EMBEDDING BMO INTO WEIGHTED BMO

ADAM OSĘKOWSKI

Abstract. A classical result of harmonic analysis asserts that if a weight w satisfies Muckenhoupt's condition A_∞, then the unweighted class BMO is contained in the weighted space $BMO(w)$. The paper identifies the norm of this embedding in the one-dimensional setting. Specifically, for any function $f \in BMO(\mathbb{R})$ and any weight $w \in A_\infty(\mathbb{R})$ of characteristic $[w]_{A_\infty}$, we have the estimate

$$\|f\|_{BMO(w)} \leq e\sqrt{2}[w]_{A_\infty}\|f\|_{BMO}.$$

The constant $e\sqrt{2} = 3.8442\ldots$ is the best possible. We also prove a sharp version of this result in which the characteristic $[w]_{A_\infty}$ of the weight is fixed. Further extensions to the theory of martingales are obtained.

1. Introduction

The principal purpose of this paper is to compare the BMO norms of a function in the weighted and the unweighted setting, under the assumption that the weight satisfies Muckenhoupt's condition A_∞. Let us start with the necessary background and notation. Suppose that f is a real-valued locally integrable function on $\mathbb{R}^n$. It belongs to BMO, the class of functions of bounded mean oscillation, if we have

$$\sup_Q \left\langle |f - \langle f\rangle_Q|\right\rangle_Q < \infty, \tag{1.1}$$

where the supremum is taken over all cubes Q in $\mathbb{R}^n$ with edges parallel to the coordinate axes and

$$\langle f\rangle_Q = \frac{1}{|Q|}\int_Q f(x)\mathrm{d}x$$

is the average of f over Q. The space BMO was introduced by John and Nirenberg in [6] and has turned out to play a fundamental role in analysis and probability. For example, many classical operators (e.g., singular integrals, wide classes of Fourier multipliers, etc.) map L^∞ into BMO; this space also behaves nicely from the viewpoint of interpolation. Let us also mention the remarkable result of Fefferman [2] which asserts that the space BMO is dual to the Hardy space H^1. We refer the interested reader to any textbook on harmonic analysis for more on the subject, its connections and applications.

It is well-known that the functions of bounded mean oscillation have very strong integrability properties (see e.g. [6]). In particular, the p-oscillations

$$||f||_{BMO^p} := \sup_Q \left\langle |f - \langle f\rangle_Q|^p\right\rangle_Q^{1/p}, \quad 1 < p < \infty, \tag{1.2}$$

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

are finite for any $f \in BMO$, and form a family of equivalent quasinorms on $BMO(\mathbb{R}^n)$. In our considerations, we will work with $\|\cdot\|_{BMO^2}$ and denote it simply by $\|\cdot\|_{BMO}$. One of the reasons we choose this particular norm is that we have the convenient identity

$$||f||_{BMO^2} = \sup_Q \left\{ \langle f^2 \rangle_Q - \langle f \rangle_Q^2 \right\}^{1/2}. \tag{1.3}$$

In what follows, we will be interested in the weighted context. Suppose that w is a weight, i.e., a positive, locally integrable function on $\mathbb{R}^d$ or $\mathcal{Q}$, depending on the context. This function gives rise to the associated measure $w\mathrm{d}x$, i.e., defined by the formula $w(Q) = \int_Q w\mathrm{d}x$. Then one defines the associated weighted BMO space as before, but replacing the norm by

$$\|f\|_{BMO(w)} = \sup_Q \left\langle \left(f - \langle f \rangle_{Q,w}\right)^2 \right\rangle_{Q,w}^{1/2},$$

where $\langle f \rangle_{Q,w} = \frac{1}{w(Q)} \int_Q fw\mathrm{d}x$ is the weighted average.

There is a natural question about those weights w, for which we have the inclusion $BMO \hookrightarrow BMO(w)$. A classical and well-known result asserts that this embedding holds true if w belongs to the Muckenhoupt's class A_∞. The latter means that

$$[w]_{A_\infty} := \sup \left\langle w \right\rangle_Q \exp\left(- \left\langle \log w \right\rangle_Q \right) < \infty, \tag{1.4}$$

where the supremum is taken over all cubes Q, with axes parallel to the axes, contained in the domain of w. (The original definition of Muckenhoupt was slightly different, the above formulation, due to Khrushchev [8], will be more convenient for our purposes). The quantity $[w]_{A_\infty}$ will be called the characteristic of the weight w. By Jensen's inequality we have $[w]_{A_\infty} \geq 1$, and equality holds for weights which are constant almost everywhere. Roughly speaking, the characteristic measures how disbalanced the weight is: the bigger $[w]_{A_\infty}$, the more oscillations of w should be expected.

One of the main results of this paper is the identification of the norm of the embedding $BMO \hookrightarrow BMO(w)$ in the one-dimensional setting. Here is the precise statement.

Theorem 1.1. *For any function $f \in BMO(\mathbb{R})$ and any weight $w \in A_\infty(\mathbb{R})$ we have the estimate*

$$\|f\|_{BMO(w)} \leq e\sqrt{2}[w]_{A_\infty} \|f\|_{BMO}. \tag{1.5}$$

The constant $e\sqrt{2} = 3.8442\ldots$ is the best possible.

Actually, we will prove a much more precise statement: we will establish a sharp version of the above inequality when restricted to the class of weights of a prescribed characteristic. To formulate the result, we need to introduce a certain special parameter. Namely, for a given $c \geq 1$, let $d = d(c)$ be a number belonging to $[0, 1)$ which satisfies

$$c(1-d)e^d = 1. \tag{1.6}$$

Such a number d exists and is unique: this follows at once from Darboux property and the fact that the function $s \mapsto (1-s)e^s$ is strictly decreasing on $[0, 1]$.

Theorem 1.2. *Let $c \geq 1$ be a fixed parameter and let $d = d(c)$ be the solution to* (1.6). *For any function $f \in BMO(\mathbb{R})$ and any weight $w \in A_\infty(\mathbb{R})$ satisfying $[w]_{A_\infty} \leq c$ we have the estimate*

$$\|f\|_{BMO(w)} \leq \frac{2^{d/2}}{1-d} \|f\|_{BMO}. \tag{1.7}$$

The constant $2^d/(1-d)$ is the best possible.

At some points we will deal with the localized versions of BMO spaces and A_∞ weights. Suppose that $\mathcal{Q} \subset \mathbb{R}^n$ is a given base cube. Then one defines the spaces $BMO(\mathcal{Q})$ and $A_\infty(\mathcal{Q})$ as above, the only difference is that in (1.1) and (1.4) one needs to take the suprema over all cubes Q contained in $\mathcal{Q}$ (with sides parallel to the axes). Theorems 1.1 and 1.2 remains valid (with unchanged constants) in the localized one-dimensional setting, i.e., when f and w are defined on an arbitrary interval contained in $\mathbb{R}$. Our reasoning, combined with the results and the arguments from [16], imply the validity of Theorems 1.1 and 1.2 also in the context of functions on the circle $\mathbb{T}$ (i.e., 1-periodic functions/weights on $\mathbb{R}$).

A few words about the proof are in order. The efficient control of the norm $\|f\|_{BMO(w)}$ requires the effective handling of the integral $\int_I f^2 w \mathrm{d}x$, where $I \subset \mathbb{R}$ is an arbitrary interval. The problem is that the functions f, w evolve 'independently' and according to their own restrictions (coming from the BMO property and the A_∞ condition), and it seems difficult to deal with them simultaneously. Quite unexpectedly, the following simple idea turned out to be successful. Namely, the use of the Young inequality allows to estimate the integral $\int_I f^2 w \mathrm{d}x$ from above by the sum $\int_I \Phi(|f|)\mathrm{d}x + \int_I \Psi(w)\mathrm{d}w$, for a wide class of functions Φ and Ψ. Now each of the summands depends on one function only and can be handled more easily with the use of the so-called Bellman function method (which is very well understood in these contexts: see [4, 9, 10, 11, 12, 13, 17]). However, it was quite surprising to the author that the functions Φ, Ψ can be chosen appropriately so that the information on the best constants is not lost on the way. We shall see that the functions Φ and Ψ above have quite involved formulas; we discovered them simply by guessing what the extremal f and w in (1.7) should be, and then choosing Φ and Ψ for which both sides of the corresponding Young inequality were the same. See Section 5 for details.

The remaining part of the paper is organized as follows. The next section contains the formulas for the functions Φ and Ψ as well as the proofs of sharp upper bounds for $\int_I \Phi(|f|)\mathrm{d}x$ and $\int_I \Psi(w)\mathrm{d}w$. Section 3 is devoted to the proofs of our main results, Theorems 1.1 and 1.2. Section 4 contains extensions of (1.5) and (1.7) to the probabilistic setting. The final part of the paper contains the description of some steps which have led us to the discovery of the key special functions Φ and Ψ used in the proof.

2. Two auxiliary inequalities

In this section we prove a sharp exponential estimate for BMO functions and a sharp logarithmic bound for A_∞ weights. Throughout, $c \geq 1$ is a fixed parameter and $d = d(c) \in [0,1)$ is the unique solution to (1.6). Let Φ, $\Psi : [0,\infty) \to [0,\infty)$ be given by

$$\Phi(t) = 2^d(1-d)\int_0^{t^2} e^{d\sqrt{s}}\mathrm{d}s$$

and

$$\Psi(t) = \begin{cases} 0 & \text{if } t \leq 2^d(1-d), \\ \displaystyle\int_{2^d(1-d)}^t \log^2\left(2\left(\frac{1-d}{s}\right)^{1/d}\right)\mathrm{d}s & \text{if } t \geq 2^d(1-d). \end{cases}$$

Both formulas for Φ and Ψ can be expressed explicitly, but we prefer the above more concise forms, which make the calculations shorter. We easily see that Φ is of class C^∞ on $(0,\infty)$, while Ψ is of class C^1 there (it is of class C^∞ if we remove the point $2^d(1-d)$).

The main results of this section can be formulated as follows.

