SHARP MAXIMAL INEQUALITY FOR NONNEGATIVE MARTINGALES

ADAM OSĘKOWSKI

Abstract. Let X be a nonnegative martingale, H be a predictable process taking values in $[-1,1]$ and let Y be an Itô integral of H with respect to X. We establish the bound

$$||\sup_{t\geq 0}|Y_t|||_1 \leq 3||\sup_{t\geq 0} X_t||_1$$

and show that the constant 3 is the best possible.

1. Introduction

The purpose of the paper is to establish a sharp inequality for stochastic integrals in which the integrator is a nonnegative martingale. Let us introduce the necessary background and notation. Assume that $(\Omega,\mathcal{F},\mathbb{P})$ is a complete probability space, filtered by a nondecreasing right-continuous family $(\mathcal{F}_t)_{t\geq 0}$ of sub-σ-algebras of $\mathcal{F}$. In addition, let $\mathcal{F}_0$ contain all the events of probability 0. Suppose that $X=(X_t)_{t\geq 0}$ is an adapted real-valued martingale, which has right-continuous paths with limits from the left and let $H=(H_t)_{t\geq 0}$ be a predictable process taking values in the interval $[-1,1]$. Let $Y=(Y_t)_{t\geq 0}$ be the Itô integral of H with respect to X, that is, for $t\geq 0$,

$$Y_t = H_0X_0 + \int_{(0,t]} H_s dX_s.$$

Furthermore, let $X^*=\sup_{t\geq 0}|X_t|$ denote the maximal function of X and let $||X||_p=\sup_{t\geq 0}||X_t||_p$ be the p-th moment of X, for $1\leq p\leq\infty$.

The martingale Y can be viewed as the result of control of the martingale X by the process H, and hence there is a natural question about the comparison of the sizes of X and Y. An excellent source of information about this subject is the survey [2], which contains, among other things, moment, weak-type and exponential estimates for Y. In the present paper we will be interested in sharp bounds for the maximal functions X^* and Y^*. There is a method, introduced by Burkholder in [3], which allows to determine the optimal values of constants in inequalities of this type. Using this technique, Burkholder proved the following.

Theorem 1.1. *If X and Y are as above, then*

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

2000 *Mathematics Subject Classification.* Primary: 60G42. Secondary: 60G44.

Key words and phrases. Martingale, maximal function, stochastic integral, martingale transform.

Partially supported by Foundation for Polish Science and MNiSW Grant N N201 397437.

where $\kappa = 2.536\ldots$ is the unique positive solution to the equation

$$\kappa = 3 - \exp\frac{1-\kappa}{2}.$$

Let us mention here two further results in this direction, which are due to the author. In [4] it was shown that if the martingale X is nonnegative (that is, $X_t \geq 0$ for all t), then the best constant in (1.1) decreases to $2 + (3e)^{-1} = 2,1226\ldots$. In [5], a related estimate was studied, with the first moment of Y replaced by the first moment of Y^*. Precisely, it was shown that if X is real valued and Y is as above, then

$$||Y^*||_1 \leq \eta ||X^*||_1, \tag{1.2}$$

with optimal η equal to $3.4351\ldots$. The precise decription of η is quite complicated and is related to solutions of certain ODE's; for details, see [5].

In the present paper we continue this line of research and prove a sharp version of (1.2) under the assumption that the martingale X is nonnegative. Thus we complete the description of the optimal constants in the maximal L^1 inequalities for stochastic integrals.

Theorem 1.2. *If X is a nonnegative martingale, H is a predictable process with values in $[-1,1]$ and Y is the stochastic integral of H with respect to X, then*

$$||Y^*||_1 \leq 3||X^*||_1 \tag{1.3}$$

and the constant 3 is the best possible. It is already the best possible if H is assumed to take values in $\{-1,1\}$.

