SHARP RESULTS IN THE INTEGRAL-FORM
JOHN–NIRENBERG INEQUALITY

L. SLAVIN AND V. VASYUNIN

ABSTRACT. We consider the strong form of the John-Nirenberg inequality for the L^2-based BMO. We construct explicit Bellman functions for the inequality in the continuous and dyadic settings and obtain the sharp constant, as well as the precise bound on the inequality's range of validity, both previously unknown. The results for the two cases are substantially different. The paper not only gives another instance in the short list of such explicit calculations, but also presents the Bellman function method as a sequence of clear steps, adaptable to a wide variety of applications.

CONTENTS

1. INTRODUCTION

In this paper, we are dealing with the space BMO that first appeared in the classical paper [3]. A crucial property of elements of BMO, the exponential decay of their distribution function, was also established in that paper; it is now known as the weak-form John–Nirenberg inequality.

4135

For an interval I and a real-valued function $\varphi \in L^1(I)$, let $\langle\varphi\rangle_I$ be the average of φ over I, $\langle\varphi\rangle_I = \frac{1}{|I|}\int_I \varphi$. For $1 \le p < \infty$, let

$$(1.1) \qquad \mathrm{BMO}(I) = \left\{\varphi \in L^1(I) : \langle|\varphi - \langle\varphi\rangle_J|^p\rangle_J \le C^p < \infty, \ \forall \text{ interval } J \subset I\right\}$$

with the best (smallest) such C being the corresponding norm of φ. The classical definition of John and Nirenberg uses $p = 1$; it is known that the norms for different p's are equivalent. For every $\varphi \in \mathrm{BMO}(I)$ and every $\lambda \in \mathbb{R}$ one has

Theorem (John, Nirenberg; weak form).

$$(1.2) \qquad\qquad \frac{1}{|I|}\left|\{s \in I : \ \varphi - \langle\varphi\rangle_I > \lambda\}\right| \le c_1 e^{-c_2\lambda/\|\varphi\|_{\mathrm{BMO}(I)}}.$$

BMO plays a major role in modern analysis (in particular, because it is dual to the Hardy space H^1 [2]). In addition, inequality (1.2) can be viewed as an accurate characterization of unbounded BMO functions. Thus, finding sharp constants c_1 and c_2 is of much interest. For the classical case $p = 1$, Korenvoskii [4] established the exact value $c_2 = 2/e$. Inequality (1.1) can be integrated to produce an equivalent statement. For $\varepsilon \ge 0$, let

$$\mathrm{BMO}_\varepsilon(I) = \{\varphi \in \mathrm{BMO}(I) : \|\varphi\| \le \varepsilon\}.$$

Then we have

Theorem (John, Nirenberg; integral form). *There exists $\varepsilon_0 > 0$ such that for every $0 \le \varepsilon < \varepsilon_0$ there is $C(\varepsilon) > 0$ such that for any function $\varphi \in \mathrm{BMO}_\varepsilon(I)$,*

$$(1.3) \qquad\qquad\qquad \langle e^\varphi\rangle_I \le C(\varepsilon)e^{\langle\varphi\rangle_I}.$$

This paper has two main objectives: the first one is to establish, for the case $p = 2$, the sharp values for ε_0 and $C(\varepsilon)$ in (1.3). We accomplish this for the continuous BMO defined above as well as its dyadic analog BMO^d, for which every subinterval J of I in definition (1.1) is an element of the dyadic lattice rooted in I.

The second objective is to showcase the tool that is at the center of the proofs. It is the Bellman function method, a powerful harmonic analysis technique developed during the past 12 years. In the important paper [1], Burkholder found what can now be understood as the first explicit harmonic analysis Bellman function. However, his language was different from ours, and the method did not appear in its present form until 1995, when a two-weight martingale transform was handled in [8] (later published as [9]). In the big paper [6], the authors define many Bellman functions, as a matter of both developing the method and solving several important problems. Many results, using different variants of the technique, have followed. However, until [13] was published in 2003, none had found their Bellman functions explicitly, instead relying on Bellman-type arguments, when one uses a substitute function with required size and concavity properties.

The list of explicit Bellman functions is still very short. Besides [13], we note the papers [5, 15, 11]; several others are in the works. While the present paper gives one of the earliest known such computations (see [10, 14]), it has taken time to bring it to print. Finding the corresponding Bellman function exactly will always yield sharp results for an inequality, but this paper also has methodical value: it is our hope that it will further a new paradigm in Bellman investigations and help bring about a new pure-Bellman template. To describe it briefly, upon choosing the Bellman variables and setting up the corresponding extremal problem, one is

to establish the finite-difference inequality(ies) codifying the concavity (convexity) of the Bellman function along the trajectories defined by the choice of variables. The inequality then is rephrased as a set of partial differential conditions, which are "sharpened" to become what we now call "the Bellman PDE". Using homogeneity inherent in the problem, one reduces the order of the PDE and finds a solution, a "candidate" Bellman function. Then, one proves that the candidate is indeed the true Bellman function, using a dyadic-type induction on scales in one direction and finding an extremal function to establish the other. We follow this template in both continuous and dyadic cases.

Surprisingly, in our case the solution of that PDE turns out to be the Bellman function for the *continuous* John–Nirenberg setup, and that takes a substantial amount of work to show. We then solve the dyadic case, using the continuous solution as a starting point. The results for the two cases turn out to be substantially different.

As its name suggests, the method has its origins in stochastic optimal control. We refer the reader to papers [7, 16], where the connection between the two incarnations of the method is explored. In fact, it was in an early version of [7] where we first saw a Bellman setup for a dyadic version of inequality (1.3). The authors then stated a formal PDE for the problem and found a majorant of its solution, in effect establishing the dyadic inequality with some suboptimal values. Each of us, independently, solved the PDE exactly, and we then pooled our efforts to proceed from this formal solution to the rigorous proof of our theorems. We would especially like to acknowledge the help of A. Volberg who formulated the problem to each of us and brought us together.

2. The Bellman setup

We use definition (1.1) with $p = 2$. The reason is that it can be rewritten as

$$\mathrm{BMO}(I) = \left\{ \varphi \in L^1(I) : \langle \varphi^2 \rangle_J - \langle \varphi \rangle_J^2 \leq C^2, \ \forall \ \text{interval} \ J \subset I \right\}$$

with the norm

$$\|\varphi\|_{\mathrm{BMO}(I)} = \left(\sup_{J \subset I} \left\{ \langle \varphi^2 \rangle_J - \langle \varphi \rangle_J^2 \right\} \right)^{1/2},$$

with the appropriate modifications for the dyadic space BMO^d. This rewriting greatly facilitates the description of the problem in terms of Bellman variables, as shown below.

As mentioned before, by $\mathrm{BMO}_\varepsilon(I)$ and $\mathrm{BMO}_\varepsilon^d(I)$ we denote the ε-ball (the ball of radius ε centered at 0) in the corresponding space. With every such ball and the set of all subintervals $J \subset I$ we associate the domain $\Omega_\varepsilon = \{x = (x_1, x_2) : x_1 \in \mathbb{R}, \ x_1^2 \leq x_2 \leq x_1^2 + \varepsilon^2\}$ as follows:

$$(2.1) \qquad\qquad (\varphi, J) \longmapsto \left(\langle \varphi \rangle_J, \langle \varphi^2 \rangle_J \right).$$

This map is well-defined because $\langle\varphi\rangle_J^2 \leq \langle\varphi^2\rangle_J$ (Cauchy inequality) and $\varphi \in$ $\mathrm{BMO}_\varepsilon(I)$ ($\mathrm{BMO}_\varepsilon^d(I)$). On Ω_ε, we define the following Bellman functions:

$$(2.2) \qquad \boldsymbol{B}_\varepsilon^+(x) = \sup_{\varphi\in\mathrm{BMO}_\varepsilon(I)} \left\{ \langle e^\varphi\rangle_I : \langle\varphi\rangle_I = x_1, \langle\varphi^2\rangle_I = x_2 \right\},$$

$$(2.3) \qquad \boldsymbol{B}_\varepsilon^-(x) = \inf_{\varphi\in\mathrm{BMO}_\varepsilon(I)} \left\{ \langle e^\varphi\rangle_I : \langle\varphi\rangle_I = x_1, \langle\varphi^2\rangle_I = x_2 \right\},$$

$$(2.4) \qquad \boldsymbol{B}_\varepsilon^{d+}(x) = \sup_{\varphi\in\mathrm{BMO}_\varepsilon^d(I)} \left\{ \langle e^\varphi\rangle_I : \langle\varphi\rangle_I = x_1, \langle\varphi^2\rangle_I = x_2 \right\},$$

$$(2.5) \qquad \boldsymbol{B}_\varepsilon^{d-}(x) = \inf_{\varphi\in\mathrm{BMO}_\varepsilon^d(I)} \left\{ \langle e^\varphi\rangle_I : \langle\varphi\rangle_I = x_1, \langle\varphi^2\rangle_I = x_2 \right\}.$$

Observe that these functions do not depend on I. The functions with "+" give the exact upper bound on $\langle e^\varphi\rangle_I$ (and so the sharp John–Nirenberg inequality), while the ones with "−" give the lower bound. While the overall lower bound (over all x) is well known ($\langle e^\varphi\rangle_I \geq e^{\langle\varphi\rangle_I}$, by Jensen's inequality), the lower Bellman functions give nontrivial results for each particular choice of x. In addition, they arise naturally in the process of solving the Bellman PDE.

