Czechoslovak Mathematical Journal

Adam Osękowski

A sharp maximal inequality for continuous martingales and their differential subordinates

Czechoslovak Mathematical Journal, Vol. 63 (2013), No. 4, 1001–1018

Persistent URL: `http://dml.cz/dmlcz/143613`

Czechoslovak Mathematical Journal, 63 (138) (2013), 1001–1018

A SHARP MAXIMAL INEQUALITY FOR CONTINUOUS MARTINGALES AND THEIR DIFFERENTIAL SUBORDINATES

Adam Osękowski, Warszawa

(Received August 1, 2012)

Abstract. Assume that X, Y are continuous-path martingales taking values in $\mathbb{R}^\nu$, $\nu \geqslant 1$, such that Y is differentially subordinate to X. The paper contains the proof of the maximal inequality

$$\|\sup_{t\geqslant 0}|Y_t|\|_1 \leqslant 2\|\sup_{t\geqslant 0}|X_t|\|_1.$$

The constant 2 is shown to be the best possible, even in the one-dimensional setting of stochastic integrals with respect to a standard Brownian motion. The proof uses Burkholder's method and rests on the construction of an appropriate special function.

Keywords: martingale; stochastic integral; maximal inequality; differential subordination

MSC 2010: 60G44, 60G46

1. Introduction

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a complete probability space, filtered by a nondecreasing right-continuous family $(\mathscr{F}_t)_{t\geqslant 0}$ of sub-σ-fields of $\mathscr{F}$. In addition, we assume that $\mathscr{F}_0$ contains all events of probability 0. Let X, Y be two adapted martingales, taking values in $\mathbb{R}^\nu$ (where ν is a fixed positive integer) with norm $|\cdot|$ and scalar product $\langle\cdot,\cdot\rangle$. As usual, we assume that paths of the processes are right-continuous, with limits from the left. The symbol $[X, X]$ will stand for the quadratic covariance process of X, given by $[X, X] = \sum_{n=1}^{\nu} [X^n, X^n]$. Here X^n denotes the n-th coordinate of X and $[X^n, X^n]$ is the usual square bracket of the real-valued martingale X^n (see Dellacherie and Meyer [7] for details). In what follows, $X^* = \sup_{t\geqslant 0}|X_t|$ will denote the maximal function of X; we also use the notation $X_t^* = \sup_{0\leqslant s\leqslant t}|X_s|$.

The research was supported by the NCN grant DEC-2012/05/B/ST1/00412.

Throughout the paper we assume that the process Y is *differentially subordinate* to X. This concept was originally introduced by Burkholder [3] in the discrete-time case: a martingale $g = (g_n)_{n\geqslant 0}$ is differentially subordinate to $f = (f_n)_{n\geqslant 0}$, if for any $n \geqslant 0$ we have $|dg_n| \leqslant |df_n|$. Here $df = (df_n)_{n\geqslant 0}$, $dg = (dg_n)_{n\geqslant 0}$ are the difference sequences of f and g, respectively, given by the equations

$$f_n = \sum_{k=0}^{n} df_k \quad \text{and} \quad g_n = \sum_{k=0}^{n} dg_k, \quad n = 0, 1, 2, \ldots.$$

This domination was extended to the continuous-time setting by Bañuelos and Wang [2] and Wang [16]. We say that Y is *differentially subordinate* to X, if the process $([X, X]_t - [Y, Y]_t)_{t\geqslant 0}$ is nondecreasing and nonnegative as a function of t. If we treat given discrete-time martingales f, g as continuous-time processes (via $X_t = f_{\lfloor t\rfloor}$ and $Y_t = g_{\lfloor t\rfloor}$, $t \geqslant 0$), we see this domination is consistent with the original definition of Burkholder.

To illustrate this notion, consider the following example. Suppose that X is an $\mathbb{R}^\nu$-valued martingale, H is a predictable process taking values in the interval $[-1, 1]$ and let Y be given as the stochastic integral $Y_t = H_0 X_0 + \int_{0+}^t H_s \mathrm{d}X_s$, $t \geqslant 0$. Then Y is differentially subordinate to X: we have

$$[X, X]_t - [Y, Y]_t = (1 - H_0^2)|X_0|^2 + \int_{0+}^t (1 - H_s^2)\mathrm{d}[X, X]_s.$$

Another example for stochastic integrals, which plays an important role in applications (see e.g. [1], [2], [8]), is the following. Suppose that B is a Brownian motion in $\mathbb{R}^\nu$ and H, K are predictable processes taking values in the matrices of dimensions $m \times \nu$ and $n \times \nu$, respectively. For any $t \geqslant 0$, define

$$X_t = \int_{0+}^t H_s \cdot \mathrm{d}B_s \quad \text{and} \quad Y_t = \int_{0+}^t K_s \cdot \mathrm{d}B_s.$$

If the Hilbert-Schmidt norms of H and K satisfy $\|K_t\|_{HS} \leqslant \|H_t\|_{HS}$ for all $t > 0$, then Y is differentially subordinate to X: this follows from the identity

$$[X, X]_t - [Y, Y]_t = \int_{0+}^t (\|H_s\|_{HS}^2 - \|K_s\|_{HS}^2)\,\mathrm{d}s.$$

The differential subordination implies many interesting inequalities comparing the sizes of X and Y. A celebrated result of Burkholder gives the following information on the L^p-norms $\|X\|_p = \sup_{t\geqslant 0} \|X_t\|_p$, $\|Y\|_p = \sup_{t\geqslant 0} \|Y_t\|_p$ (see [3], [4], [5] and [16]).

Theorem 1.1. *Suppose that X, Y are Hilbert-space-valued martingales such that Y is differentially subordinate to X. Then*

$$\|Y\|_p \leqslant (p^* - 1)\|X\|_p, \quad 1 < p < \infty, \tag{1.1}$$

where $p^* = \max\{p, p/(p-1)\}$. The constant is the best possible, even if $\mathscr{H} = \mathbb{R}$.

For $p = 1$, the above moment inequality does not hold with any finite constant, but we have the corresponding weak-type and logarithmic estimates; see [3], [10] and [15]. The bounds above have found numerous applications in many areas of mathematics (consult, for instance, [1], [2], [8] and [9]). There is a general method, invented by Burkholder, which enables one not only to establish various estimates of this type, but is also very efficient in determining the optimal constants in such inequalities. The technique rests on the construction of an appropriate special function (usually, quite complicated) and a careful use of its properties. See the survey [5] for the detailed description of the technique in the discrete-time setting and consult Wang [16] for the modification in the continuous case.

