h|)$ and so $|\beta_k(h)| \leq \beta_k(|h|)$ because of (2.18) and the fact that the matrix A has nonnegative entries. It follows from this and (2.19) that T satisfies (2.10).

Suppose now that g is an arbitrary nonnegative function satisfying $g \prec f$. Then there exist simple functions g_m, $(m = 1, 2, \ldots)$, each of the form (2.11), such that $0 \leq g_m \uparrow g$. The relation $g \prec f$ implies that $g_m \prec f$ for each m. But we may in fact assume that the sharper relation of the form (2.12) holds for each g_m: to see this, simply replace each g_m by $\gamma_m g_m$, where $(\gamma_m)_{m=1}^\infty$ is a sequence of constants satisfying $0 < \gamma_m < 1$ and $\gamma_m \uparrow 1$. Under these conditions, the proof above shows that there exist substochastic operators T_m, $(m = 1, 2, \ldots)$, satisfying

$$|T_m h| \leq T_m(|h|), \qquad (h \in L^1 + L^\infty) \tag{2.20}$$

and

$$T_m f = g_m, \tag{2.21}$$

for each $m = 1, 2, \ldots$.

The next step is to construct the desired operator T from the sequence $(T_m)_{m=1}^\infty$. So let h be an arbitrary function in $L^1 + L^\infty$ and we shall define Th. Since T_m is substochastic, Proposition 2.4 shows that $T_m h \prec h$. In that case,

$$\int_F T_m(|h|)\,d\mu \leq \int_0^{\mu(F)} T_m(|h|)^*\,ds \leq \int_0^{\mu(F)} h^*\,ds,$$

so using (2.20) we obtain

$$\left|\int_F T_m h\,d\mu\right| \leq \int_F |T_m h|\,d\mu \leq \int_F T_m(|h|)\,d\mu \leq \int_0^{\mu(F)} h^*\,ds \tag{2.22}$$

for every measurable set F and for all $m = 1, 2, \ldots$.

If F has finite measure, the integral on the right of (2.22) is finite because h is in $L^1 + L^\infty$ (cf. Theorem II.6.2). In particular, the sequence $\left(\int_F T_m h\,d\mu\right)_{m=1}^\infty$ is bounded, that is, belongs to ℓ^∞. Hence, if E is a fixed set of finite measure and if λ is a Banach limit (cf. Lemma 2.8), then we may define a set-function $\nu_h = \nu_{h,E}$ on the measurable subsets F of E by setting

$$\nu_h(F) = \lambda\left(\left(\int_F T_m h\,d\mu\right)\right). \tag{2.23}$$

Observe, from (2.22) and from the properties (2.7) and (2.8) of λ, that

$$|\nu_h(F)| \leq \nu_{|h|}(F) \leq \int_0^{\mu(F)} h^*\,ds \tag{2.24}$$

for any measurable subset F of E.

Now ν_h is evidently finitely additive because of the linearity of λ and of the integral. To see that it is in fact countably additive, let F be the union of a countable collection of pairwise-disjoint measurable subsets F_n, $(n = 1, 2, \ldots)$, of E. Then, since F has finite measure, the sets $F'_N = F \setminus \bigcup_{n=1}^N F_n$ satisfy $\mu(F'_N) \to 0$ as $N \to \infty$. By (2.24), this implies $\nu_h(F'_N) \to 0$ as $N \to \infty$. Since

$$\nu_h(F) - \sum_{n=1}^N \nu_h(F_n) = \nu_h(F) - \nu_h\left(\bigcup_{n=1}^N F_n\right) = \nu_h(F'_N),$$

we deduce that ν_h is countably additive and hence that ν_h is a bounded complex measure. Recalling that λ and the operators T_n are positive mappings, we see easily from (2.23) that $\nu_{|h|}$ is a positive measure. Moreover, it is clear from (2.24) that $\nu_{|h|}$ dominates the total variation $|\nu_h|$ of the measure ν_h. Hence,

$$|\nu_h|(E) \leq \nu_{|h|}(E) \leq \int_0^{\mu(E)} h^*\,ds. \tag{2.25}$$

A further obvious consequence of (2.24) is that $\nu_h = \nu_{h,E}$ is absolutely continuous with respect to μ (restricted to E). Hence, by the Radon-Nikodym theorem, there is a μ-integrable function, $\phi_{h,E}$ say, defined on E such that

$$\nu_{h,E}(F) = \int_F \phi_{h,E}\,d\mu \tag{2.26}$$

for every μ-measurable subset F of E.

It is not hard to check that if E_1 and E_2 are any subsets of R having finite measure, then $\phi_{h,E_1} = \phi_{h,E_2}$ μ-a.e. on $E_1 \cap E_2$. Indeed, if G is any measurable subset of $E_1 \cap E_2$, then we have from (2.23) and (2.26),

$$\int_G \phi_{h,E_1}\,d\mu = \nu_{h,E_1}(G) = \lambda\left(\left(\int_G T_m h\,d\mu\right)\right) = \nu_{h,E_2}(G) = \int_G \phi_{h,E_2}\,d\mu,$$

from which the result follows. Since R can be expressed as the union of an increasing sequence of sets of finite measure, this implies that there is a μ-measurable function ϕ_h defined on all of R, whose restriction to any set E of finite measure coincides with $\phi_{h,E}$. Hence, if we set $Th = \phi_h$, then we have from (2.23) and (2.26),

$$\int_E Th\,d\mu = \lambda\left(\left(\int_E T_m h\,d\mu\right)\right) \tag{2.27}$$

for any subset E of R of finite measure.

Now, on E, the function Th is the Radon-Nikodym derivative $\phi_{h,E}$ of $\nu_{h,E}$ with respect to μ, and so its norm in $L^1(E)$ is the total variation $|\nu_h|(E)$ of the measure ν_h. Hence, by (2.25) we have

$$\int_E |Th|\,d\mu = |\nu_h|(E) \le \int_0^{\mu(E)} h^*(s)\,ds \tag{2.28}$$

for any subset E of R of finite measure. Since (R,μ) is certainly resonant (Theorem II.2.7), it follows from this and Proposition II.3.3 that

$$\int_0^t (Th)^*(s)\,ds \le \int_0^t h^*(s)\,ds$$

for every $t > 0$. Hence $Th \prec h$. But T is evidently positive (cf. (2.27)) so we conclude from Proposition 2.4 that T is substochastic.

Notice from (2.25) and (2.28) that

$$\int_E |Th|\,d\mu = |\nu_h|(E) \le \nu_{|h|}(E) = \int_E T(|h|)\,d\mu$$

for any subset E of R having finite measure, and this implies that T satisfies (2.10).

It remains only to show that $Tf = g$ μ-a.e. But since $0 \le g_m \uparrow g$, we see from (2.27), (2.21), the monotone convergence theorem, and (2.9) that

$$\int_E Tf\,d\mu = \lambda\left(\left(\int_E T_m f\,d\mu\right)\right) = \lambda\left(\left(\int_E g_m\,d\mu\right)\right) = \int_E g\,d\mu$$

for every set E of finite measure, from which the result follows. ■

The final step is to remove the restriction that (R,μ) be nonatomic. This is fairly easy to achieve by deploying the method of retracts developed in Section II.3. The final form of the result is therefore as follows.

Theorem 2.10 (A. P. Calderón; J. V. Ryff). *Let (R,μ) be a totally σ-finite measure space, and let f and g be nonnegative μ-measurable functions on R.*

Suppose f belongs to $L^1 + L^\infty$ and $g \prec f$. Then there exists a substochastic operator T such that $Tf = g$ μ-a.e.

Proof. As in Section II.3, we may embed (R,μ) into a nonatomic totally σ-finite measure space $(\bar{R},\bar{\mu})$, and associated with this embedding are mappings $\mathscr{E}_1$ and $\mathscr{E}_2$ satisfying II.(3.11), II.(3.17), and II.(3.18). It is clear from these properties and the hypotheses on f and g that $\mathscr{E}_1 f$ and $\mathscr{E}_1 g$ are nonnegative $\bar{\mu}$-measurable functions on $\bar{R}$, that $\mathscr{E}_1 f$ belongs to $(L^1 + L^\infty)(\bar{R},\bar{\mu})$, and that $\mathscr{E}_1 g \prec \mathscr{E}_1 f$. Since $(\bar{R},\bar{\mu})$ is nonatomic, we may therefore apply Proposition 2.9 to obtain a substochastic operator (relative to $(\bar{R},\bar{\mu})$), say $\bar{T}$, such that

$$\bar{T}\mathscr{E}_1 f = \mathscr{E}_1 g \qquad \bar{\mu}\text{-a.e.} \tag{2.29}$$

For each h in $(L^1 + L^\infty)(R,\mu)$, set

$$Th = \mathscr{E}_2 \bar{T} \mathscr{E}_1 h.$$

The operator T so defined is evidently positive. Furthermore, by using II.(3.11), II.(3.18), and Proposition 2.4 applied to $\bar{T}$, we obtain

$$\begin{aligned}\int_0^t (Th)^*(s)\,ds &= \int_0^t (\mathscr{E}_2 \bar{T}\mathscr{E}_1 h)^*(s)\,ds = \int_0^t (\mathscr{E}_1\mathscr{E}_2 \bar{T}\mathscr{E}_1 h)^*(s)\,ds \\ &\le \int_0^t (\bar{T}\mathscr{E}_1 h)^*(s)\,ds \le \int_0^t (\mathscr{E}_1 h)^*(s)\,ds \\ &= \int_0^t h^*(s)\,ds,\end{aligned}$$

for all h in $(L^1 + L^\infty)(R,\mu)$ and all $t > 0$. Hence $Th \prec h$ for all h, and so, by Proposition 2.4, the operator T is substochastic. Finally, using II.(3.17) and (2.29), we have

$$Tf = \mathscr{E}_2 \bar{T}\mathscr{E}_1 f = \mathscr{E}_2\mathscr{E}_1 g = g \qquad \mu\text{-a.e.}$$

and this completes the proof. ■

Corollary 2.11. *Let f be an arbitrary function in $L^1 + L^\infty$ and suppose g is any measurable function satisfying $g \prec f$. Then there is an admissible operator T with $\|T\|_{\mathscr{A}} \le 1$ such that $Tf = g$.*

Proof. Since $|g|$ and $|f|$ are nonnegative and satisfy $|g| \prec |f|$, there is, by virtue of Theorem 2.10, a substochastic operator S such that $S|f| = |g|$. For each h in $L^1 + L^\infty$, set

$$Th = (\operatorname{sgn} g)S(h \cdot \overline{\operatorname{sgn} f}).$$

Then

$$f = (\operatorname{sgn} g)S(f \cdot \overline{\operatorname{sgn} f}) = (\operatorname{sgn} g)S|f| = (\operatorname{sgn} g)|g| = g,$$

so $Tf = g$, as desired. Since S is a contraction on L^1 and on L^∞, the same is evidently true of T. Hence, T is admissible and

$$\|T\|_{\mathscr{A}} = \max\{\|T\|_{\mathscr{B}(L^1)}, \|T\|_{\mathscr{B}(L^\infty)}\} \leq 1. \qquad \blacksquare$$

We are now in a position to establish the main result of this section.

Theorem 2.12. (A. P. Calderón). *Let X be a Banach function space over a resonant measure space. Then X is an exact interpolation space between L^1 and L^∞ if and only if X is rearrangement-invariant.*

Proof. We have seen already in Theorem 2.2 that X is an exact interpolation space if it is rearrangement-invariant. Suppose conversely that X is an exact interpolation space. To establish rearrangement-invariance, we need to show that if f is an arbitrary function in X, and if g is equimeasurable with f, $(g \sim f)$, then g belongs to X and $\|g\|_X = \|f\|_X$. But if $g \sim f$, then certainly $g \prec f$ and so, by Corollary 2.11, there is an admissible operator T with $\|T\|_{\mathscr{A}} \leq 1$ such that $Tf = g$. Since X is an exact interpolation space, the function $g = Tf$ belongs to X, and from (1.16) we have

$$\|g\|_X = \|Tf\|_X \leq \|T\|_{\mathscr{B}(X)}\|f\|_X \leq \|T\|_{\mathscr{A}}\|f\|_X \leq \|f\|_X.$$

Since f and g are equimeasurable, we also have $f \prec g$ and so we may repeat the above argument with the roles of f and g interchanged to obtain $\|f\|_X \leq \|g\|_X$. Hence, $\|f\|_X = \|g\|_X$ and the proof is complete. $\blacksquare$

Theorem 2.12 characterizes the Banach function spaces that are interpolation spaces between L^1 and L^∞ but, as it stands, does not characterize all interpolation spaces between L^1 and L^∞. Little modification is needed however to achieve such a characterization and this will be done in Section 1 of Chapter V (see Theorem V.1.17).

It is easy to check that the operator T constructed in the proof of Theorem 2.10 also satisfies property (2.10), that is, $|Th| \leq T|h|$ μ-a.e. for every h in $L^1 + L^\infty$. We shall not make use of this property here.

However, we shall make use of the observation that Corollary 2.11 remains valid when f and g are defined on *different* measure spaces. The proof carries over almost verbatim. For the application we have in mind, the function g will be the decreasing rearrangement f^* of f, so the hypothesis $g \prec f$ of Corollary 2.11 is trivially satisfied. The result we shall need is then as follows.

Corollary 2.13. *Let (R, μ) be a resonant measure space and let $\tilde{R}$ denote the interval $(0, \mu(R))$ equipped with Lebesgue measure m. Let f belong to $(L^1 + L^\infty)(R, \mu)$. Then there is an admissible operator T, with respect to the couples $(L^1, L^\infty)(R, \mu)$ and $(L^1, L^\infty)(\tilde{R}, m)$, of norm at most 1, such that $Tf = f^*$ m-a.e. on $\tilde{R}$.*

3. THE HARDY-LITTLEWOOD MAXIMAL OPERATOR

The L^p-boundedness hypothesized in Theorem 2.2 ($p = 1$ and $p = \infty$) is too strong for many potential applications and our goal in the remainder of this chapter will be to develop an interpolation theory under weaker hypotheses. The basis of this so-called *weak-type interpolation theory* is a fundamental interpolation theorem due to J. Marcinkiewicz. Before presenting the Marcinkiewicz interpolation theorem, however, it will be instructive to consider some examples of operators for which the preceding interpolation theory is inadequate and upon which the weak-type theory will be founded. We begin in this section with the Hardy-Littlewood maximal operator, and continue in the next section with a discussion of the Hilbert transform.

Definition 3.1. Let f be a locally integrable function on $\mathbf{R}^n$. The *Hardy-Littlewood maximal function Mf* of f is defined by

$$(Mf)(x) = \sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f(y)|\,dy, \qquad (x \in \mathbf{R}^n), \tag{3.1}$$

where the supremum extends over all cubes Q containing x (here, as throughout, cubes will be assumed to have their sides parallel to the coordinate axes). The operator $M : f \to Mf$ is called the *Hardy-Littlewood maximal operator.*

Observe that M is merely sublinear, rather than linear:

$$M(f+g) \leq Mf + Mg; \qquad M(\lambda f) = |\lambda| Mf.$$

It is clear that M is a contraction on L^∞:

$$\|Mf\|_{L^\infty} \leq \|f\|_{L^\infty}, \qquad (f \in L^\infty), \tag{3.2}$$

but not on L^1. In fact, it is easy to see that $(Mf)(x)$ never decays faster than $|x|^{-n}$ as $|x| \to \infty$, that is,

$$(Mf)(x) \geq \frac{c}{|x|^n}, \qquad (|x| > 1),$$

if f is nonzero on a set of positive measure. In particular, Mf is *never* integrable unless $f = 0$ a.e. In fact, Mf need not even be locally integrable, as the one-dimensional example

$$f(x) = \begin{cases} \dfrac{1}{|x|(\log|x|)^2}, & |x| \le \dfrac{1}{2}, \\ 0, & |x| > \dfrac{1}{2} \end{cases}$$

shows.

Our first objective, an estimate for the size of Mf for integrable f, will require the following elementary variant of the Vitali covering lemma.

Lemma 3.2. *Let Ω be an arbitrary measurable subset of $\boldsymbol{R}^n$ of finite measure. Let $\mathscr{F}$ be a collection of cubes Q that covers Ω. Then there exist finitely many disjoint cubes, say $Q_1, Q_2, \ldots, Q_K$, from $\mathscr{F}$ such that*

$$\sum_{k=1}^{K} |Q_k| \ge 4^{-n}|\Omega|. \tag{3.3}$$

Proof. By the inner regularity of Lebesgue measure on $\boldsymbol{R}^n$, the set Ω can be approximated arbitrarily closely from within by compact sets. Furthermore, each cube Q in $\mathscr{F}$ can be replaced by a larger open cube with the ratio of the measures arbitrarily close to 1. Hence, it will suffice to establish the lemma under the assumptions that Ω is compact and each cube Q in $\mathscr{F}$ is open, provided that (3.3) is then achieved with a constant larger than 4^{-n} (our proof will give constant 3^{-n}, which will suffice).

With these hypotheses, $\mathscr{F}$ is an open cover of the compact set Ω so there exist finitely many cubes in $\mathscr{F}$ which cover Ω. Hence, we may as well assume also that $\mathscr{F}$ is finite. Let Q_1 be the largest of the cubes in $\mathscr{F}$, let Q_2 be the largest remaining cube in $\mathscr{F}$ disjoint from Q_1, let Q_3 be the largest disjoint from Q_1 and Q_2, and so on. Since $\mathscr{F}$ is finite, this process ends after finitely many steps, yielding disjoint cubes $Q_1, Q_2, \ldots Q_K$, say. Let $\bar{Q}_k$ be the cube concentric with Q_k but with side-length three times as large, $(k = 1, 2, \ldots, K)$. We claim that the cubes $\bar{Q}_k$ cover Ω. For, if not, then there is a point x in $\Omega \setminus \bigcup_{k=1}^{K} \bar{Q}_k$ and, since $\mathscr{F}$ covers Ω, the point x is contained in a cube Q from $\mathscr{F}$. By construction, Q is not larger than Q_1 and contains the point x of $\bar{Q}_1^c$. Hence, Q is disjoint from Q_1. Therefore, again by construction, Q is not larger than Q_2 and contains the point x of $\bar{Q}_2^c$, hence is disjoint from Q_2. Continuing in this way, we see that Q is disjoint from $Q_1, Q_2, \ldots, Q_K$. But this is impossible because then the process would not have terminated after K steps. The cubes $\bar{Q}_k$, $(k = 1, 2, \ldots, K)$, therefore cover Ω. Consequently,

$$|\Omega| \le \sum_{k=1}^{K} |\bar{Q}_k| = 3^n \sum_{k=1}^{K} |Q_k|,$$

and, as we remarked above, this establishes the lemma. ■

Theorem 3.3. *If f belongs to $L^1(\mathbf{R}^n)$, then*

$$t(Mf)^*(t) \le 4^n \|f\|_{L^1}, \qquad (t > 0). \tag{3.4}$$

Proof. Suppose first that f has compact support, in which case it is clear that $(Mf)(x) = 0(|x|^{-n})$ as $|x| \to \infty$. In particular, for each $\lambda > 0$, the set $E_\lambda = \{x \in \mathbf{R}^n : (Mf)(x) > \lambda\}$ has finite measure. For each x in E_λ, the definition (3.1) shows that there is a cube Q_x containing x such that

$$\lambda|Q_x| < \int_{Q_x} |f(y)|\, dy. \tag{3.5}$$

The collection of all such cubes Q_x covers E_λ, so by Lemma 3.2 there exist finitely many disjoint cubes, say $Q_1, Q_2, \ldots, Q_K$, from this collection such that

$$\sum_{k=1}^{K} |Q_k| \ge 4^{-n}|E_\lambda|. \tag{3.6}$$

Hence, combining (3.5) and (3.6), we obtain

$$|E_\lambda| \le 4^n \sum_{k=1}^{K} |Q_k| \le \frac{4^n}{\lambda} \sum_{k=1}^{K} \int_{Q_k} |f(y)|\, dy \le \frac{4^n}{\lambda} \|f\|_{L^1}.$$

Since $|E_\lambda|$ is the distribution function of Mf, this estimate is equivalent to the desired result (3.4) for the decreasing rearrangement $(Mf)^*$ (see Figure 6).

In the general case of an integrable function f, we may select an increasing sequence of nonnegative simple functions $f_k \uparrow |f|$ a.e. It is an easy consequence of the monotone convergence theorem that $Mf_k \uparrow Mf$ a.e., and so it follows from II.(1.17) that $(Mf_k)^* \uparrow (Mf)^*$. Since $\|f_k\|_{L^1} \uparrow \|f\|_{L^1}$, we see that (3.4), which by the argument above holds for each f_k, persists also for f. This completes the proof. ■

If $g = Mf$, the estimate (3.4) shows that the areas of the rectangles lying below the graph of g^* are uniformly bounded (cf. Figure 8). This condition, that $\sup_t tg^*(t)$ be finite, is clearly weaker than integrability, being satisfied by every L^1-function, and by nonintegrable functions such as $1/|x|$. The collection of all such functions is referred to as *weak-L^1*. We shall have much to say about

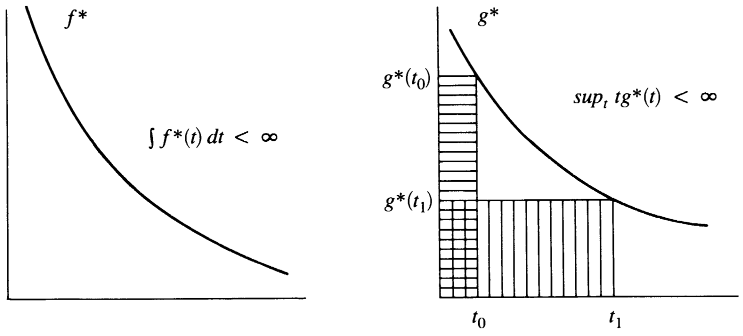


Figure 8. L^1 and weak-L^1.

such spaces later (cf. Definition V.7.13). For the present, we note simply that the maximal function Mf is in weak-L^1 whenever f belongs to L^1. This weak-type property of the maximal operator has an important application to the differentiation of integrals.

Theorem 3.4 (Lebesgue's differentiation theorem). *If f is a locally integrable function on $\mathbf{R}^n$, then*

$$\lim_{\substack{|Q|\to 0\\ Q\ni x}} \frac{1}{|Q|}\int_Q |f(y)-f(x)|\,dy = 0, \tag{3.7}$$

for almost every x in $\mathbf{R}^n$.

Proof. Because of the local nature of the result, it is clear that once it is established for integrable f, then it will follow also for all locally integrable f. Thus, we may suppose $f \in L^1(\mathbf{R}^n)$. For each x in $\mathbf{R}^n$, let

$$(\Omega f)(x) = \limsup_{\substack{|Q|\to 0\\ Q\ni x}} \left\{\frac{1}{|Q|}\int_Q |f(y)-f(x)|\,dy\right\}.$$

The theorem will be proved if we show $(\Omega f)(x) = 0$ a.e. in $\mathbf{R}^n$.

Observe that

$$(\Omega f)(x) \le (Mf)(x) + |f(x)|,$$

so, for any $t > 0$ (cf. II.(1.16)),

$$(\Omega f)^*(t) \le (Mf)^*\left(\frac{t}{2}\right) + f^*\left(\frac{t}{2}\right).$$

Hence, using the weak-type property (3.4) of M and the fact that

$$f^*\left(\frac{t}{2}\right) \le f^{**}\left(\frac{t}{2}\right) \le \left(\frac{2}{t}\right)\|f\|_{L^1},$$

we have

$$(\Omega f)^*(t) \le \frac{c}{t}\|f\|_{L^1}, \tag{3.8}$$

with c depending only on n.

It follows from the fundamental theorem of calculus that $\Omega f = \Omega(f-h)$ whenever h is a continuous function (of compact support, say) so, applying (3.8) to $f-h$ instead of f, we obtain

$$(\Omega f)^*(t) \le \frac{c}{t}\|f-h\|_{L^1}.$$

But the right-hand side may be made as small as we please since the continuous functions of compact support are dense in L^1. Hence $(\Omega f)^*(t) = 0$ for all $t > 0$, and from this we conclude that $\Omega f = 0$ a.e. in $\mathbf{R}^n$. ■

Corollary 3.5. *If f is locally integrable in $\mathbf{R}^n$, then*

$$\lim_{\substack{|Q|\to 0\\ Q\ni x}} \frac{1}{|Q|}\int_Q f(y)\,dy = f(x), \tag{3.9}$$

for almost every x in $\mathbf{R}^n$.

Corollary 3.6. *If f is locally integrable in $\mathbf{R}^n$, then*

$$|f(x)| \le (Mf)(x), \tag{3.10}$$

for almost every x in $\mathbf{R}^n$.

Before proceeding further with our analysis of the maximal function we shall need another covering lemma. In this result, the cover is constructed from the collection of *dyadic cubes*, that is, the cubes formed by means of dilations and contractions by a factor of two of the basic partition of $\mathbf{R}^n$ into unit cubes with vertices at the lattice points.

Lemma 3.7. *Let Ω be an open subset of $\mathbf{R}^n$ with finite measure. Then there is a sequence of dyadic cubes $Q_1, Q_2, \ldots$, with pairwise disjoint interiors, that covers*

Ω and satisfies

(i) $Q_k \cap \Omega^c \neq \varnothing, \qquad (k = 1, 2, \ldots);$

(ii) $|\Omega| \leq \sum_{k=1}^{\infty} |Q_k| \leq 2^n|\Omega|.$

Proof. Since Ω is open, there exists, for each x in Ω, a dyadic cube, say $Q(x)$, of smallest diameter, which contains x and has nonempty intersection with Ω^c. Subdivide $Q(x)$ into 2^n congruent dyadic subcubes and select any one, say $\tilde{Q}(x)$, that contains x. Then $\tilde{Q}(x) \subset \Omega$ and so

$$2^{-n}|Q(x)| = |\tilde{Q}(x)| = |\tilde{Q}(x) \cap \Omega| \leq |Q(x) \cap \Omega|. \tag{3.11}$$

Now, pairs of dyadic cubes have the property that if their interiors have non-empty intersection, then one cube is contained in the other. Hence, since Ω has finite measure, every point x in Ω is contained in a maximal cube from the collection $\{Q(y) : y \in \Omega\}$. There are at most countably many such maximal cubes (since there are only countably many dyadic cubes), which we list as $Q_1, Q_2, \ldots$, and they have pairwise disjoint interiors. They evidently cover Ω and so the first inequality in (ii) holds. Property (i) holds by construction. Furthermore, every Q_k satisfies an inequality of the type (3.11), and so

$$\sum_k |Q_k| \leq 2^n \sum_k |Q_k \cap \Omega| = 2^n|\Omega|.$$

This establishes the second of the inequalities in (ii) and hence completes the proof. ■

The function $f^{**}(t) = \frac{1}{t}\int_0^t f^*(s)\,ds$ is the average of f^* over the interval $(0, t)$. This is maximal among all averages of f^* over intervals containing t because f^* is decreasing. Hence, f^{**} is the Hardy-Littlewood maximal function of f^*. The next result demonstrates the equivalence of the maximal function of the decreasing rearrangement (namely, f^{**}) and the decreasing rearrangement of the maximal function (namely, $(Mf)^*$).

Theorem 3.8 *There are constants c and c', depending only on n, such that*

$$c(Mf)^*(t) \leq f^{**}(t) \leq c'(Mf)^*(t), \qquad (t > 0), \tag{3.12}$$

for every locally integrable function f on $\mathbf{R}^n$.

Proof. Fix $t > 0$. For the left-hand inequality, we may suppose $f^{**}(t) < \infty$, otherwise there is nothing to prove. In that case, by Theorem II.6.2, given $\varepsilon > 0$, there are functions g_t in L^1 and h_t in L^∞ such that $f = g_t + h_t$ and

$$\|g_t\|_{L^1} + t\|h_t\|_{L^\infty} \leq tf^{**}(t) + \varepsilon. \tag{3.13}$$

Then, by (3.4) and (3.2), for any $s > 0$,

$$(Mf)^*(s) \leq (Mg_t)^*\left(\frac{s}{2}\right) + (Mh_t)^*\left(\frac{s}{2}\right) \leq \frac{c}{s}\|g_t\|_{L^1} + \|h_t\|_{L^\infty}$$

$$\leq \frac{c}{s}(\|g_t\|_{L^1} + s\|h_t\|_{L^\infty}).$$

Putting $s = t$, using (3.13), and letting $\varepsilon \to 0$, we obtain the first of the inequalities in (3.12).

For the right-hand inequality in (3.12), we may suppose $(Mf)^*(t) < \infty$, otherwise there is nothing to prove. The lower semi-continuity of Mf ensures that the set

$$\Omega = \{x \in \mathbf{R}^n : (Mf)(x) > (Mf)^*(t)\}$$

is open, and we have $|\Omega| \leq t$ because Mf and $(Mf)^*$ are equimeasurable. Applying Lemma 3.7, we obtain a sequence of cubes $Q_1, Q_2, \ldots$, with pairwise disjoint interiors, that cover Ω and satisfy

$$Q_k \cap \Omega^c \neq \varnothing, \qquad (k = 1, 2 \ldots) \tag{3.14}$$

and

$$\sum |Q_k| \leq 2^n|\Omega| \leq 2^n t. \tag{3.15}$$

With $F = (\bigcup_k Q_k)^c$, we set

$$g = \sum_k f\chi_{Q_k}, \qquad h = f\chi_F,$$

so $f = g + h$. Then the subadditivity of $f \to f^{**}$ (Theorem II.3.4) gives

$$f^{**}(t) \leq g^{**}(t) + h^{**}(t) \leq \frac{1}{t}\|g\|_{L^1} + \|h\|_{L^\infty}. \tag{3.16}$$

Now, by (3.14), each cube Q_k contains a point of Ω^c, and at such a point the maximal function has value at most $(Mf)^*(t)$ because of the way in which Ω is defined. Thus,

$$\frac{1}{|Q_k|}\int_{Q_k} |f(y)|\,dy \leq (Mf)^*(t), \qquad (k = 1, 2, \ldots).$$

The L^1-norm of g may therefore be estimated by

$$\|g\|_{L^1} = \sum_k \int_{Q_k} |f(y)|\,dy \leq \sum_k |Q_k|(Mf)^*(t).$$

Hence, using (3.15), we have

$$\|g\|_{L^1} \leq 2^n t(Mf)^*(t). \tag{3.17}$$

On the other hand, the set F is contained in Ω^c and so the maximal function is bounded by $(Mf)^*(t)$ on F. Hence, using (3.10), we have

$$\|h\|_{L^\infty} = \|f\chi_F\|_{L^\infty} \le \|(Mf)\chi_F\|_{L^\infty} \le (Mf)^*(t).$$

Combining the last estimate with (3.16) and (3.17), we therefore obtain the right-hand inequality in (3.12). ∎

The last result simplifies the problem of establishing the boundedness of the maximal operator on rearrangement-invariant spaces. Indeed, it reduces the problem to that of establishing boundedness of f^{**}, which often is easier since f^{**} is a simple average rather than a supremum of averages. We shall elaborate on this further in Section 5 in the context of arbitrary rearrangement-invariant spaces. For the present, however, we shall content ourselves with establishing the boundedness of the maximal operator in the L^p-setting. The following inequalities, known simply as *Hardy's inequalities*, will be crucial.

Lemma 3.9 (G. H. Hardy). *Let ψ be a nonnegative measurable function on $(0,\infty)$ and suppose $-\infty < \lambda < 1$ and $1 \le q \le \infty$. Then*

$$\left\{\int_0^\infty \left(t^\lambda \frac{1}{t}\int_0^t \psi(s)\,ds\right)^q \frac{dt}{t}\right\}^{1/q} \le \frac{1}{1-\lambda}\left\{\int_0^\infty (t^\lambda \psi(t))^q \frac{dt}{t}\right\}^{1/q} \tag{3.18}$$

and

$$\left\{\int_0^\infty \left(t^{1-\lambda} \int_t^\infty \psi(s)\frac{ds}{s}\right)^q \frac{dt}{t}\right\}^{1/q} \le \frac{1}{1-\lambda}\left\{\int_0^\infty (t^{1-\lambda} \psi(t))^q \frac{dt}{t}\right\}^{1/q} \tag{3.19}$$

(with the obvious modification if $q = \infty$).

Proof. Writing $\psi(s) = s^{-\lambda/q'}s^{\lambda/q'}\psi(s)$ and applying Hölder's inequality, we obtain

$$\frac{1}{t}\int_0^t \psi(s)\,ds \le \left(\frac{1}{t}\int_0^t s^{-\lambda}\,ds\right)^{1/q'}\left(\frac{1}{t}\int_0^t s^{\lambda q/q'}\psi(s)^q\,ds\right)^{1/q}$$
$$= (1-\lambda)^{-1/q'}t^{-\lambda/q'-1/q}\left(\int_0^t s^{\lambda(q-1)}\psi(s)^q\,ds\right)^{1/q}.$$

Hence, by an interchange in the order of integration,

$$\int_0^\infty \left(t^\lambda \frac{1}{t}\int_0^t \psi(s)\,ds\right)^q \frac{dt}{t} \le (1-\lambda)^{1-q}\int_0^\infty t^{\lambda-2}\int_0^t s^{\lambda(q-1)}\psi(s)^q\,ds\,dt$$
$$= (1-\lambda)^{1-q}\int_0^\infty s^{\lambda(q-1)}\psi(s)^q \int_s^\infty t^{\lambda-2}\,dt\,ds.$$

Performing the integration over t and taking q-th roots, we obtain (3.18). The proof of (3.19) is similar and is left as an exercise. ∎

Now we can establish the L^p-boundedness of the Hardy-Littlewood maximal operator.

Theorem 3.10 (Hardy-Littlewood maximal theorem). *Let $1 < p \le \infty$ and suppose f belongs to $L^p(\mathbf{R}^n)$. Then Mf belongs to $L^p(\mathbf{R}^n)$ and*

$$\|Mf\|_{L^p} \le c\|f\|_{L^p}, \tag{3.20}$$

where c is a constant depending only on p and n.

Proof. Since $p > 1$, we may apply the first part of Lemma 3.9 with $\lambda = 1/p$ and $q = p$. In that case, we obtain from (3.12) and (3.18), with c depending only on n,

$$\|Mf\|_{L^p(\mathbf{R}^n)} = \left(\int_0^\infty (Mf)^*(t)^p\,dt\right)^{1/p}$$
$$\le c\left(\int_0^\infty \left(\frac{1}{t}\int_0^t f^*(s)\,ds\right)^p dt\right)^{1/p}$$
$$\le cp'\left(\int_0^\infty f^*(t)^p\,dt\right)^{1/p}$$
$$= cp'\|f\|_{L^p(\mathbf{R}^n)}. \quad ∎$$

The last result may be regarded as a rudimentary interpolation theorem for the maximal operator. The crucial ingredient in the proof (Theorem 3.8) asserts that

$$(Mf)^*(t) \le cf^{**}(t), \qquad (t > 0), \tag{3.21}$$

for every locally integrable function f. Encoded in this inequality are the properties that M is bounded on L^∞:

$$\|Mf\|_{L^\infty} = \sup_{t>0}(Mf)^*(t) \le c\cdot\sup_{t>0} f^{**}(t) = c\|f\|_{L^\infty},$$

and that M is bounded from L^1 into weak-L^1:

$$\|Mf\|_{\text{weak-}L^1} = \sup_{t>0} t(Mf)^*(t) \le \sup_{t>0} ctf^{**}(t)$$
$$= \sup_{t>0} c\int_0^t f^*(s)\,ds = c\|f\|_{L^1}.$$

These are the hypotheses of the interpolation theorem. The conclusion is the L^p-boundedness of the maximal operator for $1 < p < \infty$ (cf. Theorem 3.10). Note that the proof uses only (3.21) and the appropriate Hardy inequality.

We shall show in the next section that an inequality analogous to (3.21) holds for the Hilbert transform, which can then be interpolated in the same elementary fashion. Ultimately, we shall identify a large class of operators for which such interpolation is possible, the end result being the celebrated Marcinkiewicz interpolation theorem.

4. THE HILBERT TRANSFORM

The Hilbert transform is a linear operator which arises from the study of boundary values of the real and imaginary parts of analytic functions. Questions involving the Hilbert transform arise therefore from the utilization of complex methods in Fourier analysis, for example. In particular, the Hilbert transform plays the decisive role in questions of norm-convergence of Fourier series and Fourier integrals. These applications will be deferred for the time being, as will any discussion of complex methods. Instead, a direct real variable approach will be used. This will establish the existence and L^p-boundedness of the Hilbert transform with a minimum of effort and will also serve to motivate the weak-type interpolation theory to be developed in the next section.

Definition 4.1. Let f be a locally integrable function on the real line $\mathbf{R}$. The *Hilbert transform Hf* of f is defined by the principal-value integral

$$(Hf)(x) = \text{p.v.}\frac{1}{\pi}\int_{-\infty}^{\infty} f(t)\frac{dt}{x-t} \tag{4.1}$$

$$= \lim_{\varepsilon\to 0}\frac{1}{\pi}\int_{|x-t|\geq\varepsilon} f(t)\frac{dt}{x-t}, \qquad (x \in \mathbf{R}),$$

provided the limit exists a.e. The operator $H: f \to Hf$ is also referred to as the *Hilbert transform.*

Even for relatively "nice" functions, such as the continuous functions of compact support, say, it is by no means obvious that the principal-value integral in (4.1) exists a.e. Thus, establishing existence will be one of our first priorities. As a starting point, notice that if f is the characteristic function of a finite interval (a, b), then a direct integration shows that

$$H\chi_{(a,b)}(x) = \frac{1}{\pi}\log\left|\frac{x-a}{x-b}\right|, \qquad (x \neq a, b). \tag{4.2}$$

It follows that Hf exists a.e. whenever f is a step function (that is, a function which assumes only finitely many values, each being assumed on a finite disjoint union of intervals of finite length).

There is an analogy between the problems of the existence of the Hilbert transform and the existence of the limit in the Lebesgue differentiation theorem (Theorem 3.4). In the latter situation, the a.e. existence of the limit (of certain integral averages) was known for a dense subset of L^1 (the continuous functions of compact support) and the result was extended to all of L^1 by establishing control (Theorem 3.3) over the corresponding maximal operator (the Hardy-Littlewood maximal operator M). For the Hilbert transform, the dense subset of L^1 consists of the step functions, and in order to extend to all of L^1 the a.e. existence of the limit of

$$(H_\varepsilon f)(x) = \frac{1}{\pi}\int_{|x-t|\geq\varepsilon} f(t)\frac{dt}{x-t}, \qquad (x \in \mathbf{R}), \tag{4.3}$$

as $\varepsilon \to 0$, it will be necessary to consider the following maximal operator.

Definition 4.2. Let f be a locally integrable function on $\mathbf{R}$. The *maximal Hilbert transform $\mathscr{H}f$* of f is defined by

$$(\mathscr{H}f)(x) = \sup_{\varepsilon>0}|(H_\varepsilon f)(x)|, \qquad (x \in \mathbf{R}). \tag{4.4}$$

The operator $\mathscr{H}: f \to \mathscr{H}f$ is called the *maximal Hilbert transform.*

The necessary estimates for the maximal Hilbert transform will be derived from the following pair of elementary lemmas, which provide information on the distribution functions of certain rational functions.

Lemma 4.3. *Suppose a_i, b_i, $(i = 1, 2, \ldots, n)$, are real numbers satisfying $a_1 < b_1 < a_2 < b_2 < \cdots < a_n < b_n$ and let g be the rational function*

$$g(x) = \prod_{k=1}^{n}\frac{x-a_k}{x-b_k}, \qquad (x \in \mathbf{R}). \tag{4.5}$$

If $\Delta \neq 1$, then the equation $g(x) = \Delta$ has n distinct roots $r_1, r_2, \ldots, r_n$, which satisfy

$$\sum_{k=1}^{n} b_k = \sum_{k=1}^{n} r_k + (1-\Delta)^{-1}\sum_{k=1}^{n}(b_k - a_k). \tag{4.6}$$

Furthermore, if $\Delta > 1$, then

$$(\Delta - 1)|\{g > \Delta\}| = (\Delta + 1)|\{g < -\Delta\}| = \sum_{k=1}^{n}(b_k - a_k). \tag{4.7}$$

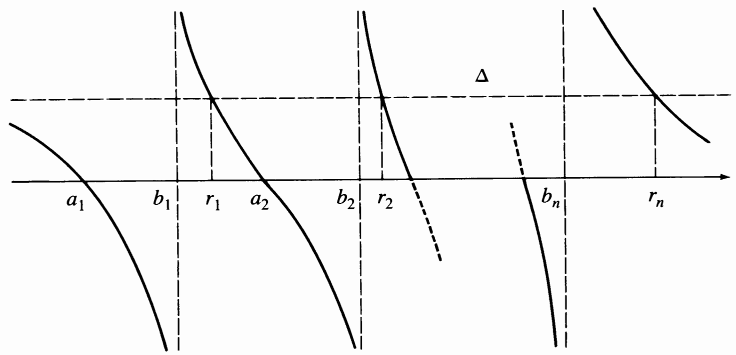


Figure 9. Graph of $g(x) = \prod_{k=1}^{n} \frac{x - a_k}{x - b_k}$.

Proof. Since g has a simple pole at each b_k, $(k = 1, 2, \dots, n)$, and $g(x) \to 1$ as $|x| \to \infty$, there are exactly n distinct solutions, say $r_1, r_2, \dots, r_n$, to the equation $g(x) = \Delta$, $(\Delta \neq 1)$ (cf. Figure 9). By (4.5), the numbers $r_1, r_2, \dots, r_n$ are the n roots of the n-th degree polynomial equation $p(x) = 0$, where

$$p(x) = \sum_{k=0}^{n} p_k x^k = \prod_{k=1}^{n} (x - a_k) - \Delta \prod_{k=1}^{n} (x - b_k). \tag{4.8}$$

The sum $\sum r_k$ of the roots is equal to $-p_{n-1}/p_n$ so, equating coefficients of x^n and of x^{n-1} in (4.8), we obtain

$$\sum_{k=1}^{n} r_k = \frac{-1}{1 - \Delta}\left(-\sum_{k=1}^{n} a_k + \Delta \sum_{k=1}^{n} b_k\right),$$

which is equivalent to (4.6).

If $\Delta > 1$, then $\{g > \Delta\} = \bigcup_{k=1}^{n} (b_k, r_k)$ (cf. Figure 9), so the first identity in (4.7) follows at once from (4.6). The second is established in similar fashion. ■

Lemma 4.4. *Suppose $t_1 < t_2 < \cdots < t_n$ and $m_1, m_2, \dots, m_n \geq 0$. Let*

$$h(x) = \sum_{k=1}^{n} \frac{m_k}{x - t_k}, \qquad (x \in \mathbf{R}). \tag{4.9}$$

If $\lambda \neq 0$, then the equation $h(x) = \lambda$ has n distinct roots $r_1, r_2, \dots, r_n$, which satisfy

$$\sum_{k=1}^{n} r_k = \sum_{k=1}^{n} t_k + \lambda^{-1} \sum_{k=1}^{n} m_k. \tag{4.10}$$

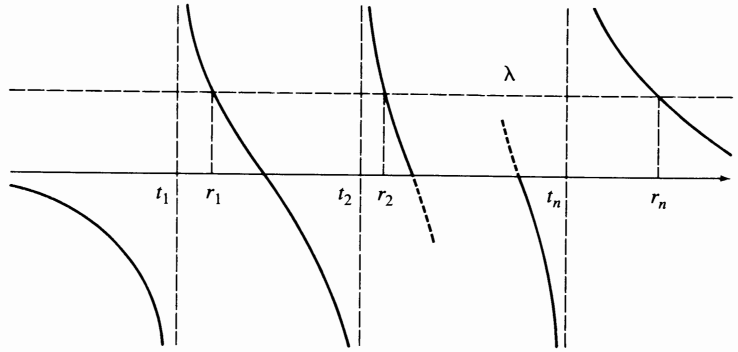


Figure 10. Graph of $h(x) = \sum_{k=1}^{n} \frac{m_k}{x - t_k}$.

Furthermore, if $\lambda > 0$ then

$$|\{h > \lambda\}| = |\{h < -\lambda\}| = \lambda^{-1} \sum_{k=1}^{n} m_k. \tag{4.11}$$

Proof. The function h has a simple pole at each t_k and tends to 0 as $|x| \to \infty$. Hence, there are exactly n solutions, say $r_1, r_2, \dots, r_n$, to $h(x) = \lambda$, $(\lambda \neq 0)$ (cf. Figure 10). Since $h(x) = \lambda$ reduces to the polynomial equation

$$\sum_{k=1}^{n} m_k \left(\prod_{\substack{j=1 \\ j \neq k}}^{n} (x - t_j)\right) - \lambda \prod_{j=1}^{n} (x - t_j) = 0,$$

the remaining assertions may be established by arguing exactly as in the proof of the last lemma. ■

It has already been observed that the Hilbert transform $H\chi_E$ exists if E is a finite union of disjoint intervals. With the aid of Lemma 4.3, we can now make a further observation: the distribution function of $H\chi_E$ depends only on the measure of E and not on the way in which E happens to be distributed over the real line. This is the content of Lemma 4.5 below. The result is valid for arbitrary sets E of finite measure (details are sketched in Exercise 10) but we shall not need the result in this generality.

Lemma 4.5 (E. M. Stein & G. Weiss). Let E be the union of finitely many disjoint intervals, each of finite length. Then

$$\mu_{H\chi_E}(\lambda) = \frac{2|E|}{\sinh(\pi\lambda)}, \qquad (\lambda > 0). \tag{4.12}$$

Proof. We may express E in the form $E = \bigcup_{j=1}^{n} (a_j, b_j)$, where the a_j and $b_j, (j = 1, 2, \ldots, n)$, satisfy the hypotheses of Lemma 4.3. It follows from (4.2) that

$$(H\chi_E)(x) = \frac{1}{\pi} \log \left| \prod_{j=1}^{n} \frac{x - a_j}{x - b_j} \right|. \tag{4.13}$$

Fix $\lambda > 0$ and set $F = \{|H\chi_E| > \lambda\}$ so $|F| = \mu_{H\chi_E}(\lambda)$. Then F may be decomposed into the disjoint union

$$F = \{|g| > e^{\pi\lambda}\} \cup \{|g| < e^{-\pi\lambda}\} = F_1 \cup F_2,$$

say, where g is the rational function defined by (4.5). Applying Lemma 4.3 to g, we obtain from (4.7),

$$|F_1| = |\{g > e^{\pi\lambda}\}| + |\{g < -e^{\pi\lambda}\}| = \frac{|E|}{e^{\pi\lambda} - 1} + \frac{|E|}{e^{\pi\lambda} + 1} = \frac{|E|}{\sinh(\pi\lambda)}.$$

By considering the rational function $1/g$ instead of g, and applying the analogous version of Lemma 4.3, we obtain a similar estimate for $|F_2|$. Since $|F| = |F_1| + |F_2|$, this establishes (4.12). ■

Part (a) of the next result shows that the maximal Hilbert transform (for simple functions of bounded support) satisfies the same kind of weak-type inequality that was established for the Hardy-Littlewood maximal operator in Theorem 3.3. Thus, the operators M and $\mathscr{H}$ behave in much the same way with respect to the L^1-norm. On L^∞, however, the situation is quite different. Indeed, the operator M is bounded on L^∞, but, as is clear from (4.2), the maximal Hilbert transform of a bounded function need not itself be bounded. Thus, in addition to an estimate on the growth of $\mathscr{H}$ on L^1, we shall also need an estimate of the growth on L^∞. Part (b) of the next result is a step in this direction.

Proposition 4.6. (a) *If f is a nonnegative simple function with bounded support, then*

$$t(\mathscr{H}f)^*(t) \le \frac{64}{\pi} \|f\|_{L^1}, \qquad (t > 0). \tag{4.14}$$

(b) *If F is a finite union of disjoint intervals, each of finite length, and $f = \chi_F$, then*

$$(\mathscr{H}f)^*(t) \le \frac{2}{\pi} \sinh^{-1}\left(\frac{64|F|}{t}\right), \qquad (t > 0). \tag{4.15}$$

Proof. Let $\lambda > 0$. In either case (a) or (b), we may use (4.4) to write

$$E = \{\mathscr{H}f > \lambda\} = \left\{\sup_{\varepsilon > 0} H_\varepsilon f > \lambda\right\} \cup \left\{\sup_{\varepsilon > 0} (-H_\varepsilon f) > \lambda\right\} = E_+ \cup E_-,$$

say. Let Ω be any compact subset of E_+. For each $x \in \Omega$, there is a finite interval, say I_x, centered at x, such that

$$\int_{(I_x)^c} f(t) \frac{dt}{x - t} > \pi\lambda.$$

The intervals I_x, $(x \in \Omega)$, cover Ω so, by Lemma 3.2, we may select from this family finitely many disjoint intervals, say $I_1, I_2, \ldots, I_n$, with centers $x_1 < x_2 < \cdots < x_n$, such that

$$|\Omega| \le 4 \sum_{j=1}^{n} |I_j| \tag{4.16}$$

and

$$\int_{(I_j)^c} f(t) \frac{dt}{x_j - t} > \pi\lambda, \qquad (j = 1, 2, \ldots, n). \tag{4.17}$$

Turning to the proof of part (a), we note that there is nothing to prove if $f = 0$ a.e., so we assume $\|f\|_{L^1} > 0$. Furthermore, since (4.17) involves only finitely many inequalities, there exists $\varepsilon > 0$ such that

$$\int_{(I_j)^c} f(t) \frac{dt}{x_j - t} > \pi\lambda + \varepsilon, \qquad (j = 1, 2, \ldots, n). \tag{4.18}$$

Let $\delta = \min\{|I_j| : j = 1, 2, \ldots, n\}$ and let $[-N, N]$ be any interval containing the support of f and each of the intervals I_j. Then we may construct a partition $(t_k)_{k=1}^{K}$ of the interval $[-N, N]$ that contains the endpoints of each I_j, $(j = 1, 2, \ldots, n)$, and

$$|t_{k+1} - t_k| < \frac{\varepsilon\delta^2}{4\|f\|_{L^1}}, \qquad (k = 1, 2, \ldots, K - 1). \tag{4.19}$$

Let

$$m_k = \int_{t_k}^{t_{k+1}} f(t)\, dt, \qquad (k = 1, 2, \ldots, K - 1),$$

and

$$A_j = \{k : (t_k, t_{k+1}) \not\subset I_j\}, \qquad (j = 1, 2, \ldots, n).$$

Set

$$g_j(y) = \sum_{k \in A_j} \frac{m_k}{y - t_k}, \qquad (j = 1, 2, \ldots, n;\ y \in \mathbf{R}).$$

Since $(I_j)^c \cap [-N, N] = \bigcup_{k \in A_j} [t_k, t_{k+1}]$, we have from (4.19),

$$\begin{aligned}
\left| g_j(x_j) - \int_{(I_j)^c} f(t) \frac{dt}{x_j - t} \right|
&\le \sum_{k \in A_j} \left| \int_{t_k}^{t_{k+1}} f(t) \left\{ \frac{1}{x_j - t_k} - \frac{1}{x_j - t} \right\} dt \right| \\
&\le \sum_{k \in A_j} \int_{t_k}^{t_{k+1}} |f(t)| \frac{|t_{k+1} - t_k|}{(|I_j|/2)^2} dt \\
&< \frac{\varepsilon}{\|f\|_{L^1}} \sum_{k \in A_j} \int_{t_k}^{t_{k+1}} |f(t)|\, dt \le \varepsilon.
\end{aligned}$$

This, together with (4.18), shows that $g_j(x_j) > \pi\lambda$. But g_j is decreasing on I_j so in fact $g_j(y) > \pi\lambda$ on the entire left-hand half of I_j. Hence,

$$\frac{1}{2} \sum_{j=1}^{n} |I_j| \le \sum_{j=1}^{n} |\{g_j > \pi\lambda\} \cap I_j|. \tag{4.20}$$

Now define

$$h(y) = \sum_{k=1}^{K} \frac{m_k}{y - t_k}, \qquad h_j(y) = \sum_{k \notin A_j} \frac{m_k}{y - t_k}, \qquad (j = 1, 2, \ldots, n).$$

Since $g_j = h - h_j$, the inequality $g_j(y) > \pi\lambda$ implies either $h(y) > \pi\lambda/2$ or $h_j(y) < -\pi\lambda/2$. Hence,

$$|\{g_j > \pi\lambda\} \cap I_j| \le \left| \left\{ h > \frac{\pi\lambda}{2} \right\} \cap I_j \right| + \left| \left\{ h_j < \frac{-\pi\lambda}{2} \right\} \right|. \tag{4.21}$$

Summing over j and using (4.16) and (4.20), we obtain

$$|\Omega| \le 8 \left\{ \left| \left\{ h > \frac{\pi\lambda}{2} \right\} \right| + \sum_{j=1}^{n} \left| \left\{ h_j < \frac{-\pi\lambda}{2} \right\} \right| \right\}.$$

Now applying Lemma 4.4 to h and to h_j, we find

$$|\Omega| \le 8 \left\{ \frac{2}{\pi\lambda} \sum_k m_k + \sum_j \frac{2}{\pi\lambda} \sum_{k \notin A_j} m_k \right\} = \frac{32}{\pi\lambda} \sum_k m_k = \frac{32\|f\|_{L^1}}{\pi\lambda}.$$

This estimate holds for all compact subsets Ω of E_+, hence also for E_+ itself by the inner regularity of Lebesgue measure. A similar argument establishes

the same estimate for E_-. Hence,

$$\mu_{\mathscr{H}f}(\lambda) = |E| = |E_+| + |E_-| \le \frac{64\|f\|_{L^1}}{\pi\lambda}, \qquad (\lambda > 0).$$

The desired estimate (4.14) follows easily from this one by using the definition of the decreasing rearrangement. This establishes part (a).

The proof of part (b) is similar. Suppose $\lambda > 0$, and proceed as in the first part of the proof to construct intervals $I_1, I_2, \ldots, I_n$ satisfying (4.16) and (4.17). Let

$$g_j(y) = \int_{(I_j)^c} f(t) \frac{dt}{y - t}, \qquad (j = 1, 2, \ldots, n;\ y \in \mathbf{R}).$$

Then (4.17) shows that $g_j(x_j) > \pi\lambda$ so, since g_j decreases on I_j, the estimate (4.20) holds for g_j. Now $f = \chi_F$, where F is a finite disjoint union of intervals so, writing

$$g_j = \pi H(\chi_F) - \pi H(\chi_{F \cap I_j}) = h - h_j,$$

say, we obtain (4.21) exactly as before. It follows from Lemma 4.5 that

$$\mu_h\left(\frac{\pi\lambda}{2}\right) = \frac{2|F|}{\sinh\left(\frac{\pi\lambda}{2}\right)}, \qquad \mu_{h_j}\left(\frac{\pi\lambda}{2}\right) = \frac{2|F \cap I_j|}{\sinh\left(\frac{\pi\lambda}{2}\right)},$$

and therefore from (4.16), (4.20), and (4.21) that

$$|\Omega| \le \frac{32|F|}{\sinh\left(\frac{\pi\lambda}{2}\right)}.$$

Hence, as before, we obtain

$$\mu_{\mathscr{H}f}(\lambda) = |E| \le \frac{64|F|}{\sinh\left(\frac{\pi\lambda}{2}\right)}, \qquad (\lambda > 0),$$

and this is equivalent to (4.15). ■

For each measurable function f on $(0, \infty)$ and each $t > 0$, let

$$\begin{aligned}
(Sf)(t) &= \int_0^\infty \min\left(1, \frac{s}{t}\right) f(s) \frac{ds}{s} \\
&= \frac{1}{t} \int_0^t f(s)\, ds + \int_t^\infty f(s) \frac{ds}{s}.
\end{aligned} \tag{4.22}$$

It is clear that S is linear. Moreover, if $0 < t < t'$. then

$$\min\left(1, \frac{s}{t'}\right) \leq \min\left(1, \frac{s}{t}\right) \leq \frac{t'}{t} \cdot \min\left(1, \frac{s}{t'}\right), \qquad (s > 0). \tag{4.23}$$

If f is nonnegative, it follows from the first of these inequalities that $(Sf)(t)$ is a decreasing function of t, and the two inequalities taken together show that $(Sf)(t)$ is finite for any one value of $t > 0$ if and only if it is finite for all values of $t > 0$.

The operator S will often be applied to the decreasing rearrangement f^* of a function f defined on some other measure space. Since $S(f^*)$ is itself decreasing, we have

$$[S(f^*)]^* = S(f^*). \tag{4.24}$$

Furthermore, a simple computation involving Fubini's theorem shows that

$$[S(f^*)]^{**} = S(f^{**}). \tag{4.25}$$

We shall meet the operator S again in the next section, where its properties will be developed in greater generality. For our present purposes, its importance stems from the fact that it dominates the maximal Hilbert transform, in the following sense.

Theorem 4.7. *Let f be a locally integrable function on* $\mathbf{R}$, *and suppose*

$$S(f^*)(1) = \int_0^1 f^*(s)\,ds + \int_1^\infty f^*(s)\frac{ds}{s} < \infty. \tag{4.26}$$

Then

$$(\mathscr{H}f)^*(t) \leq cS(f^*)(t), \qquad (0 < t < \infty), \tag{4.27}$$

where c is a constant independent of f and t.

Proof. Suppose first that $f = \chi_F$, with F a finite union of disjoint intervals of finite length. In that case, the estimate (4.15) holds for $\mathscr{H}f$. Now,

$$\sinh^{-1}(x) = \log(x + (1 + x^2)^{1/2}) \leq \log(2x + 1).$$

If $x \geq 1$, then $2x + 1 \leq 3x \leq e^2 x$, so $\log(2x + 1) \leq 2(1 + \log x)$. On the other hand, if $0 < x < 1$, then $\log(2x + 1) \leq 2x$. Hence, from (4.15),

$$(\mathscr{H}\chi_F)^*(t) \leq \frac{4}{\pi}\begin{cases} 1 + \log\left(\dfrac{64|F|}{t}\right), & 0 < t < 64|F|, \\ \dfrac{64|F|}{t}, & 64|F| \leq t < \infty. \end{cases} \tag{4.28}$$

For each $t > 0$, the right-hand side may be written

$$\frac{4}{\pi}\int_0^{64|F|} \min\left(1, \frac{s}{t}\right)\frac{ds}{s} = \frac{4}{\pi}\int_0^{|F|} \min\left(1, \frac{64s}{t}\right)\frac{ds}{s} \leq \frac{256}{\pi}\int_0^{|F|} \min\left(1, \frac{s}{t}\right)\frac{ds}{s}.$$

But the integral on the right is exactly

$$S(\chi_{(0,|F|)})(t) = S(\chi_F^*)(t),$$

so we have from (4.28),

$$(\mathscr{H}\chi_F)^*(t) \leq cS(\chi_F^*)(t), \qquad (t > 0). \tag{4.29}$$

Next, suppose F is an arbitrary bounded measurable subset of $\mathbf{R}$. Then there is a sequence of sets F_n, $(n = 1, 2, \ldots)$, each of which is a finite union of disjoint intervals, such that the symmetric difference $F \Delta F_n$ satisfies

$$\lim_{n\to\infty} |F \Delta F_n| = 0. \tag{4.30}$$

Since F is bounded, there is a finite interval I containing F. By replacing each F_n by $F_n \cap I$ (which does not affect the convergence in (4.30)), we may assume that each F_n is also contained in I. Since $\int |\chi_F - \chi_{F_n}| = |F \Delta F_n|$, we see from (4.30) that $\chi_{F_n} \to \chi_F$ in L^1. By selecting a subsequence, if necessary, we may therefore assume that $\chi_{F_n}(x) \to \chi_F(x)$ a.e. If $\varepsilon > 0$, then by the dominated convergence theorem,

$$|(H_\varepsilon\chi_F)(x)| = \lim_{n\to\infty} |(H_\varepsilon\chi_{F_n})(x)| \leq \liminf_{n\to\infty}(\mathscr{H}\chi_{F_n})(x).$$

This holds for all $\varepsilon > 0$, so, passing to the supremum, we have

$$\mathscr{H}\chi_F \leq \liminf \mathscr{H}\chi_{F_n},$$

and hence $(\mathscr{H}\chi_F)^* \leq \liminf(\mathscr{H}\chi_{F_n})^*$. But each F_n is a finite union of intervals, so we may apply (4.29) to each of these sets to obtain

$$(\mathscr{H}\chi_F)^*(t) \leq c \cdot \liminf_{n\to\infty} S(\chi_{F_n}^*)(t). \tag{4.31}$$

Now, for each $n = 1, 2, \ldots,$

$$\int_0^{|I|} |\chi_F^*(s) - \chi_{F_n}^*(s)|\,ds = |(0, |F|)\,\Delta\,(0, |F_n|)| \leq |F \Delta F_n|.$$

Hence, for each $t > 0$,

$$|S(\chi_F^*)(t) - S(\chi_{F_n}^*)(t)| = \left|\int_0^{|I|} (\chi_F^* - \chi_{F_n}^*)(s) \cdot \min\left(1, \frac{s}{t}\right)\frac{ds}{s}\right|$$
$$\leq \frac{1}{t}\int_0^{|I|} |\chi_F^* - \chi_{F_n}^*|(s)\,ds \leq \frac{|F \Delta F_n|}{t}.$$

Thus, it follows from (4.30) that $S(\chi_{F_n}^*)(t) \to S(\chi_F^*)(t)$, for all $t > 0$. In particular, we deduce from (4.31) that (4.29) holds for the bounded measurable set F.

Now suppose that f is a nonnegative simple function with bounded support. Then f may be expressed in the form $f = \sum_{j=1}^J a_j \chi_{F_j}$, where $a_j \geq 0$, each F_j is a bounded measurable set, and $F_1 \subset F_2 \subset \cdots \subset F_J$ (cf. Example I.1.6(b)). Then $f^* = \sum a_j \chi_{(0,|F_j|)}$. Using the subadditivity of the operators $\mathscr{H}$ and $g \to g^{**}$ (Theorem II.3.4), the linearity of S, and the estimate (4.29) applied to each F_j, we obtain

$$(\mathscr{H}f)^{**}(t) \leq \sum_{j=1}^J a_j (\mathscr{H}\chi_{F_j})^{**}(t) \leq c \sum_{j=1}^J a_j [S(\chi_{F_j}^*)]^{**}(t)$$

$$= c\left[S\left(\sum_{j=1}^J a_j \chi_{(0,|F_j|)}\right)\right]^{**}(t) = c[S(f^*)]^{**}(t).$$

Hence, by (4.25),

$$(\mathscr{H}f)^{**}(t) \leq cS(f^{**})(t), \qquad (t > 0). \tag{4.32}$$

Now fix $t > 0$ and let $E = \{f > f^*(t)\}$. If

$$g = [f - f^*(t)]\chi_E, \qquad h = f^*(t)\chi_E + f\chi_{E^c},$$

then $f = g + h$ and, for all $s > 0$,

$$g^*(s) = [f^*(s) - f^*(t)]_+, \qquad h^*(s) = \min[f^*(s), f^*(t)].$$

Moreover, since f is a nonnegative simple function with bounded support, so are g and h. Hence, Proposition 4.6(a) may be applied to g to give

$$(\mathscr{H}g)^*\left(\frac{t}{2}\right) \leq \frac{c}{t}\|g\|_{L^1} = \frac{c}{t}\int_0^t [f^*(s) - f^*(t)]\,ds, \tag{4.33}$$

and (4.32) may be applied to h to give

$$(\mathscr{H}h)^{**}(t) \leq cS(h^{**})(t) = c\left[\frac{1}{t}\int_0^t h^{**}(s)\,ds + \int_t^\infty h^{**}(s)\frac{ds}{s}\right]. \tag{4.34}$$

An interchange in the order of integration shows that the second integral is equal to

$$\int_t^\infty \left(\frac{1}{s}\int_0^s h^*(u)\,du\right)\frac{ds}{s} = h^{**}(t) + \int_t^\infty h^*(s)\frac{ds}{s}.$$

Furthermore, since $h^*(s)$ is constant and equal to $f^*(t)$ for $0 < s \leq t$, we have $h^{**}(s) = f^*(t)$ for $0 < s \leq t$. Hence, it follows from (4.34) that

$$(\mathscr{H}h)^{**}(t) \leq c\left[f^*(t) + \int_t^\infty f^*(s)\frac{ds}{s}\right]. \tag{4.35}$$

Noting that $(\mathscr{H}h)^*(t/2) \leq (\mathscr{H}h)^{**}(t/2) \leq 2(\mathscr{H}h)^{**}(t)$, and using (4.33) and (4.35), we obtain

$$(\mathscr{H}f)^*(t) \leq (\mathscr{H}g)^*\left(\frac{t}{2}\right) + (\mathscr{H}h)^*\left(\frac{t}{2}\right)$$

$$\leq c\left\{\frac{1}{t}\int_0^t [f^*(s) - f^*(t)]\,ds + f^*(t) + \int_t^\infty f^*(s)\frac{ds}{s}\right\}$$

$$= cS(f^*)(t).$$

This establishes the conclusion (4.27) for all nonnegative simple functions of bounded support.

Next, let f be an arbitrary nonnegative measurable function for which (4.26) holds. Then $S(f^*)(t) < \infty$ for all $t > 0$. Let $(f_n)_{n=1}^\infty$ be a sequence of nonnegative simple functions, each with bounded support, such that $f_n \uparrow f$ a.e. Then $f_n^* \uparrow f^*$ (II.(1.17)) and so, by the monotone convergence theorem, $S(f_n^*)(t) \uparrow S(f^*)(t)$ for all $t > 0$. For each $\varepsilon > 0$, we have from the Hardy-Littlewood inequality II.(2.3),

$$\begin{aligned}\int_{|y| \geq \varepsilon} \left|\frac{f(x-y)}{y}\right| dy &\leq 2\int_0^\infty f^*(s) \cdot \min\left(\frac{1}{\varepsilon}, \frac{1}{s}\right) ds \\ &= 2S(f^*)(\varepsilon) < \infty.\end{aligned} \tag{4.36}$$

Hence, it follows from the dominated convergence theorem that

$$|(H_\varepsilon f)(x)| = \lim_{n\to\infty} |(H_\varepsilon f_n)(x)| \leq \liminf_{n\to\infty} (\mathscr{H}f_n)(x).$$

Taking the supremum over ε, passing to the decreasing rearrangements, and applying (4.27) to each f_n, we therefore obtain

$$(\mathscr{H}f)^*(t) \leq c \cdot \liminf_{n\to\infty} S(f_n^*)(t) = cS(f^*)(t).$$

Hence, (4.27) holds for nonnegative f.

Finally, if f assumes both positive and negative values, then we may apply (4.27) to each of the positive and negative parts f_+ and f_- of f. Since $(f_+)^* \leq f^*$ and $(f_-)^* \leq f^*$, we obtain

$$(\mathscr{H}f)^*(t) \leq (\mathscr{H}f_+)^*\left(\frac{t}{2}\right) + (\mathscr{H}f_-)^*\left(\frac{t}{2}\right) \leq c\left\{S(f_+^*)\left(\frac{t}{2}\right) + S(f_-^*)\left(\frac{t}{2}\right)\right\}$$

$$\leq 2cS(f^*)\left(\frac{t}{2}\right) \leq 4cS(f^*)(t),$$

and this completes the proof. ■

It follows immediately from (4.27) that the maximal Hilbert transform $\mathscr{H}f$ is finite a.e., for all f in the class defined by (4.26). For the same class of functions, we can now establish the existence of the Hilbert transform Hf.

Theorem 4.8. *If f satisfies (4.26), then the principal-value integral*

$$(Hf)(x) = \text{p.v.}\frac{1}{\pi}\int_{-\infty}^{\infty} f(t)\frac{dt}{x-t}$$

exists a.e. *Furthermore,*

$$(Hf)^*(t) \le cS(f^*)(t), \qquad (t > 0), \tag{4.37}$$

for some constant c independent of f and t.

Proof. Let $g = f\chi_{(-2N,2N)}$ and $h = f - g$, where N is an arbitrary positive integer. For all ε satisfying $0 < \varepsilon < N$ and for all $x \in (-N, N)$, we have $(H_\varepsilon h)(x) = (H_N h)(x)$ and (cf. (4.36))

$$|(H_N h)(x)| \le \frac{1}{\pi}\int_{|y|>N}\left|\frac{f(x-y)}{y}\right| dy \le \frac{2}{\pi}S(f^*)(N) < \infty.$$

Hence, $(Hh)(x) = \lim_{\varepsilon\to 0}(H_\varepsilon h)(x)$ exists. Notice that g belongs to L^1 because

$$\int_{-\infty}^{\infty} |g(x)|\,dx = \int_{-2N}^{2N} |f(t)|\,dt \le \int_0^{4N} f^*(t)\,dt \le 4NS(f^*)(4N) < \infty.$$

For each integrable function k, let

$$(\Omega k)(x) = \left|\limsup_{\varepsilon\to 0+}(H_\varepsilon k)(x) - \liminf_{\varepsilon\to 0+}(H_\varepsilon k)(x)\right|.$$

Then $\Omega k \le 2\mathscr{H}k$ and $\Omega(g - \phi) = \Omega(g)$ a.e. for any step function ϕ because the Hilbert transform $H\phi = \lim_{\varepsilon\to 0} H_\varepsilon\phi$ is known to exist a.e. for step functions. Hence, by (4.27),

$$(\Omega g)^*(t) \le 2(\mathscr{H}(g-\phi))^*(t) \le 2cS(g-\phi)^*(t) \le \frac{2c}{t}\|g - \phi\|_{L^1}.$$

The step functions are dense in L^1 so the right-hand side may be made as small as we please. It follows that $\Omega g(x) = 0$ a.e. and hence that $\lim_{\varepsilon\to 0} H_\varepsilon g(x)$ exists a.e. Combined with the result for h, this establishes the existence of $(Hf)(x)$ a.e. in $(-N, N)$ and hence, since N is arbitrary, a.e. in $\mathbf{R}$. Since $|Hf| \le \mathscr{H}f$, the estimate (4.37) follows directly from (4.27). ■

By Hölder's inequality,

$$S(f^*)(1) = \int_0^\infty f^*(s)\cdot\min(1, s)\frac{ds}{s}$$
$$\le \|f^*\|_{L^p(0,\infty)}\left\|\min\left(\frac{1}{s}, 1\right)\right\|_{L^{p'}(0,\infty)}.$$

Hence, if $1 \le p < \infty$, the quantity $S(f^*)(1)$ is finite for every f in $L^p(\mathbf{R})$. The preceding theorem therefore establishes the existence of the Hilbert transform on every L^p with $1 \le p < \infty$. It also leads to the following result.

Theorem 4.9. (a) (**M. Riesz**) *If $1 < p < \infty$ and f belongs to $L^p(\mathbf{R})$, then Hf belongs to $L^p(\mathbf{R})$ and*

$$\|Hf\|_{L^p} \le c_p\|f\|_{L^p}, \qquad (f \in L^p), \tag{4.38}$$

where c_p depends only on p.
(b) (**A. N. Kolmogorov**) *If f belongs to $L^1(\mathbf{R})$, then*

$$\sup_{t>0} t(Hf)^*(t) \le c\|f\|_{L^1}, \qquad (f \in L^1), \tag{4.39}$$

for some absolute constant c.

Proof. From Theorem 4.8,

$$\|Hf\|_{L^p(\mathbf{R})} = \|(Hf)^*\|_{L^p(0,\infty)} \le c\|S(f^*)\|_{L^p(0,\infty)}$$
$$\le c\left\{\left\|\frac{1}{t}\int_0^t f^*(s)\,ds\right\|_{L^p(0,\infty)} + \left\|\int_t^\infty f^*(s)\frac{ds}{s}\right\|_{L^p(0,\infty)}\right\}.$$

If $p > 1$, we apply Hardy's inequality (3.18) (with $q = \lambda^{-1} = p$) to obtain

$$\left\|\frac{1}{t}\int_0^t f^*(s)\,ds\right\|_{L^p(0,\infty)} \le p'\|f^*\|_{L^p(0,\infty)} = p'\|f\|_{L^p(\mathbf{R})}.$$

Similarly, if $p < \infty$, we may apply Hardy's inequality (3.19) (with $q = (1-\lambda)^{-1} = p$) to obtain

$$\left\|\int_t^\infty f^*(s)\frac{ds}{s}\right\|_{L^p(0,\infty)} \le p\|f^*\|_{L^p(0,\infty)} = p\|f\|_{L^p(\mathbf{R})}.$$

Hence, if $1 < p < \infty$, these estimates may be combined to produce (4.38) (with $c_p = (p + p')c$).

In case $p = 1$, the estimate (4.37) yields

$$t(Hf)^*(t) \le ctS(f^*)(t) = c\int_0^\infty f^*(s) \cdot \min\left(1, \frac{t}{s}\right) ds$$

$$\le c\int_0^\infty f^*(s)\, ds = c\|f\|_{L^1(\mathbf{R})},$$

from which (4.39) follows. ■

Of course, the results (4.38) and (4.39) hold also for the maximal Hilbert transform because the same estimate in terms of $S(f^*)$ holds for both H and $\mathscr{H}$ (cf. (4.27) and (4.37)).

In Theorem 3.8, we were able to estimate the Hardy-Littlewood maximal function Mf from above and below by f^{**}. Theorem 4.8 provides an estimate of the Hilbert transform Hf from above by the function $S(f^*)$. The corresponding lower estimate is false. However, there is the following substitute, which will prove useful later in describing the action of H on arbitrary rearrangement-invariant spaces.

Proposition 4.10. *If $S(f^*)(1) < \infty$, then there is a function g equimeasurable with f such that*

$$S(f^*)(t) \le 2\pi(Hg)^*(t), \qquad (t > 0). \tag{4.40}$$

Proof. If

$$g(x) = \begin{cases} 0, & \text{if } x \ge 0, \\ f^*(-x), & \text{if } x < 0, \end{cases}$$

then clearly f and g are equimeasurable. In particular, $S(g^*)(1) = S(f^*)(1) < \infty$, so $(Hg)(x)$ exists a.e. by Theorem 4.8. If $x > 0$, then

$$(Hg)(x) = \frac{1}{\pi}\int_0^\infty f^*(u)\frac{du}{x+u}$$

$$\ge \frac{1}{2\pi}\int_0^\infty f^*(u) \cdot \min\left(\frac{1}{x}, \frac{1}{u}\right) du = \frac{1}{2\pi}S(f^*)(x).$$

Hence,

$$|Hg(x)| \ge \begin{cases} \dfrac{1}{2\pi}S(f^*)(x), & x > 0, \\ 0, & x < 0, \end{cases}$$

so, taking decreasing rearrangements, we obtain (4.40). ■

5. OPERATORS OF JOINT WEAK TYPE $(p_0, q_0; p_1, q_1)$

We have seen in the last two sections that the fundamental inequalities

$$(Mf)^*(t) \le c\frac{1}{t}\int_0^t f^*(s)\, ds, \qquad (0 < t < \infty), \tag{5.1}$$

and

$$(Hf)^*(t) \le c\left(\frac{1}{t}\int_0^t f^*(s)\, ds + \int_t^\infty f^*(s)\frac{ds}{s}\right), \qquad (0 < t < \infty) \tag{5.2}$$

for the Hardy-Littlewood maximal operator M (Theorem 3.3) and the Hilbert transform H (Theorem 4.8) provide easily the L^p-boundedness of these operators for $1 < p < \infty$ (Theorems 3.10 and 4.9). Indeed, all that is required beyond (5.1) and (5.2) is the L^p-boundedness of the averaging operator $\frac{1}{t}\int_0^t(\cdot)\, ds$ and of its adjoint $\int_t^\infty(\cdot)\frac{ds}{s}$, and this is easily resolved with the Hardy inequalities (3.18) and (3.19).

This rudimentary interpolation principle can be developed into a worthwhile interpolation theory. The operators to be interpolated will, by hypothesis, satisfy inequalities analogous to (5.1) or (5.2). Their boundedness from L^p to L^q, for intermediate values of p and q, will then follow as before by applying the appropriate Hardy inequalities. The resulting theorem is essentially the classical Marcinkiewicz interpolation theorem in disguise, but in a form that is easily applicable in the more general context of rearrangement-invariant spaces.

The appropriate generalization of the operators on the right-hand sides of (5.1) and (5.2) is as follows.

Definition 5.1. Suppose $1 \le p_0 < p_1 \le \infty$, $1 \le q_0, q_1 \le \infty$ and $q_0 \ne q_1$. Let σ denote the *interpolation segment*

$$\sigma = \left[\left(\frac{1}{p_0}, \frac{1}{q_0}\right), \left(\frac{1}{p_1}, \frac{1}{q_1}\right)\right],$$

that is, the line segment in the unit square $\{(x, y): 0 \le x, y \le 1\}$ with endpoints $(1/p_i, 1/q_i)$, $(i = 0, 1)$. Let m denote the slope

$$m = \frac{1/q_0 - 1/q_1}{1/p_0 - 1/p_1} \tag{5.3}$$

of the line segment σ. For each measurable function f on $(0, \infty)$ and each

$t > 0$, let

$$(S_\sigma f)(t) = \int_0^\infty f(s) \cdot \min\left(\frac{s^{1/p_0}}{t^{1/q_0}}, \frac{s^{1/p_1}}{t^{1/q_1}}\right) \frac{ds}{s}, \tag{5.4}$$

or equivalently,

$$(S_\sigma f)(t) = t^{-1/q_0} \int_0^{t^m} s^{1/p_0} f(s) \frac{ds}{s} + t^{-1/q_1} \int_{t^m}^\infty s^{1/p_1} f(s) \frac{ds}{s}. \tag{5.5}$$

The operator $S_\sigma : f \to S_\sigma f$ is the *Calderón operator* associated with the interpolation segment σ.

Observe that the operator S defined in the last section by (4.22) is precisely the Calderón operator S_σ associated with the diagonal segment $\sigma = [(1, 1), (0, 0)]$ of the unit square. Here are some simple properties of the S_σ-operator.

Proposition 5.2. *If f is a nonnegative measurable function on $(0, \infty)$, then, for each $t, u > 0$,*

$$(S_\sigma f)(t) \le \max\left\{\left(\frac{u}{t}\right)^{1/q_0}, \left(\frac{u}{t}\right)^{1/q_1}\right\} (S_\sigma f)(u). \tag{5.6}$$

In particular, $S_\sigma f$ is decreasing and, for each $t > 0$,

$$(S_\sigma f)(t) = (S_\sigma f)^*(t) \le S_\sigma(f^*)(t). \tag{5.7}$$

Proof. The inequality (5.6) follows directly from (5.4) and the obvious estimate

$$\min_{j=0,1} \left\{\frac{s^{1/p_j}}{t^{1/q_j}}\right\} \le \min_{j=0,1} \left\{\frac{s^{1/p_j}}{u^{1/q_j}}\right\} \cdot \max_{i=0,1} \left\{\frac{u}{t}\right\}^{1/q_i}.$$

The fact that $S_\sigma f$ is decreasing is an immediate consequence of (5.4) and so the first equality in (5.7) is evident. For the remaining assertion in (5.7), note that, for each $t > 0$, the kernel

$$k_t(s) = \frac{1}{s} \cdot \min_{j=0,1} \left\{\frac{s^{1/p_j}}{t^{1/q_j}}\right\} = \min_{j=0,1} \left\{\frac{s^{1/p_j - 1}}{t^{1/q_j}}\right\}$$

is a decreasing function of s. Hence, the Hardy-Littlewood inequality II.(2.3) gives the desired result, namely,

$$(S_\sigma f)(t) = \int_0^\infty f(s) k_t(s)\, ds \le \int_0^\infty f^*(s) k_t(s)\, ds = S_\sigma(f^*)(t). \qquad \blacksquare$$

The Hardy-Littlewood maximal operator M is subadditive rather than additive, and so the interpolation theory should be developed with such operators in mind. In fact, we shall be able to treat the following kinds of operators.

Definition 5.3. Let (R_0, μ_0) and (R_1, μ_1) be totally σ-finite measure spaces. Let T be an operator whose domain is some linear subspace of $\mathcal{M}_0(R_0, \mu_0)$ and whose range is contained in the μ_1-measurable functions on R_1. Then T is said to be *quasilinear* if there is a constant $k \ge 1$ such that the relations

$$|T(f + g)| \le k(|Tf| + |Tg|), \qquad |T(\lambda f)| = |\lambda| \cdot |Tf| \tag{5.8}$$

hold μ_1-a.e. on R_1 for all f and g in the domain of T and for all scalars λ. If (5.8) holds with $k = 1$, then T is said to be *sublinear*.

Definition 5.4. Suppose $1 \le p_0 < p_1 \le \infty$ and $1 \le q_0$, $q_1 \le \infty$ with $q_0 \ne q_1$. Let T be a quasilinear operator (with respect to measure spaces (R_0, μ_0) and (R_1, μ_1)) and suppose Tf is defined for all μ_0-measurable functions f on R_0 for which

$$S_\sigma(f^*)(1) = \int_0^1 s^{1/p_0} f^*(s) \frac{ds}{s} + \int_1^\infty s^{1/p_1} f^*(s) \frac{ds}{s} < \infty. \tag{5.9}$$

Then T is said to be of *joint weak type* $(p_0, q_0; p_1, q_1)$ if there is a constant c such that

$$(Tf)^*(t) \le c S_\sigma(f^*)(t), \qquad (0 < t < \infty), \tag{5.10}$$

for all f satisfying (5.9).

The following result is an immediate consequence of (5.7) and the preceding definition.

Proposition 5.5. *The operator S_σ is of joint weak type $(p_0, q_0; p_1, q_1)$ relative to the measure spaces $(R_0, \mu_0) = (R_1, \mu_1) = (\mathbf{R}^+, m)$.*

Theorems 3.8, 4.7, and 4.8 provide the following examples of operators of joint weak type.

Theorem 5.6. *The Hardy-Littlewood maximal operator M, the Hilbert transform H and the maximal Hilbert transform $\mathscr{H}$ are operators of joint weak type $(1, 1; \infty, \infty)$.*

We come now to the fundamental interpolation theorem describing the action of operators of joint weak type on rearrangement-invariant spaces.

Theorem 5.7 (A. P. Calderón). *Suppose $1 \le p_0 < p_1 \le \infty$ and $1 \le q_0$, $q_1 \le \infty$ with $q_0 \ne q_1$. Let X and Y be rearrangement-invariant Banach function spaces over resonant measure spaces (R_0, μ_0) and (R_1, μ_1), respectively, and suppose $S_\sigma(f^*)(1) < \infty$ for every f in X. Then the following conditions are equivalent:*

(i) *every linear operator of joint weak type $(p_0, q_0; p_1, q_1)$ is bounded from X to Y;*
(ii) *every quasilinear operator of joint weak type $(p_0, q_0; p_1, q_1)$ is bounded from X to Y;*
(iii) *if f belongs to X and g belongs to $\mathscr{M}_0(R_1, \mu_1)$, then the inequality $g^* \le S_\sigma(f^*)$ implies g belongs to Y and $\|g\|_Y \le c\|f\|_X$, where c is a constant independent of f and g.*

Proof. It is clear that (ii) implies (i) because every linear operator is quasilinear. We show next that (iii) implies (ii). Let T be a quasilinear operator of joint weak type $(p_0, q_0; p_1, q_1)$. Then Tf is defined whenever $S(f^*)(1) < \infty$ and hence, by hypothesis, whenever f belongs to X. By (5.10), the operator T satisfies $(Tf)^* \le cS_\sigma(f^*) = S_\sigma(cf^*)$, so property (iii) shows that Tf belongs to Y and $\|Tf\|_Y \le c\|f\|_X$. Hence, T is a bounded operator from X to Y and property (ii) holds.

The proof will thus be complete if we show that (i) implies (iii). To this end, suppose f belongs to X and g is a μ_1-measurable function on R_1 for which

$$g^*(t) \le S_\sigma(f^*)(t), \qquad (0 < t < \infty). \tag{5.11}$$

Then there is a measurable function h on $\mathbf{R}^+$, satisfying $|h| \le 1$, such that

$$g^*(t) = h(t)S_\sigma(f^*)(t), \qquad (0 < t < \infty). \tag{5.12}$$

Let H denote the operation of multiplication by the function h.

It is clear from (5.9) and II.(6.8) that any μ_0-measurable function k on R_0 with $S_\sigma(k^*)(1)$ finite belongs to $(L^1 + L^\infty)(R_0, \mu_0)$. In particular, the function f belongs to X, which in turn is contained in $(L^1 + L^\infty)(R_0, \mu_0)$. By Corollary 2.13, there is an admissible operator T_1, of norm at most 1, with respect to the couples $(L^1, L^\infty)(R_0, \mu_0)$ and $(L^1, L^\infty)(\tilde{R}_0, m)$, such that

$$T_1(|f|) = f^* \tag{5.13}$$

m-a.e. on $\tilde{R}_0 = (0, \mu_0(R_0))$.

If q_0, $q_1 < \infty$, it follows from (5.6) (with $u = 1$) that $S_\sigma(f^*)(t) \to 0$ as $t \to \infty$. If one of q_0, q_1 is infinite, the same result is derived directly from (5.5). Hence, in any case, (5.11) shows that $g^*(t) \to 0$ as $t \to \infty$. By Corollary II.7.6, there is a measure-preserving transformation σ from the support, say E, of g onto the interval $(0, \mu_1(E))$ such that $|g| = g^* \circ \sigma$ on E. For each measurable function F on $\mathbf{R}^+$, define a function $T_2 F$ on R_1 by

$$(T_2 F)(x) = \begin{cases} (F \circ \sigma)(x), & (x \in E), \\ 0, & (x \in R_1 \setminus E). \end{cases}$$

Since, by Proposition II.7.2, the function $F \circ \sigma$ restricted to E is equimeasurable with F restricted to $(0, \mu_1(E))$, it is clear that

$$(T_2 F)^*(t) \le F^*(t), \qquad (F \in \mathscr{M}_0(\mathbf{R}^+), 0 < t < \infty). \tag{5.14}$$

Furthermore, for the function g, we have

$$T_2(g^*) = |g| \qquad \mu_1\text{-a.e. on } R_1. \tag{5.15}$$

With these preliminaries in place, it is clear that the operator T defined by $T = T_2 H S_\sigma T_1$ is linear and, by (5.12), (5.13), and (5.15), satisfies

$$T(|f|) = |g| \qquad \mu_1\text{-a.e. on } R_1. \tag{5.16}$$

We claim that T is of joint weak type $(p_0, q_0; q_1, q_1)$ with respect to the measure spaces (R_0, μ_0) and (R_1, μ_1). Indeed, let k be any μ_0-measurable function on R_0 for which $S_\sigma(k^*)(1)$ is finite. Then, as we observed above, the function k belongs to $(L^1 + L^\infty)(R_0, \mu_0)$. Since T_1 is admissible, we see, exactly as in Proposition 2.1, that $T_1 k \prec k^*$. It is routine to verify that, for each $t > 0$, the functional $k \to S_\sigma(k^*)(t)$ is a rearrangement-invariant Banach function norm (in fact, a Lorentz Λ-norm), so it follows from Corollary II.4.7 that $S_\sigma(T_1 k)(t) \le S_\sigma(k^*)(t)$, for all $t > 0$. Multiplying by h (which satisfies $|h| \le 1$) and taking decreasing rearrangements, we therefore have

$$(HS_\sigma T_1 k)^*(t) \le S_\sigma(k^*)(t), \qquad (0 < t < \infty).$$

Finally, applying (5.14), we obtain

$$(Tk)^*(t) = [T_2(HS_\sigma T_1 k)]^*(t) \le S_\sigma(k^*)(t), \qquad (0 < t < \infty).$$

Hence, T is of joint weak type $(p_0, q_0; p_1, q_1)$. By property (i), the operator T is bounded from X to Y, that is, there is a constant c such that

$$\|Tk\|_Y \le c\|k\|_X, \qquad (k \in X).$$

Since f, hence $|f|$, belongs to X we may apply this in conjunction with (5.16) to find that $|g|$, hence g, belongs to Y and $\|g\|_Y \le c\|f\|_X$. Hence, property (iii) holds and the proof is complete. ■

Theorem 5.7 characterizes the pairs (X, Y) of rearrangement-invariant spaces for which every operator of joint weak type $(p_0, q_0; p_1, q_1)$ is bounded from X into Y. In the special case where $X = Y$ and $p_i = q_i$, $(i = 0, 1)$, there is an elegant formulation of Calderón's theorem, due to D. W. Boyd. The basic idea is to carry the space X over to its representation $\bar{X}$ as a space of functions on $(0, \infty)$ (the Luxemburg representation theorem) and to characterize the S_σ-condition in Calderón's theorem in terms of the growth of the operator norms of the dilation operators $f(t) \to f(st)$ on $\bar{X}$ as $s \to 0$ and $s \to \infty$. It can be shown that the latter growth conditions depend only on the values of two parameters $(\underline{\alpha}, \bar{\alpha})$—the Boyd indices of X—satisfying $0 \le \underline{\alpha} \le \bar{\alpha} \le 1$. The indices gauge, in some sense, the "position" of the space X on the scale between L^1 and L^∞. Thus, as we shall see, the indices of L^p are both equal to $1/p$, the indices of $L^p \cap L^q$, $(p < q)$, are $\underline{\alpha} = 1/q$, $\bar{\alpha} = 1/p$, and so on.

We shall need some elementary facts from the theory of subadditive functions. A real-valued function ω on $(-\infty, \infty)$ is *subadditive* if

$$\omega(s + t) \le \omega(s) + \omega(t), \qquad (-\infty < s, t < \infty).$$

Lemma 5.8. *Let ω be an increasing subadditive function on $(-\infty, \infty)$ for which $\omega(0) = 0$. Then*

$$-\omega(-s) \le \omega(s), \qquad (-\infty < s < \infty). \tag{5.17}$$

Furthermore, $\omega(s)/s$ tends to a finite limit $\alpha > 0$ as $s \to \infty$ and

$$\alpha = \lim_{s \to \infty} \frac{\omega(s)}{s} = \inf_{s > 0} \frac{\omega(s)}{s}. \tag{5.18}$$

Proof. The estimate (5.17) follows immediately from the subadditivity of ω and the fact that $\omega(0) = 0$:

$$0 = \omega(0) = \omega(s + (-s)) \le \omega(s) + \omega(-s).$$

Let $\alpha = \inf_{s>0} \omega(s)/s$. It is clear that $0 \le \alpha \le \omega(1)$ so, in particular, α is finite. Let $\varepsilon > 0$, choose $t > 0$ satisfying $\alpha \le \omega(t)/t < \alpha + \varepsilon$, and select a positive integer N such that $(1 + 1/N)\omega(t)/t < \alpha + \varepsilon$. To each $s \ge Nt$, there corresponds an integer $n \ge N$ such that $nt \le s < (n + 1)t$. Using the latter inequality and the fact that ω is increasing and subadditive, we obtain $\omega(s) \le (n + 1)\omega(t)$, and so

$$\frac{\omega(s)}{s} \le \frac{(n + 1)\omega(t)}{nt} \le \left(1 + \frac{1}{N}\right)\frac{\omega(t)}{t} < \alpha + \varepsilon.$$

Hence, $\omega(s)/s \to \alpha$ as $s \to \infty$. ■

A nonnegative function ψ on $(0, \infty)$ is said to be *submultiplicative* if

$$\psi(st) \le \psi(s)\psi(t), \qquad (0 < s, t < \infty).$$

To each submultiplicative function ψ on $(0, \infty)$, we associate a function ω on $(-\infty, \infty)$ defined by

$$\omega(s) = \log \psi(e^s), \qquad (-\infty < s < \infty). \tag{5.19}$$

It is clear that ω is subadditive. Thus, if ψ is increasing and $\psi(1) = 1$, then ω satisfies the hypotheses of Lemma 5.8. Let α be the corresponding limit defined by (5.18) and set $\bar{\alpha}(\psi) = \alpha$. Then $0 \le \bar{\alpha} < \infty$ and, by (5.18),

$$\bar{\alpha}(\psi) = \lim_{t \to \infty} \frac{\log \psi(t)}{\log t} = \inf_{t > 1} \frac{\log \psi(t)}{\log t}. \tag{5.20}$$

Lemma 5.9. *Let ψ be an increasing submultiplicative function on $(0, \infty)$ for which $\psi(1) = 1$, and let a be an arbitrary positive number. Then $\bar{\alpha}(\psi) < a$ if and only if*

$$\int_1^\infty t^{-a}\psi(t)\frac{dt}{t} < \infty. \tag{5.21}$$

Proof. If $\bar{\alpha}(\psi) < a$, then there exists $\varepsilon > 0$ such that $\bar{\alpha}(\psi) < a - \varepsilon$. By (5.20), there exists $T > 1$ such that $\log \psi(t)/\log t < a - \varepsilon$ for all $t \ge T$. Then $\psi(t) < t^{a-\varepsilon}$ for $t \ge T$, so

$$\int_1^\infty t^{-a}\psi(t)\frac{dt}{t} \le \psi(T)\int_1^T t^{-1-a}\,dt + \int_T^\infty t^{-1-\varepsilon}\,dt < \infty.$$

Conversely, if (5.21) holds, then $s^{-a}\psi(s) < 1$ for some $s > 1$. Then, by (5.20),

$$\bar{\alpha}(\psi) \le \frac{\log \psi(s)}{\log s} < \frac{\log(s^a)}{\log s} = a,$$

as desired. ■

For the remainder of this section, we shall denote by $X = X(\rho)$ a rearrangement-invariant Banach function space over an infinite, nonatomic, totally σ-finite measure space (R, μ). In this case, the Luxemburg representation theorem (Theorem II.4.10) provides a (unique) rearrangement-invariant Banach function norm $\bar{\rho}$ over $(\mathbf{R}^+, m)$, defined by (cf. II.(4.12))

$$\bar{\rho}(h) = \sup\left\{\int_0^\infty g^*(t)h^*(t)\,dt : \rho'(g) \le 1\right\}, \tag{5.22}$$

such that

$$\rho(f) = \bar{\rho}(f^*) \tag{5.23}$$

for all $f \in \mathscr{M}_0^+(R,\mu)$. The rearrangement-invariant Banach function space generated by $\bar{\rho}$ is denoted by $\bar{X}$.

Definition 5.10. For each $t > 0$, let E_t denote the *dilation operator* defined on $\mathscr{M}_0(\mathbf{R}^+, m)$ by

$$(E_t f)(s) = f(st), \qquad (0 < s < \infty). \tag{5.24}$$

With X and $\bar{X}$ as above, let $h_X(t)$ denote the operator norm of $E_{1/t}$ as an operator from $\bar{X}$ to $\bar{X}$. Thus,

$$h_X(t) = \|E_{1/t}\|_{\mathscr{B}(\bar{X})}, \qquad (0 < t < \infty). \tag{5.25}$$

It is not immediately clear that E_t is bounded on $\bar{X}$. The next result establishes this fact and provides an estimate for the operator norm.

Proposition 5.11. *For each $t > 0$, the operator E_t is bounded from $\bar{X}$ to $\bar{X}$. The function h_X is increasing and submultiplicative on $(0,\infty)$, satisfies $h_X(1) = 1$, and*

$$h_X(t) \le \max(1,t), \qquad (0 < t < \infty). \tag{5.26}$$

Morever, if X' denotes the associate space of X, then

$$h_X(t) = t h_{X'}\left(\frac{1}{t}\right), \qquad (0 < t < \infty). \tag{5.27}$$

Proof. The dilation operator $E_{1/t}$ is a contraction on $L^\infty(\mathbf{R}^+, m)$ and, since

$$\int_0^\infty |(E_{1/t}f)(s)|\,ds = \int_0^\infty \left|f\left(\frac{s}{t}\right)\right| ds = t\int_0^\infty |f(s)|\,ds, \tag{5.28}$$

it is a bounded operator on $L^1(\mathbf{R}^+, m)$ with operator norm t. Hence, Theorem 2.2 shows that $E_{1/t}$ is bounded on $\bar{X}$ with operator norm at most $\max(1,t)$. This establishes (5.26). The assertion $h_X(1) = 1$ is obvious since E_1 is the identity operator.

It is easily verified that $(E_t f)^* = E_t(f^*)$ so it follows from II.(4.6) that

$$\|E_{1/t}f\|_{\bar{X}} = \sup\left\{\int_0^\infty f^*\left(\frac{s}{t}\right) g^*(s)\,ds : \|g\|_{(\bar{X})'} \le 1\right\}, \tag{5.29}$$

for any f in $\bar{X}$. From this it is clear that $h_X(t)$ increases with t. The submultiplicativity of h_X is an immediate consequence of the fact that $E_{st} = E_s E_t$ for all $s, t > 0$.

Finally, a change of variables in (5.29) leads to the estimate

$$\|E_{1/t}f\|_{\bar{X}} \le t\|f\|_{\bar{X}}\|E_t g\|_{(\bar{X})'} \le t\|f\|_{\bar{X}} h_{X'}\left(\frac{1}{t}\right),$$

where, in the last inequality, we have used the fact that $(\bar{X})' = (X')^-$ (cf. Theorem II.4.10). Hence, $h_X(t) \le t h_{X'}(1/t)$. Replacing X by X', t by $1/t$, and recalling from Theorem I.2.7 that $X'' = X$, we obtain also $h_{X'}(1/t) \le (1/t)h_X(t)$. This, together with the previous estimate, yields (5.27). ■

Definition 5.12. Let X be a rearrangement-invariant Banach function space over an infinite, nonatomic, totally σ-finite measure space. The *Boyd indices* of X are the numbers $\underline{\alpha}_X$ and $\bar{\alpha}_X$ defined by

$$\underline{\alpha}_X = \sup_{0<t<1} \frac{\log h_X(t)}{\log t}, \qquad \bar{\alpha}_X = \inf_{1<t<\infty} \frac{\log h_X(t)}{\log t}. \tag{5.30}$$

Proposition 5.13. *The indices $\underline{\alpha} = \underline{\alpha}_X$ and $\bar{\alpha} = \bar{\alpha}_X$ of X are given by the limits*

$$\underline{\alpha}_X = \lim_{t\to 0+} \frac{\log h_X(t)}{\log t}, \qquad \bar{\alpha}_X = \lim_{t\to\infty} \frac{\log h_X(t)}{\log t}, \tag{5.31}$$

and they satisfy

$$0 \le \underline{\alpha} \le \bar{\alpha} \le 1. \tag{5.32}$$

Furthermore, the indices $\underline{\alpha}' = \underline{\alpha}_{X'}$ and $\bar{\alpha}' = \bar{\alpha}_{X'}$ of the associate space X' are given by

$$\underline{\alpha}' = 1 - \bar{\alpha}, \qquad \bar{\alpha}' = 1 - \underline{\alpha}. \tag{5.33}$$

Proof. The identity

$$\frac{\log h_X(t)}{\log t} = 1 - \frac{\log h_{X'}(1/t)}{\log(1/t)}, \qquad (0 < t < \infty) \tag{5.34}$$

is an immediate consequence of (5.27). The relations (5.33) follow at once from this and (5.30).

The properties established for h_X in Proposition 5.11 show that the upper index $\bar{\alpha}$ coincides with the constant $\bar{\alpha}(h_X)$ defined by (5.20), and from this the second of the identities in (5.31) follows. The first identity in (5.31) now results from the second by appealing to (5.33) and (5.34).

That $\bar{\alpha} \le 1$ follows from (5.26) and (5.30). Applying this result to X', we have $\underline{\alpha} = 1 - \bar{\alpha}' \ge 0$. Thus, it remains only to show that $\underline{\alpha} \le \bar{\alpha}$. Since h is submultiplicative, we have $1 = h_X(1) \le h_X(t)h_X(1/t)$. Hence, for all $t > 1$,

$$\frac{\log h_X(1/t)}{\log(1/t)} = \frac{\log\left(\dfrac{1}{h_X(1/t)}\right)}{\log t} \le \frac{\log h_X(t)}{\log t}.$$

Letting $t \to \infty$, we obtain from (5.30) that $\underline{\alpha} \le \bar{\alpha}$, as desired. ■

In the case under consideration ($p_i = q_i$), the Calderón operator S_σ splits into two operators of the types defined below.

Definition 5.14. Let P_a, $(0 < a \le 1)$, be the integral operator defined on $\mathscr{M}_0(\mathbf{R}^+, m)$ by

$$(P_a f)(t) = t^{-a}\int_0^t s^a f(s)\frac{ds}{s}, \qquad (0 < t < \infty). \tag{5.35}$$

Similarly, let Q_a, $(0 \le a < 1)$, be the integral operator defined on $\mathscr{M}_0(\mathbf{R}^+, m)$ by

$$(Q_a f)(t) = t^{-a}\int_t^\infty s^a f(s)\frac{ds}{s}, \qquad (0 < t < \infty). \tag{5.36}$$

Note that Q_b is the formal adjoint of P_a when $a + b = 1$. In other words, as an interchange in the order of integration shows,

$$\int_0^\infty (P_a f)(t)g(t)\,dt = \int_0^\infty f(t)(Q_b g)(t)\,dt, \tag{5.37}$$

for all f and g for which the integrals exist.

Theorem 5.15. *The operator P_a is bounded on $\bar{X}$ if and only if $a > \bar{\alpha}_X$, and Q_a is bounded on $\bar{X}$ if and only if $a < \underline{\alpha}_X$.*

Proof. Suppose first that P_a is bounded on $\bar{X}$. Let f and g be functions in $\bar{X}$ and $\bar{X}'$, respectively, such that

$$\|f\|_{\bar{X}} \le 1, \qquad \|g\|_{\bar{X}'} \le 1. \tag{5.38}$$

Then $\int_0^\infty f^*(s/t)g^*(s)\,ds$ decreases with t and so, for each fixed $t > 0$, we have

$$\begin{aligned}\int_0^\infty f^*\left(\frac{s}{t}\right)g^*(s)\,ds &= at^a\left(\int_0^\infty f^*\left(\frac{s}{t}\right)g^*(s)\,ds\right)\left(\int_0^{1/t} u^{a-1}\,du\right)\\ &\le at^a\int_0^{1/t}\left(\int_0^\infty f^*(su)g^*(s)\,ds\right)u^{a-1}\,du\\ &= at^a\int_0^\infty g^*(s)\left(\int_0^{1/t} f^*(su)u^a\frac{du}{u}\right)ds.\end{aligned}$$

If $t > 1$, we may extend the range of integration in the inner integral to $0 \le u \le 1$. Then, by a change of variables, we obtain from (5.38),

$$\begin{aligned}\int_0^\infty f^*\left(\frac{s}{t}\right)g^*(s)\,ds &\le at^a\int_0^\infty g^*(s)\left(s^{-a}\int_0^s f^*(v)v^a\frac{dv}{v}\right)ds\\ &= at^a\int_0^\infty g^*(s)(P_a f^*)(s)\,ds\\ &\le at^a\|P_a\|_{\mathscr{B}(\bar{X})}.\end{aligned}$$

Taking the supremum over all f and g satisfying (5.38), we therefore have

$$h_X(t) = \|E_{1/t}\|_{\mathscr{B}(\bar{X})} \le at^a\|P_a\|_{\mathscr{B}(\bar{X})}, \qquad (t > 1).$$

Hence,

$$\frac{\log h_X(t)}{\log t} \le a + \frac{\log(a\|P_a\|)}{\log t} \to a,$$

as $t \to \infty$, and so it follows from (5.31) that $\bar{\alpha}_X \le a$. Thus, we have shown that

$$\|P_a\| = \|P_a\|_{\mathscr{B}(\bar{X})} < \infty \quad\Rightarrow\quad a \ge \bar{\alpha}_X. \tag{5.39}$$

We still need to obtain the strict inequality $a > \bar{\alpha}_X$. Choose $\varepsilon > 0$ sufficiently small so that $\varepsilon\|P_a\| < 1$. Then the operator $I - \varepsilon P_a$ belongs to $\mathscr{B}(\bar{X})$, is invertible, and

$$(I - \varepsilon P_a)^{-1} = \sum_{n=0}^\infty \varepsilon^n P_a^n, \tag{5.40}$$

where the convergence is in the norm of $\mathscr{B}(\bar{X})$. The operator

$$T = P_a(I - \varepsilon P_a)^{-1} = \sum_{n=0}^\infty \varepsilon^n P_a^{n+1} \tag{5.41}$$

is therefore also in $\mathscr{B}(\bar{X})$. We claim that the iterate P_a^{n+1} of P_a may be written in the closed form

$$(P_a^{n+1}f)(t) = \int_0^1 f(st)\frac{(\log 1/s)^n}{n!}s^{a-1}\,ds. \tag{5.42}$$

The proof proceeds by induction on n. The case $n = 0$ follows immediately from the definition of P_a, so suppose (5.42) holds for $n = 0, 1, 2, \ldots, N$. Then

$$\begin{aligned}(P_a^{N+1}f)(t) = P_a(P_a^N f)(t) &= \int_0^1 P_a^N f(rt)r^{a-1}\,dr\\ &= \int_0^1\left(\int_0^1 f(srt)\frac{(\log 1/s)^N}{N!}s^{a-1}\,ds\right)r^{a-1}\,dr,\end{aligned}$$

so making the change of variable $u = rs$, we have

$$(P_a^{N+1}f)(t) = \int_0^1 \left(\int_0^r f(ut) \frac{(\log r/u)^N}{N!} u^{a-1}\, du \right) \frac{dr}{r}.$$

Interchanging the order of integration and making the change of variable $v = r/u$, we obtain

$$\begin{aligned}(P_a^{N+1}f)(t) &= \int_0^1 \left(\int_u^1 \frac{(\log r/u)^N}{N!} \frac{dr}{r} \right) f(ut)u^{a-1}\, du \\ &= \int_0^1 \left(\int_1^{1/u} \frac{(\log v)^N}{N!} \frac{dv}{v} \right) f(ut)u^{a-1}\, du \\ &= \int_0^1 \frac{(\log 1/u)^{N+1}}{(N+1)!} f(ut)u^{a-1}\, du.\end{aligned}$$

This completes the induction and hence establishes (5.42) for all n.

Combining (5.41), (5.42), and using Beppo Levi's theorem, we obtain, for nonnegative functions f in $\bar{X}$,

$$(Tf)(t) = \int_0^1 \left(\sum_{n=0}^{\infty} \frac{(\varepsilon \log 1/s)^n}{n!} \right) f(st)s^{a-1}\, ds = \int_0^1 f(st)s^{a-\varepsilon-1}\, ds.$$

By the usual device of splitting a function into its positive and negative parts, we obtain this identity for all f in $\bar{X}$. Hence, $T = P_{a-\varepsilon}$. Since $T \in \mathscr{B}(\bar{X})$, we may therefore apply (5.39) to obtain $a - \varepsilon \geq \bar{\alpha}_X$. Hence, $a > \bar{\alpha}_X$, as desired.

Suppose conversely that $a > \bar{\alpha}_X$. Then Lemma 5.9 shows that

$$\int_0^1 \|E_s\|_{\mathscr{B}(\bar{X})} s^{a-1}\, ds = \int_1^{\infty} t^{-a} h_X(t) \frac{dt}{t} < \infty. \tag{5.43}$$

Hence, if f and g satisfy (5.38),

$$\begin{aligned}\left| \int_0^{\infty} (P_a f)(t)g(t)\, dt \right| &\leq \int_0^{\infty} \left(\int_0^1 |f(st)|s^{a-1}\, ds \right) |g(t)|\, dt \\ &= \int_0^1 \left(\int_0^{\infty} |f(st)g(t)|\, dt \right) s^{a-1}\, ds \\ &\leq \int_0^1 \|E_s\|_{\mathscr{B}(\bar{X})} s^{a-1}\, ds.\end{aligned}$$

Using (5.43) and taking the supremum over all f and g satisfying (5.38), we find that P_a is a bounded operator on $\bar{X}$. Thus we have shown that P_a is bounded on $\bar{X}$ if and only if $a > \bar{\alpha}_X$.

If now $0 \leq a < 1$, it follows from (5.37) that Q_a is bounded on $\bar{X}$ if and only if P_{1-a} is bounded on $\bar{X}'$. By the result we have just proved, this occurs if and only if $1 - a > \bar{\alpha}_{X'}$. But $\bar{\alpha}_{X'} = 1 - \underline{\alpha}_X$, by (5.33). Hence Q_a is bounded on $\bar{X}$ if and only if $a < \underline{\alpha}_X$. ■

Theorem 5.16 (D. W. Boyd). *Suppose $1 \leq p < q \leq \infty$. Let X be a rearrangement-invariant Banach function space over an infinite, nonatomic, totally σ-finite measure space (R, μ). Then every linear (or quasilinear) operator T of joint weak type $(p, p; q, q)$ is bounded from X to X if and only if the indices $\underline{\alpha}$ and $\bar{\alpha}$ of X satisfy*

$$\frac{1}{q} < \underline{\alpha} \leq \bar{\alpha} < \frac{1}{p}. \tag{5.44}$$

Proof. The interpolation segment here is

$$\sigma = \left[\left(\frac{1}{p}, \frac{1}{p} \right), \left(\frac{1}{q}, \frac{1}{q} \right) \right]$$

so the corresponding Calderón operator S_σ is given by

$$S_\sigma = P_{1/p} + Q_{1/q}.$$

Suppose first that the indices of X satisfy (5.44). Then, by Theorem 5.15, the operators $P_{1/p}$ and $Q_{1/q}$ are bounded on $\bar{X}$. Thus, if f belongs to X and g to $\mathscr{M}_0(R, \mu)$ with $g^* \leq S_\sigma(f^*)$, we see from (5.29) that

$$\|g\|_X = \|g^*\|_{\bar{X}} \leq (\|P_{1/p}\| + \|Q_{1/q}\|)\|f^*\|_{\bar{X}} = c\|f\|_X.$$

This shows that X satisfies property (iii) of Theorem 5.7. Hence, that result establishes the desired interpolation property for X.

Conversely, suppose X satisfies property (iii) of Theorem 5.7. We shall show that S_σ is a bounded operator on $\bar{X}$. If h belongs to X, then both $h^*(t)$ and $S_\sigma(h^*)(t)$ tend to 0 as $t \to \infty$. Hence, by Corollary II.7.8, there are μ-measurable functions f and g on R such that $f^* = h^*$ and $g^* = S_\sigma(h^*)$. It follows from (5.22) and (5.23) that f belongs to X. Hence, by property (iii) of Theorem 5.7, we see that g belongs to X and $\|g\|_X \leq c\|f\|_X$. Using (5.7), we obtain

$$\begin{aligned}\|S_\sigma(h)\|_{\bar{X}} &\leq \|S_\sigma(h^*)\|_{\bar{X}} = \|g^*\|_{\bar{X}} = \|g\|_X \\ &\leq c\|f\|_X = c\|f^*\|_{\bar{X}} = c\|h^*\|_{\bar{X}} = c\|h\|_{\bar{X}},\end{aligned}$$

which shows that S_σ is bounded on $\bar{X}$. This implies that both $P_{1/p}$ and $Q_{1/q}$ are bounded on $\bar{X}$ and hence, by Theorem 5.15, that the indices of X satisfy (5.44). ■

We can now present the following elegant characterization of the rearrangement-invariant spaces on which the Hardy-Littlewood maximal operator and the Hilbert transform are bounded operators.

Theorem 5.17 (G. G. Lorentz & T. Shimogaki). *Let X be a rearrangement-invariant Banach function space on $\mathbf{R}^n$. Then the Hardy-Littlewood maximal operator M is bounded on X if and only if the upper index of X satisfies $\bar{\alpha}_X < 1$.*

Proof. It follows from Theorem 3.8 that M is bounded on X if and only if P_1 is bounded on $\bar{X}$. This, by Theorem 5.15, occurs if and only if $\bar{\alpha}_X < 1$. ■

Theorem 5.18 (D. W. Boyd). *Let X be a rearrangement-invariant Banach function space on $\mathbf{R}$. Then the Hilbert transform H is bounded on X if and only if the indices of X satisfy*

$$0 < \underline{\alpha}_X \leq \bar{\alpha}_X < 1.$$

Proof. If $0 < \underline{\alpha}_X \leq \bar{\alpha}_X < 1$, then, by Theorem 5.15, the operators P_1 and Q_0 are bounded on $\bar{X}$. In that case, Theorem 4.8 shows that H is bounded on X. Conversely, if H is bounded on X, then Proposition 4.10 shows that to each f in X, there corresponds a function g equimeasurable with f such that $S(f^*) \leq 2(Hg)^*$. Then

$$\|S(f^*)\|_{\bar{X}} \leq 2\|(Hg)^*\|_{\bar{X}} = 2\|Hg\|_X \leq 2c\|g\|_X = 2c\|f^*\|_{\bar{X}}$$

for all $f \in X$. As in the preceding proof, this is enough to show that S is bounded on $\bar{X}$ and hence that P_1 and Q_0 are bounded on $\bar{X}$. By Theorem 5.15, the indices of X must satisfy $\bar{\alpha}_X < 1$ and $\underline{\alpha}_X > 0$, and this completes the proof. ■

Although we have developed the foregoing theory of the Boyd indices only in the context of infinite nonatomic measure spaces, only minor modifications in the definitions are needed to produce a comparable theory for the other types of resonant measure spaces, namely, finite nonatomic measure spaces (such as the unit circle $\mathbf{T}$) or atomic measure spaces all of whose atoms have equal measure (such as the integers $\mathbf{Z}$). We shall address these matters in more detail in the next section in connection with an analysis of the Fourier transform.

6. NORM-CONVERGENCE OF FOURIER SERIES

Many of the ideas developed in preceding sections will be brought together here as we consider the question of norm-convergence of Fourier series in rearrangement-invariant spaces. In order to provide the proper framework for dealing with Fourier series on the unit circle $\mathbf{T}$, we shall first need to interpret some of our previous results for rearrangement-invariant spaces in terms of trigonometric polynomials instead of simple functions. This merely focuses attention on the order ideal X_b of X rather than on X itself. The next step (Lemmas 6.6 and 6.7) is to translate the question of norm-convergence of Fourier series to one of boundedness of a "multiplier" operator $\mathscr{m}$ (which multiplies the sequence of Fourier coefficients of a function by the fixed sequence $n \to -i \cdot \operatorname{sgn}(n)$). The multiplier operator can be described as a principal-value integral, which is nothing more than the periodic analogue of the Hilbert transform discussed in Section 4. The resulting operator is sometimes referred to as the *periodic Hilbert transform*, but its more usual name, derived from its role in harmonic function theory, is the *conjugate-function operator*.

Many of the properties of the conjugate-function operator can be derived from their counterparts for the Hilbert transform established in Section 4. This may be seen in the proof of Theorem 6.8, where we majorize the corresponding maximal operator with the appropriate Calderón operator and thereby establish the joint weak type $(1, 1; \infty, \infty)$ estimate and the a.e. existence of the principal-value integral. This enables us (in Theorem 6.10) to express norm-convergence of Fourier series in terms of boundedness of the Calderón operator and hence, as in Section 5, in terms of the Boyd indices of the space. The only detail to be attended to here is an appropriate definition of the Boyd indices for function spaces on the circle (they were defined previously only with respect to infinite nonatomic measure spaces). This, however, presents little difficulty.

We shall denote by $\mathbf{T}$ the unit circle $\{e^{i\theta}: -\pi < \theta \leq \pi\}$, with multiplication as the group operation, and by $\mathbf{Z}$ the additive group of integers.

