

# THE BEST CONSTANT IN THE DAVIS INEQUALITY FOR THE EXPECTATION OF THE MARTINGALE SQUARE FUNCTION

DONALD L. BURKHOLDER

**ABSTRACT.** A method is introduced for the simultaneous study of the square function and the maximal function of a martingale that can yield sharp norm inequalities between the two. One application is that the expectation of the square function of a martingale is not greater than  $\sqrt{3}$  times the expectation of the maximal function. This gives the best constant for one side of the Davis two-sided inequality. The martingale may take its values in any real or complex Hilbert space. The elementary discrete-time case leads quickly to the analogous results for local martingales  $M$  indexed by  $[0, \infty)$ . Some earlier inequalities are also improved and, closely related, the Lévy martingale is embedded in a large family of submartingales.

## 1. INTRODUCTION

Square-function inequalities abound in harmonic analysis, in both measure-based and noncommutative probability theory, and elsewhere, as can be seen, for example, in Stein [24, 25], Dellacherie and Meyer [8], Carlen and Krée [5], and Pisier and Xu [23]. Our aim here is to introduce a new approach to the simultaneous study of the square function and the maximal function of a martingale in the measure setting, and, with its use, to obtain the sharp form of an inequality due to Davis [7] together with sharper versions of some earlier inequalities.

We begin with martingales  $f$  indexed by the set of nonnegative integers. The results for such martingales then lead to the analogous results of Section 6 for local martingales  $M$  indexed by  $[0, \infty)$ . To help clarify the scope of the results,  $f$  will sometimes have its values in  $\mathbf{R}$ , sometimes in  $\mathbf{H}$ , and sometimes in  $\mathbf{B}$ , where  $\mathbf{H}$  will always denote a Hilbert space and  $\mathbf{B}$  a Banach space. For all three cases, the norm will be denoted by  $|\cdot|$ . For the second and third cases, the scalar field can be either  $\mathbf{R}$  or  $\mathbf{C}$ . The results are new even for real-valued martingales.

We recall some definitions. Let  $(\Omega, \mathcal{F}, P)$  be a probability space,  $(\mathcal{F}_n)_{n \geq 0}$  a nondecreasing sequence of sub- $\sigma$ -algebras of  $\mathcal{F}$ ,  $d_n : \Omega \rightarrow \mathbf{B}$  a (Bochner)  $\mathcal{F}_n$ -measurable and integrable function, and

$$(1.1) \quad f_n = \sum_{k=0}^n d_k.$$

---

Received by the editors March 2, 2001.

2000 *Mathematics Subject Classification.* Primary 60G44, 60G42; Secondary 60J65, 42B25.

*Key words and phrases.* Square function, maximal function, martingale.

The sequence  $(f_n)_{n \geq 0}$ , denoted by  $f$ , is a *martingale* if  $E[d_n | \mathcal{F}_{n-1}] = 0$  for all  $n \geq 1$ . That this conditional expectation is equal (except on a set of measure zero) to the origin of the Banach space is equivalent to  $\int_A d_n dP = 0$  for all  $A$  in  $\mathcal{F}_{n-1}$ . The martingale  $f$  is *simple* if each  $f_n$  is a simple function and there is an integer  $m$  such that  $f_n = f_m$  for all  $n \geq m$ . Any  $\mathbf{B}$ -valued function of a simple martingale has a finite range and is therefore integrable. In some of what follows  $(\Omega, \mathcal{F}, P)$  needs to be sufficiently rich: we can and do assume that it is nonatomic.

Let  $S_n(f)$  and  $f_n^*$  be defined on  $\Omega$  by

$$S_n(f) = \left[ \sum_{k=0}^n |d_k(\cdot)|^2 \right]^{1/2} \quad \text{and} \quad f_n^* = \sup_{0 \leq m \leq n} |f_m(\cdot)|.$$

Then  $S(f)$ , the *square function* of  $f$ , and  $f^*$ , the *maximal function*, are their respective pointwise limits as  $n \rightarrow \infty$ . If  $0 < p < \infty$ , let  $\|f\|_p = \sup_{n \geq 0} \|f_n\|_p$ . Doob's maximal inequality [10], the martingale version of the Hardy-Littlewood inequality, is  $\|f\|_p \leq \|f^*\|_p \leq q \|f\|_p$  where  $q = p/(p-1)$  and  $p > 1$ .

Inequalities between  $\|S(f)\|_p$  and  $\|f^*\|_p$  or  $\|f\|_p$  for several families of martingales, real-valued in this paragraph, such as the partial sums of Rademacher series and the partial sums of Haar series, inequalities that depend on whether  $0 < p < 1$ ,  $p = 1$ , or  $1 < p < \infty$ , were proved (without using the martingale concept or word) by Khintchine [16], Littlewood [18], Paley [22], Marcinkiewicz [20], and Marcinkiewicz and Zygmund [21]. For many applications, however, it is important to know, for a given  $p$ , whether or not these and related inequalities hold for much larger families of martingales. For example, consider

$$(1.2) \quad c_p \|f^*\|_p \leq \|S(f)\|_p \leq C_p \|f^*\|_p.$$

If  $p \geq 1$ , then this inequality holds for *all* martingales. For  $p > 1$ , this follows at once from the two-sided inequality between  $\|S(f)\|_p$  and  $\|f\|_p$  in [1] and the Doob maximal inequality. Later it was shown [4] that if  $p \leq 1$ , then (1.2) holds for a large family of martingales that includes the special families mentioned above and others, but that for  $p < 1$ , the choice of the positive real numbers  $c_p$  and  $C_p$  for this family must depend on more than  $p$ . In particular, if  $p < 1$ , then (1.2) does not hold for all martingales, not even for all martingales of the form (1.1) where  $\{d_k : k \geq 0\}$  is a stochastically independent family of mean-zero functions, known already in the 1930s by Marcinkiewicz and Zygmund [21]. After the work of [1] and [4], there remained the question for  $p = 1$  of whether or not (1.2) holds for all martingales. Davis [7] proved that it does. The proof rests on a clever decomposition of the martingale  $f$ . Later, using the martingale version [12, p. 149] of the Fefferman duality theorem  $(H^1)^* = BMO$  (see [11, 12]), Garsia [13, 14] proved that  $2 + \sqrt{5}$  is a permissible choice for  $C_1$  in (1.2). Although both of these proofs for the case  $p = 1$  are of interest, neither yields the best constant. Here we prove that  $\sqrt{3}$  is the best constant by using an entirely different approach.

The following theorem is our key result for martingales  $f$  with index set the nonnegative integers. It yields the analogous but more general theorem of Section 6 for local martingales  $M$  with index set  $[0, \infty)$ .

**Theorem 1.1.** *If  $f$  is a martingale with values in a Hilbert space  $\mathbf{H}$ , then*

$$(1.3) \quad E[S(f)] \leq \sqrt{3} E[f^*].$$

The constant  $\sqrt{3}$  is the best possible. If  $1 \leq p \leq 2$ , then

$$(1.4) \quad \|S(f)\|_p \leq \gamma_p \|f^*\|_p$$

where the positive real number  $\gamma_p$  is given by

$$(1.5) \quad \gamma_p^2 = 1 + \sup_{y>1} \frac{(2y-1)(1-y^{p-2})}{y^p-1}.$$

So if  $0 < \gamma < \sqrt{3}$ , then there is a martingale  $f$  such that  $E[S(f)] > \gamma E[f^*]$ . If  $p = 1$ , then (1.5) simplifies and gives  $\gamma_1 = \sqrt{3}$  so (1.4) contains (1.3). Also,  $\gamma_2 = 1$  so for these two cases (1.5) gives the best constant. We do not know whether or not  $\gamma_p$  is the best constant for  $1 < p < 2$ , but some computation suggests that it is not far away. Another interesting question is whether or not  $\gamma_p$  can be expressed as a function of  $p$  in a more pleasing way.

