SHARP $L^2 \log L$ INEQUALITIES FOR THE HAAR SYSTEM AND MARTINGALE TRANSFORMS

ADAM OŚKOWSKI

Abstract. Let $(h_n)_{n \geq 0}$ be the Haar system of functions on $[0, 1]$. The paper contains the proof of the estimate

$$\int_0^1 \left| \sum_{k=0}^n \varepsilon_k a_k h_k \right|^2 \log \left| \sum_{k=0}^n \varepsilon_k a_k h_k \right| ds \leq \int_0^1 \left| \sum_{k=0}^n a_k h_k \right|^2 \log \left| e^2 \sum_{k=0}^n a_k h_k \right| ds,$$

for $n = 0, 1, 2, \ldots$. Here $(a_n)_{n \geq 0}$ is an arbitrary sequence with values in a given Hilbert space $H$ and $(\varepsilon_n)_{n \geq 0}$ is a sequence of signs. The constant $e^2$ appearing on the right is shown to be the best possible. This result is generalized to the sharp inequality

$$E|g_n|^2 \log |g_n| \leq E|f_n|^2 \log (e^2|f_n|), \quad n = 0, 1, 2, \ldots,$$

where $(f_n)_{n \geq 0}$ is an arbitrary martingale with values in $H$ and $(g_n)_{n \geq 0}$ is its transform by a predictable sequence with values in $\{-1, 1\}$. As an application, we obtain the two-sided bound for the martingale square function $S(f)$:

$$E|f_n|^2 \log (e^{-2}|f_n|) \leq ES_n^2(f) \log S_n(f) \leq E|f_n|^2 \log (e^2|f_n|), \quad n = 0, 1, 2, \ldots.$$

1. Introduction

Let $h = (h_n)_{n \geq 0}$ be the Haar system, i.e., the collection of functions given by

$h_0 = [0, 1], \quad h_1 = [0, 1/2) - [1/2, 1),$
$h_2 = [0, 1/4) - [1/4, 1/2), \quad h_3 = [1/2, 3/4) - [3/4, 1),$
$h_4 = [0, 1/8) - [1/8, 1/4), \quad h_5 = [1/4, 3/8) - [3/8, 1/2),$
$h_6 = [1/2, 5/8) - [5/8, 3/4), \quad h_7 = [3/4, 7/8) - [7/8, 1)$

and so on. Here we have identified a set with its indicator function. A classical result of Schauder [12] states that the Haar system forms a basis of $L^p = L^p(0, 1)$, $1 \leq p < \infty$ (throughout, the underlying measure will be the Lebesgue measure). Using an inequality of Paley [11], Marcinkiewicz [7] proved that the Haar system is an unconditional basis provided $1 < p < \infty$. That is, there is a universal finite constant $c_p$ such that

$$c_p^{-1} \left\| \sum_{k=0}^n a_k h_k \right\|_{L^p(0,1)} \leq \left\| \sum_{k=0}^n \varepsilon_k a_k h_k \right\|_{L^p(0,1)} \leq c_p \left\| \sum_{k=0}^n a_k h_k \right\|_{L^p(0,1)}$$

for any $n$ and any $a_k \in \mathbb{R}$, $\varepsilon_k \in \{-1, 1\}$, $k = 0, 1, 2, \ldots, n$. This result is a starting point for numerous extensions and applications: in particular, it has led to the development of the theory of singular integrals, stochastic integrals, stimulated the studies on the geometry of Banach spaces and has been extended to other areas of

2000 Mathematics Subject Classification. Primary: 60G42. Secondary: 26D15.

Key words and phrases. Haar system, martingale, square function, best constants.
mathematics. In particular, the inequality (1.1) has a natural counterpart in the martingale theory. Suppose that \((Ω, F, P)\) is a probability space, equipped with a nondecreasing sequence \((F_n)_{n \geq 0}\) of sub-\(σ\)-algebras of \(F\). Let \(f = (f_n)_{n \geq 0}\) be an adapted real-valued martingale and let \(df = (df_n)_{n \geq 0}\) stand for its difference sequence, given by \(df_0 = f_0\) and \(df_n = f_n - f_{n-1}\) for \(n \geq 1\). So, the differences \(df_n\) are \(F_n\)-measurable and integrable, and the martingale property amounts to saying that for each \(n \geq 1\), \(E(df_n|F_{n-1}) = 0\). Given a deterministic sequence \(ε = (ε_n)_{n \geq 0}\) of signs, we define \(g = (g_n)_{n \geq 0}\), the associated transform of \(f\), by

\[ g_n = \sum_{k=0}^{n} ε_k df_k, \quad n = 0, 1, 2, \ldots \]

