Bernoulli **26**(3), 2020, 1912–1926
https://doi.org/10.3150/19-BEJ1175

On the best constant in the martingale version of Fefferman's inequality

ADAM OSĘKOWSKI

Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland. E-mail: ados@mimuw.edu.pl

Let $X = (X_t)_{t\geq 0} \in H^1$ and $Y = (Y_t)_{t\geq 0} \in \mathrm{BMO}$ be arbitrary continuous-path martingales. The paper contains the proof of the inequality

$$\mathbb{E}\int_0^\infty \left|d\langle X, Y\rangle_t\right| \leq \sqrt{2}\|X\|_{H^1}\|Y\|_{\mathrm{BMO}_2},$$

and the constant $\sqrt{2}$ is shown to be the best possible. The proof rests on the construction of a certain special function, enjoying appropriate size and concavity conditions.

Keywords: BMO; best constants; duality; martingale; maximal

1. Introduction

A classical result of Fefferman asserts that the space BMO, the class of functions of bounded mean oscillation, can be identified with the dual space of H^1. This fact was originally obtained in Fefferman [4] in the analytic setting, and was later extended to the probabilistic context by Getoor and Sharpe [5]. The purpose of this paper is to provide the best constant in the corresponding martingale inequality.

We start with the introduction of the basic notion and notation. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, equipped with a right-continuous filtration $(\mathcal{F}_t)_{t\geq 0}$ such that $\mathcal{F}_0$ contains all the sets of probability zero. Throughout, we assume that any local martingale adapted to this filtration is continuous: for example, it is well known that the Brownian filtration enjoys this property. Furthermore, for technical reasons, we will assume that the probability space is not too small: there exists a Brownian motion on it. For any adapted martingale $X = (X_t)_{t\geq 0}$, we denote the corresponding square bracket by $\langle X, X\rangle$: see Dellacherie and Meyer [3] for the definition. The maximal function of X is given by $X^* = \sup_{t\geq 0}|X_t|$ and we use the notation $X_t^* = \sup_{s\in[0,t]}|X_s|$ (where $t \geq 0$) for the corresponding truncated version. The Hardy norm of X is given by $\|X\|_{H^1} = \mathbb{E}X^*$, and the space H^1 is defined as the class of all martingales whose Hardy norms are finite. The space H^1 is a Banach space, and the aforementioned result of Fefferman, Getoor and Sharpe asserts that its dual is equal to BMO, the definition of which we now recall. Following Getoor and Sharpe [5], given $1 \leq p < \infty$, a uniformly integrable martingale $Y = (Y_t)_{t\geq 0}$ belongs to the class BMO_p, if

$$\|Y\|_{\mathrm{BMO}_p} = \sup_{t\geq 0}\left\|\mathbb{E}\left[|Y_\infty - Y_t|^p|\mathcal{F}_t\right]^{1/p}\right\|_\infty < \infty.$$

1350-7265

It can be shown that all the seminorms $\|\cdot\|_{\mathrm{BMO}_p}$ are equivalent, and hence all the classes BMO_p coincide. Therefore, we may skip the lower index and write BMO for the class. Throughout the paper, we will work with the norm $\|\cdot\|_{\mathrm{BMO}_2}$ denoted, for notational simplicity, by $\|\cdot\|_{\mathrm{BMO}}$: the reason for the choice of the parameter $p = 2$ comes from the identity

$$\|Y\|^2_{\mathrm{BMO}_2} = \sup_{t\geq 0} \operatorname{essup}\big(\mathbb{E}\big(Y^2_\infty|\mathcal{F}_t\big) - Y_t^2\big),$$

which makes the bounded mean oscillation easier to handle (see below).

BMO functions and martingales have quite strong integrability properties and there is a natural question about the identification of best constants in the corresponding estimates. This subject has been studied in the literature. We mention here the works of Ivanisvili et al. [6], Korenovskii [8], Osękowski [9,10], Slavin and Vasyunin [11,12] and Vasyunin and Volberg [14]. We also refer the interested reader to the survey of Kazamaki [7], which presents an overview of various related objects from a slightly wider perspective.

In this paper, we continue this line of research and determine the best constant in the martingale version of $H^1 - \mathrm{BMO}$ duality estimate. Here is the main result.

Theorem 1.1. *Let $X = (X_t)_{t\geq 0} \in H^1$ and $Y = (Y_t)_{t\geq 0} \in \mathrm{BMO}$ be arbitrary continuous-path martingales. Then the random variable $\langle X, Y\rangle_\infty$ is integrable and*

$$\mathbb{E}\langle X, Y\rangle_\infty \leq \sqrt{2}\|X\|_{H^1}\|Y\|_{\mathrm{BMO}}. \tag{1}$$

The constant $\sqrt{2}$ is the best possible: for each $\varepsilon > 0$, there exists a probability space and a pair X, Y as above for which

$$\mathbb{E}\langle X, Y\rangle_\infty > (\sqrt{2} - \varepsilon)\|X\|_{H^1}\|Y\|_{\mathrm{BMO}}. \tag{2}$$

Actually, our argumentation will allow us to establish the stronger bound

$$\mathbb{E}\int_0^\infty \big|d\langle X, Y\rangle_t\big| \leq \sqrt{2}\|X\|_{H^1}\|Y\|_{\mathrm{BMO}} \tag{3}$$

(which is also sharp, in the light of (2)). Our approach will exploit the so-called Bellman function (or Burkholder's) method: the estimate (1) will be deduced from the existence of a certain special function, enjoying appropriate size and concavity requirements. This technique is presented in the next section, and our main result is established in Section 3.

2. On the method of proof

2.1. Geometric properties of BMO martingales

We start with certain geometrical interpretation of BMO martingales, which will be helpful in our considerations below (see also Stolyarov and Zatitsky [13] which handles this topic from a wider perspective). Our first observation is that BMO martingales can be identified with two-dimensional martingales with values in appropriate parabolic domains. To be precise, suppose that $Y = (Y_t)_{t\geq 0}$ is a BMO martingale satisfying $\|Y\|_{\mathrm{BMO}} \leq c$, and let $Z = (Z_t)_{t\geq 0} = (\mathbb{E}(Y_\infty^2|\mathcal{F}_t))_{t\geq 0}$ be the associated auxiliary martingale. Then the martingale pair (Y, Z) takes values in the parabolic domain

$$P_c = \{(y, z) : y^2 \leq z \leq y^2 + c^2\}.$$

Indeed, we have $Z_t - Y_t^2 \leq \|Y\|_{\mathrm{BMO}}^2$ (by the definition of BMO norm) and $Z_t \geq Y_t^2$ (by Jensen's inequality). Furthermore, with probability one, the terminal random variable $(Y_\infty, Z_\infty) = (Y_\infty, Y_\infty^2)$ takes values in the lower boundary of P_c. A beautiful fact is that the above implication can be reversed: any uniformly integrable martingale (Y, Z) taking values in P_c and terminating at the lower boundary of this domain gives rise to the BMO martingale of norm not exceeding c. This process is just the first coordinate of the pair (Y, Z).