Theorem 2.1. *Let $I \subset \mathbb{R}$ be an arbitrary interval. If $f \in BMO$ satisfies $\|f\|_{BMO} \leq 1$ and $\langle f\rangle_I = 0$, then*

$$\frac{1}{|I|}\int_I \Phi(|f|)\,ds \leq \frac{2^d}{1-d}. \tag{2.1}$$

Theorem 2.2. *Let $I \subset \mathbb{R}$ be an arbitrary interval. For any A_∞ weight w satisfying $\langle w\rangle_I = 1$ and $[w]_{A_\infty} \leq c$, we have*

$$\frac{1}{|I|}\int_I \Psi(w)\,ds \leq \frac{d2^d}{(1-d)^2}. \tag{2.2}$$

In the case $c = 1$ we have $d(c) = 0$ and the inequalities (2.1), (2.2) are trivial. Indeed, then $\Phi(t) = t^2$ and the first estimate is equivalent to $\langle f^2\rangle_I \leq 1$ (which holds since $\langle f^2\rangle_I \leq \|f\|_{BMO}^2$); concerning (2.2), the condition $[w]_{A_\infty} \leq 1$ implies that w is constant (and hence $w \equiv 1$, because of the assumption $\langle w\rangle_I = 1$) and the inequality is equivalent to $0 \leq 0$. Therefore, from now on, we assume that c is bigger than 1; by (1.6), the parameter $d(c)$ must be strictly positive.

We will handle the above theorems in two separate subsections below.

2.1. **An exponential estimate for BMO functions.** Consider the auxiliary function $F : [0,\infty) \to [0,\infty)$ defined by

$$F(t) = e^t \int_t^\infty e^{-s}\Phi(s)\mathrm{d}s.$$

The Bellman function $B : \{(x,y) \in \mathbb{R}^2 : x^2 \leq y \leq x^2+1\} \to \mathbb{R}$ is given by the formula

$$B(x,y) = \begin{cases} \frac{y}{2}F(0) & \text{if } 2|x| < y,\, |x| \leq 1,\\ \sqrt{x^2-y+1}\,\Phi\left(|x| + \sqrt{x^2-y+1} - 1\right) & \\ \qquad +(1-\sqrt{x^2-y+1})F\left(|x| + \sqrt{x^2-y+1} - 1\right) & \text{otherwise.}\end{cases}$$

It is easy to check that B is continuous on its domain and of class C^1 in the interior.

Lemma 2.3. *The function B satisfies $B_y \geq 0$ and is locally concave, i.e., concave along any line segment contained in the domain.*

Proof. If $2|x| \leq y$ and $|x| \leq 1$, then $B_y(x,y) = F(0)/2$. For remaining (x,y), we use the equation $F' = F - \Phi$ and compute that

$$B_y(x,y) = \frac{-\Phi - \Phi' + F}{2} = \frac{1}{2}\exp(|x| + \sqrt{x^2-y+1} - 1)\int_{|x|+\sqrt{x^2-y+1}-1}^\infty e^{-s}\Phi''(s)\mathrm{d}s$$

(the functions Φ, Φ' and F are evaluated at the point $|x| + \sqrt{x^2-y+1} - 1$ and the last equality follows from integration by parts). The expression on the right is obviously nonnegative, so the first part of the lemma is established. To show the second part, we start with the observation that B is linear on $\{(x,y) : x^2 \leq y \leq x^2+1,\, 2|x| \leq y, |x| \leq 1\}$. Hence it is enough to prove that for (x,y) belonging to the interior of the remaining part of the domain, the Hessian matrix $D^2B(x,y)$ is seminegative-definite. To this end, note that the function B is linear along the line segment of slope $a = 2x\big(1 + \sqrt{x^2 - y + 1}/|x|\big)$ passing through (x, y). Therefore, we get

$$B_{xx}(x,y) + 2aB_{xy}(x,y) + a^2 B_{yy}(x,y) = 0. \tag{2.3}$$

Now, let us compute the partial derivatives B_{xy} and B_{yy}, differentiating the formula for B_y, obtained above, with respect to x and y. Integration by parts reveals that

$$\begin{aligned} &B_{xy}(x,y) \\ &= \frac{1}{2}\left(\frac{x}{|x|} + \frac{x}{\sqrt{x^2-y+1}}\right)\exp(|x| + \sqrt{x^2-y+1} - 1)\int_{|x|+\sqrt{x^2-y+1}-1}^{\infty} e^{-s}\Phi'''(s)\mathrm{d}s \end{aligned}$$

and

$$B_{yy}(x,y) = -\frac{1}{4\sqrt{x^2-y+1}}\exp(|x| + \sqrt{x^2-y+1} - 1)\int_{|x|+\sqrt{x^2-y+1}-1}^{\infty} e^{-s}\Phi'''(s)\mathrm{d}s.$$

Consequently, we obtain $aB_{yy}(x,y) + B_{xy}(x,y) = 0$, which combined with (2.3) yields $aB_{xy}(x,y) + B_{xx}(x,y) = 0$ and implies that the Hessian matrix $D^2B(x,y)$ has determinant zero. It remains to note that $B_{yy}(x,y) \le 0$, which is evident from the above formula. $\square$

Note that $B(x, x^2) = \Phi(|x|)$: the Bellman function coincides with Φ on the lower part of its domain. This observation, combined with the lemma above, yields the validity of the BMO estimate (2.1). This follows from Theorem 4.13 in [15], which can be used to study much more general estimates, far beyond the BMO context.

For the reader's convenience, we have decided to present below how to deduce (2.1). The starting point is the following auxiliary lemma, which can be found in [12] (consult Lemma 4c there).

Lemma 2.4. *Fix $\varepsilon < 1$. Then for every interval I and every $f : I \to \mathbb{R}$ with $||f||_{BMO(I)} \le \varepsilon$, there exists a splitting $I = I_- \cup I_+$ such that the whole straight-line segment with the endpoints $(\langle f\rangle_{I_\pm}, \langle f^2\rangle_{I_\pm})$ is contained within Ω. Moreover, the splitting parameter $\alpha = |I_+|/|I|$ can be chosen uniformly (with respect to f and I) separated from 0 and 1.*

We proceed to the estimate for BMO functions.

Proof of (2.1). Let $\varepsilon \in (0,1)$ be a fixed parameter and define $\tilde f = \varepsilon f$; then $||\tilde f||_{BMO(I)} \le \varepsilon$. Consider the family $\{\mathcal{I}^n\}_{n\ge 0}$ of partitions of I, generated by the inductive use of Lemma 2.4. We start with $\mathcal{I}^0 = \{I\}$; then, given $\mathcal{I}^n = \{I^{n,1}, I^{n,2}, \dots, I^{n,2^n}\}$, we split each $I^{n,k}$ according to Lemma 2.4, applied to the function $\tilde f$, and put

$$\mathcal{I}^{n+1} = \big\{I_-^{n,1}, I_+^{n,1}, I_-^{n,2}, I_+^{n,2}, \dots, I_-^{n,2^n}, I_+^{n,2^n}\big\}.$$

Next, we define functional sequences $(f_n)_{n\ge 0}$ and $(g_n)_{n\ge 0}$ by the formulas

$$f_n(x) = \langle \tilde f\rangle_{I^n(x)} \quad \text{and} \quad g_n(x) = \langle \tilde f^2\rangle_{I^n(x)},$$

where $I^n(x) \in \mathcal{I}^n$ is an interval containing x (if there are two such intervals, we pick the one which has x as its right endpoint). An important observation, which is the consequence of the fact that we work with $||\cdot||_{BMO^2}$-norm, is that for each n the pair (f_n, g_n) takes values in $\{(x,y) : x^2 \le y \le x^2 + \varepsilon^2\}$. Indeed, for any $J \in \mathcal{I}^n$ we have

$$0 \le \langle \tilde f^2\rangle_J - \langle \tilde f\rangle_J^2 \le \varepsilon^2,$$

where the left bound is due to Schwarz inequality, and the right follows from (1.3) and the assumption $||\tilde{f}||_{BMO} \leq \varepsilon$. Now, we will show that for any $n \geq 0$ and any $1 \leq k \leq 2^n$,

$$\int_{I^{n,k}} B(f_n(s), g_n(s))\mathrm{d}s \geq \int_{I^{n,k}} B(f_{n+1}(s), g_{n+1}(s))\mathrm{d}s. \tag{2.4}$$

To do this, note that f_n, g_n are constant on $I^{n,k}$, while f_{n+1}, g_{n+1} are constant on the intervals $I_\pm^{n,k}$ into which $I^{n,k}$ splits; therefore, dividing both sides by $|I^{n,k}|$, we see that the above estimate is equivalent to

$$\begin{aligned} B\big(\langle \tilde{f}\rangle_{I^{n,k}}, \langle \tilde{f}^2\rangle_{I^{n,k}}\big) \geq \frac{|I_-^{n,k}|}{|I^{n,k}|} B\left(\langle \tilde{f}\rangle_{I_-^{n,k}}, \langle \tilde{f}^2\rangle|_{I_-^{n,k}}\right) \\ + \frac{|I_+^{n,k}|}{|I^{n,k}|} B\left(\langle \tilde{f}\rangle_{I_+^{n,k}}, \langle \tilde{f}^2\rangle_{I_+^{n,k}}\right). \end{aligned}$$