In fact, we will restrict ourselves to the discrete-time version of the result above. Using standard approximation theorems due to Bichteler [1], one easily obtains the above continuous time version (see [3] for analogous argumentation). Let us reformulate our problem in this new setting. Let $f = (f_n)_{n\geq 0}$ be a discrete-time real-valued martingale, with a difference sequence $(df_n)_{n\geq 0}$ given by $df_0 = f_0$ and $df_n = f_n - f_{n-1}$ for $n \geq 1$. Let $v = (v_n)_{n\geq 0}$ be a predictable sequence taking values in $[-1,1]$ and let $g = (g_n)_{n\geq 0}$ be a transform of f by v: that is, assume that $dg_n = v_n df_n$ for $n \geq 0$. In the particular case when v_n is deterministic and takes values in $\{-1,1\}$, we will write $dg_n = \pm df_n$. If we have $dg_n = \pm df_n$ for all n, we will say that g is a ± 1-transform of f. We will use the notation $f^* = \sup_{n\geq 0}|f_n|$ and $||f||_p = \sup_n ||f_n||_p$, analogous to the one used in the continuous-time setting.

The discrete-time version of Theorem 1.2 can be stated as follows.

Theorem 1.3. *Assume that f is a nonnegative martingale and g is its transform by a predictable sequence bounded in absolute value by 1. Then*

$$||g^*||_1 \leq 3||f^*||_1 \tag{1.4}$$

and the constant 3 is the best possible. It is already the best possible even if g is assumed to be a ± 1-transform of f.

A few words about the proof and the organization of the paper. We will make a heavy use of Burkholder's technique, described in the next section. The method turns the problem of proving a given maximal inequality for martingales into the problem of finding an upper solution to a certain nonlinear problem (see also [2] for discussion in the case of non-maximal inequallities). Comparing to the results mentioned above, the estimate (1.4) turns out to be much more difficult. Namely, in most of the inequalities where the Burkholder's method has been succesfully implemented, the corresponding nonlinear problems were two-dimensional; in the present paper we will have to find a special function of three variables, and this makes a significant complication. The function is constructed in Section 3, and the final part of the paper concerns the optimality of the constant 3.

2. Burkholder's method

Our goal is to determine the least β such that

$$||g^*||_1 \leq \beta ||f^*||_1 \tag{2.1}$$

for f and g as in the statement of Theorem 1.3. Let us start with some reductions. First, by Lebesgue's monotone convergence theorem, it suffices to show that

$$||g_n^*||_1 \leq \beta ||f_n^*||_1 \tag{2.2}$$

for all nonnegative integers n. Here $f_n^* = \sup_{k\leq n} f_k$ and $g_n^* = \sup_{k\leq n} |g_k|$. The second reduction is that we may assume that f is simple: for any n the random variable f_n takes only a finite number of values. Moreover, adding a small $\varepsilon > 0$ if necessary, we may consider only these f, for which f_0 is strictly positive with probability 1. The final observation is that we may restrict ourselves to ± 1 transforms. To see this, let us use the following modification of Lemma A.1 from [2].

Lemma 2.1. *Let g be the transform of a nonnegative martingale f by a real-valued predictable sequence v uniformly bounded in absolute value by* 1. *Then there exist nonnegative martingales $F^j = (F_n^j)_{n\geq 0}$ and Borel measurable functions $\varphi_j : [-1, 1] \to \{-1, 1\}$ such that, for $j \geq 1$ and $n \geq 0$,*

$$f_n = F_{2n+1}^j, \qquad f_n^* = (F_{2n+1}^j)^*,$$

$$g_n = \sum_{j=1}^{\infty} 2^{-j} \varphi_j(v_0) G_{2n+1}^j,$$

where G^j is the transform of F^j by $\varepsilon = (\varepsilon_k)_{k\geq 0}$ with $\varepsilon_k = (-1)^k$.

The proof of this fact goes along the same lines as in [2], and hence is omitted. Now, if we have (2.2) for ± 1 transforms, and g is a transform of f as in the lemma above, then

$$||g_n^*||_1 \leq \sum_{j=1}^{\infty} 2^{-j} ||(G_{2n+1}^j)^*||_1 \leq \beta \sum_{j=1}^{\infty} 2^{-j} ||(F_{2n+1}^j)^*||_1 = \beta ||f_n^*||_1,$$

as needed.

