INTEGRAL INEQUALITIES FOR CONVEX FUNCTIONS OF OPERATORS ON MARTINGALES

D. L. BURKHOLDER
UNIVERSITY OF ILLINOIS
and
B. J. DAVIS and R. F. GUNDY
RUTGERS UNIVERSITY

1. Introduction

Let $\mathscr{M}$ be a family of martingales on a probability space $(\Omega, \mathscr{A}, P)$ and let Φ be a nonnegative function on $[0, \infty]$. The general question underlying both [2] and the present work may be stated as follows: *If U and V are operators on $\mathscr{M}$ with values in the set of nonnegative $\mathscr{A}$ measurable functions on Ω, under what further conditions does*

$$\lambda^{p_0} P(Vf > \lambda) \leqq c\|Uf\|_{p_0}^{p_0}, \qquad \lambda > 0, f \in \mathscr{M}, \tag{1.1}$$

imply $E\Phi(Vf) \leqq cE\Phi(Uf), f \in \mathscr{M}$? Here E denotes expectation, integration over Ω with respect to P, and the letter c denotes a positive real number, not necessarily the same number from line to line. In most applications, the first inequality can be proved easily for only one particular value of p_0, usually for $p_0 = 2$, although it is the second inequality that is really needed. Therefore, it is important to know conditions under which the second follows from the first.

In [2], the function Φ may be any nondecreasing function that satisfies a mild growth condition. The above question is then answered by suitably restricting the martingale f. In this paper, Φ is restricted to be convex, but no conditions are placed on the martingale f.

We state our main results in Section 2. Here, we mention one special but important application. If $f = (f_1, f_2, \cdots)$ is a martingale, we write

$$\begin{aligned} f_n &= \sum_{k=1}^{n} d_k, \qquad n \geqq 1,\\ f^* &= \sup_n |f_n|, \\ S(f) &= \left(\sum_{k=1}^{\infty} d_k^2\right)^{1/2} \end{aligned} \tag{1.2}$$

The maximal function f^* and the square function $S(f)$ are closely linked.

This research was supported in part by the National Science Foundation under Grants GP-23835 and GP-19222.

THEOREM 1.1. *Suppose that Φ is a convex function from $[0, \infty)$ into $[0, \infty)$ satisfying $\Phi(0) = 0$ and the growth condition*

$$\Phi(2\lambda) \leqq c\Phi(\lambda), \qquad \lambda > 0; \tag{1.3}$$

set $\Phi(\infty) = \lim_{\lambda\to\infty} \Phi(\lambda)$. Then

$$cE\Phi(S(f)) \leqq E\Phi(f^*) \leqq CE\Phi(S(f)) \tag{1.4}$$

for all martingales f. The choice of c and C depends only on $c_{(1.3)}$, the growth parameter of Φ.

This result is already known for convex powers $\Phi(\lambda) = \lambda^p$; see [1], for the case where $1 < p < \infty$ and [3] for $p = 1$. (Also, the inequality holds for $0 < p < 1$ if f is restricted; see [2].) However, Theorem 1.1 gives new information about the quadratic variation of right continuous martingales $X = \{X(t), 0 \leqq t \leqq 1\}$. Defining S_j^2 to be the usual approximation to the quadratic variation of X, we show that $\{S_j\}$ converges in L_1 if and only if the maximal function of X is integrable (see Theorem 5.1).

Condition (1.3) is a necessary condition for (1.4) to hold for all martingales f on a nonatomic probability space (see Remark 6.2).

In this paper, we use some of the methods developed in [2] together with the decomposition of martingales introduced in [3]. However, this paper may be read independently, or nearly independently, of [2] and [3]. One new tool of special interest is Theorem 3.2, which may be stated as follows. Let $z_1, z_2, \cdots$ be a sequence of nonnegative $\mathscr{A}$ measurable functions, $\mathscr{B}_1, \mathscr{B}_2, \cdots$ a monotone sequence of sub-σ-fields of $\mathscr{A}$, and Φ a convex function as in Theorem 1.1. Then

$$E\Phi\left(\sum_{k=1}^{\infty} E(z_k \mid \mathscr{B}_k)\right) \leqq cE\Phi\left(\sum_{k=1}^{\infty} z_k\right), \tag{1.5}$$

and the choice of c depends only on $c_{(1.3)}$.

2. Main results

Let $(\Omega, \mathscr{A}, P)$ be a probability space and let $\mathscr{A}_0, \mathscr{A}_1, \cdots$ be a nondecreasing sequence of sub-σ-fields of $\mathscr{A}$. Let $\mathscr{M}$ be the set of all martingales $f = (f_1, f_2, \cdots)$ relative to $\mathscr{A}_1, \mathscr{A}_2, \cdots$. We consider operators T defined on $\mathscr{M}$ with values in the set of nonnegative $\mathscr{A}$ measurable functions. Examples of such operators are

$$\begin{aligned} f^* &= \sup_{1 \leqq n < \infty} |f_n|, \\ S(f) &= \left[\sum_{k=1}^{\infty} d_k^2\right]^{1/2}, \\ s(f) &= \left[\sum_{k=1}^{\infty} E(d_k^2 \mid \mathscr{A}_{k-1})\right]^{1/2}, \end{aligned} \tag{2.1}$$

where $d = (d_1, d_2, \cdots)$ is the difference sequence of f:

$$f_n = \sum_{k=1}^{n} d_k, \qquad n \geqq 1. \tag{2.2}$$

Other examples are discussed below.

An operator T is *quasilinear* if

$$T(f + g) \leqq \gamma(Tf + Tg) \tag{2.3}$$

for some real number $\gamma \geqq 1$ and all f and g in $\mathscr{M}$. An operator T is *symmetric* if $T(-f) = Tf$ and is *local* if $Tf = 0$ on the set $\{s(f) = 0\}$, $f \in \mathscr{M}$. For example, the operators S, s, and $f \to f^*$ satisfy these conditions.

A stopping time τ is a function from Ω into $\{0, 1, \cdots, \infty\}$ such that $\{\tau \leqq n\} \in \mathscr{A}_n$, $n \geqq 0$. If μ and ν are stopping times, let ${}^{\mu}f^{\nu}$ denote the sequence

$$\sum_{k=1}^{n} I(\mu < k \leqq \nu)\, d_k, \qquad n \geqq 1, \tag{2.4}$$

where $I(A)$ is the indicator function of the set A. If f is in $\mathscr{M}$, then ${}^{\mu}f^{\nu}$, f started at μ and stopped at ν, is also in $\mathscr{M}$. Write f^{ν} for ${}^{0}f^{\nu}$; in particular, f^n is the martingale f stopped at n.

If T is an operator, let $T_n f = Tf^n$, $0 \leqq n \leqq \infty$,

$$\begin{aligned} T^*f &= \sup_{1 \leqq n < \infty} T_n f, \\ T^{**}f &= T^*f \vee Tf. \end{aligned} \tag{2.5}$$

The operator T is *measurable* if $T_n f$ is A_n measurable for $n \geqq 1$, $f \in \mathscr{M}$. For example, $S_n(f) = (\sum_{k=1}^{n} d_k^2)^{1/2}$ so that $S = S^* = S^{**}$ and S is measurable.