Until now, a typical Bellman function paper would first establish a dyadic result and then modify the proof to cover the continuous case. A notable feature of our result is that we first find a family of "continuous" Bellman functions and then choose appropriate members of that family to deal with the dyadic case.

3. Main results

Throughout the paper, we will mark results about the continuous case with index "c" and their dyadic analogs with index "d."

Theorem 1c. *Let $\varepsilon_0 = 1$. For every $0 \leq \varepsilon < \varepsilon_0$, let*

$$(3.1) \qquad C(\varepsilon) = \frac{e^{-\varepsilon}}{1-\varepsilon}.$$

Then, for any $\varphi \in \mathrm{BMO}_\varepsilon(I)$,

$$(3.2) \qquad \langle e^\varphi\rangle_I \leq C(\varepsilon)e^{\langle\varphi\rangle_I}.$$

Moreover, ε_0 and $C(\varepsilon)$ are sharp.

Theorem 1d. *Let $\varepsilon_0^d = \sqrt{2}\log 2$. For every $0 \leq \varepsilon < \varepsilon_0^d$, let*

$$(3.3) \qquad C^d(\varepsilon) = \frac{e^{-\frac{\varepsilon}{\sqrt{2}}}}{2 - e^{\frac{\varepsilon}{\sqrt{2}}}}.$$

Then, for any $\varphi \in \mathrm{BMO}_\varepsilon^d(I)$,

$$(3.4) \qquad \langle e^\varphi\rangle_I \leq C^d(\varepsilon)e^{\langle\varphi\rangle_I}.$$

Moreover, ε_0^d and $C^d(\varepsilon)$ are sharp.

In our proofs, we will repeatedly use the function

$$H(t) = (1-t)e^t.$$

The following proposition summarizes the simple properties of H we will need; the proof is omitted.

Proposition 1.

(a) H *is strictly increasing for* $t < 0$ *and strictly decreasing for* $t > 0$;

(b) $H(-t) > H(t)$ *for* $t > 0$.

Theorems 1c and 1d are immediate consequences of Proposition 1 and the following results for the Bellman functions (2.2)–(2.5). Let

$$(3.5) \qquad B_\delta^\pm(x) = \frac{H\left(\pm\sqrt{\delta^2 + x_1^2 - x_2}\,\right)}{H(\pm\delta)}\, e^{x_1}.$$

Theorem 2c. *If* $0 \le \varepsilon < 1$, *then*

$$\boldsymbol{B}_\varepsilon^+(x) = B_\varepsilon^+(x).$$

If $\varepsilon \ge 1$, *then*

$$\boldsymbol{B}_\varepsilon^+(x) = \begin{cases} e^{x_1} & \text{if } x_2 = x_1^2, \\ +\infty & \text{if } x_2 > x_1^2. \end{cases}$$

In addition,

$$\boldsymbol{B}_\varepsilon^-(x) = B_\varepsilon^-(x), \qquad \forall \varepsilon \ge 0.$$

Theorem 2d. *If* $0 \le \varepsilon < \sqrt{2}\log 2$, *then*

$$\boldsymbol{B}_\varepsilon^{d+}(x) = B_{\delta^+(\varepsilon)}^+(x),$$

where $\delta = \delta^+(\varepsilon)$ *is the unique solution of the equation*

$$(3.6) \qquad (1 - \sqrt{\delta^2 - \varepsilon^2})e^{\sqrt{\delta^2 - \varepsilon^2}}\left(2 - e^{\varepsilon/\sqrt{2}}\right) - (1 - \delta)e^{\delta - \varepsilon/\sqrt{2}} = 0$$

in the interval $(\varepsilon, 1)$.

If $\varepsilon \ge \sqrt{2}\log 2$, *then*

$$\boldsymbol{B}_\varepsilon^{d+}(x) = \begin{cases} e^{x_1} & \text{if } x_2 = x_1^2, \\ +\infty & \text{if } x_2 > x_1^2. \end{cases}$$

In addition,

$$\boldsymbol{B}_\varepsilon^{d-}(x) = B_{\delta^-(\varepsilon)}^-(x), \qquad \forall \varepsilon \ge 0,$$

where $\delta = \delta^-(\varepsilon)$ *is the unique solution of the equation*

$$(3.7) \qquad (1 + \sqrt{\delta^2 - \varepsilon^2})e^{-\sqrt{\delta^2 - \varepsilon^2}}\left(2 - e^{-\varepsilon/\sqrt{2}}\right) - (1 + \delta)e^{-\delta + \varepsilon/\sqrt{2}} = 0$$

in the interval $\left(\varepsilon, \frac{3}{2\sqrt{2}}\varepsilon\right)$.

Indeed, Proposition 1 implies that the functions B_δ^+, for any $\delta \ge \varepsilon$, assume their maxima on Ω_ε on its upper boundary, i.e. when $x_2 = x_1^2 + \varepsilon^2$. Then we have

$$B_\varepsilon^+(x) \le \frac{H(0)}{H(\varepsilon)} = \frac{e^{-\varepsilon}}{1 - \varepsilon}\, e^{x_1}$$

and

$$B_{\delta^+(\varepsilon)}^+(x) \le \frac{H(\sqrt{\delta^+(\varepsilon)^2 - \varepsilon^2}\,)}{H(\delta^+(\varepsilon))}\, e^{x_1} = \frac{e^{-\frac{\varepsilon}{\sqrt{2}}}}{2 - e^{\frac{\varepsilon}{\sqrt{2}}}}\, e^{x_1},$$

giving (3.2) and (3.4) with the sharp constants (3.1) and (3.3).

The rest of the paper is devoted to proving Theorems 2c and 2d. We will first consider the continuous case and then the dyadic one.

4. The continuous case

We split the proof of Theorem 2c into two parts.

Lemma 1c. *For every $x \in \Omega_\varepsilon$,*

$$(4.1) \qquad \boldsymbol{B}_\varepsilon^+(x) \geq B_\varepsilon^+(x), \qquad \boldsymbol{B}_\varepsilon^-(x) \leq B_\varepsilon^-(x),$$

where $0 \leq \varepsilon < 1$ for B^+ and $\varepsilon \geq 0$ for B^-.

We prove each of inequalities (4.1) by explicitly finding a function φ for every point $x \in \Omega_\varepsilon$ such that $\left(\langle \varphi \rangle_I, \langle \varphi^2 \rangle_I\right) = (x_1, x_2)$ and

$$\langle e^\varphi \rangle_I = B_\varepsilon(x_1, x_2).$$

Here B_ε stands for B_ε^+ or B_ε^-, as appropriate, and the result will then follow from the definition of $\boldsymbol{B}_\varepsilon^\pm$.

Proof. Since $x_2 = x_1^2$ occurs if and only if $\varphi = x_1 = \text{const}$, it is clear that $\boldsymbol{B}_0^\pm(x) = B_0^\pm(x) = e^{x_1}$. So we only need to consider $\varepsilon > 0$.