The first consequence of the above identification is that for each $(y, z) \in P_c$ there is a BMO martingale Y with $\|Y\|_{\mathrm{BMO}} \leq c$ such that $\mathbb{E}Y_\infty = y$ and $\mathbb{E}Y_\infty^2 = z$. To see this, note there exists a line segment I passing through (y, z), with endpoints (y^-, z^-), (y^+, z^+) lying at the lower boundary of P_c (i.e., satisfying $z^- = (y^-)^2$ and $z^+ = (y^+)^2$), and entirely contained within P_c. Let (Y, Z) be an arbitrary martingale starting from (y, z), taking values in I and stopped upon reaching one of the endpoints $(y^\pm, z^\pm)$ (which occurs with probability 1). For instance, we can take Y to be a one-dimensional Brownian motion starting from y and stopped when visiting $y^\pm$, and Z to be an appropriate affine transformation of Y. Then Y is the desired BMO martingale.

As the second consequence, it is easy to explain an important operation, the so-called "splicing" of BMO martingales, whose origins go the works of Burkholder [1] concerning martingale transforms. Namely, suppose that Y^-, Y^+ are two martingales satisfying $\|Y^\pm\|_{\mathrm{BMO}} \leq c$. Denote $\mathbb{E}Y^\pm = y^\pm$, $\mathbb{E}Z^\pm = \mathbb{E}(Y^\pm)^2 = z^\pm$; the above discussion implies that the points $(y^\pm, z^\pm)$ lie within P_c. Suppose that the *entire* line segment I with endpoints (y^-, z^-), (y^+, z^+) lies inside P_c, and let $(y, z) = \alpha(y^-, z^-) + (1 - \alpha)(y^+, z^+)$ be an arbitrary point from this segment $(0 \leq \alpha \leq 1)$. Let $B = (B^1, B^2)$ be an arbitrary martingale started at (y, z), taking in the segment I, and stopped upon reaching one of the endpoints $(y^\pm, z^\pm)$ at time τ. (For instance, if $y^- \neq y^+$, we can take B^1 to be the one-dimensional Brownian motion started at y and stopped when reaching $y^\pm$, and B^2 to be an appropriate affine transformation of B^1). Now we can "glue" the martingales Y^- and Y^+ along B. Namely, consider the martingale (Y, Z) given by $(Y_t, Z_t) = (B_t^1, B_t^2)$ when $t \leq \tau$, and extended to the full time interval by the requirements:

- On the set $\{(Y_\tau, Z_\tau) = (y^-, z^-)\}$, the conditional distribution of $((Y_t, Z_t))_{t\geq\tau}$ coincides with the distribution of (Y^-, Z^-).
- On the set $\{(Y_\tau, Z_\tau) = (y^+, z^+)\}$, the conditional distribution of $((Y_t, Z_t))_{t\geq\tau}$ coincides with the distribution of (Y^+, Z^+).

Thus we have obtained a martingale which equals B on $[0, \tau]$ and then behaves as Y^- or Y^+, depending on whether it left I through (y^-, z^-) or (y^+, z^+). It follows directly from the above construction (and the assumption $I \subset P_c$) that the martingale Y – the first coordinate of the "spliced process" (Y, Z) – is a BMO martingale satisfying $\mathbb{E}Y_\infty = y$, $\mathbb{E}Y_\infty^2 = z$ and $\|Y\|_{\mathrm{BMO}} \leq c$. Moreover, if $Y^\pm$ were assumed to be bounded, then so is Y.

We would like to point out that the above splicing can be carried out simultaneously with the presence of additional processes. For example, suppose that $Y^\pm$ are two BMO martingales satisfying $\|Y^\pm\|_{\mathrm{BMO}} \leq c$, $\mathbb{E}Y_\infty^\pm = y^\pm$, $\mathbb{E}(Y_\infty^\pm)^2 = z^\pm$ and assume that the line segment with endpoints

$(y^{\pm}, z^{\pm})$ is contained within P_c. Next, suppose that $X^{\pm}$ are two uniformly integrable martingales satisfying $\mathbb{E}X^{\pm}_{\infty} = x^{\pm}$ and let $\alpha \in [0, 1]$ be a fixed parameter. Then we can splice the pairs $(X^{\pm}, Y^{\pm})$ into one pair (X, Y) such that $\mathbb{E}X_{\infty} = \alpha x^{-} + (1-\alpha)x^{+}$, $\mathbb{E}Y_{\infty} = \alpha y^{-} + (1-\alpha)y^{+}$ and $\mathbb{E}Y^2_{\infty} = \alpha z^{-} + (1-\alpha)z^{+}$. Indeed, it suffices to add to the "splicing process" $B = (B^1, B^2)$ an extra coordinate responsible for gluing of X.