There is another, very interesting direction in which the results can be extended. In [6] Burkholder modified his technique so that it could be used to study maximal inequalities for stochastic integrals. As an application, he proved the following result, which can be regarded as a version of (1.1) for $p = 1$.

Theorem 1.2. *Suppose that X is a real-valued martingale and Y is the stochastic integral, with respect to X, of some predictable real-valued process H taking values in $[-1, 1]$. Then we have the sharp estimate*

$$\|Y\|_1 \leqslant \gamma \|X^*\|_1, \tag{1.2}$$

where $\gamma = 2.536\ldots$ is the unique positive number satisfying $\gamma = 3 - \exp \frac{1-\gamma}{2}$.

This result was strengthened by the author to the case in which the first moment of Y is replaced by the first moment of its maximal function.

Theorem 1.3. *Under the assumptions of the above theorem, we have the sharp inequality*

$$\|Y^*\|_1 \leqslant 3.4351\ldots\|X^*\|_1.$$

The precise description of the above constant involves an analysis of a complicated system of ODE's. For the details, we refer the reader to [11].

We would like to point out here that both the theorems above are valid for real-valued martingales X, Y such that Y is differentially subordinate to X. However, this is no longer true when X, Y are assumed to take values in $\mathbb{R}^2$ (cf. [13]).

We will be interested in the sharp version of Theorem 1.3 for continuous-path martingales. In general, the best constants in non-maximal inequalities for differentially subordinated martingales do not change when we pass to this more restrictive setting. See, e.g., Section 15 in [3] for the justification of this phenomenon. However, if we study the maximal estimates, the best constants may be different: for example, the passage to continuous-time martingales reduces the constant γ in (1.2) to $\sqrt{2}$ (see [12]).

Our main result can be stated as follows.

Theorem 1.4. *Suppose that X, Y are continuous-path $\mathbb{R}^\nu$-valued martingales such that Y is differentially subordinate to X. Then*

$$\|Y^*\|_1 \leqslant 2\|X^*\|_1 \tag{1.3}$$

and the constant is the best possible.

In fact, the constant 2 is optimal even in the one-dimensional setting of stochastic integrals. More precisely, we will prove that for any $\kappa < 2$ there is a stopped Brownian motion X in $\mathbb{R}$ and a predictable process H with values in $\{-1, 1\}$ such that the stochastic integral

$$Y_t = \int_{0+}^t H_s \mathrm{d}X_s, \quad t \geqslant 0,$$

satisfies $\|Y^*\|_1 > \kappa\|X^*\|_1$.

The paper is organized as follows. Our approach exploits Burkholder's method; in the next section we introduce the special function corresponding to (1.3), and in Section 3 we complete the proof of this estimate. Section 4 concerns the optimality of the constant 2, and in the final part of the paper we sketch some steps which lead to the discovery of the special function.

2. A special function

A key role in the proof of Theorem 1.4 is played by a special function U defined on the set

$$D = \{(x, y, z, w) \in \mathbb{R}^\nu \times \mathbb{R}^\nu \times (0, \infty) \times (0, \infty) \colon |x| \leqslant z\}.$$

To introduce this function, we distinguish the subdomains $D_1 - D_4$ of D, given by

$$\begin{aligned}
D_1 &= \{(x, y, z, w) \in D \colon w \leqslant z\},\\
D_2 &= \{(x, y, z, w) \in D \colon |x| + |y| < z < w\},\\
D_3 &= \{(x, y, z, w) \in D \colon z \leqslant |x| + |y| < w\},\\
D_4 &= \{(x, y, z, w) \in D \colon z < w \leqslant |x| + |y|\}.
\end{aligned}$$

Now, for $(x,y,z,w) \in D$, we define $U(x,y,z,w)$ by

$$\begin{cases} \dfrac{|y|^2-|x|^2-z^2}{2z} & \text{if } \ (x,y,z,w) \in D_1, \\ \dfrac{|y|^2-|x|^2+z^2}{2z} \cdot \mathrm{e}^{1-w/z} + w - 2z & \text{if } \ (x,y,z,w) \in D_2, \\ (z-|x|)\exp\Big(\dfrac{|x|+|y|-w}{z}\Big) + w - 2z & \text{if } \ (x,y,z,w) \in D_3, \\ \dfrac{(|y|-w+z)^2-|x|^2-3z^2}{2z} + w & \text{if } \ (x,y,z,w) \in D_4. \end{cases}$$

Lemma 2.1. *The function U enjoys the following properties.*

(i) *It is continuous on D. Furthermore, for fixed w and z, the function $U(\cdot,\cdot,w,z)\colon (x,y) \mapsto U(x,y,z,w)$ is of class C^1 on the set $\{(x,y) \in \mathbb{R}^\nu \times \mathbb{R}^\nu\colon\ 0 < |x| < z\}$.*

(ii) *For any $(x,y,z,w) \in D$ we have the majorization*

$$U(x,y,z,w) \geqslant w - 2z. \tag{2.1}$$

(iii) *For any $w,z > 0$ satisfying $w \neq z$ and any $x,y \in \mathbb{R}^\nu$ such that $|x| < z$ we have*

$$U_z(x,y,z,w) \leqslant 0. \tag{2.2}$$

P r o o f. (i) This is straightforward and reduced to a tedious verification that the appropriate limits of U and its partial derivatives match at the common boundaries of D_1, D_2, D_3 and D_4. We leave the details to the reader.

(ii) If $(x,y,z,w) \in D_1$, then we use the bounds $|y| \geqslant 0$ and $|x| \leqslant z$ to obtain

$$\frac{|y|^2-|x|^2-z^2}{2z} \geqslant -z \geqslant w - 2z.$$

If (x,y,z,w) lies in D_2, then the majorization follows immediately from the obvious estimate $|y|^2-|x|^2+z^2 \geqslant 0$. If $(x,y,z,w) \in D_3$, then (2.1) is trivial. Finally, for $(x,y,z,w) \in D_4$ it suffices to apply the inequalities $(|y|-w+z)^2 \geqslant 0$ and $|x| \leqslant z$ to get the assertion.

