MAXIMAL INEQUALITIES FOR STOCHASTIC INTEGRALS

ADAM OSĘKOWSKI

ABSTRACT. Let X be a continuous-time martingale and H be a predictable process taking values in $[-1,1]$. Let Y denote the stochastic integral of H with respect to X. The paper contains the proof of sharp bound for one-sided maximal function of Y by the p-th moment of X. A discrete-time version of this inequality is also established.

1. Introduction

Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete 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 events of probability 0. Let $X = (X_t)_{t\geq 0}$ be an adapted real-valued right-continuous semimartingale with left limits. Let Y be the Itô integral of H with respect to X, that is,

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

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 sizes of X and Y^*. Let us describe some related results from the literature. In [3], Burkholder invented a method of proving maximal inequalities for martingales and used it to obtain the following sharp estimate.

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

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

where $\gamma = 2,536\ldots$ is the unique solution of the equation

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

The constant is the best possible.

Then it was shown by the author in [4], then 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 [5] contains the further study in this direction and, in particular, the proof of the following fact.

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

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

Partially supported by Foundation for Polish Science.

Theorem 1.2. *If X is a martingale and Y is as above, then*

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

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^*. Let

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

Here is our main result.

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

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

The constant C_p is the best possible. Furthermore, for $p \leq 1$ the inequality does not hold in general with any finite C_p.

In fact, the emphasis is put on the discrete-time version of the theorem above. Suppose $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, filtered by $(\mathcal{F}_n)_{n\geq 0}$. Let $f = (f_n)_{n\geq 0}$ be an adapted martingale and $g = (g_n)_{n\geq 0}$ be its transform by a predictable sequence $v = (v_n)_{n\geq 0}$ bounded in absolute value by 1. That is, we have

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

and 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 $f^* = \sup_k f_k$. Here is a discrete-time version of Theorem 1.3.

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*

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

For $p \leq 1$, the inequality does not hold in general with any finite C_p.

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 Section 4. The final part of the paper concerns the optimality of the constant C_p.

2. Some reductions and on the method of proof

Using approximation arguments of Bichteler [1], 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 $f_N = f_{N+1} = \ldots$ with probability 1. Clearly, if f and g are simple, then the almost sure limits f_∞ and g_∞ exist and are finite.

The key reduction is that it suffices to work with ± 1 transforms only. Recall Lemma A.1 from [2].

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 there exist martingales $F^j = (F^j_n)_{n\geq 0}$ and Borel measurable functions $\phi_j : [-1,1] \to \{-1,1\}$ such that, for $j \geq 1$ and $n \geq 0$,*

$$f_n = F^j_{2n+1} \quad \text{and} \quad g_n = \sum_{j=1}^{\infty} 2^{-j}\phi_j(v_0)G^j_{2n+1},$$

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

To see how the lemma works in our setting, suppose we have established (1.4) for ± 1 transforms. Lemma 2.1 gives us the processes F^j and the functions ϕ_j, $j \geq 1$. For any $j \geq 1$, conditionally on $\mathcal{F}_0$, the sequence $\phi_j(v_0)G^j$ is a ± 1 transform of F^j and hence we may write

$$\begin{aligned} ||g^*||_1 &\leq \left|\left| \sum_{j=1}^{\infty} 2^{-j} \sup_n \left(\phi_j(v_0) G^j_{2n+1} \right) \right|\right|_1 \\ &\leq \sum_{j=1}^{\infty} 2^{-j} \left|\left| \left(\phi_j(v_0) G^j \right)^* \right|\right|_1 \\ &\leq C_p \sum_{j=1}^{\infty} 2^{-j} ||F^j||_p \\ &= C_p ||f||_p, \end{aligned}$$

as needed.

Now we will describe Burkholder's method, introduced in [3], which will be used to establish our results. Let

$$\mathcal{A} = \{(x,y,z) \in \mathbb{R}^3 : y \leq z\},$$

fix a real number C and let $V : \mathcal{A} \to \mathbb{R}$ be a given function (not necessarily measurable). Suppose we want to show that

$$\mathbb{E}V(f_\infty, g_\infty, g^*_\infty) \leq C \tag{2.1}$$

for all simple martingales f, g such that g is a ± 1 transform of f. The tool to handle this problem is the class $\mathcal{U}(V,C)$, which consists of functions $U : \mathcal{A} \to \mathbb{R}$ satisfying the following three conditions.