2.2. Associated special functions

The proof of our main result will rest on properties of a certain special function associated with (1). We are interested in the best constant C in the inequality

$$\mathbb{E}\langle X, Y\rangle_{\infty} \le C\mathbb{E}X^{*}\|Y\|_{\mathrm{BMO}}, \tag{4}$$

where X, Y are continuous-path martingales belonging to H^1 and BMO, respectively. Consider the domain

$$D = \left\{(x_1, x_2, y, z) \in \mathbb{R} \times (0, \infty) \times \mathbb{R} \times \mathbb{R} : |x_1| \le x_2,\, y^2 \le z \le y^2 + 1\right\}.$$

In what follows, for a given C^2 function $U : D \to \mathbb{R}$, the symbol $D^2_{x_1,y,z}U$ stands for the Hessian matrix of U considered as a function of x_1, y and z, that is,

$$D^2_{x_1,y,z}U = \begin{bmatrix} U_{x_1x_1} & U_{x_1y} & U_{x_1z} \\ U_{yx_1} & U_{yy} & U_{yz} \\ U_{zx_1} & U_{zy} & U_{zz} \end{bmatrix}.$$

Theorem 2.1. *Suppose that a C^2 function $U : D \to \mathbb{R}$ satisfies the conditions*

$$U(x_1, x_2, 0, z) \le 0 \quad \textit{for all } (x_1, x_2; 0, z) \in D, \tag{5}$$

$$U(x_1, x_2, y, z) \ge x_1 y - Cx_2 \quad \textit{for all } (x_1, x_2; y, z) \in D, \tag{6}$$

$$x_2 \mapsto U(x_1, x_2, y, z) \quad \textit{is decreasing for all } x_1, y, z. \tag{7}$$

Furthermore, assume that the Hessian matrix $D^2_{x_1,y,z}U$ is negative semidefinite on D. Then the inequality (4) *holds true.*

Proof. By homogeneity, it is enough to study this estimate under the additional assumption $\|Y\|_{\mathrm{BMO}} \le 1$. Furthermore, we may assume that $Y_0 = 0$, replacing Y with $Y - Y_0$ if necessary. Fix $\varepsilon > 0$. The argument rests on Itô's formula applied to the composition of U with the semimartingale $S = (X, X^* \vee \varepsilon, Y, Z)$. As the result, for any stopping time τ,

$$U(S_\tau) = I_0 + I_1 + I_2 + \frac{1}{2}I_3,$$

where

$$I_0 = U(S_0),$$

$$I_1 = \int_0^\tau U_{x_1}(S_s)\,\mathrm{d}X_s + \int_0^\tau U_y(S_s)\,\mathrm{d}Y_s + \int_0^\tau U_z(S_s)\,\mathrm{d}Z_s,$$
$$I_2 = \int_0^\tau U_{x_2}(S_s)\,\mathrm{d}\bigl(X_s^* \vee \varepsilon\bigr),$$
$$I_3 = \int_0^\tau D^2_{x_1,y,z} U(S_s)\,\mathrm{d}[X,Y,Z].$$

Here, in I_3 we have used the abbreviated notation for the sum of all the nontrivial second-order terms, that is,

$$I_3 = \int_0^\tau U_{x_1x_1}(S_s)\,\mathrm{d}[X,X]_s + 2\int_0^\tau U_{x_1y_1}(S_s)\,\mathrm{d}[X,Y]_s + 2\int_0^\tau U_{x_1y_2}(S_s)\,\mathrm{d}[X,Z]_s + \cdots.$$

Before we proceed, let us emphasize here that this is the place where we use the fact (assumed in the introductory section) that all adapted martingales have continuous paths. Thanks to this condition, there is no term involving the jumps of the processes which seems to be difficult to control efficiently.

Let us analyze the behavior of I_0, I_1, I_2 and I_3. By (5), we get $U(S_0) \le 0$. The integrals appearing in I_1 are local martingales; if $(\tau_n)_{n\ge 0}$ is a localizing sequence for each of them and we put $\tau = \tau_n$, then I_1 has expectation zero. The term I_2 is nonpositive by (7), since the process $(X_t^* \vee \varepsilon)_{t\ge 0}$ is nondecreasing. Finally, we have $I_3 \le 0$, which is a direct consequence of the fact that $D^2_{x_1,y,z} \le 0$ and a simple approximation of integrals in I_3 by Riemann sums. Putting all the observations together, we see that $\mathbb{E}U(S_{\tau_n}) \le 0$, which by (6) yields

$$\mathbb{E}X_{\tau_n}Y_{\tau_n} \le C\mathbb{E}\bigl(X^*_{\tau_n} \vee \varepsilon\bigr) \le C\bigl(\mathbb{E}X^* + \varepsilon\bigr).$$

The process $XY - \langle X,Y\rangle$ is a local martingale, so performing another localization if necessary (i.e., replacing τ_n with $\tau_n \wedge \sigma_n$, where $(\sigma_n)_{n\ge 0}$ is some localizing sequence for $XY - \langle X,Y\rangle$), the above estimate yields

$$\mathbb{E}\langle X,Y\rangle_{\tau_n} \le C\bigl(\mathbb{E}X^* + \varepsilon\bigr), \quad n = 0,1,2,\ldots. \tag{8}$$

It remains to carry out an appropriate limiting procedure. It will be shown below (see Remark 3.1), that the assumptions $X \in H^1$ and $Y \in \mathrm{BMO}$ imply that $\mathbb{E}\int_0^\infty |d\langle X,Y\rangle_t| < \infty$, which gives us an integrable majorant of the sequence $(\langle X,Y\rangle_{\tau_n})_{n=0}^\infty$. Therefore, by Lebesgue's dominated convergence theorem, (8) gives

$$\mathbb{E}\langle X,Y\rangle_\infty \le C\bigl(\mathbb{E}X^* + \varepsilon\bigr),$$

and it remains to let $\varepsilon \to 0$ to get the claim. □

Remark 2.2. The following simple and informal observation plays an important role when one tries to *guess* the structure of extremal martingales (i.e., those for which equality, or almost equality is attained in (4)). Namely, one should try to ensure that all the intermediate inequalities (encountered in the course of the above proof) actually become equalities. In particular, this suggests that the pair (X, Y) should be chosen so that the integral

$$I_3(t) = \int_0^t D^2_{x_1,y,z} U(S_s)\, \mathrm{d}[X, Y, Z]_s$$

is constant as a function of time: see the proof of Theorem 2.1. This has two important consequences: first, the Hessian $D^2_{x_1,y,z}U$ should be degenerate, and second, the process S should evolve – locally – along the direction of degeneration of $D^2_{x_1,y,z}U$. As we shall see, this will lead us later to a nontrivial interplay between extremal X and Y in Fefferman's inequality.

Theorem 2.1 shows that the estimate (4) follows once we have constructed a function satisfying appropriate size and concavity-type conditions. Interestingly, a reverse implication is also true: the validity of (4) yields the existence of a certain special function. This fact will play a fundamental role in the proof of sharpness of (1), as it will enable us to shorten the argument and avoid many technical issues arising in the analysis of explicit examples. This approach is not new: for example, it was used by Burkholder [2] to get lower bounds for the constants in martingale transform estimates and to give a new proof of the sharpness of Doob's maximal inequalities.

