PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 85, Number 3, July 1982

THE BEST CONSTANT IN BURKHOLDER'S WEAK-L^1 INEQUALITY FOR THE MARTINGALE SQUARE FUNCTION

DAVID C. COX

ABSTRACT. Let $Y_1, Y_2, \ldots$ be a martingale with difference sequence $X_1 = Y_1$, $X_i = Y_i - Y_{i-1}$, $i \geq 2$. We give a new proof of the inequality

$$P\left(\sum_{i\geq 1} X_i^2 \geq \lambda^2\right) \leq \lambda^{-1} C \sup_{i\geq 1} E|Y_i|,$$

for all $y > 0$, and show that the best constant is $C = e^{1/2}$.

Introduction. Let $Y_1, Y_2, \ldots$ be a martingale, i.e. $E(Y_i|Y_1, \ldots, _{i-1}) = Y_{i-1}$, for $i = 2, 3, \ldots$, and define its difference sequence by $X_1 = Y_1$, $X_i = Y_i - Y_{i-1}$, for $i = 2, 3, \ldots$. Thus, $E(X_i|X_1, \ldots, X_{i-1}) = 0$, for $i = 2, 3, \ldots$. By the weak-L^1 inequality for the martingale square function, we mean the inequality

$$P\left(\sum_{i\geq 1} X_i^2 \geq \lambda^2\right) \leq \lambda^{-1} C \sup_{i\geq 1} E|Y_i|, \tag{1}$$

for all $y > 0$, where C is an absolute constant. Inequality (1) was first proved by Burkholder [**2**], without an explicit estimate of C. In [**3**], he showed that one can take $C = 3$, and in [**4**] improved that estimate to $C = 2$. Bollobas [**1**] gave a different proof that $C = 2$ works in (1) and showed by example that the best value of C is at least $3/2$.

In this paper, it is shown that $C = e^{1/2} = 1.648721271\ldots$ is the best constant in (1). More precisely, we prove that the inequality

$$P\left(\sum_{i=1}^{n} X_i^2 \geq 1\right) \leq \left(\frac{n}{n-1}\right)^{(n-1)/2} E|Y_n|, \quad n = 2, 3, \ldots, \tag{2}$$

is sharp. The proof of (2) is by induction, using methods from the theory of moments, together with the device of conditioning. For each n, an explicit example realizing equality in (2) is constructed.

The methods of this paper are applicable to a number of other martingale inequalities and, where applicable, provide sharp constants. The complexity of the necessary computations is, however, often a drawback.

1. Preliminaries. The results of this paper are based on the following elementary lemma, a special case of a general theorem in the theory of moments [**5**]. Its proof is straightforward.

LEMMA 1.1. *Let f be a real-valued measurable function defined on a measurable set $A \subseteq R$. Define the moment space M as the set of all points $(EX, Ef(X))$, where*

Received by the editors May 1, 1981 and, in revised form, September 25, 1981.
1980 *Mathematics Subject Classification*. Primary 60E15, 60G42; Secondary 42B25.

X is a random variable taking values in A for which $E|X|$ and $E|f(X)|$ are finite. Then M is precisely the convex hull of the graph of f.

COROLLARY 1.1. *Let f be as in Lemma 1.1, and assume $\inf A < 0 < \sup A$. Then the number $\inf\{Ef(X):\ X$ takes values in A with $EX = 0\}$ is the height, at location $x = 0$, of the lower boundary of the convex hull of the graph of f. The infimum can thus be determined by examining inequalities of the form $f(x) \geq ax+b$, for all $x \in A$.*

