The Annals of Probability
1995, Vol. 23, No. 1, 223–235

SHARP INEQUALITIES FOR THE DISTRIBUTION OF A STOCHASTIC INTEGRAL IN WHICH THE INTEGRATOR IS A BOUNDED SUBMARTINGALE

BY WILLIAM HAMMACK

University of Illinois

We obtain a sharp probability bound on the maximal function of a strong subordinate of a bounded submartingale. An analogous inequality also holds for stochastic integrals in which the integrator is a bounded submartingale and the integrand is a bounded predictable process.

1. Introduction. Let $(\Omega, \mathscr{F}, P)$ be a complete probability space with a right-continuous filtration $(\mathscr{F}_t)_{t \geq 0}$ where $\mathscr{F}_0$ contains all P-null sets. Suppose X is an adapted right-continuous real-valued submartingale with left limits and H is a predictable process with values in the closed unit ball of $\mathbb{R}^\nu$, where ν is a positive integer. We may then define an adapted right-continuous process Y with left limits by

$$Y_t = H_0 X_0 + \int_{(0,t]} H_s \, dX_s.$$

Suppose that $\|X\|_\infty \leq 1$, where $\|X\|_\infty = \sup_{t \geq 0} \|X_t\|_\infty$. What can be said about the size of Y? In particular, can we say anything about the distribution of its maximal function $Y^* = \sup_{t \geq 0} |Y_t|$? Here, for $y, k \in \mathbb{R}^\nu$, we will denote the Euclidean norm of y by $|y|$ and the inner product of y and k by $y \cdot k$.

It is clear that where X and H start will play a significant role in the distribution of Y^*: for example, on any probability space $(\Omega, \mathscr{F}, P)$, for all $\omega \in \Omega$ let

$$X_t(\omega) = \begin{cases} -1, & 0 \leq t < 1, \\ 1, & 1 \leq t, \end{cases} \qquad H_t(\omega) = \begin{cases} -1, & t = 0, \\ 1, & t > 0, \end{cases}$$

thus

$$Y_t(\omega) = \begin{cases} 1, & 0 \leq t < 1, \\ 3, & 1 \leq t, \end{cases}$$

and $P(Y^* \geq 3) = 1$. However, as we shall see, if $X_0 \equiv 0$, then, for any H satisfying the initial hypotheses, $P(Y^* \geq 3) \leq 5/9$. Thus for each $\lambda > 0$, we would like to find a function $U_\lambda : [-1, 1] \times \mathbb{R}^\nu \to [0, 1]$ such that for any H as above,

$$P(Y^* \geq \lambda) \leq \mathbf{E} U_\lambda(X_0, Y_0) \tag{1.1}$$

Received July 1993; revised February 1994.
AMS 1991 *subject classifications*. Primary 60G42, 60H05; secondary 60E15.
Key words and phrases. Martingale, submartingale, maximal inequality, differential subordination, strong subordination, stochastic integral.

and the U_λ is a sharp estimate in the sense that given any $\lambda > 0$, $x \in [-1,1]$ and $y \in \mathbb{R}^\nu$ with $|y| \le |x|$, then for all $\beta < U_\lambda(x,y)$ there exist H and X as above with $X_0 \equiv x$, $X_0 H_0 = y$ and

$$P(Y^* \ge \lambda) > \beta. \tag{1.2}$$

As a specific result, we shall show that for any X and Y as above, if $\lambda > 4$, then

$$P(Y^* \ge \lambda) \le \gamma e^{-\gamma/4}, \tag{1.3}$$

where $\gamma = (8 + \sqrt{2})e/12$, and this inequality is sharp. In contrast, if in addition we require X to be a martingale, it is known [Theorem 8.1 of Burkholder (1991)] that for $\lambda > 2$, $P(Y^* \ge \lambda) \le \alpha e^{-\lambda}$, where $\alpha = e^2/4$.

We shall make heavy use of the techniques developed by Burkholder (1991), who used them for the martingale case. The first step, after describing the U_λ, is to establish an inequality similar to (1.1) but more general for discrete-time submartingales. We will then give its implications for stochastic integrals.

2. The majorants U_λ. For $\lambda > 4$, define the following subsets of $[-1,1] \times \mathbb{R}^\nu$:

$$\begin{aligned}
A_\lambda &= \{(x,y): |y| \ge \lambda - 1 + x\},\\
B_\lambda &= \{(x,y): \lambda - 3 - x \le |y| < \lambda - 1 + x\},\\
C_\lambda &= \{(x,y): 1 - x \le |y| < \lambda - 3 - x\},\\
D_\lambda &= \{(x,y): 0 \le |y| < 1 - x\},
\end{aligned}$$

and let $U_\lambda: [-1,1] \times \mathbb{R}^\nu \to \mathbb{R}$ be

$$U_\lambda(x,y) = \begin{cases} 1, & \text{if } (x,y) \in A_\lambda,\\ \dfrac{2-2x}{1+\lambda-x-|y|}, & \text{if } (x,y) \in B_\lambda,\\ \left(\dfrac{1-x}{2}\right)\exp\left(\dfrac{3+x+|y|-\lambda}{4}\right), & \text{if } (x,y) \in C_\lambda,\\ \left(\dfrac{2}{3} - \dfrac{2+2x-|y|}{6\sqrt{2}}\sqrt{1+x+|y|}\right)\exp\left(1+\dfrac{\lambda}{4}\right), & \text{if } (x,y) \in D_\lambda. \end{cases}$$