We need more definitions. For any $(x, y, z) \in \mathbb{R}^3$ such that $y^2 \le z \le y^2 + 1$, let $M(x, y, z)$ denote the class of all pairs (X, Y) of *bounded* martingales satisfying $\|Y\|_{\mathrm{BMO}} \le 1$, $\mathbb{E}X_\infty = x$, $\mathbb{E}Y_\infty = y$ and $\mathbb{E}Y_\infty^2 = z$. By the arguments in Section 2.1, we know that each class $M(x, y, z)$ is nonempty. Now, introduce an abstract function $U^0 : D \to \mathbb{R}$, given by

$$U^0(x_1, x_2, y, z) = \sup\bigl\{\mathbb{E}\bigl(X_\infty Y_\infty - C\bigl(X^* \vee x_2\bigr)\bigr) : (X, Y) \in M(x_1, y, z)\bigr\}.$$

Here, the probability space and the filtration are allowed to vary.

The crucial properties of U^0 are studied in the two theorems below.

Theorem 2.3. *Suppose that the estimate* (4) *holds true for any martingales* $X \in H^1$ *and* $Y \in$ BMO. *Then the function* U^0 *is finite on* D *and satisfies* (5), (6) *and* (7).

Proof. To see that U^0 is finite, pick an arbitrary point (x_1, x_2, y, z) from D and any pair $(X, Y) \in M(x_1, y, z)$. Then

$$\mathbb{E}\bigl(X_\infty Y_\infty - C\bigl(X^* \vee x_2\bigr)\bigr) \le \mathbb{E}\bigl(X_\infty Y_\infty - CX^*\bigr) = x_1 y + \mathbb{E}\bigl(\langle X, Y\rangle_\infty - CX^*\bigr),$$

since X and Y are bounded. By (4), the latter expression is bounded from above by $x_1 y$ and hence also $U^0(x_1, x_2, y, z) \le x_1 y < \infty$, since $(X, Y) \in M(x_1, y, z)$ was chosen arbitrarily. The estimate (5) also follows from the assumed inequality (4). Indeed, pick $(X, Y) \in M(x_1, 0, z)$. Then $\mathbb{E}X_\infty Y_\infty = \mathbb{E}\langle X, Y\rangle_\infty \le C\mathbb{E}X^* \le C\mathbb{E}(X^* \vee x_2)$, that is,

$$\mathbb{E}\bigl(X_\infty Y_\infty - C\bigl(X^* \vee x_2\bigr)\bigr) \le 0.$$

Taking the supremum over all X, Y, we get (5). To prove (6), let $X \equiv x_1$ be a constant martingale and let Y be an arbitrary BMO martingale satisfying $\|Y\|_{\mathrm{BMO}} \le 1$, $\mathbb{E}Y_\infty = y$ and $\mathbb{E}Y_\infty^2 = z$. Then

$(X, Y) \in M(x_1, y, z)$, and hence

$$U^0(x_1, x_2, y, z) \geq \mathbb{E}\big(X_\infty Y_\infty - C\big(X^* \vee x_2\big)\big) = x_1 y - Cx_2,$$

which is (6). The condition (7) is obvious from the very definition of U^0. □

It turns out that the function U^0 also satisfies a certain version of the last condition of Theorem 2.1. However, in general, U^0 does not have to be of class C^2, and hence we need to express the condition $D^2_{x_1,y,z} U^0 \leq 0$ in a slightly different manner which does not refer to any differentiability of the special function. Recall the notation

$$P_1 = \big\{(y, z) : y^2 \leq z \leq y^2 + 1\big\}$$

for the parabolic region.

Theorem 2.4. *Let $(y^\pm, z^\pm)$ be two points such that the whole line segment joining them is contained within P_1. Assume further that $\alpha \in (0, 1)$ and $x_1^\pm$ are two real numbers satisfying $|x_1^-| \leq |x_1^+|$. Set $x_1 = \alpha x_1^- + (1-\alpha)x_1^+$, $y = \alpha y^- + (1-\alpha)y^+$ and $z = \alpha z^- + (1-\alpha)z^+$. Then for any $x_2 \geq |x_1|$ we have*

$$\begin{aligned} &U^0(x_1, x_2, y, z) \\ &\quad \geq \alpha U^0\big(x_1^-, \big|x_1^+\big| \vee x_2, y^-, z^-\big) + (1-\alpha) U^0\big(x_1^+, \big|x_1^+\big| \vee x_2, y^+, z^+\big). \end{aligned} \tag{9}$$

Proof. Fix arbitrary two pairs $(X^\pm, Y^\pm)$ belonging to $M(x_1^\pm, y^\pm, z^\pm)$. Since the line segment with endpoints $(y^\pm, z^\pm)$ is contained in P_1, the arguments from Section 2.1 imply that these martingales can be spliced into one martingale $(X, Y) \in M(x, y, z)$, at some time τ. However, note that the maximal function X^* does not splice: on $X_\tau = x_1^\pm$, the conditional distribution of $(X^*)_{t \geq \tau}$ *does not* coincide with the distribution of $(X^\pm)^*$. However, we have an appropriate estimate: since $|x_1^-| \leq |x_1^+|$, we have

$$X^* = \max\Big\{\sup_{t\in[0,\tau]} |X_t|, \sup_{t\geq\tau} |X_t|\Big\} \leq \big|x_1^+\big| \vee \sup_{t\geq\tau} |X_t|,$$

which implies

$$\begin{aligned} U^0(x_1, x_2, y, z) &\geq \mathbb{E}\big(X_\infty Y_\infty - C\big(X^*_\infty \vee x_2\big)\big) \\ &= \alpha \mathbb{E}\big(X^-_\infty Y^-_\infty - C\big(\big(X^-_\infty\big)^* \vee \big|x_1^+\big| \vee x_2\big)\big) \\ &\quad + (1-\alpha)\mathbb{E}\big(X^+_\infty Y^+_\infty - C\big(\big(X^+_\infty\big)^* \vee \big|x_1^+\big| \vee x_2\big)\big). \end{aligned}$$

Taking the supremum over all $(X^\pm, Y^\pm)$ as above, we get the claim. □

We conclude this section by providing certain additional properties of U^0, which will be quite helpful later.