To study (2.2) for ± 1 transforms, we introduce the class $\mathcal{U}(\beta)$ which consists of those functions $U : [0, \infty) \times \mathbb{R} \times (0, \infty) \times (0, \infty) \to \mathbb{R}$, which satisfy

$$U(x, y, z, w) = U(x, y, x \vee z, |y| \vee w), \tag{2.3}$$

$$U(x, y, x, |y|) \leq 0 \quad \text{ if } x \geq |y| > 0, \tag{2.4}$$

$$|y| \vee w - \beta x \vee z \leq U(x, y, z, w), \tag{2.5}$$

and, furthermore, for any z, $w > 0$, $x \in [0, z]$, $y \in [-w, w]$, $\varepsilon \in \{-1, 1\}$, $\alpha \in (0, 1)$ and t_1, $t_2 \geq -x$ such that $\alpha t_1 + (1 - \alpha)t_2 = 0$,

$$\alpha U(x + t_1, y + \varepsilon t_1, z, w) + (1 - \alpha)U(x + t_2, y + \varepsilon t_2, z, w) \leq U(x, y, z, w). \tag{2.6}$$

By straightforward induction, the latter condition implies the following. If x, y, z, w and ε are as above, then for any simple mean-zero random variable T,

$$\mathbb{E}U(x+T, y+\varepsilon T, z, w) \leq U(x,y,z,w). \tag{2.7}$$

Let us describe the connection between the class $\mathcal{U}(\beta)$ and the estimate (2.2).

Theorem 2.2. *If $\mathcal{U}(\beta)$ is nonempty, then* (2.2) *is valid.*

Proof. Let $U \in \mathcal{U}(\beta)$. The key fact is that the sequence $(U(f_k, g_k, f_k^*, g_k^*))_{k\geq 0}$ is a supermartingale. Indeed, we have

$$\begin{aligned}\mathbb{E}[U(f_k, g_k, f_k^*, g_k^*)|\mathcal{F}_{k-1}] &= \mathbb{E}[U(f_k, g_k, f_{k-1}^*, g_{k-1}^*)|\mathcal{F}_{k-1}]\\ &= \mathbb{E}[U(f_{k-1}+df_k, g_{k-1}+\varepsilon_k df_k, f_{k-1}^*, g_{k-1}^*)|\mathcal{F}_{k-1}]\\ &\leq U(f_{k-1}, g_{k-1}, f_{k-1}^*, g_{k-1}^*),\end{aligned}$$

where in the first passage we have used (2.3) and in the latter we have exploited (2.7) conditionally on $\mathcal{F}_{k-1}$. Therefore, by (2.5) and then (2.4),

$$\mathbb{E}\big\{g_n^* - \beta f_n^*\big\} \leq \mathbb{E}U(f_n, g_n, f_n^*, g_n^*) \leq \mathbb{E}U(f_0, g_0, f_0^*, g_0^*) \leq 0. \quad \square$$

We have the following result in the reverse direction, Theorem 2.2 from [3], which will be useful in providing the lower bound for the constant β. For $x \geq 0$ and $y \in \mathbb{R}$, let $\mathcal{M}(x,y)$ be the class of all pairs (f,g) of simple martingales such that f is nonnegative and starts from x, g starts from y and $dg_n = \pm df_n$ for all $n \geq 1$.

Theorem 2.3. *Suppose that the inequality* (2.2) *holds for all n and all pairs (f,g) such that f is a simple nonnegative martingale and g is its ± 1-transform. Then the class $\mathcal{U}(\beta)$ is nonempty. Furthermore, let $U_0 : [0,\infty)\times\mathbb{R}\times(0,\infty)\times(0,\infty) \to \mathbb{R}$ be given by*

$$U_0(x,y,z,w) = \sup\{\mathbb{E}(g_n^* \vee w - \beta f_n^* \vee z)\}, \tag{2.8}$$

where the supremum is taken over all n and all pairs $(f,g) \in \mathcal{M}(x,y)$. Then U_0 is the least element in $\mathcal{U}(\beta)$.

Thus an equivalent reformulation of our goal is to determine the least β for which the class $\mathcal{U}(\beta)$ is nonempty. This is the purpose of the next section.

3. Proof of (1.4)

This section contains of two parts. First we shall exhibit a special function which belongs to $\mathcal{U}(3)$ (and thus establish (1.4)); in the second we will sketch some steps which lead to its discovery.