The key to Theorem 1.1 is the function  $U$  defined by (2.6) in Lemma 2.1, a function that also throws new light on one of the most intriguing properties of a real Brownian motion  $B$ , one discovered by Lévy [17]:  $(B_t^2 - t)_{t \geq 0}$  is a martingale. Using  $U$ , we can embed the Lévy martingale in a large family of submartingales, equivalently, its negative in a family of supermartingales. Suppose that  $B$  starts at zero (for simplicity),  $1 \leq p \leq 2$ ,  $\gamma \geq \gamma_p$ ,  $\alpha = \frac{p}{2}\gamma^{p-2}$ , and  $z > 0$ . Then, as a special case of Theorem 7.1, the process  $Y$  defined by

$$Y_t = U(B_t, t, B_t^* \vee z) = \alpha[(1 - \gamma^2)(B_t^* \vee z)^p - (B_t^2 - t)(B_t^* \vee z)^{p-2}]$$

is a supermartingale. If  $p = 2$  and  $\gamma = 1$ , then  $\alpha = 1$  and  $Y$  is the martingale given by  $Y_t = -(B_t^2 - t)$ .

Section 2 contains the first steps to the proof in Section 3 of the inequality (1.4). Similarly, Section 4 provides a part of the proof in Section 5 that no smaller number than  $\sqrt{3}$  suffices for the case  $p = 1$ . The probability background needed is minimal for the first five sections and Section 8: just a few properties of conditional expectations. Sections 6 and 7 require a little more. Section 6 contains the continuous-time local martingale version of Theorem 1.1. In Section 7, we embed the Lévy martingale  $(B_t^2 - t)_{t \geq 0}$  in a large family of submartingales. In Section 8, we describe some aspects of our search for the function  $U$  of Lemma 2.1.

## 2. FIRST STEPS

The following theorem, which is of interest in its own right, will be used in the proof of the inequality (1.4), a proof that is no simpler in the special case of real-valued martingales and  $p = 1$ .

**Theorem 2.1.** *Let  $\mathbf{B}$  be a Banach space and suppose that  $U$  and  $V$  are functions from  $\mathbf{B} \times [0, \infty) \times (0, \infty)$  into  $\mathbf{R}$  satisfying*

$$(2.1) \quad V(x, t, z) \leq U(x, t, z),$$

$$(2.2) \quad U(x, t, z) = U(x, t, |x| \vee z),$$

*and the further condition that if  $|x| \leq z$  and  $d : \Omega \rightarrow \mathbf{B}$  is a simple  $\mathcal{F}$ -measurable function with  $E[d] = 0$ , then*

$$(2.3) \quad E[U(x + d, t + |d|^2, z)] \leq U(x, t, z).$$

*Under these three conditions,*

$$(2.4) \quad E[V(f_n, t - |x|^2 + S_n^2(f), f_n^* \vee z)] \leq U(x, t, z)$$

for all nonnegative integers  $n$  and simple martingales  $f$  with  $f_0 = x$ .

The rather mysterious condition (2.2) gives a tool for handling maximal functions. It has already played a role in [3] for a different problem that requires a different function  $U$ . Condition (2.3) is a kind of concavity condition.

*Proof.* Let  $Y_n = U(f_n, t - |x|^2 + S_n^2(f), f_n^* \vee z)$ . By (2.1), it is enough to prove that  $E[Y_n] \leq U(x, t, z)$ . By (2.2),  $Y_0 = U(x, t, |x| \vee z) = U(x, t, z)$  so  $E[Y_0] = U(x, t, z)$ . If  $n \geq 1$ , then  $f_n^* = |f_n| \vee f_{n-1}^*$  so, by (2.2), the function  $Y_n$  is equal to

$$U(f_{n-1} + d_n, t - |x|^2 + S_{n-1}^2(f) + |d_n|^2, f_{n-1}^* \vee z).$$

We can assume that the filtration  $(\mathcal{F}_n)_{n \geq 0}$  is generated by the simple martingale  $f$ . Therefore,  $\mathcal{F}_{n-1}$  is generated by a finite partition of  $\Omega$  and on each partition set  $f_{n-1}$ ,  $f_{n-1}^*$ , and  $S_{n-1}^2(f)$  are constant. Because  $|f_{n-1}| \leq f_{n-1}^* \vee z$ , condition (2.3) implies that

$$(2.5) \quad E[Y_n | \mathcal{F}_{n-1}] \leq Y_{n-1}.$$

Taking the expectation of each side of this inequality, we see that

$$E[Y_n] \leq E[Y_{n-1}] \leq \cdots \leq E[Y_0]$$

so  $E[Y_n] \leq U(x, t, z)$  as desired.  $\square$

Let us now consider the following pair of functions  $U$  and  $V$ , actually a family of pairs, one of which will be used in the proof (see Section 3) of the inequality (1.4).

**Lemma 2.1.** *Let  $\mathbf{H}$  be a Hilbert space,  $\gamma > 0$ ,  $1 \leq p \leq 2$ , and  $\alpha = \frac{p}{2}\gamma^{p-2}$ . Define functions  $U, V : \mathbf{H} \times [0, \infty) \times (0, \infty) \rightarrow \mathbf{R}$  by*

$$(2.6) \quad U(x, t, z) = \alpha(|x| \vee z)^p \left[ \frac{t - |x|^2}{(|x| \vee z)^2} - \gamma^2 + 1 \right],$$

$$(2.7) \quad V(x, t, z) = t^{p/2} - \gamma^p(|x| \vee z)^p.$$

If  $\gamma \geq \gamma_p$ , where  $\gamma_p$  is the positive number defined in (1.5), then  $U$  and  $V$  satisfy (2.1), (2.2), and (2.3).

*Proof.* To check that the majorization condition (2.1) holds, define the functions  $u$  and  $v$  on  $\{(x, t) \in \mathbf{H} \times [0, \infty) : |x| \leq 1\}$  by

$$u(x, t) = \alpha(t - |x|^2 - \gamma^2 + 1),$$

$$v(x, t) = t^{p/2} - \gamma^p.$$

The concave function  $v(x, \cdot)$  is majorized by the affine function  $u(x, \cdot)$  because they have the same derivative at  $\gamma^2$  and  $v(x, \gamma^2) \leq u(x, \gamma^2)$ . This implies that for  $(x, t, z) \in \mathbf{H} \times [0, \infty) \times (0, \infty)$ ,

$$\begin{aligned} V(x, t, z) &= (|x| \vee z)^p v\left(\frac{x}{|x| \vee z}, \frac{t}{(|x| \vee z)^2}\right) \\ &\leq (|x| \vee z)^p u\left(\frac{x}{|x| \vee z}, \frac{t}{(|x| \vee z)^2}\right) = U(x, t, z). \end{aligned}$$

Therefore, (2.1) holds. The replacement of  $z$  by  $|x| \vee z$  does not change  $U$ , so condition (2.2) also holds. Because  $U(x, t, z) = z^p U(x/z, t/z^2, 1)$ , it is enough to

check (2.3) for  $z = 1$ . Fix  $x \in \mathbf{H}$  satisfying  $|x| \leq 1$  and define  $\Phi$  and  $\Psi$  on  $\mathbf{H}$  by

$$\begin{aligned}\Phi(s) &= \frac{1}{\alpha} U(x + s, t + |s|^2, 1), \\ \Psi(s) &= t - |x|^2 - 2x \cdot s - \gamma^2 + 1,\end{aligned}$$

where  $x \cdot s$  denotes the real part of the inner product of  $x$  and  $s$ . If  $|x + s| \leq 1$ , then  $\Psi(s) = \Phi(s)$  because

$$t - |x|^2 - 2x \cdot s - \gamma^2 + 1 = t + |s|^2 - |x + s|^2 - \gamma^2 + 1.$$

If  $|x + s| > 1$ , then  $|x + s| \vee 1 = |x + s|$  so in this case  $\Psi(s) - \Phi(s)$  is given by

$$(t - |x|^2 - 2x \cdot s - \gamma^2 + 1) - |x + s|^p \left[ \frac{t + |s|^2 - |x + s|^2}{|x + s|^2} - \gamma^2 + 1 \right],$$

which is nondecreasing in  $t$  and, at  $t = 0$ , is given by

$$\gamma^2(|x + s|^p - 1) + (1 - |x|^2 - 2x \cdot s - |s|^2|x + s|^{p-2}).$$

If this expression is nonnegative, then  $\Psi(s) \geq \Phi(s)$ . So  $\Psi$  majorizes  $\Phi$  on  $\mathbf{H}$  if the inequality  $\gamma^2 \geq r(x, y)$  holds for all  $y \in \mathbf{H}$  satisfying  $|y| > 1$ , where

$$r(x, y) = \frac{-1 + |x|^2 + 2x \cdot (y - x) + |y - x|^2|y|^{p-2}}{|y|^p - 1}.$$

Here  $y - x$  replaces  $s$ . A simple calculation shows that  $r(y/|y|, y) \geq r(x, y)$ :

$$(|y| - 1)[r(y/|y|, y) - r(x, y)] = (1 - |y|^{-1})(-1 + |x|^2 + 2|y| - 2x \cdot y),$$

so, by the Cauchy-Schwarz inequality and the assumption that  $|x| \leq 1 < |y|$ ,

$$|y|[r(y/|y|, y) - r(x, y)] \geq (1 - |x|)[2|y| - (1 + |x|)] \geq 0.$$

Therefore, the majorization of  $\Phi$  by  $\Psi$  follows from  $\gamma^2 \geq \sup_{|y| > 1} r(y/|y|, y)$ , which is easily seen to be equivalent to  $\gamma^2 \geq \gamma_p^2$ .