Lemma 2.5. *Let* (x_1, x_2, y, z) *be an arbitrary point from* D. *Then for any* $\lambda > 0$ *and* $a \in \mathbb{R}$ *we have the identities*

$$U^0(\lambda x_1, \lambda x_2, y, z) = \lambda U^0(x_1, x_2, y, z), \tag{10}$$

$$U^0\big(x_1, x_2, y + a, z + 2ay + a^2\big) = ax_1 + U^0(x_1, x_2, y, z) \tag{11}$$

and

$$U^0(x_1, x_2, y, z) = U^0(-x_1, x_2, -y, z). \tag{12}$$

In addition, for any $\delta > 0$ *we have*

$$U^0(x_1, x_2 + \delta, y, z) \geq U^0(x_1, x_2, y, z) - C\delta. \tag{13}$$

Proof. To show (10), pick an arbitrary pair $(X, Y) \in M(x_1, y, z)$. Then we have $(\lambda X, Y) \in M(\lambda x_1, y, z)$, so

$$\begin{aligned} U^0(\lambda x_1, \lambda x_2, y, z) &\geq \mathbb{E}\big(\lambda X_\infty Y_\infty - C\big[\big(\lambda X^*\big) \vee (\lambda x_2)\big]\big) \\ &= \lambda \mathbb{E}\big(X_\infty Y_\infty - C\big(X^* \vee x_2\big)\big). \end{aligned}$$

Hence $U^0(\lambda x_1, \lambda x_2, y, z) \geq \lambda U^0(x_1, x_2, y, z)$, since (X, Y) was arbitrary; the reverse bound follows by replacing x_1, x_2, λ by λx_1, λx_2 and λ^{-1}, respectively. The proof of (11) is similar: for any pair $(X, Y) \in M(x_1, y, z)$ we have $(X, Y + a) \in M(x_1, y + a, z + 2ay + a^2)$ (indeed we have $\mathbb{E}(Y_\infty + a)^2 = \mathbb{E}Y_\infty^2 + 2a\mathbb{E}Y_\infty + a^2 = z + 2ay + a^2$). Consequently,

$$\begin{aligned} U^0\big(x_1, x_2, y + a, z + 2ay + a^2\big) &\geq \mathbb{E}\big(X_\infty(Y_\infty + a) - C\big(X^* \vee x_2\big)\big) \\ &= ax_1 + \mathbb{E}\big(X_\infty Y_\infty - C\big(X^* \vee x_2\big)\big), \end{aligned}$$

and hence $U^0(x_1, x_2, y + a, z + 2ay + a^2) \geq ax_1 + U^0(x_1, x_2, y, z)$, by taking the supremum over all (X, Y). Replacing y, a with $y + a$ and $-a$, respectively, we get the reverse bound. The identity (12) follows directly from the definition of U^0 and the equivalence $(X, Y) \in M(x_1, y, z)$ iff $(-X, -Y) \in M(-x_1, -y, z)$. Finally, to show (13), pick an arbitrary $(X, Y) \in M(x_1, y, z)$. Since $a \vee (b + c) \leq a \vee b + c$ for any nonnegative numbers a, b, c, we may write

$$U^0(x_1, x_2 + \delta, y, z) \geq \mathbb{E}\big(X_\infty Y_\infty - CX^* \vee (x_2 + \delta)\big) \geq \mathbb{E}\big(X_\infty Y_\infty - C\big(\big(X^* \vee x_2\big) - \delta\big)\big)$$

and it remains to take the supremum over all (X, Y) as above. □

3. Proof of Theorem 1.1

The contents of this section is split into three parts. First, we will prove the validity of (1), by providing an explicit formula for the special function U as in Theorem 2.1. Then we will show how to modify the approach to get the stronger estimate (3). The main technical difficulty lies in the proof of the sharpness of (1), which is done in Section 3.3. The construction of an appropriate pair (X, Y) for which the equality is (almost) attained is very complicated; the process has a complicated fractal-type behavior. However, it turns out that Theorem 2.3 enables us to avoid most of these issues.

3.1. A special function and its properties

Let $U : D \to \mathbb{R}$ be given by

$$U(x_1, x_2; y, z) = x_1 y - \frac{1}{2\sqrt{2}}\left(\frac{x_1^2}{x_2} + 2\big(y^2 - z\big)x_2 + 3x_2\right).$$

Let us check that this object meets all the requirements of Theorem 2.1, with $C = \sqrt{2}$. First, U has the appropriate regularity (it is even of class C^∞). The inequality (5) is straightforward: since $(x_1, x_2, 0, z) \in D$, we have $z \le 1$ and

$$U(x_1, x_2; 0, z) = -\frac{1}{2\sqrt{2}}\left(\frac{x_1^2}{x_2} - 2zx_2 + 3x_2\right) \le -\frac{1}{2\sqrt{2}}(-2x_2 + 3x_2) \le 0.$$

The majorization (6) is equivalent to $x_2 - \frac{x_1^2}{x_2} + 2(z - y^2)x_2 \ge 0$, which follows immediately from the estimates defining the domain D. To show (7), it is enough to prove that the partial derivative with respect to x_2 is nonpositive: but we have

$$U_{x_2}(x_1, x_2, y, z) = \frac{1}{2\sqrt{2}}\left(\frac{x_1^2}{x_2^2} - 2\big(y^2 - z\big) - 3\right) \le 0.$$

Finally, the condition on the "partial Hessian" $D^2_{x_1,y,z}U$ is also very easy to check: a direct computation reveals that

$$D^2_{x_1,y,z}U(x_1, x_2, y, z) = \begin{bmatrix} -\frac{1}{\sqrt{2}x_2} & 1 & 0 \\ 1 & -\sqrt{2}x_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{14}$$

and this matrix is obviously negative semidefinite. This gives the proof of (1).