2. Main results. We start by proving inequality (2) in the special case $n = 2$. Define

$$\phi_1(t) = \inf\{E|t+X|:\ EX = 0,\ t^2 + X^2 \geq 1 \text{ a.e.}\}.$$

Since $E|t+X| \geq |t|$ for $EX = 0$, we have $\phi_1(t) \geq |t|$. Thus, $\phi_1(t) = |t|$ for $t^2 \geq \frac{1}{2}$, as we see by taking $P(X = \pm\max(1-t^2,\ 0)^{1/2}) = \frac{1}{2}$. Thus, we may assume $t^2 < \frac{1}{2}$. By Corollary 1.1, we must determine the lower boundary of the convex hull of the graph of $f(x) = |x+t|$, for $|x| \geq (1-t^2)^{1/2}$. The required lower boundary has equation $y = (1-t^2)^{1/2} + t(1-t^2)^{-1/2}x$. Thus, $\phi_1(t) = (1-t^2)^{1/2}$. We have $E|X+t| = (1-t^2)^{1/2}$, $EX = 0$ and $X^2+t^2 \geq 1$ when $P(X = \pm(1-t^2)^{1/2}) = \frac{1}{2}$. These results are based on the inequality

$$|t+x| \geq (1-t^2)^{1/2} + t(1-t^2)^{-1/2}x \tag{3}$$

which is valid for $t^2 < \frac{1}{2}$ and $x^2 + t^2 \geq 1$. Inequality (3) easily extends to

$$|t+x| \geq a_1(t) + b_1(t)1_{[1,\infty)}(t^2+x^2) + c_1(t)x, \tag{4}$$

valid for all t, x. Here, 1_A denotes the function which is 1 on the set A and 0 elsewhere. The coefficients in (4) are given by

$$\begin{aligned} a_1(t) &= t^2(1-t^2)^{-1/2}, \quad \text{if } t^2 < \tfrac{1}{2},\\ &\qquad |t|, \quad \text{if } t^2 \geq \tfrac{1}{2}.\\ b_1(t) &= (1-2t^2)(1-t^2)^{-1/2}, \quad \text{if } t^2 < \tfrac{1}{2},\\ &\qquad 0, \quad \text{if } t^2 \geq \tfrac{1}{2}.\\ c_1(t) &= t(1-t^2)^{-1/2}, \quad \text{if } t^2 < \tfrac{1}{2},\\ &\qquad \operatorname{sgn} t, \quad \text{if } t^2 \geq \tfrac{1}{2}. \end{aligned} \tag{5}$$

Hence, if $EX = 0$, we have

$$E|t+X| \geq a_1(t) + b_1(t)P(t^2+X^2 \geq 1) \geq \phi_1(t)P(t^2+X^2 \geq 1),$$

since $a_1(t) \geq 0$ and $a_1(t) + b_1(t) = \phi_1(t)$. We summarize these results in

THEOREM 2.1. *Let $K(t)$ be the best constant in the inequality*

$$P(t^2+X^2 \geq 1) \leq K(t)E|t+X| \tag{6}$$

for $EX = 0$. Then

$$\begin{aligned} K(t) &= (1-t^2)^{-1/2}, \quad \textit{if } t^2 < \tfrac{1}{2},\\ &\qquad |t|^{-1}, \quad \textit{if } t^2 \geq \tfrac{1}{2}. \end{aligned}$$

Moreover, equality is attained in (6) by a random variable X for which $P(t^2+X^2 \geq 1) = 1$.

COROLLARY 2.1. *Let* $E(X_2|X_1) = 0$. *Then*

$$P(X_1^2 + X_2^2 \geq 1) \leq 2^{1/2} E|X_1 + X_2|,$$

with equality for $X_1 \equiv 2^{-1/2}$, $P(X_2 = \pm 2^{-1/2}) = \frac{1}{2}$.

PROOF. Since $K(t) \leq 2^{1/2}$ for all t, we have, from Theorem 2.1,

$$P(X_1^2 + X_2^2 \geq 1 | X_1 = x) \leq 2^{1/2} E(|X_1 + X_2| \, |X_1 = x) \quad \text{for all } x.$$

Integrating with respect to the distribution of X_1 gives the result.