While for the case $0 < \lambda \le 4$, define

$$\begin{aligned}
A_\lambda &= \{(x,y): |y| \ge \lambda - 1 + z\},\\
B_\lambda &= \{(x,y): 1 - x \le |y| < \lambda - 1 + x\},\\
C_\lambda &= \{(x,y): \lambda - 3 - x \le |y| < 1 - x \text{ and } |y| < \lambda - 1 + x\},\\
D_\lambda &= \{(x,y): 0 \le |y| < \lambda - 3 - x\},
\end{aligned}$$

and define U_λ by

$$U_\lambda(x,y) = \begin{cases} 1 & \text{if } (x,y) \in A_\lambda, \\ \dfrac{2-2x}{1+\lambda-x-|y|}, & \text{if } (x,y) \in B_\lambda, \\ 1 - \dfrac{(\lambda-1+x-|y|)(\lambda-1+x+|y|)}{\lambda^2}, & \text{if } (x,y) \in C_\lambda, \\ 1 - \left(\dfrac{2\lambda-4}{\lambda}\right)^2\left(\dfrac{1}{3} + \dfrac{2+2x-|y|}{3(\lambda-2)^{3/2}}\sqrt{1+x+|y|}\right), & \\ & \text{if } (x,y) \in D_\lambda. \end{cases}$$

Note in the case $\lambda \le 2$, $D_\lambda = \varnothing$. Also for each $\lambda > 0$, U_λ maps into $[0,1]$ and is continuous on $[-1,1] \times \mathbb{R}^\nu$ except at those points where $x = 1$ and $|y| = \lambda$. Further for $|y| \ge \lambda$, $U_\lambda(x,y) = 1$.

Note also that for fixed $(x,y) \in [-1,1] \times \mathbb{R}^\nu$, the map $\lambda \mapsto U_\lambda(x,y)$ is left-continuous on $(0,\infty)$ [this can be seen by considering first the case $(1-x) > |y|$ and then considering the case $(1-x) \le |y|$ first for $\lambda > 4$ and then for $\lambda \le 4$].

LEMMA 2.1. *For $\lambda > 0$, let $\varphi_\lambda: [-1,1] \times \mathbb{R}^\nu \to \mathbb{R}$ and $\psi_\lambda: [-1,1] \times \mathbb{R}^\nu \to \mathbb{R}^\nu$ be defined to be the derivatives of U_λ with respect to x and y, respectively, on the interiors of A_λ, B_λ, C_λ and D_λ extended continuously to the whole of these sets. Then whenever x and $x+h$ are in $[-1,1]$, y and k are in $\mathbb{R}^\nu$ and $|k| \le |h|$, we have that*

$$|\psi_\lambda(x,y)| \le -\varphi_\lambda(x,y), \tag{2.1}$$

$$U_\lambda(x+h, y+k) \le U_\lambda(x,y) + \varphi_\lambda(x,y)h + \psi_\lambda(x,y)\cdot k, \tag{2.2}$$

and further, with $S_\lambda = \{(x,y) \in [-1,1] \times R^\nu: |y| \ne \lambda - 1 + x\}$,

$$\varphi_\lambda \text{ and } \psi_\lambda \text{ are continuous on } S_\lambda. \tag{2.3}$$

PROOF. For $\lambda > 4$, we have that

$$\varphi_\lambda(x,y) = \begin{cases} 0, & \text{if } (x,y) \in A_\lambda \\ & \quad \text{or } (x,y) = (-1,0), \\ -\dfrac{2\lambda - 2|y|}{(1+\lambda-x-|y|)^2}, & \text{if } (x,y) \in B_\lambda, \\ -\left(\dfrac{3+x}{8}\right)\exp\left(\dfrac{3+x+|y|-\lambda}{4}\right), & \text{if } (x,y) \in C_\lambda, \\ -\left(\dfrac{2+2x+|y|}{4\sqrt{2}\sqrt{1+x+|y|}}\right)\exp\left(1-\dfrac{\lambda}{4}\right), & \text{if } (x,y) \in D_\lambda \setminus (-1,0), \end{cases}$$

$$\psi_\lambda(x,y)=\begin{cases}0, & \text{if } (x,y)\in A_\lambda \\ & \quad\text{or } (|x|,y)=(1,0),\\ \dfrac{2-2x}{(1+\lambda-x-|y|)^2}y', & \text{if } (x,y)\in B_\lambda,\\ \left(\dfrac{1-x}{8}\right)\exp\left(\dfrac{3+x+|y|-\lambda}{4}\right)y', & \text{if } (x,y)\in C_\lambda\setminus(1,0)\\ \left(\dfrac{1}{4\sqrt{2}\sqrt{1+x+|y|}}\right)\exp\left(1-\dfrac{\lambda}{4}\right)y, & \text{if } (x,y)\in D_\lambda\setminus(-1,0)\end{cases}$$

where $y'=y/|y|$.

While for $0<\lambda\le 4$, we have that

$$\varphi_\lambda(x,y)=\begin{cases}0, & \text{if } (x,y)\in A_\lambda \text{ or } (x,y)=(-1,0),\\ -\dfrac{2\lambda-2|y|}{(1+\lambda-x-|y|)^2}, & \text{if } (x,y)\in B_\lambda,\\ -\dfrac{2\lambda+2x-2}{\lambda^2}, & \text{if } (x,y)\in C_\lambda,\\ -\dfrac{(4+4x+2|y|)\sqrt{\lambda-2}}{\lambda^2\sqrt{1+x+|y|}}, & \text{if } (x,y)\in D_\lambda\setminus(-1,0)\end{cases}$$

$$\psi_\lambda(x,y)=\begin{cases}0, & \text{if } (x,y)\in A_\lambda \text{ or } (|x|,y)=(1,0),\\ \dfrac{2-2x}{(1+\lambda-x-|y|)^2}y', & \text{if } (x,y)\in B_\lambda \text{ and } (x,y)\ne(1,0),\\ \dfrac{2y}{\lambda^2}, & \text{if } (x,y)\in C_\lambda,\\ \dfrac{2y\sqrt{\lambda-2}}{\lambda^2\sqrt{1+x+|y|}}, & \text{if } (x,y)\in D_\lambda \text{ and } (x,y)\ne(-1,0).\end{cases}$$

In all cases, (2.1) and (2.3) are satisfied. To show (2.2), let x, y, h and k satisfy the assumptions in the lemma and fix $\lambda>0$. Since $|k|\le|h|$, we may assume $h\ne 0$. Suppose for some $t\in\mathbb{R}$, $x+th=1$ and $|y+tk|=\lambda$. Then, if $\delta\in\mathbb{R}$ and $|x+(t+\delta)h|\le 1$, it follows that $\delta h\le 0$ and

$$\begin{aligned}|y+(t+\delta)k| &\ge |y+tk|-|\delta k|\\ &\ge |y+tk|+\delta h=\lambda-1+x+(t+\delta)h;\end{aligned}$$

hence $(x+(t+\delta)h, y+(t+\delta)k)\in A_\lambda$. In particular, both (x,y) and $(x+h, y+k)\in A_\lambda$ and (2.2) is satisfied. Thus we will assume in the following that if $x+th=1$ for some t, then $|y+tk|\ne\lambda$.