(iii) It is easy to check that the assumptions on x,y,z and w imply the existence of the partial derivative U_z. If (x,y,z,w) belongs to D_1, then

$$U_z(x,y,z,w) = \frac{|x|^2-|y|^2-z^2}{2z^2}$$

is nonpositive. When $(x,y,z,w)\in D_2$, then we derive that

$$\begin{aligned}U_z(x,y,z,w)&=\Big[\frac{|x|^2-|y|^2}{2z^2}\Big(1-\frac{w}{z}\Big)+\frac12\Big(1+\frac{w}{z}\Big)\Big]\cdot \mathrm{e}^{1-w/z}-2\\&\leqslant\Big[\frac{-z^2}{2z^2}\Big(1-\frac{w}{z}\Big)+\frac12\Big(1+\frac{w}{z}\Big)\Big]\cdot \mathrm{e}^{1-w/z}-2\\&=\frac{w}{z}\mathrm{e}^{1-w/z}-2<0.\end{aligned}$$

Now suppose that $(x,y,z,w)\in D_3$. Then

$$\begin{aligned}U_z(x,y,z,w)&=\exp\Big(\frac{|x|+|y|-w}{z}\Big)\cdot\Big(1-(z-|x|)\frac{|x|+|y|-w}{z^2}\Big)-2\\&\leqslant\exp\Big(\frac{|x|+|y|-w}{z}\Big)\cdot\Big(1-\frac{|x|+|y|-w}{z}\Big)-2<0.\end{aligned}$$

Finally, when $(x,y,z,w)\in D_4$, then

$$U_z(x,y,z,w)=-\frac{(|y|-w)^2}{2z^2}+\frac{|x|^2}{2z^2}-1<-\frac12$$

and we are done. $\square$

To prove the next property, let us introduce an auxiliary function $c\colon\ D\to[0,\infty)$ given by

$$c(x,y,z,w)=\begin{cases}z^{-1} & \text{if }\ (x,y,z,w)\in D_1,\\ z^{-1}\cdot \mathrm{e}^{1-w/z} & \text{if }\ (x,y,z,w)\in D_2,\\ z^{-1}\cdot\exp\Big(\dfrac{|x|+|y|-w}{z}\Big) & \text{if }\ (x,y,z,w)\in D_3,\\ z^{-1} & \text{if }\ (x,y,z,w)\in D_4.\end{cases}$$

Lemma 2.2. *Let* $\mathtt{x}=(x,y,z,w)$ *be a point belonging to the interior of one of the sets* D_1, D_2, D_3 *or* D_4, *satisfying* $|x|\cdot|y|\neq 0$. *Then for any* $h,k\in\mathbb{R}^\nu$ *we have*

$$\langle U_{xx}(\mathtt{x})h,h\rangle+2\langle U_{xy}(\mathtt{x})h,k\rangle+\langle U_{yy}(\mathtt{x})k,k\rangle\leqslant c(\mathtt{x})(|k|^2-|h|^2). \tag{2.3}$$

P r o o f. If $\mathtt{x}$ belongs to the interior of D_1 or D_2, the claim is evident; in fact, then both sides of (2.3) are equal. The most technical part corresponds to the domain D_3. A little computation gives that the left-hand side of (2.3) is equal to $c(\mathtt{x})(|k|^2-|h|^2)+I+II$, where

$$\begin{aligned}I&=\Big(\frac{\langle y,k\rangle^2}{|y|^2}-|k|^2\Big)\cdot\frac{|x|+|y|-z}{2z|y|}\cdot\exp\Big(\frac{|x|+|y|-w}{z}\Big),\\ II&=-\frac{|x|}{2z^2}\Big(\frac{\langle x,h\rangle}{|x|}-\frac{\langle y,k\rangle}{|y|}\Big)^2\cdot\exp\Big(\frac{|x|+|y|-w}{z}\Big),\end{aligned}$$

and it suffices to note that both the terms above are nonpositive. Finally, if (x, y, z, w) lies in the interior of D_4, then we rewrite the definition of $U(x, y, z, w)$ in the form

$$U(x, y, z, w) = \frac{|y|^2 - |x|^2 - 2(w - z)|y| + (w - z)^2 - 3z^2}{2z} + w.$$

If the term $-2(w - z)|y|$ were absent in the numerator, then we would have equality in (2.3). However, the function $(x, y) \mapsto -(w-z)|y|/z$ is concave on $\mathbb{R}^\nu \times \mathbb{R}^\nu$, because of the inequality $w > z$. This yields (2.3) and completes the proof. □

The final fact concerning the function U is the following.

Lemma 2.3. *For any $(x, y, z, w) \in D$ such that $0 < |y| \leqslant |x|$ we have*

$$U(x, y, |x|, |y|) \leqslant 0. \tag{2.4}$$

Proof. This is straightforward: for x, y as above, we have $(x, y, |x|, |y|) \in D_1$ and hence $U(x, y, |x|, |y|) = (|y|^2 - 2|x|^2)/(2|x|) \leqslant 0$. □

3. Proof of the inequality (1.3)

For the reader's convenience, we have split this section into two parts. In the first part we present a slight modification of the function U, and then, in the other part, we use its properties to establish the inequality (1.3).

3.1. A mollified function. The general idea of the proof of (1.3) is to prove that the process $U(X, Y, X^*, Y^*)$ is a supermartingale. To show this, it is natural to try to apply Itô's formula and use the inequality (2.3) together with the differential subordination to control the finite variation term. However, things are a little bit more complicated since the function U does not have the necessary smoothness and the direct application of Itô's formula is not permitted. To overcome this difficulty, we use a standard mollification argument. Pick a radial function $g\colon \mathbb{R}^\nu \times \mathbb{R}^\nu \times \mathbb{R} \times \mathbb{R} \to [0, \infty)$ of class C^∞, supported on the unit ball B of $\mathbb{R}^\nu \times \mathbb{R}^\nu \times \mathbb{R} \times \mathbb{R}$, satisfying $\int_B g = 1$. For a fixed $\delta > 0$ and $(x, y, z, w) \in D$ such that $|x| > \delta$ and $w > 3\delta$, define

$$U^\delta(x, y, z, w) = \int_B U(x + \delta u, y + \delta v, z + 2\delta + \delta r, w - 2\delta + \delta s) g(u, v, r, s)\, \mathrm{d}u\, \mathrm{d}v\, \mathrm{d}r\, \mathrm{d}s.$$

