BULLETIN OF THE POLISH
ACADEMY OF SCIENCES
MATHEMATICS
Vol. 58, No. 3, 2010

PROBABILITY THEORY AND STOCHASTIC PROCESSES

Maximal Inequalities for Stochastic Integrals

by

Adam OSĘKOWSKI

Presented by Stanisław KWAPIEŃ

Summary. We find the optimal universal constant C_p $(1 < p \leq \infty)$ in the following inequality. If $X = (X_t)_{t \geq 0}$ is a martingale and $Y = (\int_0^t H_s\, dX_s)_{t\geq 0}$ for some predictable process H taking values in $[-1,1]$, then

$$\mathbb{E}|\sup_{t\geq 0} Y_t| \leq C_p \|X\|_p.$$

1. Introduction. Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, equipped with a nondecreasing right-continuous family $(\mathcal{F}_t)_{t\geq 0}$ of sub-σ-fields of $\mathcal{F}$. In addition, assume that $\mathcal{F}_0$ contains all the sets of probability 0. Let $X = (X_t)_{t\geq 0}$ be an adapted real-valued right-continuous martingale with left limits. Let Y be the Itô integral of H with respect to X, that is,

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

Here H is a predictable process with values in $[-1, 1]$. For $p \in [1,\infty]$, let $\|X\|_p = \sup_{t\geq 0} \|X_t\|_p$. Furthermore, let $X^* = \sup_{t\geq 0} X_t$ and $|X|^* = \sup_{t\geq 0} |X_t|$.

The purpose of this paper is to compare the moments of X and Y^*. In [B2], Burkholder developed a method to obtain the following sharp estimate.

THEOREM 1.1. *If X is a martingale and Y is as above, then*

$$\|Y\|_1 \leq \gamma \big\| |X|^* \big\|_1, \tag{1.1}$$

2010 *Mathematics Subject Classification*: Primary 60G42; Secondary 60G44.
Key words and phrases: martingale, maximal function, stochastic integral, martingale transform, norm inequality.

DOI: 10.4064/ba58-3-10

where $\gamma = 2.536\ldots$ *is the unique solution of the equation*

$$\gamma - 3 = -\exp\left(\frac{1-\gamma}{2}\right).$$

The constant is the best possible.

It was shown in [O1] that if X is assumed to be a nonnegative supermartingale, then the optimal constant in (1.1) decreases to $2 + (3e)^{-1} = 2.1226\ldots$. The paper [O2] contains the following fact.

THEOREM 1.2. *If* X *and* Y *are as above, then*

$$\|Y^*\|_1 \le \beta \big\| |X|^* \big\|_1, \tag{1.2}$$

where $\beta = 2.0856\ldots$ *is the positive solution to the equation*

$$2\log\left(\frac{8}{3} - \beta_0\right) = 1 - \beta_0.$$

Furthermore, if X *is assumed to be nonnegative, then the optimal constant in* (1.2) *decreases to* $14/9 = 1.5555\ldots$.

In the present paper we continue this line of research and provide new sharp bounds for the first moment of Y^* by $\|X\|_p$ for $p > 1$. If $p = 1$, then there is no finite constant C_1 such that $\|Y^*\|_1 \le C_1\|X\|_1$, even when $Y = X$. For example, take $X_t = e^{\alpha W_t - \alpha^2 t/2}$, where W is the Wiener process; then $\mathbb{E}X^* = \infty$ and $\mathbb{E}|X_t| = \mathbb{E}X_t = 1$ for all t. Let

$$C_p = \begin{cases} \Gamma\left(\dfrac{2p-1}{p-1}\right)^{1-1/p} & \text{if } 1 < p \le 2, \\ \left(2^{p/(p-1)} - \dfrac{p}{p-1}\displaystyle\int_1^2 s^{1/(p-1)} e^{s-2}\, ds\right)^{1-1/p} & \text{if } 2 < p < \infty, \\ 1 + e^{-1} & \text{if } p = \infty. \end{cases}$$

Here is our main result.

THEOREM 1.3. *Suppose* X *is a martingale and* Y *is as above. If* $1 < p \le \infty$, *then*

$$\|Y^*\|_1 \le C_p \|X\|_p. \tag{1.3}$$

The constant C_p *is the best possible.*

By the approximation arguments of Bichteler [Bi], the theorem above is a quick consequence of its discrete-time version, which we will prove next. Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, filtered by $(\mathcal{F}_n)_{n\ge 0}$. Let $f = (f_n)_{n\ge 0}$ be an adapted martingale and $g = (g_n)_{n\ge 0}$ be its transform by a predictable sequence $v = (v_n)_{n\ge 0}$ bounded in absolute value by 1. That is, we have

$$g_n = \sum_{k=0}^{n} v_k df_k, \qquad n = 0, 1, 2, \ldots,$$

where $df_0 = f_0$ and $df_k = f_k - f_{k-1}$ for $k \geq 1$. Here by predictability of v we mean that v_0 is $\mathcal{F}_0$-measurable and for any $k \geq 1$, v_k is measurable with respect to $\mathcal{F}_{k-1}$. In the particular case when each v_k is deterministic and takes values in the set $\{-1, 1\}$, we will say that g is a ± 1 *transform* of f.

Denote $f_n^* = \max_{k \leq n} f_k$ and $f^* = \sup_k f_k$.

THEOREM 1.4. *Suppose f, g are martingales such that g is a transform of f by a predictable sequence bounded in absolute value by* 1. *If $1 < p \leq \infty$, then*

$$\|g^*\|_1 \leq C_p \|f\|_p. \tag{1.4}$$

A few words about the organization of the paper. The proof of our result is based on Burkholder's technique, which exploits properties of certain special functions; the method is described in the next section. Section 3 contains the proof of (1.3) and (1.4) for $p \in (1, 2]$, while the case $p \in (2, \infty]$ is postponed to the final part of the paper, Section 4.