Clearly, this is equivalent to saying that the difference sequence of \(g\) is given by \(dg_n = ε_n df_n\). Note that the sequence \(g = (g_n)_{n \geq 0}\) is again an adapted martingale. Actually, this is still true if we allow the following more general class of the transforming sequences. Namely, suppose that \(ε = (ε_n)_{n \geq 0}\) is a sequence of random signs. Then we have the estimate (1.1), and (1.2). These include the weak-type \((p, p)\) inequalities (cf. [3, 13]), exponential bounds ([6]), logarithmic estimates ([9]) and many others: see the monograph [10] for the detailed exposition on the subject. The purpose of this paper is to continue this line of research. Our main result is the following sharp \(L^2\) bound for the Haar system and martingale transforms.

**Theorem 1.1.** Let \(f\) be a martingale taking values in a Hilbert space \(H\) and let \(g\) be its transform by a predictable sequence of signs. Then we have the estimate

\[ \mathbb{E}|g_n|^2 \log |g_n| \leq \mathbb{E}|f_n|^2 \log (c^2 |f_n|), \quad n = 0, 1, 2, \ldots \]

This inequality is already sharp for the Haar system: for any \(κ < e^2\) there exists a positive integer \(n\), real numbers \(a_0, a_1, \ldots, a_n\) and signs \(ε_0, ε_1, \ldots, ε_n\) such that

\[ \int_0^1 \left| \sum_{k=0}^{n} ε_k a_k h_k \right|^2 \log \left| \sum_{k=0}^{n} ε_k a_k h_k \right| dx > \int_0^1 \left| \sum_{k=0}^{n} a_k h_k \right|^2 \log \left( \sum_{k=0}^{n} a_k h_k \right) dx. \]

Actually, it will be clear from the proof that the estimate (1.3) holds true for any martingales \(f, g\) satisfying the condition \(|df_n| = |dg_n|\) almost surely for all
Indeed, if $x \in S$ bound for the martingale square function $S(f) = (S_n(f))_{n \geq 0}$, defined by

$$S_n(f) = \left( \sum_{k=0}^{n} |df_k|^2 \right)^{1/2}, \quad n = 0, 1, 2, \ldots .$$

The result can be stated as follows.

**Theorem 1.2.** Let $f$ be a martingale taking values in a Hilbert space $\mathcal{H}$. Then for any $n = 0, 1, 2, \ldots$ we have

$$\mathbb{E}|f_n|^2 \log(e^{-2}|f_n|) \leq \mathbb{E}S_n^2(f) \log S_n(f) \leq \mathbb{E}|f_n|^2 \log(e^2|f_n|).$$

The left inequality is sharp: the constant $e^{-2}$ cannot be replaced by a larger number.

Unfortunately, we have been unable to identify the optimal constant in the right inequality (1.4), but some calculations show that it should not be far from $e^2$. We will not pursue farther in this direction.

A few words about the proof and the organization of the paper are in order. Our reasoning will rest on Burkholder’s method (cf. [3], [4], [6]): the estimate (1.3) will be deduced from the existence of a certain special function, enjoying appropriate majorization and concavity requirements. The proof of the estimates (1.3) and (1.4) can be found in the next section. Section 3 is devoted to the optimality of the constants in these inequalities.

2. Proofs of (1.3) and (1.4)

As we have announced in the preceding section, the proofs of the inequalities (1.3) and (1.4) will exploit the properties of a certain special function. Let $D = \mathcal{H} \times \mathcal{H} \setminus \{(x, y) : |x|/|y| = 0\}$ and consider $U : D \to \mathbb{R}$, given by