With the norm

$$\|f\|_{L^1} = \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(e^{i\theta})|\, d\theta,$$

the space $L^1 = L^1(\mathbf{T})$ of integrable functions on $\mathbf{T}$ is a Banach algebra under the *convolution* multiplication

$$(f * g)(e^{i\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{i(\theta-\phi)}) g(e^{i\phi})\, d\phi, \qquad (f, g \in L^1(\mathbf{T})).$$

The function $f * g$ is measurable and, by Fubini's theorem, belongs to L^1 and satisfies

$$\|f * g\|_{L^1} \leq \|f\|_{L^1} \|g\|_{L^1}. \tag{6.1}$$

A further appeal to Fubini's theorem shows that $f * g = g * f$ a.e. on $\mathbf{T}$.

Lemma 6.1. *If f and g belong to $L^1(\mathbf{T})$, then*

$$f * g \prec \|g\|_{L^1} f. \tag{6.2}$$

In particular,

$$(f * g)^*(t) \leq \|g\|_{L^1} f^{**}(t), \qquad (0 < t < 1). \tag{6.3}$$

Moreover, if X is a rearrangement-invariant Banach function space and f belongs to X, then $f * g$ belongs to X and

$$\|f * g\|_X \leq \|g\|_{L^1} \|f\|_X. \tag{6.4}$$

Proof. For fixed g in L^1, let T be the linear operator defined by $Tf = g * f$. Inequality (6.1) shows that T is bounded on L^1, with operator norm at most $\|g\|_{L^1}$, and it is routine to verify that T is bounded on L^∞, also with norm at most $\|g\|_{L^1}$. Hence, the desired estimates (6.2) and (6.3) follow at once from Proposition 2.1 and II.3.2. Now using (6.2) and Theorem 2.2, we obtain (6.4). ■

Definition 6.2. If f belongs to $L^1(\mathbf{T})$, then for each integer n, the *n-th Fourier coefficient* of f is defined by

$$\hat{f}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{i\theta}) e^{-in\theta} \, d\theta. \tag{6.5}$$

The resulting formal trigonometric series $\sum_{-\infty}^{\infty} \hat{f}(n) e^{in\theta}$ is called the *Fourier series* of f; the correspondence is written as

$$f \sim \sum_{n=-\infty}^{\infty} \hat{f}(n) e^{in\theta}.$$

The *N-th partial sum* of the Fourier series is denoted by $S_N f$:

$$(S_N f)(e^{i\theta}) = \sum_{n=-N}^{N} \hat{f}(n) e^{in\theta}. \tag{6.6}$$

Our goal is to characterize the rearrangement-invariant spaces X for which the Fourier series of each function f in X converges in the norm of X to f, that is,

$$\lim_{N \to \infty} \|S_N f - f\|_X = 0, \qquad (f \in X). \tag{6.7}$$

In this case, we shall say simply that *Fourier series converge in norm in X.*

Any sum of the form

$$P(e^{i\theta}) = \sum_{n=-N}^{N} a_n e^{in\theta}$$

is referred to as a *trigonometric polynomial* (of degree N if a_N or a_{-N} is nonzero). Clearly, $S_N P = P$ for any trigonometric polynomial P of degree at most N.

The subspace of X for which Fourier series converge in norm is necessarily separable because the continuous embedding of L^∞ in X allows for approximation of $S_N f$ by trigonometric polynomials with rational coefficients. The next lemma equates this subspace with the closure X_b of the simple functions in X, provided the fundamental function φ_X tends to zero at the origin. The latter condition merely rules out the space $X = L^\infty$. We denote by T_ϕ the *translation operator*:

$$T_\phi f(e^{i\theta}) = f(e^{i(\theta+\phi)}), \qquad (-\pi < \theta, \phi \leq \pi),$$

and by $\omega(f, \cdot)_X$ the *X-modulus of continuity* of f:

$$\omega(f, t)_X = \sup_{|\phi| \leq t} \|T_\phi f - f\|_X, \qquad (0 \leq t \leq \pi).$$

Lemma 6.3. *Let X be a rearrangement-invariant space and let X_b be the closure in X of the simple functions. The following assertions are equivalent:*

(i) $\varphi_X(0+) = 0$;
(ii) *X_b is the closure in X of the continuous functions;*
(iii) *X_b is the closure in X of the trigonometric polynomials;*
(iv) *translation is continuous in X_b, that is,*

$$\lim_{\phi \to 0} \|T_\phi f - f\|_X = 0, \qquad (f \in X_b);$$

(v)
$$\lim_{t \to 0+} \omega(f, t)_X = 0, \qquad (f \in X_b).$$

Proof. To show that (i) implies (ii), it will suffice to prove that each simple function f in X may be approximated arbitrarily closely in X by continuous functions. To this end, if $\varepsilon > 0$, we may choose a subset E of $\mathbf{T}$ of measure at most ε and a continuous function g on $\mathbf{T}$ satisfying $\|g\|_{L^\infty} \leq \|f\|_{L^\infty}$ and $|f - g| \leq \varepsilon$ on the complement E^c of E. Then

$$\begin{aligned} \|f - g\|_X &\leq \|f\chi_E\|_X + \|(f - g)\chi_{E^c}\|_X + \|g\chi_E\|_X \\ &\leq 2\|f\|_{L^\infty} \varphi_X(\varepsilon) + \varepsilon \varphi_X(1), \end{aligned}$$

which is small when ε is small, provided (i) holds.

That (ii) implies (iii) is an immediate consequence of the fact that the trigonometric polynomials are uniformly dense in the space of continuous functions.

Suppose now that (iii) holds. Let $f \in X_b$ and choose $\varepsilon > 0$. There is a trigonometric polynomial P such that $\|f - P\|_X < \varepsilon/4$ and then by uniform continuity of P we may choose $\delta > 0$ such that

$$\|T_\phi P - P\|_{L^\infty} < \frac{\varepsilon}{2\varphi_X(1)}, \qquad (|\phi| \le \delta).$$

Since X is rearrangement-invariant, hence translation-invariant, we therefore have

$$\begin{aligned}\|T_\phi f - f\|_X &\le \|T_\phi(f - P) + (T_\phi P - P) + (P - f)\|_X \\ &\le 2\|f - P\|_X + \|T_\phi P - P\|_{L^\infty}\varphi_X(1) \le \varepsilon,\end{aligned}$$

whenever $|\phi| \le \delta$. Hence, translation is continuous in X_b, that is, (iv) holds.

That (iv) implies (v) is clear, so to complete the proof we need only show that (v) implies (i). However, this is immediate if the fundamental function is written in the form

$$\varphi_X(2\phi) = \|T_\phi\chi - \chi\|_X \le \omega(\chi, |\phi|)_X,$$

with χ the characteristic function of an appropriate interval. ■

Remark 6.4. In view of the previous lemma and the remarks preceding it, we shall assume for the remainder of this section that $\varphi_X(0+) = 0$.

Lemma 6.5. *If f belongs to X_b and $g \prec f$, then g belongs to X_b. In particular, X_b is rearrangement-invariant.*

Proof. If f is bounded, then g is bounded and hence belongs to X_b. If f is unbounded, then $\lim_{t\to 0+} f^*(t) = \infty$. We show that g has absolutely continuous norm and hence belongs to X_b. If $E_n \downarrow \varnothing$ and $t_n = |E_n|$, let $\tilde{E}_n = \{|f| \ge f^*(t_n)\}$ and $\tilde{t}_n = |\tilde{E}_n|$. Then $g\chi_{E_n} \prec f\chi_{\tilde{E}_n}$ so Corollary II.4.7 shows that $\|g\chi_{E_n}\|_X \le \|f\chi_{\tilde{E}_n}\|_X$. But f has absolutely continuous norm and $\tilde{t}_n \to 0$ (since $f^*(t) \to \infty$ as $t \to 0$) so this shows that g also has absolutely continuous norm.

With these preliminaries in place, we can begin our discussion of norm-convergence of Fourier series.

Lemma 6.6. *A necessary and sufficient condition that Fourier series converge in norm in X_b is that the operator norms of the partial-sum operators S_N be uniformly bounded:*

$$\|S_N\|_{\mathscr{B}(X_b)} \le M, \qquad (N = 0, 1, 2, \ldots). \tag{6.8}$$

Proof. Suppose (6.8) holds and let f belong to X_b. If $\varepsilon > 0$, Lemma 6.3 provides a trigonometric polynomial P such that $\|f - P\|_X < \varepsilon/(M + 1)$. Then, for every $N \ge \text{degree}(P)$,

$$\begin{aligned}\|f - S_N f\|_X &\le \|f - P\|_X + \|S_N(f - P)\|_X \\ &\le (1 + M)\|f - P\|_X < \varepsilon.\end{aligned}$$

Hence, $S_N f \to f$ in norm in X_b.

Conversely, if Fourier series converge in norm in X_b, then $\{S_N f\}$ is a norm bounded subset of X_b for each f in X_b. It follows from the uniform boundedness principle that the operator norms of the S_N are uniformly bounded. ■

We next reformulate condition (6.8) in terms of the boundedness of the "multiplier" operator $\mathscr{m}$ which multiplies the Fourier coefficients of the operand by the fixed sequence $-i \cdot \operatorname{sgn}(n)$. The operator $\mathscr{m}$ is defined initially on trigonometric polynomials $P(e^{i\theta}) = \sum_{-N}^{N} a_n e^{in\theta}$ by

$$\mathscr{m}P(e^{i\theta}) = \sum_{-N}^{N} -i \cdot \operatorname{sgn}(n) a_n e^{in\theta}.$$

Lemma 6.7. *The estimate* (6.8) *holds if and only if there is a constant c such that*

$$\|\mathscr{m}(P)\|_X \le c\|P\|_X$$

for all trigonometric polynomials P. In that case, $\mathscr{m}$ has a unique extension to a bounded linear operator on all of X_b, which satisfies $\|\mathscr{m}\|_{\mathscr{B}(X_b)} \le c$ and

$$(\mathscr{m}f)\hat{\ }(n) = -i \cdot \operatorname{sgn}(n)\hat{f}(n), \qquad (f \in X_b, n \in \mathbf{Z}). \tag{6.9}$$

Proof. The operator T defined on trigonometric polynomials P by

$$T(P)(e^{i\theta}) = \sum_{n=0}^{N} a_n e^{in\theta}, \qquad \text{where } P(e^{i\theta}) = \sum_{-N}^{N} a_n e^{in\theta},$$

satisfies

$$T(P) = \frac{1}{2}\hat{P}(0) + \frac{1}{2}[P + i\mathscr{m}(P)].$$

Since $|\hat{P}(0)| \le \|P\|_{L^1} \le c\|P\|_X$, it follows that $\mathscr{m}$ is bounded on X_b if and only if T is.

If P is a trigonometric polynomial of degree N, then

$$T(P)(e^{i\theta}) = e^{iN\theta} S_N(e^{-iN(\cdot)}P(e^{i(\cdot)}))(e^{i\theta}).$$

Hence, if (6.8) holds,

$$\|T(P)\|_X \le \|S_N(e^{-iN(\cdot)}P(e^{i(\cdot)}))\|_X \le M\|P\|_X.$$

Since the trigonometric polynomials are dense in X_b, we conclude that T (hence $\mathscr{m}$) has a unique norm-preserving extension to X_b.

Conversely, if $\mathscr{m}$ (hence T) is bounded on X_b, then since

$$S_N(P)(e^{i\theta}) = e^{-iN\theta}T(e^{iN(\cdot)}P)(e^{i\theta}) - e^{i(N+1)\theta}T(e^{-i(N+1)(\cdot)}P)(e^{i\theta}),$$

the partial sum operators S_N have uniformly bounded operator norms.

It remains only to establish (6.9). If f belongs to X_b, choose any sequence of trigonometric polynomials P_j that converges to f in X_b. Then $\mathscr{m}(P_j) \to \mathscr{m}(f)$ in X_b because $\mathscr{m}$ has a bounded extension on X_b. Noting that

$$|\hat{g}(n)| \le \|g\|_{L^1} \le c\|g\|_X, \qquad (n \in \mathbf{Z}), \tag{6.10}$$

for any g in X_b, and applying this first to $g = \mathscr{m}f - \mathscr{m}P_j$, then to $g = f - P_j$, we obtain

$$(\mathscr{m}f)\hat{\ }(n) = \lim_{j\to\infty} \mathscr{m}(P_j)\hat{\ }(n) = -i\cdot \operatorname{sgn}(n)\cdot \lim_{j\to\infty}\hat{P}_j(n) = -i\cdot\operatorname{sgn}(n)\hat{f}(n).$$

This establishes (6.9) and hence completes the proof. ∎

Next, we consider the principal-value integral

$$\tilde{f}(e^{i\theta}) = \frac{1}{2\pi}\lim_{\varepsilon\to 0+}\int_{\varepsilon<|\phi|\le\pi} f(e^{i(\theta-\phi)})\cot\left(\frac{\phi}{2}\right)d\phi. \tag{6.11}$$

The operator $f \to \tilde{f}$, which is known as the *conjugate-function operator*, is a periodic analogue of the Hilbert transform, which we discussed in Section 4. The a.e.-existence of the principal-value integral will be established in the same way, via an analysis of the corresponding maximal operator, that is, the *maximal conjugate-function operator*

$$(\mathscr{C}f)(e^{i\theta}) = \sup_{0<\varepsilon\le\pi}\frac{1}{2\pi}\left|\int_{\varepsilon<|\phi|\le\pi} f(e^{i(\theta-\phi)})\cot\left(\frac{\phi}{2}\right)d\phi\right|. \tag{6.12}$$

As with the Hilbert transform, the key result is the domination of the maximal operator by the appropriate Calderón operator S:

$$S(f^*)(t) = \frac{1}{t}\int_0^t f^*(s)\,ds + \int_t^1 f^*(s)\frac{ds}{s}.$$

Theorem 6.8. *If f belongs to $L^1(\mathbf{T})$, then*

$$(\mathscr{C}f)^*(t) \le cS(f^*)(t), \qquad (0<t<1), \tag{6.13}$$

for some constant c independent of f. Furthermore, the conjugate function $\tilde{f}$ exists a.e. *and satisfies*

$$(\tilde{f})^*(t) \le cS(f^*)(t), \qquad (0<t<1). \tag{6.14}$$

Proof. We shall exploit the corresponding results for the Hilbert transform developed in Section 4. If $f \in L^1(\mathbf{T})$, let $F(t) = f(e^{it})$ be its periodic extension to the interval $[-2\pi, 2\pi]$ and define F to be zero elsewhere on the real line. By (4.3), the Hilbert transform HF of F is the limit as $\varepsilon\to 0$ of

$$(H_\varepsilon F)(t) = \frac{1}{\pi}\int_{\varepsilon\le|s|} F(t-s)\frac{ds}{s}.$$

Let

$$k(s) = \begin{cases} \dfrac{1}{s} - \dfrac{1}{2}\cot\left(\dfrac{s}{2}\right), & 0<|s|<\pi,\\ 0, & \text{elsewhere.}\end{cases}$$

The function k is continuous on $(-\pi,\pi)$ and is bounded on the real line by $1/\pi$. Hence, if $0<\varepsilon<\pi$ and $|\theta|\le\pi$, we have

$$\begin{aligned}\frac{1}{2\pi}&\left|\int_{\varepsilon<|\phi|\le\pi} f(e^{i(\theta-\phi)})\cot\left(\frac{\phi}{2}\right)d\phi\right|\\ &= \frac{1}{\pi}\left|\int_{\varepsilon<|\phi|\le\pi} F(\theta-\phi)\left(\frac{1}{\phi} - k(\phi)\right)d\phi\right|\\ &\le |H_\varepsilon F)(\theta) - (H_\pi F)(\theta)| + \frac{1}{\pi}\int_{-\pi}^{\pi}|f(e^{i(\theta-\phi)})|\cdot|k(\phi)|\,d\phi.\end{aligned}$$

Taking the supremum over all such ε, we obtain

$$(\mathscr{C}f)(e^{i\theta}) \le 2[(\mathscr{H}F)(\theta) + (|f| * |k|)(e^{i\theta})].$$

Passing to the decreasing rearrangements, we use (II.1.16), (4.27), and (6.3) to conclude that

$$(\mathscr{C}f)^*(t) \le 2\left[(\mathscr{H}F)^*\left(\frac{t}{2}\right) + f^{**}\left(\frac{t}{2}\right)\right] \le c\left[S(F^*)\left(\frac{t}{2}\right) + f^{**}\left(\frac{t}{2}\right)\right].$$

But $F^*(s) = f^*(s/4\pi)$, for $0<s<4\pi$, and is zero otherwise, so, with the appropriate changes of variables, the last estimate reduces to (6.13).

It remains only to establish the a.e.-existence of the principal-value integral for $\tilde{f}$ because then we shall have $|\tilde{f}| \le \mathscr{C}f$ a.e. and (6.14) will follow directly from (6.13). If

$$(\Gamma f)(e^{i\theta}) = \inf_{\varepsilon>0}\left\{\sup_{0<\delta_1<\delta_2<\varepsilon}\left|\int_{\delta_1<|\phi|<\delta_2} f(e^{i(\theta-\phi)})\cot\left(\frac{\phi}{2}\right)d\phi\right|\right\}, \qquad (6.15)$$

then $\tilde{f}(e^{i\theta})$ exists if and only if $(\Gamma f)(e^{i\theta}) = 0$.

If g is smooth, say

$$|g(e^{i(\theta+\phi)}) - g(e^{i(\theta-\phi)})| \leq M|\phi|, \qquad (\theta, \phi \in (-\pi, \pi]),$$

for example, then

$$\left|\int_{\delta_1<|\phi|<\delta_2} g(e^{i(\theta-\phi)})\cot\left(\frac{\phi}{2}\right)d\phi\right|$$

$$= \left|\int_{\delta_1<\phi<\delta_2}[g(e^{i(\theta+\phi)}) - g(e^{i(\theta-\phi)})]\cot\left(\frac{\phi}{2}\right)d\phi\right|$$

$$\leq M\int_0^{\delta_2}\phi\cot\left(\frac{\phi}{2}\right)d\phi \to 0 \qquad \text{as } \delta_2 \to 0,$$

and so $\Gamma g = 0$ everywhere. The subadditivity of Γ therefore gives $\Gamma f \leq \Gamma(f-g)$ and $\Gamma(f-g) \leq \Gamma f$. Hence, $\Gamma f = \Gamma(f-g)$ for all smooth g.

In order to show that $\Gamma f = 0$ a.e. for arbitrary $f \in L^1$, we shall show that $(\Gamma f)^*(t) = 0$ for $0 < t < 1$. To this end, fix t and choose a smooth g (a trigonometric polynomial for example) such that $\|f-g\|_{L^1} < \varepsilon t/2c$, where c is the constant in (6.13). Using (6.13) and the estimate $\Gamma f \leq 2\mathscr{C}f$, we obtain

$$(\Gamma f)^*(t) = \Gamma(f-g)^*(t) \leq 2[\mathscr{C}(f-g)]^*(t)$$

$$\leq 2cS(f-g)^*(t) \leq \frac{2c}{t}\|f-g\|_{L^1} < \varepsilon.$$

Since ε is an arbitrary positive number, we conclude that $(\Gamma f)^*(t) = 0$ for $0 < t < 1$. As we remarked above, this shows that $\Gamma f = 0$ a.e. and establishes the existence of the principal-value integral. ■

Now we can describe the multiplier operator m in terms of the conjugate-function operator. We begin with trigonometric polynomials.

Lemma 6.9. *If P is a trigonometric polynomial, then*

$$mP = \tilde{P}. \qquad (6.16)$$

Proof. The identity (6.16) need only be verified for the exponentials $\{e^{iNt}: N \in \mathbf{Z}\}$ since these form a basis for the trigonometric polynomials and since

both operators in question are linear. The case $N = 0$ is trivial. Moreover, the proof for $N < 0$ will follow that for $N > 0$ by a change of variables. Hence, we may assume $N > 0$. In that case, we have

$$(e^{iN(\cdot)})^{\sim}(e^{i\theta}) = \frac{1}{2\pi}\,\text{p.v.}\int_{-\pi}^{\pi} e^{iN(\theta-\phi)}\cot\left(\frac{\phi}{2}\right)d\phi$$

$$= \frac{1}{2\pi}\left\{\text{p.v.}\int_{-\pi}^{\pi} e^{-iN\phi}\cot\left(\frac{\phi}{2}\right)d\phi\right\}e^{iN\theta}$$

$$= -i\left\{\frac{1}{2\pi}\,\text{p.v.}\int_{-\pi}^{\pi}\sin(N\phi)\cos\left(\frac{\phi}{2}\right)d\phi\right\}e^{iN\theta}$$

$$= -i\cdot\operatorname{sgn}(N)e^{iN\theta} = m(e^{iN(\cdot)})(e^{i\theta}),$$

because of the following elementary identity:

$$\sin(N\phi)\cot\left(\frac{\phi}{2}\right) = 1 + 2\sum_{k=1}^{N-1}\cos(k\phi) + \cos(N\phi). \qquad ■$$

Theorem 6.10. *Suppose X is a rearrangement-invariant Banach function space on $\mathbf{T}$ whose fundamental function satisfies $\varphi_X(0+) = 0$. Let X_b be the closure of the simple functions in X. Then the following conditions are equivalent:*

(i) *Fourier series converge in norm in X_b;*
(ii) *the partial-sum operators S_N are uniformly bounded on X_b;*
(iii) *the multiplier operator m is bounded on X_b;*
(iv) *the conjugate-function operator is bounded on X_b;*
(v) *the Calderón operator*

$$Sf^*(t) = \int_0^1 f^*(s)\cdot\min\left(1, \frac{s}{t}\right)\frac{ds}{s}$$

is bounded on $(X_b)^{-}$, the Luxemburg representation of X_b on the interval $[0,1]$.

If any one of these conditions holds, then $mf = \tilde{f}$ a.e. for all f in X_b.

Proof. Lemmas 6.6 and 6.7 show that (i), (ii), and (iii) are equivalent. That (iv) implies (iii) follows at once from Lemmas 6.7 and 6.9. We now establish the reverse implication. For each $f \in X_b$, let

$$(\Lambda f)(e^{i\theta}) = |\tilde{f}(e^{i\theta}) - mf(e^{i\theta})|.$$

The hypothesis that m is bounded on X_b gives, for each $g \in X_b$,

$$\begin{aligned}(\Lambda g)^*(2t) &\leq (\tilde{g})^*(t) + (mg)^*(t)\\ &\leq c\left\{\frac{\|g\|_{L^1}}{t} + \frac{\|g\|_X}{\varphi_X(t)}\right\} \qquad (6.17)\\ &\leq c \cdot \max\left(\frac{1}{t}, \frac{1}{\varphi_X(t)}\right)\|g\|_X.\end{aligned}$$

Now $\Lambda(f) = \Lambda(f - P)$ for all trigonometric polynomials P. The latter are dense in X_b, so we obtain from (6.17),

$$(\Lambda f)^*(2t) = [\Lambda(f - P)]^*(2t) \leq c \cdot \max\left(\frac{1}{t}, \frac{1}{\varphi_X(t)}\right)\|f - P\|_X,$$

which may be made as small as we please. Hence, $\Lambda f = 0$ a.e. and accordingly $\tilde{f} = mf$ a.e.

That (v) implies (iv) follows directly from Theorem 6.8. Hence, it remains only to show that (iv) implies (v). Given any function f in $\bar{X}_b$, we claim there is a corresponding function f_1 in X_b that satisfies $f_1^* \leq f^*$ and

$$\int_0^{1/2} f^*(s)\min\left(1, \frac{s}{t}\right)\frac{ds}{s} \leq c(\tilde{f}_1)^*(t), \qquad \left(0 < t < \frac{1}{4}\right). \qquad (6.18)$$

The corresponding integral over $[1/2, 1]$ can be majorized by a multiple of the L^1-norm of f, as can the entire integral over $[0, 1]$ when $t \geq 1/4$. Hence, using (6.18), we have

$$(Sf)^*(t) \leq c[(\tilde{f}_1)^*(t) + \|f\|_{L^1}], \qquad (0 < t < 1).$$

Applying the norm of $\bar{X}$ to both sides and using the hypothesis that the conjugate-function operator is bounded on X_b, we therefore obtain

$$\begin{aligned}\|Sf\|_{\bar{X}} &\leq c[\|\tilde{f}_1\|_X + \|f\|_{\bar{X}}]\\ &\leq c[\|f_1\|_X + \|f\|_{\bar{X}}] \leq c\|f\|_{\bar{X}}\end{aligned}$$

Hence, S is bounded on $\bar{X}$, as desired.

To complete the proof, we have only to establish the claim (6.18). For this, set $f_1(e^{it}) = f^*(-t/\pi)\chi_{(-\pi/2, 0)}(t)$. Clearly, $f_1^* \leq f^*$ and if $0 < t < \pi/2$, then

$$\begin{aligned}\tilde{f}_1(e^{it}) &= \frac{1}{2\pi}\text{p.v.}\int_{-\pi/2}^{0} f^*\left(-\frac{s}{\pi}\right)\cot\left(\frac{t - s}{2}\right)ds\\ &= \frac{1}{2}\text{p.v.}\int_0^{1/2} f^*(u)\cot\left(\frac{t + \pi u}{2}\right)du.\end{aligned}$$

Now, $\cot(x/2) \geq 2/(\pi x)$ for $0 < x < \pi/2$ and

$$\frac{1}{t + \pi u} \geq \frac{1}{2\pi}\min\left(\frac{1}{t}, \frac{1}{u}\right), \qquad \left(0 < u < \frac{1}{2}, 0 < t < \frac{\pi}{2}\right).$$

Incorporating these estimates into the preceding one, we obtain (6.18) with $= 2\pi^2$. ■

The preceding result has the following concise formulation in terms of the Boyd indices of the space involved. Slight modifications (discussed below) are needed to define the indices for the circle (they were defined only for *infinite* nonatomic measure spaces in the preceding section).

Corollary 6.11. *Let X be a separable rearrangement-invariant Banach function space on* $\mathbf{T}$. *Fourier series converge in norm in X if and only if the Boyd indices of X satisfy*

$$0 < \underline{\alpha}_X, \bar{\alpha}_X < 1.$$

We remark that for nonseparable spaces, a similar result holds with X_b in place of X.

We now indicate the modifications necessary to define the Boyd indices on the circle (or any finite nonatomic measure space). By the Luxemburg representation theorem, we may assume that the rearrangement-invariant space X is defined on the interval $[0, 1]$. The appropriate choice of dilation operator E_t is in this case given by

$$(E_t f)(s) = \begin{cases} f(st), & 0 \leq s \leq \min(1, 1/t),\\ 0, & \min(1, 1/t) \leq s \leq 1.\end{cases} \qquad (6.19)$$

As before, set

$$h_X(t) = \|E_{1/t}\|_{\mathscr{B}(X)}.$$

Then $h_{L^\infty}(t) \leq 1$ and $h_{L^1}(t) \leq t$ so Proposition 2.1 shows that $E_{1/t}f \prec \max(1, t)f$. The interpolation theorem (Theorem 2.2) then gives the estimate $h_X(t) \leq \max(1, t)$. It follows from (6.19) that $(E_t f)^* \leq E_t(f^*)$ so h_X is given by

$$h_X(t) = \sup\left\{\int_0^1 E_{1/t}f^*(s)g^*(s)\,ds : \|f\|_X \leq 1, \|g\|_{X'} \leq 1\right\}.$$

This shows in particular that h_X is nondecreasing. It is also submultiplicative because $E_{1/st} = E_{1/t} \circ E_{1/s}$, for $0 < s \leq t$. We may therefore appeal to the general theory of submultiplicative functions to define the Boyd indices $\underline{\alpha}_X$ and $\bar{\alpha}_X$ of X (cf. Definition 5.12).

The properties of the indices can now be established exactly as before. The following identity is crucial:

$$h_X(t) = t h_{X'}\left(\frac{1}{t}\right), \qquad (t > 0), \tag{6.20}$$

and is established as follows. We have

$$\int_0^1 E_{1/t}(f^*)(s) g^*(s)\,ds = \int_0^{t\wedge 1} f^*\left(\frac{s}{t}\right) g^*(s)\,ds$$
$$= t \int_0^{1\wedge 1/t} f^*(u) g^*(ut)\,du$$
$$\le t \int_0^1 f^*(u) E_t g^*(u)\,du,$$

from which it follows that $h_X(t) \le t h_{X'}(1/t)$. Applying the same result to X' gives the reverse inequality and we obtain (6.20). With the averaging operators

$$(P_a f)(t) = t^{-a} \int_0^t s^a f(s) \frac{ds}{s}, \qquad (Q_a f)(t) = t^a \int_t^1 s^{-a} f(s) \frac{ds}{s},$$

identity (5.37) remains intact and Theorem 5.15 is proved exactly as before.

The modifications necessary to define the Boyd indices with respect to discrete measure spaces (such as the integers **Z**) are presented in Exercise 15.

7. THEOREMS OF LORENTZ AND SHIMOGAKI

We conclude this chapter with two fundamental results of G. G. Lorentz and T. Shimogaki involving rearrangements of functions and the Hardy-Littlewood-Pólya relation. The first (Theorem 7.4) deals with decreasing rearrangements and differences of functions, establishing the relation

$$f^* - g^* \prec f - g.$$

The second (Theorem 7.7) provides a certain "splitting" of measurable functions with respect to the Hardy-Littlewood-Pòlya relation. It will be used in Chapter V to characterize certain abstract interpolation spaces.

As in Section 2, we shall denote by **a**, **b**, etc., n-dimensional vectors $\mathbf{a} = (a_1, a_2, \ldots, a_n)$, $\mathbf{b} = (b_1, b_2, \ldots, b_n)$, and by $\mathbf{a}^*$ and $\mathbf{b}^*$, etc., their decreasing rearrangements. Recall from (2.3) that two such vectors **a** and **b** satisfy the Hardy-Littlewood-Pòlya relation $\mathbf{b} \prec \mathbf{a}$ if, for each $m = 1, 2, \ldots, n$,

$$\sum_{i=1}^m b_i^* \le \sum_{i=1}^m a_i^*. \tag{7.1}$$

We begin with the following elementary result for two-dimensional vectors.

Lemma 7.1. *Suppose u, v, x, and y are real numbers satisfying*

$$v \le u, \qquad x \le y. \tag{7.2}$$

Then

$$(u - y, v - x) \prec (u - x, v - y). \tag{7.3}$$

Proof. It follows from (7.1) (with $m = 1$ and with $m = 2$) that the desired conclusion (7.3) is equivalent to the pair of inequalities

$$\max\{|u - y|, |v - x|\} \le \max\{|u - x|, |v - y|\} \tag{7.4}$$

and

$$|u - y| + |v - x| \le |u - x| + |v - y|. \tag{7.5}$$

The validity of these inequalities is easily demonstrated by checking them in all possible cases of the inequalities (7.2). ■

The Hardy-Littlewood-Pólya theorem (Theorem 2.7) asserts that if **a** and **b** have nonnegative entries, then $\mathbf{b} \prec \mathbf{a}$ if and only if there is a substochastic matrix A such that $\mathbf{b} = A\mathbf{a}$. If E is a subset of $\{1, 2, \ldots, n\}$, we shall denote its characteristic function by $\chi = \chi_E$. The complement of E will be denoted by E^c.

Lemma 7.2. *Let E be a subset of $\{1, 2, \ldots, n\}$ and let* **b** *and* **c** *be n-dimensional vectors. If $\mathbf{b}\chi_E \prec \mathbf{c}\chi_E$, then $\mathbf{c}\chi_{E^c} + \mathbf{b}\chi_E \prec \mathbf{c}$.*

Proof. Since the decreasing rearrangement depends only on the absolute values of the vector entries, it will be enough to establish the lemma for nonnegative vectors **b** and **c**. If E is empty, the result holds trivially, so suppose E is of the form $E = \{n_1, n_2, \ldots, n_p\}$, where $1 \le n_1 < \cdots < n_p \le n$. Let $\mathbf{b}'$ and $\mathbf{c}'$ be the restrictions to E of **b** and **c**, respectively, regarded as p-dimensional vectors. Thus,

$$b_i' = b_{n_i}, \qquad c_i' = c_{n_i}, \qquad (i = 1, 2, \ldots, p).$$

The hypothesis on **b** and **c** implies that $\mathbf{b}' \prec \mathbf{c}'$. Hence, the Hardy-Littlewood-Pólya theorem (Theorem 2.7) provides a p-dimensional substochastic matrix $P = (p_{st})$ such that $P(\mathbf{c}') = \mathbf{b}'$. Now define an n-dimensional substochastic matrix $Q = (q_{ij})$ by setting

$$q_{ij} = \begin{cases} p_{st}, & \text{if } (i, j) \in E \times E, \text{ say } (i, j) = (n_s, n_t), \\ 1, & \text{if } i = j \text{ and } i \notin E, \\ 0, & \text{otherwise.} \end{cases}$$

If i is not in E, then $(Q\mathbf{c})_i = c_i$. On the other hand, if i is in E, say $i = n_s$, then

$$(Q\mathbf{c})_i = \sum_{j=1}^{n} q_{ij} c_j = \sum_{t=1}^{p} p_{st} c_{n_t} = \sum_{t=1}^{p} p_{st} c'_t = b'_s = b_i.$$

Hence, $Q\mathbf{c} = \mathbf{c}\chi_{E^c} + \mathbf{b}\chi_E$. The desired result can therefore be reformulated as $Q\mathbf{c} \prec \mathbf{c}$, in which form it follows immediately from the Hardy-Littlewood-Pólya theorem since Q is substochastic. ■

Lemma 7.3. *If* $\mathbf{a}$ *and* $\mathbf{b}$ *are n-dimensional vectors with nonnegative entries, then*

$$\mathbf{a}^* - \mathbf{b}^* \prec \mathbf{a} - \mathbf{b}. \tag{7.6}$$

Proof. It will suffice to establish (7.6) for $\mathbf{a}$ and $\mathbf{b}$ with $\mathbf{a}$ decreasing. Indeed, for arbitrary nonnegative vectors $\mathbf{A}$ and $\mathbf{B}$, let σ be a permutation of $\{1,2,\ldots,n\}$ so that $\mathbf{a} = \mathbf{A} \circ \sigma = \mathbf{A}^*$ is decreasing and set $\mathbf{b} = \mathbf{B} \circ \sigma$. By assumption, the estimate (7.6) then holds for $\mathbf{a}$ and $\mathbf{b}$. Clearly, $\mathbf{a}^* = \mathbf{A}^*$ and $\mathbf{b}^* = \mathbf{B}^*$ so the left-hand side of (7.6) is simply $\mathbf{A}^* - \mathbf{B}^*$. On the other hand, $\mathbf{a} - \mathbf{b}$ is given by $\mathbf{A} \circ \sigma - \mathbf{B} \circ \sigma = (\mathbf{A} - \mathbf{B}) \circ \sigma$ so $(\mathbf{a} - \mathbf{b})^* = (\mathbf{A} - \mathbf{B})^*$. Hence, (7.6) holds also with $\mathbf{A}$ and $\mathbf{B}$ in place of $\mathbf{a}$ and $\mathbf{b}$, respectively.

We may therefore assume in what follows that $\mathbf{a}$ is decreasing, that is, $\mathbf{a} = \mathbf{a}^*$. Let $\mathbf{c}^{(0)}$ be the vector $\mathbf{a} - \mathbf{b} = \mathbf{a}^* - \mathbf{b}$. The conclusion (7.6) will be obtained by constructing vectors $\mathbf{c}^{(1)}, \mathbf{c}^{(2)}, \ldots, \mathbf{c}^{(n)}$ such that

$$\mathbf{a}^* - \mathbf{b}^* = \mathbf{c}^{(n)} \prec \mathbf{c}^{(n-1)} \prec \cdots \prec \mathbf{c}^{(1)} \prec \mathbf{c}^{(0)} = \mathbf{a} - \mathbf{b}.$$

Let ν be any permutation of $\{1,2,\ldots,n\}$ for which $b_i^* = b_{\nu(i)}$, $i = 1,2,\ldots,n$. To construct $\mathbf{c}^{(1)}$, let $E_1 = \{1, \nu(1)\}$. The vector $\mathbf{c}^{(0)}$ has components in the 1- and $\nu(1)$-positions equal to

$$a_1^* - b_1 \qquad \text{and} \qquad a_{\nu(1)}^* - b_{\nu(1)},$$

respectively. To form $\mathbf{c}^{(1)}$, we replace these by

$$a_1^* - b_{\nu(1)} \qquad \text{and} \qquad a_{\nu(1)}^* - b_1,$$

respectively, while leaving all other coefficients of $\mathbf{c}^{(0)}$ unchanged. Since $b_{\nu(1)} = b_1^* \geq b_1$ and $a_{\nu(1)}^* \leq a_1^*$, Lemma 7.1 shows that

$$(a_1^* - b_{\nu(1)}, a_{\nu(1)}^* - b_1) \prec (a_1^* - b_1, a_{\nu(1)}^* - b_{\nu(1)}),$$

or, equivalently, that

$$\mathbf{c}^{(1)}\chi_{E_1} \prec \mathbf{c}^{(0)}\chi_{E_1}.$$

Hence, by Lemma 7.2,

$$\mathbf{c}^{(1)} = \mathbf{c}^{(1)}\chi_{E_1^c} + \mathbf{c}^{(1)}\chi_{E_1} \prec \mathbf{c}^{(1)}\chi_{E_1^c} + \mathbf{c}^{(0)}\chi_{E_1} = \mathbf{c}^{(0)}.$$

Note that the first coefficient of $\mathbf{c}^{(1)}$ is equal to the first coefficient of $\mathbf{a}^* - \mathbf{b}^*$.

Now let us repeat the above procedure but starting with $\mathbf{c}^{(1)}$ instead of $\mathbf{c}^{(0)}$. The second coefficient of $\mathbf{c}^{(1)}$ is equal to $a_2^* - b_j$ for some $j \neq \nu(1)$. On the other hand, for some J, $(2 \leq J \leq n)$, the J-th coefficient of $\mathbf{c}^{(1)}$ is equal to $a_J^* - b_{\nu(2)}$. In forming $\mathbf{c}^{(2)}$, we replace these coefficients

$$a_2^* - b_j \qquad \text{and} \qquad a_J^* - b_{\nu(2)}$$

by

$$a_2^* - b_{\nu(2)} \qquad \text{and} \qquad a_J^* - b_j,$$

respectively, while leaving all other components of $\mathbf{c}^{(1)}$ unchanged. Since $b_{\nu(2)} = b_2^* \geq b_j$ (because $j \neq \nu(1)$) and $a_J^* \leq a_2^*$ (because $J \geq 2$), we may apply Lemmas 7.1 and 7.2 exactly as before to find that the new vector $\mathbf{c}^{(2)}$ produced in this way satisfies $\mathbf{c}^{(2)} \prec \mathbf{c}^{(1)}$ and that the first two components of $\mathbf{c}^{(2)}$ are equal to $a_1^* - b_1^*$ and $a_2^* - b_2^*$.

Continuing in this way, we obtain vectors

$$\mathbf{c}^{(n)} \prec \mathbf{c}^{(n-1)} \prec \cdots \prec \mathbf{c}^{(2)} \prec \mathbf{c}^{(1)} \prec \mathbf{c}^{(0)},$$

with $\mathbf{c}^{(n)} = \mathbf{a}^* - \mathbf{b}^*$. Since $\mathbf{c}^{(0)} = \mathbf{a} - \mathbf{b}$, the proof is complete. ■

It is now relatively simple to extend the preceding result to more general measure spaces.

Theorem 7.4 (G. G. Lorentz & T. Shimogaki). *Let f and g be nonnegative integrable functions on a totally σ-finite measure space. Then*

$$f^* - g^* \prec f - g. \tag{7.7}$$

Proof. Suppose first that the underlying measure space (R, μ) is nonatomic and let $\varepsilon > 0$. Then there exist simple functions F and G satisfying $0 \leq F \leq f$, $0 \leq G \leq g$, and

$$\int_R (f - F)\,d\mu \leq \varepsilon, \qquad \int_R (g - G)\,d\mu \leq \varepsilon, \tag{7.8}$$

with all sets of constancy of F and G having common measure, say k. It follows directly from Lemma 7.3 that

$$\int_0^t (F^* - G^*)^*(s)\,ds \leq \int_0^t (F - G)^*(s)\,ds, \tag{7.9}$$

whenever t is of the form $t = mk$ with $m = 0, 1, 2, \ldots$. But each of the integrals is linear for the intermediate values of t so (7.9) in fact holds for all $t > 0$.

Since $0 \le F \le f$, we have $F^* \le f^*$ and so

$$\int_0^\infty (f^* - F^*)^*(s)\,ds = \int_0^\infty |f^* - F^*|(s)\,ds = \int_0^\infty (f^*(s) - F^*(s))\,ds$$
$$= \int_R |f|\,d\mu - \int_R |F|\,d\mu \le \int_R |f - F|\,d\mu.$$

There is a similar estimate for g. In conjunction with (7.8), they give

$$\int_0^\infty (f^* - F^*)^*(s)\,ds \le \varepsilon, \qquad \int_0^\infty (g^* - G^*)^*(s)\,ds \le \varepsilon. \tag{7.10}$$

Using (7.8), (7.9), and (7.10), together with the subadditivity of the map $h \to h^{**}$ (Theorem II.3.4), we obtain for any $t > 0$,

$$\begin{aligned}\int_0^t (f^* - g^*)^* &\le \int_0^t (f^* - F^*)^* + \int_0^t (F^* - G^*)^* + \int_0^t (G^* - g^*)^* \\ &\le \int_0^t (F^* - G^*)^* + 2\varepsilon \le \int_0^t (F - G)^* + 2\varepsilon \\ &\le \int_0^t (F - f)^* + \int_0^t (f - g)^* + \int_0^t (g - G)^* + 2\varepsilon \\ &\le \int_0^t (f - g)^* + 4\varepsilon.\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, this establishes the desired result (7.7) for f and g on nonatomic measure spaces. The analogous result for an arbitrary totally σ-finite measure space follows easily from this one by embedding the measure space into a nonatomic one (cf. the method of retracts in Section 3 of Chapter II). ■

Our next objective (Theorem 7.7) is based on the following refinement of the Hardy-Littlewood-Pólya-Calderón-Ryff results (Theorems 2.7 and 2.10). The proof makes essential use of a construction due to Lorentz and Shimogaki.

Lemma 7.5. *Let f and g be nonnegative decreasing functions in $(L^1 + L^\infty)(\mathbf{R}^+)$. Suppose, in addition, that g is a step function and that $g \prec f$. Then there exists a substochastic operator T such that $Tf = g$ and with the property that Th is decreasing whenever h is.*

Proof. Write

$$g = \sum_{j=1}^n b_j \chi_{(t_{j-1}, t_j)},$$

where $b_1 \ge b_2 \ge \cdots \ge b_n \ge 0$ and $0 = t_0 < t_1 < \cdots < t_n < \infty$. Define a linear operator T_0 by

$$T_0 h = \sum_{j=1}^n h_{I_j} \chi_{I_j},$$

where h_{I_j} is the average of h over the interval $I_j = (t_{j-1}, t_j)$. It follows from Proposition II.3.7 and Proposition 2.4 that T_0 is substochastic. Clearly, $T_0 h$ is a step function relative to the intervals I_j, which is decreasing if h is. Furthermore, it is easy to verify that

$$g \prec T_0 f \prec f, \qquad |T_0 h| \le T_0 |h| \quad \text{(for all } h\text{)}.$$

Since composition preserves these properties, the lemma will be proved if we show that there are finitely many substochastic operators T_j, each of which has the properties that $T_j h$ is a step function relative to the intervals I_j, that $T_j h$ is decreasing whenever h is, that $|T_j h| \le T_j |h|$ for all h, and such that

$$T = T_n \circ \cdots \circ T_1 \circ T_o \qquad \text{and} \qquad Tf = g. \tag{7.11}$$

Suppose $1 \le j \le n$ and that $T_1, T_2, \ldots, T_{j-1}$ have been defined in this way. We show how T_j is determined. Set

$$f_0 = T_0 f, \qquad f_k = T_k f_{k-1}, \qquad (k = 1, 2, \ldots, j-1).$$

By the induction hypothesis, the functions f_k are step functions relative to the intervals I_j and satisfy

$$g \prec f_{j-1} \prec \cdots \prec f_1 \prec f_0 \prec f.$$

Let n_j be the largest integer such that

$$g = f_{j-1} \qquad \text{on } (0, t_{n_j}) \tag{7.12}$$

and let n_j' be the largest integer such that f_{j-1} is constant on $E_j = (t_{n_j}, t_{n_j'})$. If $n_j' = n$, then there exists a nonnegative decreasing function ω such that $\omega f_{j-1} = g$. In this case, (7.11) holds with $T_j h = \omega h$ and $T_{j+1}, T_{j+2}, \ldots, T_n$ equal to the identity operator. Similarly, if $n_j = n$, then $T_{j-1} \circ \cdots \circ T_0 f = g$ and again (7.11) holds if we set the remaining operators equal to the identity. In the remaining case where both $n_j < n$ and $n_j' < n$, we construct T_j with the desired properties and in such a way that at least one of n_j or n_j' increases by at least one. By induction, the proof will then be complete since there are at most n steps.

Let $E'_j = (t_{n'_j}, t_{n'_j+1})$ and set $K = |E_j|/|E'_j|$. Let

$$p = \min\left\{\frac{1}{K+1}, \frac{a-b}{a-a'}\right\}, \tag{7.13}$$

where a is the value of f_{j-1} on E_j, a' is the value of f_{j-1} on E'_j, and b is the largest value of g on E_j, that is, the value of g on (t_{n_j}, t_{n_j+1}). Note that $a' < a$ because of the way n'_j was defined. Since $g \prec f_{j-1}$ and (7.12) holds, we also have $b < a$. It follows that $0 < p < 1$ and $0 < Kp < 1$.

Define the operator T_j by

$$T_j h = \begin{cases} h, & (x \notin E_j \cup E'_j), \\ (1-p)h_{E_j} + p h_{E'_j}, & (x \in E_j), \\ (1-Kp)h_{E'_j} + Kp h_{E_j}, & (x \in E'_j). \end{cases} \tag{7.14}$$

Note that T_j changes the values of h only on the sets E_j and E'_j. Since $p \leq 1/(K+1)$, a computation shows that $T_j h$ is decreasing whenever h is. It is obvious that T_j is linear, positive, and satisfies $|T_j h| \leq T_j|h|$. Furthermore, simple calculations show that T_j is a contraction on L^1 and L^∞ so that T_j is substochastic. Hence, by Proposition 2.4, the function f_j defined by $f_j = T_j f_{j-1}$ satisfies $f_j \prec f_{j-1}$. The condition $p \leq (a-b)/(a-a')$, which follows from (7.13), guarantees that $g \prec f_j$.

Finally, if in (7.13) we have $p = (a-b)/(a-a')$, then $g = f_j$ on $(0, t_{n_j+1})$ so that $n_{j+1} \geq n_j + 1$. If, on the other hand, we have $p = 1/(K+1)$ in (7.13), then f_j is constant on $E_j \cup E'_j = E_{j+1}$ so $n'_{j+1} = n'_j + 1$. In either case, at least one of n_j or n'_j increases. With this the proof is complete. ■

Remark 7.6. Suppose g and f are nonnegative decreasing functions in $(L^1 + L^\infty)(\mathbf{R}^+)$ and $g \prec f$. Then the Calderón-Ryff theorem (Theorem 2.10) shows that $Tf = g$ a.e. for some substochastic operator T. The proof proceeds by a limiting argument from the finite-dimensional result of Hardy-Littlewood-Pólya (Theorem 2.7). By starting instead with the preceding lemma, the property that Th decrease whenever h does can be carried along through the limiting argument, so that the resulting operator T in the Calderón-Ryff theorem inherits the property that Th is decreasing whenever h is.

Our basic goal here is to obtain a decomposition with respect to the Hardy-Littlewood-Pólya relation as follows. If $g \prec f_1 + f_2$, we wish to find a representation $g = g_1 + g_2$ of g that satisfies $g_k \prec f_k$ for $k = 1, 2$. In fact, we shall derive the following slightly more general result.

Theorem 7.7. *Suppose f_1 and f_2 are nonnegative decreasing functions in $(L^1 + L^\infty)(\mathbf{R}^+)$ and that g is a function of the form*

$$g = \sum_{j=1}^{\infty} b_j \chi_{(0,t_j)}, \qquad (t_j > 0, b_j > 0, j = 1, 2, \ldots).$$

If C, C_1, C_2 are nonnegative constants with $C = C_1 + C_2$ and

$$\int_0^t g(s)\,ds \leq C + \int_0^t [f_1(s) + f_2(s)]\,ds, \qquad (t > 0), \tag{7.15}$$

then, for $k = 1, 2$, there exist numbers $\theta_k(j)$ in $[0, 1]$ with

$$\theta_1(j) + \theta_2(j) = 1, \qquad (j = 1, 2, \ldots) \tag{7.16}$$

and corresponding nonnegative decreasing functions

$$g_k = \sum_{j=1}^{\infty} \theta_k(j) b_j \chi_{(0,t_j)}, \tag{7.17}$$

which satisfy $g_1 + g_2 = g$ and

$$\int_0^t g_k(s)\,ds \leq C_k + \int_0^t f_k(s)\,ds, \qquad (t > 0, k = 1, 2). \tag{7.18}$$

Proof. Suppose first that $C = 0$. Let $f = f_1 + f_2$ and, for each $N = 1, 2, \ldots,$ let

$$g^{(N)} = \sum_{j=1}^{N} b_j \chi_{(0,t_j)}. \tag{7.19}$$

Then, by (7.15), we have $g^{(N)} \prec g \prec f$ and so, according to Lemma 7.5, there is a substochastic operator S_N (with the additional property that $S_N h$ decreases whenever h does) such that $S_N f = g^{(N)}$. Denote by A_0 the averaging operator

$$A_0 h = \sum_{j=1}^{N} h_{I_j} \chi_{I_j},$$

where I_j are the intervals of constancy of $g^{(N)}$. Then $T_N = A_0 \circ S_N$ is a substochastic operator (with the property that $T_N h$ decreases whenever h does) such that $T_N h$ is constant on each of the intervals I_j, $(j = 1, 2, \ldots, N)$. Consequently, the step-functions $g_k^{(N)} = T_N f_k$, $(k = 1, 2)$, satisfy

$$g_k^{(N)} \prec f_k, \qquad (k = 1, 2), \qquad g_1^{(N)} + g_2^{(N)} = g^{(N)} \tag{7.20}$$

and may be expressed in the form

$$g_k^{(N)} = \sum_{j=1}^{N} \theta_k^{(N)}(j) b_j \chi_{(0,t_j)},$$

where $\theta_1^{(N)}(j) + \theta_2^{(N)}(j) = 1$. Now letting N vary, we use a standard diagonalization argument to obtain a subsequence $\theta_k^{(N_m)}(j)$ which, for each $j = 1, 2, \dots$ and each $k = 1, 2$, converges to a limit $\theta_k(j)$, say, as $m \to \infty$. Defining g_k by (7.17), we note that $g_k^{(N_m)} \to g_k$. Hence, from (7.20) and the dominated convergence theorem, we see that (7.18) holds (with $C_1 = C_2 = 0$).

The proof is similar in the case where $C \neq 0$. With $g^{(N)}$ defined again by (7.19), we claim that the estimate

$$\int_0^t g^{(N)}(s)\,ds \le C \cdot \min\left(\frac{t}{K}, 1\right) + \int_0^t f(s)\,ds \tag{7.21}$$

holds for all $t > 0$, where $K = \inf\{t_j : j = 1, 2, \dots, N\} > 0$. Indeed, (7.21) holds for all $t \ge K$ by virtue of the hypothesis (7.15) so we need only verify (7.21) on the interval $[0, K]$. As we just remarked, it is true for $t = K$ and it is trivially true for $t = 0$. But $g^{(N)}$ is constant on $[0, K]$ so the left-hand side of (7.21) is linear there. The right-hand side is concave. Since (7.21) is true at the endpoints, it is therefore true in all of $[0, K]$, and this establishes the claim made above. Now observe that the minimum in (7.21) is the integral of a characteristic function so that (7.21) may be rewritten in the form

$$g^{(N)} \prec \left[\frac{C_1}{K}\chi_{(0,K)} + f_1\right] + \left[\frac{C_2}{K}\chi_{(0,K)} + f_2\right].$$

Since this is an estimate of the form (7.15) with $C = 0$, we may apply the result established in the first part of the proof. From this, the desired result (7.18) follows immediately. ■

EXERCISES AND FURTHER RESULTS FOR CHAPTER 3

1. A compatible couple (X_0, X_1) with the property that $X_0 \cap X_1$ is dense in X_0 and in X_1 is called a *conjugate couple*. For such couples, the dual spaces X_0^* and X_1^* also form a compatible couple (X_0^*, X_1^*). (HINT: Both X_0^* and X_1^* embed continuously in $(X_0 \cap X_1)^*$).

2. Let (X_0, X_1) be a conjugate couple of Banach spaces. Then $(X_0 + X_1)^*$ is isometrically isomorphic to $X_0^* \cap X_1^*$. (HINT: Bounded linear functionals on X_0, X_1, and $X_0 + X_1$ may be identified with their restrictions to $X_0 \cap X_1$).

3. Let (X_0, X_1) be a conjugate couple of *complex* Banach spaces. Then $(X_0 \cap X_1)^*$ is isometrically isomorphic to $X_0^* + X_1^*$. (HINT: Embed $X_0 \cap X_1$ as a closed subspace of the direct sum $X_0 \oplus X_1$ and use the Hahn-Banach theorem).

4. Let X_0 and X_1 be an arbitrary pair of Banach function spaces over the same totally σ-finite measure space. Then $X_0 \cap X_1$ is a Banach function space.

5. Let X_0 and X_1 be complex Banach function spaces over the same totally σ-finite measure space. If at least one of X_0 and X_1 has absolutely continuous norm, then the associate space $(X_0 \cap X_1)'$ coincides with $X_0' + X_1'$, with identical norms. In particular, $X_0' + X_1'$ is a Banach function space. (HINT: Exercise 3 and Corollary I.4.3).

6. If $1 \le p_0, p_1 \le \infty$, the complex spaces $L^{p_0} \cap L^{p_1}$ and $L^{p_0} + L^{p_1}$ are rearrangement-invariant Banach function spaces and

$$(L^{p_0} \cap L^{p_1})' = L^{p_0'} + L^{p_1'}; \qquad (L^{p_0} + L^{p_1})' = L^{p_0'} \cap L^{p_1'},$$

isometrically.

7. Suppose $1 \le p_0 \le p \le p_1 \le \infty$, so $1/p = (1-\theta)/p_0 + \theta/p_1$ for some $0 \le \theta \le 1$. Then the complex Lebesgue spaces satisfy

$$\|f\|_p \le \|f\|_{p_0}^{1-\theta} \|f\|_{p_1}^{\theta}, \qquad (f \in L^{p_0} \cap L^{p_1})$$

and

$$\inf_{f = f_0 + f_1} (\|f_0\|_{p_0} + t\|f_1\|_{p_1}) \le t^{\theta} \|f\|_p, \qquad (f \in L^p).$$

8. Let $P_y(x) = y/[\pi(x^2 + y^2)]$ and $\phi_\lambda(x) = (1/2\lambda)\chi_{(-\lambda,\lambda)}(x)$ denote the *Poisson kernel* and the *box kernel*, respectively, for the upper half plane $\mathbf{U}$, where $y, \lambda > 0$ and $-\infty < x < \infty$. The Poisson kernel may be expressed as a "convex" combination of the functions in the box kernel by means of the formula

$$P_y(x) = \int_0^\infty \phi_\lambda(x)\,d\nu_y(\lambda), \qquad (-\infty < x < \infty),$$

where $d\nu_y(\lambda) = -2\lambda P_y'(\lambda)\,d\lambda$ is a positive measure on $\mathbf{R}^+$ of mass 1. If f is a locally integrable function on $\mathbf{R}$, the *centered Hardy-Littlewood maximal function* $M_c f$ of f:

$$M_c f(x) = \sup_{\lambda > 0} \frac{1}{2\lambda} \int_{x-\lambda}^{x+\lambda} |f(t)|\,dt = \sup_{\lambda > 0} |\check{f}| * \phi_\lambda,$$

(where $\check{f}(x) = f(-x)$) and the *radial maximal function* Rf of f:

$$Rf(x) = \sup_{y > 0} \left| \int_{-\infty}^{\infty} f(x-t) P_y(t)\,dt \right| = \sup_{y > 0} |f * P_y(x)|,$$

satisfy $Rf(x) \le M_c f(x)$ a.e. on $\mathbf{R}$.

9. (a) For each $y > 0$, P_y is a nonnegative, symmetrically decreasing function on $\mathbf{R}$ with $\int_{\mathbf{R}} P_y(x)\,dx = 1$; moreover, $h(x, y) = P_y(x)$ is a harmonic function on $\mathbf{U}$.
(b) If $1 \le p \le \infty$ and $f \in L^p(\mathbf{R})$, then the *Poisson integral* $u(x, y) = (f * P_y)(x)$ is harmonic in $\mathbf{U}$.
(c) The *nontangential maximal operator* of f defined by

$$Nf(x) = \sup_{(t,y) \in \Gamma_x} |u(t, y)|, \qquad \Gamma_x = \{(t, y) : |x - t| \le y\},$$

satisfies $Nf \le cMf$ a.e., where M is the Hardy-Littlewood maximal operator.

(d) If $1 \le p \le \infty$ and $f \in L^p(\mathbf{R})$, then

$$\lim_{\substack{y\to 0+\\ (t,y)\in\Gamma_x}} u(t,y) = f(x) \qquad \text{a.e.}$$

and if $p < \infty$, then (cf. also Exercise 16)

$$\lim_{y\to 0+} \|u(\cdot,y) - f\|_p = 0.$$

10. If $E \subset \mathbf{R}$ has finite Lebesgue measure, the identity

$$(H\chi_E)^*(t) = \frac{1}{\pi}\sinh^{-1}\left(\frac{2|E|}{t}\right)$$

may be established as follows: If $E_n \Delta E \to \varnothing$ and each E_n is a finite disjoint union of intervals, then by Lemma 4.5,

$$(H\chi_E)^{**}(t) = \lim_{n\to\infty}(H\chi_{E_n})^{**}(t) = \frac{1}{t}\int_0^t \frac{1}{\pi}\sinh^{-1}\left(\frac{2|E|}{s}\right)ds.$$

The desired result follows by differentiation.

11. If $S(f^*)(1) < \infty$, where S is the Calderón operator given by (4.22), the *conjugate Poisson kernel* $Q_y(x) = x/[\pi(x^2+y^2)]$, $(y > 0)$, generates a harmonic function $v(x,y) = (f * Q_y)(x)$ in $\mathbf{U}$. Since

$$|v(x,y) - H_y f(x)| \le (|f| * P_y)(x) + \frac{2}{\pi}Mf(x)$$

the maximal operator $\tilde{N}f(x) = \sup_{(t,y)\in\Gamma_x}|v(t,y)|$ satisfies $\tilde{N}f \le \mathscr{H}f + 3Mf$ a.e., and so there exists $c > 0$ such that

$$(\tilde{N}f)^*(t) \le cS(f^*)(t), \qquad (t > 0).$$

It follows that $\lim_{\substack{(t,y)\in\Gamma_x\\ y\to 0+}} v(t,y) = Hf(x)$ a.e. Furthermore, if $f \in L^p(\mathbf{R})$ for some $1 < p < \infty$, then $v(x,y) = (Hf * P_y)(x)$. Hence, the function $F(x,y) = u(x,y) + iv(x,y)$ is analytic in $\mathbf{U}$ with (nontangential) boundary values $f(x) + iHf(x)$.

12. (a) The Poisson kernel P_r for the unit circle $\mathbf{T}$

$$P_r(\theta) = \frac{1-r^2}{1-2r\cos\theta + r^2} = \mathrm{Re}\left\{\frac{1+re^{i\theta}}{1-re^{i\theta}}\right\}, \qquad (0 \le r < 1)$$

is a nonnegative symmetric decreasing function with $\int_{\mathbf{T}} P_r\,d\theta = 1$. It has the equivalent representation

$$P_r(\theta) = 1 + 2\sum_{n=1}^{\infty} r^n\cos(n\theta), \qquad \text{so } \hat{P}_r(n) = r^{|n|}, \qquad (n \in \mathbf{Z})$$

and

$$(f * P_r)(e^{i\theta}) = \sum_{n=-\infty}^{\infty} r^{|n|}\hat{f}(n)e^{in\theta}, \qquad (e^{i\theta} \in \mathbf{T}).$$

In particular, $u(re^{i\theta}) = (f * P_r)(e^{i\theta})$ is harmonic in $\mathbf{D}$.

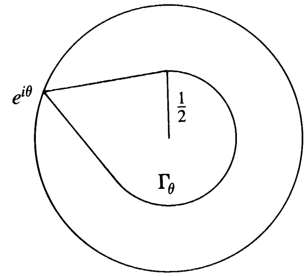
(b) The *Hardy-Littlewood maximal operator* M is defined on $L^1(\mathbf{T})$ by $Mf(e^{i\theta}) = \sup(1/m(I))\int_I |f|\,dm$, where m is normalized Lebesgue measure on $\mathbf{T}$ and the supremum extends over all intervals (subarcs) of $\mathbf{T}$ containing $e^{i\theta}$. As in Theorem 3.8, there are constants c and c' such that $c(Mf)^* \le f^{**} \le c'(Mf)^*$ on the interval $(0,1)$.

(c) Let Γ_θ denote the *Stoltz domain* shown in the figure below. The *nontangential maximal function* Nf of an integrable function on $\mathbf{T}$ is defined by

$$Nf(e^{i\theta}) = \sup_{re^{i\phi}\in\Gamma_\theta} |(f * P_r)(e^{i\phi})|.$$

The *nontangential maximal operator* N satisfies $Nf \le Mf$ a.e.. Furthermore, for each $f \in L^1(\mathbf{T})$,

$$\lim_{\substack{re^{i\phi}\to e^{i\theta}\\ re^{i\phi}\in\Gamma_\theta}} (f * P_r)(e^{i\phi}) = f(e^{i\theta}) \qquad \text{a.e.}$$



(d) The *conjugate Poisson kernel* Q_r is given by

$$Q_r(\theta) = \frac{2r\sin\theta}{1-2r\cos\theta + r^2} = \mathrm{Im}\left\{\frac{1+re^{i\theta}}{1-re^{i\theta}}\right\}, \qquad (0 \le r < 1).$$

If $f \in L^p(\mathbf{T})$, $(1 < p < \infty)$, then

$$\lim_{\substack{r\to 1-\\ re^{i\phi}\in\Gamma_\theta}} v(re^{i\phi}) = \tilde{f}(e^{i\theta}) \qquad \text{a.e.}$$

Hence, $v(re^{i\theta}) = (f * Q_r)(e^{i\theta})$ is harmonic in $\mathbf{D}$ and coincides with $(\tilde{f} * P_r)(e^{i\theta})$. Furthermore, if $u = f * P_r$, then $F = u + iv$ is analytic in $\mathbf{D}$ and $\hat{F}(n) = 0$ for all $n < 0$ (HINT: Exercise 11).

13. If $1 \le p \le q \le \infty$, the lower and upper Boyd indices of $(L^p \cap L^q)(\mathbf{R})$, and of $(L^p + L^q)(\mathbf{R})$, are equal to $1/q$ and $1/p$, respectively.

14. Let X be a rearrangement-invariant Banach function space on $\mathbf{R}$ with fundamental function φ_X. Let M_X be the submultiplicative function $M_X(t) = \sup_s\{\varphi_X(st)/\varphi_X(s)\}$, $(t > 0)$, and define the *lower and upper fundamental indices* of X by

$$\underline{\beta}_X = \sup_{0<t<1}\frac{\log M_X(t)}{\log t}, \qquad \bar{\beta}_X = \inf_{t>1}\frac{\log M_X(t)}{\log t}.$$

Then:

(a) $\underline{\beta}_X = \lim_{t\to 0+} \dfrac{\log M_X(t)}{\log t}, \qquad \bar{\beta}_X = \lim_{t\to\infty} \dfrac{\log M_X(t)}{\log t}$;
(b) $M_X(t) \le h_X(t), \qquad (t > 0)$;
(c) $\underline{\beta}_{X'} = 1 - \bar{\beta}_X, \qquad \bar{\beta}_{X'} = 1 - \underline{\beta}_X$;
(d) $0 \le \underline{\alpha}_X \le \underline{\beta}_X \le \bar{\beta}_X \le \bar{\alpha}_X \le 1$;
(e) X, $\Lambda(X)$, and $M(X)$ have identical fundamental indices;
(f) There is a rearrangement-invariant space X for which $\varphi_X(t) = t^{1/2}$ (and so $\underline{\beta}_X = \bar{\beta}_X = \frac{1}{2}$) but $\underline{\alpha}_X = 0$ and $\bar{\alpha}_X = 1$ (T. Shimogaki [2]).

15. (D. W. Boyd [5]) Let D denote the collection of decreasing sequences on the set $\mathbf{N} = \{1, 2, \ldots\}$. Let X be a rearrangement-invariant Banach function space on the integers $\mathbf{Z}$ and define

$$g_X(m) = \sup\{\|E_m f\|_X : f \in D, \|f\|_X \le 1\}, \qquad (m \in \mathbf{N}),$$

where E_m is the dilation operator: $E_m f(n) = f(mn)$, $(n \in \mathbf{Z})$. Then g_X is decreasing, submultiplicative, and satisfies $0 < g_X(m) \le 1$. The *upper Boyd index* of X is defined by

$$\bar{\alpha}_X \equiv \sup_{m\in\mathbf{N}} -\frac{\log g_X(m)}{\log m} = \lim_{m\to\infty} -\frac{\log g_X(m)}{\log m},$$

and the *lower Boyd index* by $\underline{\alpha}_X = 1 - \bar{\alpha}_{X'}$. The latter index may also be defined as follows. Let F_m, $(m \in \mathbf{N})$, be the dilation operator given by $F_m f(n) = f([(n-1)/m] + 1)$, $(n \in \mathbf{Z})$, where $[k]$ denotes the integer part of k. If

$$k_X(m) = \sup\{\|F_m f\|_X : \|f\|_X \le 1\} \qquad (m \in \mathbf{N}),$$

then $mh_{X'}(m) = k_X(m)$, $(m \in \mathbf{N})$, and

$$\underline{\alpha}_X \equiv \inf_{m\in\mathbf{N}} \frac{\log k_X(m)}{\log m} = \lim_{m\to\infty} \frac{\log k_X(m)}{\log m}.$$

16. Let (k_n) be a sequence of bounded measurable functions on $\mathbf{T} \times \mathbf{T}$ and set $\kappa_n(e^{i\theta}) = \operatorname{ess\,sup}_\phi |k_n(e^{i(\theta+\phi)}, e^{i\phi})|$. Then (k_n) is said to be an *approximate identity* if

(i) $\sup_n \int_{\mathbf{T}} \kappa_n(e^{i\theta})\, d\theta < \infty$;
(ii) $\int_{\mathbf{T}} k_n(e^{i\theta}, e^{i\phi})\, d\theta = 1, \qquad$ for all n and ϕ;
(iii) for each $\varepsilon > 0$ and $\delta > 0$, there exists N such that $\int_{\delta < |\theta| < \pi} \kappa_n(e^{i\theta})\, d\theta < \varepsilon$, whenever $n \ge N$.

Let X be a rearrangement-invariant space for which $\varphi_X(0+) = 0$. If $\mathscr{K}_n f(e^{i\phi}) = \int_{\mathbf{T}} k_n(e^{i\theta}, e^{i\phi}) f(e^{i\theta})\, d\theta$, then for each $f \in X_b$,

$$\lim_{n\to\infty} \|\mathscr{K}_n f - f\|_X = 0.$$

The same is true of families indexed by continuous parameters, such as $k_r(e^{i\theta}, e^{i\phi}) = P_r(\theta - \phi)$, $(0 \le r < 1)$, for example.

17. The trigonometric identity

$$D_N(e^{i\theta}) \equiv \sum_{n=-N}^{N} e^{in\theta} = \frac{\sin(N + \frac{1}{2})\theta}{\sin\frac{1}{2}\theta}$$

gives the relation $S_N f(e^{i\theta}) = (f * D_N)(e^{i\theta})$ for the N-th partial sum of the Fourier series of f. Hence, with the *Féjer kernel* K_N defined by

$$K_N(e^{i\theta}) \equiv \sum_{n=-N}^{N} \left(1 - \frac{|n|}{N+1}\right) e^{in\theta} = \frac{1}{N+1} \left\{\frac{\sin\frac{1}{2}(N+1)\theta}{\sin\frac{1}{2}\theta}\right\}^2,$$

the Féjer sum $\sigma_N f \equiv f * K_N$ is expressed as the *N-th Césaro mean*

$$\sigma_N f = \frac{1}{N+1}(S_0 f + S_1 f + \cdots + S_N f)$$

of the Fourier series of f. The kernel $k_N(e^{i\theta}, e^{i\phi}) = K_N(e^{i(\theta-\phi)})$ is an approximate identity. It follows that the trigonometric polynomials are dense in X_b whenever $\varphi_X(0+) = 0$ (see Exercise 16).

18. (Jackson's theorem for rearrangement-invariant spaces X). For each $N = 1, 2, \ldots$, define the *Jackson kernel* L_N by

$$L_N(e^{i\theta}) = \frac{1}{\lambda_n} \left\{\frac{\sin\frac{1}{2}n\theta}{\sin\frac{1}{2}\theta}\right\}^4,$$

where $n = [N/2] + 1$ and λ_N is chosen so that $\int_{\mathbf{T}} L_N = 1$.
(a) L_N is a nonnegative even trigonometric polynomial of degree $2(n-1)$.
(b) The constants λ_n satisfy $\lambda_n \approx n^3$ and the first moments of the functions L_N satisfy $\int_0^\pi t L_N(t)\, dt \approx 1/N$.
(c) There is a constant M such that, for all $f \in X$,

$$\inf_{\deg(p) \le N} \|f - p\|_X \le M\omega\left(f, \frac{1}{N}\right)_X,$$

that is, the error of approximation of functions in X by trigonometric polynomials is controlled by the modulus of continuity in X (HINT: Let $p = f * L_N$).

19. Suppose $0 \le b < a \le 1$ and let $\sigma = [(a, a), (b, b)]$. Then

$$P_a \circ Q_b = Q_b \circ P_a = \frac{1}{a+b}(P_a + Q_b) = S_\sigma,$$

(cf. Definitions 5.1 and 5.14).

NOTES FOR CHAPTER 3

Much of the material in §1 may be found in the pioneering work of N. Aronszajn & E. Gagliardo [1]. In finite dimensions, Propositions 2.1, 2.4, and Theorem 2.7 were known in some form to R. F. Muirhead [1] and G. H. Hardy, J. E. Littlewood & G. Pólya [1]; see A. W. Marshall & I. Olkin [1, Chapters 1,2]. Theorem 2.10 was established independently by J. V. Ryff [1] (in the case of a finite interval) and A. P. Calderón [3] (for general measure spaces). To each nonnegative measurable function f there corresponds a substochastic operator T (induced by the measure-preserving transformation of Theorem II.7.5) such that $Tf^* = f$ a.e. Although f^* may not in general be obtained from f by composition with a measure-preserving transformation (Example II.7.7), Ryff observes that the adjoint of the operator T nevertheless carries f^* onto f. An alternative proof of Theorem 2.10 could be based on this device, although some kind of limiting argument would again be necessary in order to pass from the functions with $f^*(+\infty) = 0$ for which Ryff's theorem is valid (Corollary II.7.6) to all of $L^1 + L^\infty$. We have preferred to follow Calderón's argument in detail, including a proof of the finite-dimensional case, not only for the historical perspective the latter provides but more importantly because it serves as a model for the related Lorentz-Shimogaki results of §7.

That rearrangement-invariant Banach function spaces X are interpolation spaces for (L^1, L^∞) (Theorem 2.2) was established for Orlicz spaces by W. Orlicz [2] (for linear operators) and [4] (for Lipschitz operators); for arbitrary spaces X and linear integral operators, the result is due to G. G. Lorentz [4]. In full generality, the result was established by A. P. Calderón [3] (for quasilinear operators); G. G. Lorentz & T. Shimogaki [1] established the corresponding result for Lipschitz operators. The complete characterization of the interpolation spaces (Theorem 2.12) is due to A. P. Calderón [3]. A similar result was established independently by B. S. Mitjagin [1], in which the notion of rearrangement is restricted to that resulting from composition with measure-preserving transformations (the converse of Theorem 2.2 is trivial in that case). Interpolation between pairs of Λ-spaces has been investigated by G. G. Lorentz & T. Shimogaki [2] and R. Sharpley [3]; corresponding results for M_ϕ spaces appear in R. Sharpley [4].

G. H. Hardy & J. E. Littlewood, in the classic paper [2] introducing the maximal operator M, established the integral inequalities $\int [\Phi(Mf)]^* \leq c\int \Phi(M(f^*)) \sim c\int \Phi(f^{**})$ for increasing Φ and, via Hardy's inequalities, proved the L^p-boundedness of M for $p > 1$ (Theorem 3.10). F. Riesz [2], used the "rising sun lemma" and previous work of R. M. Gabriel [1] to obtain the stronger pointwise estimate $(Mf)^* \leq cf^{**}$ (cf. Theorem 3.8) from which the weak-type estimate (Theorem 3.3) follows immediately; in n dimensions, the weak-type estimate was established by N. Wiener [1]. The reverse estimate $f^{**} \leq c(Mf)^*$ to Riesz' inequality (cf. Theorem 3.8) was established in one dimension by C. Herz [1], using martingale methods; the covering lemma (Lemma 3.7) and n-dimensional form of Theorem 3.8 were established by C. Bennett & R. Sharpley [1]. That Lebesgue's differentiation theorem follows from the weak-type estimate has a remarkable converse in the theorem of E. M. Stein on limits of sequences of operators (Chapter IV.5). Excellent accounts of differentiation theory may be found in the monographs of M. de Guzman [1] and E. M. Stein [5].

The L^p-boundedness $(1 < p < \infty)$ of the Hilbert transform (Theorem 4.9) was established by M. Riesz [2]. The notion of "weak-type" dates back to A. N. Kolmogorov [1], who obtained such an estimate for the conjugate function. "Real-variable" methods originated with L. H. Loomis [1]. The presentation given in the text is based on that of E. M. Stein & G. Weiss [3], [4], in which the interpolation of operators of restricted weak type is introduced (cf. Chapter IV.5); interestingly, a result closely related to the fundamental Lemma 4.5 was established more than a century earlier by G. Boole [1]. The weak-type inequality (4.27) is due to C. Bennett and K. Rudnick [1]; the integrated form $(Hf)^{**} \leq cS(f^{**})$ was known previously to R. O'Neil & G. Weiss [1] and A. P. Calderón [3]. Proposition 4.10 and similar results may be found in R. Sharpley [7].

The notion of joint weak type, formalized from the Calderón theory by C. Bennett & K. Rudnick [1], allows for a unified treatment of interpolation of weak-type operators, regardless of whether a separate weak-type estimate exists at L^∞. In particular, the characterization by A. P. Calderón [3] of weak interpolation spaces (Theorem 5.7) then reformulates directly (in the diagonal case) to a statement in terms of the Boyd indices (Theorem 5.16) with no exceptional case when the lower index vanishes. The indices were introduced by D. W. Boyd in the series of papers [2], [3], [4], [5]. Results equivalent to Theorem 5.17 on the maximal operator were established earlier, however, by G. G. Lorentz [5] and T. Shimogaki [1]. Theorem 5.18 on the Hilbert transform is due to D. W. Boyd [1]; its variant for the conjugate-function operator was established by F. Féher, D. Gaspar, & H. Johnen [2]; related results for norm-convergence of Fourier series are in [1] (cf. Theorem 6.11). Their techniques are connected to some extent with the characterization by C. Bennett [1] of the Boyd indices of function spaces on the unit circle in terms of "dilation" operators fashioned from Mobiüs transformations. Weak-type interpolation theorems involving more general kinds of endpoint spaces than considered in the Calderón-Boyd theorems have been obtained by S. G. Krein & E. M. Semenov [1] and R. Sharpley [1].

Theorem 7.4, on the decreasing rearrangement of differences, is due to G. G. Lorentz & T. Shimogaki [1]. The remaining results of §7, leading to Theorem 7.7, are based on Proposition 3 of G. G. Lorentz & T. Shimogaki [1], [2], but are reworked to make explicit the linearity of the selection process (cf. C. Bennett & R. Sharpley [3]).

4 The Classical Interpolation Theorems

Having established various interpolation principles of a rather general nature, we now narrow the focus somewhat and explore the ramifications of these principles in specific classes of rearrangement-invariant spaces.

In section 1, we return to the L^p-spaces, and thereby to the beginnings of the subject, with a discussion of the *M. Riesz convexity theorem*. The result is formulated initially in the finite-dimensional setting of matrix transformations and their associated bilinear forms; the passage to more general kinds of measure spaces is achieved by a straightforward approximation argument. Although the subsequent function-theoretic proof of this result due to O. V. Thorin (presented in Section 2) has become standard in more recent texts, we chose to include Riesz' original proof here not only for the historical perspective it provides but also because of a conviction that it contains the germs of ideas that are yet to be fully exploited.

The *Riesz-Thorin theorem* is presented in Section 2, together with a multilinear version of the theorem and a variant which is applicable to compact operators. Further extensions, due to E. M. Stein, are derived in Section 3. Here, the interpolation extends over an "analytic family" of operators T_z that depend in analytic fashion on a complex variable z. One manifestation of this result allows the interpolation to extend over L^p-spaces in which the underlying measures are also allowed to vary.

In Section 4, we return to real-variable methods and to another of the cornerstones of the classical theory: the *Marcinkiewicz interpolation theorem.* Much of the underlying structure needed here has already been developed in Chapter III in our discussion of the Calderón S_σ-operator and the interpolation of operators of joint weak type $(p_0, q_0; p_1, q_1)$ on arbitrary rearrangement-invariant Banach function spaces. Instead of direct specialization to the L^p-spaces, it is better to resort to a more general two-parameter family of spaces—the *Lorentz $L^{p,q}$-spaces*—in terms of which the Marcinkiewicz theorem may be conveniently formulated.

Interpolation of operators of *restricted weak type* is taken up in Section 5, where the basic Stein-Weiss interpolation theorem is established. The discussion includes a comparison of the notions of weak type and restricted weak type, culminating in the theorem of K. H. Moon. Section 5 concludes with *Stein's theorem* on limits of sequences of operators.

In the Marcinkiewicz theorem and its variants discussed thus far, attention has been confined to the interior $0 < \theta < 1$ of the interpolation segment in question. In Section 6, we consider the limiting cases $\theta \to 1$ and $\theta \to 0$ of these results. Although the underlying structure remains the same, the Lorentz $L^{p,q}$-spaces no longer suffice to describe the mapping properties involved. It is not difficult, however, to determine appropriate substitutes. Our discussion begins with the classical *Zygmund spaces* $L\log L$ and $L_{\exp}$. Both have a simple Lorentz-space structure, a fact which was known in some form to Hardy and Littlewood but which seems to have been all but ignored in the subsequent literature. It is, however, precisely this property that links them with the $L^{p,q}$-spaces, and points the way to the consolidation of both families into the single larger family of *Lorentz-Zygmund spaces* $L^{p,q}(\log L)^\alpha$. One obtains in this way a unified treatment of the entire interpolation segment $0 \le \theta \le 1$.

Two further extensions of the weak-type theory are presented in Section 7. The first involves the interpolation of bilinear operators that satisfy n initial weak-type conditions. The second involves a more general notion of "weak type" which makes use of the Lorentz spaces $\Lambda(X)$ and $M(Y)$, where X and Y are arbitrary rearrangement-invariant spaces. This effectively removes the classical L^p-emphasis and paves the way for the development of a natural theory of weak-type interpolation for general rearrangement-invariant spaces.

The final section (Section 8) contains a brief exposition of another important class of rearrangement-invariant spaces: the *Orlicz spaces.* While there is some intersection between this class and the class of Lorentz-type spaces (in the Zygmund spaces, for example), neither fully contains the other. We do not attempt a detailed discussion of the interpolation properties of Orlicz spaces, contenting ourselves instead with the computation of the Boyd indices in terms of the generating Young's function.

1. THE RIESZ CONVEXITY THEOREM

Throughout this section, the L^p-spaces considered may be real or complex. In the complex case, somewhat more precise results can be obtained by using function-theoretic methods due to Thorin, and these will be discussed in the next section. In either case, the following terminology will be useful.

Definition 1.1. Let (R, μ) and (S, ν) be totally σ-finite measure spaces and let T be a linear operator defined on all μ-simple functions on R and taking values in the ν-measurable functions on S (as usual, functions that coincide a.e. on either measure space are identified). Suppose $1 \le p, q \le \infty$. If there is a constant M such that

$$\|Tf\|_{L^q(\nu)} \le M\|f\|_{L^p(\mu)}, \tag{1.1}$$

for all μ-simple functions f on R, then T is said to be of *strong type* (p, q). The least constant M for which (1.1) holds is called the *strong-type* (p, q) *norm* of T (or simply the *norm* of T if no confusion can arise).

The following elementary result exhibits a structure typical of all the L^p-interpolation theorems. We consider here only positive integral operators, that is, operators T of the form

$$(Tf)(y) = \int_R f(x)A(x, y)\,d\mu(x), \qquad (y \in S) \tag{1.2}$$

with A measurable and nonnegative.

Theorem 1.2. *Suppose* $1 \le p_k, q_k \le \infty$, $(k = 0, 1)$, *and* $0 \le \theta \le 1$. Let

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \qquad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}. \tag{1.3}$$

Let T be a positive integral operator and suppose that T is of strong types (p_0, q_0) and (p_1, q_1), with respective strong-type norms M_0 and M_1. Then T is of strong type (p, q) and its strong-type norm M_θ satisfies

$$M_\theta \le M_0^{1-\theta} M_1^\theta. \tag{1.4}$$

Proof. In order to establish (1.4), it will suffice to show that

$$\left|\int_S (Tf)g\,dv\right| \le M_0^{1-\theta}M_1^{\theta}, \tag{1.5}$$

for all simple functions f and g satisfying

$$\|f\|_{L^p(\mu)} \le 1, \qquad \|g\|_{L^{q'}(v)} \le 1 \tag{1.6}$$

(here, as throughout, q' denotes the complementary exponent to $q: 1/q + 1/q' = 1$). Indeed, the supremum of the left-hand side of (1.5) over all such g is the L^q-norm of Tf, and then the supremum of that quantity over all such f is the strong-type (p, q) norm M_θ of T.

Since T is positive, it follows from (1.2) that

$$\left|\int_S (Tf)g\,dv\right| \le \int_S T(|f|)|g|\,dv,$$

so the problem reduces to establishing (1.5) merely for nonnegative simple functions f and g satisfying (1.6). In that case, however, the functions

$$f_k = f^{p/p_k}, \qquad g_k = g^{q'/q_k'}, \qquad (k = 0, 1)$$

are well-defined and satisfy

$$\|f_k\|_{p_k} = \|f\|_p^{p/p_k} \le 1, \qquad \|g_k\|_{q_k'} = \|g\|_{q'}^{q'/q_k'} \le 1. \tag{1.7}$$

Using (1.3) and applying Hölder's inequality, we have

$$\begin{aligned}
\int_S (Tf)g\,dv &= \int_S\int_R (Af_0g_0)^{1-\theta}(Af_1g_1)^{\theta}\,d\mu\,dv \\
&\le \left(\int_S\int_R Af_0g_0\,d\mu\,dv\right)^{1-\theta}\left(\int_S\int_R Af_1g_1\,d\mu\,dv\right)^{\theta} \\
&= \left(\int_S (Tf_0)g_0\,dv\right)^{1-\theta}\left(\int_S (Tf_1)g_1\,dv\right)^{\theta}.
\end{aligned}$$

Hence, by Hölder's inequality and the strong-type hypotheses on T,

$$\begin{aligned}
\int_S (Tf)g\,dv &\le (\|Tf_0\|_{q_0}\|g_0\|_{q_0'})^{1-\theta}(\|Tf_1\|_{q_1}\|g_1\|_{q_1'})^{\theta} \\
&\le (M_0\|f_0\|_{p_0}\|g_0\|_{q_0'})^{1-\theta}(M_1\|f_1\|_{p_1}\|g_1\|_{q_1'})^{\theta}.
\end{aligned}$$

An appeal to (1.7) now shows that the right-hand side does not exceed $M_0^{1-\theta}M_1^{\theta}$, and this completes the proof. ■

It is easy to see that the preceding result does not hold for arbitrary linear

operators. Here is a counterexample for *real* L^p-spaces in two dimensions (see Exercise 2 for the complex case).

Example 1.3. Let A be the matrix

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

so that the operator T in (1.2) is the matrix transformation

$$T:(x_1, x_2) \to (x_1 + x_2, x_1 - x_2), \qquad ((x_1, x_2) \in \mathbf{R}^2).$$

Let $p_0 = \infty$, $q_0 = 1$, and $p_1 = q_1 = 2$. Elementary computations show that the corresponding strong-type norms are $M_0 = 2$ and $M_1 = 2^{1/2}$.

Now choose $\theta = 1/2$ so that (1.3) gives $p = 4$ and $q = 4/3$. If (1.4) were to hold, we should have

$$\|(x_1 + x_2, x_1 - x_2)\|_{4/3} \le 2^{3/4}\|(x_1, x_2)\|_4, \tag{1.8}$$

for all real x_1 and x_2. However, calculation shows this to be false with $x_1 = 1$ and $x_2 = 2$ for example (cf. also Exercise 1).

It is clear already from the preceding results that the interpolation structure involves the *reciprocals* of the exponents p, q, etc., rather than the exponents themselves. Accordingly, and we shall use this convention throughout, we set

$$\alpha_k = \frac{1}{p_k}, \qquad \beta_k = \frac{1}{q_k}, \qquad (k = 0, 1), \qquad \alpha = \frac{1}{p}, \qquad \beta = \frac{1}{q}. \tag{1.9}$$

With this notation, the hypotheses of Theorem 1.2 exhibit (α, β) as an intermediate point of the straight line segment whose endpoints are (α_0, β_0) and (α_1, β_1). With the strong-type (p, q) norm M regarded as a function of the point (α, β), the theorem asserts that $\log M(\alpha, \beta)$ defines a convex surface over the unit square

$$\square = \{(\alpha, \beta): 0 \le \alpha, \beta \le 1\}.$$

The more usual terminology is that $M(\alpha, \beta)$ is *logarithmically convex* in the unit square $\square$. By Theorem 1.2, this is true of all positive integral operators (and, in fact, all positive operators; cf. Exercise 5). Example 1.3, on the other hand, shows that the logarithmic convexity can fail if the operator is not positive. Notice, however, that the line segment in Example 1.3 extends outside the *lower triangle*

$$\triangle = \{(\alpha, \beta): 0 \le \beta \le \alpha \le 1\}$$

of the unit square $\square$. The remarkable theorem of M. Riesz asserts that for *every* linear operator the logarithmic convexity persists everywhere inside the lower triangle $\triangle$.

We begin with the finite-dimensional case. Let $A = (a_{ij})_{i,j=1}^{m;n}$ denote an $m \times n$ matrix with real or complex coefficients a_{ij}. Let $\mathbf{x} = (x_1, x_2, \ldots, x_m)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ denote arbitrary m- and n-dimensional real or complex vectors. We denote also by A the bilinear form corresponding to the matrix A, namely,

$$A(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} x_i y_j. \tag{1.10}$$

If, in addition, we write

$$X_j = \sum_{i=1}^{m} a_{ij} x_i, \qquad Y_i = \sum_{j=1}^{n} a_{ij} y_j, \tag{1.11}$$

for $j = 1, 2, \ldots, n$ and $i = 1, 2, \ldots, m$, respectively, then clearly

$$A(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{n} X_j y_j = \sum_{i=1}^{m} x_i Y_i. \tag{1.12}$$

Suppose $1 \leq p, q \leq \infty$ so that the point (α, β), with α and β as in (1.9), lies in the unit square $\square$. The maximum $M(\alpha, \beta)$ of the bilinear form A is defined by

$$M(\alpha, \beta) = \max\{|A(\mathbf{x}, \mathbf{y})| : \|\mathbf{x}\|_p \leq 1, \|\mathbf{y}\|_{q'} \leq 1\}. \tag{1.13}$$

Consequently, $M(\alpha, \beta)$ is the least constant for which

$$|A(\mathbf{x}, \mathbf{y})| \leq M(\alpha, \beta) \|\mathbf{x}\|_p \|\mathbf{y}\|_{q'} \tag{1.14}$$

holds for all $\mathbf{x}$ and $\mathbf{y}$. The estimates

$$|A(\mathbf{x}, \mathbf{y})| \leq \|\mathbf{X}\|_q \|\mathbf{y}\|_{q'} \tag{1.15}$$

and

$$|A(\mathbf{x}, \mathbf{y})| \leq \|\mathbf{x}\|_p \|\mathbf{Y}\|_{p'} \tag{1.16}$$

follow directly from (1.12) and Hölder's inequality.

Lemma 1.4. *The maximum $M = M(\alpha, \beta)$ of the bilinear form A is given by*

$$M = \max_{\|\mathbf{x}\|_p \leq 1} \|\mathbf{X}\|_q = \max_{\|\mathbf{y}\|_{q'} \leq 1} \|\mathbf{Y}\|_{p'}. \tag{1.17}$$

Furthermore, if $\mathbf{x}$ and $\mathbf{y}$ satisfy

$$|A(\mathbf{x}, \mathbf{y})| = M, \qquad (\|\mathbf{x}\|_p \leq 1, \|\mathbf{y}\|_{q'} \leq 1), \tag{1.18}$$

then

$$|X_j| = M |y_j|^{q'-1}, \qquad |Y_i| = M |x_i|^{p-1}, \tag{1.19}$$

for $j = 1, 2, \ldots, n$ and $i = 1, 2, \ldots, m$ respectively.

Proof. Let $N = \max\{\|\mathbf{X}\|_q : \|\mathbf{x}\|_p \leq 1\}$. By applying (1.15) to any pair of vectors $\mathbf{x}$, $\mathbf{y}$ satisfying (1.18), we see at once that $M \leq N$. Conversely, given any $\mathbf{x}$ with $\|\mathbf{x}\|_p \leq 1$ (then $\mathbf{X}$ is determined also), we can construct $\mathbf{y}$ (which can be normalized so that $\|\mathbf{y}\|_{q'} = 1$) so that equality occurs in the Hölder inequality (1.15), that is, $|A(\mathbf{x}, \mathbf{y})| = \|\mathbf{X}\|_q$. Applying (1.14), we obtain $\|\mathbf{X}\|_q \leq M$. This holds for all $\mathbf{x}$ in the unit ball of ℓ^p so we deduce that $N \leq M$. Hence, $N = M$ and the first of the identities in (1.17) is established. The proof of the second is similar.

Suppose now that (1.18) holds. Since $\mathbf{x}$ and $\mathbf{y}$ maximize A, they in fact lie on the surfaces of their respective unit balls. From (1.17), it follows that equality occurs in each of the Hölder inequalities (1.15) and (1.16). Such equality can occur, however, only when

$$|X_j|^q = c |y_j|^{q'}, \qquad (j = 1, 2, \ldots, n)$$

in the case of (1.15), and

$$|x_i|^p = d |Y_i|^{p'}, \qquad (i = 1, 2, \ldots, m)$$

in the case of (1.16), where c and d are constants independent of i and j. Substituting these values into the equalities (1.15) and (1.16), and using the fact that $\mathbf{x}$ and $\mathbf{y}$ have norm 1, we obtain the values $c = M^q$ and $d = M^{-p'}$, from which the desired identities (1.19) result. ■

Theorem 1.5 (M. Riesz). *The maximum $M(\alpha, \beta)$ of an arbitrary bilinear form is logarithmically convex in the lower triangle $\triangle$ of the unit square.*

Proof. Observe first that to establish convexity of a continuous function f on a closed interval, it is enough to show that in every subinterval $[c, d]$ there exists *one* value of θ, with $0 < \theta < 1$, such that

$$f((1-\theta)c + \theta d) \leq (1-\theta) f(c) + \theta f(d).$$

In the case at hand, let (α_0, β_0) and (α_1, β_1) be arbitrary points in $\triangle$ and let (α, β) lie on the line segment joining them so that

$$\alpha = (1-\theta)\alpha_0 + \theta\alpha_1, \qquad \beta = (1-\theta)\beta_0 + \theta\beta_1. \tag{1.20}$$

We shall employ the notation (1.9) and abbreviate $M(\alpha_k, \beta_k)$ by M_k,

$(k = 0, 1)$, and $M(\alpha, \beta)$ by M. The desired logarithmic convexity resides in the estimate

$$M \le M_0^{1-\theta} M_1^{\theta} \tag{1.21}$$

which, in view of the remark made above, need only be established for one value of θ satisfying $0 < \theta < 1$. The appropriate choice of θ will be made later.

Let us proceed under the assumption that all the parameters involved are finite. Only routine modifications to the argument are necessary to accommodate the exceptional cases. We begin by choosing $\mathbf{x}$ and $\mathbf{y}$ with

$$\|\mathbf{x}\|_p = 1, \qquad \|\mathbf{y}\|_{q'} = 1 \tag{1.22}$$

that maximize the form $A: |A(\mathbf{x}, \mathbf{y})| = M$. Then, by Lemma 1.4, we have

$$M\Big(\sum_i |x_i|^{(p-1)p_0'}\Big)^{1/p_0'} = \Big(\sum_i |Y_i|^{p_0'}\Big)^{1/p_0'} \le M_0\Big(\sum_j |y_j|^{q_0'}\Big)^{1/q_0'}$$

and

$$M\Big(\sum_j |y_j|^{(q'-1)q_1}\Big)^{1/q_1} = \Big(\sum_j |X_j|^{q_1}\Big)^{1/q_1} \le M_1\Big(\sum_i |x_i|^{p_1}\Big)^{1/p_1}.$$

Raising the first of these estimates to the power $1 - \theta$, the second to the power θ, and multiplying, we obtain

$$M\Big(\sum_i |x_i|^{(p-1)p_0'}\Big)^{(1-\theta)/p_0'}\Big(\sum_j |y_j|^{(q'-1)q_1}\Big)^{\theta/q_1}$$
$$\le M_0^{1-\theta} M_1^{\theta}\Big(\sum_i |x_i|^{p_1}\Big)^{\theta/p_1}\Big(\sum_j |y_j|^{q_0'}\Big)^{(1-\theta)/q_0'}. \tag{1.23}$$

To obtain (1.21) from (1.23), we shall need to eliminate the terms in x_i and y_j. For the sum on the right involving x_i, for example, the strategy will be to apply Hölder's inequality to obtain two new sums in $|x_i|$, one to the power $(p-1)p_0'$ (to cancel with the sum on the left of (1.23)) and one to the power p (which can then be estimated by (1.22)). The terms in y_j will be treated similarly. Hence, it will be necessary to express p_1 and q_0' in the form

$$p_1 = \lambda(p-1)p_0' + (1-\lambda)p, \qquad q_0' = \mu(q'-1)q_1 + (1-\mu)q' \tag{1.24}$$

with

$$0 \le \lambda \le 1, \qquad 0 \le \mu \le 1. \tag{1.25}$$

Solving for λ and μ, we obtain

$$\lambda = \left[\frac{1-\theta}{\theta}\right]\frac{p_1}{p_0'}, \qquad \mu = \left[\frac{\theta}{1-\theta}\right]\frac{q_0'}{q_1} \tag{1.26}$$

but there is no guarantee as yet that (1.25) holds. Assuming for the moment that (1.25) is indeed valid, we may apply Hölder's inequality and use (1.22) to obtain

$$\sum_i |x_i|^{p_1} \le \Big(\sum_i |x_i|^{(p-1)p_0'}\Big)^{\lambda}\Big(\sum_i |x_i|^{p}\Big)^{1-\lambda} = \Big(\sum_i |x_i|^{(p-1)p_0'}\Big)^{\lambda}$$

and

$$\sum_j |y_j|^{q_0'} \le \Big(\sum_j |y_j|^{(q'-1)q_1}\Big)^{\mu}\Big(\sum_j |y_j|^{q'}\Big)^{1-\mu} = \Big(\sum_j |y_j|^{(q'-1)q_1}\Big)^{\mu}.$$

Using these estimates in the right-hand side of (1.23) and appealing to the identities (1.26), we see easily that the desired result (1.21) is produced.

It remains only to check the validity of (1.25). The positivity of λ and μ is clear from (1.26) so we need only show that they do not exceed 1. By (1.26), this is equivalent to the requirement

$$\frac{q_0'}{q_1} \le \frac{1-\theta}{\theta} \le \frac{p_0'}{p_1}. \tag{1.27}$$

However, since (α_k, β_k), $(k = 0, 1)$, belong to the lower triangle $\triangle$, we have $p_0 \le q_0$ (so $q_0' \le p_0'$) and $p_1 \le q_1$. Hence, $q_0'/q_1 \le p_0'/p_1$ and so, since $(1-\theta)/\theta$ assumes all positive values as θ ranges from 0 to 1, it is clear that (1.25) must be satisfied for some value of θ satisfying $0 < \theta < 1$. The proof above now shows that (1.21) holds for this particular value of θ. By the remark made at the beginning of the proof, this establishes (1.21) for all θ and hence completes the proof. ■

The preceding results have straightforward analogues in the case where the ℓ^p- and ℓ^q-norms are replaced by weighted norms. Let $\boldsymbol{\rho} = (\rho_1, \rho_2, \ldots, \rho_m)$ and $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ be fixed sequences of positive numbers. Let

$$\|\mathbf{x}\|_{\ell^p_\rho} = \Big(\sum_{i=1}^{m} |x_i|^p \rho_i\Big)^{1/p}, \qquad \|\mathbf{y}\|_{\ell^q_\sigma} = \Big(\sum_{j=1}^{n} |y_j|^q \sigma_j\Big)^{1/q} \tag{1.28}$$

when p and q are finite, and take these to be the usual ℓ^∞-norms when p and q are infinite.

Keeping the basic definitions (1.10), (1.11), and (1.12) of A, $\mathbf{X}$, and $\mathbf{Y}$ as before, we now compute the maximum of the form A with respect to the weighted norms (1.28). Thus, following (1.13), we set

$$M^*(\alpha, \beta) = \max\{|A(\mathbf{x}, \mathbf{y})| : \|\mathbf{x}\|_{\ell^p_\rho} \le 1, \|\mathbf{y}\|_{\ell^{q'}_\sigma} \le 1\}. \tag{1.29}$$

Using exactly the same arguments as before, but with M replaced by M^*, X_j by X_j/σ_j, Y_i by Y_i/ρ_i, and all the norms by their weighted analogues, we arrive without difficulty at the following generalization of Theorem 1.5.

Theorem 1.6. *The maximum $M^*(\alpha,\beta)$ of an arbitrary bilinear form A, computed with respect to weighted norms, is logarithmically convex in the lower triangle $\triangle$ of the unit square.*

With Theorem 1.6, we can now pass to the main interpolation theorem for strong-type operators defined with respect to a pair of measure spaces (R,μ) and (S,ν), as in Definition 1.1.

Theorem 1.7 (Riesz convexity theorem). *Suppose $1 \le p_k \le q_k \le \infty$, $(k=0,1)$, and let T be a linear operator of strong types (p_k,q_k) with respective strong-type norms M_k, $(k=0,1)$. Suppose $0 \le \theta \le 1$ and define p and q by*

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \qquad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}. \tag{1.30}$$

Then T is of strong type (p,q) and its strong-type (p,q) norm M_θ satisfies

$$M_\theta \le M_0^{1-\theta} M_1^\theta. \tag{1.31}$$

Proof. It will suffice to show that

$$\left|\int_S (Tf)g\,d\nu\right| \le M_0^{1-\theta} M_1^\theta \tag{1.32}$$

for all simple functions f and g satisfying

$$\|f\|_{L^p(\mu)} = \|g\|_{L^{q'}(\nu)} = 1. \tag{1.33}$$

Such functions f and g may be expressed in the form

$$f = \sum_{i=1}^m x_i \chi_{F_i}, \qquad g = \sum_{j=1}^n y_j \chi_{G_j}, \tag{1.34}$$

where the F_i are disjoint subsets of R of finite μ-measure, the G_j are disjoint subsets of S of finite ν-measure, and x_i and y_j, by virtue of (1.33), satisfy

$$\sum_{i=1}^m |x_i|^p \mu(F_i) = \sum_{j=1}^n |y_j|^{q'} \nu(G_j) = 1$$

in case p and q' are finite, or

$$\max_i |x_i| = \max_j |y_j| = 1$$

if these parameters are infinite. In any case, if we set

$$\rho_i = \mu(F_i), \qquad (i = 1,2,\ldots,m), \qquad \sigma_j = \nu(G_j), \qquad (j = 1,2,\ldots,n),$$

then (cf. (1.28)) these requirements may be written

$$\|\mathbf{x}\|_{\ell^p_\rho} = \|\mathbf{y}\|_{\ell^{q'}_\sigma} = 1. \tag{1.35}$$

Using (1.34) and the linearity of T, we have

$$\int_S (Tf)g\,d\nu = \sum_{i,j} x_i y_j \int_S (T\chi_{F_i})\chi_{G_j}\,d\nu = A(\mathbf{x},\mathbf{y}), \tag{1.36}$$

where A is the bilinear form with coefficients

$$a_{ij} = \int_S (T\chi_{F_i})\chi_{G_j}\,d\nu, \qquad (i = 1,\ldots,m;\ j = 1,\ldots,n). \tag{1.37}$$

With the standard notation (1.9), we write M^* for $M^*(\alpha,\beta)$ and M_k^* for $M_k^*(\alpha_k,\beta_k)$, $(k=0,1)$. Then it follows from (1.29), (1.35), and (1.36) that

$$\left|\int_S (Tf)g\,d\nu\right| \le M^*. \tag{1.38}$$

However, the hypotheses on p_k and q_k imply that the line segment with endpoints (α_k,β_k), $(k=0,1)$, lies in the lower triangle $\triangle$ and hence Theorem 1.6 shows that $M^* \le (M_0^*)^{1-\theta}(M_1^*)^\theta$. This and (1.38) together will produce the desired result (1.32), provided we show that

$$M_k^* \le M_k, \qquad (k=0,1). \tag{1.39}$$

To this end, recall that M_k is the strong-type (p_k,q_k) norm of T. Hence, as in (1.36) and (1.37), we have, for arbitrary vectors $\mathbf{w} = (w_1,\ldots,w_m)$ and $\mathbf{z} = (z_1,\ldots,z_n)$,

$$\begin{aligned}\left|\sum_{i=1}^m \sum_{j=1}^n a_{ij} w_i z_j\right| &= \left|\int_S \left[T\Big(\sum_i w_i \chi_{F_i}\Big)\right]\Big(\sum_j z_j \chi_{G_j}\Big)\,d\nu\right| \\ &\le \left\|T\Big(\sum_i w_i \chi_{F_i}\Big)\right\|_{L^{q_k}(\nu)} \left\|\sum_j z_j \chi_{G_j}\right\|_{L^{q_k'}(\nu)} \\ &\le M_k \left\|\sum_i w_i \chi_{F_i}\right\|_{L^{p_k}(\mu)} \left\|\sum_j z_j \chi_{G_j}\right\|_{L^{q_k'}(\nu)} \\ &= M_k \|\mathbf{w}\|_{\ell^{p_k}_\rho} \|\mathbf{z}\|_{\ell^{q_k'}_\sigma}, \qquad (k=0,1).\end{aligned}$$

An appeal to (1.29) now establishes the desired result (1.39) and hence completes the proof. ■

The operator T in Theorem 1.7 is defined *a priori* only on the simple functions but it does of course have a unique norm-preserving extension to all of L^p provided p is finite. The hypotheses can also be formulated in terms of operators that are defined on the full L^p-spaces instead of just the simple functions (Corollary 1.8 below). In the language of the abstract interpolation theory introduced at the beginning of Chapter III, this version of the theorem exhibits (L^p, L^q) as an exact interpolation pair for the couples (L^{p_0}, L^{p_1}) and (L^{q_0}, L^{q_1}). In this connection, notice that if $1 \le p_0 \le p \le p_1 \le \infty$, so $1/p = (1-\theta)/p_0 + \theta/p_1$ for some θ with $0 \le \theta \le 1$, then it follows easily from Hölder's inequality that

$$\|f\|_p \le \|f\|_{p_0}^{1-\theta}\|f\|_{p_1}^{\theta}, \tag{1.40}$$

and hence that

$$L^{p_0} \cap L^{p_1} \hookrightarrow L^p \hookrightarrow L^{p_0} + L^{p_1} \tag{1.41}$$

(cf. Exercises III.6,7). In other words, L^p is an intermediate space for the couple (L^{p_0}, L^{p_1}).

Corollary 1.8. *Let p_k, q_k, $(k = 0, 1)$, θ, p, and q be as in the statement of Theorem 1.7. Let T be an admissible operator for the couples (L^{p_0}, L^{p_1}), (L^{q_0}, L^{q_1}) and let M_k denote the operator norm of T considered as a bounded operator from L^{p_k} into L^{q_k}:*

$$\|Tf\|_{L^{q_k}(\nu)} \le M_k\|f\|_{L^{p_k}(\mu)}, \qquad (f \in L^{p_k}(\mu)). \tag{1.42}$$

Then T is a bounded operator from L^p into L^q:

$$\|Tf\|_{L^q(\nu)} \le M\|f\|_{L^p(\mu)}, \qquad (f \in L^p(\mu)) \tag{1.43}$$

and its operator norm M satisfies

$$M \le M_0^{1-\theta}M_1^{\theta}. \tag{1.44}$$

Proof. When p is finite, the result follows immediately from Theorem 1.7 since then the simple functions are dense in L^p. Suppose therefore that p is infinite. If $\theta = 0$ or $\theta = 1$, there is nothing to prove. Hence, we may assume $0 < \theta < 1$. In that case, we have $p_0 = p = p_1 = \infty$ and it follows from (1.40) and (1.42) that

$$\begin{aligned}\|Tf\|_q &\le \|Tf\|_{q_0}^{1-\theta}\|Tf\|_{q_1}^{\theta} \le (M_0\|f\|_\infty)^{1-\theta}(M_1\|f\|_\infty)^{\theta}\\ &= M_0^{1-\theta}M_1^{\theta}\|f\|_\infty.\end{aligned}$$

This establishes (1.43) and (1.44) in the case at hand and hence completes the proof. ■

2. THE RIESZ-THORIN CONVEXITY THEOREM

Example 1.3 of the previous section shows that the Riesz convexity theorem does not extend to the "upper triangle" of the unit square, at least for real L^p-spaces. The question of its validity in the entire unit square for complex L^p-spaces was raised by M. Riesz in his original 1926 work on the convexity theorem, and an affirmative answer was provided by O. V. Thorin in 1939. Thorin's ingenious proof, which we present below, drew subsequent tribute from J. E. Littlewood as "the most impudent idea in analysis". It is based on the following elementary convexity property of analytic functions. We denote by Ω the strip $\{z = x + iy : 0 < x < 1\}$ in the complex plane and by $\bar{\Omega}$ its closure.

Lemma 2.1 (Hadamard three-lines theorem). *Let F be a bounded continuous function on $\bar{\Omega}$, analytic in Ω. Then the function*

$$M_\theta = \sup\{|F(\theta + iy)| : -\infty < y < \infty\}$$

satisfies

$$M_\theta \le M_0^{1-\theta}M_1^{\theta}, \qquad (0 \le \theta \le 1). \tag{2.1}$$

Proof. We need to show that

$$|F(\theta + iy)| \le M_0^{1-\theta}M_1^{\theta}$$

for all real y. It is enough to do so for all pairs of constants larger than M_0, M_1, so we may assume that neither M_0 nor M_1 is zero. Then, by considering the function $F(z)/(M_0^{1-z}M_1^{z})$, it is clear that the result need only be established in the case $M_0 = M_1 = 1$.

Under these hypotheses, the analytic function F is bounded by 1 on the boundary $\partial\Omega$ of Ω and is bounded (by a constant K, say) on Ω itself. We need to show that

$$|F(z)| \le 1, \qquad (z \in \Omega). \tag{2.2}$$

For each $\varepsilon > 0$, let F_ε be the analytic function given by $F_\varepsilon(z) = F(z)/(1 + \varepsilon z)$. Then

$$|F_\varepsilon(z)| \le \frac{|F(z)|}{1 + \varepsilon x} \le 1, \qquad (z = x + iy \in \partial\Omega)$$

and

$$|F_\varepsilon(z)| \le \frac{|F(z)|}{\varepsilon|y|} \le \frac{K}{\varepsilon|y|}, \qquad (z = x + iy \in \Omega). \tag{2.3}$$

These estimates show that $|F_\varepsilon(z)| \le 1$ on the boundary ∂R of the rectangle R whose vertices lie at $\pm iK/\varepsilon$ and $1 \pm iK/\varepsilon$. The maximum modulus theorem shows that the same estimate persists throughout all of R. Since (2.3) provides the same estimate again on $\Omega \setminus R$, we conclude that $|F(z)| \le 1 + \varepsilon|z|$ on all of Ω. Letting $\varepsilon \to 0$, we obtain the desired result (2.2). ■

The three-lines theorem will be used to establish Thorin's result for complex L^p-spaces (part (a) of the following theorem). For completeness, we include the corresponding result for real L^p-spaces (part (b)). In the lower triangle $\triangle$ of the unit square, this is merely a restatement of the Riesz convexity theorem (Theorem 1.7). Elsewhere in the unit square, the result for real L^p-spaces is obtained from Thorin's result by "complexification".

Theorem 2.2 (Riesz-Thorin convexity theorem). *Suppose*

$$1 \le p_0, p_1, q_0, q_1 \le \infty.$$

Suppose $0 \le \theta \le 1$ *and let*

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \qquad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}. \tag{2.4}$$

(a) *(complex spaces) Let T be a linear operator of strong types (p_0, q_0) and (p_1, q_1) with respective strong-type norms M_0 and M_1. Then T is of strong type (p, q) and its strong-type (p, q) norm M_θ satisfies*

$$M_\theta \le M_0^{1-\theta} M_1^\theta. \tag{2.5}$$

(b) *(real spaces) For real spaces, the result of part* (a) *remains valid if $p_k \le q_k$, $(k = 0, 1)$. Otherwise, it holds with* (2.5) *replaced by*

$$M_\theta \le 2M_0^{1-\theta} M_1^\theta. \tag{2.6}$$

Proof. (a) For each complex number z, let

$$\alpha(z) = \frac{1-z}{p_0} + \frac{z}{p_1}, \qquad \beta(z) = \frac{1-z}{q_0} + \frac{z}{q_1}. \tag{2.7}$$

Then from (2.4) we have

$$\alpha(\theta) = \frac{1}{p}, \qquad \beta(\theta) = \frac{1}{q}. \tag{2.8}$$

The theorem will be proved if we show that

$$\left| \int_S (Tf) g \, dv \right| \le M_0^{1-\theta} M_1^\theta \tag{2.9}$$

for all simple functions f and g satisfying

$$\|f\|_{L^p(\mu)} = \|g\|_{L^{q'}(v)} = 1. \tag{2.10}$$

Suppose first that both p and q' are finite. Then such functions f and g may be expressed in the form

$$f = \sum_{j=1}^{J} a_j \chi_{A_j}, \qquad g = \sum_{k=1}^{K} b_k \chi_{B_k}, \tag{2.11}$$

where the A_j are disjoint subsets of R of finite μ-measure, the B_k are disjoint subsets of S of finite v-measure, and the coefficients a_j and b_k, by virtue of (2.10), satisfy

$$\sum_{j=1}^{J} |a_j|^p \mu(A_j) = \sum_{k=1}^{K} |b_k|^{q'} v(B_k) = 1. \tag{2.12}$$

For each complex number z, set

$$\begin{aligned} f_z &= |f|^{\alpha(z)/\alpha(\theta)} e^{i \cdot \arg f}, \\ g_z &= |g|^{(1-\beta(z))/(1-\beta(\theta))} e^{i \cdot \arg g} \end{aligned} \tag{2.13}$$

and define $F(z)$ by

$$F(z) = \int_S (Tf_z) g_z \, dv. \tag{2.14}$$

Since $F(\theta)$ coincides with the integral in (2.9), that estimate will follow at once from Lemma 2.1 if we show that F is a bounded analytic function on $\bar{\Omega}$ and

$$|F(iy)| \le M_0, \qquad |F(1+iy)| \le M_1, \qquad (-\infty < y < \infty). \tag{2.15}$$

The analyticity is clear because, by linearity of T,

$$\begin{aligned} F(z) = \sum_{j=1}^{J} \sum_{k=1}^{K} & |a_j|^{\alpha(z)/\alpha(\theta)} |b_k|^{(1-\beta(z))/(1-\beta(\theta))} \\ & \times \int_S (T\chi_{A_j}) \chi_{B_k} e^{i(\arg a_j + \arg b_k)} \, dv. \end{aligned}$$

This expresses F as a finite linear combination of pure exponentials so that F is in fact entire. It also reveals that F is bounded because, as an inspection of (2.7) shows, the real parts of $\alpha(z)$ and $\beta(z)$ are bounded on $\bar{\Omega}$. Hence, all that remains is to establish (2.15).

Applying Hölder's inequality to (2.14) and using the fact that T has strong-type (p_0, q_0) norm M_0, we obtain

$$\begin{aligned} |F(iy)| &\le \|Tf_{iy}\|_{L^{q_0}(v)} \|g_{iy}\|_{L^{q_0'}(v)} \\ &\le M_0 \|f_{iy}\|_{L^{p_0}(\mu)} \|g_{iy}\|_{L^{q_0'}(v)} \end{aligned} \tag{2.16}$$

for all real y. From (2.7) and (2.8), we have $\operatorname{Re}\alpha(iy) = 1/p$ and $\alpha(\theta) = 1/p$. Hence, by (2.11), (2.12), and (2.13),

$$\begin{aligned} \|f_{iy}\|_{L^{p_0}(\mu)}^{p_0} &= \sum_{j=1}^{J} \big||a_j|^{\alpha(iy)/\alpha(\theta)}\big|^{p_0}\mu(A_j) \\ &= \sum_{j=1}^{J} |a_j|^p \mu(A_j) = 1. \end{aligned} \tag{2.17}$$

Similarly,

$$\begin{aligned} \|g_{iy}\|_{L^{q_0'}(\nu)}^{q_0'} &= \sum_{k=1}^{K} \big||b_k|^{(1-\beta(iy))/(1-\beta(\theta))}\big|^{q_0'}\nu(B_k) \\ &= \sum_{k=1}^{K} |b_k|^{q'}\nu(B_k) = 1. \end{aligned} \tag{2.18}$$

These relations, in conjunction with (2.16), produce the first of the estimates in (2.15). The second follows by an analogous argument using the hypothesis that T has strong-type (p_1, q_1) norm equal to M_1.

This completes the proof for finite p and q. In the remaining cases, we may assume $0 < \theta < 1$ since otherwise there is nothing to prove. Then $p = q = \infty$ can occur only if $p_k = q_k = \infty$, $(k = 0, 1)$, and again there is nothing to prove. If $p < \infty$ and $q = \infty$ (so $q_0 = q = q_1 = \infty$), define f_z as before in (2.13) but now take $g_z = g$ for all z. With this modification, the proof proceeds exactly as before. The case where $p = \infty$ and $q < \infty$ is disposed of in similar fashion and with this the proof of part (a) is complete.