2. Some reductions and the method of proof. Using approximation arguments of Bichteler [Bi], it suffices to focus on the discrete-time setting. Now, with no loss of generality, we may assume that in (1.4) we deal with *simple* sequences f and g. By simplicity of f we mean that for any integer n, the random variable f_n takes only a finite number of values and there exists a deterministic number N such that $f_N = f_{N+1} = \cdots$ with probability 1. Clearly, if f and g are simple, then the almost sure limits f_∞ and g_∞ exist and are finite. Next, we may assume that $g_0 \geq 0$ almost surely, which gives $|g^*| = g^*$. Indeed, it suffices to replace v_0 by $\operatorname{sgn} f_0$ if necessary; then $|g^*|$ increases, so we obtain a stronger estimate to prove.

The key reduction is that it suffices to work with ± 1 transforms only. Recall Lemma A.1 from [B1].

LEMMA 2.1. *Let g be the transform of a martingale f by a real-valued predictable sequence v uniformly bounded in absolute value by* 1. *Then for each $j \geq 1$ there exist martingales $F^j = (F_n^j)_{n \geq 0}$ and $G^j = (G_n^j)_{n \geq 0}$ such that for $j \geq 1$ and $n \geq 0$,*

$$f_n = F_{2n+1}^j \quad \textit{and} \quad g_n = \sum_{j=1}^{\infty} 2^{-j} G_{2n+1}^j,$$

and G^j is a ± 1 transform of F^j.

To see how the lemma works in our setting, suppose we have established (1.4) for ± 1 transforms. Now, if g is a transform of f, then Lemma 2.1

gives us the processes F^j and G^j, for which we may write

$$\|g^*\|_1 = \Big\| \sup_n \sum_{j=1}^{\infty} 2^{-j} G^j_{2n+1} \Big\|_1 \le \sum_{j=1}^{\infty} 2^{-j} \|G^{j*}\|_1$$
$$\le C_p \sum_{j=1}^{\infty} 2^{-j} \|F^j\|_p = C_p \|f\|_p,$$

as needed.

Observe that in the proof of (1.4) we may assume that p is finite. Let $\mathcal{A} = \{(x,y,z) \in \mathbb{R}^3 : y \le z\}$ and define $V_p : \mathcal{A} \to \mathbb{R}$ by

$$V_p(x,y,z) = \begin{cases} y \vee z - |x|^p + \gamma_p(0) & \text{if } 1 < p \le 2, \\ y \vee z - |x|^p + M_p & \text{if } 2 < p < \infty, \end{cases}$$

where γ_p is given by (3.1) and M_p is introduced in (4.2) below. It is enough to show that

$$\mathbb{E} V_p(f_\infty, g_\infty, g^*_\infty) \le 0 \tag{2.1}$$

for all simple martingales f, g such that g is a ± 1 transform of f. This follows from a standard homogenization procedure. Indeed: for $1 < p \le 2$, apply (2.1) to the martingales f/λ, g/λ, where $\lambda > 0$ is fixed. This yields

$$\mathbb{E} g^*_\infty \le \lambda^{1-p} \mathbb{E}|f_\infty|^p - \lambda \gamma_p(0).$$

Now the choice

$$\lambda = \left(-\frac{p-1}{\gamma_p(0)} \right)^{1/p} \|f\|_p$$

gives (1.4). For $p > 2$ the reasoning is the same.

The estimate (2.1) will be achieved if we find a function $U : \mathcal{A} \to \mathbb{R}$ with the following three properties.

1° For any $\varepsilon \in \{-1,1\}$ and $(x,y,z) \in \mathcal{A}$ there is a number $c = c(\varepsilon,x,y,z)$ such that for all $d \in \mathbb{R}$,

$$U(x + \varepsilon d, y + d, (y+d) \vee z) \le U(x,y,z) + cd.$$

2° $U(x,y,z) \ge V_p(x,y,z)$ for all (x,y,z).

3° $U(x,y,y) \le 0$ for all x, y such that $x = |y|$.

The class of all functions U satisfying 1°–3° will be denoted by $\mathcal{U}(V_p)$.

Sometimes it is convenient to replace 1° with the following equivalent condition (see [B2]):

1°′ For any $\varepsilon \in \{-1,1\}$, $(x,y,z) \in \mathcal{A}$ and any simple centered random variable T, we have

$$\mathbb{E} U(x + \varepsilon T, y + T, (y+T) \vee z) \le U(x,y,z).$$

The relation between the inequality (2.1) and the class $\mathcal{U}(V_p)$ is described in the following fact.

THEOREM 2.2. *If the class $\mathcal{U}(V_p)$ is nonempty, then the inequality* (2.1) *holds for any simple f, g such that g is a ± 1 transform of f.*

Proof. Take $U \in \mathcal{U}(V_p)$ and simple f, g such that g is a ± 1 transform of f. The process $(U(f_n, g_n, g_n^*))_{n\geq 0}$ is a supermartingale: indeed, the inequality $\mathbb{E}[U(f_n, g_n, g_n^*) \,|\, \mathcal{F}_{n-1}] \leq U(f_{n-1}, g_{n-1}, g_{n-1}^*)$, $n \geq 1$, follows from the conditional form of $1^{\circ\prime}$, with $x = f_{n-1}$, $y = g_{n-1}$, $z = g_{n-1}^*$, $T = dg_n$ and $\varepsilon \in \{-1, 1\}$ such that $dg_n = \varepsilon df_n$. Consequently, using 2° and then 3°, one gets

$$\mathbb{E}V_p(f_\infty, g_\infty, g_\infty^*) \leq \mathbb{E}U(f_\infty, g_\infty, g_\infty^*) \leq \mathbb{E}U(f_0, g_0, g_0^*) \leq 0. \ \blacksquare$$

Thus the problem of proving a given martingale inequality (2.1) is reduced to the problem of constructing a function with properties 1°, 2° and 3°.