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

Though the definition makes perfect sense also in the case $|x|/|y| = 0$, we have decided to exclude these pairs $(x, y)$ from the domain of $U$; this will guarantee that the function $U$ is smooth, which in turn will allow us to avoid unpleasant technicalities. We will also need the auxiliary functions $\varphi, \psi : D \to \mathcal{H}$, defined by

$$\varphi(x, y) = -5x - 2x \log \frac{|x| + |y|}{2}, \quad \psi(x, y) = y - 2|x|y' + 2y \log \frac{|x| + |y|}{2},$$

where $y' = y/|y|$. Then we have $\varphi(x, y) = U_x(x, y)$ and $\psi(x, y) = U_y(x, y)$ on $D$. Indeed, if $x \neq 0$, then for any $h \in \mathcal{H}$ we have

$$\lim_{t \to 0} \frac{|x + th| - |x|}{t} = x' \cdot h,$$

where $\cdot$ denotes the scalar product in $\mathcal{H}$. Consequently,

$$\lim_{t \to 0} \frac{U(x + th, y) - U(x, y)}{t} = \left( -5|x| - 2|x| \log \frac{|x| + |y|}{2} \right) x' \cdot h = \varphi(x, y) \cdot h,$$

that is, $\varphi(x, y) = U_x(x, y)$. The identity $\psi(x, y) = U_y(x, y)$ is verified similarly.

The crucial properties of the above objects are studied in a lemma below.
**Lemma 2.1.** (i) For any \( x \in \mathcal{H} \setminus \{0\} \) we have
\[
U(x, \pm x) \leq 0.
\]
(ii) For any \((x, y) \in D\) we have the majorization
\[
U(x, y) \geq |y|^2 \log |y| - |x|^2 \log(e^2|x|).
\]
(iii) For any \( x, y, h, k \in \mathcal{H} \) such that \( |k| = |h| \), \((x, y) \in D\) and \((x+h, y+k) \in D\), we have
\[
U(x+h, y+k) \leq U(x, y) + \varphi(x, y) \cdot h + \psi(x, y) \cdot k.
\]

**Proof.** (i) This is evident: \( U(x, \pm x) = -2|x|^2 \leq 0 \).
(ii) Of course, it is enough to show the majorization for \( \mathcal{H} = \mathbb{R} \) and for positive \( x, y \) only (simply introduce the new variables \( x := |x| \) and \( y := |y| \)). Fix \( x > 0 \) and define \( F : (0, \infty) \to \mathbb{R} \) by the formula
\[
F(y) = y^2 - xy - 2x^2 + (y^2 - x^2) \log \frac{x+y}{2} - y^2 \log y + x^2 \log(e^2x).
\]
A straightforward differentiation yields
\[
F'(y) = 2y - x + 2y \log \frac{x+y}{2} - x \quad \text{and} \quad F''(y) = 2 \log \frac{x+y}{2y} + \frac{2y}{x+y}.
\]
Since \( a - 2 \log a > 0 \) for any \( a > 0 \), we see that \( F \) is a convex function. Furthermore, we have \( F'(x) = F(x) = 0 \); this shows that \( F \) is nonnegative, which is precisely the desired bound (2.2).

(iii) By continuity, it is enough to prove the assertion under the addition assumption that for each \( t \), both vectors \( x + th \) and \( y + tk \) are nonzero. To see this, simply pick a vector \( v \neq 0 \) orthogonal to the subspace generated by \( x, y, h \) and \( k \). Then the vectors \( \tilde{h} = h + v, \tilde{k} = k + v \) satisfy \( |\tilde{h}| = |\tilde{k}| \) and \( |x + t\tilde{h}|, |y + t\tilde{k}| \neq 0 \) for all \( t \); having proved (2.3) for \( x, y, \tilde{h} \) and \( \tilde{k} \), we let \( v \to 0 \) to obtain the claim in the general case.

Now we apply a well-known procedure of proving the inequality (2.3) (see e.g. [6]). Namely, for a fixed \( x, y, h, k \) as above, introduce the function
\[
G(t) = G_{x,y,h,k}(t) = U(x + th, y + tk), \quad t \in \mathbb{R}.
\]
Then (2.3) is equivalent to \( G(1) \leq G(0) + G'(0) \), and hence we will be done if we show that \( G \) is concave. The assumption \( |x + th||y + tk| \neq 0 \) guarantees that the function \( G \) is twice differentiable and hence we must prove that \( G''(t) \leq 0 \) for all \( t \in \mathbb{R} \). Actually, it is enough to consider the case \( t = 0 \) only, because of the translation property \( G_{x,y,h,k}(s+t) = G_{x+sh,y+sk,h,k}(t) \). A little tedious calculation gives
\[
G''(0) = \frac{2|x|}{|y|} \left[ -|k|^2 + (y' \cdot k)^2 \right] + 2 \left[ -|h|^2 - (x'\tilde{h})(y'\tilde{k}) \right] \leq 0,
\]
since both expressions in the square brackets are nonpositive. This completes the proof of the lemma.

