EXTRAPOLATION AND INTERPOLATION OF QUASI-LINEAR OPERATORS ON MARTINGALES

BY

D. L. BURKHOLDER and R. F. GUNDY

University of Illinois, Urbana, Illinois (1)

Rutgers University, New Brunswick, New Jersey, and the Hebrew University of Jerusalem (2)

Contents

---

(1) The research of the first-named author was supported in part by the Center for Advanced Study of the University of Illinois and by the National Science Foundation under grant GP–8727.

(2) The research of the second-named author was supported in part by the Research Council of Rutgers University while the author was in residence at the University of Illinois (summer 1968) and the Hebrew University of Jerusalem (academic year 1968–69) and by the National Science Foundation under grant GP–8056.

Both authors acknowledge the generous hospitality of the Institute of Mathematics, Hebrew University, and the Department of Mathematics, Westfield College, University of London.

1. Introduction

In this paper we introduce a new method to obtain one-sided and two-sided integral inequalities for a class of quasi-linear operators. Some of our assumptions are similar to those of the Marcinkiewicz interpolation theorem. However, in contrast to the Marcinkiewicz theorem, the operators that we study here are local in a certain sense and are usually most conveniently defined on martingales. In fact, the suitable choice of starting and stopping times for martingales, together with the systematic use of maximal functions and maximal operators, is central to our method.

Before describing our results in detail, we consider a few simple applications. We begin with an application to classical orthogonal series.

Let $\chi_0, \chi_1, \ldots$ be the complete orthonormal system of Haar functions on the Lebesgue unit interval. Let $\sum_{k=0}^{\infty} a_k \chi_k$ be the Haar–Fourier series of an integrable function f and $S(f) = [\sum_{k=0}^{\infty} (a_k \chi_k)^2]^{\frac{1}{2}}$. Then

$$c_p \|S(f)\|_p \leqslant \|f\|_p \leqslant C_p \|S(f)\|_p, \quad 1 < p < \infty. \tag{1.1}$$

This inequality is due to R. E. A. C. Paley [14], who stated it in an equivalent form for Walsh series; the Haar series version (1.1) was noted by Marcinkiewicz [11]. Inequality (1.1) should be compared with the inequality

$$c_p \|f^*\|_p \leqslant \|f\|_p \leqslant \|f^*\|_p, \quad 1 < p < \infty, \tag{1.2}$$

where $f^* = \sup_n |\sum_{k=0}^{n} a_k \chi_k|$, which follows from the maximal inequality of Hardy and Littlewood [10]. The two inequalities imply that

$$c_p \|S(f)\|_p \leqslant \|f^*\|_p \leqslant C_p \|S(f)\|_p, \tag{1.3}$$

for $1 < p < \infty$. Although it is known that neither (1.1) nor (1.2) hold in general for $0 < p \leqslant 1$, our results reveal a quite different picture for the last inequality: from the fact that (1.3) holds for $1 < p < \infty$, we are able to show that it holds for the entire range $0 < p < \infty$. This extrapolation effect is typical of our method. Even more is true: the fact that (1.3) holds for two values of p is enough to imply that (1.3) holds for all p.

The next example has many of the same elements. Suppose that $X = \{X(t), 0 \leqslant t < \infty\}$ is standard Brownian motion (see Section 7) and τ is a stopping time of X. Let X^τ be the process X stopped at τ: $X^\tau(t) = X(\tau \wedge t)$, $0 \leqslant t < \infty$. Its maximal function is defined by

$$(X^\tau)^* = \sup_{0 \leqslant t < \infty} |X^\tau(t)|.$$

Let b be a positive real number and consider the stopping time $\tau \wedge b$. Then the inequalities

$$c_p \|(\tau \wedge b)^{\frac{1}{2}}\|_p \leq \|X(\tau \wedge b)\|_p \leq C_p\|(\tau \wedge b)^{\frac{1}{2}}\|_p \tag{1.1'}$$

and

$$c_p \|(X^{\tau \wedge b})^*\|_p \leq \|X(\tau \wedge b)\|_p \leq \|(X^{\tau \wedge b})^*\|_p \tag{1.2'}$$

are known to hold for all $1 < p < \infty$. The first follows from the results of Millar [13], and the second is a standard martingale maximal inequality (see Doob [5], Chapter VII, Theorem 3.4 and page 354). If we combine these inequalities as before, and use the monotone convergence theorem, we obtain

$$c_p \|\tau^{\frac{1}{2}}\|_p \leq \|(X^{\tau})^*\|_p \leq C_p \|\tau^{\frac{1}{2}}\|_p \tag{1.3'}$$

for $1 < p < \infty$. As in the previous example, it is known that neither (1.1′) nor (1.2′) can be extended to the interval $0 < p \leq 1$. However, again the last inequality is different: our method shows that (1.3′) is, in fact, valid for the entire range $0 < p < \infty$.

Related integral inequalities for stopped random walk and sums of independent random variables are given in Section 5.

Both of the above examples may be considered from a common viewpoint. Let $f = (f_1, f_2, \dots)$ be a martingale on some probability space and $d = (d_1, d_2, \dots)$ its difference sequence, so that

$$f_n = \sum_{k=1}^{n} d_k, \quad n \geq 1.$$

Let f^* denote the maximal function of the sequence $f: f^* = \sup_n |f_n|$. The maximal function is related to the function $S(f) = (\sum_{k=1}^{\infty} d_k^2)^{\frac{1}{2}}$ by the inequality

$$c_p \|S(f)\|_p \leq \|f^*\|_p \leq C_p \|S(f)\|_p, \quad 1 < p < \infty. \tag{1.4}$$

(See Theorem 9 of [1] and Theorem 3.4 of Doob [5], Chapter VII.) We obtain new information about this inequality in two directions. For a special class of martingales, our extrapolation method allows us to extend this inequality to the range $0 < p < \infty$. In particular, this extension implies (1.3) and (1.3′). In a second direction, the operator $S: f \to S(f)$ may be replaced by other operators. An interesting class of such operators, which we call operators of matrix type, is defined as follows. Let (a_{jk}) be a matrix of real numbers such that

$$c \leq \sum_{j=1}^{\infty} a_{jk}^2 \leq C, \quad k \geq 1,$$

where c and C are positive real numbers. Define the operator M by

$$Mf = [\sum_{j=1}^{\infty} (\limsup_{n \to \infty} |\sum_{k=1}^{n} a_{jk} d_k|)^2]^{\frac{1}{2}}.$$

Clearly, S is an operator of matrix type with (a_{jk}) the identity matrix. Another example of an operator of matrix type is the "Littlewood–Paley" operator

$$L(f) = [f_1^2 + \sum_{k=1}^{\infty} (f_k - \sigma_k)^2/k]^{\frac{1}{2}},$$

where $\sigma_k = \sum_{j=1}^{k} f_j/k$. This operator has been studied and used in connection with martingales by T. Tsuchikura [16] and E. M. Stein [15]. Let n be a positive integer and f^n the martingale f stopped at $n: f^n = (f_1, \ldots, f_{n-1}, f_n, f_n, \ldots)$. Define the maximal operator M^* by

$$M^* f = \sup_{1 \leqslant n < \infty} M f^n.$$

Notice, for example, that $S^* = S$; also $f \to f^*$ is the maximal operator associated with $f \to \lim \sup_n |f_n|$, which is another example of an operator of matrix type. We show that for any operator M of matrix type,

$$c_p \|M^* f\|_p \leqslant \|f^*\|_p \leqslant C_p \|M^* f\|_p, \quad 1 < p < \infty,$$

for all martingales f. For martingales in a special class, our method allows us to extend this inequality to the entire range $0 < p < \infty$.

We also obtain similar inequalities for more general operators. An interesting example of an operator that is not of matrix type is

$$s(f) = [\sum_{k=1}^{\infty} E(d_k^2 \,|\, \mathcal{A}_{k-1})]^{\frac{1}{2}}.$$

This operator is useful in the study of random walk since it often happens that $s(f^\tau) = \tau^{\frac{1}{2}}$, where τ is a stopping time and f^τ is the random walk f stopped at τ.

The L_p-norm inequalities described in the above examples are special cases of more general integral inequalities. Inequality (1.3′), for example, is a consequence of the inequality

$$c \int_\Omega \Phi(\tau^{\frac{1}{2}}) \leqslant \int_\Omega \Phi[(X^\tau)^*] \leqslant C \int_\Omega \Phi(\tau^{\frac{1}{2}}).$$

Here Φ is any nondecreasing absolutely continuous function satisfying a growth condition; the choice of c and C depend only on the rate of growth of Φ.

Finally, the assumptions of most of our theorems cannot be substantially weakened. This is supported by a number of remarks and examples, some of which are contained in Section 8.

2. Preliminaries

Notation. Let $(\Omega, \mathcal{A}, P)$ be a probability space. If $\mathcal{B}$ is a sub-σ-field of $\mathcal{A}$ and f is an integrable or nonnegative $\mathcal{A}$-measurable function, recall that $E(f\,|\,\mathcal{B})$, the conditional expectation of f given $\mathcal{B}$, is any $\mathcal{B}$-measurable function g satisfying

$$\int_B f = \int_B g, \quad B \in \mathcal{B}.$$

Such a function g always exists and is unique up to a set of measure zero. We usually do not distinguish between functions equal almost everywhere.

Let $\mathcal{A}_0, \mathcal{A}_1, \ldots$ be a nondecreasing sequence of sub-σ-fields of $\mathcal{A}$, $f = (f_1, f_2, \ldots)$ a sequence of real functions on Ω, and $d = (d_1, d_2, \ldots)$ the difference sequence of f so that

$$f_n = \sum_{k=1}^{n} d_k, \quad n \geq 1.$$

Recall that f is a *martingale* (relative to $\mathcal{A}_1, \mathcal{A}_2, \ldots$) if d_k is $\mathcal{A}_k$-measurable and integrable, $k \geq 1$, and

$$E(d_k\,|\,\mathcal{A}_{k-1}) = 0, \quad k \geq 2.$$

The sequence f is a *martingale transform* (relative to $\mathcal{A}_0, \mathcal{A}_1, \ldots$) if

$$f_n = \sum_{k=1}^{n} d_k = \sum_{k=1}^{n} v_k x_k, \quad n \geq 1,$$

where v_k is $\mathcal{A}_{k-1}$-measurable, $k \geq 1$, and $x = (x_1, x_2, \ldots)$ is a martingale difference sequence relative to $\mathcal{A}_1, \mathcal{A}_2, \ldots$.

A martingale transform f is also a martingale if each d_k is integrable, in which case,

$$E(d_k\,|\,\mathcal{A}_{k-1}) = v_k E(x_k\,|\,\mathcal{A}_{k-1}) = 0, \quad k \geq 2.$$

A stopping time is a function ν from Ω into $\{0, 1, \ldots, \infty\}$ such that the indicator functions $I(\nu \leq k)$ are $\mathcal{A}_k$-measurable, $k \geq 0$. (If $A \subset \Omega$, $I(A)$ denotes the function on Ω taking the value 1 on A and the value 0 off A.) The martingale transform f *stopped at* ν, denoted by $f^\nu = (f_1^\nu, f_2^\nu, \ldots)$, is defined by

$$f_n^\nu = \sum_{k=1}^{n} I(\nu \geq k)\, d_k, \quad n \geq 1.$$

The martingale transform f *started at* μ, where μ is a stopping time, is denoted by ${}^\mu f$ where

$${}^\mu f_n = \sum_{k=1}^{n} I(\mu < k)\, d_k, \quad n \geq 1.$$

Finally, f *started at* μ *and stopped at* ν is written as ${}^\mu f^\nu$,

$$^{\mu}f_n^{\nu} = \sum_{k=1}^{n} I(\mu < k \leqslant \nu)\, d_k, \quad n \geqslant 1.$$

Notice that $I(\nu \geqslant k)$ is $\mathcal{A}_{k-1}$-measurable so that f^{ν} is also a martingale transform. The same is true of $^{\mu}f$ and $^{\mu}f^{\nu}$. The following relations are easily verified: $f = f^{\nu} + {}^{\nu}f$, $(f^{\mu})^{\nu} = f^{\mu\wedge\nu}$, $^{\mu}({}^{\nu}f) = {}^{\mu\vee\nu}f$ and $f^{\mu} - f^{\nu} = {}^{\nu}f^{\mu} - {}^{\mu}f^{\nu}$, where $\vee$ and $\wedge$ denote the usual max and min operations. If $\{\mu = n\} = \Omega$ for some n, $0 \leqslant n \leqslant \infty$, we write $^{n}f^{\nu}$ for $^{\mu}f^{\nu}$; then $^{0}f^{\nu} = f^{\nu}$ and $f^{\infty} = f$.

Throughout the paper $\mathcal{N}$ denotes the collection of all martingale transforms relative to $\mathcal{A}_0, \mathcal{A}_1, \ldots$. Let $x = (x_1, x_2, \ldots)$ be a fixed martingale difference sequence relative to $\mathcal{A}_1, \mathcal{A}_2 \ldots$; we denote by $\mathcal{M}$ the subcollection of $\mathcal{N}$ consisting of all martingale transforms of x.

Our principal aim is to study certain operators T defined on $\mathcal{M}$ or $\mathcal{N}$ with values in the set of nonnegative $\mathcal{A}$-measurable functions on Ω. Three important examples of such operators are

$$f^* = \sup_{1 \leqslant n < \infty} |f_n|,$$

$$S(f) = \Big(\sum_{k=1}^{\infty} d_k^2\Big)^{\frac{1}{2}},$$

$$s(f) = \Big[\sum_{k=1}^{\infty} E(d_k^2 \mid \mathcal{A}_{k-1})\Big]^{\frac{1}{2}}.$$

We adopt the following notation:

$$T_n f = T f^n, \quad 1 \leqslant n \leqslant \infty,$$

$$T^* f = \sup_{1 \leqslant n < \infty} T_n f,$$

$$T^{**} f = \sup_{1 \leqslant n \leqslant \infty} T_n f = T^* f \vee T f.$$

In some cases, $T = T^* = T^{**}$; for example, S and s have this property. However, $Tf = \limsup_{n\to\infty} |f_n|$ does not since it can happen that $Tf < f^* = T^* f$.

We use the notation

$$\|f_n\|_p = \left[\int_{\Omega} |f_n|^p\right]^{1/p}, \quad 0 < p < \infty,$$

even if the integral is infinite. Also, it is convenient to let

$$\|f\|_p = \sup_{1 \leqslant n < \infty} \|f_n\|_p.$$

If $\|f\|_p$ is finite, then the sequence f *is* L_p *bounded.*

The letter c, with or without subscripts, denotes a positive real number, not necessarily the same from line to line. The letter C is also used for the same purpose.

Assumptions. In this section are collected the conditions we sometimes impose on martingale transforms and operators. Recall that $\mathcal{M}$ is the set of all transforms of the martingale difference sequence $x=(x_1, x_2, \ldots)$. This set is closed under addition: if f and g belong to $\mathcal{M}$, then $f \pm g=(f_1 \pm g_1, f_2 \pm g_2, \ldots)$ also belong to $\mathcal{M}$. Moreover, $\mathcal{M}$ has the even more important property of being closed under starting and stopping; that is, if μ and ν are stopping times and f belongs to $\mathcal{M}$, then ${}^{\mu}f^{\nu}$ also belongs to $\mathcal{M}$.

Let $0<\delta\leqslant 1$ and $\varrho\geqslant 2$. We say that condition A holds if, for all $k\geqslant 1$,

A1. $E(|x_k|\,|\mathcal{A}_{k-1})\geqslant\delta$,

A2. $E(x_k{}^2\,|\mathcal{A}_{k-1})=1$,

A3. $E(|x_k|^{\varrho}\,|\mathcal{A}_{k-1})\leqslant c$.

Note that these conditions are redundant in some cases. If A2 holds and $\varrho=2$, A3 imposes no extra restriction and c_{A3}, the c in A3, may be taken to be 1. If A2 and A3 hold with $\varrho>2$, then A1 holds, which follows from Hölder's inequality. For further discussion of these conditions, see Section 8.

Now consider an operator T from $\mathcal{M}$ (or $\mathcal{N}$) into the nonnegative $\mathcal{A}$-measurable functions. The operator T satisfies condition B if, for $\gamma\geqslant 1$,

B1. T is quasi-linear: $T(f+g)\leqslant\gamma(Tf+Tg)$;

B2. T is local: $Tf=0$ on the set $\{s(f)=0\}$;

B3. T is symmetric: $T(-f)=Tf$.

Nonnegativity and symmetry are not essential: if T does not satisfy these conditions, it can be replaced, without loss of generality for our results, by $\tilde{T}f=|Tf|\vee|T(-f)|$.

Note that B1 and B3 imply that $T(f-g)\leqslant\gamma(Tf+Tg)$. Also, if T satisfies condition B, then so do T^* and T^{**}.

There is another local condition that is sometimes satisfied: $Tf=0$ on the set where $f=0$. This is more restrictive than B2 since $f=0$ on the set where $s(f)=0$ but not always the other way around: note that $d_k^2=0$ almost everywhere on the set $A=\{E(d_k^2\,|\mathcal{A}_{k-1})=0\}$ since

$$\int_A d_k^2=\int_A E(d_k^2\,|\mathcal{A}_{k-1})=0.$$

The operators $f\to f^*$, S, and s are sublinear $(\gamma=1)$ and satisfy condition B.

Let $0<p_1<p_2\leqslant\varrho$, where ϱ is the same as in A3. The operator T satisfies condition R if, for all $\lambda>0$ and $f\in\mathcal{M}$, R1. $\lambda^{p_1} P(f^* > \lambda) \leq c \|Tf\|_{p_1}^{p_1}$;

R2. $\lambda^{p_2} P(Tf > \lambda) \leq c \|f^*\|_{p_2}^{p_2}$.

Let $0 < \pi_1 < \varrho$. The operator T satisfies condition L if, for all $\lambda > 0$ and $f \in \mathfrak{M}$,

L1. $\lambda^{\pi_1} P(Tf > \lambda) \leq c \|f^*\|_{\pi_1}^{\pi_1}$;

L2. Condition R holds;

L3. $T_n f$ is $\mathcal{A}_n$-measurable, $n \geq 1$.(1)

Preliminary lemmas. Here we collect some inequalities, remarks, and lemmas.

If f is a martingale, and $\lambda > 0$, then

$$\lambda^p P(f^* > \lambda) \leq \|f\|_p^p, \quad 1 \leq p < \infty; \tag{2.1}$$

$$\|f\|_p \leq \|f^*\|_p \leq q \|f\|_p, \qquad p^{-1} + q^{-1} = 1, \quad 1 < p < \infty; \tag{2.2}$$

$$\lambda P(S(f) > \lambda) \leq c \|f\|_1; \tag{2.3}$$

$$\lambda P(f^* > \lambda) \leq c \|S(f)\|_1; \tag{2.4}$$

$$c_p \|S(f)\|_p \leq \|f\|_p \leq C_p \|S(f)\|_p, \quad 1 < p < \infty. \tag{2.5}$$

For (2.1) and (2.2), see Doob [5]; and for (2.3), (2.4), and (2.5), see [1].

By these inequalities, the operator S satisfies condition R with $p_1 = 1$ and $p_2 = 2$, and condition L with $\pi_1 = 1$.

Suppose T satisfies condition B. If μ and ν are stopping times, then

$$I(\mu < k \leq \nu) \leq I(\mu < \nu), \quad k \geq 1,$$

implying that

$$s({}^{\mu}f^{\nu}) \leq I(\mu < \nu) s(f)$$

(as usual, $0 \cdot \infty = 0$). Therefore, by the local condition B2,

$$T({}^{\mu}f^{\nu}) = 0 \quad \text{on } \{\mu \geq \nu\}. \tag{2.6}$$

In turn, (2.6) implies that

$$T(f^{\mu} - f^{\nu}) = 0 \quad \text{on } \{\mu = \nu\} \tag{2.7}$$

since

$$T(f^{\mu} - f^{\nu}) = T({}^{\nu}f^{\mu} - {}^{\mu}f^{\nu}) \leq \gamma [T({}^{\nu}f^{\mu}) + T({}^{\mu}f^{\nu})].$$

Recall that $T_n f = T f^n$, $1 \leq n \leq \infty$. If T is local, we extend this definition in a consistent way for $n = 0$ by setting $T_0 f = 0$, since $s(f^0) = 0$. Now define $T_\tau f$ for any stopping time τ as follows:

$$T_\tau f = T_n f \text{ on } \{\tau = n\}, \quad 0 \leq n \leq \infty.$$

It can happen that $T_\tau f \neq T f^\tau$; however, we do have the following double inequality.