Now suppose $(x, y) \in S_\lambda$. Define G on $\{t \in \mathbb{R}: |x + th| \le 1\}$ by

$$G(t) = U_\lambda(x + th, y + tk).$$

Since we are assuming $(x + th, |y + tk|) \neq (1, \lambda)$ for all t, we have that G is continuous, $G'(0^+)$ exists and (2.2) is equivalent to $G(1) \le G(0) + G'(0^+)$. Thus it suffices to show that G is concave. Since $U_\lambda(x, y) \le 1$, with equality if and only if $(x, y) \in A_\lambda$, it suffices by the continuity of G to show that G is concave on the set of t such that $|y + tk| < \lambda - 1 + x + th$. Because G' is continuous for such t, it suffices to show that $G''(t_0) \le 0$ for those t_0 having a neighborhood U in $\mathbb{R}$ such that the function $t \mapsto (x + th, y + tk)$ maps U into exactly one of B_λ, C_λ and D_λ. After a translation, we may assume $t_0 = 0$ and it suffices to show $G''(0) \le 0$.

If $(x, y) \in B_\lambda$, then for all $\lambda > 0$,

$$G''(0) = \frac{-2}{(1 + \lambda - x - |y|)^3}(G_1 + G_2),$$

where

$$G_1 = 2h^2\left(1 + y' \cdot \frac{k}{h}\right)\left(\lambda - y' \cdot \frac{k}{h} + xy' \cdot \frac{k}{h} - |y|\right),$$

$$G_2 = (1 - x)(\lambda + 1 - x - |y|)\left(\frac{(y \cdot k)^2 - |y|^2|k|^2}{|y|^3}\right).$$

Thus to show $G''(0) \le 0$, it suffices to show $G_1 + G_2 \ge 0$. Let $\theta = y' \cdot k/h$, so $|\theta| \le 1$. By the definition of B_λ, $G_1 \ge 0$ and $G_2 \le 0$. After we divide through by h^2 and note that $|y| \ge 1 - x$, it suffices for us to show that

$$2(1 + \theta)(\lambda - \theta + x\theta - |y|) - (\lambda + 1 - x - |y|)(1 - \theta^2) \ge 0$$

or, equivalently, that

$$(1 + \theta)^2(\lambda - 1 + x - |y|) \ge 0,$$

an inequality which follows from $|y| < \lambda - 1 + x$ on B_λ.

For the case $\lambda > 4$ and $(x, y) \in C_\lambda$, we have that

$$G''(0) = -\exp\left(\frac{3 + x + |y| - \lambda}{4}\right)\left(\frac{G_1 + G_2}{32}\right),$$

where

$$G_1 = h^2\left(1 + y' \cdot \frac{k}{h}\right)\left(7 + x - y' \cdot \frac{k}{h} + xy' \cdot \frac{k}{h}\right),$$

$$G_2 = 4(1 - x)\left(\frac{(y \cdot k)^2 - |y|^2|k|^2}{|y|^3}\right).$$

As in the previous case, it suffices to show $G_1 + G_2 \ge 0$. By using the same steps as before, we only need to show that

$$(1+\theta)(7-x-\theta+x\theta)-4(1-\theta^2)\geq 0.$$

Since the left side simplifies to $(1+\theta)^2(3+x)$, this is clear.

If $\lambda \leq 4$ and $(x,y)\in C_\lambda$, then

$$G''(0)=\frac{-2h^2}{\lambda^2}\left(1-\frac{|k|^2}{h^2}\right),$$

which is nonpositive from our assumptions on h and k.

Finally for $(x,y)\in D_\lambda$ (so that $\lambda\geq 2$),

$$G''(0)=-\gamma_\lambda\left(\frac{G_1+G_2}{(1+x+|y|)^{3/2}}\right),$$

where

$$G_1=h^2\left(1+y'\cdot\frac{k}{h}\right)\left(2|y|+\left(1-y'\cdot\frac{k}{h}\right)(2+2x+|y|)\right),$$

$$G_2=2|y|(1+x+|y|)\left(\frac{(y\cdot k)^2-|y|^2|k|^2}{|y|^3}\right),$$

$$\gamma_\lambda=\begin{cases}\exp(1-\lambda/4)/(8\sqrt{2}), & \text{for } \lambda>4,\\ \sqrt{\lambda-2}/\lambda^2, & \text{otherwise.}\end{cases}$$

Thus $\gamma_\lambda\geq 0$ and, using the methods of the previous cases, it suffices to show that

$$(1+\theta)(2|y|+(1-\theta)(2+2x+|y|))-2(1+x+|y|)(1-\theta^2)\geq 0.$$

However, note that the left side simplifies to $(1+\theta)^2|y|$.

