ON THE BEST CONSTANT IN THE ESTIMATE RELATED TO $H^1 - BMO$ DUALITY

ADAM OSĘKOWSKI

Abstract. Let $I \subset \mathbb{R}$ be an interval and let f, φ be arbitrary elements of $H^1(I)$ and $BMO(I)$, respectively, with $\int_I \varphi = 0$. The paper contains the proof of the estimate

$$\int_I f\varphi \leq \sqrt{2}\|f\|_{H^1(I)}\|\varphi\|_{BMO(I)}$$

and it is shown that $\sqrt{2}$ cannot be replaced by a smaller universal constant. The argument rests on the existence of a special function enjoying appropriate size and concavity requirements.

1. Introduction

Suppose that f is a real-valued locally integrable function defined on some fixed interval $I \subset \mathbb{R}$. It maximal function $\mathcal{M}f : I \to [0, \infty)$ is defined by

$$\mathcal{M}f(x) = \sup \left|\langle f\rangle_{[a,b]}\right|,$$

where the supremum is taken over all intervals $[a, b] \subseteq I$ containing x and $\langle f\rangle_{[a,b]} = \frac{1}{b-a}\int_a^b f$ stands for the average of f over $[a, b]$ (all the integrals considered in this paper will be taken with respect to the Lebesgue measure). Given $1 \leq p < \infty$, if the maximal function $\mathcal{M}f$ lies in $L^p(I)$, then f is said to belong to the Hardy space $H^p(I)$ and we define $\|f\|_{H^p(I)} = \|\mathcal{M}f\|_{L^p(I)}$. By the classical inequality of Hardy and Littlewood, if p is strictly bigger than 1, then we have $H^p(I) = L^p(I)$ and therefore the dual space of $H^p(I)$ is equal to $L^q(I)$, $1/p + 1/q = 1$. In the boundary case $p = 1$, the celebrated result of Fefferman [5] asserts that $(H^p(I))^* = BMO(I)$, the class of functions of bounded mean oscillation. The latter space consists of all (equivalence classes of) functions $\varphi : I \to \mathbb{R}$ satisfying

$$\|\varphi\|_{BMO(I)} := \sup_{[a,b]\subseteq I} \left\langle \left(\varphi - \langle\varphi\rangle_{[a,b]}\right)^2\right\rangle_{[a,b]}^{1/2} = \sup_{[a,b]\subset I} \left(\langle\varphi^2\rangle_{[a,b]} - \langle\varphi\rangle_{[a,b]}^2\right)^{1/2} < \infty.$$

The class BMO was introduced by John and Nirenberg in [9], and it has turned out to be one of the fundamental spaces in harmonic analysis and probability theory. One of important features of this object is that in many contexts, it is a convenient replacement for the space L^∞. For example, many classical operators in analysis do not map L^∞ into L^∞, but are bounded on BMO (cf. [8]); another important example comes from the interpolation theory: there is an appropriate version of Marcinkiewicz' theorem, which describes L^p spaces ($1 < p < \infty$) as interpolation spaces between L^1 and BMO (cf. [2]).

2010 *Mathematics Subject Classification.* Primary: 42A05, 42B35, 49K20.
Key words and phrases. BMO, Hardy space, Bellman function.
Research supported by Narodowe Centrum Nauki (Poland), grant no. DEC-2014/14/E/ST1/00532.

It is not difficult to see that the aforementioned $H^1 - BMO$ duality is equivalent to the existence of a universal constant C such that

$$\int_I f\varphi \leq C\|f\|_{H^1(I)}\|\varphi\|_{BMO(I)}, \tag{1.1}$$

provided $\int_I \varphi = 0$. The purpose of this paper is to identify the optimal constant in the estimate (1.1).

Theorem 1.1. *The inequality* (1.1) *holds with* $C = \sqrt{2}$. *The constant is the best possible.*

The inequality (1.1) and its variants have been studied by a number of authors, mostly in the context of square-function-based H^1. For example, Getoor and Sharpe [7] established a probabilistic analogue, with the same constant $\sqrt{2}$. See also [4] for a version in the context of $\mathcal{E}$-martingales and [3, 6, 12] for other probabilistic variants. Consult also [15] for an analytic estimate involving Triebel-Lizorkin spaces.

Our approach rests on the Bellman function method, a powerful technique used widely in analysis and probability theory. This approach has its origins in the theory of optimal control (cf. [1]) and, roughly speaking, it reduces the problem of proving a given estimate to the existence of a certain special function, enjoying appropriate size and concavity conditions. For an overview of the general method, see e.g. [10, 11]; consult also the works [13, 14, 16] for a version specified to BMO-estimates.

The next section contains the proof of (1.1) with $C = \sqrt{2}$. The final part of the paper is devoted to the optimality of the constant; this is accomplished by constructing appropriate examples.

2. Proof of (1.1)

Throughout this section, ε is a fixed positive number (which eventually will be sent to zero). Introduce the domain $\mathcal{D} = \{(u, v, y, z) : v > 0,\ v \geq |u|,\ 0 \leq z - y^2 \leq 1\}$ and consider the function $B : \mathcal{D} \to \mathbb{R}$ given by the formula

$$B(u, v, y, z) = uy - \frac{v}{2\sqrt{2(1-\varepsilon)}}\left(\frac{u^2}{v^2} + 2(y^2 - z) + 3\right).$$

This function has a certain concavity-type property, which will be studied in two lemmas below.