To finish the proof that  $U$  satisfies the condition (2.3) if  $\gamma \geq \gamma_p$ , we use the majorization of  $\Phi$  by  $\Psi$  on  $\mathbf{H}$  to obtain

$$(2.8) \quad E[\Phi(d)] \leq E[\Psi(d)] = \Psi(E[d]) = \Psi(0) = \Phi(0)$$

for all simple  $\mathbf{H}$ -valued  $\mathcal{F}$ -measurable functions  $d$  with  $E[d] = 0$ . This gives (2.3) for  $z = 1$  and completes the proof of Lemma 2.1.  $\square$

**Lemma 2.2.** *Let  $\mathbf{B}$  be a Banach space,  $1 \leq p < \infty$ , and  $1 \leq \gamma < \infty$ . (i) If the inequality  $\|S(f)\|_p \leq \gamma \|f^*\|_p$  holds for all  $\mathbf{B}$ -valued martingales  $f$  with  $f_0 = 0$ , then it holds for all  $\mathbf{B}$ -valued martingales  $f$ . (ii) Let  $n$  be a positive integer and  $\varepsilon$  a positive number. If  $f$  is a  $\mathbf{B}$ -valued martingale with  $\|f\|_p$  finite and  $f_0 = 0$ , then there is a simple martingale  $g$  such that  $g_0 = 0$  and*

$$\|S_n(g) - S_n(f)\|_p \vee \|g_n^* - f_n^*\|_p < \varepsilon.$$

*Proof.* (i) Let  $f$  be a  $\mathbf{B}$ -valued martingale on  $(\Omega, \mathcal{F}, P)$  and  $\mu$  Lebesgue measure on the Borel sets of  $[0, 1]$ . Define  $r$  by  $r(s) = 1$  if  $s \in [0, 1/2)$  and  $r(s) = -1$  if  $s \in [1/2, 1]$ . On the product probability space, define a martingale difference sequence  $D_0, D_1, \dots$  by letting  $D_0(s, \omega) = 0$  and  $D_k(s, \omega) = r(s)d_{k-1}(\omega)$  for  $k \geq 1$ . The martingale  $F$  with this difference sequence satisfies  $F_0 = 0$ ,  $S(F)(s, \omega) = S(f)(\omega)$ , and  $F^*(s, \omega) = f^*(\omega)$ . This yields the desired conclusion.

(ii) Here let  $f$  be a  $\mathbf{B}$ -valued martingale with  $f_0 = 0$ ,  $a_k$  a simple  $\mathcal{F}_k$ -measurable function such that  $\|a_k - d_k\|_p \leq \varepsilon/2n$  if  $k = 1, \dots, n$ , and  $a_k = 0$  if  $k = 0$  or

$k > n$ . Let  $e_0 = 0$  and  $e_k = a_k - E[a_k | \mathcal{G}_{k-1}]$  if  $k \geq 1$ , where  $\mathcal{G}_k$  is the smallest  $\sigma$ -algebra with respect to which the functions  $a_0, \dots, a_k$  are measurable. Then  $e_k$  is a simple function and  $(e_k)_{k \geq 0}$  is the difference sequence of a simple martingale  $g$  satisfying  $g_0 = 0$ . Therefore,  $\|e_0 - d_0\|_p = 0$  and if  $1 \leq k \leq n$ , then  $\mathcal{G}_{k-1} \subset \mathcal{F}_{k-1}$  so  $E[d_k | \mathcal{G}_{k-1}] = 0$  and, by the contraction property of conditional expectations,

$$\|e_k - d_k\|_p = \|a_k - d_k - E[a_k - d_k | \mathcal{G}_{k-1}]\|_p \leq 2\|a_k - d_k\|_p \leq \varepsilon/n.$$

By Minkowski's inequality,

$$|S_n(g) - S_n(f)| \leq \left[ \sum_{k=0}^n (|e_k| - |d_k|)^2 \right]^{1/2} \leq \sum_{k=0}^n |e_k - d_k|.$$

Similarly,

$$|g_n^* - f_n^*| \leq \sup_{0 \leq m \leq n} |g_m - f_m| \leq \sum_{k=0}^n |e_k - d_k|.$$

So  $\|S_n(g) - S_n(f)\|_p \vee \|g_n^* - f_n^*\|_p \leq 2 \sum_{k=1}^n \|a_k - d_k\|_p < \varepsilon$ , the desired result.  $\square$

### 3. PROOF OF THE INEQUALITY (1.4)

Let  $f$  be a martingale with values in  $\mathbf{H}$ . By the monotone convergence theorem, it is enough to show that

$$(3.1) \quad \|S_n(f)\|_p \leq \gamma_p \|f_n^*\|_p$$

for all nonnegative integers  $n$ . By Lemma 2.2, it is enough to prove this inequality for fixed  $n$  and all simple martingales  $f$  with  $f_0 = 0$ . If  $\gamma \geq \gamma_p$ , then the functions  $U$  and  $V$  of Lemma 2.1 satisfy conditions (2.1), (2.2), and (2.3) of Theorem 2.1, so if  $\gamma = \gamma_p$ ,  $t = 0$ , and  $f$  is a simple martingale with  $f_0 = 0$ , then

$$E[V(f_n, S_n^2(f), f_n^* \vee z)] \leq U(0, 0, z) = \alpha(1 - \gamma_p^2)z^p \leq 0.$$

Thus, by the definition of  $V$ ,  $\|S_n(f)\|_p \leq \gamma_p \|f_n^* \vee z\|_p$ . Let  $z \downarrow 0$  to see that (3.1) holds for all simple martingales  $f$  with  $f_0 = 0$ , hence for all martingales  $f$ . This completes the proof of the inequality (1.4).

### 4. A COMPANION TO THEOREM 2.1

Here the martingales have their values in  $\mathbf{B}$ ,  $\Omega = [0, 1)$ , and  $P$  is Lebesgue measure on the  $\sigma$ -algebra of Borel sets of  $[0, 1)$ . Denote the pointwise limit of a simple martingale  $f$  by  $f_\infty$  and the set of all simple martingales  $f$  with  $f_0 = x$  by  $\mathcal{M}(x)$ . The filtration  $(\mathcal{F}_n)_{n \geq 0}$  varies over all possibilities as does  $f$ .

**Lemma 4.1.** *Let  $V : \mathbf{B} \times [0, \infty) \times (0, \infty) \rightarrow \mathbf{R}$  satisfy*

$$(4.1) \quad V(x, t, z) = V(x, t, |x| \vee z)$$

*and define  $U : \mathbf{B} \times [0, \infty) \times (0, \infty) \rightarrow (-\infty, \infty]$  by*

$$(4.2) \quad U(x, t, z) = \sup\{E[V(f_\infty, t - |x|^2 + S^2(f), f^* \vee z)] : f \in \mathcal{M}(x)\}.$$

*Then the pair  $(U, V)$  satisfies (2.1), (2.2), and (2.3).*

*Proof.* If  $f_n = x$  for all  $n$ , then  $f$  is in  $\mathcal{M}(x)$  and the expectation in (4.2) is  $V(x, t, z)$ . Therefore, the majorization condition (2.1) holds. Replacing  $z$  by  $|x| \vee z$  does not change  $U(x, t, z)$  so (2.2) also holds. To see that (2.3) holds, fix  $x \in \mathbf{B}$  with  $|x| \leq z$ . Let  $d$  be a  $\mathbf{B}$ -valued simple function with  $E[d] = 0$  and  $P(d = s_j) = p_j \in (0, 1]$  for  $1 \leq j \leq m$ , where  $\sum_{j=1}^m p_j = 1$ . Choose  $b_j \in \mathbf{R}$  so that

$$(4.3) \quad U(x + s_j, t + |s_j|^2, z) > b_j.$$

Then, by the definition of  $U$ , there is an  $f^j \in \mathcal{M}(x + s_j)$  satisfying

$$(4.4) \quad E[V(f_\infty^j, t + |s_j|^2 - |x + s_j|^2 + S^2(f^j), (f^j)^* \vee z)] > b_j.$$

