Weighted inequalities for the martingale square and maximal functions

Adam Osękowski

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

Abstract

Let W be a weight, i.e., a uniformly integrable, continuous-path martingale, and let W^* denote the associated maximal function. We show that if X is an arbitrary càdlàg martingale and X^*, $[X]$ denote its maximal and square functions, then

$$||[X]^{1/2}||_{L^p(W)} \leq \gamma_p ||X^*||_{L^p(W^*)}, \qquad 1 \leq p \leq 2,$$

where

$$\gamma_p^2 = 1 + \sup_{t>1} \frac{(2t-1)(1-t^{p-2})}{t^p - 1}.$$

The estimate is sharp for $p \in \{1, 2\}$. Furthermore, it is proved that if $p > 2$, then the above weighted inequality does not hold with any finite constant γ_p depending only on p.

Keywords: Martingale, maximal function, square function, best constant
2010 MSC: 60G44, 60G42.

1. Introduction

Square function and maximal inequalities for martingales are of fundamental importance to the whole probability theory, and their extensions and applications to various areas of mathematics can be found in numerous papers in the literature. The purpose of this paper is to investigate weighted inequalities involving the square and maximal function of an arbitrary càdlàg martingale. For closely related results, see e.g. Burkholder (1989), Kazamaki (1994), Osękowski (2012), Pisier & Xu (1997) and references therein.

Let us start with introducing the background and notation. Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, filtered by $(\mathcal{F}_t)_{t\geq 0}$, a nondecreasing

Email address: ados@mimuw.edu.pl (Adam Osękowski)
URL: www.mimuw.edu.pl/∼ados (Adam Osękowski)

family of sub-σ-fields of $\mathcal{F}$, such that $\mathcal{F}_0$ contains all the events of probability 0. Let X be an adapted martingale taking values in a separable real Hilbert space $\mathcal{H}$ with norm $|\cdot|$ and inner product $\langle\cdot,\cdot\rangle$; with no loss of generality, we may and will assume that $\mathcal{H}=\ell_2$. We impose the usual regularity conditions on X and assume that its paths are right-continuous and have limits from the left. Let $X^*=\sup_{s\geq 0}|X_s|$ stand for the maximal function of X; we will also use the notation $X_t^*=\sup_{0\leq s\leq t}|X_s|$ for the associated truncated maximal function. Next, let $[X]=([X]_t)_{t\geq 0}$ denote the square function (or square bracket) of X: see e.g. Dellacherie & Meyer (1982) for details when $\mathcal{H}=\mathbb{R}$, and extend the definition to the general case by setting $[X]_t=\sum_{j=1}^\infty [X^j]_t$, where X^j stands for the j-th coordinate of X.

The inequalities involving X, $[X]$ and X^* have their roots in the classical works of Khintchine (1923), Littlewood (1930), Marcinkiewicz (1937) and Paley (1932) concerning the properties of Haar and Rademacher systems. For an overview of the results in this direction and their extensions to the martingale setting, we refer the interested reader to the survey by Burkholder (1989). The proof of the inequality

$$c_p||X||_p \leq ||[X]^{1/2}||_p \leq C_p||X||_p, \qquad \text{if } 1<p<\infty, \tag{1}$$

valid for all real-valued martingales, can be extracted from Burkholder (1966) (that paper contains the discrete-time version of the above estimate, but the passage to the general setting is a matter of straightforward approximation arguments). Later, it was proved (cf. Burkholder (1989)) that the estimate holds with $c_p^{-1}=C_p=p^*-1$, where $p^*=\max\{p,p/(p-1)\}$. It turns out that the constant c_p is optimal for $p\geq 2$, C_p is the best for $1<p\leq 2$ and the proof carries over to the case of martingales taking values in a separable Hilbert space. It can be shown that the right inequality (1) does not hold for general martingales if $p\leq 1$ and nor does the left one if $p<1$. The paper by Osękowski (2005) proves that $c_1=1/2$ is the best. In the remaining cases the optimal constants c_p and C_p are not known.

Let us now turn our attention to related maximal inequalities. If $p>1$, then (1) combined with Doob's maximal inequality implies the existence of some finite c_p^*, C_p^* such that for any martingale X,

$$c_p^*||X^*||_p \leq ||[X]_\infty^{1/2}||_p \leq C_p^*||X^*||_p. \tag{2}$$

On the other hand, neither of the inequalities holds for $p<1$ without additional assumptions on X. The limit case $p=1$ was studied by Davis (1970), who proved the bound using a clever decomposition of the martingale. Then Burkholder (2002) proved that the optimal choice for the constant C_1^* is $\sqrt{3}$. In the other cases (except for $p=2$, when $c_2^*=1/2$ and $C_2^*=1$) the optimal values of c_p^* and C_p^* seem to be unknown at the moment.