It turns out that the implication can be reversed. For V_p as above, consider $U_0 : \mathcal{A} \to \mathbb{R}$ given by

$$U_0(x, y, z) = \sup \mathbb{E}V_p(f_\infty, g_\infty, g_\infty^* \vee z),$$

where the supremum is taken over the class $M(x, y)$ of all pairs (f, g) of simple martingales such that $(f_0, g_0) = (x, y)$ and $dg_n = \pm df_n$ for all $n \geq 1$ (that is, there is a deterministic $v = (v_n)_{n\geq 1}$ taking values in $\{-1, 1\}$ such that $dg_n = v_n df_n$, $n \geq 1$).

THEOREM 2.3. *If* (2.1) *is valid, then the class $\mathcal{U}(V_p)$ is nonempty and U_0 is its least element.*

For the proof, one needs to slightly modify the argument used in [B2] (see Theorem 2.2 there). Theorem 2.3 will be quite useful in the proof of the optimality of the constants C_p. In the next two sections we will construct appropriate special functions.

3. The proof of (1.4) **for** $1 < p \leq 2$**.** We start by defining a function $\gamma_p : [0, \infty) \to (-\infty, 0]$ by

$$\gamma_p(t) = -\exp(pt^{p-1}) \int_t^\infty \exp(-ps^{p-1})\, ds. \tag{3.1}$$

Since

$$\gamma_p(t) = -\int_0^\infty \exp\Big\{-p(p-1)\int_0^s (t+u)^{p-2}\, du\Big\}\, ds,$$

the function γ_p is nonincreasing on $[0, \infty)$. Let $G_p : (-\infty, \gamma_p(0)] \to [0, \infty)$ denote the inverse of the function $t \mapsto \gamma_p(t) - t$, $t \geq 0$. We will need the following estimate.

Lemma 3.1. *We have* $G_pG_p'' + (p-2)(G_p')^2 \le 0$.

Proof. The inequality to be proved is equivalent to $(G_p/G_p')' \ge p-1$. Since $\gamma_p'(t) = p(p-1)t^{p-2}\gamma_p(t) + 1$, we obtain

$$G_p'(x) = (\gamma_p(G_p(x)) - 1)^{-1} = [p(p-1)G_p^{p-2}(x)(x + G_p(x))]^{-1}$$

and

$$1 + G_p'(x) = \frac{\gamma_p(G_p(x))}{p(p-1)G_p^{p-2}(x)\gamma_p(G_p(x))}.$$

Therefore

$$\left(\frac{G_p(x)}{G_p'(x)}\right)' = [p(p-1)G_p^{p-1}(x)(x+G_p(x))]' = p-1+\frac{G_p(x)\gamma_p'(G_p(x))}{\gamma_p(G(x))} \ge p-1,$$

because $G_p(x) \ge 0$ and $\gamma_p(G_p(x)) < 0$, $\gamma_p'(G_p(x)) \le 0$. ■

Now we are ready to introduce a special function. Let

$$D_1 = \{(x,y,z) \in \mathcal{A} : y - z - |x| \ge \gamma_p(0)\},$$
$$D_2 = \{(x,y,z) \in \mathcal{A} : y - z - |x| < \gamma_p(0) \text{ and } |x| \ge G_p(y - z - |x|)\},$$
$$D_0 = \mathcal{A} \setminus (D_1 \cup D_2).$$

Let $U_p : \mathcal{A} \to \mathbb{R}$ be given by

$$U_p(x,y,z) = \begin{cases} -\dfrac{(y-z)^2 - x^2}{2\gamma_p(0)} + \dfrac{\gamma_p(0)}{2} + y & \text{on } D_1, \\ z + \gamma_p(0) + (p-1)G_p(y-z-|x|)^p & \\ \qquad - p|x|G_p(y-z-|x|)^{p-1} & \text{on } D_2, \\ z - |x|^p + \gamma_p(0) & \text{on } D_0. \end{cases}$$

We will now verify that U_p belongs to $\mathcal{U}(V_p)$ and thus establish (1.4). To do this, it suffices to show the following fact.

Lemma 3.2.

(i) *The function* U_p *is of class* C^1 *in the interior of* $\mathcal{A}$.

(ii) *For any* $\varepsilon \in \{-1,1\}$ *and* $(x,y,z) \in \mathcal{A}$, *the function* $F = F_{\varepsilon,x,y,z} : (-\infty, z-y] \to \mathbb{R}$, *given by* $F(t) = U_p(x+\varepsilon t, y+t, z)$, *is concave.*

(iii) *For any* $\varepsilon \in \{-1,1\}$ *and* $x, y, h \in \mathbb{R}$,

$$U_p(x+\varepsilon t, y+t, (y+t)\vee y) \le U_p(x,y,y) + \varepsilon U_{px}(x,y,y)t + t. \tag{3.2}$$

(iv) *We have*

$$U_p(x,y,z) \ge V_p(x,y,z) \quad \textit{for } (x,y,z) \in \mathcal{A}. \tag{3.3}$$

(v) *We have*

$$\sup U_p(x,y,y) = 0, \tag{3.4}$$

where the supremum is taken over all x, y *satisfying* $|x| = |y|$.