We have shown that equality is attained in (2), when $n = 2$, by a martingale for which $X_1^2 + X_2^2 \geq 1$ a.e. It seems reasonable to conjecture that, in proving (2) for general n, one may assume $\sum_{i=1}^n X_i^2 \geq 1$ a.e. This conjecture is proved in the appendix. We therefore define, for $n \geq 1$,

$$\phi_n(t) = \inf \left\{ E\left| t + \sum_{i=1}^n X_i \right| : EX_1 = 0, E(X_i | X_1, \ldots, X_{i-1} = 0), \right.$$
$$\left. \text{for } i = 2, \ldots, n, \; t^2 + \sum_{i=1}^n X_i^2 \geq 1 \text{ a.e.} \right\}.$$

THEOREM 2.2.

$$\phi_n(t) = \begin{cases} (n-1)^{(n-1)/2}(1-t^2)^{n/2}(n-(n+1)t^2)^{-(n-1)/2}, & \text{if } t^2 < \frac{1}{2}, \\ |t|, & \text{if } t^2 \geq \frac{1}{2}. \end{cases}$$

PROOF. We have already established the result for $n = 1$, so assume, by induction, that it holds for some $n \geq 1$. First, for $t^2 \geq \frac{1}{2}$, it is clear that $\phi_{n+1}(t) = |t|$: just take $P(X_1 = \pm \max(1 - t^2, 0)^{1/2}) = \frac{1}{2}$, while $X_2 \equiv \cdots \equiv X_{n+1} \equiv 0$. Thus, assume $t^2 < \frac{1}{2}$, and, since $\phi_{n+1}(-t) = \phi_{n+1}(t)$, $t \geq 0$. Observe that $t^2 + \sum_{i=1}^{n+1} X_i^2 \geq 1$ if and only if $(t + X_1)^2 + \sum_{i=2}^{n+1} X_i^2 \geq 1 + 2tX_1$. Conditioning on the variable $X_1 = X$ we find $\phi_{n+1}(t) = \inf\{Ef(X): \; EX = 0\}$, where

$$f(x) = \begin{cases} (1+2tx)^{1/2}\phi_n((t+x)(1+2tx)^{-1/2}), & \text{if } x_- \leq x \leq x_+, \\ |t+x|, & \text{elsewhere.} \end{cases} \tag{7}$$

Here, $x_- < 0$ and $x_+ > 0$ are the roots of the quadratic equation $(t+x)^2 = (1+2tx)/2$.

Let $t < s \leq 2^{-1/2}$ and denote by $x_+(t, x) > 0$ and $x_-(t, x) < 0$ the roots of $(t+x)^2 = s^2(1+2tx)$. Both roots lie in the interval $[x_-, x_+]$. Let X_s be a random variable taking only the two values $x_+(t, s)$, $x_-(t, s)$, with $EX_s = 0$. We find

$$Ef(X_s) = (1-t^2)(1-2t^2+s^2t^2)^{-1/2}\phi_n(s). \tag{8}$$

Differentiating, we find that (8) is minimized for $t < s \leq 2^{-1/2}$ when

$$s^2 = s_n^2(t) = (1 + (n-2)t^2)(n+1-2t^2)^{-1}.$$

Straightforward but tedious calculations using the properties of $x_+(t, s)$ and $x_-(t, s)$, and the relation

$$\frac{d}{ds}(1-t^2)(1-2t^2+s^2t^2)^{-1/2}\phi_n(s)|_{s=s_n(t)} = 0,$$

establish the following identities:

$$f'(x_+(t, s_n(t))) = f'(x_-(t, s_n(t))), \tag{10}$$

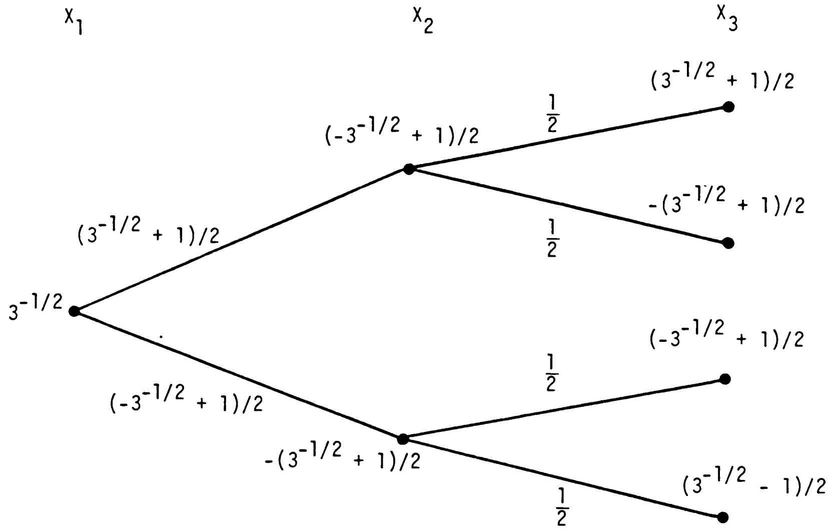

FIGURE 1. A martingale difference sequence attaining equality in (2), for $n=3$

$$(x_+(t,s_n(t)) - x_-(t,s_n(t)))^{-1}(f(x_+(t,s_n(t))) - f(x_-(t,s_n(t)))) = f'(x_+(t,s_n(t))) = f'(x_-(t,s_n(t))). \tag{11}$$

Thus, there is a common tangent to the graph of $f(x)$ with contact points at $x = x_\pm(t,s_n(t))$. Its equation is

$$y = (1-2t^2+s_n^2(t)t^2)^{-1/2}\phi_n(s_n(t))(tx+1-t^2). \tag{12}$$

Again, some long-winded calculations show that $f(x) \geq y$ for all x. Hence, if $EX=0$ we have

$$Ef(X) \geq (1-t^2)(1-2t^2+s_n^2(t)t^2)^{-1/2}\phi_n(s_n(t)). \tag{13}$$

Equality is attained by a random variable supported by the points $x_\pm(t,s_n(t))$. The right-hand side of (13) is easily shown to equal $\phi_{n+1}(t)$ as given in the statement of the theorem.

THEOREM 2.3. *Let $X_1,\dots,X_n$ be a martingale difference sequence. Then,*

$$P\left(\sum_{i=1}^n X_i^2 \geq 1\right) \leq \left(\frac{n}{n-1}\right)^{(n-1)/2} E\left|\sum_{i=1}^n X_i\right|, \tag{14}$$

and this inequality is sharp.

PROOF. As shown in the appendix, we may assume $\sum_{i=1}^n X_i^2 \geq 1$ a.e. Setting $\phi'_{n-1}(t)=0$ we find that the minimum value of $\phi_{n-1}(t)$ is $((n-1)/n)^{(n-1)/2}$, attained for $t^2=\frac{1}{n}$. Hence,

$$P\left(\sum_{i=1}^n X_i^2 \geq 1 \,\middle|\, X_1 = x\right) \leq \left(\frac{n}{n-1}\right)^{(n-1)/2} E\left(\left|\sum_{i=1}^n X_i\right| \,\middle|\, X_1 = x\right)$$

for all x. Integrating with respect to the distribution of X_1 gives (14). The sharpness of (14) is implied by the proof of Theorem 2.2. An example attaining equality in (14) can be constructed as follows. Let $t^2 < \frac{1}{2}$ and define $t^{(1)} = t$, $t^{(i)} = s_{n-i}(t^{(i-1)})$, for $i = 2, \ldots, n$. We have $t^{(1)} < t^{(2)} < \cdots < t^{(n-1)} = 2^{-1/2} < t^{(n)} = 1$. Let the random variable Z_i have expected value 0 and take only the two values $x_\pm(t^{(i)}, t^{(i+1)})/t^{(i)}$. Let $Z_1, \ldots, Z_{n-1}$ be independent. Finally, define $X_1 \equiv t$, $X_{i+1} = (\sum_{k=1}^{i} X_k)Z_i$, for $i = 1, \ldots, n-1$. Then $X_1, \ldots, X_n$ is a martingale difference sequence, which attains equality in (14) for $t = n^{-1/2}$.

EXAMPLE 2.1. Let $n = 3$, $t = 3^{-1/2}$. Then $t^{(1)} = 3^{-1/2}$, $t^{(2)} = 2^{-1/2}$, $t^{(3)} = 1$. We find

$$x_\pm(t^{(1)}, t^{(2)}) = (-3^{-1/2} \pm 1)/2 \quad \text{and} \quad x_\pm(t^{(2)}, t^{(3)}) = \pm 2^{-1/2}.$$

The sequence X_1, X_2, X_3 is shown schematically in Figure 1. It is easily checked that $X_1^2 + X_2^2 + X_3^2 \equiv 1$, while $E|X_1 + X_2 + X_3| = \frac{2}{3}$.