In this paper, we will be interested in *weighted* estimates involving X^* and $[X]$. In a probabilistic context, the word "weight" refers to an adapted, uniformly integrable, nonnegative and continuous-path martingale $W=(W_t)_{t\geq 0}$.

In Kazamaki (1994), it was proved that if $W = \mathcal{E}(M)$ is the exponential process of some uniformly integrable martingale $M \in BMO$, then

$$C^{-1}||X^*||_{L^1(W)} \leq ||[X]_\infty^{1/2}||_{L^1(W)} \leq C||X^*||_{L^1(W)}$$

for all continuous-path martingales X and some positive finite constant C not depending on X. Here and below, $||\xi||_{L^p(W)} = (\mathbb{E}|\xi|^p W_\infty)^{1/p}$ denotes the usual weighted p-th norm of a random variable ξ, $1 \leq p < \infty$. The above result of Kazamaki implies in particular that if W satisfies the A_∞ condition

$$||W||_{A_\infty} := \sup_\tau \left|\left|Y_\tau \exp\left[-\mathbb{E}\left(\log W_\infty \middle| \mathcal{F}_\tau\right)\right]\right|\right|_\infty < \infty$$

(the supremum being taken over all adapted stopping times τ), then the weighted Davis' inequality holds true. The problem we will deal with is to study a version of the right inequality in (2) without any structural assumptions on the weight. It is clear that if W is arbitrary, then the right estimate in (2) does not hold in general with any finite C. For example, consider the Brownian motion X started at 1 and stopped upon leaving the interval $[0, 2]$. Then $X^* \leq 2$, while $[X]_\infty$ is an unbounded random variable; hence for any p there exists a positive, integrable random variable W_∞ for which $\mathbb{E}[X]_\infty^{p/2} W_\infty = \infty$ and $\mathbb{E}(X^*)^p W_\infty < \infty$. This variable gives rise to a uniformly integrable weight W for which our desired estimate is not true.

Thus, we need to modify the inequality. Motivated by related results of Fefferman & Stein (1971) from harmonic analysis, we will replace the weight W on the right by its maximal function. Here is the precise statement.

Theorem 1.1. *Suppose that W is a weight. Then for $1 \leq p \leq 2$ and any martingale X we have*

$$||[X]_\infty^{1/2}||_{L^p(W)} \leq \gamma_p ||X^*||_{L^p(W^*)}, \tag{3}$$

where $\gamma_p > 1$ is given by the equation

$$\gamma_p^2 = 1 + \sup_{t>1} \frac{(2t-1)(1-t^{p-2})}{t^p - 1}.$$

The inequality is sharp for $p \in \{1, 2\}$ (we have $\gamma_1 = \sqrt{3}$ and $\gamma_2 = 1$).

The above theorem studies the weighted inequalities for $p \in [1, 2]$. What about p larger than 2? What is a little unexpected, then (3) does not hold, even if X is assumed to have continuous trajectories.

Theorem 1.2. *Let $p > 2$. Then for any c there is a weight W and a continuous-path martingale X such that*

$$||[X]_\infty^{1/2}||_{L^p(W)} > c||X^*||_{L^p(W^*)}.$$

The remainder of this paper is organized as follows. Theorem 1.1 is established in the next section. The counterexamples described in Theorem 1.2 are constructed in the final part of this note.

For clarity, we should stress here that throughout the paper, we impose the continuity condition on the paths of W, while X will only be assumed to be càdlàg.

2. Proof of Theorem 1.1

Since the constants $\gamma_1 = \sqrt{3}$ and $\gamma_2 = 1$ are already the best possible in the unweighted setting (see Burkholder (2002)), all we need is to establish the estimate (3). In the proof of this inequality we will exploit the properties of a certain special function of five variables. For a fixed positive integer d, introduce the domain

$$\mathcal{D} = \{(x, y, z, w, v) \in \mathbb{R}^d \times [0, \infty) \times (0, \infty) \times [0, \infty)^2 : |x| \le z,\ w \le v\}.$$

Fix $p \in [1, 2]$ and consider the function $U_p : \mathcal{D} \to \mathbb{R}$ given by

$$U_p(x, y, z, w, v) = \frac{yw - |x|^2 v - (\gamma_p^2 - 1)z^2 v}{z^{2-p}}.$$

Let us establish two important technical facts about this object.