This bound follows from the local concavity of B and the fact that the whole line segment with endpoints $\big(\langle \tilde{f}\rangle_{I_\pm^{n,k}}, \langle \tilde{f}^2\rangle_{I_\pm^{n,k}}\big)$ is contained in $\{(x,y) : x^2 \leq y \leq x^2+1\}$ (which is guaranteed by Lemma 2.4). Summing (2.4) over all $k = 1, 2, \ldots, 2^n$, we get

$$\int_I B(f_n(s), g_n(s))\mathrm{d}s \geq \int_I B(f_{n+1}(s), g_{n+1}(s))\mathrm{d}s$$

and hence, by induction,

$$\int_I B(f_0(s), g_0(s))\mathrm{d}s \geq \int_I B(f_n(s), g_n(s))\mathrm{d}s \tag{2.5}$$

for any $n = 0, 1, 2, \ldots$. But $f_0 \equiv 0$, since f is assumed to have vanishing integral; furthermore, we have $g_0(s) \leq f_0(s)^2 + 1 = 1$ and hence by the first part of Lemma 2.3,

$$B(f_0(s), g_0(s)) \leq B(0,1) = \frac{1}{2}F(0).$$

To deal with the right-hand side of (2.5), let n go to infinity. Since the splitting ratio of Lemma 2.4 is bounded away from 0 and 1, we see that the diameter of the partition $\mathcal{I}^n$ (i.e., $\sup_{1\leq k\leq 2^n} |I^{n,k}|$) tends to 0. Consequently, by Lebesgue's differentiation theorem, we have $f_n \to \varepsilon f$ and $g_n \to \varepsilon^2 f^2$ almost everywhere on I. Combining the above facts with Fatou's lemma, we see that (2.5) leads to

$$\frac{1}{|I|}\int_I B(\varepsilon f(s), \varepsilon^2 f^2(s))\mathrm{d}s = \frac{1}{|I|}\int_I B(\tilde{f}(s), \tilde{f}^2(s))\mathrm{d}s \leq \frac{1}{2}F(0),$$

or, since $B(x, x^2) = \Phi(|x|)$,

$$\frac{1}{|I|}\int_I \Phi(\varepsilon|f(s)|)\mathrm{d}s \leq \frac{1}{2}F(0).$$

Letting $\varepsilon \to 1$ and using continuity of Φ and Fatou's lemma again, gives

$$\frac{1}{|I|}\int_I \Phi(|f(s)|)\mathrm{d}s \leq \frac{1}{2}F(0).$$

It remains to compute directly that

$$\frac{1}{2}F(0) = \frac{1}{2}\int_0^\infty e^{-s}\Phi(s)\mathrm{d}s = \frac{1}{2}\int_0^\infty e^{-s}\Phi'(s)\mathrm{d}s = 2^d(1-d)\int_0^\infty se^{-s(1-d)}\mathrm{d}s = \frac{2^d}{1-d}.$$

The claim is proved. $\square$

2.2. **A logarithmic bound for A_∞ weights.** Consider the logarithmic domain

$$D^c_\infty = \{(x,y) \in (0,\infty) \times \mathbb{R} : \log(x/c) \leq y \leq \log x\}$$

and the Bellman function $U = U^c : D^c_\infty \to \mathbb{R}$ uniquely determined by the equality

$$\begin{aligned} &U\Big(x(1-t), -t+\log(x/c)\Big) \\ &\qquad = \frac{t}{d}\cdot\Psi(x(1-d)) + \left(1-\frac{t}{d}\right)\frac{(x(1-d))^{1/d}}{d}\int_{x(1-d)}^\infty \Psi(s)s^{-1/d-1}\mathrm{d}s \\ &\qquad = \Psi(x(1-d)) + \left(1-\frac{t}{d}\right)(x(1-d))^{1/d}\int_{x(1-d)}^\infty \Psi'(s)s^{-1/d}\mathrm{d}s, \end{aligned}$$

for $x>0$ and $t\in[0,d]$. To understand the above definition, observe that on the upper boundary of the set D^c_∞, i.e., for $y=\log x$, one takes $U(x,y)=\Psi(x)$ (this corresponds to the choice $x := x/(1-d)$ and $t=d$ in the above formula for U). On the lower boundary $y=\log(x/c)$, we take

$$U(x,y) = \frac{(x(1-d))^{1/d}}{d}\int_{x(1-d)}^\infty \Psi(s)s^{-1/d-1}\mathrm{d}s$$

(take $t=0$ in the formula for U). To understand the behavior of U in the interior of D^c_∞, we 'foliate' the domain, splitting it into the union of pairwise disjoint line segments which start from the lower boundary, are tangent to it and go to the left: see Figure 1 below. The definition of U states that the function is linear along any leaf of the foliation. So, for any point $(x(1-t), -t+\log(x/c))$ lying in the interior of D^c_∞, the function U is linear along a small line segment of slope $1/x$ passing through this point.

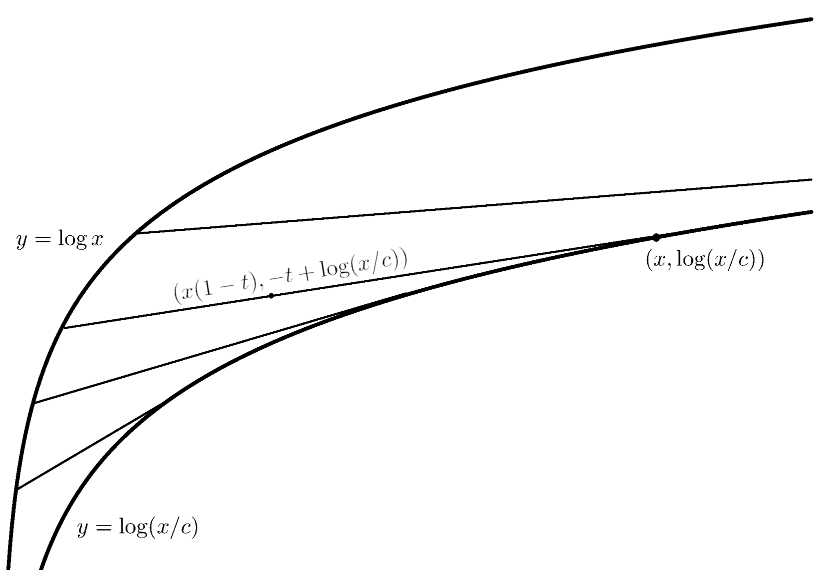

FIGURE 1. The function U is linear along the leaves of the foliation.

Note that by standard theorems on implicit functions, we get that U is of class C^1. Furthermore, it is of class C^2 outside the leaf of the foliation which starts from the point $(2^d, \log(2^d/c))$ (this particular leaf corresponds to the choice $x = 2^d$ in the formula for U, and the point $2^d(1-d)$ is the only one at which Ψ is not twice differentiable).

In the lemma below, we prove further important properties of U.

Lemma 2.5. *The function U satisfies $U_y \leq 0$ and is locally concave.*

Proof. In the calculations below, all the partial derivatives are evaluated at an arbitrary point $(x(1-t), -t+\log(x/c))$. We assume that this point lies in the interior of D_∞^c and is not contained in the particular leaf of the foliation just described above. This guarantees that U is of class C^2 in some neighborhood of the point. Differentiating formula for U with respect to x, we obtain

$$\begin{aligned}(2.6)\qquad &(1-t)U_x + \frac{1}{x}U_y\\ &= (1-d)\left[\frac{t}{d}\Psi'(x(1-d)) + \left(1-\frac{t}{d}\right)\frac{(x(1-d))^{1/d-1}}{d}\int_{x(1-d)}^\infty \Psi'(s)s^{-1/d}\mathrm{d}s\right],\end{aligned}$$

while the differentiation with respect to t yields

$$-xU_x - U_y = -\frac{(x(1-d))^{1/d}}{d}\int_{x(1-d)}^\infty \Psi'(s)s^{-1/d}\mathrm{d}s. \tag{2.7}$$

If we multiply the first identity by x, the second by $1-t$ and add, we obtain

$$\begin{aligned} tU_y &= \frac{tx(1-d)}{d}\left[\Psi'(x(1-d)) + \left(1-\frac{1}{d}\right)(x(1-d))^{1/d-1}\int_{x(1-d)}^\infty \Psi'(s)s^{-1/d}\mathrm{d}s\right]\\ &= -\frac{t(x(1-d))^{1/d}}{d}\int_{x(1-d)}^\infty \Psi''(s)s^{1-1/d}\mathrm{d}s \leq 0\end{aligned}$$

(the second equality follows from integration by parts). This gives the first part of the lemma. To show the local concavity of U, multiply (2.6) by x and add (2.7) to get

$$-txU_x = \frac{tx(1-d)}{d}\left[\Psi'(x(1-d)) - \frac{(x(1-d))^{1/d-1}}{d}\int_{x(1-d)}^\infty \Psi'(s)s^{-1/d}\mathrm{d}s\right],$$

that is,

$$U_x = \frac{(1-d)}{d}\left[-\Psi'(x(1-d)) + \frac{(x(1-d))^{1/d-1}}{d}\int_{x(1-d)}^\infty \Psi'(s)s^{-1/d}\mathrm{d}s\right]. \tag{2.8}$$

Differentiation of both sides with respect to t gives

$$-xU_{xx} - U_{xy} = 0. \tag{2.9}$$

But, as we have noted above, the function U is linear along a short line segment of slope $1/x$ passing through $(x(1-t), -t+\log(x/c))$. This gives $x^2U_{xx} + 2xU_{xy} + U_{yy} = 0$, which combined with (2.9) implies $xU_{xy} + U_{yy} = 0$ and hence $\det D^2U = 0$. Therefore, we will be done if we prove that $U_{xx} \leq 0$. To this end, let us differentiate (2.8) with respect to x

to obtain

$$\begin{aligned}(1-t)U_{xx}+\frac{1}{x}U_{xy}&=\frac{(1-d)^2}{d}\Bigg[\frac{1}{d}\left(\frac{1}{d}-1\right)(x(1-d))^{1/d-2}\int_{x(1-d)}^\infty \Psi'(s)s^{-1/d}\mathrm{d}s\\&\qquad\qquad-\frac{\Psi'(x(1-d))}{dx(1-d)}-\Psi''(x(1-d))\Bigg]\\&=\frac{(1-d)^2}{d}\left[\frac{(x(1-d))^{1/d-2}}{d}\int_{x(1-d)}^\infty \Psi''(s)s^{1-1/d}\mathrm{d}s-\Psi''(x(1-d))\right],\end{aligned}$$

where the last passage follows from integration by parts. Combining this identity with (2.9), we get

$$-txU_{xx}=\frac{x(1-d)^2}{d}\left[\frac{(x(1-d))^{1/d-2}}{d}\int_{x(1-d)}^\infty \Psi''(s)s^{1-1/d}\mathrm{d}s-\Psi''(x(1-d))\right]$$