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 $d \in \mathbb{R}$,

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

2° $U(x,y,z) \geq V(x,y,z)$ for all (x,y,z).

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

Sometimes it is convenient to replace 1° with the following equivalent condition (see [3]):

1°' For any $\varepsilon \in \{-1,1\}$, $(x,y,z) \in \mathcal{A}$ and any simple centered random variable T, we have

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

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

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

Proof. Take simple f, g such that g is a ±1 transform of f. The process $(U(f_n, g_n, g_n^*))$ is a supermartingale: the inequality $\mathbb{E}\big[U(f_n,g_n,g_n^*)|\mathcal{F}_{n-1}\big] \leq U(f_{n-1}, g_{n-1}, g_{n-1}^*)$, $n \geq 1$, follows from the conditional form of 1°', 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

$$\mathbb{E}V(f_\infty, g_\infty, g_\infty^*) \leq \mathbb{E}U(f_\infty, g_\infty, g_\infty^*) \leq \mathbb{E}U(f_0,g_0,g_0^*) \leq C. \quad \square$$

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

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

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

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 $v = (v_n)_{n\geq1}$ 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 $\mathcal{U}(V,C)$ is nonempty and U_0 is its least element.*

For the proof, one needs to modify slightly the argumentation used in [3] (see Theorem 2.2 there). This fact will be quite useful in the proof of the optimality of the constants C_p.

3. The proof of (1.4) for $1 < p \leq 2$

We start from defining a function $\gamma_p : [0,\infty) \to (-\infty, 0]$ by

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

Lemma 3.1. *The function γ_p is nonincreasing.*

Proof. The inequality $\gamma_p'(t) \leq 0$ is equivalent to

$$t^{2-p}\exp(-pt^{p-1}) - p(p-1)\int_t^\infty \exp(-ps^{p-1})ds \leq 0.$$

It suffices to note that the left-hand side tends to 0 as $t \to \infty$, and its derivative equals $(2-p)t^{1-p}\exp(-pt^{p-1}) \geq 0$. $\square$

Let $G_p : (-\infty, \gamma_p(0)] \to [0, \infty)$ denote the inverse to the function $t \mapsto \gamma_p(t) - t$, $t \geq 0$ (by the previous lemma, the function is invertible). We will need the following estimate.

Lemma 3.2. *We have $G_p G_p'' + (p-2)(G_p')^2 \leq 0$.*

Proof. An easy computation shows that

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

and

$$G_p''(x) = -(G_p'(x))^2 \left[\frac{p-2}{G_p(x)} + p(p-1)G_p(x)^{p-2} + \frac{1}{G_p(x) + x}\right].$$

Therefore the desired inequality reads, after some manipulations,

$$G_p(x)G_p''(x) + (p-2)(G_p'(x))^2 = -\frac{G_p(x)G_p'(x)(G_p'(x)+1)}{G_p(x) + x} \leq 0. \tag{3.2}$$

We have $G_p(x) \geq 0$. Furthermore, as proved in the previous lemma, we have $\gamma_p' \leq 0$. This implies $G_p'(x) \leq 0$, $G_p'(x) \geq -1$ and $G_p(x) + x \leq 0$, see the formula for G_p' above. This establishes (3.2). □

Now we are ready to introduce the main object in this section. Let $U_p : \mathcal{A} \to \mathbb{R}$ be given by

$$U_p(x, y, z) = -\frac{(y-z)^2 - x^2}{2\gamma_p(0)} - \frac{\gamma_p(0)}{2} + y$$

if $(x, y, z) \in D_1 = \{(x, y, z) \in \mathcal{A} : y - z - |x| \geq \gamma_p(0)\}$,

$$U_p(x, y, z) = z + (p-1)G_p(y - z - |x|)^p - p|x|G_p(y - z - |x|)^{p-1}$$

if $(x, y, z) \in D_2 = \{(x, y, z) \in \mathcal{A} : y - z - |x| < \gamma_p(0) \text{ and } |x| \geq G_p(y - z - |x|)\}$, and

$$U_p(x, y, z) = z - |x|^p,$$

for $(x, y, z) \in D_0 = \mathcal{A} \setminus (D_1 \cup D_2)$.