We have now shown (2.2) holds for $(x,y)\in S_\lambda$. For $(x,y)\notin S_\lambda$, $y\neq 0$, we have that $(x,y)\in A_\lambda$ and for all $\delta>0,(x,(1+\delta)y)\in A_\lambda\cap S_\lambda$; hence,

$$U_\lambda(x+h,(1+\delta)y+k)\leq U_\lambda(x,(1+\delta)y)$$

(recall φ_λ and ψ_λ are both zero on A_λ). By the continuity of U_λ, letting δ to to zero gives (2.2). Finally if $(x,0)\notin S_\lambda$, let y_1 be any nonzero element of $\mathbb{R}^\nu$. Then for all $\delta>0,(x,\delta y_1)\in A_\lambda\cap S_\lambda$; hence $U_\lambda(x+h,\delta y_1+k)\leq U_\lambda(x,\delta y_1)$. Letting δ go to zero gives (2.2) and completes the proof. □

3. Discrete submartingales. Let $f=(f_n)_{n\geq 0}$ be a real-valued submartingale relative to a filtration $(\mathscr{F}_n)_{n\geq 0}$ on a probability space $(\Omega,\mathscr{F},P)$ with difference sequence $(d_n)_{n\geq 0}$ and let $g=(g_n)_{n\geq 0}$ be a $\mathbb{R}^\nu$-valued process adapted to $(\mathscr{F}_n)_{n\geq 0}$ with difference sequence $(e_n)_{n\geq 0}$, where ν is a positive integer. We say that g is conditionally differentially subordinate to f if $|\mathbf{E}(e_n|\mathscr{F}_{n-1})|\leq|\mathbf{E}(d_n|\mathscr{F}_{n-1})|$ for all $n\geq 1$ and g is differentially subordinate to f if $|e_n|\leq|d_n|$. If g is both conditionally differentially subordinate to f and, for $n\geq 0$, differentially subordinate to f, we say g is strongly subordinate to

f. For the next theorem we will require g to be differentially subordinate to f for $n \ge 1$.

THEOREM 3.1. *If f is a submartingale relative to a filtration $(\mathscr{F}_n)_{n\ge 0}$ such that $\|f\|_\infty \le 1$ and g is an adapted process that is conditionally differentially subordinate and, for $n \ge 1$, differentially subordinate to f, then for all $\lambda > 0$,*

$$P(g^* \ge \lambda) \le \mathbf{E}U_\lambda(f_0, g_0), \tag{3.1}$$

where $g^* = \sup_{n\ge 0}|g_n|$ and $\|f\|_\infty = \sup_{n\ge 0}\|f_n\|_\infty$.

PROOF. Fix $\lambda > 0$. It suffices to show that

$$P(|g_n| \ge \lambda) \le \mathbf{E}U_\lambda(f_0, g_0) \tag{3.2}$$

since if (3.2) holds for all $\lambda > 0$, then for all $0 < \varepsilon < \lambda$, with $\tau = \inf\{n \ge 0: |g_n| \ge \lambda - \varepsilon\}$, τ is a stopping time, $f^\tau = (f_{\tau\wedge n})_{n\ge 0}$ is a submartingale, $\|f_\tau\|_\infty \le 1$ and $g^\tau = (g_{\tau\wedge n})_{n\ge 0}$ is conditionally differentially subordinate and, for $n \ge 1$, differentially subordinate to f^τ. We then have

$$P\Big(\sup_{m\le n}|g_m| \ge \lambda - \varepsilon\Big) = P(|g_{\tau\wedge n}| \ge \lambda - \varepsilon) \le \mathbf{E}U_{\lambda-\varepsilon}(f_0, g_0).$$

Letting $n \to \infty$ then implies $P(g^* \ge \lambda) \le \mathbf{E}U_{\lambda-\varepsilon}(f_0, g_0)$. (3.1) now follows by the left-continuity of U_λ as a function of λ.

Since $U_\lambda(x, y) = 1$ for $|y| \ge \lambda$ and $U_\lambda \ge 0$ we have that $P(|g_n| \ge \lambda) \le \mathbf{E}U_\lambda(f_n, g_n)$. Thus to prove (3.2), it suffices to show $\mathbf{E}U_\lambda(f_j, g_j) \le \mathbf{E}U_\lambda(f_{j-1}, g_{j-1})$ holds for $1 \le j \le n$.

Let φ_λ and ψ_λ be as in Lemma 2.1. Since, by assumption, $|e_j| \le |d_j|$, (2.2) implies

$$U_\lambda(f_j, g_j) \le U_\lambda(f_{j-1}, g_{j-1}) + \varphi_\lambda(f_{j-1}, g_{j-1})d_j + \psi_\lambda(f_{j-1}, g_{j-1})\cdot e_j.$$

Assuming for now the integrability of $\varphi_\lambda(f_{j-1}, g_{j-1})d_j$ and $\psi_\lambda(f_{j-1}, g_{j-1})\cdot e_j$, taking the conditional expectations relative to $\mathscr{F}_{j-1}$ of both sides of (3.3) gives

$$\mathbf{E}\big(U_\lambda(f_j, g_j)|\mathscr{F}_{j-1}\big) \le U_\lambda(f_{j-1}, g_{j-1}) + \varphi_\lambda(f_{j-1}, g_{j-1})\mathbf{E}\big(d_j|\mathscr{F}_{j-1}\big) + \psi_\lambda(f_{j-1}, g_{j-1})\cdot\mathbf{E}\big(e_j|\mathscr{F}_{j-1}\big).$$

Since f is a submartingale, $\mathbf{E}(d_j|\mathscr{F}_{j-1}) \ge 0$. It then follows from (2.1) and the assumption that g is conditionally differentially subordinate to f that

$$|\psi_\lambda(f_{j-1}, g_{j-1})\cdot\mathbf{E}\big(e_j|\mathscr{F}_{j-1}\big)| \le -\varphi_\lambda(f_{j-1}, g_{j-1})\mathbf{E}\big(d_j|\mathscr{F}_{j-1}\big).$$

Hence

$$\varphi_\lambda(f_{j-1}, g_{j-1})\mathbf{E}\big(d_j|\mathscr{F}_{j-1}\big) + \psi_\lambda(f_{j-1}, g_{j-1})\cdot\mathbf{E}\big(e_j|\mathscr{F}_{j-1}\big) \le 0$$

and so

$$\mathbf{E}\big(U_\lambda(f_j, g_j)|\mathscr{F}_{j-1}\big) \le U_\lambda(f_{j-1}, g_{j-1}).$$

Taking expectations of both sides then gives $\mathbf{E}U_\lambda(f_j, g_j) \le \mathbf{E}U_\lambda(f_{j-1}, g_{j-1})$.