This function is of class C^∞ and inherits all the crucial properties of U. First of all, the somewhat surprising summand 2δ on the third coordinate guarantees that U^δ is well-defined: we have $|x+\delta u| \leqslant |z+2\delta+\delta r|$ and hence $(x+\delta u, y+\delta v, z+2\delta+\delta r, w-2\delta+\delta s)$ falls into the domain of U. By (2.1), we have the majorization

$$\begin{aligned} U^\delta(x,y,z,w) &\geqslant \int_B [(w-2\delta+\delta s)-2(z+2\delta+\delta r)]g(u,v,r,s)\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}r\,\mathrm{d}s \\ &= w-2z-6\delta, \end{aligned} \tag{3.1}$$

where on the last line we have used the fact that g is radial and has integral 1. Furthermore, we have

$$U^\delta_z \leqslant 0 \tag{3.2}$$

on the domain of U^δ, which follows directly from (2.2) by integration by parts. There is a version of this inequality for the partial derivative U_w: if δ is sufficiently small, then for any $(x,y,z,w) \in D$ such that $|x| > 3\delta$ and $|y| = w > 3\delta$ we have

$$U^\delta_w(x,y,z,w) \leqslant 0. \tag{3.3}$$

To show this, we use integration by parts to get

$$U^\delta_w(x,y,z,w) = \int_B U_w(x+\delta u, y+\delta v, z+2\delta+\delta r, w-2\delta+\delta s)g(u,v,r,s)\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}r\,\mathrm{d}s.$$

Now, if $w-2\delta+\delta s < z+2\delta+\delta r$, then the integrand vanishes (because then we have $(x+\delta u, y+\delta v, z+2\delta+\delta r, w-2\delta+\delta s) \in D_1$ and the function U restricted to D_1 does not depend on w). If $w-2\delta+\delta s > z+2\delta+\delta r$, then $|x+\delta u|+|y+\delta v| > 3\delta-\delta+|y|-\delta \geqslant |w-2\delta+\delta s|$, so the point $(x+\delta u, y+\delta v, z+2\delta+\delta r, w-2\delta+\delta s)$ belongs to the interior of D_4. Therefore,

$$\begin{aligned} &U_w(x+\delta u, y+\delta v, z+2\delta+\delta r, w-2\delta+\delta s) \\ &\qquad = \frac{(w-2\delta+\delta s)-|y+\delta v|}{z+2\delta+\delta r} \leqslant \frac{w-2\delta+\delta s-w+\delta v}{z+2\delta+\delta r} < 0 \end{aligned}$$

and (3.3) is established. Finally, the function U^δ inherits the property (2.3). To see this, fix $\mathtt{x} = (x,y,z,w)$ belonging to the domain of U^δ. A combination of Lemma 2.1 (i) with integration by parts gives

$$U^\delta_{xx}(\mathtt{x}) = \int_B U_{xx}(x+\delta u, y+\delta v, z+2\delta+\delta r, w-2\delta+\delta s)g(u,v,r,s)\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}r\,\mathrm{d}s$$

and similar formulas for the remaining second-order partial derivatives of U^δ. Thus,

$$\langle U^\delta_{xx}(\mathtt{x})h,h\rangle + 2\langle U^\delta_{xy}(\mathtt{x})h,k\rangle + \langle U^\delta_{yy}(\mathtt{x})k,k\rangle \leqslant c^\delta(\mathtt{x})(|k|^2-|h|^2), \tag{3.4}$$

where c^δ is a nonnegative function given by

$$c^\delta(\mathrm{x}) = \int_B c(x+\delta u, y+\delta v, z+2\delta+\delta r, w-2\delta+\delta s)g(u,v,r,s)\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}r\,\mathrm{d}s.$$

Equipped with the function U^δ, we turn to the assertion of Theorem 1.4.

3.2. Proof of the inequality (1.3)**.** With no loss of generality we may and do assume that $\|X^*\|_1$ is finite, since otherwise there is nothing to prove. Furthermore, we may restrict ourselves to the setting in which X and Y are bounded away from 0. Indeed, if this is not the case, then we fix a small positive number a and consider the $\mathbb{R}^{\nu+1}$-valued martingales $\overline{X}$, $\overline{Y}$ given by $\overline{X}_t = (X_t, a)$, $\overline{Y}_t = (Y_t, a)$ for $t \geqslant 0$. These new processes are bounded away from 0 and inherit the differential subordination. Having proved (1.3) for $\overline{X}$ and $\overline{Y}$, we let $a \to 0$ and obtain the desired estimate for the initial pair. Thus, from now on, we assume that $\inf_{t\geqslant 0} |X_t|$ and $\inf_{t\geqslant 0} |Y_t|$ are larger than a certain deterministic constant $\varepsilon > 0$. Fix a large positive integer N and consider the stopping time $\tau_N = \inf\{t \geqslant 0\colon\ |X_t| + |Y_t| \geqslant N\}$. Pick $\delta \in (0, \varepsilon/3)$, and apply Itô's formula (cf. Revuz and Yor [14]) to U^δ composed with the process $\mathscr{Z} = (X, Y, X^*, Y^*)$ to get

$$U^\delta(\mathscr{Z}_{\tau_N\wedge t}) - U^\delta(\mathscr{Z}_0) = I_1 + I_2 + I_3/2, \tag{3.5}$$

where

$$\begin{aligned}
I_1 &= \int_{0+}^{\tau_N\wedge t} U_x^\delta(\mathscr{Z}_s)\cdot \mathrm{d}X_s + \int_{0+}^{\tau_N\wedge t} U_y^\delta(\mathscr{Z}_s)\cdot \mathrm{d}Y_s,\\
I_2 &= \int_{0+}^{\tau_N\wedge t} U_z^\delta(\mathscr{Z}_s)\,\mathrm{d}X_s^* + \int_{0+}^{\tau_N\wedge t} U_w^\delta(\mathscr{Z}_s)\,\mathrm{d}Y_s^*,\\
I_3 &= \int_{0+}^{\tau_N\wedge t} U_{xx}^\delta(\mathscr{Z}_s)\,\mathrm{d}[X,X]_s + 2\int_{0+}^{\tau_N\wedge t} U_{xy}^\delta(\mathscr{Z}_s)\,\mathrm{d}[X,Y]_s\\
&\quad + \int_{0+}^{\tau_N\wedge t} U_{yy}^\delta(\mathscr{Z}_s)\,\mathrm{d}[Y,Y]_s.
\end{aligned}$$