Lemma 2.1. *(i) For any $x \in \mathbb{R}^d \setminus \{0\}$ and any $w \ge 0$ we have*

$$U_p(x, |x|^2, |x|, w, w) \le 0. \tag{4}$$

(ii) For any $(x, y, z, w, v) \in \mathcal{D}$ we have

$$U_p(x, y, z, w, v) \ge \frac{2\gamma_p^{2-p}}{p}\big(y^{p/2} w - \gamma_p^p z^p v\big). \tag{5}$$

Proof. The inequality (4) is clear: we have $U_p(x, |x|^2, |x|, w, w) = -(\gamma_p^2 - 1)|x|^p w \le 0$. The inequality (5) reads

$$yw - |x|^2 v - (\gamma_p^2 - 1)z^2 v \ge \frac{2\gamma_p^{2-p}}{p}\left(y^{p/2} z^{2-p} w - \gamma_p^p z^2 v\right).$$

By continuity, we may assume that $w > 0$. Furthermore, the left-hand side is a nonincreasing function of $|x|$, while the right-hand side does not depend on $|x|$; so, it is enough to show the bound for $|x| = z$. If we substitute $y = \gamma_p^2 z^2 y'$ and $t = v/w \ge 1$, the inequality becomes

$$y' - t - \frac{2}{p}\left((y')^{p/2} - t\right) \ge 0.$$

But $2/p \ge 1$, so

$$y' - t - \frac{2}{p}\left((y')^{p/2} - t\right) \ge y' - 1 - \frac{2}{p}\left((y')^{p/2} - 1\right) \ge 0,$$

where the latter bound follows from the mean-value property for the concave function $s \mapsto s^{p/2}$.

Lemma 2.2. *For any* $(x, y, z, w, v) \in \mathcal{D}$ *and any* $h \in \mathbb{R}^d$ *we have*

$$U_p(x+h, |y|^2 + |h|^2, |x+h| \vee z, w, v) \leq U_p(x, y, z, w, v) - \frac{2\langle x, h\rangle v}{z^{2-p}}. \quad (6)$$

Proof. If $|x+h| \leq z$, then both sides are equal. If $|x+h| > z$, then the inequality can be rewritten in the equivalent form

$$\begin{aligned}&(y + |h|^2)wz^{2-p} - \gamma_p^2|x+h|^2vz^{2-p}\\ &\quad \leq yw|x+h|^{2-p} - (|x|^2 + 2\langle x, h\rangle)v|x+h|^{2-p} - (\gamma_p^2 - 1)z^2v|x+h|^{2-p}.\end{aligned}$$

Since $|x+h|^{2-p} \geq z^{2-p}$, it is enough to show the bound for $y = 0$. Then the right-hand side does not depend on w, while the left-hand side is a nondecreasing function of this variable; hence we may restrict ourselves to the largest value of w, i.e., $w = v$. After these two reductions, the inequality becomes

$$\left(|h|^2 - \gamma_p^2|x+h|^2\right) z^{2-p} \leq \left[-\left(|x|^2 + 2\langle x, h\rangle\right) - (\gamma_p^2 - 1)z^2\right] |x+h|^{2-p},$$

or, after the substitution $t = |x+h|/z$,

$$\gamma_p^2 \geq \frac{|h|^2z^{-2}(1 - t^{2-p}) + t^{2-p}(t^2 - 1)}{t^2 - t^{2-p}}.$$

But $|h|^2z^{-2} \geq (|x+h| - |x|)^2z^{-2} \geq (t-1)^2$ (since $|x| \leq z$), so we will be done if we show that

$$\gamma_p^2 \geq \frac{(t-1)^2(1 - t^{2-p}) + t^{2-p}(t^2 - 1)}{t^2 - t^{2-p}} = 1 + \frac{(2t-1)(1 - t^{p-2})}{t^p - 1}.$$

But this follows directly from the definition of the constant γ_p.

In the proof of (3), we will also require the following basic fact concerning stochastic processes (cf. Dellacherie & Meyer (1982)). Namely, for any semimartingale X there exists a unique continuous local martingale part X^c of X satisfying

$$[X]_t = |X_0|^2 + [X^c]_t + \sum_{0<s\leq t} |\triangle X_s|^2$$

for all $t \geq 0$. Here $|\triangle X_s| = X_s - X_{s-}$ denotes the jump of X at time s. Furthermore, we have $[X^c] = [X]^c$, where the expression on the right is the pathwise continuous part of $[X]$.