We are now ready to state some of our main results. In the following theorem $\Delta^* = \sup_{1 \leqq n < \infty} \Delta_n$, where $\Delta_n = T({}^{n-1}f^n)$. Throughout the paper, if $f = (f_1, f_2, \cdots)$ is *any* sequence of functions, then f^* denotes the maximal function of f.

THEOREM 2.1. *Let $0 < p_0 < \infty$. Suppose that T is a local, quasilinear, symmetric, and measurable operator on $\mathscr{M}$ such that*

$$\lambda^{p_0} P(Tf > \lambda) \leqq c \|f^*\|_{p_0}^{p_0} \tag{2.6}$$

for all $\lambda > 0$ and f in $\mathscr{M}$. If Φ is a convex function as in Theorem 1.1, then

$$E\Phi(T^{**}f) \leqq cE\Phi(f^*) \tag{2.7}$$

for all f in $\mathscr{M}$ provided that

$$E\Phi(\Delta^*) \leqq cE\Phi(f^*), \qquad f \in \mathscr{M}, \tag{2.8}$$

$$E\Phi(Tf) \leqq cE\Phi\left(\sum_{k=1}^{\infty} |d_k|\right), \qquad f \in \mathscr{M}. \tag{2.9}$$

The choice of $c_{(2.7)}$ depends only on the parameters of the assumptions, that is, on the quasilinearity constant γ, p_0, $c_{(1.3)}$, $c_{(2.6)}$, $c_{(2.8)}$ and $c_{(2.9)}$.

REMARK 2.1. Conditions (2.8) and (2.9) are necessary for (2.7). This follows from $\Delta^* \leqq 2\gamma T^{**}f$, from $Tf \leqq T^{**}f$, from $f^* \leqq \Sigma_{k=1}^{\infty}|d_k|$, from (1.3), and from the fact that Φ is nondecreasing. To prove the first inequality, notice that

$$\Delta_n = T({}^{n-1}f^n) = T(f^n - f^{n-1}) \leqq \gamma(Tf^n + Tf^{n-1}) \leqq 2\gamma T^*f. \tag{2.10}$$

THEOREM 2.2. *Let $0 < p_0 < \infty$. Suppose that T is a local, quasilinear, symmetric, and measurable operator on $\mathscr{M}$ such that*

$$\lambda^{p_0}P(f^* > \lambda) \leqq c\|T^*f\|_{p_0}^{p_0} \tag{2.11}$$

for all $\lambda > 0$ and f in $\mathscr{M}$. If Φ is a convex function as in Theorem 1.1, then

$$E\Phi(f^*) \leqq cE\Phi(T^*f) \tag{2.12}$$

for all f in $\mathscr{M}$ provided that

$$E\Phi(d^*) \leqq cE\Phi(T^*f), \qquad f \in \mathscr{M}, \tag{2.13}$$

$$Tf \leqq c\sum_{k=1}^{\infty}|d_k|, \qquad f \in \mathscr{M}. \tag{2.14}$$

The choice of $c_{(2.12)}$ depends only on γ, p_0, $c_{(1.3)}$, $c_{(2.11)}$, $c_{(2.13)}$, and $c_{(2.14)}$.

Since $d^* \leqq S(f) \leqq \Sigma_{k=1}^{\infty}|d_k|$, the operator S satisfies (2.13) and (2.14) and it is elementary to check that S satisfies all of the conditions of Theorems 2.1 and 2.2 with $p_0 = 2$. Therefore, Theorem 1.1 follows from inequalities (2.7) and (2.12). More generally, consider any operator M of matrix type:

$$Mf = \left[\sum_{j=1}^{\infty}\left(\limsup_{n\to\infty}\Big|\sum_{k=1}^{n} a_{j,k}d_k\Big|\right)^2\right]^{1/2}. \tag{2.15}$$

Here $a_{j,k}$ is an $\mathscr{A}_{k-1}$ measurable function and

$$c \leqq \sum_{j=1}^{\infty} a_{j,k}^2 \leqq C, \qquad k \geqq 1. \tag{2.16}$$

THEOREM 2.3. *If M is an operator of matrix type and Φ is a convex function as in Theorem 1.1, then*

$$cE\Phi(M^{**}f) \leqq E\Phi(f^*) \leqq CE\Phi(M^*f) \tag{2.17}$$

for all f in $\mathscr{M}$. The choice of c and C depends only on $c_{(1.3)}$, $c_{(2.16)}$, and $C_{(2.16)}$.

This generalizes the first part of Theorem 6.1 in [2], which should be consulted for further discussion and examples.

PROOF. We prove the left side first, and, indeed, a little more. Letting

$$M^{***}f = \left[\sum_{j=1}^{\infty}\left(\sup_{1\leqq n<\infty}\Big|\sum_{k=1}^{n} a_{j,k}d_k\Big|\right)^2\right]^{1/2}, \tag{2.18}$$

we have that

$$E\Phi(M^{***}f) \leqq cE\Phi(f^*). \tag{2.19}$$

Only the right side of (2.16) is needed for (2.19). Let $T = M^{***}$; then T is local, sublinear ($\gamma = 1$), symmetric, and measurable. Condition (2.9) holds; in fact, the stronger condition (2.14) holds, since

$$(2.20)\qquad \begin{aligned}(Tf)^2 &\leqq \sum_{j=1}^{\infty}\left(\sum_{k=1}^{\infty}|a_{j,k}\,d_k|\right)^2\\ &\leqq \sum_{j=1}^{\infty}\left(\sum_{k=1}^{\infty}a_{j,k}^2|d_k|\right)\left(\sum_{k=1}^{\infty}|d_k|\right) \leqq C\left(\sum_{k=1}^{\infty}|d_k|\right)^2\end{aligned}$$

by the right side of (2.16). Therefore, (2.8) holds also, since $\Delta_n = T({}^{n-1}f^n) \leqq c|d_n| \leqq cf^*$ by (2.14). Finally, using the fact that $\{\sum_{k=1}^n a_{j,k}\,d_k,\ n \geqq 1\}$ is a martingale, we have that

$$(2.21)\qquad \begin{aligned}\|Tf\|_2^2 &= \sum_{j=1}^{\infty} E\left(\sup_n \left|\sum_{k=1}^{n} a_{j,k}\,d_k\right|\right)^2 \leqq 4\sum_{j=1}^{\infty}\sup_n E\left(\sum_{k=1}^{n} a_{j,k}\,d_k\right)^2\\ &= 4\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} Ea_{j,k}^2\,d_k^2 \leqq C\|S(f)\|_2^2 \leqq C\|f^*\|_2^2.\end{aligned}$$

This implies (2.6) for $p_0 = 2$, and (2.19) follows from Theorem 2.1.

Turning to the proof of the right side of (2.17), we note that M satisfies (2.14), since $Mf \leqq M^{***}f$. By the left side of (2.16),

$$(2.22)\qquad c|d_n| \leqq \left(\sum_{j=1}^{\infty}|a_{j,n}\,d_n|^2\right)^{1/2} = M(f^n - f^{n-1}) \leqq 2M^*f$$

so that $d^* \leqq cM^*f$ and (2.13) is satisfied. Also,

$$(2.23)\qquad \begin{aligned}\|M_n f\|_2^2 &= E\sum_{j=1}^{\infty}\left(\sum_{k=1}^{n} a_{j,k}\,d_k\right)^2\\ &= \sum_{j=1}^{\infty}\sum_{k=1}^{n} Ea_{j,k}^2\,d_k^2 \geqq c\|S_n(f)\|_2^2 = c\|f_n\|_2^2,\end{aligned}$$

so that

$$(2.24)\qquad \lambda^2 P(f^* > \lambda) \leqq \sup_n \|f_n\|_2^2 \leqq c\|M^*f\|_2^2,$$

which is (2.11) with $p_0 = 2$. The right side of (2.17) now follows from Theorem 2.2.