Thus it remains to show the integrability of $\varphi_\lambda(f_{j-1}, g_{j-1})d_j$ and $\psi_\lambda(f_{j-1}, g_{j-1}) \cdot e_j$. Since $|\psi_\lambda| \le -\varphi_\lambda$ and we are assuming $|e_j| \le |d_j|$, it suffices to show that $\varphi_\lambda(f_{j-1}, g_{j-1})d_j$ is integrable. We will do this for $\lambda > 4$, the case $\lambda \le 4$ being essentially the same.

For $(x, y) \notin B_\lambda$ and for (x, y) satisfying $|y| < \lambda - 1, |\varphi_\lambda|$ is bounded by a constant depending only on λ. Since $\|f\|_\infty \le 1$ implies $|d_j| \le 2$ a.s., it suffices to show $\varphi_\lambda(f_{j-1}, g_{j-1})d_j$ is integrable on the set $\{(f_{j-1}, g_{j-1}) \in B_\lambda, |g_{j-1}| \ge \lambda - 1\}$.

Fix $N \ge 1$ and let $R_N = \{(x, y) \in B_\lambda \colon \lambda - 2^{1-N} \le |y| < \lambda - 2^{-N}\}$. Then for $(x, y) \in R_N$,

$$|\varphi_\lambda(x, y)| \le |\varphi_\lambda(1, y)| = \frac{2}{\lambda - |y|} < 2^{N+1}.$$

Since $(x, y) \in B_\lambda$ and $|y| \ge \lambda - 2^{1-N}$ together imply that $x \ge 1 - 2^{1-N}$, we have that $(f_{j-1}, g_{j-1}) \in R_N$ implies that $d_j \le 2^{1-N}$ a.s. Further, since f is a submartingale,

$$\int \mathbf{1}_{R_N}(f_{j-1}, g_{j-1})d_j \ge 0.$$

Hence

$$\int \mathbf{1}_{R_N}(f_{j-1}, g_{j-1})|d_j| \le 2\int \mathbf{1}_{R_N}(f_{j-1}, g_{j-1})d_j^+.$$

Thus

$$\begin{aligned}\mathbf{E}\big(|\mathbf{1}_{R_N}(f_{j-1}, g_{j-1})\varphi_\lambda(f_{j-1}, g_{j-1})d_j|\big) &\le 2^{N+2}\mathbf{E}\big(\mathbf{1}_{R_N}(f_{j-1}, g_{j-1})d_j^+\big)\\ &\le 8P\big((f_{j-1}, g_{j-1}) \in R_N\big).\end{aligned}$$

Since $\bigcup_{N=1}^\infty R_N = \{(x, y) \in B_\lambda \colon |y| \ge \lambda - 1\}$, it follows that

$$\int |\mathbf{1}_{B_\lambda \cap \{|y| \ge \lambda - 1\}}(f_{j-1}, g_{j-1})\varphi_\lambda(f_{j-1}, g_{j-1})d_j| \le 8,$$

thus finishing the proof. □

The following corollary will be useful in extending Theorem 3.1 to stochastic integrals.

COROLLARY 3.1. *Under the assumption of Theorem* 3.1, *if in addition* $|e_0| \le |d_0|$, *so that g is strongly subordinate to f, then*

$$\sup_{\lambda > 0} \lambda P(g^* \ge \lambda) \le \frac{8 + \sqrt{2}}{3}.$$

PROOF. Since $(8 + \sqrt{2})/3 > 3$, it suffices to consider $\lambda \ge 3$. It follows from the partial derivative of U_λ with respect to x being nonpositive for all λ, that for all y, $U_\lambda(x, y) \le U_\lambda(-1, y)$. By considering $U_\lambda(-1, y)$, we see that for $|y| \le 1, U_\lambda(-1, y)$ is maximized whenever $|y| = 1$. Let $y_0 \in \mathbb{R}^\nu$ satisfy $|y_0| = 1$.

Then

$$U_\lambda(-1, y_0) = \begin{cases} \left(\dfrac{2}{3} + \dfrac{1}{6\sqrt{2}}\right)\exp\left(1 - \dfrac{\lambda}{4}\right), & \text{if } \lambda > 4, \\ 1 - \left(\dfrac{2\lambda - 4}{\lambda}\right)^2 \left(\dfrac{1}{3} - \dfrac{1}{3(\lambda - 2)^{3/2}}\right), & \text{if } 3 < \lambda \le 4. \end{cases}$$

As a function of λ, $\lambda U_\lambda(-1, y_0)$ is increasing on the interval $[3, 4)$, decreasing on the interval $[4, \infty)$ and continuous on $[3, \infty)$. Hence

$$\lambda P(g^* \ge \lambda) \le \lambda \mathbf{E} U_\lambda(f_0, g_0) \le \lambda U_\lambda(-1, y_0) \le 4U_4(-1, y_0) = \frac{8 + \sqrt{2}}{3}. \quad \square$$