Proof of (3). Let us start with some reductions. First, it is enough to show the estimate for $\mathbb{R}^d$-valued martingales X, where d is an arbitrary positive integer (more precisely, to show the claim for martingales taking values in the finite-dimensional subspace spanned by the first d coordinates of $\mathcal{H}$); indeed, then letting $d \to \infty$ yields the desired statement in full generality. Next, we may assume that X takes values in $\mathbb{R}^d \setminus \{0\}$; indeed, otherwise we fix a small $\varepsilon > 0$ and consider the martingale $\tilde{X} = (X, \varepsilon)$; this process takes values in

$\mathbb{R}^{d+1} \setminus \{0\}$, and having established (3) for $\tilde{X}$, we recover the estimate for the initial process X by letting $\varepsilon \to 0$.

The proof rests on application of Itô's formula to the composition of U_p and the process $A_t = (X_t, [X]_t, X_t^*, W_t, W_t^*)$, $t \geq 0$. Some remarks are in order. The above reduction to finite-dimensional subspace of $\mathcal{H}$ guarantees that the application of Itô's formula is permitted (for some problems arising in the context of processes taking values in infinite-dimensional spaces, see Métivier (1982)). Next, the assumption $X \in \mathbb{R}^d \setminus \{0\}$ implies that $X^* > 0$ almost surely and hence A takes values in the domain of U_p. Finally, note that the formula for U_p makes sense on the whole $\mathbb{R}^d \times \mathbb{R} \times (0, \infty) \times \mathbb{R}^2$ and defines a C^∞ function there; hence, in particular, U_p has the required regularity.

So, for any $t \geq 0$ we have

$$U_p(A_t) = I_0 + I_1 + I_2 + I_3 + I_4,$$

where

$$\begin{aligned}
I_0 &= U_p(A_0), \\
I_1 &= \int_0^t U_{px}(A_{s-}) \cdot \mathrm{d}X_s + \int_0^t U_{pw}(A_{s-}) \mathrm{d}W_s, \\
I_2 &= \int_0^t U_{py}(A_{s-}) \mathrm{d}[X]_s^c + \frac{1}{2} \sum_{j,k=1}^{d} \int_0^t U_{px_jx_k}(A_{s-}) \mathrm{d}[X^j, X^k]_s^c, \\
I_3 &= \int_0^t U_{pz}(A_{s-}) \mathrm{d}(X_s^*)^c + \int_0^t U_{pv}(A_{s-}) \mathrm{d}W_s^*, \\
I_4 &= \sum_{0<s\leq t} \Big[U_p(A_s) - U_p(A_{s-}) - \langle U_{px}(A_{s-}), \Delta X_s \rangle \Big]
\end{aligned} \tag{7}$$

and U_{px} denotes the gradient of U_p with respect to the variables x_1, x_2, $\ldots$, x_d. All the other second-order terms vanish, since the processes $[X]$, X^*, W^* are nondecreasing or the appropriate partial derivatives are zero. Let us handle the terms I_0 through I_4 separately. By (4), we see that $I_0 = U_p(A_0) = U_p(X_0, |X_0|^2, |X_0|, W_0, W_0) \leq 0$. The process I_1 is a local martingale starting from 0, by the properties of stochastic integrals. The term I_2 equals

$$\int_0^t W_{s-}(X_{s-}^*)^{p-2} - W_{s-}^*(X_{s-}^*)^{p-2} \mathrm{d}[X]_s^c \leq 0.$$

Both integrals in I_3 are nonpositive: indeed, for $(x, y, z, w, v) \in \mathcal{D}$ we have

$$U_{pv}(x, y, z, w, v) = -\frac{|x|^2 + (\gamma_p^2 - 1)z^2}{z^{2-p}} \leq 0$$

and

$$\begin{aligned}
U_{pz}(x, y, z, w, v) &= (p-2)ywz^{p-3} + (2-p)|x|^2wz^{p-3} - p(\gamma_p^2 - 1)vz^{p-1} \\
&\leq (2-p)|x|^2wz^{p-3} - p(\gamma_p^2 - 1)vz^{p-1} \\
&\leq (2-p)vz^{p-1} - p(\gamma_p^2 - 1)vz^{p-1} = pvz^{p-1}\left(\frac{2}{p} - \gamma_p^2\right) \leq 0.
\end{aligned}$$