and hence, to prove that $U_{xx}\leq 0$, we must check whether

$$\frac{u^{1/d-2}}{d}\int_u^\infty \Psi''(s)s^{1-1/d}\mathrm{d}s\geq \Psi''(u) \tag{2.10}$$

for all $u>0$, $u\neq 2^d(1-d)$ (the latter requirement comes from the fact that Ψ'' does not exist at $2^d(1-d)$) . If $u<2^d(1-d)$, then the inequality is trivial, since $\Psi''(u)=0$ and the integral is positive. If $u>2^d(1-d)$, we compute explicitly that $\Psi''(u)=-2\log\left(2\left(\frac{1-d}{u}\right)^{1/d}\right)\cdot(du)^{-1}$ and

$$\begin{aligned}\int_u^\infty \Psi''(s)s^{1-1/d}\mathrm{d}s&=-\frac{2}{d}\int_u^\infty\log\left(2\left(\frac{1-d}{s}\right)^{1/d}\right)s^{-1/d}\mathrm{d}s\\&=\frac{2^d}{(1-d)^{1/d}}\int_0^{2((1-d)/u)^{1/d}}(-\log s)(s^{1-d})'\mathrm{d}s\\&=\frac{2u^{1-1/d}}{1-d}\left[-\log\left(2\left(\frac{1-d}{u}\right)^{1/d}\right)+\frac{1}{1-d}\right],\end{aligned}$$

where in the last passage we have exploited the integration by parts. Therefore, (2.10) is equivalent to

$$\frac{1}{1-d}\left[-\log\left(2\left(\frac{1-d}{u}\right)^{1/d}\right)+\frac{1}{1-d}\right]\geq-\log\left(2\left(\frac{1-d}{u}\right)^{1/d}\right),$$

which is obvious (we have $1/(1-d)\geq 1$). This completes the proof. □

As we have already noted above, $U(x,\log x)=\Psi(x)$. Together with the above lemma, this gives the validity of (2.2): this is due to Theorem 4.13 in [15]. For the sake of completeness, we provide the detailed proof of the weighted inequality. We will need the following technical fact, a version of Lemma 2.4 for A_∞ weights. See Lemma 4_∞ in [17].

Lemma 2.6. *For any $\varepsilon>c$ and an arbitrary weight on I with $[w]_{A_\infty(I)}\leq c$ there exists a splitting $I=I^-\cup I^+$, $|I^\pm|=\alpha_\pm|I|$, such that the entire interval with the endpoints $p^\pm=(\langle w\rangle_{I^\pm},\langle\log w\rangle_{I^\pm})$ is in D_∞^ε. Moreover, the splitting parameters $\alpha_\pm$ can be chosen bounded away from 0 and 1 uniformly with respect to w and, therefore, with respect to I as well.*

Equipped with the above facts, we proceed to our estimate for A_∞ weights. The argumentation is quite similar to that used in the proof of the BMO estimate (2.1).

Proof of (2.2). Fix an A_∞ weight w as in the statement. Let $\varepsilon > c$ be an auxiliary parameter.

Step 1. Consider the family $\{\mathcal{I}^n\}_{n\geq 0}$ of partitions of I, generated by the inductive use of Lemma 2.6. Namely, we put $\mathcal{I}^0 = \{I\}$ and then, given $\mathcal{I}^n = \{I^{n,1}, I^{n,2}, \dots, I^{n,2^n}\}$, we split each $I^{n,k}$ according to Lemma 2.6, applied to the function $w|_{I^{n,k}}$ and the parameter ε. Finally, put

$$\mathcal{I}^{n+1} = \left\{I_-^{n,1}, I_+^{n,1}, I_-^{n,2}, I_+^{n,2}, \dots, I_-^{n,2^n}, I_+^{n,2^n}\right\}.$$

Next, we define the sequences $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ of functions on I by

$$f_n(x) = \langle w\rangle_{I^n(x)}, \qquad g_n(t) = \langle \log w\rangle_{I^n(x)},$$

where $I^n(x) \in \mathcal{I}^n$ is an interval containing x (as previously, if there are two such intervals, we pick the one which has x as its right endpoint). Since $[w]_{A_\infty} \leq c$, we have $(f_n, g_n) \in D_\infty^c$ almost everywhere for each n.

Step 2. Let $U = U^\varepsilon$ be the Bellman function constructed above, corresponding to the parameter ε. Then for any nonnegative integer n and any $I^{n,k} \in \mathcal{I}^n$ we have

$$\int_{I^{n,k}} U^\varepsilon(f_{n+1}, g_{n+1})\,\mathrm{d}s \leq \int_{I^{n,k}} U^\varepsilon(f_n, g_n)\,\mathrm{d}s. \tag{2.11}$$

This follows from the local concavity of U^ε: the pair (f_n, g_n) is constant on $I^{n,k}$, say, equal to $p = (\langle w\rangle_{I^{n,k}}, \langle \log w\rangle_{I^{n,k}})$ there, while (f_{n+1}, g_{n+1}) takes two values on this interval: $p_\pm = (\langle w\rangle_{I_\pm^{n,k}}, \langle \log w\rangle_{I_\pm^{n,k}}))$. By Lemma 2.6, the entire interval with the endpoints $p_\pm$ is contained within D_∞^ε, and hence U^ε is concave along this interval.

Step 3. Summing (2.11) over k, we get

$$\int_I U^\varepsilon(f_{n+1}, g_{n+1})\,\mathrm{d}s \leq \int_I U^\varepsilon(f_n, g_n)\,\mathrm{d}s.$$

Consequently, for any nonnegative integer n we have

$$\frac{1}{|I|}\int_I U^\varepsilon(f_n, g_n)\,\mathrm{d}s \leq \frac{1}{|I|}\int_I U^\varepsilon(f_0, g_0)\,\mathrm{d}s.$$

But $f_0 \equiv \langle w\rangle_I = 1$ and $g_0 \geq \log(f_0/\varepsilon) = -\log\varepsilon$, so by the first part of Lemma 2.5, we have $U(f_0, g_0) \leq U(1, -\log\varepsilon)$. Similarly, we have $g_n \leq \log f_n$, so $U(f_n, g_n) \geq U(f_n, \log f_n) = \Psi(f_n)$ and the above inequality yields

$$\frac{1}{|I|}\int_I \Psi(f_n)\mathrm{d}s \leq U^\varepsilon\left(1, -\log\varepsilon\right). \tag{2.12}$$

However, recall that the splitting ratios $\alpha_\pm$ of Lemma 2.6 were bounded away from 0 and 1. Therefore, the diameter of $\mathcal{I}^n$ tends to 0 as $n \to \infty$ and Lebesgue's differentiation theorem yields $f_n \to w$ almost everywhere on I. By Fatou's lemma, (2.12) yields

$$\frac{1}{|I|}\int_I \Psi(w)\mathrm{d}s \leq U^\varepsilon\left(1, -\log\varepsilon\right),$$

and letting $\varepsilon \to c$ gives

$$\frac{1}{|I|}\int_I \Psi(w)\mathrm{d}s \leq U^c(1, -\log c) = (1-d)^{1/d}\int_{2^d(1-d)}^\infty \Psi'(s)s^{-1/d}\mathrm{d}s.$$

It remains to compute that the right hand side equals

$$\int_{2^d(1-d)}^{\infty} -\log\left(2\left(\frac{1-d}{s}\right)^{1/d}\right)\left(\frac{1-d}{s}\right)^{1/d}\mathrm{d}s = d2^{d-1}\int_0^1 (\log s)^2 (s^{1-d})'\mathrm{d}s,$$

which is $d2^d/(1-d)^2$, by the integration by parts. $\square$

3. Proof of BMO estimates

Equipped with the estimates (2.1) and (2.2), we are ready for the comparison of BMO norms in the weighted and the unweighted settings. The starting lemma below is, essentially, the Young inequality for the functions $t \mapsto \Phi(\sqrt{t})$ and Ψ. However, we need to provide the proof, since, formally, these functions are not Young functions (the first of them has non-vanishing derivative at zero, the second is zero on a nontrivial interval). Hence the classical Young inequality does not apply directly. Nevertheless, the argument is standard and easy.

Lemma 3.1. *For any $x \in \mathbb{R}$ and $y > 0$ we have the estimate*

$$x^2 y \leq \Phi(|x|) + \Psi(y). \tag{3.1}$$

Proof. Fix $x \in \mathbb{R}$ and consider the function $F(y) = x^2 y - \Psi(y)$. We compute that

$$F'(y) = x^2 - \Psi'(y) = \begin{cases} x^2 & \text{if } y \leq 2^d(1-d), \\ x^2 - \log^2\left(2\left(\frac{1-d}{y}\right)^{1/d}\right) & \text{if } y > 2^d(1-d), \end{cases}$$

so F attains its maximal value for y_0 satisfying $x^2 = \log^2\left(2\left(\frac{1-d}{y_0}\right)^{1/d}\right)$, i.e., for $y_0 = 2^d(1-d)\exp(|x|d)$. We have

$$F(y_0) = x^2 y_0 - \int_{2^d(1-d)}^{y_0} \log^2\left(2\left(\frac{1-d}{s}\right)^{1/d}\right)\mathrm{d}s = \Phi(|x|),$$

which can be computed after some straightforward manipulations. $\square$

We turn our attention to our main result.