(b) If T is a linear operator of strong type (r, s) with strong-type norm equal to N, then T extends to a (complex-) linear operator $T'(f + ig) = Tf + iTg$ of strong type (r, s) relative to the complex spaces, and its strong-type norm N' satisfies $N \le N' \le 2N$. With this observation, part (b) of the theorem follows directly from part (a). Of course, if the additional hypotheses $p_k \le q_k$ are satisfied, then Theorem 1.7 provides the sharper estimate (2.5) in place of (2.6). ■

For completeness, we include here the corresponding statement for operators defined on the full L^p-spaces rather than just on the simple functions. The proof is similar to that of Corollary 1.8.

Corollary 2.3. *Let p_k, q_k, θ, p, and q be as in the statement of the previous theorem.*

(a) (*complex spaces*) *Let T be an admissible operator for the couples (L^{p_0}, L^{p_1}), (L^{q_0}, L^{q_1}) and let M_k denote the operator norm of T considered as a bounded operator from L^{p_k} into L^{q_k}, for $k = 0, 1$. Thus,*

$$\|Tf\|_{L^{q_k}(\nu)} \le M_k \|f\|_{L^{p_k}(\mu)}, \qquad (f \in L^{p_k}(\mu),\ k = 0, 1). \tag{2.19}$$

Then T is a bounded operator from L^p into L^q:

$$\|Tf\|_{L^q(\nu)} \le M \|f\|_{L^p(\mu)}, \qquad (f \in L^p(\mu)) \tag{2.20}$$

and its operator norm M satisfies

$$M \le M_0^{1-\theta} M_1^{\theta}. \tag{2.21}$$

(b) (*real spaces*) *For real spaces, the result of part* (a) *remains valid if $p_k \le q_k$, $(k = 0, 1)$. Otherwise, it holds with* (2.21) *replaced by*

$$M \le 2M_0^{1-\theta} M_1^{\theta}. \tag{2.22}$$

We shall use the Riesz-Thorin theorem to derive Young's inequality for the convolution product and the Hausdorff-Young theorem for the Fourier transform. The first is not hard to establish directly but the second is considerably more difficult to prove without interpolation.

Recall from Section III.6 that the convolution $f * g$ of two functions f and g in $L^1(\mathbf{T})$ is given by

$$(f * g)(e^{it}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{is})g(e^{i(t-s)})\,ds, \qquad (e^{it} \in \mathbf{T}). \tag{2.23}$$

Theorem 2.4 (Young's inequality). *Suppose $1 \le p, q \le \infty$ and $1/r = 1/p + 1/q - 1 \ge 0$. If $f \in L^p$ and $g \in L^q$, then $f * g \in L^r$ and*

$$\|f * g\|_r \le \|f\|_p \|g\|_q. \tag{2.24}$$

Proof. For fixed f in L^1, the estimates

$$\|f * g\|_1 \le \|f\|_1 \|g\|_1, \qquad \|f * g\|_\infty \le \|f\|_1 \|g\|_\infty,$$

which follow immediately from (2.23), exhibit the linear operator $g \to f * g$ as a bounded operator from L^1 into L^1 and from L^∞ into L^∞, with operator norm in each case at most $\|f\|_1$. Hence, Corollary 2.3 shows that the operator is bounded from L^q into L^q with norm at most $\|f\|_1$:

$$\|f * g\|_q \le \|f\|_1 \|g\|_q, \qquad (f \in L^1,\ g \in L^q).$$

This, together with the estimate

$$\|f * g\|_\infty \le \|f\|_{q'} \|g\|_q, \qquad (f \in L^{q'},\ g \in L^q),$$

which follows by applying Hölder's inequality to (2.23), shows that, for fixed

g in L^q, the linear operator $f \to f * g$ is bounded from L^1 into L^q and from $L^{q'}$ into L^∞, with operator norm in each case at most $\|g\|_q$. The hypotheses imply that $1/p = (1-\theta)/1 + \theta/q'$ and $1/r = (1-\theta)/q + \theta/\infty$, with $\theta = q/p'$ satisfying $0 \le \theta \le 1$. Hence, Corollary 2.3 shows that $f \to f * g$ is a bounded operator from L^p into L^r with norm at most $\|g\|_q$. This is precisely the content of (2.24). ■

Recall from Definition III.6.2 that the n-th Fourier coefficient of a function f in $L^1(\mathbf{T})$ is given by

$$\hat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{it})e^{-int}\,dt, \qquad (n \in \mathbf{Z}). \tag{2.25}$$

Hence, there is the trivial estimate

$$|\hat{f}(n)| \le \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(e^{it})|\,dt = \|f\|_{L^1}, \qquad (n \in \mathbf{Z}) \tag{2.26}$$

for functions f in $L^1(\mathbf{T})$.

There is a dual situation for sequences $\mathbf{c} = (c_n)_{n=-\infty}^{\infty}$ in $\ell^1(\mathbf{Z})$. For any such sequence $\mathbf{c}$, the series $\sum c_n e^{int}$ converges absolutely, hence uniformly, to a continuous function, say f, on $\mathbf{T}$. Clearly,

$$|f(e^{it})| \le \sum_{n=-\infty}^{\infty} |c_n| = \|\mathbf{c}\|_{\ell^1}, \qquad (e^{it} \in \mathbf{T}). \tag{2.27}$$

For square-integrable functions, there is the following well known result [HS, p. 245].

Theorem 2.5 (a) (**Parseval**). *If $f \in L^2(\mathbf{T})$, then $\hat{f} \in \ell^2(\mathbf{Z})$ and*

$$\sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(e^{it})|^2\,dt. \tag{2.28}$$

(b) (**F. Riesz-Fischer**) *If $\mathbf{c} = (c_n)_{n=-\infty}^{\infty} \in \ell^2(\mathbf{Z})$, then there is a function f in $L^2(\mathbf{T})$ with $\hat{f}(n) = c_n$ for all n. In particular,* (2.28) *holds for this f.*

Interpolating now between the L^1- and L^2-results, we obtain the following L^p-estimate for the Fourier transform.

Theorem 2.6 (**Hausdorff-Young theorem**). *Suppose $1 \le p \le 2$ and $1/p + 1/p' = 1$.*
(a) *If $f \in L^p(\mathbf{T})$, then $\hat{f} \in \ell^{p'}(\mathbf{Z})$ and*

$$\|\hat{f}\|_{p'} \le \|f\|_p. \tag{2.29}$$

(b) *If $\mathbf{c} = (c_n)_{n=-\infty}^{\infty} \in \ell^p(\mathbf{Z})$, then the function f with $\hat{f} = \mathbf{c}$ belongs to $L^{p'}(\mathbf{T})$ and*

$$\|f\|_{p'} \le \|\mathbf{c}\|_p. \tag{2.30}$$

Proof. The estimates (2.26) and (2.28) show that the Fourier transform $f \to \hat{f}$ is bounded from L^1 into ℓ^∞ and from L^2 into ℓ^2, with operator norm in each case at most 1. If $\theta = 2/p'$, then $0 \le \theta \le 1$ and we have $1/p = (1-\theta)/1 + \theta/2$ and $1/p' = (1-\theta)/\infty + \theta/2$. Hence, it follows from Corollary 2.3 that the Fourier transform maps L^p into $\ell^{p'}$, with operator norm at most 1. This is the content of (2.29).
(b) The proof of part (b) is similar, once the operator in question is properly defined. The operator $\mathscr{F}_1$ given by

$$\mathscr{F}_1 : \mathbf{c} \to f(e^{it}) = \sum_{n=-\infty}^{\infty} c_n e^{int}$$

is well defined on ℓ^1 into L^∞ (cf. (2.27)). On the other hand, the Riesz-Fischer theorem exhibits an isometry, say $\mathscr{F}_2$, from ℓ^2 onto L^2 with the property that $(\mathscr{F}_2\mathbf{c})\hat{}\,(n) = c_n$ for all n. Thus, if $\mathbf{c} \in \ell^1$, then $\mathscr{F}_1\mathbf{c}$ and $\mathscr{F}_2\mathbf{c}$ have the same Fourier coefficients and hence coincide a.e. [HS, p. 249]. Since ℓ^1 is dense in ℓ^2, the operator $\mathscr{F}_2$ is therefore the unique extension of $\mathscr{F}_1$ to an isometry of ℓ^2 into L^2. This extension is the *Fourier transform* defined with respect to the group $\mathbf{Z}$ of integers. Interpolating as in part (a), we obtain (2.30). ■

Young's inequality and the Hausdorff-Young theorem are valid when $\mathbf{T}$ or $\mathbf{Z}$ is replaced by the real line $\mathbf{R}$ or indeed any locally compact abelian group. The proofs, by interpolation, are the same as above. Young's inequality (2.24) and the Hausdorff-Young inequalities (2.29), (2.30) are sharp for $\mathbf{T}$ and $\mathbf{Z}$ (use constant functions f on $\mathbf{T}$, and sequences $\mathbf{c}$ on $\mathbf{Z}$ that have only one nonzero coefficient) but not for $\mathbf{R}$ (cf. Exercises 13,14).

In the proof of Young's inequality, two distinct interpolations were used (one for each of the "variables" f and g). The following "multilinear" version of the Riesz-Thorin theorem achieves the same result with a single interpolation in the two variables simultaneously.

We consider *multilinear* operators T, that is, operators $(f_1, f_2, \ldots, f_n) \to T(f_1, f_2, \ldots, f_n)$, linear in each f_j, defined for each n-tuple of simple functions f_j on measure spaces (R_j, μ_j), $(j = 1, 2, \ldots, n)$, and taking values in a measure space (S, ν). Such an operator is of *strong type* $(p_1, p_2, \ldots, p_n; q)$ if

$$\|T(f_1, f_2, \ldots, f_n)\|_{L^q(\nu)} \le M\|f_1\|_{L^{p_1}(\mu_1)} \cdots \|f_n\|_{L^{p_n}(\mu_n)}$$

for all simple $f_1, f_2, \ldots, f_n$. We present the multilinear interpolation theorem only for complex spaces. The proof is entirely analogous to that of Theorem 2.2(a) and is omitted.

Theorem 2.7. *Suppose* $1 \le p_{0,j}, p_{1,j} \le \infty$, $(j = 1,2,\ldots,n)$, *and* $1 \le q_0$, $q_1 \le \infty$. *Suppose* $0 \le \theta \le 1$ *and let*

$$\frac{1}{p_j} = \frac{1-\theta}{p_{0,j}} + \frac{\theta}{p_{1,j}}, \qquad (j = 1,\ldots,n), \qquad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}. \tag{2.31}$$

Let T be a multilinear operator of strong types $(p_{0,1}, p_{0,2}, \ldots, p_{0,n}; q_0)$ and $(p_{1,1}, p_{1,2}, \ldots, p_{1,n}; q_1)$ with respective strong-type norms M_0 and M_1. Then T is of strong type $(p_1, p_2, \ldots, p_n; q)$ and its strong-type norm M satisfies

$$M \le M_0^{1-\theta} M_1^{\theta}. \tag{2.32}$$

We conclude this section with a version of the Riesz-Thorin theorem involving compact operators. The result is typical of a number of interpolation theorems for compact operators in that compactness at one of the endpoints of the interpolation segment implies compactness throughout the interior of the segment. We present the result here only for finite measure spaces; the modifications required in the case of infinite measure spaces are left as an exercise.

If $T: X \to Y$ is a bounded linear operator between Banach spaces X and Y, then T is *compact* if it maps the unit ball of X into a set with compact closure in Y. It is not difficult to show that *finite rank operators* (i.e., those with finite dimensional range) are compact, and that norm-limits of compact operators are compact. We shall be concerned with Banach spaces of measurable functions. In this setting, it is natural to construct finite-rank operators by composing a given operator with averaging operators of the kind described in the following lemma.

Lemma 2.8. *Let (S, ν) be a finite measure space and suppose $g_1, g_2, \ldots, g_m$ belong to $L^\infty(\nu)$. Then, to each $\varepsilon > 0$, there corresponds a partition of S into finitely many disjoint subsets $E_1, E_2, \ldots, E_K$* of S, *each of positive measure, such that the averaging operator*

$$\mathscr{E} g = \sum_{k=1}^{K} \left(\frac{1}{\nu(E_k)} \int_{E_k} g \, d\nu \right) \chi_{E_k} \tag{2.33}$$

satisfies

$$\|\mathscr{E} g_j - g_j\|_{L^\infty(\nu)} \le \varepsilon, \qquad (j = 1,2,\ldots,m). \tag{2.34}$$

Proof. Let $\varepsilon > 0$ and put

$$N = (2/\varepsilon)\sup\{\|g_j\|_\infty : j = 1,2,\ldots,m\}.$$

For each $j = 1,2,\ldots,m$, the $2N+1$ sets

$$\{\eta\varepsilon/2 \le \mathrm{Re}(g_j) < (\eta+1)\varepsilon/2\}, \qquad (|\eta| \le N)$$

comprise a partition of S. For each j, there is a similar partition with $2N+1$ members corresponding to the imaginary part of g_j. The common refinement of all these $2m$ partitions taken together will therefore be a partition with at most $(2N+1)^{2m}$ members. These are essentially the sets we want except that some of them may have measure zero. However, we may coalesce all such sets into a single set of measure zero and then combine it with any one of the remaining sets of positive measure. The result is a collection of pairwise disjoint ν-measurable subsets $E_1, E_2, \ldots, E_K$ of S, where $K \le (2N+1)^{2m}$, such that each E_k has positive measure and $\bigcup_{k=1}^{K} E_k = S$.

Fix $j = 1,2,\ldots,m$. For almost all x in E_k we have, because of the defining property of the original partitions,

$$|g_j(x) - g_j(y)| \le |\mathrm{Re}(g_j(x) - g_j(y))| + |\mathrm{Im}(g_j(x) - g_j(y))|$$
$$\le \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon,$$

for almost all y in E_k. Hence, the averaging operator $\mathscr{E}$ defined by (2.33) satisfies

$$|\mathscr{E} g_j(x) - g_j(x)| \le \frac{1}{\nu(E_k)} \int_{E_k} |g_j(y) - g_j(x)| \, d\nu(y) \le \varepsilon,$$

for almost all x in E_k, $(k = 1,2,\ldots,K)$. Since S is the union of the sets E_k, this establishes (2.34) and hence completes the proof. ■

Theorem 2.9. *Let (R, μ) and (S, ν) be finite measure spaces. Suppose $1 \le p_j, q_j \le \infty$, $(j = 0,1)$, and let T be a linear operator that satisfies*

$$T: L^{p_0}(\mu) \to L^{q_0}(\nu) \qquad \textit{boundedly} \tag{2.35}$$

and

$$T: L^{p_1}(\mu) \to L^{q_1}(\nu) \qquad \textit{compactly.} \tag{2.36}$$

If $0 < \theta < 1$ and p, q are defined by

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \qquad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1},$$

then

$$T: L^{p}(\mu) \to L^{q}(\nu) \qquad \textit{compactly.} \tag{2.37}$$

Proof. The Riesz-Thorin theorem (Corollary 2.3) shows that T is a bounded operator from L^p into L^q so it remains only to establish the compactness. For each $\varepsilon > 0$, we shall construct a finite-rank operator T_ε from L^p to L^q for which

$$\|T - T_\varepsilon\|_{\mathscr{B}(L^p, L^q)} \le \{2\|T\|_{\mathscr{B}(L^{p_0}, L^{q_0})}\}^{1-\theta} \varepsilon^{\theta}. \tag{2.38}$$

This will exhibit T as a norm-limit of finite-rank operators and hence establish its compactness.

It follows from (2.36) that the set $V = \{Tf : \|f\|_{p_1} \le 1\}$ has compact closure in L^{q_1} and so, in particular, is totally bounded in L^{q_1}. Let $\{g_j : j = 1, 2, \ldots, m\}$ be an ε'-net for V with $\varepsilon' = \varepsilon/(4 + \nu(S))$. If $q_1 < \infty$, then $L^{q_1}(\nu)$ has absolutely continuous norm and so we may choose a ν-measurable subset E of S and a positive number M such that

$$\|g_j \chi_{E^c}\|_{q_1} < \varepsilon' \qquad (j = 1, 2, \ldots, m) \tag{2.39}$$

and

$$\|g_j \chi_E\|_\infty \le M \qquad (j = 1, 2, \ldots, m). \tag{2.40}$$

In the case $q_1 = \infty$, this may be achieved by choosing $E = S$ and M a uniform essential bound for all the functions g_j.

By Lemma 2.8, there are disjoint subsets $E_1, E_2, \ldots, E_K$ of E, each of positive measure and with union equal to E, such that the averaging operator

$$\mathscr{E}g = \sum_{k=1}^{K} \left(\frac{1}{\nu(E_k)} \int_{E_k} g \, d\nu \right) \chi_{E_k} \tag{2.41}$$

satisfies

$$\|\mathscr{E}g_j - g_j \chi_E\|_{L^\infty(\nu)} < \varepsilon', \qquad (j = 1, 2, \ldots, m). \tag{2.42}$$

Now let g be an arbitrary member of V. Then there exists j with $1 \le j \le m$ such that

$$\|g - g_j\|_{q_1} < \varepsilon'. \tag{2.43}$$

The averaging operator $\mathscr{E}$ is a contraction on L^{q_1} (cf. Theorem II.4.8) so

$$\begin{aligned} \|\mathscr{E}g - g\|_{q_1} &\le \|\mathscr{E}(g - g_j)\|_{q_1} + \|\mathscr{E}g_j - g_j\|_{q_1} + \|g_j - g\|_{q_1} \\ &\le 2\|g - g_j\|_{q_1} + \|\mathscr{E}g_j - g_j\|_{q_1}. \end{aligned} \tag{2.44}$$

The first term is easily estimated from (2.43). For the second, we have from (2.39) and (2.42),

$$\begin{aligned} \|\mathscr{E}g_j - g_j\|_{q_1} &\le \|g_j \chi_{E^c}\|_{q_1} + \gamma \|\mathscr{E}g_j - g_j \chi_E\|_\infty \\ &< \varepsilon'(1 + \gamma), \end{aligned}$$

where $\gamma = 1$ if $q_1 = \infty$ and $\gamma = \nu(E)^{1/q_1}$ otherwise. Hence, from (2.44), we have

$$\|\mathscr{E}g - g\|_{q_1} \le \varepsilon'(3 + \gamma) \le \varepsilon'(4 + \nu(S)) = \varepsilon. \tag{2.45}$$

The operator $T_\varepsilon = \mathscr{E} \circ T$ is a finite-rank operator which, by virtue of (2.45), satisfies

$$\begin{aligned} \|T - T_\varepsilon\|_{\mathscr{B}(L^{p_1}, L^{q_1})} &= \sup_{\|f\|_{p_1} \le 1} \|Tf - \mathscr{E}Tf\|_{q_1} \\ &= \sup_{g \in V} \|g - \mathscr{E}g\|_{q_1} \le \varepsilon. \end{aligned}$$

By the Riesz-Thorin theorem, we therefore have

$$\begin{aligned} \|T - T_\varepsilon\|_{\mathscr{B}(L^p, L^q)} &\le \|T - T_\varepsilon\|^{1-\theta}_{\mathscr{B}(L^{p_0}, L^{q_0})} \|T - T_\varepsilon\|^{\theta}_{\mathscr{B}(L^{p_1}, L^{q_1})} \\ &\le \{2\|T\|_{\mathscr{B}(L^{p_0}, L^{q_0})}\}^{1-\theta} \varepsilon^\theta. \end{aligned}$$

This establishes (2.38) and hence completes the proof. ■

3. ANALYTIC FAMILIES OF OPERATORS

We turn now to a generalization of the Riesz-Thorin theorem in which the single operator T is replaced by a family of operators T_z that depend analytically on a complex variable z. Once the appropriate notion of such "analytic families" has been defined (Definition 3.2), the corresponding interpolation theorem can be established by suitable modification of Thorin's proof. We shall present a more comprehensive result (Theorem 3.3) in which the underlying three-lines theorem is replaced by a more general form of the maximum principle (Lemma 3.1). As we shall see, this is nothing more than the classical Poisson-Jensen formula for the open disk D transferred via conformal mapping to the strip

$$\Omega = \{z = x + iy : 0 < x < 1\}.$$

The conformal mapping in question is given by

$$z = h(\zeta) = \frac{1}{\pi i} \log \left\{ i \frac{1 + \zeta}{1 - \zeta} \right\}, \tag{3.1}$$

the inverse of which is

$$\zeta = h^{-1}(z) = \frac{e^{\pi i z} - i}{e^{\pi i z} + i}. \tag{3.2}$$

Note that h is the composition of the mappings $\zeta \to i(1 + \zeta)/(1 - \zeta)$, which maps the open unit disk D conformally onto the upper half-plane, and $w \to (1/\pi i)\log w$, which maps the upper half-plane conformally onto the strip Ω. Hence, h is a conformal mapping of D onto Ω.

It is easy to check that the upper semicircle $\{e^{i\phi} : 0 < \phi < \pi\}$ is mapped by h onto the right-hand boundary $\{1 + iy : -\infty < y < \infty\}$ of Ω, and that the

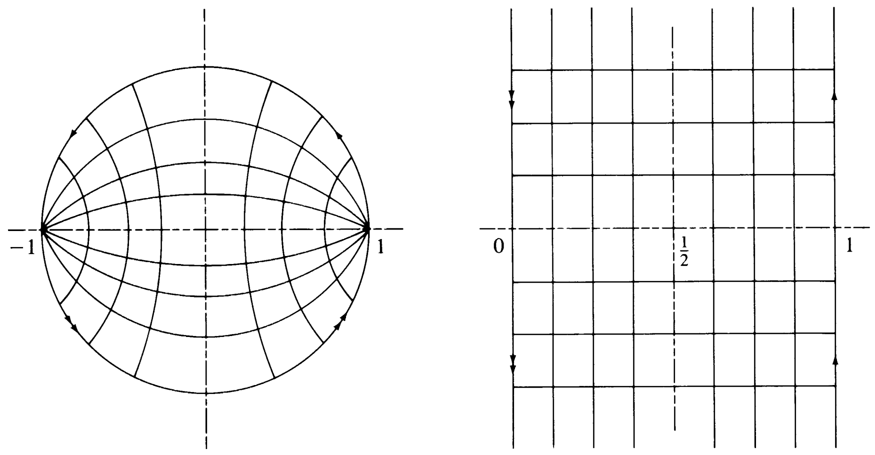


Figure 11: The conformal map $h(\zeta) = \frac{1}{\pi i}\log\left\{i\frac{1+\zeta}{1-\zeta}\right\}$.

lower semicircle $\{e^{i\phi}: -\pi < \phi < 0\}$ is mapped onto the left-hand boundary $\{iy: -\infty < y < \infty\}$. Hence, h maps $\partial D \setminus \{-1, 1\}$ onto $\partial\Omega$ (cf. Figure 11).

Vertical and horizontal lines in the strip have, as their inverse images under h, circular arcs centered on the imaginary and real axes, respectively. In particular, the midlines $\{1/2 + iy: -\infty < y < \infty\}$ and $\{x: 0 < x < 1\}$ of the strip correspond to the diameters of the disk on the real and imaginary axes, respectively. As in the three-lines theorem, we shall be interested in the values of an analytic function along the latter segment $\{x: 0 < x < 1\}$. Setting $re^{i\theta} = h^{-1}(x)$, we obtain from (3.2),

$$re^{i\theta} = \frac{e^{\pi i x} - i}{e^{\pi i x} + i} = -i\frac{\cos \pi x}{1 + \sin \pi x}. \tag{3.3}$$

With this we can compute the value of the *Poisson kernel* $P_z(\phi)$ corresponding to the point $z = re^{i\theta}$ in the disk:

$$P_z(\phi) = \operatorname{Re}\left\{\frac{e^{i\phi} + re^{i\theta}}{e^{i\phi} - re^{i\theta}}\right\} = \frac{1 - r^2}{1 - 2r\cos(\theta - \phi) + r^2}. \tag{3.4}$$

Substituting the value for $re^{i\theta}$ from (3.3), we therefore obtain

$$\begin{aligned}\frac{1 - r^2}{1 - 2r\cos(\theta - \phi) + r^2} &= \operatorname{Re}\left\{\frac{(1 + \sin \pi x)e^{i\phi} - i\cos \pi x}{(1 + \sin \pi x)e^{i\phi} + i\cos \pi x}\right\} \\ &= \frac{\sin \pi x}{1 + \cos \pi x \sin \phi}. \end{aligned} \tag{3.5}$$

Lemma 3.1. *Let F be a continuous function on $\bar{\Omega}$ that is analytic on Ω. Suppose F satisfies*

$$\sup_{z = x + iy \in \bar{\Omega}} e^{-a|y|}\log|F(z)| < \infty \tag{3.6}$$

for some constant $a < \pi$. If $0 < x < 1$, then

$$\log|F(x)| \le \frac{\sin \pi x}{2}\int_{-\infty}^{\infty}\left\{\frac{\log|F(iy)|}{\cosh \pi y - \cos \pi x} + \frac{\log|F(1 + iy)|}{\cosh \pi y + \cos \pi x}\right\} dy. \tag{3.7}$$

Proof. With h defined by (3.1), the function $G(\zeta) = F(h(\zeta))$ is evidently analytic in D and continuous on $\bar{D} \setminus \{-1, 1\}$. Hence, if $0 < r < R < 1$ and $\zeta = re^{i\theta}$, we may apply the Poisson-Jensen formula to obtain

$$\log|G(\zeta)| \le \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{R^2 - r^2}{R^2 - 2Rr\cos(\theta - \phi) + r^2}\log|G(Re^{i\phi})|\, d\phi. \tag{3.8}$$

We wish to pass to the limit as $R \to 1$ so that the integral will reflect the boundary values of G. To this end, we observe from (3.1) and (3.6) that

$$\begin{aligned}\log|G(Re^{i\phi})| &\le c \cdot \exp\{a|\operatorname{Im} h(Re^{i\phi})|\} \\ &\le c \cdot \exp\left\{\frac{a}{\pi}\left|\log\left|\frac{1 + Re^{i\phi}}{1 - Re^{i\phi}}\right|\right|\right\} \\ &\le c\{|1 + Re^{i\phi}|^{-a/\pi} + |1 - Re^{i\phi}|^{-a/\pi}\}. \end{aligned} \tag{3.9}$$

Now $|1 \pm Re^{i\phi}|^2 = (1 \pm R\cos\phi)^2 + R^2\sin^2\phi$. If $0 < R \le 1/2$, the first term has value at least $1/4 \ge (\sin^2\phi)/4$, whereas if $R > 1/2$, the second term is at least $(\sin^2\phi)/4$. Hence, in either case, $|1 \pm Re^{i\phi}|^2 \ge (\sin^2\phi)/4$, so we have from (3.9)

$$\log|G(Re^{i\phi})| \le c|\sin\phi|^{-a/\pi}, \qquad (0 < R < 1).$$

Since $a/\pi < 1$ by hypothesis, the function on the right is integrable. Hence, we may apply the dominated convergence theorem to pass to the limit (as $R \to 1$) inside the integral sign in (3.8) and obtain

$$\log|G(re^{i\theta})| \le \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1 - r^2}{1 - 2r\cos(\theta - \phi) + r^2}\log|G(e^{i\phi})|\, d\phi. \tag{3.10}$$

Suppose now that $0 < x < 1$ and $re^{i\theta} = h^{-1}(x)$. Then from (3.5) and (3.10),

$$\log|F(x)| \le \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\sin \pi x}{1 + \cos \pi x \sin\phi}\log|F(h(e^{i\phi}))|\, d\phi. \tag{3.11}$$

All that remains is to make a suitable change of variables. Let us consider the upper semicircle $\{e^{i\phi}: 0 < \phi < \pi\}$ first. In this case, we see from (3.1) that

$h(e^{i\phi}) = 1 + iy$ with $y = (-1/\pi)\log(\cot(\phi/2))$ (in particular, y runs from $-\infty$ to ∞ as ϕ runs from 0 to π). It follows that if $t = \tan(\phi/2) = e^{\pi y}$, then

$$\sin\phi = \frac{2t}{1+t^2} = \operatorname{sech}\pi y, \qquad \cos\phi = \frac{1-t^2}{1+t^2} = -\tanh\pi y.$$

Differentiating the first identity, we obtain

$$\cos\phi\, d\phi = -\pi\operatorname{sech}\pi y \tanh\pi y\, dy$$

and substituting the value for $\cos\phi$ from the second identity, we have $d\phi = \pi\operatorname{sech}\pi y\, dy$. Hence,

$$\frac{1}{2\pi}\int_0^{\pi} \frac{\sin\pi x}{1+\cos\pi x\sin\phi}\log|F(h(e^{i\phi}))|\, d\phi$$

$$= \frac{\sin\pi x}{2}\int_{-\infty}^{\infty} \frac{\log|F(1+iy)|}{\cosh\pi y + \cos\pi x}\, dy. \tag{3.12}$$

The argument for the lower semicircle $\{e^{i\phi}: -\pi < \phi < 0\}$ is similar. In this case, we have $h(e^{i\phi}) = iy$ with $y = (-1/\pi)\log(-\cot(\phi/2))$ (so y runs from ∞ to $-\infty$ as ϕ runs from $-\pi$ to 0). Arguing as before, we obtain

$$\sin\phi = -\operatorname{sech}\pi y, \qquad \cos\phi = -\tanh\pi y, \qquad d\phi = -\pi\operatorname{sech}\pi y\, dy,$$

so that

$$\frac{1}{2\pi}\int_{-\pi}^{0} \frac{\sin\pi x}{1+\cos\pi x\sin\phi}\log|F(h(e^{i\phi}))|\, d\phi$$

$$= \frac{\sin\pi x}{2}\int_{-\infty}^{\infty} \frac{\log|F(iy)|}{\cosh\pi y - \cos\pi x}\, dy.$$

When taken together with (3.12), this identity and (3.11) produce the desired conclusion (3.7). ■

If F is bounded in $\bar{\Omega}$, as in the hypothesis of the three-lines theorem (Lemma 2.1), then the terms $\log|F(iy)|$ and $\log|F(1+iy)|$ in (3.7) can be estimated in terms of $\log M_0$ and $\log M_1$, respectively (in the notation of Lemma 2.1). The remaining integrals in (3.7) can then be evaluated explicitly so that the estimate (3.7) produces the logarithmic convexity (2.1) as asserted by the three-lines theorem. Details are provided in Exercise 19.

Definition 3.2. Let (R, μ) and (S, ν) be totally σ-finite measure spaces. Suppose that to each z in the strip $\bar{\Omega}$ there is associated a linear operator T_z, defined on the space of μ-simple functions f on R and taking values in the ν-measurable functions g on S, in such a way that $(T_z f)g$ is ν-integrable on S whenever f and g are simple. Then $\{T_z\}_{z\in\bar{\Omega}}$ is said to be an *analytic family* if $z \to \int_S (T_z f)g\, d\nu$ is analytic in Ω and continuous on $\bar{\Omega}$ whenever f and g are simple.

The analytic family $\{T_z\}$ is of *admissible growth* if there is a constant $a < \pi$ such that, for each simple f and g, the function

$$z \to e^{-a|\operatorname{Im} z|}\log\left|\int_S (T_z f)g\, d\nu\right|$$

is bounded above in $\bar{\Omega}$.

Theorem 3.3 (E. M. Stein). *Let $\{T_z\}_{z\in\bar{\Omega}}$ be an analytic family of operators of admissible growth. Let $1 \le p_0, p_1, q_0, q_1 \le \infty$ and suppose*

$$\|T_{iy}f\|_{q_0} \le M_0(y)\|f\|_{p_0},$$
$$\|T_{1+iy}f\|_{q_1} \le M_1(y)\|f\|_{p_1} \tag{3.13}$$

for all y, $-\infty < y < \infty$, and all simple functions f, where $M_0(y)$ and $M_1(y)$ are independent of f and satisfy

$$\sup_{-\infty<y<\infty} e^{-b|y|}\log M_j(y) < \infty, \qquad (j = 0, 1) \tag{3.14}$$

for some constant $b < \pi$. Suppose $0 \le \theta \le 1$ and let

$$\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \qquad \frac{1}{q_\theta} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}. \tag{3.15}$$

Then

$$\|T_\theta f\|_{q_\theta} \le M_\theta\|f\|_{p_\theta} \tag{3.16}$$

for all simple functions f, where

$$M_\theta = \exp\left\{\frac{\sin\pi\theta}{2}\int_{-\infty}^{\infty}\left[\frac{\log M_0(y)}{\cosh\pi y - \cos\pi\theta} + \frac{\log M_1(y)}{\cosh\pi y + \cos\pi\theta}\right]dy\right\}. \tag{3.17}$$

Proof. Let f and g be simple functions on R and S, respectively, satisfying

$$\|f\|_{p_\theta} = 1 = \|g\|_{q'_\theta}. \tag{3.18}$$

As in the proof of the Riesz-Thorin theorem (Theorem 2.2), write

$$f = \sum_{j=1}^{J} a_j\chi_{A_j}, \qquad g = \sum_{k=1}^{K} b_k\chi_{B_k}$$

and let

$$F(z) = \int_S (T_z f_z) g_z \, dv$$
$$= \sum_{j=1}^{J} \sum_{k=1}^{K} |a_j|^{\alpha(z)/\alpha(\theta)} |b_k|^{(1-\beta(z))/(1-\beta(\theta))} \gamma_{jk}(z),$$

where f_z and g_z are defined by (2.13) and

$$\gamma_{jk}(z) = e^{i(\arg f + \arg g)} \int_S (T_z \chi_{A_j}) \chi_{B_k} \, dv.$$

Since $\{T_z\}$ is an analytic family, the function $F(z)$ is analytic on Ω and continuous on $\bar{\Omega}$. Since $\{T_z\}$ is of admissible growth, F satisfies the remaining hypothesis (3.6) of Lemma 3.1. Furthermore, using Hölder's inequality and (3.13), we have

$$|F(iy)| \le \|T_{iy} f_{iy}\|_{q_0} \|g_{iy}\|_{q_0'} \le M_0(y) \|f_{iy}\|_{p_0} \|g_{iy}\|_{q_0'}.$$

However, exactly as in (2.17) and (2.18), we see that these norms of f_{iy} and g_{iy} are equal to 1, so we conclude that $|F(iy)| \le M_0(y)$. Similarly, $|F(1+iy)| \le M_1(y)$ so now an appeal to Lemma 3.1 gives

$$\left| \int_S (T_\theta f) g \, dv \right| = |F(\theta)| \le M_\theta,$$

with M_θ defined by (3.17). This holds for all simple f and g satisfying (3.18) so we conclude that

$$\|T_\theta f\|_{q_\theta} = \sup\left\{ \left| \int_S (T_\theta f) g \, dv \right| : g \text{ simple}, \|g\|_{q_\theta'} = 1 \right\} \le M_\theta$$

for all simple functions f of norm 1 in L^{p_θ}. Clearly, this is equivalent to (3.16) and so the proof is complete. ■

We shall derive from Theorem 3.3 a version of the Riesz-Thorin theorem in which the underlying measures are allowed to vary (Theorem 3.6 below). The measures in question will be furnished by multiplying the fixed measure μ on the space R by *positive weights* ω, that is, μ-measurable functions ω on R for which $0 < \omega(x) < \infty$ μ-a.e. If ω is a positive weight, the *weighted Lebesgue space* $L^p_\omega(d\mu)$ consists of those μ-measurable functions f on R for which $f\omega \in L^p(d\mu)$. We write

$$\|f\|_{L^p_\omega} = \|f\omega\|_{L^p} = \begin{cases} \left(\int_R |f|^p \omega^p \, d\mu \right)^{1/p}, & (1 \le p < \infty), \\ \operatorname*{ess\,sup}_R |f\omega|, & (p = \infty). \end{cases} \tag{3.19}$$

Note that the sets of measure zero for the measures $d\mu$ and $\omega^p \, d\mu$ are the same so that the weighted space $L^p_\omega(d\mu)$ coincides with the usual Lebesgue space $L^p(\omega^p \, d\mu)$ with respect to the measure $\omega^p \, d\mu$. In particular, L^p_ω is a Banach space under the norm in (3.19).

There may of course be μ-simple functions that do not belong to L^p_ω. Nevertheless, there is the following result.

Lemma 3.4. *Let (R, μ) be a totally σ-finite measure space and let ω be a positive weight. If $1 \le p < \infty$, then the μ-simple functions in $L^p_\omega(d\mu)$ are dense in $L^p_\omega(d\mu)$.*

Proof. Let λ be the measure $\lambda = \omega^p \, d\mu$ so we may view $L^p_\omega(d\mu)$ as the Lebesgue space $L^p(d\lambda)$. Since p is finite, the λ-simple functions are of course dense in $L^p(d\lambda)$. Hence, in order to prove the lemma, it will suffice to show that the set of μ-simple functions in $L^p(d\lambda)$ and the set of λ-simple functions have the same closure in $L^p(d\lambda)$.

One direction is obvious since if A is a set with $\mu(A) < \infty$ and $\chi_A \in L^p(d\lambda)$, then $\lambda(A) < \infty$. Hence, every μ-simple function in $L^p(d\lambda)$ is λ-simple.

Conversely, suppose f is λ-simple. To complete the proof, we need only show that f can be approximated arbitrarily closely in $L^p(d\lambda)$ by μ-simple functions in $L^p(d\lambda)$. Since f is λ-simple, it has a representation

$$f = \sum_{k=1}^{K} a_k \chi_{A_k}$$

with $\lambda(A_k)$ finite for each k. Now (R, μ) is totally σ-finite so there is an increasing sequence of sets R_N, $(N = 1, 2, \ldots)$, each of finite μ-measure, whose union is all of R. For each $N = 1, 2, \ldots$, let

$$f_N = f\chi_{R_N} = \sum_{k=1}^{K} a_k \chi_{A_k \cap R_N}.$$

Clearly, each f_N is μ-simple (because $\mu(R_N) < \infty$) and λ-simple (because f is), hence is a μ-simple function in $L^p(d\lambda)$. Furthermore,

$$\|f - f_N\|^p_{L^p(d\lambda)} = \sum_{k=1}^{K} |a_k|^p \lambda(A_k \cap R_N^c) \to 0, \qquad (N \to \infty)$$

because $\lambda(A_k) < \infty$ and $A_k \cap R_N^c \downarrow \varnothing$ as $N \to \infty$. ■

Lemma 3.5. *Let (R, μ) be a totally σ-finite measure space and let ω_0, ω_1 be positive weights. Suppose $1 \le p_0, p_1 \le \infty$ and $0 \le \theta \le 1$. Define p and ω by*

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \qquad \omega = \omega_0^{1-\theta} \omega_1^\theta.$$

Then

$$L^{p_0}_{\omega_0} \cap L^{p_1}_{\omega_1} \hookrightarrow L^p_\omega \hookrightarrow L^{p_0}_{\omega_0} + L^{p_1}_{\omega_1}. \tag{3.20}$$

Proof. We shall present proofs only for the cases $1 \le p_0 < p < p_1 < \infty$ (so, in particular, $0 < \theta < 1$). The remaining cases require only routine modifications to the proof below.

The first embedding in (3.20) involves only a simple application of Hölder's inequality. Indeed, if f belongs to the intersection on the left of (3.20), then

$$\int_R |f|^p \omega^p \, d\mu = \int_R (|f|\omega_0)^{(1-\theta)p}(|f|\omega_1)^{\theta p} \, d\mu$$
$$\le \left(\int_R (|f|\omega_0)^{p_0} \, d\mu\right)^{(1-\theta)p/p_0}\left(\int_R (|f|\omega_1)^{p_1} \, d\mu\right)^{\theta p/p_1}.$$

Hence,

$$\|f\|_{L^p_\omega} \le \|f\|^{1-\theta}_{L^{p_0}_{\omega_0}} \|f\|^{\theta}_{L^{p_1}_{\omega_1}}, \tag{3.21}$$

from which the desired result follows.

The second embedding in (3.20) is a little more involved because it requires a decomposition of $f \in L^p_\omega$ into a sum of two parts. This, however, is achieved by setting $E = \{|f\omega_0|^{p_0} < |f\omega_1|^{p_1}\}$ and putting $g = f\chi_E$ and $h = f\chi_{E^c}$ so $f = g + h$. Then using the defining property of E we have

$$\int_R |g\omega_0|^{p_0} \, d\mu = \int_E |f\omega_0|^{p_0} \, d\mu = \int_E |f\omega_0|^{p_0(1-\eta)}|f\omega_0|^{p_0\eta} \, d\mu$$
$$\le \int_E |f\omega_0|^{p_0(1-\eta)}|f\omega_1|^{p_1\eta} \, d\mu.$$

This holds for any $\eta > 0$ but we shall choose the value η $(=\theta p/p_1)$ for which $p = (1-\eta)p_0 + \eta p_1$. Then $\omega^p = \omega_0^{(1-\eta)p_0}\omega_1^{\eta p_1}$ and so

$$\int_R |g\omega_0|^{p_0} \, d\mu \le \int_R |f\omega|^p \, d\mu.$$

A similar argument for h gives

$$\int_R |h\omega_1|^{p_1} \, d\mu = \int_{E^c} |f\omega_1|^{p_1} \, d\mu = \int_{E^c} |f\omega_1|^{p_1\eta}|f\omega_1|^{p_1(1-\eta)} \, d\mu$$
$$\le \int_{E^c} |f\omega_1|^{p_1\eta}|f\omega_0|^{p_0(1-\eta)} \, d\mu \le \int_R |f\omega|^p \, d\mu.$$

Combining this with the result for g, we therefore have

$$\|f\|_{L^{p_0}_{\omega_0} + L^{p_1}_{\omega_1}} \le \|g\|_{L^{p_0}_{\omega_0}} + \|h\|_{L^{p_1}_{\omega_1}} \le \|f\|^{p/p_0}_{L^p_\omega} + \|f\|^{p/p_1}_{L^p_\omega}.$$

This inequality is not homogeneous. Applying it to λf instead of f, we obtain

$$\|f\|_{L^{p_0}_{\omega_0} + L^{p_1}_{\omega_1}} \le \lambda^{p/p_0 - 1}\|f\|^{p/p_0}_{L^p_\omega} + \lambda^{p/p_1 - 1}\|f\|^{p/p_1}_{L^p_\omega}$$

and a routine differentiation shows that the right-hand side is minimized by the value

$$\lambda = \left[\frac{1-\theta}{\theta}\right]^{\frac{\theta(1-\theta)}{\theta-\eta}} \|f\|^{-1}_{L^p_\omega}.$$

For this value of λ, the preceding estimate then reduces to

$$\|f\|_{L^{p_0}_{\omega_0} + L^{p_1}_{\omega_1}} \le \left\{\left[\frac{1-\theta}{\theta}\right]^{\theta} + \left[\frac{\theta}{1-\theta}\right]^{1-\theta}\right\}\|f\|_{L^p_\omega}, \tag{3.22}$$

and with this the proof is complete. ■

Theorem 3.6 (E. M. Stein). *Let (R,μ) and (S,ν) be totally σ-finite measure spaces and let T be a linear operator defined on the μ-simple functions on R and taking values in the ν-measurable functions on S. Suppose that u_i, v_i are positive weights on R and S, respectively, and that $1 \le p_i, q_i \le \infty$, $(i = 0, 1)$. Suppose*

$$\|(Tf)v_i\|_{q_i} \le M_i \|fu_i\|_{p_i}, \qquad (i = 0, 1) \tag{3.23}$$

for all μ-simple functions f. Let $0 \le \theta \le 1$ and define

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \qquad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1} \tag{3.24}$$

and

$$u = u_0^{1-\theta}u_1^{\theta}, \qquad v = v_0^{1-\theta}v_1^{\theta}. \tag{3.25}$$

Then, if $p < \infty$, the operator T has a unique extension to a bounded linear operator from L^p_u into L^q_v which satisfies

$$\|(Tf)v\|_q \le M_0^{1-\theta}M_1^{\theta}\|fu\|_p, \tag{3.26}$$

for all f in L^p_u.

Proof. We present the proof only for the case where both p_0 and p_1 are finite, the proof in the remaining cases being similar. The hypotheses (3.23) on T and a standard argument using Lemma 3.4 (with $p = p_0$, $w = u_0$, and with $p = p_1$, $w = u_1$) then show that T has a unique extension to a bounded linear operator from $L^{p_0}_{u_0} + L^{p_1}_{u_1}$ into $L^{q_0}_{v_0} + L^{q_1}_{v_1}$:

$$\|Th\|_{L^{q_0}_{v_0} + L^{q_1}_{v_1}} \le c\|h\|_{L^{p_0}_{u_0} + L^{p_1}_{u_1}}. \tag{3.27}$$

Notice that the theorem follows directly from Lemma 3.4 if $\theta = 0$ or $\theta = 1$ so we may assume here that $0 < \theta < 1$. Let f be any function in L^p_u for which

$$\|fu\|_p = 1. \tag{3.28}$$

Then, by duality, in order to establish (3.26) it will suffice to show that

$$\left|\int_S (Tf)gv\,dv\right| \le M_0^{1-\theta}M_1^{\theta}, \tag{3.29}$$

for all g in $L^{q'}$ with

$$\|g\|_{q'} = 1. \tag{3.30}$$

As usual, because of the monotone convergence theorem, we may restrict attention to functions g that are ν-simple. For the same reason, we may assume henceforth that v_0 and v_1 are ν-simple. With

$$\alpha(z) = \frac{1-z}{p_0} + \frac{z}{p_1}, \qquad \beta(z) = \frac{1-z}{q_0} + \frac{z}{q_1}, \tag{3.31}$$

define H_z in the strip $\bar{\Omega} = \{z = x + iy : 0 \le x \le 1\}$ by

$$\begin{aligned} H_z &= (u_0^{-1}|fu|^{\alpha(z)/\alpha(\theta)})^{1-z}(u_1^{-1}|fu|^{\alpha(z)/\alpha(\theta)})^{z}e^{i\cdot\arg(fu)} \\ &= \frac{\{|fu|^{\alpha(z)/\alpha(\theta)}e^{i\cdot\arg(fu)}\}}{\{u_0^{1-z}u_1^{z}\}}. \end{aligned} \tag{3.32}$$

Then, with $u_x = u_0^{1-x}u_1^{x}$ and $p_x = 1/\alpha(x)$, we see from (3.28), (3.31), and (3.32) that

$$\int_R |(H_z)u_x|^{p_x}\,d\mu = \int_R |fu|^{\alpha(x)p_x/\alpha(\theta)}\,d\mu = \int_R |fu|^{p}\,d\mu = 1.$$

Hence, H_{x+iy} belongs to $L_{u_x}^{p_x}$ and so, by Lemma 3.5, to $L_{u_0}^{p_0} + L_{u_1}^{p_1}$. In particular, it follows from (3.27) that $T(H_z)$ is a well-defined member of $L_{v_0}^{q_0} + L_{v_1}^{q_1}$ for every z in $\bar{\Omega}$. Accordingly, we may define F on $\bar{\Omega}$ by

$$F(z) = \int_S T(H_z)|g|^{(1-\beta(z))/(1-\beta(\theta))}e^{i\cdot\arg g}v_0^{1-z}v_1^{z}\,dv \tag{3.33}$$

and observe that the desired inequality (3.29) may be reformulated as

$$|F(\theta)| \le M_0^{1-\theta}M_1^{\theta}. \tag{3.34}$$

We should like to establish this exactly as in the proof of the Riesz-Thorin theorem, that is, by appealing to the three-lines theorem. The only real difficulty is in establishing the analyticity of F. While g, v_0, and v_1 are simple (so that the corresponding terms in the integral are pure exponentials, hence entire functions), the function H_z is not. An approximation argument is in order.

Let $(\phi_n)_{n=1}^{\infty}$ and $(\psi_n)_{n=1}^{\infty}$ be sequences of μ-simple functions such that

$$0 \le \phi_n \uparrow u_0^{-1}|fu|^{\alpha(0)/\alpha(\theta)}, \qquad 0 \le \psi_n \uparrow u_1^{-1}|fu|^{\alpha(1)/\alpha(\theta)}.$$

As in (3.32) and (3.33), set

$$(H_n)_z = \phi_n^{1-z}\psi_n^{z}e^{i\cdot\arg(fu)} \tag{3.35}$$

and

$$F_n(z) = \int_S T((H_n)_z)|g|^{(1-\beta(z))/(1-\beta(\theta))}e^{i\cdot\arg g}v_0^{1-z}v_1^{z}\,dv. \tag{3.36}$$

Each function F_n is a finite sum of pure exponentials and hence is entire. Furthermore, using Hölder's inequality, the estimate (3.30) for g, the hypothesis (3.23) on T, and the estimate (3.28) for f, we obtain

$$\begin{aligned} |F_n(iy)| &\le \|T((H_n)_{iy})v_0\|_{q_0} \le M_0\|\phi_n u_0\|_{p_0} \\ &\le M_0\|[fu]^{\alpha(0)/\alpha(\theta)}\|_{p_0} = M_0. \end{aligned}$$

Similarly, we have $|F_n(1+iy)| \le M_1$ and so, from the three-lines theorem (Lemma 2.1), we obtain

$$|F_n(x+iy)| \le M_0^{1-x}M_1^{x}, \qquad (n = 1, 2, \ldots) \tag{3.37}$$

for all $z = x + iy$ in $\bar{\Omega}$.

Clearly, the desired result (3.34) will follow from (3.37) if we show that $F_n(\theta) \to F(\theta)$ as $n \to \infty$. To this end, note from (3.32) and (3.35) that $|(H_n)_\theta| \uparrow |H_\theta|$ and so, multiplying by $e^{i\cdot\arg(fu)}$, we have $(H_n)_\theta \to H_\theta$ a.e. Since $|(H_n)_\theta| \le |H_\theta|$ for all n, and H_θ belongs to L_u^p, we may apply the dominated convergence theorem to conclude that $(H_n)_\theta \to H_\theta$ in L_u^p. By Lemma 3.5, we deduce that $(H_n)_\theta \to H_\theta$ in $L_{u_0}^{p_0} + L_{u_1}^{p_1}$ and hence by (3.27) that $T((H_n)_\theta) \to T(H_\theta)$ in $L_{v_0}^{q_0} + L_{v_1}^{q_1}$. Finally, since g is simple we obtain from (3.33) and (3.36) that

$$\begin{aligned} |F_n(\theta) - F(\theta)| &\le \int_S |T[(H_n)_\theta - H_\theta]||g|v\,dv \\ &\le c\|T[(H_n)_\theta - H_\theta]\|_{L_{v_0}^{q_0} + L_{v_1}^{q_1}} \to 0 \end{aligned}$$

as $n \to \infty$. This shows that $F_n(\theta) \to F(\theta)$ as $n \to \infty$ and hence completes the proof. ■

Remarks 3.7. (i) Essentially the same argument shows that $F_n(z) \to F(z)$, $(z \in \bar{\Omega})$. Consequently, since $\{F_n\}$ is uniformly bounded (cf. (3.37)), it constitutes a normal family and so F is necessarily analytic in Ω. The desired inequality $|F(z)| \le M_0^{1-x}M_1^{x}$, $(0 \le x = \operatorname{Re} z \le 1)$, would then follow directly from (3.37).

(ii) The result could also have been derived from Theorem 3.3 by considering the analytic family $T_z f = v_0^{1-z}v_1^{z}T(f/[u_0^{1-z}u_1^{z}])$. Establishing the analyticity, however, would require much the same argument as the one given above in the proof of Theorem 3.6. The full power of Lemma 3.1 would not be needed in this particular application—as in the proof above, the weaker three-lines theorem would suffice.

4. THE MARCINKIEWICZ INTERPOLATION THEOREM

Consider the averaging operator A defined on $L^1(0,1)$ by

$$(Af)(t) = \frac{1}{t}\int_0^t f(s)\,ds, \qquad (0 < t < 1). \tag{4.1}$$

It follows at once from Hardy's inequality (Lemma III.3.9) that A is a bounded linear operator on L^p (into itself) for $1 < p < \infty$. If we wished to establish this result by appealing to the Riesz-Thorin interpolation theorem, we would first need to verify that A is bounded on L^∞ and on L^1. The L^∞-boundedness follows immediately from (4.1), but A is not bounded on L^1, as may easily be seen by considering a decreasing function of the form $f(s) = s^{-1}(\log s)^{-2}$ near the origin and observing that Tf fails to be integrable there. Thus, even in this simple case, the Riesz-Thorin theorem does not apply.

The desired interpolation can still be accomplished, but by a quite different technique introduced by J. Marcinkiewicz in 1939. The Marcinkiewicz interpolation theorem is best formulated in the larger context of a two-parameter family of spaces $L^{p,q}$ (the Lorentz spaces), which generalize the Lebesgue spaces L^p. Our first task therefore will be to define the Lorentz spaces and derive some of their elementary properties.

Definition 4.1. Let (R,μ) be a totally σ-finite measure space and suppose $0 < p,\ q \le \infty$. The *Lorentz space* $L^{p,q} = L^{p,q}(R,\mu)$ consists of all f in $\mathcal{M}_0(R,\mu)$ for which the quantity

$$\|f\|_{p,q} = \begin{cases} \left\{\displaystyle\int_0^\infty [t^{1/p}f^*(t)]^q \frac{dt}{t}\right\}^{1/q}, & (0 < q < \infty), \\ \displaystyle\sup_{0<t<\infty} \{t^{1/p}f^*(t)\}, & (q = \infty), \end{cases} \tag{4.2}$$

is finite.

It is clear from Proposition II.1.8 that the Lorentz space $L^{p,p}$, $(0 < p \le \infty)$, coincides with the Lebesgue space L^p, and

$$\|f\|_{p,p} = \|f\|_p, \qquad (f \in L^p). \tag{4.3}$$

Note also that the space $L^{\infty,q}$, for finite q, is trivial, in the sense that it contains only the zero-function.

The next result shows that, for any fixed p, the Lorentz spaces $L^{p,q}$ increase as the secondary exponent q increases.

Proposition 4.2. *Suppose* $0 < p \le \infty$ *and* $0 < q \le r \le \infty$. *Then*

$$\|f\|_{p,r} \le c\|f\|_{p,q}, \tag{4.4}$$

for all f *in* $\mathcal{M}_0(R,\mu)$, *where* c *is a constant depending only on* p, q, *and* r. *In particular, there is the embedding* $L^{p,q} \hookrightarrow L^{p,r}$.

Proof. We may assume $p < \infty$ and $q < r$ since in the other cases there is nothing to prove. Using the fact that f^* is decreasing, we have

$$\begin{aligned} t^{1/p}f^*(t) &= \left\{\frac{p}{q}\int_0^t [s^{1/p}f^*(t)]^q \frac{ds}{s}\right\}^{1/q} \\ &\le \left\{\frac{p}{q}\int_0^t [s^{1/p}f^*(s)]^q \frac{ds}{s}\right\}^{1/q} \\ &\le \left(\frac{p}{q}\right)^{1/q} \|f\|_{p,q}. \end{aligned}$$

Hence, taking the supremum over all $t > 0$, we obtain

$$\|f\|_{p,\infty} \le \left(\frac{p}{q}\right)^{1/q} \|f\|_{p,q}. \tag{4.5}$$

This establishes (4.4) in the case $r = \infty$.

In the remaining case where $r < \infty$, we have

$$\begin{aligned} \|f\|_{p,r} &= \left\{\int_0^\infty [t^{1/p}f^*(t)]^{r-q+q} \frac{dt}{t}\right\}^{1/r} \\ &\le \|f\|_{p,\infty}^{1-q/r} \|f\|_{p,q}^{q/r}. \end{aligned}$$

When combined with (4.5), this produces (4.4) with the constant $c = (p/q)^{(r-q)/rq}$. ■

Inclusion relations among $L^{p,q}$-spaces, with p varying, are like those for the Lebesgue spaces L^p, in that they depend on the structure of the underlying measure space. The secondary exponent q is not involved. Thus, if $0 < p < r \le \infty$ and $0 < q, s \le \infty$, then $L^{r,s} \hookrightarrow L^{p,q}$ on finite measure spaces and $\ell^{p,q} \hookrightarrow \ell^{r,s}$ on discrete measure spaces (the details are left as an exercise).

We next wish to determine for which values of p and q the Lorentz space $L^{p,q}$ may be regarded as a rearrangement-invariant Banach function space. The functional $f \to \|f\|_{p,q}$ is not always a norm, even when $p, q \geq 1$. However, there is the following result.

Theorem 4.3. *Suppose $1 \leq q \leq p < \infty$ or $p = q = \infty$. Then $(L^{p,q}, \|\cdot\|_{p,q})$ is a rearrangement-invariant Banach function space, with upper and lower Boyd indices both equal to $1/p$.*

Proof. The result is clear when $p = q = 1$ or $p = q = \infty$ since then $L^{p,q}$ reduces to the Lebesgue spaces L^1 and L^∞, respectively. Hence, we may assume that $1 < p < \infty$ and $1 \leq q \leq p$. In that case, with $q' = q/(q-1)$, we have from (4.2) and II.(4.6),

$$\begin{aligned} \|f + g\|_{p,q} &= \left\{ \int_0^\infty [t^{1/p - 1/q}(f+g)^*(t)]^q \, dt \right\}^{1/q} \\ &= \sup\left\{ \int_0^\infty t^{1/p - 1/q}(f+g)^*(t) h^*(t)\, dt : \|h\|_{L^{q'}(0,\infty)} = 1 \right\}. \end{aligned} \tag{4.6}$$

The hypothesis $q \leq p$ implies that $t^{1/p - 1/q}$, hence also $t^{1/p-1/q}h^*(t)$, is decreasing. Hence, since $(f+g)^* \prec f^* + g^*$ (Theorem II.3.4), we may apply Hardy's lemma (Proposition II.3.6) and then Hölder's inequality to obtain

$$\begin{aligned} \int_0^\infty t^{1/p-1/q}(f+g)^*(t)h^*(t)\,dt &\leq \int_0^\infty t^{1/p-1/q} f^*(t)h^*(t)\,dt \\ &\quad + \int_0^\infty t^{1/p-1/q} g^*(t) h^*(t)\,dt \\ &\leq \left\{ \int_0^\infty t^{q/p-1} f^*(t)^q\, dt \right\}^{1/q} \|h\|_{q'} \\ &\quad + \left\{ \int_0^\infty t^{q/p-1} g^*(t)^q \, dt\right\}^{1/q} \|h\|_{q'} \\ &= \|f\|_{p,q} + \|g\|_{p,q}, \end{aligned}$$

since h has norm 1 in $L^{q'}(0,\infty)$. This, together with (4.6), establishes the triangle inequality for $\|\cdot\|_{p,q}$. The remaining properties of a Banach function norm (cf. Definition I.1.1) are easy to verify for $\|\cdot\|_{p,q}$, and the rearrangement-invariance is obvious.

To compute the Boyd indices, let $t > 0$ and consider the dilation operator

$(E_t g)(s) = g(st)$ (cf. Definition III.5.10). If $q < \infty$, a change of variables in (4.2) gives

$$\begin{aligned} \|E_t(f^*)\|_{p,q} &= \left\{ \int_0^\infty [s^{1/p} f^*(st)]^q \frac{ds}{s} \right\}^{1/q} \\ &= \left\{ \int_0^\infty [t^{-1/p} u^{1/p} f^*(u)]^q \frac{du}{u} \right\}^{1/q} \\ &= t^{-1/p} \|f\|_{p,q}. \end{aligned}$$

The operator norm $h(t)$ of $E_{1/t}$ is therefore equal to $t^{1/p}$. It now follows at once from Definition III.5.12 that both Boyd indices are equal to $1/p$. The proof is similar for $q = \infty$. ■

Although the restriction $q \leq p$ in the previous result is necessary, it can be circumvented in the case $p > 1$ by replacing $\|\cdot\|_{p,q}$ with an equivalent functional which is a norm for all $q \geq 1$. The trick is simply to replace f^* with f^{**} in the definition (4.2) of $\|f\|_{p,q}$.

Definition 4.4. Suppose $1 < p \leq \infty$ and $0 < q \leq \infty$. If $f \in \mathcal{M}_0(R, \mu)$, let

$$\|f\|_{(p,q)} = \begin{cases} \left\{ \int_0^\infty [t^{1/p} f^{**}(t)]^q \dfrac{dt}{t} \right\}^{1/q} & (0 < q < \infty) \\ \sup\limits_{0<t<\infty} \{t^{1/p} f^{**}(t)\} & (q = \infty). \end{cases} \tag{4.7}$$

Lemma 4.5. *If $1 < p \leq \infty$ and $1 \leq q \leq \infty$, then*

$$\|f\|_{p,q} \leq \|f\|_{(p,q)} \leq p'\|f\|_{p,q} \tag{4.8}$$

for all $f \in \mathcal{M}_0(R,\mu)$, where $p' = p/(p-1)$. In particular, $L^{p,q}$ consists of all f for which $\|f\|_{(p,q)}$ is finite.

Proof. The first inequality in (4.8) is an immediate consequence of the definitions (4.2) and (4.7), and the fact that $f^* \leq f^{**}$. The second follows directly from Hardy's inequality III.(3.18). ■

Since $f \to f^{**}$ is subadditive (Theorem II.3.4), the triangle inequality for $\|\cdot\|_{(p,q)}$ follows immediately from Minkowski's inequality. Hence, we have the following result.

Theorem 4.6. *If $1 < p < \infty$, $1 \leq q \leq \infty$ or if $p = q = \infty$, then $(L^{p,q}, \|\cdot\|_{(p,q)})$*

is a rearrangement-invariant Banach function space with lower and upper Boyd indices both equal to $1/p$.

It is worth noting that if $1 \leq p < \infty$, the space $L^{p,1}$, equipped with the norm $\|\cdot\|_{p,1}$, coincides with the Lorentz space $\Lambda(L^p)$, and (cf. II.(5.18))

$$\|f\|_{p,1} = \int_0^\infty t^{1/p} f^*(t) \frac{dt}{t} = p\|f\|_{\Lambda(L^p)}. \tag{4.9}$$

On the other hand, if $1 < p \leq \infty$, the space $L^{p,\infty}$, equipped with the modified norm $\|\cdot\|_{(p,\infty)}$, coincides with the Lorentz space $M(L^p)$, and (cf. II.(5.17))

$$\|f\|_{(p,\infty)} = \sup_{t>0} t^{1/p} f^{**}(t) = \|f\|_{M(L^p)}. \tag{4.10}$$

In particular, if $1 < p < \infty$, then Theorem II.5.13 shows that $L^{p,1}$ and $L^{p,\infty}$, when suitably normed, are respectively the smallest and the largest of all rearrangement-invariant spaces having the same fundamental function as L^p. These spaces will play a particularly important role in the weak-type interpolation theory to be discussed later. First, however, let us determine the associate and dual spaces of the Lorentz spaces $L^{p,q}$.

Theorem 4.7. *Let (R, μ) be a resonant measure space and suppose $1 < p < \infty$, $1 \leq q \leq \infty$ (or $p = q = 1$ or $p = q = \infty$). Then the associate space of $L^{p,q}(R, \mu)$ is, up to equivalence of norms, the Lorentz space $L^{p',q'}(R, \mu)$, where $1/p + 1/p' = 1/q + 1/q' = 1$.*

Proof. If $p = q = 1$ or if $p = q = \infty$, the result merely reiterates the well-known fact that the Lebesgue spaces L^1 and L^∞ are mutually associate. Hence, we may assume that $1 < p < \infty$ and $1 \leq q \leq \infty$. By Hölder's inequality, we have

$$\begin{aligned} \left| \int_R fg\, d\mu \right| &\leq \int_0^\infty f^*(t) g^*(t)\, dt \leq \int_0^\infty f^{**}(t) g^{**}(t)\, dt \\ &= \int_0^\infty [t^{1/p} f^{**}(t)][t^{1/p'} g^{**}(t)] \frac{dt}{t} \\ &\leq \|f\|_{(p,q)} \|g\|_{(p',q')}. \end{aligned}$$

Hence, taking the supremum over all f of norm at most 1, we obtain

$$\|g\|_{(L^{p,q})'} \leq \|g\|_{(p',q')}, \qquad (g \in L^{p',q'}). \tag{4.11}$$

To establish an inequality of this type in the opposite direction, it suffices by the Luxemburg representation theorem (Theorem II.4.10) to do so for the measure space $(\mathbf{R}^+, m)$ and functions g on $\mathbf{R}^+$ for which $g = g^*$. Suppose first that $1 < q < \infty$ and that $g = g^*$ is a simple function. Let

$$f(t) = \int_{t/2}^\infty \phi(s) \frac{ds}{s}, \qquad (t > 0),$$

where

$$\phi(s) = s^{q'/p' - 1} g^*(s)^{q' - 1}, \qquad (s > 0).$$

Then $f = f^*$ and, by (4.2),

$$\begin{aligned} \|g\|_{p',q'}^{q'} &= \int_0^\infty \phi(s) g^*(s)\, ds \\ &\leq c \int_0^\infty \left\{ \int_{s/2}^s \phi(t) \frac{dt}{t} \right\} g^*(s)\, ds \\ &\leq c \int_0^\infty f^*(s) g^*(s)\, ds \\ &\leq c \|f\|_{L^{(p,q)}} \|g\|_{(L^{p,q})'}, \end{aligned} \tag{4.12}$$

where c is a constant depending only on p and q. However, using Lemma 4.5 and Hardy's inequality (Lemma III.3.9), we have

$$\begin{aligned} \|f\|_{(p,q)} \leq c\|f\|_{p,q} &= c \left\{ \int_0^\infty \left[t^{1/p} \int_{t/2}^\infty \phi(s) \frac{ds}{s} \right]^q \frac{dt}{t} \right\}^{1/q} \\ &\leq c \left\{ \int_0^\infty [t^{1/p} \phi(t)]^q \frac{dt}{t} \right\}^{1/q} \\ &= c \left\{ \int_0^\infty [t^{1/p'} g^*(t)]^{q'} \frac{dt}{t} \right\}^{1/q} = c\|g\|_{p',q'}^{q'/q}. \end{aligned}$$

Combining this with (4.12), we obtain

$$\|g\|_{p',q'} \leq \|g\|_{(L^{p,q})'}, \tag{4.13}$$

with c depending only on p and q. This holds for all simple $g\,(=g^*)$ and hence, by the Fatou property and rearrangement-invariance, for all g in $(L^{p,q})'$. Hence, (4.11) and (4.13) together complete the proof for the case $1 < q < \infty$. The proof for $q = 1$ and $q = \infty$ is similar. ∎

Corollary 4.8. *If (R, μ) is resonant and if $1 < p < \infty$, $1 \leq q < \infty$ (or $p = q = 1$), then the dual space $(L^{p,q})^*$ of $L^{p,q}(R, \mu)$ can be identified with the associate space $L^{p',q'}$ (with equivalent norms).*

Proof. For $p = q = 1$, the result merely asserts that L^∞ is identifiable as the dual of L^1. In the remaining case where $1 < p < \infty$, it is easy to see directly from (4.7) that $L^{p,q}$ has absolutely continuous norm when $1 \le q < \infty$. Hence, the desired result follows at once from the preceding theorem and Corollary I.4.3. ∎

With these preliminaries disposed of, let us now return to the main topic of weak-type interpolation. Recall from Definition III.5.1 that if $1 \le p_0 < p_1 \le \infty$, $1 \le q_0, q_1 \le \infty$ with $q_0 \ne q_1$, and if m denotes the slope

$$m = \frac{\dfrac{1}{q_0} - \dfrac{1}{q_1}}{\dfrac{1}{p_0} - \dfrac{1}{p_1}} \tag{4.14}$$

of the interpolation segment $\sigma = [(1/p_0, 1/q_0), (1/p_1, 1/q_1)]$, then the Calderón operator S_σ is defined by

$$(S_\sigma f)(t) = t^{-1/q_0} \int_0^{t^m} s^{1/p_0} f(s) \frac{ds}{s} + t^{-1/q_1} \int_{t^m}^\infty s^{1/p_1} f(s) \frac{ds}{s} \tag{4.15}$$

for each $t > 0$. Furthermore (cf. Definition III.5.4), an operator T is of joint weak type $(p_0, q_0; p_1, q_1)$ if

$$(Tf)^*(t) \le M S_\sigma(f^*)(t), \qquad (t > 0), \tag{4.16}$$

for all f for which the right-hand side is finite, the least value of M being the weak-type $(p_0, q_0; p_1, q_1)$ norm of T.

The Lorentz space structure will enable us to separate this joint weak-type property into two individual weak-type conditions for the pairs (p_0, q_0) and (p_1, q_1) (at least when p_0 and p_1 are finite).

Definition 4.9. Let (R, μ) and (S, ν) be totally σ-finite measure spaces and suppose $1 \le p < \infty$ and $1 \le q \le \infty$. Let T be an operator defined on $L^{p,1}(R, \mu)$ and taking values in $\mathscr{M}_0(S, \nu)$. Then T is said to be of *weak type* (p, q) if it is a bounded operator from $L^{p,1}(R, \mu)$ into $L^{q,\infty}(S, \nu)$, that is, if there is a constant M such that

$$\|Tf\|_{q,\infty} \le M \|f\|_{p,1} \tag{4.17}$$

for all f in $L^{p,1}$. The least constant M for which (4.17) holds is called the *weak-type* (p, q) *norm* of T.

Explicitly, in terms of the decreasing rearrangement, the estimate (4.17)

asserts that

$$(Tf)^*(t) \le M t^{-1/q} \|f\|_{p,1}, \qquad (t > 0),$$

or, in terms of the distribution function, that

$$\nu\{|Tf| > \lambda\} \le [M \lambda^{-1} \|f\|_{p,1}]^q, \qquad (\lambda > 0).$$

If p is finite, then a linear operator T of strong type (p, q), defined initially on the simple functions, has a unique linear extension to a bounded operator from L^p to L^q. Since Proposition 4.2 shows that $L^{p,1} \hookrightarrow L^p$ and $L^q \hookrightarrow L^{q,\infty}$, it follows that T is also of weak type (p, q).

Lemma 4.10. *Suppose* $1 \le p_0 < p_1 < \infty$ *and* $1 \le q_0, q_1 \le \infty$ *with* $q_0 \ne q_1$. *If* S_σ *is defined as in* (4.15), *then*

$$t^{1/q_i} S_\sigma(f^*)(t) \le \int_0^\infty s^{1/p_i} f^*(s) \frac{ds}{s}, \qquad (t > 0), \tag{4.18}$$

for all f *in* $L^{p_i,1}(\mathbf{R}^+)$, $(i = 0, 1)$. *In particular,* S_σ *is of weak types* (p_0, q_0) *and* (p_1, q_1).

Proof. From (4.15), we have

$$t^{1/q_0} S_\sigma(f^*)(t) = \int_0^{t^m} s^{1/p_0} f^*(s) \frac{ds}{s} + t^{1/q_0 - 1/q_1} \int_{t^m}^\infty s^{1/p_1} f^*(s) \frac{ds}{s}, \tag{4.19}$$

where m is given by (4.14). The hypothesis $p_0 < p_1$ implies that

$$t^{1/q_0 - 1/q_1} = (t^m)^{1/p_0 - 1/p_1} \le s^{1/p_0 - 1/p_1}$$

whenever $s \ge t^m$. Hence, applying this estimate to the second term on the right of (4.19), then combining the two integrals, we obtain (4.18) in the case $i = 0$. The proof for $i = 1$ is similar. Since III.(5.7) shows that $[S_\sigma(f)]^* \le S_\sigma(f^*)$, the weak-type assertions follow at once from (4.18). ∎

Theorem 4.11. *Suppose* $1 \le p_0 < p_1 < \infty$ *and* $1 \le q_0, q_1 \le \infty$ *with* $q_0 \ne q_1$. *Then a quasilinear operator* T *is of joint weak type* $(p_0, q_0; p_1, q_1)$ *if and only if it is of weak types* (p_0, q_0) *and* (p_1, q_1).

Proof. Suppose T is of weak types (p_0, q_0) and (p_1, q_1). Let $f \in (L^{p_0,1} + L^{p_1,1})(R, \mu)$ and fix $t > 0$. With m given by (4.14), define f_0 and f_1 on R by

$$\begin{aligned} f_1(x) &= \min[|f(x)|, f^*(t^m)] \cdot \operatorname{sgn} f(x) \\ f_0(x) &= f(x) - f_1(x) = [|f(x)| - f^*(t^m)]^+ \cdot \operatorname{sgn} f(x). \end{aligned} \tag{4.20}$$

Then

$$f_1^*(s) = \min[f^*(s), f^*(t^m)], \qquad (s > 0)$$

and so, by (4.2),

$$\|f_1\|_{p_1,1} = p_1 t^{m/p_1} f^*(t^m) + \int_{t^m}^{\infty} s^{1/p_1} f^*(s) \frac{ds}{s}. \tag{4.21}$$

Similarly,

$$f_0^*(s) = [f^*(s) - f^*(t^m)]^+, \qquad (s > 0)$$

and

$$\|f_0\|_{p_0,1} = \int_0^{t^m} s^{1/p_0} f^*(s) \frac{ds}{s} - p_0 t^{m/p_0} f^*(t^m). \tag{4.22}$$

Now suppose that T is quasilinear with constant k (cf. III.(5.8)). Since $f = f_0 + f_1$, we have from Proposition II.1.7,

$$\begin{aligned}(Tf)^*(t) &\le [k(|Tf_0| + |Tf_1|)]^*(t)\\ &\le k\left[(Tf_0)^*\left(\frac{t}{2}\right) + (Tf_1)^*\left(\frac{t}{2}\right)\right].\end{aligned}$$

Furthermore, by (4.17), the weak-type hypotheses on T give

$$(Tf_i)^*\left(\frac{t}{2}\right) \le \left(\frac{t}{2}\right)^{-1/q_i} M_i \|f_i\|_{p_i,1}, \qquad (i = 0, 1). \tag{4.23}$$

Combining these estimates, we obtain

$$(Tf)^*(t) \le c\left\{\frac{t^{-1/q_0}}{p_0}\|f_0\|_{p_0,1} + \frac{t^{-1/q_1}}{p_1}\|f_1\|_{p_1,1}\right\} \tag{4.24}$$

with $c = k \cdot \max_i p_i M_i 2^{1/q_i}$. Incorporating (4.21), (4.22) into (4.24) and observing that the terms in $f^*(t^m)$ cancel, we find that the right-hand side of (4.24) is majorized by $cS_\sigma(f^*)(t)$. Hence, T is of joint weak type $(p_0, q_0; p_1, q_1)$.

The converse is an immediate consequence of (4.16) and Lemma 4.10. ■

Remarks 4.12. (a) If T is of weak type $(p_0, q_0; \infty, q_1)$, with $p_0 < \infty$, we can conclude from the argument above only that T is of weak type (p_0, q_0).
(b) A simple modification of the proof above shows that if T is of weak type (p_0, q_0), $(p_0 < \infty)$, and *strong* type (∞, q_1), then T is of joint weak type $(p_0, q_0; \infty, q_1)$. The relation (4.21) is replaced by the property that $\|f_1\|_\infty = f^*(t^m)$ and then the strong-type (∞, q_1) hypothesis is invoked in the form

$$T: L^\infty \to L^{q_1} \hookrightarrow L^{q_1,\infty},$$

so that the case $i = 1$ of (4.23) may be replaced by the estimate

$$(Tf_1)^*\left(\frac{t}{2}\right) \le \left(\frac{t}{2}\right)^{-1/q_1} M_1 \|f_1\|_\infty.$$

Now we can present the main result of this section, the Marcinkiewicz interpolation theorem. The following formulation in terms of Lorentz spaces is due to A. P. Calderón.

Theorem 4.13 (Marcinkiewicz interpolation theorem). *Suppose* $1 \le p_0 < p_1 < \infty$ *and* $1 \le q_0, q_1 \le \infty$ *with* $q_0 \ne q_1$. *Let* $0 < \theta < 1$ *and define* p *and* q *by*

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \qquad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}. \tag{4.25}$$

Let T *be a quasilinear operator defined on* $(L^{p_0,1} + L^{p_1,1})(R, \mu)$ *and taking values in* $\mathcal{M}_0(S, \nu)$, *where* (R, μ) *and* (S, ν) *are totally* σ*-finite measure spaces. Suppose* T *is of weak types* (p_0, q_0) *and* (p_1, q_1), *with respective weak-type norms* M_0 *and* M_1. *If* $1 \le r \le \infty$, *then*

$$T: L^{p,r} \to L^{q,r}. \tag{4.26}$$

Precisely, there is a constant c, *depending only on* p_0, q_0, p_1, q_1, *and* r, *such that, for all* f *in* $L^{p,r}$,

$$\|Tf\|_{q,r} \le \frac{c}{\theta(1-\theta)} \max(M_0, M_1) \|f\|_{p,r}. \tag{4.27}$$

Proof. Note first that with m defined by (4.14) and p, q by (4.25), we have

$$\frac{1}{m}\left[\frac{1}{q} - \frac{1}{q_0}\right] = \frac{1}{p} - \frac{1}{p_0}, \qquad \frac{1}{m}\left[\frac{1}{q} - \frac{1}{q_1}\right] = \frac{1}{p} - \frac{1}{p_1}. \tag{4.28}$$

Because of Theorem 4.11, we know that T is of joint weak type $(p_0, q_0; p_1, q_1)$ with weak-type norm $M \le c \cdot \max(M_0, M_1)$, where c depends only on p_0, q_0, p_1, q_1. Hence, using (4.16) and (4.2), we have, in the case where $r < \infty$,

$$\|Tf\|_{q,r} \le M\left\{\int_0^\infty [t^{1/q} S_\sigma(f^*)(t)]^r \frac{dt}{t}\right\}^{1/r},$$

with S_σ defined as in (4.15). Applying Minkowski's inequality, we obtain

$$\begin{aligned}\|Tf\|_{q,r} \le M\Bigg[&\left\{\int_0^\infty \left[t^{1/q - 1/q_0} \int_0^{t^m} s^{1/p_0} f^*(s) \frac{ds}{s}\right]^r \frac{dt}{t}\right\}^{1/r}\\ &+ \left\{\int_0^\infty \left[t^{1/q - 1/q_1} \int_{t^m}^\infty s^{1/p_1} f^*(s) \frac{ds}{s}\right]^r \frac{dt}{t}\right\}^{1/r}\Bigg].\end{aligned}$$

Making the change of variables $t^m = u$ in each of the integrals and using the relations (4.28), we therefore have

$$\|Tf\|_{q,r} \le M|m|^{-1/r}\left[\left\{\int_0^\infty\left[u^{1/p-1/p_0}\int_0^u s^{1/p_0}f^*(s)\frac{ds}{s}\right]^r\frac{du}{u}\right\}^{1/r} + \left\{\int_0^\infty\left[u^{1/p-1/p_1}\int_u^\infty s^{1/p_1}f^*(s)\frac{ds}{s}\right]^r\frac{du}{u}\right\}^{1/r}\right].$$

When Hardy's inequalities III.(3.18) and III.(3.19) are applied to the first and second terms, respectively, the resulting estimate reduces to

$$\begin{aligned}\|Tf\|_{q,r} &\le M|m|^{-1/r}\left[c_1\left\{\int_0^\infty[u^{1/p}f^*(u)]^r\frac{du}{u}\right\}^{1/r} + c_2\left\{\int_0^\infty[u^{1/p}f^*(u)]^r\frac{du}{u}\right\}^{1/r}\right] \\ &= M|m|^{-1/r}(c_1+c_2)\|f\|_{p,r},\end{aligned}$$

with

$$\frac{1}{c_1} = \frac{1}{p_0} - \frac{1}{p} = \theta\left(\frac{1}{p_0} - \frac{1}{p_1}\right), \qquad \frac{1}{c_2} = \frac{1}{p} - \frac{1}{p_1} = (1-\theta)\left(\frac{1}{p_0} - \frac{1}{p_1}\right).$$

This establishes (4.27) for r finite. The proof for $r = \infty$ is similar. ∎

The original theorem of Marcinkiewicz, formulated for Lebesgue spaces rather than the more general Lorentz spaces, is essentially as follows.

Corollary 4.14. *With parameters as above, suppose in addition that $p_i \le q_i$, $(i = 0, 1)$. If T is of weak types (p_0, q_0) and (p_1, q_1), with respective norms M_0 and M_1, then*

$$T: L^p \to L^q. \tag{4.29}$$

Precisely, there is a constant $c = c(p_0, q_0, p_1, q_1)$ such that

$$\|Tf\|_q \le \frac{c}{\theta(1-\theta)}\max(M_0, M_1)\|f\|_p \tag{4.30}$$

for all f in L^p.

Proof. Since $p_i \le q_i$, $(i = 0, 1)$, it follows from (4.25) that $p \le q$. Taking $r = p$ in (4.26), we conclude that

$$T: L^{p,p} = L^p \to L^{q,p}.$$

However, since $p \le q$, Proposition 4.2 shows that $L^{q,p}$ is continuously embedded in $L^{q,q} = L^q$, and from this the result follows. ∎

Remarks 4.15 (a) Theorem 4.13 holds also in the case $p_1 = \infty$ provided the weak-type hypothesis "weak type (p_1, q_1)" is replaced by "strong type (p_1, q_1)". In that case (cf. Remark (4.12)(b)), the operator is of joint weak type $(p_o, q_o; \infty, q_1)$ and the proof of Theorem 4.13 carries over almost verbatim. As an example, consider the Hardy-Littlewood maximal operator M (Definition III.3.1), which is of weak type $(1, 1)$ (Theorem III.3.3) and of strong type (∞, ∞). Corollary 4.14 therefore establishes the L^p-boundedness $(1 < p < \infty)$ as asserted by the Hardy-Littlewood maximal theorem (Theorem III.3.10).
(b) More generally, Theorem 4.13 holds also in the case $p_1 = \infty$ when the separate weak-type hypotheses are replaced by the joint weak-type $(p_0, q_0; p_1, q_1)$ hypothesis. This is clear from an examination of the proof, which depends solely on the properties of the associated S_σ-operator. In this more general form, Theorem 4.13 may be regarded as a specialization to Lorentz $L^{p,q}$-spaces of the general Marcinkiewicz-type theorem for rearrangement-invariant spaces established in the previous chapter (Theorem III.5.7). The advantage of using the joint weak-type condition rather than the pair of individual weak-type conditions can be seen in connection with operators such as the Hilbert transform H (Definition III.4.1). This operator is of joint weak type $(1, 1; \infty, \infty)$ so, as we remarked above, the more general version of Corollary 4.14 for the joint weak-type operators can be applied to establish the L^p-boundedness of H for $1 < p < \infty$, as asserted by the M. Riesz theorem (Theorem III.4.9). The Hilbert transform is, of course, of weak type $(1, 1)$ (Theorem III.4.9) but not of strong type (∞, ∞), hence cannot be interpolated directly by the results above (one remedy is to use weak type $(1, 1)$ together with the strong-type $(2, 2)$ property of the Hilbert transform, so that Corollary 4.14 then yields the L^p-boundedness for $1 < p < 2$; a duality argument then produces the result for $2 < p < \infty$). It is clear that a meaningful notion of "weak type (∞, ∞)" would be desirable. We shall have more to say about this in Section V.7.
(c) The proof of the Marcinkiewicz interpolation theorem presented above for real spaces carries over in the obvious way for complex spaces, modulo an adjustment to the constant c in (4.27). The L^p-version (Corollary 4.14), producing a strong-type conclusion from weak-type hypotheses, is often more useful in the applications than the Riesz-Thorin theorem (Theorem 2.2). The additional restriction $p_i \le q_i$, $(i = 0, 1)$ in Corollary 4.14 rarely limits the applicability of the Marcinkiewicz theorem in practice.
(d) By a simple modification of the Calderón operator, the quantity

$\max(M_0, M_1)$ in (4.27) and (4.30) can be replaced by $M_0^{1-\theta}M_1^{\theta}$ (cf. Theorem V.2.5).

Recall from Section 2 that the Fourier transform $f \to \hat{f}$ is of strong types $(1, \infty)$ and $(2, 2)$, hence, by the Riesz-Thorin theorem, also of strong type (p, p'), whenever $1 < p < 2$ and $1/p + 1/p' = 1$. This is the classical Hausdorff-Young theorem (Theorem 2.6). Since the Fourier transform is *a fortiori* of weak types $(1, \infty)$ and $(2, 2)$, we may instead apply the Marcinkiewicz interpolation theorem (Theorem 4.13 with $r = p$) to conclude that the Fourier transform gives a bounded map of L^p into $L^{p',p}$. The latter space is properly contained in $L^{p'}$ since $p < p'$ when $1 < p < 2$. This sharper version of the Hausdorff-Young theorem was first established (by other means) by Paley.

Theorem 4.16 (R. E. A. C. Paley). *Suppose $1 < p < 2$ and $1/p + 1/p' = 1$.*
(a) *If $f \in L^p(\mathbf{T})$, then $\hat{f} \in \ell^{p',p}(\mathbf{Z})$:*

$$\left\{\sum_{n=1}^{\infty} [n^{1/p'}c_n^*]^p \frac{1}{n}\right\}^{1/p} \le c_p \left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(e^{it})|^p\,dt\right\}^{1/p}, \tag{4.31}$$

where $(c_n^*)_{n=1}^{\infty}$ is the decreasing rearrangement of the sequence $(\hat{f}(n))_{n=-\infty}^{\infty}$ of Fourier coefficients of f.
(b) *If $\mathbf{c} = (c_n)_{n=-\infty}^{\infty} \in \ell^p(\mathbf{Z})$, then the function f on $\mathbf{T}$ with $\hat{f} = \mathbf{c}$ belongs to $L^{p',p}(\mathbf{T})$ and*

$$\left\{\int_0^1 [t^{1/p'}f^*(t)]^p \frac{dt}{t}\right\}^{1/p} \le c_p \left\{\sum_{n=-\infty}^{\infty} |c_n|^p\right\}^{1/p}. \tag{4.32}$$

We conclude this section with one further application of the Marcinkiewicz interpolation theorem.

Definition 4.17. If $0 < \alpha < n$, the *fractional integral $I_\alpha f$ of order α* of a measurable function f on $\mathbf{R}^n$ is defined by

$$(I_\alpha f)(x) = \int_{\mathbf{R}^n} f(y)\phi(x-y)\,dy, \qquad (x \in \mathbf{R}^n), \tag{4.33}$$

where

$$\phi(y) = c(\alpha)|y|^{\alpha-n}, \qquad c(\alpha) = \Gamma\left(\frac{n-\alpha}{2}\right)\left[\pi^{n/2}2^{\alpha}\Gamma\left(\frac{\alpha}{2}\right)\right]^{-1}. \tag{4.34}$$

Theorem 4.18. (a) *The fractional integral operator I_α, $(0 < \alpha < n)$, is of weak types $(1, n/(n-\alpha))$ and $(n/\alpha, \infty)$.*

(b) (*Hardy-Littlewood-Sobolev theorem of fractional integration*) *Suppose $0 < \alpha < n$ and $1 < p < q < \infty$ with $1/q = 1/p - \alpha/n$. If $1 \le r \le \infty$, then*

$$\|I_\alpha f\|_{q,r} \le c\|f\|_{p,r}, \qquad (f \in L^{p,r}) \tag{4.35}$$

so that I_α is a bounded linear operator from $L^{p,r}$ into $L^{q,r}$. In particular, I_α is bounded from L^p to L^q.

Proof. The distribution function m_ϕ of ϕ is given by

$$m_\phi(\lambda) = c_n\left[\frac{\lambda}{c(\alpha)}\right]^{n/(\alpha-n)}, \qquad (0 < \lambda < \infty),$$

where c_n is the volume of the unit ball in $\mathbf{R}^n$. Hence, the decreasing rearrangement ϕ^* of ϕ has the form

$$\phi^*(t) = ct^{\alpha/n-1}, \qquad (0 < t < \infty), \tag{4.36}$$

where c depends only on α and n. Notice in particular that

$$\phi \in L^{n/(n-\alpha),\infty}. \tag{4.37}$$

For each $f \in L^1(\mathbf{R}^n)$, the convolution operator

$$(Tg)(x) = (f * g)(x) = \int_{\mathbf{R}^n} f(x-y)g(y)\,dy, \qquad (x \in \mathbf{R}^n)$$

satisfies

$$\|Tg\|_1 \le M\|g\|_1, \qquad \|Tg\|_\infty \le M\|g\|_\infty$$

with $M = \|f\|_1$. Hence, by Theorem III.2.2, we also have

$$\|Tg\|_X \le \|f\|_1\|g\|_X, \qquad (g \in X)$$

for any rearrangement-invariant space X on $\mathbf{R}^n$. By (4.37), we may select $g = \phi$ and $X = L^{n/(n-\alpha),\infty}$ to obtain

$$\|I_\alpha f\|_{n/(n-\alpha),\infty} = \|Tg\|_X \le \|\phi\|_{n/(n-\alpha),\infty}\|f\|_1, \qquad (f \in L^1),$$

which shows that I_α is of weak type $(1, n/(n-\alpha))$.

Next, since the decreasing rearrangement of any translate of ϕ is equal to ϕ^* itself, we may apply the Hardy-Littlewood inequality II.(2.3) to (4.33) and use (4.36) to obtain

$$|(I_\alpha f)(x)| \le c\int_0^\infty f^*(t)t^{\alpha/n-1}\,dt = c\|f\|_{n/\alpha,1}, \qquad (f \in L^{n/\alpha,1}).$$

This shows that I_α is of weak type $(n/\alpha, \infty)$ and hence completes the proof of part (a).

The assertion (4.35) now follows from the weak-type estimates in part (a) and the Marcinkiewicz interpolation theorem (Theorem 4.13). The boundedness from L^p to L^q follows as in Corollary 4.14. ■

5. RESTRICTED WEAK TYPE AND A.E. CONVERGENCE

With the main interpolation theorems now in place, it is worth pausing here to take a closer look at the notion of "weak type" itself. An operator T is of *weak type* (p,q) if it is bounded from $L^{p,1}$ to $L^{q,\infty}$, that is, if

$$\|Tf\|_{q,\infty} \le c\|f\|_{p,1}, \qquad (f \in L^{p,1}). \tag{5.1}$$

Establishing such a result for a given operator T requires verification of (5.1) for all f in a dense subset of $L^{p,1}$—the simple functions, for example. The discussion below of the notion of *restricted weak type* will show, however, that in many cases the estimate (5.1) need only be verified for characteristic functions $f = \chi_E$ of arbitrary sets of finite measure. This can make the Marcinkiewicz interpolation theorem much easier to apply in practice.

We shall also consider connections between the a.e. convergence of a sequence of operators T_n and the size of the corresponding maximal operator $\mathscr{T}f = \sup_n |T_n f|$. We have already seen how the weak-type $(1,1)$ property of the Hardy-Littlewood maximal operator implies the a.e. convergence of an associated sequence of averages (the Lebesgue differentiation theorem). Similarly, a weak-type property of the maximal Hilbert transform is responsible for the a.e. convergence of the principal-value integral defining the Hilbert transform. We shall be interested here in a converse—a remarkable result due to E. M. Stein which shows, under suitable conditions, that the a.e. finiteness of the maximal operator $\mathscr{T}$ implies that $\mathscr{T}$ satisfies an appropriate weak-type condition.

Maximal operators, although nonlinear, are incorporated into the discussion below as examples of sublinear operators that assume only nonnegative values. Recalling Definition III.5.3, we see that such operators T, defined on a subspace D of $\mathscr{M}_0(R,\mu)$ and taking values in the ν-measurable functions on S, satisfy

$$0 \le T(f+g) \le Tf + Tg \quad \nu\text{-a.e.}; \qquad T(\lambda f) = |\lambda| T(f) \quad \nu\text{-a.e.}, \tag{5.2}$$

for all $f, g \in D$ and all constants λ. We shall refer to them as *nonnegative sublinear operators*[1].

[1] This is not to be confused with the notion of a positive operator, which requires $Tf \ge 0$ merely when $f \ge 0$.

The following result is well known for linear operators. With the property

$$|Tf - Tg| \le |T(f-g)| = T(f-g), \qquad (f,g \in D), \tag{5.3}$$

which is an immediate consequence of (5.2), essentially the same proof establishes the result also for nonnegative sublinear operators.

Lemma 5.1. *Let X and Y be rearrangement-invariant Banach function spaces over resonant measure spaces (R,μ) and (S,ν), respectively. Let T be a linear (respectively, nonnegative sublinear) operator defined on a dense linear subspace D of X and taking values in Y. If*

$$\|Tf\|_Y \le M\|f\|_X \tag{5.4}$$

for all $f \in D$, then T has a unique extension to a linear (respectively, nonnegative sublinear) operator from X to Y, for which (5.4) holds for all $f \in X$.

Definition 5.2. Suppose $1 \le p < \infty$ and $1 \le q \le \infty$. Let T be an operator defined on the μ-simple functions on R and taking values in the ν-measurable functions on S. Then T is said to be of *restricted weak type* (p,q) if there is a constant M such that

$$t^{1/q}(T\chi_E)^*(t) \le M\mu(E)^{1/p}, \qquad (t > 0), \tag{5.5}$$

for all μ-measurable subsets E of R.

The estimate (5.5) can also be formulated in terms of the distribution function as follows:

$$\nu\{|T\chi_E| > \lambda\} \le \left(\frac{M}{\lambda}\mu(E)^{1/p}\right)^q, \qquad (\lambda > 0). \tag{5.6}$$

Of course, (5.5) is nothing more than the weak-type inequality (5.1) with $f = \chi_E$. Hence, every operator of weak type (p,q) is also of restricted weak type (p,q). The following result provides a partial converse.

Theorem 5.3. *Suppose $1 \le p < \infty$ and $1 < q \le \infty$. Let T be a linear (respectively, nonnegative sublinear) operator defined on the simple functions and suppose T is of restricted weak type (p,q). Then T has a unique extension to a linear (respectively, nonnegative sublinear) operator of weak type (p,q).*

Proof. Suppose first that f is a nonnegative simple function and write

$$f = \sum_{j=1}^{J} a_j \chi_{A_j},$$

where $A_1 \supset A_2 \supset \cdots \supset A_J$ and $a_j > 0$, $(j = 1, 2, \ldots, J)$. Then, with $t_j = \mu(A_j)$, we have

$$f^* = \sum_{j=1}^{J} a_j \chi_{[0,t_j)}.$$

Since $q > 1$, Theorem 4.6 shows that $\|\cdot\|_{(q,\infty)}$ is indeed a norm. Hence, using the triangle inequality in conjunction with the (sub-)linearity of T, we have

$$\|Tf\|_{(q,\infty)} \leq \left\| \sum_{j=1}^{J} T(a_j \chi_{A_j}) \right\|_{(q,\infty)} \leq \sum_{j=1}^{J} a_j \|T\chi_{A_j}\|_{(q,\infty)}.$$

Since T is of restricted weak type (p, q), we may therefore apply (5.5) and (4.8) to obtain

$$\|Tf\|_{(q,\infty)} \leq \sum_{j=1}^{J} a_j M q' t_j^{1/p} \leq \frac{Mq'}{p} \|f\|_{p,1}.$$

A similar estimate

$$\|Tf\|_{(q,\infty)} \leq cM\|f\|_{p,1}, \tag{5.7}$$

with c depending only on p and q, can now be obtained for arbitrary simple functions f by the usual device of decomposing f into its real and imaginary parts, then each of these into its positive and negative parts. The simple functions, however, are dense in $L^{p,1}$ so Lemma 5.1 shows that T has a unique extension to all of $L^{p,1}$ with (5.7) holding for all f in that space. The extension is thus of weak type (p, q) and the proof is complete. ■

The next lemma may be regarded as a crude interpolation theorem for restricted weak-type operators.

Lemma 5.4. *Suppose $1 \leq p_0 < p_1 < \infty$ and $1 \leq q_0, q_1 \leq \infty$. Suppose $0 < \theta < 1$ and set*

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \qquad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}. \tag{5.8}$$

If a sublinear operator T, defined on the simple functions, is of restricted weak types (p_j, q_j), $(j = 0, 1)$, then T is of restricted weak type (p, q).

Proof. The hypotheses, in the format of (5.5), combine to give

$$(T\chi_E)^*(t) \leq \max_{j=0,1}(M_j) \cdot \min_{j=0,1}\left(\frac{\mu(E)^{1/p_j}}{t^{1/q_j}}\right), \qquad (t > 0),$$

for all μ-measurable subsets E of R. The result will therefore be established if we show that the function Φ given by

$$\Phi(s,t) = \min_{j=0,1}\left(\frac{s^{1/p_j}}{t^{1/q_j}}\right) = \begin{cases} \dfrac{s^{1/p_0}}{t^{1/q_0}}, & 0 < s \leq t^m \\[2ex] \dfrac{s^{1/p_1}}{t^{1/q_1}}, & s \geq t^m, \end{cases}$$

where $m = (1/q_0 - 1/q_1)/(1/p_0 - 1/p_1)$, satisfies the inequality

$$\Phi(s,t) \leq \frac{s^{1/p}}{t^{1/q}}, \qquad (s, t > 0).$$

But this estimate follows at once from the elementary identities

$$\frac{1}{m}\left(\frac{1}{q} - \frac{1}{q_0}\right) = \frac{1}{p} - \frac{1}{p_0}, \qquad \frac{1}{m}\left(\frac{1}{q} - \frac{1}{q_1}\right) = \frac{1}{p} - \frac{1}{p_1},$$

which themselves are immediate consequences of (5.8) and the definition of m. ■

Now we can present the main interpolation theorem for restricted weak-type operators.

Theorem 5.5 (E. M. Stein & G. Weiss). *Suppose $1 \leq p_0 < p_1 < \infty$ and $1 \leq q_0, q_1 \leq \infty$ with $q_0 \neq q_1$. Suppose further that $0 < \theta < 1$ and define p and q by (5.8). Let T be a linear (respectively, nonnegative sublinear) operator, defined on the simple functions, and suppose that T is of restricted weak types (p_0, q_0) and (p_1, q_1). If $1 \leq r \leq \infty$, then T has a unique extension to a linear (respectively, nonnegative sublinear) operator, again denoted by T, which is bounded from $L^{p,r}$ into $L^{q,r}$, that is,*

$$T : L^{p,r} \to L^{q,r}. \tag{5.9}$$

If, in addition, the inequalities $p_j \leq q_j$ $(j = 0, 1)$ hold, then T is of strong type (p, q).

Proof. If q_0 and q_1 are both greater than 1, then the result follows at once from Theorem 5.3 and the Marcinkiewicz interpolation theorem (Theorem 4.13). Otherwise, choose θ_0 and θ_1 with $0 < \theta_0 < \theta < \theta_1 < 1$ and define $\tilde{p}_j, \tilde{q}_j$, $(j = 0, 1)$, by

$$\frac{1}{\tilde{p}_j} = \frac{1-\theta_j}{p_0} + \frac{\theta_j}{p_1}, \qquad \frac{1}{\tilde{q}_j} = \frac{1-\theta_j}{q_0} + \frac{\theta_j}{q_1}.$$

Then by Lemma 5.4, the operator T is of restricted weak type $(\tilde{p}_j, \tilde{q}_j)$, $(j = 0, 1)$. We have

$$\frac{1}{p} = \frac{1-\tilde{\theta}}{\tilde{p}_0} + \frac{\tilde{\theta}}{\tilde{p}_1}, \qquad \frac{1}{q} = \frac{1-\tilde{\theta}}{\tilde{q}_0} + \frac{\tilde{\theta}}{\tilde{q}_1},$$

with $\tilde{\theta} = (\theta - \theta_0)/(\theta_1 - \theta_0)$. Hence, since $0 < \tilde{\theta} < 1$ and $\tilde{q}_j > 1$, $(j = 0, 1)$, we argue as in the first part of the proof and obtain the desired conclusion from the Marcinkiewicz interpolation theorem. The final assertion, that T is of strong type (p, q) if $p_j \le q_j$, $(j = 0, 1)$, follows exactly as in Corollary 4.14. ■

The next result shows that for certain maximal operators the conclusion of Theorem 5.3 remains valid even when $p = q = 1$.

Theorem 5.6 (K. H. Moon). *Let $(g_m)_{m=1}^{\infty}$ be a sequence of integrable functions on $\mathbf{R}^n$ (or $\mathbf{T}^n$) and define the maximal operator $\mathscr{T}$ by*

$$(\mathscr{T}f)(x) = \sup_m |(f * g_m)(x)|, \qquad (x \in \mathbf{R}^n). \tag{5.10}$$

If $\mathscr{T}$ is of restricted weak type $(1, 1)$, then $\mathscr{T}$ is also of weak type $(1, 1)$.

Proof. We show first that the operators $\mathscr{T}_N$, defined for $N = 1, 2, \ldots$ on $\mathbf{R}^n$ by

$$(\mathscr{T}_N f)(x) = \max_{1 \le m \le N} |(f * g)(x)|, \qquad (x \in \mathbf{R}^n), \tag{5.11}$$

are uniformly of weak type $(1, 1)$, that is,

$$|\{\mathscr{T}_N f > \lambda\}| \le \frac{c\|f\|_1}{\lambda}, \qquad (\lambda > 0, f \in L^1), \tag{5.12}$$

with c a constant depending only on $\mathscr{T}$.

To this end, fix $\lambda > 0$ and set $\varepsilon = \lambda/6$. We consider first the case where f is a nonnegative simple function with bounded support and $\alpha \equiv \|f\|_\infty > 0$. For each $m = 1, 2, \ldots, N$, let h_m be a continuous function of compact support such that

$$\|g_m - h_m\|_1 < \frac{\varepsilon}{\alpha}. \tag{5.13}$$

With

$$(H_N f)(x) = \max_{1 \le m \le N} |(f * h_m)(x)|, \qquad (x \in \mathbf{R}^n), \tag{5.14}$$

we therefore obtain, for every bounded function ϕ,

$$(\mathscr{T}_N\phi)(x) \le (H_N\phi)(x) + \max_{1 \le m \le N} \|g_m - h_m\|_1 \|\phi\|_\infty$$

$$\le (H_N\phi)(x) + \frac{\varepsilon\|\phi\|_\infty}{\alpha}.$$

Similarly, H_N can be estimated by $\mathscr{T}_N$, and so

$$|(\mathscr{T}_N\phi)(x) - (H_N\phi)(x)| \le \frac{\varepsilon\|\phi\|_\infty}{\alpha}, \qquad (\phi \in L^\infty). \tag{5.15}$$

The uniform continuity of the functions h_m guarantees the existence of $\delta > 0$ such that, whenever $|x - y| < \delta$,

$$\max_{1 \le m \le N} |h_m(x) - h_m(y)| < \frac{\varepsilon}{\|f\|_1}. \tag{5.16}$$

With this value of δ, we express f as a finite sum $f = \sum_k \alpha_k \chi_{E_k}$, with $\alpha_k > 0$ and the sets E_k measurable, pairwise disjoint, and satisfying $\operatorname{diam} E_k < \delta$. For each k, select a measurable subset F_k of E_k with

$$|F_k| = \left(\frac{\alpha_k}{\alpha}\right)|E_k| \tag{5.17}$$

and set $F = \bigcup_k F_k$. Clearly,

$$\alpha|F| = \alpha\sum_k |F_k| = \sum_k \alpha_k|E_k| = \|f\|_1. \tag{5.18}$$

Using (5.17), (5.18), and the mean-value theorem for integrals, we find numbers $t_k' \in E_k$ and $t_k'' \in F_k$ such that

$$|\alpha(\chi_F * h_m)(x) - (f * h_m)(x)|$$

$$\le \sum_k \left|\alpha \int_{F_k} h_m(x-t)\,dt - \alpha_k \int_{E_k} h_m(x-t)\,dt\right|$$

$$\le \sum_k \left|\alpha|F_k|h_m(x - t_k') - \alpha_k|E_k|h_m(x - t_k'')\right|$$

$$= \|f\|_1 \sup_k |h_m(x - t_k') - h_m(x - t_k'')| \le \varepsilon,$$

the last inequality following from (5.16). Together with (5.15), this gives

$$(\mathscr{T}_N f)(x) \le (H_N f)(x) + \frac{\varepsilon}{\alpha}\|f\|_\infty = \max_{1 \le m \le N} |(f * h_m)(x)| + \varepsilon$$

$$\le \alpha(H_N\chi_F)(x) + 2\varepsilon \le \alpha\left\{(\mathscr{T}_N\chi_F)(x) + \frac{\varepsilon}{\alpha}\right\} + 2\varepsilon \le \alpha(\mathscr{T}\chi_F)(x) + \frac{\lambda}{2}.$$

It follows that $\{\mathscr{T}_N f > \lambda\} \subset \{\mathscr{T}\chi_F > \lambda/2\alpha\}$. Hence, since $\mathscr{T}$ is of restricted weak type $(1,1)$, we see using (5.6) and (5.18) that

$$|\{\mathscr{T}_N f > \lambda\}| \le M\left(\frac{2\alpha}{\lambda}\right)|F| = \frac{2M\|f\|_1}{\lambda}.$$

This establishes (5.12) for f nonnegative, simple, and with bounded support.

For an arbitrary simple function f of bounded support, we may split f into its real and imaginary parts, then each of these into its positive and negative parts. Applying the preceding result to each of these, and using the sublinearity of $\mathscr{T}_N$ together with property (1.4) of Proposition I.1.3, we obtain

$$|\{\mathscr{T}_N f > \lambda\}| \le \frac{16M\|f\|_1}{\lambda}, \qquad (\lambda > 0). \tag{5.19}$$

Now suppose that f is an arbitrary integrable function on $\mathbf{R}^n$ and let $\varepsilon > 0$ be arbitrary. Then there is a simple function h with bounded support such that $|h| \le |f|$ and

$$\|f - h\|_1 \le \frac{\lambda\varepsilon}{2(K+1)N}, \qquad K = \max_{1\le m\le N} \|g_m\|_1.$$

It follows from the definition of $\mathscr{T}_N$ that

$$\{\mathscr{T}_N f > \lambda\} \subset \left\{\mathscr{T}_N h > \frac{\lambda}{2}\right\} \cup \left(\bigcup_{m=1}^{N} \left\{|(f-h) * g_m| > \frac{\lambda}{2}\right\}\right). \tag{5.20}$$

But

$$\begin{aligned}\left|\left\{|(f-h) * g_m| > \frac{\lambda}{2}\right\}\right| &\le \frac{2}{\lambda}\|(f-h) * g_m\|_1\\ &\le \frac{2K}{\lambda}\|f-h\|_1 < \frac{\varepsilon}{N}.\end{aligned}$$

Hence, using this estimate in conjunction with (5.20), and using (5.19) applied to h, we obtain $|\{\mathscr{T}_N f > \lambda\}| \le 32M\|h\|_1/\lambda + \varepsilon$. But $|h| \le |f|$ and $\varepsilon > 0$ is arbitrary, so this establishes (5.12).

Finally, by property (1.5) of Proposition I.1.3, we may let $N \to \infty$, observe that $\mathscr{T}_N f \uparrow \mathscr{T} f$ a.e., and hence obtain the desired result

$$|\{\mathscr{T} f > \lambda\}| \le c\frac{\|f\|_1}{\lambda}, \qquad (\lambda > 0). \qquad \blacksquare$$

We turn now to Stein's theorem on limits of sequences of operators. On a finite measure space (R,μ), the space $\mathscr{M}_0(R,\mu)$ of measurable functions that are finite μ-a.e. on R is a complete metric linear space under the topology of convergence in measure (cf. Exercise I.1). We consider sequences $(T_n)_{n=1}^{\infty}$ of linear operators that map a Banach space X continuously into $\mathscr{M}_0$. The N-th *maximal operator* $\mathscr{T}_N$, given by

$$(\mathscr{T}_N f)(x) \equiv \max_{1\le n\le N} |(T_n f)(x)|, \qquad (N = 1, 2, \ldots) \tag{5.21}$$

also maps X continuously into $\mathscr{M}_0$ and is nonnegative sublinear:

$$\begin{aligned} 0 \le \mathscr{T}_N(f+g) &\le \mathscr{T}_N f + \mathscr{T}_N g \qquad \mu\text{-a.e.},\\ \mathscr{T}_N(\lambda f) &= |\lambda|\mathscr{T}_N f \qquad \mu\text{-a.e.},\end{aligned} \tag{5.22}$$

for all $f, g \in X$ and all constants λ. The *maximal operator* $\mathscr{T}$,

$$(\mathscr{T} f)(x) \equiv \sup_{n=1,2,\ldots} |(T_n f)(x)|, \tag{5.23}$$

for which $(\mathscr{T}_N f)(x) \uparrow (\mathscr{T} f)(x)$ μ-a.e. as $N \uparrow \infty$, is evidently a nonnegative sublinear operator from X into $\mathscr{M}$, the space of μ-measurable functions on R.

If $\mathscr{T} f$ is finite μ-a.e. for each $f \in X$, that is, is T maps into $\mathscr{M}_0$ rather than $\mathscr{M}$, then since $\mu(R)$ is finite, the distribution function $\mu\{\mathscr{T} f > \lambda\}$ decays to zero as $\lambda \to \infty$, for each individual $f \in X$. The next result establishes a uniform rate of decay.

Theorem 5.7 (Banach's principle). *Let X be a Banach space, let (R,μ) be a finite measure space, and let $(T_n)_{n=1}^{\infty}$ be a sequence of linear operators each of which maps X continuously into $\mathscr{M}_0(R,\mu)$. If the maximal operator $\mathscr{T}$ takes its values in $\mathscr{M}_0$, then there exists a positive decreasing function $c(\lambda)$ on $(0,\infty)$, tending to zero as λ tends to ∞, such that*

$$\mu\{x \in R : (\mathscr{T} f)(x) > \lambda\|f\|_X\} \le c(\lambda) \tag{5.24}$$

for all $\lambda > 0$ and all $f \in X$.

Proof. Fix $\varepsilon > 0$. To each $f \in X$, there corresponds a positive integer k_f such that $\mu\{\mathscr{T} f > k_f\} \le \varepsilon$. Hence, X may be expressed as the union of an increasing family of sets X_k as follows:

$$X = \bigcup_{k=1}^{\infty} \{f \in X : \mu\{Tf > k\} \le \varepsilon\} = \bigcup_{k=1}^{\infty} X_k.$$

Furthermore, since $(\mathscr{T}_N f) \uparrow \mathscr{T} f$ μ-a.e., each X_k may be expressed as the intersection of a decreasing family of sets $X_{k,N}$ in the following way:

$$X_k = \bigcap_{N=1}^{\infty} \{f \in X : \mu\{\mathscr{T}_N f > k\} \le \varepsilon\} = \bigcap_{N=1}^{\infty} X_{k,N}.$$

The continuity of $\mathscr{T}_N$ ensures that each $X_{k,N}$, hence each X_k, is closed in X. By the Baire category theorem, there is a positive integer K such that X_K has nonempty interior. In particular, for some $F \in \mathrm{X}$ and some $\delta > 0$, we have

$$f \in X, \qquad \|f\|_X \leq 1 \quad \Rightarrow \quad \mu\{\mathscr{T}(F + \delta f) > K\} \leq \varepsilon. \tag{5.25}$$

However, the sublinearity of $\mathscr{T}$ gives $\delta \mathscr{T} f \leq \mathscr{T} F + \mathscr{T}(F + \delta f)$, so

$$\mu\left\{\mathscr{T} f > \frac{2K}{\delta}\right\} \leq \mu\{\mathscr{T} F > K\} + \mu\{\mathscr{T}(F + \delta f) > K\}.$$

If $\|f\|_X \leq 1$, we see from (5.25) that the right-hand side has value at most 2ε.
To summarize, if we set

$$c(\lambda) = \sup_{\|f\|_X \leq 1} \mu\{\mathscr{T} f > \lambda\}, \qquad (\lambda > 0),$$

then clearly $c(\lambda)$ is nonnegative and decreasing, and we have shown that to each $\varepsilon > 0$, there corresponds $\delta > 0$ such that $c(\lambda) \leq 2\varepsilon$ whenever $\lambda \geq 2K/\delta$. It follows that $c(\lambda) \to 0$ as $\lambda \to \infty$ and with this the proof is complete. ∎

The following simple application of Banach's principle illustrates once again how a.e. convergence is governed by the associated maximal operator.

Corollary 5.8. *If the maximal operator $\mathscr{T}$ takes its values in $\mathscr{M}_0$, then the subset X_0 of elements $f \in X$ for which*

$$\lim_{n \to \infty} (T_n f)(x) \qquad \textit{exists} \qquad \mu\text{-a.e.} \tag{5.26}$$

is a closed linear subspace of X. In particular, if (5.26) *holds on a dense subset of X, then it holds on all of X.*

Proof. Let f belong to the closure of X_0 in X, and let $(f_n)_{n=1}^{\infty}$ be any sequence of elements of X_0 that converges in X to f. For each $g \in X$, the quantity

$$(\mathscr{R} g)(x) = \limsup_{m,n \to \infty} |(T_m g)(x) - (T_n g)(x)|$$

is majorized by $2\mathscr{T} g(x)$. Therefore, by Banach's principle,

$$\mu\{\mathscr{R} g > \lambda \|g\|_X\} \leq c\left(\frac{\lambda}{2}\right), \qquad (\lambda > 0).$$

But $\mathscr{R} f = \mathscr{R}(f - f_n)$ so, applying the previous estimate with $g = f - f_n$, we have, for any $\varepsilon > 0$,

$$\mu\{\mathscr{R} f > \varepsilon\} \leq c\left(\frac{\varepsilon}{2\|f - f_n\|_X}\right) \to 0 \qquad \text{as} \qquad n \to \infty.$$

It follows that $\mathscr{R} f = 0$ μ-a.e. and hence that $f \in X_0$. ∎

The context for Stein's theorem is the finite measure space consisting of the unit circle $\mathbf{T}$ furnished with normalized Lebesgue measure $dm = d\theta/2\pi$. It is worth noting, however, that the group structure of $\mathbf{T}$ also plays an important role here. If E is a subset of $\mathbf{T}$ and t an arbitrary real number, we denote by E_t the *translate* through e^{it} of E, that is,

$$E_t = \{e^{iy} : e^{i(y-t)} \in E\}.$$

Similarly, if f is a function defined on $\mathbf{T}$, we define its *translate* f_t through e^{it} by $f_t(e^{iy}) = f(e^{i(y-t)})$. An operator T, whose domain and range both are closed under translations, is said to *commute with translations* if $T(f_t) = (Tf)_t$ for all f in its domain and for all real t.