Lemma 2.1. *Let α_-, α_+ be two positive numbers satisfying $\alpha_- + \alpha_+ = 1$ and $\alpha_- \leq \varepsilon$. Assume further that (u, v, y, z), (u_-, v_-, y_-, z_-) and (u_+, v_+, y_+, z_+) are elements of $\mathcal{D}$ such that $u = \alpha_- u_- + \alpha_+ u_+$, $y = \alpha_- y_- + \alpha_+ y_+$, $z = \alpha_- z_- + \alpha_+ z_+$ and $v_\pm = \max\{|u_\pm|, v\}$. If in addition we have $|u_-| \leq |u_+|$, then*

$$B(u, v, y, z) \geq \alpha_- B(u_-, v_-, y_-, z_-) + \alpha_+ B(u_+, v_+, y_+, z_+). \tag{2.1}$$

In the proof of the above statement, we will use the notation $d_\pm = u_\pm - u$ and $e_\pm = y_\pm - y$. We split the analysis into two major cases.

Proof of Lemma 2.1 for $|u_+| \leq v$. Then both $|u_-|$ and $|u_+|$ are not bigger than v, so $v_- = v_+ = v$ and the estimate (2.1) is equivalent to

$$0 \geq \alpha_- d_- e_- + \alpha_+ d_+ e_+ - \frac{v}{2\sqrt{2(1-\varepsilon)}}\left[\frac{\alpha_- d_-^2 + \alpha_+ d_+^2}{v^2} + 2(\alpha_- e_-^2 + \alpha_+ e_+^2)\right].$$

But this is easy: observe that

$$\frac{v}{2\sqrt{2(1-\varepsilon)}}\left(\frac{\alpha_-d_-^2}{v^2}+2\alpha_-e_-^2\right)\geq \alpha_-\cdot\frac{2\sqrt{2}}{2\sqrt{2(1-\varepsilon)}}|d_-e_-|\geq \alpha_-d_-e_-$$

and similarly,

$$\frac{v}{2\sqrt{2(1-\varepsilon)}}\left(\frac{\alpha_+d_+^2}{v^2}+2\alpha_+e_+^2\right)\geq \alpha_+\cdot\frac{2\sqrt{2}}{2\sqrt{2(1-\varepsilon)}}|d_+e_+|\geq \alpha_+d_+e_+.$$

Summing these two estimates, we get the assertion. □

Proof of Lemma 2.1 for $|u_+|>v$. Here the reasoning will be a bit more technical. We start from the observation that it is enough to show the claim under the additional assumption $|u_-|\leq v$. Indeed, if $|u_-|>v$, then replacing v with the bigger quantity $|u_-|$ does not change the right-hand side of (2.1) and decreases the left (we have $B_v(u,v,y,z)=-(2\sqrt{2(1-\varepsilon)})^{-1}\left[-(u/v)^2+2(y^2-z)+3\right]\leq 0$, due to the estimates $y^2-z\geq -1$ and $0\leq |u/v|\leq 1$). For $|u_-|\leq v<|u_+|$, the inequality (2.1) takes the form

$$\begin{aligned} &uy-\frac{v}{2\sqrt{2(1-\varepsilon)}}\left[\frac{u^2}{v^2}+2(y^2-z)+3\right]\\ &\geq \alpha_-\left\{u_-y_- - \frac{v}{2\sqrt{2(1-\varepsilon)}}\left[\frac{u_-^2}{v^2}+2(y_-^2-z_-)+3\right]\right\}\\ &\quad+\alpha_+\left\{u_+y_+-\frac{|u_+|}{2\sqrt{2(1-\varepsilon)}}\left[1+2(y_+^2-z_+)+3\right]\right\}. \end{aligned}\tag{2.2}$$

Let us look at the terms involving the variables y and z. Note that

$$\begin{aligned} &-2(y^2-z)v+\alpha_-\cdot 2(y_-^2-z)v+\alpha_+\cdot 2(y_+^2-z_+)|u_+|\\ &=2v\big[-(y^2-z)+\alpha_-(y_-^2-z_-)+\alpha_+(y_+^2-z_+)\big]+\alpha_+\cdot 2(y_+^2-z_+)(|u_+|-v)\\ &\geq 2v\big(\alpha_-e_-^2+\alpha_+e_+^2\big)-\alpha_+\cdot 2(|u_+|-v)=\frac{2\alpha_+}{\alpha_-}\cdot ve_+^2-2\alpha_+(|u_+|-v). \end{aligned}$$

Here in the last line we have used the identity $\alpha_-e_-+\alpha_+e_+=0$ (which implies $\alpha_-e_-^2+\alpha_+e_+^2=\frac{\alpha_+}{\alpha_-}e_+^2$). We come back to (2.2). Observe that

$$uy-\alpha_-u_-y_--\alpha_+u_+y_+=-\alpha_-d_-e_--\alpha_+d_+e_+=-\frac{\alpha_+}{\alpha_-}d_+e_+$$

and

$$-\frac{v}{2\sqrt{2(1-\varepsilon)}}\cdot\left(\frac{u^2}{v^2}+3\right)+\frac{\alpha_-v}{2\sqrt{2(1-\varepsilon)}}\cdot\left(\frac{u_-^2}{v^2}+3\right)=\frac{-\alpha_+u_+^2+\frac{\alpha_+}{\alpha_-}d_+^2-3\alpha_+v^2}{2\sqrt{2(1-\varepsilon)}v}.$$

Therefore, the estimate (2.2) will follow if we prove that

$$-\frac{\alpha_+}{\alpha_-}d_+e_++\frac{\frac{2\alpha_+}{\alpha_-}\cdot ve_+^2-2\alpha_+(|u_+|-v)}{2\sqrt{2(1-\varepsilon)}}+\frac{-\alpha_+u_+^2+\frac{\alpha_+}{\alpha_-}d_+^2-3\alpha_+v^2}{2\sqrt{2(1-\varepsilon)}v}+\frac{4\alpha_+|u_+|}{2\sqrt{2(1-\varepsilon)}}\geq 0.$$