3.1. **An element of $\mathcal{U}(3)$.** Let $S = \{(x,y,w) : x \in [0,1],\ |y| \leq w\}$ and consider the following subsets of S:

$$\begin{aligned}D_1 &= \{(x,y,w) \in S : |y| \leq x\},\\ D_2 &= \{(x,y,w) \in S : x \leq |y| \leq x+w-1\},\\ D_3 &= \{(x,y,w) \in S : x+w-1 < |y| \leq w\}.\end{aligned}$$

Let $u : S \to \mathbb{R}$ be given as follows. First, if $w \geq 1$, then $u(x,y,w)$ equals

$$\begin{cases} \frac{2}{3}\exp[\frac{1}{2}(1-w)]\big\{2+(2x+|y|-2)(-x+|y|+1)^{1/2}\big\}+w-3 & \text{on } D_1,\\ 2x\exp[\frac{1}{2}(-x+|y|-w+1)]+w-3 & \text{on } D_2,\\ 2x - x\log(x-|y|+w)+w-3 & \text{on } D_3 \end{cases}$$

(with the convention $0\log 0 = 0$). If $w < 1$, then we set $u(x,y,w) = u(x,y,1)$.

The key properties of the function u are studied in the lemma below.

Lemma 3.1. *Let $w > 0$ be fixed.*

(i) If $y \in [-w-1, w]$, then the function $G = G_{y,w} : [0,1] \to \mathbb{R}$, given by $G(t) = u(t, y+t, w \vee |y+t|)$, is concave and nondecreasing.

(ii) The function $J = J_w : \mathbb{R} \to \mathbb{R}$, given by $J(t) = u(1, t, w \vee |t|)$, is even and convex.

(iii) We have $u(x,y,w) \geq w - 3$ on S.

Proof. (i) Since $u(t, y+t, w \vee |y+t|) = u(t, y+t, 1 \vee |y+t|)$ for $w < 1$, we may assume that $w \geq 1$. We will consider three cases separately.

1° If $y \geq w - 1$, then

$$G(t) = \begin{cases} 2t - t\log(-y+w) + w - 3 & \text{for } t \leq w - y, \\ 3t - t\log t + y - 3 & \text{for } t > w - y \end{cases}$$

is a concave and nondecreasing function.

2° If $y \in [0, w-1)$, then $G(t) = 2t\exp[\frac{1}{2}(y - w + 1)] + w - 3$ is linear and nondecreasing.

3° Suppose that $y \in [-w-1, 0)$. Then $G(t)$ equals

$$\begin{cases} t - t\log t - y - 3 & \text{if } t < -w - y, \\ 2t - t\log(2t + y + w) + w - 3 & \text{if } -w - y \leq t \leq \frac{-y-w+1}{2}, \\ 2t\exp[\frac{1}{2}(-2t - y - w + 1)] + w - 3 & \text{if } \frac{-y-w+1}{2} < t < -y/2, \\ \frac{2}{3}\exp[\frac{1}{2}(1-w)]\times & \\ \times\{2 + (t - y - 2)(-2t - y + 1)^{1/2}\} + w - 3 & \text{if } -y/2 \leq t \leq -y, \\ \frac{2}{3}\exp[\frac{1}{2}(1-w)]\{2 + (3t + y - 2)(y+1)^{1/2}\} + w - 3 & \text{if } -y < t \leq 1. \end{cases}$$

It is evident that G is concave in the interiors of the intervals above; morevore, one easily checks that the one-sided derivatives of G match at the endpoints of these intervals, and this yields the concavity on whole $[0,1]$. The monotonicity of G follows from the fact that $G'(1-) < 0$ for $y > -1$, and $G'(1-) = 0$ for remaining values of y.

(ii) As previously, we may assume that $w \geq 1$. Clearly, $J(t) = J(-t)$ for all t. Furthermore, it is easy to check that J is convex on $[-w, w]$, linear on the halflines $(-\infty, -w]$ and $[w, \infty)$, and $J'(w-) = J'(w+) = 1$, $J'(-w-) = J'(-w+) = -1$. This proves the claim.