Lemma 5.9. (a) *If E and F are measurable subsets of $\mathbf{T}$, then there exists a translate F_t of F such that*

$$m(E \cap F_t) = m(E)m(F). \tag{5.27}$$

(b) *If E is a subset of $\mathbf{T}$ with positive measure, then there exist n translates of E whose union has measure exceeding $1/2$, provided*

$$n > \frac{\log 2}{m(E)}. \tag{5.28}$$

Proof. (a) The function

$$e^{it} \to m(E \cap F_t) = \frac{1}{2\pi} \int_0^{2\pi} \chi_E(e^{ix}) \chi_F(e^{i(x-t)})\, dx,$$

being the convolution of bounded functions, is continuous. Furthermore, an integration gives

$$\frac{1}{2\pi} \int_0^{2\pi} m(E \cap F_t)\, dt = m(E)m(F),$$

so the result follows from the mean-value theorem for integrals.
(b) Let e^{it_j}, $(j = 0, 1, \ldots, n-1)$, be arbitrary points of $\mathbf{T}$ and consider the corresponding translates E_{t_j} of E. If $F = E^c$, the complement of E, then $F_{t_j} = (E_{t_j})^c$. Hence, with $E' = \bigcup_{j=0}^{n-1} E_{t_j}$, we have

$$m(E') = 1 - m((E')^c) = 1 - m\left(\bigcap_{j=0}^{n-1} E_{t_j}^c\right) = 1 - m\left(\bigcap_{j=0}^{n-1} F_{t_j}\right).$$

By repeated application of part (a), we may select e^{it_j}, $(j = 0, 1, \ldots, n-1)$, so that

$$m\left(\bigcap_{j=0}^{n-1} F_{t_j}\right) = \prod_{j=0}^{n-1} m(F_{t_j}) = m(F)^n = (1 - m(E))^n.$$

Hence,

$$m(E') = 1 - (1 - m(E))^n.$$

Since $1 - x \le e^{-x}$, $(x \ge 0)$, it follows that

$$m(E') \ge 1 - e^{-n \cdot m(E)} \ge \frac{1}{2},$$

provided $n \cdot m(E) > \log 2$. ■

Theorem 5.10. (E. M. Stein). *Suppose $1 \le p \le 2$ and let $(T_k)_{k=1}^\infty$ be a sequence of linear operators each of which commutes with translations and maps $L^p(\mathbf{T})$ continuously into $\mathscr{M}_0(\mathbf{T})$. If the maximal operator $\mathscr{T}$ has the property that $\mathscr{T}f(e^{ix}) < \infty$ a.e. for all $f \in L^p$, then $\mathscr{T}$ is a bounded operator from L^p into $L^{p,\infty}$. In particular, $\mathscr{T}$ is of weak type (p, p).*

Proof. The assertion is that if $\mathscr{T}$ maps L^p into $\mathscr{M}_0$, then in fact $\mathscr{T}$ maps L^p boundedly into $L^{p,\infty}$. Thus we need to show that

$$m\{\mathscr{T}f > \lambda\} \le \frac{c_p}{\lambda^p}, \qquad (\lambda > 0) \tag{5.29}$$

for all f in the unit ball of L^p, with c_p independent of f and λ.

Fix f with $\|f\|_p \le 1$ and $\lambda > 0$. We may suppose that $E = \{\mathscr{T}f > \lambda\}$ has positive measure or else there is nothing to prove. By Lemma 5.9(b), there is an integer n satisfying

$$n \cdot m(E) > \log 2 \tag{5.30}$$

and n points e^{it_j}, $(j = 0, 1, \dots, n-1)$, of $\mathbf{T}$ such that the union $E' = \bigcup_{j=0}^{n-1} E_{t_j}$ of the translates E_{t_j} of E satisfies

$$m(E') \ge \frac{1}{2}. \tag{5.31}$$

For each $j = 0, 1, \dots, n-1$, let ϕ_j be the (Rademacher) function

$$\phi_j(\omega) = \operatorname{sgn} \sin(2^{j+1}\pi\omega), \qquad (0 \le \omega \le 1),$$

and let

$$F(e^{ix}, \omega) = \frac{1}{M} \sum_{j=0}^{n-1} \phi_j(\omega) f_{t_j}(e^{ix}), \qquad (e^{ix} \in \mathbf{T}, 0 \le \omega \le 1),$$

where the f_{t_j} are the translates of f through e^{it_j} and M is a constant to be determined later. As ω varies over $[0, 1]$, we obtain in this way 2^n functions

of the form $1/M \sum_j \pm f_{t_j}(e^{ix})$. The idea now is to force sufficiently many of these functions to belong to the unit ball of L^p without, on the other hand, having to choose M too large. The argument is as follows. Since $p \le 2$, successive applications of Hölder's inequality and Parseval's theorem (the ϕ_j are evidently orthogonal) give

$$\begin{aligned}\int_0^1 |F(e^{ix}, \omega)|^p \, d\omega &\le \left\{\int_0^1 |F(e^{ix}, \omega)|^2 \, d\omega\right\}^{p/2} \\ &\le \left\{\sum_{j=0}^{n-1} \left|\frac{f_{t_j}(e^{ix})}{M}\right|^2\right\}^{p/2} \\ &\le \frac{1}{M^p} \sum_{j=0}^{n-1} |f_{t_j}(e^{ix})|^p.\end{aligned}$$

Integrating with respect to x and changing the order of integration, we therefore obtain

$$\int_0^1 \frac{1}{2\pi} \int_0^{2\pi} |F(e^{ix}, \omega)|^p \, dx \, d\omega \le \frac{n}{M^p} \|f\|_p^p \le \frac{n}{M^p}.$$

Hence, if M is chosen so that

$$M^p \ge 4n, \tag{5.32}$$

then the set $\Gamma = \{\omega : \|F(\cdot, \omega)\|_p \le 1\}$ will have measure

$$|\Gamma| \ge \frac{3}{4}. \tag{5.33}$$

Next, we examine the behavior of $\mathscr{T}F$ on E'. Each e^{ix} in E' belongs to some translate E_{t_ν}, $(\nu = 0, 1, \dots, n-1)$, of E and so $e^{iy} = e^{i(x - t_\nu)}$ belongs to E. It is clear that $\mathscr{T}$ commutes with translations because each T_k does. It follows that

$$\mathscr{T}(f_{t_\nu})(e^{ix}) = (\mathscr{T}f)_{t_\nu}(e^{ix}) = (\mathscr{T}f)(e^{iy}) > \lambda$$

and hence that

$$|T_k(f_{t_\nu})(e^{ix})| > \lambda \tag{5.34}$$

for some positive integer k (also depending on x). We write

$$T_k F(e^{ix}, \omega) = \frac{1}{M} \sum_{\substack{j=0 \\ j \ne \nu}}^{n-1} \phi_j(\omega) T_k f_{t_j}(e^{ix}) + \frac{1}{M} \phi_\nu(\omega) T_k f_{t_\nu}(e^{ix})$$

and denote by Γ_x the subset of $\{\omega : 0 \le \omega \le 1\}$ for which the sign of the first term on the right (the sum) is either zero or equal to the sign of the second term.

By the symmetry of the functions ϕ_j, it is clear that $|\Gamma_x| \geq 1/2$. However, we see also from (5.34) that

$$|T_k F(e^{ix}, \omega)| > \frac{\lambda}{M}, \qquad (\omega \in \Gamma_x).$$

Hence, for each $x \in E'$, we have

$$\mathscr{T} F(e^{ix}, \omega) > \frac{\lambda}{M}, \qquad (\omega \in \Gamma_x), \tag{5.35}$$

and, by (5.33),

$$|\Gamma_x \cap \Gamma| = |\Gamma_x| + |\Gamma| - |\Gamma_x \cup \Gamma| \geq \frac{1}{2} + \frac{3}{4} - 1 = \frac{1}{4}. \tag{5.36}$$

Now set

$$\chi(e^{ix}, \omega) = \begin{cases} 1 & \text{if } (\mathscr{T} F)(e^{ix}, \omega) > \dfrac{\lambda}{M} \\ 0 & \text{otherwise.} \end{cases}$$

Then, by (5.31), (5.35), and (5.36),

$$\int_{E'} \int_F \chi(e^{ix}, \omega)\, d\omega\, dm \geq \int_{E'} \int_{\Gamma \cap \Gamma_x} d\omega\, dm = \int_{E'} |\Gamma \cap \Gamma_x|\, dm > \frac{1}{8}$$

and so, interchanging the order of integration, we see that

$$\int_{E'} \chi(e^{ix}, \omega_0)\, dm > \frac{1}{8}$$

for some $\omega_0 \in \Gamma$. Hence, $m\{e^{ix} : (\mathscr{T} F)(e^{ix}, \omega_0) > \lambda/M\} > 1/8$.

To summarize, we have constructed a function $F(e^{ix}) \equiv F(e^{ix}, \omega_0)$ in the unit ball of L^p, which satisfies

$$m\left\{\mathscr{T} F > \frac{\lambda}{M}\right\} > \frac{1}{8}, \tag{5.37}$$

provided M and n are chosen according to (5.30) and (5.32), that is,

$$\frac{\log 2}{m(E)} < n \leq \frac{M^p}{4}. \tag{5.38}$$

However, by Banach's principle (Theorem 5.7), there exists $\lambda_0 > 0$ such that

$$m\{\mathscr{T} g > \tau\} \leq \frac{1}{8}, \qquad (\tau \geq \lambda_0) \tag{5.39}$$

for all g in the unit ball of L^p. Hence, choosing n to be the integer part of $(2\log 2)/m(E)$, and $M = (4n)^{1/p}$ in accordance with (5.38), we see from (5.37) and (5.39) that $\lambda/M \leq \lambda_0$, that is,

$$m(E) \leq \frac{2\log 2}{n} = \frac{8\log 2}{M^p} \leq 8\log 2 \cdot \left(\frac{\lambda_0}{\lambda}\right)^p.$$

This establishes (5.39) and hence completes the proof. ■

6. *L*log*L* AND $L_{\exp}$

The Marcinkiewicz interpolation theorem describes the mapping properties of weak-type operators on the interior $0 < \theta < 1$ of the interpolation segment in question. The limiting cases $\theta \to 0$ and $\theta \to 1$ are also of interest, however, as we shall see in this section. The results have their most natural formulation on finite measure spaces (R, μ). For convenience, we shall assume $\mu(R) = 1$.

So as not to obscure the main ideas with notation, and also to maintain a reasonable historical perspective, we shall begin with the spaces $L\log L$ and $L_{\exp}$, which were introduced by A. Zygmund (and E. C. Titchmarsh) in 1928. The Marcinkiewicz interpolation theorem shows that operators of joint weak type $(1, 1; \infty, \infty)$ are bounded on L^p for $1 < p < \infty$. The spaces $L\log L$ and $L_{\exp}$ enter the picture in the limiting cases $p \to 1$ and $p \to \infty$ of this and other related results.

Definition 6.1. The *Zygmund space* $L\log L$ consists of all μ-measurable functions f on R for which

$$\int_R |f(x)| \log^+ |f(x)|\, dx < \infty \tag{6.1}$$

(here, $\log^+ x = \max(\log x, 0)$). The *Zygmund space* $L_{\exp}$ consists of all μ-measurable functions f on R for which there is a constant $\lambda = \lambda(f) > 0$ such that

$$\int_R \exp(\lambda |f(x)|)\, dx < \infty. \tag{6.2}$$

The quantities in (6.1) and (6.2) are evidently far from satisfying the properties of a norm. The expressions introduced in the next lemma, defined in terms of the decreasing rearrangement, will prove more amenable in this regard.

Lemma 6.2. *Suppose* $\mu(R) = 1$. *Then a* μ*-measurable function* f *on* R

satisfies (6.1) *if and only if*

$$\int_0^1 f^*(t)\log\left(\frac{1}{t}\right)dt < \infty. \tag{6.3}$$

On the other hand, a μ-measurable function f on R satisfies (6.2) *for some $\lambda > 0$ if and only if there is a constant $c = c(f) > 0$ such that*

$$f^*(t) \le c\left(1 + \log\frac{1}{t}\right), \qquad (0 < t < 1). \tag{6.4}$$

Proof. Since f and f^* are equimeasurable and $x \to \log^+ x$ is a nonnegative, continuous, increasing function, a standard argument (cf. Exercise II.3) shows that (6.1) is equivalent to

$$\int_0^1 f^*(t)\log^+ f^*(t)\,dt < \infty. \tag{6.5}$$

Similarly, (6.2) is equivalent to

$$\int_0^1 \exp(\lambda f^*(t))\,dt < \infty. \tag{6.6}$$

Hence, the lemma will be established if we show the equivalence of (6.3) and (6.5), and of (6.4) and (6.6).

Suppose first that (6.3) holds. Evidently, f is integrable and so $f^*(t) \le f^{**}(t) \le \|f\|_1/t$. Hence, assuming $\|f\|_1 > 0$ (otherwise there is nothing to prove), we have

$$\begin{aligned}\int_0^1 f^*(t)\log^+ f^*(t)\,dt &\le \int_0^1 f^*(t)\log^+\left\{\frac{\|f\|_1}{t}\right\}dt \\ &= \int_0^{\min(\|f\|_1,1)} f^*(t)\log\left\{\frac{\|f\|_1}{t}\right\}dt \\ &\le \int_0^1 f^*(t)\log\left(\frac{1}{t}\right)dt + \|f\|_1\left|\log\|f\|_1\right|.\end{aligned}$$

Since (6.3) holds and $\|f\|_1 > 0$, this establishes (6.5).

Conversely, suppose (6.5) holds. Let $E = \{f^*(t) > t^{-1/2}\}$ and $F = [0,1]\setminus E$ (either of which may be empty). Then

$$\begin{aligned}\int_0^1 f^*(t)\log\left(\frac{1}{t}\right)dt &\le \int_E f^*(t)\log(f^*(t)^2)\,dt + \int_F t^{-1/2}\log\left(\frac{1}{t}\right)dt \\ &\le 2\int_0^1 f^*(t)\log^+ f^*(t)\,dt + \int_0^1 t^{-1/2}\log\left(\frac{1}{t}\right)dt.\end{aligned}$$

The second integral on the right is obviously convergent, and the first converges because of (6.5). Hence, (6.3) holds.

Turning now to the equivalence of (6.4) and (6.6), we suppose first that (6.4) holds for some constant $c > 0$. Then

$$\begin{aligned}\int_0^1 \exp(\lambda f^*(t))\,dt &\le \int_0^1 \exp\left[c\lambda\left(1 + \log\frac{1}{t}\right)\right]dt \\ &= e^{c\lambda}\int_0^1 t^{-c\lambda}\,dt\end{aligned}$$

and so (6.6) holds for any constant $\lambda < 1/c$.

Conversely, suppose (6.6) holds with constant $\lambda > 0$ and set

$$K = \int_0^1 \exp(\lambda f^*(t))\,dt < \infty.$$

By Jensen's inequality,

$$\begin{aligned}\exp(\lambda f^*(t)) \le \exp(\lambda f^{**}(t)) &= \exp\left(\frac{1}{t}\int_0^t \lambda f^*(s)\,ds\right) \\ &\le \frac{1}{t}\int_0^t \exp(\lambda f^*(s))\,ds \le \frac{K}{t}.\end{aligned}$$

Hence,

$$f^*(t) \le \frac{1}{\lambda}\log\left(\frac{K}{t}\right) \le \frac{1}{\lambda}\left(1 + \log\frac{1}{t}\right)\cdot\max(\log K, 1)$$

and so (6.4) holds with constant $c = \lambda^{-1}\max(\log K, 1)$. The proof is now complete. ■

An integration by parts in (6.3) shows that

$$\int_0^1 f^*(t)\log\left(\frac{1}{t}\right)dt = \int_0^1 f^{**}(t)\,dt. \tag{6.7}$$

The latter quantity, involving the subadditive function $f \to f^{**}$, satisfies the triangle inequality and so may be used directly to define a norm on $L\log L$ (cf. Definition 6.3 below). On the other hand, the expression (6.4) for the space $L_{\exp}$ involves f^* rather than the subadditive f^{**}. This presents no problem, however, because (6.4) is *equivalent* to the corresponding inequality for f^{**}, namely

$$f^{**}(t) \le c\left(1 + \log\frac{1}{t}\right), \qquad (0 < t < 1). \tag{6.8}$$

Indeed, (6.8) implies (6.4) because $f^* \leq f^{**}$, whereas, in the opposite direction, an integration of (6.4) gives

$$f^{**}(t) = \frac{1}{t}\int_0^t f^*(s)\,ds \leq \frac{1}{t}\int_0^t c\left(1 + \log\frac{1}{s}\right)ds \leq 2c\left(1 + \log\frac{1}{t}\right).$$

Hence, (6.4) (with constant c) implies (6.8) (with constant $2c$), and so (6.4) and (6.8) are equivalent. We use the relation (6.8) to define a norm on $L_{\exp}$ as follows.

Definition 6.3. Suppose $\mu(R) = 1$ and let f be a μ-measurable function on R. Set

$$\|f\|_{L\log L} \equiv \int_0^1 f^*(t)\log\left(\frac{1}{t}\right)dt = \int_0^1 f^{**}(t)\,dt \tag{6.9}$$

and

$$\|f\|_{L_{\exp}} \equiv \sup_{0<t<1} \frac{f^{**}(t)}{1 + \log\frac{1}{t}}. \tag{6.10}$$

It follows from Lemma 6.2 and the observations made above that $L\log L$ and $L_{\exp}$ consist of all μ-measurable functions f on R for which the respective quantities in (6.9) and (6.10) are finite. Since $f \to f^{**}$ is subadditive, it is easy to prove directly that these quantities define norms under which $L\log L$ and $L_{\exp}$ are rearrangement-invariant Banach function spaces.

The same result could also be obtained from Theorem II.5.13 with the observation that $L\log L$ is precisely the Lorentz Λ-space with concave fundamental function

$$\varphi(t) = t\left(1 + \log\frac{1}{t}\right)$$

(cf. II.(5.18)), whereas $L_{\exp}$ is none other than the Lorentz M-space with quasiconcave fundamental function

$$\varphi(t) = \frac{1}{1 + \log\frac{1}{t}}$$

(cf. II.(5.13)). In any event, we have the following result.

Theorem 6.4 *Suppose $\mu(R) = 1$. Then $(L\log L, \|\cdot\|_{L\log L})$ and $(L_{\exp}, \|\cdot\|_{L_{\exp}})$ are rearrangement-invariant Banach function spaces.*

The spaces $L\log L$ and $L_{\exp}$ are related to one another, and to the Lebesgue spaces, as follows.

Theorem 6.5 *Suppose $\mu(R) = 1$. The continuous embeddings*

$$L^\infty \hookrightarrow L_{\exp} \hookrightarrow L^p \hookrightarrow L\log L \hookrightarrow L^1 \tag{6.11}$$

hold for all p satisfying $1 < p < \infty$. The spaces $L\log L$ and $L_{\exp}$ are mutually associate (up to equivalence of norms) and $L_{\exp}$ may be identified (again, up to equivalence of norms) with the Banach space dual of $L\log L$. Both Boyd indices of $L\log L$ are equal to 1, and both those of $L_{\exp}$ are equal to 0.

Proof. We show first that $L_{\exp}$ is equivalent to the associate space of $L\log L$. For convenience, set $X = L\log L$ and $Y = L_{\exp}$. If g belongs to the associate space X' of X, then by Hölder's inequality, for $0 < t < 1$,

$$tg^{**}(t) = \int_0^1 \chi_{[0,t]}(s)g^*(s)\,ds \leq \|\chi_{[0,t]}\|_X \|g\|_{X'}.$$

The norm of the characteristic function can be computed directly from (6.9): its value is $t(1 + \log 1/t)$. Hence, dividing both sides by this value and taking the supremum over t, we see from (6.10) that

$$\|g\|_Y \leq \|g\|_{X'}, \qquad (g \in X'). \tag{6.12}$$

Conversely, suppose that $g \in Y$. Then, for any $f \in X$,

$$\begin{aligned}\int_0^1 f^*(t)g^*(t)\,dt &\leq \int_0^1 f^*(t)\left(1 + \log\frac{1}{t}\right)\frac{g^{**}(t)}{\left(1 + \log\frac{1}{t}\right)}\,dt \\ &\leq \|g\|_Y \int_0^1 f^*(t)\left(1 + \log\frac{1}{t}\right)dt \\ &\leq \|g\|_Y \left\{\int_0^1 f^{**}(t)\,dt + \int_0^1 f^*(t)\log\frac{1}{t}\,dt\right\} \\ &= 2\|f\|_X \|g\|_Y.\end{aligned}$$

Hence, taking the supremum over all f in the unit ball of X, we obtain

$$\|g\|_{X'} \leq 2\|g\|_Y, \qquad (g \in Y). \tag{6.13}$$

The estimates (6.12) and (6.13) together show that $Y = L_{\exp}$ is equivalent to the associate space of $X = L\log L$ and hence, by the Lorentz-Luxemburg theorem (Theorem I.2.7), that $L\log L$ and $L_{\exp}$ are mutually associate.

Using (6.9) it is routine to check that $L\log L$ has absolutely continuous norm. By Corollary I.4.3, its dual space is therefore canonically isometrically isomorphic to its associate space, that is, the space $L_{\exp}$, up to equivalence of norms.

The first and the last of the continuous embeddings in (6.11) follow immediately from the definitions (6.9) and (6.10). To see that L^p is embedded in $L\log L$ when $p > 1$, we simply apply Hölder's inequality and then Hardy's inequality III.(3.18) (with $\lambda^{-1} = q = p$) to (6.9) to obtain

$$\|f\|_{L\log L} \le \left(\int_0^1 f^{**}(t)^p\,dt\right)^{1/p} \le p'\left(\int_0^1 f^*(t)\,dt\right)^{1/p} = p'\|f\|_p, \tag{6.14}$$

where $p' = p/(p-1)$. The remaining embedding of $L_{\exp}$ into L^p $(p < \infty)$ now follows from this one by passing to the associate spaces.

It remains only to compute the indices. If $t > 1$, the dilation operator $E_{1/t}$ is defined on functions g on $(0,1)$ by $(E_{1/t}g)(s) = g(s/t)$, $(0 < s < 1)$. Thus, if $f \in L_{\exp}$, we have from (6.10),

$$\begin{aligned}[E_{1/t}(f^*)]^{**}(s) = f^{**}\left(\frac{s}{t}\right) &\le \left(1 + \log\frac{t}{s}\right)\|f\|_{L_{\exp}} \\ &\le (1+\log t)\left(1 + \log\frac{1}{s}\right)\|f\|_{L_{\exp}}.\end{aligned}$$

Dividing both sides by $(1 + \log 1/s)$ and taking the supremum over all $s < 1$, we see again from (6.10) that the norm $h(t)$ of $E_{1/t}$ in $L_{\exp}$ satisfies $h(t) \le (1 + \log t)$, $(t > 1)$. It follows therefore from Definition III.5.12 that the upper index $\bar{\alpha}$ of $L_{\exp}$ satisfies

$$\bar{\alpha} = \inf_{t>1} \frac{\log h(t)}{\log t} \le \inf_{t>1} \frac{\log(1+\log t)}{\log t} = 0.$$

This, together with Proposition III.5.13, shows that $0 \le \underline{\alpha} \le \bar{\alpha} \le 0$ and we conclude that both indices of $L_{\exp}$ are equal to 0. Since $L\log L$ is the associate space of $L_{\exp}$, its indices are complementary to those of $L_{\exp}$ (Theorem III.5.13) and hence are equal to 1. ■

Now let us explore some of the elementary interpolation properties of $L\log L$ and $L_{\exp}$.

Theorem 6.6 *Let T be a quasilinear operator relative to two finite measure spaces (R, μ) and (S, ν).*
(a) *Suppose $1 < p, q \le \infty$. If T is of joint weak type $(1, 1; p, q)$, then T is a bounded operator from $L\log L(R, \mu)$ into $L^1(S, \nu)$:*

$$T: L\log L \to L^1. \tag{6.15}$$

(b) *Suppose $1 \le p, q < \infty$. If T is of joint weak type $(p, q; \infty, \infty)$, then T is a bounded operator from $L^\infty(R, \mu)$ into $L_{\exp}(S, \nu)$:*

$$T: L^\infty \to L_{\exp}. \tag{6.16}$$

Proof. Without loss of generality, we may assume that $\mu(R) = \nu(S) = 1$. In case (a), the operator T of joint weak type $(1, 1; p, q)$ satisfies

$$(Tf)^*(t) \le c\left\{\frac{1}{t}\int_0^{t^m} f^*(s)\,ds + t^{-1/q}\int_{t^m}^1 s^{1/p} f^*(s)\frac{ds}{s}\right\} \tag{6.17}$$

for $0 < t < 1$, where c is a constant and

$$m = \frac{\dfrac{1}{q} - 1}{\dfrac{1}{p} - 1} \tag{6.18}$$

(cf. Definition III.5.4). Note that the hypotheses imply that $0 < m < \infty$.

A simple change of variables gives

$$\begin{aligned}\int_0^1 \frac{1}{t}\int_0^{t^m} f^*(s)\,ds\,dt &= \frac{1}{m}\int_0^1 \frac{1}{u}\int_0^u f^*(s)\,ds\,du \\ &= \frac{1}{m}\int_0^1 f^{**}(u)\,du = \frac{1}{m}\|f\|_{L\log L}.\end{aligned}$$

On the other hand, interchanging the order of integration and using (6.18), we have

$$\begin{aligned}&\int_0^1 t^{-1/q}\int_{t^m}^1 s^{1/p-1} f^*(s)\,ds\,dt \\ &\quad= \int_0^1 s^{1/p-1} f^*(s)\,ds \int_0^{s^{1/m}} t^{-1/q}\,dt \\ &\quad= \frac{1}{1-\dfrac{1}{q}}\int_0^1 f^*(s)\,ds \le \frac{1}{1-\dfrac{1}{q}}\|f\|_{L\log L}.\end{aligned}$$

Hence, combining these estimates with (6.17), we obtain

$$\|Tf\|_1 = \int_0^1 (Tf)^*(t)\,dt \le c\left\{\frac{2-\dfrac{1}{p}}{1-\dfrac{1}{q}}\right\}\|f\|_{L\log L}, \tag{6.19}$$

and this establishes part (a).

In case (b), the operator T of joint weak type $(p, q; \infty, \infty)$ satisfies

$$(Tf)^*(t) \le c\left\{t^{-1/q}\int_0^{t^m} s^{1/p} f^*(s)\frac{ds}{s} + \int_{t^m}^1 f^*(s)\frac{ds}{s}\right\} \tag{6.20}$$

for $0 < t < 1$, where

$$m = \frac{p}{q}. \tag{6.21}$$

If $f \in L^\infty$, we therefore have

$$\begin{aligned}(Tf)^*(t) &\le c\|f\|_\infty\left\{t^{-1/q}\int_0^{t^m} s^{1/p-1}\,ds + \int_{t^m}^1 \frac{ds}{s}\right\}\\ &= c\|f\|_\infty\left\{pt^{-1/q+m/p} + \log\left(\frac{1}{t^m}\right)\right\}\\ &= c\|f\|_\infty\left(p + m\log\frac{1}{t}\right) \le cp\|f\|_\infty\left(1 + \log\frac{1}{t}\right).\end{aligned}$$

Exactly as in the discussion preceding Definition 6.3, we thus obtain

$$(Tf)^{**}(t) \le 2cp\|f\|_\infty\left(1 + \log\frac{1}{t}\right)$$

and hence, from (6.10),

$$\|Tf\|_{L_{\exp}} \le 2cp\|f\|_\infty. \tag{6.22}$$

This establishes part (b) and hence completes the proof. ■

The Hardy-Littlewood maximal operator M on the unit circle $\mathbf{T}$ is of joint weak type $(1, 1; \infty, \infty)$, so both (6.15) and (6.16) are valid for M. The latter property (6.16) is without interest here since we know already the stronger property $T: L^\infty \to L^\infty$. What is more interesting is that (6.15) has a converse.

Theorem 6.7. *If $f \in L^1(\mathbf{T})$, then its Hardy-Littlewood maximal function Mf is integrable if and only if $f \in L\log L$. Furthermore, there are constants c_1, c_2 such that*

$$c_1\|f\|_{L\log L} \le \|Mf\|_{L^1} \le c_2\|f\|_{L\log L} \tag{6.23}$$

for all $f \in L\log L$.

Proof. Recall from (6.9) that

$$\|f\|_{L\log L} = \int_0^1 f^*(t)\log\frac{1}{t}\,dt = \int_0^1 f^{**}(t)\,dt. \tag{6.24}$$

By Theorem III.3.8, the quantity f^{**} is equivalent to $(Mf)^*$, so (6.23) follows at once from (6.24). ■

This result demonstrates the intrinsic nature of the space $L\log L$: it arises quite naturally as soon as one investigates the integrability of certain maximal functions. There are related results for other kinds of maximal operators. Consider, for example, the nontangential maximal operator N (cf. Exercise III.9,12). Since the Poisson kernel P_r can be uniformly approximated by convex combinations of functions of the form $2\pi\,\chi_I/|I|$, with I an interval of the form $I = [-a, a]$, $0 < a \le \pi$, it is easy to see that $Nf \le cMf$ a.e. for all $f \in L^1$ (cf. Exercise III.9). Hence, N is also of joint weak type $(1, 1; \infty, \infty)$. Using Theorem 6.6 exactly as before, we deduce that $N: L\log L \to L^1$. Although the converse is false, there is a result in this direction for nonnegative functions f. To see this, we need only observe that for each interval $I = [-a, a]$, $0 < a \le \pi$, there is a value of r, $0 \le r < 1$, such that $\chi_I/|I| \le P_r$ (if $0 < a \le \pi/2$, take r to be the unique solution of $2r/(r^2 + 1) = \cos a$, and if $\pi/2 \le a \le \pi$, take $r = 0$). With this observation and some computation, it can be shown that $Mf \le cNf$ a.e. for all $f \ge 0$. The following result is then an immediate consequence of Theorem 6.7.

Corollary 6.8. *If $f \in L\log L(\mathbf{T})$, then its nontangential maximal function $Nf \in L^1(\mathbf{T})$. Conversely, if $f \ge 0$, and $Nf \in L^1$, then $f \in L\log L$.*

The conjugate-function operator $f \to \tilde{f}$ is another important example of an operator of joint weak type $(1, 1; \infty, \infty)$ (cf. Theorem III.6.8). In this case, Theorem 6.6 yields the following classical result.

Corollary 6.9 (A. Zygmund). (a) *If $f \in L\log L(\mathbf{T})$, then its conjugate function $\tilde{f}$ belongs to $L^1(\mathbf{T})$ and*

$$\|\tilde{f}\|_{L^1} \le c\|f\|_{L\log L}, \qquad (f \in L\log L). \tag{6.25}$$

(b) *If $f \in L^\infty(\mathbf{T})$, then its conjugate function $\tilde{f}$ belongs to $L_{\exp}$ and*

$$\|\tilde{f}\|_{L_{\exp}} \le c\|f\|_{L^\infty}, \qquad (f \in L^\infty). \tag{6.26}$$

Here again there is a partial converse to the $L\log L$ result, for nonnegative functions. Although this is not difficult to prove directly, we shall expedite matters by appealing to a theorem of Hardy and Littlewood (cf. Theorem V.6.14) to the effect that if both f and $\tilde{f}$ are integrable, then so is Nf. The following result is then an immediate consequence of Corollary 6.8.

Corollary 6.10 (M. Riesz). *If both f and its conjugate function $\tilde{f}$ belong to $L^1(\mathbf{T})$, and if $f \ge 0$ a.e., then $f \in L\log L$.*

Finally, let us remark that analogous results hold also for the maximal conjugate-function operator $\mathscr{C}$ (cf. III.(6.12)). Indeed, Theorem III.6.8 shows that $\mathscr{C}$ is of joint weak type $(1, 1; \infty, \infty)$ so the analogue of Zygmund's theorem (Corollary 6.9) follows as before from Theorem 6.6. The converse, for non-negative f, is an immediate consequence of Corollary 6.10 since $|\tilde{f}| \le \mathscr{C}f$ a.e.

We turn now from the special case of operators of joint weak type $(1, 1; \infty, \infty)$ (and the spaces $L\log L$ and $L_{\exp}$) to the general case of operators of weak type $(p_0, q_0; p_1, q_1)$. We begin with the following generalizations of $L\log L$ and $L_{\exp}$.

Definition 6.11. Let (R, μ) be a finite measure space. If $0 < p < \infty$ and $-\infty < \alpha < \infty$, the *Zygmund space* $L^p(\log L)^\alpha$ consists of all μ-measurable functions f on R for which

$$\int_R [|f(x)| \log^\alpha(2 + |f(x)|)]^p \, d\mu < \infty. \tag{6.27}$$

If $\alpha \ge 0$, the *Zygmund space* $L^\alpha_{\exp}$ consists of all μ-measurable functions f on R for which there is a constant $\lambda = \lambda(f) > 0$ such that

$$\int_R \exp[\lambda |f(x)|]^{1/\alpha} \, d\mu < \infty \tag{6.28}$$

(if $\alpha = 0$, (6.28) is interpreted to mean that f is bounded; thus, $L^0_{\exp} = L^\infty$).

A comparison with Definition 6.1 shows that in this notation

$$L^1(\log L)^1 = L\log L, \qquad L^1_{\exp} = L_{\exp}. \tag{6.29}$$

As in Lemma 6.2, there are the following characterizations of the Zygmund spaces in terms of the decreasing rearrangement f^*. The proof is similar to that of Lemma 6.2 and is omitted.

Lemma 6.12. *Suppose $\mu(R) = 1$. If $0 < p < \infty$ and $-\infty < \alpha < \infty$, then a μ-measurable function f on R belongs to the Zygmund space $L^p(\log L)^\alpha$ if and only if*

$$\left(\int_0^1 [(1 - \log t)^\alpha f^*(t)]^p \, dt \right)^{1/p} < \infty. \tag{6.30}$$

If $\alpha \ge 0$, then a μ-measurable function f on R belongs to the Zygmund space $L^\alpha_{\exp}$ if and only if

$$\sup_{0<t<1} (1 - \log t)^{-\alpha} f^*(t) < \infty. \tag{6.31}$$

The expressions (6.30) and (6.31) for the Zygmund spaces are similar to those used to define the norms of the Lorentz spaces $L^{q,p}$ (Definition 4.1) except that they contain a logarithmic weighting factor $(1 - \log t)^\alpha$ instead of a power $t^{1/q}$. By incorporating both kinds of weighting factors into the same expression, we shall be able to treat both classes of spaces simultaneously. We make the following definition.

Definition 6.13. Suppose $\mu(R) = 1$ and $0 < p,\ q \le \infty$, $-\infty < \alpha < \infty$. The *Lorentz-Zygmund space* $L^{p,q}(\log L)^\alpha$ consists of all μ-measurable functions f on R for which

$$\|f\|_{p,q;\alpha} = \begin{cases} \left(\int_0^1 [t^{1/p}(1 - \log t)^\alpha f^*(t)]^q \dfrac{dt}{t} \right)^{1/q}, & (0 < q < \infty), \\ \sup\limits_{0<t<1} [t^{1/p}(1 - \log t)^\alpha f^*(t)], & (q = \infty), \end{cases} \tag{6.32}$$

is finite.

Clearly,

$$L^{p,q}(\log L)^0 = L^{p,q}, \qquad (0 < p, q \le \infty) \tag{6.33}$$

and, in view of Lemma 6.12,

$$L^{p,p}(\log L)^\alpha = L^p(\log L)^\alpha, \qquad (0 < p < \infty,\ -\infty < \alpha < \infty) \tag{6.34}$$

and

$$L^{\infty,\infty}(\log L)^{-\alpha} = L^\alpha_{\exp}, \qquad (\alpha \ge 0). \tag{6.35}$$

Standard arguments provide a generalization of the Marcinkiewicz interpolation theorem (Theorem 4.13) for Lorentz-Zygmund spaces (the conclusion $T: L^{p,r} \to L^{q,r}$ being replaced by $T: L^{p,r}(\log L)^\alpha \to L^{q,r}(\log L)^\alpha$, $(-\infty < \alpha < \infty)$). The behavior at the endpoints $\theta = 0$ and $\theta = 1$ of the interpolation segment is less obvious but nevertheless quite simple to describe:

Theorem 6.14 (C. Bennett & K. Rudnick). *Let T be a quasilinear operator relative to a pair of finite measure spaces (R, μ) and (S, ν). Suppose $0 < p_0 < p_1 \le \infty$ and $0 < q_0 < q_1 \le \infty$ and that T is a quasilinear operator of joint weak type $(p_0, q_0; p_1, q_1)$. Suppose $1 \le a \le b \le \infty$ and $-\infty < \alpha,\ \beta < \infty$. Then*

$$T: L^{p_0,a}(\log L)^{\alpha+1} \to L^{q_0,b}(\log L)^\beta \tag{6.36}$$

provided $\alpha + 1/a = \beta + 1/b > 0$. *On the other hand,*

$$T: L^{p_1,a}(\log L)^{\alpha+1} \to L^{q_1,b}(\log L)^{\beta} \tag{6.37}$$

provided $\alpha + 1/a = \beta + 1/b < 0$.

The proof of Theorem 6.14 is similar to that of Theorem 6.6 (and all other Marcinkiewicz-type theorems we have established) in that the weak-type operator T satisfies $(Tf)^* \le cS_\sigma(f^*)$ for the appropriate Calderón operator S_σ and this reduces to the desired estimates (6.36) and (6.37) by applying suitable generalizations of Hardy's inequalities. Some special cases of Theorem 6.14 are worth noting (cf. also Exercise 20). The following result is essentially due to A. Zygmund.

Corollary 6.15. *Suppose* $1 \le p_0 < p_1 \le \infty$ *and let* T *be of joint weak type* $(p_0, p_0; p_1, p_1)$. *Then*

$$T: L^{p_0}(\log L)^{\alpha+1} \to L^{p_0}(\log L)^{\alpha}, \qquad \left(\alpha > \frac{-1}{p_0}\right) \tag{6.38}$$

and, if $p_1 < \infty$,

$$T: L^{p_1}(\log L)^{\alpha+1} \to L^{p_1}(\log L)^{\alpha}, \qquad \left(\alpha < \frac{-1}{p_1}\right) \tag{6.39}$$

or, if $p_1 = \infty$,

$$T: L^{\alpha}_{\exp} \to L^{\alpha+1}_{\exp}, \qquad (\alpha \ge 0). \tag{6.40}$$

The fractional integral operator I_λ of order λ, $(0 < \lambda < n)$, on $\mathbf{R}^n$ (Definition 4.17) is of joint weak type $(1, n/(n-\lambda); n/\lambda, \infty)$ (Theorem 4.18). In defining the analogous operators on the unit circle $\mathbf{T}$, additional care must be exercised to ensure that the fractional integral itself is a periodic function. This is the case with the *Weyl fractional integral* f_λ of f of order λ, $(0 < \lambda < 1)$, given by

$$f_\lambda(e^{ix}) = \frac{1}{\Gamma(\lambda)} \int_{-\infty}^{x} f(e^{it})(x-t)^{\lambda-1}\, dt \tag{6.41}$$

for functions f whose mean value $f_I = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{it})\,dt$ over the circle is equal to zero. For arbitrary integrable f, we may set $I_\lambda(f) = (f - f_I)_\lambda + f_I$ and thus obtain a linear operator defined on all of L^1. Essentially the same arguments as before now show that I_λ is of joint weak type $(1, 1/(1-\lambda); 1/\lambda, \infty)$ and for such operators Theorem 6.14 provides the following kinds of estimates.

Corollary 6.16. *Suppose* $1 < p < \infty$ *and* $1/p + 1/p' = 1$. *Let* T *be a quasi-*

linear operator (relative to finite measure spaces) of joint weak type $(1, p; p', \infty)$.
(a) (**R. O'Neil**) *If* $\alpha \ge 1$, *then*

$$T: L(\log L)^{\alpha} \to L^{p,1}(\log L)^{\alpha-1} \tag{6.42}$$

and, if $0 \le \alpha < 1$, *then*

$$T: L(\log L)^{\alpha} \to L^{p,1/\alpha}. \tag{6.43}$$

(b) *If* $\alpha < 1/p$, *then*

$$T: L^{p'}(\log L)^{\alpha} \to L^{1/p-\alpha}_{\exp} \tag{6.44}$$

and, if $1 \le q \le \infty$, *then*

$$T: L^{p',q} \to L^{1/q'}_{\exp}. \tag{6.45}$$

7. FURTHER EXTENSIONS OF THE WEAK-TYPE THEORY

The weak-type theory developed in preceding sections can be generalized in a number of ways. We shall concentrate on two of them, one for bilinear operators satisfying n initial weak-type estimates, and a second in which the weak-type (p,q) conditions themselves are generalized by replacing the associated spaces $L^{p,1}$ and $L^{q,\infty}$ with Lorentz Λ and M spaces. The key to both is the construction of an appropriate generalization of the Calderón operator S_σ.

We begin with the bilinear theory. Let (R,μ), (S,ν) and (W,ω) be totally σ-finite measure spaces. We consider operators T defined for all pairs (f,g), where f is μ-simple on R and g is ν-simple on S, taking values in the ω-measurable functions on W. We shall assume further that T is linear in each of the variables f and g separately. Such operators will be referred to simply as *bilinear operators.*

Definition 7.1. Suppose $1 \le p, q, r \le \infty$. A bilinear operator T is said to be of *restricted weak type* $(p,q;r)$ if there is a constant M such that

$$\sup_{0<t<\infty} \{t^{1/r}[T(\chi_E, \chi_F)]^{**}(t)\} \le M\mu(E)^{1/p}\nu(F)^{1/q} \tag{7.1}$$

for all μ-measurable subsets E of R and all ν-measurable subsets F of S that, in each case, have finite measure.

We shall express f and g in the form

$$f = \sum_{j=1}^{J} a_j\chi_{E_j}, \qquad g = \sum_{k=1}^{K} b_k\chi_{F_k},$$

where $a_j, b_k > 0$, the E_j are subsets of R of finite μ-measure and satisfy $E_1 \supset E_2 \supset \cdots \supset E_J$, and the F_k are subsets of S of finite ν-measure and satisfy $F_1 \supset F_2 \supset \cdots \supset F_K$. Then

$$f^* = \sum_{j=1}^{J} a_j \chi_{(0,\mu(E_j))}, \qquad g^* = \sum_{k=1}^{K} b_k \chi_{(0,\nu(F_k))}.$$

Hence, if T satisfies (7.1) and if both p and q are finite, then

$$\begin{aligned} t^{1/r} T(f,g)^{**}(t) &\le \sum_j \sum_k a_j b_k t^{1/r} T(\chi_{E_j}, \chi_{F_k})^{**}(t) \\ &\le M\Big(\sum_j a_j \mu(E_j)^{1/p}\Big)\Big(\sum_k b_k \nu(F_k)^{1/q}\Big) \\ &\le M \|f\|_{L^{p,1}} \|g\|_{L^{q,1}}. \end{aligned}$$

Applying a standard limiting argument, we obtain the following analogue of Theorem 5.3.

Proposition 7.2. *Suppose $1 \le p$, $q < \infty$ and $1 < r \le \infty$. Then T is of restricted weak type $(p,q;r)$ if and only if T extends uniquely to a bounded bilinear operator from $L^{p,1} \times L^{q,1}$ into $L^{r,\infty}$:*

$$T: L^{p,1} \times L^{q,1} \to L^{r,\infty}.$$

Beginning with n initial estimates of the form

$$t^{1/r_i} T(\chi_E, \chi_F)^{**}(t) \le M_i \mu(E)^{1/p_i} \nu(F)^{1/q_i}, \qquad (i = 1, \ldots, n) \tag{7.2}$$

valid for all sets E and F of finite measure, we shall establish the interpolated result

$$\|T(f,g)\|_{L^{r,c}} \le c \|f\|_{L^{p,a}} \|g\|_{L^{q,b}}, \tag{7.3}$$

where the point $(1/p,\ 1/q,\ 1/r)$ lies in the convex hull of the points $(1/p_i,\ 1/q_i,\ 1/r_i)$, $i = 1, 2, \ldots, n$.

The appropriate Calderón operator is given by

$$S_\sigma(f,g)(t) = \int_0^\infty \int_0^\infty f(u) g(v) \min_{1 \le i \le n} \left\{ \frac{u^{1/p_i} v^{1/q_i}}{t^{1/r_i}} \right\} \frac{du}{u} \frac{dv}{v}, \tag{7.4}$$

where σ is the set of points

$$\sigma = \left\{ \left(\frac{1}{p_i}, \frac{1}{q_i}, \frac{1}{r_i} \right) \right\}_{i=1}^{n}. \tag{7.5}$$

The Hardy-Littlewood inequality II.(2.3) applied twice to (7.4) gives

$$|S_\sigma(f,g)| \le S_\sigma(|f|, |g|) \le S_\sigma(f^*, |g|) \le S_\sigma(f^*, g^*), \tag{7.6}$$

which, among other things, demonstrates that only nonnegative decreasing functions need be considered when estimating S_σ on rearrangement-invariant spaces.

Lemma 7.3. *Let k be any integer satisfying $1 \le k \le n$ and suppose $1 \le p_k, q_k < \infty$ and $1 \le r_k \le \infty$. Suppose $(1/p_k, 1/q_k, 1/r_k) \in \sigma$. If E and F have finite measure, then*

$$t^{1/r_k} S_\sigma(\chi_E, \chi_F)^*(t) \le p_k q_k \mu(E)^{1/p_k} \nu(F)^{1/q_k}. \tag{7.7}$$

If, in addition, $r_k > 1$ holds, then S_σ is of restricted weak type $(p_k, q_k; r_k)$.

Proof. From (7.6) and (7.4), we have

$$\begin{aligned} S_\sigma(\chi_E, \chi_F)^*(t) &\le S_\sigma(\chi_{(0,\mu(E))}, \chi_{(0,\nu(F))})(t) \\ &\le \int_0^{\nu(F)} \int_0^{\mu(E)} \left\{ \frac{u^{1/p_k} v^{1/q_k}}{t^{1/r_k}} \right\} \frac{du}{u} \frac{dv}{v} \\ &= p_k \mu(E)^{1/p_k} q_k \nu(F)^{1/q_k} t^{-1/r_k}, \end{aligned}$$

from which (7.7) follows. If $r_k > 1$, it follows from (4.8) and (7.7) that

$$\begin{aligned} \|S_\sigma(\chi_E, \chi_F)\|_{L^{(r_k,\infty)}} &\le \frac{r_k}{r_k - 1} \|S_\sigma(\chi_E, \chi_F)\|_{L^{r_k,\infty}} \\ &\le \frac{p_k q_k r_k}{r_k - 1} \mu(E)^{1/p_k} \nu(F)^{1/q_k}, \end{aligned}$$

which shows that S_σ is of restricted weak type $(p_k, q_k; r_k)$. ■

Theorem 7.4. *Suppose $1 \le p_i, q_i \le \infty$ and $1 \le r_i \le \infty$ for $i = 1, 2, \ldots, n$. Define σ as in (7.5). If a bilinear operator T satisfies the n initial estimates in (7.2), then*

$$T(f,g)^*(t) \le M S_\sigma(f^*, g^*)(t), \qquad (t > 0), \tag{7.8}$$

for all simple functions f and g, where

$$M = \max_{1 \le i \le n} M_i. \tag{7.9}$$

Proof. Since (7.2) holds for $i = 1, 2, \ldots, n$, we have

$$\begin{aligned} T(\chi_E, \chi_F)^{**}(t) &\le \max_{1 \le i \le n} M_i \cdot \min_{1 \le i \le n} \frac{\mu(E)^{1/p_i} \nu(F)^{1/q_i}}{t^{1/r_i}} \\ &\le M S_\sigma(\chi_E^*, \chi_F^*)(t), \end{aligned} \tag{7.10}$$

where the last estimate holds because

$$\min_i \frac{u^{1/p_i - 1} v^{1/q_i - 1}}{t^{1/r_i}}$$

is decreasing in both u and v. A standard argument (similar to the one following Definition 7.1) now allows us to pass to simple functions f and g in (7.10), and from this the desired result (7.8) follows at once. ■

Corollary 7.5. *With the notation of Theorem 7.4, let X, Y, and Z be rearrangement-invariant spaces for which S_σ is a bounded map*

$$S_\sigma : X \times Y \to Z.$$

If T is of restricted weak types $(p_i, q_i; r_i)$ for $i = 1, 2, \ldots, n$, then

$$\|T(f,g)\|_Z \le c\|f\|_X \|g\|_Y \tag{7.11}$$

for all simple f and g.

Rather than list numerous versions of Corollary 7.5 for the various special classes of function spaces, we shall content ourselves with illustrations in a few specific applications.

Consider first bilinear operators C that satisfy the following inequalities:

$$\|C(f,g)\|_1 \le \|f\|_1 \|g\|_1 \tag{7.12}$$

and

$$\|C(f,g)\|_\infty \le \begin{cases} \|f\|_\infty \|g\|_1 \\ \|f\|_1 \|g\|_\infty \end{cases}. \tag{7.13}$$

The convolution multiplication $C(f,g) = f * g$ on the unit circle $\mathbf{T}$ (cf. Section III.6) is the canonical example of such an operator so we refer to all such operators C as *convolution operators*. It follows from (7.12) that

$$C(f,g)^{**}(t) = \frac{1}{t}\int_0^t C(f,g)^*(s)\,ds \le \frac{1}{t}\|f\|_1 \|g\|_1.$$

Hence, by choosing $f = \chi_E$ and $g = \chi_F$, we see that C is of restricted weak type $(1, 1; 1)$.

From the first of the estimates in (7.13), we have

$$C(f,g)^{**}(t) \le \|C(f,g)\|_\infty \le \|f\|_\infty \|g\|_1,$$

which implies that C is of restricted weak type $(\infty, 1; \infty)$.

Finally, the second estimate in (7.13) gives

$$C(f,g)^{**}(t) \le \|C(f,g)\|_\infty \le \|f\|_1 \|g\|_\infty,$$

and this implies that T is of restricted weak type $(1, \infty; \infty)$.

Theorem 7.4 therefore shows that

$$C(f,g)^{**}(t) \le S_\sigma(f^*, g^*)(t), \qquad (0 \le t \le 1), \tag{7.14}$$

for each pair of simple functions f and g, where

$$\sigma = \{(1,1,1), (0,1,0), (1,0,0)\}. \tag{7.15}$$

In that case, we may write

$$\begin{aligned} S_\sigma(f^*, g^*)(t) &= \int_0^1\int_0^1 f^*(u) g^*(v) \min\left\{\frac{uv}{t}, u, v\right\} \frac{du}{u}\frac{dv}{v} \\ &= t f^{**}(t) g^{**}(t) \\ &\quad + \int_t^1 [f^{**}(u) g^*(u) + f^*(u) g^{**}(u)]\,du \\ &\le 4 \int_{t/2}^1 [u f^{**}(u) g^{**}(u)] \frac{du}{u} \end{aligned} \tag{7.16}$$

Theorem 7.6. *Suppose $1 \le p, q, r, a, b, c \le \infty$ and*

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1, \qquad \frac{1}{c} = \frac{1}{a} + \frac{1}{b}. \tag{7.17}$$

Then any convolution operator C satisfies

$$\|C(f,g)\|_{L^{r,c}} \le c\|f\|_{L^{p,a}} \|g\|_{L^{q,b}}. \tag{7.18}$$

Proof. We assume that r and c are finite. Only routine modifications are necessary to the argument below to accommodate the cases where these parameters are infinite.

Observe that the first of the relations in (7.17) merely asserts that the point $(1/p, 1/q, 1/r)$ lies in the convex hull of the points of σ given by (7.15). Using (7.17), and applying first Hardy's inequality then Hölder's inequality to the estimate (7.16), we obtain

$$\begin{aligned} \|S_\sigma(f^*, g^*)\|_{r,c} &= c\left\{\int_0^1 \left[t^{1/r} \int_{t/2}^1 [u f^{**}(u) g^{**}(u)] \frac{du}{u}\right]^c \frac{dt}{t}\right\}^{1/c} \\ &\le c\left\{\int_0^1 [t^{1+1/r} f^{**}(t) g^{**}(t)]^c \frac{dt}{t}\right\}^{1/c} \le c\|f\|_{p,a} \|g\|_{q,b}. \end{aligned}$$

The desired result (7.18) now follows at once from this one and the previous estimate (7.14). ■

A similar analysis can be performed for *tensor-product operators* $T(f, g)$. Based on the familiar tensor product

$$(f, g) \to f(u)g(v) \equiv (f \otimes g)(u, v),$$

such operators satisfy

$$\|T(f, g)\|_1 \leq \|f\|_1 \|g\|_1 \tag{7.19}$$

and

$$\|T(f, g)\|_\infty \leq \|f\|_\infty \|g\|_\infty. \tag{7.20}$$

Clearly, tensor-product operators T are of restricted weak types $(1, 1; 1)$ and $(\infty, \infty; \infty)$ and hence, by Theorem 7.4, satisfy

$$\begin{aligned} T(f, g)^{**}(t) &\leq S_\sigma(f^*, g^*)(t) \\ &= \int_0^\infty \int_0^\infty f^*(u)g^*(v) \min\left\{\frac{uv}{t}, 1\right\} \frac{du}{u}\frac{dv}{v}, \end{aligned} \tag{7.21}$$

where σ is the set

$$\sigma = \{(0, 0, 0), (1, 1, 1)\}. \tag{7.22}$$

Theorem 7.7. *Suppose $1 < p < \infty$ and $1 \leq a, b, c \leq \infty$ with*

$$\frac{1}{a} + \frac{1}{b} = \frac{1}{c} + 1. \tag{7.23}$$

Then any tensor-product operator T satisfies

$$\|T(f, g)\|_{L^{p,c}} \leq c\|f\|_{L^{p,a}} \|g\|_{L^{p,b}}. \tag{7.24}$$

Proof. With σ as in (7.22), a simple computation involving (7.21) shows that

$$S_\sigma(f^*, g^*)(t) = \int_0^\infty S(f^*)\left(\frac{t}{s}\right) g^*(s) \frac{ds}{s}, \tag{7.25}$$

where S is the one-variable Calderón operator for the segment $(1, 1; 0, 0)$:

$$(Sf)(t) = \int_0^\infty f(s) \min\left(\frac{s}{t}, 1\right) \frac{ds}{s}, \qquad (t > 0).$$

Hence, $t^{1/p} S_\sigma(f^*, g^*)(t)$ may be viewed as the convolution of the functions

$F(s) = s^{1/p} S(f^*)(s)$ and $G(s) = s^{1/p} g^*(s)$ with respect to the multiplicative group of positive real numbers (whose Haar measure is ds/s). We therefore have from Young's inequality (Theorem 2.4, or more generally, Theorem 7.6), together with (7.21) and (7.25),

$$\begin{aligned} \|T(f, g)\|_{L^{p,c}} &\leq \|S_\sigma(f^*, g^*)\|_{L^{p,c}} = \|F * G\|_{L^c(\frac{dt}{t})} \\ &\leq \|F\|_{L^a(\frac{dt}{t})} \|G\|_{L^b(\frac{dt}{t})} \\ &= \|S(f^*)\|_{L^{p,a}} \|g\|_{L^{p,b}}. \end{aligned}$$

The operator S is bounded on $L^{p,a}$ since $1 < p < \infty$ (Proposition III.5.5 and Theorem 4.13). Hence, the last estimate reduces to (7.24). This completes the proof. ■

We conclude this section with a version of the weak-type theory in which the weak-type conditions themselves are generalized to rearrangement-invariant spaces.

Definition 7.8. Let X and Y be rearrangement-invariant spaces with respective fundamental functions φ_X and φ_Y (Section II.5). A sublinear operator T is said to be of *weak type* (X, Y) if T maps the Lorentz space $\Lambda(X)$ into the measurable functions $\mathscr{M}$ and there is a constant $c > 0$ such that

$$\sup_{t>0}\{(Tf)^*(t)\varphi_Y(t)\} \leq c\int_0^\infty f^*(s)\, d\varphi_X(s), \qquad (f \in \Lambda(X)). \tag{7.26}$$

Denoting by $M^*(Y)$ the set of functions g for which

$$\|g\|_{M^*(Y)} \equiv \|g^* \varphi_Y\|_\infty < \infty,$$

we may then rewrite (7.26) in the form

$$\|Tf\|_{M^*(Y)} \leq c\|f\|_{\Lambda(X)}, \qquad (f \in \Lambda(X)). \tag{7.27}$$

When $X = L^p$, $(p < \infty)$, and $Y = L^q$, the above notion of weak type (X, Y) coincides with the standard notion of weak type (p, q). Moreover, for any rearrangement-invariant space Z, we have from Theorem II.5.13,

$$\Lambda(Z) \hookrightarrow Z \hookrightarrow M(Z) \hookrightarrow M^*(Z).$$

Hence, any bounded operator from X into Y is also of weak type (X, Y).

Given two pairs of rearrangement-invariant spaces (X_j, Y_j), $j = 0, 1$, let

$$\Psi(s, t) = \min_{j=0,1}\left\{\frac{\varphi_{X_j}(s)}{\varphi_{Y_j}(t)}\right\}, \qquad (s, t > 0). \tag{7.28}$$

As a function of s, $\Psi(s,t)$ is increasing and $\Psi(s,t)/s$ is decreasing. Furthermore, $\Psi(s,t)$ is an absolutely continuous function of s on each interval of the form (ε,∞), $(\varepsilon>0)$, and

$$\Phi(s,t)=\frac{\partial\Psi}{\partial s}(s,t)\le\frac{\Psi(s,t)}{s}\qquad\text{a.e.}$$

Define the *Calderón operator* S_σ for the interpolation segment $\sigma=[(X_0,Y_0),(X_1,Y_1)]$ by

$$S_\sigma f(\cdot)=\int_0^\infty f(s)\,d\Psi(s,\cdot).$$

If either of the fundamental functions of X_0 and X_1 is continuous from the right at the origin, that is, if $\Psi(0+,\cdot)=0$, then S_σ is given by the improper integral

$$S_\sigma f(t)=\int_0^\infty f(s)\Phi(s,t)\,ds,\qquad(t>0).\tag{7.29}$$

As with previous versions of the Calderón operator, it is routine to show that $S_\sigma(f^*)$ is a decreasing function and that $(S_\sigma f)^*(t)\le S_\sigma(f^*)(t)$ for all $t>0$. The next result shows that S_σ satisfies the appropriate weak-type conditions at the endpoints of the interpolation segment σ.

Lemma 7.9. *If $\Psi(0+,\cdot)=0$, then S_σ is of weak types (X_0,Y_0) and (X_1,Y_1).*

Proof. For each ε, $N>0$, an integration by parts gives

$$\begin{aligned}
&\int_\varepsilon^N f^*(s)\,d\Psi(s,t)\\
&\quad=\int_\varepsilon^N \Psi(s,t)\,d[-f^*(s)]+f^*(s)\Psi(s,t)\Big|_{s=\varepsilon}^{N}\\
&\quad\le\frac{1}{\varphi_{Y_j}(t)}\int_\varepsilon^N \varphi_{X_j}(s)\,d[-f^*(s)]+f^*(s)\Psi(s,t)\Big|_{s=\varepsilon}^{N}\\
&\quad\le\int_0^\infty f^*(s)\frac{d\varphi_{X_j}(s)}{\varphi_{Y_j}(t)},\qquad(j=0,1).
\end{aligned}$$

Since Ψ is right-continuous (in s) at the origin, the result follows by letting $\varepsilon\to0+$ and $N\to\infty$. ■

We shall say that a pair (X,Y) of rearrangement-invariant Banach function spaces has the *weak-interpolation property* with respect to σ if every linear operator of weak types (X_j,Y_j), $(j=0,1)$, has a unique extension to a bounded linear operator from X into Y.

Theorem 7.10. *Suppose $\Psi(0+,\cdot)=0$. A necessary and sufficient condition for a pair (X,Y) to have the weak-interpolation property for σ is that S_σ be bounded from X to Y. Furthermore, every sublinear operator T of weak types (X_j,Y_j), $(j=0,1)$, satisfies*

$$(Tf)^*(t)\le MS_\sigma(f^*)(t),\qquad(t>0),\tag{7.30}$$

for all f in $\Lambda(X_0)+\Lambda(X_1)$, where M is the larger of the two weak-type norms.

In order to prove Theorem 7.10, we shall need the following lemma.

Lemma 7.11. *Suppose f is a nonnegative decreasing function on $(0,\infty)$ and let $t>0$. Then there exist nonnegative decreasing functions f_0 and f_1 such that $f=f_0+f_1$ and*

$$S_\sigma(f_j)(t)=\frac{\|f_j\|_{\Lambda(X_j)}}{\varphi_{Y_j}(t)},\qquad(j=0,1).\tag{7.31}$$

Proof. Since $\Psi(0+,\cdot)=0$, we may assume without loss of generality that $\varphi_{X_1}(0+)=0$. Set

$$E=\left\{s:\frac{\varphi_{X_1}(s)}{\varphi_{X_0}(s)}>\frac{\varphi_{Y_1}(t)}{\varphi_{Y_0}(t)}\right\}.\tag{7.32}$$

Then E is a countable disjoint union $\bigcup_k(a_k,b_k)$ of intervals, and, by continuity, equality is attained in the defining inequality in (7.32) at each of the endpoints of the intervals (a_k,b_k). Define g_k by

$$g_k(s)=\begin{cases}f(a_k)-f(b_k), & 0<s<a_k,\\ f(s)-f(b_k), & a_k\le s\le b_k,\\ 0, & b_k<s,\end{cases}$$

and set

$$f_0=\sum_k g_k,\qquad f_1=f-f_0.$$

It is clear that both f_0 and $f_1\,(=\lim_{n\to\infty}(f-\sum_{k=1}^n g_k))$ are nonnegative decreasing since the g_k are and the intervals (a_k,b_k) are disjoint. We have

$$\begin{aligned}
S_\sigma(g_k)(t)&=g_k(a_k)\frac{\varphi_{X_0}(a_k)}{\varphi_{Y_0}(t)}+\int_{a_k}^{b_k}g_k(s)\frac{d\varphi_{X_0}(s)}{\varphi_{Y_0}(t)}\\
&=\frac{1}{\varphi_{Y_0}(t)}\int_0^\infty g_k(s)\,d\varphi_{X_0}(s).
\end{aligned}$$

Hence, a summation over k gives

$$S_\sigma(f_0)(t) = \frac{1}{\varphi_{Y_0}(t)} \|f_0\|_{\Lambda(X_0)}.$$

Now consider the case $j = 1$. The complement F of the set E is closed and we have

$$\Phi(s,t) = \frac{\varphi'_{X_1}(s)}{\varphi_{Y_1}(t)} \qquad \text{a.e. on } F.$$

Hence,

$$S_\sigma(f_1\chi_F)(t) = \int_F f_1(s) \frac{\varphi'_{X_1}(s)}{\varphi_{Y_1}(t)}\, ds. \tag{7.33}$$

Also, since f_1 is constant on each of the intervals (a_k, b_k),

$$S_\sigma(f_1\chi_E)(t) = \sum_k f_1(a_k) \int_{a_k}^{b_k} \Phi(s,t)\, ds = \sum_k f_1(a_k)\Psi(s,t)\Big|_{s=a_k}^{b_k}$$

$$= \sum_k f_1(a_k) \frac{\varphi_{X_1}(s)}{\varphi_{Y_1}(t)}\Big|_{s=a_k}^{b_k} = \frac{\int_E f_1(s)\, d\varphi_{X_1}(s)}{\varphi_{Y_1}(t)}.$$

Hence, combining this estimate with the one in (7.33), we obtain (7.31) in the case $j = 1$. ■

Proof of Theorem 7.10. The necessity follows immediately from Lemma 7.9. For the sufficiency, suppose that S_σ is bounded from X into Y. Then, for all f in X and all $t > 0$,

$$\varphi_Y\left(\frac{t}{2}\right) S_\sigma f^*\left(\frac{t}{2}\right) \le \|S_\sigma f^*\|_Y \le c\|f\|_X. \tag{7.34}$$

Applying Lemma 7.11 to the function f^* and the value $t/2$, we obtain nonnegative decreasing functions g_0 and g_1 with sum equal to f^* and

$$S_\sigma g_j\left(\frac{t}{2}\right) = \frac{\|g_j\|_{\Lambda(X_j)}}{\varphi_{Y_j}\left(\frac{t}{2}\right)}, \qquad (j = 0, 1). \tag{7.35}$$

The estimates (7.34) and (7.35) together show that f^* belongs to $\Lambda(X_0) + \Lambda(X_1)$. By Corollary II.7.6, there exists a measure-preserving transformation τ such that $|f| = f^* \circ \tau$. Thus, if we define f_j, $(j = 0, 1)$, by $f_j = (g_j \circ \tau)\operatorname{sgn} f$, then $f = f_0 + f_1$ and $f_j^* = g_j$. Hence, if T is an operator of weak types (X_j, Y_j), $(j = 0, 1)$, we see from inequality II.1.16 that

$$(Tf)^*(t) \le (Tf_0)^*\left(\frac{t}{2}\right) + (Tf_1)^*\left(\frac{t}{2}\right)$$

$$\le M\left\{\frac{\|f_0\|_{\Lambda(X_0)}}{\varphi_{Y_0}\left(\frac{t}{2}\right)} + \frac{\|f_1\|_{\Lambda(X_1)}}{\varphi_{Y_1}\left(\frac{t}{2}\right)}\right\}. \tag{7.36}$$

Appealing to (7.35) and using the fact that $f_j^* = g_j$, we therefore obtain

$$(Tf)^*(t) \le M[S_\sigma g_0(t) + S_\sigma g_1(t)] \le M S_\sigma f^*(t),$$

which establishes (7.31). Finally, by hypothesis,

$$\|Tf\|_Y \le M\|S_\sigma f^*\|_Y \le c\|f^*\|_X \le c\|f\|_X,$$

so that T is bounded from X to Y. This shows that (X, Y) has the weak-interpolation property and hence completes the proof. ■

We remark that one can also develop a corresponding theory of interpolation of operators of *restricted* weak type (X, Y) in this general setting.

8. ORLICZ SPACES

We conclude this chapter with a brief look at another interesting class of rearrangement-invariant spaces: the *Orlicz spaces.* In view of the results of Chapter II, it will suffice to consider measure spaces that are either a finite or infinite interval $[0, a)$, $(0 \le a \le \infty)$, with Lebesgue measure, or the discrete set $\{1, 2, \ldots\}$ with counting measure. Since the differences between these various cases involve detail rather than principle, we shall confine our attention here exclusively to the case of a finite interval.

Definition 8.1. Let $\phi: [0, \infty) \to [0, \infty]$ be increasing and left-continuous, with $\phi(0) = 0$. Suppose on $(0, \infty)$ that ϕ is neither identically zero nor identically infinite[1]. Then the function Φ defined by

$$\Phi(s) = \int_0^s \phi(u)\, du, \qquad (s \ge 0) \tag{8.1}$$

is said to be a *Young's function.*

[1] In particular, if ϕ assumes infinite values, then there exists s_∞ with $0 < s_\infty < \infty$ such that $\phi(s) < \infty$ for $0 \le s < s_\infty$ and $\phi(s) = \infty$ for $s > s_\infty$. If $\phi(s_\infty) = \infty$, the left-continuity at s_∞ is interpreted to mean that $\phi(s) \uparrow \infty$ as $s \to s_\infty^-$.

Note that a Young's function is convex on the interval where it is finite.

Definition 8.2. Let Φ be a Young's function. The *Orlicz class* $P(\Phi)$ consists of all measurable functions f on the interval $[0, a]$ for which the functional

$$M^{\Phi}(f) = \int_0^a \Phi(|f(x)|)\,dx \tag{8.2}$$

is finite.

Examples 8.3. (a) If $\phi(u) = pu^{p-1}$, where $1 \le p < \infty$, then $\Phi(s) = s^p$. The Orlicz class $P(\Phi)$ in this case is the Lebesgue space L^p.
(b) If $\phi(u) = 0$, $(0 \le u \le 1)$, and $\phi(u) = 1 + \log u$, $(1 < u < \infty)$, then $\Phi(s) = s\log^+ s$, $(0 \le s \le \infty)$. The Orlicz class $P(\Phi)$ in this case is the Zygmund space $L\log L$ (Definition 6.1).
(c) If $\phi(u) = 0$, $(0 \le u \le 1)$, and $\phi(u) = \infty$, $(u > 1)$, then $\Phi = \phi$ and the Orlicz class $P(\Phi)$ is the unit ball of L^∞.
(d) If

$$\phi(u) = \begin{cases} 0, & u = 0, \\ 1, & 0 < u \le 1, \\ e^{u-1}, & 1 < u < \infty, \end{cases}$$

then

$$\Phi(s) = \begin{cases} s, & 0 \le s \le 1, \\ e^{s-1}, & 1 < s < \infty. \end{cases}$$

The Orlicz class $P(\Phi)$, which consists of all f for which $\exp(|f|)$ is integrable, is then a proper subset of the Zygmund space $L_{\exp}$ (Definition 6.1).
(e) The Zygmund spaces $L^p(\log L)^\alpha$, $(\alpha \ge 1/p)$, (Definition 6.11) are realizable as Orlicz classes $P(\Phi)$ for suitable choice of Φ.

The Orlicz classes $P(\Phi)$ in parts (c) and (d) of the last example are not linear spaces; this is clear in part (c), and in part (d) we note that $P(\Phi)$ contains, for example, the function $f(x) = \log^+(x^{-1/2})$ but not the function $2f$. As we shall see, the problem is that their Young's functions grow too fast.

Definition 8.4. A Young's function Φ is said to satisfy the Δ_2-*condition* if there exist $s_0 > 0$ and $c > 0$ such that

$$\Phi(2s) \le c\Phi(s) < \infty, \qquad (s_0 \le s < \infty). \tag{8.3}$$

It is easy to verify that the Young's functions in Examples 2.3(a) and (b) satisfy the Δ_2-condition but those in parts (c) and (d) do not.

Proposition 8.5. *Let Φ be a Young's function. Then the Orlicz class $P(\Phi)$ is a linear subspace of $\mathscr{M}_0([0, a])$ if and only if Φ satisfies the Δ_2-condition.*

Proof. Suppose Φ satisfies the Δ_2-condition and let s_0 and c be the constants prescribed by Definition 8.4. We claim that

$$f \in P(\Phi) \quad \Rightarrow \quad 2f \in P(\Phi). \tag{8.4}$$

To see this, let $E = \{x \in [0, a] : |f(x)| \ge s_0\}$ and $F = E^c$. Then

$$\int_0^a \Phi(2|f(x)|)\,dx \le \int_E + \int_F \Phi(2|f(x)|)\,dx$$
$$\le c\left\{\int_E \Phi(|f(x)|)\,dx + a\Phi(s_0)\right\} < \infty,$$

which establishes (8.4).

Suppose now that $f \in P(\Phi)$ and let α be any scalar. Choosing an integer n with $2^n \ge |\alpha|$, we see by iterating (8.4) that $2^n f \in P(\Phi)$. Hence, since Φ is increasing, $M^{\Phi}(\alpha f) \le M^{\Phi}(2^n f) < \infty$, and so $\alpha f \in P(\Phi)$.

On the other hand, if f and g belong to $P(\Phi)$, the convexity of Φ and the fact that Φ is increasing combine to give

$$M^{\Phi}(f + g) = M^{\Phi}\left(\frac{1}{2}(2f) + \frac{1}{2}(2g)\right) \le \frac{1}{2}M^{\Phi}(2f) + \frac{1}{2}M^{\Phi}(2g).$$

The latter quantity is finite by virtue of (8.4), so we conclude that $f + g$ belongs to $P(\Phi)$. Hence, $P(\Phi)$ is a linear subspace.

Suppose conversely that $P(\Phi)$ is a linear subspace. It is clear that $\Phi(s)$ is finite for all s since otherwise, if Φ first becomes infinite beyond the value $s = s_\infty$, say, then the constant function $f \equiv (3s_\infty)/4$ would belong to $P(\Phi)$ but $2f$ would not.

If the Δ_2-condition fails, then there is a sequence $s_n \uparrow \infty$ such that

$$\Phi(2s_n) > 2^n\Phi(s_n) > 0, \qquad (n = 1, 2, \ldots).$$

Choose disjoint subsets E_n of $[0, a]$ with measure

$$|E_n| \equiv t_n = \frac{a\Phi(s_1)}{2^n\Phi(s_n)}, \qquad (n = 1, 2, \ldots).$$

This is possible because $t_n \le 2^{-n}a$ and so $\sum t_n \le a$. The function f defined by

$$f(x) = \sum_{n=1}^{\infty} s_n \chi_{E_n}(x), \qquad (0 \le x \le a)$$

then belongs to $P(\Phi)$ because

$$\int_0^a \Phi(|f(x)|)\,dx = \sum_1^\infty \Phi(s_n)t_n \le a\Phi(s_1)\sum_1^\infty 2^{-n} = a\Phi(s_1) < \infty.$$

The function $2f$, on the other hand, does not belong to $P(\Phi)$ because

$$\int_0^a \Phi(2|f(x)|)\,dx = \sum_1^\infty \Phi(2s_n)t_n \ge a\sum_1^\infty \Phi(s_1) = \infty.$$

This contradicts the fact that $P(\Phi)$ is a linear space and hence shows that Φ satisfies the Δ_2-condition. ■

Although the Orlicz class $P(\Phi)$ may fail to be linear, it is always a convex set that contains the origin. It is natural therefore to use the associated Minkowski functional in order to generate a linear space with a norm topology. The following approach, due to W. A. J. Luxemburg, begins with the construction of the appropriate Banach function norm.

Definition 8.6. If Φ is a Young's function, the corresponding *Luxemburg norm* ρ^Φ is defined by

$$\rho^\Phi(f) = \inf\{k^{-1} : M^\Phi(kf) \le 1\}, \qquad (f \in \mathscr{M}_0^+([0,a])). \tag{8.5}$$

Since Φ is left-continuous, it follows from the monotone convergence theorem that the infimum in (8.5), if positive, is actually attained. Hence, if $\rho^\Phi(f) > 0$, then

$$\rho^\Phi(f) = \min\{k^{-1} : M^\Phi(kf) \le 1\}, \qquad (f \in \mathscr{M}_0^+([0,a])). \tag{8.6}$$

In order to show that ρ^Φ is indeed a Banach function norm, we shall need the following preliminaries.

Lemma 8.7. *If Φ is a Young's function, then*

$$f = 0 \quad \text{a.e.} \quad \Leftrightarrow \quad M^\Phi(kf) \le 1 \qquad \text{for all } k > 0.$$

Proof. The forward implication is obvious. In the reverse direction, suppose that $M^\Phi(kf) \le 1$ for all $k > 0$, but, for some $\varepsilon > 0$, we have $|f| \ge \varepsilon$ on a set E of positive measure. Then $M^\Phi(kf) \ge \int_E \Phi(k\varepsilon)\,dx = \Phi(k\varepsilon)|E|$. Since $\Phi(s) \uparrow \infty$ as $s \uparrow \infty$, we therefore obtain the contradiction that $M^\Phi(kf) \uparrow \infty$ as $k \uparrow \infty$. ■

Lemma 8.8. *If Φ is a Young's function, then*

$$\rho^\Phi(|f|) \le 1 \quad \Rightarrow \quad M^\Phi(f) \le \rho^\Phi(|f|) \tag{8.7}$$

and

$$\rho^\Phi(|f|) > 1 \quad \Rightarrow \quad \rho^\Phi(|f|) \le M^\Phi(f). \tag{8.8}$$

Consequently,

$$\rho^\Phi(|f|) \le 1 \quad \Leftrightarrow \quad M^\Phi(f) \le 1. \tag{8.9}$$

Proof. It suffices to consider $f \ge 0$. Suppose $\rho^\Phi(f) \le 1$. If $\rho^\Phi(f) = 0$, it follows from (8.5) and Lemma 8.7 that $f = 0$ a.e. and hence that $M^\Phi(f) = 0$. On the other hand, if $\rho^\Phi(f)$ is positive, then (8.6) shows that $M^\Phi(k_0 f) \le 1$ with $k_0 = 1/\rho^\Phi(f)$. Since $k_0 \ge 1$ and $\Phi(s)/s$ is increasing, we have $k_0\Phi(s) \le \Phi(k_0 s)$. Hence, $k_0 M^\Phi(f) \le M^\Phi(k_0 f) \le 1$, and this establishes (8.7).

Now suppose $\rho^\Phi(f) > 1$. For each γ with $1 < \gamma < \rho^\Phi(f)$, it follows from (8.5) that $M^\Phi(f/\gamma) > 1$. Since $\Phi(s)/s$ increases and $\gamma > 1$, we have $\gamma\Phi(s/\gamma) \le \Phi(s)$ and so $1 < M^\Phi(f/\gamma) \le M^\Phi(f)/\gamma$. This shows that $M^\Phi(f) > \gamma$, and hence, because of the arbitrary nature of $\gamma < \rho^\Phi(f)$, that (8.8) holds.

The remaining assertion in (8.9) follows at once from those in (8.7) and (8.8). ■

Theorem 8.9. *If Φ is a Young's function, then ρ^Φ is a rearrangement-invariant Banach function norm.*

Proof. We need to verify properties (P1)–(P5) of Definition I.1.1 for ρ^Φ. That $\rho^\Phi(f) = 0 \Leftrightarrow f = 0$ a.e. follows from (8.5) and Lemma 8.7. Homogeneity of ρ^Φ is obvious. Hence, (P1) will be established once we prove the triangle inequality. Let f and g be nonzero functions in $\mathscr{M}_0^+$ with $\gamma = \rho^\Phi(f) + \rho^\Phi(g)$ finite. Let $\alpha = \rho^\Phi(f)/\gamma$ and $\beta = \rho^\Phi(g)/\gamma$, so $\alpha, \beta > 0$ and $\alpha + \beta = 1$. By (8.5), the constants $M^\Phi(f/\rho^\Phi(f))$ and $M^\Phi(g/\rho^\Phi(g))$ do not exceed 1; in particular, $\Phi(f/\rho^\Phi(f))$ and $\Phi(g/\rho^\Phi(g))$ are finite a.e. Since Φ is convex on the interval where it is finite, we therefore have

$$\begin{aligned} M^\Phi\left(\frac{f+g}{\gamma}\right) &= M^\Phi\left(\frac{\alpha f}{\rho^\Phi(f)} + \frac{\beta g}{\rho^\Phi(g)}\right) \\ &\le \alpha M^\Phi\left(\frac{f}{\rho^\Phi(f)}\right) + \beta M^\Phi\left(\frac{g}{\rho^\Phi(g)}\right) \\ &\le \alpha + \beta = 1. \end{aligned}$$

Hence, we conclude from (8.5) that

$$\rho^\Phi(f+g) \le \gamma \le \rho^\Phi(f) + \rho^\Phi(g),$$

and, with this, property (P1) is established.

To establish the lattice property (P2), consider measurable functions f and g with $0 \le g \le f$ a.e. and $0 < \rho^\Phi(f) < \infty$. Then

$$M^\Phi\left(\frac{g}{\rho^\Phi(f)}\right) \le M^\Phi\left(\frac{f}{\rho^\Phi(f)}\right) \le 1,$$

and (8.5) shows that $\rho^\Phi(g) \le \rho^\Phi(f)$.

For the Fatou property (P3), suppose $0 \le f_n \uparrow f$ a.e. Then, by (P2), the sequence $\rho^\Phi(f_n)$ is increasing. Let $\alpha_n = \rho^\Phi(f_n)$ and put $\alpha = \sup_n \alpha_n$. Since $\rho^\Phi(f) \ge \alpha_n$ for each n, it follows that $\rho^\Phi(f) \ge \alpha$. We need to show that equality holds. This is clear for $\alpha = 0$ or $\alpha = \infty$, so we may assume that $0 < \alpha_n \le \alpha < \infty$, $(n = 1, 2, \ldots)$. In that case,

$$M^\Phi\left(\frac{f_n}{\alpha}\right) \le M^\Phi\left(\frac{f_n}{\alpha_n}\right) \le 1,$$

and the monotone convergence theorem shows that the quantity on the left converges to $M^\Phi(f/\alpha)$. Hence, $M^\Phi(f/\alpha) \le 1$ and now the desired conclusion $\rho^\Phi(f) \le \alpha$ follows from (8.5).

Property (P4) requires that $\rho^\Phi(\chi_E)$ be finite for any measurable subset of $[0, a]$. Let b denote the measure of E. We may assume $b > 0$. By hypothesis, the Young's function Φ is not identically infinite on $(0, \infty)$, and is continuous on the interval where it is finite. Since $\Phi(0) = 0$, it follows that there is a number $k > 0$ for which $\Phi(k) \le 1/b$. Then $M^\Phi(k\chi_E) = b\Phi(k) \le 1$, and it follows from (8.5) that $\rho^\Phi(\chi_E) \le 1/k < \infty$.

For property (P5), again let E be a subset of $[0, a]$ of measure $b > 0$, say. It suffices to consider measurable functions $f \ge 0$ on $[0, a]$ with $0 < \rho^\Phi(f) < \infty$. With $k = 1/\rho^\Phi(f)$, Jensen's inequality gives

$$\Phi\left(\frac{1}{b}\int_E kf(x)\,dx\right) \le \frac{1}{b}\int_E \Phi(kf(x))\,dx \le \frac{1}{b}M^\Phi(kf) \le \frac{1}{b}.$$

Hence, since Φ increases to ∞, there is a constant $c = c(\Phi, b)$ such that $(1/b)\int_E kf(x)\,dx \le c$, that is $\int_E f \le bc/k = bc\rho^\Phi(f)$. This establishes property (P5), and hence completes the proof that ρ^Φ is a Banach function norm.

The rearrangement-invariance follows from (8.5) and the fact that $M^\Phi(f) = M^\Phi(g)$ whenever f and g are equimeasurable. The latter property need only be established with $g = f^*$; that case is easily verified for simple functions f, and the passage to general f is then achieved by means of a limiting argument using the monotone convergence theorem. The proof is now complete. ■

Definition 8.10. Let Φ be a Young's function. The *Orlicz space* L^Φ is the rearrangement-invariant Banach function space generated by ρ^Φ. Thus L^Φ

consists of those measurable functions f on $[0, a]$ for which $\rho^\Phi(|f|) < \infty$. The *Luxemburg norm* on L^Φ is given by

$$\|f\|_{L^\Phi} \equiv \rho^\Phi(|f|), \qquad (f \in L^\Phi). \tag{8.10}$$

In Examples 8.3(a) and (c), it is easy to see directly that the Orlicz spaces L^Φ are precisely the L^p-spaces, and that the Luxemburg norm (8.10) coincides with the L^p-norm. On the other hand, while the Orlicz space L^Φ in part (b) evidently coincides with the Zygmund space $L\log L$, the exact relationship between the Luxemburg norm in (8.10) and the Hardy-Littlewood norm in (6.9) is not clear. However, because of Corollary I.1.9 and Lemma 6.2, we can assert that the two norms are equivalent. A similar remark applies to the space $L_{\exp}$.

Recall from Theorem 6.5 that the spaces $L\log L$ and $L_{\exp}$ are mutually associate. The following notion will enable us to derive such relationships directly from the Orlicz-space structure.

Definition 8.11. Let Φ be a Young's function, represented as in (8.1) as the indefinite integral of Φ. Let

$$\psi(v) = \inf\{u : \phi(u) \ge v\}, \qquad (0 \le v \le \infty). \tag{8.11}$$

Then the function

$$\Psi(t) = \int_0^t \psi(v)\,dv, \qquad (0 \le t \le \infty) \tag{8.12}$$

is called the *complementary Young's function* of Φ.

The function ψ has the same properties as ϕ: it is increasing, left-continuous, vanishes at the origin, and on $(0, \infty)$ is neither identically zero nor identically infinite. Hence, the complementary Young's function Ψ defined by (8.12) is indeed a Young's function in the sense of Definition 8.1. The function ψ is called the *left-continuous inverse* of ϕ. It is easy to verify that ϕ is the left-continuous inverse of ψ:

$$\phi(u) = \inf\{v : \psi(v) \ge u\}, \qquad (0 \le u \le \infty), \tag{8.13}$$

and hence that Φ is the complementary Young's function of Ψ. Thus, we may refer to Φ and Ψ simply as *complementary Young's functions*. The following basic result is known as *Young's inequality*.

Theorem 8.12 (W. H. Young). *Let Φ and Ψ be complementary Young's functions. Then*

$$st \le \Phi(s) + \Psi(t), \qquad (0 \le s, t < \infty). \tag{8.14}$$

Furthermore, equality holds in (8.14) *if and only if either* $t = \phi(s)$ *or* $s = \psi(t)$ *holds.*

Proof. Fix $s, t \geq 0$. We may assume that $\Phi(s)$ and $\Psi(t)$ are finite since otherwise there is nothing to prove. In that case, ϕ and ψ are finite-valued on the respective intervals $[0, s)$ and $[0, t)$. We distinguish the three cases $t = \phi(s)$, $t > \phi(s)$, and $t < \phi(s)$.
Case I: $t = \phi(s)$. Let $P = \{u_0, u_1, \ldots, u_n\}$, where $0 = u_0 < u_1 < \ldots < u_n = s$, be a partition of $[0, s]$. Let $v_i = \phi(u_i)$, $i = 1, 2, \ldots, n$. Since ϕ is increasing, the set $Q = \{v_0, v_1, \ldots, v_n\}$ is a partition of $[0, t]$ (it may be that $v_i = v_{i+1}$ for some values of i, but this will not affect the values of the Riemann sums below).

Let $\bar{\psi}$ denote the function ψ modified at each of its jump discontinuities so as to be right-continuous rather than left-continuous. Thus, at each point v where ψ jumps, we replace its value $\psi(v)$ by $\psi(v+)$. Since ψ is increasing and therefore has at most countably many jump discontinuities, the integral of $\bar{\psi}$ will remain the same as that of ψ.

Select a lower Riemann sum for ϕ and an upper Riemann sum for $\bar{\psi}$ as follows:

$$L(\phi) = \sum_{i=1}^{n} \phi(u_{i-1})(u_i - u_{i-1}), \qquad U(\bar{\psi}) = \sum_{i=1}^{n} \psi(v_i+)(v_i - v_{i-1}).$$

Since $\phi(u_i) < v_i + \varepsilon$ for each $\varepsilon > 0$, it follows from (8.11) that $u_i \leq \psi(v_i + \varepsilon)$. Letting $\varepsilon \to 0+$, we conclude that $u_i \leq \psi(v_i+)$. Hence,

$$\begin{aligned} L(\Phi) + U(\bar{\psi}) &\geq \sum_{i=1}^{n} \{\phi(u_{i-1})(u_i - u_{i-1}) + u_i(\phi(u_i) - \phi(u_{i-1}))\} \\ &= -u_0\phi(u_0) + u_n\phi(u_n) = s\phi(s) = st. \end{aligned}$$

Next, select upper and lower sums for ϕ and ψ, respectively, as follows:

$$U(\phi) = \sum_{i=1}^{n} \phi(u_i)(u_i - u_{i-1}), \qquad L(\psi) = \sum_{i=1}^{n} \psi(v_{i-1})(v_i - v_{i-1}).$$

Since $v_{i-1} = \phi(u_{i-1})$, it follows at once from (8.11) that $\psi(v_{i-1}) \leq u_{i-1}$. Hence,

$$\begin{aligned} U(\phi) + L(\psi) &\leq \sum_{i=1}^{n} \{\phi(u_i)(u_i - u_{i-1}) + u_{i-1}(\phi(u_i) - \phi(u_{i-1})\} \\ &= -u_0\phi(u_0) + u_n\phi(u_n) = s\phi(s) = st. \end{aligned}$$

We therefore have

$$U(\phi) + L(\psi) \leq st \leq L(\phi) + U(\bar{\psi}).$$

Now let the mesh of the partition P tend to 0. In order to have the mesh

of Q also tend to zero, it will be necessary to add points to the partition in the intervals of constancy of ψ; clearly, this does not affect the value of the corresponding Riemann sum. Since ϕ, ψ, and $\bar{\psi}$ are increasing, their Riemann integrals exist as the limits of Riemann sums. Hence, we obtain from the estimates above

$$\int_0^s \phi(u)\,du + \int_0^t \psi(v)\,dv \leq st \leq \int_0^s \phi(u)\,du + \int_0^t \bar{\psi}(v)\,dv.$$

However, as we remarked above, the integrals of ψ and $\bar{\psi}$ coincide. Hence, equality obtains in the preceding estimates, and therefore also in (8.14).
Case II: $t > \phi(s)$. Then with $t_0 = \phi(s)$, the result from Case I shows that $\Phi(s) + \Psi(t_0) = st_0$. Hence,

$$\begin{aligned} \Phi(s) + \Psi(t) &= \Phi(s) + \Psi(t_0) + \int_{t_0}^{t} \psi(v)\,dv \\ &\geq st_0 + (t - t_0)\psi(t_0+) \geq st_0 + (t - t_0)s = st. \end{aligned}$$

Equality can occur only if ψ assumes the constant value s on the interval (t_0, t). But then the left continuity of ψ ensures that $\psi(t) = s$.
Case III: $t < \phi(s)$. It follows from (8.13) that $\psi(t) < s$, and now the desired result follows from the argument in Case II with ϕ and ψ interchanged. This completes the proof. ■

With Φ and Ψ complementary Young's functions, consider the functional

$$f \to \sup\left\{\int_0^a |f(x)g(x)|\,dx : M^{\Psi}(g) \leq 1\right\}. \tag{8.15}$$

By Lemma 8.8, this is the norm of f in the associate space $(L^{\Psi})'$ of L^{Ψ}, and hence is a rearrangement-invariant Banach function norm in its own right. We make the following definition.

Definition 8.13. Let Φ and Ψ be complementary Young's functions. Denote by L_{Φ} the space of measurable functions f on $[0, a]$ for which the functional in (8.15) is finite. The *Orlicz norm* on L_{Φ} is given by

$$\begin{aligned} \|f\|_{L_{\Phi}} &\equiv \sup\left\{\int_0^a |f(x)g(x)|\,dx : M^{\Psi}(g) \leq 1\right\} \\ &= \sup\left\{\int_0^a |f(x)g(x)|\,dx : \|g\|_{L^{\Psi}} \leq 1\right\} \\ &= \|f\|_{(L^{\Psi})'}. \end{aligned} \tag{8.16}$$

Theorem 8.14. *If Φ is a Young's function, then the spaces L^Φ and L_Φ coincide, and the Luxemburg and Orlicz norms are equivalent:*

$$\|f\|_{L^\Phi} \le \|f\|_{L_\Phi} \le 2\|f\|_{L^\Phi}, \qquad (f \in L^\Phi). \tag{8.17}$$

Proof. Let Ψ be the complementary Young's function of Φ. Suppose first that f is a nonzero function in L^Φ and let k be the reciprocal of its Luxemburg norm. Then it follows from (8.6) and (8.10) that $M^\Phi(kf) \le 1$. Thus, if g is any measurable function for which $M^\Psi(g) \le 1$, Young's inequality (8.14) gives

$$\int_0^a |kf(x)g(x)|\,dx \le M^\Phi(kf) + M^\Psi(g) \le 2,$$

and so

$$\int_0^a |fg|\,dx \le \frac{2}{k} = 2\|f\|_{L^\Phi}.$$

Taking the supremum over all such g, and using (8.16), we therefore obtain the second of the inequalities in (8.17).

It will suffice to establish the first estimate in (8.17) for nonnegative simple functions f, since both L^Φ and L_Φ are Banach function spaces and hence satisfy the Fatou property. Suppose therefore that f is such a function and that its Orlicz norm is nonzero. In view of (8.5) and (8.10), the proof will be complete if we show that

$$M^\Phi(kf) \le 1, \qquad \text{where } k = \frac{1}{\|f\|_{L_\Phi}}. \tag{8.18}$$

Suppose first that Φ (and hence ϕ) is everywhere finite on $[0,\infty)$. Then the fact that f is simple implies that $M^\Phi(kf)$ is finite. It also implies that the function

$$g(x) = \phi(kf(x)), \tag{8.19}$$

is simple. Theorem 8.12 shows that

$$\Psi(g(x)) + \Phi(kf(x)) = kf(x)g(x),$$

for all x, and so

$$M^\Phi(kf) + M^\Psi(g) = \int_0^a kfg\,dx. \tag{8.20}$$

In particular, $M^\Psi(g)$ is finite since the other two terms are. Since L^Ψ is the associate space of L_Φ, we obtain from Hölder's inequality,

$$\int_0^a kfg\,dx \le \|kf\|_{L_\Phi}\|g\|_{L^\Psi} = \|g\|_{L^\Psi}. \tag{8.21}$$

However, Lemma 8.8 shows that the norm on the right of (8.21) is majorized by $\max\{1, M^\Psi(g)\}$. Hence, (8.20) and (8.21) give

$$M^\Phi(kf) + M^\Psi(g) \le 1 + M^\Psi(g). \tag{8.22}$$

Since $M^\Psi(g)$ is finite, we may subtract it from both sides of (8.22) and thereby obtain the desired estimate (8.18). This completes the proof in the case where Φ is finite-valued.

If Φ assumes infinite values, set $s_\infty = \inf\{s : \Phi(s) = \infty\}$. Since, by definition, a Young's function cannot be identically infinite on $(0,\infty)$, we have $0 < s_\infty < \infty$. The function ψ is then bounded by s_∞, and so Ψ satisfies

$$\Psi(t) \le ts_\infty, \tag{8.23}$$

for all $t > 0$. We claim that

$$kf(x) \le s_\infty \qquad \text{a.e. on } [0,a]. \tag{8.24}$$

Indeed, if this were not the case, then, for some $\varepsilon > 0$, we should have $kf > s_\infty + \varepsilon$ on a set E of positive measure b, say. Let $h = (bs_\infty)^{-1}\chi_E$. Then, by (8.23), $M^\Psi(h) = b\Psi(1/(bs_\infty)) \le 1$. However, for this function h,

$$\int_0^a fh\,dx \ge \frac{1}{bs_\infty}\int_E f\,dx \ge \frac{s_\infty + \varepsilon}{ks_\infty} > \frac{1}{k} = \|f\|_{L_\Phi},$$

which is impossible since the Orlicz norm on L_Φ is the supremum of all integrals $\int fh\,dx$ with $M^\Psi(h) \le 1$. Hence, (8.24) holds.

Now let $\gamma < 1$. Since $\phi(\gamma s_\infty) < \infty$, it follows from (8.24) that $M^\Phi(\gamma kf) < \infty$ and that $g(x) = \phi(\gamma kf(x))$ is simple. We therefore obtain (8.20), (8.21), and (8.22) exactly as before but with γf replacing f. In particular, we have $M^\Phi(\gamma kf) \le 1$ for all $\gamma < 1$. An appeal to the monotone convergence theorem therefore establishes the desired result (8.18). ∎

Corollary 8.15. *Let Φ and Ψ be complementary Young's functions. Then L^Φ, equipped with the Luxemburg norm, has for its associate space the space L^Ψ equipped with the Orlicz norm.*

The Boyd indices of an Orlicz space L^Φ, as with any rearrangement-invariant Banach function space, are computed in terms of the norms of the dilation operators acting on the space (Definition III.5.12). We turn now to the problem of describing the indices directly in terms of the Young's function Φ itself.

The description will be in terms of the *right-continuous inverse* Φ^{-1} of Φ, which is defined by

$$\Phi^{-1}(t) = \sup\{s : \Phi(s) \le t\}, \qquad (0 \le t < \infty). \tag{8.25}$$

If

$$s_0 = \sup\{s : \Phi(s) = 0\}, \qquad s_\infty = \inf\{s : \Phi(s) = \infty\}, \tag{8.26}$$

then Φ is continuous and strictly increasing on $[s_0, s_\infty)$. Hence,

$$s = \Phi^{-1}(t) \quad \Leftrightarrow \quad t = \Phi(s), \qquad (s_0 < s < s_\infty). \tag{8.27}$$

In general, one checks easily that

$$\Phi(\Phi^{-1}(t)) \le t \le \Phi^{-1}(\Phi(t)), \qquad (0 \le t < \infty). \tag{8.28}$$

We need two lemmas.

Lemma 8.16. *Let Φ and Ψ be complementary Young's functions. Then*

$$w \le \Phi^{-1}(w)\Psi^{-1}(w) \le 2w, \qquad (0 \le w < \infty). \tag{8.29}$$

Proof. Since both Φ^{-1} and Ψ^{-1} are continuous on $[0, \infty)$, hence in particular at the origin, it will suffice to establish (8.29) for $0 < w < \infty$. Set $s = \Phi^{-1}(w)$ and $t = \Psi^{-1}(w)$. Then, by (8.27), we have $\Phi(s) = w = \Psi(t)$, and the second estimate in (8.29) follows at once from Young's inequality (8.14).

The first estimate will follow from the inequality

$$\min\left\{\frac{\phi(u)}{v}, \frac{\psi(v)}{u}\right\} \le 1, \tag{8.30}$$

which is valid for all $u, v > 0$. To establish (8.30), simply note that if $\phi(u) \ge v$, then (8.11) shows that $\psi(v) \le u$. Hence, with s and t as above, we obtain from (8.30),

$$w = \min\{\Phi(s), \Psi(t)\} \le \min\{s\phi(s), t\psi(t)\} \le st,$$

as desired. ∎

Lemma 8.17. *Let Φ be a Young's function. Then the fundamental function of L^Φ, equipped with the Luxemburg norm, is given by*

$$\varphi(b) = \varphi_{L^\Phi}(b) = \frac{1}{\Phi^{-1}(1/b)}, \qquad 0 < b \le a. \tag{8.31}$$

In particular, if s_∞, defined by (8.26), and $\Phi(s_\infty)$ are finite, then $\varphi(b)$ assumes the constant value $1/s_\infty$ for $0 < b \le 1/\Phi(s_\infty)$.

Proof. If $0 < b \le a$, and $k > 0$, we have

$$M^\Phi(k\chi_{[0,b]}) = \int_0^b \Phi(k)\,dx = b\Phi(k).$$

Hence, by (8.6) and (8.25),

$$\varphi_{L^\Phi}(b) = \|\chi_{[0,b]}\|_{L^\Phi} = \inf\{k^{-1} : b\Phi(k) \le 1\}$$
$$= \left[\sup\left\{k : \Phi(k) \le \frac{1}{b}\right\}\right]^{-1} = \frac{1}{\Phi^{-1}(1/b)}. \qquad ∎$$

Theorem 8.18. *Let Φ be a Young's function and set*

$$g(t) = \limsup_{s\to\infty}\left\{\frac{\Phi^{-1}(s)}{\Phi^{-1}(s/t)}\right\}, \qquad (0 < t < \infty). \tag{8.32}$$

Then the Boyd indices $\underline{\alpha} = \underline{\alpha}_\Phi$ and $\bar{\alpha} = \bar{\alpha}_\Phi$ of L^Φ are given by

$$\underline{\alpha} = \lim_{t\to 0+}\frac{\log g(t)}{\log t}, \qquad \bar{\alpha} = \lim_{t\to\infty}\frac{\log g(t)}{\log t}. \tag{8.33}$$

Proof. We can immediately dispense with the case where the value s_∞ defined by (8.26) is finite. For then $L^\Phi = L^\infty$, whose indices we know already to be equal to 0. However, Φ^{-1} increases monotonically to the value s_∞ so we see from (8.32) that $g(t) = 1$ for all $t > 0$. Hence, $\underline{\alpha} = \bar{\alpha} = 0$ is also the value for the indices predicted by (8.33). In what follows, we may therefore assume that Φ is everywhere finite on $(0, \infty)$.

Let $\|\cdot\|$ denote the Luxemburg norm on L^Φ. First, let us establish the identity for $\underline{\alpha}$ in (8.33). Suppose $0 < t < 1$. Recall from Definition III.5.10 that in computing the Boyd indices we use the operator norm $h(t)$ of the dilation operator $E_{1/t}$ acting on rearrangements of functions in L^Φ. Thus, on the interval $[0, a]$, we have

$$(E_{1/t}f^*)(s) = \begin{cases} f^*\left(\dfrac{s}{t}\right), & 0 \le s \le ta, \\ 0, & ta < s \le a, \end{cases} \tag{8.34}$$

and

$$h(t) = \sup\left\{\frac{\|E_{1/t}f^*\|}{\|f^*\|} : f \ne 0,\ f \in L^\Phi\right\}. \tag{8.35}$$

Let

$$k(t) = \sup\left\{\frac{\Phi^{-1}(s)}{\Phi^{-1}(s/t)} : \frac{1}{a} \le s < \infty\right\}. \tag{8.36}$$

We shall show that

$$k(t) \le h(t) \le k(2t), \qquad \left(0 < t < \frac{1}{2}\right). \tag{8.37}$$

Applying (8.34) and (8.35) with $f^* = \chi_{[0,b]}$, $(0 < b \le a)$, we obtain from (8.31) and (8.36), for $0 < t < 1$,

$$h(t) \ge \sup\left\{\frac{\|\chi_{[0,tb]}\|}{\|\chi_{[0,b]}\|} : 0 < b \le a\right\}$$
$$= \sup\left\{\frac{\Phi^{-1}(1/b)}{\Phi^{-1}(1/tb)} : 0 < b \le a\right\} = k(t),$$

which establishes the first of the estimates in (8.37).

For the other, note from (8.36) that

$$\frac{\Phi^{-1}(s)}{k(t)} \le \Phi^{-1}\left(\frac{s}{t}\right), \qquad \left(\frac{1}{a} \le s < \infty, 0 < t < 1\right).$$

Since Φ is increasing, an application of (8.28) gives

$$\Phi\left(\frac{\Phi^{-1}(s)}{k(t)}\right) \le \frac{s}{t}, \qquad \left(\frac{1}{a} \le s < \infty, 0 < t < 1\right).$$

Hence, putting $s = \Phi(u)$ (so $u \le \Phi^{-1}(s)$ by (8.25)), we obtain

$$\Phi\left(\frac{u}{k(t)}\right) \le \frac{\Phi(u)}{t}, \qquad \left(\Phi^{-1}\left(\frac{1}{a}\right) < u < \infty, 0 < t < 1\right).$$

Thus, if $c(f) = \sup\{x : 0 \le x < a, f^*(x) > \Phi^{-1}(1/a)\}$, then for $0 < t < 1$,

$$\Phi\left(\frac{f^*(v)}{k(t)}\right) \le \begin{cases} \dfrac{1}{t}\Phi(f^*(v)), & 0 \le v < c(f), \\[2ex] \Phi\left(\dfrac{\Phi^{-1}(1/a)}{k(t)}\right) \le \dfrac{1}{ta}, & c(f) \le v \le a. \end{cases}$$

Hence, if $0 < t < 1/2$, we have from (8.34),

$$M^\Phi\left(\frac{E_{1/t}f^*}{k(2t)}\right) \le \int_0^{ta} \Phi\left(\frac{f^*(s/t)}{k(2t)}\right) ds = \int_0^a \Phi\left(\frac{f^*(v)}{k(2t)}\right) t\, dv$$
$$\le \int_0^{c(f)} \frac{1}{2}\Phi(f^*(v))\, dv + \int_{c(f)}^a \frac{1}{2a} dv$$
$$\le \frac{1}{2}M^\Phi(f^*) + \frac{1}{2}.$$

By Lemma 8.8, this gives $\|E_{1/t}f^*/k(2t)\| \le 1$ if $\|f\| \le 1$, and so we see from (8.35) that the second of the inequalities in (8.37) also holds. Thus, (8.37) is established.

Now using (8.37) and Proposition III.5.13, and noting the obvious fact that $k(t) \ge g(t)$, we see that the lower Boyd index satisfies

$$\underline{\alpha} = \lim_{t \to 0+} \frac{\log h(t)}{\log t} = \lim_{t \to 0+} \frac{\log k(t)}{\log t} \le \lim_{t \to 0+} \frac{\log g(t)}{\log t}. \tag{8.38}$$

The estimate in (8.33) for $\underline{\alpha}$ will therefore be established if we show that

$$g(t) \ge t^{\underline{\alpha}}, \qquad (0 < t \le 1). \tag{8.39}$$

By III.(5.30), we have $t^{\underline{\alpha}} \le h(t) \le t^{\underline{\alpha} - \varepsilon(t)}$ for all $t \le 1$, where $\varepsilon(t) \to 0+$ as $t \to 0+$. Hence, by (8.37),

$$\left(\frac{t}{2}\right)^{\underline{\alpha}} \le k(t) \le t^{\underline{\alpha} - \varepsilon(t)}, \qquad (0 < t \le 1). \tag{8.40}$$

For each $s, t > 0$, let $\Delta(s,t) = \Phi^{-1}(s)/\Phi^{-1}(s/t)$.

Fix t with $0 < t \le 1$. For each positive integer m, it follows from (8.36) that there exists $s_m \ge 1/a$ such that $tk(t^m) \le \Delta(s_m, t^m) \le k(t^m)$, and hence by (8.40) that

$$2^{-\alpha}t^{m\underline{\alpha} - 1} \le \Delta(s_m, t^m), \qquad (m = 1, 2, \ldots).$$

Using these estimates, we obtain

$$\prod_{n=m}^{2m-1} \Delta\left(\frac{s_{2m}}{t^n}, t\right) = \prod_{n=m}^{2m-1} \frac{\Phi^{-1}(s_{2m}/t^n)}{\Phi^{-1}(s_{2m}/t^{n+1})} = \frac{\Phi^{-1}(s_{2m}/t^m)}{\Phi^{-1}(s_{2m}/t^{2m})}$$
$$= \frac{\Delta(s_{2m}, t^{2m})}{\Delta(s_{2m}, t^m)} \ge \frac{2^{-\underline{\alpha}}t^{2m\underline{\alpha} - 1}}{t^{m(\underline{\alpha} - \varepsilon(t^m))}}$$
$$= 2^{-\underline{\alpha}}t^{m\underline{\alpha} - 1 + m\varepsilon(t^m)}.$$

Thus, if n_m is chosen so that

$$\Delta(s_{2m}t^{-n_m}, t) = \max\{\Delta(s_{2m}t^{-n}, t) : m \le n < 2m\},$$

then this term dominates each of those in the product above and so we see from the preceding estimate that

$$\Delta(s_{2m}t^{-n_m}, t) \ge 2^{-\underline{\alpha}/m}t^{\underline{\alpha} - \frac{1}{m} + \varepsilon(t^m)}.$$

Now $s_{2m} \ge 1/a$ and $n_m \ge m$ so $s_{2m}t^{-n_m} \ge 1/(at^m) \to \infty$ as $m \to \infty$, since $0 < t < 1$. Hence, from (8.32) and the preceding estimate,

$$g(t) = \limsup_{s \to \infty} \Delta(s,t) \ge \limsup_{m \to \infty} \Delta(s_{2m}t^{-n_m}, t) = t^{\underline{\alpha}}.$$

This establishes (8.39) and hence also the first identity in (8.33).

We may therefore use the counterpart of the first identity in (8.33) for the complementary Young's function Ψ to determine the lower index $\underline{\alpha}_\Psi$ of L^Ψ. With the aid of Lemma 8.16, it is then easy to show that $1 - \underline{\alpha}_\Psi$ is equal to the expression on the right of the second identity in (8.33). But III.(5.33) reveals that $1 - \underline{\alpha}_\Psi = \bar{\alpha}_\Phi$, because L^Ψ is the associate space of L^Φ. This establishes the second identity in (8.33) and hence completes the proof. ■

EXERCISES AND FURTHER RESULTS FOR CHAPTER 4

1. (M. Riesz [1]) Let A be the bilinear form corresponding to the matrix

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Consider the line segment with endpoints $(0, 1)$ and $(1/2, 1/2)$, a generic point of which is $(\theta/2, 1 - \theta/2)$ with $0 \le 0 \le 1$. The maximum M_θ (computed with respect to *real* vectors) of A at such a point is the maximum of value of the expression

$$\frac{\{|x_1 + x_2|^{2/(2-\theta)} + |x_1 - x_2|^{2/(2-\theta)}\}^{(2-\theta)/2}}{\{|x_1|^{2/\theta} + |x_2|^{2/\theta}\}^{\theta/2}}, \tag{1}$$

as x_1, x_2 vary over all real numbers. Let m_θ denote the *minimum* of this expression as x_1, x_2 vary. Then $m_i = M_i, (i = 0, 1)$, and $\log m_\theta$ is *concave* for $0 \le \theta \le 1$. Hence, $\log M_\theta$ cannot be convex on the interval $0 \le \theta \le 1$ (this was established by direct computation in Example 1.3) (HINT: Apply the Riesz convexity theorem to the adjoint transformation on the complementary line segment, whose endpoints are $(1, 0)$ and $(1/2, 1/2)$).

2. (M. Riesz [1]) Let M_θ be the maximum of the bilinear form A of the preceding exercise, computed this time with respect to *complex* vectors. Then $M_\theta = 2^{3/2-\theta}$ for $0 \le \theta \le 1$. In particular, $\log M_\theta$ is linear, hence certainly convex for $0 \le \theta \le 1$ (HINT: The maximum of the expression (1) occurs when $x_1 : x_2 = 1 : \pm i$).

3. (M. Riesz [1]) Let A be an $m \times n$ matrix with real coefficients a_{ij}. Let (α, β) be a point in the unit square and consider complex vectors $\mathbf{x}$ and $\mathbf{y}$ which furnish the maximum $M(\alpha, \beta)$ of the bilinear form A:

$$|A(\mathbf{x}, \mathbf{y})| = M(\alpha, \beta)\|\mathbf{x}\|_{1/\alpha}\|\mathbf{y}\|_{1/(1-\beta)}.$$

(a) If (α, β) lies in the lower triangle $(\alpha > \beta)$, then (up to an arbitrary constant factor of modulus one) any maximizing vectors are real.
(b) If (α, β) lies on the diagonal $(\alpha = \beta)$, then real maximizing vectors $\mathbf{x}$ and $\mathbf{y}$ exist but non-real ones may also exist.
(c) If (α, β) lies in the upper triangle $(\alpha < \beta)$, then it may be the case that no real vectors $\mathbf{x}$ and $\mathbf{y}$ furnish the maximum.

4. (a) Let f be a continuous function on a closed interval I. Suppose that to each subinterval $[a, b]$ of I there corresponds a value θ, $0 < \theta < 1$, such that

$$f((1-\theta)a + \theta b) \le (1-\theta)f(a) + \theta f(b).$$

Then f is convex on I.

(b) A positive function f is logarithmically convex on an interval I if $\log f$ is convex on I. Thus, f is logarithmically convex on I if and only if $e^{\alpha x}f(x)$ is convex on I for all real α.

5. For positive operators ($f \ge 0 \Rightarrow Tf \ge 0$), the Riesz convexity theorem is valid in the entire unit square (that is, the hypotheses of Theorem 1.7 may be relaxed to $1 \le p_k, q_k \le \infty$, $(k = 0,1)$) (HINT: Extend Theorem 1.5 to the unit square (cf. Theorem 1.2), then extend to general measure spaces as in the proof of Theorem 1.7).

6. (A. P. Calderón & A. Zygmund [2]). The Riesz-Thorin theorem (Theorem 2.2) remains valid in the entire first quadrant, that is, for $0 < p_k, q_k \le \infty$, $(k = 0, 1)$. Indeed, with the notation of Theorem 2.2, choose c with $0 < c < q_k$, $(k = 0, 1)$, so that $q_k/c > 1$ and $q/c > 1$.
(a) To establish the result, it suffices to show that

$$I \equiv \int_S |Tf|^c g \, dv \le M_0^{c(1-\theta)} M_1^{c\theta}, \tag{2}$$

for all simple f and all nonnegative simple g satisfying

$$\|f\|_p = 1 = \|g\|_{q/(q-c)}.$$

(b) Modify (2.13) and (2.14) by taking f_z as before but setting

$$g_z = g^{(1-c\beta(z))/(1-c\beta(\theta))}, \qquad F(z) = \int_S |Tf_z|^c |g_z| \, dv.$$

Then $F(\theta) = I$ and, as in the proof of Theorem 2.2,

$$F(iy) \le M_0^c, \qquad F(1 + iy) \le M_1^c, \qquad (-\infty < y < \infty).$$

(c) At each point s of the measure space (S, v), the function $G_s(z) = (Tf_z)(g_z)^{1/c}$ is a finite sum of exponentials, hence is analytic. Consequently, by the Poisson-Jensen formula,

$$F(\theta) = \int_S \exp\{c\log|G_s(\theta)|\} \, dv(s)$$

$$\le \int_S \left\{\exp \int_{-\infty}^{\infty} c\log|G_s(iy)| \frac{P_\theta^0(y)}{1-\theta} dy\right\}^{1-\theta}$$

$$\times \left\{\exp \int_{-\infty}^{\infty} c\log|G_s(1+iy)| \frac{P_\theta^1(y)}{\theta} dy\right\}^{\theta} dv(s).$$

Using Jensen's inequality, deduce that

$$F(\theta) \le \left\{\int_{-\infty}^{\infty} F(iy) \frac{P_\theta^0(y)}{1-\theta} dy\right\}^{1-\theta} \left\{\int_{-\infty}^{\infty} F(1+iy) \frac{P_\theta^1(y)}{\theta} dy\right\}^{\theta},$$

and derive the desired result (2).

7. Establish Young's convolution inequality (2.24)

(i) using multilinear interpolation (Theorem 2.7);

(ii) directly, without using interpolation, by writing

$$f(e^{is})g(e^{i(t-s)}) = (f^{p/r}g^{q/r})(f^{p(1/p-1/r)})(g^{q(1/q-1/r)})$$

and applying a three-term Hölder inequality.

8. Show that the Hausdorff-Young inequality (2.29), for p' an even integer, follows directly from Young's convolution inequality (2.24). Conversely, Young's convolution inequality (2.24), in the special case $p, q, r' \le 2$, follows directly from the Hausdorff-Young theorem.

9. An orthonormal system (ϕ_n) in $L^2(a, b)$ is *uniformly bounded* if there is a constant M such that $|\phi_n(x)| \le M$ for all n and for all x in (a, b). Use the proof of Theorem 2.6 to establish the following Hausdorff-Young inequalities (valid for $1 < p \le 2$, with $p' = p/(p-1)$):

$$\left(\sum_n |\langle f, \phi_n \rangle|^{p'}\right)^{1/p'} \le M^{(2-p)/p}\left(\int_a^b |f(x)|^p\,dx\right)^{1/p}, \qquad (f \in L^p) \tag{3}$$

and

$$\left(\int_a^b |f(x)|^{p'}\,dx\right)^{1/p'} \le M^{(2-p)/p}\left(\sum_{n=0}^{\infty} |c_n|^p\right)^{1/p}, \qquad (c \in \ell^p) \tag{4}$$

where, in the latter inequality, $\langle f, \phi_n \rangle = c_n$ for all n. These results were established prior to the M. Riesz convexity theorem by F. Riesz [1].

10. (S. Verblunsky [1]) Let (ϕ_n) be a complete orthonormal system in $L^2(a, b)$, uniformly bounded by a constant M.

(a) For equality to occur in (3), it is necessary that

$$f(x) = \sum_{k=1}^{K} c_k \phi_{n_k}, \qquad (n_1 < n_2 < \cdots < n_K). \tag{5}$$

For such functions, equality occurs in (3) if and only if

(i) $|c_{n_1}| = |c_{n_2}| = \cdots = |c_{n_K}|$;
(ii) $|f(x)|$ is constant in a set E of measure $1/(KM^2)$ and $f = 0$ outside E.

(b) For equality to occur in (4), it is necessary that only finitely many c_n's, say $c_{n_1}, c_{n_2}, \ldots, c_{n_K}$, are distinct from 0 and satisfy (i). The function f is then of the form (5), and a necessary and sufficient condition that equality hold in (4) is that f satisfy (ii) (cf. also A. P. Calderón & A. Zygmund [1]).

11. In the case of the trigonometric system, the preceding result reduces to the theorem of G. H. Hardy & J. E. Littlewood [1] on the cases of equality in the classical Hausdorff-Young inequalities:

(a) Equality holds in (2.29) if and only if f is a constant multiple of e^{int} for some integer n.
(b) Equality holds in (2.30) if and only if the sequence **c** has at most one coefficient different from zero.

12. The Fourier transform $\hat{f}$ of a function f in $L^1(\mathbf{R})$ is defined by

$$\hat{f}(y) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x y}\,dx, \qquad (y \in \mathbf{R}). \tag{6}$$

The Fourier transform $f \to \hat{f}$ extends to an admissible operator of norm at most one for the couples (L^1, L^2) and (L^∞, L^2) and

$$\|\hat{f}\|_{p'} \le \|f\|_p, \qquad (f \in L^p(\mathbf{R})). \tag{7}$$

This version of the Hausdorff-Young theorem, for the real line $\mathbf{R}$, is due to E. C. Titchmarsh [1].