Appendix. The purpose of this appendix is to show that, in proving the fundamental inequality (2), we may assume that $\sum_{i=1}^n X_i^2 \geq 1$ a.e. First, we observe that the proof of Theorem 2.2 is implicitly based on an inequality of the form

$$\left| t + \sum_{i=1}^n x_i \right| \geq \phi_n(t) + \sum_{i=1}^n x_i \psi_{n,i}(t;\, x_1, \ldots, x_{i-1}), \tag{15}$$

valid as long as $t^2 + \sum_{i=1}^n x_i^2 \geq 1$. For $n = 1$, $\psi_{1,1}(t) = c_1(t)$ is given by (5). In general, it is straightforward to show that the functions $\psi_{n,i}$ may be defined as follows:

$$\begin{aligned} \psi_{n,1}(t) = t(1-t^2)^{-1}\phi_n(t), &\quad \text{if } t^2 < \tfrac{1}{2}, \\ \operatorname{sgn} t, &\quad \text{if } t^2 \geq \tfrac{1}{2}. \end{aligned}$$

$$\begin{aligned} &\psi_{n+1,i}(t; x_1, \ldots, x_{i-1}) \\ &\quad = \psi_{n,i-1}((t+x_1)(1+2tx_1)^{-1/2}; x_2(1+2tx_1)^{-1/2}, \\ &\qquad\qquad \ldots, x_{i-1}(1+2tx_1)^{-1/2}) \\ &\qquad\qquad \text{if } t^2 < \tfrac{1}{2} \text{ and } x_- < x_1 < x_+, \\ &\quad\ \operatorname{sgn}\left(t + \sum_{j=1}^{i-1} x_j\right), \quad \text{elsewhere.} \end{aligned}$$

The first relation follows from (12) and holds for $n = 1, 2, \ldots$. The second is valid for $n = 1, 2, \ldots$ and $i = 2, \ldots, n+1$. To establish the desired assumption, it suffices to extend (15) to an inequality

$$\begin{aligned} \left| t + \sum_{i=1}^n x_i \right| &\geq a_n(t) + b_n(t) 1_{[1,\infty)}\left(t^2 + \sum_{i=1}^n x_i^2 \right) \\ &\quad + \sum_{i=1}^n x_i \psi_{n,i}(t; x_1, \ldots, x_{i-1}), \end{aligned} \tag{16}$$

valid for all $t, x_1, \ldots, x_n$, with $a_n(t) \geq 0$ for all t, and $a_n(t) + b_n(t) = \phi_n(t)$. For, (16) then immediately implies

$$E\left| t + \sum_{i=1}^n X_i \right| \geq \phi_n(t) P\left(t^2 + \sum_{i=1}^n X_i^2 \geq 1 \right),$$

whenever $EX_1 = 0$, $E(X_i|X_1,\dots,X_{i-1}) = 0$, $i = 2,\dots,n$. For $t^2 \le \frac{1}{2}$, $a_n(t)$ may be defined by the inductive scheme

$$a_1(t) = t^2(1-t^2)^{-1/2}, \tag{17}$$

$$\begin{aligned} a_{n+1}(t) = \inf\{&(1+2tx)^{1/2}a_n((t+x)(1+2tx)^{-1/2}) \\ &- t(1-t^2)^{-1}\phi_{n+1}(t)x\colon x_- < x < x_+\},\ n = 1,2,\dots. \end{aligned} \tag{18}$$

We must prove that (17) and (18) define a sequence of nonnegative functions.

LEMMA A.1. *For each $n = 1,2,\dots$, the "functional" equation*

$$\begin{aligned} f_n(t) = \inf\{&(1+2tx)^{1/2}f_n((t+x)(1+2tx)^{-1/2}) \\ &- t(1-t^2)^{-1}\phi_n(t)x\colon x_- < x < x_+\} \end{aligned} \tag{19}$$

has a unique differentiable solution defined for $t^2 < \frac{1}{2}$ and satisfying $f_n(0) = 0$.