Proof of (1.7). Fix a function $f \in BMO$ and an A_∞ weight w satisfying $[w]_{A_\infty} \leq c$. Let $d = d(c)$ be the parameter given by (1.6) and let I be an arbitrary subinterval of $\mathbb{R}$. By (3.1), we may write

$$\frac{1}{w(I)}\int_I \left(\frac{f - \langle f\rangle_I}{\|f\|_{BMO}}\right)^2 w\mathrm{d}s \leq \frac{1}{|I|}\int_I \Phi\left(\frac{|f - \langle f\rangle_I|}{\|f\|_{BMO}}\right)\mathrm{d}s + \frac{1}{|I|}\int_I \Psi\left(\frac{w}{\langle w\rangle_I}\right)\mathrm{d}s.$$

The function $(f - \langle f\rangle_I)/\|f\|_{BMO}$, restricted to I, has integral zero and the BMO norm less or equal to 1; in addition, $w/\langle w\rangle_I$, considered as a weight on I, has average 1 and characteristic not exceeding c. Therefore, applying (2.1) and (2.2), we obtain

$$\frac{1}{w(I)}\int_I \left(\frac{f - \langle f\rangle_I}{\|f\|_{BMO}}\right)^2 w\mathrm{d}s \leq \frac{2^d}{1-d} + \frac{d2^d}{(1-d)^2} = \frac{2^d}{(1-d)^2}. \tag{3.2}$$

Now we are ready for (1.7). The crucial fact is that for any $a \in \mathbb{R}$,

$$\frac{1}{w(I)}\int_I \left(f - \frac{1}{w(I)}\int_I fw\right)^2 w\mathrm{d}s \leq \frac{1}{w(I)}\int_I (f-a)^2\, w\mathrm{d}s,$$

so in particular, plugging $a = \langle f\rangle_I$ and applying (3.2), we get

$$\frac{1}{w(I)}\int_I\left(f-\frac{1}{w(I)}\int_I fw\right)^2 w\mathrm{d}s \leq \frac{2^d}{(1-d)^2}\|f\|_{BMO}^2.$$

The proof is complete, since I was arbitrary. $\square$

Sharpness, the localized case. Suppose that $c \geq 1$ is a fixed parameter and let $d = d(c)$ be given by (1.6). Let $I = [-1,1]$ and introduce the functions $f, w : I \to \mathbb{R}$ by the formulas

$$f(u) = \begin{cases} \log(2(1+u)) & \text{if } u < -1/2, \\ 0 & \text{if } -1/2 \leq u \leq 1/2, \\ -\log(2(1-u)) & \text{if } u > 1/2, \end{cases} \qquad w(u) = (1-d)(1-|u|)^{-d}. \tag{3.3}$$

Since f is odd and w is even, we have $\int_I fw = 0$. Furthermore, we easily compute that $w(I) = 2$ and

$$\begin{aligned}\int_I f^2 w\mathrm{d}s &= 2\int_{1/2}^1 \big[\log(2(1-s))\big]^2(1-d)(1-s)^{-d}\mathrm{d}s \\ &= 2^d(1-d)\int_0^1 (\log s)^2 s^{-d}\mathrm{d}s = \frac{2^{d+1}}{(1-d)^2}.\end{aligned}$$

This implies

$$\|f\|_{BMO(w)}^2 \geq \frac{1}{w(I)}\int_I f^2 w\mathrm{d}s - \left(\frac{1}{w(I)}\int_I fw\mathrm{d}s\right)^2 = \frac{2^d}{(1-d)^2}.$$

Therefore, we will be done if we show that $\|f\|_{BMO} \leq 1$ and $[w]_{A_\infty} \leq c$. Let us handle these two properties separately.

The inequality $\|f\|_{BMO} \leq 1$. We need to prove that for any $-1 \leq a < b \leq 1$, the point $\big(\langle f\rangle_{[a,b]}, \langle f^2\rangle_{[a,b]}\big)$ lies in the set $\{(x,y) : x^2 \leq y \leq x^2+1\}$. By the Schwarz inequality we have $\langle f^2\rangle_{[a,b]} \geq \langle f\rangle_{[a,b]}^2$, so we only need to check that any such point lies on or below the upper parabola $y = x^2+1$. For clarity, we split the verification into several steps.

Step 1. Suppose first that $a \geq 1/2$ and $b = 1$. A direct computation reveals that

$$\frac{1}{1-a}\int_a^1 \log(2(1-u))\mathrm{d}u = 1 - \log(2(1-a)) \tag{3.4}$$

and

$$\frac{1}{1-a}\int_a^1 \big(\log(2(1-u))\big)^2\mathrm{d}u = \big(\log(2(1-a))\big)^2 - 2\log(2(1-a)) + 2,$$

so $\langle f^2\rangle_{[a,1]} - \langle f\rangle_{[a,1]}^2 = 1$.

Step 2. Now suppose that $1/2 \leq a < b \leq 1$. We have the identity

$$\frac{1}{1-a}\int_a^1 f\mathrm{d}u = \frac{1}{1-a}\left(\int_a^b f\mathrm{d}u + \int_b^1 f\mathrm{d}u\right) = \frac{b-a}{1-a}\cdot\frac{1}{b-a}\int_a^b f\mathrm{d}u + \frac{1-b}{1-a}\cdot\frac{1}{1-b}\int_b^1 f\mathrm{d}u$$

and similarly for $\frac{1}{1-a}\int_a^1 f^2\mathrm{d}u$. In other words, we have

$$\big(\langle f\rangle_{[a,1]}, \langle f^2\rangle_{[a,1]}\big) = \frac{b-a}{1-a}\big(\langle f\rangle_{[a,b]}, \langle f^2\rangle_{[a,b]}\big) + \frac{1-b}{1-a}\big(\langle f\rangle_{[b,1]}, \langle f^2\rangle_{[b,1]}\big),$$

that is, the point $\big(\langle f\rangle_{[a,1]},\langle f^2\rangle_{[a,1]}\big)$ lies on the line segment with the endpoints equal to $\big(\langle f\rangle_{[a,b]},\langle f^2\rangle_{[a,b]}\big)$ and $\big(\langle f\rangle_{[b,1]},\langle f^2\rangle_{[b,1]}\big)$. But, as we have just shown above, the first and the third of these points belong to the parabola $y=x^2+1$; consequently, the second point must lie below. Let us also make an observation which will be useful later. Namely, by (3.4), we have $\langle f\rangle_{[b,1]}\geq\langle f\rangle_{[\frac12,1]}=1$ and hence

$$\langle f\rangle_{[\frac12,b]}=\frac{1}{2b-1}\langle f\rangle_{[\frac12,1]}-\frac{2-2b}{2b-1}\langle f\rangle_{[b,1]}\leq 1. \tag{3.5}$$

Step 3. Now suppose that $0\leq a<b\leq 1$. If $b\leq 1/2$, then there is nothing to prove: we have $\langle f\rangle_{[a,b]}=\langle f^2\rangle_{[a,b]}=0$. Therefore, let us assume that $b>1/2$. Arguing as above, we observe that

$$\begin{aligned}\big(\langle f\rangle_{[a,b]},\langle f^2\rangle_{[a,b]}\big)&=\frac{\frac12-a}{b-a}\big(\langle f\rangle_{[a,\frac12]},\langle f^2\rangle_{[a,\frac12]}\big)+\frac{b-\frac12}{b-a}\big(\langle f\rangle_{[\frac12,b]},\langle f^2\rangle_{[\frac12,b]}\big)\\&=\frac{b-\frac12}{b-a}\big(\langle f\rangle_{[\frac12,b]},\langle f^2\rangle_{[\frac12,b]}\big).\end{aligned}$$

But as we have shown in the previous step, the point $\big(\langle f\rangle_{[\frac12,b]},\langle f^2\rangle_{[\frac12,b]}\big)$ lies below the parabola $y=x^2+1$; furthermore, by (3.5), we have $\langle f\rangle_{[\frac12,b]}\leq 1$. Consequently,

$$\begin{aligned}\langle f^2\rangle_{[a,b]}=\frac{b-\frac12}{b-a}\langle f^2\rangle_{[\frac12,b]}&\leq\frac{b-\frac12}{b-a}\big[\langle f\rangle^2_{[\frac12,b]}+1\big]\\&\leq\left(\frac{b-\frac12}{b-a}\right)^2\langle f\rangle^2_{[\frac12,b]}+\frac{b-\frac12}{b-a}\left(\frac{\frac12-a}{b-a}+1\right)\\&=\langle f\rangle^2_{[a,b]}+\left(-\frac{\frac12-a}{b-a}+1\right)\left(\frac{\frac12-a}{b-a}+1\right)\leq\langle f\rangle^2_{[a,b]}+1,\end{aligned}$$

as desired. As a by-product of the above reasoning, note that

$$\langle f^2\rangle_{[0,b]}=\left(1-\frac{1}{2b}\right)\langle f^2\rangle_{[\frac12,b]}\leq\frac12\langle f^2\rangle_{[\frac12,b]}\leq\frac12\big(\langle f\rangle^2_{[\frac12,b]}+1\big)\leq 1.$$

Step 4. The only possibility which needs to be considered is $-1\leq a<0<b\leq 1$ (the case $-1\leq a<b\leq 0$ follows by the symmetry of f). We write

$$\big(\langle f\rangle_{[a,b]},\langle f^2\rangle_{[a,b]}\big)=\frac{-a}{b-a}\big(\langle f\rangle_{[a,0]},\langle f^2\rangle_{[a,0]}\big)+\frac{b}{b-a}\big(\langle f\rangle_{[0,b]},\langle f^2\rangle_{[0,b]}\big).$$

By the symmetry of f and the last line of the previous step, both points on the right lie on or below the line $\{(x,y):y=1\}$, and hence so does $\big(\langle f\rangle_{[a,b]},\langle f^2\rangle_{[a,b]}\big)$. It remains to note that this line is tangent to the parabola $y=x^2+1$.

The inequality $[w]_{A_\infty}\leq c$. There are lots of similarities with the above verification of the BMO property of f. We need to check that for each $-1\leq a<b\leq 1$ the point $\big(\langle w\rangle_{[a,b]},\langle\log w\rangle_{[a,b]}\big)$ lies in the set $\{(x,y)\in(0,\infty)\times\mathbb{R}:\log(x/c)\leq y\leq\log x\}$. By Jensen's inequality we have $\langle\log w\rangle_{[a,b]}\leq\log\langle w\rangle_{[a,b]}$, so we only need to check that the above point lies on or above the logarithmic curve $y=\log(x/c)$, the lower boundary of the set. As previously, it is convenient to split the argumentation.