We will now study the properties of the function U_p. They will be needed to establish the validity of the conditions 1°, 2° and 3°.

Lemma 3.3. *(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, $h \in \mathbb{R}$,

$$U_p(x + \varepsilon t, y + t, (y + t) \vee y) \leq U_p(x, y, y) + \varepsilon U_{px}(x, y, y)t + t. \tag{3.3}$$

(iv) We have

$$U_p(x, y, z) \geq z - |x|^p \quad \text{for } (x, y, z) \in \mathcal{A}. \tag{3.4}$$

(v) We have

$$\sup U_p(x, y, y) = -\gamma_p(0), \tag{3.5}$$

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 virtue 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 $U_{px}(0,y,z)=0$ 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) belongs to D_1^o. If $(x,y,z)\in D_2^o$, then

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

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 3.2,

$$\begin{aligned} F''(0) &\leq 4p(p-1)G_p^{p-3}(x_0)[G_p(x_0)G_p'(x_0)(G_p'(x_0)+1) \\ &\qquad + (G_p(x_0)+x_0)((p-2)(G_p'(x_0))^2 + G_p(x_0)G_p''(x_0))] \\ &= 0, \end{aligned} \tag{3.6}$$

where in the latter passage we have used the equality from (3.2). 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 (3.6), with $x_0=-s$.

(iv) Again, it suffices to deal only with nonnegative x. On the set D_0 both sides of (3.4) are equal. To prove the majorization on D_2, let $\Phi(s) = -s^p$ for $s\geq 0$. Observe that

$$U_p(x,y,z) = z + \Phi(G_p(y-z-x)) + \Phi'(G_p(y-z-x))(x - G_p(y-z-x)),$$

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

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

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 $-\gamma_p(0)$), is the supremum we are searching for. □

Now we are ready to prove the inequality (1.4).

Proof of (1.4). Let f, g be as in the statement. Using standard approximation argument, we may assume that both martingales are simple and that $||f||_p > 0$. Let $V_p : \mathcal{A} \to \mathbb{R}$ be given by $V_p(x, y, z) = z - |x|^p$. We shall show that U_p belongs to the class $\mathcal{U}(V_p, -\gamma_p(0))$. By Lemma 3.3 (ii) and (iii), U_p has the property 1°. The parts (iv) and (v) of this lemma imply the validity of the conditions 2° and 3°, respectively. Thus, applying Theorem 2.2 to the martingales f/λ, g/λ, where $\lambda > 0$ is fixed, yields

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

Now the choice

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

gives (1.4). □

Sharpness. As shown by Peskir [6], the following 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 of B satisfying $\tau \in L^{p/2}$. In consequence, the estimate (1.4) is also sharp, even if $X = Y$.

It remains to show that the inequality (1.4) fails to hold for $p \leq 1$. This is due to the fact that $C_p \to \infty$ as $p \to 1+$. Indeed, if the estimate was valid for some $p \leq 1$ and $C_p < \infty$, then for any $p' > 1$ we would have $||g^*||_1 \leq C_p||f||_{p'}$; this cannot be true if p' is sufficiently close to 1. □

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

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

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 3.2.

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 $G_p'(x) \leq 1$, $x - G_p(x) < 0$. □

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

$$U_p(x,y,z) = z + H(x, y - z + (p-1)p^{-1/(p-1)})$$

if $(x,y,z) \in D_1 = \{(x,y,z) \in \mathcal{A} : y - z \geq \gamma_p(x),\ x + y - z \leq 0\}$,

$$U_p(x,y,z) = z + (p-1)G_p(|x| + y - z)^p - p|x|G_p(|x| + y - z)^{p-1}$$

if $(x,y,z) \in D_2 = \{(x,y,z) \in \mathcal{A} : y - z \geq \gamma_p(x),\ x + y - z > 0\}$, and

$$U_p(x,y,z) = z - |x|^p$$

if $(x,y,z) \in D_0 = \mathcal{A} \setminus (D_1 \cup D_2)$.