EXAMPLE 2.1. Here is an operator satisfying the conditions of Theorem 2.1, but not all the conditions of Theorem 2.2:

$$(2.25)\qquad f \to \left[\sum_{k=1}^{\infty}[E(|d_k|\,|\,\mathscr{A}_{k-1})]^2\right]^{1/2}.$$

In fact, the conclusion of Theorem 2.2 need not hold. These statements are proved in Section 6, where an application of this operator to the theory of random walk is presented.

3. Preliminary results

The next section contains the proofs of Theorems 2.1 and 2.2. Here we develop the necessary tools, some of which are interesting in their own right.

THEOREM 3.1. *Let $0 < p_0 < \infty$. Suppose that U and V are local, quasilinear, symmetric, and measurable operators on $\mathscr{M}$ such that*

$$\lambda^{p_0} P(Vf > \lambda) \leqq c\|U^*f\|_{p_0}^{p_0} \tag{3.1}$$

for all $\lambda > 0$ and f in $\mathscr{M}$. Let Φ be a convex function as in Theorem 1.1 and f a martingale in $\mathscr{M}$. Suppose that w_n is an $\mathscr{A}_{n-1}$ measurable function satisfying

$$U({}^{n-1}f^n) \leqq w_n \tag{3.2}$$

and $\Delta_n = V({}^{n-1}f^n)$, $n \geqq 1$.

Then

$$E\Phi(V^*f) \leqq cE\Phi(U^*f) + cE\Phi(\Delta^*) + cE\Phi(w^*) \tag{3.3}$$

with the choice of c depending only on γ_U and γ_V, the quasilinearity constants of U and V, respectively, and on $c_{(1.3)}$, $c_{(3.1)}$, and p_0.

Actually, much less than the convexity of Φ is needed for this theorem; the precise condition is described below. We need the following lemma. Let $\gamma = \gamma_U \vee \gamma_V$.

LEMMA 3.1. *Suppose that U and V satisfy*

$$Uf \leqq Vf, \qquad f \in \mathscr{M}, \tag{3.4}$$

in addition to the conditions of Theorem 3.1. Let $\alpha \geqq 1$ and $\beta > \gamma^6$. Then

$$P(V^*f > \lambda) \leqq cP(cU^*f > \lambda) + cP(c\Delta^* > \lambda) + cP(w^* > \lambda) \tag{3.5}$$

for all $\lambda > 0$ satisfying

$$P(V^*f > \lambda) \leqq \alpha P(V^*f > \beta\lambda). \tag{3.6}$$

The choise of $c_{(3.5)}$ depends only on α, β, γ, p_0, and $c_{(3.1)}$.

The proof is similar to, but simpler than, the proof of Theorem 4.1 in [2]. We need two elementary facts. First, if T is a local, quasilinear, and symmetric operator on $\mathscr{M}$ and τ is a stopping time, let $T_\tau f = T_n f$ on the set $\{\tau = n\}$, $0 \leqq n \leqq \infty$. Then

$$\gamma^{-1} T_\tau f \leqq Tf^\tau \leqq \gamma T_\tau f. \tag{3.7}$$

This is Lemma 2.1 in [2]. Second, if h is a nonnegative $\mathscr{A}$ measurable function, $A \in \mathscr{A}$, and a and b are positive numbers such that

$$\int_A h \geqq 2aP(A) \quad \text{and} \quad \int_A h^2 \leqq b^2 P(A), \tag{3.8}$$

then

$$P(A) \leqq \left(\frac{b}{a}\right)^2 P(h > a). \tag{3.9}$$

See Chapter 5, Section 8.26 of Zygmund [5]. In our application, $A = \{V^*f > \lambda\}$.

PROOF. We first examine the case in which $U({}^{n-1}f^n)$ is uniformly bounded.

(a) *Let* $\alpha \geqq 1$ *and* $\beta > \gamma^4$. *Suppose that* (3.6) *and*

$$U({}^{n-1}f^n) \leqq \lambda, \qquad n \geqq 1, \tag{3.10}$$

are satisfied. Then

$$P(V^*f > \lambda) \leqq cP(cU^*f > \lambda) + cP(c\Delta^* > \lambda). \tag{3.11}$$

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

$$P(V^*f > \lambda) \leqq 2\alpha P(V^*f > \beta\lambda, \Delta^* \leqq \theta\lambda) \tag{3.12}$$

or

$$P(V^*f > \lambda) \leqq 2\alpha P(\Delta^* > \theta\lambda). \tag{3.13}$$

The latter possibility leads directly to (3.11); assume the former. Let

$$\begin{aligned} \mu &= \inf\{n: V_n f > \lambda\}, \\ \nu &= \inf\{n: V_n f > \beta\lambda\}, \end{aligned} \tag{3.14}$$

and $g = {}^{\mu}f^{\nu}$. Since V is a measurable operator, μ and ν are stopping times. If n is a positive integer, then

$$V_n f = V(f^{n-1} + {}^{n-1}f^n) \leqq \gamma(V_{n-1}f + \Delta_n), \tag{3.15}$$

so that $V_\mu f \leqq \gamma(\lambda + \Delta^*)$ on the set where μ is finite. On $\{V^*f > \beta\lambda, \Delta^* \leqq \theta\lambda\}$, $\mu \leqq \nu < \infty$,

$$\beta\lambda < V_\nu f \leqq \gamma V(f^\mu + g) \leqq \gamma^3(V_\mu f + Vg) \leqq \gamma^4(\lambda + \theta\lambda + Vg), \tag{3.16}$$

and so

$$Vg > (\beta\gamma^{-4} - 1 - \theta)\lambda = \theta\lambda. \tag{3.17}$$

Therefore, by (3.12),

$$P(Vg > \theta\lambda) \geqq P(V^*f > \beta\lambda, \Delta^* \leqq \theta\lambda) \geqq cP(V^*f > \lambda). \tag{3.18}$$

Since U is a local operator,

$$\{V^*f \leqq \lambda\} = \{\mu = \infty\} \subset \{s(g) = 0\} \subset \{U^*g = 0\}, \tag{3.19}$$

so that, by (3.1), we have the lower estimate

$$\begin{aligned} \int_{\{V^*f > \lambda\}} (U^*g)^{p_0} &= \|U^*g\|_{p_0}^{p_0} \\ &\geqq c(\theta\lambda)^{p_0}P(Vg > \theta\lambda) \geqq c\lambda^{p_0}P(V^*f > \lambda). \end{aligned} \tag{3.20}$$