Proof. (i) This is straightforward: U_p is of class C^1 in the interior of D_0, D_1 and D_2, so the claim reduces to tedious verification that the partial derivatives U_{px}, U_{py} and U_{pz} match at the common boundaries of D_0, D_1 and D_2.

(ii) In view of (i), it suffices to show that $F''(t) \leq 0$ for those t for which the second derivative exists. In view of the translation property $F_{\varepsilon,x,y,z}(u) = F_{\varepsilon,x+\varepsilon s,y+s,z}(u-s)$, valid for all u and s, it suffices to check $F''(t) \leq 0$ only for $t = 0$. Furthermore, since we have $U_{px}(0,y,z) = 0$ and $U_p(x,y,z) = U_p(-x,y,z)$, we may restrict ourselves to $x > 0$.

If $\varepsilon = 1$, then we easily verify that $F''(0) = 0$ if (x,y,z) lies in the interior $(D_1 \cup D_2)^o$ of $D_1 \cup D_2$ and $F''(0) = -p(p-1)x^{p-2} \leq 0$ if $(x,y,z) \in D_0^o$. Thus it remains to check the case $\varepsilon = -1$. We start from the observation that $F''(0) = 0$ if $(x,y,z) \in D_1^o$. If $(x,y,z) \in D_2^o$, then

$$F''(0) = 4p(p-1)G_p^{p-3}[G_pG_p'(G_p'+1) + (G_p - x)((p-2)(G_p')^2 + G_pG_p'')],$$

where all the functions on the right are evaluated at $x_0 = y - z - x$. Since $y \leq z$, we have $x \leq -x_0$ and, in view of Lemma 3.1,

$$\begin{aligned}(3.5)\qquad F''(0) &\leq 4p(p-1)G_p^{p-3}(x_0)[G_p(x_0)G_p'(x_0)(G_p'(x_0)+1) \\ &\qquad + (G_p(x_0)+x_0)((p-2)(G_p'(x_0))^2 + G_p(x_0)G_p''(x_0))] \\ &= 0.\end{aligned}$$

Here in the last step we have used the equality

$$G_p(x)G_p''(x) + (p-2)(G_p'(x))^2 = -\frac{G_p(x)G_p'(x)(G_p'(x)+1)}{G_p(x)+x},$$

which can be easily extracted from the proof of Lemma 3.1. Thus we are done with D_2^o. Finally, if (x,y,z) belongs to the interior of D_0, then $F''(0) = -p(p-1)x^{p-2} \leq 0$.

(iii) We may assume that $x \geq 0$, due to the symmetry of the function U_p. Note that $U_{py}(x,y-,y) = 1$; therefore, if $t \leq 0$, then the estimate follows from the concavity of U_p along the lines of slope ± 1, established in the previous part. If $t > 0$, then

$$U_p(x+\varepsilon t, y+t, (y+t)\vee y) = U_p(x, y+t, y+t) = y + t + U_p(x+\varepsilon t, 0, 0),$$

and hence we will be done if we show that the function $s \mapsto U_p(s,0,0)$ is concave on $[0,\infty)$. However, its second derivative equals $1/\gamma_p(0) < 0$ for $s < \gamma_p(0)$ and

$$\begin{aligned}p(p-1)G_p^{p-3}(-s)&[(G_p(-s)-s)((p-2)(G_p'(-s))^2 + G_p(-s)^{p-2}G_p''(-s)) \\ &+ G_p(-s)G_p'(-s)(G_p'(-s)+2)] \\ &= p(p-1)G_p(-s)^{p-2}G_p'(-s) \leq 0\end{aligned}$$

for $s > \gamma_p(0)$. Here we have used the equality from (3.5), with $x_0 = -s$.

(iv) Again, it suffices to deal only with nonnegative x. On the set D_0 both sides of (3.3) are equal. To prove the majorization on D_2, let $\Phi(s) = \gamma_p(0) - s^p$ for $s \geq 0$. Observe that

$$\begin{aligned} U_p(x,y,z) = z + \Phi(G_p(y-z-x)) \\ + \Phi'(G_p(y-z-x))(x - G_p(y-z-x)), \end{aligned}$$

which, by concavity of Φ, is not smaller than $z + \Phi(x)$. Finally, the estimate for $(x,y,z) \in D_1$ is a consequence of the fact that

$$U_{py}(x,y-,z) = \frac{\gamma_p(0) - (y-z)}{\gamma_p(0)} \geq 0,$$

so

$$U_p(x,y,z) - V_p(x,y,z) \geq U_p(x,y_0,z) - V_p(x,y_0,z) \geq 0.$$

Here $(x,y_0,z) \in \partial D_2$ and the latter bound follows from the majorization on D_2, which we have just established.