**Proof of (1.3).** Pick two adapted martingales \( f, g \) satisfying the condition \( |df_k| = |dg_k| \) for any \( k \geq 0 \) and fix a nonnegative integer \( n \). By adding a small vector \( v \), orthogonal to the ranges of \( f \) and \( g \) (as in the proof of Lemma 2.1 (iii) above), we may assume that \( |fx_k||gk| \neq 0 \) with probability 1 for all \( k \geq 0 \). For the sake of clarity, it is convenient to split the reasoning into two parts.
Step 1. Integrability conditions. If \( \mathbb{E}[|f_n|^2 \log(e^2|f_n|)] = \infty \), then there is nothing to prove. Suppose then that \( \mathbb{E}[|f_n|^2 \log(e^2|f_n|)] < \infty \); then also \( \mathbb{E}[|f_n|^2 \log_+(e^2|f_n|)] < \infty \), and since the function \( \Phi(t) = t^2 \log_+ |t| \) is convex on \( \mathbb{R} \), we conclude that

\[ \mathbb{E}[|f_k|^2 \log_+(e^2|f_k|)] < \infty \quad \text{for all} \quad k \leq n. \]

This in turn implies that for each \( k \leq n \) we have \( \mathbb{E}[|df_k|^2 \log_+ |df_k|] = \mathbb{E}[|df_k|^2 \log_+ |df_k|] < \infty \), due to the simple pointwise bound

\[ |f_k - f_{k-1}|^2 \log_+ |f_k - f_{k-1}| \leq 4|f_k|^2 \log_+(2|f_k|) + 4|f_{k-1}|^2 \log_+(2|f_{k-1}|). \]

This finally gives that \( \mathbb{E}|g_k|^2 \log |g_k| < \infty \) for all \( k \leq n \): apply the estimate

\[
\mathbb{E} \left[ \sum_{k=0}^{k} \log \sum_{\ell=0}^{k} |dg_{\ell}| \right] \leq \mathbb{E} \left( \max_{0 \leq \ell \leq k} |dg_{\ell}| \right)^2 \log_+ \left( \max_{0 \leq \ell \leq k} |dg_{\ell}| \right) \leq k^2 \mathbb{E} \sum_{\ell=0}^{k} |dg_{\ell}|^2 \log_+(k|dg_{\ell}|).
\]

Step 2. Proof of the \( L^2 \log L \) inequality. The key observation is that the sequence \( (U(f_k, g_k))_{k=0}^{n-1} \) is a supermartingale. Indeed, its integrability follows easily from the facts proved in the preceding step, and the supermartingale property follows from the part (iii) of Lemma 2.1. To see this, fix 0 ≤ k < n and note that

\[
U(f_{k+1}, g_{k+1}) = U(f_k + df_{k+1}, g_k + dg_{k+1}) \\
\leq U(f_k, g_k) + \varphi(f_k, g_k) \cdot df_{k+1} + \psi(f_k, g_k) \cdot dg_{k+1}.
\]

Applying to both sides the conditional expectation with respect to \( \mathcal{F}_k \) yields the desired bound \( \mathbb{E}[U(f_{k+1}, g_{k+1}) | \mathcal{F}_k] \leq U(f_k, g_k) \). Thus, by (2.1) and (2.2), we get

\[
0 \geq \mathbb{E}U(f_0, g_0) \geq \mathbb{E}U(f_n, g_n) \geq \mathbb{E}|g_n|^2 \log |g_n| - \mathbb{E}|f_n|^2 \log(e^2|f_n|),
\]

which is (1.3).

Proof of (1.4). Consider the Hilbert space \( \mathbb{H} = L^2(\mathcal{H}) \). We can treat an \( \mathcal{H} \)-valued martingale \( f \) as an \( \mathbb{H} \)-valued sequence, embedding it onto the first coordinate: \( f_n \sim (f_0, 0, 0, \ldots) \in \mathbb{H} \). To handle the square function, consider the martingale \( g_n = (df_0, df_1, df_2, \ldots, df_n, 0, 0, \ldots), n = 0, 1, 2, \ldots \). Then \( |df_n|_{\mathbb{H}} = |dg_n|_{\mathbb{H}} \) with probability 1 and hence, by the estimate (1.3) just established above,

\[
\mathbb{E}[|f_n|^2 \log(e^2|f_n|)] \leq \mathbb{E}|g_n|^2 \log |g_n|_{\mathbb{H}} \leq \mathbb{E}[|f_n|^2 \log(e^2|f_n|)], \quad n = 0, 1, 2, \ldots.
\]

It remains to observe that \( |g_n|_{\mathbb{H}} = S_n(f) \) for all \( n \). This proves the inequality.