Take $I = [0,1]$, $a \in (0,1]$, $b \in \mathbb{R}$, $\gamma \in \mathbb{R}\backslash\{0\}$. Let

$$\varphi_{a,b,\gamma}(t) = \begin{cases} \gamma \log \frac{a}{t} + b & \text{for } 0 \leq t \leq a, \\ b & \text{for } a \leq t \leq 1. \end{cases}$$

Let us calculate the BMO norm of $\varphi_{a,b,\gamma}$. To simplify calculations, let $l(t) = \log(a/t)$ and observe that

$$\int (\gamma l(t) + b) \, dt = (\gamma + b)t + \gamma t l(t) + C$$

and

$$\int (\gamma l(t) + b)^2 \, dt = (2\gamma^2 + 2\gamma b + b^2)t + \gamma^2 t l^2(t) + 2\gamma(b + \gamma)t l(t) + C.$$

Take an interval $[c, d] \subset I$. We have the following trichotomy:

(1) $0 \leq c < d \leq a \leq 1$. In this case

$$\langle \varphi \rangle_{[c,d]} = \gamma + b + \gamma \frac{dl(d) - cl(c)}{d - c}$$

and

$$\langle \varphi^2 \rangle_{[c,d]} = 2\gamma^2 + 2\gamma b + b^2 + 2\gamma(b + \gamma)\frac{dl(d) - cl(c)}{d - c} + \gamma^2 \frac{dl^2(d) - cl^2(c)}{d - c}.$$

Therefore,

$$\langle \varphi^2 \rangle_{[c,d]} - \langle \varphi \rangle_{[c,d]}^2 = \gamma^2 + \frac{\gamma^2}{(d-c)^2}\left[(dl^2(d) - cl^2(c))(d - c) - (dl(d) - cl(c))^2\right]$$

$$= \gamma^2 - \frac{\gamma^2 cd}{(d-c)^2}\left[l(d) - l(c)\right]^2 \leq \gamma^2.$$

(2) $0 \leq c \leq a \leq d \leq 1$. In this case

$$\langle \varphi \rangle_{[c,d]} = \frac{-\gamma cl(c) + (b + \gamma)(a - c) + b(d - a)}{d - c} = \gamma \frac{-cl(c) + a - c}{d - c} + b$$

and

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

$$= \frac{\gamma}{d-c}\left[2(b+\gamma)(a-c) - \gamma cl^2(c) - 2(b+\gamma)cl(c)\right] + b^2,$$

so

$$\langle\varphi^2\rangle_{[c,d]} - \langle\varphi\rangle^2_{[c,d]} = \frac{\gamma^2}{d-c}\left[2(a-c) - cl^2(c) - 2cl(c)\right]$$

$$- \frac{\gamma^2}{(d-c)^2}\left[(a-c)^2 - 2c(a-c)l(c) + c^2l^2(c)\right]$$

$$= \frac{\gamma^2}{(d-c)^2}\left[2(a-c)(d-c) - (a-c)^2 - cdl^2(c) - 2c(d-a)l(c)\right]$$

$$\leq \gamma^2\frac{a-c}{d-c}\left[2 - \frac{a-c}{d-c}\right] \leq \gamma^2,$$

since $d \geq a$ and $\log(a/c) \geq 0$ if $a \geq c$. The last inequality follows from the fact that the vertex of the parabola $(x, x(2-x))$ is at $(1,1)$.

(3) $0 \leq a \leq c < d \leq 1$. In this case,

$$\langle\varphi^2\rangle_{[c,d]} - \langle\varphi\rangle^2_{[c,d]} = b^2 - b^2 = 0.$$

We have shown that $\varphi_{a,b,\gamma} \in \mathrm{BMO}_{|\gamma|}(I)$. Also, using Case 2 above with $c = 0$, $d = 1$, we get $\langle\varphi_{a,b,\gamma}\rangle_I = \gamma a + b$ and $\langle\varphi^2_{a,b,\gamma}\rangle_I = 2\gamma^2 a + 2\gamma ab + b^2$. Finally,

$$\langle e^{\varphi_{a,b,\gamma}}\rangle_I = \int_0^a e^b\left(\frac{a}{t}\right)^\gamma dt + \int_a^1 e^b\, dt = \begin{cases} \dfrac{1-\gamma+a\gamma}{1-\gamma}\, e^b & \text{if } \gamma < 1, \\ \infty & \text{if } \gamma \geq 1. \end{cases}$$

Since $\boldsymbol{B}_\varepsilon(x_1, x_1^2) = B_\varepsilon(x_1, x_1^2) = e^{x_1}$ for all ε, we only need to consider the points $x \in \Omega_\varepsilon$ with $x_2 > x_1^2$. Then we can set $a = 1 - \frac{1}{|\gamma|}\sqrt{\gamma^2 + x_1^2 - x_2}$ and $b = x_1 - \gamma a$, which yields $\langle\varphi_{a,b,\gamma}\rangle_I = x_1$, $\langle\varphi^2_{a,b,\gamma}\rangle_I = x_2$. Now, if we put $\gamma = \varepsilon \geq 1$, we get $\boldsymbol{B}_\varepsilon^+(x) = \infty$. For $\gamma = \varepsilon \in (0,1)$, we get

$$\boldsymbol{B}_\varepsilon^+(x) \geq \langle e^{\varphi_{a,b,\gamma}}\rangle_I = \frac{1 - \sqrt{\varepsilon^2 + x_1^2 - x_2}}{1-\varepsilon}\exp\left(x_1 + \sqrt{\varepsilon^2 + x_1^2 - x_2} - \varepsilon\right) = B_\varepsilon^+(x).$$

If we set $\gamma = -\varepsilon \in (-\infty, 0)$, we obtain

$$\boldsymbol{B}_\varepsilon^-(x) \leq \langle e^{\varphi_{a,b,\gamma}}\rangle_I = \frac{1 + \sqrt{\varepsilon^2 + x_1^2 - x_2}}{1+\varepsilon}\exp\left(x_1 - \sqrt{\varepsilon^2 + x_1^2 - x_2} + \varepsilon\right)$$

$$= B_\varepsilon^-(x). \qquad \square$$

Lemma 2c. *For every $x \in \Omega_\varepsilon$,*

(4.2) $$\boldsymbol{B}_\varepsilon^+(x) \leq B_\varepsilon^+(x), \qquad \boldsymbol{B}_\varepsilon^-(x) \geq B_\varepsilon^-(x),$$

where $0 \leq \varepsilon < 1$ for B^+ and $\varepsilon \geq 0$ for B^-.

Proof. To establish (4.2), we first prove that $\boldsymbol{B}_\varepsilon^+(x) \leq B_{\varepsilon_1}^+(x)$, $\boldsymbol{B}_\varepsilon^-(x) \geq B_{\varepsilon_1}^-(x)$, $\forall\varepsilon_1 > \varepsilon$, $\forall x \in \Omega_\varepsilon$, and take the limit as $\varepsilon_1 \to \varepsilon$. (Observe that B_ε^+ and B_ε^- are continuous in ε.) We need the following two results; their proofs will be postponed until the end of the proof of Lemma 2c.

Lemma 3c. *The function B_ε^+ is locally concave and the function B_ε^- locally convex in Ω_ε, i.e.*

$$(4.3) \qquad \begin{aligned} B_\varepsilon^+(\alpha_- x^- + \alpha_+ x^+) &\geq \alpha_- B_\varepsilon(x^-) + \alpha_+ B_\varepsilon(x^+), \\ B_\varepsilon^-(\alpha_- x^- + \alpha_+ x^+) &\leq \alpha_- B_\varepsilon(x^-) + \alpha_+ B_\varepsilon(x^+) \end{aligned}$$

for any straight-line segment with the endpoints $x^\pm$ that lies entirely in Ω_ε and any pair of nonnegative numbers $\alpha_\pm$ such that $\alpha_- + \alpha_+ = 1$.

Lemma 4c. *Fix ε. Take any $\varepsilon_1 > \varepsilon$. Then for every interval I and every $\varphi \in \mathrm{BMO}_\varepsilon(I)$, there exists a splitting $I = I_- \cup I_+$ such that the whole straight-line segment with the endpoints $x^\pm = \left(\langle \varphi \rangle_{I_\pm}, \langle \varphi^2 \rangle_{I_\pm} \right)$ is inside Ω_{ε_1}. Moreover, the splitting parameter $\alpha_+ = |I_+|/|I|$ can be chosen uniformly (with respect to φ and I) separated from 0 and 1.*

Assuming these lemmas for the moment, take $\varphi \in \mathrm{BMO}_\varepsilon(I)$. Take any $\varepsilon_1 > \varepsilon$. Observe that $\varphi \in \mathrm{BMO}_\varepsilon(J)$ for any subinterval J of I. Split I according to the rule from Lemma 4c. Let $I^{0,0} = I$, $I^{1,0} = I_-$, and $I^{1,1} = I_+$. Now split I_- and I_+ according to the rule from Lemma 4c and continue this splitting process. By $I^{n,m}$ we denote the intervals of the n-th generation as follows: $I^{n,2k} = I_-^{n-1,k}$ and $I^{n,2k+1} = I_+^{n-1,k}$, so the second index runs from 0 to $2^n - 1$. We call the quasi-dyadic lattice so obtained $D_\varphi = D_\varphi(I)$. Let $x^{n,m} = \left(\langle \varphi \rangle_{I^{n,m}}, \langle \varphi^2 \rangle_{I^{n,m}} \right)$. Since Lemma 4c provides for the value of α_+ uniformly separated from 0 and 1 on every step, we have

$$\max_{k=0,1,\dots,2^n-1} \left\{ |I^{n,k}| \right\} \longrightarrow 0 \text{ as } n \to \infty.$$