Let  $a_0 = 0$  and  $a_j = \sum_{i=1}^j p_i$ . With  $I_j = [a_{j-1}, a_j]$ , define  $\varphi_j : I_j \rightarrow [0, 1]$  by  $\varphi_j(\omega) = (\omega - a_{j-1})/p_j$ . Let  $f$  be the martingale with  $f_0 = x$  such that if  $\omega \in I_j$ ,  $1 \leq j \leq m$ , and  $n \geq 0$ , then  $f_{n+1}(\omega) = f_n^j \circ \varphi_j(\omega)$ . This martingale belongs to  $\mathcal{M}(x)$ . On  $I_j$  its difference sequence satisfies  $d_0 = x, d_1 = s_j, d_2 = d_1^j \circ \varphi_j, \dots$ , and  $f_\infty = f_\infty^j \circ \varphi_j$ . Because  $|x| \leq z$ , it also satisfies  $f^* \vee z = ((f^j)^* \vee z) \circ \varphi_j$  on  $I_j$ . So by (4.2),  $U(x, t, z) \geq E[V(f_\infty, t - |x|^2 + S^2(f), f^* \vee z)]$ , which is equal to

$$\begin{aligned} & \sum_{j=1}^m \int_{I_j} V(f_\infty, t + |s_j|^2 + |d_2|^2 + \dots, f^* \vee z) dP \\ &= \sum_{j=1}^m \int_{I_j} V(f_\infty^j \circ \varphi_j, t + |s_j|^2 - |x + s_j|^2 + S^2(f^j) \circ \varphi_j, ((f^j)^* \vee z) \circ \varphi_j) dP \\ &= \sum_{j=1}^m p_j E[V(f_\infty^j, t + |s_j|^2 - |x + s_j|^2 + S^2(f^j), (f^j)^* \vee z)]. \end{aligned}$$

By (4.4) and (4.3),  $U(x, t, z) \geq \sup \sum_{j=1}^m p_j b_j = \sum_{j=1}^m p_j U(x + s_j, t + |s_j|^2, z)$ , which is equal to  $E[U(x + d, t + |d|^2, z)]$ , the supremum being taken over all the  $b_j$  satisfying (4.3). Therefore, the pair  $(U, V)$  does satisfy (2.1), (2.2), and (2.3).  $\square$

## 5. THE BEST POSSIBLE CONSTANT IN THE DAVIS INEQUALITY

We shall show that  $\sqrt{3}$  is the best possible constant for the inequality (1.3) by showing that it is already the best if  $\mathbf{H} = \mathbf{R}$ . Suppose that  $\gamma \in [1, \infty)$  and  $V : \mathbf{R} \times [0, \infty) \times (0, \infty) \rightarrow \mathbf{R}$  is given by  $V(x, t, z) = \sqrt{t} - \gamma(|x| \vee z)$ . So (4.1) is satisfied by  $V$ . By Lemma 4.1, the pair  $(U, V)$ , where  $U$  is defined by (4.2), satisfies (2.1), (2.2), and (2.3). Clearly,  $U(x, t, z) = \lambda U(x/\lambda, t/\lambda^2, z/\lambda)$  for all  $\lambda > 0$ , a scaling property that, by (4.2),  $U$  inherits from  $V$ . Let  $u$  and  $v$  be defined on  $[-1, 1] \times [0, \infty)$  by  $u(x, t) = U(x, t, 1)$  and  $v(x, t) = V(x, t, 1)$ . Because  $V(x, \cdot, z)$  is nondecreasing, (4.2) implies the same for  $U(x, \cdot, z)$  and  $u(x, \cdot)$ , so limits from the left exist. We shall show that

$$(5.1) \quad u(1, 1-) \geq u(0, 2-) + u(1, 1-).$$

Let  $r > 0$  and  $0 < s < 1$ . Then (2.2), (2.3), and scaling imply that

$$\begin{aligned} u(1, s) &\geq \frac{r}{r+1} u(0, s+1) + \frac{1}{1+r} U(1+r, s+r^2, 1+r) \\ &= \frac{r}{r+1} u(0, s+1) + u(1, (s+r^2)/(1+r)^2). \end{aligned}$$

The map  $r \mapsto (s+r^2)/(1+r)^2$  is strictly increasing on  $(s, \infty)$  and has the limit 1 as  $r \rightarrow \infty$ . Therefore,  $u(1, s) \geq u(0, s+1) + u(1, 1-)$ . Now let  $s \uparrow 1$  to obtain (5.1).

Suppose that  $E[S(f)] \leq \gamma E[f^*]$  for all real-valued simple martingales  $f$  with  $f_0 = 0$ . If  $t = 0$  and  $f \in \mathcal{M}(0)$ , then the expectation on the right side of (4.2) is

$$E[S(f)] - \gamma E[f^* \vee z] \leq E[S(f)] - \gamma E[f^*] \leq 0,$$

and by the definition of  $u$ , this implies that  $u(0, 0) = U(0, 0, 1) \leq 0$ . Because  $U$  satisfies (2.3) and  $u(x, t) = u(-x, t)$ , which also follows from the definition of  $u$ ,

$$0 \geq u(0, 0) \geq \frac{1}{2}[u(1, 1) + u(-1, 1)] = u(1, 1) \geq u(1, 1-) \geq v(1, 1-) = 1 - \gamma.$$

This implies that  $u(1, 1-)$  is finite. So (5.1) yields  $u(0, 2-) \leq 0$  and, by (2.3),

$$0 \geq u(0, 2-) \geq \frac{1}{2}[u(1, 3-) + u(-1, 3-)] = u(1, 3-) \geq v(1, 3-) = \sqrt{3} - \gamma.$$

Therefore,  $\gamma \geq \sqrt{3}$  and this completes the proof of Theorem 1.1.

## 6. FROM DISCRETE TIME TO CONTINUOUS TIME

Except for its second sentence, Theorem 1.1 follows at once from Theorem 6.1 below: any martingale  $f$  indexed by the nonnegative integers can be embedded in a right-continuous martingale  $M$  with  $M^* = f^*$  and  $S(M) = S(f)$ . In the opposite direction, Theorem 6.1 follows easily from Theorem 1.1 and the Hilbert-space version of a convergence result of Doléans [9]. Here  $(\Omega, \mathcal{F}, P)$  is a complete probability space and  $(\mathcal{F}_t)_{t \geq 0}$  a right-continuous filtration such that  $\mathcal{F}_0$  contains all  $A \in \mathcal{F}$  with  $P(A) = 0$ . We first consider right-continuous martingales  $M$  indexed by  $[0, \infty)$ . The function  $M_t : \Omega \rightarrow \mathbf{H}$  is (Bochner)  $\mathcal{F}_t$ -measurable and integrable, and  $E[M_t | \mathcal{F}_s] = M_s$  for all  $s, t \in [0, \infty)$  with  $s \leq t$ . In the proofs we can assume that the real or complex Hilbert space  $\mathbf{H}$  is separable.

**Lemma 6.1.** *Let  $M$  be a right-continuous martingale with values in  $\mathbf{H}$ . There is a family of nonnegative functions  $(S_t(M))_{t \geq 0}$  such that  $S_0(M) = |M_0|$ , the function  $S_t(M)$  is  $\mathcal{F}_t$ -measurable for all  $t \geq 0$ , the function  $t \mapsto S_t(M)(\omega)$  is right-continuous and nondecreasing on  $[0, \infty)$  for all  $\omega \in \Omega$ , and  $S_t^2(M)$  is the limit in probability of  $(Q_{n,t})_{n \geq 0}$  as  $n \rightarrow \infty$  for all  $t \geq 0$ , where*

$$(6.1) \quad Q_{n,t} = |M_0|^2 + \sum_{k=0}^{2^n-1} |M_{t(k+1)2^{-n}} - M_{tk2^{-n}}|^2.$$

This was proved for real-valued martingales by Doléans in [9]. Her proof rests on Theorem 8 of [1]: if  $f$  is a discrete-time real-valued martingale and  $S(f)$  is its square function, then for all  $\lambda > 0$ ,

$$(6.2) \quad \lambda P(S(f) \geq \lambda) \leq c \|f\|_1.$$

This inequality carries over to Hilbert space with 2 as an admissible constant [2]. (In the real-valued case, the best constant is  $\sqrt{e}$  as Cox [6] proved in the 1980s. Littlewood's letter reproduced on page 18 of [19] indicates that Bollobás and Littlewood knew in 1975 that  $\sqrt{e}$  was the likely value.) Using the Hilbert-space version of (6.2), we see that the Doléans proof of the lemma for the real-valued case carries over with only a few simple changes, for example, with the substitution of the real part of the inner product for a product at some places.