(1) The L1 part of condition L is a temporary assumption only; in Remark 8.3, we show that it is not needed to obtain the results of this paper and can be eliminated.

LEMMA 2.1. *Let T be an operator satisfying condition* B. *If τ is a stopping time, then*

$$\gamma^{-1}T_\tau f \leqslant Tf^\tau \leqslant \gamma T_\tau f.$$

Proof. Let $0 \leqslant n \leqslant \infty$; on the set $\{\tau = n\}$,

$$Tf^\tau = T[f^n + (f^\tau - f^n)] \leqslant \gamma[Tf^n + T(f^\tau - f^n)] = \gamma T_n f = \gamma T_\tau f.$$

Here we have used B1 and (2.7). The proof of the left-hand side is omitted since it is similar.

Remark 2.1. It can happen that T does not satisfy R1 but T^{**} does. (Consider $Tf = \limsup_n |f_n|$.) In such a case there would be no loss in replacing T by T^{**} provided R2 is satisfied by T^{**}. In this connection, the following fact is useful: If T satisfies condition B and the measurability condition L3, then T^{**} satisfies R2 whenever T does. To see this, let $\tau(\omega) = \inf\{1 \leqslant n \leqslant \infty : (T_n f)(\omega) > \lambda\}$, $\omega \in \Omega$, where $\inf \varnothing = \infty$. Then τ is a stopping time by L3, and

$$\lambda^{p_2} P(T^{**}f > \lambda) = \lambda^{p_2} P(T_\tau f > \lambda) \leqslant \lambda^{p_2} P(\gamma T f^\tau > \lambda) \leqslant c\|(f^\tau)^*\|_{p_2}^{p_2} \leqslant c\|f^*\|_{p_2}^{p_2}.$$

LEMMA 2.2. *Let $0 < p \leqslant 2$. If f is a martingale transform in $\mathcal{N}$, then*

$$\|f^*\|_p \leqslant c_p \|s(f)\|_p.$$

The choice of c_p depends only on p.

See Section 5 for other results about $s(f)$. This one can be obtained directly and is needed in case the number p_2 in R2 is less than 2.

Proof. We may assume that $\|s(f)\|_p < \infty$. If $p = 2$, by (2.2) and the orthogonality of the difference sequence d of f, we have that $\|f^*\|_2 \leqslant 2\|f\|_2 = 2\|S(f)\|_2 = 2\|s(f)\|_2$. Now let $0 < p < 2$. Since

$$s_n(f) = s(f^n) = [\sum_{k=1}^{n} E(d_k^2 \,|\, \mathcal{A}_{k-1})]^{\frac{1}{2}}$$

is $\mathcal{A}_{n-1}$-measurable,

$$\tau = \inf\{0 \leqslant n < \infty : s_{n+1}(f) > \lambda\}$$

is a stopping time. Let $g = f^\tau$. Then $s(g) = s_\tau(f) \leqslant \lambda$, $s(g) \leqslant s(f)$, and

$$P(f^* > \lambda) \leqslant P(s(f) > \lambda) + P(f^* > \lambda,\, s(f) \leqslant \lambda).$$

Since $g = f$ on $\{\tau = \infty\} = \{s(f) \leqslant \lambda\}$, the last probability is equal to

$$P(g^* > \lambda,\, s(f) \leqslant \lambda) \leqslant P(g^* > \lambda) \leqslant \lambda^{-2}\|g\|_2^2 = \lambda^{-2}\int_{\{s(f)>\lambda\}} s(g)^2$$

$$+\lambda^{-2}\int_{\{s(f)\leqslant\lambda\}} s(g)^2 \leqslant P(s(f) > \lambda) + \lambda^{-2}\int_{\{s(f)\leqslant\lambda\}} s(f)^2.$$

Therefore,
$$\|f^*\|_p^p = p\int_0^\infty \lambda^{p-1}P(f^*>\lambda)\,d\lambda \leqslant 2\|s(f)\|_p^p + p\int_\Omega s(f)^2\int_{s(f)}^\infty \lambda^{p-3}\,d\lambda$$
$$= \|s(f)\|_p^p\,[2+p/(2-p)] = c_p^p\,\|s(f)\|_p^p.$$

In Section 5, we improve this upper bound by showing that c_p^p may be chosen to be bounded on the interval $0<p\leqslant 2$.

Remark 2.2. If T is an operator on $\mathcal{N}$ satisfying condition B such that, for all $f\in\mathcal{N}$,
$$\|Tf\|_2 \leqslant c\|f\|_2,$$
then
$$\|Tf\|_p \leqslant c_p\|s(f)\|_p, \quad 0<p\leqslant 2,$$
by an argument similar to the proof of Lemma 2.2.

The underlying idea of the next lemma is well known. See Zygmund [17; Chapter V, 8. 26].

LEMMA 2.3. *Suppose that g is a nonnegative $\mathcal{A}$-measurable function, α, β, p_1, p_2 are positive real numbers with $p_1<p_2$, $A\in\mathcal{A}$,*
$$\int_A g^{p_1} \geqslant \alpha^{p_1}P(A),$$
and, for all $\eta>0$,
$$\eta^{p_2}P(g>\eta, A) \leqslant \beta^{p_2}P(A).$$
Then,
$$P(g>\theta\alpha, A) \geqslant \left[(1-\theta^{p_1})\,\frac{p_2-p_1}{p_2}\left(\frac{\alpha}{\beta}\right)^{p_1}\right]^{\frac{p_2}{p_2-p_1}} P(A), \quad 0<\theta<1.$$

Proof. It is sufficient to prove this for $A=\Omega$, $p_1=1$, and $p_2=p>1$.

For any positive real number B,
$$\alpha \leqslant \int_\Omega g = \int_0^\infty P(g>\eta)\,d\eta \leqslant \int_0^{\theta\alpha} d\eta + \int_0^{B\alpha} P(g>\theta\alpha)\,d\eta + \int_{B\alpha}^\infty \beta^p\eta^{-p}\,d\eta$$
$$\leqslant \theta\alpha + B\alpha P(g>\theta\alpha) + \frac{1}{p-1}\left(\frac{\beta}{\alpha}\right)^p \frac{\alpha}{B^{p-1}}.$$

Let
$$B = \left[\frac{1}{1-\theta}\cdot\frac{p}{p-1}\left(\frac{\beta}{\alpha}\right)^p\right]^{\frac{1}{p-1}}.$$

Then
$$P(g>\theta\alpha) \geqslant \frac{1}{B}\left[1-\theta-\frac{1}{p-1}\left(\frac{\beta}{\alpha}\right)^p\frac{1}{B^{p-1}}\right] = \left[(1-\theta)\,\frac{p-1}{p}\,\frac{\alpha}{\beta}\right]^{\frac{p}{p-1}}.$$

The next lemma shows how the variability of f^* and, more generally, Tf is controlled by the variability of $s(f)$.

LEMMA 2.4. *Suppose condition* A *holds. If* $f \in \mathcal{M}$, $\lambda > 0$, *and* $A \in \mathcal{A}_m$ *for some* $m \geqslant 0$ *such that*

$$A \subset \{s_m(f) = 0,\ \lambda \leqslant s(f) \leqslant 2\lambda\}, \tag{2.8}$$

then

$$P(f^* > c\lambda,\ A) \geqslant cP(A), \tag{2.9}$$

$$\int_A |f^*|^p \leqslant c\lambda^p P(A), \quad 0 < p \leqslant \varrho. \tag{2.10}$$

If T *is any operator satisfying conditions* B *and* R, *then*

$$P(Tf > c\lambda,\ A) \geqslant cP(A). \tag{2.11}$$

The choice of $c_{(2.9)}$ *depends only on* δ; *that of* $c_{(2.10)}$ *only on* p *and* c_{A3}; *and that of* $c_{(2.11)}$ *only on the parameters of* A, B, *and* R.

Proof. We may assume in the proof that

$$s(f) = 0 \text{ off } A. \tag{2.12}$$

For consider $g = {}^m f^\nu$ with ν the stopping time defined by

$$\nu = m \text{ off } A, \quad \nu = \infty \text{ on } A.$$

Then

$$s(g) = 0 \text{ off } A, \quad g = f \text{ on } A,$$

so that g satisfies not only (2.8) but also (2.12). If g satisfies the conclusions of the lemma, then so does f. This is clear for the first two, (2.9) and (2.10). For the third, note that $\nu = \infty$ and $s(f^m) = 0$ on the set $\{Tg > C\lambda,\ A\}$, so that

$$C\lambda < Tg = T[f^\nu - f^m] \leqslant \gamma[Tf^\nu + Tf^m] = \gamma Tf^\nu \leqslant \gamma^2 T_\nu f = \gamma^2 Tf.$$

Therefore, letting $c = \gamma^{-2} C$, we have that

$$P(Tf > c\lambda,\ A) \geqslant P(Tg > C\lambda,\ A)$$

and we may assume (2.12).

As usual, write

$$f_n = \sum_{k=1}^n d_k = \sum_{k=1}^n v_k x_k, \quad n \geqslant 1,$$

and note that, under condition A 2,

$$s(f) = \Big(\sum_{k=1}^{\infty} v_k^2\Big)^{\frac{1}{2}}$$

so that

$$s(f)\,S(f) \geq \sum_{k=1}^{\infty} |v_k d_k| = \sum_{k=1}^{\infty} v_k^2\,|x_k|.$$

Therefore, by A1, A2, and (2.8),

$$2\lambda \int_A S(f) \geq \int_A \sum_{k=m+1}^{\infty} v_k^2\,|x_k| \geq \delta \int_A \sum_{k=m+1}^{\infty} v_k^2 = \delta \int_A s(f)^2 \geq \delta\lambda^2 P(A).$$

Also, we have that

$$\int_A S(f)^2 = \int_A \sum_{k=m+1}^{\infty} v_k^2 x_k^2 = \int_A \sum_{k=m+1}^{\infty} v_k^2 = \int_A s(f)^2 \leq 4\lambda^2 P(A).$$

The conditions of Lemma 2.3 are satisfied by $g = S(f)$, $\alpha = 2^{-1}\delta\lambda$, $\beta = 2\lambda$, $p_1 = 1$, $p_2 = 2$, $\theta = 2^{-1}$, so that

$$P(S(f) > 2^{-2}\delta\lambda, A) \geq 2^{-8}\delta^2 P(A).$$

Therefore, by (2.12), (2.3), and (2.2),

$$\int_A f^* = \|f^*\|_1 \geq \|f\|_1 \geq c\lambda\, P(S(f) > c\lambda) \geq c\lambda\, P(A),$$

$$\int_A |f^*|^2 \leq \|f^*\|_2^2 \leq 4\,\|f\|_2^2 = 4\,\|s(f)\|_2^2 = 4\int_A s(f)^2 \leq 16\,\lambda^2 P(A).$$

Another application of Lemma 2.3 now gives (2.9).

If $2 \leq p \leq \varrho$, then by (2.5), (2.12), and the inequality $E(|x_k|^p \mid \mathcal{A}_{k-1}) \leq c$, which follows from A3, we have that

$$\int_A |f^*|^p \leq \|f^*\|_p^p \leq c\,\|S(f)\|_p^p = c\int_A \left(\sum_{k=1}^{\infty} v_k^2 x_k^2\right)^{p/2} \leq c\int_A s(f)^{p-2} \sum_{k=1}^{\infty} v_k^2\,|x_k|^p$$

$$\leq c\lambda^{p-2} \sum_{k=m+1}^{\infty} \int_A v_k^2 \leq c\lambda^{p-2} \int_A s(f)^2 \leq c\lambda^p P(A).$$

If $0 < p \leq 2$, then by Lemma 2.2,

$$\int_A |f^*|^p \leq \|f^*\|_p^p \leq c\,\|s(f)\|_p^p = c\int_A s(f)^p \leq c\lambda^p P(A).$$

This proves (2.10).

Using (2.12), B2, R1, and (2.9), we have that

$$\int_A |Tf|^{p_1} = \|Tf\|_{p_1}^{p_1} \geqslant c\lambda^{p_1} P(f^* > c\lambda) \geqslant c\lambda^{p_1} P(A).$$

By R2, (2.12), and (2.10), for $\eta > 0$,

$$\eta^{p_2} P(Tf > \eta) \leqslant c \|f^*\|_{p_2}^{p_2} = c\int_A |f^*|^{p_2} \leqslant c\lambda^{p_2} P(A).$$

Therefore, (2.11) follows after another application of Lemma 2.3. This completes the proof.

The following two theorems provide upper bounds for stopped martingales.

THEOREM 2.1. *Suppose that condition* A *holds and* $0 < p \leqslant \varrho$. *If* $f \in \mathcal{M}$, $f_n = \sum_{k=1}^n v_k x_k$, $n \geqslant 1$, $v = (v_1, v_2, \ldots)$ *is uniformly bounded by a positive real number* b, *and* τ *is the stopping time defined by*

$$\tau = \inf\{n\colon |f_n| > b\},$$

then

$$\|(f^\tau)^*\|_p \leqslant cb[P(s(f) > 0)]^{1/p} \leqslant cb.$$

The choice of c *depends only on* p *and the parameters of* A.

Condition A cannot be substantially weakened. See Example 8.3.

THEOREM 2.2. *Suppose that condition* A *holds and* $0 < p \leqslant \varrho$. *Let* T *be an operator satisfying conditions* B, R, *and* L3. *If* $f \in \mathcal{M}$, $f_n = \sum_{k=1}^n v_k x_k$, $n \geqslant 1$, $v = (v_1, v_2, \ldots)$ *is uniformly bounded by a positive real number* b, *and* τ *is the stopping time defined by*

$$\tau = \inf\{n\colon T_n f > b\},$$

then

$$\|(f^\tau)^*\|_p \leqslant cb[P(s(f) > 0)]^{1/p} \leqslant cb.$$

The choice of c *depends only on* p *and the parameters of* A, B, *and* R.

Theorem 2.1 follows immediately from Theorem 2.2: let T be the operator defined by $Tf = f^*$. In this case $T_n f = (f^n)^*$ and B, R, and L3 are satisfied with $\gamma = 1$, $p_1 = 1$, $p_2 = 2$, $c_{R1} = c_{R2} = 1$.

Proof of Theorem 2.2. Let N be a positive integer. We must show that

$$\|(f^{\tau \wedge N})^*\|_p \leqslant cb\,[P(s(f) > 0)]^{1/p}$$

with c not depending on N.

Let $\lambda = 2\gamma^2 b/\beta$ where $\beta = c_{(2.11)} < 1$. Note that $v^* \leqslant b < \lambda$. Define the multiplier sequence w:

$$w_k = v_k, \quad 1 \leqslant k \leqslant N, \qquad w_k = \lambda I(s_k(f) > 0), \quad k > N;$$

let g be the corresponding transform: $g_n = \sum_{k=1}^n w_k x_k$, $n \geqslant 1$. Then $g \in \mathcal{M}$, $f^n = g^n$, $1 \leqslant n \leqslant N$, and on the set $\{s(f) > 0\}$,

$$s_n(g) = (\sum_{k=1}^n w_k^2)^{\frac{1}{2}} \to \infty$$

as $n \to \infty$. Define a sequence of stopping times as follows. Let

$$\mu_0 = \inf \{n \geqslant 0 \colon s_{n+1}(f) > 0\}.$$

This is a stopping time. Note that $\{\mu_0 < \infty\} = \{s(f) > 0\}$ and $s(g^{\mu_0}) = 0$. If $j \geqslant 1$ and μ_{j-1} is a stopping time, let

$$\mu_j = \inf \{n \colon s({}^m g^n) > \lambda\}$$

on the set where $\mu_{j-1} = m$, $m \geqslant 0$, and let $\mu_j = \infty$ on the set where $\mu_{j-1} = \infty$. Then μ_j is a stopping time satisfying

$$\mu_{j-1} < \mu_j < \infty \text{ on } \{s(f) > 0\},$$
$$\mu_j = \infty \text{ on } \{s(f) = 0\}.$$

Let $h_j = (h_{j1}, h_{j2}, \ldots)$ denote the martingale transform ${}^{\mu_{j-1}}g^{\mu_j}$, $j \geqslant 1$. Since $w^* \leqslant \lambda$ and $s({}^m g^n) \leqslant s({}^m g^{n-1}) + |w_n|$, we have that

$$\lambda \leqslant s(h_j) \leqslant 2\lambda \quad \text{on} \quad \{s(f) > 0\},$$

$j \geqslant 1$. Also, $h_j^* = 0$ on $\{s(f) = 0\}$.

Now let
$$\sigma = \inf \{j \colon T^*_{\mu_j} g > b\},$$

$\mu_\infty = \infty$, and $\nu = \mu_\sigma$. Note that $\sigma \geqslant 1$: by Lemma 2.1 and B2, $T^*_{\mu_0} g \leqslant \gamma T^* g^{\mu_0} = 0$ since $s(g^{\mu_0}) = 0$. Also, $\tau \wedge N \leqslant \nu$: if $n \leqslant N$, then on the set $\{\nu = n\}$, we have that $b < T^*_n g = T^* g^n = T^* f^n = T^*_n f$ and $\tau \leqslant n$. Therefore

$$(f^{\tau \wedge N})^* = (g^{\tau \wedge N})^* \leqslant (g^\nu)^* \leqslant \sum_{j=1}^\infty I(\sigma \geqslant j)\, h_j^*,$$

since on $\{\sigma = j\}$,

$$(g^{\mu_j})^* = (g^{\mu_0} + {}^{\mu_0}g^{\mu_1} + \ldots + {}^{\mu_{j-1}}g^{\mu_j})^* = (0 + h_1 + \ldots + h_j)^* \leqslant h_1^* + \ldots + h_j^*, \qquad j \geqslant 1.$$

Consequently,
$$\|(f^{\tau \wedge N})^*\|_p^{p \wedge 1} \leqslant \sum_{j=1}^\infty \|I(\sigma \geqslant j)\, h_j^*\|_p^{p \wedge 1}$$

for $0 < p < \infty$. From now on assume that $0 < p \leqslant \varrho$.