4. Inequalities for stochastic integrals.

THEOREM 4.1. *Let $(\Omega, \mathscr{F}, P)$ be a complete probability space with a right-continuous filtration $(\mathscr{F}_t)_{t \ge 0}$, where $\mathscr{F}_0$ contains all P-null sets. Suppose X is an adapted right-continuous submartingale with left limits such that $\|X\|_\infty \le 1$ and H is a predictable process with values in the closed unit ball of $\mathbb{R}^\nu$. Then with*

$$Y_t = H_0 X_0 + \int_{(0,t]} H_s \, dX_s,$$

we have that, for $\lambda > 0$,

$$P(Y^* \ge \lambda) \le \mathbf{E} U_\lambda(X_0, Y_0). \tag{4.1}$$

PROOF. First consider $\tilde{Y}$ of the form

$$\tilde{Y}_t = H_0 X_0 + \sum_{j=1}^{n} a_k \left[X_{\tau_j \wedge t} - X_{\tau_{j-1} \wedge t} \right], \tag{4.2}$$

where $a_1, \dots, a_n$ are in the closed unit ball of $\mathbb{R}^\nu$ and $0 = \tau_0 \le \tau_1 \le \cdots \le \tau_n$ are stopping times taking only finitely many values, all of them finite. Let t be an upper bound for τ_n. Let $\tau = \inf\{s \in [0, t]: |\tilde{Y}_s| \ge \lambda\}$. By the right-continuity of X, on $\{\tau < \infty\}$, $\tilde{Y}_\tau \ge \lambda$. For $j = 0, 1, \cdots, n$, let $f_j = X_{\tau_j \wedge \tau}$ and $g_j = \tilde{Y}_{\tau_j \wedge \tau}$. Since X is bounded we can apply Doob's optional sampling theorem to get that f is a submartingale. Since $\|f\|_\infty \le 1$ and g is strongly subordinate to f we can apply Theorem 3.1 and get

$$P(\tilde{Y}^* \ge \lambda) = P(g^* \ge \lambda) \le \mathbf{E} U_\lambda(f_0, g_0) = \mathbf{E} U_\lambda(X_0, Y_0).$$

Thus any $\tilde{Y}$ of the form in (4.2) satisfies (4.1).

In particular with $\nu = 1$, this and Corollary 3.1 show that X is an $L^{1,\infty}$-integrator in the sense of Bichteler (1981).

By the additivity of the integral and Theorem 4.1 of Bichteler (1981), there exist Y^n of the form in (4.2) such that $P(\limsup (Y^n - Y)^* > 0) = 0$. It follows that for all $0 < \varepsilon < \lambda$, there exists an n such that $P((Y^n - Y)^* > \varepsilon) < \varepsilon$ and since the Y^n satisfy (4.1), we have

$$P(Y^* \geq \lambda) \leq P((Y^n)^* \geq \lambda - \varepsilon) + \varepsilon$$
$$\leq \mathbf{E}U_{\lambda-\varepsilon}(X_0, Y_0) + \varepsilon.$$

Letting ε go to zero now gives (4.1) by the left-continuity of U_λ as a function of λ. □

REMARKS. As in the proof of Corollary 3.1, if $y \in \mathbb{R}^\nu$ satisfies $|y| = 1$, then by Theorem 4.1, for $\lambda > 4$ we have that

$$P(Y^* \geq \lambda) \leq \mathbf{E}U_\lambda(X_0, Y_0) \leq U_\lambda(-1, y) = \gamma e^{-\lambda/4},$$

where $\gamma = (8 + \sqrt{2})e/12$. This gives inequality (1.3). The sharpness will follow from Theorem 5.1 with $x_0 = -1$ and $y_0 = 1$.

5. Sharpness of the U_λ. We will now show that the inequality in Theorem 3.1 is sharp even for the case in which g is assumed to be a ± 1-transform of the submartingale f. This will yield the sharpness of the inequality in Theorem 5.1. Since U_λ as a function of y depends only on $|y|$, we may assume $\nu = 1$, that is, we will consider U_λ as a function from $[-1, 1] \times \mathbb{R}$ into $[0, 1]$.

THEOREM 5.1. *Let $\lambda > 0$ and $(x_0, y_0) \in [-1, 1] \times [0, \infty)$. Then for all $\beta < U_\lambda(x_0, y_0)$, there exist constants $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ with $\varepsilon_j = \pm 1$ for all j and a probability space $(\Omega, \mathscr{F}, P)$ having a family of random variables $d_1, d_2, \dots, d_n$ such that with $f_0 \equiv x_0, g_0 \equiv y_0$ and, for $1 \leq m \leq n, f_m = f_{m-1} + d_m$ and $g_m = g_{m-1} + \varepsilon_m d_m$, we have the following*:

$$(5.1)\quad \begin{aligned} &(f_m)_{0 \leq m \leq n} \textit{ is a submartingale relative to the filtration } (\mathscr{F}_j)_{1 \leq j \leq n},\\ &\textit{where } \mathscr{F}_j \textit{ is the } \sigma\textit{-algebra generated by } (f_0, f_1, \dots, f_j)\\ &\textit{and for } 0 \leq m \leq n, \|f_m\|_\infty \leq 1;\end{aligned}$$

$$(5.2)\qquad P(g^* \geq \lambda) > \beta.$$

PROOF. Fix $\lambda > 0$. If $x_0 = 1$, let $(\Omega, \mathscr{F}, P)$ be any probability space and let $f_0 \equiv 1$, $g_0 \equiv y_0$. Then $P(g_0 \geq \lambda) = U_\lambda(x_0, y_0)$. Thus we may assume $x_0 < 1$.

For $(x_0, y_0) \in A_\lambda \cup B_\lambda$, we can actually achieve $U_\lambda(x_0, y_0)$, for example, let $(\Omega, \mathscr{F}, P)$ be the unit interval $[0, 1]$ with Lebesgue measure. For $(x_0, y_0) \in A_\lambda$, let $d_1 \equiv 1 - x_0$ and $\varepsilon_1 = 1$. Then (5.1) is satisfied and $g_1 \equiv y_0 + 1 - x_0 \geq \lambda$; hence $P(g^* \geq \lambda) = 1 = U_\lambda(x_0, y_0)$.