In the following,  $M$  is a local martingale. Localization, introduced by Itô and Watanabe [15], plays an important role in martingale theory and its applications. A convenient reference is Dellacherie and Meyer [8]. A right-continuous process  $M$  with values in  $\mathbf{H}$  and with  $M_t$  measurable with respect to  $\mathcal{F}_t$  as above is a local



martingale if there are stopping times  $\tau_1 \leq \tau_2 \leq \dots \leq \infty$  such that  $\lim_{n \rightarrow \infty} \tau_n = \infty$  with probability 1 and the process  $M^n$  defined by

$$(6.3) \quad M_t^n = M_{t \wedge \tau_n} I_{\{\tau_n > 0\}}$$

is a uniformly integrable martingale for all  $n \geq 1$ . For example, every right-continuous martingale is a local martingale. Let  $M^*$  denote the maximal function of  $M$ :  $M^*(\omega) = \sup_{t \geq 0} |M_t(\omega)|$ . Note that with probability 1,  $\lim_{n \rightarrow \infty} M_t^n = M_t$ , so if  $0 \leq s < t$  and  $E[M^*]$  is finite, then by the dominated convergence theorem for conditional expectations and the martingale property of  $M^n$ ,

$$E[M_t | \mathcal{F}_s] = \lim_{n \rightarrow \infty} E[M_t^n | \mathcal{F}_s] = \lim_{n \rightarrow \infty} M_s^n = M_s$$

with probability 1. Therefore, the finiteness of  $E[M^*]$  implies that  $M$  is a right-continuous martingale. In the real case, the process  $(S_t^2(M))_{t \geq 0}$  of Lemma 6.1 is a version of the square bracket process  $[M, M]$ ; see Chapter VIII of [8]. In the following theorem,  $S(M)$  denotes the pointwise limit of  $S_t(M)$  as  $t \rightarrow \infty$ .

**Theorem 6.1.** *Let  $M$  be a local martingale with values in  $\mathbf{H}$ . Then*

$$(6.4) \quad E[S(M)] \leq \sqrt{3} E[M^*]$$

*and the constant  $\sqrt{3}$  is the best possible. If  $1 \leq p \leq 2$ , then*

$$(6.5) \quad \|S(M)\|_p \leq \gamma_p \|M^*\|_p$$

*where the positive real number  $\gamma_p$  is defined in (1.5).*

*Proof.* That  $\sqrt{3}$  is a lower bound for the admissible constants  $C$  in the inequality  $E[S(M)] \leq C E[M^*]$  follows from Theorem 1.1. Because  $\gamma_1 = \sqrt{3}$ , it remains to prove (6.5). In the proof, we can and do assume that  $E[M^*]$  is finite, and hence that  $M$  is a right-continuous martingale. If  $t > 0$  and  $f^n$  is the martingale defined by  $f_k^n = M_{tk/2^n}$  for  $k = 0, \dots, 2^n$  and by  $f_k^n = M_t$  for  $k > 2^n$ , then  $S^2(f^n) = Q_{n,t}$ . So (6.5) holds because, by Lemma 6.1, Fatou's Lemma, and Theorem 1.1,

$$\begin{aligned} \|S_t(M)\|_p &\leq \liminf_{n \rightarrow \infty} \|\sqrt{Q_{n,t}}\|_p = \liminf_{n \rightarrow \infty} \|S(f^n)\|_p \\ &\leq \gamma_p \liminf_{n \rightarrow \infty} \|(f^n)^*\|_p \leq \gamma_p \|M^*\|_p. \end{aligned}$$

□

## 7. A FAMILY OF SUPERMARTINGALES

As mentioned in Section 1, by using the function  $U$  of (2.6), we can embed the Lévy martingale  $(B_t^2 - t)_{t \geq 0}$  in a large family of submartingales. This is a special case of Theorem 7.1 below in which  $U$  again plays the leading role.

**Lemma 7.1.** *Let  $1 \leq p \leq 2$ ,  $\gamma \geq \gamma_p$ ,  $\alpha = \frac{p}{2}\gamma^{p-2}$ ,  $t \geq 0$ , and  $z > 0$ . Let  $f$  be a martingale with values in  $\mathbf{H}$  such that  $\|f^*\|_p$  is finite. If  $Y$  is defined by  $Y_n = U(f_n, t - |f_0|^2 + S_n^2(f), f_n^* \vee z)$ , then  $Y$  is a supermartingale.*

*Proof.* The function  $U$  satisfies condition (2.3) for all integrable  $d$  with  $E[d] = 0$ , not just for simple  $d$ : in the proof of Lemma 2.1, the function  $\Psi$  is affine so (2.8) continues to hold. Because  $U$  majorizes  $V$ ,

$$E[Y_n] \geq E[V(f_n, t - |f_0|^2 + S_n^2(f), f_n^* \vee z)] \geq -\gamma^p E[(f_n^* \vee z)^p] > -\infty.$$

The proof that  $E[Y_n] < \infty$  and  $E[Y_{n+1}|\mathcal{F}_n] \leq Y_n$  for all  $n \geq 0$  is by induction. If  $n = 0$ , then

$$(7.1) \quad E[Y_0] = E[U(f_0, t, |f_0| \vee z)] \leq \alpha t E[(|f_0| \vee z)^{p-2}] \leq \alpha t z^{p-2} < \infty.$$

By slightly modifying the proof of Theorem 2.1, we obtain  $E[Y_1|\mathcal{F}_0] \leq Y_0$ , which gives  $E[Y_1] \leq E[Y_0] < \infty$ , and so on. Therefore,  $Y$  is a supermartingale.  $\square$

**Theorem 7.1.** *Let  $1 \leq p \leq 2$ ,  $\gamma \geq \gamma_p$ ,  $\alpha = \frac{p}{2}\gamma^{p-2}$ ,  $y \geq 0$ , and  $z > 0$ . If  $M$  is a local martingale with values in  $\mathbf{H}$  such that  $\|M^*\|_p$  is finite, then  $Y$  defined by*

$$\begin{aligned} Y_t &= U(M_t, y - |M_0|^2 + S_t^2(M), M_t^* \vee z) \\ &= \alpha[(1 - \gamma^2)(M_t^* \vee z)^p - (|M_t|^2 - y + |M_0|^2 - S_t^2(M))(M_t^* \vee z)^{p-2}] \end{aligned}$$

*is a supermartingale. If  $p = 2$ ,  $\gamma = 1$ ,  $y = 0$ , and  $M_0 = 0$ , then  $\alpha = 1$  and  $Y$  is the martingale given by  $Y_t = -(|M_t|^2 - S_t^2(M))$ .*

*Proof.* Because  $\|M^*\|_p$  is finite, the local martingale  $M$  is a right-continuous martingale. Let  $0 \leq s < t$ , where we assume at first that  $s/t$  is a dyadically rational number: for all large  $n$ , the nonnegative number  $2^n s/t$  is an integer, which we denote by  $k_n$ . Let  $f^n$  be the martingale defined in the proof of Theorem 6.1. This martingale satisfies the conditions of Lemma 7.1 because  $(f^n)^* \leq M^*$ . Therefore,  $Y^n$ , defined by  $Y_k^n = U(f_k^n, y - |f_0^n|^2 + S_k^2(f^n), (f^n)^* \vee z)$ , is a supermartingale. By Lemma 6.1,  $\lim_{n \rightarrow \infty} Y_{2^n}^n = Y_t$  and  $\lim_{n \rightarrow \infty} Y_{k_n}^n = Y_s$ , both limits in probability, but with probability 1 for  $n$  in some infinite set  $N$  of positive integers. This supermartingale also has the property that  $\inf_{n,k} Y_k^n \geq -\gamma^p (M^* \vee z)^p$ , an integrable function. Therefore, by the Fatou Lemma for conditional expectations,

$$E[Y_t|\mathcal{F}_s] \leq \liminf_{N \ni n \rightarrow \infty} E[Y_{2^n}^n|\mathcal{F}_s] \leq \lim_{N \ni n \rightarrow \infty} Y_{k_n}^n = Y_s$$

with probability 1. The continuity of  $U$  and the right-continuity of both  $M$  and the filtration now yields the inequality  $E[Y_t|\mathcal{F}_s] \leq Y_s$  for all  $s \in [0, t)$ . As in the discrete-time case of Lemma 7.1,  $-\gamma^p E[(M^* \vee z)^p] \leq E[Y_t] \leq E[Y_0] \leq \alpha y z^{p-2}$ .  $\square$

This leads to a slightly different proof of (6.5). Assume that  $y = 0$ ,  $M_0 = 0$ , and  $\|M^*\| < \infty$ . Then  $E[V(M_t, S_t^2(M), M_t^* \vee z)] \leq E[Y_t] \leq 0$  which yields (6.5).