3.2. On the estimate (3)

The above function can be modified to yield the stronger bound (3). To understand the passage, we start with complicating the above proof a little bit, adding an extra dimension to the special function. Namely, denote the term $x_1 y$ in the formula for U as an additional variable and consider the function of five variables

$$\tilde{U}(w, x_1, x_2, y, z) = w - \frac{1}{2\sqrt{2}}\left(\frac{x_1^2}{x_2} + 2\big(y^2 - z\big)x_2 + 3x_2\right),$$

defined on $\mathbb{R} \times D$. This function inherits most of the structural properties from U: we have the estimates $\tilde{U}(0, x_1, x_2, 0, z) \leq 0$, $U(w, x_1, x_2, y, z) \geq w - \sqrt{2}x_2$ and the section $x_2 \mapsto U(w, x_1, x_2, y, z)$ is nonincreasing for each w, x_1, y and z. The only condition which changes significantly, is the Hessian requirement: we have

$$\begin{bmatrix} \tilde{U}_{x_1x_1} & \tilde{U}_{x_1y} + \tilde{U}_w & \tilde{U}_{x_1z} \\ \tilde{U}_{yx_1} + \tilde{U}_w & \tilde{U}_{yy} & \tilde{U}_{yz} \\ \tilde{U}_{zx_1} & \tilde{U}_{zy} & \tilde{U}_{zz} \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}x_2} & 1 & 0 \\ 1 & -\sqrt{2}x_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \leq 0 \tag{15}$$

at each point of the domain. Now we repeat the reasoning from Section 2. We assume that $\|Y\|_{\mathrm{BMO}} \leq 1$ and apply Itô's formula to the composition of $\tilde{U}$ and the semimartingale $S = ((\langle X, Y\rangle_t, X_t, X_t^* \vee \varepsilon, Y_t, Z_t))_{t\geq 0}$. All the calculations are essentially the same, the additional first-order term $\int_0^t U_w(S_s)\, d\langle X, Y\rangle_s$ is handled with the use of the modified Hessian requirement above. Thus, we obtain

$$\mathbb{E}\langle X, Y\rangle_{\tau_n} \leq \sqrt{2}\mathbb{E}\big(X^* \vee \varepsilon\big)$$

for some localizing sequence $(\tau_n)_{n\geq 0}$, and then limiting arguments yield (1).

To get (3), we observe that the following stronger version of (15) is valid:

$$\begin{bmatrix} \tilde{U}_{x_1x_1} & \tilde{U}_{x_1y} \pm \tilde{U}_w & \tilde{U}_{x_1z} \\ \tilde{U}_{yx_1} \pm \tilde{U}_w & \tilde{U}_{yy} & \tilde{U}_{yz} \\ \tilde{U}_{zx_1} & \tilde{U}_{zy} & \tilde{U}_{zz} \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}x_2} & \pm 1 & 0 \\ \pm 1 & -\sqrt{2}x_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \leq 0.$$

Consequently, the whole above proof carries over if we compose $\tilde{U}$ with the semimartingale $S = ((\int_0^t |d\langle X, Y\rangle_s|, X_t, X_t^* \vee \varepsilon, Y_t, Z_t))_{t\geq 0}$. We get

$$\mathbb{E}\int_0^{\tau_n} \big|d\langle X, Y\rangle_s\big| \leq \sqrt{2}\mathbb{E}\big(X^* \vee \varepsilon\big)$$

for an appropriate localizing sequence $(\tau_n)_{n\geq 0}$. Letting $n \to \infty$ and then $\varepsilon \to 0$ yields the desired stronger bound, by Lebesgue's monotone convergence theorem.

Remark 3.1. The above proof shows that if $X \in H^1$ and $Y \in \mathrm{BMO}$, then the random variable $\int_0^\infty |d\langle X, Y\rangle_s|$ is integrable.

3.3. On the optimality of the constant $\sqrt{2}$

Suppose that Fefferman's inequality (4) holds with some constant C. By Theorem 2.3, this gives rise to the associated special function U^0. Our plan is to exploit the structural properties of this function to get the estimate $C \geq \sqrt{2}$: this will be done in Lemmas 3.2 and 3.4 below. The reader can proceed directly to these statements, however, we feel that the parameters appearing there are quite mysterious and it is convenient to shed some light on them, explaining first the geometry of the extremal martingales.

We start with the description of the BMO martingale Y satisfying $\|Y\|_{\mathrm{BMO}} \leq 1$. By the discussion in Section 2, such a martingale is uniquely determined by a uniformly integrable martingale (Y, Z) taking values in the parabolic region P_1 and terminating at the lower boundary of this domain. We assume that (Y, Z) starts from $(0, 1)$ and Y is a one-dimensional Brownian motion, stopped appropriately, as the following evolution scheme indicates. Set $\tau_0 \equiv 0$.

Rule I. Suppose that for some n, the variable (Y_{τ_n}, Z_{τ_n}) lies at the upper boundary of P_1: we have $Y_{\tau_n} = y_0$, $Z_{\tau_n} = z_0$ for some y_0, z_0 satisfying $z_0 = y_0^2 + 1$. Then (Y, Z) starts moving along the line tangent to the upper boundary, until Y gets to $y_0 - \frac{1}{2}$ or to $y_0 + \delta$: the corresponding stopping time is denoted by τ_{n+1}. If the first possibility happens, proceed to Rule II. If the second scenario occurs, then $(Y_{\tau_{n+1}}, Z_{\tau_{n+1}})$ has found itself at the point $(y_0 + \delta, z_0 + 2y_0\delta)$. The line passing through this point and $(1 + 2\delta, (1 + 2\delta)^2 + 1)$ is tangent to the upper boundary, and the process starts moving along this new line, until Y gets to the point $y_0 + 2\delta - \frac{1}{2}$ or to $y_0 + 2\delta$: denote the corresponding stopping time by τ_{n+2}. If the first possibility happens, we go to Rule II; otherwise, $(Y_{\tau_{n+2}}, Z_{\tau_{n+2}})$ is again at the upper boundary and we may go to the beginning of Rule I.

Rule II. Suppose that for some n, the random variable (Y_{τ_n}, Z_{τ_n}) has found itself at a point (y_0, z_0) with $z_0 - y_0^2 = \frac{3}{4}$. This point lies at the line segment connecting $(y_0 + \frac{1}{2}, (y_0 + \frac{1}{2})^2)$ and $(y_0 - \frac{1}{2}, (y_0 - \frac{1}{2})^2 + 1)$ (actually, this line is tangent to the upper boundary at the latter point). The process moves along this segment until it reaches one of its endpoints: denote the associated stopping time by τ_{n+1}. If it gets to $(y_0 + \frac{1}{2}, (y_0 + \frac{1}{2})^2)$ first, then $(Y_{\tau_{n+1}}, Z_{\tau_{n+1}})$ belongs to the lower boundary and the evolution stops ultimately. If the martingale reaches the second endpoint, then it is on the upper boundary and we proceed to Rule I.