Step 1. We first consider the case $a\geq 0$ and $b=1$. Since

$$\langle w\rangle_{[a,1]}=(1-a)^{-d}\quad\text{ and }\quad\langle\log w\rangle_{[a,1]}=d+\log\frac{1-d}{(1-a)^d},$$

we see that $\langle \log w\rangle_{[a,1]} - \log\langle w\rangle_{[a,1]} = d + \log(1-d) = -\log c$. Here in the last equality we have used (1.6). Therefore, the point $\big(\langle w\rangle_{[a,1]}, \langle \log w\rangle_{[a,1]}\big)$ lies at the logarithmic curve $y = \log(x/c)$.

Step 2. Now suppose that $0 \le a < b \le 1$. We write the identity

$$\big(\langle w\rangle_{[a,1]}, \langle \log w\rangle_{[a,1]}\big) = \frac{b-a}{1-a}\big(\langle w\rangle_{[a,b]}, \langle \log w\rangle_{[a,b]}\big) + \frac{1-b}{1-a}\big(\langle w\rangle_{[b,1]}, \langle \log w\rangle_{[b,1]}\big), \tag{3.6}$$

which guarantees that the three points involved are colinear. As we checked in the previous step, the point on the left and the second point on the right lie on the curve $y = \log(x/c)$. By the concavity of the logarithmic function, the third point must lie above the curve.

Step 3. Now we will study the final case $a \le 0 < b$ (the possibility $a < b < 0$ follows by the symmetry of w). We start with an observation. Let us specify $a = 0$ and rewrite (3.6) in the form

$$(1, -\log c) = b\big(\langle w\rangle_{[0,b]}, \langle \log w\rangle_{[0,b]}\big) + (1-b)\left((1-b)^{-d}, d + \log\frac{1-d}{(1-b)^d}\right).$$

We can infer the following information about the location of the point $\big(\langle w\rangle_{[0,b]}, \langle \log w\rangle_{[0,b]}\big)$. See Figure 2 below. First, note that both $(1, -\log c)$ and $\left((1-b)^{-d}, d + \log\frac{1-d}{(1-b)^d}\right)$ lie

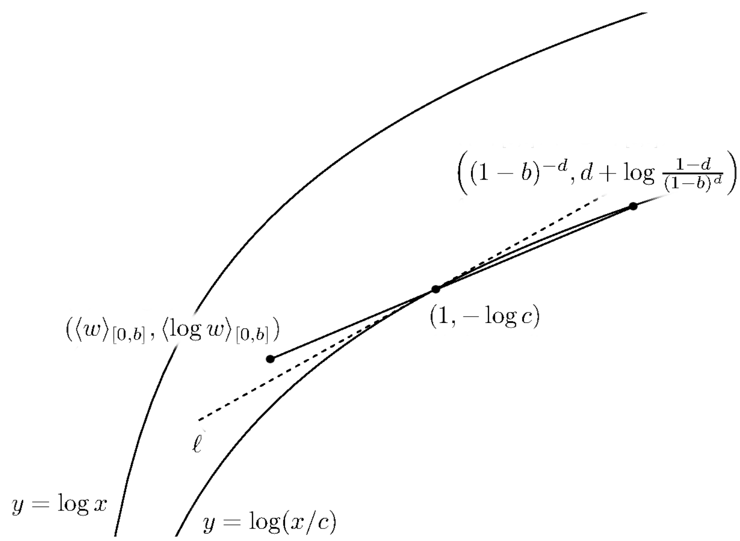

FIGURE 2. The location of the point $\big(\langle w\rangle_{[0,b]}, \langle \log w\rangle_{[0,b]}\big)$.

on the curve $y = \log(x/c)$ and the second point has a bigger x-coordinate. Consequently, $\langle w\rangle_{[0,b]} \le 1$ and the slope of the line joining the above three points is smaller than the slope of a line ℓ tangent to the curve $y = \log(x/c)$ at $(1, -\log c)$. This in particular implies that the point $\big(\langle w\rangle_{[0,b]}, \langle \log w\rangle_{[0,b]}\big)$ lies above ℓ. We return to the general case and write

$$\big(\langle w\rangle_{[a,b]}, \langle \log w\rangle_{[a,b]}\big) = \frac{-a}{b-a}\big(\langle w\rangle_{[a,0]}, \langle \log w\rangle_{[a,0]}\big) + \frac{b}{b-a}\big(\langle w\rangle_{[0,b]}, \langle \log w\rangle_{[0,b]}\big).$$

It follows from the symmetry of w and the above observation that both points on the right lie above ℓ, and hence so does the point on the left. Therefore it must also lie above the curve $y = \log(x/c)$, by the concavity of the logarithmic function.

The proof is complete. $\square$

Sharpness, the general case. Let $c \geq 1$ and $d = d(c)$ be as previously, and let $\varepsilon > 0$ be arbitrary. The idea is to apply the transference result from [16]. First, we use a discretization argument (see e.g. Section 4 in [16]) to obtain a pair $(f^\varepsilon, w^\varepsilon)$ of functions on $[-1,1]$, which takes values in a finite set, $\|f^\varepsilon\|_{BMO} \leq 1+\varepsilon$, $[w^\varepsilon]_{A_\infty} \leq c+\varepsilon$ and $\|f^\varepsilon\|_{BMO(w^\varepsilon)} \geq 2^{d/2}/(1-d) - \varepsilon$. Next, we use the periodization argument developed in [16] (see Section 2 there): it allows us to obtain a pair $(\tilde{f}^\varepsilon, \tilde{w}^e)$ defined on the whole real line, such that $\|\tilde{f}^\varepsilon\|_{BMO} \leq 1+2\varepsilon$, $[\tilde{w}^\varepsilon]_{A_\infty} \leq c+2\varepsilon$ and $\|\tilde{f}^\varepsilon\|_{BMO(\tilde{w}^\varepsilon)} \geq \|f^\varepsilon\|_{BMO(w^\varepsilon)}$. This yields the desired sharpness, since ε was arbitrary. $\square$

Proof of Theorem 1.1. The inequality (1.5) follows directly from (1.7) by optimization of the constant. Namely, suppose that $c \geq 1$ and let $d = d(c)$ be given by (1.6). We have $d < 1$, so

$$\frac{2^{d/2}}{1-d} = c(e\sqrt{2})^d \leq e\sqrt{2}c.$$

The fact that $e\sqrt{2}$ is optimal follows immediately from the sharpness of (1.7) established above and the equality $\lim_{c\to\infty} d(c) = 1$: the examples constructed above show that

$$\sup_{f,\,w} \frac{\|f\|_{BMO(w)}}{\|f\|_{BMO}[w]_{A_\infty}} \geq \sup_{c\geq 1} \frac{2^{d(c)/2}}{(1-d(c))c} = \sup_{c\geq 1}(e\sqrt{2})^{d(c)} = e\sqrt{2}. \quad \square$$

4. Inequalities for BMO martingales

In this section we will extend Theorems 1.1 and 1.2 to the martingale context. 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. We will assume in addition that any local martingale adapted to this filtration is continuous: for instance, this requirement is satisfied for the Brownian filtration. 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 [1] for the definition. Following Getoor and Sharpe [3], given $1 \leq p < \infty$, a uniformly integrable martingale $X = (X_t)_{t\geq 0}$ belongs to the class BMO_p, if

$$\|X\|_{BMO_p} = \sup_{T\geq 0} \left\| \mathbb{E}\Big[|X_\infty - X_T|^p \,\big|\mathcal{F}_T \Big]^{1/p} \right\|_\infty < \infty.$$

This is precisely the probabilistic counterpart of the oscillations (1.2). In analogy to the analytic setting, it can be shown that all the seminorms $\|\cdot\|_{BMO_p}$ are equivalent and hence all the probabilistic classes BMO_p coincide. It will be convenient for us to work with the L^2-based seminorm $\|\cdot\|_{BMO_2}$, denoted, for notational simplicity, by $\|\cdot\|_{BMO}$. Note that we have the following probabilistic version of (1.3):

$$\|X\|_{BMO}^2 = \sup_{T\geq 0} \operatorname{essup}\Big(\mathbb{E}(X_\infty^2|\mathcal{F}_T) - X_T^2 \Big). \tag{4.1}$$

We also need to introduce the stochastic version of the A_∞ theory. Any integrable and positive random variable W is called a weight, and it gives rise to the associated uniformly integrable martingale $(W_t)_{t\geq 0} = (\mathbb{E}(W|\mathcal{F}_t))_{t\geq 0}$. The weight W is said to satisfy

Muckenhoupt's condition A_∞, if its characteristic

$$[W]_{A_\infty} = \sup_{T\geq 0} \left\| W_T \exp\Big(-\mathbb{E}\big(\log W|\mathcal{F}_T\big)\Big)\right\|_\infty$$

is finite. See [5, 7] for more on the subject. The associated weighted BMO space is defined as the collection of all uniformly integrable martingales $X=(X_t)_{t\geq 0}$, for which

$$\|X\|_{BMO(W)} = \sup_{T\geq 0}\left\| \mathbb{E}_W\Big[\big|X_\infty - \mathbb{E}_W\big(X_\infty|\mathcal{F}_T\big)\big|^2\,\big|\mathcal{F}_T\Big]^{1/2}\right\|_\infty < \infty.$$

Here $\mathbb{E}_W(\cdot|\mathcal{F}_T)$ is the conditional expectation with respect to the measure $W\mathrm{d}\mathbb{P}$, i.e., $\mathbb{E}_W(\xi|\mathcal{F}_T)=\mathbb{E}(\xi W|\mathcal{F}_T)/W_T$ for any $W\mathrm{d}\mathbb{P}$-measurable random variable ξ.

We will prove the following.

Theorem 4.1. *For any BMO martingale X and any probabilistic weight $W\in A_\infty$ we have*

$$\|X\|_{BMO(W)} \leq e\sqrt{2}[W]_{A_\infty}\|X\|_{BMO}. \tag{4.2}$$

The constant $e\sqrt{2}=3.8442\ldots$ is the best possible.

Theorem 4.2. *Let $c\geq 1$ be a fixed parameter and let $d=d(c)$ be the solution to* (1.6)*. For any BMO martingale X and any probabilistic weight $W\in A_\infty$ satisfying $[W]_{A_\infty}\leq c$ we have*

$$\|X\|_{BMO(W)} \leq \frac{2^{d/2}}{1-d}\|X\|_{BMO}. \tag{4.3}$$

The constant $2^d/(1-d)$ is the best possible.