## 8. ON THE SEARCH FOR A SUITABLE MAJORANT

Here we describe some of the steps that led to the discovery of the function  $U$  defined in Lemma 2.1. It is the most convenient majorant of the function  $V$  of that lemma, but is not the least majorant of  $V$  that satisfies (2.2) and (2.3). Because we have already proved the main results of the paper, not all of the results contained in this section are proved in detail. The following lemma can also be used for related problems.

Let  $V : \mathbf{B} \times [0, \infty) \times (0, \infty) \rightarrow \mathbf{R}$  satisfy  $V(x, t, x) = V(x, t, |x| \vee z)$  as in Lemma 4.1. Then the function  $U$  defined there majorizes  $V$  as well as satisfying the two other conditions of Theorem 2.1. For its use in Section 5, no explicit analytic expression for this function is needed. The following iterative approach is helpful in the search for such an expression. Let  $U^0 = V$ . If  $n$  is a nonnegative integer and  $U^n$  satisfies  $U^n(x, t, x) = U^n(x, t, |x| \vee z)$ , let

$$(8.1) \quad U^{n+1}(x, t, z) = \sup E[U^n(x + d, t + |d|^2, |x| \vee |x + d| \vee z)],$$

where the supremum is over the family of all simple functions  $d$  satisfying  $E[d] = 0$ . The right side of (8.1) is not changed if  $z$  is replaced by  $|x| \vee z$ , so  $U^{n+1}(x, t, z) = U^{n+1}(x, t, |x| \vee z)$ . Take  $d = 0$  in the expectation to see that  $U^{n+1} \geq U^n$ . By induction,  $V = U^0 \leq U^1 \leq U^2 \leq \dots$  and each  $U^n$  satisfies (2.2).

**Lemma 8.1.** *The function  $U^\infty : \mathbf{B} \times [0, \infty) \times (0, \infty) \rightarrow (-\infty, \infty]$  defined by*

$$U^\infty(x, t, z) = \lim_{n \rightarrow \infty} U^n(x, t, z)$$

*satisfies (2.1), (2.2), and (2.3), and is the least such function.*

*Proof.* Clearly, both (2.1) and (2.2) are satisfied by  $U^\infty$ . If  $|x| \leq z$ , then

$$\begin{aligned} E[U^\infty(x + d, t + |d|^2, z)] &= E[U^\infty(x + d, t + |d|^2, |x| \vee |x + d| \vee z)] \\ &= \lim_{n \rightarrow \infty} E[U^n(x + d, t + |d|^2, |x| \vee |x + d| \vee z)] \\ &\leq \lim_{n \rightarrow \infty} U^{n+1}(x, t, z) = U^\infty(x, t, z), \end{aligned}$$

so (2.3) is also satisfied by  $U^\infty$ . The monotone convergence theorem can be used here because  $d$  is a simple function so  $U^0(x + d, t + |d|^2, z)$  is integrable. Let  $W$  be a function with the same domain as  $V$  such that (2.1), (2.2), and (2.3) are satisfied. The transformation (8.1) applied to both sides of  $V \leq W$  gives  $U^1 \leq W$  and, by induction,  $U^n \leq W$  for all  $n$ . Therefore,  $U^\infty \leq W$ .  $\square$

It is easy to see that the majorant defined in Lemma 4.1 is also the least. Consequently, it must be equal to  $U^\infty$ , but we shall not use that here. In fact, this section does not require any knowledge of the results of the previous sections.

If  $V$  satisfies a scaling property, then the problem of finding a suitable  $U$  is reduced in complexity. Such is the case with  $V : \mathbf{R} \times [0, \infty) \times (0, \infty) \rightarrow \mathbf{R}$  defined by

$$(8.2) \quad V(x, t, z) = \sqrt{t} - \gamma(|x| \vee z),$$

which satisfies

$$(8.3) \quad V(x, t, z) = |\lambda|^{-1} V(\lambda x, \lambda^2 t, |\lambda| z) \quad \text{for } \lambda \neq 0.$$

One of the majorants that we shall find for this case leads to the majorant (2.6) that we use for the Hilbert-space case with  $1 \leq p \leq 2$ .

**Lemma 8.2.** *Let  $v$  be defined on  $[-1, 1] \times [0, \infty)$  by  $v(x, t) = \sqrt{t} - \gamma$ . There is a majorant  $U : \mathbf{R} \times [0, \infty) \times (0, \infty) \rightarrow \mathbf{R}$  of  $V$  satisfying (2.2) and (2.3) if and only if there is a majorant  $u : [-1, 1] \times [0, \infty) \rightarrow \mathbf{R}$  of  $v$  such that*

$$(8.4) \quad u(x, t) = u(-x, t),$$

$$(8.5) \quad u(x, t) \geq \lambda u(x + h, t + h^2) + (1 - \lambda)u(x + k, t + k^2) \quad \text{if } 0 \leq \lambda \leq 1, \\ |x + h| \leq 1, \quad |x + k| \leq 1, \quad \text{and } \lambda h + (1 - \lambda)k = 0,$$

$$(8.6) \quad u(x, t) \geq u(x - h, t + h^2) + u(1, 1)h \quad \text{if } 0 < h \leq 1 + x,$$

$$(8.7) \quad s \mapsto u(x, s^2) \quad \text{is convex on } \mathbf{R}.$$

There are several other properties of a finite majorant  $u$  of  $v$  that are needed later and which follow immediately from the four above:

$$(8.8) \quad u(x, \cdot) \text{ is continuous and nondecreasing on } [0, \infty],$$

$$(8.9) \quad u(\cdot, t) \text{ is concave on } [-1, 1],$$

$$(8.10) \quad u(1, 1) \leq 0.$$

Property (8.7) implies (8.8). Therefore, the right side of (8.5) is greater than or equal to  $\lambda u(x + h, t) + (1 - \lambda)u(x + k, t)$  so (8.9) holds. If  $x = t = 1$  and  $h = 2$  in (8.6), then  $u(1, 1) \geq u(-1, 1 + 2^2) + 2u(1, 1) \geq u(-1, 1) + 2u(1, 1) = 3u(1, 1)$  and this yields (8.10). Here we have used (8.8) once more and (8.4).

*Proof.* The “only if” part of the proof is not difficult. A majorant  $W$  of  $V$  that maps into  $\mathbf{R}$  and satisfies (2.2) and (2.3) can be replaced by such a majorant  $U$  that, in addition, satisfies the scaling property (8.3) that  $V$  satisfies, for example, the majorant given by  $U(x, t, z) = \inf_{\alpha \neq 0} |\alpha|^{-1} W(\alpha x, \alpha^2 t, |\alpha|z)$ . So  $U^\infty$ , the result of iteration starting with  $V$ , has this scaling property because of its minimality: if  $W = U^\infty$ , then  $U^\infty \leq U \leq W$ . With this choice, the function  $u$  defined on  $[-1, 1] \times [0, \infty)$  by  $u(x, t) = U(x, t, 1)$  satisfies the desired properties: (8.4) and (8.5) follow at once from the corresponding properties of  $U$ . The convexity property (8.7) follows from the convexity of the function  $s \mapsto V(x, s^2, z)$  which carries over by induction to  $s \mapsto U^\infty(x, s^2, z)$ . If  $0 < h \leq 1 + x$  and  $r > 1$ , then by (2.2), (2.3), and the scaling property of  $U$ ,

$$u(x, t) \geq \frac{hr}{r - x + h} u(1, (t + (r - x)^2)/r^2) + \frac{r - x}{r - x + h} u(x - h, t + h^2).$$

The continuity of  $u(1, \cdot)$  gives the equality of the limit of the right side as  $r \rightarrow \infty$  with the right side of (8.6). This completes the “only if” part of the proof.