Here is the analogue of Lemma 3.3.

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, $h \in \mathbb{R}$,

$$U_p(x + \varepsilon t, y + t, (y+t) \vee y) \leq U_p(x,y,y) + \varepsilon U_{px}(x,y,y)t + t. \tag{4.2}$$

(iv) We have

$$U_p(x,y,z) \geq z - |x|^p \quad \textit{for } (x,y,z) \in \mathcal{A}. \tag{4.3}$$

(v) We have

$$M_p = \sup U_p(x,y,y) = \frac{p-1}{p^{p/(p-1)}}\left[2^{p/(p-1)} - \frac{p}{p-1}\int_1^2 s^{1/(p-1)}e^{s-2}ds\right], \tag{4.4}$$

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

Proof. (i) Straightforward.

(ii) We proceed as in the proof of part (ii) in Lemma 3.3 and check $F''(0) \leq 0$ 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 $F''(0) = -p(p-1)x^{p-2} \leq 0$ 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 $F''(0) \leq 0$; 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 page 17 in [2]. If $(x, y, z) \in D_2^o$, then

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

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

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

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

(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 3.3, we see that it suffices to show that the function $s \mapsto U_p(s, 0, 0)$, $s > 0$, is concave. However, its second derivative at s equals

$$-p(p-1)G_p^{p-2}(s)G_p'(s) \leq 0 \tag{4.5}$$

and we are done.

(iv) The majorization can be proved in the same manner as in the Lemma 3.3, using the concave function $\Phi(s) = -s^p$, $s \geq 0$. The details are left to the reader.

(v) Observe that

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

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

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

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

$$\begin{aligned} t_0 &= \gamma_p((2/p)^{1/(p-1)}) + (2/p)^{1/(p-1)} \\ &= p(p-1)e^{-2}\int_{p^{-1/(p-1)}}^{(p/2)^{-1/(p-1)}} s^{p-1}\exp(ps^{p-1})ds \\ &= p^{-1/(p-1)}\int_1^2 s^{1/(p-1)}e^{s-2}ds \end{aligned} \tag{4.6}$$

and, as one easily checks, the maximum is equal to M_p. This completes the proof. □

Proof of the inequality (1.4). It suffices to establish the estimate for finite p, as $\lim_{p\to\infty} C_p = C_\infty$. We proceed as in the proof of (1.4). By Lemma 4.3, the function U_p belongs to the class $U_p \in \mathcal{U}(V_p, M_p)$, where $V_p(x,y,z) = z - |x|^p$. Therefore, by Theorem 2.2, for any $\lambda > 0$,

$$||g^*||_1 \leq \lambda^{1-p}||f||_p^p + \lambda M_p,$$

and taking $\lambda = (p-1)^{1/p} M_p^{-1/p}||f||_p$ gives (1.4). □

5. Sharpness

The case $p < \infty$. We have, by Young's inequality,

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

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

$$||g^*||_1 \leq ||f||_p^p + C \tag{5.1}$$

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 (5.1) is valid, equals M_p.

Suppose then, that (5.1) 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°, 2° and 3°. Observe that for any $(x,y,z) \in \mathcal{A}$ and $t \in \mathbb{R}$,

$$U_0(x,y,z) = t + U_0(x, y-t, z-t). \tag{5.2}$$

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

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

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

$$U_0(0,y,y) \geq y + (p-1)p^{-p/(p-1)} = U_p(0,y,y). \tag{5.3}$$

In view of (5.2), 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 (5.2), $U_0(-\delta,\delta,\delta) = \delta + U_0(-\delta,0,0)$. Furthermore, by 2°, $U_0(d,-d,0) \geq -d^p$, so the above estimate yields

$$U_0(0,0,0) \geq \frac{d}{d+\delta}(\delta + U_0(-\delta,0,0)) - \frac{\delta}{d+\delta}|d|^p. \tag{5.4}$$

Similarly, one uses the property 1°' and then 2°, and gets

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

Combining this with (5.4), subtracting $U_0(0,0,0)$ from both sides of the obtained estimate, dividing throughout 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.