Fix $j \geq 1$ and let

$$A_m = \{\mu_{j-1} = m,\ T_m^* g \leq b\}, \quad m \geq 0.$$

Clearly,
$$\{\sigma \geq j, s(f) > 0\} = \{T^*_{\mu_{j-1}} g \leq b, s(f) > 0\} = \bigcup_{m=0}^{\infty} A_m.$$

Note that
$$A_m \subset \{s_m(h_j) = 0, \lambda \leq s(h_j) \leq 2\lambda\},$$

and $A_m \in \mathcal{A}_m$ by L3. Applying Lemma 2.4 to h_j, we have that

$$\|I(\sigma \geq j)\, h_j^*\|_p^p = \|I(\sigma \geq j, s(f) > 0)\, h_j^*\|_p^p = \sum_{m=0}^{\infty} \int_{A_m} |h_j^*|^p$$
$$\leq c\lambda^p \sum_{m=0}^{\infty} P(A_m) = c\lambda^p P(\sigma \geq j, s(f) > 0).$$

On the set where $T^*_{\mu_j} g \leq b$ and $\mu_j < \infty$,

$$Th_j = T[g^{\mu_j} - g^{\mu_{j-1}}] \leq \gamma[Tg^{\mu_j} + Tg^{\mu_{j-1}}] \leq \gamma^2[T_{\mu_j} g + T_{\mu_{j-1}} g) \leq 2\gamma^2 T^*_{\mu_j} g \leq 2\gamma^2 b = \beta\lambda.$$

Therefore, applying Lemma 2.4 to h_j a second time, we obtain

$$P(\sigma \geq j+1, s(f) > 0) = P(T^*_{\mu_j} g \leq b, s(f) > 0) = \sum_{m=0}^{\infty} P(\mu_{j-1} = m, T^*_{\mu_j} g \leq b)$$
$$\leq \sum_{m=0}^{\infty} P(Th_j \leq \beta\lambda, A_m) \leq (1-\beta) \sum_{m=0}^{\infty} P(A_m) = (1-\beta)\, P(\sigma \geq j, s(f) > 0).$$

By induction, for all $j \geq 1$,

$$P(\sigma \geq j, s(f) > 0) \leq (1-\beta)^{j-1} P(\sigma \geq 1, s(f) > 0) = (1-\beta)^{j-1} P(s(f) > 0).$$

Accordingly, for $0 < p \leq 1$,

$$\|(f^{\tau \wedge N})^*\|_p^p \leq \sum_{j=1}^{\infty} \|I(\sigma \geq j)\, h_j^*\|_p^p \leq \sum_{j=1}^{\infty} c\lambda^p (1-\beta)^{j-1} P(s(f) > 0) = c\lambda^p P(s(f) > 0),$$

and for $1 \leq p \leq \varrho$,

$$\|(f^{\tau \wedge N})^*\|_p \leq \sum_{j=1}^{\infty} \|I(\sigma \geq j)\, h_j^*\|_p \leq \sum_{j=1}^{\infty} [c\lambda^p (1-\beta)^{j-1} P(s(f) > 0)]^{1/p} = c\lambda [P(s(f) > 0)]^{1/p}.$$

This completes the proof of Theorem 2.2.

LEMMA 2.5. *Suppose that conditions* A1 *and* A2 *hold. Then for all* $\lambda > 0$ *and* $f \in \mathcal{M}$,

$$P(v^* > \lambda) \leq cP(cd^* > \lambda)$$

with the choice of c *depending only on* δ.

Recall that d is the difference sequence of f and $d_k = v_k x_k$, $k \geq 1$; as usual v^* and d^* are the maximal functions of the sequences v and d, respectively.

Proof. Let $\tau = \inf \{k\colon |v_k| > \lambda\}$ and $A_k = \{\tau = k\}$, $k \geq 1$. *Then* $A_k \in \mathcal{A}_{k-1}$ and, by A1,

$$\int_{A_k} |x_k| \geq \delta P(A_k);$$

by A2,

$$\int_{A_k} x_k^2 = P(A_k);$$

hence by Lemma 2.3,

$$P(|x_k| > c, A_k) \geq cP(A_k).$$

Therefore,

$$P(v^* > \lambda) = \sum_{k=1}^{\infty} P(A_k) \leq c \sum_{k=1}^{\infty} P(|x_k| > c, A_k)$$

$$\leq c \sum_{k=1}^{\infty} P(|d_k| > c\lambda, A_k) \leq cP(cd^* > \lambda).$$

LEMMA 2.6. *Let ψ be a nonnegative measurable function on the real line satisfying*

$$\int_{-\infty}^{a} \psi(t)\, dt < \infty$$

for some real number a. If

$$B = \{t : \psi(t) < \alpha\psi(t+1)\}$$

for a real number $\alpha > 1$, then

$$\int_{-\infty}^{\infty} \psi(t)\, dt \leq \frac{\alpha}{\alpha - 1} \int_B \psi(t+1)\, dt.$$

Proof. For each real λ, let

$$A_\lambda = \{t \leq \lambda : \psi(t) \geq \alpha\psi(t+1)\},$$
$$B_\lambda = \{t \leq \lambda : \psi(t) < \alpha\psi(t+1)\}.$$

Then

$$\alpha \int_{-\infty}^{\lambda} \psi(t)\, dt \leq \alpha \int_{-\infty}^{\lambda+1} \psi(t)\, dt$$

$$= \alpha \int_{A_\lambda} \psi(t+1)\, dt + \alpha \int_{B_\lambda} \psi(t+1)\, dt \leq \int_{-\infty}^{\lambda} \psi(t)\, dt + \alpha \int_B \psi(t+1)\, dt. \qquad (2.13)$$

If $\int_{-\infty}^{\lambda} \psi(t)\, dt < \infty$, then

$$\int_{-\infty}^{\lambda} \psi(t)\, dt \leq \frac{\alpha}{\alpha - 1} \int_B \psi(t+1)\, dt.$$

Therefore, the desired result holds if $\int_{-\infty}^{\lambda}\psi(t)\,dt<\infty$ for all real λ. If, on the contrary $\int_{-\infty}^{\lambda}\psi(t)\,dt<\infty$ but $\int_{-\infty}^{\lambda+1}\psi(t)\,dt=\infty$ for some λ (by our assumption, the only other possibility), inequality (2.13) shows that $\int_B\psi(t+1)\,dt=\infty$, so the desired result holds in any case.

3. The right-hand side

In this section, we prove that

$$\|f^*\|_p \leqslant c\|T^*f\|_p, \quad f\in\mathcal{M}, \tag{3.1}$$

under conditions A, B, and R for $0<p<\infty$ as specified in (3.13). In fact, we prove a stronger inequality (Theorem 3.3).

If it were true that $P(f^*>\lambda)\leqslant cP(cT^*f>\lambda)$ for all $\lambda>0$, then (3.1) would follow easily from the formula

$$\|f^*\|_p^p=p\int_0^\infty \lambda^{p-1}P(f^*>\lambda)\,d\lambda.$$

However, even in simple examples, there may be no such inequality between distribution functions over the entire interval $0<\lambda<\infty$: consider $f_n=\sum_{k=1}^n x_k/k$ where $x=(x_1,x_2,\ldots)$ is an independent sequence such that $x_k=\pm 1$ with equal probability. Then $S(f)=(\sum_{k=1}^n 1/k^2)^{\frac{1}{2}}$ and $P(cS(f)>\lambda)=0$ for all large λ. On the other hand, $P(f^*>\lambda)>0$ for all $\lambda>0$. In spite of such examples, it turns out that distribution function inequalities do exist for sufficiently many values of λ to allow us to use integral formulas such as the one above. A substitute for a full strength inequality between distribution functions is provided by the following theorem. This, in conjunction with Lemma 2.6, leads to integral inequalities such as (3.1).

THEOREM 3.1. *Suppose that conditions* A, B, *and* R *hold. Let* $\alpha\geqslant 1$ *and* $\beta>1$. *Then*

$$P(f^*>\lambda)\leqslant cP(cT^*f>\lambda)+cP(cd^*>\lambda) \tag{3.2}$$

for all f *in* $\mathcal{M}$ *and* $\lambda>0$ *satisfying*

$$P(f^*>\lambda)\leqslant \alpha P(f^*>\beta\lambda). \tag{3.3}$$

The choice of c *depends only on* α, β, *and the parameters of* A, B, *and* R. *Furthermore, this choice may be made so that, with* β *and the other parameters fixed, the function* $\alpha\to c$ *is nondecreasing.*

Recall that d is the difference sequence of f. If $d^*\leqslant cT^*f$, as sometimes happens, then (3.2) simplifies to

$$P(f^*>\lambda)\leqslant cP(cT^*f>\lambda).$$

Proof. The last assertion is obvious since if $1 \leqslant \alpha_1 < \alpha_2 < \infty$, then the set of pairs (f, λ) satisfying (3.3) for $\alpha = \alpha_1$ is a subset of the set of pairs (f, λ) satisfying (3.3) for $\alpha = \alpha_2$. Therefore, if c is suitable for $\alpha = \alpha_2$ it is also suitable for $\alpha = \alpha_1$.

Let $\lambda > 0$. We first prove (3.2) for all f in $\mathcal{M}$ satisfying (3.3) *and*

$$v^* \leqslant \lambda. \tag{3.4}$$

Here $d_k = v_k x_k$, $k \geqslant 1$, as usual.

Let $\theta = (\beta - 1)/2$. Then either

$$P(f^* > \lambda) \leqslant 2\alpha P(f^* > \beta\lambda,\ d^* \leqslant \theta\lambda) \tag{3.5}$$

or

$$P(f^* > \lambda) \leqslant 2\alpha P(d^* > \theta\lambda), \tag{3.6}$$

since otherwise, (3.3) would not hold. If (3.6) is satisfied, then (3.2) holds trivially. Therefore, from now on we suppose that (3.5) is satisfied.

Define stopping times μ and ν as follows:

$$\mu = \inf\{n: |f_n| > \lambda\},$$
$$\nu = \inf\{n: |f_n| > \beta\lambda\}.$$

Then $\{\mu < \infty\} = \{f^* > \lambda\}$, $\{\nu < \infty\} = \{f^* > \beta\lambda\}$, and $\mu \leqslant \nu$. Let $g = {}^{\mu}f^{\nu}$. Then $g \in \mathcal{M}$ and

$$P(g^* > \theta\lambda) \geqslant P(f^* > \beta\lambda,\ d^* \leqslant \theta\lambda)$$

since, on the latter set, $\mu \leqslant \nu < \infty$, and

$$g^* \geqslant |f_\nu - f_\mu| \geqslant |f_\nu| - |f_\mu| > \beta\lambda - (\lambda + d^*) \geqslant \beta\lambda - (\lambda + \theta\lambda) = \theta\lambda.$$

Therefore, by (3.5),

$$P(g^* > \theta\lambda) \geqslant cP(f^* > \lambda).$$

We now wish to apply Lemma 2.3 to the function Tg on the set $A = \{f^* > \lambda\}$. To do this, we establish upper and lower estimates as follows. By the local condition B2,

$$\{Tg = 0\} \supset \{s(g) = 0\} \supset \{\mu = \infty\} = \{f^* \leqslant \lambda\}.$$

Therefore, by R1 and the preceding paragraph,

$$\int_{\{f^* > \lambda\}} |Tg|^{p_1} = \|Tg\|_{p_1}^{p_1} \geqslant c(\theta\lambda)^{p_1} P(g^* > \theta\lambda) \geqslant c\lambda^{p_1} P(f^* > \lambda),$$

so that the lower (p_1) estimate holds.

Let $b = 2\beta\lambda$ and $\tau = \inf\{n: |({}^{\mu}f)_n| > b\}$. Then $\tau \geqslant \nu$: since $\tau \geqslant \mu$, we see that $\tau \geqslant \nu$ on the sets $\{\mu = \nu\}$ and $\{\tau = \infty\}$. Also, $\tau \geqslant \nu$ on $\{\mu < \nu,\ \tau < \infty\}$, since

$$|f_\tau| = |f_\mu + ({}^{\mu}f)_\tau| \geqslant |({}^{\mu}f)_\tau| - |f_\mu| > b - \beta\lambda = \beta\lambda.$$

Note that the multiplier sequence defining ${}^{\mu}f$ is uniformly bounded by b, using (3.4) and $\lambda < b$.

Therefore, by Theorem 2.1 with $p=p_2$,

$$\|g^*\|_{p_2}=\|({}^\mu f^\nu)^*\|_{p_2}\leqslant\|({}^\mu f^\tau)^*\|_{p_2}\leqslant cb[P(s({}^\mu f)>0)]^{1/p_2}$$
$$\leqslant cb[P(\mu<\infty)]^{1/p_2}=c\lambda[P(f^*>\lambda)]^{1/p_2}.$$

This leads to the upper (p_2) estimate: for all $\eta>0$, we have, by R 2, that

$$\eta^{p_2}P(Tg>\eta, f^*>\lambda)\leqslant\eta^{p_2}P(Tg>\eta)\leqslant c\|g^*\|_{p_2}^{p_2}\leqslant c\lambda^{p_2}P(f^*>\lambda).$$

Applying Lemma 2.3 to Tg, we obtain

$$P(Tg>c\lambda,\ f^*>\lambda)\geqslant cP(f^*>\lambda).$$

Therefore,
$$P(f^*>\lambda)\leqslant cP(Tg>c\lambda)\leqslant cP(cT^{**}f>\lambda),$$

using
$$Tg=T(f^\nu-f^\mu)\leqslant\gamma[Tf^\nu+Tf^\mu]\leqslant\gamma^2[T_\nu f+T_\mu f]\leqslant 2\gamma^2T^{**}f.$$

In summary, we have shown that for all f and λ such that (3.3) and (3.4) hold, we have the inequality (3.2) with T^{**} in place of T^*. Fix such an f and λ; there is a positive integer N such that for all $n>N$,

$$P((f^n)^*>\lambda)\leqslant 2\alpha P((f^n)^*>\beta\lambda).$$

Note that $T^{**}f^n\leqslant T^*f$. If we now apply what we have already proved to $T^{**}f^n$ with α replaced by 2α, we obtain

$$P((f^n)^*>\lambda)\leqslant cP(cT^{**}f^n>\lambda)+cP(cd^*>\lambda)\leqslant cP(cT^*f>\lambda)+cP(cd^*>\lambda).$$

Finally, since the above inequality holds with c independent of n for $n>N$, we may let $n\to\infty$ to obtain (3.2) under assumption (3.4).

We now eliminate assumption (3.4). Consider any f in $\mathcal{M}$ satisfying (3.3). Let

$$\sigma=\inf\{n\geqslant 0:\ |v_{n+1}|>\lambda\}.$$

Since v_{n+1} is $\mathcal{A}_n$-measurable, σ is a stopping time and $h=f^\sigma$ belongs to $\mathcal{M}$. Note that h satisfies (3.4). Either

$$P(f^*>\lambda)\leqslant 2\alpha P(f^*>\beta\lambda,\ v^*\leqslant\lambda)\tag{3.7}$$

or
$$P(f^*>\lambda)\leqslant 2\alpha P(v^*>\lambda),\tag{3.8}$$

since otherwise, f would not satisfy (3.3). If (3.8) is satisfied, then (3.2) holds by Lemma 2.5.

From now on suppose that (3.7) holds. From this, and the fact that $h^* \leqslant f^*$ with equality on the set $\{v^* \leqslant \lambda\}$, it follows that

$$P(h^* > \lambda) \leqslant P(f^* > \lambda) \leqslant 2\alpha P(f^* > \beta\lambda,\ v^* \leqslant \lambda) = 2\alpha P(h^* > \beta\lambda,\ v^* \leqslant \lambda) \leqslant 2\alpha P(h^* > \beta\lambda). \quad (3.9)$$

So h satisfies (3.3) with α replaced by 2α. Therefore, by what we have already proved,

$$P(h^* > \lambda) \leqslant cP(cT^*h > \lambda) + cP(cd^* > \lambda).$$

Here we have used the fact that the difference sequence of h has a maximal function no greater than d^*. By Lemma 2.1 applied to T^*,

$$T^*h \leqslant \gamma T^*_\sigma f \leqslant \gamma T^* f.$$

Therefore, using part of (3.9), we have that

$$P(f^* > \lambda) \leqslant 2\alpha P(h^* > \beta\lambda) \leqslant 2\alpha P(h^* > \lambda) \leqslant cP(cT^*f > \lambda) + cP(cd^* > \lambda).$$

This completes the proof of Theorem 3.1.

We now turn to integral inequalities. Consider a function Φ on $[0, \infty]$ such that

$$\Phi(b) = \int_0^b \varphi(\lambda)\, d\lambda, \quad 0 \leqslant b \leqslant \infty,$$

for some nonnegative measurable function φ on $(0, \infty)$ satisfying

$$\varphi(2\lambda) \leqslant c\varphi(\lambda), \quad \lambda > 0. \quad (3.10)$$

We also assume that $\Phi(1) < \infty$. (This together with (3.10) implies that $\Phi(b) < \infty$, $0 \leqslant b < \infty$.) For example, if $0 < p < \infty$, $\Phi(b) = b^p$ defines such a function. Also, many Orlicz spaces may be determined by such functions; for example, the space $L \log L$ is determined by $\Phi(b) = (b+1)\log(b+1)$. If a is real and positive, let k be the smallest nonnegative integer such that $a \leqslant 2^k$; since Φ is nondecreasing, we have that

$$\Phi(ab) \leqslant \Phi(2^k b) = \int_0^{2^k b} \varphi(\lambda)\, d\lambda = 2^k \int_0^b \varphi(2^k\lambda)\, d\lambda \leqslant 2^k c^k \Phi(b) \quad (3.11)$$

using $\varphi(2^k\lambda) \leqslant c^k\varphi(\lambda)$, which follows from (3.10). From this it follows, for example, that $\int_\Omega \Phi(f^*)$ and $\int_\Omega \Phi(af^*)$ are simultaneously finite or infinite. Also, by Fubini's theorem, we have the integral formula

$$\int_\Omega \Phi(f^*) = \int_\Omega \int_0^{f^*} \varphi(\lambda)\, d\lambda = \int_0^\infty \varphi(\lambda)\, P(f^* > \lambda)\, d\lambda.$$

THEOREM 3.2. *Suppose that conditions* A, B, *and* R *hold. Let* Φ *be as above and* $f \in \mathcal{M}$. *Then*

$$\int_\Omega \Phi(f^*) \leqslant c\int_\Omega \Phi(T^*f) + c\int_\Omega \Phi(d^*). \tag{3.12}$$

The choice of c *depends only on* $c_{(3.10)}$ *and the parameters of* A, B, *and* R. *This choice may be made so that, if the latter are fixed, the function* $c_{(3.10)} \to c$ *is nondecreasing.*

If, in addition, we assume that for a specific function Φ,

RΦ. $\quad \int_\Omega \Phi(d^*) \leqslant c\int_\Omega \Phi(T^*f)$

for all $f \in \mathcal{M}$, the inequality (3.12) may be simplied as follows:

THEOREM 3.3. *Suppose that conditions* A, B, R, *and* RΦ *hold. Then, for all* $f \in \mathcal{M}$,

$$\int_\Omega \Phi(f^*) \leqslant c\int_\Omega \Phi(T^*f).$$

The choice of c *depends only on* $c_{(3.10)}$ *and the parameters of* A, B, R, *and* RΦ. *This choice may be made so that, if the latter are fixed, the function* $c_{(3.10)} \to c$ *is nondecreasing.*

In particular, if conditions A, B, and R hold, and, for some p, $0 < p < \infty$,

$$\|d^*\|_p \leqslant c_p \|T^*f\|_p \tag{3.13}$$

for all $f \in \mathcal{M}$, then we have (3.1) as mentioned at the beginning of this section.

Some regularity assumption such as condition A is required in the theorems of this section and their left-hand analogues in Section 4. See the examples in Section 8.

Proof of Theorem 3.2. In preparation for using Lemma 2.6, we define

$$\psi(t) = ae^{at}\varphi(e^{at})P(f^* > e^{at}),$$
$$B = \{t: \psi(t) < 2\psi(t+1)\}.$$

Here we take $a = \log 2$ so that $e^{a(t+1)} = 2e^{at}$. The assumption of Lemma 2.6 is satisfied by ψ since

$$\int_{-\infty}^0 \psi(t)\,dt = \int_0^1 \varphi(\lambda)\,P(f^* > \lambda)\,d\lambda \leqslant \int_0^1 \varphi(\lambda)\,dy = \Phi(1) < \infty,$$

so that, by Lemma 2.6,

$$\int_{-\infty}^{\infty} \psi(t)\,dt \leqslant 2\int_B \psi(t+1)\,dt. \tag{3.14}$$

Now, note that if $t \in B$, then $\lambda = e^{at}$ satisfies

$$a\lambda\varphi(\lambda)P(f^*>\lambda) < 4a\lambda\varphi(2\lambda)P(f^*>2\lambda) \leqslant 4\,c_{(3.10)}a\lambda\varphi(\lambda)P(f^*>2\lambda).$$

In particular the above inequality implies that $\varphi(\lambda)$ is positive and finite, so that $P(f^*>\lambda) \leqslant \alpha P(f^*>2\lambda)$ with $\alpha = 4\,c_{(3.10)}$. With this choice of α,

$$2\psi(t+1) \leqslant \alpha\psi(t)$$

since $P(f^*>2\lambda) \leqslant P(f^*>\lambda)$. Therefore, by (3.14), Theorem 3.1, and (3.11),

$$\int_\Omega \Phi(f^*) = \int_{-\infty}^{\infty} \psi(t)\,dt \leqslant 2\int_B \psi(t+1)\,dt \leqslant \alpha\int_B \psi(t)\,dt$$

$$\leqslant \alpha\int_B ae^{at}\varphi(e^{at})\,[cP(cT^*f>e^{at}) + cP(cd^*>e^{at})]\,dt$$

$$\leqslant c\int_\Omega \Phi(cT^*f) + c\int_\Omega \Phi(cd^*) \leqslant c\int_\Omega \Phi(T^*f) + c\int_\Omega \Phi(d^*).$$

The assertions about the choice of c are evident, once the above argument is examined. This completes the proof.