For the “if” part of the proof, suppose that  $u$  is a finite majorant of  $v$  satisfying (8.4), ..., (8.7). Define  $U$  on  $\mathbf{R} \times [0, \infty) \times (0, \infty)$  by

$$U(x, t, z) = (|x| \vee z) u\left(\frac{x}{|x| \vee z}, \frac{t}{(|x| \vee z)^2}\right).$$

It follows at once that  $U$  is a finite majorant of  $V$  and satisfies (2.2). To see that (2.3) also holds, it is enough, by the scaling property, to prove this for  $z = 1$ . Fix  $(x, t) \in [-1, 1] \times [0, \infty)$  and define  $\Phi$  on  $\mathbf{R}$  by  $\Phi(s) = U(x + s, t + s^2, 1)$ . As we shall show, there is an affine majorant  $\Psi$  of  $\Phi$  satisfying  $\Psi(0) = \Phi(0)$ . This implies (2.3): if  $d$  is a simple function with  $E[d] = 0$ , then  $E[\Phi(d)] \leq E[\Psi(d)] = \Psi(E[d]) = \Psi(0) = \Phi(0) = U(x, t, 1)$ . If  $x \in [-1, 1)$ , let  $R = \sup_{0 < k \leq 1-x} [\Phi(k) - \Phi(0)]/k$ ; clearly,  $R > -\infty$ . If  $x \in (-1, 1]$ , let  $L = \inf_{0 < h \leq 1+x} [\Phi(0) - \Phi(-h)]/h$ ; here  $L < \infty$ . If  $x \in (-1, 1)$ , then  $R \leq L$  by (8.5) so  $(R + L)/2$ , denoted in this case by  $\beta$ , is finite. If  $x = 1$ , let  $\beta$  denote  $L$ , which is greater than or equal to  $u(1, 1)$  by (8.6). If  $x = -1$ , let  $\beta = R$ ; here, by symmetry,  $\beta \leq -u(1, 1)$ . So  $\beta$  is finite for all  $x \in [-1, 1]$  and  $t \geq 0$ . Define  $\Psi$  on  $\mathbf{R}$  by

$$\Psi(s) = \Phi(0) + \beta s.$$

Then  $\Psi$  does majorize  $\Phi$  at least on the set  $[-1 - x, 1 - x]$  as easily follows from the definitions above. To prove that

$$(8.11) \quad \Phi(s) \leq \Psi(s)$$

holds also on the complement of this interval, we will use the convexity of  $\Phi$  on the interval  $[1-x, \infty)$  and on  $(-\infty, -1-x]$ . To prove the convexity on  $[1-x, \infty)$ , let  $s_1$  and  $s_2$  belong to this interval and  $s = \sum \alpha_k s_k$ , where the  $\alpha_k$  are nonnegative numbers adding to 1. Let  $w$  be the function defined by  $w(s) = u(1, s^2)$ , a convex function on  $\mathbf{R}$  by (8.7), and  $r_k = [(t + s_k^2)/(x + s_k)^2]^{1/2}$ . By (2.2), scaling, and the convexity of both  $w$  and the function  $s \mapsto \sqrt{t + s^2}$ ,

$$\begin{aligned}
\sum \alpha_k \Phi(s_k) &= \sum \alpha_k (x + s_k) w(r_k) \\
&= (x + s) \sum \alpha_k \frac{x + s_k}{x + s} w(r_k) \\
&\geq (x + s) w \left( \sum \alpha_k \frac{x + s_k}{x + s} r_k \right) \\
&= (x + s) w \left( \frac{1}{x + s} \sum \alpha_k \sqrt{t + s_k^2} \right) \\
&\geq (x + s) w \left( \frac{1}{x + s} \sqrt{t + s^2} \right) \\
&= (x + s) u \left( 1, \frac{t + s^2}{(x + s)^2} \right) \\
&= \Phi(s).
\end{aligned}$$

So  $\Phi$  is convex on  $[1-x, \infty)$ . Using this convexity, we see that

$$\frac{\Phi(s) - \Phi(1-x)}{s - (1-x)} \uparrow u(1, 1) \quad \text{as } s \uparrow \infty.$$

In view of (8.10), this implies that  $\Phi(s) - \Phi(1-x) \leq 0$  for all  $s > 1-x$ . Because we have already proved that (8.11) holds on the interval  $[-1-x, 1-x]$ , we have that for  $s > 1-x$ ,  $\Phi(1-x) - \Phi(0) \leq \Psi(1-x) - \Psi(0) \leq \beta s$ . So for  $s > 1-x$ ,

$$\begin{aligned}
\Phi(s) &= \Phi(0) + [\Phi(s) - \Phi(1-x)] + [\Phi(1-x) - \Phi(0)] \\
&\leq \Phi(0) + \beta s = \Psi(s).
\end{aligned}$$

Similarly,  $\Phi(s) \leq \Psi(s)$  for all  $s < -1-x$ . Therefore, (8.11) holds for all  $s \in \mathbf{R}$  and the proof is complete.  $\square$

We continue with the special case of  $V$  as given by (8.2). The reduction of the problem of finding a majorant  $U$  of  $V$  on  $\mathbf{R} \times [0, \infty) \times (0, \infty)$  to finding a majorant  $u$  of  $v$  on  $[-1, 1] \times (0, \infty)$  permits an easier computer search for a good approximation to such a  $u$ , one that will give some information about the least majorant  $U$  of  $V$  even if the grid chosen is quite sparse, say

$$\{(j\delta, k\delta^2) : j = -m, \dots, m; k = 0, 1, \dots, 8m^2\}$$

with  $\delta = 1/m$  and  $m = 6$ . By starting with  $u^0$ , the restriction of  $v$  to this grid, the  $n^{\text{th}}$  iteration  $u^n$  for moderate  $n$  suggests the following possibilities:

(i) Let  $t = k\delta^2$ . If  $x = 0$ , then  $u^\infty(x, t) = u^\infty(1, t+1)$ . If  $j \in \{1, \dots, m-1\}$  and  $x = j\delta$ , then

$$(8.12) \quad u^\infty(x, t) = \frac{1-x}{1-x+\delta} u^\infty(x-\delta, t+\delta^2) + \frac{\delta}{1-x+\delta} u^\infty(1, t+(1-x)^2).$$

(ii) There is a number  $\theta(m)$  such that if  $\gamma < \theta(m)$ , then  $u^\infty(1, 1) \geq u^n(1, 1) > 0$ , which implies that  $u(1, 1) > 0$ . So, in this case, (8.10) and Lemma 8.2 imply that

$U$  does not map into  $\mathbf{R}$ . On the other hand, if  $\gamma \geq \theta(\infty)$ , the limit of  $\theta(m)$  as  $m \rightarrow \infty$ , then  $U$  is finite.

(iii) Suppose that  $U$  is the least majorant of  $V$  where  $\gamma$  is large enough so that  $U$  is finite. Let  $B$  be defined on  $[0, \infty)$  by  $B(t) = u(1, t)$ . Then the approximation of  $B(k\delta^2)$  by  $u^n(1, k\delta^2)$  suggests that

$$(8.13) \quad B(t) = \sqrt{t} - \gamma \quad \text{if } t > \gamma^2 \quad \text{and} \quad B(t) = \alpha(t - \gamma^2) \quad \text{if } 0 \leq t \leq \gamma^2,$$

where  $\alpha$  is the number that makes  $B$  differentiable at  $\gamma^2$ , that is,  $\alpha = (2\gamma)^{-1}$ .

All of this leads to the boundary value problem of finding a function  $u$  on the set  $[-1, 1] \times [0, \infty)$  satisfying (8.4), (8.5), (8.6), (8.7), as well as the boundary condition  $u(1, t) = B(t)$  for all  $t \geq 0$ .