To obtain an upper estimate, consider U^*g. Since $\mu \leqq \nu$, we have $U^*g = U^*(f^\nu - f^\mu) \leqq 2\gamma^2 U_\nu^* f$. On $\{\nu = \infty\}$, by (3.4) and the definition of ν, $U_\nu^* f \leqq V_\nu^* f \leqq \beta\lambda$. On $\{\nu = n\}$, n a positive integer,

$$U_\nu^* f = U_n^* f = U^*(f^{n-1} + {}^{n-1}f^n), \tag{3.21}$$
$$U_{n-1}^* f \leqq V_{n-1}^* f \leqq \beta\lambda;$$

by (3.10), $U^*({}^{n-1}f^n) = U({}^{n-1}f^n) \leqq \lambda$, so that $U_\nu^* f \leqq \gamma(\beta\lambda + \lambda)$. Therefore, $U^* g \leqq 2\gamma^3(\beta + 1)\lambda$, and we have the upper estimate

$$\int_{\{V^*f > \lambda\}} (U^*g)^{2p_0} = \|U^*g\|_{2p_0}^{2p_0} \tag{3.22}$$
$$\leqq c\lambda^{2p_0} P(V^*f > \lambda).$$

Applying (3.9) with $A = \{V^*f > \lambda\}$ and $h = (U^*g)^{p_0}$, we obtain

$$P(V^*f > \lambda) \leqq cP([U^*g]^{p_0} > c\lambda^{p_0}). \tag{3.23}$$

Since $U^*g \leqq 2\gamma^2 U^*f$, inequality (3.11) follows.

(b) We now complete the proof of Lemma 3.1 by reducing the general case to that considered in (a). It follows from (3.6) that either

$$P(V^*f > \lambda) \leqq 2\alpha P(V^*f > \beta\lambda, w^* \leqq \lambda) \tag{3.24}$$

or

$$P(V^*f > \lambda) \leqq 2\alpha P(w^* > \lambda). \tag{3.25}$$

The latter possibility implies (3.5). Assume the former and let $h = f^\sigma$, where σ is the stopping time

$$\sigma = \inf\{n \geqq 0: w_{n+1} > \lambda\}. \tag{3.26}$$

Note that, by (3.2),

$$U({}^{n-1}h^n) = U({}^{n-1}f^{n\wedge\sigma}) \leqq \gamma U_\sigma({}^{n-1}f^n) \leqq \gamma w_\sigma \leqq \gamma\lambda, \tag{3.27}$$

so that h satisfies (3.10) with λ replaced by $\lambda_0 = \gamma\lambda$. We now show that h satisfies (3.6) with λ replaced by λ_0, α by 2α, and β by $\beta_0 = \beta\gamma^{-2}$. Since $V^*h \leqq \gamma V_\sigma^* f \leqq \gamma V^* f$,

$$\begin{aligned} P(V^*h > \lambda_0) &\leqq P(V^*f > \lambda) \\ &\leqq 2\alpha P(V^*f > \beta\lambda, w^* \leqq \lambda) \\ &= 2\alpha P(V_\sigma^* f > \beta\lambda, \sigma = \infty) \\ &\leqq 2\alpha P(\gamma V^*h > \beta\lambda) \\ &= 2\alpha P(V^*h > \beta_0\lambda_0). \end{aligned} \tag{3.28}$$

Therefore, by (a), $P(V^*h > \lambda_0) \leqq cP(cU^*h > \lambda_0) + cP(c\Delta_0^* > \lambda_0)$. Here, $\Delta_{0,n} = V({}^{n-1}h^n) \leqq \gamma\Delta_n$ and $U^*h \leqq \gamma U^*f$ so that

$$\begin{aligned} P(V^*f > \lambda) &\leqq 2\alpha P(V^*h > \beta_0\lambda_0) \leqq 2\alpha P(V^*h > \lambda_0) \\ &\leqq cP(cU^*f > \lambda) + cP(c\Delta^* > \lambda). \end{aligned} \tag{3.29}$$

This completes the proof of Lemma 3.1.

LEMMA 3.2. *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 $B = \{t: \psi(t) < r\psi(t+1)\}$ for a real number $r > 1$, then*

$$\int_{-\infty}^{\infty} \psi(t)\,dt \leqq \frac{r}{r-1}\int_{B} \psi(t+1)\,dt. \tag{3.30}$$

The proof of this is straightforward; see [2].

Now consider a convex function Φ from $[0, \infty)$ into $[0, \infty)$ with $\Phi(0) = 0$. As usual, set $\Phi(\infty) = \lim_{\lambda\to\infty} \Phi(\lambda)$. It follows that there is a nonnegative nondecreasing function φ on $(0, \infty)$ such that

$$\Phi(b) = \int_0^b \varphi(\lambda)\,d\lambda, \qquad 0 \leqq b \leqq \infty. \tag{3.31}$$

(See Zygmund [5], Chapter I, Section 10.11.) Furthermore, if Φ satisfies the growth condition (1.3), then φ satisfies

$$\varphi(2\lambda) \leqq c\varphi(\lambda), \qquad \lambda > 0. \tag{3.32}$$

This follows from

$$\begin{aligned} 2\lambda\varphi(2\lambda) &\leqq \int_{2\lambda}^{4\lambda} \varphi(t)\,dt \leqq \Phi(4\lambda) \\ &\leqq c^2\Phi(\lambda) = c^2 \int_0^{\lambda} \varphi(t)\,dt \leqq c^2\lambda\varphi(\lambda). \end{aligned} \tag{3.33}$$

The converse also holds [2]. Since Φ is nondecreasing,

$$\Phi\left(\frac{1}{2}[a+b]\right) \leqq \Phi(a) + \Phi(b). \tag{3.34}$$

Therefore, the growth condition (1.3) implies that

$$\Phi(a+b) \leqq c[\Phi(a) + \Phi(b)] \tag{3.35}$$

for all a and b in $[0, \infty]$.

PROOF OF THEOREM 3.1. In the proof we may assume that relation (3.4) holds. For if (3.4) does not hold, replace U by U^* and Vf by $U^*f \vee Vf$. Then the new pair of operators satisfies both (3.4) and the conditions of Theorem 3.1. Furthermore, by (3.35), the inequality (3.3) for the new pair implies (3.3) for the original pair.

The proof is now similar to the proof of Theorem 3.2 in [2]. Let k be the least positive integer j such that $2^j > \gamma^6$, $\beta = 2^k$, and $b = \log \beta$. Let

$$\begin{aligned} \psi(t) &= b\beta^t\varphi(\beta^t)P(V^*f > \beta^t), \\ B &= \{t: \psi(t) < 2\psi(t+1)\}, \end{aligned} \tag{3.36}$$

and notice that

$$\int_{-\infty}^{0} \psi(t)\,dt = \int_0^1 \varphi(\lambda)P(V^*f > \lambda)\,d\lambda \leqq \Phi(1) < \infty, \tag{3.37}$$

so that, by Lemma 3.2, (3.31), and Fubini's theorem,

$$E\Phi(V^*f) = \int_0^\infty \varphi(\lambda)P(V^*f > \lambda)\,d\lambda = \int_{-\infty}^\infty \psi(t)\,dt \tag{3.38}$$
$$\leqq 2\int_B \psi(t+1)\,dt.$$

If $t \in B$, then $\lambda = \beta^t$ satisfies

$$b\lambda\varphi(\lambda)P(V^*f > \lambda) < 2b\beta\lambda\varphi(\beta\lambda)P(V^*f > \beta\lambda). \tag{3.39}$$

In particular, $\varphi(\beta\lambda) = \varphi(2^k\lambda) \leqq c^k\varphi(\lambda)$ by the growth condition, so that (3.6) holds with $\alpha = 2\beta c^k$. Also, $2\psi(t+1) \leqq \alpha\psi(t)$, since $P(V^*f > \beta\lambda) \leqq P(V^*f > \lambda)$. Therefore, by Lemma 3.1,

$$E\Phi(V^*f) \leqq \alpha\int_B \psi(t)\,dt \tag{3.40}$$
$$\leqq \alpha\int_{-\infty}^\infty b\beta^t\varphi(\beta^t)G(\beta^t)\,dt,$$

where $G(\lambda)$ is the right side of (3.5). This implies (3.3).