4. The left-hand side

In this section, we prove integral inequalities analogous to those in Section 3. In particular, if $0<p<\infty$, then

$$\|T^{**}f^*\|_p \leqslant c\|f^*\|_p$$

for all $f \in \mathcal{M}$, under conditions A, B, L, and (4.4). Our discussion is briefer here because the proofs have much the same pattern as those of Section 3. The principal changes are as follows: (a) The function d^* is replaced by Δ^*, the maximal function of the sequence $\Delta = (\Delta_1, \Delta_2, \ldots)$ defined by $\Delta_n = T(^{n-1}f^n)$, $n \geqslant 1$. (b) Instead of the sequence f_n, $n \geqslant 1$, the sequence $T_n f$, $n \geqslant 1$, is used to define stopping times.

THEOREM 4.1. *Suppose that conditions* A, B, *and* L *hold. Let* $\alpha \geqslant 1$ *and* $\beta > \gamma^9$. *Then*

$$P(T^{**}f>\lambda) \leqslant cP(cf^*>\lambda) + cP(c\Delta^*>\lambda)$$

for all f *in* $\mathcal{M}$ *and* $\lambda > 0$ *satisfying*

$$P(T^{**}f>\lambda) \leqslant \alpha P(T^{**}f>\beta\lambda).$$

The choice of c *depends only on* α, β, *and the parameters of* A, B, *and* L. *Furthermore, this choice may be made so that, with* β *and the other parameters fixed, the function* $\alpha \to c$ *is non-decreasing.*

Proof. We proceed in steps, always assuming A, B, and L.

(i) *Let* $\alpha \geqslant 1$ *and* $\beta > \gamma^4$. *Then*

$$P(T^*f > \lambda) \leqslant cP(cf^* > \lambda) + cP(c\Delta^* > \lambda) \tag{4.1}$$

for all f in $\mathfrak{M}$ and $\lambda > 0$ satisfying $v^* \leqslant \lambda$ and

$$P(T^*f > \lambda) \leqslant \alpha P(T^*f > \beta\lambda). \tag{4.2}$$

Let $\theta = (\beta\gamma^{-4} - 1)/2$. Either

$$P(T^*f > \lambda) \leqslant 2\alpha P(T^*f > \beta\lambda,\ \Delta^* \leqslant \theta\lambda)$$

or

$$P(T^*f > \lambda) \leqslant 2\alpha P(\Delta^* > \theta\lambda).$$

The latter possibility leads directly to (4.1); therefore we assume the former.

Let

$$\mu = \inf\{n\colon T_n f > \lambda\},$$
$$\nu = \inf\{n\colon T_n f > \beta\lambda\},$$

and $g = {}^{\mu}f^{\nu}$. By L3, μ and ν are stopping times. On the set where μ is finite,

$$T_\mu f \leqslant \gamma(\lambda + \Delta^*).$$

To see this, let n be a positive integer. Then, on $\{\mu = n\}$,

$$T_\mu f = T_n f = Tf^n = T(f^{n-1} + {}^{n-1}f^n) \leqslant \gamma(T_{n-1}f + \Delta_n) \leqslant \gamma(\lambda + \Delta^*).$$

On the set where $T^*f > \beta\lambda$ and $\Delta^* \leqslant \theta\lambda$, we have that $\mu \leqslant \nu < \infty$, and

$$\beta\lambda < T_\nu f \leqslant \gamma T f^\nu = \gamma T(f^\mu + g) \leqslant \gamma^2(Tf^\mu + Tg) \leqslant \gamma^3(T_\mu f + Tg) \leqslant \gamma^4(\lambda + \theta\lambda + Tg)$$

or

$$Tg > (\beta\gamma^{-4} - 1 - \theta)\lambda = \theta\lambda.$$

Therefore,

$$P(Tg > \theta\lambda) \geqslant P(T^*f > \beta\lambda,\ \Delta^* \leqslant \theta\lambda) \geqslant cP(T^*f > \lambda),$$

so that, by L1, we have the lower estimate

$$\int_{\{T^*f > \lambda\}} |g^*|^{\pi_1} = \|g^*\|_{\pi_1}^{\pi_1} \geqslant c(\theta\lambda)^{\pi_1} P(Tg > \theta\lambda) \geqslant c\lambda^{\pi_1} P(T^*f > \lambda).$$

Now we compute the upper estimate. Let $b = 2\beta\gamma^3\lambda$ and $\tau = \inf\{n\colon T_n({}^{\mu}f) > b\}$. Then $\tau \geqslant \nu$, for on $\{\tau < \infty\}$,

$$b < T_\tau({}^{\mu}f) \leqslant \gamma T({}^{\mu}f^{\tau}) \leqslant \gamma^2(Tf^\tau + Tf^\mu) \leqslant \gamma^3(T_\tau f + T_\mu f) \leqslant 2\gamma^3 T_\tau^* f,$$

or

$$T_\tau^* f > \beta\lambda,$$

which implies that $\tau \geqslant \nu$.

Let $\pi_2=\varrho$. Then, by Theorem 2.2, for all $\eta>0$,

$$\eta^{\pi_2}P(g^*>\eta,\, T^*f>\lambda)\leqslant\eta^{\pi_2}P(g^*>\eta)\leqslant\|g^*\|_{\pi_2}^{\pi_2}=\|({}^{\mu}f^{\nu})^*\|_{\pi_2}^{\pi_2}$$
$$\leqslant\|({}^{\mu}f^{\tau})^*\|_{\pi_2}^{\pi_2}\leqslant cb^{\pi_2}P(s({}^{\mu}f)>0)\leqslant c\lambda^{\pi_2}P(\mu<\infty)=c\lambda^{\pi_2}P(T^*f>\lambda).$$

Using the lower and upper estimates just obtained, we may apply Lemma 2.3 to g^*:

$$P(g^*>c\lambda,\ T^*f>\lambda)\geqslant cP(T^*f>\lambda).$$

Since $g^*\leqslant 2f^*$, we have that

$$P(T^*f>\lambda)\leqslant cP(cf^*>\lambda).$$

This completes the proof of (i).

We now eliminate the assumption that $v^*\leqslant\lambda$.

(ii) *Let* $\alpha\geqslant 1$ *and* $\beta>\gamma^6$. *Then* (4.1) *holds for all* f *in* $\mathfrak{M}$ *and* $\lambda>0$ *satisfying* (4.2).

Let
$$\sigma=\{n\geqslant 0\colon |v_{n+1}|>\lambda\}$$

and $h=f^{\sigma}$. We now show that except for one case easily handled separately h satisfies (4.2) with α replaced by 2α, β by $\beta_0=\beta\gamma^{-2}$, and λ by $\lambda_0=\gamma\lambda$. Certainly, the multiplier sequence defining h is uniformly bounded by $\lambda\leqslant\lambda_0$. Also $\beta_0>\gamma^4$.

Either
$$P(T^*f>\lambda)\leqslant 2\alpha P(T^*f>\beta\lambda,\ v^*\leqslant\lambda)$$
or
$$P(T^*f>\lambda)\leqslant 2\alpha P(v^*>\lambda).$$

Using $d^*\leqslant 2f^*$ and Lemma 2.5, we have that $P(v^*>\lambda)\leqslant cP(cf^*>\lambda)$, so the latter possibility implies (4.1). Assume the former. Then, since $T^*h\leqslant\gamma T^*_{\sigma}f\leqslant\gamma T^*f$, we have that

$$P(T^*h>\lambda_0)\leqslant P(T^*f>\lambda)\leqslant 2\alpha P(T^*f>\beta\lambda,\ v^*\leqslant\lambda)=2\alpha P(T^*_{\sigma}f>\beta\lambda,\ \sigma=\infty)$$
$$\leqslant 2\alpha P(\gamma T^*h>\beta\lambda)=2\alpha P(T^*h>\beta_0\lambda_0).$$

Therefore, by (i),
$$P(T^*h>\lambda_0)\leqslant cP(ch^*>\lambda_0)+cP(c\Delta_0^*>\lambda_0).$$

Here $\Delta_0=(\Delta_{01},\Delta_{02},\ldots)$ is defined by $\Delta_{0n}=T({}^{n-1}h^n)$, $n\geqslant 1$. Note that

$$\Delta_{0n}=T({}^{n-1}f^{n\wedge\sigma})\leqslant\gamma T_{\sigma}({}^{n-1}f^n)\leqslant\gamma\Delta_n$$

since
$$T_{\sigma}({}^{n-1}f^n)=0\quad\text{on }\{\sigma<n\},$$
$$=\Delta_n\quad\text{on }\{\sigma\geqslant n\}.$$

Hence, $\Delta_0^*\leqslant\gamma\Delta^*$. Also, $h^*\leqslant f^*$, so we have that

$$P(T^*f>\lambda) \leqslant 2\alpha P(T^*h>\beta_0\lambda_0) \leqslant 2\alpha P(T^*h>\lambda_0) \leqslant cP(cf^*>\lambda)+cP(c\Delta^*>\lambda).$$

This completes the proof of (ii).

The proof of Theorem 4.1 may now be completed as follows. Either

$$P(T^{**}f>\lambda) \leqslant 2\alpha P(T^{**}f>\beta\lambda,\ f^*<\infty)$$

or

$$P(T^{**}f>\lambda) \leqslant 2\alpha P(f^*=\infty).$$

In the latter case, $P(f^*=\infty) \leqslant P(f^*>\lambda)$ and the desired inequality holds. Assume the former. By Lemma 4.1, proved below, we have that

$$P(T^*f>\lambda) \leqslant P(T^{**}f>\lambda) \leqslant 2\alpha P(T^{**}f>\beta\lambda,\ f^*<\infty) \leqslant 2\alpha P(\gamma^3 T^*f>\beta\lambda).$$

Applying (ii) with α replaced by 2α and β by $\beta\gamma^{-3}>\gamma^6$, we have that

$$P(T^{**}f>\lambda) \leqslant 2\alpha P(T^*f>\beta\gamma^{-3}\lambda) \leqslant 2\alpha P(T^*f>\lambda) \leqslant cP(cf^*>\lambda)+cP(c\Delta^*>\lambda),$$

the desired inequality.

LEMMA 4.1. *If conditions* A, B, *and either* L1 *or* R2 *hold, then, for any f in $\mathcal{M}$,*

$$Tf \leqslant \gamma^3 T^*f$$

on the set where $f^*<\infty$.

Accordingly, $T^{**}f \leqslant \gamma^3 T^*f$ on the same set. However, this inequality need not hold on the set $\{f^*=\infty\}$. Consider the operator

$$Tf = \limsup_{n\to\infty}\ ({}^n f)^*, \quad f \in \mathcal{M}.$$

Then

$$T^*f=0, \quad f\in\mathcal{M},$$

since

$$T_n f = \limsup_{m\to\infty}\ ({}^m f^n)^* = 0$$

for all positive integers n. On the other hand,

$$Tf=\infty \quad \text{on } \{f^*=\infty\}.$$

Therefore, although T satisfies the conditions of the lemma, $Tf \leqslant \gamma^3 T^*f$ fails to hold on $\{f^*=\infty\}$. (Under A1 and A2, there *are* f in $\mathcal{M}$ satisfying $f^*=\infty$ almost everywhere. See Corollary 5.6.)

Proof. Let $\lambda>0$,

$$\tau = \inf\{n \geqslant 0\colon s_{n+1}(f)>\lambda\},$$

and $g=f^\tau$. Then $s(g)\leqslant\lambda$ and, by the proof of (2.10), we have that

$$\|g^*\|_p \leq c_p \lambda, \quad 0 < p \leq \varrho.$$

Since $\varrho \geq 2$, by classical martingale theory, g converges almost everywhere. Therefore,

$$2g^* \geq ({}^n g)^* \to 0$$

almost everywhere, and, by the Lebesgue dominated convergence theorem, $\|({}^n g)^*\|_p \to 0$ as $n \to \infty$, for $0 < p \leq \varrho$. Now let $p = \pi_1$ or $p = p_2$ depending on whether L1 or R2 is satisfied. Let n_k be a positive integer such that

$$\|({}^{n_k} g)^*\|_p \leq 4^{-k}, \quad k \geq 1.$$

By L1 or R2,
$$2^{-k}[P(T({}^{n_k} g) > 2^{-k})]^{1/p} \leq c\, 4^{-k}, \quad k \geq 1,$$

implying that
$$T({}^{n_k} g) \to 0$$

almost everywhere as $k \to \infty$. Therefore,

$$Tg = T(g^{n_k} + {}^{n_k} g) \leq \gamma[Tf^{\tau \wedge n_k} + T({}^{n_k} g)] \leq \gamma^2[T_{\tau \wedge n_k} f + T({}^{n_k} g)] \leq \gamma^2[T^* f + T({}^{n_k} g)],$$

which implies that $Tg \leq \gamma^2 T^* f$. Consequently, on $\{\tau = \infty\} = \{s(f) \leq \lambda\}$,

$$Tf = T(g + {}^\tau f) \leq \gamma[Tg + T({}^\tau f)] = \gamma Tg \leq \gamma^3 T^* f.$$

Letting $\lambda \to \infty$, we see that this inequality holds on the set $\{s(f) < \infty\}$. But, by [7] this set is equivalent to $\{f^* < \infty\}$. (For another proof of this fact, using part (ii) of the proof of Theorem 4.1, see Corollary 5.6.) This completes the proof of Lemma 4.1.

THEOREM 4.2. *Suppose that conditions* A, B, *and* L *hold. Let* Φ *be as in Theorem 3.2 and* $f \in \mathfrak{M}$. *Then*

$$\int_\Omega \Phi(T^{**} f) \leq c \int_\Omega \Phi(f^*) + c \int_\Omega \Phi(\Delta^*). \tag{4.3}$$

The choice of c depends only on $c_{(3.10)}$ *and the parameters of* A, B, *and* L. *This choice may be made so that, if the latter are fixed, the function* $c_{(3.10)} \to c$ *is nondecreasing.*

The proof of Theorem 4.2 is similar to the proof of Theorem 3.2 and is omitted. One small change is to take $a = k \log 2$ with k the least positive integer satisfying $2^k > \gamma^9$. Then $e^{a(t+1)} = 2^k e^{at}$ and Theorem 4.1 is applied with $\beta = 2^k$.

If we assume that for a specific function Φ,

LΦ. $$\int_\Omega \Phi(\Delta^*) \leq c \int_\Omega \Phi(f^*)$$

for all $f \in \mathfrak{M}$, then inequality (4.3) may be simplified as follows:

THEOREM 4.3. *Suppose that conditions* A, B, L, *and* LΦ *hold. Then, for all* $f \in \mathcal{M}$,

$$\int_\Omega \Phi(T^{**}f) \leqslant c \int_\Omega \Phi(f^*).$$

The choice of c depends only on $c_{(3.10)}$ *and the parameters of* A, B, L, *and* LΦ. *This choice may be made so that, if the latter are fixed, the function* $c_{(3.10)} \to c$ *is nondecreasing.*

In particular, if conditions A, B, and L hold, and, for some p, $0 < p < \infty$,

$$\|\Delta^*\|_p \leqslant c_p \|f^*\|_p \tag{4.4}$$

for all $f \in \mathcal{M}$, then we have $$\|T^{**}f\|_p \leqslant c_p \|f^*\|_p$$

as mentioned at the beginning of this section.

5. The operators S and s

We now examine some applications of the theorems of Sections 3 and 4. This section is devoted to the two operators

$$S(f) = (\sum_{k=1}^{\infty} d_k^2)^{\frac{1}{2}}$$

and $$s(f) = [\sum_{k=1}^{\infty} E(d_k^2 \mid \mathcal{A}_{k-1})]^{\frac{1}{2}}.$$

THEOREM 5.1. *Let* $0 < p < \infty$. *If* A1 *and* A2 *are satisfied, then*

$$c_{p\delta} \|S(f)\|_p \leqslant \|f^*\|_p \leqslant C_{p\delta} \|S(f)\|_p$$

for all f in $\mathcal{M}$. *The choice of* $c_{p\delta}$ *and* $C_{p\delta}$ *depends only on p and* δ *and may be made so that, for fixed* δ, *the functions* $p \to C_{p\delta}^p$ *and* $p \to 1/c_{p\delta}^p$ *are nondecreasing.*

THEOREM 5.2. *Let* Φ *be as in Section 3. If* A1 *and* A2 *are satisfied, then*

$$c \int_\Omega \Phi(S(f)) \leqslant \int_\Omega \Phi(f^*) \leqslant C \int_\Omega \Phi(S(f))$$

for all f in $\mathcal{M}$. *The choice of c and C depends only on* $c_{(3.10)}$ *and* δ *and may be made so that, for fixed* δ, *the functions* $c_{(3.10)} \to C$ *and* $c_{(3.10)} \to 1/c$ *are nondecreasing.*

Theorem 5.1. is a consequence of Theorem 5.2, and both are special cases of Theorems 3.3 and 4.3. We have shown in Section 2 that the operator S satisfies B, R, and L. Also, the operator S satisfies RΦ and LΦ for *every* Φ since $d^* \leqslant S(f)$ and $\Delta^* = d^* \leqslant 2f^*$.

We now turn to the corresponding theorems for the operator s. Here, though the inequalities are similar, there are significant differences. The contrast between S and s, under conditions A1 and A2, may be summarized as follows: (a) The Φ-inequalities are two-sided for the operator S (Theorem 5.2), but only the left-hand side holds for s (Theorem 5.4). (b) Two-sided L_p-norm inequalities are valid for S in the interval $0<p<\infty$ (Theorem 5.1), but for s, they hold only in the range $0<p\leqslant 2$ (Theorem 5.3).

Notice that in the following theorem, assertions (i) and (ii) hold for all f in $\mathcal{N}$ (the natural domain of the operator s); we do not assume any part of condition A.

THEOREM 5.3.

(i) *Let* $2\leqslant p<\infty$. *Then, for all* $f\in\mathcal{N}$,

$$\|s(f)\|_p\leqslant c_p\|f\|_p.$$

The choice of c_p depends only on p and may be made so that the function $p\to c_p$ is bounded on each compact subinterval of $[2, \infty)$.

(ii) *Let* $0<p\leqslant 2$. *Then, for all* $f\in\mathcal{N}$,

$$\|f^*\|_p\leqslant C_p\|s(f)\|_p.$$

The choice of C_p depends only on p and may be made so that the function $p\to C_p^p$ is bounded on $(0, 2]$.

(iii) *Let* $0<p\leqslant 2$. *If* A1 *and* A2 *are satisfied, then, for all* $f\in\mathcal{M}$,

$$\|s(f)\|_p\leqslant c_{p\delta}\|f^*\|_p.$$

The choice of $c_{p\delta}$ depends only on p and δ and may be made so that, for fixed δ, the function $p\to c_{p\delta}^p$ is bounded on $(0, 2]$.

(iv) *Let* $2\leqslant p\leqslant\varrho$. *If condition* A *holds, then, for all* $f\in\mathcal{M}$,

$$\|f^*\|_p\leqslant C\|s(f)\|_p.$$

The choice of C depends only on p, δ, and c_{A3}, and may be made so that, for fixed δ and c_{A3}, the function $p\to C$ is bounded on $[2, \varrho]$.