The proof is an adaptation of the analytic argumentation presented in the previous two sections. In particular, we will need appropriate versions of the inequalities (2.1) and (2.2).

Theorem 4.3. *Let $T\geq 0$ be fixed. Then for any BMO martingale $X=(X_t)_{t\geq 0}$ satisfying $\|X\|_{BMO}\leq 1$ and $\mathbb{E}(X|\mathcal{F}_T)=0$, we have*

$$\mathbb{E}\Big[\Phi\big(|X_\infty|\big)\big|\mathcal{F}_T\Big] \leq \frac{2^d}{1-d}.$$

Proof. By a standard limiting argument, we may assume that $\|X\|_{BMO}$ is strictly less than 1: say, $\|X\|_{BMO}^2 = 1-\varepsilon$ for some $\varepsilon>0$. Consider an auxiliary martingale $Y_t=\mathbb{E}(X_\infty^2|\mathcal{F}_t)$, $t\geq 0$. By (4.1) and the assumption $\|X\|_{BMO}=1-\varepsilon$, the pair (X,Y) takes values in the parabolic domain $D_{BMO}=\{(x,y): x^2\leq y\leq x^2+(1-\varepsilon)^2\}$. Let B be the Bellman function introduced in Section 2. We apply Itô's formula to the process $B(X,Y)$. Formally this is not permitted, since B is only of class C^1. This obstacle is handled by performing an appropriate mollification argument (see e.g. [14], formula (5.3)). Namely, let g be a nonnegative C^∞ function on $\mathbb{R}^2$, supported on the unit ball and satisfying $\int_{\mathbb{R}^2} g=1$. Given $\delta\in(0,\varepsilon/2)$, we consider the function $B^\delta: D_{BMO}\to\mathbb{R}$ given by the convolution-type expression

$$B^\delta(x,y)=\int_{[-1,1]^2} B(x-\delta u, y+\delta-2x\delta u+\delta^2u^2-\delta v)g(u,v)\mathrm{d}u\mathrm{d}v.$$

Note that the integrand is well-defined: we have

$$\big(y+\delta-2x\delta u+\delta^2u^2-\delta v\big)-(x-\delta u)^2 = y-x^2+\delta-\delta v\in[0,1]$$

for $v \in [-1,1]$. The function B^δ is of class C^∞ and, by the very definition, it inherits the local concavity. Therefore, the application of Itô's formula gives, for any $t \geq T$,

$$B^\delta(X_t,Y_t) = B^\delta(X_T,Y_T) + I_1 + I_2/2, \tag{4.4}$$

where

$$I_1 = \int_t^T B_x^\delta(X_s,Y_s)\mathrm{d}X_s + \int_t^T B_y^\delta(X_s,Y_s)\mathrm{d}Y_s,$$
$$I_2 = \int_t^T B_{xx}^\delta(X_s,Y_s)\mathrm{d}\langle X,X\rangle_s + 2\int_t^T B_{xy}^\delta(X_s,Y_s)\mathrm{d}\langle X,Y\rangle_s + \int_t^T B_{yy}^\delta(X_s,Y_s)\mathrm{d}\langle Y,Y\rangle_s.$$

Observe that $\mathbb{E}(I_1|\mathcal{F}_T) = 0$, by the properties of stochastic integrals with respect to martingales. Furthermore, since B^δ is locally concave, the term I_2 is nonpositive: this can be seen by approximating the integrals with Riemann sums. Consequently, taking the conditional expectation in (4.4), we obtain $\mathbb{E}(B^\delta(X_t,Y_t)|\mathcal{F}_T) \leq B^\delta(X_T,Y_T)$. Now we let $\delta \to 0$: since B is continuous, we have $B^\delta \to B$ pointwise and therefore Fatou's lemma yields $\mathbb{E}(B(X_t,Y_t)|\mathcal{F}_T) \leq B(X_T,Y_T)$. But by Lemma 2.3 and the inequality $X_t^2 \leq Y_t \leq X_t^2+1$, we conclude that

$$B(X_t,Y_t) \geq B(X_t,X_t^2) = \Phi(|X_t|) \quad \text{and} \quad B(X_T,Y_T) = B(0,Y_T) \leq B(0,1) = \frac{2^d}{1-d}.$$

Thus we have proved the estimate

$$\mathbb{E}\Big[\Phi(|X_t|)|\mathcal{F}_T\Big] \leq \frac{2^d}{1-d}$$

and it remains to let $t \to \infty$ and apply Fatou's lemma. □

The martingale version of (2.2) is the following.

Theorem 4.4. *Let $T \geq 0$ be fixed. Then for any A_∞ weight W satisfying $[W]_{A_\infty} \leq c$ and $W_T \equiv 1$, we have*

$$\mathbb{E}\Big[\Psi(W)\big|\mathcal{F}_T\Big] \leq \frac{d2^d}{(1-d)^2}.$$

Proof. The argumentation is the same as above and rests on the application of the Itô formula to $U(W,V)$, where $V = (V_t)_{t\geq 0}$ is the martingale given by $V_t = \mathbb{E}\big[\log W|\mathcal{F}_t\big]$. The details are left to the reader. □

Proof of (4.2) *and* (4.3). Pick arbitrary $T \geq 0$. By (3.1), we may write

$$\mathbb{E}\left[\frac{(X_\infty - X_T)^2}{\|X\|_{BMO}^2}\cdot\frac{W_\infty}{W_T}\Big|\mathcal{F}_T\right] \leq \mathbb{E}\left[\Phi\left(\frac{|X_\infty - X_T|}{\|X\|_{BMO}}\right)\Big|\mathcal{F}_T\right] + \mathbb{E}\left[\Psi\left(\frac{W_\infty}{W_T}\right)\Big|\mathcal{F}_T\right]. \tag{4.5}$$

Observe that the martingale

$$\tilde{X}_t = \begin{cases} \dfrac{X_t - X_T}{\|X\|_{BMO}} & \text{if } t > T,\\ 0 & \text{if } t \leq T \end{cases}$$

satisfies the assumptions of Theorem 4.3 and hence

$$\mathbb{E}\left[\Phi\left(\frac{|X_\infty - X_T|}{\|X\|_{BMO}}\right)\Big|\mathcal{F}_T\right] = \mathbb{E}\Big[\Phi\big(|\tilde{X}_\infty|\big)\big|\mathcal{F}_T\Big] \leq \frac{2^d}{1-d}.$$

Similarly, the weight $\tilde{W} = (\tilde{W}_t)_{t\geq 0}$ given by

$$\tilde{W}_t = \begin{cases} W_t/W_T & \text{if } t > T, \\ 1 & \text{if } t \leq T \end{cases}$$

has all the properties required in Theorem 4.4, so

$$\mathbb{E}\left[\Psi\left(\frac{W_\infty}{W_T}\right)\Big|\mathcal{F}_T\right] = \mathbb{E}\Big[\Psi(\tilde{W})\big|\mathcal{F}_T\Big] \leq \frac{d2^d}{(1-d)^2}.$$

Plugging the above two estimates into (4.5) yields

$$\mathbb{E}_W\Big[(X_\infty - X_T)^2|\mathcal{F}_T\Big] = \mathbb{E}\left[\frac{(X_\infty - X_T)^2}{\|X\|_{BMO}^2}\cdot\frac{W}{W_T}\Big|\mathcal{F}_T\right]\cdot\|X\|_{BMO}^2 \leq \frac{2^d}{(1-d)^2}\|X\|_{BMO}^2.$$

But for any $\mathcal{F}_T$ measurable random variable ξ we have,

$$\mathbb{E}_W\Big[\big(X_\infty - \mathbb{E}_W(X_\infty|\mathcal{F}_T)\big)^2|\mathcal{F}_T\Big] \leq \mathbb{E}_W\Big[\big(X_\infty - \xi\big)^2|\mathcal{F}_T\Big],$$

so taking $\xi = X_T$ and using the previous estimate, we get

$$\mathbb{E}_W\Big[\big(X_\infty - \mathbb{E}_W(X_\infty|\mathcal{F}_T)\big)^2|\mathcal{F}_T\Big] \leq \frac{2^d}{(1-d)^2}\|X\|_{BMO}^2.$$

Since T was arbitrary, the estimate (4.3) follows. The inequality (4.2) is an immediate consequence, by optimizing over c: see the proof in the analytic setting. □

Sharpness. The idea is very simple and natural: we will show that the pair f, w constructed in the proof of the sharpness of (1.7) can be reinterpreted as a pair X, W of martingales which have the required properties. To this end, consider the probability space $([-1,1],\mathcal{B}([-1,1]),|\cdot|/2)$, equipped with the filtration $(\mathcal{F}_t)_{t\geq 0}$, where for each t, $\mathcal{F}_t$ is generated by all sets of measure zero, the interval $[0,1]$ and all intervals $[a,b]$ with $-t \leq a < b \leq t$. Then the calculations for f above give that the martingale $X_t = \mathbb{E}(f|\mathcal{F}_t)$, $t \geq 0$, satisfies $\|X\|_{BMO} \leq 1$: this follows directly from the identities

$$\mathbb{E}(f|\mathcal{F}_t) = \langle f\rangle_{[-1,-t]}\chi_{[-1,-t]} + f\chi_{(-t,t)} + \langle f\rangle_{[t,1]}\chi_{[t,1]}$$

and

$$\mathbb{E}(f^2|\mathcal{F}_t) = \langle f^2\rangle_{[-1,-t]}\chi_{[-1,-t]} + f^2\chi_{(-t,t)} + \langle f^2\rangle_{[t,1]}\chi_{[t,1]}.$$

Similarly, the weight $W = \mathbb{E}(w|\mathcal{F}_t)$, $t \geq 0$, satisfies the A_∞ condition with $[W]_{A_\infty} \leq c$. Furthermore, we have $W_0 \equiv 1$ and $\mathbb{E}_W(X_\infty^2|\mathcal{F}_0) = 2^d/(1-d)^2$. Unfortunately, we do not have the identity $\mathbb{E}_W(X_\infty|\mathcal{F}_0) = 0$ (this would allow us to finish the proof), but only

$$\mathbb{E}_W(X_\infty|\mathcal{F}_0) = -\frac{1}{2}\chi_{[-1,0)} + \frac{1}{2}\chi_{[0,1]}.$$