REMARK 3.1. It is clear that Theorem 3.1 holds for any Φ satisfying (3.31), (3.32), and $\Phi(1) < \infty$, for some nonnegative function φ. If $0 < p < \infty$, the power function $\Phi(\lambda) = \lambda^p$ satisfies these conditions so convexity is not required. Convexity is required in the next theorem.

THEOREM 3.2. *Let Φ be a convex function as in Theorem 1.1 and let $z_1, z_2, \cdots$ be a sequence of nonnegative $\mathscr{A}$ measurable functions. Then*

$$E\Phi\left(\sum_{k=1}^\infty E(z_k \mid \mathscr{A}_{k-1})\right) \leqq cE\Phi\left(\sum_{k=1}^\infty z_k\right), \tag{3.41}$$

and the choice of c depends only on $c_{(1.3)}$.

The function z_k need not be $\mathscr{A}_k$ measurable.

PROOF. Let $w_k = E(z_k \mid \mathscr{A}_{k-1})$, let $W_n = \sum_{k=1}^n w_k$, and let $Z_n = \sum_{k=1}^n z_k$, for $1 \leqq n \leqq \infty$.

For each integer j, let

$$\begin{aligned} \mu_j &= \inf\{n: W_n > 2^j \text{ or } w_{n+1} > 2^{j-1}\},\\ \nu_j &= \inf\{n: W_n > 2^{j+1} \text{ or } w_{n+1} > 2^{j-1}\}. \end{aligned} \tag{3.42}$$

Then $\mu_j \leqq \nu_j \leqq \mu_{j+1}$ and letting

$$\begin{aligned} A_j &= \{W_\infty > 2^{j+1}, w^* \leqq 2^{j-1}\},\\ B_j &= \{W_\infty > 2^j\}, \end{aligned} \tag{3.43}$$

we have that $A_j \subset B_j$, both

$$W^j = \sum_{k=1}^\infty I(\mu_j < k \leqq \nu_j)\,w_k \tag{3.44}$$

and

$$Z^j = \sum_{k=1}^{\infty} I(\mu_j < k \leqq \nu_j)\, z_k \tag{3.45}$$

vanish off B_j, and

$$W^j \geqq 2^{j-1} I(A_j). \tag{3.46}$$

As a consequence, we have that if $\mathscr{B}_j$ is the smallest σ-field containing B_j and its complement, then, on B_j,

$$\begin{aligned} E(Z^j \mid \mathscr{B}_j)P(B_j) &= \int_{B_j} Z^j \, dP \\ &= EZ^j = \sum_{k=1}^{\infty} E[I(\mu_j < k \leqq \nu_j)E(z_k \mid \mathscr{A}_{k-1})] \\ &= EW^j \geqq 2^{j-1}P(A_j). \end{aligned} \tag{3.47}$$

Here, we have used (3.46) and the fact that $I(\mu_j < k \leqq \nu_j)$ is $\mathscr{A}_{k-1}$ measurable.

Note that, by (3.31) and (1.3),

$$\begin{aligned} E\Phi(W_\infty) &= \sum_{j=-\infty}^{\infty} \int_{2^j}^{2^{j+1}} \varphi(\lambda)P(W_\infty > \lambda)\, d\lambda \\ &\leqq \sum_{j=-\infty}^{\infty} \Phi(2^{j+1})P(B_j) \leqq c \sum_{j=-\infty}^{\infty} \psi(j), \end{aligned} \tag{3.48}$$

where $\psi(j) = \Phi(2^j)P(B_j)$. By the obvious analogue of Lemma 3.2 (and which is implied by it),

$$\sum_{j=-\infty}^{\infty} \psi(j) \leqq 2 \sum_{j \in J} \psi(j+1), \tag{3.49}$$

where $J = \{j : \psi(j) < 2\psi(j+1)\}$. If $j \in J$, then $P(B_j) \leqq \alpha P(B_{j+1})$ and $2\psi(j+1) \leqq \alpha\psi(j)$ with $\alpha = 2c_{(1.3)}$, so that either

$$P(B_j) \leqq 2\alpha P(A_j) \tag{3.50}$$

or

$$P(B_j) \leqq 2\alpha P(w^* > 2^{j-1}). \tag{3.51}$$

Therefore, $J \subset J_1 \cup J_2$, where J_1 and J_2 correspond to the two possibilities, and

$$E\Phi(W_\infty) \leqq c \sum_{j \in J_1} \psi(j) + c \sum_{j \in J_2} \psi(j). \tag{3.52}$$

Suppose that $j \in J_1$. Then, by (3.47),

$$2^j I(B_j) \leqq 2^j I(B_j)\, 2\alpha \frac{P(A_j)}{P(B_j)} \leqq 4\alpha E(Z^j \mid \mathscr{B}_j), \tag{3.53}$$

and, using Jensen's inequality for conditional expectations, we have that

$$\begin{aligned}(3.54)\qquad \Phi(2^j)I(B_j) &\leqq \Phi\big(4\alpha E(Z^j\,|\,\mathscr{B}_j)\big)\\ &\leqq c\Phi\big(E(Z^j\,|\,\mathscr{B}_j)\big) \leqq cE[\Phi(Z^j)\,|\,\mathscr{B}_j].\end{aligned}$$

Taking expectations of both sides, we obtain $\Phi(2^j)P(B_j) \leqq cE\Phi(Z^j)$.

We now need to use the superadditivity of the convex function Φ: if $a_1, a_2, \cdots$ are nonnegative numbers, then

$$(3.55)\qquad \sum_{k=1}^{\infty} \Phi(a_k) \leqq \Phi\left(\sum_{k=1}^{\infty} a_k\right).$$

This is an immediate consequence of the easy special case: if a and b are nonnegative and finite, then

$$(3.56)\qquad \Phi(a) \leqq \Phi(a + b) - \Phi(b).$$

By superadditivity,

$$(3.57)\qquad \sum_{j\in J_1} \psi(j) \leqq cE \sum_{j\in J_1} \Phi(Z^j) \leqq cE\Phi\left(\sum_{j=-\infty}^{\infty} Z^j\right) \leqq cE\Phi(Z_\infty).$$

Considering the other sum in the bound on $E\Phi(W_\infty)$, we have that

$$\begin{aligned}(3.58)\qquad \sum_{j\in J_2} \psi(j) &\leqq 2\alpha \sum_{j\in J_2} \Phi(2^j)P(w^* > 2^{j-1})\\ &\leqq 2\alpha \sum_{j=-\infty}^{\infty} \int_{2^j}^{2^{j+1}} \varphi(\lambda)P(4w^* > 2^{j+1})\, d\lambda\\ &\leqq 2\alpha E\Phi(4w^*) \leqq cE\Phi(w^*).\end{aligned}$$

Using $\Phi(w^*) \leqq \Sigma_{k=1}^{\infty} \Phi(w_k)$, Jensen's inequality, and superadditivity, we obtain

$$(3.59)\qquad E\Phi(w^*) \leqq \sum_{k=1}^{\infty} E\Phi(w_k) \leqq \sum_{k=1}^{\infty} E\Phi(z_k) \leqq E\Phi(Z_\infty).$$

This completes the proof of Theorem 3.2.