Part (iii) of Theorem 5.3 is a special case of the following:

THEOREM 5.4. *Suppose that conditions* A1 *and* A2 *hold. Let Φ be as in Section* 3. *Then for all* $f\in\mathcal{M}$,

$$\int_\Omega \Phi(s(f)) \leqslant c \int_\Omega \Phi(f^*).$$

The choice of c depends only on $c_{(3.10)}$ and δ and may be made so that for fixed δ, the function $c_{(3.10)} \to c$ is nondecreasing.

Condition A2 alone is not sufficient to imply the conclusions of Theorem 5.4 and parts (iii) and (iv) of Theorem 5.3. See Example 8.2.

Proof of Theorem 5.3. (i) If j is a positive integer and f is in $\mathcal{N}$, then

$$\|s(f)\|_{2j} \leqslant j^{\frac{1}{2}} \|S(f)\|_{2j}. \tag{5.1}$$

This is certainly true if $j=1$. Let $j>1$ and suppose that $\|S(f)\|_{2j} < \infty$. Then, letting $v_k^2 = E(d_k^2 \mid \mathcal{A}_{k-1})$, we have that $\|v_k^2\|_j \leqslant \|d_k^2\|_j \leqslant \|s(f)\|_{2j}^2 < \infty$, by the norm diminishing property of conditional expectations, and therefore all of the following integrals are finite. Let n be a positive integer and

$$K = \{1, \ldots, n\} \times \ldots \times \{1, \ldots, n\} \ (j \text{ factors}).$$

If $k = (k_1, \ldots, k_j) \in K$, let $|k| = \max(k_1, \ldots, k_j)$. Then letting

$$A_i = \{k \in K : |k| = k_i\},$$

we have that
$$K \subset A_1 \cup \ldots \cup A_j$$

and
$$\int_\Omega \Big|\sum_{k=1}^n v_k^2\Big|^j \leqslant \sum_{i=1}^j \sum_{k \in A_i} \int_\Omega v_{k_1}^2 \ldots v_{k_j}^2$$

$$= \sum_{i=1}^j \sum_{k \in A_i} \int_\Omega d_{k_i}^2 \prod_{\substack{m=1 \\ m \neq i}}^j v_{k_m}^2 \leqslant \sum_{i=1}^j \int_\Omega \sum_{k \in K} d_{k_i}^2 \prod_{\substack{m=1 \\ m \neq i}}^j v_{k_m}^2$$

$$= j \int_\Omega \Big|\sum_{k=1}^n v_k^2\Big|^{j-1} \Big|\sum_{k=1}^n d_k^2\Big| \leqslant j \left[\int_\Omega \Big|\sum_{k=1}^n v_k^2\Big|^j\right]^{(j-1)/j} \left[\int_\Omega \Big|\sum_{k=1}^n d_k^2\Big|^j\right]^{1/j}.$$

Therefore, we have obtained

$$\|s_n(f)\|_{2j}^2 \leqslant j \|S_n(f)\|_{2j}^2.$$

and (5.1) follows by the monotone convergence theorem. Combining (2.5) and (5.1), we have that

$$\|s(f)\|_{2j} \leqslant c_{2j} \|f\|_{2j} \tag{5.2}$$

for each positive integer j. To finish the proof of (i), we apply the Riesz–Thorin interpolation theorem in the form given by Calderón and Zygmund [2].

First we remark that (5.2) is true for complex martingales provided d_k^2 is replaced by $|d_k|^2$ in the definition of $s(f)$. This can be seen in two ways: with obvious modifications, all of our results so far carry over to the complex case; or (5.2) for the real case directly implies an inequality of the same form for the complex case.

Now consider the operator T defined on complex $L_2 = L_2(\Omega, \mathcal{A}, P)$ as follows. If $f_\infty \in L_2$, then $Tf_\infty = s(f)$ where $f = (f_1, f_2, \ldots)$ is the complex martingale defined by

$$f_n = E(f_\infty \mid \mathcal{A}_n), \quad n \geq 1.$$

We use the fact that

$$\|f\|_p \leq \|f_\infty\|_p, \quad 1 < p < \infty,$$

with equality holding if f_∞ is measurable relative to $\mathcal{A}_\infty$, the smallest σ-field containing $\bigcup_{k=1}^\infty \mathcal{A}_k$. By the complex version of (5.2),

$$\|Tf_\infty\|_{2j} \leq c_{2j}\|f_\infty\|_{2j}, \quad j = 1, 2, \ldots.$$

Clearly, T satisfies

$$T(f_\infty + g_\infty) \leq Tf_\infty + Tg_\infty, \quad f_\infty, g_\infty \in L_2 \tag{5.3}$$

and the other conditions of the interpolation theorem of Calderón and Zygmund. Therefore, if $2 \leq p \leq 2j$, we have that

$$\|Tf_\infty\|_p \leq c_p \|f_\infty\|_p, \quad f_\infty \in L_p, \tag{5.4}$$

$$c_p \leq c_2 \vee c_{2j}.$$

In summary, if f is in $\mathcal{N}$, $2 \leq p < \infty$, $\|f\|_p < \infty$, and f_∞ denotes the almost everywhere limit of f, then

$$\|s(f)\|_p = \|Tf_\infty\|_p \leq c_p\|f_\infty\|_p = c_p\|f\|_p$$

and the function $p \to c_p$ is bounded on each compact subinterval of $[2, \infty)$. This completes the proof of (i).

(ii) We may assume in the proof of (ii) that $\mathcal{A}_\infty = \mathcal{A}$. Then T, the operator defined in the proof of (i), is an isometry in L_2, where now it is enough to consider real L_2. Therefore, by (5.3),

$$\begin{aligned} 2(f_\infty, g_\infty) &= \|f_\infty + g_\infty\|_2^2 - \|f_\infty\|_2^2 - \|g_\infty\|_2^2 \\ &= \|T(f_\infty + g_\infty)\|_2^2 - \|Tf_\infty\|_2^2 - \|Tg_\infty\|_2^2 \leq 2(Tf_\infty, Tg_\infty) \end{aligned}$$

for all f_∞ and g_∞ in L_2.

Now let $1 < p \leq 2 \leq q$, $p^{-1} + q^{-1} = 1$. If $f_\infty \in L_2$ and $B = \{g_\infty \in L_2 : \|g_\infty\|_q \leq 1\}$, then

$$\|f_\infty\|_p = \sup_{g_\infty \in B} (f_\infty, g_\infty) \leq \sup_{g_\infty \in B} (Tf_\infty, Tg_\infty) \leq \sup_{g_\infty \in B} \|Tf_\infty\|_p \|Tg_\infty\|_q \leq c_q \|Tf_\infty\|_p,$$

by (5.4). If f is in $\mathcal{N}$ and $\|f\|_2 = \|s(f)\|_2 < \infty$, then f converges almost everywhere to a function f_∞ in L_2. Therefore, using (2.2), we have that

$$\|f^*\|_p \leqslant q\|f\|_p = q\|f_\infty\|_p \leqslant qc_q\|Tf_\infty\|_p = qc_q\|s(f)\|_p. \tag{5.5}$$

Now we show that (5.5) holds without the assumption that $\|s(f)\|_2 < \infty$. This will complete the proof of (ii) since, by (i), the function $p \to qc_q$ is bounded on compact subintervals of $(1, 2]$ and, by the proof of Lemma 2.2,

$$\|f^*\|_p^p \leqslant 4\,\|s(f)\|_p^p, \quad 0 < p \leqslant 4/3.$$

Let f belong to $\mathcal{N}$ and $1 < p \leqslant 2 \leqslant q$, $p^{-1} + q^{-1} = 1$. To prove (5.5), we may assume that $\|s(f)\|_p < \infty$; then $E(d_k^2 \mid \mathcal{A}_{k-1}) < \infty$, $k \geqslant 1$. Let m be a positive integer. If $1 \leqslant k \leqslant m$, let

$$\begin{aligned} e_{mk} &= d_k \quad \text{on} \quad \{E(d_k^2 \mid \mathcal{A}_{k-1}) \leqslant m\}, \\ &= d_k/[E(d_k^2 \mid \mathcal{A}_{k-1})]^{\frac{1}{2}} \quad \text{elsewhere;} \end{aligned}$$

if $k > m$, let $e_{mk} = 0$. Then $e_m = (e_{m1}, e_{m2}, \ldots)$ is a martingale difference sequence relative to $\mathcal{A}_1, \mathcal{A}_2, \ldots$, and the martingale $g_m = (g_{m1}, g_{m2}, \ldots)$ with this difference sequence satisfies $\|s(g_m)\|_2 < \infty$. Therefore,

$$\|(g_m)^*\|_p \leqslant qc_q\|s(g_m)\|_p \leqslant qc_q\|s(f)\|_p,$$

using the fact that $|e_{mk}|$ increases to $|d_k|$ as $m \to \infty$. Since $\lim g_{mn} = f_n$,

$$f^* \leqslant \liminf_{m\to\infty} (g_m)^*.$$

Therefore, by Fatou's lemma,

$$\|f^*\|_p \leqslant \liminf_{m\to\infty} \|(g_m)^*\|_p \leqslant qc_q\|s(f)\|_p.$$

This completes the proof of (ii).

(iii) This part of Theorem 5.3 is an immediate corollary of Theorem 5.4 with $\Phi(b) = b^p$, $0 < p \leqslant 2$. Here $c_{(3.10)} = 2^{p-1}$ and $c_{p\delta}^p \leqslant c_{2\delta}^2$, according to the final assertion of Theorem 5.4.

(iv) This part follows from Theorem 3.3 applied to the operator s and the function $\Phi(b) = b^p$, $2 \leqslant p \leqslant \varrho$. As mentioned in Section 2, s satisfies condition B. Condition R1 is satisfied with $p_1 = 1$ by Lemma 2.2; condition R2 holds for $p_2 = 2$ by (i) of this theorem. Now let us assume that condition A holds in order to check RΦ with $\Phi(b) = b^p$ for some p, $2 \leqslant p \leqslant \varrho$. Then

$$\begin{aligned} \|d^*\|_p^p &\leqslant \int_\Omega \sum_{k=1}^\infty |v_k|^p |x_k|^p = \int_\Omega \sum_{k=1}^\infty |v_k|^p E(|x_k|^p \mid \mathcal{A}_{k-1}) \leqslant c_{A3}^{p/\varrho} \int_\Omega \sum_{k=1}^\infty |v_k|^p \\ &\leqslant c_{A3} \int_\Omega \Big(\sum_{k=1}^\infty v_k^2\Big)^{p/2} = c_{A3}\|s(f)\|_p^p. \end{aligned}$$

Therefore, RΦ holds with $c_{R\Phi} = c_{A3}$ for every p, $2 \leqslant p \leqslant \varrho$. Part (iv) of Theorem 5.3 now follows from Theorem 3.3.

Proof of Theorem 5.4. We assume that conditions A1 and A2 hold and apply Theorem 4.3 to the operator s. This operator satisfies B and the measurability condition L3, and, as shown in the proof of Theorem 5.3 (iv), condition L2 holds. Therefore, only L1 and LΦ must be checked; we show that L1 holds with $\pi_1 = 1$.

Define an operator T on the set of all martingales f relative to $\mathcal{A}_1, \mathcal{A}_2, \ldots$ by

$$Tf = \left[\sum_{k=1}^{\infty} E(|d_k| \,|\, \mathcal{A}_{k-1})^2\right]^{\frac{1}{2}}.$$

Then T satisfies conditions (1) and (3) in the definition of class $\mathcal{B}$ mappings in [8]:

$$T(f+g) \leqslant Tf + Tg, \quad \|Tf\|_2 \leqslant \|s(f)\|_2 = \|f\|_2,$$

and

$$\|Tf\|_1 \leqslant \left\|\sum_{k=1}^{\infty} E(|d_k| \,|\, \mathcal{A}_{k-1})\right\|_1 = \left\|\sum_{k=1}^{\infty} |d_k|\right\|_1.$$

Since $Tf \leqslant s(f)$, T satisfies B2, but not the local condition (2) of class $\mathcal{B}$ mappings. However, the results of [8] remain valid if condition (2) of class $\mathcal{B}$ mappings is replaced by condition B2 of the present paper. In fact, consider the martingale $a = f - f^t = {}^t f$ as defined on page 137 of [8]. It is clear that $I(t=\infty) \leqslant I(s(a)=0)$ so that $I(Ta > \lambda) \leqslant I(Ta > 0) \leqslant I(t < \infty)$ for *any* operator that is local according to B2. Therefore, B2 may be used as the local condition in place of (2) since we have the required estimate:

$$P(Ta > \lambda) \leqslant P(t < \infty) \leqslant c\|f\|_1/\lambda$$

by the definition of the stopping time t.

In particular, we conclude that the operator T defined above satisfies the inequality

$$\lambda P(Tf > \lambda) \leqslant c\|f\|_1,$$

by the proposition on page 136 of [8].

Now consider the operator s. Since A2 holds, $s(f) = [\sum_{k=1}^{\infty} v_k^2]^{\frac{1}{2}}$, where, as usual, $d_k = v_k x_k$, $k \geqslant 1$. By A1,

$$\delta|v_k| \leqslant |v_k| E(|x_k| \,|\, \mathcal{A}_{k-1}) = E(|d_k| \,|\, \mathcal{A}_{k-1}),$$

so that $\delta s(f) \leqslant Tf$ for all f in $\mathcal{M}$. Therefore,

$$\lambda P(s(f) > \lambda) \leqslant \lambda P(\delta^{-1} Tf > \lambda) \leqslant c\|f\|_1,$$

so that L1 holds with $\pi_1 = 1$. Finally, we note that $\Delta^* = v^*$ in this case, so that, by Lemma 2.5,

$$P(\Delta^* > \lambda) \leqslant cP(cd^* > \lambda) \leqslant cP(cf^* > \lambda),$$

which in turn implies LΦ for any Φ:

$$\int_\Omega \Phi(\Delta^*) \leqslant c \int_\Omega \Phi(f^*).$$

Therefore, by Theorem 4.3, we obtain

$$\int_\Omega \Phi(s(f)) \leqslant c \int_\Omega \Phi(f^*).$$

This completes the proof of Theorem 5.4.

Some applications of S and s. Just a few are mentioned here. The next two sections contain others.

Random walk. One interesting special case of Theorem 5.3 is the following. Let $x = (x_1, x_2, \ldots)$ be an independent sequence of random variables, each with expectation zero and variance one, and such that

$$\|x_k\|_1 \geqslant \delta > 0, \quad k \geqslant 1.$$

Then, for any stopping time τ,

$$c_{p\delta} \|\tau^{\frac{1}{2}}\|_p \leqslant \|(X^\tau)^*\|_p \leqslant C_p \|\tau^{\frac{1}{2}}\|_p, \quad 0 < p \leqslant 2,$$

where X^τ is the martingale of partial sums $X_n = \sum_{k=1}^n x_k$ stopped at τ:

$$X_n^\tau = \sum_{k=1}^n I(\tau \geqslant k)\, x_k.$$

The following corollary contains a variation of the Wald equation for the expectation of a sum of a random number of random variables:

COROLLARY 5.1. *Suppose that the martingale difference sequence $x = (x_1, x_2, \ldots)$ satisfies* A2 *and $\int_\Omega x_1 = 0$. If τ is a stopping time such that $\|\tau^{\frac{1}{2}}\|_1 < \infty$, then*

$$\int_\Omega \sum_{k=1}^{\tau} x_k = 0. \tag{5.6}$$

Note that, although we have assumed condition A2, we have relaxed the usual requirement that τ be integrable.

Proof. Let $f = (f_1, f_2, \ldots)$ be defined by

$$f_n = \sum_{k=1}^{n} I(\tau \geqslant k)\, x_k.$$

Since τ is finite almost everywhere,

$$\lim_{n\to\infty} f_n = \sum_{k=1}^{\tau} x_k.$$

Also, f is a martingale and $$\int_\Omega f_n = 0, \quad n \geqslant 1.$$

By A2, $s(f) = \tau^{\frac{1}{2}}$, and, by Lemma 2.2,

$$\int_\Omega f^* \leqslant c \,\|\tau^{\frac{1}{2}}\|_1 < \infty.$$

Therefore, by the Lebesgue dominated convergence theorem,

$$\int_\Omega \sum_{k=1}^{\tau} x_k = \lim_{n\to\infty} \int_\Omega f_n = 0,$$

the desired result.

Corollary 5.2. *Suppose that $x = (x_1, x_2, \ldots)$ satisfies the conditions of Corollary 5.1 and that τ is the stopping time defined by*

$$\tau = \inf\{n \geqslant 1 : \sum_{k=1}^{n} x_k \geqslant 0\}.$$

Then $\tau^{\frac{1}{2}}$ is not integrable.

Proof. Suppose that $\tau^{\frac{1}{2}}$ is integrable. Then

$$\sum_{k=1}^{\tau} x_k$$

is nonnegative almost everywhere and is positive on a set of positive measure (on the set $\{x_1 > 0\}$ and possibly elsewhere). Therefore,

$$0 < \int_\Omega \sum_{k=1}^{\tau} x_k,$$

which contradicts the conclusion of Corollary 5.1. Thus, $\tau^{\frac{1}{2}}$ is not integrable.

The following inequality is due to Khintchine [17; Chapter V, 8.5]. If $a_1, a_2, \ldots$ is a sequence of real numbers, and $x_1, x_2, \ldots$ is an independent sequence of random variables such that $P(x_k = 1) = P(x_k = -1) = \frac{1}{2}$, then, for every $n \geqslant 1$,

$$c_p \left(\sum_{k=1}^{n} a_k^2\right)^{p/2} \leqslant \int_\Omega \left|\sum_{k=1}^{n} a_k x_k\right|^p \leqslant C_p \left(\sum_{k=1}^{n} a_k^2\right)^{p/2}, \quad 0 < p < \infty. \tag{5.7}$$

The next corollary generalizes Khintchine's inequality (even for the above random variable sequence). If x is a random variable, we say that x is symmetrically distributed if x and $-x$ have the same distribution.

COROLLARY 5.3. *Let $x_1, x_2, \ldots$ be an independent sequence of symmetrically distributed random variables and Φ a function as in Section 3. Then, for every $n \geqslant 1$,*

$$c\int_\Omega \Phi\left(\left[\sum_{k=1}^n x_k^2\right]^{\frac{1}{2}}\right) \leqslant \int_\Omega \Phi\left(\left|\sum_{k=1}^n x_k\right|\right) \leqslant C\int_\Omega \Phi\left(\left[\sum_{k=1}^n x_k^2\right]^{\frac{1}{2}}\right),$$

where the choice of c and C depends only on the function Φ.

Proof. Note that we do not impose any integrability conditions on x_k, $k \geqslant 1$, so that $f = (f_1, f_2, \ldots)$, $f_n = \sum_{k=1}^n x_k$, is not necessarily a martingale. However, by Remark 8.2, f has the same distribution as a martingale transform of an independent sequence of symmetrically distributed random variables taking the values ± 1. Furthermore, by Lévy's inequality (see [5], page 106),

$$P(|f_n| > \lambda) \leqslant P((f^n)^* > \lambda) \leqslant 2\,P(|f_n| > \lambda),$$

so that

$$\int_\Omega \Phi(|f_n|) \leqslant \int_\Omega \Phi(|f^n|^*) \leqslant 2\int_\Omega \Phi(|f_n|).$$

This, together with Theorem 5.2, implies the desired result.

Neither side of the inequality in Corollary 5.3 holds if symmetry is replaced by the assumption that each x_k has expectation zero. (For a counterexample with $\Phi(b) = b^p$, $0 < p < 1$, see Example 8.1.) However, we have the following generalization of an inequality of Marcinkiewicz and Zygmund [12; Theorem 5].