This completes the description of the evolution of Y. Note that the process (Y, Z) is self-similar in the sense that for any fixed parameter $a \in \mathbb{R}$, the evolution of $(Y + a, Z + 2aY + a^2)$ follows the same rules (the only thing which changes is the starting position). The construction of X depends heavily on Remark 2.2 above. As stated there, we should find X so that the process $S = (X, X^*, Y, Z)$ moves along degeneration directions of $D^2_{x_1,y,z}U$. By (14), we see that this will be the case if we have the equality

$$dX_t = \sqrt{2}X_t^* \, dY_t \tag{16}$$

for each t; we add the initial condition $X_0 \equiv 1$ and may hope that the pair (X, Y) will lead us to the best constant $\sqrt{2}$ in (1).

However, the rigorous verification that $\lim_{\delta \to 0} \mathbb{E}\langle X, Y\rangle_\infty / (\|X\|_{H^1}\|Y\|_{\mathrm{BMO}}) = \sqrt{2}$ would be very elaborate, in particular it would require the computation of the H^1 norm of X. Fortunately, the two evolution rules given above can be reformulated in the language of the abstract function U^0 and then some simple manipulations yield the optimality of the constant $\sqrt{2}$. Rule I gives the following estimate: see Remark 3.3 below for the careful explanation.

Lemma 3.2. *We have the inequality*

$$U^0(1, 1, 0, 1) \geq (2 + \sqrt{2})U^0\left(1 - \frac{\sqrt{2}}{2}, 1, -\frac{1}{2}, 1\right) + \frac{1}{2 - \sqrt{2}}.$$

Proof. For the sake of convenience, we split the reasoning into three parts.

Step 1. By (9), (13) and finally (10), we have

$$\begin{aligned}
&U^0(1,1,0,1)\\
&\geq \frac{1}{1+2\delta}U^0(1+\sqrt{2}\delta, 1+\sqrt{2}\delta, \delta, 1) + \frac{2\delta}{1+2\delta}U^0\left(1-\frac{\sqrt{2}}{2}, 1+\sqrt{2}\delta, -\frac{1}{2}, 1\right)\\
&\geq \frac{1}{1+2\delta}U^0(1+\sqrt{2}\delta, 1+\sqrt{2}\delta, \delta, 1) + \frac{2\delta}{1+2\delta}U^0\left(1-\frac{\sqrt{2}}{2}, 1, -\frac{1}{2}, 1\right) + o(\delta)\\
&= \frac{1+\sqrt{2}\delta}{1+2\delta}U^0(1,1,\delta,1) + \frac{2\delta}{1+2\delta}U^0\left(1-\frac{\sqrt{2}}{2}, 1, -\frac{1}{2}, 1\right) + o(\delta).
\end{aligned}$$

Step 2. An application of (11) (with $x_1 = x_2 = z = 1$, $y = -\delta$ and $a = 2\delta$) gives

$$U^0(1,1,\delta,1) = 2\delta + U^0(1,1,-\delta,1). \tag{17}$$

Step 3. The second term in (17) is handled by (9) and (13) again:

$$\begin{aligned}
&U^0(1,1,-\delta,1)\\
&\geq (1-2\delta)U^0\left(1+\frac{\sqrt{2}\delta}{1-2\delta}, 1+\frac{\sqrt{2}\delta}{1-2\delta}, 0, 1\right) + 2\delta U^0\left(1-\frac{\sqrt{2}}{2}, 1, -\frac{1}{2}, 1\right) + o(\delta)\\
&\geq (1-2\delta)\left(1+\frac{\sqrt{2}\delta}{1-2\delta}\right)U^0(1,1,0,1) + 2\delta U^0\left(1-\frac{\sqrt{2}}{2}, 1, -\frac{1}{2}, 1\right) + o(\delta).
\end{aligned}$$

Putting all the above facts together, we obtain the estimate

$$\begin{aligned}
U^0(1,1,0,1) \geq{}& \frac{1+\sqrt{2}\delta}{1+2\delta}\cdot(1-2\delta)\left(1+\frac{\sqrt{2}\delta}{1-2\delta}\right)U^0(1,1,0,1)\\
&+\left[\frac{2\delta}{1+2\delta} + \frac{1+\sqrt{2}\delta}{1+2\delta}\cdot 2\delta\right]U^0\left(1-\frac{\sqrt{2}}{2}, 1, -\frac{1}{2}, 1\right)\\
&+\frac{1+\sqrt{2}\delta}{1+2\delta}\cdot 2\delta + o(\delta).
\end{aligned}$$

Moving the first term on the right to the left-hand side, dividing throughout by δ and letting $\delta \to 0$ yields the claim. □

Remark 3.3. The first inequality in Step 1 above corresponds to the first part of Rule I, applied to $(y_0, z_0) = (0, 1)$. Indeed: the martingale (Y, Z) starting from $(0, 1)$ must move horizontally (this is the tangent direction) until it gets to $(\delta, 1)$ or $(-\frac{1}{2}, 1)$, which happens with probabilities $(1+2\delta)^{-1}$ and $2\delta(1+2\delta)^{-1}$, respectively. At the same time, according to (16), the martingale X moves proportionally to Y, so it gets to $1+\sqrt{2}\delta$ or $1-\sqrt{2}/2$ (we may treat X^* as 1: the error here disappears as $\delta \to 0$). The equality of Step 2 corresponds, in a sense, to the self-similarity of the process Y. As the second part of Rule I states, if (Y, Z) got to the point $(\delta, 1)$, then it should move along the line segment joining $(2\delta, (2\delta)^2 + 1)$ and $(2\delta - 1/2, (2\delta)^2 - 2\delta + 1)$. The idea behind (17) is to look at the evolution of the pair $(Y - 2\delta, 1)$, corresponding to shifted martingale $Y - 2\delta$. Then, by the above requirement, this new pair starts from $(-\delta, 1)$ and moves until it gets to $(0, 1)$ or to $(-1/2, 1)$, which gives the estimate for U^0 appearing in Step 3.

We turn our attention to the translation of Rule II into the language of U^0.

Lemma 3.4. *We have the inequality*

$$U^0\left(1 - \frac{\sqrt{2}}{2}, 1, -\frac{1}{2}, 1\right) \geq \frac{U^0(1, 1, 0, 1)}{2 + \sqrt{2}} - \frac{C}{\sqrt{2}} + \frac{1}{2},$$

where C is the constant in the assumed inequality (4).