LEMMA 3.3. *Under the conditions of Theorem* 2.1,

$$(3.60)\qquad E\Phi(T^{**}f) \leqq cE\Phi\left(\sum_{k=1}^{\infty} |d_k|\right), \qquad f \in \mathscr{M}.$$

PROOF. Since $\Phi(T^{**}f) \leqq \Phi(T^*f) + \Phi(Tf)$, we need to prove only

$$(3.61)\qquad E\Phi(T^*f) \leqq cE\Phi\left(\sum_{k=1}^{\infty} |d_k|\right),$$

and this only for martingales in $\mathscr{M}$ satisfying $f = f^N$ for some positive integer N. Except for the trivial case in which Φ vanishes identically, $\Phi(\infty) = \infty$, and this implies that $T^*f < \infty$: by (2.9),

$$(3.62)\qquad E\Phi(T^*f) \leqq E \sum_{n=1}^{N} \Phi(T_nf) \leqq \sum_{n=1}^{N} cE\Phi\left(\sum_{k=1}^{n} |d_k|\right) < \infty,$$

assuming, as we may, that the right side of (3.61) is finite.

Let $\tau_0 = 0$; if $j \geqq 1$ and τ_{j-1} is a stopping time, let

$$\tau_j = \inf\{n > \tau_{j-1}: T_n f > 2\gamma^3 T_{\tau_{j-1}} f\}. \tag{3.63}$$

Then τ_j is a stopping time and the finiteness of T^*f imples that $\Omega = \bigcup_{j=0}^{\infty} A_j$, where

$$A_j = \{\tau_j < \infty, \tau_{j+1} = \infty\}. \tag{3.64}$$

On A_0, $T^*f = 0$. Let $j \geqq 1$. Then, on A_j, and in fact on $\{\tau_j < \infty\}$,

$$\begin{aligned} 2\gamma^3 T_{\tau_{j-1}} f < T_{\tau_j} f &\leqq \gamma T(f^{\tau_{j-1}} + h_j) \\ &\leqq \gamma^3[T_{\tau_{j-1}} f + Th_j], \end{aligned} \tag{3.65}$$

where h_j is the martingale f started at τ_{j-1} and stopped at τ_j. Therefore, on A_j, $T_{\tau_{j-1}} f \leqq Th_j$ and, since $\tau_{j+1} = \infty$ implies that $T^*f \leqq 2\gamma^3 T_{\tau_j} f$, we have

$$T^*f \leqq 2\gamma^6[T_{\tau_{j-1}} f + Th_j] \leqq 4\gamma^6 Th_j. \tag{3.66}$$

Therefore, by (1.3), (2.9), and (3.55), we have

$$\begin{aligned} E\Phi(T^*f) &= \sum_{j=1}^{\infty} \int_{A_j} \Phi(T^*f) \leqq c \sum_{j=1}^{\infty} \int_{A_j} \Phi(Th_j) \\ &\leqq c \sum_{j=1}^{\infty} E\Phi\left(\sum_{k=1}^{\infty} I(\tau_{j-1} < k \leqq \tau_j)|d_k|\right) \leqq cE\Phi\left(\sum_{k=1}^{\infty} |d_k|\right), \end{aligned} \tag{3.67}$$

proving (3.61) and completing the proof of the lemma.

4. Proofs of the main results

PROOF OF THEOREM 2.1. If f is in $\mathcal{M}$, then $f = g + h$, where g and h in $\mathcal{M}$ are defined by

$$\begin{aligned} g_n &= \sum_{k=1}^{n} a_k = \sum_{k=1}^{n} [y_k - E(y_k | \mathcal{A}_{k-1})], \\ h_n &= \sum_{k=1}^{n} b_k = \sum_{k=1}^{n} [z_k + E(y_k | \mathcal{A}_{k-1})], \end{aligned} \tag{4.1}$$

with

$$\begin{aligned} y_k &= d_k I(|d_k| \leqq 2d_{k-1}^*), \\ z_k &= d_k I(|d_k| > 2d_{k-1}^*), \end{aligned} \tag{4.2}$$

and $d_k^* = \sup_{0 \leqq j \leqq k} |d_j|$ with $d_0 = 0$. This decomposition was introduced in [3]. Note that $|y_k| \leqq 2d_{k-1}^*$ so that $y_1 = 0$ and

$$|a_k| \leqq 4d_{k-1}^*, \tag{4.3}$$

a bound which is $\mathcal{A}_{k-1}$ measurable. For $k \geqq 2$, $E(d_k | \mathcal{A}_{k-1}) = 0$, which implies that $E(y_k | \mathcal{A}_{k-1}) = -E(z_k | \mathcal{A}_{k-1})$, so that

$$\sum_{k=1}^{\infty} |b_k| \leqq \sum_{k=1}^{\infty} |z_k| + \sum_{k=1}^{\infty} E(|z_k| \,|\, \mathscr{A}_{k-1}). \tag{4.4}$$

On the set $\{|d_k| > 2d_{k-1}^*\}$,

$$|d_k| + 2d_{k-1}^* \leqq 2|d_k| = 2d_k^*. \tag{4.5}$$

Therefore,

$$\sum_{k=1}^{\infty} |z_k| \leqq \sum_{k=1}^{\infty} 2(d_k^* - d_{k-1}^*) = 2d^* \leqq 4f^*, \tag{4.6}$$

and, by (3.35) and Theorem 3.2, we have

$$\begin{aligned} E\Phi\left(\sum_{k=1}^{\infty} |b_k|\right) &\leqq cE\Phi\left(\sum_{k=1}^{\infty} |z_k|\right) \\ &\leqq cE\Phi(d^*) \leqq cE\Phi(f^*). \end{aligned} \tag{4.7}$$