(iii) By (i), it suffices to establish the majorization only on the set $\{(x,y,w) : x \in \{0,1\} \text{ or } |y| = w\}$. Moreover, since $u(x,y,w) = u(x,-y,w)$, we may assume that $y \geq 0$. Now if $x = 0$, then both sides are equal. If $y = w$, then the inequality reduces to $2x - x\log x \geq 0$, which is trivial. Finally, if $x = 1$, we use (ii) to obtain

$$u(1,y,w) \geq u(1,0,w) = \frac{4}{3}\exp\left[\frac{1}{2}(1-w)\right] + w - 3 \geq w - 3. \quad \square$$

Let $U : [0,\infty) \times \mathbb{R} \times (0,\infty) \times (0,\infty) \to \mathbb{R}$ be defined by

$$U(x,y,z,w) = (x \vee z)u\left(\frac{x}{x \vee z}, \frac{y}{x \vee z}, \frac{|y| \vee w}{x \vee z}\right). \tag{3.1}$$

Lemma 3.2. *The function U belongs to the class $\mathcal{U}(3)$.*

Proof. The condition (2.3) follows directly from the definition. To check (2.4), note that by Lemma 3.1 (ii),

$$U(x,y,x,|y|) = xu(1,y/x,|y|/x) \le xu(1,1,1) = 0.$$

The majorization (2.5) is an immediate consequence of lemma 3.1 (iii). It remains to establish (2.6). We will prove a slightly stronger statement: for x, y, z, w, ε as in (2.6), there is a linear function Ψ such that $\Phi(t) := U(x+t, y+\varepsilon t, z, w) \le \Psi(t)$ for all $t \ge -x$ and $\Psi(0) = \Phi(0)$. This will follow from the two conditions below:

$$\Phi \text{ is concave and nondecreasing on } [-x, -x+z], \tag{3.2}$$

$$\Phi \text{ is convex and nonincreasing on } [-x+z, \infty). \tag{3.3}$$

To prove these properties, observe that we may assume that $\varepsilon = 1$, since U is symmetric with respect to the variable y. Now the condition (3.2) is an immediate consequence of Lemma 3.1 (i). To prove (3.3), let us first use the second part of that lemma. Suppose that t_1, $t_2 \ge -x+z$ and α_1, $\alpha_2 \in (0,1)$ satisfy $\alpha_1 + \alpha_2 = 1$. For $\alpha_i' = \alpha_i(x+t_i)/(\alpha_1(x+t_1) + \alpha_2(x+t_2))$,

$$\begin{aligned}
\alpha_1\Phi(t_1) + \alpha_2\Phi(t_2) &= (\alpha_1(x+t_1) + \alpha_2(x+t_2))\left[\alpha_1' J\left(\frac{y+t_1}{x+t_1}\right) + \alpha_2' J\left(\frac{y+t_2}{x+t_2}\right)\right] \\
&\ge (\alpha_1(x+t_1) + \alpha_2(x+t_2)) J\left(\alpha_1' \frac{y+t_1}{x+t_1} + \alpha_2' \frac{y+t_2}{x+t_2}\right) \\
&= (\alpha_1(x+t_1) + \alpha_2(x+t_2)) J\left(\frac{y+\alpha_1 t_1 + \alpha_2 t_2}{\alpha_1(x+t_1) + \alpha_2(x+t_2)}\right) \\
&= \Phi(\alpha_1 t_1 + \alpha_2 t_2).
\end{aligned}$$

Thus Φ is convex; therefore, for any $t \ge -x+z$,

$$\Phi'(t+) \le \lim_{s\to\infty} \frac{\Phi(s)}{s} = \lim_{s\to\infty} \frac{(x+s)u(1, \frac{y+s}{x+s}, \frac{y+s}{x+s})}{s} = u(1,1,1) = 0,$$

which completes the proof. $\square$

3.2. **On the search of a suitable function.** Now we will present a reasoning which leads to the optimal constant $\beta = 3$ and produces the special function U. We start with the function U_0 defined by (2.8). For any x, y, z, w we have

$$U_0(x,y,z,w) = U_0(x,-y,z,w), \tag{3.4}$$

$$U_0(\lambda x, \lambda y, \lambda z, \lambda w) = \lambda U_0(x,y,z,w), \qquad \lambda > 0, \tag{3.5}$$

and

$$U_0(0,y,z,w) = w - \beta z. \tag{3.6}$$

The first two properties follow immediately from the definition of U_0 and the fact that $(f,g) \in \mathcal{M}(x,y)$ implies $(f,-g) \in \mathcal{M}(x,-y)$ and $(\lambda f, \lambda g) \in \mathcal{M}(\lambda x, \lambda y)$; the third holds since $\mathcal{M}(0,y)$ contains only the constant pair $(0,y)$. We shall search for U in the class of functions satisfying these three conditions. In fact, we shall find $u : (x,y,w) \mapsto U(x,y,1,w)$ (for $x \in [0,1]$ and $y \in [0,w]$) and recover U using (3.1) and (3.4). It is convenient to split the remaining part into two steps.