After some straightforward manipulations, this inequality becomes

$$-\frac{\alpha_-}{v}(|u_+|-v)^2+2ve_+^2+\frac{d_+^2}{v}\geq 2\sqrt{2(1-\varepsilon)}d_+e_+.$$

Note that $0 \leq |u_+| - v \leq |u| + |d_+| - v \leq |d_+|$; furthermore, we have $\alpha_- \leq \varepsilon$, as we assume in the statement of the lemma. Consequently, the left-hand side above is not smaller than

$$(1-\varepsilon)\frac{d_+^2}{v} + 2ve_+^2 \geq 2\sqrt{2(1-\varepsilon)}d_+e_+,$$

which is the desired claim. □

Now we will establish the key geometric splitting lemma.

Lemma 2.2. *Let I be an arbitrary interval, let $f : I \to \mathbb{R}$ be an integrable function and pick $\varphi \in BMO(I)$ with $\|\varphi\|_{BMO} \leq 1$. In addition, let v be a positive number satisfying $|\langle f \rangle_I| \leq v$. Then there is a splitting $I = I_- \cup I_+$ with $\alpha_\pm := |I_\pm|/|I| \geq \varepsilon$ such that*

$$\begin{aligned} &B\left(\langle f\rangle_I, v, \langle\varphi\rangle_I, \langle\varphi^2\rangle_I\right) \\ &\geq \alpha_- B\left(\langle f\rangle_{I_-}, v_-, \langle\varphi\rangle_{I_-}, \langle\varphi^2\rangle_{I_-}\right) + \alpha_+ B\left(\langle f\rangle_{I_+}, v_+, \langle\varphi\rangle_{I_+}, \langle\varphi^2\rangle_{I_+}\right), \end{aligned} \tag{2.3}$$

where $v_\pm = \max\{v, |\langle f\rangle_{I_\pm}|\}$.

Remark 2.3. The assumptions $|\langle f\rangle_I| \leq v$ and $\|\varphi\|_{BMO} \leq 1$ imply that the points $\left(\langle f\rangle_I, v, \langle\varphi\rangle_I, \langle\varphi^2\rangle_I\right)$ and $\left(\langle f\rangle_{I_\pm}, v_\pm, \langle\varphi\rangle_{I_\pm}, \langle\varphi^2\rangle_{I_\pm}\right)$ lie in the domain of B.

Proof of Lemma 2.2. We may and do assume that $I = [0,1]$, by a simple affine change of variables. Furthermore, passing from f to $-f$ if necessary, we may assume that $\int_0^1 f \geq 0$. We consider separately three cases.

Case 1. Suppose that $\langle f\rangle_{[0,\varepsilon]}$ and $\langle f\rangle_{[1-\varepsilon,1]}$ are both not smaller or both not bigger than $\langle f\rangle_{[0,1]}$. Then there is a number $a \in [\varepsilon, 1-\varepsilon]$ such that $\langle f\rangle_{[0,a]} = \langle f\rangle_{[a,1]} = \langle f\rangle_{[0,1]}$. (This can be easily shown by Darboux property of the continuous function $F(t) = \langle f\rangle_{[0,t]} - \langle f\rangle_{[t,1]}$, $t \in [\varepsilon, 1-\varepsilon]$). Set $I_- = [0,a]$ and $I_+ = [a,1]$. Then $\alpha_\pm \geq \varepsilon$, $\langle f\rangle_{I_-} = \langle f\rangle_{I_+} = \langle f\rangle_I$, $v_- = v_+ = v$ and the desired inequality (2.3) is equivalent to

$$-(\langle\varphi\rangle_I^2 - \langle\varphi^2\rangle_I) \geq -\alpha_-(\langle\varphi\rangle_{I_-}^2 - \langle\varphi^2\rangle_{I_-}) - \alpha_+(\langle\varphi\rangle_{I_+}^2 - \langle\varphi^2\rangle_{I_+}).$$

But this follows at once from the convexity of the function $t \mapsto t^2$, since $\langle\varphi\rangle_I = \alpha_-\langle\varphi\rangle_{I_-} + \alpha_+\langle\varphi\rangle_{I_+}$ and $\langle\varphi^2\rangle_I = \alpha_-\langle\varphi^2\rangle_{I_-} + \alpha_+\langle\varphi^2\rangle_{I_+}$.

Case 2. Now, in contrast to the previous situation, we assume that $\langle f\rangle_{[0,1]}$ lies between $\langle f\rangle_{[0,\varepsilon]}$ and $\langle f\rangle_{[1-\varepsilon,1]}$. Replacing f with $f(2-\cdot)$ if necessary, we may assume that $\langle f\rangle_{[0,\varepsilon]} \leq \langle f\rangle_{[0,1]} \leq \langle f\rangle_{[1-\varepsilon,1]}$. Now, if $|\langle f\rangle_{[0,\varepsilon]}| \leq |\langle f\rangle_{[\varepsilon,1]}|$, then we take $I_- = [0,\varepsilon]$ and $I_+ = [\varepsilon,1]$. Then (2.3) follows from (2.1), applied to $\alpha_- = \varepsilon$, $\alpha_+ = 1-\varepsilon$, $u_- = \langle f\rangle_{[0,\varepsilon]}$, $u_+ = \langle f\rangle_{[\varepsilon,1]}$, $u = \langle f\rangle_I$, with $y_\pm$, y, $z_\pm$ and z defined as analogous averages for φ and φ^2. One proceeds similarly if $|\langle f\rangle_{[1-\varepsilon,1]}| \leq |\langle f\rangle_{[0,1-\varepsilon]}|$, taking $I_- = [1-\varepsilon,1]$ and $I_+ = [0,\varepsilon]$ and using (2.1), with the same choices for $\alpha_\pm$, u, y and z, but for u_-, y_-, z_- one takes the averages of f, φ and φ^2 over $[1-\varepsilon,1]$, and for u_+, y_+, z_+ - the averages over $[0,1-\varepsilon]$.