This implies, by Lemma 3.3, that $E\Phi(T^{**}h) \leqq cE\Phi(f^*)$. To complete the proof of the theorem, we show that

$$E\Phi(T^{**}g) \leqq cE\Phi(f^*). \tag{4.8}$$

Let $Uf = f^*$ and $V = T$. Note that $U({}^{n-1}g^n) = |a_n| \leqq 4d_{n-1}^*$, so that, by Theorem 3.1 and (2.8) applied to g,

$$E\Phi(T^*g) \leqq cE\Phi(g^*) + cE\Phi(4d^*). \tag{4.9}$$

Since $d^* \leqq 2f^*$, the last term is dominated by $cE\Phi(f^*)$. Since $g^* = (f - h)^* \leqq f^* + h^*$ and $h^* \leqq \Sigma_{k=1}^{\infty} |b_k|$, we have, by (4.7), that

$$E\Phi(g^*) \leqq cE\Phi(f^*). \tag{4.10}$$

Therefore,

$$E\Phi(T^*g) \leqq cE\Phi(f^*), \tag{4.11}$$

which is not quite (4.8). To complete the proof, we may assume that Φ does not vanish identically and that $E\Phi(f^*) < \infty$. Then, by (4.10), $g^* < \infty$ and this implies that

$$Tg \leqq \gamma^3 T^*g \tag{4.12}$$

as we now show. Clearly, (4.11) and (4.12) imply (4.8). Let $\lambda > 0$,

$$\tau = \inf\{n: |g_n| > \lambda \text{ or } d_n^* > \lambda\}, \tag{4.13}$$

and $G = g^\tau$. Then, by (4.3) and the definition of τ, $G^* \leqq 5\lambda$, so that G converges almost everywhere and there exist $n_1 < n_2 < \cdots$ such that

$$\|({}^{n_j}G)^*\|_{p_0}^{p_0} \leqq (2^{-j})^{p_0+1}, \qquad j \geqq 1. \tag{4.14}$$

Therefore, by (2.6),

$$P(T({}^{n_j}G) > 2^{-j}) \leqq c2^{-j} \tag{4.15}$$

and $\lim_{j\to\infty} T({}^{n_j}G) = 0$ almost everywhere. Since

$$TG = T(g^{\tau\wedge n_j} + {}^{n_j}G) \leqq \gamma^2[T_{\tau\wedge n_j}g + T({}^{n_j}G)], \tag{4.16}$$

it follows that $TG \leqq \gamma^2 T^*g$.

Moreover,

$$Tg = T(G + {}^{\tau}g) \leqq \gamma[TG + T({}^{\tau}g)], \tag{4.17}$$

so that $Tg \leqq \gamma^3 T^*g$ on the set

$$\{T({}^{\tau}g) = 0\} \supset \{\tau = \infty\} = \{g^* \leqq \lambda, d^* \leqq \lambda\}, \tag{4.18}$$

which converges almost everywhere to Ω as λ increases. This gives (4.12) and completes the proof of the theorem.

Proof of Theorem 2.2. Here let $U = T$ and $Vf = f^*$. Again, for any f in $\mathscr{M}$, let $f = g + h$, where g and h are defined as in (4.1). By (2.14) applied to ${}^{n-1}g^n$ and by (4.3),

$$U({}^{n-1}g^n) = T({}^{n-1}g^n) \leqq c|a_n| \tag{4.19}$$

and

$$V({}^{n-1}g^n) = |a_n| \leqq 4d^*_{n-1}. \tag{4.20}$$

Therefore, by Theorem 3.1 applied to g, we have

$$E\Phi(g^*) \leqq cE\Phi(T^*g) + cE\Phi(d^*). \tag{4.21}$$

By (2.14) applied to h,

$$T^*g \leqq \gamma(T^*f + T^*h) \leqq \gamma T^*f + c\sum_{k=1}^{\infty}|b_k|, \tag{4.22}$$

and so, by (4.7), we obtain $E\Phi(T^*g) \leqq cE\Phi(T^*f) + cE\Phi(d^*)$.

Since $h^* \leqq \Sigma_{k=1}^{\infty}|b_k|$, we also have $E\Phi(h^*) \leqq cE\Phi(d^*)$. An application of (2.13) now gives

$$E\Phi(f^*) \leqq cE\Phi(g^*) + cE\Phi(h^*) \leqq cE\Phi(T^*f), \tag{4.23}$$

completing the proof of the theorem.

5. Quadratic variation of right continuous martingales

Let $X = \{X(t), 0 \leqq t \leqq 1\}$ be a right continuous martingale and

$$S_j = \left[X(t_{j,1})^2 + \sum_{k=2}^{\infty}\left(X(t_{j,k}) - X(t_{j,k-1})\right)^2\right]^{1/2}, \tag{5.1}$$

where $0 = t_{j,1} \leqq t_{j,2} \leqq \cdots \leqq 1$ for $j \geqq 1$. Note that $S_j = S(f_j)$, where $f_j = (f_{j,1}, f_{j,2}, \cdots)$ is the martingale defined by $f_{j,n} = X(t_{j,n})$. We assume that $t_{j,k} = 1$ for $k \geqq k_j$, and

$$\sup_{2\leqq k<\infty}(t_{j,k} - t_{j,k-1}) \to 0 \tag{5.2}$$

as j increases. Let X^* denote the maximal function of the process X: $X^* = \sup_{0 \leqq t \leqq 1} |X(t)|$.

THEOREM 5.1. *The sequence $\{S_j\}$ converges in L_1 if and only if X^* is integrable.*

We need the following elementary lemma.

LEMMA 5.1. *Let $Y \geqq 0$ be an integrable random variable. Then there is a convex function Φ from $[0, \infty)$ into $[0, \infty)$, with $\Phi(0) = 0$, satisfying the growth condition* (1.3),

$$\lim_{\lambda \to \infty} \frac{\Phi(\lambda)}{\lambda} = \infty, \tag{5.3}$$

and

$$E\Phi(Y) < \infty. \tag{5.4}$$

PROOF OF THEOREM 5.1. Suppose that X^* is integrable. Choose Φ to satisfy Lemma 5.1 with $Y = X^*$. By Theorem 1.1, we then have

$$E\Phi(S_j) \leqq cE\Phi(f_j^*) \leqq cE\Phi(X^*), \tag{5.5}$$

which implies the uniform integrability of $\{S_j\}$. Since, by a result of Doléans [4], $\{S_j\}$ converges in probability, the convergence in L_1 follows.

To go the other way, we note that if $\{S_j\}$ is uniformly integrable or converges in L_1, then $\sup_j ES_j < \infty$. Therefore, since $X^* = \lim_{j \to \infty} f_j^*$ by right continuity, we have

$$EX^* \leqq \liminf_{j \to \infty} Ef_j^* \leqq c \sup_j ES_j < \infty. \tag{5.6}$$

Here, we have used the right side of (1.4) for the special case of Φ the identity function, the case treated in [3].

PROOF OF LEMMA 5.1. Choose $0 = a_0 < a_1 < a_2 < \cdots$ to satisfy (a) $a_j - a_{j-1} > 2a_{j-1}$ for $j \geqq 1$ and (b) $E(YI(Y \geqq a_j)) \leqq 2^{-j}EY$ for $j \geqq 0$. Let $\Phi(\lambda) = \int_0^\lambda \varphi(t)\,dt$ be defined by $\varphi = j$ on $[a_{j-1}, a_j)$. Since φ is increasing, Φ is convex, and since $\lim_{\lambda \to \infty} \varphi(\lambda) = \infty$ equation (5.3) is satisfied. By (a), if $\lambda \in [a_{j-1}, a_j)$ then $2\lambda \leqq a_{j+1}$, so that $\varphi(2\lambda) \leqq (j + 1/j)\varphi(\lambda) \leqq 2\varphi(\lambda)$ for all $\lambda > 0$, and thus

$$\Phi(2\lambda) = \int_0^{2\lambda} \varphi(t)\,dt = 2\int_0^{\lambda} \varphi(2t)\,dt \leqq 4\int_0^{\lambda} \varphi(t)\,dt = 4\Phi(\lambda). \tag{5.7}$$

Also note that $\Phi(\lambda) \leqq j\lambda$ if $\lambda \leqq a_j$, since $\varphi(\lambda) \leqq j$ if $\lambda \leqq a_j$.