But this difficulty can be easily overcome: we may enlarge the above probability space and the filtration (to, say $(\mathcal{F}_t)_{t\geq -1}$) such that $\mathcal{F}_{-1}$ is a trivial σ-algebra, i.e., generated by all sets of measure 0. Then the 'extended' martingale $X = (X_t)_{t\geq -1}$ still satisfies $\|X\|_{BMO} \leq 1$. Indeed, it suffices to note that for $t \in [-1,0)$ we have

$$\mathbb{E}(X_\infty^2|\mathcal{F}_t) = \mathbb{E}\big[\mathbb{E}(X_\infty^2|\mathcal{F}_0)|\mathcal{F}_t\big] = \mathbb{E}\big[1|\mathcal{F}_t\big] = 1 \leq 1 + \mathbb{E}\big(X_\infty|\mathcal{F}_t\big)^2.$$

Similarly, the weight $W=(W_t)_{t\geq -1}$ still satisfies $[W]_{A_\infty}\leq c$ and $\mathbb{E}[W|\mathcal{F}_{-1}]=1$: this follows immediately from the fact that W_0 is already a constant random variable (equal to 1). Taking all the above observations into account, we get

$$\mathbb{E}_W(X_\infty^2|\mathcal{F}_{-1})=\mathbb{E}_W\Big[\mathbb{E}_W(X_\infty^2|\mathcal{F}_0)\big|\mathcal{F}_{-1}\Big]=\frac{2^d}{(1-d)^2},\qquad \mathbb{E}_W(X_\infty|\mathcal{F}_{-1})=0$$

and hence $\|X\|_{BMO(W)}\geq 2^{d/2}/(1-d)^2$. This gives the sharpness of (4.3), for the estimate (4.2) just repeat the reasoning used in the analytic setting. $\square$

5. On the discovery of Φ and Ψ

Now we will describe informally some steps which lead to the exponential function Φ and the logarithmic function Ψ used above. Suppose that we are interested in the sharp bound for $\|f\|_{BMO(w)}$ in the localized setting, in which both f and w are given on the interval $[-1,1]$. Let us recall the argument which has led us to (1.7). The key point is the Young inequality (3.1):

$$x^2y\leq \Phi(|x|)+\Psi(y),\qquad x\in\mathbb{R},\ y>0, \tag{5.1}$$

which implies that for any subinterval $I\subseteq[-1,1]$ we have

$$\begin{aligned}\frac{1}{w(I)}\int_I\left(\frac{f-\langle f\rangle_I}{\|f\|_{BMO}}\right)^2 w\mathrm{d}s&\leq \frac{1}{|I|}\int_I\Phi\left(\frac{|f-\langle f\rangle_I|}{\|f\|_{BMO}}\right)\mathrm{d}s+\frac{1}{|I|}\int_I\Psi\left(\frac{w}{\langle w\rangle_I}\right)\mathrm{d}s\\&\leq C_\Phi+C'_{\Psi,c}.\end{aligned}\tag{5.2}$$

Here

$$C_\Phi=\sup\left\{\frac{1}{|I|}\int_I\Phi(|f|)\mathrm{d}s\,:\,\langle f\rangle_I=0,\ \|f\|_{BMO(I)}\leq 1\right\}\tag{5.3}$$

and

$$C'_{\Psi,c}=\sup\left\{\frac{1}{|I|}\int_I\Psi(w)\mathrm{d}s\,:\,\langle w\rangle_I=1,\ [w]_{A_\infty(I)}\leq c\right\}.\tag{5.4}$$

(By a standard affine transformation argument, the definitions of C_Φ and $C'_{\Psi,c}$ do not depend on I). This in particular gives

$$\frac{1}{w(I)}\int_I\left(\frac{f-\langle f\rangle_I}{\|f\|_{BMO}}\right)^2 w\mathrm{d}s-\left(\frac{1}{w(I)}\int_I\frac{f-\langle f\rangle_I}{\|f\|_{BMO}}w\mathrm{d}s\right)^2\leq C_\Phi+C'_{\Psi,c}.\tag{5.5}$$

To ensure that the bound for $\|f\|_{BMO(w)}$ we obtain in this manner is sharp, we need to find an interval I, a function f and a weight w for which all the intermediate estimates become equalities (or almost equalities, up to an arbitrary positive error term). The first observation is that we may assume that $I=[-1,1]$, by the affine transformation argument already mentioned above. Second, substituting $f:=(f-\langle f\rangle_I)/\|f\|_{BMO}$, we see that we may restrict our search to functions of integral zero and BMO norm equal to 1. To guarantee that there is no loss when passing from (5.2) to (5.5), we need to find f, w such that $\int_I fw=0$: it seems plausible to search for an odd function f and an even weight w: then this vanishing condition will be automatic. Now we look at the second inequality in (5.2). The quantities of the form (5.3) and (5.4), for various choices of Φ and Ψ, have been studied in many papers in the literature; for instance, convenient references are [12], [13] and [17] (but the full list of papers is much, much longer). A little thought and experimentation, motivated by these examples, leads us to the functions f

and w given by (3.3) (actually, up to an affine transformation, the function f appears in [13, formula (8.2)]; concerning the weight, see the second half of Appendix 2 in [17]). Now it remains to find Φ and Ψ so that the first estimate in (5.2) becomes an equality. Looking back at (5.1), we see that for any fixed $s \in [0,1]$, we want the function

$$\xi : y \mapsto f(s)^2 y - \Phi(|f(s)|) - \Psi(y)$$

to attain its maximal value equal to zero at $y = w(s)$. In particular, plugging $s = 0$, we obtain $0 = \xi(w(0)) = -\Phi(0) - \Psi(1-d)$. We may assume that $\Phi(0) = 0$ and $\Psi(t) = 0$ for $t \in [0, 1-d]$: adding a constant κ to Φ and subtracting it from Ψ does not change the overall argument (5.2) (the quantities C_Φ and $C'_{\Psi,c}$ increase/decrease by κ, respectively). Next, note that the derivative of ξ vanishes for $\Psi'(y) = f(s)^2$, which leads us to the equation $\Psi'(w(s)) = f(s)^2$. If $s \in [0, 1/2]$, this is equivalent to saying that $\Psi'(t) = 0$ for $t \in [1-d, 2^d(1-d)]$; for $s \in (1/2, 1]$ we obtain

$$\Psi'(s) = \log^2\left(2\left(\frac{1-d}{s}\right)^{1/d}\right).$$

Together with the condition $\Psi(t) = 0$ for $t \in [0, 1-d]$, this yields the special function Ψ used in the previous sections. The equation $0 = \xi(w(s)) = f(s)^2 w(s) - \Phi(|f(s)|) - \Psi(w(s))$ leads to the exponential function Φ.

Acknowledgments

The author would like to thank anonymous Referees for the careful reading of the paper and many helpful suggestions.

References

[1] C. Dellacherie and P.-A. Meyer, *Probabilities and potential B: Theory of martingales*, North Holland, Amsterdam, 1982.
[2] C. Fefferman, *Characterizations of bounded mean oscillation*, *Bull. Amer. Math. Soc.* **77**, 1971, 587–588.
[3] R. K. Getoor and M. J. Sharpe, *Conformal martingales*, Invent. Math. **16** (1972), 271–308.
[4] P. Ivanisvili, N. Osipov, D. Stolyarov, V. Vasyunin and P. Zatitskiy, *Bellman function for extremal problems in BMO*, Trans. Amer. Math. Soc. **368** (2016), 3415–3468.
[5] M. Izumisawa and N. Kazamaki, *Weighted norm inequalities for martingales*, Tohoku Math. J. 29 (1977), 115–124.
[6] F. John and L. Nirenberg, *On functions of bounded mean oscillation*, *Comm. Pure and Appl. Math.* **14**, 1961, 415–426.
[7] N. Kazamaki, *Continuous exponential martingales and BMO*, Lecture Notes in Math. 1579, Springer-Verlag, Berlin, Heidelberg, 1994.
[8] S. V. Khrushchev, *A description of weights satisfying the A_∞ condition of Muckenhoupt*, Proc. Amer. Math. Soc. **90** (1984), 253–257.
[9] A. Osękowski, *Sharp maximal estimates for BMO martingales*, Osaka J. Math. **52** (2015), 1125-1142.
[10] A. Osękowski, *A sharp estimate for Muckenhoupt class A_∞ and BMO*, Positivity **23** (2019), 711–725.
[11] A. Osękowski, *On the best constant in the martingale version of Fefferman's inequality*, Bernoulli **26** (2020), 1912–1926.
[12] L. Slavin and V. Vasyunin, *Sharp results in the integral-form John-Nirenberg inequality*, *Trans. Amer. Math. Soc.* **363**, 2011, 4135–4169.
[13] L. Slavin and V. Vasyunin, *Sharp L^p estimates on BMO*, Indiana Math. J. **61** (2012), 1051–1110.
[14] L. Slavin and P. Zatitskii, *Dimension-free estimates for semigroup BMO and A_p*, available at https://arxiv.org/pdf/1908.02602.pdf.
[15] D. M. Stolyarov and P. B. Zatitskiy, *Theory of locally concave functions and its applications to sharp estimates of integral functionals*, Adv. Math. **291** (2016), 228–273.

[16] D. M. Stolyarov and P. B. Zatitskiy, *Sharp transference principle for BMO and A_p*, available at https://arxiv.org/pdf/1908.09497.pdf.
[17] V. Vasyunin, *The exact constant in the inverse Hölder inequality for Muckenhoupt weights (Russian)*, Algebra i Analiz **15** (2003), 73–117; translation in St. Petersburg Math. J. **15** (2004), pp. 49–79.

Department of Mathematics, Informatics and Mechanics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland
Email address: `A.Osekowski@mimuw.edu.pl`