Proof. By (9), we obtain

$$\begin{aligned}
&U^0\left(1 - \frac{\sqrt{2}}{2}, 1, -\frac{1}{2}, 1\right) \\
&\quad \geq \frac{\sqrt{2}}{1+\sqrt{2}} U^0(1, 1, 0, 0) + \frac{1}{1+\sqrt{2}} U^0\left(-\frac{\sqrt{2}}{2}, 1, -\frac{1+\sqrt{2}}{2}, 1 + \sqrt{2}\right) \\
&\quad \geq -\frac{C\sqrt{2}}{1+\sqrt{2}} + \frac{1}{1+\sqrt{2}} U^0\left(-\frac{\sqrt{2}}{2}, 1, -\frac{1+\sqrt{2}}{2}, 1 + \sqrt{2}\right),
\end{aligned}$$

where in the last passage we have exploited (6). To handle the second term on the right, we use (9) again:

$$\begin{aligned}
&U^0\left(-\frac{\sqrt{2}}{2}, 1, -\frac{1+\sqrt{2}}{2}, 1 + \sqrt{2}\right) \\
&\quad \geq (\sqrt{2} - 1) U^0\left(-1 + \frac{\sqrt{2}}{2}, 1, \frac{1}{2} - \sqrt{2}, 3 - \sqrt{2}\right) + (2 - \sqrt{2}) U^0(-1, 1, -\sqrt{2}, 3).
\end{aligned}$$

However, both terms on the right can be simplified, by (11); namely,

$$\begin{aligned}
U^0\left(-1 + \frac{\sqrt{2}}{2}, 1, \frac{1}{2} - \sqrt{2}, 3 - \sqrt{2}\right) &= -\left(-1 + \frac{\sqrt{2}}{2}\right)\sqrt{2} + U^0\left(-1 + \frac{\sqrt{2}}{2}, 1, \frac{1}{2}, 1\right) \\
&= -\left(-1 + \frac{\sqrt{2}}{2}\right)\sqrt{2} + U^0\left(1 - \frac{\sqrt{2}}{2}, 1, -\frac{1}{2}, 1\right)
\end{aligned}$$

and

$$U^0(-1, 1, -\sqrt{2}, 3) = U^0(-1, 1, 0, 1) + \sqrt{2} = U^0(1, 1, 0, 1) + \sqrt{2}.$$

Putting all the facts together, we obtain the desired estimate. □

Now the proof of the sharpness follows easily: the combination of the above two lemmas yields

$$U^0(1,1,0,1) \geq U^0(1,1,0,1) + (2+\sqrt{2})\left(-\frac{C}{\sqrt{2}} + \frac{1}{2}\right) + \frac{1}{2-\sqrt{2}},$$

which is equivalent to $C \geq \sqrt{2}$. This shows that the constant $\sqrt{2}$ we obtained in (1) is indeed the best possible.

Acknowledgements

The author would like to thank anonymous referees for the careful reading of the first version of the paper and many valuable comments and suggestions. The research was supported by Narodowe Centrum Nauki (Poland), grant no. DEC-2014/14/E/ST1/00532.

References

[1] Burkholder, D.L. (1984). Boundary value problems and sharp inequalities for martingale transforms. *Ann. Probab.* **12** 647–702. MR0744226

[2] Burkholder, D.L. (1991). Explorations in martingale theory and its applications. In *École d'Été de Probabilités de Saint-Flour XIX—1989*. *Lecture Notes in Math.* **1464** 1–66. Berlin: Springer. MR1108183 https://doi.org/10.1007/BFb0085167

[3] Dellacherie, C. and Meyer, P.-A. (1982). *Probabilities and Potential. B: Theory of Martingales. North-Holland Mathematics Studies* **72**. Amsterdam: North-Holland. MR0745449

[4] Fefferman, C. (1971). Characterizations of bounded mean oscillation. *Bull. Amer. Math. Soc.* **77** 587–588. MR0280994 https://doi.org/10.1090/S0002-9904-1971-12763-5

[5] Getoor, R.K. and Sharpe, M.J. (1972). Conformal martingales. *Invent. Math.* **16** 271–308. MR0305473 https://doi.org/10.1007/BF01425714

[6] Ivanisvili, P., Osipov, N.N., Stolyarov, D.M., Vasyunin, V.I. and Zatitskiy, P.B. (2016). Bellman function for extremal problems in BMO. *Trans. Amer. Math. Soc.* **368** 3415–3468. MR3451882 https://doi.org/10.1090/tran/6460

[7] Kazamaki, N. (1994). *Continuous Exponential Martingales and BMO. Lecture Notes in Math.* **1579**. Berlin: Springer. MR1299529 https://doi.org/10.1007/BFb0073585

[8] Korenovskiĭ, A.A. (1992). The connection between mean oscillations and exact exponents of summability of functions. *Math. USSR-Sb.* **71** 561–567.

[9] Osękowski, A. (2015). Sharp maximal estimates for BMO martingales. *Osaka J. Math.* **52** 1125–1142. MR3426632

[10] Osękowski, A. (2015). Weak-L^∞ inequalities for BMO functions. *New York J. Math.* **21** 699–713. MR3386542

[11] Slavin, L. and Vasyunin, V. (2011). Sharp results in the integral-form John–Nirenberg inequality. *Trans. Amer. Math. Soc.* **363** 4135–4169. MR2792983 https://doi.org/10.1090/S0002-9947-2011-05112-3

[12] Slavin, L. and Vasyunin, V. (2012). Sharp L^p estimates on BMO. *Indiana Univ. Math. J.* **61** 1051–1110. MR3071693 https://doi.org/10.1512/iumj.2012.61.4651

[13] Stolyarov, D.M. and Zatitskiy, P.B. (2016). Theory of locally concave functions and its applications to sharp estimates of integral functionals. *Adv. Math.* **291** 228–273. MR3459018 https://doi.org/10.1016/j.aim.2015.11.048

[14] Vasyunin, V. and Volberg, A. (2014). Sharp constants in the classical weak form of the John–Nirenberg inequality. *Proc. Lond. Math. Soc.* (3) **108** 1417–1434. MR3218314 https://doi.org/10.1112/plms/pdt063

Received July 2019 and revised November 2019