Let us analyze the terms I_1–I_3. We have $\mathbb{E}I_1 = 0$, since both the stochastic integrals are martingales. Next, $I_2 \leqslant 0$: by (3.2), we have $U_z(\mathscr{Z}_s) \leqslant 0$ and hence the first integral in I_2 is nonpositive. Furthermore, for any $\omega \in \Omega$, the second summand in I_2 is the Lebesgue-Stieltjes integral of $U_w^\delta(\mathscr{Z}_s(\omega))$ with respect to the continuous nondecreasing function $s \mapsto Y_s^*(\omega)$. Clearly, the support of the measure generated by this function is contained in the set $\{s\colon\ |Y_s(\omega)| = Y_s^*(\omega)\}$, on which the integrand is nonpositive (see (3.3)). This shows that the second integral, and hence the whole I_2, is nonpositive. To deal with I_3, fix $0 \leqslant s_0 < s_1 \leqslant t$. For any $l \geqslant 0$, let $(\eta_i^l)_{1\leqslant i\leqslant i_l}$ be a nondecreasing sequence of stopping times with $\eta_0^l = s_0$, $\eta_{i_l}^l = s_1$

such that $\lim_{l\to\infty} \max_{1\leqslant i\leqslant i_l-1} |\eta_{i+1}^l - \eta_i^l| = 0$. Keeping l fixed, we apply, for each $i = 0, 1, 2, \ldots, i_l$, the property (3.4) to $x = X_{s_0}$, $y = Y_{s_0}$, $z = X_{s_0}^*$, $w = Y_{s_0}^*$ and $h = h_i^l = X_{\tau_N\wedge\eta_{i+1}^l} - X_{\tau_N\wedge\eta_i^l}$, $k = k_i^l = Y_{\tau_N\wedge\eta_{i+1}^l} - Y_{\tau_N\wedge\eta_i^l}$. We sum the obtained $i_l + 1$ inequalities and let $l \to \infty$. Using the notation $[S,T]_s^u = [S,T]_u - [S,T]_s$, we may write the result in the form

$$\begin{aligned}\sum_{m=1}^{\nu}\sum_{n=1}^{\nu}[U_{x_m x_n}^\delta(\mathscr{Z}_{s_0})[X^m, X^n]_{\tau_N\wedge s_0}^{\tau_N\wedge s_1} &+ 2U_{x_m y_n}^\delta(\mathscr{Z}_{s_0})[X^m, Y^n]_{\tau_N\wedge s_0}^{\tau_N\wedge s_1} \\ &+ U_{y_m y_n}^\delta(\mathscr{Z}_{s_0})[Y^m, Y^n]_{\tau_N\wedge s_0}^{\tau_N\wedge s_1}] \\ \leqslant c^\delta(\mathscr{Z}_{s_0})\{[Y,Y]_{\tau_N\wedge s_0}^{\tau_N\wedge s_1} &- [X,X]_{\tau_N\wedge s_0}^{\tau_N\wedge s_1}\} \leqslant 0,\end{aligned}$$

where the last inequality is due to the differential subordination. Thus $I_3 \leqslant 0$, using a standard approximation of integrals by discrete sums. Plugging all the above facts into (3.5) and taking expectation of both sides, we obtain $\mathbb{E}\{U^\delta(\mathscr{Z}_{\tau_N\wedge t}) - U^\delta(\mathscr{Z}_0)\} \leqslant 0$, or

$$\mathbb{E}U^\delta(\mathscr{Z}_{\tau_N\wedge t})1_{\{\tau_N>0\}} \leqslant \mathbb{E}U^\delta(\mathscr{Z}_0)1_{\{\tau_N>0\}}.$$

An application of (3.1) gives

$$\mathbb{E}(Y_{\tau_N\wedge t}^* - 2X_{\tau_N\wedge t}^* - 6\delta)1_{\{\tau_N>0\}} \leqslant \mathbb{E}U^\delta(\mathscr{Z}_0)1_{\{\tau_N>0\}}.$$

By the continuity of U, if we let $\delta \to 0$, then $U^\delta(\mathscr{Z}_0)$ converges to $U(\mathscr{Z}_0) = U(X_0, Y_0, |X_0|, |Y_0|)$, which is nonpositive (see (2.4)). Therefore, by Lebesgue's dominated convergence theorem,

$$\mathbb{E}Y_{\tau_N\wedge t}^* 1_{\{\tau_N>0\}} \leqslant 2\mathbb{E}X_{\tau_N\wedge t}^* 1_{\{\tau_N>0\}}.$$

Finally, letting N go to infinity yields (1.3), in light of Lebesgue's monotone convergence theorem.

4. Sharpness

Now we will construct an appropriate example to show that the constant 2 is optimal in (1.3). The construction consists of two steps. Let K be a large even integer and assume that B is a standard one-dimensional Brownian motion starting from 1.

Step 1. Introduce a nondecreasing sequence $(\tau_n)_{n=0}^K$ of stopping times given by $\tau_0 \equiv 0$ and, inductively,

$$\tau_{n+1} = \inf\{t \geqslant \tau_n\colon\ B_t \in \{0, n+2\}\}, \quad n = 0, 1, 2, \ldots, K-1.$$

Define X and Y by

$$X_t = B_t \quad \text{and} \quad Y_t = \sum_{n=0}^{K-1} (-1)^n (B_{\tau_{n+1}\wedge t} - B_{\tau_n \wedge t})$$

for $t \in [0, \tau_K]$. We see that (X, Y) starts from the point $(1, 0)$ and moves along the line segment of slope 1, joining the points $(0, -1)$ and $(2, 1)$. If the process reaches the point $(0, -1)$, it stops (because, directly from the definition, we have $\tau_1 = \tau_2 = \ldots = \tau_K$); if the pair gets to the point $(2, 1)$ first, then it starts to evolve along the line segment joining $(0, 3)$ and $(3, 0)$ (note that the slope switches to -1). If (X, Y) visits $(0, 3)$, the process stops (by similar reasons as above); if it gets to the other endpoint of the line segment, then the pair begins to move along the line segment with endpoints $(0, -3)$ and $(4, 1)$, and so on. The first stage ends at time τ_K, when (X, Y) reaches $(1 + K, 0)$ or visits the line $x = 0$. Observe that if $Y_{\tau_K} \neq 0$ (so (X_{τ_K}, Y_{τ_K}) lands on the y-axis), then $Y^*_{\tau_K} = |Y_{\tau_K}| \in [X^*_{\tau_K} - 1, X^*_{\tau_K} + 1]$, directly from the construction.