(v) We have

$$U_p(x,y,y) = U_p(|x|,0,0) + y \leq U_p(|x|,0,0) + |x|.$$

As shown in the proof of (iii), $s \mapsto U_p(s,0,0)$, $s \geq 0$, is concave, hence so is the function $s \mapsto U_p(s,0,0) + s$, $s \geq 0$. It suffices to note that its derivative vanishes at $-\gamma_p(0)$, so the value at this point (which is equal to 0) is the supremum we are searching for. ■

Sharpness. As shown by Peskir [P], the Doob-type bound

$$\|B_\tau^*\|_1 \leq \Gamma\left(\frac{2p-1}{p-1}\right)^{1-1/p} \|B_\tau\|_p, \quad 1 < p \leq 2,$$

is sharp. Here B is a Brownian motion (not necessarily starting from 0) and τ is a stopping time for B satisfying $\tau \in L^{p/2}$. Consequently, the estimate (1.4) is also sharp, even if $X = Y$. ■

4. The proof of (1.4) **for** $p > 2$**.** Suppose that p is finite. Let $\gamma_p : [0,\infty) \to (-\infty, 0)$ be given by

$$\begin{aligned} \gamma_p(t) &= \exp(-pt^{p-1})\Big[-\int_{p^{-1/(p-1)}}^{t} \exp(ps^{p-1})\,ds - p^{-1/(p-1)}e\Big] \\ &= -t + p(p-1)\exp(-pt^{p-1}) \int_{p^{-1/(p-1)}}^{t} s^{p-1}\exp(ps^{p-1})\,ds \end{aligned}$$

if $t > p^{-1/(p-1)}$, and

$$\gamma_p(t) = (p-2)(t - p^{-1/(p-1)}) - p^{-1/(p-1)}$$

if $t \in [0, p^{-1/(p-1)}]$. We start with the following straightforward fact.

LEMMA 4.1. *The function γ_p is of class C^1 and nondecreasing.*

Proof. The first assertion can be verified easily. To prove the second one, note that it suffices to show $\gamma_p'(t) \geq 0$ for $t \geq p^{-1/(p-1)}$. Equivalently, $\gamma_p'(t) \geq 0$ reads

$$t^{2-p}\exp(pt^{p-1}) - p(p-1)\int_{p^{-1/(p-1)}}^{t}\exp(ps^{p-1})\,ds - p^{(p-2)/(p-1)}(p-1)e \leq 0.$$

However, the inequality is true for $t = p^{-1/(p-1)}$ and the derivative of the left-hand side equals $(2-p)t^{1-p}\exp(pt^{p-1}) \leq 0$. This completes the proof. ∎

Let $G_p : [0,\infty) \to [p^{-1/(p-1)},\infty)$ be the inverse to the function $t \mapsto \gamma_p(t) + t$, $t \geq p^{-1/(p-1)}$ (the function is invertible, by the previous fact). We have the following version of Lemma 3.1.

LEMMA 4.2. *We have $G_pG_p'' + (p-2)(G_p')^2 \geq 0$.*

Proof. It can be verified that

$$G_p(x)G_p''(x) + (p-2)(G_p'(x))^2 = \frac{G_p(x)G_p'(x)(G_p'(x)-1)}{x - G_p(x)}, \tag{4.1}$$

and this is nonnegative: it follows from the very definition of G_p that $G_p(x) \geq 0$, $G_p'(x) \geq 0$ and $G_p'(x) \leq 1$, $x - G_p(x) < 0$. ∎

Define

$$M_p = \frac{p-1}{p^{p/(p-1)}}\left[2^{p/(p-1)} - \frac{p}{p-1}\int_1^2 s^{1/(p-1)}e^{s-2}\,ds\right]. \tag{4.2}$$

Let $H_p : \mathbb{R}^2 \to \mathbb{R}$ be given by

$$H_p(x,y) = (p-1)^{1-p}(-(p-1)|x| + |y|)(|x| + |y|)^{p-1}$$

and put

$$\begin{aligned} D_1 &= \{(x,y,z) \in \mathcal{A} : y - z \geq \gamma_p(x),\ x + y - z \leq 0\},\\ D_2 &= \{(x,y,z) \in \mathcal{A} : y - z \geq \gamma_p(x),\ x + y - z > 0\},\\ D_0 &= \mathcal{A} \setminus (D_1 \cup D_2). \end{aligned}$$

Introduce $U_p : \mathcal{A} \to \mathbb{R}$ by

$$U_p(x,y,z) = \begin{cases} z + H_p(x, y - z + (p-1)p^{-1/(p-1)}) - M_p & \text{on } D_1,\\ z - M_p + (p-1)G_p(|x| + y - z)^p & \\ \qquad - p|x|G_p(|x| + y - z)^{p-1} & \text{on } D_2,\\ z - |x|^p - M_p & \text{on } D_0. \end{cases}$$

Here is the analogue of Lemma 3.2. Again, once we show it, we will be done with the proof of (1.4).

LEMMA 4.3.

(i) *The function* U_p *is of class* C^1.

(ii) *For any* $\varepsilon \in \{-1,1\}$ *and* $(x,y,z) \in \mathcal{A}$, *the function* $F = F_{\varepsilon,x,y,z} : (-\infty, z-y] \to \mathbb{R}$, *given by* $F(t) = U_p(x+\varepsilon t, y+t, z)$, *is concave.*

(iii) *For any* $\varepsilon \in \{-1,1\}$ *and* $x,y,h \in \mathbb{R}$,

$$U_p(x+\varepsilon t, y+t, (y+t)\vee y) \le U_p(x,y,y) + \varepsilon U_{px}(x,y,y)t + t. \tag{4.3}$$

(iv) *We have*

$$U_p(x,y,z) \ge V_p(x,y,z) \quad \text{for } (x,y,z) \in \mathcal{A}. \tag{4.4}$$

(v) *We have*

$$\sup U_p(x,y,y) = 0, \tag{4.5}$$

where the supremum is taken over all x,y *satisfying* $|x| = |y|$.

Proof. (i) Straightforward.

(ii) We proceed as in the proof of Lemma 3.2(ii) and check $F''(0) \le 0$ for $x > 0$ and (x,y,z) lying in the interior of some D_i.