Therefore,

$$\begin{aligned} E\Phi(Y) &= \sum_{j=1}^{\infty} E\Phi(YI(a_{j-1} \leqq Y < a_j)) \\ &\leqq \sum_{j=2}^{\infty} \frac{E(jYI(Y \geqq a_{j-1}))}{a_{j-1}} + 1 \\ &\leqq \sum_{j=2}^{\infty} \frac{j2^{-j+1}EY}{a_1} + 1 < \infty, \end{aligned} \tag{5.8}$$

completing the proof.

6. Further applications and remarks

Let

$$(6.1) \qquad r(f) = \left[\sum_{k=1}^{\infty} [E(|d_k| \mid \mathscr{A}_{k-1})]^2\right]^{1/2}$$

for f in $\mathscr{M}$. This is the operator mentioned at the end of Section 2. Let Φ be a convex function as in Theorem 1.1. Then

$$(6.2) \qquad E\Phi(r(f)) \leqq cE\Phi(f^*), \qquad f \in \mathscr{M},$$

and the choice of c depends only on $c_{(1.3)}$. We check that the conditions of Theorem 2.1 are satisfied. Since $r(f) \leqq s(f)$, the operator r is local and satisfies (2.6) for $p_0 = 2$:

$$(6.3) \qquad \begin{aligned} \|r(f)\|_2^2 \leqq \|s(f)\|_2^2 &= \sum_{k=1}^{\infty} E\, d_k^2 \\ &\leqq \sup_n \|f_n\|_2^2 \leqq \|f^*\|_2^2. \end{aligned}$$

Clearly, r is sublinear, symmetric, and measurable. Since

$$(6.4) \qquad r(f) \leqq \sum_{k=1}^{\infty} E(|d_k| \mid \mathscr{A}_{k-1}),$$

we have, by Theorem 3.2, that condition (2.9) is satisfied. To check (2.8) note that

$$(6.5) \qquad \begin{aligned} r({}^{n-1}f^n) &= E(|d_n| \mid \mathscr{A}_{n-1}) \\ &\leqq d_{n-1}^* + E(d_n^* - d_{n-1}^* \mid \mathscr{A}_{n-1}) \leqq d^* + \sum_{k=1}^{\infty} E(d_k^* - d_{k-1}^* \mid \mathscr{A}_{k-1}) \end{aligned}$$

so (2.8) follows from Theorem 3.2 and the fact that $d^* \leqq 2f^*$. Therefore, (6.2) follows from Theorem 2.1.

The formula (6.2) has a simple consequence in the theory of random walk. Let X be a martingale in $\mathscr{M}$ and write $X_n = \sum_{k=1}^{n} x_k$ for $n \geqq 1$. Assume that there is a positive number δ such that

$$(6.6) \qquad E(|x_k| \mid \mathscr{A}_{k-1}) \geqq \delta, \qquad k \geqq 1.$$

For example, this condition is satisfied if $x = (x_1, x_2, \cdots)$ is an independent sequence of identically distributed random variables with $Ex_1 = 0$, $E|x_1| > 0$, and $\mathscr{A}_k$ is the smallest σ-field with respect to which $x_0, \cdots, x_k$ are measurable ($x_0 = 0$), $k \geqq 0$. Let τ be a stopping time and $f = X^\tau$. Then

$$(6.7) \qquad E(|d_k| \mid \mathscr{A}_{k-1}) = I(\tau \geqq k) E(|x_k| \mid \mathscr{A}_{k-1}) \geqq \delta I(\tau \geqq k)$$

so that $r(f) \geqq \delta\tau^{1/2}$ and, by (6.2),

$$(6.8) \qquad E\Phi(\tau^{1/2}) \leqq cE\Phi[(X^\tau)^*].$$

The choice of c in (6.8) depends only on δ and $c_{(1.3)}$. See Section 5 in [2] for related results.

REMARK 6.1. Here we show that the conclusion of Theorem 2.2 need not hold for the operator r. Consider $\Phi(\lambda) = \lambda$ and an independent random variable sequence $d = (d_1, d_2, \cdots)$ satisfying $P(d_k = -1) = P(d_k = 1) = (2k)^{-1}$ and $P(d_k = 0) = 1 - k^{-1}$. We assume that $\mathscr{A}_k$ is generated by $d_0, \cdots, d_k$ with $d_0 = 0$. Then

$$E(|d_k| \,|\, \mathscr{A}_{k-1}) = E|d_k| = k^{-1} \text{ so that } Er^*(f) = \left(\sum_{k=1}^{\infty} k^{-2}\right)^{1/2} < \infty. \tag{6.9}$$

Suppose that (2.12) holds here. Then $Ef^* < \infty$, which implies that f converges almost everywhere. Hence, d converges almost everywhere to 0. But this contradicts the fact, which follows from the Borel-Cantelli lemma, that, with probability one, $|d_k| = 1$ for infinitely many k. If $n \geqq 2$, then $r({}^{n-1}f^n) = n^{-1}$ but $|d_n| = 0$ on a set of positive measure. Therefore, r fails to satisfy (2.14).

REMARK 6.2. Suppose that $(\Omega, \mathscr{A}, P)$ is nonatomic. Then the growth condition $\Phi(2\lambda) \leqq c\Phi(\lambda)$ is a necessary condition for either side of the conclusion of Theorem 1.1. For example, suppose that the left side holds and h is a martingale satisfying $\|S(h)\|_\infty > 2$ and $\|h^*\|_\infty = 1$ so that $\alpha = P(S(h) > 2) > 0$. Then, for $\lambda > 0$ and $f = \lambda h$, we have

$$\begin{aligned} \alpha\Phi(2\lambda) &= \Phi(2\lambda)P(S(f) > 2\lambda) \\ &\leqq E\Phi(S(f)) \leqq cE\Phi(f^*) \leqq c\Phi(\lambda). \end{aligned} \tag{6.10}$$

Such an h exists. Let $x = (x_1, x_2, \cdots)$ be an independent sequence satisfying $P(x_k = -1) = P(x_k = 1) = \frac{1}{2}$, $k \geqq 1$, X the martingale with difference sequence x, and $\tau = \inf\{n: |X_n| = 2\}$. Then $h = \frac{1}{2}X^\tau$ satisfies $h^* = 1$ and $\|S(h)\|_\infty = \infty$.

REFERENCES

[1] D. L. BURKHOLDER, "Martingale transforms," *Ann. Math. Statist.*, Vol. 37 (1966), pp. 1494–1504.

[2] D. L. BURKHOLDER and R. F. GUNDY, "Extrapolation and interpolation of quasi-linear operators on martingales," *Acta Math.*, Vol. 124 (1970), pp. 249–304.

[3] B. J. DAVIS, "On the integrability of the martingale square function," *Israel J. Math.*, Vol. 8 (1970) pp. 187–190.

[4] C. DOLÉANS, "Variation quadratique des martingales continues à droite," *Ann. Math. Statist.*, Vol. 40 (1969), pp. 284–289.

[5] A. ZYGMUND, *Trigonometric Series*, Vol. I, Cambridge, Cambridge University Press, 1959.