Case 3. It remains to consider the case in which $\langle f\rangle_{[0,\varepsilon]} \leq \langle f\rangle_{[0,1]} \leq \langle f\rangle_{[1-\varepsilon,1]}$, $|\langle f\rangle_{[0,\varepsilon]}| > |\langle f\rangle_{[\varepsilon,1]}|$ and $|\langle f\rangle_{[1-\varepsilon,1]}| > |\langle f\rangle_{[0,1-\varepsilon]}|$. If these conditions hold, then there is a number $a \in [\varepsilon, 1-\varepsilon]$ such that $\langle f\rangle_{[0,a]} + \langle f\rangle_{[a,1]} = 0$. To see this, recall that $\langle f\rangle_{[0,\varepsilon]} < \langle f\rangle_{[0,1]}$, which implies $\langle f\rangle_{[0,\varepsilon]} < \langle f\rangle_{[\varepsilon,1]}$ and hence $\langle f\rangle_{[0,\varepsilon]} < 0$ (because $|\langle f\rangle_{[0,\varepsilon]}| > |\langle f\rangle_{[\varepsilon,1]}|$, as we have assumed above). Thus

$$F(\varepsilon) = \langle f\rangle_{[0,\varepsilon]} + \langle f\rangle_{[\varepsilon,1]} = -|\langle f\rangle_{[0,\varepsilon]}| + \langle f\rangle_{[\varepsilon,1]} < 0.$$

Analogous reasoning shows that $F(1-\varepsilon) > 0$: we have $\langle f\rangle_{[1-\varepsilon,1]} > \langle f\rangle_{[0,1]} \geq 0$, so

$$F(1-\varepsilon) = \langle f\rangle_{[0,1-\varepsilon]} + \langle f\rangle_{[1-\varepsilon,1]} = \langle f\rangle_{[0,1-\varepsilon]} + |\langle f\rangle_{[1-\varepsilon,1]}| > 0.$$

This guarantees the existence of the parameter a, by the Darboux property of the function $t \mapsto \langle f\rangle_{[0,t]} + \langle f\rangle_{[t,1]}$. Set $I_- = [0,a]$ and $I_+ = [a,1]$; then $\alpha_\pm \geq \varepsilon$. Note that $\langle f\rangle_{I_-} = -\langle f\rangle_{I_+}$, by the very definition of parameter a. As in the proof of the previous lemma, we may assume that $|\langle f\rangle_{I_-}| \leq v$, since otherwise we replace v with $|\langle f\rangle_{I_-}|$, decreasing the left-hand side and not changing the right. But then $|\langle f\rangle_{I_\pm}| \leq v$ and the inequality was proved at the beginning of Lemma 2.1. □

Equipped with the above splitting lemma, we are ready for the proof of our main inequality.

Proof of (1.1) *with* $C = \sqrt{2}$. By homogeneity, we may restrict ourselves to $\|\varphi\|_{BMO} \leq 1$. In addition, by straightforward approximation, we may assume that f and φ are bounded, and $\langle f\rangle_I \neq 0$. Our next step is to construct inductively an increasing set of partitions of I: first we set $\mathcal{I}_0 = \{I\}$, and then, for any $n \geq 0$ and any $J \in \mathcal{I}_n$, we apply Lemma 2.2 to J and the functions $f|_J$, φ_J, obtaining its splitting $J_- \cup J_+$. All the intervals J_-, J_+, corresponding to different choices of $J \in \mathcal{I}_n$, are put into $\mathcal{I}_{n+1}$. Note that the condition $\alpha_\pm := |I_\pm|/|I| \geq \varepsilon$ appearing in the assertion of Lemma 2.2 guarantees that the diameter of $\mathcal{I}_n$ converges to zero as $n \to \infty$. By a straightforward induction, (2.3) implies that for any n we have

$$\sum_{J\in\mathcal{I}_n} \frac{|J|}{|I|} B\big(\langle f\rangle_J, v_J, \langle\varphi\rangle_J, \langle\varphi^2\rangle_J\big) \leq B\big(\langle f\rangle_I, |\langle f\rangle_I|, \langle\varphi\rangle_I, \langle\varphi^2\rangle_I\big),$$

where $v_J = \max\Big\{|\langle f\rangle_K| \;:\; K \in \mathcal{I}_j \text{ for some } j,\, J \subseteq K\Big\}$ is the 'truncated' maximal function of f, associated with the partitions $\mathcal{I}_0$, $\mathcal{I}_1$, …, $\mathcal{I}_n$. (Note that the assumption $\langle f\rangle_I \neq 0$, imposed at the beginning of the proof, guarantees that $v_J > 0$ for each J: this is required for the use of (2.3)). Now, observe that

$$B\big(\langle f\rangle_I, |\langle f\rangle_I|, \langle\varphi\rangle_I, \langle\varphi^2\rangle_I\big) = B\big(\langle f\rangle_I, |\langle f\rangle_I|, 0, \langle\varphi^2\rangle_I\big) \leq 0,$$

since $\langle\varphi^2\rangle_I \leq \langle\varphi\rangle_I^2 + 1 = 1$. Furthermore, for any $(u,v,y,z) \in D$ we have $y^2 - z \leq 0$, so

$$B(u,v,y,z) \geq uv - \frac{v}{2\sqrt{2(1-\varepsilon)}} \cdot 4 = uv - \sqrt{\frac{2}{1-\varepsilon}}v.$$