13. By contrast with the situation for the unit circle $\mathbf{T}$ or the integers $\mathbf{Z}$ (cf. Exercise 11), equality is never attained in (7) except when $f \equiv 0$ (E. Hewitt & I. I. Hirschmann [1]). In fact, the operator norm itself is strictly less than 1:

$$\|\hat{f}\|_{p'} \le A_p \|f\|_p, \qquad (f \in L^p(\mathbf{R})), \tag{8}$$

where

$$A_p = \left(\frac{p^{1/p}}{p'^{1/p'}}\right)^{1/2}, \qquad \left(1 < p < 2, \frac{1}{p} + \frac{1}{p'} = 1\right). \tag{9}$$

Equality occurs in (8) for the Gaussian functions $f(x) = e^{-\alpha x^2}$, $(\alpha > 0)$. This sharp form of the Hausdorff-Young theorem is due to K. I. Babenko [1], who proved the result for p' even, and to W. Beckner [1] in the general case. In $\mathbf{R}^n$, the best constant in (8) is $(A_p)^n$ instead of A_p.

14. (W. Beckner [1]) The sharp form of Young's convolution inequality for $\mathbf{R}^n$ is

$$\|f * g\|_r \le (A_p A_q A_{r'})^n \|f\|_p \|g\|_q, \qquad (f \in L^p, g \in L^q), \tag{10}$$

where $1 \le p, q \le \infty$, $1/r = 1/p + 1/q - 1 \ge 0$, and the constants A_p, A_q, and $A_{r'}$ are defined as in (9). Equality is attained for Gaussian functions. Show that (10) follows directly from (8) in the special case $p, q, r' \le 2$ (HINT: cf. Exercise 8).

15. (S. K. Pichorides [1]) The best constant A_p in the M. Riesz inequality

$$\|\tilde{f}\|_p \le A_p \|f\|_p, \qquad (1 < p < \infty)$$

for the conjugate-function operator $f \to \tilde{f}$ is given by

$$A_p = \begin{cases} \tan\left(\dfrac{\pi}{2p}\right), & \text{if } 1 < p \le 2, \\[2ex] \cot\left(\dfrac{\pi}{2p}\right), & \text{if } 2 < p \le \infty. \end{cases}$$

T. Gokhberg & N. Krupnik [1] had established this result previously for $p = 2^n$, $(n = 1, 2, \ldots)$. The same constants provide sharp Hilbert transform inequalities on the real line.

16. (B. Davis [1]) The best constant Θ in the Kolmogorov weak type (1, 1) inequality

$$\|\tilde{f}\|_{L^{1,\infty}} \equiv \sup_{0 < t < 1} t\tilde{f}^*(t) \le \Theta \|f\|_{L^1(\mathbf{T})}$$

for the conjugate-function operator $f \to \tilde{f}$ is given by

$$\Theta^{-1} = \frac{2}{\pi^2} \int_0^\pi \left|\log\left|\cot\frac{\theta}{2}\right|\right| d\theta.$$

Show that the function $f(e^{i\theta})$ which is the boundary value of the real part of $F(z) = (2/\pi)\log((1+z)/(1-z))$ is extremal. Show also that Θ is given by the quotient

$$\Theta = \frac{1^{-2} + 3^{-2} + 5^{-2} + 7^{-2} + \cdots}{1^{-2} - 3^{-2} + 5^{-2} - 7^{-2} + \cdots}.$$

Davis [2], using Brownian motion techniques as in [1], obtains the best constants $\Theta_p = \|(\sin\theta)^{-1}\|_p$ for the Kolmogorov inequalities $\|\tilde{f}\|_p \le \Theta_p\|f\|_1$, $(0 < p < 1)$. A. Baernstein [2] has supplied non-probabilistic proofs of these results.

17. (S. K. Pichorides [1]) In the Zygmund conjugate function inequality

$$\|\tilde{f}\|_1 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |\tilde{f}(e^{i\theta})|\,d\theta \le \frac{A}{2\pi}\int_{-\pi}^{\pi} |f|\log^+|f| + B,$$

the only admissible values of A are $A > 2/\pi$ (and $B = B(A)$).

18. (C. Bennett [6]) The best constant A in the Zygmund conjugate function inequality

$$\|\tilde{f}\|_1 \le A\|f\|_{L\log L} \equiv A\int_0^1 f^{**}(t)\,dt$$

is given by $A = \Theta^{-1}$, where Θ is defined as in Exercise 16. Find an extremal function (HINT: cf. Exercise 16).

19. (a) The Poisson kernel for the strip

$$\bar{\Omega} = \{z = x + iy : 0 \le x \le 1, -\infty < y < \infty\},$$

corresponding to a point $z = x + iy$ of Ω, is given by the pair of functions

$$P_z^0 = \frac{(\sin\pi x)/2}{\cosh\pi(y-t) - \cos\pi x},$$
$$P_z^1(t) = \frac{(\sin\pi x)/2}{\cosh\pi(y-t) + \cos\pi x}, \tag{11}$$

corresponding to the left- and right-hand boundaries $\{it : -\infty < t < \infty\}$ and $\{1 + it : -\infty < t < \infty\}$, respectively, of Ω (HINT: Transform the Poisson kernel for the unit circle (cf. Exercise III.12) under the conformal mapping in (3.1)).

(b) With the identities

$$\int_{-\infty}^{\infty} P_z^0(t)\,dt = 1 - x, \qquad \int_{-\infty}^{\infty} P_z^1(t)\,dt = x, \qquad (z = x + iy), \tag{12}$$

the three-lines theorem (Lemma 2.1) follows from Lemma 3.1.

20. (C. Bennett & K. Rudnick [1]) The Lorentz-Zygmund space $\ell^{p,q}(\log\ell)^\alpha$ con-

sists of all sequences $(f_n)_{n=-\infty}^{\infty}$ for which

$$\|f\| = \left\{\sum_{n=1}^{\infty} [n^{1/p}(1 + \log n)^\alpha f_n^*]^q \frac{1}{n}\right\}^{1/q}$$

is finite (with the obvious modification if $q = \infty$). Theorem 6.14 holds also whenever (R, μ) or (S, ν) is replaced by the discrete measure space $\mathbf{Z}$; the only change necessary is that $\alpha + 1/a = \beta + 1/b$ in (6.36) must be negative instead of positive (and vice versa in (6.37)) when (R, μ) is replaced by $\mathbf{Z}$. The following classical results (G. H. Hardy & J. E. Littlewood [3, 5]; A. Zygmund [1, 2]) for the Fourier transform $\mathscr{F}$ are immediate corollaries.

(a) Suppose $f \in L(\log L)^\alpha(\mathbf{T})$, $(\alpha > 0)$, and let (c_n) denote its sequence of Fourier coefficients. Then

(i) $\sum_{n=-\infty}^{\infty} \exp\{-\lambda/|c_n|^{1/\alpha}\} < \infty$ for all $\lambda > 0$;
(ii) there are constants A_α and B_α independent of f such that

$$\sum_{n=1}^{\infty} \frac{1}{n}(\log n)^{\alpha-1} c_n^* \le A_\alpha \int_0^{2\pi} |f|(\log^+|f|)^\alpha + B_\alpha;$$

(iii) if, in addition, $\alpha \le 1$, then $\sum_{n=1}^{\infty} \frac{1}{n}(c_n^*)^{1/\alpha} < \infty$

(HINT: Reformulate each result as $\mathscr{F}: L^{1,1}(\log L)^\alpha \to X$, where (i) $X = \ell_0^{\infty,\infty}(\log\ell)^\alpha$, (ii) $X = \ell^{\infty,1}(\log\ell)^{\alpha-1}$, (iii) $X = \ell^{\infty,1/\alpha}$, the subscript "0" denoting the closure of the simple functions in the space involved).

(b) Suppose $\alpha > 0$.

(i) If $\mathbf{c} = (c_n) \in \ell(\log\ell)^{-\alpha}$, then $\mathbf{c}$ is the sequence of Fourier coefficients of a function f for which

$$\int_0^{2\pi} \exp\{\lambda|f|^{1/\alpha}\} < \infty \tag{*}$$

for all $\lambda > 0$;

(ii) If $c_n^* = O(\frac{1}{n}\log^{\alpha-1} n)$ as $n \to \infty$, then $\mathbf{c}$ is the sequence of Fourier coefficients of a function f for which (*) holds for some $\lambda > 0$;

(iii) Suppose $1 < p < \infty$ and $1/p + 1/p' = 1$. If $\sum_{n=1}^{\infty} n^{p-1} c_n^{*p} < \infty$, then $\mathbf{c}$ is the sequence of Fourier coefficients of a function f for which $\int_0^{2\pi} \exp\{\lambda|f|^{p'}\} < \infty$ for all $\lambda > 0$

(HINT: The results reformulate as:

(i) $\mathscr{F}': \ell^{1,1}(\log\ell)^{-\alpha} \to L_0^{\infty,\infty}(\log L)^{-\alpha}$;
(ii) $\mathscr{F}': \ell^{1,\infty}(\log\ell)^{1-\alpha} \to L^{\infty,\infty}(\log L)^{-\alpha}$;
(iii) $\mathscr{F}': \ell^{1,1/(1-\alpha)} \to L^{\infty,\infty}(\log L)^{-\alpha}$),

where $\mathscr{F}': \mathbf{c} \to f$ is the Fourier transform from $\mathbf{Z}$ to $\mathbf{T}$).

21. (R. Sharpley [1]) Let X be a rearrangement-invariant Banach function space, and let $\Lambda_a(X)$ consist of those f for which

$$|f|_{\Lambda_a(X)} = \begin{cases} \left\{\int [f^*(t)\varphi_X(t)]^a \dfrac{dt}{t}\right\}^{1/a}, & (a < \infty), \\ \sup_t [f^*(t)\varphi_X(t)], & (a = \infty), \end{cases}$$

is finite. Suppose that the fundamental indices (cf. Exercise III.14) of X lie strictly between 0 and 1.

(a) $\Lambda_a(X)$ is a rearrangement-invariant Banach function space when equipped with the norm $f \to |f^{**}|_{\Lambda_a(X)}$. The space $Y = \Lambda_a(X)$ has fundamental function equal to φ_X and its indices satisfy

$$\underline{\alpha}_Y = \underline{\beta}_Y = \underline{\beta}_X, \qquad \overline{\alpha}_Y = \overline{\beta}_Y = \overline{\beta}_X.$$

(b) $\Lambda_1(X) = \Lambda(X)$ and $\Lambda_\infty(X) = M(X)$ with equivalent norms. The Lorentz-Zygmund spaces of the preceding exercise may also be realized as $\Lambda_a(X)$ spaces.
(c) If $1 \le a \le b \le \infty$, then $\Lambda_a(X) \hookrightarrow \Lambda_b(X)$.
(d) The associate space of $\Lambda_a(X)$ (and the dual space if $a < \infty$) is equivalent to $\Lambda_{a'}(X')$, where $1/a + 1/a' = 1$.
(e) An operator T is of *weak type* (X, Y) if it is bounded from $\Lambda_1(X)$ into $\Lambda_\infty(Y)$. It is of *restricted weak type* (X, Y) if

$$\varphi_Y(t)(T\chi_E)^*(t) \le c\varphi_X(\mu(E)), \qquad (t > 0)$$

for each set E of finite measure. These conditions are equivalent if $\underline{\beta}_X > 0$ and $\overline{\beta}_Y < 1$.

22. (R. Sharpley [1]) Suppose X_0, X_1, X, Y_0, Y_1, Y are rearrangement-invariant Banach function spaces with fundamental indices strictly between 0 and 1. Let

$$F(s,t) = \frac{\varphi_Y(t)}{\varphi_X(s)} \min\left\{\frac{\varphi_{X_0}(s)}{\varphi_{Y_0}(t)}, \frac{\varphi_{X_1}(s)}{\varphi_{Y_1}(t)}\right\}.$$

Let W_σ denote the collection of pairs (X, Y) that have the weak-interpolation property with respect to $\sigma = [(X_0, Y_0), (X_1, Y_1)]$.
(a) $(\Lambda(X), M(Y)) \in W_\sigma$ if and only if $\sup_s \sup_t F(s,t) < \infty$, and this condition is necessary in order that (X, Y) belong to W_σ.
(b) $(M(X), \Lambda(Y)) \in W_\sigma$ if and only if $\iint F(s,t)\, ds/s\, dt/t < \infty$. This condition is sufficient for (X, Y) to belong to W_σ, and in particular for (X, Y) to have the strong interpolation property with respect to σ.
(c) $(\Lambda(X), \Lambda(Y)) \in W_\sigma$ if and only if $\sup_s \int F(s,t)\, dt/t < \infty$.
(d) $(M(X), M(Y)) \in W_\sigma$ if and only if $\sup_t \int F(s,t)\, ds/s < \infty$.
(e) $(\Lambda_a(X), \Lambda_a(Y)) \in W_\sigma$, $(1 \le a \le \infty)$, if and only if both conditions in (c) and (d) hold.

NOTES FOR CHAPTER 4

The results of §1 and their proofs are taken virtually unchanged from the seminal paper of M. Riesz [1]. R. E. A. C. Paley [1] gave an alternative proof of the convexity theorem (Theorem 1.5) in which the extrema of the forms are

computed directly by differentiation instead of relying on the cases of equality in Hölder's inequality. Another proof, given by L. C. Young [1], invokes Minkowski's theory of convex bodies and in particular the notion of "reciprocal polars". More recently, J. Peetre [6] has based an interpolation method on Riesz' proof.

Thorin's convexity theorem appears in O. V. Thorin [1], [2]; cf. also J. D. Tamarkin & A. Zygmund [1], R. Salem & A. Zygmund [1] and R. Salem [1]. Thorin's use of complex methods inspired several generalizations of the interpolation theorem based on geometric properties of analytic and subharmonic functions; cf. A. P. Calderón & A. Zygmund [1] (multilinear interpolation and interpolation of H^p-spaces, based on Phragmen-Lindelöf principles), [2] (extension from the unit square to the first quadrant), [3] (extension to sublinear operators, using subharmonic function theory), and G. Weiss [1] (extension to sublinear operators on H^p-spaces). G. Bennett [1] extends the result to the third quadrant! In this interpretation, T is of strong type (p, q) for negative p and q when its adjoint is of strong type (p', q') for the conjugate exponents $p' = p/(p-1)$ and $q' = q/(q-1)$. The characteristic sets, that is, the subsets of the unit square (or first quadrant) that can occur as the sets $\{(1/p, 1/q) : T \text{ is of strong type } (p,q)\}$ for linear operators T have been characterized by S. D. Riemenschneider [1], [2] (cf. also P. P. Zabrieko & M. A. Krasnosel'skii [1] and G. Bennett [1, pp. 10–11]).

Theorem 2.9 on the interpolation of compact operators is due to M. A. Krasnosel'skii [1]. Generalizations for rearrangement-invariant spaces may be found in G. G. Lorentz & T. Shimogaki [3] and R. Sharpley [2]. Interpolation of analytic families (Theorem 3.3) and interpolation with change of measures (Theorem 3.6) are due to E. M. Stein [1], although special cases of these results were also obtained independently by I. I. Hirschman [1]. Interpolation with change of measures for sublinear operators was investigated by E. M. Stein & G. Weiss [2]. The results of A. P. Calderón & A. Zygmund [3] on interpolation on H^p-spaces were extended to analytic families by E. M. Stein & G. Weiss [1].

Thorin's method of proof is the basis for the *complex method of interpolation*, introduced independently by A. P. Calderón [1], [2], S. G. Krein [1], [2], and J. L. Lions [1], [2]. For a given couple (X_0, X_1), consider the space of $(X_0 + X_1)$-valued functions $F(z)$ that are analytic in the strip $0 < \mathrm{Re}(z) < 1$, bounded on its closure, and assume values in X_j on $\mathrm{Re}(z) = j$, $(j = 0, 1)$. For $0 < \theta < 1$, the interpolation space $[X_0, X_1]_\theta$ consists of the values $F(\theta)$ as F varies over all such functions described above. The complex method has been extended by R. R. Coifman, M. Cwikel, R. Rochberg, Y. Sagher & G. Weiss [1], [2] and independently by L. J. Nikolova [1] (and her joint work with S. G. Krein [1]) to treat the case of interpolation of infinitely many Banach spaces. R. Rochberg & G. Weiss [1] have shown that the corresponding methods which utilize the derivatives $F'(\theta)$ have fundamental applications in the theory of commutator integrals. More generally, N. J. Kalton [1] has established a commutator theorem couched in terms of the Boyd indices to the effect that weak-type operators "almost commute" with all symmetric functionals which "almost commute" with multiplication operators.

J. Marcinkiewicz [1] announced his now celebrated interpolation theorem in 1939, shortly before his untimely death. The theorem in complete (off-diagonal) form with proof and numerous applications was supplied by A. Zygmund [3] in 1956 (cf. also the note in W. A. J. Luxemburg [2]). Various formulations in terms of Lorentz spaces (which were introduced by G. G. Lorentz [1], [2] in 1950) were given by A. P. Calderón [3], R. A. Hunt [1], [2], R. A. Hunt & G. Weiss [1], R. O'Neil [1], R. O'Neil & G. Weiss [1], and E. T. Oklander [1]; the treatment adopted in the text, in terms of the S_σ-operator, which is the one with the capacity for greatest generalization, is that of A. P. Calderón [3] (see also R. A. Hunt & G. Weiss [1]).

Restricted weak-type operators were introduced by E. M. Stein & G. Weiss [3]. Theorem 5.6 is in K. H. Moon [1]. For Stein's theorem on limits of sequences of operators (Theorem 5.10), see E. M. Stein [2] and the monograph by A. Garsia [1]; the result has far reaching applications to factorization of operators in Banach spaces; see B. Maurey [1], the survey article by J. E. Gilbert [1], and the monograph by M. de Guzmán [1].

The spaces $L\log L$ and $L_{\exp}$ were introduced in 1928 by A. Zygmund [1] and independently by E. C. Titchmarsh [2], [3], in connection with the integrability of the conjugate function. G. H. Hardy & J. E. Littlewood [2] supplied the "Lorentz-type" norm (cf. Definition 6.3) for $L\log L$ (see the Notes for Chapter III), which is the basis for the subsequent development of the Lorentz-Zygmund spaces by C. Bennett & K. Rudnick [1] (cf. also C. Bennett [2]); necessity of the interpolation conditions for these spaces was demonstrated by R. Sharpley [7]. Theorem 6.7 is due to G. H. Hardy and J. E. Littlewood [2]; see E. M. Stein [4] for a local analogue in $\mathbf{R}^n$. For Corollaries 6.9 and 6.10, see A. Zygmund [1], [4, Chapter VII and p. 381]; see E. M. Stein [4] for analogous results on the line, and C. S. Stanton [1] for a martingale version.

The multilinear weak-type theory presented here, which generalizes the Calderón theory of preceding sections, was developed by R. Sharpley [5], [6]. The extension of the linear theory given in Theorem 7.10 is due to R. Sharpley [1].

The exposition of the theory of Orlicz spaces in this section follows closely the account in the thesis of W. A. J. Luxemburg [1]. Theorem 8.14 and Corollary 8.15 are established there. A further comprehensive account of the theory of Orlicz spaces is given by M. A. Krasnosel'skii & Ya. B. Rutickii [1]. J. Gustavsson & J. Peetre [1] have made a detailed analysis of interpolation of Orlicz spaces and Orlicz classes; cf. also I. B. Simonenko [1], [2]. The identification of the Boyd indices of an Orlicz space with the reciprocals of the "Matuzewska-Orlicz exponents" of the space (Theorem 8.18) is due to D. W. Boyd [6]. There is some intersection between the families of Orlicz spaces and Lorentz spaces (in the Zygmund spaces $L^p(\log L)^\alpha$ (cf. §6), for example). Conditions under which a Lorentz space may be regarded as an Orlicz space are given by G. G. Lorentz [7].

5 The K-Method

The K-*method*, or *real method*, of interpolation may be regarded as a lifting of the Marcinkiewicz interpolation theorem from its classical context in spaces of measurable functions to an abstract Banach space setting. Since the essence of the K-interpolation structure resides in the K-functional itself, the primary goal of the subject (and of this chapter) is to determine in concrete terms the K-functionals for a variety of specific pairs of Banach spaces.

We begin by deriving some elementary properties of the K-interpolation method, such as the basic interpolation theorems and the invariance of the K-functional under *Gagliardo completion.* The K-functional for the couple (L^1, L^∞) is identified in Theorem 1.6 as $tf^{**}(t)$. The k-method generates all interpolation spaces between L^1 and L^∞ (Theorem 1.17) and therefore serves as a prototype for interpolation in more general settings.

Section 2 contains the *reiteration theorem* for the K-method. One form of this result describes the K-functional (and hence the interpolation spaces) between a pair of spaces that are themselves K-interpolation spaces of a given couple. For example, it provides the K-interpolation spaces between L^p and L^q, $(1 < p < q < \infty)$, in terms of the K-functional $tf^{**}(t)$ for L^1 and L^∞. Section 2 also contains a brief treatment of the J-method and its equivalence with the K-method. The section concludes with a result of T. H. Wolff, which allows the "patching" together of two interpolation scales.

Section 3 considers the question of which interpolation spaces for a given (Gagliardo) couple are generated by the K-method (or the k-method). Elementary considerations show that a kind of *monotonicity* is necessary; Theorem 3.7 shows that it is also sufficient. The key property is that every Gagliardo couple is *divisible* (Theorem 3.6); the proof requires a basic decomposition due to M. Cwikel (Theorem 3.4) but otherwise proceeds by lifting a primitive form of divisibility for the couple (L^1, L^∞).

In section 4, we consider smoothness spaces, namely, the Besov spaces $B^p_{\alpha,q}$ and the Sobolev spaces W^p_r. A description of the K-functional for the couple (L^p, W^p_r) is obtained (Theorem 4.12) in terms of the r-th order L^p-modulus of smoothness. The K-interpolation spaces are thus identified as the intermediate Besov spaces for this couple (Corollary 4.13). The K-interpolation spaces between certain pairs of Besov spaces and/or Sobolev spaces are determined in Theorem 4.17. The section concludes with a derivation of various embedding theorems.

Section 5 deals with interpolation between the Sobolev spaces W^1_k and W^∞_k for a positive integer k. The Whitney covering lemma (Lemma 5.1) and maximal function arguments are used to establish the DeVore-Scherer theorem (Theorem 5.12), which describes the K-functional for the couples (W^1_k, W^∞_k).

The Hardy space H^1 is considered in Section 6. Several characterizations of $\operatorname{Re} H^1$ are obtained, in terms of the Hilbert transform, the nontangential maximal operator, and the atomic decomposition (Theorem 6.14). The K-functional for the pair $(\operatorname{Re} H^1, L^\infty)$ is described in terms of the nontangential maximal function (Theorem 6.15). Fefferman's duality theorem, which identifies the dual of $\operatorname{Re} H^1$ as the space BMO of functions of bounded mean oscillation, is established in Theorem 6.17.

The notion of oscillation is studied in greater detail in section 7. A fundamental inequality (Theorem 7.3) estimates the quantity $f^{**} - f^*$ in terms of the oscillation of f, and this provides a simple proof of the John-Nirenberg lemma (Corollary 7.7) for functions in BMO. It also leads to the construction of a (nonlinear) space W, which serves as the space weak-L^∞ in the Marcinkiewicz interpolation theory. In bounded euclidean space, the space W may also be characterized as the rearrangement-invariant hull of BMO (Theorem 7.10). The boundedness of the Hardy-Littlewood maximal operator on BMO is established in Theorem 7.18. A more precise formulation of this result (Theorem 7.20) characterizes BLO-functions as (modulo bounded functions) the Hardy-Littlewood maximal functions of BMO-functions.

The main result of Section 8 is the identification of the K-functional for the

pair (L^1, BMO) with the functional $tf^{\#*}(t)$, where $f^\#$ is the Fefferman-Stein sharp function of f. This identifies (Theorem 8.11) the (θ, q)-interpolation spaces between L^1 and BMO as the Lorentz $L^{p,q}$-spaces, $(\theta = 1 - 1/p)$.

Sections 9 and 10 are concerned with interpolation between Hardy spaces H^1 and H^∞. Section 9 is preparatory, and describes Jones' constructive solution of the equation $\bar{\partial} f = \mu$, where μ is a Carleson measure (Theorem 9.5). The K-functional for the pair (H^1, H^∞) is described in terms of the nontangential maximal operator N (Theorem 10.1 and Corollary 10.2). The remainder of the section is devoted to a description of the interpolation spaces between H^1 and H^∞. The main result here is Theorem 10.6, which, under the hypothesis $NG \prec NF$, constructs an admissible operator that maps F to G. It is thus the analogue for (H^1, H^∞) of the Calderón-Ryff result (Theorem III.2.10) for (L^1, L^∞).

1. THE *K*-METHOD

The notion of a compatible couple (X_0, X_1) of Banach spaces X_0 and X_1 was introduced in Section 1 of Chapter III. Recall from Definition III.1.1 that this requires the existence of a Hausdorff topological vector space $\mathscr{X}$ in which each of X_0 and X_1 is continuously embedded. For each compatible couple (X_0, X_1), the sum $X_0 + X_1$ and intersection $X_0 \cap X_1$ are Banach spaces under the norms

$$\|f\|_{X_0+X_1} = \inf\{\|f_0\|_{X_0} + \|f_1\|_{X_1} : f = f_0 + f_1\} \tag{1.1}$$

and

$$\|f\|_{X_0 \cap X_1} = \max\{\|f\|_{X_0}, \|f\|_{X_1}\}, \tag{1.2}$$

respectively (cf. Theorem III.1.3).

The Peetre K- and J-functionals are constructed from these expressions by introducing a positive weighting factor t, as follows:

Definition 1.1. Let (X_0, X_1) be a compatible couple of Banach spaces.

(a) The *K-functional* is defined for each $f \in X_0 + X_1$ and $t > 0$ by

$$K(f, t; X_0, X_1) = \inf\{\|f_0\|_{X_0} + t\|f_1\|_{X_1} : f = f_0 + f_1\}, \tag{1.3}$$

where the infimum extends over all representations $f = f_0 + f_1$ of f with $f_0 \in X_0$ and $f_1 \in X_1$.

(b) The *J-functional* is defined for each $f \in X_0 \cap X_1$ and $t > 0$ by

$$J(f, t; X_0, X_1) = \max\{\|f\|_{X_0}, t\|f\|_{X_1}\}. \tag{1.4}$$

The K-functional will be of primary interest to us in this section. If tX_1

denotes the space X_1 with norm $f \to t\|f\|_{X_1}$, then

$$K(f,t;X_0,X_1) = \|f\|_{X_0+tX_1}, \qquad (f \in X_0 + X_1, t > 0). \tag{1.5}$$

Hence, since

$$\min(1,t)\|f\|_{X_0+X_1} \le K(f,t;X_0,X_1) \le \max(1,t)\|f\|_{X_0+X_1}, \tag{1.6}$$

the functionals $f \to K(f,t;X_0,X_1)$, $(t>0)$, define a family of mutually equivalent *norms* on $X_0 + X_1$.

Since every f in X_0 has the trivial representation $f = f + 0$ as a member of $X_0 + X_1$, it follows immediately from (1.3) that

$$K(f,t;X_0,X_1) \le \|f\|_{X_0}, \qquad (f \in X_0, t > 0). \tag{1.7}$$

Similarly, for f in X_1, we have

$$K(f,t;X_0,X_1) \le t\|f\|_{X_1}, \qquad (f \in X_1, t > 0). \tag{1.8}$$

Hence, if f belongs to $X_0 \cap X_1$, we see from (1.2), (1.7), and (1.8) that

$$K(f,t;X_0,X_1) \le \min(1,t)\|f\|_{X_0 \cap X_1}, \qquad (f \in X_0 \cap X_1, t > 0). \tag{1.9}$$

Proposition 1.2. *For each f in $X_0 + X_1$, the K-functional $K(f,t;X_0,X_1)$ is a nonnegative concave function of $t > 0$, and*

$$t^{-1}K(f,t;X_0,X_1) = K(f,t^{-1};X_1,X_0). \tag{1.10}$$

In particular, $K(f,t;X_0,X_1)$ is increasing on $(0,\infty)$ and $t^{-1}K(f,t;X_0,X_1)$ is decreasing.

Proof. The only property not immediately obvious from the definition is the concavity of $K(f,t) = K(f,t;X_0,X_1)$. To establish this, suppose $t_1, t_2 > 0$ and let t be the convex combination $t = \alpha_1 t_1 + \alpha_2 t_2$, where $\alpha_1, \alpha_2 > 0$ and $\alpha_1 + \alpha_2 = 1$. For each decomposition $f = f_0 + f_1$ of f with $f_i \in X_i$, $(i = 0, 1)$, we have from (1.3),

$$\begin{aligned}\alpha_1 K(f,t_1) + \alpha_2 K(f,t_2) &\le (\alpha_1 + \alpha_2)\|f_0\|_{X_0} + (\alpha_1 t_1 + \alpha_2 t_2)\|f_1\|_{X_1} \\ &= \|f_0\|_{X_0} + t\|f_1\|_{X_1}.\end{aligned}$$

Taking the infimum over all such decompositions $f = f_0 + f_1$ of f, we therefore obtain $\alpha_1 K(f,t_1) + \alpha_2 K(f,t_2) \le K(f,t)$. Hence, $t \to K(f,t)$ is concave, as desired. ■

The space consisting of those f for which $K(f,t;X_0,X_1)$ is a bounded function of t will be of particular interest. Notice from (1.7) that it contains X_0.

We use the following notation, which is suggested by formally allowing $t \to \infty$ in (1.5).

Definition 1.3. Let (X_0, X_1) be a compatible couple and set $K(f,t) = K(f,t;X_0,X_1)$. The space $X_0 + \infty X_1$ consists of all f in $X_0 + X_1$ for which the norm

$$\|f\|_{X_0+\infty X_1} = \sup_{0<t<\infty} K(f,t) = \lim_{t\to\infty} K(f,t) \tag{1.11}$$

is finite.

The space $X_1 + \infty X_0$ of course consists of those f for which $K(f,t;X_1,X_0)$ is bounded. Observe, however, from (1.10) that we may interchange the roles of X_0 and X_1 to obtain the following description of the norm in terms of $K(f,t) = K(f,t;X_0,X_1)$:

$$\|f\|_{X_1+\infty X_0} = \sup_{0<t<\infty} t^{-1}K(f,t) = \lim_{t\to 0} t^{-1}K(f,t). \tag{1.12}$$

Next, we obtain the following useful characterizations of the spaces $X_0 + \infty X_1$ and $X_1 + \infty X_0$.

Theorem 1.4. *Let (X_0, X_1) be a compatible couple. An element f of $X_0 + X_1$ belongs to $X_0 + \infty X_1$ if and only if there is a sequence $(g_n)_{n=1}^\infty$ of elements of X_0 for which*

$$\sup_n \|g_n\|_{X_0} < \infty, \qquad \lim_{n\to\infty} \|f - g_n\|_{X_0+X_1} = 0. \tag{1.13}$$

Furthermore, the norm of f in $X_0 + \infty X_1$ is given by

$$\|f\|_{X_0+\infty X_1} = \inf\left\{\sup_n \|g_n\|_{X_0} : (g_n)_{n=1}^\infty \text{ satisfies (1.13)}\right\}. \tag{1.14}$$

The analogous statements, with X_0 and X_1 interchanged, characterize the space $X_1 + \infty X_0$.

Proof. Let M denote the expression on the right of (1.14) (which, *a priori*, may be infinite). Suppose first that f belongs to $X_0 + \infty X_1$ and let $\varepsilon > 0$. By (1.11), the K-functional is bounded by $\|f\|_{X_0+\infty X_1}$ so, for each $n = 1, 2, \ldots$, there exists a decomposition $f = g_n + h_n$ of f, with $g_n \in X_0$ and $h_n \in X_1$, such that

$$\|g_n\|_{X_0} + n\|h_n\|_{X_1} \le \|f\|_{X_0+\infty X_1} + \varepsilon, \qquad (n = 1, 2, \ldots).$$

In particular,

$$\sup\|g_n\|_{X_0} \le \|f\|_{X_0+\infty X_1} + \varepsilon \tag{1.15}$$

and

$$\|f - g_n\|_{X_0+X_1} = \|h_n\|_{X_0+X_1} \le \|h_n\|_{X_1} \le \frac{1}{n}(\|f\|_{X_0+\infty X_1} + \varepsilon),$$

which tends to 0 as $n \to \infty$. Hence, (g_n) satisfies (1.13). Furthermore, since $\varepsilon > 0$ is arbitrary, it follows from (1.15) and the definition of M that

$$M \le \|f\|_{X_0+\infty X_1}. \tag{1.16}$$

Conversely, suppose now that M is finite. The proof will be complete if we establish the reverse inequality to (1.16). In view of (1.11), this amounts to showing that

$$K(f,t) \le M \tag{1.17}$$

for all $t \ge 1$, say.

Fix $t \ge 1$ and let $\varepsilon > 0$. Since M is finite, there is a sequence $(g_n)_{n=1}^{\infty}$ in X_0 for which

$$\|g_n\|_{X_0} < M + \varepsilon, \qquad (n = 1, 2, \ldots),$$
$$\lim_{n\to\infty} \|f - g_n\|_{X_0+X_1} = 0. \tag{1.18}$$

Hence, for a suitably large value of m,

$$\|g_m\|_{X_0} < M + \varepsilon, \qquad \|f - g_m\|_{X_0+X_1} < \frac{\varepsilon}{t}. \tag{1.19}$$

Then, by (1.6) and (1.7),

$$K(f,t) \le K(f - g_m, t) + K(g_m, t) \le t\|f - g_m\|_{X_0+X_1} + \|g_m\|_{X_0},$$

which, by (1.19), does not exceed $M + 2\varepsilon$. Since $\varepsilon > 0$ is arbitrary, this establishes (1.17) and hence completes the proof. ■

The closure $\overline{X_0}^{X_0+X_1}$ of X_0 in $X_0 + X_1$ consists of those f that can be approximated in the norm of $X_0 + X_1$ by a sequence $(g_n)_{n=1}^{\infty}$ in X_0. The norms in $X_0 + X_1$ of the elements g_n will of course be uniformly bounded in n. In the preceding theorem, there is the additional restriction that the *norms in* X_0 of the elements g_n be bounded. The resulting space $X_0 + \infty X_1$, which obviously is contained in the closure $\overline{X_0}^{X_0+X_1}$, is known as the *Gagliardo completion* of X_0 in $X_0 + X_1$.

It follows easily from Definition 1.3 that $X_0 + \infty X_1$ is a Banach space. From (1.7) and (1.11), we have

$$\|f\|_{X_0+\infty X_1} \le \|f\|_{X_0}, \qquad (f \in X_0) \tag{1.20}$$

and hence, in view of the remarks made above,

$$X_0 \hookrightarrow X_0 + \infty X_1 \hookrightarrow \overline{X_0}^{X_0+X_1}. \tag{1.21}$$

Similarly,

$$\|f\|_{X_1+\infty X_0} \le \|f\|_{X_1}, \qquad (f \in X_1) \tag{1.22}$$

and

$$X_1 \hookrightarrow X_1 + \infty X_0 \hookrightarrow \overline{X_1}^{X_0+X_1}. \tag{1.23}$$

The importance of the Gagliardo completion stems from the following property.

Theorem 1.5. *If* (X_0, X_1) *is a compatible couple, then*

$$K(f,t; X_0, X_1) = K(f,t; X_0 + \infty X_1, X_1 + \infty X_0), \qquad (t > 0) \tag{1.24}$$

for all f *in* $X_0 + X_1$.

Proof. Denote the left- and right-hand sides of (1.24) by $K(f,t)$ and $\bar{K}(f,t)$, respectively. If $f \in X_0 + X_1$ and $f = g + h$ with $g \in X_0$ and $h \in X_1$, it follows directly from (1.20) and (1.22) that

$$\bar{K}(f,t) \le \|g\|_{X_0+\infty X_1} + t\|h\|_{X_1+\infty X_0} \le \|g\|_{X_0} + t\|h\|_{X_1}.$$

Taking the infimum over all such representations $f = g + h$, we obtain

$$\bar{K}(f,t) \le K(f,t). \tag{1.25}$$

Conversely, suppose $f = g + h$ with $g \in X_0 + \infty X_1$ and $h \in X_1 + \infty X_0$. Given $\varepsilon > 0$, we obtain from Theorem 1.4 sequences $(g_n)_{n=1}^{\infty}$ in X_0 and $(h_n)_{n=1}^{\infty}$ in X_1 such that

$$\|g_n\|_{X_0} \le \|g\|_{X_0+\infty X_1} + \varepsilon,$$
$$\|h_n\|_{X_1} \le \|h\|_{X_1+\infty X_0} + \varepsilon \tag{1.26}$$

for $n = 1, 2, \ldots,$ and

$$\lim_{n\to\infty} \|g - g_n\|_{X_0+X_1} = 0 = \lim_{n\to\infty} \|h - h_n\|_{X_0+X_1}.$$

Hence, choosing m sufficiently large and setting $f_m = g_m + h_m$, we have

$$\|f - f_m\|_{X_0+X_1} < \frac{\varepsilon}{\max(1,t)}. \tag{1.27}$$

Then, by (1.6),

$$\begin{aligned} K(f,t) &\le K(f - f_m, t) + K(f_m, t) \\ &\le \max(1,t)\|f - f_m\|_{X_0+X_1} + \|g_m\|_{X_0} + t\|h_m\|_{X_1}. \end{aligned}$$

Using (1.26) and (1.27), we therefore obtain

$$K(f,t) \le \|g\|_{X_0+\infty X_1} + t\|h\|_{X_1+\infty X_0} + 3\varepsilon.$$

Hence, letting $\varepsilon \to 0$ and then passing to the infimum over all such representations $f = g + h$ of f, we conclude that $K(f,t) \le \bar{K}(f,t)$. This, in conjunction with (1.25), establishes (1.24) and hence completes the proof. ■

Before proceeding further, let us pause to illustrate some of these ideas with a specific example. Consider the compatible couple consisting of the Lebesgue spaces L^1 and L^∞ defined with respect to a totally σ-finite measure space (R,μ). Note that the space $(L^1 + L^\infty)(R,\mu)$ considered previously (cf. Definition II.6.1) is indeed the sum of the Banach spaces $L^1(R,\mu)$ and $L^\infty(R,\mu)$ in the sense of Definition III.1.2. Moreover, under the auspices of Theorem II.6.2, we have in effect already computed the K-functional for this couple. The result may be reformulated as follows.

Theorem 1.6. *Let (R,μ) be a totally σ-finite measure space. Then, for each f in $(L^1 + L^\infty)(R,\mu)$,*

$$K(f,t;L^1,L^\infty) = \int_0^t f^*(s)\,ds = tf^{**}(t), \qquad (t > 0). \tag{1.28}$$

Directly from (1.11) and (1.28), we see that

$$\|f\|_{L^1+\infty L^\infty} = \lim_{t\to\infty} \int_0^t f^*(s)\,ds = \|f\|_{L^1}, \tag{1.29}$$

and from (1.12) and (1.28),

$$\|f\|_{L^\infty+\infty L^1} = \lim_{t\to 0} \frac{1}{t}\int_0^t f^*(s)\,ds = \|f\|_{L^\infty}. \tag{1.30}$$

The Gagliardo completions of L^1 and L^∞ in $L^1 + L^\infty$ therefore coincide with the original spaces L^1 and L^∞ themselves. Consequently, Theorem 1.5 provides nothing new in these cases.

Consider now, however, the spaces C of continuous functions and L^1 of integrable functions on the closed interval $[0,1]$, say. Since $L^1 + C = L^1$, the Gagliardo completion $L^1 + \infty C$ of L^1 in $L^1 + C$ is simply L^1 again. On the other hand, Theorem 1.4 identifies the Gagliardo completion $C + \infty L^1$ of C in $L^1 + C = L^1$ as the space of integrable functions on $[0,1]$ that can be approximated in the L^1-norm by a sequence of uniformly bounded continuous functions. It is not hard to see that this is precisely the space L^∞:

$$C + \infty L^1 = L^\infty. \tag{1.31}$$

Hence, Theorem 1.5 provides a description of the K-functional for the pair (L^1, C):

$$K(f,t;L^1,C) = K(f,t;L^1,L^\infty) = \int_0^t f^*(s)\,ds, \qquad (t > 0). \tag{1.32}$$

The space C cannot be retrieved from $K(f,t;L^1,C)$ by imposing any kind of growth condition. Instead, it is the Gagliardo completion of C, namely, the space L^∞, that arises in this way (cf. (1.11)).

We have seen in Section IV.4 that the Marcinkiewicz interpolation theorem has a natural formulation in terms of the Lorentz $L^{p,q}$-spaces. Now, if $1 < p < \infty$, $1 \le q \le \infty$, then $L^{p,q}$ is an intermediate space between L^1 and L^∞. Furthermore, Theorem 1.6 shows that the $L^{p,q}$-norm (cf. Definition IV.4.4) can be defined entirely in terms of the K-functional for (L^1, L^∞). It is therefore a simple matter to impose an entirely analogous structure on any compatible couple (X_0, X_1) by defining a two-parameter family of intermediate spaces as follows:

Definition 1.7. Let (X_0, X_1) be a compatible couple and suppose $0 < \theta < 1$, $1 \le q < \infty$ or $0 \le \theta \le 1$, $q = \infty$. The space $(X_0, X_1)_{\theta,q}$ consists of all f in $X_0 + X_1$ for which the functional

$$\|f\|_{\theta,q} = \begin{cases} \left\{\int_0^\infty [t^{-\theta}K(f,t)]^q \frac{dt}{t}\right\}^{1/q}, & 0 < \theta < 1,\ 1 \le q < \infty, \\ \sup_{t>0} t^{-\theta}K(f,t), & 0 \le \theta \le 1,\ q = \infty. \end{cases} \tag{1.33}$$

is finite (here, as usual, $K(f,t) = K(f,t;X_0,X_1)$).

Note from Definitions 1.3 and 1.7 that the $(0,\infty)$ and $(1,\infty)$ spaces so defined are simply the Gagliardo completions of X_0 and X_1, respectively (with identical norms):

$$(X_0, X_1)_{0,\infty} = X_0 + \infty X_1, \qquad (X_0, X_1)_{1,\infty} = X_1 + \infty X_0. \tag{1.34}$$

Proposition 1.8. *Let (X_0, X_1) be a compatible couple of Banach spaces and suppose $0 < \theta < 1$, $1 \le q < \infty$ or $0 \le \theta \le 1$, $q = \infty$. Then $(X_0, X_1)_{\theta,q}$, equipped with the norm (1.33), is a Banach space intermediate between X_0 and X_1:*

$$X_0 \cap X_1 \hookrightarrow (X_0, X_1)_{\theta,q} \hookrightarrow X_0 + X_1. \tag{1.35}$$

Proof. That the expression in (1.33) indeed defines a norm is evident from Minkowski's inequality and the subadditivity of the K-functional. To establish completeness, it will suffice to show that any absolutely summable series in $(X_0, X_1)_{\theta,q}$ is summable. Suppose therefore that $(f_n)_{n=1}^{\infty}$ is a sequence of elements of $X_0 + X_1$ with $\sum_n \|f_n\|_{\theta,q}$ finite. Since $L_q(dt/t)$ is complete, the series $t^{-\theta}\sum_n K(f_n, t)$ converges in $L^q(dt/t)$ and hence (since each of the K-functionals is an increasing function) is finite for all $t > 0$. Now $X_0 + X_1$ is also complete, so this implies that $K(\sum f_n, t) \le \sum K(f_n, t)$. Hence, multiplying by $t^{-\theta}$ and applying the $L^q(dt/t)$-norm to each side, we obtain

$$\|\sum f_n\|_{\theta,q} \le \sum \|f_n\|_{\theta,q} < \infty,$$

and this establishes the completeness of $(X_0, X_1)_{\theta,q}$.

The first of the embeddings in (1.35) follows immediately by substituting the estimate (1.9) into (1.33). The second follows in similar fashion by using (1.6) instead of (1.9). ■

As a consequence of Theorem 1.6 (and Definitions 1.7 and IV.4.4), we have the following result.

Theorem 1.9. *If $0 < \theta < 1$ and $1 \le q \le \infty$, then*

$$(L^1, L^\infty)_{\theta,q} = L^{p,q}, \tag{1.36}$$

where $1/p = 1 - \theta$.

Given that the structure of the spaces $(X_0, X_1)_{\theta,q}$ is modeled on that of the $L^{p,q}$-spaces, it is to be expected that they will satisfy similar inclusion relations. Indeed, essentially the same argument used to prove Proposition IV.4.2 also serves to establish the following result.

Proposition 1.10. *If $0 < \theta < 1$ and $1 \le q \le r \le \infty$, then*

$$(X_0, X_1)_{\theta,q} \hookrightarrow (X_0, X_1)_{\theta,r}. \tag{1.37}$$

Now let us examine the interpolation structure of the (θ, q)-spaces. Consider an admissible operator T with respect to two compatible couples (X_0, X_1) and (Y_0, Y_1). Thus (cf. Definition III.1.5), T maps $X_0 + X_1$ into $Y_0 + Y_1$ and the restriction of T to X_i is a bounded operator from X_i into Y_i with norm M_i, say $(i = 0, 1)$. The importance of the K-functional in interpolation theory stems from the following fundamental inequality.

Theorem 1.11. *Let T be an admissible linear operator with respect to compatible couples (X_0, X_1) and (Y_0, Y_1), Then*

$$K(Tf, t; Y_0, Y_1) \le M_0 K(f, tM_1/M_0; X_0, X_1) \tag{1.38}$$

for all f in $X_0 + X_1$ and all $t > 0$.

Proof. The admissible operator T satisfies

$$\|Tf_i\|_{Y_i} \le M_i \|f\|_{X_i}, \qquad (f_i \in X_i, i = 0, 1). \tag{1.39}$$

If $f \in X_0 + X_1$ and $f = f_0 + f_1$ is any decomposition of f with $f_i \in X_i$, $(i = 0, 1)$, then $Tf = Tf_0 + Tf_1$ and $Tf_i \in Y_i$, $(i = 0, 1)$. Hence, by (1.39),

$$\begin{aligned} K(Tf, t; Y_0, Y_1) &\le \|Tf_0\|_{Y_0} + t\|Tf_1\|_{Y_1} \\ &\le M_0\left(\|f_0\|_{X_0} + t\frac{M_1}{M_0}\|f_1\|_{X_1}\right). \end{aligned}$$

Taking the infimum over all such representations $f = f_0 + f_1$ of f, we obtain (1.38). ■

Applying this result to the (θ, q)-spaces, we obtain the following basic interpolation theorem.

Theorem 1.12. *Let (X_0, X_1) and (Y_0, Y_1) be compatible couples and let $0 < \theta < 1$, $1 \le q < \infty$ or $0 \le \theta \le 1$, $q = \infty$. Let T be an admissible linear operator with respect to (X_0, X_1) and (Y_0, Y_1):*

$$\|Tf_i\|_{Y_i} \le M_i \|f_i\|_{X_i}, \qquad (f \in X_i, i = 0, 1). \tag{1.40}$$

Then

$$T:(X_0, X_1)_{\theta,q} \to (Y_0, Y_1)_{\theta,q}, \tag{1.41}$$

and, in fact,

$$\|Tf\|_{\theta,q} \le M_0^{1-\theta} M_1^{\theta} \|f\|_{\theta,q} \tag{1.42}$$

for all f in $(X_0, X_1)_{\theta,q}$.

Proof. The desired estimate (1.42) follows by applying the (θ, q)-norm (cf. Definition 1.7) to each side of (1.38) and making a change of variables in the right-hand side. ■

It is not necessary to limit consideration to the (θ, q)-norms. More general interpolation spaces may be constructed in a number of ways. In order to build on the known structure associated with the seminal pair (L^1, L^∞), it will be convenient to define the interpolation spaces for more general couples (X_0, X_1) in terms of the k-functional, which is defined as follows.

Definition 1.13. Let (X_0, X_1) be a compatible couple of Banach spaces. The K-functional $K(f, t; X_0, X_1)$ is a nonnegative concave function of $t \ge 0$ (Proposition 1.2) and hence may be represented in the form

$$K(f, t; X_0, X_1) = K(f, 0+; X_0, X_1) + \int_0^t k(f, s; X_0, X_1)\, ds, \tag{1.43}$$

where the *k-functional* $k(f, s; X_0, X_1)$ is a uniquely defined nonnegative, decreasing, right-continuous function of $s > 0$.

The values of $k(f, s)$ at its jumps could have been defined otherwise; we select right-continuity so as to have the following result, which follows at once from (1.28) and (1.43).

Proposition 1.14. *The k-functional for the couple (L^1, L^∞) is given by*

$$k(f, t; L^1, L^\infty) = f^*(t), \qquad (t > 0;\ f \in L^1 + L^\infty). \tag{1.44}$$

Notice, also from (1.28), that $K(f, 0+; L^1, L^\infty) = 0$ for all f in $L^1 + L^\infty$. The following more general result holds. When (X_0, X_1) is a compatible couple, we shall denote by X_0^o the closure of $X_0 \cap X_1$ in X_0:

$$X_0^o = \overline{(X_0 \cap X_1)}^{X_0}.$$

Proposition 1.15. *Let (X_0, X_1) be a compatible couple and let $f \in X_0 + X_1$. The following assertions are equivalent:*

(i) $$K(f, 0+; X_0, X_1) = 0;$$

(ii) $$K(f, t; X_0, X_1) = \int_0^t k(f, s; X_0, X_1)\, ds, \qquad (t > 0);$$

(iii) $$f \in X_0^o + X_1 = \overline{(X_0 \cap X_1)}^{X_0} + X_1;$$

(iv) $$f \in \bar{X}_1^{X_0 + X_1}.$$

In particular, property (i) *holds for all $f \in X_0 + X_1$ if and only if $X_0 \cap X_1$ is dense in X_0.*

Proof. The equivalence of (i) and (ii) is immediate from (1.43). To see that (i) implies (iv), suppose $K(f, 0+) = 0$ and suppose without loss of generality that $f \ne 0$. For each $n = 1, 2, \ldots$, there exist $g_n \in X_0$ and $h_n \in X_1$ such that $f = g_n + h_n$ and

$$\|g_n\|_{X_0} + 2^{-n}\|h_n\|_{X_1} \le 2K(f, 2^{-n}).$$

In particular, $g_n \to 0$ in X_0, hence in $X_0 + X_1$, as $n \to \infty$. This shows that $f - h_n = g_n \to 0$ in $X_0 + X_1$ as $n \to \infty$, and hence that property iv) holds.

Next, we show that iii) implies i). Suppose f satisfies iii) and let $\varepsilon > 0$. If $f = g + h$, with $g \in X_0^o$ and $h \in X_1$, there exists $g' \in X_0 \cap X_1$ whose distance in X_0 from g is at most ε. Writing $f = (g - g') + (g' + h)$, we have, for any $t > 0$,

$$K(f, t) \le \|g - g'\|_{X_0} + t\|g' + h\|_{X_1}.$$

Letting $t \to 0$, we conclude that $K(f, 0+) \le \varepsilon$ and hence, since ε is arbitrary, that (i) holds.

The same construction shows also that (iii) implies (iv). Indeed, the function $f' = g' + h$ belongs to X_1. But $f - f' = g - g'$, so $f - f'$ has norm in X_0, hence also in $X_0 + X_1$, at most ε. Since ε is arbitrary, we conclude that iv) holds.

To show that (iv) implies (iii), suppose (iv) holds and select elements $h_n \in X_1$ such that

$$\|f - h_n\|_{X_0 + X_1} < 2^{-n}, \qquad (n = 1, 2, \ldots).$$

Each $f - h_n$ belongs to $X_0 + X_1$ so there exist $p_n \in X_0$ and $q_n \in X_1$ such that $f - h_n = p_n + q_n$ and

$$\|p_n\|_{X_0} + \|q_n\|_{X_1} \le 2\|f - h_n\|_{X_0 + X_1} < 2^{1-n}.$$

Write $f = p_1 + (q_1 + h_1)$. Clearly $q_1 + h_1 \in X_1$. The proof will therefore be completed by showing that $p_1 \in X_0^o$. For this, we merely observe from the preceding estimate that

$$p_1 = \sum_{n=1}^{\infty} (p_n - p_{n+1}) \qquad \text{in } X_0,$$

and that each $p_n - p_{n+1}$ belongs to $X_0 \cap X_1$ (because of the identity

$$p_n - p_{n+1} = q_{n+1} - q_n + h_{n+1} - h_n,$$

the left-hand side of which belongs to X_0, the right to X_1).

Finally, X_0^o is a closed subspace of X_0 and it is easy to show that $X_0^o + X_1$ coincides with $X_0 + X_1$ if and only if X_0^o coincides with X_0, that is, if and only if $X_0 \cap X_1$ is dense in X_0. The last statement of the theorem therefore follows from the equivalence of (i) and (iii) already established. This completes the proof. ■

Definition 1.16. A *Riesz-Fischer norm* is a rearrangement-invariant functional $\bar{\rho}: \mathscr{M}_0^+(\mathbf{R}^+, m) \to [0, \infty]$ which satisfies properties (P1), (P2), (P4), (P5) of Definition I.1.1 and

$$\bar{\rho}\left(\sum_n f_n\right) \leq \sum_n \bar{\rho}(f_n), \tag{1.45}$$

for all sequences (f_n) in $\mathscr{M}_0^+(\mathbf{R}^+, m)$.

Riesz-Fischer norms ρ are defined on other totally σ-finite measure spaces (R, μ) by

$$\rho(f) = \bar{\rho}(f^*), \tag{1.46}$$

where $\bar{\rho}$ is a Riesz-Fischer norm on $(\mathbf{R}^+, m)$. The space $X = X(\rho)$ of functions f in $\mathscr{M}_0(R, \mu)$ for which $\rho(|f|) < \infty$ is called a *Riesz-Fischer space.* The norm on X is given by $\|f\|_X = \rho(|f|)$.

Every rearrangement-invariant Banach function space on a resonant measure space (R, μ) is a Riesz-Fischer space. Indeed, the Luxemburg representation theorem (Theorem II.4.10) provides a representation of the form (1.46) for the norm, and the Fatou property of Banach function norms implies the Riesz-Fischer property (1.45) (Theorem I.1.6).

Perhaps the simplest example of a Riesz-Fischer space that is not a Banach function space is the sequence space c_0, which is generated by the norm

$$\rho(c) = \begin{cases} \sup_n |c_n| & \text{if } \lim_{n\to\infty} c_n = 0; \\ \infty & \text{otherwise.} \end{cases}$$

The space c_0 is complete and hence has the Riesz-Fischer property, but examination of the finite sequences which are equal to 1 in their first n entries and 0 thereafter shows that c_0 does not have the Fatou property. Nevertheless, the space c_0 is an interpolation space between ℓ^1 and ℓ^∞, as is clear from Proposition III.2.1.

Since the Lorentz-Luxemburg property (Theorem I.2.7) may fail in Riesz-Fischer spaces, there is no guarantee that the monotonicity property

$$g \prec f, \qquad f \in X \quad \Rightarrow \quad g \in X, \qquad \|g\|_X \leq \|f\|_X, \tag{1.47}$$

holds either. We shall refer to spaces in which (1.47) holds as *monotone Riesz-Fischer spaces.* Corollary II.4.7 shows that every rearrangement-invariant Banach function space is a monotone Riesz-Fischer space. The importance of the monotone Riesz-Fischer spaces is seen in the next result.

Theorem 1.17. *The exact interpolation spaces between L^1 and L^∞ (with respect to a resonant measure space) are precisely the monotone Riesz-Fischer spaces.*

Proof. The proof follows exactly the same lines as that of Theorem III.2.12. ■

Using the monotone Riesz-Fischer norms and the k-functional $k(f, t)$, we may generate intermediate spaces for general couples (X_0, X_1) in the following way.

Definition 1.18. Let (X_0, X_1) be a compatible couple and let ρ be a monotone Riesz-Fischer norm on $(\mathbf{R}^+, m)$. The space $(X_0, X_1)_\rho$ consists of those $f \in X_0^o + X_1$ for which

$$\|f\|_{(X_0, X_1)_\rho} = \rho(k(f, \cdot)) < \infty. \tag{1.48}$$

By Proposition 1.14 and Theorem 1.17, the exact interpolation spaces between L^1 and L^∞ are precisely the spaces

$$X = L^\rho = (L^1, L^\infty)_\rho \tag{1.49}$$

generated by monotone Riesz-Fischer norms ρ. The next result shows that interpolation spaces for more general couples may also be generated by this method.

Theorem 1.19. *Let (X_0, X_1) be a compatible couple and let ρ be a monotone*

Riesz-Fischer norm on $(\mathbf{R}^+, m)$. *If* (Y_0, Y_1) *is any other compatible couple, then the pair* $((X_0, X_1)_\rho, (Y_0, Y_1)_\rho)$ *has the interpolation property with respect to the couples* (X_0, X_1) *and* (Y_0, Y_1).

Proof. It is routine to verify that $X = (X_0, X_1)_\rho$ is a Banach space. Only the triangle inequality requires comment. If f and g belong to $X_\rho \equiv (X_0, X_1)_\rho$, then both belong to $X_0^o + X_1$ and so Proposition 1.15 shows that $K(f, 0+) = K(g, 0+) = 0$. The subadditivity of the K-functional therefore shows that $K(f+g; 0+) = 0$ and that

$$\int_0^t k(f+g, s)\,ds \le \int_0^t k(f, s)\,ds + \int_0^t k(g, s)\,ds, \qquad (t > 0),$$

or, equivalently, that $k(f+g, \cdot) \prec k(f, \cdot) + k(g, \cdot)$. The monotonicity (1.47) and subadditivity of ρ therefore give

$$\rho(k(f+g, \cdot)) \le \rho(k(f, \cdot) + k(g, \cdot)) \le \rho(k(f, \cdot)) + \rho(k(g, \cdot)),$$

which establishes the triangle inequality in X_ρ.

The proof that X_ρ is intermediate between X_0 and X_1 is similar. If $f \in X_0 \cap X_1$, it follows from (1.9) that

$$k(f, \cdot) \prec \chi_{(0,1)}(\cdot)\|f\|_{X_0 \cap X_1}.$$

Hence, with $c = \rho(\chi_{(0,1)})$, which is finite because of property (P4) of Definition I.1.1, we conclude that

$$\|f\|_{X_\rho} \le c\|f\|_{X_0 \cap X_1}, \qquad (f \in X_0 \cap X_1). \tag{1.50}$$

On the other hand, if $f \in X_\rho$, then (1.6) gives

$$\chi_{(0,1)}(\cdot)\|f\|_{X_0 + X_1} \prec k(f, \cdot),$$

and hence, with c as above,

$$c\|f\|_{X_0 + X_1} \le \|f\|_{X_\rho}. \tag{1.51}$$

The estimates (1.50) and (1.51) together show that X is an intermediate space for (X_0, X_1).

Next, suppose T is an admissible operator with respect to the couples (X_0, X_1) and (Y_0, Y_1). It follows from (1.38) and the concavity of $K(f, \cdot)$ that

$$K(Tf, t; Y_0, Y_1) \le \max(M_0, M_1) K(f, t; X_0, X_1), \qquad (t > 0). \tag{1.52}$$

If $f \in X_\rho$, then $K(f, 0+) = 0$ and so (1.52) shows that the same is true of Tf: $K(Tf, 0+) = 0$. Hence, $Tf \in Y_0^o + Y_1$, and (1.52) assumes the equivalent form

$$k(Tf, \cdot; Y_0, Y_1) \prec \max(M_0, M_1) k(f, \cdot; X_0, X_1). \tag{1.53}$$

Once again, the monotonicity (1.47) gives the desired interpolation estimate

$$\|Tf\|_{(Y_0, Y_1)_\rho} \le \max(M_0, M_1)\|f\|_{(X_0, X_1)_\rho}. \tag{1.54}$$

In particular, X_ρ is an interpolation space between X_0 and X_1. This completes the proof. ■

The converse, whether all interpolation spaces are of the type $(X_0, X_1)_\rho$, for suitable ρ, will be taken up in Section 3.

As yet, we have identified the K-functional in concrete terms for only one couple, namely, (L^1, L^∞) (cf. Theorem 1.6). Much of the remainder of this chapter will be devoted to determining the K-functional in many other situations that occur commonly in practice. We make a start in the next section with the so-called *reiteration theorem.* This general result characterizes the K-functional of a couple only in terms of the K-functional of some larger "universal" couple, but nevertheless is quite useful in practice. It provides, for example, descriptions of the K-functionals for the pairs (L^p, L^q), $(1 \le p, q \le \infty)$, in terms of the K-functional for (L^1, L^∞), namely, the functional $tf^{**}(t)$.

2. STRUCTURE THEOREMS FOR THE (θ, q)-SPACES

Let (X_0, X_1) be a compatible couple and consider two interpolation spaces

$$\bar{X}_{\theta_0} = (X_0, X_1)_{\theta_0, q_0}, \qquad \bar{X}_{\theta_1} = (X_0, X_1)_{\theta_1, q_1}, \tag{2.1}$$

where $0 < \theta_0 < \theta_1 < 1$ and $1 \le q_0, q_1 \le \infty$. Then $(\bar{X}_{\theta_0}, \bar{X}_{\theta_1})$ is itself a compatible couple and so we may form the corresponding interpolation spaces $(\bar{X}_{\theta_0}, \bar{X}_{\theta_1})_{\theta, q}$ for $0 < \theta < 1$ and $1 \le q \le \infty$. The *reiteration theorem* (or *stability theorem*), our first objective in this section, will show that these interpolation spaces can in fact be obtained as interpolation spaces from the original couple (X_0, X_1), as follows:

$$(\bar{X}_{\theta_0}, \bar{X}_{\theta_1})_{\theta, q} = (X_0, X_1)_{\theta', q}, \qquad (\theta' = (1-\theta)\theta_0 + \theta\theta_1). \tag{2.2}$$

The interpolation process is thus stable in the sense that reiteration of it leads to the same interpolation spaces as before.

The first step will be to relate the K-functionals of the underlying couples (X_0, X_1) and $(\bar{X}_{\theta_0}, \bar{X}_{\theta_1})$. We shall write

$$K(f, t) = K(f, t; X_0, X_1), \qquad \bar{K}(f, t) = K(f, t; \bar{X}_{\theta_0}, \bar{X}_{\theta_1}). \tag{2.3}$$

Theorem 2.1 (T. Holmstedt). *Let* (X_0, X_1) *be a compatible couple and suppose* $0 < \theta_0 < \theta_1 < 1$ *and* $1 \le q_0, q_1 \le \infty$. *Let* $\delta = \theta_1 - \theta_0$. *Then, with the* *notation of* (2.1) *and* (2.3),

$$\bar{K}(f,t^\delta) \sim \left\{\int_0^t [s^{-\theta_0}K(f,s)]^{q_0}\frac{ds}{s}\right\}^{1/q_0} + t\left\{\int_t^\infty [s^{-\theta_1}K(f,s)]^{q_1}\frac{ds}{s}\right\}^{1/q_1}, \tag{2.4}$$

for all f in $\bar{X}_{\theta_0} + \bar{X}_{\theta_1}$ and all $t > 0$; if q_0 or q_1 is infinite, the corresponding integral in (2.4) *is replaced by the supremum in the usual way.*

Proof. For $j = 0$ and 1, let

$$P_j g(t) = \left\{\int_0^t [s^{-\theta_j}K(g,s)]^{q_j}\frac{ds}{s}\right\}^{1/q_j} \tag{2.5}$$

and

$$Q_j g(t) = \left\{\int_t^\infty [s^{-\theta_j}K(g,s)]^{q_j}\frac{ds}{s}\right\}^{1/q_j} \tag{2.6}$$

(with the obvious modification if $q_j = \infty$). The desired result (2.4) is then expressible in the form

$$\bar{K}(f,t^\delta) \sim P_0 f(t) + t^\delta Q_1 f(t). \tag{2.7}$$

Suppose first that $f \in \bar{X}_{\theta_0} + \bar{X}_{\theta_1}$ and fix $t > 0$. Let $f = g + h$ be any representation of f with $g \in \bar{X}_{\theta_0}$ and $h \in \bar{X}_{\theta_1}$. Then, using the subadditivity of P_0 and Q_1 together with the relations (2.1), we obtain

$$P_0 f(t) + t^\delta Q_1 f(t) \le \|g\|_{\bar{X}_{\theta_0}} + P_0 h(t) + t^\delta(Q_1 g(t) + \|h\|_{\bar{X}_{\theta_1}}). \tag{2.8}$$

However, by (2.1) and Proposition 1.10, we have $\bar{X}_{\theta_0} \hookrightarrow (X_0, X_1)_{\theta_0,\infty}$ and so

$$K(g,s) \le cs^{\theta_0}\|g\|_{\bar{X}_{\theta_0}}, \qquad (s > 0). \tag{2.9}$$

Substituting this into (2.6) and performing the integration, we find that

$$Q_1 g(t) \le ct^{-\delta}\|g\|_{\bar{X}_{\theta_0}}.$$

Similarly,

$$P_0 h(t) \le ct^\delta\|h\|_{\bar{X}_{\theta_1}},$$

so these estimates combine with (2.8) to give

$$P_0 f(t) + t^\delta Q_1 f(t) \le c(\|g\|_{\bar{X}_{\theta_0}} + t^\delta\|h\|_{\bar{X}_{\theta_1}}).$$

Hence, passing to the infimum over all such representations $f = g + h$ of f, we conclude that

$$P_0 f(t) + t^\delta Q_1 f(t) \le c\bar{K}(f,t^\delta). \tag{2.10}$$

Conversely, suppose $f \in X_0 + X_1$ and that $P_0 f(t)$ and $Q_1 f(t)$ are finite. We shall show that $f \in \bar{X}_{\theta_0} + \bar{X}_{\theta_1}$ and

$$\bar{K}(f,t^\delta) \le c(P_0 f(t) + t^\delta Q_1 f(t)). \tag{2.11}$$

This, together with (2.10), will establish the desired equivalence (2.7) and hence complete the proof.

Let $f = g + h$ be a representation of f with $g \in X_0$, $h \in X_1$, and

$$\|g\|_{X_0} + t\|h\|_{X_1} \le 2K(f,t).$$

Then, by (1.7),

$$K(g,s) \le \|g\|_{X_0} \le 2K(f,t), \qquad (s > 0) \tag{2.12}$$

and, by (1.8),

$$K(h,s) \le s\|h\|_{X_1} \le \frac{2s}{t}K(f,t), \qquad (s > 0). \tag{2.13}$$

Substituting (2.12) into (2.6), we obtain

$$Q_0 g(t) = \left\{\int_t^\infty [s^{-\theta_0}K(g,s)]^{q_0}\frac{ds}{s}\right\}^{1/q_0}$$
$$\le ct^{-\theta_0}K(f,t) \le cP_0 f(t).$$

Similarly, from (2.13) and (2.5),

$$P_0 h(t) = \left\{\int_0^t [s^{-\theta_0}K(h,s)]^{q_0}\frac{ds}{s}\right\}^{1/q_0}$$
$$\le ct^{-\theta_0}K(f,t) \le cP_0 f(t).$$

Since $g = f - h$, so $P_0 g \le P_0 f + P_0 h$, these estimates give

$$\|g\|_{\bar{X}_{\theta_0}} \le (P_0 + Q_0)g(t) \le cP_0 f(t) < \infty, \tag{2.14}$$

which shows, in particular, that $g \in \bar{X}_{\theta_0}$.

Similarly, one shows that both $P_1 h$ and $Q_1 g$ are majorized by $Q_1 f$. Hence,

$$\|h\|_{\bar{X}_{\theta_1}} \le (P_1 + Q_1)h(t) \le cQ_1 f(t) < \infty, \tag{2.15}$$

and so $h \in \bar{X}_{\theta_1}$. Since $f = g + h$, it now follows from (2.14) and (2.15) that

$$\bar{K}(f,t^\delta) \le \|g\|_{\bar{X}_{\theta_0}} + t^\delta\|h\|_{\bar{X}_{\theta_1}} \le c(P_0 f(t) + t^\delta Q_1 f(t)),$$

which is the desired estimate (2.11). ∎

Definition 2.2. Suppose $0 \le \theta \le 1$. An intermediate space X of a compatible couple (X_0, X_1) is said to be of *class* θ if

$$(X_0, X_1)_{\theta,1} \hookrightarrow X \hookrightarrow (X_0, X_1)_{\theta,\infty}, \qquad (0 < \theta < 1), \tag{2.16}$$

or

$$X_0 \hookrightarrow X \hookrightarrow (X_0, X_1)_{0,\infty} \equiv X_0 + \infty X_1, \qquad (\theta = 0), \tag{2.17}$$

or

$$X_1 \hookrightarrow X \hookrightarrow (X_0, X_1)_{1,\infty} \equiv X_1 + \infty X_0, \qquad (\theta = 1). \tag{2.18}$$

The following result corresponds to the extreme cases $\theta_0 = 0$ and $\theta_1 = 1$ of Theorem 2.1. As usual, we write $K(f, t)$ for $K(f, t; X_0, X_1)$.

Corollary 2.3. *Let (X_0, X_1) be a compatible couple and let $\bar{X}_0$ and $\bar{X}_1$ be intermediate spaces of (X_0, X_1) of class 0 and 1, respectively. Suppose $0 < \theta < 1$, $1 \le q \le \infty$, and let $X_{\theta,q} = (X_0, X_1)_{\theta,q}$. Then*

$$K(f, t^\theta; \bar{X}_0, X_{\theta,q}) \sim t^\theta \left\{ \int_t^\infty [s^{-\theta} K(f, s)]^q \frac{ds}{s} \right\}^{1/q} \tag{2.19}$$

and

$$K(f, t^{1-\theta}; X_{\theta,q}, \bar{X}_1) \sim \left\{ \int_0^t [s^{-\theta} K(f, s)]^q \frac{ds}{s} \right\}^{1/q} \tag{2.20}$$

(with obvious modification if $q = \infty$).

Proof. The assertion (2.19) with $\bar{X}_0 = X_0$, that is,

$$K(f, t^\theta; X_0, X_{\theta,q}) \sim t^\theta \left\{ \int_t^\infty [s^{-\theta} K(f, s)]^q \frac{ds}{s} \right\}^{1/q} \tag{2.21}$$

can be established by essentially the same argument used in the proof of the preceding theorem. Next, if $\tilde{X}_0$ and $\tilde{X}_1$ denote the Gagliardo completions $X_0 + \infty X_1$ and $X_1 + \infty X_0$, respectively, then the K-functionals for the couples (X_0, X_1) and $(\tilde{X}_0, \tilde{X}_1)$ coincide (Theorem 1.5). Hence, in particular,

$$\tilde{X}_{\theta,q} = (\tilde{X}_0, \tilde{X}_1)_{\theta,q} = (X_0, X_1)_{\theta,q} = X_{\theta,q}.$$

Applying (2.21) to the couple $(\tilde{X}_0, \tilde{X}_1)$ instead of (X_0, X_1), we therefore obtain

$$\begin{aligned} K(f, t^\theta; \tilde{X}_0, X_{\theta,q}) &= K(f, t^\theta; \tilde{X}_0, \tilde{X}_{\theta,q}) \\ &\sim t^\theta \left\{ \int_t^\infty [s^{-\theta} K(f, s; \tilde{X}_0, \tilde{X}_1)]^q \frac{ds}{s} \right\}^{1/q} \\ &\sim t^\theta \left\{ \int_t^\infty [s^{-\theta} K(f, s)]^q \frac{ds}{s} \right\}^{1/q}. \end{aligned} \tag{2.22}$$

Thus, (2.21) and (2.22) show that (2.19) holds if $\bar{X}_0$ is equal to X_0 or to $\tilde{X}_0 = X_0 + \infty X_1$. Hence, because of the embeddings (2.17), the assertion (2.19) holds whenever $\bar{X}_0$ is of class 0. A similar argument shows that (2.20) holds for any $\bar{X}_1$ of class 1. ■

Now we can present the reiteration theorem. Note that the rather rigid requirement imposed in (2.1) that $\bar{X}_{\theta_j}$, $(j = 0, 1)$, be an interpolation space is not necessary. All that is required is that $\bar{X}_{\theta_j}$ be an intermediate space of class θ_j, $(j = 0, 1)$.

Theorem 2.4. (Reiteration theorem). *Let (X_0, X_1) be a compatible couple and suppose $0 \le \theta_0 < \theta_1 \le 1$. Let $\bar{X}_{\theta_j}$ be an intermediate space of (X_0, X_1) of class θ_j, $(j = 0, 1)$. If $0 < \theta < 1$ and $1 \le q \le \infty$, then*

$$(\bar{X}_{\theta_0}, \bar{X}_{\theta_1})_{\theta,q} = (X_0, X_1)_{\theta',q}, \tag{2.23}$$

with equivalent norms, where

$$\theta' = (1 - \theta)\theta_0 + \theta\theta_1. \tag{2.24}$$

Proof. Suppose first that $0 < \theta_0 < \theta_1 < 1$ and let $\delta = \theta_1 - \theta_0$, By (2.16),

$$(X_0, X_1)_{\theta_j,1} \hookrightarrow \bar{X}_{\theta_j} \hookrightarrow (X_0, X_1)_{\theta_j,\infty}, \qquad (j = 0, 1). \tag{2.25}$$

Applying (2.4) (with $q_0 = q_1 = \infty$) and using the second of the embeddings in (2.25), we obtain

$$\begin{aligned} t^{-\theta_0} K(f, t) &\le c K(f, t^\delta; (X_0, X_1)_{\theta_0,\infty}, (X_0, X_1)_{\theta_1,\infty}) \\ &\le c K(f, t^\delta; \bar{X}_{\theta_0}, \bar{X}_{\theta_1}) = c\bar{K}(f, t^\delta). \end{aligned}$$

If q is finite, this, together with (2.24), gives

$$\begin{aligned} \|f\|^q_{(X_0,X_1)_{\theta',q}} &= \int_0^\infty [t^{-\theta'} K(f, t)]^q \frac{dt}{t} \\ &\le c \int_0^\infty [t^{\theta_0 - \theta'} \bar{K}(f, t^\delta)]^q \frac{dt}{t} \\ &= c \int_0^\infty [s^{(\theta_0 - \theta')/(\theta_1 - \theta_0)} \bar{K}(f, s)]^q \frac{ds}{s} \\ &= c \int_0^\infty [s^{-\theta} \bar{K}(f, s)]^q \frac{ds}{s} \\ &= c \|f\|^q_{(\bar{X}_{\theta_0}, \bar{X}_{\theta_1})_{\theta,q}}. \end{aligned}$$

Hence, by taking q-th roots, we see that the space on the left of (2.23) is continuously embedded in the one on the right. The proof for $q = \infty$ is similar.

To complete the proof, we need to establish the embedding in the opposite direction, that is, we need to show that

$$\|f\|_{(\bar{X}_{\theta_0}, \bar{X}_{\theta_1})_{\theta,q}} \le c\|f\|_{(X_0,X_1)_{\theta',q}}. \tag{2.26}$$

Using the first of the embeddings in (2.25) and applying (2.4) (with $q_0 = q_1 = 1$), we obtain

$$\begin{aligned}\bar{K}(f,t^\delta) &= K(f,t^\delta; \bar{X}_{\theta_0}, \bar{X}_{\theta_1}) \\ &\le cK(f,t^\delta; (X_0,X_1)_{\theta_0,1}, (X_0,X_1)_{\theta_1,1}) \\ &\le c\left\{\int_0^t s^{-\theta_0}K(f,s)\frac{ds}{s} + t^\delta \int_t^\infty s^{-\theta_1}K(f,s)\frac{ds}{s}\right\}.\end{aligned}$$

With a change of variables and a call to Minkowski's inequality, this gives (for $q < \infty$),

$$\begin{aligned}\left\{\int_0^\infty [t^{-\theta}\bar{K}(f,t)]^q \frac{dt}{t}\right\}^{1/q} &\le c\left\{\left(\int_0^\infty \left[t^{-\delta\theta}\int_0^t s^{-\theta_0}K(f,s)\frac{ds}{s}\right]^q \frac{dt}{t}\right)^{1/q}\right. \\ &\quad \left.+\left(\int_0^\infty \left[t^{\delta(1-\theta)}\int_t^\infty s^{-\theta_1}K(f,s)\frac{ds}{s}\right]^q \frac{dt}{t}\right)^{1/q}\right\}.\end{aligned}$$

An application of Hardy's inequalities (Lemma III.3.9) and an appeal to (2.24) now shows that each of the terms on the right is majorized by

$$c\left\{\int_0^\infty [t^{-\theta'}K(f,t)]^q \frac{dt}{t}\right\}^{1/q},$$

thus establishing (2.26) for finite q. The proof for $q = \infty$ is similar.

This completes the proof in the case $0 < \theta_0 < \theta_1 < 1$. Essentially the same argument works also when $0 = \theta_0 < \theta_1 < 1$, except that (2.17) is used in place of (2.16), and (2.19) in place of (2.4). Similarly, when $0 < \theta_0 < \theta_1 = 1$, we appeal to (2.18) and (2.20), respectively. The remaining case when $\theta_0 = 0$ and $\theta_1 = 1$ follows easily from Theorem 1.5. ■

Via the reiteration theorem, the basic interpolation theorem (Theorem 1.12) can be generalized as follows:

Theorem 2.5. *Let (X_0, X_1) and (Y_0, Y_1) be compatible couples and suppose $0 \le \theta_0 < \theta_1 \le 1, 0 \le \psi_0, \psi_1 \le 1$ with $\psi_0 \ne \psi_1$. Let $\bar{X}_{\theta_j}$ and $\bar{Y}_{\psi_j}$ be intermediate spaces of (X_0, X_1) and (Y_0, Y_1) of class θ_j and ψ_j, respectively, $(j = 0, 1)$. Let T be a linear operator satisfying*

$$\|Tf\|_{\bar{Y}_{\psi_j}} \le M_j\|f\|_{\bar{X}_{\theta_j}}, \qquad (j = 0, 1). \tag{2.27}$$

If $0 < \theta < 1$ and $1 \le q \le \infty$, then

$$\|Tf\|_{(Y_0,Y_1)_{\psi',q}} \le cM_0^{1-\theta}M_1^\theta\|f\|_{(X_0,X_1)_{\theta',q}}, \tag{2.28}$$

where

$$(\theta', \psi') = (1-\theta)(\theta_0, \psi_0) + \theta(\theta_1, \psi_1). \tag{2.29}$$

Proof. The hypotheses (2.27) together with Theorem 1.11 give

$$K(Tf, t; \bar{Y}_{\psi_0}, \bar{Y}_{\psi_1}) \le M_0K\left(f, \frac{M_1}{M_0}t; \bar{X}_{\theta_0}, \bar{X}_{\theta_1}\right). \tag{2.30}$$

Hence, exactly as in Theorem 1.12, we obtain

$$\|Tf\|_{(\bar{Y}_{\psi_0}, \bar{Y}_{\psi_1})_{\theta,q}} \le M_0^{1-\theta}M_1^\theta\|f\|_{(\bar{X}_{\theta_0}, \bar{X}_{\theta_1})_{\theta,q}}.$$

The desired result (2.28) follows immediately from this one by using (2.23) and (2.29) to identify the spaces involved. ■

Remark 2.6. The Calderón operator S_σ for the interpolation segment

$$\sigma = [(1-\theta_0, 1-\psi_0), (1-\theta_1, 1-\psi_1)]$$

(cf. Definition III.5.1) is, as might be expected, lurking in the background here. Indeed, since $\bar{Y}_{\psi_j} \hookrightarrow (Y_0, Y_1)_{\psi_j,\infty}$ for $j = 0, 1$, the K-functional on the left of (2.30) may be estimated from below (via (2.4)) in terms of the K-functional for the pair (Y_0, Y_1). Similarly, if $0 < \theta_j < 1$, the embedding $(X_0, X_1)_{\theta_j,1} \hookrightarrow \bar{X}_{\theta_j}$, $(j = 0, 1)$, gives rise to an estimate of the K-functional on the right of (2.30) from above in terms of the K-functional for the pair (X_0, X_1). The resulting inequality is

$$\frac{K(Tf, t; Y_0, Y_1)}{t} \le c\int_0^\infty \min_{j=0,1}\left\{M_j\frac{s^{1-\theta_j}}{t^{1-\psi_j}}\right\}\frac{K(f, s; X_0, X_1)}{s}\frac{ds}{s}. \tag{2.31}$$

The right-hand side is simply the S_σ-operator applied to $K(f, s; X_0, X_1)/s$ (cf. III.(5.4)). Multiplying each side by $t^{1-\psi'}$, taking the $L^q(dt/t)$-norm, then using Hardy's inequalities to simplify the right-hand side, we obtain (2.28) as before. This is exactly the procedure used previously to establish the Marcinkiewicz interpolation theorem (Theorem IV.4.13). It illustrates once again how closely the abstract K-method is modeled on the classical Marcinkiewicz theory.

Further structure theorems for the (θ, q)-spaces can be obtained with the aid of the J-functional $J(f,t) = J(f,t;X_0,X_1)$. The estimates

$$\min(1,t)\|f\|_{X_0\cap X_1} \le J(f,t) \le \max(1,t)\|f\|_{X_0\cap X_1} \tag{2.32}$$

and

$$K(f,t) \le \min(\|f\|_{X_0}, t\|f\|_{X_1}) \le \min\left(1,\frac{t}{s}\right)J(f,s), \tag{2.33}$$

valid for all $s,t>0$ and all f in $X_0\cap X_1$, follow at once from Definition 1.1.

In the *J-method* of interpolation, the interpolation spaces consist of certain elements f of X_0+X_1 that are attainable from $X_0\cap X_1$ by means of a Bochner-integral representation

$$f = \int_0^\infty u_s\frac{ds}{s}, \tag{2.34}$$

where u_s belongs to $X_0\cap X_1$ for each s, the function $s\to u_s$ is locally $(X_0\cap X_1)$-integrable on $(0,\infty)$, and the Bochner integral (2.34) is understood to converge in the norm of X_0+X_1. The precise definition is as follows:

***Definition* 2.7.** Let (X_0,X_1) be a compatible couple and suppose $0<\theta<1, 1\le q\le\infty$. The space $(X_0,X_1)_{\theta,q;J}$ consists of all f in X_0+X_1 that are representable in the form (2.34) and for which the functional

$$\|f\|_{\theta,q;J} = \inf\begin{cases}\left\{\int_0^\infty [s^{-\theta}J(u_s,s)]^q\frac{ds}{s}\right\}^{1/q}, & (q<\infty),\\ \sup_{s>0} s^{-\theta}J(u_s,s), & (q=\infty),\end{cases} \tag{2.35}$$

is finite (the infimum extending over all representations (2.34) of f).

The next result shows that the (θ,q)-spaces generated by the J- and K-methods are the same.

Theorem 2.8. (Equivalence theorem). *Let (X_0,X_1) be a compatible couple and suppose $0<\theta<1, 1\le q\le\infty$. Then there are constants $c_1,c_2>0$ such that, for all f in X_0+X_1,*

$$c_1\|f\|_{\theta,q} \le \|f\|_{\theta,q;J} \le c_2\|f\|_{\theta,q}. \tag{2.36}$$

In particular,

$$(X_0,X_1)_{\theta,q;J} = (X_0,X_1)_{\theta,q}, \tag{2.37}$$

with equivalent norms.

Proof. Suppose first that f belongs to $(X_0,X_1)_{\theta,q;J}$. Then for any representation of the form (2.34), we may use the subadditivity of the K-functional and the second of the estimates in (2.33) to obtain

$$K(f,t) \le \int_0^\infty K(u_s,t)\frac{ds}{s} \le \int_0^\infty J(u_s,s)\min\left(1,\frac{t}{s}\right)\frac{ds}{s}. \tag{2.38}$$

The rest is routine. Multiply both sides by $t^{-\theta}$, take $L^q(dt/t)$-norms, and use Hardy's inequalities to reduce the right-hand side. Passing to the infimum over all representations (2.34) of f gives the first of the inequalities in (2.36).

For the remaining inequality in (2.36), we may assume that $0<\|f\|_{\theta,q}<\infty$, otherwise there is nothing to prove. In that case,

$$\lim_{t\to 0} K(f,t) = 0 = \lim_{t\to\infty} t^{-1}K(f,t). \tag{2.39}$$

For each integer j, represent f in the form $f = g_j + h_j$ with

$$\|g_j\|_{X_0} + 2^j\|h_j\|_{X_1} < 2K(f,2^j). \tag{2.40}$$

Then the element $a_j = g_{j+1} - g_j = h_j - h_{j+1}$ belongs to $X_0\cap X_1$ and

$$\begin{aligned}\left\|f - \sum_{-M}^{N-1} a_j\right\|_{X_0+X_1} &\le \|f-g_N\|_{X_1} + \|g_{-M}\|_{X_0}\\ &\le 2[2^{-N}K(f,2^N) + K(f,2^{-M})],\end{aligned} \tag{2.41}$$

which, by virtue of (2.39), tends to 0 as M and N tend to ∞. Hence, the function

$$u_s = (\log 2)^{-1}\sum_{-\infty}^{\infty} a_j\chi_{(2^j,2^{j+1})}(s) \tag{2.42}$$

represents f in the form (2.34). If $2^j<s<2^{j+1}$, then (2.40) and elementary properties of the K-functional (cf. Proposition 1.2) give

$$\begin{aligned}J(u_s,s) &\le (\log 2)^{-1}\max(\|a_j\|_{X_0}, 2^{j+1}\|a_j\|_{X_1})\\ &\le 4(\|g_{j+1}-g_j\|_{X_0} + 2^j\|h_j - h_{j+1}\|_{X_1})\\ &\le 8[K(f,2^{j+1}) + K(f,2^j)]\\ &\le 24K(f,2^j) \le 24K(f,s).\end{aligned} \tag{2.43}$$

The second of the inequalities in (2.36) now follows at once. ■

The representation constructed in the preceding proof can also be used to establish the density of $X_0\cap X_1$ in certain of the (θ,q)-spaces as follows:

Theorem 2.9 (Density theorem). *Let (X_0,X_1) be a compatible couple and suppose $0<\theta<1, 1\le q<\infty$. Then $X_0\cap X_1$ is dense in $(X_0,X_1)_{\theta,q}$.*

Proof. In view of the equivalence theorem, it will suffice to show that $X_0 \cap X_1$ is dense in $(X_0, X_1)_{\theta,q;J}$. For an arbitrary element f in the latter space, let u_s be the representation defined by (2.42) and, for each integer n, let

$$u_s^{(n)} = u_s \chi_{(2^{-n}, 2^n)}(s).$$

Then the elements $f^{(n)}$ defined by

$$f^{(n)} = \int_0^\infty u_s^{(n)} \frac{ds}{s} = \sum_{-n+1}^{n} a_j$$

belong to $X_0 \cap X_1$ and, by (2.41), satisfy

$$\lim_{n\to\infty} \|f - f^{(n)}\|_{X_0 + X_1} = 0.$$

Since q is finite, we have

$$\|f - f^{(n)}\|_{\theta,q;J} \le \left\{ \int_0^\infty [s^{-\theta} J(u_s - u_s^{(n)}, s)]^q \frac{ds}{s} \right\}^{1/q}. \tag{2.44}$$

But $J(u_s - u_s^{(n)}, s)$ is equal to 0 if $2^{-n} < s < 2^n$ and equal to $J(u_s, s)$ otherwise. In particular, for each fixed s, the functional $J(u_s - u_s^{(n)}, s)$ tends to 0 as $n \to \infty$ and is majorized for all n by $J(u_s, s)$. Hence, by the dominated convergence theorem, the right-hand side of (2.44) tends to 0 as n tends to infinity. This completes the proof. ■

Recall from Definition 2.2 that an intermediate space X is of class θ if it is continuously embedded between the $(\theta, 1)$ and (θ, ∞) interpolation spaces. The latter embedding is clearly equivalent to the requirement that

$$K(f, t; X_0, X_1) \le ct^\theta \|f\|_X, \qquad (t > 0). \tag{2.45}$$

The former embedding has the following equivalent formulation.

Proposition 2.10. *Let (X_0, X_1) be a compatible couple and suppose $0 < \theta < 1$. Then the estimate*

$$\|f\|_X \le c\|f\|_{\theta,1} \tag{2.46}$$

holds for some constant c and all f in $(X_0, X_1)_{\theta,1}$ if and only if

$$\|f\|_X \le c\|f\|_{X_0}^{1-\theta} \|f\|_{X_1}^{\theta} \tag{2.47}$$

holds for some constant c and all f in $X_0 \cap X_1$.

Proof. Suppose first that (2.47) holds. Then for any u in $X_0 \cap X_1$ and any $s > 0$,

$$\|u\|_X \le cs^{-\theta} \|u\|_{X_0}^{1-\theta} (s\|u\|_{X_1})^\theta \le cs^{-\theta} J(u, s). \tag{2.48}$$

If $f \in (X_0, X_1)_{\theta,1}$, then for each $\varepsilon > 0$ we may represent f as in (2.34) with

$$\int_0^\infty s^{-\theta} J(u_s, s) \frac{ds}{s} \le \|f\|_{\theta,1} + \varepsilon.$$

Together with (2.48), this gives

$$\|f\|_X \le \int_0^\infty \|u_s\|_X \frac{ds}{s} \le c \int_0^\infty s^{-\theta} J(u_s, s) \frac{ds}{s} \le c(\|f\|_{\theta,1} + \varepsilon),$$

and hence, since $\varepsilon > 0$ is arbitrary, establishes (2.46).

Conversely, suppose (2.46) holds. For each f in $X_0 \cap X_1$ and each $s > 0$ the estimate (2.33) gives

$$\begin{aligned} \|f\|_{\theta,1} &\le \left\{ \int_0^\infty t^{-\theta} \min\left(1, \frac{t}{s}\right) \frac{dt}{t} \right\} J(f, s) \\ &= [\theta(1-\theta)s^\theta]J(f, s). \end{aligned} \tag{2.49}$$

Moreover, if $s = \|f\|_{X_0}/\|f\|_{X_1}$, then

$$s^{-\theta} J(f, s) = \|f\|_{X_0}^{1-\theta} \|f\|_{X_1}^{\theta}.$$

Hence, with this choice of s, (2.49) and (2.46) combine to give the desired estimate (2.47). ■

We conclude this section with the following variant of the reiteration theorem which allows the "patching" together of two interpolation scales.

Theorem 2.11 (T. H. Wolff). *Let X_2 and X_3 be intermediate spaces of a compatible couple (X_1, X_4). Let $0 < \phi, \psi < 1$ and $1 \le q, r \le \infty$ and suppose that*

$$X_2 = (X_1, X_3)_{\phi,q}, \qquad X_3 = (X_2, X_4)_{\psi,r}. \tag{2.50}$$

Then (up to equivalence of norms)

$$X_2 = (X_1, X_4)_{\rho,q}, \qquad X_3 = (X_1, X_4)_{\theta,r}, \tag{2.51}$$

where

$$\rho = \frac{\phi\psi}{1 - \phi + \phi\psi}, \qquad \theta = \frac{\psi}{1 - \phi + \phi\psi}. \tag{2.52}$$

Proof. We establish only the first of the identities in (2.51) since the proof of the second is similar. By the reiteration theorem (Theorem 2.4), it will suffice to show that X_3 is of class θ for the couple (X_1, X_4). Thus (cf. (2.45) and (2.47)), we have only to show that

$$K(f,t) = K(f,t; X_1, X_4) \le ct^{\theta}\|f\|_{X_3}, \qquad (f \in X_3) \tag{2.53}$$

and

$$\|f\|_{X_3} \le c\|f\|_{X_1}^{1-\theta}\|f\|_{X_4}^{\theta}, \qquad (f \in X_1 \cap X_4). \tag{2.54}$$

To establish (2.54), note first from the hypotheses (2.50) and Proposition 2.10 that

$$\|f\|_{X_2} \le c\|f\|_{X_1}^{1-\phi}\|f\|_{X_3}^{\phi}, \qquad \|f\|_{X_3} \le c\|f\|_{X_2}^{1-\psi}\|f\|_{X_4}^{\psi},$$

valid for f in $X_1 \cap X_3$ and $X_2 \cap X_4$, respectively. In particular, each is valid for all f in $X_1 \cap X_4$. Hence, substituting the first of these inequalities into the second, we obtain

$$\|f\|_{X_3} \le c\|f\|_{X_1}^{(1-\phi)(1-\psi)}\|f\|_{X_3}^{\phi(1-\psi)}\|f\|_{X_4}^{\psi},$$

which is equivalent to (2.54).

To establish (2.53), let $f \in X_3$ and fix $t > 0$. We claim that there is a decomposition of f that satisfies

(i) $$f = g_0 + h_0 + f_1, \qquad g_0 \in X_1, h_0 \in X_4, f_1 \in X_3;$$

(ii) $$\|f_1\|_{X_3} \le \frac{1}{2}\|f\|_{X_3}; \tag{2.55}$$

(iii) $$\|g_0\|_{X_1} + t\|h_0\|_{X_4} \le ct^{\theta}\|f\|_{X_3}.$$

Once this is established, it may be applied inductively to generate sequences $(g_j) \subset X_1, (h_j) \subset X_4$, and $(f_j) \subset X_3$ that satisfy

$$f_j = g_j + h_j + f_{j+1};$$

$$\|f_{j+1}\|_{X_3} \le \frac{1}{2}\|f_j\|_{X_3} \le 2^{-j}\|f\|_{X_3};$$

$$\|g_j\|_{X_1} + t\|h_j\|_{X_4} \le ct^{\theta}\|f_j\|_{X_3} \le c2^{-j}t^{\theta}\|f\|_{X_3}.$$

In particular, for all n, we shall have

$$f = \sum_{j=0}^{n} g_j + \sum_{j=0}^{n} h_j + f_{n+1}, \qquad \lim_{n\to\infty} \|f_{n+1}\|_{X_3} = 0, \tag{2.56}$$

and

$$\sum_{j=0}^{n} \|g_j\|_{X_1} + t\sum_{j=0}^{n} \|h_j\|_{X_4} \le ct^{\theta}\|f\|_{X_3}. \tag{2.57}$$

Since X_1 and X_4 are complete, the estimate (2.57) together with the Riesz-Fischer property shows that $\sum g_j$ converges in X_1 (to an element g, say), that $\sum h_j$ converges in X_4 (to an element h, say), and

$$\|g\|_{X_1} \le \sum_j \|g_j\|_{X_1}, \qquad \|h\|_{X_4} \le \sum_j \|h_j\|_{X_4}. \tag{2.58}$$

In particular, we have $g = \sum g_j$ and $h = \sum h_j$ in $X_1 + X_4$. It follows from (2.56) that $f = \sum(g_j + h_j)$ in X_3, hence also in $X_1 + X_4$, and so we conclude that $f = g + h$ in $X_1 + X_4$. Therefore, by (2.57) and (2.58),

$$K(f,t; X_1, X_4) \le \|g\|_{X_1} + t\|h\|_{X_4} \le ct^{\theta}\|f\|_{X_3},$$

which establishes the desired result (2.53).

It remains only to verify the inductive step (2.55). Since X_3 is of class ψ for (X_2, X_4), there is, for each $u > 0$, a decomposition $f = g' + h_0$ with

$$\|g'\|_{X_2} + u\|h_0\|_{X_4} \le c_1u^{\psi}\|f\|_{X_3}. \tag{2.59}$$

But $g' \in X_2$, which is of class ϕ for (X_1, X_3), so there is, for each $v > 0$, a decomposition $g' = g_0 + f_1$ with

$$\|g_0\|_{X_1} + v\|f_1\|_{X_3} \le c_2v^{\phi}\|g'\|_{X_2}. \tag{2.60}$$

Combining (2.59) and (2.60), we obtain

$$\|g_0\|_{X_1} \le c_1c_2u^{\psi}v^{\phi}\|f\|_{X_3}, \qquad t\|h_0\|_{X_4} \le c_1tu^{\psi-1}\|f\|_{X_3} \tag{2.61}$$

and

$$\|f_1\|_{X_3} \le c_1c_2v^{\phi-1}u^{\psi}\|f\|_{X_3}. \tag{2.62}$$

Choosing $v = (2c_1c_2u^{\psi})^{1/(1-\phi)}$, we see that (2.62) yields (2.55)(ii). With this choice of v and with $u = t^{(1-\theta)/(1-\psi)}$, addition of the two estimates in (2.61) gives (2.55)(iii). This completes the proof. ■

3. MONOTONE INTERPOLATION SPACES

Although no simple characterization is known of all of the interpolation spaces of an arbitrary couple (X_0, X_1), it is possible to describe a large class of interpolation spaces—the so-called *monotone* interpolation spaces—and this will be our goal in the present section. For some couples, the monotone interpolation spaces do exhaust all interpolation spaces, but this is not true in general.

An intermediate space X of (X_0, X_1) is said to be *monotone* if it has the property

$$K(g,t) \le K(f,t), \qquad f \in X \quad \Rightarrow \quad g \in X, \qquad \|g\|_X \le \|f\|_X, \tag{3.1}$$

where $K(f,t) = K(f,t; X_0, X_1)$ is the K-functional for the couple (X_0, X_1). Clearly, every monotone intermediate space is an interpolation space between X_0 and X_1.

In the special case of interpolation between L^1 and L^∞, the notion of monotonicity in (3.1) accords with that given previously in (1.47). For L^1 and L^∞, however, much more is true: *every* interpolation space is monotone (Theorem 1.17). The structure of monotone interpolation spaces between X_0 and X_1 may therefore be regarded as a "lifting" to (X_0, X_1) of the entire interpolation structure of (L^1, L^∞).

The next result shows that monotone Riesz-Fischer norms ρ generate monotone interpolation spaces between X_0 and X_1.

Theorem 3.1. *Let (X_0, X_1) be a compatible couple and let ρ be a monotone Riesz-Fischer norm on $(\mathbf{R}^+, m)$. Then $(X_0, X_1)_\rho$ is a monotone interpolation space with respect to (X_0, X_1).*

Proof. The interpolation property was established in Theorem 1.19. It remains therefore only to establish monotonicity. Suppose $K(g,t) \le K(f,t)$ and $f \in X_\rho$. Since $f \in X_0^o + X_1$ (Definition 1.18), Proposition 1.15 shows that $K(f, 0+) = 0$. The hypothesis therefore gives $K(g, 0+) = 0$ and $k(g,\cdot) \prec k(f,\cdot)$. Using the monotonicity (1.47) of ρ, we therefore have $\|g\|_{X_\rho} \le \|f\|_{X_\rho}$, which establishes the monotonicity property (3.1) for X_ρ. ■

In general, not every interpolation space of a couple (X_0, X_1) can be realized in terms of the K-functional. Indeed, even X_0 and X_1 may not be describable in this way. Perhaps the simplest way to ensure that the endpoint spaces are describable by the K-functional is to insist that they coincide with their Gagliardo completions.

Definition 3.2. A compatible couple (X_0, X_1) is said to be a *Gagliardo couple* if $X_0 = X_0 + \infty X_1$ and $X_1 = X_1 + \infty X_0$.

Henceforth, we shall restrict our attention to Gagliardo couples (X_0, X_1). In that case, both X_0 and X_1 are interpolation spaces generated by suitable Riesz-Fischer norms. In fact, the following result holds.

Proposition 3.3. *Let (X_0, X_1) be a Gagliardo couple and let ρ and σ be the function norms in L^1 and L^∞, respectively*:

$$\rho(f) = \int_0^\infty f(t)\,dt, \qquad \sigma(f) = \sup_{t>0} f(t).$$

Then

$$X_0^o = (X_0, X_1)_\rho, \qquad X_1^o = (X_1, X_0)_\rho, \tag{3.2}$$

and

$$X_0 = (X_1, X_0)_\sigma, \qquad X_1 = (X_0, X_1)_\sigma. \tag{3.3}$$

Proof. Only the first of the identities in (3.2) and (3.3) need be established since the second in each case follows by interchanging the roles of X_0 and X_1.

Suppose first that $f \in X_0^o$. Proposition 1.15 shows that $K(f, 0+) = 0$. Then

$$\begin{aligned}\|f\|_{(X_0,X_1)_\rho} &= \rho(k(f)) = \int_0^\infty k(f,s)\,ds \\ &= \sup_{t>0} K(f,t) = \|f\|_{X_0+\infty X_1} = \|f\|_{X_0},\end{aligned}$$

the last identity because (X_0, X_1) is a Gagliardo couple. Conversely, if $f \in (X_0, X_1)_\rho$, then, by definition, $f \in X_0^o + X_1$ and the preceding identities show that $f \in X_0 + \infty X_1 = X_0$. Hence, $f \in (X_0^o + X_1) \cap X_0 = X_0^o$. This establishes (3.2).

To establish (3.3), note first that the obvious identity

$$K(f,t; X_1, X_0) = tK(f, t^{-1}; X_0, X_1),$$

upon differentiation, leads to the relation

$$k(f,t; X_1, X_0) = K(f, t^{-1}; X_0, X_1) - t^{-1}k(f, t^{-1}; X_0, X_1).$$

Thus, if $f \in (X_1, X_0)_\sigma$ (hence also $f \in X_1^o + X_0$ by Definition 1.18), Proposition 1.15 shows that $K(f, 0+; X_1, X_0)$ vanishes and the first of the identities above reveals that $k(f,s; X_0, X_1) \downarrow 0$ as $s \to \infty$ (since $k(f,s) \le K(f,s)/s$). Hence, by the monotone convergence theorem,

$$\begin{aligned}\|f\|_{X_0} = \|f\|_{X_0+\infty X_1} &= \sup_{t>0} K(f,t; X_0, X_1) \\ &= \sup_{t>0}\left\{K(f,0+) + \int_0^t [k(f,s) - k(f,t)]\,ds\right\} \\ &= \sup_{t>0}\{K(f,t; X_0, X_1) - tk(f,t; X_0, X_1)\} \\ &= \sup_{t>0} k(f,t; X_1, X_0) = \|f\|_{(X_1,X_0)_\sigma}.\end{aligned}$$

These identities show also that if $f \in X_0 = X_0 + \infty X_1$, then $f \in (X_1, X_0)_\sigma$ with identical norms. This completes the proof. ■

We shall need the following lemma, which provides a useful decomposition of an element with respect to its K-functional.

Theorem 3.4 (M. Cwikel). *Let (X_0, X_1) be a Gagliardo couple and suppose $f \in X_0 + X_1$. Then there is a sequence $(a_j)_{j=-\infty}^{\infty}$ of elements of $X_0 + X_1$ such that $f = \sum_j a_j$ (convergence in $X_0 + X_1$) for which the estimates*

$$K(f,t) \le \sum_j \min(\|a_j\|_{X_0}, t\|a_j\|_{X_1}) \le 24K(f,t) \tag{3.4}$$

hold for all $t > 0$. Furthermore, the elements a_j can be chosen so that one of the following conditions holds:

(I) *$a_j \in X_0 \cap X_1$ for all j;*

(II) *there is an index Q such that $a_j = 0$ if $j < Q$ and $a_j \in X_0 \cap X_1$ if $j > Q$; the element $a_Q \in X_0$;*

(III) *there is an index P such that $a_j = 0$ if $j > P$ and $a_j \in X_0 \cap X_1$ if $j < P$; the element $a_P \in X_1$;*

(IV) *there are indices P and Q with $P > Q$ such that $a_j \in X_0 \cap X_1$ if $Q < j < P$, $a_j = 0$ if $j < Q$ or $j > P$, and $a_Q \in X_0$, $a_P \in X_1$.*

(***Remark***: If the exceptional element a_Q in (II) does not belong to X_1, the quantity $\|a_Q\|_{X_1}$ in (3.4) is infinite; the Q-summand in the central term of (3.4) is thus equal to $\|a_Q\|_{X_0}$. A similar remark applies to the exceptional elements a_Q and a_P in (III) and (IV)).

Proof. Fix $f \in X_0 + X_1$ and, for convenience, denote $K(f,t)$ by $K(t)$. We begin by assuming that

$$\text{(a)}\quad \lim_{s\to 0} K(s) = 0, \qquad \text{(b)}\quad \lim_{s\to 0} \frac{K(s)}{s} = \infty, \tag{3.5}$$

and

$$\text{(a)}\quad \lim_{s\to \infty} K(s) = \infty, \qquad \text{(b)}\quad \lim_{s\to \infty} \frac{K(s)}{s} = 0. \tag{3.6}$$

Our first objective is to construct a bi-infinite sequence $(t_j)_{-\infty}^{\infty}$ in $(0,\infty)$ as follows. Set $t_0 = 1$. Then $t_1 < t_2 < \cdots$ are defined inductively by the requirement that $t_j > t_{j-1}$, $(j = 1, 2, \ldots)$, be the smallest number for which both

$$2K(t_{j-1}) \le K(t_j) \tag{3.7}$$

and

$$2\frac{K(t_j)}{t_j} \le \frac{K(t_{j-1})}{t_{j-1}} \tag{3.8}$$

hold. Their existence is guaranteed by the hypothesis (3.6), which in fact shows that $t_j \uparrow \infty$ as $j \uparrow \infty$. Observe also that, by continuity of the K-functional, equality must hold in one of (3.7) and (3.8). The sequence $t_{-1} > t_{-2} > \cdots > 0$ is generated in similar fashion from the requirements (3.7) and (3.8) for negative j. In this case, (3.5) guarantees that $t_j \downarrow 0$ as $j \downarrow -\infty$.

For each integer j, we may represent f in the form $f = g_j + h_j$ with $g_j \in X_0$ and $h_j \in X_1$ so that

$$\|g_j\|_{X_0} + t_j\|h_j\|_{X_1} \le 2K(t_j). \tag{3.9}$$

If

$$a_j = g_j - g_{j-1} = h_{j-1} - h_j, \tag{3.10}$$

then clearly $a_j \in X_0 \cap X_1$ for all j, as asserted in case (I) above. We shall use the notation

$$\Psi_j(t) = \min\{\|a_j\|_{X_0}, t\|a_j\|_{X_1}\}.$$

From (3.5)(a), (3.6)(b), and (3.9), we see that

$$\begin{aligned} \left\| f - \sum_{-M+1}^{N} a_j \right\|_{X_0+X_1} &= \|f - g_N + g_{-M}\|_{X_0+X_1} \\ &\le \|g_{-M}\|_{X_0} + \|f - g_N\|_{X_1} \\ &\le 2\left[K(t_{-M}) + \frac{K(t_N)}{t_N} \right], \end{aligned}$$

which tends to 0 as $M, N \to \infty$. This shows that $f = \sum_j a_j$ in $X_0 + X_1$ and establishes the left-hand inequality in (3.4):

$$K(f,t) \le \sum_j K(a_j, t) \le \sum_j \Psi_j(t).$$

Hence, it remains only to establish the right-hand inequality in (3.4). Fix $t > 0$ and let j^* be the index for which $t_{j^*-1} < t \le t_{j^*}$ holds. We split the range of summation $-\infty < j < \infty$ into three parts:

Case 1. $j \le j^* - 1$. Since $K(s)$ is increasing, we have from (3.9),

$$\Psi_j(t) \le \|a_j\|_{X_0} \le \|g_j\|_{X_0} + \|g_{j-1}\|_{X_0} \le 4K(t_j).$$

But repeated application of (3.7) gives

$$K(t_j) \le \frac{1}{2}K(t_{j+1}) \le 2^{j+1-j^*}K(t_{j^*-1}) \le 2^{j+1-j^*}K(t).$$

Hence,

$$\Psi_j(t) \le 8 \cdot 2^{j-j^*}K(t), \qquad (j \le j^* - 1). \tag{3.11}$$

Case 2. $j^* + 1 \le j$. This time we use (3.9), repeated application of (3.8), and the fact that $K(s)/s$ decreases, to see that

$$\begin{aligned}\Psi_j(t) &\le t\|a_j\|_{X_1} \le t(\|h_{j-1}\|_{X_1} + \|h_j\|_{X_1})\\ &\le 4t\frac{K(t_{j-1})}{t_{j-1}} \le 8t \cdot 2^{-2}\frac{K(t_{j-2})}{t_{j-2}}\\ &\le 8t \cdot 2^{j^*-j}\frac{K(t_{j^*})}{t_{j^*}} \le 8 \cdot 2^{j^*-j}K(t), \qquad (j \ge j^* + 1). \qquad (3.12)\end{aligned}$$

Case 3. $j = j^*$. We know that equality holds in at least one of (3.7) or (3.8). If there is equality in (3.7), then, as before,

$$\Psi_{j^*}(t) \le \|a_{j^*}\|_{X_0} \le 4K(t_{j^*}) = 8K(t_{j^*-1}) \le 8K(t).$$

If, on the other hand, equality holds in (3.8), then

$$\Psi_{j^*}(t) \le t\|a_{j^*}\|_{X_1} \le 4t\frac{K(t_{j^*-1})}{t_{j^*-1}} = 8t\frac{K(t_{j^*})}{t_{j^*}} \le 8K(t).$$

Hence, in either case,

$$\Psi_{j^*}(t) \le 8K(t). \qquad (3.13)$$

It follows therefore from (3.11), (3.12), and (3.13) that

$$\begin{aligned}\sum_{-\infty}^{\infty} \Psi_j(t) &\le 8\left\{\sum_{j<j^*} 2^{j-j^*} + \sum_{j>j^*} 2^{j^*-j} + 1\right\}K(t)\\ &= 24K(t), \qquad (3.14)\end{aligned}$$

as desired.

We conclude by indicating the modifications necessary if the assumptions (3.5) and (3.6) do not hold.

Suppose first that (3.5) holds but that (3.6)(a) fails (note that (3.6)(b) must hold in this case). Then $K(s)$ is bounded, and, since (X_0, X_1) is a Gagliardo couple, it follows that f belongs to X_0. The inductive procedure of selecting the t_j (as $j \to \infty$) then terminates at the J-th step, say, because it is impossible to find $t_{J+1} > t_J$ satisfying (3.7). Hence,

$$K(s) < 2K(t_J), \qquad (s > 0). \qquad (3.15)$$

We select g_j and h_j, $(j \le J)$, as before satisfying (3.9) but when $j = J$, in place of (3.10) we set

$$a_J = f - g_J = h_J. \qquad (3.16)$$

Note that a_J still belongs to $X_0 \cap X_1$.

Suppose now that $t \le t_J$ (so $j^* \le J$). We obtain (3.11) in case 1 exactly as before. Case 2 now covers the range $j^* + 1 \le j \le J$ and for these values of j we obtain (3.12) as before. Case 3 also proceeds as before unless $j^* = J$. Then $j = j^* = J$ so from (3.16) and (3.9) we have

$$\Psi_J(t) \le t\|h_J\|_{X_1} \le 2t\frac{K(t_J)}{t_J} \le 2K(t)$$

since $t \le t_J$ and $K(s)/s$ is decreasing. Hence, (3.13) persists in this case and we obtain the desired conclusion (3.14) as before.

For the remaining values of $t > t_J$, we argue as follows. Letting $s \to \infty$ in (3.15), we see that $\|f\|_{X_0} \le 2K(t_J)$ because X_0 is equal to its Gagliardo completion. Hence, using (3.16) and (3.9) we obtain

$$\Psi_J(t) \le \|a_J\|_{X_0} \le \|f\|_{X_0} + \|g_J\|_{X_0} \le 4K(t_J) \le 4K(t).$$

The terms $\Psi_j(t)$ for all other values of $j < J$ are estimated exactly as in case 1 above and from this we proceed directly to the conclusion (3.14).

Next suppose that (3.5) holds but (3.6)(b) fails (then (3.6)(a) must hold). Once again, the inductive procedure defining the t_j terminates at the P-th step, say, this time because it is impossible to select $t_{P+1} > t_P$ satisfying (3.8). Hence,

$$2\frac{K(s)}{s} > \frac{K(t_P)}{t_P}, \qquad (s > 0). \qquad (3.17)$$

There is no guarantee this time that f belongs to X_0. As before, we select g_j and h_j, $(j \le P)$, satisfying (3.9) but when $j = P$, in place of (3.10) we set

$$a_P = f - g_P = h_P. \qquad (3.18)$$

Note that $a_P \in X_1$, as in condition (III) of the statement of the theorem. We have from (3.17) and (3.18),

$$\Psi_P(t) \le t\|h_P\|_{X_1} \le 2t\frac{K(t_P)}{t_P} \le 4K(t), \qquad (t > 0).$$

The estimates for $\Psi_j(t)$, $(j < P)$, proceed much as before and we arrive at the conclusion (3.14) again in this case. We omit the details.

The modifications are similar if one of (3.5)(a) or (3.5)(b) fails (both cannot fail simultaneously) except that it is the series in negative j that terminates. Since in any case, one of (3.5)(a) or (b) must hold and one of (3.6)(a) or (b) must hold, it is now routine to verify that one of conditions (I), (II), (III), and (IV) above must be satisfied by the a_j. This completes the proof. ■

Definition 3.5. A compatible couple (X_0, X_1) is said to be *divisible* if there is a constant $c > 0$ such that, whenever $f \in X_0 + X_1$ and ω_k, $(k = 1, 2, \ldots)$, are nonnegative concave functions on $(0, \infty)$ with $\sum_k \omega_k(1) < \infty$ and

$$K(f, t) \le \sum_k \omega_k(t), \qquad (t > 0), \tag{3.19}$$

there exist elements $f_k \in X_0 + X_1$ for which $f = \sum_k f_k$ in $X_0 + X_1$ and

$$K(f_k, t) \le c\omega_k(t), \qquad (k = 1, 2, \ldots, t > 0). \tag{3.20}$$

Theorem 3.6 **(Ju.A. Brudnyi & N. Ja. Krugljak).** *Every Gagliardo couple is divisible.*

Proof. Let (X_0, X_1) be a Gagliardo couple and fix f in $X_0 + X_1$. Let ω_k, $k = 1, 2, \ldots$, be nonnegative concave functions on $(0, \infty)$ for which (3.19) holds. By induction, we shall construct $f_1, f_2, \ldots$ for which $f = \sum_k f_k$ and (3.20) holds.