With this notation, for a given $\varphi \in \mathrm{BMO}_\varepsilon(J)$ let us now introduce two sequences of step functions, $\varphi_n(s) = x_1^{n,k}$ and $s_n(s) = x_2^{n,k} - (x_1^{n,k})^2$ for $s \in I^{n,k}$. Note that $\varphi_n - \langle \varphi \rangle_I$ is the partial sum of the expansion of the function $\varphi - \langle \varphi \rangle_I$ with respect to the orthonormal family of the generalized Haar functions related to $D_\varphi(I)$:

$$h_J = \begin{cases} \left(\dfrac{|J_+|}{|J||J_-|} \right)^{1/2} & \text{on } J_-, \\[2ex] -\left(\dfrac{|J_-|}{|J||J_+|} \right)^{1/2} & \text{on } J_+. \end{cases}$$

It is clear that because the lengths of intervals $I^{n,k}$ go to zero as $n \to \infty$, the family $\{h_J\}_{J \in D_\varphi}$ forms a basis in $L_0^2(I) = \{\psi \in L^2(I) : \langle \psi \rangle_I = 0\}$. So $\varphi_n \to \varphi$ in the L^2-norm, and since

$$\|\varphi - \varphi_n\|_{L^2}^2 = \int_I |\varphi(s) - \varphi_n(s)|^2 ds = \sum_{I^{n,k}} \int_{I^{n,k}} |\varphi(s) - \varphi_n(s)|^2 ds$$

$$= \sum_{I^{n,k}} |I^{n,k}| \left(x_2^{n,k} - (x_1^{n,k})^2 \right) = \int_I s_n(s)\, ds,$$

we can choose a subsequence n_j such that $\varphi_{n_j}(s) \to \varphi(s)$ and $s_{n_j}(s) \to 0$ almost everywhere on I.

Now, using the statement about B^+ from Lemma 3c repeatedly, we get

$$
\begin{aligned}
B^+_{\varepsilon_1}(x^{0,0}) \;\geq\; & \frac{|I^{1,0}|}{|I^{0,0}|}B_{\varepsilon_1}(x^{1,0}) + \frac{|I^{1,1}|}{|I^{0,0}|}B^+_{\varepsilon_1}(x^{1,1}) \\[2mm]
\geq\; & \frac{|I^{1,0}|}{|I^{0,0}|}\frac{|I^{2,0}|}{|I^{1,0}|}B^+_{\varepsilon_1}(x^{2,0}) + \frac{|I^{1,0}|}{|I^{0,0}|}\frac{|I^{2,1}|}{|I^{1,0}|}B^+_{\varepsilon_1}(x^{2,1}) \\[2mm]
& + \frac{|I^{1,1}|}{|I^{0,0}|}\frac{|I^{2,2}|}{|I^{1,1}|}B^+_{\varepsilon_1}(x^{2,2}) + \frac{|I^{1,1}|}{|I^{0,0}|}\frac{|I^{2,3}|}{|I^{1,1}|}B^+_{\varepsilon_1}(x^{2,3}) \\[2mm]
=\; & \frac{|I^{2,0}|}{|I^{0,0}|}B^+_{\varepsilon_1}(x^{2,0}) + \frac{|I^{2,1}|}{|I^{0,0}|}B^+_{\varepsilon_1}(x^{2,1}) + \frac{|I^{2,2}|}{|I^{0,0}|}B^+_{\varepsilon_1}(x^{2,2}) + \frac{|I^{2,3}|}{|I^{0,0}|}B^+_{\varepsilon_1}(x^{2,3}) \\[2mm]
\geq\; & \frac{1}{|I^{0,0}|}\sum_{k=0}^{2^n-1}|I^{n,k}|B^+_{\varepsilon_1}(x^{n,k}) = \frac{1}{|I|}\int_I e^{\varphi_n(s)}b_+(s_n(s))\,ds,
\end{aligned}
$$

(4.4)

where

$$
b_+(t) = \frac{H\!\left(\sqrt{\varepsilon_1^2 - t}\,\right)}{H(\varepsilon_1)}.
$$

The last equality is just the statement $B^+_{\varepsilon_1}(x^{n,k}) = e^{\varphi_n(s)}b_+(s_n(s))$, for $s \in I^{n,k}$.

Likewise, applying the corresponding statement from Lemma 3c repeatedly, we obtain

$$
(4.5) \qquad B^-_{\varepsilon_1}(x^{0,0}) \leq \frac{1}{|I^{0,0}|}\sum_{k=0}^{2^n-1}|I^{n,k}|B^-_{\varepsilon_1}(x^{n,k}) = \frac{1}{|I|}\int_I e^{\varphi_n(s)}b_-(s_n(s))\,ds.
$$

Here

$$
b_-(t) = \frac{H\!\left(-\sqrt{\varepsilon_1^2 - t}\,\right)}{H(-\varepsilon_1)}.
$$

For functions φ bounded from above we can pass to the limit in (4.4) and (4.5) using the dominated convergence theorem. Therefore, for such functions $\varphi \in \mathrm{BMO}_\varepsilon(J)$ we have the double inequality

$$
(4.6) \qquad B^-_{\varepsilon_1}(\langle\varphi\rangle_I, \langle\varphi^2\rangle_I) \leq \frac{1}{|I|}\int_I e^{\varphi(s)}\,ds \leq B^+_{\varepsilon_1}(\langle\varphi\rangle_I, \langle\varphi^2\rangle_I).
$$

It remains to approximate an arbitrary function $\psi \in \mathrm{BMO}_\varepsilon(I)$ by its cut-offs in a standard manner; namely, we take

$$
\psi_m(s) = \begin{cases} \psi(s) & \text{if } \psi(s) \leq m, \\ m & \text{if } \psi(s) > m. \end{cases}
$$

If we denote $J_1 = \{s \in J : \psi(s) \leq m\}$ and $J_2 = \{s \in J : \psi(s) > m\}$, we have the following identity:

$$
\begin{aligned}
&\left(\langle\psi^2\rangle_J - (\langle\psi\rangle_J)^2\right) - \left(\langle\psi_m^2\rangle_J - (\langle\psi_m\rangle_J)^2\right) \\
&= \frac{|J_2|}{|J|}\left(\langle\psi^2\rangle_{J_2} - (\langle\psi\rangle_{J_2})^2\right) + \frac{|J_2||J_1|}{|J|^2}\left(\langle\psi\rangle_{J_2} - m)\right)\left(\langle\psi\rangle_{J_2} + m - 2\langle\psi\rangle_{J_1}\right) \geq 0,
\end{aligned}
$$

which implies that if ψ is in $\mathrm{BMO}_\varepsilon(I)$, then so is ψ_m. Therefore, for $\varphi = \psi_m$ inequalities (4.6) hold, and we can pass to the limit as $m \to \infty$. Clearly, the averages of ψ_m converge to the averages of ψ, and the values of $B^\pm_{\varepsilon_1}(\langle\psi_m\rangle, \langle\psi_m^2\rangle)$

converge to $B_{\varepsilon_1}^{\pm}(\langle\psi\rangle, \langle\psi^2\rangle)$ because of continuity of the functions $B^{\pm}$. Due to the monotone convergence of ψ_m we can pass to the limit under the integral. Taking first the supremum and then infimum over all $\psi \in \mathrm{BMO}_{\varepsilon}(I)$ with $\langle\psi\rangle_I = x_1$ and $\langle\psi^2\rangle_I = x_2$, we obtain the inequalities

$$B_{\varepsilon_1}^+(x) \geq \boldsymbol{B}_{\varepsilon}^+(x), \qquad B_{\varepsilon_1}^-(x) \leq \boldsymbol{B}_{\varepsilon}^-(x),$$

thus proving the lemma. $\qquad\square$

Proof of Lemma 3c. To prove the lemma, we need to check that

$$(4.7) \qquad\qquad \mp\frac{\partial^2 B_{\varepsilon}^{\pm}}{\partial x_i \partial x_j}$$

is a nonnegative matrix. Direct calculation yields

$$\frac{\partial B_{\varepsilon}^{\pm}}{\partial x_1} = \frac{1 - x_1 \mp \sqrt{\varepsilon^2 + x_1^2 - x_2}}{1 \mp \varepsilon} \exp\left\{x_1 \pm \sqrt{\varepsilon^2 + x_1^2 - x_2} \mp \varepsilon\right\},$$

$$\frac{\partial B_{\varepsilon}^{\pm}}{\partial x_2} = \frac{1}{2(1 \mp \varepsilon)} \exp\left\{x_1 \pm \sqrt{\varepsilon^2 + x_1^2 - x_2} \mp \varepsilon\right\},$$

$$\frac{\partial^2 B_{\varepsilon}^{\pm}}{\partial x_1^2} = \mp\frac{\left(x_1 \pm \sqrt{\varepsilon^2 + x_1^2 - x_2}\right)^2}{\sqrt{\varepsilon^2 + x_1^2 - x_2}(1 \mp \varepsilon)} \exp\left\{x_1 \pm \sqrt{\varepsilon^2 + x_1^2 - x_2} \mp \varepsilon\right\},$$

$$\frac{\partial^2 B_{\varepsilon}^{\pm}}{\partial x_1 \partial x_2} = \pm\frac{x_1 \pm \sqrt{\varepsilon^2 + x_1^2 - x_2}}{2\sqrt{\varepsilon^2 + x_1^2 - x_2}(1 \mp \varepsilon)} \exp\left\{x_1 \pm \sqrt{\varepsilon^2 + x_1^2 - x_2} \mp \varepsilon\right\},$$

$$\frac{\partial^2 B_{\varepsilon}^{\pm}}{\partial x_2^2} = \mp\frac{1}{4\sqrt{\varepsilon^2 + x_1^2 - x_2}(1 \mp \varepsilon)} \exp\left\{x_1 \pm \sqrt{\varepsilon^2 + x_1^2 - x_2} \mp \varepsilon\right\}.$$