Step 2. Define another nondecreasing sequence $(\sigma_n)_{n \geqslant 0}$ of stopping times, given by $\sigma_0 \equiv \tau_K$ and, by induction,

$$\begin{aligned}
\sigma_{2n+1} &= \inf\{t \geqslant \sigma_{2n} \colon\ B_t \leqslant -|Y_{\tau_K}| \text{ or } B_t \geqslant 1/2\}, \\
\sigma_{2n+2} &= \inf\{t \geqslant \sigma_{2n+1} \colon\ B_t \leqslant 0 \text{ or } B_t \geqslant |Y_{\tau_K}|\}
\end{aligned}$$

for $n = 0, 1, 2, \ldots$. Clearly, $(\sigma_n)_{n \geqslant 0}$ converges almost surely to $\sigma = \inf\{t \geqslant \tau_K \colon |B_t| \geqslant |Y_{\tau_K}|\}$, which is finite with probability 1. For $t > \tau_K$, put $X_t = B_{\sigma \wedge t}$ and

$$Y_t = \sum_{n=0}^{\infty} (-1)^n (B_{\sigma_{n+1}\wedge t} - B_{\sigma_n \wedge t}) \cdot \operatorname{sgn} Y_{\tau_K}.$$

To understand what happens during the second stage, observe first that the process (X, Y) does not evolve at all when $X_{\tau_K} = 1 + K$: indeed, then we have $B_{\tau_K} = 1 + K \geqslant |Y_{\tau_k}|$ and hence $\sigma_0 = \sigma_1 = \sigma_2 = \ldots = \sigma = \tau_K$. Suppose then, that $X_{\tau_K} = 0$ and $Y_{\tau_K} > 0$ (if $Y_{\tau_K} < 0$ then the behavior of the pair (X, Y) is symmetric). We have that $((X_t, Y_t))_{t \geqslant \tau_K}$ starts from $(0, Y_{\tau_K})$ and first moves along the line segment of slope 1, which joins $(-Y_{\tau_K}, 0)$ and $(1/2, Y_{\tau_K} + 1/2)$. If (X, Y) gets to the first endpoint, it stays there forever. If the pair reaches the second endpoint, then the line segment along which the process evolves changes to the one with endpoints $(Y_{\tau_K}, 1)$ and $(0, Y_{\tau_K} + 1)$. If (X, Y) gets to $(Y_{\tau_K}, 1)$ first, then the evolution stops; otherwise, the pair starts to move along the line segment joining $(-Y_{\tau_K} + 1, 1)$ and $(1/2, Y_{\tau_K} + 3/2)$. The pattern is then repeated.

Calculation. We start with some observations which follow from the above construction. First, X is a stopped Brownian motion, Y is an integral with respect to X of a predictable process with values in $\{-1, 1\}$, and both these martingales are uniformly integrable. Second, we have

$$X^* = \max\left\{X^*_{\tau_K}, \sup_{t>\tau_K} |X_t|\right\} \leqslant \max\{X^*_{\tau_K}, |Y_{\tau_K}|\} \leqslant X^*_{\tau_K} + 1. \tag{4.1}$$

Next, a closer look at the second stage shows that the process Y does not change its sign on the interval $[\tau_K, \infty)$, so $\mathbb{E}[|Y_\sigma||\mathscr{F}_{\tau_K}] = |Y_{\tau_K}|$ by the martingale property. Finally, if $Y_{\tau_K} \neq 0$, then

$$Y^* \geqslant |Y_\sigma| + |Y_{\tau_K}| - 1/2,$$

which combined with the preceding observation yields

$$\mathbb{E}Y^* 1_{\{Y_{\tau_K} \neq 0\}} \geqslant 2\mathbb{E}|Y_{\tau_K}| 1_{\{Y_{\tau_K} \neq 0\}} - 1/2.$$

However, on $\{Y_{\tau_K \neq 0}\}$ we have $|Y_{\tau_K}| \geqslant X^*_{\tau_K} - 1$ (see the last line in the description of Step 1) and hence

$$\mathbb{E}Y^* \geqslant 2\mathbb{E}X^*_{\tau_K} 1_{\{Y_{\tau_K} \neq 0\}} - 5/2. \tag{4.2}$$

Now, directly from the elementary properties of Brownian motion, we deduce that

$$\mathbb{P}(X^*_{\tau_K} \geqslant s) = \begin{cases} 1 & \text{if } \ s \in [0, 1], \\ s^{-1} & \text{if } \ s \in [0, K+1], \\ 0 & \text{if } \ s > K+1, \end{cases}$$

and hence

$$\begin{aligned} \mathbb{E}X^*_{\tau_K} &= \int_0^\infty \mathbb{P}(X^*_{\tau_K} \geqslant s)\mathrm{d}s = 1 + \ln(K+1), \\ \mathbb{E}X^*_{\tau_K} 1_{\{Y_{\tau_K} \neq 0\}} &= \mathbb{E}X^*_{\tau_K} - (K+1)\mathbb{P}(X^*_{\tau_K} = K+1) = \ln(K+1). \end{aligned}$$

Plugging these identities into (4.2) and applying (4.1) yields

$$\frac{\mathbb{E}Y^*}{\mathbb{E}X^*} \geqslant \frac{2\ln(K+1)}{2 + \ln(K+1)} - \frac{5}{2(1 + \ln(K+1))},$$

which can be made arbitrarily close to 2 by taking sufficiently large K. This proves the desired sharpness.

5. On the search of the suitable function

Let us sketch some steps which have led to the right choice of the optimal constant 2, and the right guess of the special function U used in the proof of (1.3). We would like to stress here that the reasoning we present is informal and rests on several intuitive assumptions. For the sake of clarity, we have split this section into three parts.

5.1. Assumptions. Suppose that β is the best constant in the inequality

$$\|Y^*\|_1 \leqslant \beta \|X^*\|_1,$$

where (X, Y) runs over the class of all pairs of continuous-path *real-valued* martingales such that Y is differentially subordinate to X. Of course, this is equivalent to saying that for such (X, Y) we have

$$\mathbb{E}V(X_t, Y_t, X_t^*, Y_t^*) \leqslant 0 \quad \text{for all } t \geqslant 0,$$

where $V(x, y, z, w) = w - \beta z$. The general idea of Burkholder's method is to find a function U defined on the set $\{(x, y, z, w) \in \mathbb{R}\times\mathbb{R}\times[0,\infty)\times[0,\infty) \colon |x| \leqslant z,\ |y| \leqslant w\}$ and satisfying the following two conditions: first,

$$V(x, y, z, w) \leqslant U(x, y, z, w) \tag{5.1}$$

and second, that for all X, Y as above,

$$\mathscr{U} = (U(X_s, Y_s, X_s^*, Y_s^*))_{s\geqslant 0} \text{ is a supermartingale with } \mathscr{U}_0 \leqslant 0. \tag{5.2}$$

Clearly, the existence of such U yields the desired bound: indeed, then

$$\mathbb{E}(Y_t^* - \beta X_t^*) \leqslant \mathbb{E}\mathscr{U}_t \leqslant \mathbb{E}\mathscr{U}_0 \leqslant 0. \tag{5.3}$$