Decompose f as a sum $f = \sum_j a_j$ as in Theorem 3.4 so that (3.4) holds. We shall assume for the moment that the elements a_j satisfy condition (I) of Theorem 3.4, that is, they all belong to $X_0 \cap X_1$. Then the function

$$\Psi(t) = c\sum_j \min\{\|a_j\|_{X_0}, t\|a_j\|_{X_1}\}$$

vanishes at the origin. We make the choice $c = 1/24$. Set

$$\psi(t) = \left(\frac{d\Psi}{dt}\right)(t) = c\sum_j \|a_j\|_{X_1} \chi_{(0, t_j)},$$

where $t_j = \|a_j\|_{X_0}/\|a_j\|_{X_1}$. Then from (1.7) and (1.8) we have, for all j,

$$\begin{aligned} K(a_j, t) &\le \min\{\|a_j\|_{X_0}, t\|a_j\|_{X_1}\} \\ &= \int_0^t \|a_j\|_{X_1} \chi_{(0, t_j)}(s)\, ds. \end{aligned} \tag{3.21}$$

Hence, if $v_k = d\omega_k/dt$, then using (3.4) we may express the hypothesis (3.19) in the form

$$\int_0^t \psi(s)\, ds \le \sum_k \omega_k(0) + \int_0^t [v_1(s) + \sum_{k \ge 2} v_k(s)]\, ds. \tag{3.22}$$

By Theorem III.7.7, there exist numbers $\theta_1(j)$ and $\theta'_1(j)$ in $[0, 1]$, with $\theta_1(j) + \theta'_1(j) = 1$, and with the property that if

$$\psi_1 = c\sum_j \theta_1(j)\|a_j\|_{X_1} \chi_{(0, t_j)}$$

and

$$\psi'_1 = c\sum_j \theta'_1(j)\|a_j\|_{X_1} \chi_{(0, t_j)},$$

then

$$\int_0^t \psi_1(s)\, ds \le \omega_1(0) + \int_0^t v_1(s)\, ds = \omega_1(t) \tag{3.23}$$

and

$$\int_0^t \psi'_1(s)\, ds \le \sum_{k \ge 2} \omega_k(0) + \int_0^t \sum_{k \ge 2} v_k(s)\, ds. \tag{3.24}$$

Hence, if we now define

$$f_1 = \sum_j \theta_1(j) a_j, \qquad f'_1 = \sum_j \theta'_1(j) a_j,$$

so that

$$f = f_1 + f'_1, \tag{3.25}$$

we see from (3.21) and (3.23) that

$$cK(f_1, t) \le \int_0^t \psi_1(s)\, ds \le \omega_1(t), \tag{3.26}$$

and from (3.21) and (3.24) that

$$cK(f'_1, t) \le \int_0^t \psi'_1(s)\, ds \le \sum_{k \ge 2} \omega_k(t). \tag{3.27}$$

The next step in the inductive procedure is to repeat the preceding argument with f'_1 in place of f, and with (3.24) in place of (3.22). This yields a decomposition $f'_1 = f_2 + f'_2$, where f_2 and f'_2 have properties analogous to (3.26) and (3.27). Proceeding in this way, we obtain at the n-th stage

$$f'_{n-1} = f_n + f'_n, \tag{3.28}$$

with

$$cK(f_n, t) \le \int_0^t \psi_n(s)\, ds \le \omega_n(t) \tag{3.29}$$

and

$$cK(f'_n, t) \le \int_0^t \psi'_n(s)\, ds \le \sum_{k \ge n+1} \omega_k(t). \tag{3.30}$$

The identity

$$f = \sum_{k=1}^{n} f_k + f'_n \tag{3.31}$$

follows at once from (3.25) and (3.28) so, using (3.30), we have

$$c\left\| f - \sum_{k=1}^{n} f_k \right\|_{X_0+X_1} = cK(f'_n, 1) \leq \sum_{k \geq n+1} \omega_k(1),$$

which tends to 0 as $n \to \infty$. Hence, $f = \sum_k f_k$ in $X_0 + X_1$ and (3.29) shows that the elements f_k satisfy the desired estimate (3.20) with a constant of 24. This completes the proof in the case where the elements a_j satisfy condition (I) of Theorem 3.4.

Suppose now that the a_j satisfy condition (II) of Theorem 3.4. We may assume that the element a_Q in X_0 does not belong to X_1, and in this case its norm in X_1 is infinite. The function Ψ can be defined exactly as before but will no longer vanish at the origin (in fact, $\Psi(0) = c\|a_Q\|_{X_0}$). Its derivative ψ is given by the same expression as before, the sum extending over all $j > Q$. The estimate (3.21) remains valid for $j > Q$ but when $j = Q$ it is replaced by $K(a_Q, t) \leq \|a_Q\|_{X_0}$. The constant $\sum \omega_k(0)$ in (3.22) is now replaced by $C = \sum \omega_k(0) - \Psi(0)$, and in order to apply Theorem III.7.7, we have to determine how to represent C as a sum $C = C_1 + C_2$ of positive numbers C_1 and C_2. We do this as follows. For each $k = 1, 2, \ldots$, let

$$\theta_k(Q) = \frac{\omega_k(0)}{\sum_j \omega_j(0)}.$$

Then $\omega_k(0) - \theta_k(Q)\Psi(0) \geq 0$ for all k. Define C_1 and C_2 by

$$C_1 = \omega_1(0) - \theta_1(Q)\Psi(0), \qquad C_2 = \sum_{k \geq 2} [\omega_k(0) - \theta_k(Q)\Psi(0)]$$

and apply Theorem III.7.7. The functions ψ_1 and ψ'_1 have exactly the same form as before, the sum in each case extending over all $j > Q$. The estimates (3.23) and (3.24) contain an extra constant term $-\theta_1(Q)\Psi(0)$ and $-\sum_{k \geq 2} \theta_k(Q)\Psi(0)$, respectively, in the right-hand side. However, if f_1 and f'_1 are defined as before, with sums extending over all $j \geq Q$, then the conclusions of (3.25), (3.26), and (3.27) are valid as stated. The remainder of the induction now follows exactly the same lines as before.

Similar modifications are necessary when the a_j satisfy conditions (III) or (IV) of Theorem 3.4. The details are straightforward and we omit them. ∎

The proof of the preceding theorem is based on the Lorentz-Shimogaki

theorem (Theorem III.7.7). In fact, since the K-functional for the couple (L^1, L^∞) is $\int_0^t f^*(s)\,ds$, one sees that the Lorentz-Shimogaki theorem itself may be regarded as a rudimentary form of divisibility for the couple (L^1, L^∞).

Divisibility will be crucial in establishing the following description of monotone interpolation spaces.

Theorem 3.7. *Let X be an intermediate space of a Gagliardo couple (X_0, X_1). Then X is a monotone interpolation space of the couple (X_0, X_1) if and only one of the following holds:*

(i) *there is a monotone Riesz-Fischer norm ρ such that $X = (X_0, X_1)_\rho$ with equivalent norms;*

(ii) *there is a monotone Riesz-Fischer norm ρ such that $X = (X_1, X_0)_\rho$ with equivalent norms;*

(iii) $X = X_0 + X_1$.

Proof. That monotone Riesz-Fischer norms generate monotone interpolation spaces was established in Theorem 3.1, and it is evident that $X_0 + X_1$ is always a monotone interpolation space. It remains therefore only to establish the converse.

Let X be a monotone intermediate space of (X_0, X_1). We may assume that X contains a nonzero element, for otherwise we must have $X = X_0 \cap X_1 = \{0\}$ and then $X = (X_0, X_1)_\rho$ with ρ the norm in $L^1 \cap L^\infty$. We suppose first that X is contained in $X_0^o + X_1$ and we shall show that there is a monotone Riesz-Fischer norm ρ such that $X = (X_0, X_1)_\rho$ with equivalent norms; that is, there exist constants $c_1, c_2 > 0$ such that

$$c_1\|f\|_X \leq \rho(k(f, \cdot)) \leq c_2\|f\|_X, \qquad (f \in X), \tag{3.32}$$

where $k(f) = k(f, \cdot) = k(f, \cdot; X_0, X_1)$.

It will suffice to define ρ for decreasing functions F^*. For such functions, set

$$\rho(F^*) = \inf\left\{\sum_j \|f_j\|_X : F^* \prec \sum_j k(f_j)\right\}. \tag{3.33}$$

Each of the admissible functions f_j in (3.33) is in X and so, by hypothesis, is also in $X_0^o + X_1$. Proposition 1.15 shows therefore that $K(f_j, 0+) = 0$. Hence, if $F^* \prec \sum_j k(f_j)$, then using (1.6) we have

$$tF^{**}(t) \leq \max(1, t)\sum_j \|f_j\|_{X_0+X_1} \leq c\max(1, t)\sum_j \|f_j\|_X \tag{3.34}$$

since $X \subset X_0 + X_1$. Taking the infimum (with $t = 1$), we obtain

$$F^{**}(1) \leq c\rho(F^*), \tag{3.35}$$

which shows that $\rho(F^*) = 0$ if and only if $F = 0$ a.e. The homogeneity of ρ is obvious. The triangle inequality is proved in the same way as the Riesz-Fischer property below. Hence, ρ satisfies property (P1) of Definition I.1.1. Property (P2) is an immediate consequence of the definition (3.33) of ρ. Property (P4) follows easily from (3.33) and (1.6), and (P5) from (3.34) (with $t = |E|$). Thus, in order to show that ρ is a Riesz-Fischer norm, it remains only to establish the Riesz-Fischer property (1.45) for ρ.

To prove $\rho(\sum F_m) \leq \sum \rho(F_m)$, it will suffice to show that $\rho(\sum F_m^*) \leq \sum \rho(F_m^*)$ because $\rho(\sum F_m) = \rho((\sum F_m)^*) \leq \rho(\sum F_m^*)$ by monotonicity and the fact that $(\sum F_m)^* \prec \sum F_m^*$. Thus, let $F^* = \sum_m F_m^*$. We may assume that $\rho(F_m^*) < \infty$ for all m since otherwise there is nothing to prove. In that case, given $\varepsilon > 0$, there is, for each m, a sequence $(f_j^{(m)})$ in X such that

$$F_m^* \prec \sum_j k(f_j^{(m)}), \qquad \sum_j \|f_j^{(m)}\|_X < \rho(F_m^*) + 2^{-m}\varepsilon.$$

Hence,

$$F^* \prec \sum_m \sum_j k(f_j^{(m)}), \qquad \sum_m \sum_j \|f_j^{(m)}\|_X < \sum_m \rho(F_m^*) + \varepsilon,$$

which shows that

$$\rho(F^*) \leq \sum_m \rho(F_m^*) + \varepsilon.$$

Letting $\varepsilon \to 0$, we obtain the Riesz-Fischer estimate (1.45). The monotonicity property (1.47) for ρ follows at once from (3.33). Hence, ρ is a monotone Riesz-Fischer norm.

We show next that $X = (X_0, X_1)_\rho$ and that (3.32) holds. Suppose $f \in (X_0, X_1)_\rho$. Then $K(f, 0+) = 0$ and $\rho(k(f)) < \infty$. For given $\varepsilon > 0$, there exists a sequence (f_j) in X (so necessarily $K(f_j, 0+) = 0$) such that

$$k(f) \prec \sum_j k(f_j) \tag{3.36}$$

and

$$\sum_j \|f_j\|_X < \rho(k(f)) + \varepsilon. \tag{3.37}$$

Theorem 3.4 provides a decomposition $f = \sum_j g_j$ of f with

$$K(g_j) \leq cK(f_j) = K(cf_j)$$

Then $K(g_j, 0+) = 0$ and

$$k(g_j) \prec ck(f_j) = k(cf_j).$$

The monotonicity of X therefore gives

$$\|g_j\|_X \leq c\|f_j\|_X$$

for all j. Hence, summing over j, we have

$$\|f\|_X \leq \sum_j \|g_j\|_X \leq c\sum_j \|f_j\|_X. \tag{3.38}$$

Combining (3.37) with (3.38), we see that f belongs to X and

$$\|f\|_X \leq c(\|f\|_{(X_0,X_1)_\rho} + \varepsilon).$$

Letting $\varepsilon \to 0$, we obtain the left-hand estimate in (3.32). The right-hand estimate follows directly from (3.33) since if $f \in X$, then obviously $\rho(k(f))$ is no larger than $\|f\|_X$. Hence, we have shown that case (i) holds.

By symmetry, if the monotone interpolation space X is contained in $X_1^\circ + X_0$, then case (ii) holds. In the remaining case, the monotone interpolation space X contains an element f in $X_0 \setminus (X_0^\circ + X_1)$ and an element g in $X_1 \setminus (X_1^\circ + X_0)$. By Proposition 1.15, this can occur of course only if $X_0 \cap X_1$ is dense in neither X_0 nor X_1. Proposition 1.15 shows also that

$$K(f, 0+; X_0, X_1) > 0 \qquad \text{and} \qquad K(g, 0+; X_1, X_0) > 0. \tag{3.39}$$

Let F be any element of $X_0 = X_0 + \infty X_1$. Since its K-functional $K(F, t)$ is bounded as a function of $t > 0$, it follows from the first relation in (3.39) that $K(F, t) \leq cK(f, t) = K(cf, t)$, for all $t > 0$, for a suitable constant c. But X is monotone and $cf \in X$ so we conclude from (3.1) that F also belongs to X. This shows that X contains every element of X_0. A similar argument based on the second inequality in (3.39) shows that X contains every element of X_1. Hence, $X = X_0 + X_1$ and the proof is complete. ■

4. BESOV AND SOBOLEV SPACES

Membership in Besov and Sobolev spaces is determined by the "smoothness" of the functions concerned; the Besov norms involve *differences*, the Sobolev norms use *derivatives*. We begin with the Besov spaces.

Definition 4.1. For each $h \in \mathbf{R}^n$, the *translation operator* T_h is defined on functions f on $\mathbf{R}^n$ by $T_h f(x) = f(x + h), (x \in \mathbf{R}^n)$. The *first-difference* operator $\Delta_h \equiv \Delta_h^1$ is defined by

$$\Delta_h f(x) = (T_h - I)f(x) = f(x + h) - f(x),$$

and higher-order differences are defined inductively by

$$\Delta_h^{r+1} f(x) = \Delta_h(\Delta_h^r f)(x), \qquad (r = 1, 2, \ldots).$$

Since $T_h^k f(x) = T_{kh} f(x) = f(x + kh)$, it is clear that

$$\Delta_h^r f(x) = (T_h - I)^r f(x) = \sum_{k=0}^{r} \binom{r}{k} (-1)^{r-k} f(x + kh). \tag{4.1}$$

Definition 4.2. The space of bounded continuous functions on $\mathbf{R}^n$ is denoted by $C = C(\mathbf{R}^n)$. It is a Banach space under the supremum norm $f \to \|f\|_\infty = \sup\{|f(x)| : x \in \mathbf{R}^n\}$.

The *r-th order modulus of continuity* of a function f in $L^p(\mathbf{R}^n)$, $1 \leq p < \infty$, is defined by

$$\omega_r(f, t)_p = \sup_{|h| \leq t} \|\Delta_h^r f\|_p, \qquad (f \in L^p, t > 0). \tag{4.2}$$

When $p = \infty$, the L^p-norm is replaced by the norm in C:

$$\omega_r(f, t)_\infty = \sup_{|h| \leq t} \|\Delta_h^r f\|_\infty, \qquad (f \in C, t > 0). \tag{4.3}$$

Each modulus $\omega_r(f, t)_p$, $(1 \leq p \leq \infty)$, is a nonnegative increasing function of $t > 0$; furthermore, for each fixed t, $\omega_r(\cdot, t)_p$ is a seminorm on L^p or C. It follows from (4.1) that

$$\omega_r(f, t)_p \leq 2^r \|f\|_p. \tag{4.4}$$

Since $\Delta_{2h} = T_h^2 - I = (T_h + I)\Delta_h$, one sees also that

$$\omega_r(f, 2t)_p \leq 2^r \omega_r(f, t)_p. \tag{4.5}$$

Definition 4.3. Suppose $\alpha > 0$ and $1 \leq p, s \leq \infty$. Let r be a positive integer with $r > \alpha$. The *Besov space* $B_{\alpha,s}^p$ consists of those f in L^p (if $p < \infty$) or C (if $p = \infty$) for which the norm

$$\|f\|_{B_{\alpha,s}^p} = \begin{cases} \|f\|_p + \left\{ \int_0^\infty [t^{-\alpha} \omega_r(f, t)_p]^s \dfrac{dt}{t} \right\}^{1/s}, & (s < \infty), \\ \|f\|_p + \sup\limits_{0 < t < \infty} \{t^{-\alpha} \omega_r(f, t)_p\}, & (s = \infty), \end{cases} \tag{4.6}$$

is finite.

It follows from (4.4) that an equivalent norm results on $B_{\alpha,s}^p$ if the range of t in (4.6) is replaced by $0 < t < 1$. Equivalent norms result also from different choices of integers $r > \alpha$. This will emerge as a corollary of the following estimates, which compare moduli of continuity of different orders.

Theorem 4.4 (A. Marchaud). *If k and r are integers satisfying $0 < k < r$,*

then, for all $t > 0$,

$$2^{k-r} \omega_r(f, t)_p \leq \omega_k(f, t)_p \leq ct^k \int_t^\infty \frac{\omega_r(f, u)_p}{u^k} \frac{du}{u}. \tag{4.7}$$

Proof. The left-hand inequality follows from the estimate

$$\|\Delta_h^r f\|_p = \|\Delta_h^{r-k} \Delta_h^k f\|_p \leq 2^{r-k} \omega_k(f, t)_p,$$

which is valid for $|h| \leq t$.

The right-hand inequality in (4.7) will require the identity

$$\Delta_h^k = 2^{-k} \left\{ \Delta_{2h}^k - \sum_{j=1}^{k} \binom{k}{j} \sum_{m=0}^{j-1} \Delta_h^{k+1} T_{mh} \right\}. \tag{4.8}$$

To prove (4.8), recall that $T_h^2 = T_{2h}$ and so

$$\Delta_{2h}^k = (T_h^2 - I)^k = (T_h + I)^k \Delta_h^k. \tag{4.9}$$

But

$$(T_h + I)^k - 2^k I = \sum_{j=1}^{k} \binom{k}{j} (T_{jh} - I) = \sum_{j=1}^{k} \binom{k}{j} \sum_{m=0}^{j-1} T_{mh} \Delta_h,$$

so substituting this into (4.9) we obtain

$$\Delta_{2h}^k = \sum_{j=1}^{k} \binom{k}{j} \sum_{m=0}^{j-1} \Delta_h^{k+1} T_{mh} + 2^k \Delta_h^k,$$

which is equivalent to the desired result (4.8).

Now apply (4.8) to a function f in the appropriate space (L^p or C) and take norms on both sides of the identity. Using the translation-invariance of the norms and writing $\omega_k(f, t)$ for $\omega_k(f, t)_p$, we have

$$\omega_k(f, t) \leq 2^{-k} \left\{ \omega_k(f, 2t) + \sum_{j=1}^{k} \binom{k}{j} j \omega_{k+1}(f, t) \right\}$$

or, equivalently,

$$\omega_k(f, t) \leq \frac{k}{2} \omega_{k+1}(f, t) + 2^{-k} \omega_k(f, 2t), \qquad (t > 0). \tag{4.10}$$

The right-hand term can be estimated again by the inequality (4.10) (with t replaced by $2t$) to give

$$\omega_k(f, t) \leq \frac{k}{2} \{\omega_{k+1}(f, t) + 2^{-k} \omega_{k+1}(f, 2t)\} + 2^{-2k} \omega_k(f, 4t).$$

Iteration of this procedure yields

$$\omega_k(f,t) \le \frac{k}{2}\sum_{j=0}^{N-1}(2^j)^{-k}\omega_{k+1}(f,2^j t) + 2^{-Nk}\omega_k(f,2^N t).$$

The last term tends to zero as $N \to \infty$, and ω_{k+1} is increasing. Hence, the sum can be replaced by an integral to give

$$\omega_k(f,t) \le \frac{k}{\log 2} t^k \int_t^\infty \frac{\omega_{k+1}(f,u)}{u^k}\frac{du}{u}, \tag{4.11}$$

which is *Marchaud's inequality.*

Iterating (4.11) and using Fubini's theorem, we obtain the right-hand inequality in (4.7) for any integer $r > k$, with

$$c = \frac{(r-1)!}{(k-1)!}(\log 2)^{k-r}. \qquad \blacksquare$$

Corollary 4.5. *Suppose $\alpha > 0$ and $1 \le p, s \le \infty$. As r varies over all positive integers with $r > \alpha$, the expressions* (4.6) *define a family of mutually equivalent norms on the Besov space $B^p_{\alpha,s}$.*

Proof. Suppose k and r are integers satisfying $r > k > \alpha$. Let $\|\cdot\|$ denote the Besov norm defined by (4.6) with respect to the parameter r and let $\|\cdot\|'$ be the corresponding norm defined by (4.6) but with r replaced by k. That $\|\cdot\| \le c\|\cdot\|'$ follows from the first of the inequalities in (4.7). The reverse inequality $\|\cdot\|' \le c\|\cdot\|$ is derived in similar fashion from the second inequality in (4.7) but requires an application of Hardy's inequality (III.(3.19)) to simplify the right-hand side. $\blacksquare$

Corollary 4.5 leads to the following embedding theorem.

Theorem 4.6. *Suppose that $1 \le p \le \infty$, that $\alpha, \beta > 0$, and that $1 \le s, t \le \infty$. If $\alpha < \beta$ or if $\alpha = \beta$ and $s \le t$, then*

$$\|f\|_{B^p_{\alpha,s}} \le c\|f\|_{B^p_{\beta,t}}, \qquad (f \in B^p_{\beta,t}). \tag{4.12}$$

In particular, $B^p_{\beta,t} \hookrightarrow B^p_{\alpha,s}$.

Proof. If r is an integer larger than both α and β, then by Corollary 4.5, both Besov norms may be realized in terms of the r-th order modulus of continuity. If $\alpha = \beta$, inequality (4.12) is established as in the proof of Proposition IV.4.2 for Lorentz spaces. The case $\alpha < \beta$ follows directly from Hölder's inequality and the remark following Definition 4.3. We omit the details. $\blacksquare$

Definition 4.7. The function M_j, $(j = 1, 2, \ldots)$, defined inductively on the real line by

$$M_1 = \chi_{(0,1)}, \qquad M_{j+1} = M_1 * M_j, \qquad (j = 1, 2, \ldots),$$

is called the *B-spline* of order j.

Each B-spline M_j is a nonnegative function on $(0, j)$, bounded by 1, and with integral equal to 1. The identity

$$\Delta_h^r f(x) = \int_{-\infty}^{\infty} f^{(r)}(x + \xi h) h^r M_r(\xi)\, d\xi, \tag{4.13}$$

valid for functions f on $\mathbf{R}$ with an absolutely continuous derivative of order $r - 1$, will be crucial in what follows.

The proof of (4.13) is by induction on r. The verification for $r = 1$ is easy, so suppose (4.13) holds for all integers $1, 2, \ldots, r$. Applying Δ_h to (4.13), we obtain

$$\begin{aligned} h^{-(r+1)}\Delta_h^{r+1} f(x) &= \int_{-\infty}^{\infty} h^{-1}\Delta_h f^{(r)}(x + \xi h) M_r(\xi)\, d\xi \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f^{(r+1)}(x + \xi h + \zeta h) M_1(\zeta)\, d\zeta\, M_r(\xi)\, d\xi \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} M_1(\eta - \xi) M_r(\xi)\, d\xi\, f^{(r+1)}(x + \eta h)\, d\eta \\ &= \int_{-\infty}^{\infty} f^{(r+1)}(x + \eta h) M_{r+1}(\eta)\, d\eta. \end{aligned}$$

This establishes (4.13) with r replaced by $r + 1$, and hence completes the proof. The same result can also be derived from the identity

$$\Delta_h^r f(x) = \int_0^h \cdots \int_0^h f^{(r)}(x + \eta_1 + \cdots + \eta_r)\, d\eta_1 \cdots d\eta_r. \tag{4.14}$$

The following multi-variable notation is standard.

Definition 4.8. A *multi-index* is an n-tuple $v = (v_1, v_2, \ldots, v_n)$ of nonnegative integers. Its *length* $|v|$ is the quantity

$$|v| = \sum_{j=1}^n v_j.$$

The *differential operator* D^v is defined by

$$D^v f = D_1^{v_1} D_2^{v_2} \cdots D_n^{v_n} f,$$

where $D_j = \partial/\partial x_j$. The *order* of the differential operator D^ν is the length $|\nu|$ of the multi-index ν. If $h = (h_1, h_2, \ldots, h_n)$ belongs to $\mathbf{R}^n$, then the *power* h^ν is defined by

$$h^\nu = h_1^{\nu_1} h_2^{\nu_2} \cdots h_n^{\nu_n}.$$

For any multi-index ν, we have

$$|h^\nu| \le |h|^{|\nu|}. \tag{4.15}$$

If f is a function on $\mathbf{R}^n$ with continuous partial derivatives of order r, then

$$\Delta_h^r f(x) = \int_{-\infty}^{\infty} M_r(\xi) \sum_{|\nu|=r} \frac{r!}{\nu!} D^\nu f(x + \xi h) h^\nu \, d\xi. \tag{4.16}$$

To see this, let $g(t) = f(x + th/|h|)$ and observe that

$$\Delta_{|h|}^r g(t) = \Delta_h^r f\left(\frac{x + th}{|h|}\right).$$

The function $g^{(r)}$ is the r-th directional derivative of f in the direction h, and

$$g^{(r)}(\xi|h|) = \sum_{|\nu|=r} \frac{r!}{\nu!} D^\nu f(x + \xi h)\left(\frac{h}{|h|}\right)^\nu.$$

Applying (4.13) to the function $g(t)$ (at $t = 0$), we obtain

$$\Delta_h^r f(x) = \Delta_{|h|}^r g(0) = \int_{-\infty}^{\infty} g^{(r)}(\xi|h|) h^r M_r(\xi)\, d\xi,$$

which, together with the preceding estimate, establishes (4.16).

Definition 4.9. The *Sobolev space* W_r^p consists of those functions f on $\mathbf{R}^n$ for which all distributional derivatives of f of order at most r belong to L^p (if $p < \infty$) or C (if $p = \infty$). The space W_r^p is a Banach space under the norm

$$\|f\|_{W_r^p} = \sum_{|\nu| \le r} \|D^\nu f\|_p \tag{4.17}$$

(the L^p-norm is replaced by the norm in C if $p = \infty$).

The space $\dot{W}_r^p$ consists of those f for which the seminorm

$$|f|_{W_r^p} = \sum_{|\nu| = r} \|D^\nu f\|_p$$

is finite.

The intermediate derivatives of orders $1, 2, \ldots, r-1$ can be estimated in terms of those of orders 0 and r, as follows.

Theorem 4.10. *If $f \in \dot{W}_r^p \cap L^p$, then $f \in W_r^p$ and*

$$\|D^\nu f\|_p \le c t^{-|\nu|}(\|f\|_p + t^r |f|_{W_r^p}), \qquad (t > 0), \tag{4.18}$$

whenever $|\nu| \le r$.

Proof. By considering dilates of f, we need only establish (4.18) for $t = 1$. The cases $|\nu| = 0$ and $|\nu| = r$ are trivial so we may assume $1 \le |\nu| \le r - 1$. Consider first the one-variable case $n = 1$. By a density argument, we may assume that $f^{(\nu)}$ is continuous. Since each B-spline M^ν has integral 1, we have from (4.13) and the mean-value theorem for integrals,

$$\Delta_1^\nu f(x) = \int_{-\infty}^{\infty} f^{(\nu)}(x + \xi) M_\nu(\xi)\, d\xi = f^{(\nu)}(x + \zeta), \tag{4.19}$$

for some ζ satisfying $0 \le \zeta \le \nu$. On the other hand,

$$f^{(\nu)}(x + \zeta) = f^{(\nu)}(x) + \int_0^\zeta f^{(\nu+1)}(x + u)\, du,$$

so combining this with (4.19) we have

$$f^{(\nu)}(x) = \Delta_1^\nu f(x) - \int_0^\zeta f^{(\nu+1)}(x + u)\, du. \tag{4.20}$$

Applying L^p-norms to each side of (4.20), we obtain

$$\|f^{(\nu)}\|_p \le 2^\nu \|f\|_p + \nu \|f^{(\nu+1)}\|_p.$$

Iteration of this result gives

$$\|f^{(\nu)}\|_p \le c\{\|f\|_p + \|f^{(r)}\|_p\}, \tag{4.21}$$

which is the one-dimensional version of (4.18).

In the multi-dimensional case $n \ge 2$, consider first a multi-index ν in which all components but one (ν_i, say) are equal to zero. Let f_i denote the function of one variable obtained from f by fixing all components except the i-th. The one-dimensional result (4.21) can then be applied to f_i to give

$$\|D_i^{\nu_i} f_i\|_{L^p(\mathbf{R})} \le c\{\|f_i\|_{L^p(\mathbf{R})} + \|D_i^r f_i\|_{L^p(\mathbf{R})}\}. \tag{4.22}$$

Applying the p-norm in the remaining $n - 1$ variables to each side of (4.22), we obtain, for all r,

$$\|D_i^{\nu_i} f\|_p \le c\{\|f\|_p + \|D_i^r f\|_p\}. \tag{4.23}$$

Now let $\nu = (\nu_1, \nu_2, \ldots, \nu_n)$ be an arbitrary multi-index, let g_i denote the function $D_i^{\nu_i} \ldots D_n^{\nu_n} f$, and let $r_i = r - \sum_{m=i+1}^n \nu_m$. Then successive applications of (4.23) give

$$\|g_i\|_p \le c\{\|g_{i+1}\|_p + |f|_{W_r^p}\}.$$

Hence,

$$\|D^\nu f\|_p = \|g_1\|_p \le c\{\|g_2\|_p + |f|_{W_r^p}\} \le \cdots$$
$$\le c\{\|g_n\|_p + |f|_{W_r^p}\} \le c\{\|f\|_p + |f|_{W_r^p}\},$$

and this completes the proof. ■

We remark that if $|f|_{W_r^p}$ is nonzero, then choosing the value $t = (\|f\|_p/|f|_{W_r^p})$ in (4.18) yields the following inequality:

$$\|D^\nu f\|_p \le c\|f\|_p^{1-|\nu|/r}|f|_{W_r^p}^{|\nu|/r}. \tag{4.24}$$

The next lemma will be crucial in helping determine the K-functional for the couple (L^p, W_r^p).

Lemma 4.11. *Let $h_1, h_2, \ldots, h_r$ be arbitrary vectors in $\mathbf{R}^n$. Then*

$$\prod_{j=1}^r \Delta_{h_j} = \sum_{D\subset\{1,\ldots,r\}} (-1)^{|D|} T_{h_D^*}\Delta_{h_D}^r, \tag{4.25}$$

where the sum extends over all subsets D of $\{1,\ldots r\}$, $|D|$ is the cardinality of D, and $h_D^* = \sum_{j\in D} h_j$, $h_D = -\sum_{j\in D} j^{-1}h_j$. Furthermore, if

$$\bar\omega(f,t)_p = \sup_{\substack{|h_j|\le t\\ 1\le j\le r}} \left\| \prod_{j=1}^r \Delta_{h_j} f\right\|_p, \tag{4.26}$$

then there is a constant $c > 0$ such that

$$\omega_r(f,t)_p \le \bar\omega(f,t)_p \le c\,\omega_r(f,t)_p, \qquad (t>0). \tag{4.27}$$

Proof. For each integer k with $0 \le k \le r$,

$$\prod_{j=1}^r \Delta_{(j-k)h_j} = \prod_{j=1}^r (T_{(j-k)h_j} - I)$$
$$= \sum_{D\subset\{1,\ldots,r\}} (-1)^{r-|D|} \prod_{j\in D} T_{(j-k)h_j}$$
$$= \sum_D (-1)^{r-|D|} T_{(\Sigma_{j\in D} jh_j)} T^k_{(-\Sigma_{j\in D} h_j)}. \tag{4.28}$$

The left-hand side of (4.28) is of course equal to zero if k is nonzero. Hence, if we multiply both sides of (4.28) by $(-1)^{r-k}\binom{r}{k}$ and sum over $k = 0, 1, \ldots, r$,

we have

$$(-1)^r \prod_{j=1}^r \Delta_{jh_j} = \sum_D (-1)^{r-|D|} T_{(\Sigma_{j\in D} jh_j)} \Delta^r_{(-\Sigma_{j\in D} h_j)}.$$

Now replacing h_j by $j^{-1}h_j$, we obtain (4.25).

The right-hand inequality in (4.27) follows by applying the p-norm to both sides of (4.25) since the sum has at most 2^r terms, each h_D satisfies $|h_D| \le rt$, and since, by (4.5), the modulus satisfies $\omega_r(f, rt)_p \le (2r)^r \omega_r(f,t)_p$. The left-hand inequality in (4.27) is an immediate consequence of the definitions. ■

Now we can describe the K-functional $K(f,t) = K(f,t; L^p, W_r^p)$ for the couple (L^p, W_r^p) in terms of the modulus of smoothness $\omega_r(f,t)_p$.

Theorem 4.12. *Suppose $1 \le p \le \infty$ and let r be any positive integer. There are positive constants c_1 and c_2 such that*

$$c_1 K(f,t^r) \le \min(1,t^r)\|f\|_p + \omega_r(f,t)_p \le c_2 K(f,t^r), \tag{4.29}$$

for all f in L^p and all $t > 0$.

Proof. We begin by establishing the right-hand inequality in (4.29). If $f \in L^p$ is decomposed as $f = b + g$, with $b \in L^p$ and $g \in W_r^p$, then

$$\min(1,t^r)\|f\|_p \le \|b\|_p + t^r\|g\|_p \le \|b\|_p + t^r\|g\|_{W_r^p}.$$

Hence, by taking the infimum with respect to all such decompositions, we see that the quantity on the left is majorized by $K(f,t^r)$. Thus, it will suffice to establish a similar estimate for the other term in (4.29), namely,

$$\omega_r(f,t)_p \le cK(f,t^r). \tag{4.30}$$

If f is decomposed as before in the form $f = b + g$, then for $|h| \le t$,

$$\|\Delta_h^r f\|_p \le \|\Delta_h^r b\|_p + \|\Delta_h^r g\|_p. \tag{4.31}$$

It follows from (4.4) that

$$\|\Delta_h^r b\|_p \le 2^r\|b\|_p. \tag{4.32}$$

On the other hand, by (4.16) and Minkowski's inequality,

$$\|\Delta_h^r g\|_p \le r! \int_{-\infty}^{\infty} M_r(\xi)\,d\xi\, t^r|g|_{W_r^p} = r!\,t^r|g|_{W_r^p}. \tag{4.33}$$

Combining (4.32) with (4.33) and passing to the infimum over all decompositions $f = b + g$, we obtain the desired result (4.30) with constant $c = 2^r + r!$

Since $K(f,t^r) \leq \|f\|_p$ for all $t \geq 1$, it will suffice to establish the left-hand inequality in (4.29) for $0 < t < 1$. In that case, define a function b by the Bochner integral

$$b = (-1)^r \int_U \cdots \int_U \Delta^r_{t\Sigma u_j} f \, du_1 \ldots du_r, \tag{4.34}$$

where U is the unit cube $U = [0,1]^n$ in $\mathbf{R}^n$. If $g = f - b$, then from (4.1) we have

$$g = (-1)^{r+1} \sum_{k=1}^r (-1)^{r-k} \binom{r}{k} \int_{U^r} T_{kt(\Sigma_j u_j)} f \, du_1 \ldots du_r. \tag{4.35}$$

An application of Minkowski's inequality to (4.34) gives

$$\begin{aligned} \|b\|_p &\leq \int_{U^r} \|\Delta^r_{t\Sigma u_j} f\|_p \, du_1 \cdots du_r \\ &\leq \omega_r(f, rt)_p \leq c\omega_r(f,t)_p, \end{aligned} \tag{4.36}$$

the last inequality because of (4.5). For the function g, on the other hand, Theorem 4.10 gives

$$\|g\|_{W_r^p} \leq c\{\|g\|_p + \sum_{|\nu|=r} \|D^\nu g\|_p\}. \tag{4.37}$$

The first term on the right can be estimated as follows by using (4.35) and the translation-invariance of the p-norm:

$$\|g\|_p \leq \sum_{k=1}^r \binom{r}{k} \|f\|_p \leq 2^r \|f\|_p. \tag{4.38}$$

To estimate the remaining terms, note that for the k-th summand g_k, say, in (4.35) we have

$$|D^\nu g_k(x)| = \binom{r}{k} \left| \int_{U^r} D^\nu f(x + kt(u_1 + \cdots + u_r)) \, du_1 \ldots du_r \right|.$$

If σ_{ji} denotes the i-th component of u_j, $(1 \leq i \leq n, 1 \leq j \leq r)$, then the r derivatives in $D^\nu(g_k)$ can be integrated to give differences, thus yielding the identity

$$|D^\nu g_k(x)| = \binom{r}{k} (kt)^{-r} \left| \int_0^1 \cdots \int_0^1 \Delta^{\nu_i}_{kt\mathbf{e}_i} f(x + kt\Sigma\sigma_{ij}) \, \Pi \, d\sigma_{ij} \right|, \tag{4.39}$$

where $\mathbf{e}_i$ is the i-th coordinate vector in $\mathbf{R}^n$ and both sum and product are over the remaining $(r-1)n$ coordinates σ_{ij}, that is, $1 \leq i \leq n$ and $\nu_i + 1 \leq j \leq r$. Applying the p-norm to both sides of (4.39) and using Minkowski's inequality, we see from the estimate (4.27) in Lemma 4.11 that

$$t^r \|D^\nu g_k\|_p \leq c\omega_r(f, kt)_p.$$

Summing over $k = 1, 2, \ldots, r$ and using (4.5), we therefore have

$$t^r \|D^\nu g\|_p \leq c\omega_r(f,t)_p. \tag{4.40}$$

The estimates (4.37), (4.38) and (4.40) now combine to give

$$t^r \|g\|_{W_r^p} \leq c\{t^r \|f\|_p + \omega_r(f,t)_p\}.$$

Together with the estimate (4.36) for the function b, this gives

$$K(f,t^r) \leq \|b\|_p + t^r \|g\|_{W_r^p} \leq c\{t^r \|f\|_p + \omega_r(f,t)_p\},$$

for $0 < t < 1$, which, as we remarked above, establishes the left-hand inequality in (4.29). The proof is now complete. ∎

If we define

$$\dot{K}(f,t) = \inf\{\|f - g\|_p + t|g|_{W_r^p} : |g|_{W_r^p} < \infty\}, \tag{4.41}$$

then the proof above shows that

$$c_1 \dot{K}(f,t^r) \leq \omega_r(f,t)_p \leq c_2 \dot{K}(f,t^r), \qquad (t > 0). \tag{4.42}$$

The left-hand inequality uses (4.36) and (4.40); the right-hand inequality requires (4.4) and (4.44) below.

Having determined the K-functional between L^p and W_r^p in terms of the modulus of smoothness, it is now routine to identify the corresponding (θ, q)-interpolation spaces as Besov spaces:

Corollary 4.13. *Suppose $1 \leq p \leq \infty$ and let r be a positive integer. If $0 < \theta < 1$ and $1 \leq q \leq \infty$, then*

$$(L^p, W_r^p)_{\theta,q} = B^p_{\theta r, q}$$

with equivalent norms.

The Marchaud inequalities (4.7) relate moduli of smoothness of different orders. We shall now establish similar kinds of inequalities relating moduli of smoothness and derivatives.

Theorem 4.14. *Let j and r be integers with $0 < j < r$. Then there are constants $c_1, c_2 > 0$ such that, for all f in W_j^p,*

$$c_1 \sup_{t>0} \{t^{-j} \omega_j(f,t)_p\} \leq |f|_{W_j^p} \leq c_2 \int_0^\infty \frac{\omega_r(f,t)_p}{t^j} \frac{dt}{t}. \tag{4.43}$$

Proof. If $|h| \le t$, then, by (4.16),

$$\|\Delta_h^j f\|_p \le r! \int_{-\infty}^{\infty} M_j(\xi)|f|_{W_j^p}|h|^j\, d\xi \le r!t^j|f|_{W_j^p}, \tag{4.44}$$

from which the first estimate in (4.43) follows.

For the second, we may assume that the integral on the right of (4.43) is finite. For each integer i, there is, by virtue of (4.42), a decomposition $f = b_i + g_i$ of f such that

$$\|b_i\|_p \le c\omega_r(f, 2^i)_p, \qquad |g_i|_{W_r^p} \le c(2^i)^{-r}\omega_r(f, 2^i)_p. \tag{4.45}$$

Define a_i in W_r^p by $a_i = b_{i+1} - b_i = g_i - g_{i+1}$. By (4.45),

$$\begin{aligned}\|a_i\|_p &\le c\{\omega_r(f, 2^i)_p + \omega_r(f, 2^{i+1})_p\}\\ &\le c(2^r + 1)\omega_r(f, 2^i)_p. \end{aligned}\tag{4.46}$$

Similarly,

$$|a_i|_{W_r^p} \le 2c(2^i)^{-r}\omega_r(f, 2^i)_p. \tag{4.47}$$

Since $f = \sum a_i$, we find from Theorem 4.10 that

$$|f|_{W_j^p} \le \sum_i |a_i|_{W_j^p} \le c\sum_i \{(2^i)^{-j}\|a_i\|_p + (2^i)^{r-j}|a_i|_{W_r^p}\},$$

and hence from (4.46) and (4.47) that

$$|f|_{W_j^p} \le c\sum_i (2^i)^{-j}\omega_r(f, 2^i)_p. \tag{4.48}$$

Since $\omega_r(f, t)_p$ is an increasing function of t, we may replace the i-th summand in (4.48) by a corresponding integral over the interval $(2^i, 2^{i+1})$, then perform the summation to obtain the right-hand inequality in (4.43). ∎

Corollary 4.15. *If $\int_0^1 t^{-j}\omega_r(f, t)_p\, dt/t < \infty$, then $f \in W_j^p$ and*

$$\sum_{|\nu|=j} \omega_r(D^\nu f, t)_p \le c\int_0^t \frac{\omega_r(f, s)_p}{s^j}\frac{ds}{s}, \qquad (t > 0). \tag{4.49}$$

Proof. If $|\nu| = j < r$, Theorem 4.14 shows that

$$\|D^\nu f\|_p \le c\int_0^\infty \frac{\omega_r(f, s)_p}{s^j}\frac{ds}{s}.$$

Applying this to $\Delta_h^r f$ instead of f, where $|h| \le t$, we obtain

$$\omega_r(D^\nu f, t)_p = \sup_{|h|\le t}\|D^\nu(\Delta_h^r f)\|_p \le \sup_{|h|\le t} c\int_0^\infty \frac{\omega_r(\Delta_h^r f, s)_p}{s^j}\frac{ds}{s}.$$

But

$$\omega_r(\Delta_h^r f, s)_p \le 2^r\min\{\omega_r(f, s)_p, \omega_r(f, t)_p\}$$

so the preceding estimate gives

$$\begin{aligned}\omega_r(D^\nu f, t)_p &\le c\left\{\int_0^t \frac{\omega_r(f, s)_p}{s^j}\frac{ds}{s} + \int_t^\infty \frac{ds}{s^{j+1}}\omega_r(f, t)_p\right\}\\ &\le c\int_0^t \frac{\omega_r(f, s)_p}{s^j}\frac{ds}{s},\end{aligned}$$

since

$$\frac{\omega_r(f, t)_p}{t^j} \le c\int_{t/2}^t \frac{\omega_r(f, s)_p}{s^j}\frac{ds}{s}. \qquad ∎$$

Corollary 4.16 (Reduction theorem). *Let r be a positive integer and let $0 < \alpha < r$. Suppose $\alpha = j + \beta$, where j is an integer between 0 and r, and $\beta > 0$. Let $1 \le p \le \infty$. A function f belongs to the Besov space $B_{\alpha,q}^p$ if and only if all distributional derivatives $D^\nu f$ of f of order $|\nu| = j$ belong to the Besov space $B_{\beta,q}^p$.*

Proof. If $|h| \le t$, it follows from (4.16) that

$$\|\Delta_h^r f\|_p = \|\Delta_h^{r-j}\Delta_h^j f\|_p \le ct^j \sum_{|\nu|=j} \|\Delta_h^{r-j}D^\nu f\|_p.$$

Hence,

$$\omega_r(f, t)_p \le ct^j \sum_{|\nu|=j} \omega_{r-j}(D^\nu f, t)_p, \tag{4.50}$$

which shows that $f \in B_{\alpha,q}^p$ if its derivatives of order j belong to $B_{\beta,q}^p$.

The converse follows immediately from (4.49) and Hardy's inequality III.(3.18). ∎

We can now identify certain interpolation spaces as follows.

Theorem 4.17. *Suppose $0 < \theta < 1$ and $1 \le q \le \infty$. Then*

$$\begin{aligned}(B_{\alpha_0,q_0}^p, B_{\alpha_1,q_1}^p)_{\theta,q} &= B_{\alpha,q}^p, \qquad \alpha = (1-\theta)\alpha_0 + \theta\alpha_1,\ \alpha_0 \ne \alpha_1;\\ (W_r^p, B_{\alpha_1,q_1}^p)_{\theta,q} &= B_{\alpha,q}^p, \qquad \alpha = (1-\theta)r + \theta\alpha_1, \quad r \ne \alpha_1;\\ (W_{r_0}^p, W_{r_1}^p)_{\theta,q} &= B_{\alpha,q}^p, \qquad \alpha = (1-\theta)r_0 + \theta r_1,\ r_0 \ne r_1.\end{aligned}$$

Proof. All of these results will follow from the reiteration theorem (Theorem 2.4), once it is shown that the interpolated spaces belong to the appropriate classes θ_j (cf. Definition 2.2). However, Corollary 4.13 shows that for large r, the Besov space

$$B^p_{\alpha_j, q_j} = (L^p, W^p_r)_{\theta_j, q_j}, \qquad \theta_j = \frac{\alpha_j}{r},$$

and hence is of class α_j/r. On the other hand, Theorem 4.12 and inequality (4.43) show that the Sobolev space $W^p_{r_j}$ is of class r_j/r for (L^p, W^p_r). We omit the remaining details. ■

Theorem 4.17 could also have been established directly from Theorem 4.12 since (4.50) and (4.24) imply conditions (2.45) and (2.47), respectively, and these guarantee that

$$(L^p, W^p_r)_{\theta,1} \hookrightarrow W^p_{r_1} \hookrightarrow (L^p, W^p_r)_{\theta,\infty}, \qquad \theta = \frac{r_1}{r}.$$

We conclude this section with a variety of embedding theorems for Sobolev and Besov spaces.

Theorem 4.18. *Suppose $f \in W^p_r(\mathbf{R}^n)$, where $1 \le p \le \infty$ and $r \ge n$. Then f can be modified on a set of measure zero so as to become a continuous function that satisfies*

$$\|f\|_\infty \le c\|f\|_{W^p_r}, \tag{4.51}$$

and, in fact, for any $t > 0$,

$$\|f\|_\infty \le ct^{-n/p}\{\|f\|_p + t^r|f|_{W^p_r}\}. \tag{4.52}$$

Proof. It will suffice to establish (4.51) for $r = n$ since the right-hand side increases with r. By a simple density argument, we may also assume that f has continuous derivatives up to and including order n. We begin by showing that

$$|f(x)| \le \int_{x+Q} \sum_{|\nu| \le n} |D^\nu f(u)|\,du, \qquad (x \in \mathbf{R}^n), \tag{4.53}$$

where Q is the unit cube $\{u : 0 \le u_j \le 1\}$ in $\mathbf{R}^n$. The proof is by induction on n.

In the case $n = 1$, the mean-value theorem for integrals yields ξ with $x < \xi < x + 1$ such that

$$f(\xi) = \int_x^{x+1} f(u)\,du.$$

Hence,

$$|f(x)| = |f(\xi) + \int_\xi^x f'(u)\,du| \le \int_x^{x+1} \{|f(u)| + |f'(u)|\}\,du,$$

which is the desired result (4.53) for $n = 1$.

Now suppose that (4.53) holds for $1, 2, \ldots, n-1$, and let $x \in \mathbf{R}^n$. Let x' denote the vector consisting of the last $n-1$ coordinates of x, so that x may be written $x = (x_1, x')$. Using the induction hypothesis, first for the value 1, then for the value $n-1$, we obtain

$$\begin{aligned} |f(x_1, x')| &\le \int_{x_1}^{x_1+1} \{|f(u_1, x')| + |D_1 f(u_1, x')|\}\,du_1 \\ &\le \int_{x_1}^{x_1+1} \int_{x'+Q'} \sum_{|\nu'| \le n-1} |D^{\nu'} f(u_1, u')| + |D_1 D^{\nu'} f(u_1, u')|\,du'\,du_1 \\ &\le \int_{x+Q} \sum_{|\nu| \le n} |D^\nu f(u)|\,du, \end{aligned}$$

where Q' is the unit cube in $\mathbf{R}^{n-1}$ and ν' is the $(n-1)$-dimensional multi-index in the coordinates $x_2, \ldots, x_n$. This completes the inductive step and hence establishes (4.53) (with $r = n$).

The estimate (4.51) follows directly from (4.53) and Hölder's inequality. The remaining estimate (4.52) follows from Theorem 4.10 and (4.51) applied to the dilates $x \to f(tx)$ of f. ■

The next result provides an estimate for the decreasing rearrangement in terms of the modulus of smoothness.

Theorem 4.19. *Suppose $1 \le p \le \infty$. Then there is a constant c such that for all $f \in L^p(\mathbf{R}^n)$,*

$$f^{**}(t) \le c \int_{t^{1/n}}^\infty \frac{\omega_n(f, s)_p}{s^{n/p}} \frac{ds}{s}, \qquad (t > 0). \tag{4.54}$$

Proof. Let $f \in L^p$ and fix $t > 0$. For each $i = 0, 1, 2, \ldots$, there is, by (4.42), a decomposition $f = b_i + g_i$ for which

$$\|b_i\|_p + (2^i t)^n |g_i|_{W^p_n} \le c\omega_n(f, 2^i t)_p. \tag{4.55}$$

Define $a_i \in W^p_n$ by

$$a_i = b_{i+1} - b_i = g_i - g_{i+1}, \qquad (i = 0, 1, \ldots). \tag{4.56}$$

Then $f = b_0 + \sum_{i=0}^{\infty} a_i$ and so

$$f^{**}(t^n) \le b_0^{**}(t^n) + \sum_{i=0}^{\infty} \|a_i\|_\infty. \tag{4.57}$$

Using the estimate (4.52) with t replaced by $2^i t$, together with (4.55) and (4.56), we obtain

$$\begin{aligned} \|a_i\|_\infty &\le c(2^i t)^{-n/p}\{\|a_i\|_p + (2^i t)^n |a_i|_{W_n^p}\} \\ &\le c(2^i t)^{-n/p}\{\omega_n(f, 2^i t)_p + \omega_n(f, 2^{i+1} t)_p\} \\ &\le c(2^i t)^{-n/p}\omega_n(f, 2^i t)_p \le c \int_{2^i t}^{2^{i+1} t} \frac{\omega_n(f,s)_p}{s^{n/p}} \frac{ds}{s}. \end{aligned} \tag{4.58}$$

On the other hand, an application of Hölder's inequality gives

$$b_0^{**}(t^n) \le t^{-n/p}\|b_0\|_p \le ct^{-n/p}\omega_n(f,t)_p \le c\int_t^{2t} \frac{\omega_n(f,s)_p}{s^{-n/p}}\frac{ds}{s}.$$

Substituting this estimate and (4.58) into (4.57), we obtain the desired conclusion (4.54). ∎

Corollary 4.20. *Suppose $1 \le p < q \le \infty$ and $1 \le a \le \infty$. Let $\alpha = n(1/p - 1/q)$.*
(a) *If $q < \infty$, then there is a constant c such that*

$$\|f\|_{L^{q,a}} \le c\|f\|_{B^p_{\alpha,a}}, \qquad (f \in B^p_{\alpha,a}). \tag{4.59}$$

In particular, the Besov space $B^p_{\alpha,a}$ is continuously embedded in the Lorentz space $L^{q,a}$.
(b) *If $q = \infty$, then there is a constant c such that*

$$\|f\|_{L^\infty} \le c\|f\|_{B^p_{n/p,1}}, \qquad (f \in B^p_{n/p,1}). \tag{4.60}$$

In particular, the Besov space $B^p_{n/p,1}$ is continuously embedded in the Lebesgue space L^∞.

Proof. The inequality (4.59) is established by applying the Lorentz $L^{q,a}$-norm to (4.54) and using Hardy's inequality III.(3.19) to simplify the right-hand side. The remaining estimate (4.60) follows immediately from (4.54) by letting t tend to zero. ∎

Corollary 4.21. *Suppose $1 \le p < q < \infty$ and let r and n be positive integers with $r \ge n$. Let $\alpha = n(1/p - 1/q)$. Then there is a constant $c > 0$ such that*

$$\omega_r(f,t)_q \le c\int_0^t s^{-\alpha}\omega_r(f,s)_p \frac{ds}{s}, \qquad (0 < t < 1). \tag{4.61}$$

In particular, whenever $\beta > 0$ and $1 \le a \le \infty$, the estimate

$$\|f\|_{B^q_{\beta,a}} \le c\|f\|_{B^p_{\alpha+\beta,a}}, \qquad (f \in B^p_{\alpha+\beta,a}) \tag{4.62}$$

holds, and so the Besov space $B^p_{\alpha+\beta,a}$ is continuously embedded in the Besov space $B^q_{\beta,a}$.

Proof. Using Corollary 4.20 and (4.12), we have

$$\|f\|_q \le c\|f\|_{B^p_{\alpha,q}} \le c\left\{\int_0^1 \omega_r(f,s)_p s^{-\alpha}\frac{ds}{s} + \|f\|_p\right\}. \tag{4.63}$$

If $|h| \le t \le 1$, then (4.63) applied to $\Delta_h^r f$ gives

$$\|\Delta_h^r f\|_q \le c\left\{\int_0^t \omega_r(f,s)_p s^{-\alpha}\frac{ds}{s} + t^{-\alpha}\omega_r(f,t)_p + \|\Delta_h^r f\|_p\right\} \tag{4.64}$$

since

$$\omega_r(\Delta_h^r f, s)_p \le 2^r \min\{\omega_r(f,s)_p, \omega_r(f,t)_p\} \le 2^r\omega_r(f, \min(s,t))_p.$$

Hence, if $0 < t \le 1$,

$$\omega_r(f,t)_p \le t^{-\alpha}\omega_r(f,t)_p \le c\int_{t/2}^t \omega_r(f,s)_p s^{-\alpha}\frac{ds}{s},$$

and so (4.61) follows from (4.64).

To establish (4.62), we first use inequality (4.59) together with (4.12) to obtain

$$\|f\|_q \le c\|f\|_{B^p_{\alpha,q}} \le c\|f\|_{B^p_{\alpha+\beta,a}}. \tag{4.65}$$

The L^q-modulus of continuity can be estimated by means of Hardy's inequality (III.(3.18)) to give

$$\left\{\int_0^1 [\omega_r(f,t)_q t^{-\beta}]^a \frac{dt}{t}\right\}^{1/a} \le c\|f\|_{B^p_{\alpha+\beta,a}}, \tag{4.66}$$

provided $a < \infty$; the case $a = \infty$ is similar. Together, estimates (4.65) and (4.66) yield the desired inequality (4.62). This completes the proof. ∎

5. INTERPOLATION BETWEEN W_k^1 AND W_k^∞

The K-functional for the couple (W_k^1, W_k^∞) of Sobolev spaces was first identified by R. A. DeVore and K. Scherer, using spline-function techniques and combinatorial arguments. We shall present a maximal-function proof based on the Whitney covering lemma, which we now describe.

By a *cube* Q in $\mathbf{R}^n$, we shall always mean a set of the form

$$Q = \{x \in \mathbf{R}^n : a_i \le x_i \le a_i + h, i = 1,2,\ldots,n\}, \tag{5.1}$$

where $a = (a_1, a_2, \ldots, a_n)$ is a point in $\mathbf{R}^n$ and $h \ge 0$, that is, a "cube" in the ordinary sense of the word but whose sides are parallel to the coordinate axes. It is immaterial whether the boundary is or is not considered a part of the cube; we have included the boundary in Q (cf. (5.1)) merely to fix ideas. In this spirit, it will be convenient to say that two cubes are *essentially disjoint* if their interiors do not intersect.

The *vertices* of a cube Q, defined as in (5.1), are the 2^n points with coordinates $a_i + \varepsilon_i h$, $(i = 1,2,\ldots,n)$, where each ε_i is 0 or 1. The *center* $\operatorname{cen}(Q)$ of Q is the point with coordinates $a_i + h/2$, $(i = 1,2,\ldots,n)$. The *side-length* of Q is given by $\operatorname{len}(Q) = h$, and the *diameter* by $\operatorname{diam}(Q) = \sqrt{n}\operatorname{len}(Q) = \sqrt{n}h$. The n-dimensional Lebesgue measure of a set E will, as usual, be denoted by $|E|$; thus, in particular, the *volume* of a cube Q of side-length h is given by $|Q| = h^n$.

We shall use $d(x, y)$ or $|x - y|$ to denote the euclidean distance between two points x and y in $\mathbf{R}^n$. Thus, $d(x, E)$ and $d(E, F)$ will denote the euclidean distance between a point x and a set E or between two sets E and F in $\mathbf{R}^n$, respectively.

The countably many *dyadic cubes* are defined as follows. First, let j be any integer and consider the lattice of points ξ in $\mathbf{R}^n$ with coordinates $\xi_i = 2^j k_i$, $(i = 1,2,\ldots,n)$, where each k_i is an integer. The collection Ξ_j then consists of all cubes Q of side-length 2^j with vertices at points of this lattice. The union over all integers j of the collections Ξ_j then constitutes the collection of dyadic cubes.

Observe that if the interiors of two dyadic cubes have nonempty intersection, then one of the cubes contains the other. In particular, coverings by dyadic cubes are easily refined to yield subcoverings consisting of cubes that are essentially disjoint. The Whitney covering lemma is of this type. It provides an essentially disjoint covering of an open set Ω by dyadic cubes, where each cube in the cover is distant from the boundary of Ω by an amount roughly proportional to its diameter.

Lemma 5.1 (Whitney covering lemma). *Let Ω be an arbitrary open subset of $\mathbf{R}^n$, and let $F = \mathbf{R}^n \setminus \Omega$. Then there is a collection $\Psi = \{Q_1, Q_2, \ldots\}$ of dyadic cubes for which*

(i) $\bigcup Q_j = \Omega$;
(ii) $Q_j^o \cap Q_k^o = \varnothing$ *if* $j \ne k$;
(iii) $\operatorname{diam}(Q_j) \le d(Q_j, F) \le 4\operatorname{diam}(Q_j)$.

Proof. For each integer j, let

$$\Omega_j = \{x \in \Omega : 2^{j+1} < n^{-1/2} d(x, F) \le 2^{j+2}\},$$

and denote by Ψ_0 the collection

$$\Psi_0 = \bigcup_j \{Q \in \Xi_j : Q \cap \Omega_j \ne \varnothing\}$$

of dyadic cubes. We claim that the cubes in Ψ_0 satisfy

$$\operatorname{diam}(Q) \le d(Q, F) \le 4\operatorname{diam}(Q), \qquad (Q \in \Psi_0). \tag{5.2}$$

Indeed, if $Q \in \Psi_0$, then, for some j, Q has diameter $n^{1/2}2^j$ and contains a point, say ξ, of Ω_j. Then

$$d(Q, F) \le d(\xi, F) \le n^{1/2}2^{j+2} = 4\operatorname{diam}(Q),$$

and

$$d(Q, F) \ge d(\xi, F) - \operatorname{diam}(Q) > n^{1/2}(2^{j+1} - 2^j) = \operatorname{diam}(Q).$$

The cubes in Ψ_0 evidently cover Ω, and it follows from the first inequality in (5.2) that each is contained in Ω. Hence, property i) holds for the cubes in Ψ_0; so does property iii), of course, as we have demonstrated with the establishment of (5.2). The desired cover Ψ will now be selected from Ψ_0 by refinement to obtain cubes that are essentially disjoint.

To achieve this, begin with any cube $Q \in \Psi_0$. We claim that there is a unique maximal cube in Ψ_0 that contains Q. Indeed, if Q' and Q'' are two cubes in Ψ_0 that contain Q, then their interiors have non-empty intersection and so one of the dyadic cubes Q' or Q'' must contain the other. On the other hand, if Q' is any cube in Ψ_0 that contains Q, then (5.2) shows that $\operatorname{diam}(Q') \le 4\operatorname{diam}(Q)$. This establishes the claim. The collection Ψ of all maximal cubes of Ψ_0 now has all the desired properties i), ii), and iii) asserted in the statement of the lemma. ∎

We shall need some additional properties of the Whitney covering Ψ. Two cubes in Ψ will be said to *touch* if their boundaries have non-empty intersection. Further, we shall denote by αQ, where $\alpha > 0$, the cube concentric with Q and with side-length α times the side-length of Q.

Lemma 5.2. *Let Ψ be the Whitney covering of Ω constructed as in the preceding lemma.*
(a) *If two cubes Q_1 and Q_2 of Ψ touch, then*

$$\frac{1}{4}\operatorname{diam}(Q_2) \le \operatorname{diam}(Q_1) \le 4\operatorname{diam}(Q_2); \tag{5.3}$$

(b) *if* $Q \in \Psi$, *then at most* 12^n *cubes of* Ψ *touch* Q;
(c) *each point of* Ω *is contained in at most* 12^n *of the enlarged cubes* $Q^* = (9/8)Q$, $(Q \in \Psi)$. *In particular*

$$\sum_j \chi_{Q_j^*} \le 12^n, \qquad (\Psi = \{Q_1, Q_2, \ldots\}). \tag{5.4}$$

Proof. (a) If Q_1 and Q_2 touch, they have at least one point in common so, by Lemma 5.1,

$$\operatorname{diam}(Q_2) \le d(Q_2, F) \le d(Q_1, F) + \operatorname{diam}(Q_1) \le 5\operatorname{diam}(Q_1).$$

However, the diameters of two dyadic cubes must be in a ratio which is an integral power of 2. Hence, the last estimate shows that $\operatorname{diam}(Q_2) \le 4\operatorname{diam}(Q_1)$. This establishes the first of the estimates in (5.3); the second is established in similar fashion by interchanging the roles of Q_1 and Q_2.
(b) If $Q \in \Xi_j$, then there are 3^n cubes in Ξ_j that touch Q. But each of these cubes can contain at most 4^n cubes in Ψ whose diameters are at least $(1/4)\operatorname{diam}(Q)$. Hence, in view of part (a), at most $3^n \cdot 4^n = 12^n$ cubes in Ψ can touch Q.
(c) Let $x \in \Omega$, and let Q be a cube in Ψ that contains x. Suppose Q' is any cube in Ψ such that $Q^* = (9/8)Q'$ also contains x. We claim that Q' touches Q. Indeed, by part a), every cube in Ψ that touches Q' has diameter at least $(1/4)\operatorname{diam}(Q')$, and so the union U of all of these cubes properly contains Q^*. Hence, Q is contained in the union U, and this shows that Q' touches Q. By part b), at most 12^n cubes Q' can touch Q, so we conclude that at most 12^n cubes in Ψ contain x. The estimate (5.4) is merely a restatement of this property. ∎

With the aid of Lemmas 5.1 and 5.2, we can construct a useful partition of unity. Let Q_0 denote the cube of unit side-length centered at the origin. Let ϕ be a C^∞-function satisfying $0 \le \phi \le 1$, with $\phi = 1$ on Q_0 and $\phi = 0$ outside of the enlarged cube $(9/8)Q_0$.

Now let Ω be an open set and let $\Psi = \{Q_1, Q_2, \ldots\}$ be a Whitney cover of Ω, as before. The function ϕ has a counterpart, say ψ_j, for each cube Q_j in Ψ, defined by

$$\psi_j(x) = \phi\left(\frac{x - \operatorname{cen}(Q_j)}{\operatorname{len}(Q_j)}\right), \qquad (x \in R^n). \tag{5.5}$$

Set

$$\Phi = \sum_j \psi_j, \qquad \phi_j = \frac{\psi_j}{\Phi}. \tag{5.6}$$

Note that, by (5.4), the sum in (5.6) at any point contains at most 12^n nonzero terms; hence, Φ is well-defined. It is clear that

$$\sum_j \phi_j = 1, \tag{5.7}$$

so the collection $\{\phi_j\}$ defines the desired partition of unity.

A simple calculation shows that for any multi-index ν,

$$|D^\nu \phi_j(x)| \le c\,(\operatorname{diam}(Q_j))^{-|\nu|}, \tag{5.8}$$

where c depends only on ϕ and ν.

In the next lemma, we consider a partition of unity $\{\phi_j\}$, constructed as above with respect to a Whitney covering $\Psi = \{Q_j\}$ of an open set Ω. To each of the Whitney cubes Q_j, we associate a point x_j of the complement F with the property that

$$d(x_j, Q_j) = d(F, Q_j). \tag{5.9}$$

Lemma 5.3. *Let* $x \in \Omega$ *and let* J *be any index for which the cube* Q_J *contains* x. *Let* j *be any index for which* $\phi_j(x) \ne 0$. *Then*

$$\frac{1}{4}\operatorname{diam}(Q_J) \le \operatorname{diam}(Q_j) \le 4\operatorname{diam}(Q_J) \tag{5.10}$$

and

$$|x_j - x_J| \le 26\operatorname{diam}(Q_J). \tag{5.11}$$

Proof. Since $\phi_j(x) \ne 0$, we see, as in the proof of Lemma 5.2(c), that Q_j and Q_J touch. Hence, (5.10) follows immediately from (5.3). From Lemma 5.1 iii), (5.9), and (5.10), we have

$$\begin{aligned} \operatorname{diam}(Q_J) &\le d(Q_J, F) \le |x - x_j| \\ &\le \operatorname{diam}(Q_J) + \operatorname{diam}(Q_j) + d(Q_j, F) \\ &\le \operatorname{diam}(Q_J) + 5\operatorname{diam}(Q_j) \le 21\operatorname{diam}(Q_J). \end{aligned} \tag{5.12}$$

Similarly,

$$|x - x_J| \le \operatorname{diam}(Q_J) + d(Q_J, F) \le 5\operatorname{diam}(Q_J), \tag{5.13}$$

so, combining (5.12) and (5.13), we obtain (5.11). ∎

Recall that if f is a locally integrable function on $\mathbf{R}^n$ and if Q is a cube, then $f_Q = (1/|Q|)\int_Q f(y)\,dy$ denotes the average of f over Q.

Definition 5.4. Let f be locally integrable and let Q be a cube in $\mathbf{R}^n$. The *Hardy-Littlewood maximal function $M_Q f$ of f relative to Q*, or, more briefly, the *Q-maximal function of f*, is defined by

$$(M_Q f)(x) = \sup_{\substack{Q' \subset Q \\ Q' \ni x}} |f|_{Q'} = \sup_{\substack{Q' \subset Q \\ Q' \ni x}} \frac{1}{|Q'|} \int_{Q'} |f(y)|\, dy, \tag{5.14}$$

where the supremum extends over all cubes Q' that contain x and are contained in Q.

Differences $f(y_1) - f(y_2)$ for suitably smooth f can be controlled in terms of a Q-maximal function of the norm $|\nabla f|$ of the gradient of f. The following lemma constitutes a first step in this direction.

Lemma 5.5. *Suppose $f \in C_1$. Let Q_0 and Q_1 be cubes for which $Q_1 \subset Q_0$ and* $\operatorname{diam}(Q_1) = (1/2)\operatorname{diam}(Q_0)$. *Then*

$$|f_{Q_1} - f_{Q_0}| \le c \operatorname{diam}(Q_0) \cdot \min_{x \in Q_0} [M_{3Q_0}(|\nabla f|)](x). \tag{5.15}$$

Proof. We begin with the obvious estimate

$$|f_{Q_0} - f_{Q_1}| \le \frac{1}{|Q_0|} \frac{1}{|Q_1|} \int_{Q_0} \int_{Q_1} |f(s) - f(t)|\, dt\, ds$$

$$= \frac{1}{|Q_0|} \frac{1}{|Q_1|} \int_{Q_0} \int_{Q_1 - s} |f(s) - f(s + h)|\, dh\, ds,$$

where $Q_1 - s$ is the translated cube $\{t - s : t \in Q_1\}$. Using (4.16), Schwarz' inequality, and an interchange in the order of integration, we obtain

$$|f_{Q_0} - f_{Q_1}| \le \frac{1}{|Q_0|} \frac{1}{|Q_1|} \int_0^1 \int_{Q_0} \int_{Q_1 - s} |(\nabla f)(s + \xi h)| |h|\, dh\, ds\, d\xi.$$

Since $Q_1 \subset Q_0$, the translated cube $Q_1 - s$ is contained in the enlarged cube $Q_0^* = 2Q_0 - \operatorname{cen}(Q_0)$, for any $s \in Q_0$. Hence,

$$|f_{Q_0} - f_{Q_1}| \le \frac{\operatorname{diam}(Q_0)}{|Q_0||Q_1|} \int_0^1 \int_{Q_0^*} \int_{Q_0} |(\nabla f)(s + \xi h)|\, ds\, dh\, d\xi$$

$$\le \frac{\operatorname{diam}(Q_0)}{|Q_0||Q_1|} \int_0^1 \int_{Q_0^*} \int_{3Q_0} |(\nabla f)(u)|\, du\, dh\, d\xi$$

since, for any $h \in Q_0^*$, the set $Q_0 + \xi h$ is contained in $3Q_0$. The inner integral is majorized by $|3Q_0|[M_{3Q_0}(|\nabla f|)](y)$ for any choice of y in Q_0. Making this

estimate, performing the other two integrations, and noting that Q_1 has half the diameter of Q_0, we arrive at the desired estimate (5.15) with $c = 12^n$. ■

Theorem 5.6. *Let $f \in C_1$ and let Q be a cube. Then*

$$|f(x) - f_Q| \le c \operatorname{diam}(Q) M_{3Q}(|\nabla f|)(x), \tag{5.16}$$

for all $x \in Q$. In particular, if y_1, $y_2 \in \mathbf{R}^n$, and if Q is a cube of minimal side-length containing y_1 and y_2, then

$$|f(y_1) - f(y_2)| \le c(M_1 + M_2)|y_1 - y_2|, \tag{5.17}$$

where $M_j = M_{3Q}(|\nabla f|)(y_j)$, $(j = 1, 2)$.

Proof. To establish (5.16), first let x be a point of Q. Set $Q_0 = Q$ and construct subcubes Q_j such that each Q_j contains x, each $Q_j \subset Q_{j-1}$, and $\operatorname{diam}(Q_j) = (1/2)\operatorname{diam}(Q_{j-1})$, $(j = 1, 2, \ldots)$. Repeated use of Lemma 5.5 gives

$$|f_{Q_0} - f_{Q_j}| \le \sum_{k=1}^{j} |f_{Q_k} - f_{Q_{k-1}}|$$

$$\le c M_{3Q_{k-1}}(|\nabla f|)(x) \sum_{k=1}^{j} \operatorname{diam}(Q_{k-1})$$

$$\le c M_{3Q_0}(|\nabla f|)(x) \operatorname{diam}(Q_0).$$

But $f_{Q_j} \to f(x)$ as $j \to \infty$, so this establishes (5.16).

Suppose now that Q_0 is a cube of minimal side-length containing both y_1 and y_2. Then, by (5.16),

$$|f(y_1) - f(y_2)| \le |f(y_1) - f_{Q_0}| + |f_{Q_0} - f(y_2)|$$

$$\le c(M_1 + M_2) \operatorname{diam}(Q_0),$$

from which (5.17) follows. ■

We shall use Theorem 5.6 to estimate the difference of Taylor polynomials of a function f expanded about different points y_1 and y_2. If $f \in C_k$, the *Taylor polynomial* we have in mind is the expansion $T_y(x)$ of order $k - 1$ about the point y, which is given by

$$T_y(x) = \sum_{|\nu| \le k-1} \frac{D^\nu f(y)}{\nu!} (x - y)^\nu, \qquad (x \in \mathbf{R}^n). \tag{5.18}$$

We shall use the following notation:

$$\mathscr{D}_k f = \sum_{|\nu|=k} |D^\nu f|, \qquad \mathscr{D} f = \sum_{j=0}^{k} \mathscr{D}_k f = \sum_{|\nu|\le k} |D^\nu f|. \tag{5.19}$$

Theorem 5.7. *If $f \in C_k$ and if $y_1, y_2 \in \mathbf{R}^n$, then*

$$|T_{y_1}(x) - T_{y_2}(x)| \le c\Delta^{k-1} \cdot |y_1 - y_2| \cdot \max_{j=1,2} M(\mathscr{D}_k f)(y_j), \tag{5.20}$$

for all $x \in \mathbf{R}^n$, where

$$\Delta = \max_{j=1,2} |x - y_j|. \tag{5.21}$$

Proof. Let Q be a cube of minimal side-length containing both y_1 and y_2. Considering $T_y(x)$ as a function of y, we may use (5.17) to obtain

$$|T_{y_1}(x) - T_{y_2}(x)| \le c(M_1 + M_2)|y_1 - y_2|, \tag{5.22}$$

where $M_j = M_{3Q}(|\nabla_y T_y(x)|)(y_j)$, $(j = 1, 2)$. For each $i = 1, \ldots, n$, the product rule for differentiation gives

$$\frac{\partial}{\partial y_i} T_y(x) = \sum_{|\nu|\le k-1} \frac{D^\nu D_i f(y)}{\nu!}(x-y)^\nu - \sum_{\substack{|\nu|\le k-1 \\ \nu \ne 0}} \frac{D^\nu f(y)}{\nu!} \nu_i \frac{(x-y)^\nu}{(x_i - y_i)}.$$

The first term on the right is the Taylor polynomial of degree $k-1$ for $D_i f$, whereas the second is the Taylor polynomial of degree $k-2$. Hence,

$$\frac{\partial}{\partial y_i} T_y(x) = \sum_{|\nu|=k-1} \frac{D^\nu D_i f(y)}{\nu!}(x-y)^\nu.$$

But then

$$M_{3Q}(|\nabla_y T_y(x)|)(y_j) \le c \sum_{|\nu|=k-1} \max_{u \in 3Q} |x-u|^{k-1} \cdot \frac{1}{\nu!} \cdot M_{3Q}(\mathscr{D}_k f)(y_j).$$

Using this in conjuction with (5.22), we obtain (5.20). ∎

Corollary 5.8. *If $f \in C_k$, then*

$$|f(x) - T_y(x)| \le c[M(\mathscr{D}_k f)(x) + M(\mathscr{D}_k f)(y)] \cdot |x-y|^k. \tag{5.23}$$

Proof. Apply Theorem 5.7 with $y_1 = x$ and $y_2 = y$. ∎

Corollary 5.9. *If $f \in C_k$, then, for all multi-indices ν of order $|\nu| \le k-1$,*

$$|D^\nu T_{y_1}(x) - D^\nu T_{y_2}(x)| \le c \max_{j=1,2} M(\mathscr{D}_k f)(y_j) \cdot \Delta^{k-1-|\nu|} |y_1 - y_2|, \tag{5.24}$$

where $\Delta = \max_{j=1,2} |x - y_j|$.

Proof. Note that $D^\nu T_y(x)$ is the Taylor polynomial of degree $k-1-|\nu|$ for $D^\nu f$ about the point y. Now apply Theorem 5.7 to $D^\nu f$ with $k' = k - |\nu|$. ∎

In the main problem at hand, for each f in $W_k^1 + W_k^\infty$, we need to construct a decomposition $f = g + h$ with g in W_k^∞. To this end, fix $t > 0$ and set

$$M_0 = M(\mathscr{D} f)^*(t). \tag{5.25}$$

Denote by Ω the open set

$$\Omega = \{x \in \mathbf{R}^n : M(\mathscr{D} f)(x) > M_0\}, \tag{5.26}$$

and by F its complement in $\mathbf{R}^n$. Let $\Psi = \{Q_1, Q_2, \ldots\}$ be a Whitney covering of Ω with dyadic cubes Q_j, and let $\{\phi_1, \phi_2, \ldots\}$ be a corresponding partition of unity constructed as in (5.5) and (5.6). As in (5.9), let x_j be a point of F for which $d(x_j, Q_j) = d(F, Q_j)$, $(j = 1, 2, \ldots)$. We shall denote by $T_j(x)$ the Taylor polynomial $T_{x_j}(x)$ of f of order $k-1$ about the point x_j. Define g on $\mathbf{R}^n$ by

$$g(x) = \begin{cases} f(x), & \text{if } x \in F; \\ \displaystyle\sum_{\operatorname{diam}(Q_j) \le 1} T_j(x)\phi_j(x), & \text{if } x \in \Omega. \end{cases} \tag{5.27}$$

Note that the sum extends over only those j for which Q_j has diameter at most 1.

Theorem 5.10. *There is a constant $c > 0$ such that, for all $f \in C_k \cap W_k^1$, the corresponding function g constructed as in (5.27) satisfies*

$$\|g\|_{W_k^\infty} \le c M_0. \tag{5.28}$$

Proof. The proof proceeds by induction on the order $|\nu|$ of a multi-index ν. Specifically, we assume that $D^\nu g$ is defined a.e. for some $|\nu| \le k-1$, and we show that

$$|D^\nu g(x) - D^\nu g(y)| \le c M_0 |x - y| \qquad \text{a.e.} \tag{5.29}$$

This guarantees that $D_j D^\nu g$ belongs to L^∞, for each $j = 1, \ldots, n$, and hence that the multi-index can be increased by one in each coordinate, thus establishing the inductive step.

To prove (5.29), we suppose that x and y are points at which $D^\nu g$ is defined, and we consider several cases:

Case 1: $x, y \in F$.

We may assume that x and y are not isolated points of F since there are only countably many such points. The derivatives $D^\nu f$ and $D^\nu g$ exist at x, and $g = f$ on a sequence of points converging to x, so $D^\nu f = D^\nu g$ at x. The same is true of the point y. Using (5.17), we have

$$|D^\nu f(x) - D^\nu f(y)| \le c[M(\mathscr{D}f)(x) + M(\mathscr{D}f)(y)] \cdot |x - y|,$$

since $M(|\nabla D^\nu f|) \le M(\mathscr{D}f)$. However, x and y are in F so it follows from the definition (5.26) of the set Ω that both $M(\mathscr{D}f)(x)$ and $M(\mathscr{D}f)(y)$ are majorized by M_0. This establishes (5.29) for x and y in F.

Case 2: $x \in \Omega$, $y \in \partial F$.

Again, we may assume that y is not an isolated point of F, so that $D^\nu g(y) = D^\nu f(y)$. Since $x \in \Omega$, there is an index J for which x belongs to the Whitney cube Q_J.

Suppose first that $\operatorname{diam}(Q_J) > 4$. If j is any index for which $\phi_j(x) \ne 0$, then (5.10) shows that $\operatorname{diam}(Q_j) > 1$. Thus, it follows from the construction of g in (5.27) that $D^\nu g(x) = 0$. Hence, since $y \in F$, we obtain from (5.26) and Lemma 5.1 (iii),

$$\begin{aligned} |D^\nu g(y) - D^\nu g(x)| &= |D^\nu f(y)| \le M_0 \\ &\le cM_0 \operatorname{diam}(Q_J) \le cM_0|x - y|, \end{aligned}$$

which establishes (5.29).

Next, suppose that $1/4 \le \operatorname{diam}(Q_J) \le 4$. Since $y \in F$, we again have $|D^\nu f(y)| \le M_0$. Hence, using (5.27) and Leibniz' rule, we have

$$\begin{aligned} |D^\nu g(y) - D^\nu g(x)| &\le |D^\nu f(y)| + \left|\sum_j{}' D^\nu(T_j\phi_j)(x)\right| \\ &\le M_0 + \sum_j{}' \sum_{0\le\mu\le\nu} \binom{\nu}{\mu} |D^\mu T_j(x)||D^{\nu-\mu}\phi_j(x)|, \qquad (5.30) \end{aligned}$$

where $\sum'$ denotes the sum over those j for which $\operatorname{diam}(Q_j) \le 1$. Note that the sum $\sum'$ can be further restricted to those j for which Q_j touches Q_J by the construction of the ϕ_j and Lemma 5.2(c); in particular, the sum contains at most 12^n terms. For those j, the diameters of the cubes Q_j are comparable to $\operatorname{diam}(Q_J)$ (cf. (5.3)), which, by hypothesis, is bounded above by 4 and below by 1/4. Hence, by (5.8), the terms $|D^{\nu-\mu}\phi_j(x)|$ in (5.30) are bounded by a constant independent of x and y. For the term in (5.30) involving the Taylor

polynomial, we have from (5.18),

$$|D^\mu T_j(x)| = \left|\sum_{|\mu+\gamma|\le k-1} \frac{D^{\mu+\gamma}f(x_j)}{\gamma!}(x - x_j)^\gamma\right| \le cM_0,$$

since the points x_j all belong to F (and hence, by (5.26), yield derivatives of f that are bounded by M_0), and since, as in (5.12), the terms $|x - x_j|$ are majorized by a multiple of $\operatorname{diam}(Q_J) \le 4$. The entire right-hand side of (5.30) is thus bounded by a multiple of M_0. Since

$$|x - y| \ge d(x, F) \ge d(Q_J, F) \ge \operatorname{diam}(Q_J) \ge \frac{1}{4},$$

we arrive at (5.29) again in this case.

Finally, suppose that $\operatorname{diam}(Q_J) < 1/4$. This time, every j with $\phi_j(x) \ne 0$ gives rise to a cube Q_j which touches Q_J and hence has $\operatorname{diam}(Q_j) \le 4\operatorname{diam}(Q_J) \le 1$. The restricted sum $\sum'$ in the preceding paragraph can therefore be replaced by a sum over all indices j. The functions ϕ_j constitute a partition of unity so $\sum \phi_j(x) = 1$ and $\sum D^{\nu-\mu}\phi_j(x) = 0$ if $\mu \ne \nu$. Hence, using Leibniz' rule as before, we obtain

$$\begin{aligned} |D^\nu g(x) - D^\nu g(y)| &= \left|\sum_j \sum_{0\le\mu\le\nu} \binom{\nu}{\mu} D^\mu T_j(x) D^{\nu-\mu}\phi_j(x) - D^\nu f(y)\right| \\ &= \left|\sum_{\substack{0\le\mu\le\nu \\ \mu\ne\nu}} \binom{\nu}{\mu} \sum_j [D^\mu T_j(x) - D^\mu T_J(x)] D^{\nu-\mu}\phi_j(x) \right. \\ &\qquad \left. + \sum_j [D^\nu T_j(x) - D^\nu f(y)]\phi_j(x)\right|. \qquad (5.31) \end{aligned}$$

Using Corollary 5.9, the estimates (5.12) and (5.11) for $|x - x_j|$ and $|x_j - x_J|$, and the estimate (5.8) for the derivatives of ϕ_j, we can majorize the first term on the right of (5.31) by

$$cM_0 \sum_{\substack{0\le\mu\le\nu \\ \mu\ne\nu}} \operatorname{diam}(Q_J)^{k-|\mu|} \operatorname{diam}(Q_J)^{|\mu|-|\nu|} \le cM_0 \operatorname{diam}(Q_J),$$

since $|\nu| \le k - 1$. On the other hand, for the second term on the right of (5.31), we have

$$\begin{aligned} &|D^\nu T_j(x) - D^\nu f(y)| \\ &\quad = \left|\sum_{\substack{|\nu+\gamma|\le k-1 \\ \gamma\ne 0}} \frac{D^{\nu+\gamma}f(x_j)}{\gamma!}(x - x_j)^\gamma + D^\nu f(x_j) - D^\nu f(y)\right|. \qquad (5.32) \end{aligned}$$

Recalling that $|x - x_j| \leq 1$, we can majorize the first term on the right by $cM_0 \operatorname{diam}(Q_J)$ (as in the preceding paragraph). The second term, by virtue of the estimate (5.17) applied to $D^\nu f$, is majorized by $cM_0|y - x_j|$. However, $\operatorname{diam}(Q_J) \leq d(Q_J, F) \leq |x - y|$, so, by (5.12),

$$|y - x_j| \leq |y - x| + |x - x_j| \leq 22|y - x|.$$

The right-hand side of (5.32), hence also the left, is therefore majorized by $cM_0|x - y|$. Putting all of these estimates together with (5.31), we obtain (5.29) again in this case.

Case 3: $x \in \Omega$, $y \in F$.

Select a point $z \in \partial F$ on the line segment connecting x and y such that $d(x, F) = |x - z|$. Then certainly $|x - z| \leq |x - y|$. Using the results established in cases 1 and 2, we have

$$\begin{aligned} |D^\nu g(x) - D^\nu g(y)| &\leq |D^\nu g(x) - D^\nu g(z)| + |D^\nu g(z) - D^\nu g(y)| \\ &\leq cM_0(|x - z| + |z - y|) \leq cM_0|x - y|. \end{aligned}$$

Case 4: $x \in \Omega$, $y \in \Omega$.

It will suffice to show that

$$|D^\nu g| \leq cM_0 \qquad \text{on } \Omega, \text{ for all } \nu \text{ with } |\nu| \leq k, \tag{5.33}$$

since the line segment $[x, y]$ joining x and y either

(a) is wholly contained in Ω, in which case, with $|\nu'| \leq k - 1$,

$$|D^{\nu'} g(x) - D^{\nu'} g(y)| \leq \left| \int_{[x,y]} \nabla D^{\nu'} g \cdot u \right| \leq cM_0|x - y|,$$

or

(b) contains a point, say z, of F, in which case the desired estimate (5.29) is obtained by a simple argument involving the triangle inequality (cf. case 3).

To establish (5.33), let x be an arbitrary point of Ω and suppose as before that Q_J is a Whitney cube containing x. Let ν be a multi-index with $|\nu| \leq k$.

Suppose first that $\operatorname{diam}(Q_J) > 4$. In that case, $D^\nu g(x) = 0$ and (5.33) follows trivially.

Suppose next that $1/4 \leq \operatorname{diam}(Q_J) \leq 4$. Then

$$|D^\nu g(x)| = \left| \sum_j{}' \sum_{\mu \leq \nu} \binom{\nu}{\mu} D^\mu T_j(x) D^{\nu - \mu} \phi_j(x) \right| \leq cM_0$$

(where $\sum'$ has its previous meaning), since

$$|D^\mu T_j(x)| \leq \sum_{|\gamma + \mu| \leq k-1} \frac{|D^{\gamma + \mu} f(x_j)|}{\gamma!} |x - x_j|^{|\gamma|} \leq cM_0$$

and $|D^{\nu - \mu} \phi_j| \leq c \operatorname{diam}(Q_J)^{|\mu| - |\nu|} \leq \text{const}.$

Finally, suppose that $\operatorname{diam}(Q_J) < 1/4$. Once again, the restricted sum $\sum'$ can be replaced by a sum over all j. We have

$$D^\nu g(x) = \sum_{\substack{0 \leq \mu \leq \nu \\ \mu \neq \nu}} \binom{\nu}{\mu} \sum_j [D^\mu T_j(x) - D^\mu T_J(x)] D^{\nu - \mu} \phi_j(x) + \sum_j D^\nu T_j(x) \phi_j(x).$$

The first term on the right is estimated as before using Corollary 5.9 to obtain the upper bound

$$c \sum_{\substack{0 \leq \mu \leq \nu \\ \mu \neq \nu}} M_0 \operatorname{diam}(Q_J)^{k - |\mu|} \operatorname{diam}(Q_J)^{|\mu| - |\nu|} \leq cM_0.$$

The second term is majorized by

$$c \sum_j \sum_{|\gamma + \nu| \leq k-1} |D^{\gamma + \nu} f(x_j)| \phi_j(x) \leq cM_0,$$

since $|x - x_j| \leq 21 \operatorname{diam}(Q_J) \leq 6$, and $\sum_j \phi_j = 1$ on Ω. This completes the proof. ∎

Theorem 5.11. *There are constants c_1 and c_2 such that for all $f \in W_k^1$ and all $t > 0$,*

$$c_1 \int_0^t (\mathscr{D} f)^*(s)\, ds \leq K(f, t; W_k^1, W_k^\infty) \leq c_2 \int_0^t (\mathscr{D} f)^*(s)\, ds. \tag{5.34}$$

Proof. The first inequality follows immediately from Theorem II.3.4 and the definitions involved. For the second, suppose first that f also belongs to C_k. Define the constant M_0 as in (5.25), the set Ω as in (5.26), and construct the function g as in (5.27). Then Theorem 5.10 shows that the estimate (5.28) holds. Since $g = f$ on F, and $D^\nu f = D^\nu g$ a.e. on F, we have

$$\|f - g\|_{W_k^1} + t\|g\|_{W_k^\infty} \leq \int_\Omega |\mathscr{D}(f - g)| + ctM_0.$$

Now $|\Omega| \leq t$ and $M_0 \leq c(\mathscr{D} f)^{**}(t)$ (cf. Theorem III.3.8), so the right-hand side is majorized by

$$\int_\Omega |\mathscr{D} f| + 2ctM_0 \leq c \int_0^t (\mathscr{D} f)^*(s)\, ds,$$

and this establishes the second of the inequalities in (5.34).

For an arbitrary f in W_k^1, we approximate f with a function f_1 in C_k to error at most $\varepsilon > 0$. Then, applying the result just established to

$f_1 = (f_1 - g_1) + g_1$, we obtain

$$\|f - g_1\|_{W_k^1} + t\|g_1\|_{W_k^\infty} \leq \|f - f_1\|_{W_k^1} + c\int_0^t (\mathscr{D}f_1)^*$$

$$\leq (1 + c)\|f - f_1\|_{W_k^1} + c\int_0^t (\mathscr{D}f)^*$$

$$\leq (1 + c)\varepsilon + c\int_0^t (\mathscr{D}f)^*.$$

Since $\varepsilon > 0$ is arbitrary, this establishes the desired result. ∎

Now we are in a position to establish the main result of this section, that is, a description of the K-functional for the couple (W_k^1, W_k^∞) of Sobolev spaces. All that remains is to show that the preceding theorem holds for all f in the sum of the two spaces concerned.

Theorem 5.12 (R. A. DeVore & K. Scherer). *The estimates* (5.34) *hold for all f in $W_k^1 + W_k^\infty$, that is,*

$$K(f, t; W_k^1, W_k^\infty) \sim \int_0^t (\mathscr{D}f)^*(s)\, ds, \qquad (t > 0). \tag{5.35}$$

Proof. That the left-hand side dominates the right follows once again from the definitions involved. Thus, only the reverse inequality need be established. As before, define M_0 and Ω by (5.25) and (5.26), so that $|\Omega| \leq t$. Let $t_j = 2^{-j}t$, and choose a sequence $\varepsilon_j \downarrow 0$ (with $\varepsilon_1 = t$) such that

$$\int_0^{\varepsilon_j} (\mathscr{D}f)^* \leq t_j M_0, \qquad (j = 2, 3, \ldots). \tag{5.36}$$

Next choose a sequence $r_j \uparrow \infty$ with $r_0 = 0$, $r_{j+1} > r_j + 1$, and, if B_j denotes the ball $B_j = \{x \in \mathbf{R}^n : |x| < r_{j-1}\}$, then

$$|\Omega \cap B_j^c| \leq \varepsilon_j, \qquad (j = 1, 2, \ldots). \tag{5.37}$$

Let A_j be the annulus $A_j = \{x : r_{j-1} \leq |x| < r_{j+1}\}$, which, clearly, is disjoint from B_j. Let $\{\psi_j\}$ be a partition of unity for the annuli A_j: thus,

$$\begin{aligned} &\text{(a)} \quad \sum_{j=1}^\infty \psi_j = 1; \\ &\text{(b)} \quad \operatorname{supp}(\psi_j) \subset A_j; \\ &\text{(c)} \quad \|\psi_j\|_{C_k} \leq \text{const.} \end{aligned} \tag{5.38}$$

Set $f_j = f\psi_j$, so each $f_j \in W_k^1$ and $f = \sum_j f_j$. In particular, by Theorem 5.11,

$$K(f_j, t_j; W_k^1, W_k^\infty) \leq c\int_0^{t_j} (\mathscr{D}f_j)^*. \tag{5.39}$$

Since f_j is supported on the set A_j of finite measure, Theorem II.2.6 enables us to find a set $E_j \subset A_j$ such that $|E_j| \leq t_j$ and

$$\int_0^{t_j} (\mathscr{D}f_j)^* = \int_{E_j} \mathscr{D}f_j. \tag{5.40}$$

By (5.38)(c), there is a constant c such that $\mathscr{D}f_j \leq c\mathscr{D}f$. Hence, since $E_j \subset B_j^c$ and $\mathscr{D}f \leq M_0$ on Ω^c, we have from (5.36) and (5.37),

$$\begin{aligned} \int_{E_j} \mathscr{D}f_j &\leq \int_{E_j \cap \Omega} \mathscr{D}f + \int_{E_j \cap \Omega^c} \mathscr{D}f \leq \int_{B_j^c \cap \Omega} \mathscr{D}f + t_j M_0 \\ &\leq \int_0^{\varepsilon_j} (\mathscr{D}f)^* + t_j M_0 \leq 2t_j M_0. \end{aligned} \tag{5.41}$$

The inequalities (5.39), (5.40), and (5.41) together imply that

$$K(f_j, t_j) \leq ct_j M_0, \qquad (j = 2, 3, \ldots). \tag{5.42}$$

As in the proof of Theorem 5.11, we may construct functions g_j in W_k^∞ with $\operatorname{supp}(g_j) \subset A_j$ and

$$\|f_j - g_j\|_{W_k^1} + t_j\|g_j\|_{W_k^\infty} \leq 2K(f_j, t_j). \tag{5.43}$$

Note, in particular, from (5.42) and (5.43) that

$$\|f_j - g_j\|_{W_k^1} \leq ct_j M_0, \tag{5.44}$$

and

$$\|g_j\|_{W_k^\infty} \leq cM_0. \tag{5.45}$$

Define g in W_k^∞ by

$$g = \sum_{j \geq 2} g_j; \tag{5.46}$$

the sum converges because $\sum_j \chi_{A_j} \leq 2$, and we see from (5.45) that

$$\|g\|_{W_k^\infty} \leq cM_0. \tag{5.47}$$

Writing

$$f = \sum_{j=1}^\infty f_j = f_1 + \sum_{j=2}^\infty (f_j - g_j) + g,$$

we have

$$K(f,t) \le K(f_1,t) + \sum_{j=2}^{\infty} \|f_j - g_j\|_{W_k^1} + t\|g\|_{W_k^\infty}.$$

Using (5.34), (5.44), (5.45), and (5.47), we thus obtain

$$K(f,t) \le c\int_0^t (\mathscr{D}f_1)^* + \sum_{j=2}^{\infty} ct_j M_0 + ctM_0$$

$$\le c\int_0^t (\mathscr{D}f)^* + 2ctM_0 \le 3c\int_0^t (\mathscr{D}f)^*,$$

and this completes the proof. ∎

It is now elementary to identify the Sobolev spaces W_k^p as interpolation spaces between W_k^1 and W_k^∞:

Corollary 5.13. *Let k be a positive integer and $1 < p < \infty$. Let $\theta = 1 - 1/p$. Then, up to equivalence of norms, the Sobolev space W_k^p coincides with the interpolation space*

$$W_k^p = (W_k^1, W_k^\infty)_{\theta,p} \tag{5.48}$$

for the couple (W_k^1, W_k^∞).

Finally, by using the reiteration theorem in conjunction with Theorem 5.12 and Corollary 5.13, one can also describe the K-functional and the corresponding interpolation spaces for any of the couples (W_k^p, W_k^q), where $1 \le p < q \le \infty$.

6. RE H^1 AND BMO

In this section, we consider the Hardy space H^1 and the space BMO of functions of bounded mean oscillation. Several equivalent characterizations of H^1 are obtained, in terms of the Hilbert transform, maximal functions, and the atomic decomposition. The K-functional for the pair (H^1, L^∞) is described in terms of the decreasing rearrangement of the nontangential maximal function, thus providing a description of the K-interpolation spaces between H^1 and L^∞. The section concludes with a proof of the duality between Re H^1 and BMO.

We have confined the discussion to the one-variable setting, where a number of simplifications enable us to present the elements of the theory in a more concise fashion than is possible in several variables. The multi-variable setting is discussed in the Exercises.

Let $\mathbf{U} = \{x + iy : x \in \mathbf{R}, y > 0\}$ denote the upper half plane. The Hardy space $H^p(\mathbf{U})$, $1 \le p \le \infty$, is the Banach space of analytic functions F on $\mathbf{U}$ for which

$$\|F\|_{H^p(\mathbf{U})} = \begin{cases} \sup_{y>0} \left\{ \int_{-\infty}^{\infty} |F(x+iy)|^p \, dx \right\}^{1/p}, & (1 \le p < \infty) \\ \sup_{z \in \mathbf{U}} |F(z)|, & (p = \infty) \end{cases} \tag{6.1}$$

is finite. Recall from Exercise III.8 that the Poisson kernel P_y is given by

$$P_y(t) = \frac{1}{\pi} \frac{y}{t^2 + y^2}, \qquad (y > 0, -\infty < t < \infty).$$

Moreover, if f belongs to $(L^1 + L^\infty)(\mathbf{R})$, then $u(x,y) = P_y * f(x)$ extends f harmonically into $\mathbf{U}$. Thus, u satisfies (cf. Exercise III.9)

$$\Delta u \equiv 0 \text{ in } \mathbf{U}, \qquad \lim_{t+iy \to x} u(t,y) = f(x) \qquad \text{a.e.}$$

Here, the values $t + iy$ converge to x within the cone $\Gamma_x = \{t + iy : |x - t| < y\} \subset \mathbf{U}$. The nontangential maximal function Nf of a function f in $L^1(\mathbf{R})$ is defined by

$$Nf(x) = \sup_{t+iy \in \Gamma_x} |(P_y * f)(t)|, \qquad (x \in \mathbf{R}). \tag{6.2}$$

Lemma 6.1. *Suppose $1 \le p \le \infty$ and $F(x+iy)$ belongs to $H^p(\mathbf{U})$. Then there is a function $F(x)$ in $L^p(\mathbf{R})$ such that*

$$F(x+iy) = P_y * F(x), \qquad (x + iy \in \mathbf{U}). \tag{6.3}$$

Proof. We may assume that F is normalized to have H^p-norm equal to one. Suppose first that $1 < p \le \infty$, so the unit ball in L^p is sequentially weak*-compact. If $y_n \downarrow 0$, then the functions $F_n(x) = F(x + iy_n)$ are in the unit ball, and so some subsequence $F_{n(j)}$ converges weak* to an L^p-function F. The Poisson kernel P_y belongs to $L^1 \cap L^\infty \subset L^{p'}$ so, for each $x \in \mathbf{R}$,

$$\begin{aligned} F * P_y(x) &= \int_{\mathbf{R}} F(t) P_y(x-t) \, dt \\ &= \lim_{j \to \infty} \int_{\mathbf{R}} F_{n(j)}(t) P_y(x-t) \, dt \\ &= \lim_{j \to \infty} P_y * F_{n(j)}(x). \end{aligned} \tag{6.4}$$

But $P_y * F_{n(j)}(x)$ is the unique harmonic extension of $F(x + iy_{n(j)})$ to the set

$\{x + iy : y \geq y_{n(j)}\}$ and therefore $P_y * F_{n(j)}(x) = F(x + i(y + y_{n(j)}))$. Hence, with (6.4), we see that (6.3) holds in the case $p > 1$. If F belongs to $H^1(\mathbf{U})$, then (cf. the appendix) F may be factored in the form $F = BG$, where $B(z)$ is the Blaschke product associated with F and G is a zero-free analytic function with the same H^1-norm as F. The functions $G_1 = B\sqrt{G}$ and $G_2 = \sqrt{G}$ are analytic in $\mathbf{U}$ and satisfy

$$\|G_j\|_{H^2} = \|G\|_{H^1}^{1/2}, \qquad (j = 1, 2).$$

By the first part of the proof, there exist functions G_j in L^2 such that $G_j(x + iy) = P_y * G_j(x)$. The function $F = G_1 G_2$ belongs to L^1 and then (6.3) follows since the harmonic extension of F is unique. This completes the proof. ■

We have used the same notation F for the analytic function $F(x + iy)$, defined on $\mathbf{U}$, and its *boundary function* $F(x)$, defined a.e. on $\mathbf{R}$. It will be clear from the context which usage of F is intended. For such functions, the nontangential maximal function NF of $F(x)$, defined by (6.2), is given by

$$NF(x) = \sup_{t+iy\in\Gamma_x} |F * P_y(t)| = \sup_{z\in\Gamma_x} |F(z)|, \qquad (x \in \mathbf{R}).$$

Theorem 6.2 (G. H. Hardy & J. E. Littlewood). *There is a constant $c > 0$ such that, for all F in $H^1(\mathbf{U})$,*

$$\|NF\|_{L^1(\mathbf{R})} \leq c\|F\|_{H^1(\mathbf{U})}. \tag{6.5}$$

Proof. Suppose F belongs to $H^1(\mathbf{U})$. Then, with G_1 and G_2 as in the preceding proof, we have

$$N(F) \leq N(G_1)N(G_2) \leq cM(G_1)M(G_2) \tag{6.6}$$

(cf. Exercise III.9(c)), where M is the Hardy-Littlewood maximal operator. Integrating (6.6) and using the Cauchy-Schwarz inequality, we obtain

$$\int_{\mathbf{R}} |NF(x)|\,dx \leq c\|MG_1\|_{L^2(\mathbf{R})}\|MG_2\|_{L^2(\mathbf{R})}$$
$$\leq c\|G_1\|_{L^2(\mathbf{R})}\|G_2\|_{L^2(\mathbf{R})}, \tag{6.7}$$

since M is a bounded operator on $L^2(\mathbf{R})$. But (cf. Exercise III.9(d))

$$\|G_j\|_{L^2(\mathbf{R})} \leq \|G_j\|_{H^2(\mathbf{U})} = \|F\|_{H^1(\mathbf{U})}^{1/2},$$

and combining this with (6.7) we obtain (6.5). ■

Corollary 6.3. *Suppose $1 \leq p < \infty$. If $F(x + iy)$ is an H^p-function with*

boundary function $F(x)$, then

$$\lim_{y\downarrow 0} \|F(\cdot + iy) - F(\cdot)\|_{L^p(\mathbf{R})} = 0.$$

Proof. The maximal function NF dominates $|F(\cdot)|$ and $|F(\cdot + iy)|$, $(y > 0)$, and $F(x + iy) \to F(x)$ a.e. as $y \downarrow 0$. The result therefore follows from the dominated convergence theorem. ■

We shall see in Theorem 6.5 below that $H^1(\mathbf{U})$ is a complex Banach space. As a Banach space over the real scalars, it is therefore isometrically isomorphic to the space $\operatorname{Re}(H^1)$ of the real parts f of the boundary functions F of functions in $H^1(\mathbf{U})$, normed by

$$\|f\|_{\operatorname{Re}(H^1)} = \|F\|_{H^1(\mathbf{U})}, \qquad (f = \operatorname{Re} F). \tag{6.8}$$

If F_1 and F_2 are two functions in $H^1(\mathbf{U})$ whose boundary functions both have real part equal to f, then the Cauchy-Riemann equations show that F_1 and F_2 can differ only by an imaginary constant. This constant must be zero, however, because $F_1 - F_2$ belongs to $H^1(\mathbf{U})$. Hence, $F_1 = F_2$ and we see that the norm in (6.8) is well-defined.

Definition 6.4. The *Hardy space for $L^1(\mathbf{R})$*, denoted by $H(L^1)$, consists of those real-valued functions f in $L^1(\mathbf{R})$ whose Hilbert transforms Hf also belong to $L^1(\mathbf{R})$. The norm is given by

$$\|f\|_{H(L^1)} = \|f\|_{L^1(\mathbf{R})} + \|Hf\|_{L^1(\mathbf{R})}. \tag{6.9}$$

Theorem 6.5. *The spaces $H(L^1)$ and $\operatorname{Re}(H^1)$ are isomorphic Banach spaces (over the real scalars) with equivalent norms:*

$$\|f\|_{\operatorname{Re}(H^1)} \leq \|f\|_{H(L^1)} \leq 2\|f\|_{\operatorname{Re}(H^1)}. \tag{6.10}$$

Proof. We show first that $H(L^1)$ is complete. Suppose (f_n) is Cauchy in the norm (6.9). Then there are integrable functions f and g such that $f_n \to f$ and $Hf_n \to g$ in L^1. Theorem III.4.8 shows that the Hilbert transform of an integrable function exists a.e., and in fact both the Hilbert transform and maximal Hilbert transform are of weak type $(1, 1)$. It follows that Hf_n converges to Hf in measure on subsets of $\mathbf{R}$ of finite measure. Hence g coincides with Hf a.e., and it follows that Cauchy sequences in $H(L^1)$ converge. Hence, $H(L^1)$ is complete.

Next (cf. Exercises III.9,11), observe that the linear map

$$f \to F(x, y) = P_y * (f + iHf)(x) = (P_y + iQ_y) * f$$

is bounded from $H(L^1)$ into $H^1(\mathbf{U})$. Indeed, F is analytic because the harmonic functions $u = P_y * f$ and $v = P_y * Hf = Q_y * f$ satisfy the Cauchy-Riemann equations in $\mathbf{U}$. Moreover,

$$\|F\|_{H^1(\mathbf{U})} \le \int |F|\,dx \le \|f\|_{H(L^1)}. \tag{6.11}$$

On the other hand, the linear map on $H^1(\mathbf{U})$ which takes F to the real part f of its boundary function is bounded into $H(L^1)$:

$$\|f\|_{H(L^1)} = \|f\|_{L^1} + \|Hf\|_{L^1} \le 2\|F\|_{H^1(\mathbf{U})}.$$

The two mappings are mutually inverse, so the last estimate and (6.11), together with the definition (6.8) of the norm in $\operatorname{Re} H^1$, show that $\operatorname{Re} H^1$ and $H(L^1)$ are isomorphic as real Banach spaces and that their norms satisfy the inequalities (6.10). ■

We next establish some additional properties of the Hilbert transform. If $f \in L^p, g \in L^q, (1 < p < \infty, 1/p + 1/q = 1)$, then

$$\int_{\mathbf{R}} gHf\,dx = -\int_{\mathbf{R}} fHg\,dx. \tag{6.12}$$

This follows from the identities

$$\int g(P_y * Hf) = \int g(Q_y * f) = -\int (Q_y * g)f = -\int (P_y * Hg)f,$$

together with two applications of the dominated convergence theorem (with dominating functions $|g|N(Hf)$ and $|f|N(Hg)$).

Let F denote the analytic function obtained from f, that is, $F = P_y * (f + iHf)$, and let G denote the corresponding function for g. Then (6.12) is equivalent to

$$\int_{\mathbf{R}} \operatorname{Im}(FG)\,dx = 0, \qquad (F \in H^p, G \in H^q), \tag{6.13}$$

valid for $1 < p < \infty$ and $1/p + 1/q = 1$.

An arbitrary H^1-function may be written as a product of two H^2-functions. Hence, by applying (6.13), we deduce that the integral of the imaginary part of its boundary function vanishes. By considering the H^1-function iF instead of F, we see that the integral of the real part of F must also vanish. Hence,

$$\int_{\mathbf{R}} F\,dx = 0, \qquad (F \in H^1(\mathbf{U})). \tag{6.14}$$

The domain of definition of the Hilbert transform may be extended as follows to include the essentially bounded functions.

Definition 6.6. If $f \in L^\infty(\mathbf{R})$, let

$$\tilde{H}f(x) = \lim_{\varepsilon \to 0+} \frac{1}{\pi} \int_{-\infty}^{\infty} [k_\varepsilon(x - t) + k_1(t)] f(t)\,dt, \tag{6.15}$$

where $k_\varepsilon(t) = (1/t)\chi_{(\varepsilon,\infty)}(|t|)$.

It is clear that $\tilde{H}f$ exists a.e. and is square integrable on each interval $I = (a, b)$. Set $J = (a - 1, b + 1)$. If $x \in I$, then we may write

$$\tilde{H}f(x) = \lim_{\varepsilon \to 0+} \frac{1}{\pi} \int_{\varepsilon < |x - t| < 1} (f\chi_J)(t) \frac{dt}{x - t} + \frac{1}{\pi} \int_{-\infty}^{\infty} \kappa(x, t) f(t)\,dt \tag{6.16}$$

where $\kappa(x, t) = k_1(x - t) + k_1(t)$. Since the maximal Hilbert transform is bounded on L^2 (cf. Theorem III.4.7 together with, for example, Theorem III.5.16) and $f\chi_J$ belongs to L^2, the first term on the right hand side of (6.16) exists a.e. and belongs to $L^2(I)$. For the second term we use the estimate

$$|\kappa(x, t)| \le \begin{cases} 1, & |t| \le 2|x| + 1, \\ \dfrac{2|x|}{t^2}, & |t| > 2|x| + 1, \end{cases} \tag{6.17}$$

which follows from the fact that $|1/(x - t) + 1/t| \le 2|x|/t^2$ when $|t| > 2|x|$. Then,

$$\left| \int_{-\infty}^{\infty} \kappa(x, t) f(t)\,dt \right| \le 4(|x| + 1)\|f\|_{L^\infty}.$$

The last term, however, is bounded on I and so the second term in (6.16) also belongs to $L^2(I)$.

Next, if f belongs to $L^p \cap L^\infty$ for some finite p, then

$$\tilde{H}f(x) = Hf(x) + c(f), \tag{6.18}$$

where

$$c(f) = \frac{1}{\pi} \int_{1 < |t|} f(t) \frac{dt}{t}$$

is a bounded linear functional on L^p with

$$|c(f)| \le \frac{1}{\pi} \left(\frac{2p}{q}\right)^{1/q} \|f\|_{L^p}, \qquad \left(\frac{1}{q} = 1 - \frac{1}{p}\right).$$

The Hilbert transforms $\tilde{H}$ and H therefore agree, modulo constants, on the intersections of their domains. We may thus extend the definition of the

Hilbert transform to all of $L^1 + L^\infty$ in the following way. Let $f \in L^1 + L^\infty$ and let $f = g + b$ be a representation of f with $g \in L^1$ and $b \in L^\infty$. Set

$$\tilde{H}f = Hg + c(g) + \tilde{H}b, \qquad (f \in L^1 + L^\infty).$$

It follows from (6.18) that $\tilde{H}f$ is well-defined and does not depend on the particular decomposition $f = b + g$ of f.

Definition 6.7. Let g be a locally integrable function on $\mathbf{R}$. Then g is said to have *bounded mean oscillation* ($g \in \mathrm{BMO}$) if the seminorm given by

$$\|g\|_* = \sup_I \frac{1}{|I|} \int_I |g(x) - g_I|\,dx \tag{6.19}$$

is finite. Here, I denotes any finite subinterval of $\mathbf{R}$ and g_I is the average of g over I.

It is clear that BMO is a linear space and that if $\|g\|_*$ vanishes, then g is constant. Also, if α is an arbitrary constant, then $(g-\alpha)_I = g_I - \alpha$ and $\int_I |g - g_I| \le 2\int_I |g-\alpha|$. Hence,

$$\frac{1}{2}\|g\|_* \le \sup_I \left\{ \frac{1}{|I|} \inf_\alpha \int_I |g(x) - \alpha|\,dx \right\} \le \|g\|_*. \tag{6.20}$$

Lemma 6.8 (S. Spanne; E. M. Stein). *If f belongs to $L^\infty(\mathbf{R})$, then $\tilde{H}f$ belongs to BMO and*

$$\|\tilde{H}f\|_* \le 8\|f\|_{L^\infty}. \tag{6.21}$$

Proof. Suppose $f \in L^\infty$. Fix an interval I and let x_0 denote its center. Let $J = 2I$ be the interval with center x_0 and with twice the length of I. Write $f = b + g$ with $g = f\chi_J$ and $b = f\chi_{J^c}$. Then $g \in L^2$ so, with $\alpha_1 = c(g)$, identity (6.18) and the fact that H is an isometry on L^2 give

$$\begin{aligned}\int_I |\tilde{H}g - \alpha_1| &\le |I|^{1/2}\|Hg\|_{L^2} \\ &\le |I|^{1/2}\|g\|_{L^2} \le 2^{1/2}|I|\,\|f\|_{L^\infty}.\end{aligned} \tag{6.22}$$

To estimate $\tilde{H}b$, notice that for $x \in I$,

$$\begin{aligned}|\tilde{H}b(x) - \tilde{H}b(x_0)| &= \frac{1}{\pi}\left| \int_{J^c} f(t)\left[\frac{1}{x-t} - \frac{1}{x_0 - t}\right] dt \right| \\ &\le \frac{1}{\pi}\|f\|_{L^\infty}|x - x_0| \int_{|u|>|I|} \frac{du}{|u|\cdot|u-(x-x_0)|},\end{aligned}$$

where we have made the change of variable $u = t - x_0$. But $x \in I$ and $|u| > |I|$, so $|u - (x - x_0)| \ge |u| - |I|/2 \ge |u|/2$. Hence

$$|\tilde{H}b(x) - \tilde{H}b(x_0)| \le \frac{2}{\pi}|I|\cdot\|f\|_{L^\infty} \int_{|I|}^\infty \frac{du}{u^2} = \frac{2}{\pi}\|f\|_{L^\infty}. \tag{6.23}$$

Combining (6.22) and (6.23), we see with $\alpha = \tilde{H}b(x_0) + \alpha_1$ that

$$\int_I |\tilde{H}f(x) - \alpha|\,dx \le 4|I|\cdot\|f\|_{L^\infty}.$$

Hence, using (6.20) we obtain (6.21) with constant $c = 8$. ■

Definition 6.9. An *atom* $a(x)$ is a real-valued function on $\mathbf{R}$ for which there exists an interval I so that

$$|a| \le |I|^{-1}\chi_I \quad \text{a.e.,} \qquad \int_I a(x)\,dx = 0. \tag{6.24}$$

The *atomic Hardy space* H^1_{at} consists of all real-valued functions f on $\mathbf{R}$ for which there exist atoms a_j and coefficients α_j, $(j = 1, 2, \ldots)$, such that

$$f(x) = \sum_j \alpha_j a_j(x) \qquad \text{a.e.}$$

and $\sum_j |\alpha_j| < \infty$. The norm is defined by

$$\|f\|_{H^1_{\mathrm{at}}} = \inf_{f = \sum_j \alpha_j a_j} \sum_j |\alpha_j|, \tag{6.25}$$

where the infimum is taken over all such atomic representations of f.

Lemma 6.10. *There is a constant $c_0 > 0$ such that*

$$\|Ha\|_{L^1} \le c_0 \tag{6.26}$$

for all atoms a.

Proof. Let $a(x)$ be an atom supported in an interval I as in (6.24). We shall use the fact that $L^1 \cap L^\infty$ is a norming subspace of the dual of L^1, that is,

$$\|\phi\|_{L^1} = \sup\{|\textstyle\int \phi g\,dx| : g \in L^1 \cap L^\infty, \|g\|_{L^\infty} \le 1\}.$$

Let $g \in L^1 \cap L^\infty$ with $\|g\|_\infty \le 1$. Then, using (6.12), we have

$$|\textstyle\int (Ha)g\,dx| = |\int aHg\,dx| = |\int a\tilde{H}g\,dx|, \tag{6.27}$$

the last equation via (6.18) and the fact that atoms have mean value zero. But by (6.24),

$$|\textstyle\int a\tilde{H}g\,dx| = \left|\displaystyle\int_I a[\tilde{H}g - (\tilde{H}g)_I]\,dx\right| \le \frac{1}{|I|}\int_I |\tilde{H}g - (\tilde{H}g)_I|\,dx \le \|\tilde{H}g\|_*. \tag{6.28}$$

Lemma 6.8 shows that the right-hand side of (6.28) is bounded by $8\|g\|_\infty \le 8$. Hence, combining this with (6.27), we obtain the desired conclusion (6.26) with $c_0 = 8$. ∎

Corollary 6.11. *If f belongs to the atomic Hardy space H^1_{at}, then f belongs to* $\mathrm{Re}(H^1)$ *and*

$$\|f\|_{\mathrm{Re}(H^1)} \le c_0 \|f\|_{H^1_{\mathrm{at}}}. \tag{6.29}$$

Proof. Let $f(x) = \sum_j \alpha_j a_j(x)$ be an atomic decomposition of f. By Lemma 6.10, the sequence $\{\sum_1^N \alpha_j H(a_j)\}_{N=1}^\infty$ is Cauchy in L^1 and

$$\|Hf\|_{L^1} \le c_0 \sum_j |\alpha_j|. \tag{6.30}$$

But (6.24) shows that atoms lie in the unit ball of L^1. Hence,

$$\|f\|_{L^1} \le \sum_j |\alpha_j|. \tag{6.31}$$

Now taking the infimum in (6.30) and (6.31) over all such representations of f, we obtain the desired result

$$\|f\|_{L^1} + \|Hf\|_{L^1} \le (c_o + 1)\|f\|_{H^1_{\mathrm{at}}}. \quad ∎$$

The mean value property for harmonic functions asserts that the average of the function over the perimeter of a circle is equal to its value at the center. The following lemma is a variant of that result for squares.