For $(x_0, y_0) \in B_\lambda$, we first go to the line $y = \lambda - 1 + x$, and then use the method used above for A_λ, that is, let $d_1 = (1 - x_0)\mathbf{1}_{[0,\gamma)} - \alpha\mathbf{1}_{[\gamma,1]}$, where α satisfies $y_0 + \alpha = \lambda - 1 + x_0 - \alpha$ and $\gamma = \alpha/(1 - x_0 + \alpha)$. Note that $\mathbf{E}d_1 = 0$. Since $(x_0, y_0) \in B_\lambda$, we have that $\alpha > 0$ and $x_0 - \alpha \geq -1$. Now let $\varepsilon_1 =$

-1, $d_2 = (1 - x_0 + \alpha)\mathbf{1}_{[\gamma,1]}$ and $\varepsilon_2 = 1$. Then $f_2 \equiv 1$ and $g_2 = (y_0 + x_0 - 1)\mathbf{1}_{[0,\gamma)} + (y_0 + 2\alpha + 1 - x_0)\mathbf{1}_{[\gamma,1]}$. By the definition of α, $y_0 + 2\alpha = \lambda - 1 + x_0$; hence, $P(g_2 = \lambda) = 1 - \gamma = U_\lambda(x_0, y_0)$ and (5.1) and (5.2) are satisfied.

Turning to the case $(x_0, y_0) \in C_\lambda$, we first consider the case $\lambda > 4$. Since as N goes to ∞, $(1 - (\lambda - 3 - x_0 - y_0)(4N)^{-1})^N$ converges to $\exp((3 + x_0 + y_0 - \lambda)/4)$, we can find an $N > \lambda$ such that

$$\left(\frac{1 - x_0}{2}\right)\left(1 - \frac{\lambda - 3 - x_0 - y_0}{4N}\right)^N > \beta.$$

Let $(\Omega, \mathscr{F}, P)$ be a probability space on which are defined independent identically distributed random variables $s_1, s_2, \dots, s_N$ and a random variable q, independent of the (s_j), such that

$$P(q = 1 - x_0) = \frac{1 + x_0}{2}, \qquad P(q = -x_0 - 1) = \frac{1 - x_0}{2}$$

and, with $\alpha = (\lambda - 3 - x_0 - y_0)/(2N)$ and $\gamma = 2 - \alpha$,

$$P(s_1 = -\alpha) = \frac{\gamma}{2}, \qquad P(s_1 = \gamma) = \frac{\alpha}{2}.$$

(Note that the inequality $N > \lambda$ implies $\gamma > 0$.) Let $d_1 = q$, $\varepsilon_1 = -1$. Then

$$P(f_1 = -1, g_1 = 1 + x_0 + y_0) + P(f_1 = 1, g_1 = y_0 + x_0 - 1) = 1.$$

For $j \geq 1$, let $d_{2j} = \alpha\mathbf{1}_{\{f_{2j-1} \neq 1\}}$, $d_{2j+1} = s_j\mathbf{1}_{\{f_{2j} \neq 1\}}$, $\varepsilon_{2j} = 1$ and $\varepsilon_{2j+1} = -1$. Finally let $d_{2N+2} = 2\mathbf{1}_{\{f_{2N+1} \neq 1\}}$ and $\varepsilon_{2N+2} = 1$. Since q and the (s_j) are all independent and have expectation 0, it follows that the f_j form a submartingale. For all $j \geq 1$, $\omega \in \Omega$, $f_j(\omega) \in \{-1, -1 + \alpha, 1\}$; hence (5.1) is satisfied. For (5.2) note that by the independence of the random variables we have

$$\begin{aligned} P(g_{2N+2} = \lambda) &\geq P(f_{2N+1} = -1, g_{2N+1} = \lambda - 2) \\ &\geq P\big(q = -x_0 - 1 \text{ and for } 1 \leq j \leq N,\ s_j = -\alpha\big) \\ &= \left(\frac{1 - x_0}{2}\right)\left(\frac{\gamma}{2}\right)^N \\ &= \left(\frac{1 - x_0}{2}\right)\left(1 - \frac{\lambda - 3 - x_0 - y_0}{4N}\right)^N \\ &> \beta. \end{aligned}$$

For $(x_0, y_0) \in C_\lambda$, $\lambda \leq 4$, let $(\Omega, \mathscr{F}, P)$ be the unit interval $[0,1]$ with Lebesgue measure and define γ to be $(\lambda - 1 + x_0 - y_0)/\lambda$. Let $d_1 = \frac{1}{2}(1 - x_0 + y_0)\mathbf{1}_{[0,\gamma)} - \frac{1}{2}(\lambda - 1 + x_0 - y_0)\mathbf{1}_{[\gamma,1]}$ and $\varepsilon_1 = -1$. Then $\mathbf{E}d_1 = 0$ and on $[0,\gamma)$, (f_1, g_1) is on the line $y = x - 1$, while on $[\gamma, 1]$, it is on the line $y = \lambda - 1 + x$.

Now let $d_2 = \frac{1}{2}(1 - x_0 - y_0)\mathbf{1}_{[0,\delta\gamma)} - \frac{1}{2}(\lambda - 1 + x_0 + y_0)\mathbf{1}_{[\delta\gamma,\gamma)} + \frac{1}{2}(\lambda + 1 - x_0 - y_0)\mathbf{1}_{[\gamma,1]}$ and $\varepsilon_2 = 1$, where $\delta = (\lambda - 1 + x_0 + y_0)/\lambda$. Hence $\mathbf{E}(d_2|d_1) \geq 0$. On $[0, \delta\gamma)$, the value of (f_2, g_2) is $(1, 0)$, while on $[\delta\gamma, \gamma)$ it is on the line $-y = \lambda - 1 + x$, and on $[\gamma, 1]$, it is $(1, \lambda)$.

Let $d_3 = \frac{1}{2}\lambda \mathbf{1}_{[\delta\gamma,\gamma)}$ and $\varepsilon_3 = -1$. We then have that on the set $[\delta\gamma, 1]$, $f_3 = 1$ and $|g_3| = \lambda$. Since $1 - \delta\gamma = U_\lambda(x_0, y_0)$, it follows that (5.1) and (5.2) are satisfied.