How to find the right function? To avoid technical problems, we assume that U is of class C^2. The first observation is that the function V is homogeneous of order 1 and satisfies $V(\pm x, \pm y, z, w) = V(x, y, z, w)$; it is reasonable to expect that U also should have these properties. Thus, the problem of finding U is reduced to that of finding $(x, y, w) \mapsto U(x, y, 1, w)$, $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant w$. The next step is to look at (5.2). In contrast with (5.1), which is of nice analytic form, this condition is more difficult to capture and thus it is plausible to replace it with possibly weaker set of pointwise estimates. A glimpse at the proof of (1.3) above suggests to impose the following requirements. First, for any x, y, z, w such that $|x| \leqslant z$, $|y| \leqslant w$,

$$U_z(x, y, |x|, w) \leqslant 0, \; U_w(x, y, z, |y|) \leqslant 0. \tag{5.4}$$

The second assumption is the existence of a nonnegative function c such that for all $\mathtt{x} = (x, y, z, w)$ and all $h, k \in \mathbb{R}$,

$$U_{xx}(\mathtt{x})h^2 + 2U_{xy}(\mathtt{x})hk + U_{yy}(\mathtt{x})k^2 \leqslant c(\mathtt{x})(k^2 - h^2).$$

This is just the one-dimensional version of (2.3). If we apply it with $z = 1$ and $h = \pm k$, we obtain the following consequence: for any fixed w, the function $(x, y) \mapsto U(x, y, 1, w)$ is concave along any line segment of slope ± 1 contained in the rectangle $[0, 1] \times [0, w]$. Such concavity is a typical property of Burkholder's functions (see the survey [5]); actually, much more can be said. Namely, usually for most (x, y) there is a (small) line segment of slope 1 or -1, passing through (x, y), such that the corresponding restriction is *linear*. Motivated by the properties of the special function constructed in [11] (where Theorem 1.3 was proved), we *assume* that

$$\begin{aligned} &(x, y) \mapsto U(x, y, 1, w) \text{ is linear along the line segments of slope } -1 \\ &\qquad\qquad\qquad \text{contained in } [0, 1] \times [0, w]. \end{aligned} \tag{5.5}$$

The next step is to look at the set $\mathscr{D} = \{(x, y, z, w) \colon U(x, y, z, w) = V(x, y, z, w)\}$. Since β is the best constant in the maximal inequality, there are $t \geqslant 0$ and a pair (X, Y) of differentially subordinate martingales for which $\|Y_t^*\|_1$ and $\beta\|X_t^*\|_1$ are almost equal. This, in view of (5.3), leads to the natural conjecture that the set $\mathscr{D}$ is nonempty (we expect to have "almost" equality throughout in (5.3), which combined with (5.1) enforces $U(X_t, Y_t, X_t^*, Y_t^*) \approx V(X_t, Y_t, X_t^*, Y_t^*)$ with overwhelming probability). What can be said about the structure of $\mathscr{D}$? No point of the form $(x, y, 1, |y|)$ can belong to it: otherwise, we exploit (5.4) and find, for any $\varepsilon > 0$, a number $w > |y|$ such that

$$U(x, y, 1, w) \leqslant U(x, y, 1, |y|) + \varepsilon(w - |y|) = |y| - \beta + \varepsilon(w - |y|) < w - \beta,$$

a contradiction with (5.1). Furthermore, $\mathscr{D}$ cannot contain any point of the form $(x, y, 1, w)$ with $|x| < 1$ and $|y| < w$. Indeed, otherwise, by (5.1) and the aforementioned concavity of U along the line segments of slope ± 1 (combined with the fact that V is constant along these segments) we would obtain that the whole rectangle $[-1, 1] \times [-|y|, |y|] \times \{1\} \times \{|y|\}$ would belong to $\mathscr{D}$. In particular, the point $(x, y, 1, |y|)$ would lie in $\mathscr{D}$, which is impossible, as we have shown above. Therefore, the set $\mathscr{D}$ can only contain points of the form $(x, y, |x|, w)$ with $|y| < w$. The crucial assumption, coming from experimentation, is as follows:

$$\text{if } w > 1 \text{ and } y \leqslant w - 1, \text{ then } (1, y, 1, w) \in \mathscr{D}. \tag{5.6}$$

Finally, we impose the condition (cf. (5.4))

$$U_w(0, y, 1, |y|) = 0. \tag{5.7}$$

5.2. Deriving U on $D_2 \cup D_3$. Introduce the function $A(y, w) = U(0, y, 1, w)$, $0 \leqslant y \leqslant w$. By (5.5), if $x, y \geqslant 0$ and $1 \leqslant x + y \leqslant w$, then

$$U(x, y, 1, w) = (1 - x)A(y + x, w) + xU(1, y + x - 1, 1, w).$$

By (5.6), this is equivalent to

$$U(x, y, 1, w) = (1 - x)A(y + x, w) + x(w - \beta). \tag{5.8}$$

Since U satisfies the symmetry condition $U(x, y, z, w) = U(-x, y, z, w)$ (this is one of the assumptions), we get $U_x(0, y, 1, w) = 0$ and hence, for $1 \leqslant y \leqslant w$, we have $A_y(y, w) - A(y, w) + w - \beta = 0$. This differential equation can be easily solved: we get that

$$A(y, w) = C(w)\mathrm{e}^y + w - \beta \quad \text{for } y \in [1, w], \tag{5.9}$$

for some function C to be found. An application of (5.7) yields $C'(w)\mathrm{e}^w + 1 = 0$, so $C(w) = \mathrm{e}^{-w} + K$ for some constant K, and hence (5.8) gives

$$U(x, y, 1, w) = (1 - x)\mathrm{e}^{x+y-w} + K(1 - x)\mathrm{e}^{x+y} + w - \beta \tag{5.10}$$

provided $x \in [0, 1]$, $1 \leqslant x + y \leqslant w$. Consequently, if $(x, y, z, w) \in D_3$, then

$$U(x, y, z, w) = (z - |x|) \exp\Big(\frac{|x| + |y| - w}{z}\Big) + K(z - x) \exp\Big(\frac{|x| + |y|}{z}\Big) + w - \beta z.$$

In particular, we have $U(0, w, 1, w) = 1 + K\mathrm{e}^w + w - \beta$, so $K \geqslant 0$, since otherwise (5.1) is violated for large w. On the other hand, we derive that $U_z(1, w - 1, 1, w) = 1 + K\mathrm{e}^w - \beta$, which implies $K \leqslant 0$, since otherwise (5.4) does not hold for large w. Thus $K = 0$, and on D_3 the function U is given by

$$U(x, y, z, w) = (z - |x|) \exp\Big(\frac{|x| + |y| - w}{z}\Big) + w - \beta z.$$