Lemma 6.12. *A function harmonic on an open square and continuous on the closure has average over the perimeter equal to its average over the union of the two diagonals.*

Proof. Let S be an arbitrary square region in the plane. Without loss of generality, we shall assume that S has its sides parallel to the coordinate axes. By a simple dilation argument, we may assume that the harmonic function u and its harmonic conjugate v are continuous on the closure of S.

The square S is composed of four congruent right triangles with common vertex at the center of S. Let T denote the bottom triangle and label its left, bottom, and right edges by L, B, and R respectively. Applying Cauchy's theorem to the analytic function $u + iv$ on T and taking real parts of the integrals, we obtain

$$0 = \int_{\partial T} u\,dx - v\,dy = \int_B u - 2^{-1/2}\left\{ \int_{R\cup L} u + \int_R v - \int_L v \right\}, \tag{6.32}$$

where the integrals on the right are taken with respect to arc length. Dividing by the edge length of S, we see that the average of u over B equals the average over the two sides of T plus a certain signed average of v over those same two sides. Proceeding in the same way for the upper triangle, but using the imaginary part of the integrals for the left and right triangles, we see by adding these equations that the terms involving v cancel, leaving us with the desired result. ∎

The lemma provides us with a relatively straightforward proof of the following key result.

Theorem 6.13. *If f and Nf are integrable then f belongs to atomic H^1 and*

$$\|f\|_{H^1_{\mathrm{at}}} \le 42 \|Nf\|_{L^1}. \tag{6.33}$$

Proof. The Poisson integral $u = f * P_y$ of f is continuous in $\mathbf{U}$ so, for each integer k, the set $E_k = \{x : Nf(x) > 2^k\}$ is open in $\mathbf{R}$. In order to simplify notation, let $I(f) = f_I$ denote the average of f over an interval I. Let $F_k = \mathbf{R} \setminus E_k$, and let $\mathscr{C}_k$ be the collection of disjoint components of E_k. Then we may decompose f as a sum $f = g_k + b_k$, where

$$b_k = \sum_{I\in\mathscr{C}_k} [f - I(f)]\chi_I, \qquad g_k = f\chi_{F_k} + \sum_{I\in\mathscr{C}_k} I(f)\chi_I. \tag{6.34}$$

We claim that $|g_k| \le 7 \cdot 2^k$ a.e.. This estimate holds on F_k since $Nf \le 2^k$ there and $|f| \le Nf$ a.e. On the remaining set E_k, we show, for each fixed interval $I \in \mathscr{C}_k$, that

$$|I(f)| \le 7 \cdot 2^k. \tag{6.35}$$

Let S_ε be the open square $I \times (\varepsilon, |I| + \varepsilon)$ in $\mathbf{U}$. Applying Lemma 6.12 to S_ε and letting $\varepsilon \downarrow 0$, we see that $I(f)$ is equal to four times the average of u over the union of the two main diagonals of $S = I \times (0, |I|)$ less the sum of its averages over the three remaining sides. But the endpoints of I belong to F_k, so the diagonals, sides, and top of S all belong to the set $\Gamma = \{(z, y) \in \Gamma_x : x \in F_k\}$. Hence, the definitions of F_k and Nf show that u is bounded by 2^k on Γ, and this establishes (6.35).

The atoms for f are defined by

$$\begin{aligned} a_I &= \alpha_I^{-1}(b_k - b_{k+1})\chi_I, \\ \alpha_I &= 21 \cdot 2^k |I|, \end{aligned} \tag{6.36}$$

for each $I \in \mathscr{C}_k$ and all integers k. Since Nf is integrable, we have $b_N \to 0$ a.e. as

$N \to \infty$ and $b_M \to f$ a.e. as $M \to -\infty$. Hence,

$$\lim_{M,N\to\infty} \sum_{k=-M}^{N} (b_k - b_{k+1}) = f \qquad \text{a.e.}$$

But $g_{k+1} - g_k = b_k - b_{k+1}$ and b_{k+1} is supported in $E_{k+1} \subset E_k$, so

$$f = \sum_k \sum_{I\in\mathscr{C}_k} \alpha_I a_I.$$

Each a_I is supported in I and the estimate $\|a_I\|_\infty \le |I|^{-1}$ follows from the L^∞-estimate for the g_k:

$$\|b_k - b_{k+1}\|_\infty = \|g_{k+1} - g_k\|_\infty \le 7(2^{k+1} + 2^k) = 21 \cdot 2^k.$$

Furthermore, writing a_I in the form

$$a_I = \alpha_I^{-1}\{[f - I(f)]\chi_I - \sum_{\substack{J\subset I\\ J\in\mathscr{C}_{k+1}}} [f - J(f)]\chi_J\},$$

we see that a_I has mean value zero. Hence, each a_I is an atom.

To establish inequality (6.33) (subject to a relabeling of the index set), we observe that

$$\begin{aligned}\sum_k \sum_{I\in\mathscr{C}_k} |\alpha_I| &= 21\sum_k 2^k \sum_{I\in\mathscr{C}_k} |I| = 21\sum_k 2^k|E_k| \\ &= 21\sum_k (2^{k+1} - 2^k)|E_k|. \end{aligned} \tag{6.37}$$

Summing by parts and using the fact that $Nf > 2^k$ on E_k, we have

$$\begin{aligned}\sum_k \sum_{I\in\mathscr{C}_k} |\alpha_I| &\le 42\sum_k 2^k|E_k \setminus E_{k+1}| \\ &\le 42\int_{\mathbf{R}} Nf(x)\,dx, \end{aligned} \tag{6.38}$$

which establishes (6.33) and hence completes the proof. ■

We are now in a position to obtain several useful characterization of $\operatorname{Re} H^1$.

Theorem 6.14. *The following conditions on a real-valued integrable function f are equivalent:*

(i) *f is the real part of the boundary function of an H^1 function (that is, f belongs to $\operatorname{Re} H^1$);*
(ii) *the Hilbert transform Hf of f is integrable;*
(iii) *the nontangential maximal function Nf of f is integrable;*
(iv) *f belongs to the atomic Hardy space H^1_{at}.*

Moreover, such functions f satisfy the inequalities

$$\begin{aligned}\|f\|_{\operatorname{Re}(H^1)} &\le \|f\|_{H(L^1)} \le c_0\|f\|_{H^1_{\mathrm{at}}} \\ &\le c_1\|Nf\|_{L^1} \le c_2\|f\|_{\operatorname{Re}(H^1)}, \end{aligned} \tag{6.39}$$

where c_0, c_1, and c_2 are constants independent of f.

Proof. From left to right, these inequalities follow from Theorem 6.5, Corollary 6.11, Theorem 6.13, and Theorem 6.2, respectively. ■

The K-functional between $\operatorname{Re} H^1$ and L^∞ may now be described as follows:

Theorem 6.15. *There exist positive constants c_1, c_2 such that*

$$\begin{aligned} c_1 K(f, t; \operatorname{Re} H^1, L^\infty) &\le \int_0^t (Nf)^*(s)\,ds \\ &\le c_2 K(f, t; \operatorname{Re} H^1, L^\infty) \end{aligned} \tag{6.40}$$

for all f in $\operatorname{Re}(H^1) + L^\infty$.

Proof. The second estimate follows from the subadditivity of the operation $f \to f^{**}$ together with Theorem 6.2 and the fact that N is a bounded operator on L^∞. In fact, if $f = b + g$, where $b \in \operatorname{Re}(H^1)$ and $g \in L^\infty$, then

$$\begin{aligned}\int_0^t (Nf)^* &\le \int_0^t (Nb)^* + \int_0^t (Ng)^* \le \|Nb\|_{L^1} + t\|Ng\|_{L^\infty} \\ &\le c\|b\|_{\operatorname{Re}(H^1)} + t\|g\|_{L^\infty}, \end{aligned}$$

from which the desired result follows.

To establish the first estimate in (6.40), fix $t > 0$ and select an integer j so that $2^{j-1} < (Nf)^*(t) \le 2^j$. If $|\{Nf > 2^{j-1}\}| < \infty$, we proceed as in the proof of Theorem 6.13 (cf. (6.34)) to find that

$$b_j = \sum_{k=j}^{\infty} (b_k - b_{k+1}) = \sum_{k=j}^{\infty} \left\{ \sum_{I\in C_k} \alpha_I a_I \right\}. \tag{6.41}$$

Arguing as in (6.37), (6.38), we obtain from (6.41),

$$\|b_j\|_{H^1_{\mathrm{at}}} \le 42c\int_{E_j} Nf(x)\,dx \le 42c\int_0^t (Nf)^*(s)\,ds, \tag{6.42}$$

and then (6.39) gives

$$\|b_j\|_{\operatorname{Re}(H^1)} \le c\int_0^t (Nf)^*(s)\,ds.$$

Together with the estimate $\|g_j\|_\infty \le 14(Nf)^*(t)$, which follows from our selection of the index j, this establishes (6.40).

In the remaining case where $|\{Nf > 2^{j-1}\}| = \infty$, let j_0 be the smallest integer for which $|\{Nf > 2^{j_0}\}| < \infty$. Then

$$2^{j_0-1} \le \lim_{s\to\infty} (Nf)^*(s) \le 2^{j_0}$$

and since $f = b_{j_0} + g_{j_0}$ we have

$$K(f,t) \le K(b_{j_0},t) + K(g_{j_0},t) \le c\left(\int_0^t (Nf)^* + t2^{j_0}\right) \le c\int_0^t (Nf)^*.$$

This completes the proof. ∎

Corollary 6.16. *If* $0 < \theta < 1$, $1 \le q \le \infty$, *and* $\theta = 1 - 1/p$, *then*

$$(\mathrm{Re}(H^1), L^\infty)_{\theta,q} = L^{p,q} \tag{6.43}$$

with equivalent norms.

Proof. The result follows immediately from (6.40) since N is a bounded operator on the Lorentz spaces $L^{p,q}$ when $1 < p < \infty$:

$$\|f\|_{L^{p,q}} \le \|Nf\|_{L^{p,q}} \le c\|f\|_{L^{p,q}}.$$ ∎

Another consequence of Theorem 6.14 is that equation (6.12) may be extended as follows:

$$\int_{\mathbf{R}} gHf\,dx = -\int_{\mathbf{R}} f\tilde{H}g\,dx, \qquad f \in \mathrm{Re}(H^1),\ g \in L^\infty. \tag{6.44}$$

This is verified by establishing the result first for atoms and then passing to the limit.

We turn now to Fefferman's duality theorem for $\mathrm{Re}(H^1)$.

Theorem 6.17 (C. Fefferman). *The Banach space dual of* $\mathrm{Re}(H^1)$ *is isomorphic to BMO (modulo constants).*

Proof. Suppose first that ϕ belongs to BMO. We show that

$$\lambda_\phi(f) = \int f\phi, \qquad (f \in \mathrm{Re}(H^1)) \tag{6.45}$$

defines a bounded linear functional on $\mathrm{Re}(H^1)$: specifically, there is a constant c independent of ϕ such that

$$\|\lambda_\phi\|_{\mathrm{Re}(H^1)^*} \le c\|\phi\|_*. \tag{6.46}$$

Suppose first that a is an atom and that I is an interval for a which satisfies (6.24). Then

$$|\lambda_\phi(a)| = \left|\int_I \phi a\,dx\right| = \left|\int_I (\phi - \phi_I)a\,dx\right|$$
$$\le \frac{1}{|I|}\int_I |\phi - \phi_I|\,dx \le \|\phi\|_*. \tag{6.47}$$

By Theorem 6.14, an arbitrary function f in $\mathrm{Re}(H^1)$ has an atomic representation $f = \sum_j \alpha_j a_j$, where the a_j are atoms and $\sum_j |\alpha_j| < \infty$. For such a representation, the sequence $\{\lambda_\phi(\sum_1^n \alpha_j a_j)\}$ is Cauchy, and

$$|\lambda_\phi(f)| \le \sum_j |\alpha_j| \cdot |\lambda_\phi(a_j)| \le \sum_j |\alpha_j| \cdot \|\phi\|_*. \tag{6.48}$$

Taking the infimum over all such representations of f and again using (6.39), we obtain the inequality (6.46).

Next, we show that each bounded linear functional λ on $\mathrm{Re}(H^1)$ is of the form $\lambda = \lambda_\phi$ for some ϕ in BMO, and that

$$\|\phi\|_* \le c\|\lambda\|_{(\mathrm{Re}(H^1))^*}, \tag{6.49}$$

with a constant c independent of λ. Theorem 6.14(ii) shows that $\mathrm{Re}(H^1)$ is isomorphic to the closed subspace of elements (f, Hf) in $L^1 \oplus L^1$. By the Hahn-Banach theorem, λ has a norm-preserving extension to a bounded linear functional Λ on $L^1 \oplus L^1$. But the Banach space dual of $L^1 \oplus L^1$ is isometrically isomorphic to $L^\infty \oplus L^\infty$, so there exist L^∞-functions g and h such that

$$\Lambda(f_1, f_2) = \int (gf_1 + hf_2)\,dx, \tag{6.50}$$

and

$$\|\Lambda\| = \|\lambda\| = \|g\|_\infty + \|h\|_\infty. \tag{6.51}$$

But Λ extends λ so (6.50) and (6.44) give

$$\lambda(f) = \int f[g - \tilde{H}(h)]\,dx, \tag{6.52}$$

for all f in $\mathrm{Re}(H^1)$. The Spanne-Stein lemma (Lemma 6.8) now shows that $\phi \equiv g - \tilde{H}(h)$ belongs to BMO and we have

$$\|\phi\|_* \le 2\|g\|_{L^\infty} + c\|h\|_{L^\infty} \le c\|\lambda\|_{(\mathrm{Re}(H^1))^*}. \tag{6.53}$$

Hence, (6.52) shows that $\lambda = \lambda_\phi$ with $\phi \in$ BMO. The space BMO (modulo constants) is evidently a normed linear space under $\phi \to \|\phi\|_*$, and the estimates (6.46) and (6.53) together show that it is isomorphic to the dual of $\mathrm{Re}(H^1)$. ∎

7. BMO AND WEAK-L^∞

The space BMO of functions of bounded mean oscillation was examined in the last section through its intimate connections with H^1. Now we shall take a closer look at BMO and at the concept of oscillation in its own right. Our discussion begins in the local context of a fixed cube Q_0 in euclidean space $\mathbf{R}^n$. Recall that f_E denotes the mean value $f_E = (1/|E|)\int_E f(y)\,dy$ of an integrable function f over a set E of positive finite measure.

Definition 7.1. If f is integrable over Q_0, the *sharp function* $f^{\#}_{Q_0}$ *of* f *relative to* Q_0 is defined by

$$f^{\#}_{Q_0}(x) = \sup_{\substack{Q \subset Q_0 \\ Q \ni x}} \frac{1}{|Q|}\int_Q |f(y) - f_Q|\,dy, \qquad (x \in Q_0), \tag{7.1}$$

where the supremum extends over all cubes Q that contain x and are contained in Q_0.

The sharp function $f^{\#}_{Q_0}(x)$ measures locally, at the point x, the average oscillation of f from its mean value over cubes containing x. Certain crude estimates for the sharp function are easy to come by. For example, it is clear that

$$f^{\#}_{Q_0}(x) \le 2M_{Q_0}f(x), \qquad (x \in Q_0), \tag{7.2}$$

where M_{Q_0} is the Q_0-maximal operator introduced in Definition 5.4. In particular, since $M_{Q_0}f \le Mf$, it follows from Theorem III.3.8 that

$$(f^{\#}_{Q_0})^*(t) \le cf^{**}(t), \qquad (0 < t < \infty), \tag{7.3}$$

for every integrable function f supported on Q_0.

Neither of the inequalities (7.2) and (7.3) can be reversed since there are unbounded functions f whose sharp functions are bounded. The function $f(x) = |\log|x||$, $(-1 \le x \le 1)$, for example, has this property. Nevertheless, there is an inequality in the opposite direction to (7.3) when the quantity f^{**} is replaced by $f^{**} - f^*$. This fundamental result will be established in Theorem 7.3 with the aid of the following covering lemma. The lemma is a variant of Lemma III.3.7 and is established by essentially the same stopping-time argument but with the stopping occurring one stage earlier. Note that the cubes Q_j are not dyadic in the absolute sense but are instead dyadic with respect to the fixed cube Q_0 (that is, they are formed by repeated subdivision of Q_0 and each of its descendants into 2^n congruent subcubes).

Lemma 7.2. *Let Ω be a relatively open subset of a fixed cube Q_0 and suppose that $|\Omega| \le |Q_0|/2$. Then there is a family of cubes $Q_j, j = 1, 2, \ldots$, with pairwise disjoint interiors, such that*

(i) $|\Omega \cap Q_j| \le \frac{1}{2}|Q_j| \le |\Omega^c \cap Q_j|, \qquad (j = 1, 2, \ldots);$

(ii) $\Omega \subset \bigcup_j Q_j \subset Q_0;$

(iii) $|\Omega| \le \sum_j |Q_j| \le 2^{n+1}|\Omega|.$

Proof. For each x in Ω, we select a cube $Q(x)$ containing x, as follows. The cube Q_0 itself contains x and, by hypothesis, satisfies $|\Omega \cap Q_0| \le |Q_0|/2 \le |\Omega^c \cap Q_0|$. Since Ω is relatively open, repeated subdivision of Q_0 into 2^n congruent subcubes leads to a first occurrence of a cube, say $Q(x)$, containing x, that satisfies

$$|\Omega \cap Q(x)| \le \frac{1}{2}|Q(x)| \le |\Omega^c \cap Q(x)| \tag{7.4}$$

but whose descendant (by subdivision), say $Q'(x)$, satisfies

$$|\Omega \cap Q'(x)| > \frac{1}{2}|Q'(x)| = 2^{-n-1}|Q(x)|. \tag{7.5}$$

Let $K = \{Q(x) : x \in \Omega\}$. Each of the cubes in K is dyadic relative to Q_0, so if the interiors of two cubes in K have nonempty intersection, then one is contained in the other. Hence, each cube in K is contained in a maximal cube in K. Again because of their dyadic nature, there are at most countably many such maximal cubes in K, which we may list as $Q_1, Q_2, \ldots$ The interiors of these cubes are pairwise disjoint, and it is clear from the construction and (7.4) that properties i) and ii) hold. The first inequality in property iii) follows at once from ii), so it remains only to establish the second inequality in iii). However, from (7.5), we have

$$2^{-n-1}\sum_j |Q_j| \le \sum_j |\Omega \cap Q_j'| \le \sum_j |\Omega \cap Q_j| \le |\Omega|,$$

and this completes the proof. ■

Theorem 7.3. *Let f be an integrable function supported on a cube Q_0. Then*

$$f^{**}(t) - f^*(t) \le c(f^{\#}_{Q_0})^*(t), \qquad \left(0 < t < \frac{|Q_0|}{6}\right). \tag{7.6}$$

Proof. It is easy to verify that $|f|^{\#}_{Q_0} \le f^{\#}_{Q_0}$, so it will suffice to establish (7.6) for nonnegative functions f. In that case, fix t with $0 < t < |Q_0|/6$ and set

$$E = \{x \in Q_0 : f(x) > f^*(t)\}, \qquad F = \{x \in Q_0 : f^{\#}_{Q_0}(x) > (f^{\#}_{Q_0})^*(t)\}.$$

Each of these sets has measure at most t. Hence, there is a relatively open subset Ω of Q_0 with $|\Omega| \le 3t$ and $E \cup F \subset \Omega \subset Q_0$. The set Ω therefore has measure at most $|Q_0|/2$. By Lemma 7.2, there are cubes $Q_1, Q_2, \ldots$, with pairwise disjoint interiors, for which properties i), ii), and iii) of the lemma hold. We have

$$\begin{aligned} t\{f^{**}(t) - f^*(t)\} &= \int_E \{f(x) - f^*(t)\}\,dx \\ &= \sum_{j=1}^{\infty} \int_{E \cap Q_j} \{f(x) - f^*(t)\}\,dx \\ &\le \sum_j \int_{Q_j} |f - f_{Q_j}|\,dx + \sum_j |E \cap Q_j|\{f_{Q_j} - f^*(t)\} \\ &= A + B, \qquad \text{say.} \end{aligned} \tag{7.7}$$

Let $\sum'$ denote the sum over those j for which $f_{Q_j} > f^*(t)$. Then

$$B \le \sum{}' |E \cap Q_j|\{f_{Q_j} - f^*(t)\} \le \sum{}' |\Omega \cap Q_j|\{f_{Q_j} - f^*(t)\}.$$

Using property i) of the lemma and the fact that $f(u) \le f^*(t)$ on Ω^c, we have

$$B \le \sum{}' \int_{\Omega^c \cap Q_j} \{f_{Q_j} - f(u)\}\,du \le \sum{}' \int_{Q_j} |f_{Q_j} - f(u)|\,du \le A.$$

Hence, combining this with (7.7), we obtain

$$t\{f^{**}(t) - f^*(t)\} \le 2A. \tag{7.8}$$

Each cube Q_j, $(j \ge 1)$, meets F^c in at least one point, say x_j, and by definition of F we then have $f^{\#}_{Q_0}(x_j) \le (f^{\#}_{Q_0})^*(t)$. It follows that

$$\begin{aligned} A &= \sum_j |Q_j| \left\{ \frac{1}{|Q_j|} \int_{Q_j} |f(x) - f_{Q_j}|\,dx \right\} \\ &\le \sum_j |Q_j| f^{\#}_{Q_0}(x_j) \le \sum_j |Q_j| (f^{\#}_{Q_0})^*(t). \end{aligned}$$

Hence, by property iii) of the lemma,

$$A \le 2^{n+1}|\Omega|(f^{\#}_{Q_0})^*(t) \le 2^{n+1}(3t)(f^{\#}_{Q_0})^*(t).$$

This estimate, combined with (7.8), produces the desired result (7.6) and hence completes the proof. ∎

The quantity $f^{**} - f^*$, written in the form

$$f^{**}(t) - f^*(t) = \frac{1}{t} \int_0^t \{f^*(s) - f^*(t)\}\,ds, \tag{7.9}$$

is not so different from the expression in (7.1) defining the local sharp function. Indeed, both involve the difference between the specific function and a certain constant term, averaged over an interval or a cube. Thus, if we accept $f^{**} - f^*$ as some kind of measure of the oscillation of f^*, then the assertion of Theorem 7.3 becomes clearer: the decreasing rearrangement of a function oscillates no more than does the function itself.

Theorem 7.3 has several important consequences. We begin with the following integrated version of (7.6).

Corollary 7.4. *If f is integrable over a cube Q_0, then for $0 < t \le |Q_0|/6$,*

$$[(f - f_{Q_0})\chi_{Q_0}]^{**}(t) \le c \int_t^{|Q_0|} (f^{\#}_{Q_0})^*(s) \frac{ds}{s}. \tag{7.10}$$

Proof. The function $g = (f - f_{Q_0})\chi_{Q_0}$ is an integrable function supported on Q_0. Hence, (7.6) shows that

$$g^{**}(s) - g^*(s) \le c(g^{\#}_{Q_0})^*(s), \qquad \left(0 < s < \frac{|Q_0|}{6}\right).$$

An integration by parts gives the identity

$$g^{**}(t) - g^{**}(u) = \int_t^u \{g^{**}(s) - g^*(s)\} \frac{ds}{s}, \tag{7.11}$$

valid for $0 < t \le u \le |Q_0|/6$. Hence, combining this with the preceding estimate, we obtain

$$g^{**}(t) - g^{**}(u) \le c \int_t^u (g^{\#}_{Q_0})^*(s) \frac{ds}{s}, \qquad \left(0 < t \le u \le \frac{|Q_0|}{6}\right).$$

Taking $u = |Q_0|/6$, and observing that since g has support in Q_0,

$$g^{**}\left(\frac{|Q_0|}{6}\right) \le \frac{6}{|Q_0|} \int_0^{|Q_0|} g^*(s)\,ds = \frac{6}{|Q_0|} \int_{Q_0} |g(x)|\,dx = 6|g|_{Q_0},$$

we obtain, for $0 < t \le |Q_0|/6$,

$$g^{**}(t) \le c\left\{ \int_t^{|Q_0|} (g^{\#}_{Q_0})^*(s) \frac{ds}{s} + |g|_{Q_0} \right\}. \tag{7.12}$$

Now, $g^{\#}_{Q_0} = (f - f_{Q_0})^{\#}_{Q_0} = f^{\#}_{Q_0}$. Furthermore, we see from (7.1) that $|g|_{Q_0} = |f - f_{Q_0}|_{Q_0} \le f^{\#}_{Q_0}(y)$, for any $y \in Q_0$. Hence, since the function $f^{\#}_{Q_0}$ satisfies $f^{\#}_{Q_0} \ge |g|_{Q_0}$ on a set of measure $|Q_0|$, we also have

$$|g|_{Q_0} \le (f^{\#}_{Q_0})^*(|Q_0| -) \le c \int_{|Q_0|/6}^{|Q_0|} (f^{\#}_{Q_0})^*(s)\frac{ds}{s}.$$

Combining these estimates with (7.12), we obtain the desired result (7.10). ∎

Corollary 7.5. *Suppose $1 < p < \infty$. Then an integrable function f on Q_0 belongs to $L^p(Q_0)$ if and only if $f^{\#}_{Q_0}$ belongs to $L^p(Q_0)$. In fact, there are constants c_1 and c_2, depending on p, such that*

$$c_1\|f^{\#}_{Q_0}\|_p \le \|f\|_p \le c_2\{|f|_{Q_0} + \|f^{\#}_{Q_0}\|_p\}. \tag{7.13}$$

Proof. The first inequality in (7.13) follows at once from (7.2) and the Hardy-Littlewood maximal theorem. For the second, the argument used in the preceding proof gives

$$f^{**}(t) - f^{**}\left(\frac{|Q_0|}{6}\right) \le c\int_t^{|Q_0|} f^{\#*}_{Q_0}(s)\frac{ds}{s}, \qquad \left(0 < t < \frac{|Q_0|}{6}\right).$$

The term $f^{**}(|Q_0|/6)$ can be estimated by the Q_0-mean of $|f|$. Hence, applying Hardy's inequality III.(3.19) to the preceding estimate, we obtain the second of the inequalities in (7.13). ∎

The case $p = \infty$ is more interesting since, as we remarked earlier, the sharp function may be bounded when f is not. We make the following definition (cf. Definition 6.7).

Definition 7.6. An integrable function f on a cube Q_0 is said to be of *bounded mean oscillation* if $f^{\#}_{Q_0}$ is bounded, that is, if

$$\|f\|_{*,Q_0} = \sup_{Q \subset Q_0} \frac{1}{|Q|}\int_Q |f(y) - f_Q|\,dy \tag{7.14}$$

is finite, where the supremum extends over all cubes $Q \subset Q_0$. The space of all functions f on Q_0 of bounded mean oscillation is denoted by $\mathrm{BMO}(Q_0)$.

It is clear that the functional in (7.14) does not define a norm since it vanishes on constant functions. However, it is routine to verify that $\mathrm{BMO}(Q_0)$ is in fact a

Banach space under the norm

$$\|f\|_{\mathrm{BMO}(Q_0)} \equiv \|f\|_{*,Q_0} + \|f\|_{L^1(Q_0)}. \tag{7.15}$$

Corollary 7.7 (John-Nirenberg lemma). *Let Q_0 be a fixed cube in $\mathbf{R}^n$. Then there is a constant c such that*

$$[(f - f_Q)\chi_Q]^*(t) \le c\|f\chi_Q\|_{*,Q}\log^+\left(\frac{6|Q|}{t}\right), \qquad (t > 0), \tag{7.16}$$

for all $f \in \mathrm{BMO}(Q_0)$ and all subcubes Q of Q_0; equivalently,

$$|\{x \in Q : |f(x) - f_Q| > \lambda\}| \le 6|Q|\exp\left\{\frac{-\lambda}{c\|f\chi_Q\|_{*,Q}}\right\} \tag{7.17}$$

holds for all $\lambda > 0$.

Proof. Let $K = \|f\chi_Q\|_{*,Q}$. The estimate

$$[(f - f_Q)\chi_Q]^*(t) \le cK\log\left(\frac{|Q|}{t}\right)$$

follows immediately from (7.10), at least for $t \le |Q|/6$. If $|Q|/6 \le t \le |Q|$, then the result just established gives

$$[(f - f_Q)\chi_Q]^*(t) \le [(f - f_Q)\chi_Q]^*\left(\frac{|Q|}{6}\right)$$
$$\le cK\log 6 \le cK\log\left(\frac{6|Q|}{t}\right).$$

This shows that (7.16) holds for all $t \le |Q|$; it holds trivially for $t \ge |Q|$ since the left-hand side vanishes in that case. The estimate (7.17) is merely a reformulation of (7.16), since the distribution function and the decreasing rearrangement are mutually inverse. ∎

Corollary 7.8. *Suppose $1 \le p < \infty$. Then an integrable function f on Q_0 belongs to $\mathrm{BMO}(Q_0)$ if and only if*

$$\|f\|_{\mathrm{BMO}^p(Q_0)} \equiv \|f\|_1 + \sup_{Q \subset Q_0}\left\{\frac{1}{|Q|}\int_Q |f(y) - f_Q|^p\,dy\right\}^{1/p} \tag{7.18}$$

is finite. In fact, there is a constant c such that

$$\|f\|_{\mathrm{BMO}(Q_0)} \le \|f\|_{\mathrm{BMO}^p(Q_0)} \le c\|f\|_{\mathrm{BMO}(Q_0)}. \tag{7.19}$$

Proof. The first inequality in (7.19) follows directly from Hölder's inequality. For the second, we see from (7.16) that

$$\frac{1}{|Q|}\int_Q |f(y)-f_Q|^p\,dy = \frac{1}{|Q|}\int_0^{|Q|}[(f-f_Q)\chi_Q]^{*p}(t)\,dt$$
$$\le c\frac{\|f\|_{*,Q_0}^p}{|Q|}\int_0^{|Q|}\log^+\left(\frac{6|Q|}{t}\right)^p dt$$
$$\le 6c\|f\|_{*,Q_0}^p\left(\int_6^\infty \log^p u\frac{du}{u^2}\right) = c\|f\|_{*,Q_0}^p,$$

for any subcube Q of Q_0. Taking the supremum over all such cubes Q, taking p-th roots, and adding the L^1-norm to each side, we obtain the second inequality in (7.19). ■

The fact that BMO-functions have singularities whose rate of growth is at most logarithmic, as shown for example by (7.16), identifies them as members of the Zygmund class $L_{\exp}(Q_0)$ (cf. Lemma IV. 6.2 and Definitions IV.6.1, IV.6.3). Thus, we have the continuous embeddings, $(0 < p < \infty)$,

$$L^\infty(Q_0) \hookrightarrow \mathrm{BMO}(Q_0) \hookrightarrow L_{\exp}(Q_0) \hookrightarrow L^p(Q_0). \tag{7.20}$$

Of the four spaces in (7.20), only BMO(Q_0) fails to be rearrangement-invariant. Hence, the inclusions persist if BMO(Q_0) is replaced by its rearrangment-invariant hull, that is, the space consisting of all BMO-functions and all equimeasurable rearrangements of BMO-functions. We shall provide a simple characterization of the rearrangement-invariant hull in Theorem 7.10. The following definition will be needed.

Definition 7.9. Let (R,μ) be a totally σ-finite measure space. Denote by $W = W(R,\mu)$ the set of μ-measurable functions f on R for which f^* is everywhere finite and for which the functional

$$\|f\|_W = \sup_{t>0}[f^{**}(t) - f^*(t)] \tag{7.21}$$

is finite.

Despite the notation, $\|\cdot\|_W$ is not a norm. In fact, the set W fails to be a linear space, which perhaps is not too surprising in view of the following result.

Theorem 7.10 *Let Q_0 be a cube in $\mathbf{R}^n$. Then $W(Q_0)$ is the rearrangement-invariant hull of $BMO(Q_0)$, that is, a function f belongs to $W(Q_0)$ if and only if f is equimeasurable with some function g in* BMO(Q_0).

Proof. Suppose first that f is equimeasurable with some function g in BMO(Q_0), that is $f \sim g$. Then $f^* = g^*$ and Theorem 7.3 shows that

$$\|f\|_W = \|g\|_W = \sup_{0<t<|Q_0|}[g^{**}(t) - g^*(t)]$$
$$\le \sup_{t\le|Q_0|/6}[g^{**}(t) - g^*(t)] + g^{**}\left(\frac{|Q_0|}{6}\right)$$
$$\le c(\|g\|_{*,Q_0} + \|g\|_1) = c\|g\|_{\mathrm{BMO}(Q_0)} < \infty.$$

Hence, f belongs to $W(Q_0)$.

Conversely, suppose f belongs to $W(Q_0)$. We shall construct a function g in BMO(Q_0) that is equimeasurable with f. It will suffice to do so under the assumption that Q_0 is the unit cube $Q_0 = I^n$, where $I = [0,1]$, since the general case reduces to this one by a suitable change of variables. In that case, let

$$g(x) = f^*(x_1), \qquad (x = (x_1, x_2, \ldots, x_n) \in I^n).$$

Then $g \sim f$ and, for any subcube $Q = \prod_{i=1}^n [q_i, q_i + \alpha]$ of I^n, we have

$$\frac{1}{|Q|}\int_Q |g(x) - f^*(q_1+\alpha)|\,dx_1\,dx_2\ldots dx_n$$
$$= \frac{1}{\alpha^n}\int_{q_n}^{q_n+\alpha}\cdots\int_{q_1}^{q_1+\alpha}|f^*(x_1) - f^*(q_1+\alpha)|\,dx_1\,dx_2\ldots dx_n$$
$$= \frac{1}{\alpha}\int_{q_1}^{q_1+\alpha}[f^*(t) - f^*(q_1+\alpha)]\,dt$$
$$\le \frac{1}{q_1+\alpha}\int_0^{q_1+\alpha}[f^*(t) - f^*(q_1+\alpha)]\,dt$$
$$= f^{**}(q_1+\alpha) - f^*(q_1+\alpha) \le \|f\|_{W(Q_0)}.$$

It follows that $|Q|^{-1}\int_Q |g(x) - g_Q|\,dx \le 2\|f\|_W < \infty$. Hence, $g \in \mathrm{BMO}(Q_0)$ and the proof is complete. ■

Corollary 7.11. *The following inclusions hold for* $0 < p < \infty$:

$$L^\infty(Q_0) \subset \mathrm{BMO}(Q_0) \subset W(Q_0) \subset L_{\exp}(Q_0) \subset L^p(Q_0). \tag{7.22}$$

The space $W(Q_0)$ can be realized in another way, one which has interesting connections with interpolation theory. Recall that the K-functional for the couple (L^1, L^∞) is given by

$$K(f,t;L^1,L^\infty) = tf^{**}(t) = \int_0^t f^*(s)\,ds, \qquad (t>0), \tag{7.23}$$

for any f in $L^1 + L^\infty$. The K-functional for the reversed couple (L^∞, L^1) is therefore given by

$$K(f,t;L^\infty,L^1) = tK(f,t^{-1};L^1,L^\infty) = t\int_0^{t^{-1}} f^*(s)\,ds. \tag{7.24}$$

The fundamental quantity $f^*(t)$ may thus be realized as the derivative $K'(f,t;L^1,L^\infty)$ of the K-functional $K(f,t;L^1,L^\infty)$:

$$K'(f,t;L^1,L^\infty) = f^*(t). \tag{7.25}$$

On the other hand, the derivative of the K-functional for the reversed couple is easily computed from (7.24) as

$$K'(f,t;L^\infty,L^1) = t^{-1}[f^{**}(t^{-1}) - f^*(t^{-1})]. \tag{7.26}$$

One sees therefore that by merely reversing the order of the couple we arrive at the functional $f^{**} - f^*$ instead of the more familiar f^*. This suggests that $f^{**} - f^*$ might serve as a replacement for f^* in certain contexts. We shall see that this idea is particularly fruitful in connection with the theory of weak-type interpolation.

To this end, recall first from Definition IV.4.9 that an operator T is of weak type (p,p) if T is a bounded operator

$$T: L^{p,1} \to L^{p,\infty}. \tag{7.27}$$

The notion is meaningful only when p is finite because in the limiting case $p = \infty$ the space $L^{\infty,1}$ is trivial. However, let us now reformulate the quantities involved in (7.27) in terms of the functional $f^{**} - f^*$.

Proposition 7.12. *Let (R,μ) be a totally σ-finite measure space and suppose $1 < p < \infty$, $1 \le q \le \infty$. Then a function f with $f^{**}(+\infty) = 0$ belongs to $L^{p,q} = L^{p,q}(R,\mu)$ if and only if the quantity*

$$\begin{cases} \left\{\int_0^\infty [t^{1/p}(f^{**}(t) - f^*(t))]^q \dfrac{dt}{t}\right\}^{1/q}, & (q < \infty), \\ \sup_{t>0} [t^{1/p}(f^{**}(t) - f^*(t))], & (q = \infty), \end{cases} \tag{7.28}$$

is finite.

Proof. Letting $u \to \infty$ in the identity (7.11) (with g replaced by f) and noting that $f^{**}(+\infty) = 0$, we have

$$f^{**}(t) = \int_t^\infty [f^{**}(s) - f^*(s)]\frac{ds}{s}. \tag{7.29}$$

Hence, substituting this into the Definition IV.(4.7) of the $L^{(p,q)}$-norm and applying Hardy's inequalities, we see that $\|\cdot\|_{(p,q)}$ is majorized by a constant multiple of the expression in (7.28). The reverse inequality is obvious since f^{**} majorizes $f^{**} - f^*$. ■

Now let us re-examine the spaces involved in (7.27). If we let $p \to \infty$ in the definition IV.(4.7) of the norms in $L^{p,1}$ and $L^{p,\infty}$, we obtain $\int_0^\infty f^*(t)\,dt/t$ and $\sup_t f^*(t)$, respectively. Then (7.27) reads $T:\{0\} \to L^\infty$, which is trivial. However, taking the limit as $p \to \infty$ when the spaces $L^{p,1}$ and $L^{p,\infty}$ are equipped as in Proposition 7.12, we obtain from (7.28) the functionals $f^{**}(0) = \|f\|_\infty$ and $\sup[f^{**} - f^*] = \|f\|_W$, respectively. The resulting spaces are L^∞ and W, respectively, and the limiting case $p = \infty$ of (7.27) then reads:

$$T: L^\infty \to W. \tag{7.30}$$

We shall incorporate the condition suggested by (7.30)—the requirement that bounded functions be transformed into functions in W—into a formal definition. Attention will be restricted, however, to characteristic functions (cf. Definition 7.14).

Definition 7.13. Suppose $1 \le p \le \infty$ and let (R,μ) be a totally σ-finite measure space. The space *weak-L^p* is defined by

$$\text{weak-}L^p = \begin{cases} L^{p,\infty}, & \text{if } p < \infty; \\ W, & \text{if } p = \infty. \end{cases} \tag{7.31}$$

Thus, with weak L^p presented as in Proposition 7.12, the space $W = \text{weak-}L^\infty$ is indeed the limit of the space weak-L^p as $p \to \infty$.

Definition 7.14. A quasilinear operator T is of *restricted weak type* (∞,∞) if there is a constant c such that

$$\|T\chi_E\|_W \le c\|\chi_E\|_\infty, \tag{7.32}$$

for all characteristic functions χ_E of sets E of finite measure, that is, if

$$(T\chi_E)^{**}(t) - (T\chi_E)^*(t) \le c, \tag{7.33}$$

for all $t > 0$.

We now show that these notions produce a valid interpolation theorem. As previously, S denotes the appropriate Calderón operator

$$(Sf)(t) = \frac{1}{t}\int_0^t f(s)\,ds + \int_t^\infty f(s)\frac{ds}{s}, \qquad (t > 0). \tag{7.34}$$

Theorem 7.15. *Let T be a sublinear operator of restricted weak types $(1, 1)$ and (∞, ∞). Then*

$$(Tf)^{**}(t) \le cS(f^{**})(t), \qquad (t > 0) \tag{7.35}$$

for all simple functions f, and

$$\|Tf\|_p \le c_p\|f\|_p, \qquad (1 < p < \infty), \tag{7.36}$$

where c depends only on T and c_p only on p and T. In particular, T has a unique extension to a bounded sublinear operator on L^p, $(1 < p < \infty)$.

Proof. Let E be any μ-measurable subset of R with positive finite measure $s = \mu(E)$, and let $g = T\chi_E$. Since T is of restricted weak type $(1, 1)$, we have from Definition IV.5.2,

$$tg^*(t) \le cs, \qquad (t > 0), \tag{7.37}$$

and since T is of restricted weak type (∞, ∞), we see from (7.33) that

$$g^{**}(t) - g^*(t) \le c, \qquad (t > 0), \tag{7.38}$$

where c is a constant depending only on T. These estimates combine to give

$$g^*(t) \le 2c\left\{\left(\frac{s}{t} \wedge 1\right) + \log^+\left(\frac{s}{t}\right)\right\}, \qquad (t > 0). \tag{7.39}$$

This follows immediately from (7.37) if $t \ge s$. If, on the other hand, we have $0 < t < s$, then using (7.38) to estimate the integrand in the identity (7.11) (with u replaced by s), we obtain $g^{**}(t) \le g^{**}(s) + c\log(s/t)$. But this yields (7.39) since successive applications of (7.38) and (7.37) show that $g^{**}(s) \le g^*(s) + c \le 2c$.

The right-hand side of (7.39) is precisely $2cS(\chi_E^*)(t)$. Hence, (7.39) may be rewritten in the form

$$(T\chi_E)^*(t) \le 2cS(\chi_E^*)(t), \qquad (t > 0). \tag{7.40}$$

However, an integration of both sides, together with a simple computation, shows that (7.40) implies

$$(T\chi_E)^{**}(t) \le 2cS(\chi_E^{**})(t), \qquad (t > 0). \tag{7.41}$$

Now, with the sublinearity of T and of the operation $h \to h^{**}$, it is routine

to pass to the corresponding estimate (7.35) for simple functions f. The remaining assertions are elementary consequences of (7.35). ■

The Hilbert transform H can be interpolated directly by means of Theorem 7.15. All that is required is the Stein-Weiss estimate (cf. Exercise III.10)

$$(H\chi_E)^*(t) = \frac{1}{\pi}\sinh^{-1}\left(\frac{2|E|}{t}\right), \qquad (t > 0),$$

which implies at once that H is of restricted weak types $(1, 1)$ and (∞, ∞), and hence, by Theorem 7.15, is a bounded operator on L^p for $1 < p < \infty$.

We conclude this section with a look at the action of the maximal operator M_{Q_0} on $\mathrm{BMO}(Q_0)$.

Lemma 7.16. *Let Q_0 be a cube in $\mathbf{R}^n$ and suppose $f \in \mathrm{BMO}(Q_0)$. If Q is any subcube of Q_0, then*

$$(M_{Q_0}f)_Q \le c\|f\|_{*,Q_0} + \inf_Q M_{Q_0}f, \tag{7.42}$$

where c is a constant independent of f.

Proof. Let Q be a subcube of Q_0 and let Q' be the smallest subcube of Q_0 containing $3Q \cap Q_0$ (recall that $3Q$ is the cube concentric with Q and with diameter three times that of Q). Thus, $|Q| \le |Q'| \le \min(3^n|Q|, |Q_0|)$. Write

$$f = (f - f_{Q'})\chi_{Q'} + (f_{Q'}\chi_{Q'} + f\chi_{(Q')^c}) = g + h,$$

say.

Using the Cauchy-Schwarz inequality, the fact that M_{Q_0} is bounded on $L^2(Q_0)$, and Corollary 7.8, we have

$$\begin{aligned}(M_{Q_0}g)_Q &\le \left\{\frac{1}{|Q|}\int_Q (M_{Q_0}g)^2\,dx\right\}^{1/2} \le c|Q|^{-1/2}\left\{\int_{Q_0}|g|^2\,dx\right\}^{1/2}\\ &= c|Q|^{-1/2}\left\{\int_{Q'}|f - f_{Q'}|^2\,dx\right\}^{1/2} \le c\|f\|_{*,Q_0},\end{aligned} \tag{7.43}$$

where c is independent of f and Q.

Next, we shall show that

$$(M_{Q_0}h)_Q \le c\|f\|_{*,Q_0} + \inf_Q M_{Q_0}f, \tag{7.44}$$

which, together with (7.43), will establish the desired result (7.42). We shall in fact establish a result stronger than (7.44) by showing that the right-hand side of (7.44) majorizes $(M_{Q_0}h)(x)$ for every x in Q. This in turn will follow from the estimate

$$h_R \le c\|f\|_{*,Q_0} + \inf_Q M_{Q_0}f, \tag{7.45}$$

which we shall establish for every subcube R of Q_0 that contains the point x.

This result follows at once if R does not meet $(Q')^c$. For then $h_R = f_{Q'} \le \inf_Q M_{Q_0}f$. Suppose therefore that $R \cap (Q')^c \ne \varnothing$, and let R' be the smallest subcube of Q_0 that contains both R and Q'. Then $|R'| \le c|R|$, and we have

$$\int_R (h - h_{R'}) \le \int_{R'} |h - h_{R'}| = \int_{Q'} |f_{Q'} - f_{R'}| + \int_{R' \cap (Q')^c} |f - f_{R'}|$$

$$\le \left(\int_{Q'} + \int_{R' \cap (Q')^c}\right)|f - f_{R'}| = \int_{R'} |f - f_{R'}|.$$

Hence,

$$\frac{1}{|R|}\int_R (h - h_{R'}) \le \frac{|R'|}{|R|}\frac{1}{|R'|}\int_{R'}|f - f_{R'}| \le c\|f\|_{*,Q_0}.$$

We therefore have

$$h_R = (h - h_{R'})_R + h_{R'} \le c\|f\|_{*,Q_0} + \inf_Q M_{Q_0}f,$$

and this, as we remarked above, completes the proof. ■

The lemma implies that M_{Q_0} is a bounded operator on $\mathrm{BMO}(Q_0)$. In fact, it implies a stronger result. We make the following definition.

Definition 7.17. An integrable function f on a cube Q_0 is said to have *bounded lower oscillation* on Q_0 if there is a constant c such that

$$f_Q - \operatorname*{ess\,inf}_Q f \le c, \tag{7.46}$$

for all subcubes Q of Q_0. The set of functions f of bounded lower oscillation is denoted by $\mathrm{BLO}(Q_0)$.

Observe that (7.46) is the analogue of the expression (7.14), which defines $\mathrm{BMO}(Q_0)$, when averages over cubes are replaced by their essential infima. Clearly, every bounded function on Q_0 belongs to $\mathrm{BLO}(Q_0)$, and the condition (7.46), with $Q = Q_0$, implies that every function in $\mathrm{BLO}(Q_0)$ is bounded below. Furthermore, since

$$\frac{1}{|Q|}\int_Q |f - f_Q| \le 2\left(f_Q - \operatorname*{ess\,inf}_Q f\right), \tag{7.47}$$

for every subcube Q of Q_0, we see that $\mathrm{BLO}(Q_0) \subset \mathrm{BMO}(Q_0)$. Hence,

$$L^\infty(Q_0) \subset \mathrm{BLO}(Q_0) \subset \mathrm{BMO}(Q_0). \tag{7.48}$$

Theorem 7.18. *The Hardy-Littlewood maximal operator M_{Q_0} is a bounded operator on* $\mathrm{BMO}(Q_0)$. *Furthermore, M_{Q_0} maps* $\mathrm{BMO}(Q_0)$ *into* $\mathrm{BLO}(Q_0)$.

Proof. Lemma 7.16 shows at once that M_{Q_0} maps $\mathrm{BMO}(Q_0)$ into $\mathrm{BLO}(Q_0)$. The estimates (7.42) and (7.47) then show that

$$\|M_{Q_0}f\|_{*,Q_0} \le c\|f\|_{*,Q_0}, \qquad (f \in \mathrm{BMO}(Q_0)). \tag{7.49}$$

Since M_{Q_0} is a bounded operator on $L^2(Q_0)$, we have

$$\frac{1}{|Q_0|}\int_{Q_0} M_{Q_0}f \le \left\{\frac{1}{|Q_0|}\int_{Q_0}(M_{Q_0}f)^2\right\}^{1/2} \le c\left\{\frac{1}{|Q_0|}\int_{Q_0}|f|^2\right\}^{1/2}.$$

But f is supported on Q_0 so the latter integral may be written in terms of the decreasing rearrangement of f. With that and an application of the John-Nirenberg estimate (7.16), we obtain

$$\frac{1}{|Q_0|}\int_{Q_0} M_{Q_0}f \le c\left\{\frac{1}{|Q_0|}\int_0^{|Q_0|} f^{*2}\right\}^{1/2} \le c\{|f|_{Q_0} + \|f\|_{*,Q_0}\},$$

for every f in $\mathrm{BMO}(Q_0)$. Hence, combining this with (7.49) and (7.15), we obtain

$$\|M_{Q_0}f\|_{\mathrm{BMO}(Q_0)} \le c\|f\|_{\mathrm{BMO}(Q_0)}, \qquad (f \in \mathrm{BMO}(Q_0)). \tag{7.50}$$

■

We conclude with one final observation: modulo bounded functions, the space $\mathrm{BLO}(Q_0)$ is *exactly* the range of the maximal operator M_{Q_0} on $\mathrm{BMO}(Q_0)$. To see this, we first derive the following elementary characterization of $\mathrm{BLO}(Q_0)$.

Lemma 7.19. *A nonnegative integrable function f on Q_0 belongs to* $\mathrm{BLO}(Q_0)$ *if and only if $M_{Q_0}f - f$ is bounded. Furthermore,*

$$\|M_{Q_0}f - f\|_{L^\infty(Q_0)} = \sup_{Q \subset Q_0}\left(f_Q - \operatorname*{ess\,inf}_Q f\right). \tag{7.51}$$

Proof. Let $K(f)$ denote the expression on the right-hand side of (7.51). Suppose first that f belongs to $\mathrm{BLO}(Q_0)$. Let x be any Lebesgue point of f and let Q be any subcube of Q_0 containing x. Then $f(x) \ge \operatorname{ess\,inf}_Q f$, and so

$$f_Q - f(x) \le f_Q - \operatorname*{ess\,inf}_Q f \le K(f).$$

Taking the supremum over all subcubes Q of Q_0 containing x, and then the supremum over all Lebesgue points x, we find that $M_{Q_0}f - f$ belongs to L^∞ with norm at most $K(f)$.

Conversely, suppose $M_{Q_0}f - f$ belongs to $L^\infty(Q_0)$. Let Q be any subcube of Q_0. Any point x of Q for which

$$f(x) < f_Q - \|M_{Q_0}f - f\|_\infty \tag{7.52}$$

must satisfy

$$(M_{Q_0}f)(x) - f(x) \geq f_Q - f(x) > \|M_{Q_0}f - f\|_\infty,$$

and consequently the set of such points has measure zero. It follows that the essential infimum of f over Q is at least as large as the value on the right-hand side of (7.52), that is,

$$f_Q - \operatorname*{ess\,inf}_Q f \leq \|M_{Q_0}f - f\|_\infty.$$

Taking the supremum over all subcubes Q of Q_0, we see that $K(f)$ does not exceed the L^∞-norm of $M_{Q_0}f - f$. This, together with the previous estimate in the opposite direction, establishes (7.51). ∎

Theorem 7.20. *An integrable function f on Q_0 belongs to* $\mathrm{BLO}(Q_0)$ *if and only if there are functions g in* $\mathrm{BMO}(Q_0)$ *and h in $L^\infty(Q_0)$ such that*

$$f = M_{Q_0}g + h. \tag{7.53}$$

Proof. If (7.53) holds, then f belongs to $\mathrm{BLO}(Q_0)$ by virtue of Theorem 7.18. Conversely, if f belongs to $\mathrm{BLO}(Q_0)$, then f is bounded below and so there is a constant c such that $g = f + c$ is nonnegative. Write

$$f = M_{Q_0}g + (g - M_{Q_0}g - c) = M_{Q_0}g + h,$$

in the form of (7.53). Then g belongs to $\mathrm{BLO}(Q_0) \subset \mathrm{BMO}(Q_0)$. Moreover, since g is nonnegative and belongs to $\mathrm{BLO}(Q_0)$, Lemma 7.19 shows that $M_{Q_0}g - g$ belongs to L^∞. Hence, the function $h = g - M_{Q_0}g - c$ also belongs to L^∞. This completes the proof. ∎

8. INTERPOLATION BETWEEN L^1 AND BMO

In the preceding section, we constructed the space weak-L^∞ and established a corresponding weak-type interpolation theorem. Now we shall concentrate on the interpolation properties of BMO itself. In particular, we shall describe the K-functional for the couple (L^1, BMO), in terms of the sharp function $f^\#$ (Theorem 8.8). Although most of what we shall say remains valid in the previous local context of a cube Q_0 in $\mathbf{R}^n$, the scope of the present discussion will be enlarged to include BMO-functions defined on the whole of $\mathbf{R}^n$.

Definition 8.1. If f is locally integrable on $\mathbf{R}^n$, the *sharp function* $f^\#$ of f is defined by

$$f^\#(x) = \sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f(y) - f_Q|\,dy, \qquad (x \in \mathbf{R}^n), \tag{8.1}$$

where the supremum extends over all cubes Q that contain x.

As in the previous section, we have

$$f^\#(x) \leq 2Mf(x), \qquad (x \in \mathbf{R}^n), \tag{8.2}$$

where M is the Hardy-Littlewood maximal operator, and

$$f^{\#*}(t) \leq c f^{**}(t), \qquad (0 < t < \infty), \tag{8.3}$$

for every locally integrable function f on $\mathbf{R}^n$. There is the following analogue of Corollary 7.4.

Proposition 8.2. *If f is locally integrable over $\mathbf{R}^n$ and Q is any cube in $\mathbf{R}^n$, then*

$$[(f - f_Q)\chi_Q]^{**}(t) \leq c \int_t^{|Q|} f^{\#*}(s)\frac{ds}{s}, \qquad \left(0 < t \leq \frac{|Q|}{6}\right). \tag{8.4}$$

Proof. The function $(f - f_Q)\chi_Q$ is an integrable function supported on Q. Applying (7.10) to this function, and noting that $[(f - f_Q)\chi_Q]_Q^\# \leq f^\#$ on Q, we obtain (8.4). ∎

Corollary 8.3. *Let f be locally integrable on $\mathbf{R}^n$. If Q_0 and Q_1 are cubes with $Q_0 \subset Q_1$, then*

$$|f_{Q_0} - f_{Q_1}| \leq c \int_{|Q_0|/6}^{|Q_1|} f^{\#*}(s)\frac{ds}{s}. \tag{8.5}$$

Proof. For any cube Q we have $(\lambda\chi_Q)^{**}(t) = \lambda$, for $0 < t < |Q|$. Hence,

$$|f_{Q_0} - f_{Q_1}| = [(f_{Q_0} - f_{Q_1})\chi_{Q_0}]^{**}\left(\frac{|Q_0|}{6}\right)$$

$$\leq [(f - f_{Q_0})\chi_{Q_0}]^{**}\left(\frac{|Q_0|}{6}\right) + [(f - f_{Q_1})\chi_{Q_1}]^{**}\left(\frac{|Q_0|}{6}\right).$$

Applying (8.4) to each of the terms on the right, we obtain

$$|f_{Q_0} - f_{Q_1}| \le c \int_{|Q_0|/6}^{|Q_0|} f^{\#*}(s)\frac{ds}{s} + c \int_{|Q_0|/6}^{|Q_1|} f^{\#*}(s)\frac{ds}{s},$$

from which (8.5) follows since $|Q_0| \le |Q_1|$. ■

Definition 8.4. A locally integrable function f on $\mathbf{R}^n$ is said to be of *bounded mean oscillation* if $f^{\#}$ is bounded, that is, if

$$\|f\|_* = \sup_Q \frac{1}{|Q|}\int_Q |f(y) - f_Q|\,dy \tag{8.6}$$

is finite, where the supremum extends over all cubes Q in $\mathbf{R}^n$. The space of all functions f of bounded mean oscillation is denoted by BMO.

It is clear that the functional in (8.6) does not define a norm since $\|f\|_* = 0$ whenever f is constant. However, by factoring out the constant functions, that is, considering the quotient space BMO/**C**, we have the following result.

Proposition 8.5. BMO/**C** *is a Banach space under* $\|\cdot\|_*$.

Proof. We establish only the completeness. Let $(F_k)_{k=1}^{\infty}$ be a sequence of cosets in BMO/**C** for which

$$\sum_k \|F_k\|_* < \infty. \tag{8.7}$$

Let Q_0 be the unit cube centered at the origin. For each $k = 1, 2, \ldots$, let f_k be the BMO-function in the coset F_k that has mean value zero over Q_0. Let Q be any cube in $\mathbf{R}^n$ and let Q' be the smallest cube that contains both Q_0 and Q. Then, by Corollary 8.3 and (8.7), we have

$$\begin{aligned}\sum_{k=1}^{\infty} \frac{1}{|Q|}\int_Q |f_k(y)|\,dy &\le \frac{|Q'|}{|Q|}\sum_{k=1}^{\infty}\frac{1}{|Q'|}\int_{Q'}|f_k(y)|\,dy \\ &\le \frac{|Q'|}{|Q|}\sum_{k=1}^{\infty}\left\{\frac{1}{|Q'|}\int_{Q'}|f_k - (f_k)_{Q'}| + |(f_k)_{Q'} - (f_k)_{Q_0}|\right\} \\ &\le \frac{|Q'|}{|Q|}\sum_{k=1}^{\infty}\left\{\|f_k\|_* + c\int_{|Q_0|/6}^{|Q'|} f_k^{\#*}(s)\frac{ds}{s}\right\} \\ &\le \frac{|Q'|}{|Q|}\left[1 + \log\left(\frac{6|Q'|}{|Q_0|}\right)\right]\sum_{k=1}^{\infty}\|f_k\|_* < \infty.\end{aligned}$$

This shows that the series $\sum_k f_k$ converges in $L^1(Q)$, and hence a.e. on Q. Since Q is arbitrary, the function $f = \sum_k f_k$ is therefore well-defined a.e. on $\mathbf{R}^n$. The Beppo Levi theorem and the preceding estimate show that $f_Q = \sum_k (f_k)_Q$ for any cube Q. It it now routine to verify that f belongs to BMO and that $\|\sum_k F_k - F\|_* \to 0$ as $k \to \infty$, where F is the coset in BMO/**C** that contains f. ■

As in the previous section (cf. Corollary 7.7), there is the following basic estimate for BMO-functions.

Proposition 8.6 (John-Nirenberg lemma). *There is a constant c such that the estimate*

$$[(f - f_Q)\chi_Q]^*(t) \le c\|f\|_* \log^+\left(\frac{6|Q|}{t}\right), \qquad (t > 0) \tag{8.8}$$

holds for all $f \in$ BMO and all cubes Q; equivalently,

$$|\{x \in Q : |f(x) - f_Q| > \lambda\}| \le 6|Q| \exp\left\{\frac{-\lambda}{c\|f\|_*}\right\} \tag{8.9}$$

holds for all $\lambda > 0$.

Proof. The result follows from (8.4) in exactly the same way that Corollary 7.7 derives from (7.10). We omit the details. ■

The John-Nirenberg lemma also provides us with an analogue of Corollary 7.8.

Corollary 8.7. *Suppose $1 \le p < \infty$. Then a locally integrable function f belongs to* BMO *if and only if*

$$\sup_Q \left\{\frac{1}{|Q|}\int_Q |f(y) - f_Q|^p\,dy\right\}^{1/p} < \infty. \tag{8.10}$$

One of our primary objectives in this section is to characterize the K-functional for the couple (L^1, BMO). For the purposes of the next result, we regard BMO as a space of functions, semi-normed by $\|\cdot\|_*$.

Theorem 8.8. *There are constants c_1 and c_2, depending only on the dimension n, such that the estimates*

$$c_1 t(f^{\#})^*(t) \le K(f, t; L^1, \text{BMO}) \le c_2 t(f^{\#})^*(t) \tag{8.11}$$

hold for all $f \in (L^1 + \text{BMO})(\mathbf{R}^n)$ and all $t > 0$.

Proof. If $f = g + h$, with $g \in L^1$ and $h \in \mathrm{BMO}$, then

$$f^\# \leq g^\# + h^\# \leq g^\# + \|h\|_*.$$

Then, for any $t > 0$,

$$t(f^\#)^*(t) \leq t(g^\#)^*(t) + t\|h\|_*. \tag{8.12}$$

Using (8.2) and the weak-type (1, 1) estimate for the Hardy-Littlewood maximal operator, we find that

$$t(g^\#)^*(t) \leq 2t(Mg)^*(t) \leq c\|g\|_1.$$

Hence, combining this with (8.12), we have

$$t(f^\#)^*(t) \leq c(\|g\|_1 + t\|h\|_*).$$

Now taking the infimum over all such representations $f = g + h$ of f, we obtain the first of the inequalities in (8.11).

The second of the inequalities in (8.11) is more difficult to establish. Fix $f \in L^1 + \mathrm{BMO}$ and fix $t > 0$. We need to exhibit $g \in L^1$ and $h \in \mathrm{BMO}$ with $f = g + h$ and

$$\|g\|_1 + t\|h\|_* \leq c_2 t(f^\#)^*(t), \tag{8.13}$$

where c_2 depends only on the dimension n.

Let $\Omega = \{x \in \mathbf{R}^n : f^\#(x) > (f^\#)^*(t)\}$. Then Ω is an open set with measure $|\Omega| \leq t$. Let $F = \Omega^c$ and let $\{Q_j\}$ be a Whitney covering of Ω. Thus (cf. Lemma 5.1), the Q_j are dyadic cubes with pairwise disjoint interiors such that $\bigcup_j Q_j = \Omega$ and

$$\operatorname{diam}(Q_j) \leq d(Q_j, F) \leq 4\operatorname{diam}(Q_j), \qquad (j = 1, 2, \ldots). \tag{8.14}$$

For a fixed index j_0, let $J_0 = \{j : Q_j \cap Q_{j_0} \neq \varnothing\}$. By Lemma 5.2(a), we have

$$\frac{1}{4}\operatorname{diam}(Q_{j_0}) \leq \operatorname{diam}(Q_j) \leq 4\operatorname{diam}(Q_{j_0}), \qquad (j \in J_0), \tag{8.15}$$

and so

$$\frac{3}{2}Q_{j_0} \subset \bigcup_{j \in J_0} Q_j \subset 9Q_{j_0}. \tag{8.16}$$

The functions g and h desired in (8.13) are defined by

$$g(x) = \sum_j [f(x) - f_{Qj}]\chi_{Q_j}(x) \tag{8.17}$$

and

$$h(x) = \sum_j f_{Q_j}\chi_{Q_j} + f(x)\chi_F(x). \tag{8.18}$$

For each $j = 1, 2, \ldots$, let $Q_j' = \beta Q_j$ be the cube concentric with Q_j but with diameter $\beta = 10n^{1/2}$ times as large. Then it follows from (8.14) that

$$Q_j' \cap F \neq \varnothing, \qquad (j = 1, 2, \ldots). \tag{8.19}$$

If Q is any cube and $Q' = \beta Q$, then

$$(|f - f_Q|)_Q \leq 2\beta^n(|f - f_{Q'}|)_{Q'}. \tag{8.20}$$

If Q is one of the Whitney cubes Q_j, then (8.19) shows that Q_j' contains a point of F. At such points (by construction of the set F), the value of $f^\#$ does not exceed $(f^\#)^*(t)$. Hence, this in conjunction with (8.20) gives

$$(|f - f_{Q_j}|)_{Q_j} \leq 2\beta^n(f^\#)^*(t), \qquad (j = 1, 2, \ldots). \tag{8.21}$$

Using (8.17), (8.21) and the fact that the Whitney cubes Q_j have union equal to the set Ω of measure at most t, we therefore arrive at the following estimate for the L^1-norm of g:

$$\|g\|_{L^1} = \sum_j |Q_j|(|f - f_{Q_j}|)_{Q_j} \leq 2\beta^n t(f^\#)^*(t). \tag{8.22}$$

In view of (8.22), the desired result (8.13) will be established if we show that $\|h\|_* \leq c(f^\#)^*(t)$ with c depending only on n. Thus, we need to verify that $(|h - h_Q|)_Q \leq c(f^\#)^*(t)$ for all cubes Q. For this, it will suffice to show that to each cube Q there corresponds a constant $\alpha = \alpha_Q$ such that

$$A(Q) \equiv (|h - \alpha|)_Q \leq c(f^\#)^*(t). \tag{8.23}$$

To establish (8.23), fix Q and let $K = \{k : Q_k \cap Q \neq \varnothing\}$. On each Q_k, $(k = 1, 2, \ldots)$, the function h is constant and equal to f_{Q_k}. Hence,

$$A(Q) = \sum_{k \in K} \frac{|Q_k \cap Q|}{|Q| \cdot |Q_k|}\left|\int_{Q_k} (f - \alpha)\right| + \frac{1}{|Q|}\int_{Q \cap F} |f - \alpha|. \tag{8.24}$$

There are three cases to consider:

Case 1. Suppose K is empty. Then $Q \subset F$ and we select $\alpha = f_Q$. We have from (8.24),

$$A(Q) \leq (|f - f_Q|)_Q \leq f^{\#*}(t),$$

the latter inequality because Q contains a point of F, and at such points we have $f^\# \leq (f^\#)^*(t)$.

Case 2. Suppose K is nonempty and that there exists $j_0 \in K$ such that

$$\operatorname{diam}(Q) < \frac{1}{4}\operatorname{diam}(Q_{j_0}). \tag{8.25}$$

Then $Q \subset (3/2)Q_{j_0}$ and so (8.16) shows that $Q \cap F = \varnothing$. We claim that $K \subset J_0$. In other words, any Whitney cube Q_k that touches Q must also touch Q_{j_0}. To see this, observe from (8.25) that Q is in the interior of $(3/2)Q_{j_0}$ at a positive distance from its boundary, so any Q_k, $(k \in K)$, since it touches Q, must intersect $(3/2)Q_{j_0}$ in a set of positive measure. Hence, (8.16) shows that Q_k intersects some Q_j, $(j \in J_0)$, in a set of positive measure. Since the Whitney cubes have mutually disjoint interiors, this implies that $Q_k = Q_j$, and hence that $k \in J_0$. This establishes our claim that $K \subset J_0$.

Let Q_0' denote the cube $Q_0' = \beta Q_{j_0}$ with $\beta = 10n^{1/2}$ as before. Let $\alpha = f_{Q_0'}$ in (8.24). The second term on the right vanishes because Q does not meet F. To estimate the first term, replace $|Q_k \cap Q|$ with the larger $|Q|$, replace the index set K with the larger J_0, and use the first inequality in (8.15) to obtain

$$A(Q) \le \sum_{j \in J_0} \frac{1}{|Q_j|} \int_{Q_j} |f(x) - f_{Q_0'}|\,dx$$
$$\le \frac{4^n}{|Q_{j_0}|} \sum_{j \in J_0} \int_{Q_j} |f(x) - f_{Q_0'}|\,dx.$$

Now $|Q_{j_0}| = \beta^{-n}|Q_0'|$ and the second inclusion in (8.16) shows that the cubes Q_j, $(j \in J_0)$, are (disjoint) subsets of Q_0'. Hence

$$A(Q) \le (4\beta)^n(|f - f_{Q_0'}|)_{Q_0'} \le (4\beta)^n(f^{\#})^*(t),$$

the last inequality because Q_0' contains a point of F.

Case 3. Suppose K is nonempty and, for all $k \in K$, we have

$$\frac{1}{4}\operatorname{diam}(Q_k) \le \operatorname{diam}(Q). \tag{8.26}$$

Then

$$\bigcup_{k \in K} Q_k \subset 9Q, \tag{8.27}$$

and it follows from (8.19) that the cube $\tilde{Q} = (9\beta)Q$ meets F. From (8.24) we have

$$A(Q) \le \frac{1}{|Q|}\left\{\sum_{k \in K} \int_{Q_k} |f(x) - \alpha|\,dx + \int_{Q \cap F} |f(x) - \alpha|\,dx\right\}$$
$$\le \frac{1}{|Q|} \int_{\tilde{Q}} |f(x) - \alpha|\,dx = \frac{(9\beta)^n}{|\tilde{Q}|} \int_{\tilde{Q}} |f(x) - \alpha|\,dx.$$

Hence, choosing $\alpha = f_{\tilde{Q}}$, we obtain

$$A(Q) \le (9\beta)^n(|f - f_{\tilde{Q}}|)_{\tilde{Q}} \le (9\beta)^n(f^{\#})^*(t),$$

since $\tilde{Q}$ contains a point of F.

In all three cases, we have obtained $A(Q) \le (9\beta)^n(f^{\#})^*(t)$, for all cubes Q in $\mathbf{R}^n$. This establishes the desired inequality (8.23) and hence, as we remarked above, completes the proof. ■

In order to properly interpret the pair (L^1, BMO) as a compatible couple of Banach spaces, it is necessary to factor out the constant functions. Then $\mathrm{BMO}/\mathbf{C}$ is a Banach space (Proposition 8.5) and $L^1/\mathbf{C}$, which may be identified with L^1, is a Banach space in the natural way. We shall use $F, G, \ldots$ to denote cosets of functions in this equivalence relation, and $f, g \ldots$ to denote representative function elements of these cosets.

Lemma 8.9. *Let f be a locally integrable function on $\mathbf{R}^n$ and suppose that*

$$\int_1^\infty (f^{\#})^*(s)\frac{ds}{s} < \infty. \tag{8.28}$$

Then the limit

$$f_\infty \equiv \lim_{|Q| \to \infty} f_Q \tag{8.29}$$

exists, where the limit is taken over all cubes Q in $\mathbf{R}^n$ as the measure $|Q|$ tends to infinity.

Proof. Let Q_k denote the cube centered at the origin with side-length 2^k, $(k = 1, 2, \ldots)$. The hypothesis (8.28) together with Corollary 8.3 shows that $(f_{Q_k})_{k=1}^\infty$ is a Cauchy sequence. Denote its limit by f_∞. For each $\varepsilon > 0$, it follows from (8.28) that there is a positive integer M such that

$$c\int_{M/2}^\infty (f^{\#})^*(s)\frac{ds}{s} < \frac{\varepsilon}{3}, \tag{8.30}$$

where c is the constant in (8.5). We may choose k sufficiently large so that

$$|f_{Q_k} - f_\infty| < \frac{\varepsilon}{3} \qquad \text{and} \qquad |Q_k| > M.$$

Now let Q be any cube in $\mathbf{R}^n$ with $|Q| > M$. Let Q' be any cube that contains both Q and Q_k. Then, by (8.5), (8.30), and the choice of f_∞, we have

$$|f_Q - f_\infty| \le |f_Q - f_{Q'}| + |f_{Q'} - f_{Q_k}| + |f_{Q_k} - f_\infty|$$
$$\le \left\{\int_{|Q|}^\infty + \int_{|Q_k|}^\infty\right\}(f^{\#})^*(s)\frac{ds}{s} + |f_{Q_k} - f_\infty| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Thus, there is a number f_∞ with the following property: for each $\varepsilon > 0$, there exists $M = M(\varepsilon, f) > 0$ such that for any cube Q with $|Q| > M$, we have $|f_Q - f_\infty| < \varepsilon$. This establishes the existence of the limit in (8.29) and hence completes the proof. ■

With the aid of Lemma 8.9, we can now establish the following variant of Proposition 8.2.

Proposition 8.10. *Let f be a locally integrable function on $\mathbf{R}^n$ and suppose that*

$$\int_1^\infty (f^\#)^*(s)\frac{ds}{s} < \infty. \tag{8.31}$$

Then

$$(f - f_\infty)^{**}(t) \le c\int_t^\infty (f^\#)^*(s)\frac{ds}{s}, \qquad (0 < t < \infty). \tag{8.32}$$

Proof. Fix $t > 0$ and let $\varepsilon > 0$. By Lemma 8.9, there is a cube Q with $|Q| > 6t$ and $|f_Q - f_\infty| < \varepsilon$. Using Proposition 8.2, we have

$$\begin{aligned}[(f - f_\infty)\chi_Q]^{**}(t) &\le [(f - f_Q)\chi_Q]^{**}(t) + |f_Q - f_\infty| \\ &\le c\int_t^\infty (f^\#)^*(s)\frac{ds}{s} + \varepsilon.\end{aligned}$$

If $Q \uparrow \mathbf{R}^n$, the monotone convergence theorem shows that

$$[(f - f_\infty)\chi_Q]^{**}(t) \uparrow (f - f_\infty)^{**}(t).$$

Hence, the preceding estimate, when ε is allowed to tend to zero, provides the desired result (8.32). ■

We are now in a position to identify the interpolation spaces between the Banach spaces L^1 and BMO. Keep in mind here that BMO is interpreted as a quotient space ($=\mathrm{BMO}/\mathbf{C}$) where the constants have been factored out. We have the following result.

Theorem 8.11. *Suppose $0 < \theta < 1$ and $1 \le q \le \infty$. Then*

$$(L^1, \mathrm{BMO})_{\theta,q} = L^{p,q}, \qquad (p = 1/(1-\theta)), \tag{8.33}$$

with equivalent norms. Precisely, in each coset F in the interpolation space $(L^1, \mathrm{BMO})_{\theta,q}$, there is one and only one function f in the Lorentz space $L^{p,q}$, and this function f satisfies

$$c_1\|f\|_{L^{p,q}} \le \|F\|_{(L^1,\mathrm{BMO})_{\theta,q}} \le c_2\|f\|_{L^{p,q}}, \tag{8.34}$$

where c_1 and c_2 are independent of f and F.

Proof. Suppose F is a coset in $(L^1, \mathrm{BMO})_{\theta,q}$. Then

$$\int_0^\infty [t^{-\theta}K(F, t; L^1, \mathrm{BMO})]^q \frac{dt}{t} < \infty,$$

with the obvious modification if $q = \infty$. Now, Theorem 8.8 shows that $K(F, t; L^1, \mathrm{BMO}) \sim t(F^\#)^*(t)$. It follows therefore from the preceding estimate that

$$\int_1^\infty (F^\#)^*(t)\frac{dt}{t} < \infty,$$

and hence from Lemma 8.9 that

$$f_\infty = \lim_{|Q|\to\infty} f_Q$$

exists, for any function f in the coset F. If we now choose f to be the unique function in F for which $f_\infty = 0$, then Proposition 8.10 shows that

$$f^{**}(t) \le c\int_t^\infty (F^\#)^*(t)\frac{dt}{t}, \qquad (0 < t < \infty).$$

Since $0 < \theta < 1$, this and a simple application of Hardy's inequalities produce the first of the inequalities in (8.34).

Conversely, if $f \in L^{p,q}$, then (8.3) shows that the coset F containing f satisfies

$$(F^\#)^*(t) \le cf^{**}(t), \qquad (0 < t < \infty).$$

It follows from another application of Hardy's inequalities that F belongs to the interpolation space $(L^1, \mathrm{BMO})_{\theta,q}$ and that the second inequality in (8.34) holds. This completes the proof. ■

The analogue of Lemma 7.16 holds (with exactly the same proof) for the Hardy-Littlewood maximal operator M defined in $\mathbf{R}^n$. Thus, as in (7.42), we have

$$(Mf)_Q \le c\|f\|_* + \inf_Q Mf, \qquad (f \in \mathrm{BMO}), \tag{8.35}$$

for any cube Q in $\mathbf{R}^n$, where c depends only on n.

In the present context of $\mathbf{R}^n$, some additional complications occur which did not arise in the former setting of a finite subcube of $\mathbf{R}^n$. For example, there are functions f in BMO($\mathbf{R}^n$) for which Mf is identically infinite (take $f(x) = \log|x|$, for example). If $Mf(x)$ is finite at one point x in $\mathbf{R}^n$, however, then (8.35) shows that Mf is finite at almost every point of $\mathbf{R}^n$. Indeed, if $Mf(x) < \infty$, and if Q is a cube containing x, it follows from (8.35) that Mf is integrable over Q, hence finite a.e. on Q. But Q is an arbitrary cube containing x so we deduce that Mf is finite a.e. on $\mathbf{R}^n$. This dichotomy will be seen in several of the results which follow. Another complication is that functions in BLO($\mathbf{R}^n$) need not be bounded below. As we shall see, these difficulties can be neatly resolved by considering the following "one-sided" version of the maximal operator (note that there are no absolute values in the definition).

Definition 8.12. Let f be a locally integrable function in $\mathbf{R}^n$. The maximal function $\tilde{M}f$ of f is defined by

$$\tilde{M}f(x) = \sup_{Q \ni x} \frac{1}{|Q|} \int_Q f(y)\,dy = \sup_{Q \ni x} f_Q, \tag{8.36}$$

where the supremum extends over all cubes Q containing x.

It is clear that $f \le \tilde{M}f$ a.e. and that $|\tilde{M}f| \le Mf$. Hence, one sees easily that the proof of (8.35) carries over almost verbatim for the operator $\tilde{M}$, to yield

$$(\tilde{M}f)_Q \le c\|f\|_* + \inf_Q \tilde{M}f, \qquad (f \in \mathrm{BMO}). \tag{8.37}$$

Thus, if the space BLO($\mathbf{R}^n$) is defined in the obvious way by the requirement (7.46) (for all subcubes Q of $\mathbf{R}^n$), we see from (8.37) that

$$f \in \mathrm{BMO}(\mathbf{R}^n), \qquad \tilde{M}f \not\equiv \infty \qquad \Rightarrow \qquad \tilde{M}f \in \mathrm{BLO}(\mathbf{R}^n). \tag{8.38}$$

One checks easily that the analogue of Lemma 7.19 is valid for the operator $\tilde{M}$ and for all *real-valued* functions f, namely:

$$f \in \mathrm{BLO}(\mathbf{R}^n) \qquad \Leftrightarrow \qquad \tilde{M}f - f \in L^\infty. \tag{8.39}$$

Hence, we obtain the following analogue of Theorem 7.20, characterizing BLO (modulo bounded functions) as the range of the maximal operator $\tilde{M}$ on BMO.

Theorem 8.13. *A locally integrable function f on $\mathbf{R}^n$ belongs to* BLO *if and only if there is a* BMO*-function g, whose maximal function $\tilde{M}g$ is not identically infinite, and an L^∞-function h such that*

$$f = \tilde{M}g + h. \tag{8.40}$$

Proof. The sufficiency of (8.40) follows at once from (8.38). For the necessity, write $f \in$ BLO in the form

$$f = \tilde{M}f + (f - \tilde{M}f).$$

Then (8.40) follows with $g = f$ and $h = f - \tilde{M}f$, the latter function being bounded by virtue of (8.39). This completes the proof. ■

9. JONES' SOLUTION OF $\bar{\partial}f = \mu$

Many of the constructions that have been used in previous sections to describe the K-functionals for various couples have relied on truncation in one form or another. Although such techniques do not in general preserve analyticity, they can nevertheless still be employed in the case of spaces of analytic functions. The key is to be able to replace functions constructed by standard methods with analytic functions which satisfy similar estimates. The transition is achieved through a method due to Peter Jones, which provides constructive solutions of $\bar{\partial}$-equations with Carleson measure data. This method will be described in the present section. It will be used in a number of ways in the next section to determine the interpolation properties of the couple (H^1, H^∞) of analytic Hardy spaces.

As before, let $\mathbf{U}$ denote the upper half plane

$$\mathbf{U} = \{z = x + iy : x, y \in \mathbf{R},\ y > 0\}.$$

For $p \ge 1$, the (analytic) Hardy space $H^p(\mathbf{U})$ is the Banach space of all analytic functions F on $\mathbf{U}$ for which the norm

$$\|F\|_{H^p} = \sup_{y>0} \left\{ \int_{\mathbf{R}} |F(x+iy)|^p\,dx \right\}^{1/p}$$

(with the obvious modification if $p = \infty$) is finite.

Definition 9.1. A measure μ on the upper half plane $\mathbf{U}$ is said to be a *Carleson measure* if there is a constant $c > 0$ such that

$$|\mu|(Q) \le c|I|,$$

for each square Q of the form $Q = I \times (0, |I|)$, where I is an open interval in $\mathbf{R}$. The smallest constant c for which the inequality holds is called the *Carleson norm* of μ and will be denoted by $\|\mu\|_C$.

The following theorem illustrates the connection between Carleson measures and the Hardy space $H^1(\mathbf{U})$.

Theorem 9.2. *A necessary and sufficient condition for a measure μ to be a Carleson measure is that there exist a constant $c > 0$ such that, for every $F \in H^1(\mathbf{U})$,*

$$\iint_{\mathbf{U}} |F(z)|\, d|\mu| \le c\|F\|_{H^1}. \tag{9.1}$$

Proof. To prove sufficiency, fix an interval I with center x_0 and length d. If $F(z) = 5[(z - x_0)/d + i]^{-2}$, then $|F(z)| \ge 1$ on the square $Q = I \times (0, d)$ and $\|F\|_{H^1} \le 5\pi d$. Hence, by (9.1),

$$|\mu|(Q) \le \iint_{\mathbf{U}} |F|\, d|\mu| \le c\|F\|_{H^1} \le c|I|,$$

and so μ is a Carleson measure.

Conversely, suppose μ is a Carleson measure. Write the integral on the left hand side of (9.1) in terms of its distribution function $\lambda_{|F|}(s) = |\mu|\{z \in \mathbf{U} : |F(z)| > s\}$:

$$\iint_{\mathbf{U}} |F|\, d|\mu| = \int_0^\infty \lambda_{|F|}(s)\, ds.$$

The subset $E_s = \{x \in \mathbf{R} : NF(x) > s\}$, where NF is the non-tangential maximal function of F, is open and may therefore be expressed as a countable disjoint union of open intervals I_j. For each j, let Q_j be the square $Q_j = I_j \times (0, |I_j|)$. Then $\{z \in \mathbf{U} : |F(z)| > s\} \subset \bigcup_j Q_j$, and so

$$\begin{aligned}\lambda_{|F|}(s) &\le \sum_j |\mu|(Q_j) \le \|\mu\|_C \sum_j |I_j| \\ &= \|\mu\|_C |E_s| = \|\mu\|_C \lambda_{NF}(s).\end{aligned}$$

Hence,

$$\iint_{\mathbf{U}} |F|\, d|\mu| \le \|\mu\|_C \int_0^\infty \lambda_{NF}(s)\, ds = \|\mu\|_C \|NF\|_{L^1}. \tag{9.2}$$

But the Hardy-Littlewood theorem (Theorem 6.2) shows that the L^1-norm of NF is majorized by a multiple of the H^1-norm of F, and so (9.2) produces the desired result (9.1). ■

The partial differential operators ∂ and $\bar\partial$ are defined by the respective formulas

$$\partial f = \frac{1}{2}(f_x - if_y), \qquad \bar\partial f = \frac{1}{2}(f_x + if_y).$$

In this notation, the Laplacian $\triangle$ is equal to $4\partial\bar\partial$, and the Cauchy-Riemann conditions imply that $\bar\partial F = 0$ and $\partial F = F'$ whenever F is analytic. Moreover, it is an elementary fact from the general theory of distributions that the condition $\bar\partial F = 0$ (in the distributional sense) implies that F is analytic (cf. G. B. Folland [1]).

As we have mentioned above, our ultimate goal in this section will be to solve the equation $\bar\partial f = \mu$, where μ is a Carleson measure. To begin with, we shall need to know the *fundamental solution* for the operator $\bar\partial$, that is, the solution f for which $\bar\partial f = \delta_\zeta$, where δ_ζ is the Dirac point mass at a point ζ in $\mathbf{U}$. The next two lemmas provide this information.

Lemma 9.3. *Suppose g has first order continuous partial derivatives in an open set R with smooth boundary ∂R. Then*

$$\iint_R \bar\partial g\, dx\, dy = \frac{1}{2i} \int_{\partial R} g\, dz. \tag{9.3}$$

Proof. According to Green's theorem,

$$\iint_R (Q_x - P_y)\, dx\, dy = \int_{\partial R} P\, dx + Q\, dy,$$

so the lemma follows at once from the definition of $\bar\partial$ by setting $P = -ig/2$ and $Q = g/2$. ■

Lemma 9.4. *Let ζ be a point in $\mathbf{U}$. Then the function $f(z) = 1/[\pi(z - \zeta)]$ satisfies the distributional equation*

$$\bar\partial f = \delta_\zeta, \tag{9.4}$$

where δ_ζ is the point mass at ζ. Moreover, if G is analytic in $\mathbf{U}$, then

$$\bar\partial(Gf) = G(\zeta)\delta_\zeta. \tag{9.5}$$

Proof. Let ϕ be a test function (infinitely differentiable with compact support contained in $\mathbf{U}$), and let Q be an open square contained in $\mathbf{U}$ with the support of ϕ in its interior. Then by the definition of distributional derivatives

$$\begin{aligned}\langle \bar\partial f, \phi \rangle &= -\frac{1}{\pi} \iint_Q \frac{\bar\partial\phi(z)}{z - \zeta}\, dx\, dy \\ &= -\frac{1}{\pi} \lim_{\rho \downarrow 0} \iint_{Q_\rho} \frac{\bar\partial\phi(z)}{z - \zeta}\, dx\, dy,\end{aligned} \tag{9.6}$$

where $Q_\rho = \{z \in Q : |z - \zeta| \ge \rho\}$ (the last equality follows because the integral is absolutely convergent). Since $1/(z - \zeta)$ is analytic in Q_ρ, it follows from (9.6) and the product rule that

$$\langle \bar{\partial} f, \phi \rangle = -\frac{1}{\pi} \lim_{\rho \downarrow 0} \iint_{Q_\rho} \bar{\partial}\left(\frac{\phi(z)}{z - \zeta}\right) dx\, dy. \tag{9.7}$$

Applying Lemma 9.3 to the right hand side of (9.7) and using the fact that ϕ is supported in the interior of Q, we obtain

$$\begin{aligned}\langle \bar{\partial} f, \phi \rangle &= \frac{i}{2\pi} \lim_{\rho \downarrow 0} \int_{\partial(Q_\rho)} \frac{\phi(z)}{z - \zeta} dz \\ &= \frac{1}{2\pi} \lim_{\rho \downarrow 0} \int_0^{2\pi} \phi(\zeta + \rho e^{i\theta})\, d\theta = \phi(\zeta),\end{aligned}$$

where the last equality follows from the continuity of ϕ. This establishes (9.4).

If ϕ is a test function and G is analytic, then (9.4) and the product rule give

$$\langle \bar{\partial}(Gf), \phi \rangle = -\langle fG, \bar{\partial}\phi \rangle = -\langle f, \bar{\partial}(G\phi) \rangle = G(\zeta)\phi(\zeta),$$

which is exactly the content of (9.5). ■

With these preliminaries, we can begin to formulate Jones' constructive solution to the equation $\bar{\partial} f = \mu$, where μ is a Carleson measure. Some additional notation will be needed. When $\zeta \in \mathbf{U}$, we shall denote by $B(\zeta)$ the strip in $\mathbf{U}$ of points that lie below ζ:

$$B(\zeta) = \{w \in \mathbf{U} : 0 \le \operatorname{Im}(w) \le \operatorname{Im}(\zeta)\}, \qquad \zeta \in \mathbf{U}.$$

Define kernels $k_1(z, \zeta)$ and $k_2(z, \zeta)$ in $\mathbf{U} \times \mathbf{U}$ by

$$\begin{aligned} k_1(z,\zeta) &= \frac{2i \operatorname{Im}(\zeta)}{z - \bar{\zeta}}, \\ k_2(z,\zeta) &= \exp\left((i-1)\sqrt{\frac{z - \operatorname{Re}(\zeta)}{\operatorname{Im}(\zeta)}} + \sqrt{2}\right). \end{aligned} \tag{9.8}$$

In the latter expression, the square root has a branch cut along the negative imaginary axis. Thus, for a complex number represented in polar form $re^{i\theta}$, where $r > 0$ and $-\pi/2 \le \theta < 3\pi/2$, the square root is taken to be the complex number $\sqrt{r}e^{i\theta/2}$. In particular, the complex number $i - 1$ in (9.8) may be viewed in this way as $i - 1 = -\sqrt{2}(-i)^{1/2}$.

Since $\operatorname{Im}(z)$ and $\operatorname{Im}(\zeta)$ are nonnegative, the kernels k_1 and k_2 in (9.8) are related by

$$|k_2(z,\zeta)| \le e^{\sqrt{2}} |k_1(z,\zeta)|. \tag{9.9}$$

To see this, note first that the quantity

$$w = (i-1)\sqrt{\frac{z - \operatorname{Re}(\zeta)}{\operatorname{Im}(\zeta)}}$$

lies in the sector $3\pi/4 \le \arg w \le 5\pi/4$, so $\operatorname{Re} w \le -|w|/\sqrt{2}$. Hence, using the fact that $e^{-t} \le 2/(1 + t^2)$ for $t > 0$, we have

$$\begin{aligned}|k_2(z,\zeta)| &= \exp(\operatorname{Re} w + \sqrt{2}) \le \exp\left(-\sqrt{\frac{|z - \operatorname{Re}(\zeta)|}{\operatorname{Im}(\zeta)}} + \sqrt{2}\right) \\ &\le e^{\sqrt{2}} \frac{2}{1 + |z - \operatorname{Re}(\zeta)|/\operatorname{Im}(\zeta)} \le e^{\sqrt{2}} \frac{2\operatorname{Im}(\zeta)}{|z - \bar{\zeta}|},\end{aligned}$$

which establishes (9.9).

We also define a kernel $K(z, \zeta)$ (corresponding to a Carleson measure μ) on $\mathbf{U} \times \mathbf{U}$ by

$$K(z,\zeta) = \exp\left\{-i \iint_{B(\zeta)} \left(\frac{1}{z - \bar{w}} - \frac{1}{\zeta - \bar{w}}\right) \frac{d|\mu|(w)}{\|\mu\|_C}\right\}. \tag{9.10}$$

Note that when $z = \zeta$, each of the kernels k_1, k_2, and K has value 1.

Theorem 9.5 (P. Jones). *Let μ be a Carleson measure on $\mathbf{U}$. Then the functions defined for $j = 1, 2$ by*

$$f_j(z) = \frac{1}{\pi} \iint_{\mathbf{U}} \frac{1}{z - \zeta} k_j(z,\zeta) K(z,\zeta)\, d\mu(\zeta) \tag{9.11}$$

satisfy the distributional equation

$$\bar{\partial} f_j = \mu \tag{9.12}$$

and the estimate

$$\|f_j\|_{L^\infty(\mathbf{R})} \le c\|\mu\|_C. \tag{9.13}$$

Moreover, if I is an interval centered at a point x_0 and if $Q = I \times (0, |I|)$, then the function defined by

$$f(z; Q) = \frac{1}{\pi} \iint_Q \frac{1}{z - \zeta} k_2(z,\zeta) K(z,\zeta)\, d\mu(\zeta) \tag{9.14}$$

satisfies $\bar{\partial} f(\cdot\,; Q) = \mu_Q$ (the restriction of μ to Q) and

$$|f(x; Q)| \le c\|\mu\|_C \exp\left\{-\sqrt{\frac{|x - x_0|}{|I|}}\right\}, \qquad (x \in \mathbf{R}). \tag{9.15}$$

Theorem 9.5 will be established through a series of technical lemmas. The first (Lemma 9.6) shows that the defining integral (9.10) for the kernel $K(z, \zeta)$ is absolutely convergent for each z and ζ in $\mathbf{U}$, and consequently that $K(z,\zeta)$ is analytic in both variables. The appearance of this kernel in the integrands (9.11) ensures that the integrals converge and hence that the functions f_j are well-defined. Throughout the proof of Theorem 9.5, we may assume without loss of generality that μ is a Carleson measure of norm one.

Lemma 9.6. *For each z and ζ in* $\mathbf{U}$, *the integral*

$$I(z,\zeta) = \iint_{B(\zeta)} \left(\frac{1}{z-\bar{w}} - \frac{1}{\zeta - \bar{w}}\right) d|\mu|(w) \tag{9.16}$$

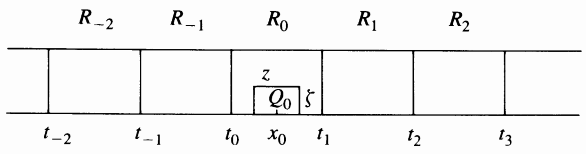
is absolutely convergent.

Proof. Let Q_0 be any closed cube whose lower edge is an interval in $\mathbf{R}$, which contains both z and ζ, and which has minimal side length subject to these constraints (see Figure 12). Let h be the side length of Q_0 and $(x_0, h/2)$ its center. Set

$$t_j = x_0 + (2j-1)h, \qquad (-\infty < j < \infty)$$

and define squares R_j by

$$R_j = (t_j, t_{j+1}) \times (0, 2h), \qquad (-\infty < j < \infty).$$



We have

$$\begin{aligned}
|I(z,\zeta)| &\le \iint_{B(\zeta)} \frac{|z-\zeta|}{|z-\bar{w}|\cdot|\zeta-\bar{w}|}\, d|\mu|(w) \\
&\le |z-\zeta| \sum_j \iint_{R_j} \frac{d|\mu|(w)}{|z-\bar{w}|\cdot|\zeta-\bar{w}|} \\
&\le |z-\zeta| \left(\frac{|\mu|(R_0)}{\operatorname{Im}(z)\operatorname{Im}(\zeta)} + \sum_{j\ne 0} \frac{|\mu|(R_j)}{(jh/2)^2}\right) \\
&\le |z-\zeta| h \left(\frac{1}{\operatorname{Im}(z)\operatorname{Im}(\zeta)} + \frac{8}{h^2}\sum_{j=1}^{\infty}\frac{1}{j^2}\right) < \infty,
\end{aligned}$$

as claimed. This completes the proof. ■

The next lemma will be used to show that f_1 and f_2 are locally integrable in $\mathbf{U}$.

Lemma 9.7. *There is a constant $c_0 > 0$ such that, for all z satisfying* $\operatorname{Im}(z) \ge 0$,

$$\iint_{\mathbf{U}} \frac{\operatorname{Im}(\zeta)}{|z-\bar{\zeta}|^2} \exp\left(-\iint_{B(\zeta)} \frac{\operatorname{Im}(w)}{|z-\bar{w}|^2}\, d|\mu|(w)\right) d|\mu|(\zeta) \le C_0. \tag{9.17}$$

Proof. Fix z with $\operatorname{Im}(z) \ge 0$. Define rectangles S_j, $|j| \le N$, by

$$S_j = \{(x,y) : |x| \le 2^N,\ 2^{j-1} < y \le 2^j\},$$

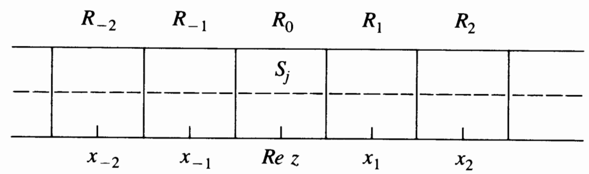
and let

$$\alpha_j = \iint_{S_j} \frac{\operatorname{Im}(\zeta)}{|z-\bar{\zeta}|^2}\, d|\mu|(\zeta). \tag{9.18}$$

Then

$$0 \le \alpha_j \le 2^j \iint_{S_j} \frac{d|\mu|(\zeta)}{|\operatorname{Re}(z) - \bar{\zeta}|^2} \le 2^j \sum_{k=-\infty}^{\infty} \iint_{R_k \cap S_j} \frac{d|\mu|(\zeta)}{|\operatorname{Re}(z) - \bar{\zeta}|^2}$$

where R_k, $(-\infty < k < \infty)$, is the square with side length 2^j and center $x_k = \operatorname{Re}(z) + k2^j$ (see Figure 13).



But then

$$\begin{aligned}
\alpha_j &\le 2^j \left(|\mu|(R_0)\frac{1}{(2^{j-1})^2} + \sum_{k\ne 0} |\mu|(R_k)\frac{1}{(k2^{j-1})^2}\right) \\
&\le \|\mu\|_C \left(4 + 2\sum_{k=1}^{\infty}\frac{4}{k^2}\right) \equiv c, \qquad (|j| \le N),
\end{aligned} \tag{9.19}$$

since we have the standing assumption that $\|\mu\|_C = 1$.

Let $S^N = \bigcup_{|j|\le N} S_j$. Then clearly $S^N \uparrow \mathbf{U}$ as $N \uparrow \infty$. To prove (9.17), we show that the quantity

$$I(N) = \iint_{S^N} \frac{\operatorname{Im}(\zeta)}{|z-\bar{\zeta}|^2} \exp\left(-\iint_{B(\zeta)} \frac{\operatorname{Im}(w)}{|z-\bar{w}|^2}\, d|\mu|(w)\right) d|\mu|(\zeta)$$

is uniformly bounded (independent of N and z). Set $\tau_{-N-1} = 0$ and $\tau_j = \sum_{-N \le k \le j} \alpha_k$. Then, with c as in (9.19),

$$I(N) \le \sum_{j=-N}^{N} \iint_{S_j} \frac{\operatorname{Im}(\zeta)}{|z-\bar{\zeta}|^2} \exp\left(-\iint_{\cup_{-N}^{j-1} S_k} \frac{\operatorname{Im}(w)}{|z-\bar{w}|^2}\, d|\mu|(w)\right) d|\mu|(\zeta)$$
$$= \sum_{j=-N}^{N} \alpha_j \exp\left(-\sum_{k=-N}^{j-1} \alpha_k\right) = \sum_{j=-N}^{N} \exp(\alpha_j)\exp(-\tau_j)\cdot \Delta\tau_j$$
$$\le e^c \int_0^\infty e^{-t}\, dt = e^c \equiv C_0,$$

since the last sum is a lower Riemann sum for e^{-t}. ∎

Lemma 9.8. *For each $\zeta \in \mathbf{U}$, let*

$$c_\zeta = \exp\left(i \iint_{B(\zeta)} \frac{1}{\zeta - \bar{w}}\, d|\mu|(w)\right). \tag{9.20}$$

Then there is a constant c such that

$$|c_\zeta| \le c, \qquad (\zeta \in \mathbf{U}). \tag{9.21}$$

Proof. We have

$$\operatorname{Re}\left(i \iint_{B(\zeta)} \frac{1}{\zeta-\bar{w}}\, d|\mu|(w)\right) = \iint_{B(\zeta)} \frac{\operatorname{Im}\zeta + \operatorname{Im} w}{|\zeta - \bar{w}|^2}\, d|\mu|(w) \le 2\iint_{\mathbf{U}} \frac{\operatorname{Im}\zeta}{|\zeta-\bar{w}|^2}\, d|\mu|(w).$$

The function $G_\zeta(w) = \operatorname{Im}(\zeta)/[w - \bar{\zeta}]^2$ has H^1-norm equal to π. Hence, by Theorem 9.2, the last term is bounded by a constant, c' say, so $|c_\zeta| \le e^{c'}$. ∎

Lemma 9.9. *Define H on $\mathbf{U}$ by*

$$H(z) = \frac{2}{\pi} \iint_{\mathbf{U}} \frac{1}{|z-\zeta|} \frac{\operatorname{Im}(\zeta)}{|z-\bar{\zeta}|} |K(z,\zeta)|\, d|\mu|(\zeta). \tag{9.22}$$

Then H is locally integrable in $\mathbf{U}$ and

$$\|H\|_{L^\infty(\mathbf{R})} \le c_0, \tag{9.23}$$

where the constant is independent of μ (with $\|\mu\|_C = 1$).

Proof. To prove that H is locally integrable in $\mathbf{U}$, we select an arbitrary square $Q = I \times (0, |I|)$ in $\mathbf{U}$ and show that

$$\frac{1}{|Q|} \iint_Q H(z)\, dx\, dy \le c_0. \tag{9.24}$$

Note that (9.23) follows immediately from (9.24).

We have from (9.10),

$$|K(z,\zeta)| \le |c_\zeta| \exp\left[\operatorname{Re}\left(-i \iint_{B(\zeta)} \frac{d|\mu|(w)}{z - \bar{w}}\right)\right], \tag{9.25}$$

with c_ζ given by (9.20). Combining this with (9.21), we obtain

$$|K(z,\zeta)| \le c \exp\left(-\iint_{B(\zeta)} \frac{\operatorname{Im}(z+w)}{|z-\bar{w}|^2}\, d|\mu|(w)\right). \tag{9.26}$$

Let $\tilde{Q}$ be the square $\tilde{Q} = (2I) \times (0, 2|I|)$ (which contains Q) and set $\tilde{Q}^c = \mathbf{U} \setminus \tilde{Q}$. Using (9.26) and Fubini's theorem, we may write

$$\iint_Q H(z)\, dx\, dy \le c \iint_{\mathbf{U}} \left(\iint_Q \frac{\operatorname{Im}(\zeta)}{|z-\zeta|\cdot|z-\bar{\zeta}|} e(z,\zeta)\, dx\, dy\right) d|\mu|(\zeta), \tag{9.27}$$

where

$$e(z,\zeta) = \exp\left(-\iint_{B(\zeta)} \frac{\operatorname{Im}(z+w)}{|z-\bar{w}|^2}\, d|\mu|(w)\right).$$

The integral on the right of (9.27) can now be split into two integrals, J_1 and J_2, say, corresponding to integration over $\tilde{Q}$ and $\tilde{Q}^c$, respectively.

Since $\operatorname{Im}(\zeta) \le |z - \bar{\zeta}|$ and the argument of the exponential is negative, we have

$$J_1 \le c \iint_{\tilde{Q}} \left(\iint_Q \frac{1}{|z-\zeta|}\, dx\, dy\right) d|\mu|(\zeta)$$
$$\le c|I| \cdot |\mu|(\tilde{Q}) \le c|Q|. \tag{9.28}$$

Here, the inner integral is evaluated by a change to polar coordinates, and the latter inequality results from the fact that $|\mu|$ is a Carleson measure.

To estimate J_2, we use the elementary inequality

$$|z - \bar{\zeta}| \le 5|z - \zeta|, \qquad (z \in Q, \zeta \in \tilde{Q}^c) \tag{9.29}$$

to obtain

$$|J_2| \le c \iint_{\tilde{Q}^c} \left[\iint_Q \frac{\operatorname{Im}(\zeta)}{|z-\bar{\zeta}|^2} e(z,\zeta)\, dx\, dy\right] d|\mu|(\zeta)$$
$$\le c \iint_Q \left[\iint_{\mathbf{U}} \frac{\operatorname{Im}(\zeta)}{|z-\bar{\zeta}|^2} e(z,\zeta)\, d|\mu|(\zeta)\right] dx\, dy.$$

Lemma 9.7 shows that the integral inside the brackets is uniformly bounded and so $|J_2| \le c|Q|$. Combining this estimate with (9.28), we obtain the desired inequality (9.24). ∎

Corollary 9.10. *If W is a bounded measurable function on* $\mathbf{U}$, *then, for each x in* $\mathbf{R}$,

$$\iint_{\mathbf{U}} \frac{\operatorname{Im}(\zeta)}{|x-\zeta|^2} |K(x,\zeta)| \cdot |W(\zeta)|\, d|\mu|(\zeta) \le c_0 \|W\|_{L^\infty(\mathbf{U})}.$$

Since $|f_1(z)| \le H(z)$, Lemma 9.9 shows that f_1 is locally integrable in $\mathbf{U}$ and that (9.13) holds (recall once again that we are assuming $\|\mu\|_C = 1$). Moreover, inequality (9.9) implies that $|f_2(z)| \le \text{const} \cdot H(z)$, and so the same statements apply also to f_2. The next lemma shows that the functions f_j solve the desired $\bar\partial$ equation.

Lemma 9.11. *The functions f_1 and f_2 defined by* (9.11) *satisfy the distributional equation*

$$\bar\partial f_j = \mu, \qquad (j = 1,2). \tag{9.30}$$

Proof. We prove (9.30) only for $f = f_2$ since the proof for f_1 is similar. Let ϕ be a test function. By definition of the distributional derivative,

$$\begin{aligned}\langle \bar\partial f_2, \phi\rangle &= -\iint_{\mathbf{U}} f_2(z)\bar\partial\phi(z)\,dx\,dy \\ &= \iint_{\mathbf{U}}\left(-\frac{1}{\pi}\iint_{\mathbf{U}} G_\zeta(z)\frac{1}{z-\zeta}\bar\partial\phi(z)\,dx\,dy\right)d\mu(\zeta), \qquad (9.31)\end{aligned}$$

where we have set

$$G_\zeta(z) = \exp\left((i-1)\sqrt{\frac{z-\operatorname{Re}(\zeta)}{\operatorname{Im}(\zeta)}} + \sqrt{2}\right)K(z,\zeta) = k_2(z,\zeta)K(z,\zeta).$$

The second equality in (9.31) is immediate since the integrand is integrable over the support of ϕ and so Fubini's theorem may be applied. Next, observe that $G_\zeta(z)$ is analytic in z. Hence, by Lemma 9.4,

$$\langle \bar\partial f_2, \phi\rangle = \iint_{\mathbf{U}} G_\zeta(\zeta)\phi(\zeta)\,d\mu(\zeta) = \iint_{\mathbf{U}} \phi(\zeta)\,d\mu(\zeta) = \langle \mu, \phi\rangle, \tag{9.32}$$

which establishes the desired result. ■

To complete the proof of Theorem 9.5, we need to verify inequality (9.15) for the function $f(\cdot\,; Q)$ defined in (9.14). Again, we are assuming that $\|\mu\|_C = 1$. Let $\tilde I = 3 \cdot I$, where $Q = I \times (0, |I|)$, and recall that x_0 is the center of I. It follows from (9.14) and (9.9) that $|f(z; Q)| \le e^{\sqrt{2}} H(z)$, where H is the function defined by (9.22). It is then clear from (9.23) that the estimate (9.15) holds for $x \in \tilde I$.

It remains therefore to verify (9.15) for $x \notin \tilde I$. In this case, with $h = |I|$, we have $|x - \zeta| \ge h$ for any $\zeta \in Q$. Hence,

$$|f(x;Q)| \le \frac{1}{\pi h}\iint_Q \exp\left(-\sqrt{\frac{|x-\operatorname{Re}(\zeta)|}{\operatorname{Im}(\zeta)}} + \sqrt{2}\right)|K(x,\zeta)|\,d|\mu|(\zeta). \tag{9.33}$$

The exponential term may be estimated as follows. We have

$$\frac{|x-\operatorname{Re}(\zeta)|}{h} \ge \frac{|x-x_0|}{h} - \frac{|x_0-\operatorname{Re}(\zeta)|}{h} \ge \frac{|x-x_0|}{h} - \frac{1}{2}.$$

Since $0 < \operatorname{Im}(\zeta) \le h$, we therefore have

$$\sqrt{\frac{|x-\operatorname{Re}(\zeta)|}{\operatorname{Im}(\zeta)}} + \frac{1}{\sqrt{2}} \ge \sqrt{\frac{|x-x_0|}{h}}.$$

Inequality (9.26) shows that $|K(x,\zeta)|$ is bounded, so the preceding estimate applied in (9.33) gives

$$\begin{aligned}|f(x;Q)| &\le \frac{c}{h}\iint_Q \exp\left(-\sqrt{\frac{|x-x_0|}{h}}\right)d|\mu|(\zeta) \\ &\le c\exp\left(-\sqrt{\frac{|x-x_0|}{h}}\right), \qquad (x \notin \tilde I).\end{aligned}$$

This establishes (9.15) and hence completes the proof. ■

10. INTERPOLATION BETWEEN H^1 AND H^∞

As a first application of the $\bar\partial$-methods developed in the last section, we obtain the following description of the K-functional for the couple (H^1, H^∞) of analytic Hardy spaces.

Theorem 10.1. *There are positive constants c_1 and c_2 such that, for each* $F \in (H^1 + H^\infty)(\mathbf{U})$,

$$c_1 K(f,t) \le \int_0^t (NF)^*(s)\,ds \le c_2 K(f,t), \qquad (t > 0). \tag{10.1}$$

Here, $K(F,t) = K(F,t; H^1(\mathbf{U}), H^\infty(\mathbf{U}))$ is the Peetre K-functional for the pair H^1 and H^∞, and N denotes the nontangential maximal operator.

Proof. The right hand inequality in (10.1) follows immediately from the subadditivity of N and of the operation $f \to f^{**}$ (cf. Theorem II.3.4). Indeed, if $F = B + G$, with $B \in H^1$ and $G \in H^\infty$, then

$$\int_0^t (NF)^* \le \int_0^t (NB + NG)^* \le \int_0^t (NB)^* + \int_0^t (NG)^*$$
$$\le \|NB\|_{L^1(\mathbf{R})} + t\|NG\|_{L^\infty(\mathbf{R})}.$$

The Hardy-Littlewood theorem (Theorem 6.2) shows that N is a bounded operator from $H^1(\mathbf{U})$ into $L^1(\mathbf{R})$, and it is clear from the definition that N is bounded from $H^\infty(\mathbf{U})$ into $L^\infty(\mathbf{R})$. Hence,

$$\int_0^t (NF)^* \le c(\|B\|_{H^1} + t\|G\|_{H^\infty}).$$

Taking the infimum over all decompositions $F = B + G$, we obtain the right hand inequality in (10.1).

For the left hand inequality, fix $t > 0$ and let

$$M = \int_0^t (NF)^*(s)\,ds.$$

We shall use Jones' theorem (Theorem 9.5) to approximate F by an H^∞-function G such that

$$\|F - G\|_{H^1} \le cM \tag{10.2}$$

and

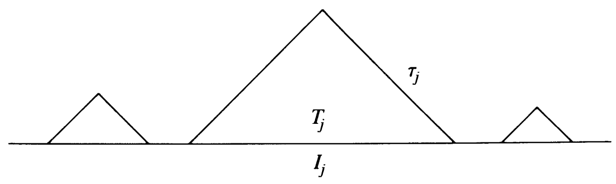
$$t\|G\|_{H^\infty} \le cM. \tag{10.3}$$

Let $\alpha = (NF)^*(t)$ and write the open set

$$\mathcal{O} = \mathcal{O}_\alpha = \{x \in \mathbf{R} : NF(x) > \alpha\}$$

as a countable disjoint union of intervals $\{I_j\}$. For each such interval I_j, construct a right triangle T_j in the upper half-plane $\mathbf{U}$ with hypotenuse I_j (see Figure 14). Let $R = R_\alpha$ denote the disjoint union $\bigcup_j T_j$, and denote by τ_j the two edges of T_j that lie in $\mathbf{U}$. Let

$$b = F\chi_R, \qquad g = F\chi_{R^c},$$



where $R^c = \mathbf{U} \setminus R$. Then $b + g = F$ and $\|g\|_{L^\infty(\mathbf{U})} \le \alpha$. Although the function g will not in general be analytic, it does have the property that $\bar\partial g$ is a Carleson measure supported on $\bigcup_j \tau_j$, with

$$\|\bar\partial g\|_C \le \alpha. \tag{10.4}$$

To see this, observe first that the analyticity of F implies that $\bar\partial g = -\bar\partial b$. Thus, if ϕ is a test function in $\mathbf{U}$,

$$\langle \bar\partial g, \phi \rangle = \langle -\bar\partial b, \phi \rangle = \sum_j \iint_{T_j} F \bar\partial \phi \,dx\,dy$$
$$= \sum_j \iint_{T_j} \bar\partial (F\phi)\,dx\,dy. \tag{10.5}$$

Applying Lemma 9.3, we have

$$\iint_{T_j} \bar\partial (F\phi)\,dx\,dy = \frac{1}{2i}\int_{\tau_j} F\phi\,dz,$$

because the integral over the lower boundary vanishes since ϕ is supported in $\mathbf{U}$. We may therefore rewrite (10.5) in the form

$$\langle \bar\partial g, \phi \rangle = \frac{1}{2i}\sum_j \int_{\tau_j} F\phi\,dz.$$

Hence, $\bar\partial g$ is a Carleson measure supported on $\bigcup_j \tau_j$, with

$$\|\bar\partial g\|_C \le \frac{1}{\sqrt{2}} \sup_{z \in \bigcup_j \tau_j} |F(z)| \le \alpha.$$

By Jones' theorem (Theorem 9.5), there is a function f_2, defined according to (9.11), such that

$$\bar\partial f_2 = \bar\partial g, \qquad \|f_2\|_{L^\infty(\mathbf{R})} \le c\alpha. \tag{10.6}$$

Let

$$G = g - f_2, \qquad B = b + f_2.$$

Then $F = B + G$ and both B and G are analytic in $\mathbf{U}$. Thus, it remains only to prove (10.2) and (10.3).

For any cube $Q = I \times (0, |I|)$, the estimate (9.24) gives

$$\frac{1}{|Q|}\iint_Q (|f_2| + g) \le c\alpha.$$

Hence,

$$t\|G\|_{H^\infty} \le c\alpha t \le c\int_0^t (NF)^* = cM,$$

which establishes (10.3). For (10.2), write $B = b + \sum_j f(\cdot; Q_j)$, where the functions $f(\cdot; Q_j)$ are given by (9.14) with $\mu = \bar{\partial} g$. Then

$$\|B\|_{H^1(\mathbf{U})} \le \int_{\mathcal{O}} |F| + \sum_j \|f(\cdot; Q_j)\|_{L^1(\mathbf{R})}. \tag{10.7}$$

But, by (9.15) (where x_j denotes the center of I_j),

$$\|f(\cdot; Q_j)\|_{L^1} \le c\alpha \int_{-\infty}^{\infty} \exp\left(-\sqrt{\frac{|x - x_j|}{|I_j|}}\right) dx \le c\alpha |I_j|.$$

Therefore, since $\mathcal{O} = \bigcup_j I_j$, we obtain

$$\sum_j \|f(\cdot; Q_j)\|_{L^1} \le c\alpha|\mathcal{O}| \le ct(NF)^*(t) \le cM. \tag{10.8}$$

It is clear from the definitions of $\mathcal{O}$ and M that

$$\int_{\mathcal{O}} NF \le \int_0^t (NF)^*(s)\, ds = M.$$

Hence, (10.7) and (10.8) combine to give (10.2), and this completes the proof. ■

The following alternative description of the K-functional will be useful in describing the interpolation spaces between H^1 and H^∞. The proof is based on the factorization technique of Hardy and Littlewood (see the proof of Theorem 6.2).

Corollary 10.2. *Let $F \in (H^1 + H^\infty)(\mathbf{U})$ and let f be the boundary function of F (the restriction of F to $\mathbf{R}$). Then there exist positive constants c_1, c_2, independent of F, such that*

$$c_1 K(F, t) \le \int_0^t f^*(s)\, ds \le c_2 K(F, t), \qquad (t > 0), \tag{10.9}$$

where $K(F, t) = K(F, t; H^1(\mathbf{U}), H^\infty(\mathbf{U}))$.

Proof. The right-hand inequality follows immediately from (10.1) since $|f| \le Nf = NF$ a.e.. The inequality may also be obtained directly from the subadditivity of $f \to f^{**}$, together with the facts

$$\int_0^t g^* \le \|g\|_{L^1} = \|G\|_{H^1}, \qquad \int_0^t h^* \le t\|h\|_{L^\infty} = t\|H\|_{H^\infty},$$

where g and h are the boundary functions of the analytic functions G and H, respectively.