Consequently, the previous estimate yields

$$\sum_{J\in\mathcal{I}_n} |J|\langle f\rangle_J\langle\varphi\rangle_J \leq |I| \cdot \sqrt{\frac{2}{1-\varepsilon}} v_J \leq \sqrt{\frac{2}{1-\varepsilon}} \int_I \mathcal{M}f.$$

Now we let $n \to \infty$: then, as we have already noted above, the diameter of $\mathcal{I}_n$ tends to zero and thus, by Lebesgue's differentiation theorem, the left-hand side converges to $\int_I f\varphi$. It remains to observe that $\varepsilon > 0$ was arbitrary to complete the proof. □

3. Sharpness

Now we will construct explicit examples showing that the constant $C = \sqrt{2}$ is optimal in (1.1). Let $q \in (0,1)$ be a fixed parameter and introduce the families $(I_n^\pm)_{n=0}^\infty$ of intervals, given by $I_n^- = \big(q^{n+1}, (q^n + q^{n+1})/2\big]$, $I_n^+ = \big((q^n + q^{n+1})/2, q^n\big]$, $n = 0, 1, 2, \ldots$.

For any positive integer N, consider $f:[0,1]\to\mathbb{R}$ given by

$$f=-q^{-N}\chi_{[0,q^N]}+\sum_{n=0}^{N-1}q^{-n}(-\chi_{I_n^-}+\chi_{I_n^+}),$$

The L^1-norm of the maximal function of f behaves as follows.

Lemma 3.1. *We have*

$$\|\mathcal{M}f\|_{L^1}\leq 1+N\left(\frac{1}{q}-1\right).$$

Proof. We will prove the pointwise estimate

$$\mathcal{M}f(x)\leq q^{-N}\chi_{[0,q^N)}(x)+\sum_{n=0}^{N-1}q^{-n-1}\chi_{[q^{n+1},q^n)}(x),$$

from which the assertion follows immediately. If $x\leq q^{N-1}$, then $\mathcal{M}f(x)\leq\|f\|_{L^\infty([0,1])}=q^{-N}$, as needed. So, suppose that $x>q^{N-1}$: then there is a unique integer $n\leq N-2$ such that $x\in I_n^-\cup I_{n^+}$. Let $[a,b]$ be an arbitrary subinterval of $[0,1]$ containing x; we will prove that $\left|\frac{1}{b-a}\int_a^b f\right|\leq q^{-n-1}$. If $a\in I_n^-\cup I_n^-$ (i.e., $a\geq q^{n+1}$), then we proceed as previously: $\left|\frac{1}{b-a}\int_a^b f\right|\leq\|f\|_{L^\infty([a,b])}\leq q^{-n-1}$, as desired. So, suppose that $a<q^{n+1}$ and write

$$\begin{aligned}\left|\frac{1}{b-a}\int_a^b f\right|&=\left|\frac{q^{n+1}-a}{b-a}\cdot\frac{1}{q^{n+1}-a}\int_a^{q^{n+1}}f+\frac{b-q^{n+1}}{b-a}\cdot\frac{1}{b-q^{n+1}}\int_{q^{n+1}}^b f\right|\\&\leq\frac{q^{n+1}-a}{b-a}\left|\frac{1}{q^{n+1}-a}\int_a^{q^{n+1}}f\right|+\frac{b-q^{n+1}}{b-a}\left|\frac{1}{b-q^{n+1}}\int_{q^{n+1}}^b f\right|.\end{aligned}$$

Since $\frac{q^{n+1}-a}{b-a}+\frac{b-q^{n+1}}{b-a}=1$, this implies

$$\left|\frac{1}{b-a}\int_a^b f\right|\leq\max\left\{\left|\frac{1}{q^{n+1}-a}\int_a^{q^{n+1}}f\right|,\left|\frac{1}{b-q^{n+1}}\int_{q^{n+1}}^b f\right|\right\}.$$

We have already proved above that $\left|\frac{1}{b-q^{n+1}}\int_{q^{n+1}}^b f\right|\leq q^{-n-1}$. To handle the second expression under the above maximum, note that

$$\frac{1}{q^{n+1}-a}\int_a^{q^{n+1}}f=\frac{1}{q^{n+1}-a}\int_a^{q^m}f,$$

where m is defined as follows: $m=N$ if $a<q^N$, and otherwise, $m\geq n+1$ is uniquely determined by the double inequality $q^{m+1}<a<q^m$. If the first possibility occurs, we note that

$$\left|\frac{1}{q^{n+1}-a}\int_a^{q^m}f\right|=\left|\frac{1}{q^{n+1}-a}\int_a^{q^N}f\right|\leq\left|\frac{1}{q^{n+1}}\int_0^{q^N}f\right|=q^{-n-1}.$$

On the other hand, if $a>q^N$, then the expression $\left|\frac{1}{q^{n+1}-a}\int_a^{q^m}f\right|$, considered as a function of $a\in[q^{m+1},q^m)$, attains its maximum for $a=(q^m+q^{m+1})/2$, with the maximal value equal to

$$\frac{1}{q^{n+1}-(q^m+q^{m+1})/2}\cdot\frac{1-q}{2}\leq\frac{1}{q^{n+1}-(q^{n+1}+q^{n+2})/2}\cdot\frac{1-q}{2}=q^{-n-1}.$$