To see the latter bound, note that

$$\gamma_p^2 \geq 1 + \lim_{t \downarrow 1} \frac{(2t-1)(1-t^{p-2})}{t^p - 1} = \frac{2}{p}.$$

Finally, each summand appearing in I_4 is nonpositive: this follows directly from the estimate (6). Putting all the above facts together, we see that if $(\tau_n)_{n\geq 1}$ is a localizing sequence for the local martingale I_1, then

$$\mathbb{E}U_p(A_{\tau_n \wedge t}) \leq 0, \qquad n = 1, 2, \ldots.$$

Combining this with (5) yields $\mathbb{E}[X]_{\tau_n\wedge t}^{p/2} W_{\tau_n \wedge t} - \gamma_p^p \mathbb{E}(X_{\tau_n\wedge t}^*)^p W_{\tau_n\wedge t}^* \leq 0$ and hence

$$\mathbb{E}[X]_{\tau_n\wedge t}^{p/2} W_{\tau_n\wedge t} \leq \gamma_p^p \mathbb{E}(X^*)^p W^*.$$

But W is a uniformly integrable martingale, so $\mathbb{E}[X]_{\tau_n\wedge t}^{p/2} W_{\tau_n\wedge t} = \mathbb{E}[X]_{\tau_n\wedge t}^{p/2} W_\infty$, by Doob's optional sampling theorem. Letting $n \to \infty$ and applying Lebesgue's monotone convergence theorem, we get

$$\mathbb{E}[X]_\infty^{p/2} W_\infty \leq \gamma_p^p \mathbb{E}(X^*)^p W^*,$$

which is exactly the claim.

3. Lack of weighted bounds for $p > 2$

Let B be a standard Brownian motion starting from 0 and let N be a fixed positive integer. Consider the stopping times

$$\tau_k = \inf\{t : B_t \in \{-1, 2^k - 1\}\}, \qquad k = 0, 1, 2, \ldots.$$

Clearly, the process $W = (1 + B_{\tau_N \wedge t})_{t\geq 0}$ is a mean-one weight bounded by 2^N. Therefore, $\mathbb{P}(W^* \geq \lambda) = 1$ if $\lambda \leq 1$, and $\mathbb{P}(W^* \geq \lambda) = 0$ if $\lambda > 2^N$. Furthermore, if $\lambda \in (1, 2^N]$, elementary properties of Brownian motion imply that

$$\mathbb{P}(W^* \geq \lambda) = \mathbb{P}(B \text{ reaches } \lambda - 1 \text{ before it visits } -1) = \frac{1}{\lambda}.$$

Consequently, we see that

$$\mathbb{E}W^* = \int_0^\infty \mathbb{P}(W^* \geq \lambda)\mathrm{d}\lambda = 1 + N \ln 2.$$

Now, consider the predictable process H given by

$$H_t = \sum_{k=1}^N (-1)^k 2^{1-k} 1_{(\tau_{k-1}, \tau_k]}(t), \qquad t \geq 0,$$

and define X as the stochastic integral

$$X_t = 1 + \int_0^t H_s \mathrm{d}B_s = 1 + \sum_{k=1}^N (-1)^k 2^{1-k}(B_{\tau_k\wedge t} - B_{\tau_{k-1}\wedge t}), \quad t \geq 0.$$

Clearly, X is a martingale. Let us gain some intuition about its behavior. It starts from 1 and if $t \in (\tau_0, \tau_1]$, then $X_t = 1 - B_t \in [0,2]$ and $X_{\tau_1} \in \{0,2\}$ almost surely. Now if $X_{\tau_1} = 2$, then $B_{\tau_1} = -1$, which means that $\tau_1 = \tau_2 = \tau_3 = \ldots = \tau_N$ and the process terminates; on the other hand, if $X_{\tau_1} = 0$, then the process continues its evolution: if $t \in (\tau_1, \tau_2]$, then $X_t = (B_t - B_{\tau_1})/2 = (B_t - 1)/2 \in [-1,1]$, and $X_{\tau_2} \in \{-1,1\}$. Now if $X_{\tau_2} = -1$, then $B_{\tau_2} = -1$ and the process stops; but if $X_{\tau_2} = 1$, then $X_t = 1 - (B_t - B_{\tau_2})/4$ on $(\tau_2, \tau_3]$. Therefore, X behaves on $(\tau_2, \tau_3]$ in the same manner as on the interval $(\tau_0, \tau_1]$ (in the sense that it moves between $[0,2]$, finishing at one of the endpoints of this interval), but it goes four times slower. Now, if $X_{\tau_3} = 2$, then the evolution stops, while for $X_{\tau_3} = 0$ we have $X_t \in [-1,1]$ for $t \in (\tau_3, \tau_4]$ (again, the movement is similar to that two steps earlier, the only difference is that now the process moves four times slower). The pattern is then repeated: on each time interval $(\tau_k, \tau_{k+1}]$ the process X behaves as on $(\tau_0, \tau_1]$ or $(\tau_1, \tau_2]$ (depending on the parity of k), at an appropriately slower rate.