To establish the left-hand inequality, write $F = BG^2$, where B is the Blaschke product formed from the zeros of F, and G is the square root of the remaining zero-free analytic function on $\mathbf{U}$. As before, let g be the boundary function of G on $\mathbf{R}$. Then $NF = Nf = (Ng)^2$, and so

$$\int_0^t (NF)^*(s)\, ds = \int_0^t (NG)^*(s)^2\, ds \le c\int_0^t (Mg)^*(s)^2\, ds, \tag{10.10}$$

since $N(g) = N(G)$ is majorized by a constant multiple of the Hardy-Littlewood maximal function $M(g)$ (cf. Exercise III.9.c). Theorem III.3.8 shows that there is a constant $c > 0$ such that $(Mg)^* \le cg^{**}$, so from (10.10) we have

$$\int_0^t (NF)^*(s)\, ds \le c\int_0^t [g^{**}(s)]^2\, ds. \tag{10.11}$$

A Hardy inequality (similar, for example, to Lemma III.3.9) applied to the right hand side of (10.11) gives

$$\int_0^t [g^{**}(s)]^2\, ds \le c\int_0^t g^*(s)^2\, ds. \tag{10.12}$$

Now $|g|^2 = |f|$ a.e., so $(g^*)^2 = f^*$. Hence, combining (10.11) and (10.12), we see from Theorem 10.1 that the left hand side of (10.9) holds. This completes the proof. ■

Since the Hilbert transform is bounded on $L^{p,q}(1 < p < \infty)$, the preceding result shows that the (θ, q)-interpolation spaces of (H^1, H^∞) may be identified with Lorentz spaces. In particular, if $0 < \theta < 1$, $1 \le q \le \infty$, and $1/p = 1 - \theta$,

$$(H^1(\mathbf{U}), H^\infty(\mathbf{U}))_{\theta,q} = L^{p,q}(\mathbf{R}).$$

We saw in Theorem III.2.10 for the couple (L^1, L^∞) that if $g \prec f$, there is an admissible operator T for which $Tf = g$. The analogous result for the couple (H^1, H^∞) will be established in Theorem 10.6 below. The following lemma will be needed.

Lemma 10.3. *Let ψ and ϕ be nonnegative measurable functions on $\mathbf{R}$. Suppose ψ assumes a constant value β on some set E of finite measure and $\phi^*(+\infty) = 0$. If $\psi \prec \phi$, there is a measurable set F with $|F| = |E|$ such that*

$$\psi\chi_E \prec \phi\chi_F \tag{10.13}$$

and

$$\psi\chi_{E^c} \prec \phi\chi_{F^c}. \tag{10.14}$$

Proof. The hypotheses $\psi \prec \phi$ and $\phi^*(+\infty) = 0$ imply $\psi^*(+\infty) = 0$. Hence, by considering the appropriate measure-preserving tranformations (cf. Lemma II.7.3, Corollary II.7.6, and Exercise II.17), it is clear that the lemma need only be established under the additional hypotheses that $\psi = \psi^*$, $\phi = \phi^*$, and E is an interval $I = (a,b) \subset (0,\infty)$.

Let $J = (c,d)$ be an interval of length $\delta = |I|$ such that

$$\int_J \phi = \int_I \psi = \beta\delta. \tag{10.15}$$

Such a choice is possible because the continuous function

$$\eta(t) = \int_t^{t+\delta} \phi - \beta\delta$$

satisfies $\eta(0) \geq 0$ and $\eta(+\infty) = -\beta\delta < 0$.

To see that (10.13) holds (with $E = I$ and $F = J$), note that $(1/t)\int_c^{c+t} \phi$ is a decreasing function of t. Its smallest value on the interval $0 < t \leq \delta$ therefore occurs at $t = \delta$ and this value is β. Hence, if $0 < t \leq \delta$,

$$\int_0^t (\phi\chi_J)^* = \int_c^{c+t} \phi \geq \beta t = \int_0^t (\psi\chi_I)^*.$$

The corresponding estimate for $t > \delta$ follows from (10.15) since the functions involved are supported in sets of measure δ.

To establish (10.14), first denote by $\tilde{\phi}$ and $\tilde{\psi}$ the decreasing rearrangements $(\phi\chi_{J^c})^*$ and $(\psi\chi_{I^c})^*$, respectively:

$$\tilde{\phi}(t) = \phi(t)\chi_{(0,c)}(t) + \phi(t+\delta)\chi_{[c,\infty)}(t)$$

$$\tilde{\psi}(t) = \psi(t)\chi_{(0,a)}(t) + \psi(t+\delta)\chi_{[a,\infty)}(t).$$

We need to show that

$$\int_0^t \tilde{\psi} \leq \int_0^t \tilde{\phi} \tag{10.16}$$

for all $t > 0$.

Case 1 $a \leq c$. If $t \leq c$, then (10.16) follows from the hypothesis $\psi \prec \phi$ and the fact that $\tilde{\psi} \leq \psi$. On the other hand, if $t \geq c$, then

$$\int_0^t \tilde{\psi} = \int_0^{t+\delta} \psi - \beta\delta \leq \int_0^{t+\delta} \phi - \beta\delta = \int_0^t \tilde{\phi}, \tag{10.17}$$

which again yields (10.16).

Case 2 $c < a$. If $0 < t < c$, then (10.16) follows from the hypothesis $\psi \prec \phi$ and the fact that $\phi = \tilde{\phi}$ and $\psi = \tilde{\psi}$ on $(0,c)$. If $t \geq a$, the desired result follows as before from (10.17). In the remaining case $c < t < a$, we have $\psi \geq \beta$ on (c,b) so

$$\int_0^t \tilde{\psi} = \int_0^t \psi \leq \int_0^{t+\delta} \psi - \beta\delta \leq \int_0^{t+\delta} \phi - \beta\delta = \int_0^t \tilde{\phi}.$$

Hence, (10.16) holds in all cases and the proof is complete. ■

Remark. If ϕ and ψ are supported in a finite interval $(0,L)$, then J may be chosen as a subset of $(0,L)$. With the construction given above, it may be impossible to choose J satisfying (10.15) because the integral of ϕ^* over $(L-\delta, L)$ may exceed $\beta\delta$. In this case, we merely replace ϕ by $\phi_0 = \phi\chi_{(0,\lambda)}$, where λ is chosen so that

$$\int_0^\lambda \phi = \int_0^L \psi.$$

Then $\phi \geq \phi_0$ and we may apply the preceding proof to ϕ_0 and ψ to obtain an interval $J \subset (0,L)$ with the desired properties.

Corollary 10.4. *Suppose ψ is of the form*

$$\psi = \sum_j \beta_j \chi_{E_j},$$

where $\beta_j \geq 0$ and $\{E_j\}$ is a collection of pairwise disjoint subsets of **R** *of finite measure. Suppose ϕ is nonnegative and*

$$\phi^*(+\infty) \equiv \lim_{t\to\infty} \phi^*(t) = 0. \tag{10.18}$$

If $\psi \prec \phi$, then there exist pairwise disjoint measurable sets F_j with $|F_j| = |E_j|$ and

$$\beta_j |E_j| \leq \int_{F_j} \phi. \tag{10.19}$$

Proof. The proof uses repeated applications of Lemma 10.3. We begin with $E = E_1$ and $\beta = \beta_1$, in which case the lemma provides a set F_1 with $|F_1| = |E_1|$ such that

$$\beta_1 \chi_{E_1} \prec \phi\chi_{F_1}, \qquad \psi_2 \prec \phi_2$$

where

$$\psi_2 = \sum_{k\geq 2} \beta_k \chi_{E_k}, \qquad \phi_2 = \phi\chi_{F_1^c}.$$

At the j-th stage, we have functions

$$\psi_j = \sum_{k\geq j} \beta_k \chi_{E_k}, \qquad \phi_j = \phi\chi_{(\cup_{i=1}^{j-1} F_i)^c},$$

which satisfy $\psi_j \prec \phi_j$. Application of Lemma 10.3 (with $E = E_j$ and $\beta = \beta_j$) produces a set F_j, disjoint from $F_1, F_2, \ldots, F_{j-1}$, such that $|F_j| = |E_j|$ and

$$\beta_j \chi_{E_j} \prec \phi \chi_{F_j}. \tag{10.20}$$

The desired estimate (10.19) follows at once from (10.20). ■

Corollary 10.5. *Suppose ψ is of the form*

$$\psi = \sum_j \beta_j \chi_{E_j},$$

where $\beta_j \geq 0$ and $\{E_j\}$ is a collection of pairwise disjoint subsets of $\mathbf{R}$ of finite measure. Let ϕ be a nonnegative function for which $\psi \prec \phi$ and let $\varepsilon > 0$. Then there exist pairwise disjoint measurable sets F_j with $|F_j| = |E_j|$ and

$$\beta_j |E_j| \leq (1 + \varepsilon) \int_{F_j} \phi. \tag{10.21}$$

Proof. The result follows from Corollary 10.4 if ϕ satisfies condition (10.18). We may suppose therefore that $\phi^*(+\infty) \equiv \alpha > 0$. The hypothesis $\psi \prec \phi$ then implies that $\psi^*(+\infty) \leq \alpha$ and so the set $E = \{x : \psi(x) > \alpha + \delta\}$ has finite measure $|E|$. Here we have chosen $\delta = \alpha\varepsilon/(2 + \varepsilon)$ so that $1 + \varepsilon = (\alpha + \delta)/(\alpha - \delta)$.

The set $F' = \{x : \phi(x) > \alpha + \delta\}$ also has finite measure. It may be, however, that $|F'| < |E|$, and in this case we shall need to augment F'. To this end, note that $\{\alpha - \delta \leq \phi(x) \leq \alpha + \delta\}$ has infinite measure so we may select a subset F'' whose measure is $|E| - |F'|$. Then $F = F' \cup F''$ has measure $|F| = |E| < \infty$ and

$$(\phi\chi_F)^* = \begin{cases} \phi^*, & \text{on } (0, |F'|), \\ \eta, & \text{on } (|F'|, |F|), \\ 0, & \text{on } (|F|, \infty), \end{cases}$$

where η is a measurable function satisfying $\alpha - \delta \leq \eta \leq \alpha + \delta$. Since ϕ^* satisfies the same estimates on $(|F'|, |F|)$, it is not difficult to check that $\psi\chi_E \prec ((\alpha + \delta)/(\alpha - \delta))\phi\chi_F = (1 + \varepsilon)\phi\chi_F$. Since these functions are supported in sets of finite measure, we may then select sets F_j (corresponding to those values of j for which $\beta_j > \alpha + \delta$) according to Corollary 10.4, and the desired estimate (10.21) results immediately from (10.19). It remains therefore only to construct the sets F_j for those indices j that correspond to values $\beta_j \leq \alpha + \delta$. But this is now elementary because $\psi \leq \alpha + \delta$ on E^c and $\{\alpha - \delta \leq \phi \leq \alpha + \delta\} \cap F^c$ has infinite measure; each F_j can therefore be selected in turn from the latter set so as to be disjoint from its predecessors and to satisfy $|F_j| = |E_j|$ and property (10.21). ■

Theorem 10.6. *Let F and G belong to $(H^1 + H^\infty)(\mathbf{U})$, and suppose that*

$$\int_0^t (NG)^*(s)\,ds \leq \int_0^t (NF)^*(s)\,ds, \qquad (t > 0). \tag{10.22}$$

Then there is a linear operator T such that $TF = G$ and

$$\|TH\|_{H^1} \leq c\|H\|_{H^1}, \qquad (H \in H^1(\mathbf{U})), \tag{10.23}$$

$$\|TH\|_{H^\infty} \leq c\|H\|_{H^\infty}, \qquad (H \in H^\infty(\mathbf{U})), \tag{10.24}$$

where c is independent of F, G, and H.

Proof. In order to construct the desired operator T we decompose G into "analytic pieces" by using a stopping-time argument and the $\bar{\partial}$-machinery developed in Theorem 9.5. Consider first the case

$$\lim_{t \to \infty} (NG)^*(t) = 0. \tag{10.25}$$

The function NG is lower semicontinuous on $\mathbf{R}$ so, for each integer n, the set

$$\mathcal{O}_n = \{x \in \mathbf{R} : NG(x) > 2^n\}$$

is open and has finite measure. By the Whitney covering lemma (cf. Lemma 5.1 and inequality (5.3)), we may write $\mathcal{O}_n$ as a union of a family Ψ_n of dyadic intervals I with disjoint interiors and satisfying the properties

$$\left.\begin{aligned} &|I| \leq \operatorname{dist}(I, \mathcal{O}_n) \leq 4|I| \\ &|I| \leq 4|J| \qquad \text{if } J \in \Psi_n \text{ touches } I. \end{aligned}\right. \tag{10.26}$$

Fix a negative integer P. We now describe a stopping criterion which will determine an appropriate subcollection $\mathscr{C}$ of these Whitney intervals.
Stage 0: Include in $\mathscr{C}$ all intervals in Ψ_P.
Stage j + 1: Let I be any interval added at the previous (j-th) stage. Let $m = m(I)$ denote the smallest integer for which

$$|\mathcal{O}_m \cap I| \leq \frac{1}{2}|I|. \tag{10.27}$$

Then $m(I) \geq n + 1$ if $I \in \Psi_n$, and

$$|\mathcal{O}_{m-1} \cap I| > \frac{1}{2}|I|. \tag{10.28}$$

At this stage we include in $\mathscr{C}$ all intervals in Ψ_m whose interiors intersect I. Recall that if J is such an interval, then the dyadic nature of the intervals implies that J is a proper subinterval of I.

We thus obtain an infinite collection $\mathscr{C}$ of dyadic intervals and an integer-valued function m on $\mathscr{C}$ with the properties:

(i) each I in $\mathscr{C}$ is a Whitney interval for some $\mathscr{O}_n$;
(ii) if $I, J \in \mathscr{C}$ and their interiors have non-empty intersection, then one of the intervals is contained in the other;
(iii) if $I, J \in \mathscr{C}$, $J \subset I$, $J \neq I$, then $m(I) < m(J)$;
(iv) if $\mathscr{C}(I) \equiv \{J \in \mathscr{C} : J \subset I, J \neq I\}$, then

$$\sum_{J \in \mathscr{C}(I)} |J| \leq |I|; \tag{10.29}$$

(v) NG is bounded by $2^{m(I)}$ on the set $E(I)$ defined by

$$E(I) = I \setminus \bigcup_{J \in \mathscr{C}(I)} J = I \setminus \mathscr{O}_{m(I)};$$

(vi) if $I \in \mathscr{C}$, then

$$|E(I)| \geq |I|/2;$$

(vii) *the collection* $\{E(I)\}_{I \in \mathscr{C}}$ is disjoint with union equal to $\mathscr{O}_P$.

Part (iv) of (10.29) is established by recursion of the inequalities (10.27) and summation of the resulting geometric series.

For each $I \in \mathscr{C}$, let $R(I) = I \times (0, 5|I|)$, and let

$$\mathbf{U}_P = \bigcup \{R(I) : I \in \Psi_P\}.$$

The sets

$$\Gamma(I) = R(I) \Big\backslash \bigcup_{J \in \mathscr{C}(I)} R(J), \qquad (I \in \mathscr{C}) \tag{10.30}$$

provide a tiling of $\mathbf{U}_P$ into regions with disjoint interiors. Since each interval J in $\mathscr{C}(I)$ is a Whitney interval, the inequalities (10.26) show that $\Gamma(I)$ is contained in the union, as x varies over $E(I)$, of the cones $\{t + iy : |x - t| < y\}$ with vertex x. Hence,

$$|G(z)| \leq 2^{m(I)}, \qquad (z \in \Gamma(I)). \tag{10.31}$$

The function $g_I = G\chi_{\Gamma(I)}$ then satisfies $|g_I(z)| \leq 2^{m(I)}\chi_{\Gamma(I)}$. Moreover, $\bar{\partial} g_I$ is absolutely continuous with respect to arclength measure on the boundary (relative to $\mathbf{U}_P$) of $\Gamma(I)$, and

$$\|\bar{\partial} g_I\|_C \leq (5 \times 5) \cdot 2^{m(I)}. \tag{10.32}$$

Set

$$\tilde{g} = \sum_{I \in \mathscr{C}} 2^{m(I)} \chi_{E(I)}.$$

We claim that NG may be estimated in terms of the "median function" $\tilde{g}$ as follows:

$$(\chi_{\mathscr{O}_N} NG)^*(t) \leq (\tilde{g})^*(t) \leq 2(NG)^*\left(\frac{t}{8}\right), \qquad (t > 0). \tag{10.33}$$

Consequently, by a simple change of variable,

$$\tilde{g} \prec 16\, NG. \tag{10.34}$$

The left hand inequality in (10.33) follows immediately from property (v) of (10.29) and the definition of $\tilde{g}$. To establish the right hand inequality in (10.33), we shall use an auxiliary (median maximal) operator $h \to h^\circ$, which is defined for $x \in \mathbf{R}$ by

$$h^\circ(x) = \sup_{I \ni x} \inf_{j \in \mathbf{Z}} \left\{ 2^j : |\{|h| > 2^j\} \cap I| \leq \frac{1}{2}|I| \right\}.$$

It follows easily from the definitions of h° and the Hardy-Littlewood maximal operator M that

$$\{h^\circ > 2^k\} \subset \left\{ M(\chi_{\{|h| > 2^k\}}) > \frac{1}{2} \right\}.$$

The distribution functions therefore satisfy

$$|\{h^\circ > 2^k\}| \leq 8\|\chi_{\{|h| > 2^k\}}\|_{L^1} \leq 8|\{|h| > 2^k\}|,$$

since M is of weak type $(1, 1)$ (Theorem III.3.3). Hence the decreasing rearrangements satisfy the inequality

$$(h^\circ)^*(t) \leq (2h)^*\left(\frac{t}{8}\right).$$

Since $\tilde{g} \leq (NG)^\circ$, the preceding inequality applied to $h = NG$ produces the right hand inequality in (10.33). Hence, both (10.33) and (10.34) are established.

Define μ to be arclength measure on the union of the boundaries (relative to $\mathbf{U}_P$) of the sets $\Gamma(I)$, $(I \in \mathscr{C})$. Then μ is a positive Carleson measure with $\|\mu\|_C \leq 225$. To see this, note that for any square $Q = I_0 \times (0, |I_0|)$, there are at most three sets $\Gamma(I)$ with $|I| \geq |I_0|$ that intersect Q. Each of these will contribute no more than $25|I_0|$ to $|\mu|(Q)$. If on the other hand $\Gamma(I)$ intersects Q and $|I| \leq |I_0|$, then $I \subset 3I_0$ and so, by property iv) of (10.29), the total contribution of these sets $\Gamma(I)$ to $\mu(Q)$ is no larger than $3 \times 2 \times 25 \times |I_0|$.

We now use Theorem 9.5 to provide the decomposition of G relative to the tiling $\{\Gamma(I)\}_{I \in \mathscr{C}}$; define

$$G_I(z) = g_I(z) - \frac{1}{\pi} \iint_{\mathbf{U}_N} \frac{1}{z - \zeta} k_2(z, \zeta) K(z, \zeta)\, d(\bar{\partial} g_I)(\zeta) \tag{10.35}$$

with k_2 given by (9.8) and K defined by (9.10) using the measure $\mu/\|\mu\|_C$. Then G_I is analytic and $\sum_{I\in\mathscr{C}} G_I$ converges uniformly on compact subsets to an analytic function G_P. A summation in (10.35) shows that G_P is given by the right-hand side of (10.35) with g_I replaced by $G\chi_{\mathbf{U}_P}$. Moreover, by (9.15),

$$\begin{aligned}\|G_I\|_{L^1} &\le \|g_I\|_{L^1} + c\|\bar\partial g_I\|_C|I| \\ &\le c2^{m(I)}|I| \le c2^{m(I)}|E(I)|. \qquad (10.36)\end{aligned}$$

It follows from (10.32) that

$$\sum_{I\in\mathscr{C}} 2^{-m(I)}\bar\partial g_I = W\mu,$$

where $\|W\|_\infty \le c$, and so Corollary 9.10 shows that

$$\sum_{I\in\mathscr{C}} 2^{-m(I)}|G_I(z)| \le c. \qquad (10.37)$$

From (10.22) and (10.34) we have $\tilde g \prec 16NF$. Corollary 10.5, applied to $\phi = 16NF$ and $\psi = (\tilde g)^*$, therefore provides disjoint measurable sets $e(I)$ with the properties:

$$|e(I)| = |E(I)|, \qquad c\int_{e(I)} NF \ge 2^{m(I)}|E(I)|. \qquad (10.38)$$

There is a measurable function $\gamma:\mathbf{R}\to\mathbf{U}$ for which $\gamma(x)$ belongs to the cone $\{t+iy:|x-t|<y\}$ with vertex x, and

$$|F(\gamma(x))| \ge \frac{1}{2}NF(x). \qquad (10.39)$$

Define the linear functional λ_I on $H^1 + H^\infty$ by

$$\lambda_I(H) = \frac{\int_{e(I)} H(\gamma(x))\omega(x)\,dx}{2^{m(I)}|E(I)|}, \qquad (H \in H^1 + H^\infty), \qquad (10.40)$$

where $\omega(x) = \operatorname{sgn}[F(\gamma(x))]$. By properties (10.38) and (10.39), there exist weights w_I such that

$$w_I\lambda_I(F) = 1 \qquad (10.41)$$

and

$$\sup_{I\in\mathscr{C}} w_I \le 2c. \qquad (10.42)$$

Let

$$T_PH(z) = \sum_{I\in\mathscr{C}} w_I\lambda_I(H)G_I(z). \qquad (10.43)$$

Then by the definition in (10.40), the inequalities (10.42), and (10.36), respectively, we obtain for $H \in H^1(\mathbf{U})$, ,

$$\begin{aligned}\|T_PH\|_{L^1} &\le c\sum_{I\in\mathscr{C}} \frac{\int_{e(I)} NH(x)\,dx}{2^{m(I)}|E(I)|}\|G_I\|_{L^1} \\ &\le c\int \sum_{I\in\mathscr{C}} \chi_{e(I)} NH\,dx \le c\|H\|_{H^1(\mathbf{U})}. \qquad (10.44)\end{aligned}$$

On the other hand, if $H \in H^\infty(\mathbf{U})$, then (10.40), (10.38), and (10.37) show, for each $e \in \mathbf{U}_P$, that

$$\begin{aligned}|T_PH(z)| &\le c\|H\|_{H^\infty}\sum_{I\in\mathscr{C}} \frac{|e(I)|}{2^{m(I)}|E(I)|}|G_I(z)| \\ &\le c\|H\|_{H^\infty}, \qquad (10.45)\end{aligned}$$

where the constant c is independent of P. The estimates (10.44) and (10.45) show that the operators T_P are uniformly bounded on both $H^1(\mathbf{U})$ and $H^\infty(\mathbf{U})$, and equation (10.41) implies $T_PF = G_P$.

Now we let $P \to -\infty$ and use a Banach limit as in Section 2 of Chapter III to construct the desired operator T. The techniques are similar so we merely outline the proof.

First, by the definition of G_P (see (10.35) and the subsequent remark) and the facts

$$\|G\chi_{\mathbf{U}_P^c}\|_{L^\infty} \le 2^P, \qquad \|\bar\partial(G\chi_{\mathbf{U}_P})\|_C \le c2^P,$$

we see from Jones' theorem (Theorem 9.5) that G_N converges uniformly to G. In fact,

$$\|G - G_P\|_{L^\infty(\mathbf{R})} \le c2^P, \qquad (P = -1, -2, \ldots). \qquad (10.46)$$

Let λ be a Banach limit. For each function $H \in H^1 + H^\infty$ and each measurable set $E \subset \mathbf{R}$, let

$$\nu_H(E) = \lambda\left(\left\{\int_E T_PH\right\}_{P=-1}^{-\infty}\right). \qquad (10.47)$$

Then ν_H is absolutely continuous with respect to Lebesgue measure. If TH denotes the Radon-Nikodym derivative, then the operator $T:H\to TH$ is linear and satisfies the estimates

$$\|TH\|_{L^1(\mathbf{R})} \le c\|H\|_{H^1(\mathbf{U})}, \qquad \|TH\|_{L^\infty(\mathbf{R})} \le c\|H\|_{H^\infty(\mathbf{U})}.$$

Since λ assigns to a convergent sequence the value of its limit, it follows from

(10.46) that $TF = G$. Hence, to complete the proof (in the case where G satisfies (10.25)), it remains only to show that TH is analytic in $\mathbf{U}$.

From (10.47), we have for each $\psi \in L^1 \cap L^\infty$,

$$\int_{\mathbf{R}} (TH)\psi = \lambda\left(\left\{\int_{\mathbf{R}} (T_P H)\psi\right\}_{P=-1}^{-\infty}\right).$$

Hence, if ψ is the Poisson kernel,

$$(TH)(z) = \lambda(\{T_P H(z)\}_{P=-1}^{-\infty}).$$

By continuity of λ on ℓ^∞, we have for each ϕ with compact support in $\mathbf{U}$,

$$\iint_{\mathbf{U}} (TH)\,\bar{\partial}\phi = \lambda\left(\left\{\iint_{\mathbf{U}} (T_P H)\,\bar{\partial}\phi\right\}_{P=-1}^{-\infty}\right) = 0,$$

which shows that TH is analytic in $\mathbf{U}$.

Finally, we consider the case where G belongs to $H^1 + H^\infty$ but fails to satisfy condition (10.25). Thus, we may assume

$$\lim_{t\to\infty} (NG)^*(t) \equiv \alpha > 0, \tag{10.48}$$

which, by (10.22), implies that $\lim_{t\to\infty} (NF)^*(t) \geq \alpha$. Hence, there is an increasing sequence $\{E_n\}_{n=1}^\infty$ of sets of finite measure with $|E_n| \uparrow \infty$ such that $NF(x) > (3/4)\alpha$ whenever $x \in \bigcup_{n=1}^\infty E_n$. It follows that there is a Borel measurable function $\psi: \mathbf{R} \to \mathbf{U}$ such that $\psi(x)$ belongs to the cone $\{(t,y): |x-t| < y\}$ and

$$|F(\psi(x))| > \frac{\alpha}{2}, \qquad \left(x \in \bigcup_{n=1}^\infty E_n\right). \tag{10.49}$$

With λ the Banach limit used above, define a linear functional γ by

$$\gamma(H) = \lambda\left(\left\{\frac{1}{|E_n|}\int_{E_n} H(\psi(x))\omega(x)\,dx\right\}_{n=1}^\infty\right), \tag{10.50}$$

where $\omega(x) = \operatorname{sgn}(F(\psi(x)))$. Since (10.49) holds and λ is a Banach limit, we have

$$\gamma(F) \geq \frac{\alpha}{2} \tag{10.51}$$

and

$$\gamma(H) \leq \lambda(\{\|H\|_{H^\infty}\}_{n=1}^\infty) \leq \|H\|_{H^\infty}. \tag{10.52}$$

Moreover, if $H \in H^1$, then

$$\lim_{n\to\infty} \frac{1}{|E_n|}\int_{E_n} NH(x)\,dx = 0$$

and so

$$\gamma(H) = 0, \qquad (H \in H^1). \tag{10.53}$$

Choose t_0 such that $(NG)^*(t) < 2\alpha$ for $t \geq t_0$. Then by Theorem 10.1, there exists $G_0 \in H^1$ such that

$$\|G_0\|_{H^1} \leq c\int_0^{t_0} (NG)^*;$$

$$\|G - G_0\|_{H^\infty} \leq c(NG)^*(t_0) \leq 2c\alpha. \tag{10.54}$$

Hence, since $(NF)^*(t) \geq \alpha$ when $t > 0$, we have

$$\int_0^t (NG_0)^* \leq \int_0^t (NG)^* + \int_0^t [N(G - G_0)]^*$$

$$\leq \int_0^t (NF)^* + ct\alpha \leq c\int_0^t (NF)^*.$$

But G_0 belongs to H^1 and hence satisfies condition (10.25). By the first part of the proof already established above, there exists an admissible operator T_0 such that $T_0 F = G_0$. Define an operator T by $T = T_0 + T_1$, where $T_1(H) = \{\gamma(H)/\gamma(F)\}(G - G_0)$. Then certainly $T(H)$ is analytic. Furthermore, by (10.51), (10.52), and (10.54),

$$\|T_1(H)\|_{H^\infty} \leq \|H\|_{H^\infty}\left(\frac{2}{\alpha}\right)(2c\alpha)$$

so T_1 is bounded on H^∞. It follows from (10.53) that H^1 is contained in the kernel of T_1 and hence we conclude that T_1 is an admissible operator for the pair (H^1, H^∞). Since $T_1 F = G - G_0$, we have $TF = G$ and the proof is complete. ■

Definition 10.7. *Let X be a monotone Riesz-Fisher space on* $\mathbf{R}$. *The Hardy space $H(X)$ for X is the Banach space consisting of all analytic functions F on* $\mathbf{U}$ *for which the norm*

$$\|F\|_{H(X)} = \|f\|_X$$

is finite, where f denotes the boundary function of F.

The next result shows that the interpolation spaces of the Hardy spaces for L^1 and L^∞ are the Hardy spaces for the interpolation spaces of the same pair.

Theorem 10.8. *Let Y be an intermediate space between H^1* ($\mathbf{U}$) *and H^∞*($\mathbf{U}$).

Then Y is an interpolation space for $H^1(\mathbf{U})$ and $H^\infty(\mathbf{U})$ if and only if Y is equivalent to the Hardy space $H(X)$ of some monotone Riesz-Fischer space X.

Proof. Theorem 10.6 shows that all interpolation spaces of (H^1, H^∞) are monotone. Since $H^1 \cap H^\infty$ is dense in H^1, it follows therefore from Theorem 3.7 that every interpolation space Y for (H^1, H^∞) is of the form $Y = (H^1, H^\infty)_\rho$, for some monotone Riesz-Fischer norm ρ. Hence, Corollary 10.2 shows that Y is equivalent to the Hardy space $H(X)$ of $X = L^\rho$. The converse is an immediate consequence of Theorem 3.1. ■

When X is a rearrangement-invariant Banach function space (that is, X possesses the Fatou property in addition to being a monotone Riesz-Fischer space), the Hardy space $H(X)$ may also be described as the space of analytic functions F on $\mathbf{U}$ for which the norm

$$\|F\|_{H(X)} = \sup_{y>0} \|F(\cdot + iy)\|_X \tag{10.55}$$

is finite. Indeed, the space so described is embedded in $(H^1 + H^\infty)(\mathbf{U})$ and so every element F in the space has a nontangential limit a.e. on $\mathbf{R}$. Since the Poisson kernel is an approximate identity, we have

$$\|F(\cdot + iy)\|_X \le \|f\|_X = \|F\|_{H(X)},$$

where f is the boundary function of F. On the other hand, by the Fatou property of X,

$$\|f\|_X \le \varliminf_{y \downarrow 0} \|F(\cdot + iy)\|_X \le \sup_{y>0} \|F(\cdot + iy)\|_X,$$

and this establishes (10.55).

EXERCISES AND FURTHER RESULTS FOR CHAPTER V

1. (L. Maligranda [1], [2], [3]). Let (X_0, X_1) be a compatible couple. Then
(a) $X_0^o + X_1^o = (X_0 + X_1)^o$;
(b) $K(f, t; X_0 + X_1, X_0 \cap X_1) \sim K(f, t; X_0, X_1) + K(f, t; X_1, X_0)$;
(c) $K(f, t; X_0 + X_1, X_i) = K(f, 1 \wedge t, X_{1-i}, X_i), (i = 0, 1)$;
(d) $L^p(0, \infty)$ is an interpolation space between $(L^1 + L^\infty)(0, \infty)$ and $(L^1 \cap L^\infty)(0, \infty)$ if and only if $p = 2$. Not all interpolation spaces of this couple can be generated by the K-method (V. I. Ovcinnikov [1]).

2. A functional $\Phi: \mathscr{M}(\mathbf{R}^+) \to [0, \infty]$ is said to be an *admissible norm* if it has the following properties:

(i) $\Phi(w) = 0 \iff w = 0$ a.e.;
(ii) if ω_j denotes the least concave majorant of $|\omega_j|$, $(j = 1, 2)$, and $\omega_1 \le \omega_2$, then $\Phi(\omega_1) \le \Phi(\omega_2)$;

(iii) $\Phi(\sum_j |w_j|) \le \sum_j \Phi(w_j)$;
(iv) $\Phi(\lambda w) = |\lambda| \Phi(w)$, for all scalars λ;
(v) if $\xi(t) = \min(1, t)$, then $0 < \Phi(\xi) < \infty$ (without loss of generality, we may normalize and assume $\Phi(\xi) = 1$);
(vi) $\Phi(w) < \infty \Rightarrow |w| < \infty$ a.e.

Let (X_0, X_1) be a compatible couple and define $(X_0, X_1)_\Phi$ to consist of all $f \in X_0 + X_1$ for which the norm

$$\|f\|_{(X_0, X_1)_\Phi} = \Phi(K(f, \cdot; X_0, X_1))$$

is finite. An intermediate space X of a Gagliardo couple (X_0, X_1) is monotone if and only if $X = (X_0, X_1)_\Phi$ for some admissible norm Φ (Ju.A. Brudnyi & N.Ja. Krugljak [1]). (HINT: Let $\Phi = \Phi_X$ be defined by

$$\Phi(\omega) = \inf\left\{\sum_j \|f_j\|_X : \omega \le \sum_j K(f_j)\right\},$$

which is the "integrated" form of (3.33); cf. the proof of Theorem 3.7.)

3. If $X_0 \cap X_1$ is dense in X_0, the spaces $(X_0, X_1)_\rho$ of Section 3 coincide with the spaces $(X_0, X_1)_\Phi$ of Exercise 2 (HINT: Define ρ from Φ by $\rho(f) = \Phi(\omega)$, where $\omega(t) = tf^{**}(t)$; define Φ from ρ by observing that every nonnegative concave function ω on $\mathbf{R}^+$ is of the form $\omega(t) = \omega(0+) + \int_0^t \omega'(s)\,ds$, where ω' is nonnegative and decreasing, and set

$$\Phi(w) = \begin{cases} \rho(\omega'), & \text{if } \omega(0+) = 0, \\ \infty, & \text{if } \omega(0+) > 0. \end{cases}$$

4. (a) Compatible couples (X_0, X_1) and (Y_0, Y_1) are said to form a *relative Calderón pair* if, whenever f and g satisfy

$$K(g, t; Y_0, Y_1) \le K(f, t; X_0, X_1), \qquad (t > 0),$$

there exists an admissible operator $T: X_i \to Y_i$, $(i = 0, 1)$, such that $Tf = g$. If this property holds with $(Y_0, Y_1) = (X_0, X_1)$, then (X_0, X_1) is said to be a *Calderón couple*. Every interpolation space of a Calderón couple is monotone.
(b) Let (X_0, X_1) be any compatible couple and $0 < \theta_0 < \theta_1 < 1$. Define p_0 and p_1 by $\theta_i = 1 - 1/p_i$, $(i = 0, 1)$. Then $(X_0, X_1)_{\theta_0, p_0}$ and $(X_0, X_1)_{\theta_1, p_1}$ form a relative Calderón pair (cf. M. Cwikel [1], [2]).

5. Denote by L^∞ the essentially bounded measurable functions on $(0, \infty)$ and by $L^\infty(1/s)$ the weighted L^∞ space with norm

$$\|w\|_{L^\infty(1/s)} = \operatorname*{ess\,sup}_{0<s<\infty} \left|\frac{w(s)}{s}\right|.$$

(a) If ω is the least concave majorant (a.e) of $|w|$ with $\omega(0) = 0$, then

$$\omega(t) \le K\left(w, t; L^\infty, L^\infty\left(\frac{1}{s}\right)\right) \le 2\omega(t), \qquad (t > 0).$$

(b) $(L^\infty, L^\infty(1/s))$ is a Gagliardo couple.
(c) If (X_0, X_1) is an arbitrary compatible couple, then $((X_0, X_1), (L^\infty, L^\infty(1/s)))$ is a relative Calderón pair (HINT: Let $\gamma > 1$ and set $t_j = \gamma^j$ $(j \in \mathbf{Z})$. By the Hahn-Banach theorem applied to $X_0 + X_1$, equipped with the norm $K(\cdot, t_j)$, select a linear functional λ_j such that $\lambda_j(f) = K(f, t_j)$ and

$$|\lambda_j(h)| \le K(h, t_j), \qquad (h \in X_0 + X_1).$$

If T_γ is defined by $T_\gamma h = \sum_j \lambda_j(h)\chi_{(t_j, t_{j+1}]}$, then $\|T_\gamma\|_{\mathscr{A}} \le 1$ and $\gamma T_\gamma f(t) \ge K(w, t; L^\infty, L^\infty(1/s))$.
(d) The set of norms, up to equivalence, of the interpolation spaces of $(L^\infty, L^\infty(1/s))$ is precisely the set of admissible norms (cf. Exercise 2). This follows directly without appeal to the Brudnyi-Krugljak theorem.

6. Let X and Y be intermediate spaces for (X_0, X_1) and (Y_0, Y_1), respectively. Then X and Y are said to be *relatively monotone* if there is a constant $c > 0$ such that

$$K(g, t; Y_0, Y_1) \le K(f, t; X_0, X_1), \qquad (t > 0) \quad \Rightarrow \quad \|g\|_Y \le c\|f\|_X.$$

If X and Y are relatively monotone, there exists an admissible function norm Φ such that

$$X \hookrightarrow (X_0, X_1)_\Phi, \qquad (Y_0, Y_1)_\Phi \hookrightarrow Y$$

(HINT: Use divisibility for (Y_0, Y_1) and take $\Phi = \Phi_X$ as in Exercise 2 (Ju.A. Brudnyi & N.Ja. Krugljak [1]).)

7. (N. Aronszajn & E. Gagliardo [1]) A mapping I which assigns to each compatible couple (X_0, X_1) an intermediate space $I(X_0, X_1)$ is called an *interpolation method* if $(I(X_0, X_1), I(Y_0, Y_1))$ is an interpolation pair for arbitrary couples (X_0, X_1) and (Y_0, Y_1). Suppose A is an intermediate space for (A_0, A_1). For each couple (X_0, X_1), let $m_A(X_0, X_1)$ denote the set of elements $f \in X_0 + X_1$ for which

$$\|f\|_{m_A} = \inf\left\{\sum_j \|T_j\|_{\mathscr{A}} \|a_j\|_A\right\}$$

is finite (where $\mathscr{A} = \mathscr{A}(X_0, X_1; A_0, A_1)$ is the space of admissible operators; cf. Definition III.1.5). The infimum extends over all representations $f = \sum_j T_j(a_j)$ for which $a_j \in A$ and $T_j \in \mathscr{A}(A_0, A_1; X_0, X_1)$.
(a) $m_A(X_0, X_1)$ is an intermediate space of (X_0, X_1) and A is continuously embedded in $m_A(A_0, A_1)$.
(b) m_A is an interpolation method.
(c) If I is an interpolation method such that A is continuously embedded in $I(A_0, A_1)$, then $m_A(X_0, X_1)$ is continuously embedded in $I(X_0, X_1)$ for all couples (X_0, X_1). In this sense m_A is a *minimal* interpolation method.

8. (N. Aronszajn & E. Gagliardo [1]) Suppose A is an intermediate space of a compatible couple (A_0, A_1). For each couple (X_0, X_1), let $M_A(X_0, X_1)$ consist of those $f \in X_0 + X_1$ for which

$$\|f\|_{M_A} = \sup\{\|Tf\|_A : \|T\|_{\mathscr{A}} \le 1\} < \infty.$$

The space $M_A(A_0, A_1)$ is continuously embedded in A and M_A is an interpolation method. Furthermore, M_A is *maximal* in the sense that if I is any other interpolation method such that $I(A_0, A_1)$ is continuously embedded in A, then $I(X_0, X_1)$ is continuously embedded in $M_A(X_0, X_1)$ for all couples (X_0, X_1) (cf. also S. Janson [1], V. I. Ovcinnikov [2]).

9. (Ju.A. Brudnyi & N.Ja. Krugljak [1]; S. Janson [1]) Let X be an intermediate space for (X_0, X_1) and let Φ be an admissible norm (cf. Exercise 2). Then Φ is equivalent to the norm of $m_X(L^\infty, L^\infty(1/s))$.

10. (J-method—Ju.A. Brudnyi & N.Ja. Krugljak [1]) Let (X_0, X_1) be a Gagliardo couple. For each $t > 0$, the functional

$$J(u, t; X_0, X_1) = \max(\|u\|_{X_0}, t\|u\|_{X_1})$$

defines a Banach space norm on $X_0 \cap X_1$. Suppose $X_0 \cap X_1$ is dense in X_j, $(j = 0, 1)$.
(a) The following converse of inequality (2.38) holds: each $f \in X_0 + X_1$ may be expressed as a Bochner integral on $X_0 + X_1$,

$$f = \int_0^\infty u_s \frac{ds}{s} \qquad (\text{convergence in } X_0 + X_1)$$

with $\{u_s\}_{s>0} \subset X_0 \cap X_1$, and for all such f,

$$\int_0^\infty J(u_s, s) \min\left(1, \frac{t}{s}\right) \frac{ds}{s} \le 100K(f, t), \qquad (t > 0).$$

(HINT: Apply Theorem 3.4 to f and let J_k denote the index set

$$J_k = \left\{j : 2^k < \frac{\|a_j\|_0}{\|a_j\|_1} \le 2^{k+1}\right\}.$$

Define

$$u_s = \sum_{k=-\infty}^{\infty} \frac{1}{\log 2}\left(\sum_{j \in J_k} a_j\right)\chi_{(2^k, 2^{k+1}]}(s).)$$

(b) Let Ψ denote the norm of an interpolation space of $(L^1(1/s), L^1(1/s^2))$. Then $(X_0, X_1)_{\Psi; J}$ is defined as the set of elements $f \in X_0 + X_1$ for which the norm

$$\|f\|_{\Psi, J} = \inf\left\{\Psi(J(u_{(\cdot)}, (\cdot))) : f = \int_0^\infty u_s \frac{ds}{s}\right\}$$

is finite. The couple $((X_0, X_1)_{\Psi, J}, (Y_0, Y_1)_{\Psi, J})$ is an interpolation pair for (X_0, X_1) and (Y_0, Y_1).
(c) It follows from (a) that there is a one-to-one correspondence between the spaces generated by the J-method and those that arise from the K-method (HINT: If

$$Sw(t) = \int_0^\infty \min\left(1, \frac{t}{s}\right) w(s) \frac{ds}{s},$$

then Ψ corresponds to Φ if and only if $\Psi \sim \Phi \circ S$; this follows from the fact that $K(w, t; L^1(1/s), L^1(1/s^2)) \sim Sw$).

11. (Reiteration: Ju. A. Brudnyi & N. Ja. Krugljak [1]; see also P. Nilsson [1]) Let (X_0, X_1) be a Gagliardo couple.

(a) If Φ_0 and Φ_1 are two admissible function norms, and if

$$E_j = \left(L^\infty, L^\infty\left(\frac{1}{s}\right)\right)_{\Phi_j}, \qquad Y_j = (X_0, X_1)_{\Phi_j}, \qquad (j = 0, 1)$$

then, from divisibility, it follows that

$$K(f, t; Y_0, Y_1) \sim K(K(f, \cdot), t; E_0, E_1),$$

where $K(f, s) = K(f, s; X_0, X_1)$.

(b) It follows from (a) that to each admissible function norm Φ there corresponds an admissible function norm Ψ such that

$$(Y_0, Y_1)_\Phi = (X_0, X_1)_\Psi.$$

(HINT: Let Ψ be the function norm for $(E_0, E_1)_\Phi$.)

12. (E. M. Stein [6]; cf. also R. A. DeVore & R. Sharpley [2]) Let Ω be an open subset of $\mathbf{R}^n$ and let ∇f denote the distributional gradient of f.

(a) There is a constant $c > 0$ such that if ∇f is locally integrable, then for each cube $Q \subset \Omega$,

$$|f_Q - f(x)| \le c \int_Q |\nabla f(y)| \cdot |x - y|^{1-n}\, dy \qquad (\text{for a.e. } x \in Q)$$

where f_Q is the average of f over Q.

(b) Let $Q(h)$ be an arbitrary cube of side length h which contains both x and $x + h$. If $(\nabla f)\chi_{Q_0}$ belongs to the Lorentz space $L^{n,1}$, then f may be redefined on a subset of Q_0 of measure zero so that f satisfies

$$|f(x + h) - f(x)| \le c \|\nabla f \chi_{Q(h)}\|_{L^{n,1}}, \qquad (Q(h) \subset Q_0).$$

(c) Part (b) gives rise to the following generalization in $\mathbf{R}^n$, $(n > 1)$, of Lebesgue's differentiation theorem: if the distributional gradient ∇f is locally in the Lorentz space $L^{n,1}(\Omega)$, then f may be redefined on a set of measure zero so that it is continuous and satisfies

$$\lim_{h \to 0} \frac{f(x + h) - f(x) - h \cdot \nabla f(x)}{|h|} = 0 \qquad (\text{for a.e. } x \in \Omega).$$

13. An open set $\Omega \subset \mathbf{R}^n$ is said to have a *minimally smooth boundary* if there is a sequence of open sets $\{\mathcal{O}_j\}$, a corresponding sequence of cones $\{C_j\}$ (all congruent to a fixed finite cone C), and positive numbers ε, N such that

(i) $\partial\Omega \subset \bigcup_j \mathcal{O}_j^\varepsilon$, where $\mathcal{O}^\varepsilon = \{x : y \in \mathcal{O} \text{ and } |x - y| < \varepsilon\}$
(ii) $x + C_j \subset \Omega$ if $x \in \mathcal{O}_j \cap \Omega$
(iii) each x is contained in at most N of the $\mathcal{O}_j$'s.

(a) (E. M. Stein [5, pp. 180–192]) There exists an extension operator E such that

$$E : W_k^p(\Omega) \to W_k^p(\mathbf{R}^n), \qquad (1 \le p \le \infty, k = 0, 1, 2 \ldots),$$

$$Ef|_\Omega = f \qquad \text{a.e.},$$

(b) (H. Johnen & K. Scherer [1]) If Ω has minimally smooth boundary and $1 \le p < \infty$ (use $C(\Omega)$ when $p = \infty$), then

$$K(f, t; L^p(\Omega), \dot{W}_k^p(\Omega)) \sim \omega_k(f, t)_p,$$

where $\dot{W}_k^p(\Omega)$ is defined by means of the Sobolev seminorm

$$|f|_{p,k} = \sum_{|v| = k} \|D^v f\|_{L^p(\Omega)}$$

and

$$\omega_k(f, t)_p = \sup_{|h| \le t} \|\Delta_h^k f\|_{L^p(\Omega(kh))},$$

with $\Omega(kh) = \{x : x + th \in \Omega, 0 \le t \le k\}$.

(c) (R. A. DeVore, S. D. Riemenschneider & R. Sharpley [1]) Marchaud's inequality (Theorem 4.4) and inequalities (4.54) and (4.61) hold for domains Ω with minimally smooth boundary (cf. R. Sharpley [8] for a proof using parts (a) and (b)). Consequently, the theory as developed in section 4 for Sobolev and Besov spaces has an analogue in such domains.

14. Let P_y denote the Poisson kernel and Q_y its conjugate (cf. Exercises 3.8, 9, 11).

(a) By uniqueness of the harmonic extension, $P_y * P_\eta = P_{(y+\eta)}$.

(b) If $t > 0$, then e^{itz} belongs to $H^\infty(\mathbf{U})$ and

$$H(e^{-itx}) = -ie^{itx}, \qquad (t > 0).$$

(c) If $f \in L^2(\mathbf{R})$, then

$$(Hf)\hat{}\,(t) = -i \operatorname{sgn} t \hat{f}(t).$$

(HINT: If $F = f + iHf$, then F is analytic in $\mathbf{U}$ (cf. Exercise III.12(d)). The function $e^{-itz}F(z)$ is analytic and its growth is regulated in $\mathbf{U}$ when $t < 0$, so Cauchy's theorem gives $\hat{F}(t) = 0$, $(t \le 0)$. Hence $(Hf)\hat{}\,(t) = i\hat{f}(t)$ if $t < 0$. For $t > 0$, use the preceding case and the change of variable $t \to -t$.)

(d) The Hilbert transform has period 4, $(H^2 = -I)$, and is an isometry on $L^2(\mathbf{R})$.

(e) For any atom a centered at the origin, it follows from Exercise II.18 that

$$\|Ha\|_1 \le (8\|Ha\|_2 \|t \cdot Ha(t)\|_2)^{1/2} \le c_0$$

(HINT: $tHa(t) = HA(t)$, where $A(x) = xa(x)$).

15. (Hardy's inequality) If a is an H^1-atom, then

$$\int_{\mathbf{R}} \left|\frac{1}{t}\hat{a}(t)\right| dt \le 3 \tag{1}$$

and so by Theorem 6.14,

$$\int_{\mathbf{R}} \left| \frac{1}{t} \hat{f}(t) \right| dt \le c \|f\|_{H^1(\mathbf{U})}.$$

(HINT: The underlying interval I (Definition 6.9) may be taken to be centered at the origin. Hence

$$|\hat{a}(t)| = \left(\frac{1}{2\pi}\right)^{1/2} \left| \int_I a(x)[e^{itx} - 1]\, dx \right| \le \left(\frac{1}{2\pi}\right)^{1/2} |t||I|. \tag{2}$$

By Plancherel's theorem,

$$\|\hat{a}\|^2_{L^2(\mathbf{R})} = \|a\|^2_{L^2(\mathbf{R})} \le |I|^{-1},$$

so

$$\int \left| \frac{1}{t} \hat{a}(t) \right| dt = \left(\int_{|t| \le |I|^{-1}} + \int_{|t| > |I|^{-1}} \right) \left| \frac{1}{t} \hat{a}(t) \right| dt$$
$$\le 1 + \left(\int_{|t| > |I|^{-1}} \frac{dt}{t^2} \right)^{1/2} \|\hat{a}\|_{L^2} \le 1 + 2^{1/2}.)$$

16. (a) Every L^1-function f may be rearranged so as to belong to $\operatorname{Re} H^1$: there exists $g \in \operatorname{Re}(H^1)$ such that $g^* = f^*$ (HINT: We may assume $f = f^*$, and so

$$\|f\|_{L^1} \le \sum_{k=-\infty}^{\infty} f^*(2^k) 2^k \le 2\|f\|_{L^1}.$$

Select ψ so that on $I_k = (2^k, 2^{k+1})$, we have $|\psi| = 1$ and $\int_{I_k} f^* \psi = 0$. Then $g = \psi f^*$ has a natural atomic decomposition with respect to the intervals I_k and $g^* = f^*$).
(b) (B. Davis [3]) Let $f \in L^1(\mathbf{T})$ and let g be the (signed) rearrangement of f (that is, $|\{g > \lambda\}| = |\{f > \lambda\}|$ for all real λ) which is decreasing on the interval $[0, 2\pi)$ (so g in general has a discontinuity at the point $e^{i0} \in \mathbf{T}$). Then some (signed) rearrangement of f is in $\operatorname{Re} H^1$ if and only if

$$\int_0^{\pi} \left| \frac{1}{\theta} \int_{-\theta}^{\theta} g(e^{i\phi})\, d\phi \right| d\theta < \infty.$$

In particular, if some (signed) rearrangement of f belongs to $\operatorname{Re} H^1$, then so does g (cf. also N.J. Kalton [1]).
(c) (A. Baernstein [2]) The signed symmetric decreasing rearrangement f^s of f (cf. Exercise II.21) satisfies $\|(f^s)^{\tilde{}}\|_p \ge \|\tilde{f}\|_p$ for $1 \le p < 2$ and the inequality is reversed for $2 < p < \infty$ ($\tilde{g}$ is the conjugate function of g). In particular, all signed rearrangements of f belong to $\operatorname{Re} H^1$ if and only if f^s does; the latter occurs if and only if $f \in L\log L(\mathbf{T})$ (M. Essén & D. Shea [1]; cf. also B. Davis [3]).

17. (C. Fefferman & E. M. Stein [1], R. Coifman & G. Weiss [1]) The characterizations of $H^1(\mathbf{U})$ given in Theorem 6.14 have analogues for $0 < p < 1$, for higher

dimensions and for spaces of homogeneous type. For $H^p(\mathbf{U})$, $(1/2 < p < 1)$, the details are as follows. If F is analytic in $\mathbf{U}$, the following conditions are equivalent:
(i) $\|F\|_{H^p(\mathbf{U})} = \sup_{y>0} (\int |F(x + iy)|^p\, dx)^{1/p} < \infty$;
(ii) $\|N(\operatorname{Re} F)\|_{L^p(\mathbf{R})} < \infty$, where $Ng(x) = \sup_{(t,y) \in \Gamma_x} |g(t + iy)|$;
(iii) there exist functions a_I with

$$\int a_I = 0, \qquad |a_I| \le |I|^{-1/p} \chi_I,$$

and coefficients λ_I with $(\sum_I |\lambda_I|^p)^{1/p} < \infty$ and

$$\operatorname{Re} F = \sum_I \lambda_I a_I \qquad \text{(as distributions).}$$

18. If $0 < \alpha < 1$, define $f^{\#}_\alpha(x)$ by

$$f^{\#}_\alpha(x) = \sup_{I \ni x} |I|^{-(\alpha+1)} \int_I |f - I(f)|$$

where $I(f)$ denotes the average of f over I.

(a) Iteration of $|I(f) - (\frac{1}{2}I)(f)| \le 2|I|^\alpha \inf_I f^{\#}_\alpha$ gives

$$|f(x) - I(f)| \le \frac{2^{\alpha+1}}{2^\alpha - 1} |I|^\alpha f^{\#}_\alpha(x), \qquad (x \in I).$$

(b) By part (a), the space $\operatorname{Lip} \alpha = \{f : |\Delta_h f(x)| \le c|h|^\alpha\}$ may be identified as the space of functions f for which $\|f^{\#}_\alpha\|_\infty < \infty$.
(c) Parts (a), (b), and the results of Exercise 17, show that the dual of $H^p(\mathbf{U})$ is (modulo constant functions) isomorphic to $\operatorname{Lip} \alpha$; $(\alpha + 1 = 1/p)$ (cf. P. Duren, B. Romberg & A. Shields [1]).

19. (Marcinkiewicz Multiplier Theorem: cf. R. Coifman & G. Weiss [1])
(a) The *Fourier multiplier* corresponding to a function $\mu \in L^\infty(\mathbf{R})$ is the linear operator m such that

$$(mf)\hat{}\,(t) = \mu(t) \hat{f}(t). \tag{3}$$

The multiplier m is bounded on $L^2(\mathbf{R})$ and $\|m\|_{\mathscr{B}(L^2)} \le \|\mu\|_{L^\infty}$. The Hilbert transform H is such a multiplier (cf. Exercise 14(c)). Clearly, m and H commute. The function μ is the *symbol* of the operator m.
(b) The *Marcinkiewicz multiplier theorem* asserts that if μ also satisfies the Hörmander condition

$$\sup_{R>0} \left\{ R \int_{R \le |x| < 2R} |\mu'(x)|^2\, dx \right\} = c_0 < \infty, \tag{4}$$

then m is a bounded operator on H^1. To see this, note first that for any atom a centered at the origin,

$$\int_{\mathbf{R}} |ma(x)|\, dx \le c, \tag{5}$$

where c depends only on $\|\mu\|_\infty$ and c_0. (HINT: Let $M = ma$ and use Exercise II.18 to obtain

$$\begin{aligned}\|M\|_{L^1} &\le (8\|M\|_{L^2}\|x\cdot M\|_{L^2})^{1/2}\\ &\le (8\|\mu\|_\infty |I|^{-1/2}\|(\mu\hat{a})'\|_{L^2})^{1/2}\end{aligned} \tag{6}$$

since $(-ix\cdot M)\hat{} = (\hat{M})'$ and a is an atom. But

$$\|(\mu\hat{a})'\|_{L^2} \le \|\mu'\hat{a}\|_{L^2} + \|\mu\|_\infty\|\hat{a}'\|_{L^2},$$

and $\|\hat{a}'\|_{L^2} = \|x\cdot a\|_{L^2} \le |I|^{1/2}$, since a is an atom centered at zero. Hence, it remains only to show that

$$\|\mu'\hat{a}\|_{L^2} \le c|I|^{1/2}. \tag{7}$$

This, however, is a consequence of the estimate (cf. Exercise 15) $|\hat{a}(t)| \le \min(1, |t|\cdot|I|)$ and the estimates

$$\begin{aligned}\|\mu'\hat{a}\|_2^2 &= \sum_{k=-\infty}^{\infty}\int_{2^k<|t|\le 2^{k+1}} |\mu'(t)\hat{a}(t)|^2\,dt\\ &\le c_0\sum_{k=-\infty}^{\infty} 2^{-k}\sup\{|\hat{a}(t)|^2 : 2^k < |t| \le 2^{k+1}\}\\ &\le 16c_0\int_0^\infty \frac{1}{s}\min(1, s|I|)^2\frac{ds}{s} = 32c_0|I|.)\end{aligned}$$

(c) If μ satisfies the Hörmander condition (4), translation invariance of m shows that (5) holds for all atoms. Hence,

$$\|mf\|_{L^1(\mathbf{R})} \le c\|f\|_{H^1(\mathbf{U})},$$

so m is bounded on $H^1(\mathbf{U})$ since m and H commute.
(d) If μ satisfies (4) and $\|\mu\|_\infty < \infty$, then m is bounded on $L^p(\mathbf{R})$, $(1 < p < \infty)$. (HINT: For $1 < p < 2$, use part (c) and interpolation. The case $2 < p < \infty$ follows by duality.)

20. (R. Sharpley [11]) The techniques of Sections 6 and 10 may be combined to show that $(\mathrm{Re}\, H^1(\mathbf{R}), L^\infty(\mathbf{R}))$ is a Calderón couple.

21. Suppose the sequence $\{z_n\} \subset \mathbf{U}$ is *uniformly separated*; that is, the z_n's are distinct and there exists $\delta > 0$ such that

$$\prod_{\substack{n\\ n\ne j}} \frac{|z_j - z_n|}{|z_j - \bar{z}_n|} \ge \delta. \tag{8}$$

(a) Let B_j denote the Blaschke product with zeros z_n, $(n \ne j)$ (cf. Appendix). Then inequality (8) in equivalent to $|B_j(z_j)| \ge \delta$ for all j.

(b) If y_n is the imaginary part of z_n, then

$$\sum_{\substack{n\\ n\ne j}} \frac{4y_n y_j}{|z_j - \bar{z}_n|^2} \le -\log|B_j(z_j)|^2 \le 2\log\left(\frac{1}{\delta}\right).$$

(HINT: Use the fact that $-\log t \ge 1 - t$, $(t > 0)$.)
(c) If $\{\beta_n\}$ is a sequence with $\|\{\beta_n\}\|_\infty \le 1$, then

$$\mu = \sum_n \beta_n y_n \delta_{z_n} \tag{9}$$

is a Carleson measure. (HINT: (cf. J.B. Garnett [1], p. 240) Let Q be a cube of side length h. If, for some j, z_j belongs to the top half of Q, then

$$\sum_{\substack{z_n\in Q\\ n\ne j}} y_n \le 10h\sum_{\substack{n\\ n\ne j}} \frac{y_n y_j}{|z_j - \bar{z}_n|^2} \le 5h\log\frac{1}{\delta}$$

by part (b). Otherwise, run a stopping-time argument on the dyadic subcubes of Q, each of which has bottom edge along $\mathbf{R}$. The stopping criteria is satisfied when (if ever) one of the z_n's belongs to the top half of the subcube. The sums of the lengths of these cubes will not exceed h.)

22. (P. Jones [1]) If μ is a Borel measure on $\mathbf{U}$, then μ is a Carleson measure if and only if μ is the weak*-limit of convex combinations of measures of the form (9), with the sequences $\{z_j\}$ uniformly separated and separation constants δ uniformly bounded away from zero.

23. (L. Carleson [1]) A sequence $(z_n) \subset \mathbf{U}$ is called an *interpolating sequence* if there is a constant $c > 0$ such that if (α_n) is a bounded sequence, then there exists an H^∞-function F such that

$$F(z_n) = \alpha_n \qquad (\text{all } n), \qquad \|F\|_{H^\infty} \le c\|(\alpha_n)\|_\infty.$$

A sequence (z_n) is an interpolating sequence if and only if it is uniformly separated (cf. Exercise 21). (HINT: (cf. P. Jones [3]) If $\{z_n\}$ is interpolating, there exist functions F_j in H^∞ such that

$$F_j(z_n) = \delta_{jn}, \qquad \|F_j\|_{H^\infty} \le c.$$

If B_j is the Blaschke product formed from the zeros of F_j, then

$$1 = |F_j(z_j)| \le c|B_j(z_j)|,$$

which implies (8). Conversely, if $\{z_n\}$ is uniformly separated, Jones' theorem (Theorem 9.5) and Exercise 21(c) yield functions f_j which satisfy $\bar\partial f_j = y_j\delta_{z_j}$ and $\sum_j |f_j(x)| \le c$, $(x \in \mathbf{R})$. Define F_j as a constant multiple of Bf_j, where B is the Blaschke product for $\{z_n\}$. Define F as the sum $\sum_j \alpha_j F_j$.)

24. (P. Jones [3]) An *analytic atom* $A(z)$ is an $H^\infty(\mathbf{U})$-function for which there is an interval $I \subset \mathbf{R}$ such that

$$|A(x)| \leq |I|^{-1} \exp\left(-\left\{\frac{|x-x_0|}{|I|}\right\}^{1/2}\right), \qquad (x \in \mathbf{R}),$$

where x_0 is the center of I. Each $H^1(\mathbf{U})$-function F has an "analytic" atomic decomposition: there exist analytic atoms A_j and coefficients λ_j such that $F = \sum_j \lambda_j A_j$ and

$$\|F\|_{H^1(\mathbf{U})} \sim \sum_j |\lambda_j|.$$

25. Suppose $g \prec f$. For each $\varepsilon > 0$, Corollary 10.5 may be used to construct a linear operator T such that $Tf = g$ and

$$\max(\|T\|_{\mathscr{B}(L^1)}, \|T\|_{\mathscr{B}(L^\infty)}) < 1 + \varepsilon.$$

(HINT: Construct $\tilde{g} = \sum \beta_j \chi_{E(j)}$ such that $\tilde{g} \prec f$ and $|g| \leq (1+\varepsilon)^{1/2}|\tilde{g}|$. Modify the operator

$$h \to \left\{\sum_j \frac{1}{|F(j)|} \int_{F(j)} h \operatorname{sgn}(f)\right\} \chi_{E(j)}.)$$

26. (I. Klemes [1]) Let $f \in \mathrm{BMO}([0,1])$. Then

$$\|f^*\|_{\mathrm{BMO}} \leq \|f\|_{\mathrm{BMO}}.$$

The corresponding result for symmetric decreasing rearrangements on the unit circle $\mathbf{T}$ is false.

27. (N. P. Korneicuk [1], B. S. Mitjagin & E. M. Semenov [1], J. Peetre [4]) When the space of Lipschitz functions Lip 1 on $[0,1]$ is furnished with the seminorm $f \to \|f'\|_\infty$, the K-functional for the pair $(C, \mathrm{Lip}\, 1)$ is given by

$$K(f,t) = \frac{1}{2}\bar{\omega}(f, 2t), \qquad (t > 0),$$

where $\bar{\omega}$ is the least concave majorant of the modulus of continuity $\omega(f,t)$ of f.

NOTES FOR CHAPTER 5

The J- and K-methods are due to J. Peetre [1], [2], whose manifold contributions have made his name almost synonymous with interpolation theory. The J- and K-methods refine and simplify earlier methods devised by E. Gagliardo and J. L. Lions & J. Peetre; P. L. Butzer & H. Berens [1] provide a good account of the evolution. In addition to the latter source, the books by J. Bergh & J. Löfström [1], S. G. Krein, Ju. A. Petunin & E. M. Semenov [1], J. L. Lions & E. Magenes [1], and H. Triebel [1] contain general treatments of interpolation, from various standpoints.

The Gagliardo completion is discussed at length by N. Aronszajn & E. Gagliardo [1] and H. Berens [1]. The systematic use of rearrangement-invariant Banach function norms ρ in connection with the k-functional is due to C. Bennett [3, 4, 5]. The reiteration theorem in the form presented in §2 is due to T. Holmstedt [1]; see also M. Cwikel [2]. Theorem 2.11 was established by T. H. Wolff [1]; see also S. Janson, P. Nilsson & J. Peetre [1]. The Schatten trace classes interpolate in much the same way as the $L^{p,q}$-spaces. Thus if $S^{p,q}$ denotes the class of compact operators on Hilbert space whose singular values lie in $\ell^{p,q}$, then

$$(S^{p_0,q_0}, S^{p_1,q_1})_{\theta,q} = S^{p,q}, \qquad \left(\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}\right)$$

(cf. J. Bergh & J. Löfström [1], B. Simon [1] for details).

That monotone interpolation spaces are generated by the K-method is announced in Ju. A. Brudnyi & N. Ja. Krugljak [1]. The "fundamental lemma" (Theorem 3.4) may have been known to them in some form. It appears in special cases in M. Cwikel & J. Peetre [1] and in general in M. Cwikel [3]. The main result (identifying monotone interpolation spaces as the spaces $(X_0, X_1)_\Phi$, where Φ is a norm of the type described in Exercises 2 through 6) is due to Ju. A. Brudnyi & N. Ja. Krugljak [1]; another proof is given by M. Cwikel [3]. The proof presented here, based on the Lorentz-Shimogaki result for (L^1, L^∞), is due to C. Bennett & R. Sharpley [3].

Calderón couples (cf. Exercise 4) have been the subject of numerous investigations. G. Sparr [1], [2] has shown that pairs of weighted L^p-spaces are Calderón couples; cf. also M. Cwikel [1] and J. Arazy & M. Cwikel [1]. Also in these works are characterizations of the interpolation spaces between L^p and L^q, which complete the earlier investigations of G. G. Lorentz & T. Shimogaki [3]. The first examples of pairs that are not Calderón couples were suggested by the work of A. A. Sedaev & E. M. Semenov [1]: the pairs (L^p, W_1^p), $(p \neq 2)$ and $(L^1 \oplus L^2, L^\infty \oplus L^\infty)$ are not Calderón couples (cf. M. Cwikel [1], [2], [3] and the references cited there).

There is an extensive literature on Besov and Sobolev spaces, for which M. H. Taibleson [1], R. A. Adams [1] and J. Peetre [5] are excellent sources. P. L. Butzer & H. Berens [1] provide a history of the characterization of various K-functionals for smoothness spaces in terms of moduli of smoothness (cf. also Exercise 27). In several variables, the result (Theorem 4.12) is due to H. Johnen & K. Scherer [1]. Much of the material in §4 derives from the latter source and the paper of R. A. DeVore, S. D. Riemenschneider & R. Sharpley [1]. The Marchaud inequalities ((4.7) and Theorem 4.4) originate with A. Marchaud [1]. Atomic decompositions of Besov spaces have been constructed by M. Frazier & B. Jawerth [1], [2] and R. A. DeVore & V. Popov [1], [2]. There is no common agreement on the definition of Besov spaces for $p < 1$; some of their interpolation properties in different contexts have been established by J. Peetre [5] and R. A. DeVore & V. Popov [1], [2].

The main result of §5 (Theorem 5.12) was established by R. A. DeVore & K. Scherer [1], using spline approximation and combinatorial covering arguments based on the Whitney covering lemma; the result is implicit also in the paper of A. P. Calderón [4]. Subsequent proofs (based also on the Whitney covering lemma) have been given by R. A. DeVore & R. Sharpley [1, Theorem 8.4] and C. P. Calderón & M. Milman [1]. The latter makes use of the Whitney extension theorem (cf. E. M. Stein [5]), which is essentially established in the proof of Theorem 5.10. The proof presented here embarks from the one-dimensional case, which is relatively straightforward, and invokes estimates on the directional derivatives to proceed to the general case.

Real variable characterizations of H^p-spaces originated with D. L. Burkholder, R. F. Gundy & M. L. Silverstein [1]; for a survey, see A. Torchinsky [1]. §6 contains a novel treatment of the H^1-theory in one variable. We do not know the exact origin of Lemma 6.12, which is apparently part of the folklore and whose existence was pointed out to us by J. Davis; there are several elementary proofs. The duality of H^1 and BMO, established by C. Fefferman [1] and C. Fefferman & E. M. Stein [1], led to the atomic characterization of H^1 by R. R. Coifman [1] and R. Latter [1]. An excellent account appears in R. R. Coifman & G. Weiss [1]. The K-functional for H^1 and L^∞ was described by C. Fefferman, N. M. Rivière & Y. Sagher [1] in terms of a "grand maximal operator". The description presented here (Theorem 6.15), in terms of the nontangential maximal operator, is due to R. Sharpley [10]. Lemma 6.8, to the effect that $H: L^\infty \to \text{BMO}$, was established independently by S. Spanne [1] and E. M. Stein [3]; subsequently, C. Fefferman [1] and C. Fefferman & E. M. Stein [1] showed that BMO $= L^\infty + H(L^\infty)$, a constructive proof of which has been given by A. Uchiyama [1] (cf. also N.Th. Varopoulos [1]). C. Bennett, R. A. DeVore & R. Sharpley [1] showed that maximal singular integral operators also map L^∞ into BMO.

The space $W = \text{weak-}L^\infty$ (Definition 7.9) was introduced by C. Bennett, R. A. DeVore & R. Sharpley [2], who also established Lemma 7.2, Theorems 7.3, 7.10, and 7.15. Another proof of the fundamental inequality (7.6) has been given by M. Milman [1], using the theory of A_p-weights; there is a related result of I. Klemes [1] in Exercise 26 (which is also asserted without

proof by A. Garsia & E. Rodemich [1]). The covering lemma (Lemma 7.2) is a variant of Lemma III.3.7, which was established by C. Bennett & R. Sharpley [1]; the derivation of the John-Nirenberg lemma (Corollary 7.7) and the weak-type inequality (7.10) are also in that paper (for the John-Nirenberg lemma, see F. John & L. Nirenberg [1] and J. B. Garnett & P. W. Jones [1]). The inequality (7.10) is a variant, formulated in terms of the decreasing rearrangement, of an earlier "good-λ inequality" of C. Fefferman & E. M. Stein [1, (4.4)], which is derived from the Calderón-Zygmund lemma rather than the covering lemma mentioned above. R. J. Bagby and D. S. Kurtz [1], [2] have generalized this procedure in connection with weighted estimates for the Hardy-Littlewood maximal operator and Calderón-Zygmund maximal singular integral operators. D. S. Kurtz [1] has established similar results for the area integral and the Littlewood-Paley g-functions. The book by A. Torchinsky contains an account of good-λ inequalities and related topics in harmonic analysis.

The space BLO was introduced by R. R. Coifman & R. Rochberg [1], who derived the representation BMO $=$ BLO $-$ BLO from a structure theorem of L. Carleson [3] for BMO-functions. The characterization in Theorem 7.20 (and Theorem 8.13) was obtained by C. Bennett [7], using the boundedness of the Hardy-Littlewood maximal operator on BMO, a result established in C. Bennett, R. A. DeVore & R. Sharpley [2]. The analogous results for the nontangential maximal operator are in C. Bennett [8]. Boundedness of the Littlewood-Paley g-function on BMO (precisely, on the subset of BMO for which $g(f) \neq \infty$) has been demonstrated by S. Wang [1]. Corresponding results for the area integral and the Littlewood-Paley functions g_λ^* have been obtained by D. S. Kurtz [2]. The K-functional between H^1 and BMO has been described by R. A. DeVore [1] and B. Jawerth [1]. Spaces defined in terms of functionals related to the sharp function have been considered by many authors; see S. Campanato [1], A. P. Calderón [4], J. Peetre [3], R. A. DeVore & R. Sharpley [1] and the references cited there.

The principal result of §8, namely, the description of the K-functional for L^1 and BMO, is due to C. Bennett & R. Sharpley [1]. The (θ, q)-interpolation spaces between L^1 and BMO had previously been described by R. Hanks [1], without explicitly determining the K-functional. Lemma 8.9, Proposition 8.10, and Theorem 8.11 are in C. Bennett & R. Sharpley [2].

The material in §9 is due almost entirely to P. Jones [3]. That paper contains also analytic decompositions of H^p-spaces (cf. Exercise 24) and constructive methods for interpolating ℓ^p data with H^p-functions. Theorem 9.2 is due to L. Carleson [1, 2]. The book by J. B. Garnett [1] contains a detailed account of interpolating sequences.

The K-functional for H^1 and H^∞ (Theorem 10.1) was determined by P. Jones [3] and its simplified form (Corollary 10.2) by R. Sharpley [9]. The monotonicity of all interpolation spaces for this couple was established by P. Jones [4]; the presentation given here incorporates the techniques of [4] with those of R. Sharpley [11]. Interpolation between H^1 and H^∞ by the complex method has been considered by S. Janson & P. W. Jones [1].

Appendix A

BLASCHKE PRODUCTS IN THE UPPER HALF PLANE U.

We briefly describe Blaschke products for the upper half plane **U**. It is assumed that the reader is familiar with the properties of Blaschke products in the unit disk **D**, as presented in J. B. Garnett [1], for example. Let Φ be the linear fractional transformation

$$z = \Phi(w) = i\frac{1-w}{1+w}, \qquad (w \in \mathbf{D}),$$

which takes **D** onto **U**, and let Ψ be the inverse mapping,

$$w = \Psi(z) = \frac{i-z}{i+z}, \qquad (z \in \mathbf{U}).$$

On the disk, the Blaschke factors (modulo rotations) are the conformal self maps ϕ_a of **D** given by

$$\phi_a(w) = -\frac{\bar{a}}{|a|}\frac{w-a}{1-\bar{a}w},$$

where $|a| < 1$. *If* $w = \Psi(z)$ and $\xi = \Phi(a)$, then

$$\phi_a(w) = \frac{|1+\xi^2|}{1+\xi^2}\frac{z-\xi}{z-\bar{\xi}}.$$

If $\{w_k\}$ is a sequence in **D**, so $z_k = \Phi(w_k)$ defines the corresponding sequence in **U**, then the Blaschke product

$$B(z) = \prod_k \frac{|1+z_k^2|}{1+z_k^2}\frac{z-z_k}{z-\bar{z}_k}, \tag{A.1}$$

converges uniformly on compact subsets of **U** precisely when

$$\sum_k (1 - |w_k|) < \infty.$$

But if $z_k = x_k + iy_k$, it follows easily from the identity

$$1 - |w_k| = \frac{4y_k}{|z_k + i|(|z_k + i| + |z_k - i|)}$$

that

$$\frac{y_k}{1 + |z_k|^2} \le 1 - |w_k| \le 4\frac{y_k}{1 + |z_k|^2}.$$

Hence $B(z)$ converges to a bounded analytic function on **U** if and only if

$$\sum_k \frac{y_k}{1 + |z_k|^2} < \infty.$$

Moreover, it follows from the corresponding properties of Blaschke products on **D** that $|B(z)| = 1$ a.e. for all real z.

If F belongs to the Hardy space $H^p(\mathbf{U})$ for some $p \ge 1$, then

$$\int_{-\pi}^{\pi} |F(\Phi(e^{i\theta}))|^p \, d\theta \le 2 \int_{-\infty}^{\infty} |F(x)|^p \, dx$$

and so $f = F \circ \Phi$ belongs to $H^p(\mathbf{D})$. If the zeros of f are denoted by $\{w_k\}$, the Poisson-Jensen formula shows that $\Sigma_k(1 - |w_k|) < \infty$. The Blaschke product $B(z)$ given in (A.1) therefore converges uniformly on compact subsets of **U** and its zeros are precisely the zeros $\{z_k\}$ of F. Hence, $G = F/B$ is a zero-free analytic function on **U** and

$$\|G\|_{H^p(\mathbf{U})} = \|F\|_{H^p(\mathbf{U})}.$$

References

[DS] N. Dunford and J. Schwartz, *Linear Operators*. Interscience, New York, 1958.
[Ha] P. R. Halmos, *Measure Theory*, Van Nostrand, New York, 1950.
[HS] E. Hewitt and K. Stromberg, *Real and Abstract Analysis*, Springer-Verlag, New York, 1965.
[Ro] H. L. Royden, *Real Analysis*, MacMillan, New York, 2nd ed., 1968.
[Ru] W. Rudin, *Functional Analysis*, McGraw-Hill, New York, 1973.

Bibliography

ADAMS, R. A.

[1] *Sobolev Spaces*, Academic Press, New York, 1975.

AMEMIYA, I.

[1] "A generalization of Riesz-Fischer's theorem," *J. Math. Soc. Japan* **5** (1953), 353–354.

ARAZY, J. & CWIKEL, M.

[1] "A new characterization of the interpolation spaces between L^p and L^q" (preprint).

ARONSZAJN, N. & GAGLIARDO, E.

[1] "Interpolation spaces and interpolation methods," *Ann. Mat. Pura Appl.* (4)**68** (1965), 51–118.

BABENKO, K. I.

[1] "An inequality in the theory of Fourier integrals," *Izv. Akad. Nauk SSSR, Ser. Mat.* **25** (1961), 531–542; *Amer. Math. Soc. Transl.* (2)**44**, 115–128.

BAERNSTEIN, A.

[1] "Integral means, univalent functions and circular symmetrization," *Acta Math.* **133** (1974), 139–169.

[2] "Some sharp inequalities for conjugate functions," *Proc. Symp. Pure Math.* **35**(1) (1979), 409–416.

BAGBY, R. J. & KURTZ, D. S.

[1] "Covering lemmas and the sharp function," *Proc. Amer. Math. Soc.* **93** (1985), 291–296.

[2] "A rearranged good λ inequality," *Trans. Amer. Math. Soc.* **293** (1986), 71–81.

BECKENBACH, E. F. & BELLMAN, R.
[1] *Inequalities*, Springer-Verlag, New York, 1965.
BECKNER, W.
[1] "Inequalities in Fourier Analysis," *Ann. of Math.* **102** (1975), 159–182.
BENNETT, C.
[1] "A pair of indices for function spaces on the circle," *Trans. Amer. Math. Soc.* **174** (1972), 289–304.
[2] "Intermediate spaces and the class $L \log^+ L$." *Ark. Mat.* **11**(1973), 215–228.
[3] "Banach function spaces and interpolation methods I," *J. Functional Anal.* **17** (1974), 409–440.
[4] "Banach function spaces and interpolation methods II," *Linear Operators and Approximation II*, ISNM **25**, Birkhäuser Verlag, Basel, 1974, 129–139.
[5] "Banach function spaces and interpolation methods III," *J. Approx. Theory* **13** (1975), 267–275.
[6] "A best constant for Zygmund's conjugate function inequality." *Proc. Amer. Math. Soc.* **56** (1976), 256–260.
[7] "Another characterization of BLO," *Proc. Amer. Math. Soc.* **85** (1982), 552–556.
[8] "Nontangential maximal functions and bounded lower oscillation." *Anniversary Volume on Approximation Theory and Functional Analysis*, ISNM **65**, Birkhäuser Verlag, Basel, 1984, 173–185.
BENNETT, C., DEVORE, R. A. & SHARPLEY, R.
[1] "Maximal singular integrals on L^∞," *Coll. Math. Soc. Janos Bolyai* **35** (1980), 233–236.
[2] "Weak-L^∞ and BMO," *Ann. of Math.* **113** (1981), 601–611.
BENNETT, C. & RUDNICK, K.
[1] "On Lorentz-Zygmund spaces." *Dissert. Math.* **175** (1980), 1–72.
BENNETT, C. & SHARPLEY, R.
[1] "Weak type inequalities for H^p and BMO," *Proc. Sympos. Pure Math.* **35** (1979), 201–229.
[2] "An inequality for the sharp function," *Quantitative Approximation*, Academic Press, New York, 1980, 1–6.
[3] "K-divisibility and a theorem of Lorentz and Shimogaki," *Proc. Amer. Math. Soc.* **96** (1986), 585–592.
BENNETT, G.
[1] "An extension of the Riesz-Thorin theorem," *Lecture Notes in Math.* **604**, 1–11, Springer-Verlag, Berlin, 1977.
BERENS, H.
[1] *Interpolationsmethoden zur Behandlung von Approximationsprozessen auf Banachräumen*, Lecture Notes in Math. **64**, Springer-Verlag, New York, 1968.
BERGH, J. & LÖFSTRÖM, J.
[1] *Interpolation Spaces. An Introduction*, Springer-Verlag, New York, 1976.

BOOLE, G.
[1] "On the comparison of transcendents, with applications to the theory of indefinite integrals," *Trans. Royal Soc.* **147** (1857), 778.
BOYD, D. W.
[1] "The Hilbert transform on rearrangement-invariant spaces," *Canad. J. Math.* **19** (1967), 599–616.
[2] "A class of operators on the Lorentz spaces $M(\phi)$," *Canad. J. Math.* **19** (1967), 839–841.
[3] "Spaces between a pair of reflexive Lebesgue spaces," *Proc. Amer. Math. Soc.* **18** (1967), 215–219.
[4] "The spectral radius of averaging operators," *Pacific J. Math.* **24** (1968), 19–28.
[5] "Indices of function spaces and their relationship to interpolation," *Canad. J. Math.* **21** (1969), 1245–1254.
[6] "Indices for the Orlicz spaces," *Pacific J. Math.* **38** (1971), 315–323.
BRUDNYI, JU. A. & KRUGLJAK, N. JA.
[1] "Real interpolation functors," *Soviet Math. Dokl.* **23** (1981), 5–8.
BURKHOLDER, D. L., GUNDY, R. F. & SILVERSTEIN, M. L.
[1] "A maximal function characterization of the class H^p," *Trans. Amer. Math. Soc.* **157** (1971), 137–153.
BUTZER, P. L. & BERENS, H.
[1] *Semi-Groups of Operators and Approximation*, Springer-Verlag, New York, 1967.
CALDERÓN, A. P.
[1] "Intermediate spaces and interpolation," *Studia Math.* (*special series*) **1** (1963), 31–34.
[2] "Intermediate spaces and interpolation, the complex method," *Studia Math.* **24** (1964), 113–190.
[3] "Spaces between L^1 and L^∞ and the theorem of Marcinkiewicz," *Studia Math.* **26** (1966), 273–299.
[4] "Estimates for singular integral operators in terms of maximal functions," *Studia Math.* **44** (1972), 563–582.
CALDERÓN, A. P. & ZYGMUND, A.
[1] "On the theorem of Hausdorff-Young and its extensions," *Ann. of Math. Studies* **25** (1950), 166–188.
[2] "A note on the interpolation of linear operations," *Studia Math.* **12** (1951), 194–204.
[3] "A note on the interpolation of sublinear operations," *Amer. J. Math.* **78** (1956), 282–288.
CALDERÓN, C. P. & MILMAN, M.
[1] "Interpolation of Sobolev spaces: the real method," *Indiana Univ. Math. J.* **32** (1983), 801–808.

CAMPANATO, S.
[1] "Proprietá de Hölderian di alcune classi di funzioni," *Ann. Scuola Norm. Sup. Pisa* **17**(3) (1963), 175–188.
CARLESON, L.
[1] "An interpolation problem for bounded analytic functions," *Amer. J. Math.* **80** (1958), 921–930.
[2] "Interpolation by bounded analytic functions and the corona problem," *Ann. of Math.* **76** (1962), 547–559.
[3] "Two remarks on H^1 and BMO," *Advances in Math.* **22** (1976), 269–277.
CHONG, K. M.
[1] "Variation reducing properties of decreasing rearrangements," *Canad. J. Math.* **27** (1975), 330–336.
CHONG, K. M. & RICE, N. M.
[1] "Equimeasurable Rearrangements of Functions," *Queen's Papers in Pure and Appl. Math.* **28**, Queen's Univ., Kingston, Ontario, 1971.
COIFMAN, R. R.
[1] "A real variable characterization of H^p," *Studia Math.* **51** (1974), 269–274.
COIFMAN, R. R., CWIKEL, M., ROCHBERG, R., SAGHER, Y. & WEISS, G.
[1] "Complex interpolation for families of Banach spaces," *Proc. Sympos. Pure Math.* **35**(2) (1979), 269–282.
[2] "The complex method for interpolation of operators acting on families of Banach spaces," *Lectures Notes in Math.* **779,** 123–153, Springer-Verlag, Berlin.
COIFMAN, R. R. & ROCHBERG, R.
[1] "Another characterization of BMO," *Proc. Amer. Math Soc.* **79** (1980), 249–254.
COIFMAN, R. R. & WEISS, G.
[1] "Extensions of Hardy spaces and their uses in analysis," *Bull. Amer Math. Soc.* **83** (1977), 569–645.
CWIKEL, M.
[1] "Monotonicity properties of interpolation spaces," *Ark. Mat.* **14** (1976), 213–236.
[2] "Monotonicity properties of interpolation spaces II," *Ark. Mat.* **19** (1981), 123–136.
[3] "K-divisibility of the K-functional and Calderón couples," *Ark. Mat.* **22** (1984), 39–62.
CWIKEL, M. & PEETRE, J.
[1] "Abstract K and J spaces," *J. Math. Pure et Appl.* **60** (1981), 1–50.
DAVIS, B.
[1] "On the weak type (1,1) inequality for conjugate functions," *Proc. Amer. Math. Soc.* **44** (1974), 307–311.
[2] "On Kolmogorov's inequalities $\|\tilde{f}\|_p \leq C_p\|f\|_1$. $0 < p < 1$," *Trans. Amer. Math. Soc.* **222** (1976), 179–192.

[3] "Hardy spaces and rearrangements," *Trans. Amer. Math. Soc.* **261** (1980), 211–233.
DAY, P. W.
[1] *Rearrangements of measurable functions*, Thesis, California Institute of Technology, 1970.
[2] "Decreasing rearrangements and doubly stochastic operators," *Trans. Amer. Math. Soc.* **178** (1973), 383–392.
DE GUZMÁN, M.
[1] *Real Variable Methods in Fourier Analysis*, North-Holland, Amsterdam, 1981.
DEVORE, R. A.
[1] "The K-functional for (H^1, BMO)." *Lecture Notes in Math.* **1070**, 66–79, Springer-Verlag, New York, 1984.
DEVORE, R. A. & POPOV, V.
[1] "Interpolation spaces and non-linear approximation" (preprint).
[2] "Interpolation of Besov spaces" (preprint).
DEVORE, R. A., RIEMENSCHNEIDER, S. D., & SHARPLEY, R. C.
[1] "Weak interpolation in Banach spaces." *J. Functional Anal.* **33** (1979), 58–94.
DEVORE, R. A. & SCHERER, K.
[1] "Interpolation of linear operators on Sobolev spaces," *Ann. of Math.* **109** (1979), 583–599.
DEVORE, R. A. & SHARPLEY, R. C.
[1] "Maximal Functions Measuring Smoothness," *Mem. Amer. Math. Soc.* **293**, Providence, 1984.
[2] "On the differentiability of functions in $\mathbf{R}^n$," *Proc. Amer. Math. Soc.* (*Shorter Notes*) **91** (1984), 326–328.
DUREN, P. L., ROMBERG, B. W. & SHIELDS, A. L.
[1] "Linear functionals on H^p spaces with $0 < p < 1$," *J. Reine Angew. Math.* **238** (1969), 32–60.
DUFF, G. F. D.
[1] "Differences, derivatives, and decreasing rearrangements," *Canad. J. Math.* **19** (1967), 1153–1178.
[2] "Integral inequalities for equimeasurable rearrangements," *Canad. J. Math.* **22** (1970), 408–430.
[3] "A general integral inequality for the derivative of an equimeasurable rearrangement," *Canad. J. Math.* **28** (1976), 793–804.
ESSÉN, M. & SHEA, D.
[1] "Some recent results on conjugate functions in the unit disk," *Progr. Math.* **11** (1981), 288–297.
FAN, K. & LORENTZ, G. G.
[1] "An integral inequality," *Amer. Math. Monthly* **61** (1954), 626–631.
FEFFERMAN, C.
[1] "Characterizations of bounded mean oscillation," *Bull. Amer. Math. Soc.* **77** (1971), 587–588.

FEFFERMAN, C., RIVIÈRE, N. & SAGHER, Y.
[1] "Interpolation between H^p spaces: The real method," *Trans. Amer. Math. Soc.* **191** (1974), 75–81.
FÉHER, F., GASPAR, D. & JOHNEN, H.
[1] "H^p spaces of several variables," *Acta Math.* **129** (1972), 137–193.
FÉHER, F., GASPAR, D. & JOHNEN, H.
[1] "Normkonvergenz von Fourier-reichen in rearrangement invarianten Banachraümen." *J. Functional Anal.* **13** (1973), 417–434.
[2] "Der Konjugiertenoperator auf rearrangement-invarianten Funktionenräumen," *Math. Z.* **134** (1973), 129–141.
FOLLAND, G. B.
[1] "On characterizations of analyticity," *Amer. Math. Monthly (8)***93** (1986), 640–641.
FRAZIER, M. & JAWERTH, B.
[1] "Decomposition of Besov spaces," *Indiana U. Math. J.* **34** (1985), 777–799.
[2] "A discrete transform and decompositions of distribution spaces" (preprint).
GABRIEL, R. M.
[1] "An additional proof of a maximal theorem of Hardy and Littlewood," *J. London Math. Soc.* **6** (1931), 163–166.
GARNETT, J. B.
[1] *Bounded Analytic Functions*. Academic Press, New York, 1981.
GARNETT, J. B. & JONES, P. W.
[1] "The distance in BMO to L^∞," *Ann. of Math.* **108** (1978), 373–393.
GARSIA, A. M.
[1] *Topics in Almost Everywhere Convergence*. Markham, Chicago, 1970.
GARSIA, A. M. & RODEMICH, E.
[1] "Monotonicity of certain functionals under rearrangement," *Ann. Inst. Fourier (Grenoble)* (2)**24** (1974), 67–116.
GILBERT, J. E.
[1] "Nikisin-Stein theory and factorization with applications," *Proc. Sympos. Pure Math.* (2)**35** (1979), 233–267.
GOKHBERG, T. & KRUPNIK, N.
[1] "Norm of the Hilbert transformation in the L^p space," *Functional Anal. and Appl.* (2)**2** (1968), 180–181.
GOULD, G. G.
[1] "On a class of integration spaces," *J. London Math. Soc.* **34** (1959), 161–172.
GROTHENDIECK, A.
[1] *Réarrangements de fonctions et inégalités de convexité dans les algèbres de von Neumann munies d'une trace*, Séminaire Bourbaki, **113**, 1955.
GUSTAVSSON, J. & PEETRE, J.
[1] "Interpolation of Orlicz spaces," *Studia Math.* **60** (1977), 33–59.
HALPERIN, I.
[1] "Function spaces," *Canad. J. Math.* **5** (1953), 273–288.

HALPERIN, I. & ELLIS, H. W.
[1] "Function spaces determined by a levelling length function," *Canad. J. Math.* **5** (1953), 576–592.
HALPERIN, I. & LUXEMBURG, W. A. J.
[1] "The Riesz-Fischer completeness theorem for function spaces and vector lattices," *Trans. Royal Soc. Canada* (3)**50** (1956), 33–39.
HANKS, R.
[1] "Interpolation by the real method between BMO, $L^\alpha(0 < \alpha < \infty)$, and $H^\alpha(0 < \alpha < \infty)$," *Indiana Univ. Math. J.* **26** (1977), 679–689.
HARDY, G. H. & LITTLEWOOD, J. E.
[1] "Some new properties of Fourier constants." *Math. Ann.* **97** (1927), 159–209.
[2] "A maximal theorem with function-theoretic applications," *Acta Math.* **54** (1930), 81–116.
[3] "Notes on the theory of series XIII: Some new properties of Fourier constants," *J. London Math. Soc.* **6** (1931), 3–9.
[4] "Some properties of conjugate functions," *J. Reine Angew. Math.* **167** (1932), 405–423.
[5] "Notes on the theory of series XVIII: On the convergence of Fourier series," *Proc. Camb. Phil. Soc.* **31** (1935), 317–323.
HARDY, G. H., LITTLEWOOD, J. E. & PÓLYA, G.
[1] *Inequalities*, Cambridge University Press, Cambridge, 1934 (2nd ed., 1952).
HERZ, C.
[1] *The Hardy-Littlewood maximal theorem*, Symposium on Harmonic Analysis, University of Warwick, 1968.
HEWITT, E. & HIRSCHMANN, I.I.
[1] "A maximum problem in harmonic analysis," *Amer. J. Math.* **76** (1954), 839–851.
HIRSCHMAN, I.I.
[1] "A note on orthogonal systems," *Pacific J. Math.* **6** (1956), 47–56.
HOLMSTEDT, T.
[1] "Interpolation of quasi-normed spaces," *Math. Scand.* **26** (1970), 177–199.
HUNT, R. A.
[1] "An extension of the Marcinkiewicz interpolation theorem to Lorentz spaces," *Bull. Amer. Math. Soc.* **70** (1964) 803–807.
[2] "On $L(p,q)$ spaces," *L'Enseignement Math.* (2)**12** (1966), 249–275.
HUNT, R. A. & WEISS, G.
[1] "The Marcinkiewicz interpolation theorem." *Proc. Amer. Math. Soc.* (*Shorter Notes*) **15** (1964), 996–998.
JAMES, R. C.
[1] "Bases and reflexivity of Banach spaces," *Ann. of Math.* **52** (1950), 518–527.
JANSON, S.
[1] "Minimal and maximal methods of interpolation." *J. Functional Anal.* **14** (1981), 50–72.

Janson, S. & Jones, P. W.
[1] "Interpolation between H^p-spaces: the complex method," *J. Functional Anal.* **48** (1982), 58–80.
Janson, S., Nilsson, P. & Peetre, J.
[1] "Notes on Wolff's note on interpolation spaces," *Proc. London Math. Soc.* **48** (1984), 283–299.
Jawerth, B.
[1] "The K-functional for H^1 and BMO," *Proc. Amer. Math. Soc.* **92** (1984), 67–71.
John, F. & Nirenberg, L.
[1] "On functions of bounded mean oscillation," *Comm. Pure Appl. Math.* **14** (1961), 415–426.
Johnen, H. & Scherer, K.
[1] "On the equivalence of the K-functional and moduli of continuity and some applications," *Lecture Notes in Math.* **571**, 119–140, Springer-Verlag, Berlin 1976.
Jones, P. W.
[1] "Carleson measures and the Fefferman-Stein decomposition of BMO(**R**)," *Ann. of Math.* **111** (1980), 197–208.
[2] "Factorization of A_p weights," *Ann. of Math.* **111** (1980), 511–530.
[3] "L^∞ estimates for the $\bar{\partial}$ problem in a half-plane," *Acta Math.* **150** (1983), 137–152.
[4] "On interpolation between H^1 and H^∞," *Lecture Notes in Math.* **1070**, 143–151, Springer-Verlag, New York, 1984.
Kalton, N. J.
[1] "Nonlinear commutators in interpolation theory" (preprint).
Klemes, I.
[1] "A mean oscillation inequality," *Proc. Amer. Math. Soc.* **93** (1985), 497–500.
Kolmogorov, A. N.
[1] "Sur les fonctions harmoniques conjuguées et les séries de Fourier," *Fundamenta Math.* **7** (1925), 23–28.
Korenblyum, B. I., Krein, S. G. & Levin, B. Ya.
[1] "On certain nonlinear questions of the theory of singular integrals," *Dokl. Akad. Nauk SSSR* **62** (1948), 17–20.
Korneicuk, N. P.
[1] "The best uniform approximation on certain classes of continuous functions," *Dokl. Akad. Nauk SSSR* **141** (1961) = *Soviet Math. Dokl.* **2** (1961), 1254–1257.
Köthe, G.
[1] *Topologische Lineare Räume*, Springer-Verlag, Berlin, 1960.
Krasnosel'skii, M. A.
[1] "On a theorem of M. Riesz," *Dokl. Akad. Nauk SSSR* **131** (1960), 246–248 = *Soviet Math. Dokl.* **1**(1960), 229–231.
Krasnosel'skii, M. A. & Ruticku, Ya. B.
[1] *Convex functions and Orlicz spaces*, GITTL, Moscow, 1958; English transl., Noordhoff, Groningen, 1961.

Krein, S. G.
[1] "On an interpolation theorem in operator theory," *Dokl. Akad. Nauk SSSR* **130** (1960), 491–494 = *Soviet Math. Dokl.* **1** (1960), 61–64.
[2] "On the concept of a normal scale of spaces," *Dokl. Akad. Nauk SSSR* **138** (1960), 510–513 = *Soviet Math. Dokl.* **1** (1960), 586–589.
Krein, S. G. & Nikolova, L. J.
[1] "Holomorphic functions in a family of Banach spaces, and interpolation," *Soviet Math. Dokl.* **21** (1980), 131–134.
Krein, S. G., Petunin, Ju. I. & Semenov, E. M.
[1] *Interpolation of Linear Operators*, Transl. Math. Monogr. **54**, Amer. Math. Soc., Providence, 1982.
Krein, S. G. & Semenov, E. M.
[1] "Interpolation of operators of weakened type," *Funkcional Anal. i Prilozen* **7** (1973), 89–90.
Kurtz, D. S.
[1] "Rearrangement inequalities for Littlewood-Paley operators" (preprint).
[2] "Littlewood-Paley operators on BMO" (preprint).
Latter, R. H.
[1] "A decomposition of $H^p(\mathbf{R}^n)$ in terms of atoms." *Studia Math.* **62** (1978), 92–101.
Lindenstrauss, J. & Tzafriri, L.
[1] *Classical Banach Spaces*. Springer-Verlag, New York, 1977.
Lions, J. L.
[1] "Une construction d'espaces d'interpolation," *C. R. Acad. Sci. Paris* **251** (1961), 1853–1855.
[2] "Sur les espaces d'interpolation: dualité," *Math. Scand.* **9** (1961), 147–177.
Lions, J. L. & Magenes, E.
[1] *Nonhomogeneous boundary value problems and applications*, Vol. **1–3**, Springer-Verlag, New York, 1972–73.
Loomis, L. H.
[1] "A note on the Hilbert transform," *Bull. Amer. Math. Soc.* **52** (1946), 1082–1086.
Lorentz, G. G.
[1] "Some new functional spaces," *Ann. of Math.* **51** (1950), 37–55.
[2] "On the theory of spaces Λ," *Pacific J. Math.* **1** (1951), 411–429
[3] "An inequality for rearrangements," *Amer. Math. Monthly* **60** (1953), 176–179.
[4] *Bernstein Polynomials*. University of Toronto Press, Toronto, 1953.
[5] "Majorants in spaces of integrable functions," *Amer. J. Math.* **77** (1955), 484–492.
[6] "Spaces of measurable functions," (*unpublished manuscript*).
[7] "Relations between function spaces," *Proc. Amer. Math. Soc.* **12** (1961), 127–132.
[8] *Approximation of Functions*. Holt, Rinehart and Winston, New York, 1966.

LORENTZ, G. G. & SHIMOGAKI, T.

[1] "Interpolation theorems for operators in function spaces," J. *Functional Anal.* **2** (1968), 31–51.

[2] "Interpolation theorems for spaces Λ," *Abstract Spaces and Approximation*, ISNM **10**, 94–98, Birkhäuser Verlag, Basel, 1969.

[3] "Interpolation theorems for the pairs of spaces (L^p, L^∞) and (L^1, L^q)," *Trans. Amer. Math. Soc.* **159** (1971), 207–221.

LORENTZ, G. G. & WERTHEIM, D. G.

[1] "Representation of linear functionals on Köthe spaces," *Canad. J. Math.* **5** (1953), 568–575.

LUXEMBURG, W. A. J.

[1] *Banach function spaces*, Ph. D. Thesis, Delft Institute of Technology, Assen (Netherlands), 1955.

[2] *The Hausdorff-Young-Riesz theorem in Orlicz spaces*, Report of the Summer Research Institute of the Canadian Mathematical Congress (1957), 14–15.

[3] "Rearrangement-invariant Banach function spaces," *Proc. Sympos. in Analysis, Queen's Papers in Pure and Appl. Math.* **10** (1967), 83–144.

LUXEMBURG, W. A. J. & ZAANEN, A. C.

[1] "Compactness of integral operators in Banach function spaces," *Math. Ann.* **149** (1963), 150–180.

[2] "Notes on Banach function spaces, I-V," *Nederl. Akad. Wetensch. Proc. Ser. A 59 = Indag. Math.* **18** (1963) 135–147; 148–153; 239–250; 251–263; 496–504.

[3] "Some examples of normed Köthe spaces," *Mat. Ann.* **162** (1966), 337–350.

[4] *Riesz spaces*, North-Holland, Amsterdam, 1971.

MALIGRANDA, L.

[1] "The K-functional for symmetric spaces," *Lecture Notes in Math.* **1070**, 169–182, Springer-Verlag, Berlin, 1984.

[2] "Interpolation between sum and intersection of Banach spaces," Preprint 308, Polish Academy of Sciences, 1984.

[3] "Indices and interpolation," *Dissert. Math.* **234** (1985), 1–49.

MARCHAUD, A.

[1] "Sur les dérivées et sur les différences des fonctions des variables réelles," *J. Math. Pure et Appl.* **6** (1927), 337–425.

MARCINKIEWICZ, J.

[1] "Sur l'interpolation d'opérations," *C. R. Acad. Sci. Paris* **208** (1939), 1272–1273.

MARSHALL, A. W. & OLKIN, I.

[1] *Inequalities: Theory of Majorization and its Applications*, Academic Press, New York, 1979.

MAUREY, B.

[1] "Théorèmes de factorisation pour les opérateurs linéaires à valeurs dans les espaces L^p," *Astérisque* **11** (1974), 1–163.

MILMAN, M.

[1] "Rearrangements of BMO functions and interpolation," *Lecture Notes in Math.* **1070**, 208–212, Springer-Verlag, New York, 1984.

MITJAGIN, B. S.

[1] "An interpolation theorem for modular spaces," *Mat. Sbornik* **66** (1965), *473–482 = Lecture Notes in Math.* **1070,** 10–23, Springer-Verlag, New York, 1984.

MITJAGIN, B. S. & SEMENOV, E. M.

[1] "Lack of interpolation of linear operators between spaces of smooth functions," *Izvestia* **41** (1977), 1289–1328 [*English transl.* **11** (1977), 1229–1266].

MOON, K. H.

[1] "On restricted weak type (1, 1)," *Proc. Amer. Math. Soc.* **42** (1974), 148–152.

MUIRHEAD, R. F.

[1] "Some methods applicable to identities and inequalities of symmetric algebraic functions of n letters," *Proc. Edinburgh Math. Soc.* **21** (1903), 144–157.

NATANSON, I. P.

[1] "On an inequality," *Dokl. Akad. Nauk SSSR* **56** (1947), 911–913.

[2] *Theorie der Funktionen einer reellen Veränderlichen.* Berlin, 1954.

NIKOLOVA, L. J.

[1] "Complex method of interpolation of some spaces," *C. R. Acad. Sci. Bulg.* **31** (1978), 1523–1526.

NILSSON, P.

[1] "Reiteration theorems for real interpolation and approximation spaces," *Ann. Mat. Pura Appl.* (4) **32** (1982), 291–330.

OKLANDER, E. T.

[1] "Interpolacion, Espacios de Lorentz y Teorema de Marcinkiewicz," *Cursos y Sem. Mat.* **20**, Univ. Buenos Aires, 1965.

O'NEIL, R.

[1] "Convolution operators and $L(p, q)$ spaces," *Duke Math. J.* **30** (1963), 129–142.

O'NEIL R. & WEISS, G.

[1] "The Hilbert transform and rearrangement of functions," *Studia Math.* **23** (1963), 189–198

ORLICZ, W.

[1] "Über eine gewisse Klasse von Räumen vom Typus B," *Bull. Int. Acad. Polon. Sci. Lett. Cl. Math. Nat.* **A** (1932), 207–220.

[2] "Ein Satz über die Erweiterung von linearen Operationen," *Studia Math.* **15** (1934), 127–140.

[3] "Über Räume (L^M)," *Bull. Int. Acad. Polon. Sci. Lett. Cl. Math. Nat.* **A** (1936), 93–107.

[4] "On a class of operations over the space of integrable functions," *Studia Math.* **14** (1953–4), 302–309.

OVCINNIKOV, V. I.

[1] "On estimates of interpolation orbits," *Mat. Sbornik* **115** (1981), 642–652 (Russian).

[2] "The method of orbits in interpolation theory," *Math. Reports* **1** (1984), 349–515.

PALEY, R. E. A. C.
[1] "A proof of a theorem on bilinear forms," *J. London Math. Soc.* **6** (1931), 226–230.

PEETRE, J.
[1] "Nouvelles propriétés d'espaces d'interpolation," *C. R. Acad. Sci. Paris* **256** (1963), 1424–1426.
[2] "A Theory of Interpolation of Normed Spaces," *Notas de Matematica*, Rio de Janeiro, **39** (1963), 1–86.
[3] "On the theory of $L_{p,\lambda}$ spaces," *J. Functional Anal.* **4** (1969), 71–87.
[4] "Exact interpolation theorems for Lipschitz continuous functions," *Ricerche Mat.* **18** (1969), 239–259.
[5] *New Thoughts on Besov Spaces*, Duke Univ. Math. Series, Duke Univ., Durham, N. C., 1976.
[6] "Two new interpolation methods based on the duality map," *Acta Math.* **143** (1979), 73–91.

PICHORIDES, S. K.
[1] "On the best values of the constants in the theorems of M. Riesz, Zygmund and Kolmogorov," *Studia Math.* **44** (1972), 165–179.

RIEMENSCHNEIDER, S. D.
[1] "The L-characteristics of linear operators on $L^{1/\alpha}([0,1])$," *J. Functional Anal.* **8** (1971), 405–421.
[2] "Linear operators on $L^{1/\alpha}(0,\infty)$ and Lorentz spaces: The Krasnosel's skii-Zabrieko characteristic sets," *Studia Math.* **49** (1973), 225–233.

RIESZ, F.
[1] "Über eine Verallgemeinerung des Parsevalschen Formel," *Math. Z.* **18** (1923), 117–124.
[2] "Sur un théorème de maximum de MM. Hardy et Littlewood," *J. London Math. Soc.* **7** (1932), 10–13.

RIESZ, M.
[1] "Sur les maxima des formes bilinéaries et sur les fonctionelles lineaires," *Acta Math.* **49** (1926), 465–497.
[2] "Sur les fonctions conjuguées," *Math. Z.* **27** (1927), 218–244.

ROCHBERG, R. & G. WEISS
[1] "Derivatives of analytic families of Banach spaces," *Ann. of Math.* **118** (1983), 315–347.

RUDERMAN, H.
[1] "Two new inequalities," *Amer. Math. Monthly* **59** (1952), 29–32.

RYFF, J. V.
[1] "Orbits of L^1-functions under doubly stochastic transformations," *Trans. Amer. Math. Soc.* **117** (1965), 92–100.
[2] "Measure preserving transformations and rearrangements," *J. Math. Anal. Appl.* **31** (1970), 449–458.

SALEM, R.
[1] "Convexity theorems," *Bull. Amer. Math. Soc.* **55** (1949), 851–860.

SALEM, R. & ZYGMUND, A.
[1] "A convexity theorem," *Proc. Nat. Acad. Sci. USA* **34** (1948), 443–447.

SCHWARZ, H. A.
[1] *Beweis des Satzes dass die Kugel kleinere Oberfläche besitzt, als jeder andere Körper gleichen Volumens*, Göttinger Nachrichten (1884), 1–13 [Werke **2**, 327–340].

SEDAEV, A. A. & SEMENOV, E. M.
[1] "On the possibility of describing interpolation spaces in terms of Peetre's K-method," *Optimizaciya* **4**(21) (1971), 98–114 [Russian].

SEMENOV, E. M.
[1] "Imbedding theorems for Banach spaces of measurable functions," *Soviet Math. Dokl.* **5** (1964), 831–834.

SHARPLEY, R.
[1] "Spaces $\Lambda_\alpha(X)$ and interpolation," *J. Functional Anal.* **11** (1972), 479–513.
[2] "Interpolation theorems for compact operators," *Indiana U. Math. J.* **22** (1973), 965–984.
[3] "Interpolation of operators for Λ_ϕ spaces," *Bull. Amer. Math. Soc.* **80** (1974), 259–261.
[4] "Characterization of intermediate spaces of M_ϕ spaces," *Linear Operators and Approximation II*, ISNM **25**, Birkhäuser Verlag, 1974, 205–214.
[5] "Interpolation of n pairs and counterexamples employing indices," *J. Approx. Theory* **13** (1975), 117–127.
[6] "Multilinear weak type interpolation of m n-tuples with applications," *Studia Math.* **60** (1977), 179–194.
[7] "Counterexamples for classical operators on Lorentz-Zygmund spaces," *Studia Math.* **68** (1980), 141–158.
[8] "Cone conditions and the modulus of continuity," *Proc. Canad. Math. Soc.* **3** (1983), 341–351.
[9] "Interpolation of H^1 and H^∞," *Anniversary Volume on Approximation Theory and Functional Analysis*, ISNM **65**, Birkhäuser Verlag, Basel, 1984, 207–211.
[10] "On the atomic decomposition of H^1 and interpolation," *Proc. Amer. Math. Soc. (Shorter Notes)* **97** (1986), 186–188.
[11] "A characterization of the interpolation spaces of H^1 and L^∞" (preprint).

SHIMOGAKI, T.
[1] "Hardy-Littlewood majorants in function spaces," *J. Math Soc. Japan* **17** (1965), 365–373.
[2] "A note on norms of compression operators on function spaces," *Proc. Japan Acad.* **46** (1970), 239–242.

SIMON, B.
[1] *Trace ideals and their applications*. Cambridge University Press, Cambridge 1979.

SIMONENKO, I. B.
[1] "Boundedness of singular integral operators in Orlicz spaces," *Dokl. Akad. Nauk SSSR* **130** (1960), 984–987.

[2] "Interpolation and extrapolation in Orlicz spaces," *Mat. Sbornik* **63** (1964), 536–553.

SPANNE, S.

[1] "Sur l'interpolation entre les espaces $L_k^{p\Phi}$," *Ann. Scuola Norm. Sup. Pisa Sci. Fis. Mat.* **20** (1966), 625–648.

SPARR, G.

[1] "Interpolation des éspaces L_w^p," *C. R. Acad. Sci. Paris* **278** (1974), 491–492.

[2] "Interpolation of weighted L_p spaces," *Studia Math.* **62** (1978), 229–271.

STANTON, C. S.

[1] "An $L\log L$ characterization of H^1," *J. Functional Anal.* **69** (1986), 409–418.

STEIN, E. M.

[1] "Interpolation of linear operators," *Trans. Amer. Math. Soc.* **83** (1956), 482–492.

[2] "On limits of sequences of operators," *Ann. of Math.* **74** (1961), 140–170.

[3] "Singular integrals, harmonic functions, and differentiability properties of functions," *Proc. Sympos. Pure Math.* **10** (1967), 316–335.

[4] "Note on the class $L\log L$," *Studia Math.* **31** (1969), 305–310.

[5] *Singular Integrals and Differentiability Properties of Functions.* Princeton University Press, Princeton, 1970.

[6] "The differentiability of functions in $\mathbf{R}^n$," *Ann. of Math.* **113** (1981), 383–385.

STEIN, E. M. & WEISS, G.

[1] "On the interpolation of analytic families of operators acting on H^p-spaces," *Tôhoku Math.* J. **9** (1957), 318–339.

[2] "Interpolation of operators with change of measures," *Trans. Amer. Math. Soc.* **87** (1958), 159–172.

[3] "An extension of a theorem of Marcinkiewicz and some of its applications," *J. Math. Mech.* **8** (1959), 263–284.

[4] *Introduction to Fourier Analysis on Euclidean Spaces.* Princeton University Press, Princeton, 1971.

STEINER, J.

[1] *Gesammelte Werke*, 1881–82, Berlin.

TAIBLESON, M. H.

[1] "On the theory of Lipschitz spaces of distributions in Euclidean n-space. I Principal properties," *J. Math. Mech.* **13** (1964), 407–479; "II. Translation invariant operators, duality, and interpolation," *ibid.* **14** (1965), 821–839.

TAMARKIN, J. D. & ZYGMUND, A.

[1] "Proof of a theorem of Thorin," *Bull. Amer. Math. Soc.* **50** (1944), 279–282.

THORIN, G. O.

[1] "An extension of a convexity theorem due to M. Riesz," *Kungl. Fys. Saell. i Lund For.* **8** (1939), # 14.

[2] "Convexity theorems," *Diss. Lund*, (1948), 1–57.

TITCHMARSH, E. C.

[1] "A contribution to the theory of Fourier transforms," *Proc. London Math. Soc.* **23** (1924), 279–289.

[2] "On conjugate functions," *Proc. London Math. Soc.* **29** (1928), 49–80.

[3] "Additional note on conjugate functions," *J. London Math. Soc.* **4** (1929), 204–206.

TORCHINSKY, A.

[1] *Real-Variable Methods in Harmonic Analysis.* Academic Press, New York, 1986.

TRIEBEL, H.

[1] *Interpolation Theory, Function Spaces, Differential Operators.* Berlin VEB, 1978.

UCHIYAMA, A.

[1] "A constructive proof of the Fefferman-Stein decomposition of BMO ($\mathbf{R}^n$)," *Acta Math.* **148** (1982), 215–241.

VAROPOULOS, N. TH.

[1] "BMO functions and the $\bar{\partial}$ equation," *Pacific J. Math.* **71** (1977), 221–273.

VERBLUNSKY, S.

[1] "Fourier constants and Lebesgue classes," *Proc. London Math. Soc.* **39** (1935), 1–31.

WANG, S.

[1] "Some properties of the Littlewood-Paley g-function," *Cont. Math.* **42** (1985), 191–202.

WEISS, G.

[1] "An interpolation theorem for sublinear operations on H^p-spaces." *Proc. Amer. Math. Soc.* **8** (1957), 92–99.

WIENER, N.

[1] "The ergodic theorem," *Duke Math. J.* **5** (1939), 1–18.

WOLFF, T. H.

[1] "A note on interpolation spaces," *Lecture Notes in Math.* **908,** Springer-Verlag, Berlin, 1982.

YOUNG, L. C.

[1] "On an inequality of Marcel Riesz," *Ann. of Math.* **40** (1939), 567–574.

ZAANEN, A. C.

[1] *Integration*, North-Holland, Amsterdam, 1967.

ZABRIEKO, P. P. & KRASNOSEL'SKII, M. A.

[1] "On the L-characteristics of operators," *Uspekhi Mat. Nauk* **19** (1964), 187–189.

ZYGMUND, A.

[1] "Sur les fonctions conjuguées," *C. R. Acad. Sci. Paris* **187** (1928), 1025–1026; *Fund. Math.* **13** (1929), 284–303.

[2] "Some points in the theory of trigonometric and power series," *Trans. Amer. Math. Soc.* **36** (1934), 586–617.

[3] "On a theorem of Marcinkiewicz concerning interpolation of operations," *J. Math. Pure et Appl.* **35** (1956), 223–248.

[4] *Trigonometric series*, Cambridge Univ. Press, 2nd ed., 1968.

Index

List of Notations

The following notation is used throughout the book.

Theorems, Lemmas, Definitions, etc., are numbered consecutively throughout each section of each chapter. Thus Theorem 2.7 refers to the seventh numbered item in section 2 of the current chapter; Theorem III.2.7 refers to the like numbered item in Chapter 3.

$\subset$, $A \subset B$: A is a subset of B
$\hookrightarrow$, $X \hookrightarrow Y$: X is continuously embedded in Y (X and Y are topological spaces with $X \subset Y$ and the inclusion map is continuous)
$\to$, $f_n \to f$: the sequence (f_n) converges to f
$\uparrow$, $f_n \uparrow f$: the sequence (f_n) is monotone increasing and $f_n \to f$
R, **C** denote respectively the fields of real and complex numbers
T, **Z** denote respectively the unit circle group and the group of integers

The following symbols are introduced in the text at the page numbers indicated.