If $\varepsilon = 1$, there is nothing to check: we have $F''(0) = 0$ if $(x,y,z) \in (D_1 \cup D_2)^o$ or $F''(0) = -p(p-1)x^{p-2} \le 0$ if $(x,y,z) \in D_0^o$. It remains to verify the case $\varepsilon = -1$. If (x,y,z) belongs to the interior of D_1, then $F''(0) \le 0$; this follows from the fact that for any $(x',y') \in \mathbb{R}^2$, the function $t \mapsto H_p(x'+t, y'-t)$ is concave (see [B1, p. 17]). If $(x,y,z) \in D_2^o$, then

$$F''(0) = 4p(p-1)G_p^{p-3}[G_pG_p'(G_p'-1) + (G_p - x)((p-2)(G_p')^2 + G_pG_p'')],$$

where all the functions on the right are evaluated at $x_0 = x+y-z$. We have $y \le z$, so $x \le x_0$ and, by Lemma 4.2,

$$\begin{aligned} F''(0) &\le 4p(p-1)G_p^{p-3}(x_0)[G_p(x_0)G_p'(x_0)(G_p'(x_0)-1) \\ &\qquad + (G_p(x_0)-x_0)((p-2)(G_p'(x_0))^2 + G_p(x_0)G_p''(x_0))] \\ &= 0, \end{aligned}$$

where we have used the equality from (4.1). Finally, if (x,y,z) belongs to the interior of D_0, then $F''(0) = -p(p-1)x^{p-2} \le 0$.

(iii) We have $U_{py}(x,y-,y) = 1$ and $U_p(x,y,y) = y + U_p(x,0,0)$. Therefore, arguing as in the proof of Lemma 3.2, we see that it suffices to show that the function $s \mapsto U_p(s,0,0)$, $s > 0$, is concave. Indeed, its second derivative at s equals

$$-p(p-1)G_p^{p-2}(s)G_p'(s) \le 0 \tag{4.6}$$

and we are done.

(iv) The majorization can be proved in the same manner as in Lemma 3.2, using the concave function $\Phi(s) = -s^p$, $s \ge 0$. The details are left to the reader.

(v) Observe that

$$U_p(x,y,y) = y + U_p(|x|,0,0) \le |x| + U_p(|x|,0,0).$$

Denoting the right-hand side by $\Psi(|x|)$, we find that Ψ is concave on $(0,\infty)$ (see the proof of (iii)) and

$$\begin{aligned}\Psi'(t) &= p(p-1)G_p'(t)G_p(t)^{p-2}(G_p(t)-t) - pG_p(t)^{p-1} + 1\\ &= -pG_p(t)^{p-1} + 2.\end{aligned}$$

Consequently, Ψ attains its maximum at the point t_0 satisfying $G_p(t_0) = (2/p)^{1/(p-1)}$, or

$$\begin{aligned} t_0 &= \gamma_p((2/p)^{1/(p-1)}) + (2/p)^{1/(p-1)} \\ &= p(p-1)e^{-2} \int_{p^{-1/(p-1)}}^{(p/2)^{-1/(p-1)}} s^{p-1}\exp(ps^{p-1})\,ds \\ &= p^{-1/(p-1)} \int_1^2 s^{1/(p-1)} e^{s-2}\,ds, \end{aligned} \tag{4.7}$$

and, as one easily checks, the maximum is equal to 0. This completes the proof. ∎

Sharpness, $2 < p < \infty$. We have, by Young's inequality,

$$c\|f\|_p \le \|f\|_p^p + p^{-p/(p-1)}(p-1)c^{p/(p-1)},$$

so if (1.4) held with some $c < C_p$, we would have

$$\|g^*\|_1 \le \|f\|_p^p + C \tag{4.8}$$

for some $C < p^{-p/(p-1)}(p-1)C_p^{p/(p-1)} = M_p$. Therefore it suffices to show that the smallest C for which (4.8) is valid equals M_p.

Suppose, then, that (4.8) holds with some universal C, and let us use Theorem 2.3 with $V = V_p$ given by $V_p(x,y,z) = z - |x|^p$. As a result, we obtain a function U_0 satisfying 1°–3°. Observe that for any $(x,y,z) \in \mathcal{A}$ and $t \in \mathbb{R}$,

$$U_0(x,y,z) = t + U_0(x, y-t, z-t). \tag{4.9}$$

This is a consequence of the fact that the function V_p also has this property, and of the very definition of U_0.

Now it is convenient to split the proof into a few parts.

Step 1. First we will show that for any y,

$$U_0(0,y,y) \ge y + (p-1)p^{-p/(p-1)} = U_p(0,y,y). \tag{4.10}$$

In view of (4.9), it suffices to prove this for $y = 0$. Let $d = p^{-1/(p-1)}$ and $\delta > 0$. Applying 1°′ to $\varepsilon = -1$, $x = y = z = 0$ and a mean-zero T taking values δ and $-d$, we obtain

$$U_0(0,0,0) \geq \frac{d}{d+\delta}U_0(-\delta,\delta,\delta) + \frac{\delta}{d+\delta}U_0(d,-d,0).$$

By (4.9), $U_0(-\delta,\delta,\delta) = \delta + U_0(-\delta,0,0)$. Furthermore, by 2°, $U_0(d,-d,0) \geq -d^p$, so the above estimate yields

$$U_0(0,0,0) \geq \frac{d}{d+\delta}(\delta + U_0(-\delta,0,0)) - \frac{\delta}{d+\delta}|d|^p. \tag{4.11}$$

Similarly, one uses property $1^{\circ\prime}$ and then 2° to get

$$\begin{aligned} U_0(-\delta,0,0) &\geq \frac{d}{d+\delta}U_0(0,\delta,\delta) + \frac{\delta}{d+\delta}U_0(-d-\delta,-d,0) \\ &\geq \frac{d}{d+\delta}(\delta + U_0(0,0,0)) - \frac{\delta}{d+\delta}(d+\delta)^p. \end{aligned}$$

Combining this with (4.11), subtracting $U_0(0,0,0)$ from both sides of the resulting estimate, dividing through by δ and letting $\delta \to 0$ leads to $U_0(0,0,0) \geq d - d^p = U_p(0,0,0)$, which is what we need.