In consequence, 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

$$U_p(0,y,y) \leq V_p(f_\infty^{\kappa,y}, g_\infty^{\kappa,y}, (g_\infty^{\kappa,y})^*) + \kappa. \tag{5.5}$$

Step 2. Let N be a positive integer and let $\delta = t_0/N$, where t_0 is given by (4.6). We will need the following auxiliary fact.

Lemma 5.1. *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 taking values in $[\gamma_p(G_p(x)), \delta]$, then*

$$\mathbb{E}U_p(x - T, y + T, (y + T) \vee y) \leq U_p(x, y, y) + R\delta^2. \tag{5.6}$$

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 (4.5). The inequality (5.6) follows immediately from the two above 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 compliment of this set, $df_{n+1} \equiv 0$. Let g be a ± 1 transform of f, given by $g_0 = f_0$ and $dg_n = -df_n$, $n = 1, 2, \ldots, N$. The key fact about the pair (f, g) is that

$$\mathbb{E}U_p(f_n, g_n, g_n^*) \leq \mathbb{E}U_p(f_{n+1}, g_{n+1}, g_{n+1}^*) + R\delta^2, \quad n = 0, 1, 2, \ldots, N - 1. \tag{5.7}$$

This is an immediate consequence of Lemma 5.1 (applied conditionally with respect to $\mathcal{F}_n$) and the fact that $U_p(f_n, g_n, g_n^*) \neq U_p(f_{n+1}, g_{n+1}, g_{n+1}^*)$ 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 $f_N \neq 0$, then $U_p(f_N, g_N, g_N^*) = V_p(f_N, g_N, g_N^*)$. Indeed, $f_N \neq 0$ implies $df_n > 0$ for some $n \geq 1$ and then, by the 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

$$\begin{aligned} M_p &= U_p(t_0, t_0, t_0) \\ &\leq \mathbb{E}U_p(f_N, g_N, g_N^*) + RN\delta^2 \\ &= \mathbb{E}V_p(f_N, g_N, g_N^*)1_{\{f_N \neq 0\}} + U_p(0, 2t_0, 2t_0)\mathbb{P}(f_N = 0) + RN\delta^2, \end{aligned} \tag{5.8}$$

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 $f_N = f_{N+1} = f_{N+2} = \ldots$ and $g_N = g_{N+1} = g_{N+2} = \ldots$ on $\{f_N \neq 0\}$, while on $\{f_N = 0\}$, let the conditional distribution of $(f_n, g_n)_{n \geq N}$ 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 (5.5) and (5.8), we have

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

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 (5.1). □

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

Theorem 5.2. *Let $U_0 : \{(x,y,z) : |x| \leq 1,\ y \leq z\} \to \mathbb{R}$ 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 $||f||_\infty \leq 1$. Then U_0 enjoys the following properties.

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

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

2° $U_0(x,y,z) \geq z$ for all (x,y,z) from the domain of U_0.

3° $U_0(x,y,y) \leq C_\infty$ for all x, y such that $|x| = |y| \in [-1,1]$.

For the proof, modify the argumentation from [3]. Note that the function U_0 satisfies (5.2) (with obvious restriction to x lying in $[-1,1]$).

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

$$U_0(0,0,0) \geq 1. \tag{5.9}$$

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

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

We have $U_0(-1,-1,0) \geq 0$ by 2°, and $U_0(\delta,\delta,\delta) = \delta + U(\delta,0,0)$ by (5.2). Thus we have

$$U_0(0,0,0) \geq \frac{\delta + U_0(\delta,0,0)}{1+\delta}. \tag{5.10}$$

Similarly, using 1° and then 2°,

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

Plug this into (5.10), subtract $U_0(0,0,0)$ from both sides, divide throughout by δ and let $\delta \to 0$. As a result, one gets (5.9).

Now fix a positive integer N and set $\delta = (1-e^{-1})/N$. For any $k = 1, 2, \ldots, N$, we have, by 1°, 2° and (5.2),

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

or, equivalently,

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

It follows by induction that

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

Letting $N \to \infty$ and using (5.9), we arrive at

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

and hence, by (5.2),

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

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

References

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

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

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

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

[5] A. Osękowski, *Sharp maximal inequality for martingales and stochastic integrals*, Electr. Comm. in Probab. **14** (2009), 17–30.

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

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

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