Now we will derive the formula for U on D_2. For $x \in [0,1]$, define $B(x,w) = U(x,0,1,w)$. Pick $x, y \geqslant 0$ with $x+y \leqslant 1$ and apply (5.5) to get that

$$U(x,y,1,w) = \frac{x}{x+y} B(x+y,w) + \frac{y}{x+y} A(x+y,w).$$

By the symmetry condition $U(\pm x, \pm y, 1, w) = U(x,y,1,w)$ we have $U_x(0,y,1,w) = U_y(x,0,1,w) = 0$, which gives the following system of partial differential equations:

$$\frac{B(y,w)}{y} - \frac{A(y,w)}{y} + A_y(y,w) = 0, \tag{5.11}$$

$$\frac{A(x,w)}{x} - \frac{B(x,w)}{x} + B_x(x,w) = 0. \tag{5.12}$$

Replacing x with y and summing the equations, we get $A_y(y,w) + B_x(y,w) = 0$, which implies that $A(y,w) + B(y,w) = \alpha$ for some constant α. Plugging this into (5.11) gives

$$A_y(y,w) = \frac{2A(y,w) - \alpha}{y}, \quad y \in [0,1].$$

It is straightforward to solve this: we get $A(y,w) = \gamma y^2 + \alpha/2$ for some constant γ. Since A is of class C^1, comparing this formula with (5.9) yields

$$A(1,w) = \gamma + \frac{\alpha}{2} = \mathrm{e}^{1-w} + w - \beta, \quad A_y(1,w) = 2\gamma = \mathrm{e}^{1-w}.$$

Therefore,

$$A(y,w) = \frac{1}{2}\mathrm{e}^{1-w} y^2 + \frac{\mathrm{e}^{1-w}}{2} + w - \beta, \quad B(x,w) = -\frac{1}{2}\mathrm{e}^{1-w} x^2 + \frac{\mathrm{e}^{1-w}}{2} + w - \beta,$$

so, by (5.10), $U(x,y,1,w) = \mathrm{e}^{1-w}(y^2 - x^2 + 1)/2 + w - \beta$. Exploiting the homogeneity of U, we obtain that on D_2

$$U(x,y,z,w) = \frac{|y|^2 - |x|^2 + z^2}{2z} \cdot \mathrm{e}^{1-w/z} + w - \beta z.$$

5.3. The formula for U on $D_1 \cup D_4$. To find the formula for U on D_1, take a point (x,y,z,w) lying on the boundary of D_1 and D_2, i.e., satisfying $w = z$. Then

$$U(x,y,z,w) = \frac{|y|^2 - |x|^2 - z^2}{2z} + (2-\beta)z.$$

We *guess* that this formula holds true for all $(x,y,z,w) \in D_1$. Then for $w < 1$ we have $U_z(1,0,1,w) = 2 - \beta$ and hence (5.4) implies $\beta \geqslant 2$. Assuming equality, we obtain the function which coincides with the special function of Section 2 on the sets

D_1, D_2 and D_3. Finally, to get the formula on D_4, the author experimented with the expression of the form

$$\frac{(y - F(w,z))^2 - x^2 - \kappa_1 z^2}{\kappa_2 z} + w - 2z,$$

with the function F and the parameters κ_1, κ_2 to be found. Expressions of this type appear in many Burkholder's functions (see [11], [12] and [13]); actually, the formulas on D_1 and D_2 are also of similar type. The unknown parameters can be derived from the fact that U is of class C^1; the luck is with us, we are led precisely to the right formula.

This completes the search.

Acknowledgment. The author would like to express his gratitude to the anonymous referee for the careful reading of the first version of the paper, and for many helpful comments and suggestions.

References

[1] *R. Bañuelos, P. J. Méndez-Hernandez*: Space-time Brownian motion and the Beurling-Ahlfors transform. Indiana Univ. Math. J. *52* (2003), 981–990.

[2] *R. Bañuelos, G. Wang*: Sharp inequalities for martingales with applications to the Beurling-Ahlfors and Riesz transforms. Duke Math. J. *80* (1995), 575–600.

[3] *D. L. Burkholder*: Boundary value problems and sharp inequalities for martingale transforms. Ann. Probab. *12* (1984), 647–702.

[4] *D. L. Burkholder*: A sharp and strict L^p-inequality for stochastic integrals. Ann. Probab. *15* (1987), 268–273.

[5] *D. L. Burkholder*: Explorations in martingale theory and its applications. Calcul des Probabilités. Ec. d'Été, Saint-Flour/Fr. 1989, Lect. Notes Math. 1464, Springer, Berlin, 1991, pp. 1–66.

[6] *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). Proc. Sympos. Appl. Math., 52, Amer. Math. Soc., Providence, RI, 1997, pp. 343–358.

[7] *C. Dellacherie, P. A. Meyer*: Probabilities and Potential. B: Theory of Martingales. Transl. from the French and prep. by J. P. Wilson, North-Holland Mathematics Studies, Amsterdam, 1982.

[8] *S. Geiss, S. Montgomery-Smith, E. Saksman*: On singular integral and martingale transforms. Trans. Am. Math. Soc. *362* (2010), 553–575.

[9] *F. L. Nazarov, A. Volberg*: Heat extension of the Beurling operator and estimates for its norm. Algebra i Analiz *15* (2003), 142–158 (In Russian.); translation in St. Petersburg Math. J. *15* (2004), 563–573.

[10] *A. Osękowski*: Sharp LlogL inequality for differentially subordinated martingales. Illinois J. Math. *52* (2008), 745–756.

[11] *A. Osękowski*: Sharp inequality for martingale maximal functions and stochastic integrals. Illinois J. Math. *54* (2010), 1133–1156.

[12] *A. Osękowski*: Maximal inequalities for continuous martingales and their differential subordinates. Proc. Am. Math. Soc. *139* (2011), 721–734.
[13] *A. Osękowski*: Maximal inequalities for martingales and their differential subordinates. To appear in J. Theor. Probab. DOI:10.1007/s10959-012-0458-8.
[14] *D. Revuz, M. Yor*: Continuous Martingales and Brownian Motion. 3rd ed., Grundlehren der Mathematischen Wissenschaften 293, Springer, Berlin, 1999.
[15] *Y. Suh*: A sharp weak type (p, p) inequality $(p > 2)$ for martingale transforms and other subordinate martingales. Trans. Am. Math. Soc. *357* (2005), 1545–1564.
[16] *G. Wang*: Differential subordination and strong differential subordination for continuous-time martingales and related sharp inequalities. Ann. Probab. *23* (1995), 522–551.

Author's address: Adam Osękowski, Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Warsaw, Poland, e-mail: ados@mimuw.edu.pl.