Therefore, the quadratic form of the matrix (4.7) is

(4.8)

$$\mp\sum_{i,j=1}^{2} \frac{\partial^2 B_{\varepsilon}^{\pm}}{\partial x_i \partial x_j} \Delta_i \Delta_j$$

$$= \frac{\left(\left(x_1 \pm \sqrt{\varepsilon^2 + x_1^2 - x_2}\right)\Delta_1 - \frac{1}{2}\Delta_2\right)^2}{\sqrt{\varepsilon^2 + x_1^2 - x_2}(1 \mp \varepsilon)} \exp\left\{x_1 \pm \sqrt{\varepsilon^2 + x_1^2 - x_2} \mp \varepsilon\right\} \geq 0,$$

which establishes the result. $\qquad\square$

Proof of Lemma 4c. We fix an interval I and a function $\varphi \in \mathrm{BMO}_{\varepsilon}(I)$. We now explicitly construct an algorithm to find the splitting $I = I_- \cup I_+$, i.e. choose the splitting parameters $\alpha_{\pm} = |I_{\pm}|/|I|$. As before, $x_1^{\pm} = \langle\varphi\rangle_{I_{\pm}}$, $x_2^{\pm} = \langle\varphi^2\rangle_{I_{\pm}}$. Also, put $x_1^0 = \langle\varphi\rangle_I$ and $x_2^0 = \langle\varphi^2\rangle_I$. Lastly, by $[s, t]$ we will denote the straight-line segment connecting two points s and t in the plane.

First, we take $\alpha_- = \alpha_+ = \frac{1}{2}$ (see Figure 1). If the whole segment $[x^-, x^+]$ is in Ω_{ε_1}, we fix this splitting. Assuming it is not the case, there exists a point x on this segment with $x_2 - x_1^2 > \varepsilon_1^2$. Observe that only one of the segments $[x^-, x^0]$ and $[x^+, x^0]$ contains such points. Call the corresponding endpoint (x^- or x^+) ξ. Its position is completely defined by the choice of α_+. Define the function ρ as follows: $\rho(\alpha_+) = \max_{x \in [\xi, x^0]}\{x_2 - x_1^2\}$. By assumption, $\rho\left(\frac{1}{2}\right) > \varepsilon_1^2$. We will now change α_+ so that ξ approaches x^0, i.e. we will increase α_+ if $\xi = x^+$ and decrease it if
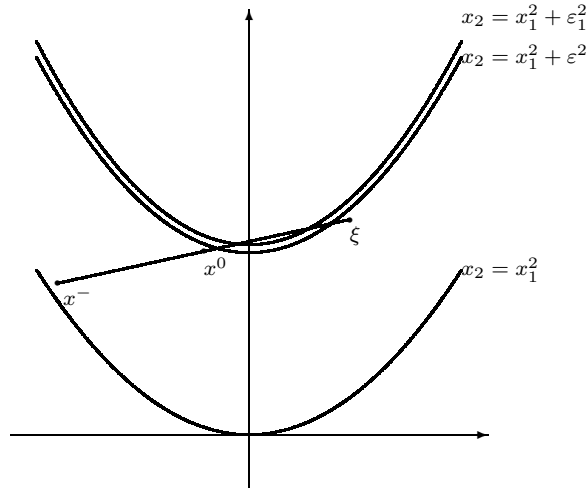
FIGURE 1. The initial splitting: $\alpha_- = \alpha_+ = \frac{1}{2}$, $\xi = x^+$.

$\xi = x^-$. We stop when $\rho(\alpha_+) = \varepsilon_1^2$ and fix that splitting. It remains to check that such a moment occurs and that the corresponding α_+ is separated from 0 and 1. Without loss of generality, assume that $\xi = x^+$. Let $I = [a, b]$. Since $\varphi \in L^2(I)$, the functions $\xi_1(\alpha_+) = \frac{1}{\alpha_+} \int_{b-|I|\alpha_+}^b \varphi(w) \, dw$ and $\xi_2(\alpha_+) = \frac{1}{\alpha_+} \int_{b-|I|\alpha_+}^b \varphi^2(w) \, dw$ are continuous on the interval $(0, 1]$ and $\xi(1) = x^0$. Therefore, ρ is continuous on $(0, 1]$. Since $\rho\left(\frac{1}{2}\right) > \varepsilon_1^2$ and $\rho(1) \leq \varepsilon^2 < \varepsilon_1^2$ (recall that $x^0 \in \Omega_\varepsilon$), we conclude that there is a point $\alpha_+ \in \left[\frac{1}{2}, 1\right]$ with $\rho(\alpha_+) = \varepsilon_1^2$ (Figure 2).

Having just proved that the desired point exists, we need to check that the corresponding α_+ is not too close to 0 or 1. If $\xi = x^+$, we have $\alpha_+ > \frac{1}{2}$ and $\xi_1 - x_1^0 = x_1^+ - x_1^0 = \alpha_-(x_1^+ - x_1^-)$. Likewise, if $\xi = x^-$, we have $\alpha_- > \frac{1}{2}$ and $\xi_1 - x_1^0 = x_1^- - x_1^0 = \alpha_+(x_1^- - x_1^+)$. Thus $|\xi_1 - x_1^0| = \min\{\alpha_\pm\}|x_1^- - x_1^+|$. For the stopping value of α_+, the straight line through the points x^-, x^+ and x^0 is tangent to the parabola $x_2 = x_1^2 + \varepsilon_1^2$ at some point y. The equation of this line is, therefore, $x_2 = 2x_1 y_1 - y_1^2 + \varepsilon_1^2$. The line intersects the graph of $x_2 = x_1^2 + \varepsilon^2$ at the points

$$x_\varepsilon^\pm = \left(y_1 \pm \sqrt{\varepsilon_1^2 - \varepsilon^2}, y_2 \pm 2y_1\sqrt{\varepsilon_1^2 - \varepsilon^2}\right)$$

and the graph of $x_2 = x_1^2$ at the points

$$x_0^\pm = (y_1 \pm \varepsilon_1, y_2 \pm 2y_1 \varepsilon_1).$$

We then have

$$[x_\varepsilon^-, x_\varepsilon^+] \subset [x^0, \xi] \subset [x^-, x^+] \subset [x_0^-, x_0^+]$$

and, therefore,

$$\begin{aligned} 2\sqrt{\varepsilon_1^2 - \varepsilon^2} &= |(x_\varepsilon^+)_1 - (x_\varepsilon^-)_1| \leq |x_1^0 - \xi_1| = \min\{\alpha_\pm\}|x_1^+ - x_1^-| \\ &\leq \min\{\alpha_\pm\}|(x_0^+)_1 - (x_0^-)_1| = \min\{\alpha_\pm\}2\varepsilon_1, \end{aligned}$$
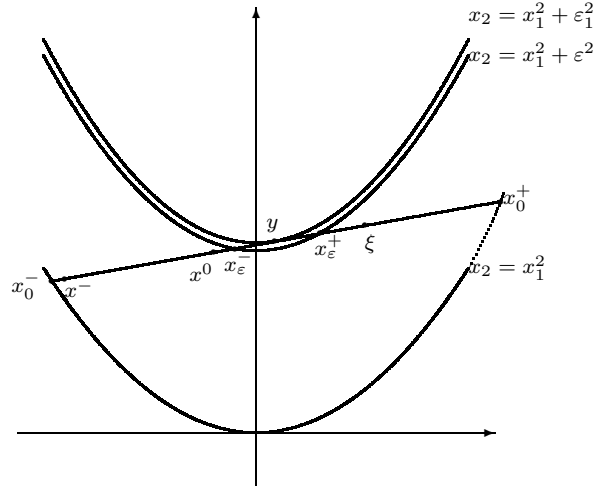
FIGURE 2. The stopping time: $[x^-, \xi]$ is tangent to the parabola $x_2 = x_1^2 + \varepsilon^2$.

which implies

$$\sqrt{1 - \left(\frac{\varepsilon}{\varepsilon_1}\right)^2} \leq \alpha_+ \leq 1 - \sqrt{1 - \left(\frac{\varepsilon}{\varepsilon_1}\right)^2}.$$

As promised, this estimate does not depend on φ or I. $\qquad \square$

4.1. **How to find the Bellman function.** We first observe that the Bellman functions $\boldsymbol{B}^\pm$ must be of the form

$$(4.9) \qquad\qquad \boldsymbol{B}_\varepsilon^\pm(x) = \exp\left\{x_1 + w_\varepsilon^\pm(x_2 - x_1^2)\right\}$$

for some positive functions $w^\pm$ on $[0, \varepsilon^2]$ such that $w_\varepsilon^\pm(0) = 0$.