3. Sharpness

We turn our attention to the optimality of the constant \( e^2 \) in the \( L^2 \log L \) inequality for the Haar system and the martingale square function. One could study this problem by constructing appropriate examples, but we have chosen a different path, which is of its own interest and connections with boundary value problems.

We start with the inequality for the Haar system. Suppose that for some constant \( \kappa > 0 \) we have

\[
(3.1) \quad \int_0^1 \left[ \sum_{k=0}^{n} \varepsilon_k a_k h_k \right]^2 \log \left[ \sum_{k=0}^{n} \varepsilon_k a_k h_k \right] ds \leq \int_0^1 \left[ \sum_{k=0}^{n} a_k h_k \right]^2 \log \left[ \kappa \sum_{k=0}^{n} a_k h_k \right] ds.
\]
for $n = 0, 1, 2, \ldots$. Consider the functions $V_n, W_n$ on $\mathbb{R}^2$, given by $V_n(x,y) = |y|^2 \log |y| - |x|^2 \log (\kappa|x|)$ and

$$(3.2) \quad W_n(x,y) = \sup \left\{ \int_0^1 V \left( x + \sum_{k=1}^n a_k h_k(s), y + \sum_{k=1}^n \varepsilon_k a_k h_k(s) \right) ds \right\},$$

where the supremum is taken over all positive integers $n$ and all sequences $a_1, a_2, \ldots, a_n \in \mathbb{R}, \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n \in \{-1, 1\}$.

**Lemma 3.1.** The function $W_n$ has the following properties.

1° We have $W_n \geq V_n$ on $\mathbb{R}^2$.

2° The function $W_n$ is concave along the lines of slope ±1.

3° We have $W_n(x,y) < \infty$ for all $x, y$.

4° For any $x, y \in \mathbb{R}$ and any $\lambda > 0$ we have the homogeneity-type property

$$(3.3) \quad W_n(\lambda x, \lambda y) = \lambda^2 W_n(x,y) + \lambda^2 \log \lambda (|y|^2 - |x|^2).$$

**Proof.** The property 1° is evident; it suffices to consider the sequence $a_1 = a_2 = \ldots = a_n = 0$ in the definition of $W_n$. To show 2°, we use the so-called “splicing” argument (see e.g. page 77 in Burkholder [5]). To be more precise, fix a line $L$ of slope 1, a point $(x,y)$ lying on it and a positive number $d$. Pick two positive integers $n, m$, and some arbitrary sequences $a_1^+, a_2^+, \ldots, a_n^+, a_1^-, a_2^-, \ldots, a_m^- \in \mathbb{R}$ and $\varepsilon_1^+, \varepsilon_2^+, \ldots, \varepsilon_n^+, \varepsilon_2^-, \ldots, \varepsilon_m^- \in \{-1, 1\}$. Let us splice the pairs of functions

$$\varphi^+ = x + d + \sum_{k=1}^n a_k^+ h_k, \quad \varphi^- = x - d + \sum_{k=1}^n a_k^- h_k$$

and

$$\psi^+ = y + d + \sum_{k=1}^n \varepsilon_k^+ a_k^+ h_k, \quad \psi^- = y - d + \sum_{k=1}^n \varepsilon_k^- a_k^- h_k$$

into one pair of functions, setting

$$(3.4) \quad (\varphi(r), \psi(r)) = \begin{cases} (\varphi^-(2r), \psi^-(2r)) & \text{if } r < 1/2, \\ (\varphi^+(2r), \psi^+(2r)) & \text{if } r \geq 1/2. \end{cases}$$