PROOF. Since $t^2 < \frac{1}{2}$, we have $x_- < 0 < x_+$. Thus, (19) can have a solution only if the infimum occurs at $x = 0$ for all t. This forces

$$\frac{d}{dx}\{(1+2tx)^{1/2}f_n((t+x)(1+2tx)^{-1/2}) - t(1-t^2)^{-1}\phi_n(t)x\}|_{x=0} = 0.$$

Thus, $f_n(t)$ must satisfy the differential equation $f'_n(t) + t(1-t^2)^{-1}f_n(t) = t(1-t^2)^{-2}\phi_n(t)$. There is a unique solution satisfying $f_n(0) = 0$, viz.,

$$f_n(t) = (1-t^2)^{1/2}\int_0^t u(1-u^2)^{-5/2}\phi_n(u)\,du. \tag{20}$$

It remains to prove that this actually satisfies (19). Using the binomial theorem, we have

$$\begin{aligned} \phi_n(t) &= (1-\tfrac{1}{n})^{(n-1)/2}(1-t^2)^{1/2}(1-t^2/(n(1-t^2)))^{-(n-1)/2} \\ &= (1-\tfrac{1}{n})^{(n-1)/2}(1-t^2)^{1/2}\sum_{k=0}^{\infty}\binom{-(n-1)/2}{k}(-t^2/(n(1-t^2)))^k \\ &= \sum_{k=0}^{\infty} c_{n,k}h_k(t) \end{aligned}$$

where $c_{n,k} \ge 0$, for all k, and $h_k(t) = t^{2k}(1-t^2)^{-k+1/2}$, for $k = 0,1,2,\dots$. Integrating in (20) we obtain

$$f_n(t) = \sum_{k=1}^{\infty}(2k)^{-1}c_{n,k-1}h_k(t).$$

Since all the coefficients $c_{n,k-1}$ are nonnegative, we can prove that f_n satisfies (19) by showing

$$\begin{aligned} (2k)^{-1}h_k(t) \le (2k)^{-1}&(1+2tx)^{1/2}h_k((t+x)(1+2tx)^{-1/2}) \\ &- t(1-t^2)^{-1}h_{k-1}(t)x \end{aligned} \tag{21}$$

for $x_- < x < x_+$. Since both sides of (21) and their derivatives are equal for $x = 0$, it suffices to prove that

$$(1+2tx)^{1/2}h_k((t+x)(1+2tx)^{-1/2}) = (t+x)^{2k}(1-t^2-x^2)^{-k+1/2}$$

is convex for $x_- < x < x_+$. This follows by calculating a second derivative.

THEOREM A.1. *The sequence $a_n(t)$ defined by* (17) *and* (18) *is nonnegative.*

PROOF. Since $\phi_n(t) \geq \phi_{n+1}(t) \geq 0$, it follows from (20) that $f_n(t) \geq f_{n+1}(t) \geq 0$. Thus, $a_n(t) \geq f_n(t) \geq 0$ implies, from (19), $a_{n+1}(t) \geq f_{n+1}(t) \geq 0$. However, $a_1(t) = t^2(1 - t^2)^{-1/2} \geq t^2(1 - t^2)^{-1/2}/2 = f_1(t)$. Thus, the theorem follows by induction.

REFERENCES

1. B. Bollobas, *Martingale inequalities*, Math. Proc. Cambridge Philos. Soc. **85** (1980), 377–382.
2. D. L. Burkholder, *Martingale transforms*, Ann. Math. Statist. **37** (1966), 1494–1504.
3. ———, *Distribution function inequalities for martingales*, Ann. Probab. **1** (1973), 19–42.
4. ———, *A sharp inequality for martingale transforms*, Ann. Probab. **7** (1979), 858–863.
5. J. H. B. Kemperman, *The general moment problem, a geometric approach*, Ann. Math. Statist. **39** (1968), 93–122.

COLUMBUS LABORATORIES, BATTELLE MEMORIAL INSTITUTE, COLUMBUS, OHIO 43201