Indeed, fix an interval I. Then $\varphi \in \mathrm{BMO}_\varepsilon(I)$ if and only if $\varphi + c \in \mathrm{BMO}_\varepsilon(I)$, where c is an arbitrary constant. Let $\tilde\varphi = \varphi + c$. We have (all averages are over I) $\langle\tilde\varphi\rangle = \langle\varphi\rangle + c$, $\langle\tilde\varphi\rangle^2 = \langle\varphi^2\rangle + 2c\langle\varphi\rangle + c^2$, and $\langle e^{\tilde\varphi}\rangle = e^c\langle e^\varphi\rangle$. Then

$$\sup_{\varphi \in \mathrm{BMO}_\varepsilon(I)} \left\{\langle e^{\tilde\varphi}\rangle : \langle\varphi\rangle = x_1, \langle\varphi\rangle^2 = x_2\right\} = e^c \sup_{\varphi \in \mathrm{BMO}_\varepsilon(I)} \left\{\langle e^\varphi\rangle : \langle\varphi\rangle = x_1, \langle\varphi^2\rangle = x_2\right\}$$

or

$$\sup_{\tilde\varphi \in \mathrm{BMO}_\varepsilon(I)} \left\{\langle e^{\tilde\varphi}\rangle : \langle\tilde\varphi\rangle = x_1 + c, \langle\tilde\varphi^2\rangle = x_2 + 2cx_1 + c^2\right\}$$

$$= e^c \sup_{\varphi \in \mathrm{BMO}_\varepsilon(I)} \left\{\langle e^\varphi\rangle : \langle\varphi\rangle = x_1, \langle\varphi^2\rangle = x_2\right\}.$$

Completely analogous statements with inf instead of sup can be made. Altogether, we get

$$\boldsymbol{B}_\varepsilon^\pm(x_1 + c, x_2 + 2cx_1 + c^2) = e^c \boldsymbol{B}_\varepsilon^\pm(x_1, x_2).$$

Setting $c = -x_1$ and omitting the index ε, we get

$$\boldsymbol{B}^\pm(0, x_2 - x_1^2) = e^{-x_1} \boldsymbol{B}^\pm(x_1, x_2).$$

Jensen's inequality ($\langle e^\varphi \rangle \geq e^{\langle \varphi \rangle}$) gives $\boldsymbol{B}^\pm(0, x_2 - x_1^2) \geq 1$. Hence, there exists a positive function $w^\pm = \log \boldsymbol{B}^\pm(0, \cdot)$ defined on the interval $[0, \varepsilon^2]$ such that (4.9) holds. Furthermore, $x_2 = x_1 = 0$ if and only if $\varphi = 0$. Thus $\boldsymbol{B}^\pm(0,0) = 1$ and $w^\pm(0) = 0$.

The successful Bellman function candidate B (we will omit the index $\pm$ when no confusion results) must be of the form (4.9). Moreover, to use the machinery of Lemma 2c, we need the statements of Lemma 3c to hold. So we want

$$(4.10) \qquad \mp \frac{\partial^2 B^\pm}{\partial x_i \partial x_j}$$

to be a nonnegative matrix.

Using (4.9), we get

$$\frac{\partial B}{\partial x_1} = (1 - 2x_1 w')B,$$

$$\frac{\partial B}{\partial x_2} = w'B,$$

$$\frac{\partial^2 B}{\partial x_1^2} = \left((1 - 2x_1 w')^2 - 2w' + 4x_1^2 w''\right)B,$$

$$\frac{\partial^2 B}{\partial x_1 \partial x_2} = \left(w'(1 - 2x_1 w') - 2x_1 w''\right)B,$$

$$\frac{\partial^2 B}{\partial x_2^2} = \left((w')^2 + w''\right)B.$$

Matrix (4.10) turns into

$$(4.11) \qquad \mp \begin{bmatrix} \dfrac{\partial^2 B^\pm}{\partial x_1^2} & \dfrac{\partial^2 B^\pm}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 B^\pm}{\partial x_1 \partial x_2} & \dfrac{\partial^2 B^\pm}{\partial x_2^2} \end{bmatrix} = \mp B^\pm \begin{bmatrix} 1 & -2x_1 \\ 0 & 1 \end{bmatrix} R \begin{bmatrix} 1 & 0 \\ -2x_1 & 1 \end{bmatrix},$$

where

$$(4.12) \qquad R = \begin{bmatrix} 1 - 2w' & w' \\ w' & (w')^2 + w'' \end{bmatrix}.$$

For the extremal function (if any) we must have equality at every step in (4.4) and (4.5) in Lemma 2c, so the matrix (4.10) has to be degenerate. Because of the representation (4.11) and (4.12), this translates into

$$(4.13) \qquad (1 - 2w')\left((w')^2 + w''\right) = (w')^2,$$

while the nonnegativity condition (4.10) is equivalent to the inequality

$$(4.14) \qquad \pm(2(w^\pm)' - 1) \geq 0.$$

We solve equation (4.13):

$$(1 - 2w')w'' = 2(w')^3,$$

$$\left(\frac{1}{2(w')^3} - \frac{1}{(w')^2} \right) w'' = 1,$$

$$\left(\frac{1}{w'} - \frac{1}{4(w')^2} \right)' = 1,$$

$$\frac{1}{w'} - \frac{1}{4(w')^2} = t + const,$$

$$-\left(1 - \frac{1}{2w'} \right)^2 = t + const.$$

This implies that the constant has to be nonpositive. We parametrize the family of possible solutions by a positive parameter δ setting $const = -\delta^2$. Then

$$\left(1 - \frac{1}{2w'} \right)^2 = \delta^2 - t$$

and

(4.15)
$$1 - \frac{1}{2w'} = \pm\sqrt{\delta^2 - t}.$$

We see that the solution is defined on the interval $[0, \delta^2]$. Condition (4.14) with "+" means that $w' \geq \frac{1}{2}$. This requires the "+" sign in (4.15), and this square root has to be strictly less than 1. Therefore, the only feasible solution for w^+ is that for $\delta < 1$. We get the solution for w^- by choosing the "−" sign in (4.15); it works for all $\delta > 0$. Thus, equation (4.15) gives

$$(w^{\pm})' = \frac{1}{2(1 \mp \sqrt{\delta^2 - t})}$$

and, taking into account that $w(0) = 0$, we obtain

$$w^{\pm}(t) = \frac{1}{2} \int_0^t \frac{1}{1 \mp \sqrt{\delta^2 - s}} \, ds = \log \frac{1 \mp \sqrt{\delta^2 - t}}{1 \mp \delta} \pm \sqrt{\delta^2 - t} \mp \delta,$$

which, together with (4.9), gives (3.5):

$$B_{\delta}^{\pm}(x) = \frac{1 \mp \sqrt{\delta^2 + x_1^2 - x_2}}{1 \mp \delta} \exp\left(x_1 \pm \sqrt{\delta^2 + x_1^2 - x_2} \mp \delta \right).$$

4.2. **How to find the extremal function.** We now show how to find the extremal function that appeared without an explanation in the proof of Lemma 1c. As mentioned in the previous section, for the extremal function there is equality at every step in the chain of inequalities (4.4). Thus in the splitting process we only proceed along the vector field defined by the kernel of the matrix (4.10). The quadratic form of that matrix is given by (4.8):

(4.16)

$$\mp \sum_{i,j=1}^2 \frac{\partial^2 B_{\delta}^{\pm}}{\partial x_i \partial x_j} \Delta_i \Delta_j$$

$$= \frac{\left(\left(x_1 \pm \sqrt{\delta^2 + x_1^2 - x_2} \right) \Delta_1 - \frac{1}{2}\Delta_2 \right)^2}{\sqrt{\delta^2 + x_1^2 - x_2} \, (1 \mp \delta)} \exp\left\{ x_1 \pm \sqrt{\delta^2 + x_1^2 - x_2} \mp \delta \right\}.$$

Hence, the trajectories along which B is a linear function are given by

$$(4.17) \qquad \left(x_1 \pm \sqrt{\delta^2 + x_1^2 - x_2} \right) dx_1 = \frac{1}{2} dx_2.$$