It is evident from the structure of the Haar system that the splice $(\varphi, \psi)$ is given by the finite sums of the form $(x - dh_1 + \sum_{k=2}^N a_k h_k, y - dh_1 + \sum_{k=2}^N \varepsilon_k a_k h_k)$, where each number $a_k$ coincides with an appropriate coefficient of $\varphi^-$ or $\varphi^+$, depending on whether the support of $h_k$ is contained in the left or the right half of the interval $[0,1]$ and, similarly, $\varepsilon_k$ is an appropriate sign coming from $\varphi^-$ or $\varphi^+$. Consequently, we may write

$$W_n(x,y) \geq \int_0^1 V(\varphi(s), \psi(s)) \, ds$$

$$= \frac{1}{2} \left[ \int_0^1 V(\varphi^-(s), \psi^-(s)) \, ds + \int_0^1 V(\varphi^+(s), \psi^+(s)) \, ds \right],$$

and taking the supremum over all $\varphi^-, \varphi^+$ yields

$$W_n(x,y) \geq (W_n(x-d, y-d) + W_n(x+d, y+d))/2.$$
To show $3^\circ$, we first apply (3.1) to obtain that $U(x, \pm x) \leq 0$ for all $x \in \mathbb{R}$. Now the finiteness of $U$ follows from the concavity along the lines of slope $\pm 1$ we have just established.

Finally, let us handle $4^\circ$. Pick an arbitrary positive number $n$ and some sequences $a_1, a_2, \ldots, a_n \in \mathbb{R}$, $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n \in \{-1, 1\}$. We have

$$W_\kappa(\lambda x, \lambda y) \geq \int_0^1 V \left( \lambda x + \sum_{k=1}^n \lambda a_k h_k(s), \lambda y + \sum_{k=1}^n \varepsilon_k \lambda a_k h_k(s) \right) \, ds$$

$$= \lambda^2 \int_0^1 V \left( \lambda x + \sum_{k=1}^n a_k h_k(s), y + \sum_{k=1}^n \varepsilon_k a_k h_k(s) \right) \, ds$$

$$+ \lambda^2 \log \lambda \int_0^1 \left[ \left( y + \sum_{k=1}^n \varepsilon_k a_k h_k(s) \right)^2 - \left( x + \sum_{k=1}^n a_k h_k(s) \right)^2 \right] \, ds.$$

However, the latter integral is equal to $|y|^2 - |x|^2$, by orthogonality of the Haar system. Taking the supremum over all $n$ and sequences $a_1, a_2, \ldots, a_n, \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ yields the estimate

$$W_\kappa(\lambda x, \lambda y) \geq \lambda^2 W_\kappa(x, y) + \lambda^2 \log \lambda (|y|^2 - |x|^2).$$

To prove the reverse bound, rewrite it in the equivalent form

$$W_\kappa \left( \frac{1}{\lambda} \lambda x, \frac{1}{\lambda} \lambda y \right) \geq \frac{1}{\lambda^2} W_\kappa(\lambda x, \lambda) + \frac{1}{\lambda^2} \log \frac{1}{\lambda} (\lambda y^2 - \lambda x^2),$$

which follows at once from the estimate we have just established.

Equipped with the above function $W_\kappa$, the optimality of the constant $e^2$ follows easily. Namely, we know that the number $W_\kappa(0, 1)$ is finite. Furthermore, applying the property $2^\circ$ twice gives

$$W_\kappa(0, 1) \geq \frac{1}{1 + 2\delta} W_\kappa(\delta, 1 + \delta) + \frac{2\delta}{1 + 2\delta} W_\kappa(-1/2, -1/2)$$

$$\geq \frac{1}{(1 + 2\delta)^2} W_\kappa(0, 1 + 2\delta) + \frac{2\delta}{(1 + 2\delta)^2} W_\kappa(1/2 + \delta, 1/2 + \delta)$$

$$+ \frac{2\delta}{1 + 2\delta} W_\kappa(-1/2, -1/2).$$

However, by $4^\circ$, we have

$$W_\kappa(0, 1 + 2\delta) = (1 + 2\delta)^2 W_\kappa(0, 1) + (1 + 2\delta)^2 \log (1 + 2\delta)$$

and, by the majorization $1^\circ$,

$$W(-1/2, -1/2) \geq -\log \kappa/4, \quad W_\kappa(1/2 + \delta, 1/2 + \delta) \geq -(1/2 + \delta)^2 \log \kappa.$$

Combining these facts with the preceding estimate yields an inequality which is equivalent to

$$\frac{2 + \delta}{2 + 2\delta} \log \kappa \geq \log \left( \frac{1 + 2\delta}{\delta} \right).$$