**Lemma 8.3.** *Let  $B : [0, \infty) \rightarrow \mathbf{R}$  be a continuous function,  $m$  a positive integer,  $\delta = 1/m$ ,  $G_m = \{(j\delta, k\delta^2) : |j| \leq m, k \geq 0\}$ , and  $(x, t) \in G_m$  with  $x \geq 0$ . Let  $f$  denote a martingale  $(f_n)_{n \geq 0}$  such that  $f_0 = x$ ,  $f_{n+1} \in \{1, -1\}$  on the set  $\{f_n = 0\}$ ,  $f_{n+1} \in \{1, (j-1)\delta\}$  on  $\{f_n = j\delta\}$  for  $0 < j < m$ ,  $f_{n+1} = 1$  on  $\{f_n = 1\}$ , and  $f_{n+1} = -1$  on  $\{f_n = -1\}$ . Then  $f_n = f_m$  for all  $n \geq m$ . If  $x = 0$ , then  $P(S^2(f) - x^2 = 1) = 1$ ; if  $0 < x < 1$  and  $0 \leq j\delta < x$ , then*

$$P(S^2(f) - x^2 = j\delta^2 + (1 - x + j\delta)^2) = \frac{1 - x}{1 - x + j\delta} \cdot \frac{\delta}{1 - x + (j+1)\delta};$$

if  $0 < x < 1$  and  $j\delta = x$ , then

$$P(S^2(f) - x^2 = j\delta^2 + 1) = 1 - x;$$

and if  $x = 1$ , then  $P(S^2(f) - x^2 = 0) = 1$ . Moreover, as  $m$  increases to infinity through the set  $\{2^n : n \geq 0\}$ , the expectation of  $B(t - x^2 + S^2(f))$  converges to

$$(8.14) \quad (1 - x) \left[ \int_{1-x}^1 \frac{B(t + y^2)}{y^2} dy + B(t + 1) \right]$$

or to  $B(t)$  for all  $(x, t) \in \bigcup_{n \geq 0} G_{2^n}$ , to (8.14) if  $0 \leq x < 1$  and to  $B(t)$  if  $x = 1$ .

The proof is straightforward and is omitted, but note that  $(x, t) \in G_{2^n}$  for all large  $n$  and that for  $0 < x < 1$  and  $\delta$  small compared to  $1 - x$ , the martingale moves to the left in small steps of size  $\delta$  each with large probability and to 1 with small probability. If it arrives at 0 before 1, it jumps to 1 or  $-1$  at the next step where it stays. In view of (i), the martingale  $f$  is a natural choice.

**Theorem 8.1.** *Let  $B$  be defined on  $[0, \infty)$  by  $B(t) = \alpha(t - \gamma^2)$ , where  $\gamma > 0$  and  $\alpha = (2\gamma)^{-1}$ . Define  $u$  on  $[-1, 1] \times [0, \infty)$  by (8.14) if  $0 \leq x < 1$ , by  $B(t)$  if  $x = 1$ , and by  $u(-x, t)$  if  $-1 \leq x < 0$ . Then, on  $[-1, 1] \times [0, \infty)$ ,*

$$(8.15) \quad u(x, t) = \alpha(t - x^2 - \gamma^2 + 1).$$

*This function  $u$  is a majorant of the function  $v$  of Lemma 8.2 and satisfies (8.4), (8.5), and (8.7) of that lemma. However, it satisfies (8.6) if and only if  $\gamma \geq \sqrt{3}$ .*

The easy proof is omitted. This function  $u$  yields the  $U$  of Lemma 2.1 for the case  $p = 1$  and  $\mathbf{H} = \mathbf{R}$ . It makes possible the sharp version of the Davis inequality in the real case. A smaller  $u$  and  $U$ , less easy to work with, can be obtained by replacing the  $B$  of Theorem 8.1 by the smaller  $B$  of (8.13). But the function  $u$  of Theorem 8.1 and its associated  $U$  also suggest experimentation with the function  $U$  of Lemma 2.1. Indeed, for  $1 \leq p \leq 2$  and martingales with values in  $\mathbf{H}$ , this  $U$  provides the key to the main results of the paper.

## REFERENCES

- [1] D. L. Burkholder, *Martingale transforms*, Ann. Math. Statist. **37** (1966), 1494–1504. MR **34**:8456
- [2] D. L. Burkholder, *Sharp inequalities for martingales and stochastic integrals*, Astérisque **157–158** (1988), 75–94. MR **90b**:60051
- [3] D. L. Burkholder, *Sharp norm comparison of martingale maximal functions and stochastic integrals*, Proceedings of the Norbert Wiener Centenary Congress, 1994 (East Lansing, MI), Proc. Sympos. Appl. Math. **52** (1997), 343–358. MR **98f**:60103
- [4] D. L. Burkholder and R. F. Gundy, *Extrapolation and interpolation of quasi-linear operators on martingales*, Acta Math. **124** (1970), 249–304. MR **55**:13567
- [5] E. A. Carlen and P. Krée, *On martingale inequalities in non-commutative stochastic analysis*, J. Funct. Anal. **158** (1998), 475–508. MR **99g**:81111
- [6] D. C. Cox, *The best constant in Burkholder’s weak- $L^1$  inequality for the martingale square function*, Proc. Amer. Math. Soc. **85** (1982), 427–433. MR **84g**:60079
- [7] B. Davis, *On the integrability of the martingale square function*, Israel J. Math. **8** (1970), 187–190. MR **42**:3863
- [8] C. Dellacherie and P.-A. Meyer, *Probabilities and potential B: Theory of martingales*, North Holland, Amsterdam, 1982. MR **85e**:60001
- [9] C. Doléans, *Variation quadratique des martingales continues à droite*, Ann. Math. Statist. **40** (1969), 284–289. MR **38**:5275
- [10] J. L. Doob, *Stochastic processes*, Wiley, New York, 1953. MR **15**:445b
- [11] C. Fefferman, *Characterization of bounded mean oscillation*, Bull. Amer. Math. Soc. **77** (1971), 587–588. MR **43**:6713
- [12] C. Fefferman and E. M. Stein,  *$H^p$  spaces of several variables*, Acta Math. **129** (1972), 137–193. MR **56**:6263
- [13] A. M. Garsia, *The Burgess Davis inequalities via Fefferman’s inequality*, Ark. Mat. **11** (1973), 229–237. MR **42**:8267
- [14] A. M. Garsia, *Martingale Inequalities: Seminar Notes on Recent Progress*, Benjamin, Reading, Massachusetts, 1973. MR **56**:6844
- [15] K. Itô and S. Watanabe, *Transformation of Markov processes by multiplicative functionals*, Ann. Inst. Fourier **15** (1965), 15–30. MR **32**:1755
- [16] A. Khintchine, *Über dyadische Brüche*, Math. Z. **18** (1923), 109–116.
- [17] P. Lévy, *Processus stochastiques et mouvement Brownien*, Gauthier-Villars, Paris, 1948. MR **10**:551a
- [18] J. E. Littlewood, *On bounded bilinear forms in an infinite number of variables*, Quart. J. Math. Oxford **1** (1930), 164–174.
- [19] J. E. Littlewood, *Littlewood’s miscellany*, edited and with a foreword by Béla Bollobás, Cambridge University Press, Cambridge–New York, 1986. MR **88d**:01036
- [20] J. Marcinkiewicz, *Quelques théorèmes sur les séries orthogonales*, Ann. Soc. Polon. Math. **16** (1937), 84–96.
- [21] J. Marcinkiewicz and A. Zygmund, *Quelques théorèmes sur les fonctions indépendantes*, Studia Math. **7** (1938), 104–120.
- [22] R. E. A. C. Paley, *A remarkable series of orthogonal functions I*, Proc. London Math. Soc. **34** (1932), 241–264.
- [23] G. Pisier and Q. Xu, *Non-commutative martingale inequalities*, Commun. Math. Phys. **189** (1997), 667–698. MR **98m**:46079
- [24] E. M. Stein, *The development of square functions in the work of A. Zygmund*, Bull. Amer. Math. Soc. **7** (1982), 359–376. MR **83i**:42001
- [25] E. M. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton University Press, Princeton, NJ, 1993. MR **95e**:42002

DEPARTMENT OF MATHEMATICS, 273 ALTGELD HALL, 1409 WEST GREEN STREET, UNIVERSITY OF ILLINOIS, URBANA, ILLINOIS 61801

*E-mail address*: donburk@math.uiuc.edu