Step 1. Assumptions. We impose the following conditions on u:

(A1) u is of class C^1.
(A2) $\lim_{y\downarrow 0} u_y(x,y,w) = 0$ for $x \in [0,1]$ and $w > 0$.

(A3) For any $w > 0$, the function $u(\cdot, \cdot, w)$ is linear along the line segments of slope 1 contained in $[0, 1] \times [0, w]$.

(A4) We have $u_w(x, w, w) = 0$ for any $x \in (0, 1]$ and $w > 0$.

(A5) We have $u_x(1, y, w) = u_y(1, y, w)$.

These conditions come from the following reasoning. It is natural to expect that the special function U will be smooth: this gives rise to (A1) and then (3.4) enforces (A2). Next, (2.6) implies that for any $w > 0$, $u(\cdot, \cdot, w)$ is concave along any line segment of slope ± 1 contained in $[0, 1] \times [0, w]$. Since u corresponds to a sharp estimate and thus is extremal in some sense, it seems reasonable to conjecture that for any $(x, y) \in (0, 1) \times (0, w)$, $u(\cdot, \cdot, w)$ is (at least, locally) linear along a line segment of slope 1 or -1 passing through (x, y). Some experiments and the formulas for the special functions appearing in [3] and [5] lead to (A3). Next, apply (2.6) to $x \in (0, 1)$, $y = w$, $z = 1$, $\varepsilon = 1$ and $t_1 = -t_2 = t > 0$ close to 0 (then we are forced to take $\alpha = 1/2$). As the result,

$$u(x, w + t, w) + u(x, w - t, w) \leq 2u(x, w, w).$$

Since $u(x, w + t, w) = u(x, w + t, w + t)$ in virtue of (2.3), letting $t \to 0$ yields $u_w(x, w, w) \leq 0$. We assume equality and arrive at (A4). Finally, apply (2.6) to $x = z = 1$, $y \in [0, w)$, $\varepsilon = -1$ and t_1, t_2 such that $y - w < t_1 < 0 < t_2$. We get

$$\frac{t_2}{t_2 - t_1} u(1 + t_1, y - t_1, w) + \frac{-t_1}{t_2 - t_1} U(1 + t_2, y - t_2, 1, w) \leq u(1, y, w),$$

or, using (2.3), (3.4) and (3.5),

$$\frac{t_2}{t_2 - t_1} u(1 + t_1, y - t_1, w) + \frac{-t_1(1 + t_2)}{t_2 - t_1} u\left(1, \frac{t_2 - y}{1 + t_2}, \frac{w}{1 + t_2}\right) \leq u(1, y, w).$$

If $t_2 - y \geq w$, then we get, by (2.3),

$$\frac{t_2}{t_2 - t_1} u(1 + t_1, y - t_1, w) + \frac{-t_1(1 + t_2)}{t_2 - t_1} u\left(1, \frac{t_2 - y}{1 + t_2}, \frac{t_2 - y}{1 + t_2}\right) \leq u(1, y, w)$$

and letting $t_2 \to \infty$ gives $u(1 + t_1, y - t_1, w) - t_1 u(1, 1, 1) \leq u(1, y, w)$. This further implies $u_x(1, y, w) - u_y(1, y, w) \geq u(1, 1, 1)$. However, $u(1, 1, 1) = U(1, 1, 1, 1) \leq 0$ by (2.4); we assume that equalities hold in the last two estimates, and this is precisely (A5).

Step 2. Deriving the formula for u. Assume first that $w \geq 1$ and denote $A_w(x) = u(x, 1, w)$, $B_w(y) = u(1, y, w)$ for $x \in [0, 1]$ and $y \in [0, w]$. By (A3), we have

$$u(x, y, w) = \frac{y}{-x + y + 1} B_w(-x + y + 1) + \frac{-x + 1}{-x + y + 1} A_w(x - y) \tag{3.7}$$

when $0 \leq y \leq x \leq 1$. An application of (A2) gives

$$\frac{B_w(1 - x) - A_w(x)}{1 - x} = A'_w(x), \qquad x \in (0, 1), \tag{3.8}$$

while the use of (A5) yields

$$\frac{B_w(y) - A_w(1 - y)}{y} = 2B'_w(y), \qquad y \in (0, 1).$$

It is easy to solve this system of differential equations: substitute $y = 1 - x$ to get $2B'_w(1 - x) - A'_w(x) = 0$, so $-2B_w(1 - x) - A_w(x) = c_1$ for some constant c_1 depending only on w. Since $B_w(0) = A_w(1) = u(1, 0, 1)$, we obtain $c_1 = -3A_w(1)$.