Introducing the variable $t = \pm\sqrt{\delta^2 + x_1^2 - x_2}$, we have $t^2 = \delta^2 + x_1^2 - x_2$ and $2t\, dt = 2x_1\, dx_1 - dx_2$. Replacing $\frac{1}{2} dx_2$ in (4.17) by $x_1\, dx_1 - t\, dt$, we get $t\, dx_1 = -t\, dt$, i.e. $t = c - x_1$ and

$$(4.18) \qquad x_2 = \delta^2 + x_1^2 - t^2 = 2cx_1 + \delta^2 - c^2.$$

The corresponding trajectories are straight lines tangent to the upper boundary $x_2 = x_1^2 + \delta^2$ of Ω_δ at the point $x = (c, c^2 + \delta^2)$. Consider the following two families of such straight-line segments:

$$\omega_\delta^+(c) = \left\{ x = (x_1, 2cx_1 + \delta^2 - c^2) : c - \delta \le x_1 \le c \right\},$$
$$\omega_\delta^-(c) = \left\{ x = (x_1, 2cx_1 + \delta^2 - c^2) : c \le x_1 \le c + \delta \right\}.$$

Each of these families covers the whole domain, i.e.

$$\Omega_\delta = \bigcup_{c \in \mathbb{R}} \omega_\delta^+(c) = \bigcup_{c \in \mathbb{R}} \omega_\delta^-(c).$$

Furthermore, B^+ is a linear function on each segment $\omega_\delta^+(c)$, while B^- is a linear function on each segment $\omega_\delta^-(c)$. Indeed, since $\sqrt{\delta^2 + x_1^2 - x_2} = |x_1 - c|$ on the line $x_2 = 2cx_1 + \delta^2 - c^2$, we have

$$B_\delta^+(x_1, 2cx_1 + \delta^2 - c^2) = \frac{1 + x_1 - c}{1 - \delta} e^{c-\delta} \text{ for } c - \delta \le x_1 \le c,$$

$$B_\delta^-(x_1, 2cx_1 + \delta^2 - c^2) = \frac{1 + x_1 - c}{1 + \delta} e^{c+\delta} \text{ for } c \le x_1 \le c + \delta.$$

Therefore, if both points $x^\pm$ are on a segment $\omega_\delta^+(c)$ or $\omega_\delta^-(c)$, we have equality in the corresponding line in (4.3) (with $\delta = \varepsilon$).[1]

Note that we have one more "acceptable trajectory": the envelope of the segments $\omega_\delta^+(c)$ (or $\omega_\delta^-(c)$), i.e. the parabola $x_2 = x_1^2 + \delta^2$.

Let x^0 be an arbitrary point inside Ω_δ. Then we make the splitting so that x^- is on the boundary $x_2 = x_1^2 + \delta^2$ and the segment $\omega_\delta^+(x_1^-)$ passes through the point x^0. Every point on that segment satisfies the equation

$$x_2 = 2x_1^- x_1 + \delta^2 - (x_1^-)^2,$$

so $x_1^- = x_1^0 + \sqrt{\delta^2 + (x_1^0)^2 - x_2^0}$. We choose the second endpoint x^+ to be the point of intersection of $\omega_\delta^+(x_1^-)$ and the lower boundary of Ω_δ, $x_2 = x_1^2$. This is equivalent to letting φ be constant on I_+. Then $x_2^+ = (x_1^+)^2 = 2x_1^- x_1^+ + \delta^2 - (x_1^-)^2$ and, hence, $x_1^+ = x_1^- - \delta$.

Assume that φ_c is the extremal function (defined on $[0, 1]$) that corresponds to the point $(c, c^2 + \delta^2)$ on the upper boundary. Then for $\varphi|_{I_-}$ we have to take

---

[1] To avoid misunderstanding, we note that $\pm$ in $x^\pm$ and in $\omega_\delta^\pm$ are independent: $x^\pm$ are two points in the domain Ω_δ whose convex combination is the point x, while $\pm$ in $\omega_\delta^\pm$ means that we consider either B^+ or B^-, as appropriate.

the function $\varphi_{x_1^-}$ rescaled to the interval I_-. So, if $I = [0,1]$, then $I_- = [0,\alpha_-]$, $I_+ = [\alpha_-, 1]$, and

$$(4.19) \qquad \varphi(t) = \begin{cases} \varphi_{x_1^-}\left(\frac{t}{\alpha_-}\right), & 0 \le t < \alpha_-, \\ x_1^+, & \alpha_- \le t \le 1. \end{cases}$$

We have defined the extremal function φ for an arbitrary point of Ω_δ under the assumption that the extremal functions φ_c for the upper boundary are known. Note that it is sufficient to find one of these functions, say φ_0, because $\varphi_c = \varphi_0 + c$. Indeed, it is clear that φ_0 and $\varphi_0 + c$ have the same BMO-norms and

$$\langle \varphi_0 + c \rangle = c, \qquad \langle (\varphi_0 + c)^2 \rangle = \langle \varphi_0^2 \rangle + 2c\langle \varphi_0 \rangle + c^2 = \delta^2 + c^2.$$

Let the point x^0 approach the point x^- along the upper boundary, i.e. let $\alpha_+ \to 0$. If we assume that the extremal function smoothly depends on the point x^0, then the function φ in (4.19) coincides up to terms of the first order in α_+ with the function $\varphi_{x_1^0}$:

$$x_1^0 = \alpha_- x_1^- + \alpha_+ x_1^+ = (1 - \alpha_+)x_1^- + \alpha_+(x_1^- - \delta) = x_1^- - \alpha_+\delta,$$
$$x_2^0 = 2x_1^- x_1^0 + \delta^2 - (x_1^-)^2 = (x_1^0)^2 - (x_1^- - x_1^0)^2 + \delta^2$$
$$= (x_1^0)^2 + (1 - \alpha_+^2)\delta^2 \approx (x_1^0)^2 + \delta^2.$$

Therefore

$$\varphi_{x_1^-}\left(\frac{t}{\alpha_-}\right) \approx \varphi_{x_1^0}(t)$$

up to terms of the first order in α_+. Since

$$\varphi_{x_1^-}\left(\frac{t}{\alpha_-}\right) = \varphi_0\left(\frac{t}{\alpha_-}\right) + x_1^- \approx \varphi_0\left(\frac{t}{1-\alpha_+}\right) + x_1^- \approx x_1^- + \varphi_0(t) + \alpha_+\varphi_0'(t)$$

and

$$\varphi_{x_1^0}(t) = \varphi_0(t) + x_1^- = x_1^- + \varphi_0(t) - \alpha_+\delta,$$

we have

$$t\varphi_0'(t) = -\delta,$$
$$\varphi_0(t) = -\delta \log t + \text{const.}$$

Condition $\langle \varphi_0 \rangle = 0$ implies

$$\varphi_0(t) = \delta\left(\log \frac{1}{t} - 1\right).$$

This yields the function we used to prove Lemma 1c.

5. The dyadic case

To prove Theorem 2d, we follow the procedure of the continuous case. Namely, we first produce extremal functions $\varphi_\pm \in \mathrm{BMO}_\varepsilon^d(I)$ with appropriate averages, for which $\langle e^{\varphi_\pm} \rangle_I = B_{\delta^\pm(\varepsilon)}^\pm$. This proves that $\boldsymbol{B}_\varepsilon^{d+} \ge B_{\delta^+(\varepsilon)}^+$ and $\boldsymbol{B}_\varepsilon^{d-} \le B_{\delta^-(\varepsilon)}^-$. Then, we use a concavity-type result similar to Lemma 3c, which allows us to run the inductive machine of Lemma 2c to prove the converse inequalities.

Lemma 1d. *For every* $x \in \Omega_\varepsilon$,

$$(5.1) \qquad \boldsymbol{B}_\varepsilon^{d+}(x) \ge B_{\delta^+(\varepsilon)}^+(x), \;\; \boldsymbol{B}_\varepsilon^{d-}(x) \le B_{\delta^-(\varepsilon)}^-(x).$$

Proof. Let $I = [0, 1]$. We prove (5.1) by explicitly finding functions $\varphi_+, \varphi_- \in$ $\mathrm{BMO}_\varepsilon^d(I)$ for every $x \in \Omega_\varepsilon$ such that $(\langle \varphi_\pm \rangle_I, \langle \varphi_\pm^2 \rangle_I) = (x_1, x_2)$ and

$$\langle e^{\varphi_+} \rangle_I = B_{\delta^+(\varepsilon)}^+(x), \ \langle e^{\varphi_-} \rangle_I = B_{\delta^-(\varepsilon)}^-(x).$$

As before, we only need to consider $\varepsilon > 0$.