Hence the claim follows. □

Now define

$$\varphi=\sum_{n=0}^\infty\left[-\frac{n}{\sqrt{2}}\left(\frac{1}{q}-1\right)\chi_{I_n^-}+\left(-\frac{n}{\sqrt{2}}\left(\frac{1}{q}-1\right)+\sqrt{2}\right)\chi_{I_n^+}\right].$$

We compute that for any nonnegative integer n,

$$\begin{aligned}\int_0^{q^n}\varphi&=\sum_{k=n}^\infty\frac{q^k-q^{k+1}}{2}\left[-\frac{k}{\sqrt{2}}\left(\frac{1}{q}-1\right)-\frac{k}{\sqrt{2}}\left(\frac{1}{q}-1\right)+\sqrt{2}\right]\\&=-\frac{(1-q)^2}{\sqrt{2}}\sum_{k=n}^\infty kq^{k-1}+\frac{1-q}{\sqrt{2}}\sum_{k=n}^\infty q^k\\&=-q^n\cdot\frac{n(1-q)}{\sqrt{2}q}\end{aligned}\tag{3.1}$$

and

$$\begin{aligned}\int_0^{q^n}\varphi^2&=\sum_{k=n}^\infty\frac{q^k-q^{k+1}}{2}\left[\left(-\frac{k}{\sqrt{2}}\left(\frac{1}{q}-1\right)\right)^2+\left(-\frac{k}{\sqrt{2}}\left(\frac{1}{q}-1\right)+\sqrt{2}\right)^2\right]\\&=\frac{(1-q)^3}{2}\sum_{k=n}^\infty k(k-1)q^{k-2}+\frac{(1-q)^2(1-3q)}{2q}\sum_{k=n}^\infty kq^{k-1}+(1-q)\sum_{k=n}^\infty q^k\\&=q^n\left[\frac{n^2(1-q)^2}{2q^2}+\frac{1+q}{2q}\right].\end{aligned}\tag{3.2}$$

We will prove the following estimate for the BMO norm of φ.

Lemma 3.2. *We have* $\|\varphi\|_{BMO}^2\leq\frac{1}{8}\left(\frac{1}{q}-1\right)^2+\frac{1+q}{2q}$.

Proof. It is convenient to split the reasoning into a few intermediate steps.

Step 1. On the approach. We will check that for any $[a,b]\subseteq[0,1]$ we have

$$\langle\varphi^2\rangle_{[a,b]}-\langle\varphi\rangle_{[a,b]}^2\leq\frac{1}{8}\left(\frac{1}{q}-1\right)^2+\frac{1+q}{2q},\tag{3.3}$$

i.e., we will show that the point $(\langle\varphi\rangle_{[a,b]},\langle\varphi^2\rangle_{[a,b]})$ lies on, or below the parabola $y=x^2+\frac{1}{8}\left(\frac{1}{q}-1\right)^2+\frac{1+q}{2q}$. We will frequently use the following simple observation, which enables us to avoid most of the technical issues. Namely, for $0\leq a<c<b\leq 1$, we have

$$(\langle\varphi\rangle_{[a,b]},\langle\varphi^2\rangle_{[a,b]})=\frac{c-a}{b-a}(\langle\varphi\rangle_{[a,c]},\langle\varphi^2\rangle_{[a,c]})+\frac{b-c}{b-a}(\langle\varphi\rangle_{[c,b]},\langle\varphi^2\rangle_{[c,b]}),\tag{3.4}$$

i.e., the point $(\langle\varphi\rangle_{[a,b]},\langle\varphi^2\rangle_{[a,b]})$ is a convex combination of the points $(\langle\varphi\rangle_{[a,c]},\langle\varphi^2\rangle_{[a,c]})$ and $(\langle\varphi\rangle_{[c,b]},\langle\varphi^2\rangle_{[c,b]})$ (in particular, all three are colinear). A similar observation holds true if $[a,b]$ is split into several intervals J_1, J_2, …, J_k: then $(\langle\varphi\rangle_{[a,b]},\langle\varphi^2\rangle_{[a,b]})$ is a convex combination of $(\langle\varphi\rangle_{J_1},\langle\varphi^2\rangle_{J_1})$, $(\langle\varphi\rangle_{J_2},\langle\varphi^2\rangle_{J_2})$, …, $(\langle\varphi\rangle_{J_k},\langle\varphi^2\rangle_{J_k})$.

Step 2. Auxiliary parameters and their geometric interpretation. For any nonnegative integer n, define the points $P_n = \left(\langle\varphi\rangle_{[0,q^n]}, \langle\varphi^2\rangle_{[0,q^n]}\right)$ and $A_n^{\pm} = \left(\langle\varphi\rangle_{I_n^{\pm}}, \langle\varphi^2\rangle_{I_n^{\pm}}\right)$. By (3.1) and (3.2), the point P_n lies on the parabola $y = x^2 + \frac{1+q}{2q}$; furthermore, since φ is constant on I_n^- and I_n^+, the points A_n^- and A_n^+ lie on the parabola $y = x^2$. A straightforward computation shows that the line P_nP_{n+1} is described by the equation

$$y = -\frac{2n+1}{\sqrt{2}}\left(\frac{1}{q}-1\right)\left(x + \frac{n}{\sqrt{2}}\left(\frac{1}{q}-1\right)\right) + \frac{n^2}{2}\left(\frac{1}{q}-1\right)^2 + \frac{1+q}{2q}$$

and is tangent to the parabola $y = x^2 + \frac{1}{8}\left(\frac{1}{q}-1\right)^2 + \frac{1+q}{2q}$. Finally, let Q_n be the intersection point of the lines $A_n^-P_{n+1}$ and $P_nA_n^+$. See Figure 1 below.