Now let $(x_0, y_0) \in D_\lambda$. Then $\lambda > 2$ since $\lambda - 3 > x_0 + y_0 \geq -1$. Note that D_λ forms a triangle with vertices at $(-1, w+1)$, $(w, 0)$ and $(-1, -(w+1))$, where $w = (\lambda - 3) \wedge 1$ and that $x_0 + y_0 < w$. Define $h: [x_0 + y_0, w] \to \mathbb{R}$ by

$$h(x) = \frac{1}{3} + \frac{2 + 2x_0 - y_0}{3(1+x)^{3/2}}\sqrt{1 + x_0 + y_0}$$

(for $x_0 + y_0 = -1$, consider h to be defined on $(x_0 + y_0, w]$). Then

$$U_\lambda(x_0, y_0) = (1 - h(w))U_\lambda(-1, w+1) + h(w)U_\lambda(w, 0).$$

Since $U_\lambda(-1, w+1) = U_\lambda(-1, -(w+1))$ and each of $(-1, w+1)$, $(w, 0)$ and $(-1, -(w+1))$ are in either A_λ or C_λ, it suffices to construct a finite length submartingale f and a ± 1-transform g, such that the pair (f, g), starts at (x_0, y_0) and ends at the vertices of D_λ with the probability close to $h(w)$ of ending at $(w, 0)$.

For the case $x_0 = -1$ and $y_0 = 0$, this is equivalent to constructing a finite length nonnegative submartingale f and a ± 1-transform g, starting at $(0,0)$ and ending at the points $(0, -1)$, $(0, 1)$ and $(1, 0)$ such that the probability of ending at $(1, 0)$ is close to $1/3$. This was done in the proof of Theorem 4.1 of Burkholder (1994) and Example 2 of Burkholder (1993), where it was shown that for all $N > 1$, on could construct such a pair (f, y) such that the probability of ending at $(1, 0)$ was $1/3 + 1/6N$.

Now suppose $(x_0, y_0) \neq (-1, 0)$. Fix $N > 0$ and let $\delta_N = (w - x_0 - y_0)/(2N)$. For $1 \leq j \leq N+1$, let

$$x_j^N = x_0 + y_0 + (2j - 2)\delta_N$$

so that $x_{N+1}^N = w$. By using the ideas of Example 2 of Burkholder (1993), we can construct a finite length submartingale f^N with $\|f^N\|_\infty \leq 1$ and a ± 1-transform of f^N, g^N such that, with $p_j^N = P(f_{3j}^N = x_j^N, g_{3j}^N = 0)$, we have for $j \geq 1$,

$$p_j^N + P\left(f_{3j}^N = 0, |g_{3j}^N| = 1 + x_j^N\right) = 1, \tag{5.3}$$

$$p_1^N = \frac{1 + x_0}{1 + x_0 + y_0}, \tag{5.4}$$

$$p_{j+1}^N = \frac{\left(1 + x_j^N\right)^2}{\left(1 + x_j^N + \delta_N\right)\left(1 + x_j^N + 2\delta_N\right)} p_j^N + \frac{\delta_N}{1 + x_j^N + 2\delta_N}. \tag{5.5}$$

Note that h satisfies the differential equation

$$h'(x) + \frac{3}{2 + 2x}h(x) = \frac{1}{2 + 2x}$$

and, since we are assuming $x_0 + y_0 > -1$, $(h(x + \delta) - h(x))/\delta$ converges uniformly to h'(x) on $[x_0 + y_0, w]$. It then follows that there exist constants

ε_N depending only on N and $x_0 + y_0$ such that $\lim_{N\to\infty} \varepsilon_N = 0$ and, for $1 \le j \le N$,

$$\left| \frac{h(x_{j+1}^N) - h(x_j^N)}{2\delta_N} + \frac{3 + 3x_j^N + 2\delta_N}{(2 + 2x_j^N + 2\delta_N)(1 + x_j^N + 2\delta_N)} h(x_j^N) - \frac{1}{2 + 2x_j^N + 4\delta_N} \right| \le \varepsilon_N$$

or, equivalently,

$$\left| h(x_{j+1}^N) - \frac{(1 - x_j^N)^2}{(1 + x_j^N + \delta_N)(1 + x_j^N + 2\delta_N)} h(x_j^N) - \frac{\delta_N}{1 + x_j^N - 2\delta_N} \right| \le 2\delta_N \varepsilon_N \le 2\delta_N \varepsilon_N. \tag{5.6}$$

Combining (5.5) and (5.6) then gives

$$|h(x_{j+1}^N) - p_{j+1}^N| \le \frac{(1 + x_j^N)^2}{(1 + x_j^N + \delta_N)(1 + x_j^N + 2\delta_N)} |h(x_j^N) - p_j^N| + 2\delta_N \varepsilon_N.$$

Since $p_1^N = h(x_1^N)$, it follows that

$$|h(w) - p_{N+1}^N| \le 2N\delta_N \varepsilon_N = (w - x_0 - y_0)\varepsilon_N.$$

Thus $\lim_{N\to\infty} p_{N+1}^N = h(w)$, which completes the proof. □

Acknowledgment. The author is very grateful to Professor D. L. Burkholder for suggesting this problem.

REFERENCES

BICHTELER, K. (1981). Stochastic integration and L^p-theory of semimartingales. *Ann. Probab.* **9** 49–89.

BURKHOLDER, D. L. (1991). Explorations in martingale theory and its applications. *Lecture Notes in Math.* **1464** 1–66. Springer, Berlin.

BURKHOLDER, D. L. (1993). Sharp probability bounds for Itô processes. In *Statistics and Probability: A Raghu Raj Bahadur Festschrift* (J. K. Ghosh, S. K. Mitra, K. R. Parthasarathy and B. L. S. Prakasa Rao, eds.) 135–145. Wiley Eastern, Singapore.

BURKHOLDER, D. L. (1994). Strong differential subordination and stochastic integration. *Ann. Probab.* **22** 995–1025.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS 61801