Letting $\delta \to 0$ we obtain $\kappa \geq e^2$. This shows that the constant $e^2$ is indeed the best possible.
The $L^2 \log L$ estimate for the square function can be handled similarly. Suppose that $\kappa > 0$ is a constant such that
\[
\mathbb{E}|f_n|^2 \log |f_n| \leq \mathbb{E}S_n^2(f) \log(\kappa S_n(f)), \quad n = 0, 1, 2, \ldots.
\]
Introduce the function $W_\kappa$ on $[0, \infty) \times \mathbb{R}$ by the formula
\[
W_\kappa(x, y) = \mathbb{E}_{V_\kappa} \left( \sqrt{x^2 - y^2 + S_n^2(f)}, f_n \right),
\]
where the supremum is taken over all $n$ and all simple martingales satisfying $f_0 \equiv y$ (a martingale is called simple if for any $n$ the variable $f_n$ takes only a finite number of values). Here, as previously, $V_\kappa(x, y) = |y|^2 \log |y| - |x|^2 \log(\kappa|x|)$.

The somewhat odd expression $\sqrt{x^2 - y^2 + S_n^2(f)}$ guarantees that the sequence $(\sqrt{x^2 - y^2 + S_n^2(f)})_{n\geq0}$ starts from $x$. An analogous reasoning to that presented above (see also Chapter 11 of [6]) yields the following statement. We omit the proof, leaving it to the interested reader.

**Lemma 3.2.** The function $W_\kappa$ has the following properties.

1° We have $W_\kappa \geq V_\kappa$ on $[0, \infty) \times \mathbb{R}$.

2° For any $(x, y) \in [0, \infty) \times \mathbb{R}$, any $\alpha \in (0, 1)$ and any $t_1, t_2 \in \mathbb{R}$ satisfying $\alpha t_1 + (1 - \alpha) t_2 = 0$, we have
\[
W_\kappa(x, y) \geq \alpha W_\kappa \left( \sqrt{x^2 + t_1^2}, y + t_1 \right) + (1 - \alpha) W_\kappa \left( \sqrt{x^2 + t_2^2}, y + t_2 \right).
\]

3° We have $W_\kappa(x, y) < \infty$ for all $x > 0$ and $y \in \mathbb{R}$.

4° For any $x, y \in \mathbb{R}$ and any $\lambda > 0$ we have the homogeneity-type property (3.3).

Equipped with this lemma, we are ready to show that the constant $\kappa$ must be at least $e^2$. Fix a parameter $\gamma > 1$ and apply the property 2° to obtain
\[
W_\kappa(\gamma^{-1}, 1) \geq \left( \frac{\gamma^2 - 1}{\gamma^2 + 1} \right)^2 W_\kappa \left( \gamma^{-1}, \frac{\gamma^2 + 1}{\gamma^2 - 1}, \frac{\gamma^2 - 1}{\gamma^2 + 1} \right) + \frac{4\gamma^2}{(\gamma^2 + 1)^2} W_\kappa \left( \frac{1 + \gamma^{-1}}{2}, \frac{1 + \gamma^{-1}}{2} \right).
\]

Next, using the homogeneity (3.3) and the majorization 1°, we get
\[
W_\kappa \left( \frac{\gamma^{-1} \gamma^2 + 1}{\gamma^2 - 1}, \frac{\gamma^2 + 1}{\gamma^2 - 1} \right) = \left( \frac{\gamma^2 + 1}{\gamma^2 - 1} \right)^2 W_\kappa(\gamma^{-1}, 1)
\]
\[
\quad + \left( \frac{\gamma^2 + 1}{\gamma^2 - 1} \right)^2 \log \left( \frac{\gamma^2 + 1}{\gamma^2 - 1} \right) (1 - \gamma^{-2}),
\]
and
\[
W_\kappa \left( \frac{1 + \gamma^{-1}}{2}, \frac{1 + \gamma^{-1}}{2} \right) \geq -\frac{1}{4} (1 + \gamma^{-2})^2 \log \kappa.
\]

Plugging these two facts into the preceding estimate and working a little bit, we arrive at the inequality equivalent to
\[
\log \kappa \geq (\gamma^2 - 1) \log \left( \frac{\gamma^2 + 1}{\gamma^2 - 1} \right).
\]

It remains to let $\gamma \to \infty$ to obtain $\log \kappa \geq 2$ or $\kappa \geq e^2$, as desired.
Acknowledgements

The author would like to thank an anonymous referee for the careful reading of the paper and several helpful suggestions. The research was partially supported by NCN grant DEC-2012/05/B/ST1/00412.

References


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

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