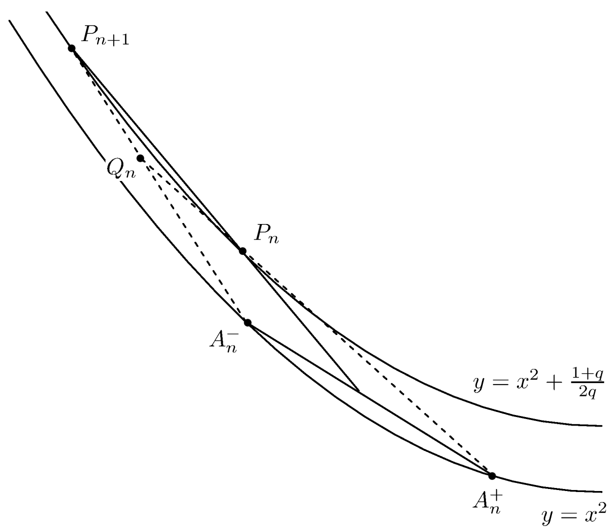

FIGURE 1. The distinguished points arising in the analysis of the BMO norm of φ. The line P_nP_{n+1} passes through the center of $A_n^-A_n^+$.

Step 3. Proof of (3.3) *for* $a = 0$. We start with the observation that the point $(\langle\varphi\rangle_{[0,(q^n+q^{n+1})/2]}, \langle\varphi^2\rangle_{[0,(q^n+q^{n+1})/2]})$ is precisely Q_n: indeed, by (3.4), it lies on the line segment $A_n^-P_{n+1}$ (since $[0,(q^n+q^{n+1})/2] = [0,q^{n+1}] \cup I_n^-$) and on the line $P_nA_n^+$ (since $[0,q^n] = [0,(q^n+q^{n+1})/2] \cup I_n^+$).

Now, suppose that $b \in I_n^-$ for some nonnegative integer n. Since $[0,(q^n+q^{n+1})/2] = [0,b] \cup [b,(q^n+q^{n+1})/2]$ and the average of φ (or φ^2) over $[b,(q^n+q^{n+1})/2]$ is the same as the average over I_n^-, (3.4) implies that the point $(\langle\varphi\rangle_{[0,b]}, \langle\varphi^2\rangle_{[0,b]})$ lies on the segment Q_nP_{n+1}, which, in turn, is located under the parabola $y = x^2 + \frac{1}{8}\left(\frac{1}{q}-1\right)^2 + \frac{1+q}{2q}$. A similar argument shows that if $b \in I_n^+$, then $(\langle\varphi\rangle_{[0,b]}, \langle\varphi^2\rangle_{[0,b]})$ belongs to the segment Q_nP_n, which again lies below the parabola.

Step 4. Proof of (3.3) *for arbitrary* $a < b$. This is the most elaborate part. Let n be an integer uniquely determined by the requirement $b \in I_n^- \cup I_n^+$ (i.e., $q^{n+1} < b \leq q^n$). We consider three cases.

1° Assume that $a \in I_m^- \cup I_m^+$ such that A_m^- lies below the line passing ℓ through A_n^+ and tangent to $y = x^2 + \frac{1+q}{2q}$. Then the point $(\langle\varphi\rangle_{[a,b]}, \langle\varphi^2\rangle_{[a,b]})$ is a convex combination of $A_m^\pm$, $A_{m-1}^\pm$, ..., $A_n^\pm$ lying below ℓ. Therefore it also lies below ℓ, and hence, in turn, also below the parabola $y = x^2 + \frac{1}{8}\left(\frac{1}{q} - 1\right)^2 + \frac{1+q}{2q}$.

2° Now suppose that $a \in I_m^- \cup I_m^+$ such that A_m^- lies on or above the line ℓ considered in the previous case. Assume additionally that the point $(A_m^- + A_m^+)/2$ lies below the line P_nP_{n+1}. Then A_m^+ lies below P_nP_{n+1} and hence, by (3.4), so does $(\langle\varphi\rangle_{[a,q^m]}, \langle\varphi^2\rangle_{[a,q^m]})$. A similar argument shows that the point $(\langle\varphi\rangle_{[q^{n+1},b]}, \langle\varphi^2\rangle_{[q^{n+1},b]})$ lies on or below P_nP_{n+1}. Therefore, the point $(\langle\varphi\rangle_{[a,b]}, \langle\varphi^2\rangle_{[a,b]})$, being the convex combination of $(\langle\varphi\rangle_{[a,q^m]}, \langle\varphi^2\rangle_{[a,q^m]})$, $(\langle\varphi\rangle_{[q^{n+1},b]}, \langle\varphi^2\rangle_{[q^{n+1},b]})$ and $A_{m-1}^\pm$, $A_{m-2}^\pm$, ..., $A_{n+1}^\pm$, lies below P_nP_{n+1} and hence also below the parabola $y = x^2 + \frac{1}{8}\left(\frac{1}{q} - 1\right)^2 + \frac{1+q}{2q}$.