Consequently, by the definition of U_0, for any $y \in \mathbb{R}$ and $\kappa > 0$ there is a pair $(f^{\kappa,y}, g^{\kappa,y}) \in M(0,y)$ satisfying

$$U_p(0,y,y) \leq V_p(f_\infty^{\kappa,y}, g_\infty^{\kappa,y}, (g_\infty^{\kappa,y})^*) + \kappa. \tag{4.12}$$

Step 2. Let N be a positive integer and let $\delta = t_0/N$, where t_0 is given by (4.7). We will need the following auxiliary fact.

Lemma 4.4. *There is a universal R such that the following holds. If $x \in [\delta, t_0]$, $y \in \mathbb{R}$ and T is a centered random variable which takes values in $[\gamma_p(G_p(x)), \delta]$, then*

$$\mathbb{E}U_p(x-T, y+T, (y+T)\vee y) \leq U_p(x,y,y) + R\delta^2. \tag{4.13}$$

Proof. We start from the observation that for any fixed $x \in [\delta, t_0]$ and $y \in \mathbb{R}$, if $t \in [-\gamma_p(G_p(x)), 0]$,

$$U_p(x-t, y+t, y) = U_p(x,y,y) - U_{px}(x,y,y)t + t.$$

For $t \in (0,\delta]$, by the concavity of $s \mapsto U_p(s,0,0)$,

$$\begin{aligned} U_p(x-t, y+t, y+t) &= y + t + U_p(x-t,0,0) \\ &\geq y + t + U_p(x,0,0) - U_{px}(x,0,0)t - R\delta^2 \\ &= U_p(x,y,y) - U_{px}(x,y,y)t + t - R\delta^2. \end{aligned}$$

Here, for example, one may take $R = -\inf_{x\in[0,t_0]} U_{pxx}(x,0,0)$, which is finite: see (4.6). The inequality (4.13) follows immediately from the above two estimates. ∎

Now consider a martingale $f = (f_n)_{n=1}^N$, starting from t_0, which satisfies the following condition: if $0 \leq n \leq N-1$, then on the set $\{f_n = t - n\delta\}$, the difference df_{n+1} takes values $-\delta$ and $-\gamma_p(G_p(f_n(\omega)))$; on the complement of this set, $df_{n+1} \equiv 0$. Let g be the ± 1 transform of f given by $g_0 = f_0$ and $dg_n = -df_n$, $n = 1, \ldots, N$. The key fact about the pair (f, g) is that

(4.14)
$$\mathbb{E}U_p(f_n, g_n, g_n^*) \le \mathbb{E}U_p(f_{n+1}, g_{n+1}, g_{n+1}^*) + R\delta^2, \qquad n = 0, 1, \ldots, N-1.$$

This is an immediate consequence of Lemma 4.4 (applied conditionally with respect to $\mathcal{F}_n$) and the fact that $U_p(f_n, g_n, g_n^*) \ne U_p(f_{n+1}, g_{n+1}, g_{n+1}^*)$ if and only if $f_n = t - n\delta$ or $g_n = t + n\delta = g_n^*$.

The next property of the pair (f, g) is that if $f_N \ne 0$, then we have $U_p(f_N, g_N, g_N^*) = V_p(f_N, g_N, g_N^*)$. Indeed, $f_N \ne 0$ implies $df_n > 0$ for some $n \ge 1$ and then, by construction,

$$g_N^* - g_N = g_n^* - g_n = -dg_n = df_n = \gamma_p(f_n) = \gamma_p(f_N).$$

Thus we may write

$$\begin{aligned} \text{(4.15)} \quad M_p &= U_p(t_0, t_0, t_0) \\ &\le \mathbb{E}U_p(f_N, g_N, g_N^*) + RN\delta^2 \\ &= \mathbb{E}V_p(f_N, g_N, g_N^*)1_{\{f_N \ne 0\}} + U_p(0, 2t_0, 2t_0)\mathbb{P}(f_N = 0) + RN\delta^2, \end{aligned}$$

since $g_N = g_N^* = 2t_0$ on $\{f_N = 0\}$.

Step 3. Now let us extend the pair (f, g) as follows. Fix $\kappa > 0$ and put $f_N = f_{N+1} = f_{N+2} = \cdots$ and $g_N = g_{N+1} = g_{N+2} = \cdots$ on $\{f_N \ne 0\}$, while on $\{f_N = 0\}$, let the conditional distribution of $(f_n, g_n)_{n \ge N}$ with respect to $\{f_N = 0\}$ be that of the pair $(f^{\kappa, 2t_0}, g^{\kappa, 2t_0})$, obtained at the end of Step 1. The process (f, g) we get consists of simple martingales and, by (4.12) and (4.15), we have

$$M_p \le \mathbb{E}V_p(f_\infty, g_\infty, g_\infty^*) + RN\delta^2 + \kappa\mathbb{P}(f_N = 0).$$

Now it suffices to note that choosing N sufficiently large and κ sufficiently small, we can make the expression $RN\delta^2 + \kappa\mathbb{P}(f_N = 0)$ arbitrarily small. This shows that M_p is indeed the smallest C which is allowed in (4.8). ■

Sharpness, $p = \infty$. We may assume that $\|X\|_\infty = 1$. The proof will be entirely based on the following version of Theorem 2.3.