Thus, $2B_w(1-x) = -A_w(x) + 3A_w(1)$, so plugging this into (3.8) and solving the differential equation gives

$$A_w(x) - A_w(1) = c_2(1-x)^{3/2} \quad \text{and} \quad B_w(y) - B_w(0) = -c_2 y^{3/2}/2 \tag{3.9}$$

for some constant c_2 depending only on w. Next, use (A3) to obtain

$$\begin{aligned} u(x,y,w) &= xB_w(-x+y+1) + (1-x)u(0,-x+y,w) \\ &= xB_w(-x+y+1) + (1-x)(w-\beta) \end{aligned} \tag{3.10}$$

for any x, y such that $x \leq y \leq x+w-1$ (the latter equality above is due to (3.6)). Applying (A5) gives the differential equation $2B_w'(y) = B_w(y) - w + \beta$ and solving it we obtain $B_w(y) = c_3 e^{y/2} + w - \beta$, for some c_3 depending only on w. By (A4), we get $c_3'(w) = -e^{-w/2}$ and hence

$$B_w(y) = 2\exp\left(\frac{y-w}{2}\right) + w - \beta. \tag{3.11}$$

Now put $w = y = 1$: then $B_1(1) = u(1,1,1) = 0$ (see the above explanation leading to (A5)) and hence we obtain $\beta = 3$. Next, note that B_w is differentiable at 1, in virtue of (A2). Comparing the left and right limits of B_w and B_w' at 1 (see (3.9) and (3.11)), we get

$$c_2 = -\frac{4}{3}\exp\left(\frac{1-w}{2}\right) \qquad \text{and} \qquad B_w(0) = \frac{4}{3}\exp\left(\frac{1-w}{2}\right) + w - 3.$$

This determines the functions A_w and B_w, and (3.7) and (3.10) give u on $y \leq x + w - 1$. The formula for u on $\{(x,y) : y > x+w-1\}$ is obtained in a similar manner, by the use of (A3) and (A4): we obtain $u(x,y,w) = 2x - x\log(x-y+w) + w - 3$. We leave the details to the reader. The final assumption is to put $u(x,y,w) = u(x,y,1)$ for $w < 1$. It is easy to check that the function we have just constructed leads to that from the previous subsection.

4. Sharpness of (1.4)

Obviously, it suffices to prove the optimality of the constant in the discrete-time case. One could try to provide an appropriate example, but this would lead to quite involved calculations; to avoid them, we take a different approach and use Theorem 2.3 instead.

Suppose that the constant $\beta > 0$ is such that the inequality (2.1) holds for any martingale f and its ± 1 transform g. Let U_0 be the function guaranteed by Theorem 2.3. The function U_0 satisfies (2.3)–(2.6) and (3.5). Furthermore, it enjoys the following property.

Lemma 4.1. *For any $x \geq 0$, $z > 0$ and y_1, $y_2 \in \mathbb{R}$, w_1, $w_2 > 0$ we have*

$$|U_0(x,y_1,z,w_1) - U_0(x,y_2,z,w_2)| \leq \max\{|y_1-y_2|, |w_1-w_2|\}.$$

Proof. By the triangle inequality, for any numbers a_0, a_1, a_2, $\ldots$, a_n,

$$|y_1+a_0| \vee |y_1+a_1| \vee \ldots \vee |y_1+a_n| \vee w_1 - |y_2+a_0| \vee |y_2+a_1| \vee \ldots \vee |y_2+a_n| \vee w_2$$
$$\leq \max\{|y_1-y_2|, |w_1-w_2|\}.$$