COROLLARY 5.4. *Let $x_1, x_2, \ldots$ be an independent sequence of random variables, each with expectation zero, and Φ a function as in Section 3. If Φ is also convex, then for every $n \geqslant 1$,*

$$c\int_\Omega \Phi\left(\left[\sum_{k=1}^n x_k^2\right]^{\frac{1}{2}}\right) \leqslant \int_\Omega \Phi\left(\left|\sum_{k=1}^n x_n\right|\right) \leqslant C\int_\Omega \Phi\left(\left[\sum_{k=1}^n x_k^2\right]^{\frac{1}{2}}\right),$$

where the choice of c and C depends only on the function Φ.

Proof. Let y be a sequence independent of x and with the same distribution. By the convexity of Φ and Jensen's inequality for conditional expectations, we have that

$$\int_\Omega \Phi\left(\left|\sum_{k=1}^n x_k\right|\right) \leqslant \int_\Omega \Phi\left(\left|\sum_{k=1}^n (x_k - y_k)\right|\right)$$

and
$$\int_\Omega \Phi\left(\left[\sum_{k=1}^n x_k^2\right]^{\frac{1}{2}}\right) \leqslant \int_\Omega \Phi\left(\left[\sum_{k=1}^n (x_k-y_k)^2\right]^{\frac{1}{2}}\right).$$

These inequalities may be reversed (up to a multiplicative constant) since, by our assumptions, for all $a \geqslant 0$, $b \geqslant 0$,
$$\Phi(a+b) \leqslant c[\Phi(a)+\Phi(b)].$$

The resulting double inequalities allow us to obtain the desired result for x from Corollary 5.3 applied to $x-y$.

Haar and Walsh series. Let $(\Omega, \mathcal{A}, P)$ be the Lebesgue unit interval and $r_0, r_1, \ldots$ the sequence of Rademacher functions: $r_k(\omega)=1$ if $\omega \in [j/2^{k+1}, (j+1)/2^{k+1})$ for some even integer j; $r_k(\omega)=-1$ otherwise. Then $x=(x_1, x_2, x_3, \ldots)=(1, r_0, r_1, \ldots)$ is a martingale difference sequence relative to $\mathcal{A}_1, \mathcal{A}_2, \ldots$, where $\mathcal{A}_n$ is the smallest σ-field with respect to which $x_1, \ldots, x_n$ are measurable. It is not difficult to see that every martingale transform relative to $\mathcal{A}_0, \mathcal{A}_1, \ldots$ (necessarily $\mathcal{A}_0=\mathcal{A}_1$) is a martingale and an element of $\mathcal{M}$, the collection corresponding to the sequence x; that is, $\mathcal{M}=\mathcal{N}$. Furthermore, x satisfies condition A and, in fact, since $|x_k|=1$, $k\geqslant 1$, there is no distinction here between the operators S and s.

To prove inequality (1.3), we consider the Walsh functions $\psi_0, \psi_1, \ldots$. Recall that if $0 \leqslant \omega < 1$ and n is a positive integer satisfying $n=2^{n_1}+\ldots+2^{n_k}$ with $n_1>n_2>\ldots>n_k\geqslant 0$, then $\psi_0(\omega)=1$ and $\psi_n(\omega)=r_{n_1}(\omega)\ldots r_{n_k}(\omega)$. If $b_0, b_1, \ldots$ is any sequence of real numbers, let
$$f_{n+1}=\sum_{k=0}^{2^n-1} b_k\psi_k, \quad n\geqslant 0.$$

Then the difference sequence d of $f=(f_1, f_2, \ldots)$ satisfies $d_k=v_kx_k$ where v_k is $\mathcal{A}_{k-1}$-measurable. For example, if $k\geqslant 1$, then
$$d_{k+2}=f_{k+2}-f_{k+1}=r_k\sum_{j=0}^{2^k-1} a_{2^k+j}\psi_j,$$

and, since the sum is a function of $r_0, \ldots, r_{k-1}$, we have that $d_{k+2}=v_{k+2}x_{k+2}$ where v_{k+2} is $\mathcal{A}_{k+1}$-measurable. Therefore, f is a martingale transform of x, and, by Theorem 5.1, we have that
$$c_p\|S(f)\|_p \leqslant \|f^*\|_p \leqslant C_p\|S(f)\|_p$$

for $0<p<\infty$. This is equivalent to (1.3) by the observation of Marcinkiewicz [11].

In [14], R. E. A. C. Paley uses the function $S(f)$ to study Walsh–Fourier series in much the same way the conjugate function is used for ordinary Fourier series. From this viewpoint, inequality (1.1) is the analogue of the L_p-norm inequality for the conjugate function, due to M. Riesz. In the same vein, we mention the following result for $S(f)$, whose analogue for the conjugate function is also due to M. Riesz [17; Chapter VII, 2.8, 2.10].

COROLLARY 5.5. *Let f be a nonnegative function belonging to $L_1[0, 1)$ and $S(f)$ its Paley function as in* (1.1). *Then $S(f)$ belongs to $L_1[0, 1)$ if and only if f belongs to $L \log L$.*

Proof. By Theorem 2 of [9], the assertion of Corollary 5.5 holds if $S(f)$ is replaced by the function f^* appearing in inequality (1.3). Since, as we have seen, this inequality holds for $p=1$, the assertion of Corollary 5.5 holds as stated.

Local convergence of martingale transforms.

COROLLARY 5.6. *Suppose that conditions* A1 *and* A2 *are satisfied. If f is in $\mathcal{M}$, then the following sets are equivalent:*

$$\{f \text{ converges}\},\quad \{s(f)<\infty\},\quad \{S(f)<\infty\},\quad \{f^*<\infty\},\quad \{\sup_n f_n<\infty\}.$$

This is a known result. We give new proofs of

$$\{f^*<\infty\}\subset\{s(f)<\infty\} \tag{5.8}$$

and

$$\{\sup_n f_n<\infty\}\subset\{f^*<\infty\}, \tag{5.9}$$

the inclusion sign to be interpreted as holding up to a set of measure zero.

The inclusion (5.8) is the main contribution of [7]. Recently, Davis [4] has shown that $\{\sup_n f_n<\infty\}\subset\{s(f)<\infty\}$. Dvoretzky has another proof [6]. The other inclusions are easier; their proofs are omitted.

For another class of equivalent sets, see Theorem 6.2.

In order to prove the local results (5.8) and (5.9), it is enough to prove the corresponding global results:

$$f^*<\infty \text{ a.e.} \quad \text{implies } s(f)<\infty \text{ a.e.}, \tag{5.10}$$

and

$$\sup_n f_n<\infty \text{ a.e.} \quad \text{implies } f^*<\infty \text{ a.e.} \tag{5.11}$$

Let us show, for example, that if (5.10) holds for all f in $\mathcal{M}$, then (5.8) holds for all f in $\mathcal{M}$. Let $f\in\mathcal{M}$ and $\tau=\inf\{n\colon |f_n|>\lambda\}$ for some positive λ. Then $g=f^\tau$ is in $\mathcal{M}$ and $g^*<\infty$ a.e. By (5.10), $s(g)<\infty$ a.e.; therefore

$$\{f^*\leqslant\lambda\}\subset\{\tau=\infty\}\subset\{s(f)=s(g)\}\subset\{s(f)<\infty\},$$

and (5.8) follows by letting $\lambda\to\infty$.

We now prove (5.10) by applying Theorem 4.1 and Lemma 2.5 to the operator s. Assume that $P(s(f)=\infty)>0$, and take $\alpha=1/P(s(f)=\infty)$, $\beta=2$. Then, for *any* $\lambda>0$,

$$P(s(f)>\lambda)\leqslant 1=\alpha P(s(f)=\infty)\leqslant\alpha P(s(f)>2\lambda).$$

Therefore, by Theorem 4.1 and Lemma 2.5,

$$0 < P(s(f) = \infty) \leqslant P(s(f) > \lambda) \leqslant cP(cf^* > \lambda).$$

Since this holds for all $\lambda > 0$, we see that $P(f^* = \infty) > 0$, which implies (5.10).

The following lemma implies (5.11) in a similar way.

LEMMA 5.1. *Suppose that the martingale difference sequence* $x = (x_1, x_2, \dots)$ *satisfies* A1, A2, *and*

$$E(x_1 \mid \mathcal{A}_0) = 0.$$

Then

$$P(\sup_n f_n > c\lambda P(f^* > \lambda)) \geqslant c[P(f^* > \lambda)]^2 \tag{5.12}$$

for all f *in* $\mathcal{M}$ *and* $\lambda > 0$. *The choice of* c *depends only on* δ.

By (5.12), if $P(f^* = \infty) > 0$, then $P(\sup_n f_n = \infty) > 0$. For the purpose of proving (5.11), the assumption that $E(x_1 \mid \mathcal{A}_0) = 0$ can be made without loss of generality.

Proof of Lemma 5.1. Using $E(x_k \mid \mathcal{A}_{k-1}) = 0$, A1, and A2, we have that, for all $A \in \mathcal{A}_{k-1}$,

$$\int_A x_k^+ = \int_A (|x_k| + x_k)/2 = \tfrac{1}{2}\int_A |x_k| \geqslant \tfrac{1}{2}\delta P(A)$$

and

$$\int_A (x_k^+)^2 \leqslant \int_A x_k^2 = P(A)$$

with $x_k^+ = x_k \vee 0$. Therefore, by Lemma 2.3,

$$P(x_k > 2b, A) = P(x_k^+ > 2b, A) \geqslant b^2 P(A), \qquad A \in \mathcal{A}_{k-1}, k \geqslant 1, \tag{5.13}$$

with $b = \delta/8$.

Let $f \in \mathcal{M}$ and $\lambda > 0$. Define stopping times σ and ν as follows:

$$\sigma = \{n \geqslant 0\colon\ |v_{n+1}| > \lambda\},$$

$$\nu = \{n\colon |f_n| > \lambda\}.$$

Let $\mu = \sigma \wedge \nu$, $g = f^\mu$, $h = f^\sigma$ and

$$\tau = \inf\{n\colon |h_n| > \lambda\}.$$

Then $\mu \leqslant \sigma \wedge \tau$. This is clearly true on the set $\{\sigma \leqslant \tau\}$ and, on its complement, $\tau = \nu$. The multiplier sequence defining h is uniformly bounded by λ. Therefore, by Theorem 2.1 applied to h,

$$\|g^*\|_2 = \|(f^\mu)^*\|_2 \leqslant \|(f^{\sigma\wedge\tau})^*\|_2 = \|(h^\tau)^*\|_2 \leqslant c\lambda,$$

with the choice of c depending only on δ. Accordingly, g converges almost everywhere to g_∞, and

$$\|g_\infty^+\|_2 \leqslant \|g^*\|_2 \leqslant c\lambda. \tag{5.14}$$

We now obtain a lower estimate. Let $\pi = P(f^* > \lambda)$; then, since $b < 1$, we have that either

$$P(|f_\mu| > b\lambda, \mu < \infty) \geqslant \tfrac{1}{2}\pi \tag{5.15}$$

or

$$P(|f_\mu| \leqslant b\lambda, |v_{\mu+1}| > \lambda, \mu < \infty) \geqslant \tfrac{1}{2}\pi. \tag{5.16}$$

(i) Suppose (5.15) holds. By the Lebesgue dominated convergence theorem,

$$\int_\Omega g_\infty = \lim_{n\to\infty} \int_\Omega g_n = 0.$$

Therefore,

$$\|g_\infty^+\|_1 = \tfrac{1}{2}\|g_\infty\|_1 \geqslant \frac{b\lambda}{2} P(|g_\infty| > b\lambda) \geqslant \frac{b\lambda}{2} P(|f_\mu| > b\lambda, \mu < \infty)$$

$$\geqslant \frac{b\lambda}{2} \cdot \frac{\pi}{2} = c\pi\lambda.$$

Using this inequality, (5.14), and Lemma 2.3, we have that

$$c\pi^2 \leqslant P(g_\infty^+ > c\pi\lambda) \leqslant P(\sup_n f_n > c\pi\lambda),$$

the desired inequality.

(ii) Suppose (5.16) holds. Then either

$$\frac{\pi}{4} \leqslant P(|f_\mu| \leqslant b\lambda, v_{\mu+1} > \lambda, \mu < \infty)$$

or

$$\frac{\pi}{4} \leqslant P(|f_\mu| \leqslant b\lambda, v_{\mu+1} < -\lambda, \mu < \infty).$$

Suppose the former, the other case being similar. Let

$$A_k = \{f_k \geqslant -b\lambda, v_{k+1} > \lambda, \mu = k\}.$$

Then, by (5.13),

$$\frac{\pi}{4} \leqslant \sum_{k=0}^\infty P(A_k) \leqslant b^{-2} \sum_{k=0}^\infty P(x_{k+1} > 2b, A_k)$$

$$\leqslant b^{-2} \sum_{k=0}^\infty P(f_k + v_{k+1} x_{k+1} > -b\lambda + 2b\lambda, \mu = k) \leqslant b^{-2} P(\sup_n f_n > b\lambda).$$

Since $\pi \leqslant 1$, we get (5.12) in this case also.

This completes the proof of Lemma 5.1.

6. Operators of matrix type

In this section, we introduce a class of operators and illustrate further the range of application of the theorems in the preceding sections.

An operator M is said to be of *matrix type* if it has the following properties:

(a) It is defined on $\mathcal{N}$, the set of all martingale transforms relative to $\mathcal{A}_0, \mathcal{A}_1, \ldots$.

(b) It can be written in the form

$$Mf = [\sum_{j=1}^{\infty} (\limsup_{n\to\infty} |\sum_{k=1}^{n} a_{jk} d_k|)^2]^{\frac{1}{2}},$$

where $d=(d_1, d_2, \ldots)$ is the difference sequence of f and $(a_{jk}; j\geqslant 1, k\geqslant 1)$ is a matrix that satisfies the following conditions. Each entry a_{jk} is a real $\mathcal{A}_{k-1}$-measurable function and, for all $k\geqslant 1$,

$$c \leqslant \sum_{j=1}^{\infty} a_{jk}^2 \leqslant C. \tag{6.1}$$

Here are a few examples of operators of matrix type:

$$f \to \limsup_{n\to\infty} |f_n|;$$

$$f \to S(f);$$

$$f \to L(f) = [f_1^2 + \sum_{k=1}^{\infty} (f_k - \sigma_k)^2/k]^{\frac{1}{2}},$$

where $\sigma_k = \sum_{j=1}^{k} f_j/k$.

To obtain S, take (a_{jk}) to be the identity matrix. The choice of the matrix is equally obvious for the first example. Clearly,

$$L(f) = [d_1^2 + \sum_{j=2}^{\infty} j^{-3} (\sum_{k=1}^{j} (k-1)\, d_k)^2]^{\frac{1}{2}}$$

and so, for the third example,

$$\begin{aligned} a_{jk} &= 1 \quad \text{if } j=1, k=1, \\ &= j^{-\frac{3}{2}}(k-1) \quad \text{if } 2\leqslant k\leqslant j, \\ &= 0 \quad \text{otherwise,} \end{aligned}$$

and elementary calculations show that (6.1) is satisfied.

If $0=\tau_0<\tau_1<\ldots$ is a sequence of stopping times, then

$$f \to [\sum_{j=1}^{\infty} (\sum_{k=1}^{\infty} I(\tau_{j-1} < k \leqslant \tau_j)\, d_k)^2]^{\frac{1}{2}}$$

is another example of an operator of matrix type.

THEOREM 6.1.

(i) *Let M be an operator of matrix type and $1 < p < \infty$. Then*

$$\alpha c_p \|M^{**}f\|_p \leqslant \|f^*\|_p \leqslant \beta C_p \sup_{1 \leqslant n < \infty} \|M_n f\|_p \tag{6.2}$$

for all martingale transforms f in $\mathfrak{N}$, with $\alpha = C_{(6.1)}^{-\frac{1}{2}}$, $\beta = c_{(6.1)}^{-\frac{1}{2}}$, and the choice of c_p and C_p depending only on p.

(ii) *Suppose that f is in $\mathfrak{N}$ and*

$$\sup_{1 \leqslant n < \infty} \|M_n f\|_1 < \infty. \tag{6.3}$$

Then f converges almost everywhere. Furthermore, if N is any operator of matrix type (N may equal M), then the sequence $\{N_n f, \mathcal{A}_n, n \geqslant 1\}$ is a submartingale converging almost everywhere to Nf (so that $N^{**}f = N^*f$), and

$$\lambda P(N^* f > \lambda) \leqslant c \sup_n \|M_n f\|_1 \tag{6.4}$$

for all $\lambda > 0$. The choice of c depends only on $c_{(6.1)}$ and $C_{(6.1)}$ and the corresponding parameters for N.

COROLLARY 6.1. *Let $1 < p < \infty$. An element $f \in \mathfrak{N}$ is L_p bounded if and only if*

$$\sup_{1 \leqslant n < \infty} \|M_n f\|_p < \infty \tag{6.5}$$

for some operator M of matrix type. If (6.5) holds, then the submartingale $\{M_n f, n \geqslant 1\}$ converges in L_p to Mf and

$$c\|Mf\|_p \leqslant \|f\|_p \leqslant C\|Mf\|_p.$$

This corollary follows immediately from Theorem 6.1 and standard facts about submartingales.

The first assertion of Corollary 6.1 is not true for $p = 1$. Also, condition (6.5) cannot be replaced by the simpler one, $\|Mf\|_p < \infty$, in the second assertion of Corollary 6.1 or in Theorem 6.1 (i). In fact, if $Mf = \limsup |f_n|$, then it can happen that $Mf = 0$ but $\sup_n \|M_n f\|_p = \infty$, $1 \leqslant p < \infty$.

THEOREM 6.2. *Let M be an operator of matrix type. If* A1 *and* A2 *are satisfied, then the two sets $\{M^* f < \infty\}$ and $\{f^* < \infty\}$ are equivalent for all f belonging to $\mathfrak{M}$.*

This theorem adds another local convergence criterion to those of Corollary 5.6: the sequence f converges almost everywhere on a set A if and only if, for some operator M of matrix type, $M^* f$ is finite almost everywhere on A.

Note that $\{Mf<\infty\}$ need not be equivalent to the set where f converges. In fact, under conditions A1 and A2, it can happen that almost everywhere $Mf=0$ but $M^*f=\infty$.

THEOREM 6.3. *Let Φ be as in Section 3, and M an operator of matrix type. If* A1 *and* A2 *are satisfied, then for all f in $\mathcal{M}$, $M^{**}f=M^*f$ and*

$$c\int_\Omega \Phi(M^*f)\leqslant \int_\Omega \Phi(f^*)\leqslant C\int_\Omega \Phi(M^*f).$$

The choice of c and C depends only on $c_{(3.10)}$, $c_{(6.1)}$, $C_{(6.1)}$ and δ, and may be made so that, with the other parameters fixed, the functions $c_{(3.10)}\to C$ and $c_{(3.10)}\to 1/c$ are nondecreasing.

In particular, if $0<p<\infty$, Theorem 6.3 implies that under conditions A1 and A2, $c\|M^*f\|_p\leqslant \|f^*\|_p\leqslant C\|M^*f\|_p$ for all f in $\mathcal{M}$.

Remark 6.1. Let f belong to $\mathcal{N}$. From Theorems 6.1 (ii) and Theorem 6.3, we have $M^{**}f=M^*f$, under either of the following conditions:

(a) For some operator N of matrix type, $\sup_n\|N_nf\|_1<\infty$.

(b) The sequence f belongs to $\mathcal{M}$ and conditions A1 and A2 hold.

However, there are operators M of matrix type such that $M^{**}f>M^*f$ on a set of positive measure for some f in $\mathcal{N}$.