In particular, it follows from the above analysis that $\mathbb{P}(X^* \leq 2) = 1$, so

$$\mathbb{E}(X^*)^p W^* \leq 2^p(1 + N \ln 2). \tag{8}$$

Next, by the properties of stochastic integrals, the square bracket of X equals

$$[X] = 1 + \int_0^\infty H_s^2 \mathrm{d}s = 1 + \sum_{k=1}^N (\tau_k - \tau_{k-1}) 4^{1-k}.$$

By the very definition of W we have $\mathbb{P}(W_\infty \in \{0, 2^N\}) = 1$, so $\mathbb{E}[X]^p W_\infty = 2^N \mathbb{E}[X]^p 1_{\{W_\infty = 2^N\}}$. Thus, it is enough to understand the behavior of $[X]$ on the set $\{W_\infty = 2^N\}$. The above analysis combined with the independence of increments and self-similarity of B implies that conditionally on the set $\{W_\infty = 2^N\}$, the variables $\tau_1 - \tau_0$, $(\tau_2 - \tau_1)/4$, ..., $(\tau_N - \tau_{N-1})/4^{N-1}$ are independent and identically distributed as τ_1. Consequently, if ξ_1, ξ_2, ..., ξ_N are N independent copies of τ_1, then, by Jensen's inequality,

$$\begin{aligned}\mathbb{E}[X]_\infty^p W_\infty &= \mathbb{E}(\xi_1 + \xi_2 + \ldots + \xi_N)^{p/2} \\ &\geq (\mathbb{E}\xi_1 + \mathbb{E}\xi_2 + \ldots + \mathbb{E}\xi_N)^{p/2} = N^{p/2}(\mathbb{E}\xi_1)^{p/2}.\end{aligned}$$

Combining this with (8), we see that if c is an arbitrary constant, then

$$\mathbb{E}[X]_\infty^p W_\infty > c\mathbb{E}(X^*)^p W^*$$

provided N is sufficiently large. This shows that there is no weighted bound of the form (3) in the case $p > 2$.

Acknowledgments

The research was supported by NCN grant DEC-2014/14/E/ST1/00532.

References

Burkholder, D. L., 1966. Martingale transforms. Ann. Math. Statist. 37, 1494–1504.

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

Burkholder, D. L., 2002. The best constant in the Davis' inequality for the expectation of the martingale square function. Trans. Amer. Soc. 354, 91–105.

Davis, B., 1970. On the integrability of the martingale square function. Israel J. Math. 8, 187–190.

Dellacherie, C. and Meyer P.-A., 1982. Probabilities and potential B: Theory of martingales. North Holland, Amsterdam.

Fefferman, C. and Stein, E. M., 1971. Some maximal inequalities. Amer. J. Math. 93, 107–115.

Kazamaki, N., 1994. Continuous exponential martingales and BMO. Lecture Notes in Math. 1579, Springer-Verlag, Berlin, Heidelberg.

Khintchine, A., 1923. Über dyadische Brüche. Math. Z. 18, 109–116.

Littlewood, J. E., 1930. On bounded bilinear forms in an infinite number of variables. Quart. J. Math. Oxford 1, 164–174.

Marcinkiewicz, J., 1937. Quelques théorèmes sur les séries orthogonales. Ann. Soc. Polon. Math., 16, 84–96.

Métivier, M., 1982. Semimartingales: A Course on Stochastic Processes. de Gruyter, Berlin.

Osękowski, A., 2005. Two inequalities for the first moment of a martingale, its square and maximal function. Bull. Polish Acad. Sci. Math. 53, 441–449.

Osękowski, A., 2012. Sharp martingale and semimartingale inequalities, Monogr. Mat. 72, Birkhäuser Basel.

Paley, R. E. A. C., 1932. A remarkable series of orthogonal functions I. Proc. London Math. Soc. 34 , 241–264.

Pisier G. and Xu, Q., 1997. Non-commutative martingale inequalities. Commun. Math. Phys. 189, 667–698.