In consequence, if f, g are martingales such that f starts from x, g starts from 0 and $dg_n = \pm df_n$ for all $n \geq 1$, then

$$\begin{aligned}
&\mathbb{E}((y_1+g)_n^* \vee w_1 - \beta f_n^*) - U_0(x, y_2, z, w_2)\\
&\leq \mathbb{E}\big[((y_1+g)_n^* \vee w_1 - \beta f_n^* \vee z) - ((y_2+g)_n^* \vee w_2 - \beta f_n^* \vee z)\big]\\
&\leq \max\{|y_1-y_2|, |w_1-w_2|\}.
\end{aligned}$$

It suffices to take supremum over f, g and n to obtain

$$U_0(x, y_1, z, w_1) - U_0(x, y_2, z, w_2) \leq \max\{|y_1-y_2|, |w_1-w_2|\},$$

and the claim follows by symmetry. □

Lemma 4.2. *Let $w > 0$ and $\delta \in (0,1)$. Then*

$$U_0(1, w, 1, w) \geq U_0(1-\delta, w+\delta, 1, w+\delta) + \delta U_0(1,1,1,1) \tag{4.1}$$

and

$$U_0(1-\delta, w+\delta, 1, w+\delta) \geq (1-\delta)U_0(1, w+2\delta, 1, w+2\delta) + \delta(w+\delta-\beta). \tag{4.2}$$

Proof. Apply (2.6) to $(x,y,z,w) := (1,w,1,w)$, $\varepsilon = -1$ and $t_1 = -\delta$, $t_2 > 0$ (the numbers α_1, α_2 are uniquely determined by t_1 and t_2). We obtain

$$\frac{t_2}{t_2+\delta}U_0(1-\delta, w+\delta, 1, w) + \frac{\delta}{t_2+\delta}U_0(1+t_2, w-t_2, 1, w) \leq U_0(1, w, 1, w),$$

or, using (2.3) and (3.5),

$$\frac{t_2}{t_2+\delta}U_0(1-\delta, w+\delta, 1, w) + \frac{\delta(1+t_2)}{t_2+\delta}U_0\left(1, \frac{w-t_2}{1+t_2}, 1, \frac{w}{1+t_2}\right) \leq U_0(1, w, 1, w).$$

If we let $t_2 \to \infty$ and use the previous lemma, together with the equality $U_0(1,-1,1,1) = U_0(1,1,1,1)$, we get (4.1). To obtain (4.2), simply apply (2.6) to $(x,y,z,w) := (1-\delta, w+\delta, 1, w+\delta)$ and $t_1 = \delta - 1$, $t_2 = \delta$. □

Now we are ready to show that the constant 3 is the best possible.

Sharpness of (1.4). Combining (4.1) and (4.2), we get

$$U_0(1, w, 1, w) \geq (1-\delta)U_0(1, w+2\delta, 1, w+2\delta) + \delta U_0(1,1,1,1) + \delta(w+\delta-\beta).$$

Substituting $F(w) = U_0(1,w,1,w) - U_0(1,1,1,1) - (w - \beta + 2)$, we rewrite the above inequality in the form $F(w) \geq (1-\delta)F(w+2\delta) - \delta^2$. This, by induction, yields

$$F(w) \geq (1-\delta)^n F(w+2n\delta) - n\delta^2.$$

Now fix $z > 1$ and take $w = 1$, $\delta = (z-1)/(2n)$ (here n must be sufficiently large so that $\delta < 1$). Letting $n \to \infty$ gives

$$\beta - 3 = F(1) \geq F(z)\exp\left(\frac{1-z}{2}\right).$$

However, by Lemma 4.1, F has at most linear growth; thus, letting $z \to \infty$, we obtain $\beta - 3 \geq 0$. This completes the proof. □

References

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

[2] D. L. Burkholder, *Explorations in martingale theory and its applications*, Ecole d'Eté de Probabilités de Saint Flour XIX-1989, Lecture Notes in Mathematics **1464** (1991), 1–66.

[3] D. L. Burkholder, *Sharp norm comparison of martingale maximal functions and stochastic integrals*, Proceedings of the Norbert Wiener Centenary Congress, 1994 (edited by V. Mandrekar and P. R. Masani), Proceedings of Symposia in Applied Mathematics **52** (1997), 343–358.

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

[5] A. Osękowski, *Sharp inequality for martingale maximal functions and stochastic integrals*, Illinois J. Math., to appear.

Department of Mathematics, Informatics and Mechanics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland

E-mail address: ados@mimuw.edu.pl