Proof of Theorem 6.1. In the proof of (i) we may assume that

$$\sum_{j=1}^{\infty} a_{jk}^2=1,\quad k\geqslant 1. \tag{6.6}$$

To see that (6.1) may be replaced by this stronger condition, let

$$b_k=[\sum_{j=1}^{\infty} a_{jk}^2]^{-\frac{1}{2}},$$

$$\hat{a}_{jk}=a_{jk}b_k,$$

and $\hat{M}$ be the operator of matrix type corresponding to $(\hat{a}_{jk})$. If f has difference sequence d, let $\hat{f}$ have difference sequence $\hat{d}$ satisfying $\hat{d}_k=b_k^{-1}d_k$. Then $\hat{a}_{jk}\hat{d}_k=a_{jk}d_k$ and $(\hat{a}_{jk})$ satisfies (6.6). If, for example, the left-hand side of (6.2) holds for $\hat{M}$, then, by (2.5),

$$\|M^{**}f\|_p=\|\hat{M}^{**}\hat{f}\|_p\leqslant c\|\hat{f}^*\|_p\leqslant c\|(\sum_{k=1}^{\infty} b_k^{-2}d_k^2)^{\frac{1}{2}}\|_p\leqslant c\alpha^{-1}\|S(f)\|_p\leqslant c\alpha^{-1}\|f^*\|_p,$$

implying that the left-hand side of (6.2) holds for M.

The proof rests on the two-sided inequality (2.5) for the operator S and Khintchine's inequality (5.7). Let $r_1, r_2, \ldots$ be an independent sequence of functions defined on the Lebesgue unit interval such that $r_k = \pm 1$ with equal probability. Recall that, if $\sum_{k=1}^{\infty} a_k^2 < \infty$, then the series $\sum_{k=1}^{\infty} a_k r_k(t)$ converges for almost all t in the unit interval.

Let
$$u_k(t, \omega) = \sum_{j=1}^{\infty} r_j(t)\, a_{jk}(\omega)$$

if the series converges, and zero otherwise. For each t, $u_k(t, \cdot)$ is $\mathcal{A}_{k-1}$-measurable. By (6.6) and Fubini's theorem, the series converges for almost all ω, for almost all t. Let n be a positive integer. Then, by Khintchine's inequality

$$\int_0^1 \left| \sum_{k=1}^{n} u_k d_k \right|^p dt = \int_0^1 \left| \sum_{j=1}^{\infty} r_j(t) \left(\sum_{k=1}^{n} a_{jk} d_k \right) \right|^p dt \leqslant c \left[\sum_{j=1}^{\infty} \left(\sum_{k=1}^{n} a_{jk} d_k \right)^2 \right]^{p/2} = c(M_n f)^p. \quad (6.7)$$

Therefore, by (2.5),

$$\|M_n f\|_p^p \geqslant c \int_0^1 \int_\Omega \left| \sum_{k=1}^{n} u_k d_k \right|^p dP\, dt \geqslant c \int_0^1 \int_\Omega \left[\sum_{k=1}^{n} u_k^2 d_k^2 \right]^{p/2} dP\, dt.$$

If $p \leqslant 2$, then, on the set where $\sum_{k=1}^{n} d_k^2 > 0$,

$$\int_0^1 \left[\sum_{k=1}^{n} u_k^2 d_k^2 \right]^{p/2} dt \geqslant \int_0^1 \left(\sum_{k=1}^{n} d_k^2 \right)^{\frac{p}{2}-1} \sum_{k=1}^{n} \left(u_k^2 \right)^{p/2} d_k^2\, dt$$
$$= \left(\sum_{k=1}^{n} d_k^2 \right)^{\frac{p}{2}-1} \sum_{k=1}^{n} d_k^2 \int_0^1 |u_k|^p dt \geqslant c S_n(f)^p, \quad (6.8)$$

since, by Khintchine's inequality,

$$\int_0^1 |u_k|^p dt \geqslant c \left[\sum_{j=1}^{\infty} a_{jk}^2 \right]^{p/2} = c.$$

If $2 \leqslant p \leqslant \infty$, then

$$\int_0^1 \left[\sum_{k=1}^{n} u_k^2 d_k^2 \right]^{p/2} dt \geqslant \left[\int_0^1 \sum_{k=1}^{n} u_k^2 d_k^2\, dt \right]^{p/2} = \left[\sum_{k=1}^{n} d_k^2 \int_0^1 u_k^2\, dt \right]^{p/2} \geqslant c S_n(f)^p. \quad (6.9)$$

Therefore, in either case,

$$\|M_n f\|_p^p \geqslant c \int_\Omega S_n(f)^p = c \|S_n(f)\|_p^p \geqslant c \|(f^n)^*\|_p^p,$$

implying the right-hand side of (6.2).

To prove the left-hand side, we may suppose that $\|f^*\|_p < \infty$. Then

$$\lim_{n\to\infty} \sum_{k=1}^{n} a_{jk} d_k$$

exists almost everywhere and is finite, $j \geqslant 1$. This follows from (6.6) and Theorem 1 of [1]: A transform of an L_1 bounded martingale converges almost everywhere on the set where its multiplier sequence is bounded. Accordingly, we have that

$$Mf = [\sum_{j=1}^{\infty} (\sum_{k=1}^{\infty} a_{jk} d_k)^2]^{\frac{1}{2}},$$

and, by Fatou's lemma,

$$Mf \leqslant \liminf_{n\to\infty} M_n f \leqslant M^* f,$$

so that $M^{**}f = M^*f$. Hence

$$\|M^{**}f\|_p^p = \int_\Omega \sup_{1\leqslant n<\infty} (M_n f)^p \leqslant c \int_\Omega \sup_{1\leqslant n<\infty} \int_0^1 \left|\sum_{k=1}^{n} u_k d_k\right|^p dt\, dP$$

$$\leqslant c \int_0^1 \int_\Omega \sup_{1\leqslant n<\infty} \left|\sum_{k=1}^{n} u_k d_k\right|^p dP\, dt \leqslant c \int_0^1 \int_\Omega \left[\sum_{k=1}^{\infty} u_k^2 d_k^2\right]^{p/2} dP\, dt$$

from the inequality in the opposite direction to (6.7), (2.2), and the right-hand side of (2.5). If $2 \leqslant p < \infty$, the inequality signs in (6.8) can be reversed. If $p \leqslant 2$, the inequality signs in (6.9) can be reversed. Therefore

$$\int_0^1 \left[\sum_{k=1}^{\infty} u_k^2 d_k^2\right]^{p/2} dt \leqslant c S(f)^p$$

and

$$\|M^{**}f\|_p^p \leqslant c \int_\Omega S(f)^p = c\, \|S(f)\|_p^p \leqslant c\, \|f^*\|_p^p.$$

This completes the proof of (i).(1)

(ii) We prove first a special case of (6.4):

$$\lambda P(S(f) > \lambda) \leqslant c \sup_n \|M_n f\|_1, \quad \lambda > 0. \tag{6.10}$$

Using (6.7) with $p = 1$ and (2.3), we have that

$$\|M_n f\|_1 \geqslant c \int_0^1 \int_\Omega \left|\sum_{k=1}^{n} u_k d_k\right| dP\, dt \geqslant c \int_0^1 \lambda P\left(\left[\sum_{k=1}^{n} u_k^2 d_k^2\right]^{\frac{1}{2}} > \lambda\right) dt$$

$$= c\lambda \int_\Omega \int_0^1 I\left(\left[\sum_{k=1}^{n} u_k^2 d_k^2\right]^{\frac{1}{2}} > \lambda\right) dt\, dP.$$

(1) For the special case $Mf = L(f)$, consult Tsuchikura [16] and Stein [15]. We are indebted to P. W. Millar for the reference [16].

By (6.8) and the inequality in the opposite direction to (6.9), we have that

$$\int_0^1 \left(\sum_{k=1}^n u_k^2 d_k^2\right)^{\frac{1}{2}} dt \geqslant cS_n(f),$$

$$\int_0^1 \left(\sum_{k=1}^n u_k^2 d_k^2\right) dt \leqslant cS_n(f)^2.$$

Therefore, by Lemma 2.3, if $\lambda < \frac{1}{2} cS_n(f)$, then

$$\int_0^1 I\left(\left[\sum_{k=1}^n u_k^2 d_k^2\right]^{\frac{1}{2}} > \lambda\right) dt \geqslant c,$$

which implies that

$$\|M_n f\|_1 \geqslant c\lambda \int_\Omega I(cS_n(f) > \lambda) = c\lambda P(cS_n(f) > \lambda).$$

Clearly, (6.10) follows by maximizing with respect to n.

Note that $$M({}^{n-1}f^n) = [\sum_{j=1}^\infty (a_{jn} d_n)^2]^{\frac{1}{2}} = |d_n| (\sum_{j=1}^\infty a_{jn}^2)^{\frac{1}{2}}.$$

Therefore, by (6.1) and the sublinearity of M,

$$|d_n| \leqslant cM({}^{n-1}f^n) \leqslant c(M_{n-1}f + M_n f),$$

and, by (6.3), d_n is integrable. Since

$$N_n f \leqslant \sum_{k=1}^n N({}^{k-1}f^k) \leqslant c \sum_{k=1}^n |d_k|,$$

this implies that $N_n f$ is integrable. Certainly, $N_n f$ is $\mathcal{A}_n$-measurable, and the submartingale inequality

$$E(N_{n+1}f \mid \mathcal{A}_n) \geqslant N_n f, \quad n \geqslant 1,$$

follows from Jensen's inequality for conditional expectations and the fact that $N_{n+1}f$ is a convex function of d_{n+1}. In particular $\{M_n f, n \geqslant 1\}$ is an L_1 bounded submartingale. (Note that $\{N_n f, n \geqslant 1\}$ need not be L_1 bounded: consider $Mf = \lim\sup |f_n|$, $Nf = S(f)$, and the example on page 1502 of [1].)

Let $\lambda > 0$. There exist martingales X, Y, and Z, in $\mathcal{N}$ such that

$$f = X + Y + Z,$$

$$\|X\|_2^2 \leqslant c\lambda \sup_n \|M_n f\|_1, \tag{6.11}$$

$$\|\sum_{k=1}^{\infty}|y_k|\|_1 \leqslant c \sup_n \|M_n f\|_1, \tag{6.12}$$

$$P(Z^*>0) \leqslant c \sup_n \|M_n f\|_1/\lambda. \tag{6.13}$$

(This is similar but not identical to the decomposition of f exhibited in [8] for the case $M_n f = |f_n|$. For the case $Mf = S(f)$, see [3].) Here x, y, and z are the corresponding difference sequences defined by

$$x_k = d_k I(\tau > k) - E(d_k I(\tau > k) \mid \mathcal{A}_{k-1}),$$

$$y_k = d_k I(\tau = k) - E(d_k I(\tau = k) \mid \mathcal{A}_{k-1}) + E(d_k \mid \mathcal{A}_{k-1}),$$

$$z_k = d_k I(\tau < k),$$

where $\tau = \mu \wedge \nu$ and the stopping times μ and ν are defined by

$$\mu = \inf \{n\colon M_n f > \lambda\},$$

$$\nu = \inf \{n\colon S_n(f) > \lambda\}.$$

(Note that the third term in the expression for y_k vanishes for $k \geqslant 2$.) Clearly, $f = X + Y + Z$ and X, Y, and Z are martingales in $\mathcal{N}$. Also, (6.13) holds since

$$P(Z^*>0) \leqslant P(\tau<\infty) \leqslant P(\mu<\infty) + P(\nu<\infty)$$
$$= P(M^*f>\lambda) + P(S(f)>\lambda) \leqslant c \sup_n \|M_n f\|_1/\lambda. \tag{6.14}$$

Here we have used (6.10) and inequality (2.1), which also holds for submartingales. Since

$$x_k^2 \leqslant 2 d_k^2 I(\tau > k) + 2 E(d_k^2 I(\tau > k) \mid \mathcal{A}_{k-1}),$$

we have that

$$\|X_n\|_2^2 = \int_\Omega \sum_{k=1}^n x_k^2 \leqslant 4 \int_\Omega \sum_{k=1}^n d_k^2 I(\tau > k) \leqslant 4 \int_\Omega S_{\nu-1}(f)^2$$
$$= 8 \int_0^\infty \eta P(S_{\nu-1}(f) > \eta)\, d\eta \leqslant c \int_0^\lambda \sup_n \|M_n f\|_1 \, d\eta = c\lambda \sup_n \|M_n f\|_1.$$

Here we have used (6.10) and the fact that $P(S_{\nu-1}(f) > \eta) = 0$ for $\eta \geqslant \lambda$. Finally, using the inequality $M_{\tau-1} f \leqslant \lambda$, we have that

$$\|\sum_{k=1}^{\infty}|y_k|\|_1 \leqslant 2 \int_\Omega \sum_{k=1}^{\infty} |d_k| I(\tau = k) + \int_\Omega |d_1| \leqslant c \int_{\{\tau<\infty\}} (M_{\tau-1} f + M_\tau f) + c \int_\Omega M_1 f$$
$$\leqslant c\lambda P(\tau < \infty) + c \sup_n \int_{\{\tau<\infty\}} M_{\tau\wedge n} f + c \int_\Omega M_1 f \leqslant c \sup_n \|M_n f\|_1.$$

Since $g = X + Y$ is an L_1 bounded martingale, any transform of g under a uniformly bounded multiplier sequence converges almost everywhere. Since $f = g$ on the set $\{Z = 0\}$, any such transform of f converges almost everywhere on the same set, which, for large λ, has small probability by (6.14). Therefore, any such transform of f converges almost everywhere on Ω. In particular, f converges almost everywhere. Furthermore, each "row" of N,

$$\{\sum_{k=1}^{n} b_{jk} d_k, n \geqslant 1\}$$

converges almost everywhere where (b_{jk}) denotes the matrix corresponding to the operator N. Accordingly, just as in (i) for M, we have $Nf \leqslant N^* f = N^{**} f$.

Using our decomposition of f and inequality (2.1) applied to the submartingale $\{N_n f, n \geqslant 1\}$, we have that

$$\begin{aligned} P(N^{**}f > 3\lambda) &\leqslant P(N^*X > \lambda) + P(N^*Y > \lambda) + P(N^*Z > \lambda) \\ &\leqslant \sup_n \|N_n X\|_2^2/\lambda^2 + \|N^*Y\|_1/\lambda + P(N^*Z > 0) \\ &\leqslant c \sup_n \|X_n\|_2^2/\lambda^2 + c\|\sum_{k=1}^{\infty} |y_k|\|_1/\lambda + P(Z^* > 0) \\ &\leqslant c \sup_n \|M_n f\|_1/\lambda. \end{aligned} \tag{6.4'}$$

This implies that (6.4) holds.

To show that $\{N_n f, n \geqslant 1\}$ converges almost everywhere to Nf, we need to show only that $\{N_n g, n \geqslant 1\}$ converges almost everywhere to Ng, where $g = X + Y$, since $N_n f = N_n g$ and $Nf = Ng$ on the set $\{Z = 0\}$. Since g^* is integrable and g converges almost everywhere, we have that $\|({}^m g)^*\|_1 \to 0$ as $m \to \infty$ by the Lebesgue dominated convergence theorem. We now use the following special case of (6.4'):

$$\eta P(N^{**}g > \eta) \leqslant c\|g\|_1, \quad \eta > 0.$$

This implies that $N^{**}g < \infty$ so that, for $m < n$,

$$|N_n g - Ng| \leqslant N({}^n g) = N({}^m g - {}^m g^n) \leqslant N({}^m g) + N_n({}^m g) \leqslant 2N^{**}({}^m g).$$

Accordingly,

$$P(\sup_{m<n} |N_n g - Ng| > \eta) \leqslant P(2N^{**}({}^m g) > \eta) \leqslant c\|{}^m g\|_1/\eta \leqslant c\|({}^m g)^*\|_1/\eta \to 0 \quad \text{as } m \to \infty$$

This implies the desired convergence and completes the proof of the theorem.

Proof of Theorem 6.2. The argument follows the pattern of the proof of (5.10). In brief, we apply Theorems 3.1 and 4.1 to M^* and use the inequalities

$$P(d^* > \lambda) \leqslant P(cM^*f > \lambda) \tag{6.15}$$

and

$$P(\Delta^* > \lambda) \leqslant P(cf^* > \lambda), \tag{6.16}$$

which hold for all $\lambda > 0$, to prove a global result that implies the assertion of Theorem 6.2.

Inequality (6.15) holds since

$$|d_n| \leqslant c[\sum_{j=1}^{\infty}(a_{jn}d_n)^2]^{\frac{1}{2}} = cM({}^{n-1}f^n) \leqslant c[M_n f + M_{n-1}f] \leqslant cM^*f;$$

inequality (6.16) also holds since

$$|\Delta_n| = M^*({}^{n-1}f^n) \leqslant c\,|d_n| \leqslant cf^*.$$

Proof of Theorem 6.3. Under A1 and A2, $M^{**}f = M^*f$ for all f in $\mathcal{M}$: this follows on the set $\{f^* < \infty\}$ by Lemma 4.1 and on the set $\{f^* = \infty\}$ by Theorem 6.2. We now apply Theorems 3.3 and 4.3. The operator M^* satisfies condition B with $\gamma = 1$; conditions R and L are satisfied with $p_1 = 3/2$, $p_2 = 2$ and $\pi_1 = 3/2$ by Theorem 6.1 (i). Finally, both RΦ and LΦ are satisfied for every Φ by (6.15) and (6.16). Therefore Theorems 3.3 and 4.3 are applicable. This concludes the proof.

7. Application to Brownian motion

To illustrate how our theorems can be useful in the study of continuous parameter martingales, we apply them to obtain new results for Brownian motion. Other applications are possible, for example, to the theory of stochastic integration (see [13]).

Let $X = \{X(t),\ 0 \leqslant t < \infty\}$ be standard Brownian motion: if $n \geqslant 2$ and $0 \leqslant t_0 < \ldots < t_n$, then

$$X(t_1) - X(t_0),\ \ldots,\ X(t_n) - X(t_{n-1})$$

are independent random variables and $X(t_k) - X(t_{k-1})$ is normally distributed with expectation zero and variance

$$\|X(t_k) - X(t_{k-1})\|_2^2 = t_k - t_{k-1},\ 1 \leqslant k \leqslant n.$$

Furthermore, for all $\omega \in \Omega$, the map $t \to X(t, \omega)$ is continuous and $X(0, \omega) = 0$.

Let $\mathcal{B}(t)$ be the smallest σ-field relative to which $X(a)$ is measurable for all $0 \leqslant a \leqslant t$. A stopping time τ of X is a function from Ω into $[0, \infty]$ such that

$$\{\tau < t\} \in \mathcal{B}(t), \quad 0 \leqslant t < \infty.$$

Let X^τ be the process X stopped at τ: $X^\tau(t) = X(\tau \wedge t)$, $0 \leqslant t < \infty$. Its maximal function is defined by

$$(X^\tau)^* = \sup_{0 \leqslant t < \infty} |X^\tau(t)|.$$

THEOREM 7.1. *Let $0<p<\infty$. If τ is a stopping time of X, then*

$$c_p\|\tau^{\frac{1}{2}}\|_p \leqslant \|(X^\tau)^*\|_p \leqslant C_p\|\tau^{\frac{1}{2}}\|_p. \tag{7.1}$$

The choice of c_p and C_p depends only on p and may be made so that the functions $p\to C_p^p$ and $p\to 1/c_p^p$ are nondecreasing.

This is an immediate consequence of the following theorem.

THEOREM 7.2. *Let Φ be as in Section 3. If τ is a stopping time of X, then*

$$c\int_\Omega \Phi(\tau^{\frac{1}{2}}) \leqslant \int_\Omega \Phi[(X^\tau)^*] \leqslant C\int_\Omega \Phi(\tau^{\frac{1}{2}}). \tag{7.2}$$

The choice of c and C depends only on $c_{(3.10)}$ and may be made so that the functions $c_{(3.10)}\to C$ and $c_{(3.10)}\to 1/c$ are nondecreasing.

Proof of Theorem 7.2. We may assume in the proof that $\tau\leqslant b$ for some positive integer b. Otherwise, replace τ by the stopping time $\tau\wedge b$ and note that if $\tau\wedge b$ satisfies (7.2) for all positive integers b, then τ also satisfies (7.2) by the monotone convergence theorem.