3° Finally, suppose that $a \in I_m^- \cup I_m^+$ such that $(A_m^- + A_m^+)/2$ lies on or above the line P_nP_{n+1}. This means that he point $(\langle\varphi\rangle_{[0,a]}, \langle\varphi^2\rangle_{[0,a]})$ lies above P_nP_{n+1} (being the convex combination of points with this property). In Step 3 we have proved that $(\langle\varphi\rangle_{[0,b]}, \langle\varphi^2\rangle_{[0,b]})$ lies on or below this line. Consequently, by (3.4), the point $(\langle\varphi\rangle_{[a,b]}, \langle\varphi^2\rangle_{[a,b]})$ lies below P_nP_{n+1} and hence also below the parabola $y = x^2 + \frac{1}{8}\left(\frac{1}{q} - 1\right)^2 + \frac{1+q}{2q}$. □

We are ready to show the optimality of the constant $\sqrt{2}$ in (1.1). We compute that

$$\begin{aligned}&\int_0^1 f\varphi \\ &= -\frac{1}{q^N}\int_0^{q^N} \varphi + \sum_{n=0}^{N-1} q^{-n} \cdot \frac{q^n - q^{n+1}}{2} \cdot \left[\frac{n}{\sqrt{2}}\left(\frac{1}{q} - 1\right) + \left(-\frac{n}{\sqrt{2}}\left(\frac{1}{q} - 1\right) + \sqrt{2}\right)\right] \\ &= \frac{N}{\sqrt{2}}\left(\frac{1}{q} - 1\right) + \frac{N}{\sqrt{2}}(1 - q),\end{aligned}$$

where in the last line we have exploited (3.1). Therefore, Lemmas 3.1 and 3.2 imply

$$\frac{\int_0^1 f\varphi}{\|\mathcal{M}f\|_{L^1}\|\varphi\|_{BMO}} \geq \frac{\frac{N}{\sqrt{2}}\left(\frac{1}{q} - 1\right) + \frac{N}{\sqrt{2}}(1 - q)}{\left(1 + N\left(\frac{1}{q} - 1\right)\right)\left(\frac{1}{8}\left(\frac{1}{q} - 1\right)^2 + \frac{1+q}{2q}\right)^{1/2}}.$$

Letting $N \to \infty$, we see that the best constant in (1.1) cannot be smaller than

$$\frac{\frac{1}{\sqrt{2}}\left(\frac{1}{q} - 1\right) + \frac{1}{\sqrt{2}}(1 - q)}{\left(\frac{1}{q} - 1\right)\left(\frac{1}{8}\left(\frac{1}{q} - 1\right)^2 + \frac{1+q}{2q}\right)^{1/2}} = \frac{q + 1}{\sqrt{2}\left(\frac{1}{8}\left(\frac{1}{q} - 1\right)^2 + \frac{1+q}{2q}\right)^{1/2}}.$$

It remains to note that the latter expression increases from zero to $\sqrt{2}$ if we let $q \to 1$. This establishes the desired sharpness.

Acknowledgment

The author would like to thank an anonymous Referee for the careful reading of the first version of the paper and several helpful suggestions.

References

[1] Bellman, R., Dynamic programming. Reprint of the 1957 edition. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 2010.

[2] Bennett, C., DeVore, R. A. and Sharpley, R.: *Weak-L^∞ and BMO*. Ann. of Math. **113**, 601–611 (1981).

[3] Chou, C. S. Sur certaines généralisations de línégalité de Fefferman. Séminaire de Probabilités XVIII, Lecture Notes in Mathematics 1059, 219–222, Springer Verlag, 1984.

[4] Choulli, T., Stricker, C. and Krawczyk, L. *On Fefferman and Burkholder-Davis-Gundy inequalities for $\mathcal{E}$-martingales.* Probab. Theory Related Fields **113** (1999), no. 4, 571–597.

[5] Fefferman, C.: *Characterization of bounded mean oscillation.* Bull. Amer. Math. Soc. **77**, 587–588 (1971).

[6] Garsia, A. M. Martingale Inequalities: Seminar Notes on Recent Progress. Mathematics Lecture Notes Series, Benjamin, Reading, 1973.

[7] Getoor, R. K. and Sharpe, M. J.. *Conformal martingales.* Invent. Math. **16**, 271–308 (1972) .

[8] Grafakos, L.: Classical Fourier Analysis. Graduate Texts in Mathematics, 2 ed., Springer, New York (2008).

[9] John, F. and Nirenberg, L.: *On functions of bounded mean oscillation.* Comm. Pure and Appl. Math. **14**, 415–426 (1961).

[10] Nazarov F. L. and Treil, S. R. *The hunt for a Bellman function: applications to estimates for singular integral operators and to other classical problems of harmonic analysis,* St. Petersburg Math. J. **8** (1997), 721–824.

[11] Nazarov, F. L., Treil S. R. and Volberg, A. *The Bellman functions and two-weight inequalities for Haar multipliers.* J. Amer. Math. Soc., **12** (1999), pp. 909–928.

[12] Pratelli, M. Sur certains espaces de martingales localement de carré intégrable. Séminaire de Probabilités X, Lecture Notes in Mathematics 511, 401–413, Springer Verlag, 1976.

[13] Slavin, L. and Vasyunin, V.: *Sharp results in the integral-form John-Nirenberg inequality.* Trans. Amer. Math. Soc. **363**, 4135–4169 (2011).

[14] Slavin, L. and Vasyunin, V.: *Sharp L^p estimates on BMO.* Indiana Math. J. **61**, 1051–1110 (2012).

[15] Slavin, L. and Volberg, A.: Bellman function and the $H^1 - BMO$ duality. Harmonic analysis, partial differential equations, and related topics, 113–126, Contemp. Math., 428, Amer. Math. Soc., Providence, RI, 2007.

[16] Vasyunin, V. and Volberg, A.: *Sharp constants in the classical weak form of the John-Nirenberg inequality.* Proc. London Math. Soc. **108**, 1417–1434 (2014).

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

Email address: A.Osekowski@mimuw.edu.pl