Theorem 4.5. *Let $U_0 : \{(x, y, z) : |x| \le 1,\ y \le z\} \to \mathbb{R}$ be given by*

$$U_0(x, y, z) = \mathbb{E}g_\infty^* \vee z,$$

where the supremum is taken over the class of all pairs $(f, g) \in M(x, y)$ such that $\|f\|_\infty \le 1$. Then U_0 enjoys the following properties:

1° *For any $\varepsilon \in \{-1, 1\}$, $x \in [-1, 1]$, $y \le z$ and any simple centered random variable T satisfying $|x + \varepsilon T| \le 1$, we have*

$$\mathbb{E}U_0(x + \varepsilon T, y + T, (y + T) \vee z) \le U_0(x, y, z).$$

2° $U_0(x,y,z) \ge z$ *for all* (x,y,z) *from the domain of* U_0.
3° $U_0(x,y,y) \le C_\infty$ *for all* x, y *such that* $|x| = |y| \in [-1,1]$.

For the proof, modify the argument from [B2]. Note that the function U_0 satisfies (4.9) (with the obvious restriction to x lying in $[-1,1]$).

Now we turn to the optimality of the constant C_∞. First we will show that

$$U_0(0,0,0) \ge 1. \tag{4.16}$$

To prove this, take $\delta \in (0,1)$ and use 1° to obtain

$$U_0(0,0,0) \ge \frac{1}{1+\delta} U_0(\delta,\delta,\delta) + \frac{\delta}{1+\delta} U_0(-1,-1,0).$$

We have $U_0(-1,-1,0) \ge 0$ by 2°, and $U_0(\delta,\delta,\delta) = \delta + U(\delta,0,0)$ by (4.9). Thus we have

$$U_0(0,0,0) \ge \frac{\delta + U_0(\delta,0,0)}{1+\delta}. \tag{4.17}$$

Similarly, using 1° and then 2° yields

$$U(\delta,0,0) \ge (1-\delta)U_0(0,\delta,\delta) + \delta U_0(1,\delta-1,0) \ge (1-\delta)[\delta + U_0(0,0,0)].$$

Plug this into (4.17), subtract $U_0(0,0,0)$ from both sides, divide through by δ and let $\delta \to 0$. As a result, one gets (4.16).

Now fix a positive integer N and set $\delta = (1-e^{-1})/N$. For any $k = 1,\ldots,N$, we have, by 1°, 2° and (4.9),

$$\begin{aligned} U_0(k\delta,0,0) &\ge \frac{\delta}{1-k\delta+\delta} U_0(1,k\delta-1,0) + \frac{1-k\delta}{1-k\delta+\delta} U_0((k-1)\delta,\delta,\delta) \\ &\ge \frac{1-k\delta}{1-k\delta+\delta}[\delta + U_0((k-1)\delta,0,0)], \end{aligned}$$

or, equivalently,

$$\frac{U_0(k\delta,0,0)}{1-k\delta} \ge \frac{U_0((k-1)\delta,0,0)}{1-(k-1)\delta} + \frac{\delta}{1-(k-1)\delta}.$$

It follows by induction that

$$eU_0(1-e^{-1},0,0) = \frac{U_0(N\delta,0,0)}{1-N\delta} \ge U_0(0,0,0) + \sum_{k=1}^{N} \frac{\delta}{1-(k-1)\delta}.$$

Letting $N \to \infty$ and using (4.16), we arrive at

$$eU_0(1-e^{-1},0,0) \ge 1 + \int_0^{1-e^{-1}} \frac{dx}{1-x} = 2,$$

and hence, by (4.9),

$$U_0(1-e^{-1},1-e^{-1},1-e^{-1}) = 1-e^{-1} + U_0(1-e^{-1},0,0) \ge 1+e^{-1}.$$

It suffices to apply 3° to complete the proof. ∎

Acknowledgments. The author would like to thank the referee for comments and suggestions, which improved the presentation of the paper. The research was partially supported by Foundation for Polish Science and MNiSW Grant N N201 397437.

References

[Bi] K. Bichteler, *Stochastic integration and L^p-theory of semimartingales*, Ann. Probab. 9 (1980), 49–89.

[B1] D. L. Burkholder, *Explorations in martingale theory and its applications*, in: École d'Été de Probabilités de Saint-Flour XIX—1989, Lecture Notes in Math. 1464, Springer, Berlin, 1991, 1–66.

[B2] —, *Sharp norm comparison of martingale maximal functions and stochastic integrals*, in: Proceedings of the Norbert Wiener Centenary Congress, 1994 (East Lansing, MI, 1994), Proc. Sympos. Appl. Math. 52, Amer. Math. Soc., Providence, RI, 1997, 343–358.

[O1] A. Osękowski, *Sharp maximal inequality for stochastic integrals*, Proc. Amer. Math. Soc. 136 (2008), 2951–2958.

[O2] —, *Sharp maximal inequality for martingales and stochastic integrals*, Electron. Comm. Probab. 14 (2009), 17–30.

[P] G. Peskir, *The best Doob-type bounds for the maximum of Brownian paths*, Progr. Probab. 43 (1998), 287–296.

Adam Osękowski
Department of Mathematics, Informatics and Mechanics
University of Warsaw
Banacha 2
02-097 Warszawa, Poland
E-mail: ados@mimuw.edu.pl

Received September 3, 2010;
received in final form October 21, 2010 (7778)