For each positive integer j, let

$$\tau_j=\inf\{t>\tau : t=k/2^j \text{ for some } k=1,2,\dots\}$$

and
$$Q_j(t)=\sum_{1\leqslant k\leqslant 2^j t}\left[X\left(\frac{k}{2^j}\right)-X\left(\frac{k-1}{2^j}\right)\right]^2.$$

Then,
$$\tau<\tau_j\leqslant\tau+2^{-j}<b+1. \tag{7.3}$$

If t is a dyadically rational number in $[0, b+1]$, then

$$\lim_{j\to\infty} Q_j(t)=t$$

almost everywhere by a theorem of Lévy [5; Theorem 2.3 of Chapter VIII]. Therefore, since the limit function is continuous on $[0, b+1]$ and each Q_j is nondecreasing, we have that almost everywhere $\lim_j Q_j(t)=t$ uniformly for t in $[0, b+1]$. Accordingly, by (7.3),

$$\lim_{j\to\infty} Q_j(\tau_j)=\tau \tag{7.4}$$

almost everywhere. Furthermore,

$$\int_\Omega \Phi[\sup_j Q_j(\tau_j)^{\frac{1}{2}}]<\infty. \tag{7.5}$$

This holds for $\Phi(a)=a^p$ since $\sup_j Q_j(\tau_j)\leqslant\sup_j Q_j(b+1)$ and

$$(X(b+1)^2, Q_1(b+1), Q_2(b+1), \ldots)$$

is a reversed martingale so that

$$\|\sup_j Q_j(b+1)\|_p \leqslant q\|X(b+1)^2\|_p < \infty, \quad p^{-1}+q^{-1}=1, \quad 1<p<\infty.$$

But this implies that (7.5) holds for general Φ by the inequality

$$\Phi(a) \leqslant ca^{p_0}, \quad a \geqslant 1, \tag{7.6}$$

where $p_0 = \log_2 c_{(3.10)} + 1$ and $c_{(7.6)} = 2^{p_0}\Phi(1)$. Inequality (7.6) is an immediate consequence of (3.11).

We are now almost ready to apply Theorem 5.2. Let j be a positive integer, $\mathcal{A}_{jk} = \mathcal{B}(k/2^j)$, $k \geqslant 0$, and $x_{jk} = 2^{j/2}[X(k/2^j) - X([k-1]/2^j)]$, $k \geqslant 1$. Then $x_j = (x_{j1}, x_{j2}, \ldots)$ is a martingale difference sequence satisfying conditions A1 and A2 with $\delta = (2/\pi)^{\frac{1}{2}}$. Since $2^j\tau_j$ has its values in $\{0, 1, \ldots, \infty\}$ and

$$\{2^j\tau_j \leqslant k\} = \{\tau < k/2^j\} \in \mathcal{A}_{jk}, \quad k \geqslant 0,$$

$2^j\tau_j$ is a stopping time relative to $\mathcal{A}_{j0}, \mathcal{A}_{j1}, \ldots$. Therefore,

$$v_{jk} = I(2^j\tau_j \geqslant k)/2^{j/2}$$

is $\mathcal{A}_{j,k-1}$-measurable and $f_j = (f_{j1}, f_{j2}, \ldots)$ defined by $f_{jn} = \sum_{k=1}^n v_{jk}x_{jk}$ is a martingale transform relative to $\mathcal{A}_{j0}, \mathcal{A}_{j1} \ldots$. Note that $S(f_j) = Q_j(\tau_j)^{\frac{1}{2}}$ and, by uniform continuity,

$$(X^\tau)^* \leqslant \liminf_{j\to\infty} f_j^*.$$

Therefore, by Fatou's lemma, Theorem 5.2, and the Lebesgue dominated convergence theorem,

$$\int_\Omega \Phi[(X^\tau)^*] \leqslant \liminf_{j\to\infty} \int_\Omega \Phi(f_j^*) \leqslant c \liminf_{j\to\infty} \int_\Omega \Phi[S(f_j)]$$

$$= c \liminf_{j\to\infty} \int_\Omega \Phi[Q_j(\tau_j)^{\frac{1}{2}}] = c\int_\Omega \Phi(\tau^{\frac{1}{2}}).$$

In particular, for the constant stopping time $b+1$,

$$\int_\Omega \Phi[(X^{b+1})^*] \leqslant c\Phi[(b+1)^{\frac{1}{2}}] < \infty.$$

Again, by uniform continuity,

$$(X^{b+1})^* \geqslant (X^{\tau+2^{-j}})^* \to (X^\tau)^*$$

almost everywhere as $j \to \infty$. Therefore, by Theorem 5.2 and the Lebesgue dominated convergence theorem,

$$\int_\Omega \Phi(\tau^{\frac{1}{2}}) = \lim_{j\to\infty} \int_\Omega \Phi[Q_j(\tau_j)^{\frac{1}{2}}] \leqslant c \lim_{j\to\infty} \int_\Omega \Phi(f_j^*) \leqslant c \lim_{j\to\infty} \int_\Omega \Phi[(X^{\tau+2^{-j}})^*] = c \int_\Omega \Phi[(X^\tau)^*].$$

This completes the proof of Theorem 7.2.

Remark 7.1. Part of Theorem 7.1 follows immediately from Lemma 2.2. Let $0 < p \leqslant 2$ and note that f_j, defined above, satisfies $s(f_j) = \tau_j^{\frac{1}{2}}$. Therefore, by Lemma 2.2,

$$\|(X^\tau)^*\|_p \leqslant \liminf_{j\to\infty} \|f_j^*\|_p \leqslant c \liminf_{j\to\infty} \|\tau_j^{\frac{1}{2}}\|_p \leqslant c \liminf_{j\to\infty} (\|\tau^{\frac{1}{2}}\|_p + 2^{-j/2}) \leqslant c \|\tau^{\frac{1}{2}}\|_p.$$

COROLLARY 7.1. *If τ is a stopping time of X such that $\tau^{\frac{1}{2}}$ is integrable, then*

$$\int_\Omega X(\tau) = 0.$$

The proof is omitted since it is similar to the proof of Corollary 5.1.

8. Further remarks and examples

This section contains remarks and counterexamples that give additional information about the significance and precision of our assumptions.

Remark 8.1. If f is a martingale transform such that $E(d_k^2 \mid \mathcal{A}_{k-1}) < \infty$, $k \geqslant 1$, then there is a martingale transform $\hat{f}$ defined on (possibly) another probability space such that $\hat{f}_n = \sum_{k=1}^n \hat{v}_k \hat{x}_k$, $n \geqslant 1$, where $\hat{x}$ satisfies A2, and the distribution of $\hat{f}$ is the same as the distribution of f.

It is sufficient to construct $\hat{f}$ in the following special case. There is an independent sequence $r = (r_1, r_2, \ldots)$, independent of $\mathcal{A}_\infty$, such that r_k takes only the values ± 1 with equal probability. In this case, let

$$\hat{x}_k = \begin{cases} d_k/[E(d_k^2 \mid \mathcal{A}_{k-1})]^{\frac{1}{2}} & \text{on } \{E(d_k^2 \mid \mathcal{A}_{k-1}) > 0\}, \\ r_k & \text{on } \{E(d_k^2 \mid \mathcal{A}_{k-1}) = 0\}, \end{cases}$$

$$\hat{v}_k = [E(d_k^2 \mid \mathcal{A}_{k-1})]^{\frac{1}{2}},$$

and $\hat{\mathcal{A}}_k$ be the σ-field generated by $\mathcal{A}_k$ and $(r_1, \ldots, r_k)$. This gives a martingale transform $\hat{f}$ with $\hat{x}$ satisfying A2 and $\hat{f} = f$.

Remark 8.2. Let $x_1, x_2, \ldots$ be an independent sequence of symmetrically distributed random variables and $f = (f_1, f_2, \ldots)$, $f_n = \sum_{k=1}^n x_k$, $n \geqslant 1$. There is a martingale transform $\hat{f}$,

$\hat{f}_n = \sum_{k=1}^n \hat{v}_k \hat{x}_k$, $n \geq 1$, defined on (possibly) another probability space, such that $\hat{f}$ has the same distribution as f. Furthermore, the sequence $\hat{x}$ has the property $|\hat{x}_k| = 1$, $k \geq 1$, and so satisfies condition A.

As in Remark 8.1, it is sufficient to construct $\hat{f}$ for the special case in which a sequence r exists with the same properties as in the above example. In this case, let

$$\hat{x}_k = \begin{cases} x_k/|x_k| & \text{on } \{x_k \neq 0\}, \\ r_k & \text{on } \{x_k = 0\}, \end{cases}$$

$$\hat{v}_k = |x_k|.$$

and $\hat{\mathcal{A}}_{k-1}$ be the σ-field generated by $(x_1, \ldots, x_{k-1}, |x_k|)$ and $(r_1, \ldots, r_{k-1})$. The above assertion is easily checked.

Example 8.1. Without condition A, Theorems 3.3 and 4.3 do not hold. Consider the operator S and the function $\Phi(b) = b^p$, $0 < p < 1$. Neither c_p nor C_p exists such that the inequality

$$c_p \|S(f)\|_p \leq \|f^*\|_p \leq C_p \|S(f)\|_p \tag{8.1}$$

holds for all martingale transforms f.

The following example of Marcinkiewicz and Zygmund [12] shows that the right-hand side of (8.1) fails. Let j be a positive integer and $d_j = (d_{j1}, d_{j2}, \ldots)$ an independent sequence such that

$$P(d_{jk} = 1) = 1 - (j+1)^{-1},$$

$$P(d_{jk} = -j) = (j+1)^{-1}.$$

Let $f_j = (f_{j1}, f_{j2}, \ldots)$ be the martingale defined by $f_{jn} = \sum_{k=1}^n d_{jk}$. By an elementary calculation, we have that

$$\lim_{m \to \infty} \lim_{j \to \infty} \|f_j^m\|_p / \|S(f_j^m)\|_p = \infty.$$

This shows that, for $0 < p < 1$, there is no C_p such that the right-hand side holds for all martingale transforms f.

Also, the left-hand side of (8.1) fails for $0 < p < 1$: consider the transform of each f_j by the multiplier sequence $v = (1, -1, 1, -1, \ldots)$.

Example 8.2. If we assume only condition A2, then the double inequality

$$c_p \|s(f)\|_p \leq \|f^*\|_p \leq C_p \|s(f)\|_p, \quad f \in \mathcal{M},$$

holds for $p = 2$ but fails for any $p \neq 2$. The right-hand side fails, in general, for $p > 2$ since it can happen that $s(f)=\|x_1\|_2=1$ but $\|f^*\|_p=\|x_1\|_p=\infty$. The left-hand side fails for $0<p<2$: Consider the sequences d_j defined in Example 8.1. Let $x_j=(x_{j1}, x_{j2}, \ldots)$ be defined by

$$x_{jk}=d_{jk}/j^{\frac{1}{2}};$$

then each sequence x_j satisfies condition A2. Define stopping times τ_j by

$$\tau_j=\inf\{k: x_{jk}=-j^{\frac{1}{2}}\},$$

and martingale transforms $f_j=(f_{j1}, f_{j2}, \ldots)$ by

$$f_{jn}=\sum_{k=1}^{n} I(\tau_j\geqslant k)\, x_{jk}.$$

Note that $s(f_j^m)=(\tau_j\wedge m)^{\frac{1}{2}}$, and, by an elementary calculation,

$$\lim_{m\to\infty}\lim_{j\to\infty}\|(\tau_j\wedge m)^{\frac{1}{2}}\|_p/\|(f_j^m)^*\|_p=\infty.$$

This shows that, for $0<p<2$, there is no c_p such that the left-hand side holds.

Example 8.3. If, in Theorem 2.1, condition A is replaced by the weaker condition

$$E(|x_k|\,|\mathcal{A}_{k-1})\geqslant c,\quad E(|x_k|^{\varrho}\,|\mathcal{A}_{k-1})\leqslant C, \tag{8.2}$$

for some ϱ, $1<\varrho<2$, then the conclusion of that theorem no longer follows. To see this let

$$x_k=y_k+z_k,\quad k\geqslant 1,$$

where $y_1, y_2, \ldots, z_1, z_2, \ldots$ are independent random variables satisfying

$$P(y_k=k^{(1-\varrho)/\varrho})=1-(k+1)^{-1},$$
$$P(y_k=-k^{1/\varrho})=(k+1)^{-1},$$
$$P(z_k=1)=P(z_k=-1)=\tfrac{1}{2}.$$

Let $\mathcal{A}_0=\{\emptyset,\Omega\}$ and $\mathcal{A}_k$ be the σ-field generated by $x_1, \ldots, x_k$, $k\geqslant 1$. Then $x=(x_1, x_2, \ldots)$ is a martingale difference sequence satisfying condition (8.2). However, for this x the conclusion of Theorem 2.1 does not hold. Let

$$a_k=(-1)^k/k^{(1/\varrho)-\varepsilon},\quad k\geqslant 1,$$

with $\varepsilon>0$ satisfying $(1/\varrho)-\varepsilon>\frac{1}{2}$, so that $\sum_{k=1}^{\infty}a_k^2<\infty$. Now choose $\lambda>1$ satisfying

$$\sum_{k=1}^{\infty}a_k^2<\lambda^2/2,$$

and let $b=2\lambda$. Let $f=g+h$ be the transform of x defined by

$$f_n=\sum_{k=1}^{n} v_k x_k=\sum_{k=1}^{n} v_k y_k+\sum_{k=1}^{n} v_k z_k=g_n+h_n, \quad n\geqslant 1,$$

with

$$v_k=\begin{cases}0 & \text{if} \quad 1\leqslant k\leqslant m,\\ a_k & \text{if} \quad k>m,\end{cases}$$

where m is a positive integer such that $m^{\varepsilon}>2b+1$. Note that $v^*\leqslant 1$. Let $A=\{h^*\leqslant\lambda\}$. Since

$$P(h^*>\lambda)\leqslant\sum_{k=m+1}^{\infty} a_k^2/\lambda^2<\tfrac{1}{2},$$

we have that $P(A)>\frac{1}{2}$. Now consider the stopping times

$$\tau=\inf\{n:|f_n|>b\},$$

$$\nu=\inf\{k>m: y_k=-k^{1/\varrho}\}.$$

By the Borel–Cantelli lemma, $\nu<\infty$ almost everywhere. Moreover, $\tau\leqslant\nu$ since, on the set $\{\nu=k\}$,

$$|d_k|=|v_k x_k|\geqslant|a_k y_k|-|a_k z_k|\geqslant k^{\varepsilon}-1>m^{\varepsilon}-1>2b, \quad k>m.$$

On the set $A, \tau\geqslant\nu$ since, if $\omega\in A$ and $n<\nu(\omega)$, then

$$|f_n(\omega)|\leqslant|g_n(\omega)|+|h_n(\omega)|=\Big|\sum_{k=m+1}^{n}(-1)^k/k^{1-\varepsilon}\Big|+\Big|\sum_{k=m+1}^{n} a_k z_k(\omega)\Big|\leqslant 1+\lambda<2\lambda=b.$$

Therefore, if $0<p\leqslant\varrho$, then

$$2\|(f^{\tau})^*\|_p\geqslant\left[\int_A|d_\nu|^p\right]^{1/p}\geqslant(m^{\varepsilon}-1)2^{-1/p}.$$

Accordingly, $\|(f^{\tau})^*\|_p/b$ is large for large m; for the above x, the conclusion of Theorem 2.1 does not hold.

Remark 8.3. By the proof of Theorem 5.4, we know that if conditions A1 and A2 are satisfied, then, for all $f\in\mathcal{M}$ and $\lambda>0$,

$$\lambda P(s(f)>\lambda)\leqslant c\|f\|_1.$$

From this follows a more general inequality: if A holds and T is an operator satisfying B and R2 with $p_2>1$, then

$$\lambda P(Tf>\lambda)\leqslant c\|f\|_1 \tag{8.3}$$

for all $f \in \mathcal{M}$ and $\lambda > 0$. To see this, let $\lambda > 0$, $\tau = \inf\{n \geq 0: s_{n+1}(f) > \lambda\}$, and $f = f^\tau + {}^\tau f = g + h$. Then $s(g) \leq \lambda$, $s(g) \leq s(f)$, and, by Theorem 5.3,

$$\|g^*\|_{p_2}^{p_2} \leq c\,\|s(g)\|_{p_2}^{p_2} = c\int_0^\lambda \eta^{p_2-1} P(s(g) > \eta)\,d\eta \leq c\int_0^\lambda \eta^{p_2-2}\,\|f\|_1\,d\eta = c\lambda^{p_2-1}\,\|f\|_1.$$

Also,
$$P(s(h) > 0) \leq P(\tau < \infty) = P(s(f) > \lambda) \leq c\|f\|_1/\lambda,$$

so that
$$P(Tf > 2\gamma\lambda) \leq P(Tg > \lambda) + P(Th > \lambda) \leq c\|g^*\|_{p_2}^{p_2}/\lambda^{p_2} + P(s(h) > 0) \leq c\|f\|_1/\lambda.$$

Therefore, to obtain the results of Section 4, we do not have to assume condition L1 explicitly and the L1 part of condition L can be eliminated: if T is an operator satisfying R2 with $p_2 < \varrho$, then L1 is satisfied with $\pi_1 = p_2$; if A holds and T is an operator satisfying B and R2 with $p_2 = \varrho$, then, by (8.3), L1 is satisfied with $\pi_1 = 1$.

A companion result to (8.3) is the following. Let $0 < p < p_2$. If A holds and T is an operator satisfying B and R2, then

$$\|Tf\|_p \leq c\|f^*\|_p$$

for all $f \in \mathcal{M}$, with the choice of c depending only on p and the parameters of A, B, and R2. The proof is similar to that of Lemma 2.2 and is omitted.

References

[1]. Burkholder, D. L., Martingale transforms. *Ann. Math. Statist.*, 37 (1966), 1494–1504.

[2]. Calderón, A. P. & Zygmund, A., A note on the interpolation of sublinear operations. *Amer. J. Math.*, 78 (1956), 282–288.

[3]. Chow, Y. S., Convergence of sums of squares of martingale differences. *Ann. Math. Statist.*, 39 (1968), 123–133.

[4]. Davis, B., Divergence properties of some martingale transforms. *Ann. Math. Statist.*, 40 (1969), 1852–1854.

[5]. Doob, J. L., *Stochastic processes.* New York, 1953.

[6]. Dvoretzky, A., Personal communication.

[7]. Gundy, R. F., The martingale version of a theorem of Marcinkiewicz and Zygmund. *Ann. Math. Statist.*, 38 (1967), 725–734.

[8]. —— A decomposition for L^1-bounded martingales. *Ann. Math. Statist.*, 39 (1968), 134–138.

[9]. —— On the class $L \log L$, martingales, and singular integrals. *Studia Math.*, 33 (1969), 109–118.

[10]. Hardy, G. H., and Littlewood, J. E., A maximal theorem with function-theoretic applications. *Acta Math.*, 54 (1930), 81–116.

[11]. Marcinkiewicz, J., Quelques théorèmes sur les séries orthogonales. *Ann. Soc. Polon. Math.*, 16 (1937), 84–96 (pages 307–318 of the *Collected Papers*.)

[12]. Marcinkiewicz, J. & Zygmund, A., Quelques théorèmes sur les fonctions indépendantes. *Studia Math.*, 7 (1938), 104–120 (pages 374–388 of the *Collected Papers* of Marcinkiewicz).

[13]. MILLAR, P. W., Martingale integrals. *Trans. Amer. Math. Soc.*, 133 (1968), 145–166.

[14]. PALEY, R. E. A. C., A remarkable series of orthogonal functions. *Proc. London Math. Soc.*, 34 (1932), 241–279.

[15]. STEIN, E. M., *Topics in harmonic analysis related to the Littlewood-Paley theory*. Ann. Math. Studies 63. Princeton, 1970.

[16]. TSUCHIKURA, T., Sample properties of martingales and their arithmetic means. *Tôhoku Math. J.*, 20 (1968), 400–415.

[17]. ZYGMUND, A., *Trigonometric series*, 2nd edition. Cambridge, 1959.

Received September 8, 1969