Fix $\varepsilon > 0$. Let the function φ_0 be defined on $I = (0, 1]$ as follows:

$$\varphi_0\big|_{(2^{-(k+1)}, 2^{-k}]} = (k-1)a, \ k = 0, 1, ...,$$

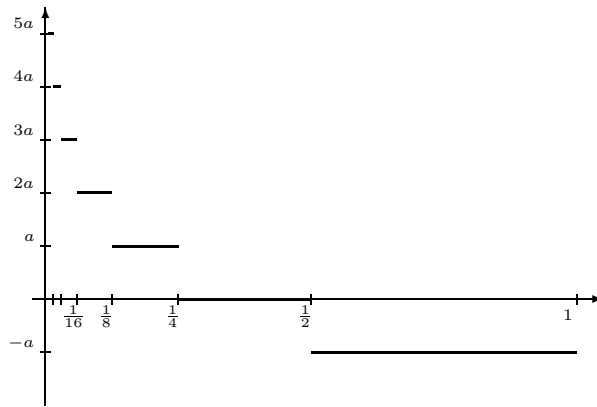with the constant a to be determined later (see Figure 3). We now calculate the

FIGURE 3. The function φ_0.

BMO^d norm of φ_0 and choose a so that $\|\varphi_0\|_{\mathrm{BMO}^d} = \varepsilon$. The only dyadic intervals on which φ_0 is not constant and, hence, $\langle \varphi_0^2 \rangle_I - \langle \varphi_0 \rangle_I^2 \neq 0$, are the ones with 0 as their left endpoint. Let $I_n = (0, 2^{-n}]$. Then

$$\langle \varphi_0 \rangle_{I_n} = 2^n \int_0^{1/2^n} \varphi_0(s)\, ds = 2^n \sum_{k=n-1}^\infty \frac{ka}{2^{k+2}} = \frac{a}{4} 2^n \left(\frac{1}{2}\right)^{n-2} n = an$$

and

$$\langle \varphi_0^2 \rangle_{I_n} = 2^n \int_0^{1/2^n} \varphi_0^2(s)\, ds = 2^n \sum_{k=n-1}^\infty \frac{k^2 a^2}{2^{k+2}} = \frac{a^2}{4} 2^n \left(\frac{1}{2}\right)^{n-2} (n^2 + 2) = a^2(n^2 + 2),$$

where we have used the identities

$$\sum_{k=N-1}^\infty k \left(\frac{1}{2}\right)^k = \left(\frac{1}{2}\right)^{N-2} N, \ \sum_{k=N-1}^\infty k^2 \left(\frac{1}{2}\right)^k = \left(\frac{1}{2}\right)^{N-2} (N^2 + 2).$$

Then

$$\|\varphi_0\|_{\mathrm{BMO}^d}^2 = \sup_{J\, \mathrm{dyad} \subset I} \left\{ \langle \varphi_0^2 \rangle_J - \langle \varphi_0 \rangle_J^2 \right\}$$

$$= \sup_n \left\{ \langle \varphi_0^2 \rangle_{I_n} - \langle \varphi_0 \rangle_{I_n}^2 \right\} = \sup_n \left\{ a^2(n^2 + 2) - a^2 n^2 \right\} = 2a^2.$$

Setting $\|\varphi_0\|_{\mathrm{BMO}^d} = \varepsilon$, we get $a = \varepsilon/\sqrt{2}$. Now,

$$\langle e^{\varphi_0} \rangle_I = \sum_{k=-1}^{\infty} \frac{e^{ka}}{2^{k+2}} = \sum_{k=-1}^{\infty} \frac{1}{4} \left(\frac{e^a}{2} \right)^k.$$

The latter sum converges if and only if $e^a < 2$, i.e. $a < \log 2$. In that case,

(5.2)
$$\langle e^{\varphi_0} \rangle_I = \frac{e^{-\varepsilon/\sqrt{2}}}{2 - e^{\varepsilon/\sqrt{2}}}.$$

Likewise,

$$\langle e^{-\varphi_0} \rangle_I = \frac{e^{\varepsilon/\sqrt{2}}}{2 - e^{-\varepsilon/\sqrt{2}}}$$

for arbitrary $\varepsilon > 0$.

Remark 5.1. We pause to observe that the calculation for $\langle e^{\varphi_0} \rangle$ has an immediate implication for the threshold ε_0^d from Theorem 1d. Namely,

$$\varepsilon_0^d \leq \sqrt{2} \log 2.$$

We now use φ_0 to construct the desired functions $\varphi_\pm$. Let

$$r_1 = \sqrt{\delta^2 - \varepsilon^2}, \ r_2 = \sqrt{\delta^2 - x_2 + x_1^2}, \ \beta = r_2 - r_1, \ \gamma = r_2 - \delta, \ \alpha = \frac{\delta - r_2}{\delta - r_1}.$$

Here δ will mean either δ^+ or δ^-, depending on the context. Define $\tilde{\varphi}_\pm$ on I by

$$\tilde{\varphi}_\pm(t) = x_1 \pm \tilde{\psi}(t), \ \text{where} \ \tilde{\psi}(t) = \begin{cases} \varphi_0\left(\frac{t}{\alpha}\right) + \beta & \text{for } 0 < t < \alpha, \\ \gamma & \text{for } \alpha < t < 1. \end{cases}$$

Observe that $\langle \varphi_0 \rangle_I = 0$, $\langle \varphi_0^2 \rangle_I = \varepsilon^2$. Since $\alpha = \gamma/(\gamma - \beta)$, we have $\langle \tilde{\psi} \rangle_I = \beta\alpha + \gamma(1 - \alpha) = 0$, and so $\langle \tilde{\varphi}_\pm \rangle_I = x_1$. Also, $\langle \tilde{\psi}^2 \rangle_{(0,\alpha)} = \langle \varphi_0^2 \rangle_I + 2\langle \varphi_0 \rangle_I \beta + \beta^2 = \varepsilon^2 + \beta^2$, and we get $\langle \tilde{\varphi}_\pm^2 \rangle_I = x_1^2 + (\varepsilon^2 + \beta^2)\alpha + \gamma^2(1 - \alpha) = x_2$. It remains to calculate $\langle e^{\tilde{\varphi}_\pm} \rangle_I$. In our notation, equations (3.6) and (3.7) can be rewritten (for the appropriate δ's) as

$$\frac{e^{-\varepsilon/\sqrt{2}}}{2 - e^{\varepsilon/\sqrt{2}}} = \frac{1 - r_1}{1 - \delta} e^{r_1 - \delta} = \left(1 + \frac{\delta - r_2}{\alpha(1 - \delta)} \right) e^{r_1 - \delta},$$

$$\frac{e^{\varepsilon/\sqrt{2}}}{2 - e^{-\varepsilon/\sqrt{2}}} = \frac{1 + r_1}{1 + \delta} e^{-r_1 + \delta} = \left(1 - \frac{\delta - r_2}{\alpha(1 + \delta)} \right) e^{-r_1 + \delta}.$$

Therefore, using (5.2) we get

$$\langle e^{\tilde{\varphi}_+} \rangle_I = e^{x_1} \langle e^{\tilde{\psi}} \rangle_I = \frac{e^{x_1 + \beta - \varepsilon/\sqrt{2}}}{2 - e^{\varepsilon/\sqrt{2}}} \alpha + e^{x_1 + \gamma}(1 - \alpha)$$

$$= \left(1 + \frac{\delta - r_2}{1 - \delta} \right) \exp(x_1 + r_2 - \delta) = \frac{1 - r_2}{1 - \delta} \exp(x_1 + r_2 - \delta) = B_{\delta^+(\varepsilon)}^+(x).$$

Similarly,

$$\langle e^{\tilde{\varphi}_-} \rangle_I = B_{\delta^-(\varepsilon)}^-(x).$$

We observe that $\tilde{\psi}$ (and so $\tilde{\varphi}_\pm$) does not, in general, belong to $\mathrm{BMO}_\varepsilon^d$, since the jumps in the scaled function φ_0 are not at dyadic nodes for an arbitrary α. We overcome this problem by constructing a rearrangement of $\tilde{\psi}$ that belongs to $\mathrm{BMO}_\varepsilon^d$, while preserving the necessary averages. Namely, let α_n be the n-th digit in the

dyadic representation of α (we will assume this representation is infinite, completing the sequence with zeros if needed). We define ψ as follows:

$$(5.3) \qquad \psi(t) = \sum_{k=1}^{\infty} \left\{ \alpha_k \left[\varphi_0(2^k t - 1) + \beta \right] + (1 - \alpha_k)\gamma \right\} \chi_{(2^{-k}, 2^{-k+1})}(t).$$

Naturally, we set

$$\varphi_{\pm} = x_1 \pm \psi.$$

Then, for any function μ,

$$\langle \mu \circ \psi \rangle_I = \sum_{k=1}^{\infty} \left\{ \alpha_k \langle \mu(\varphi_0(2^k t - 1) + \beta) \rangle_{(2^{-k}, 2^{-k+1})} + \mu(\gamma)(1 - \alpha_k) \right\} 2^{-k}$$

$$= \sum_{k=1}^{\infty} \left\{ \langle \mu(\varphi_0(t) + \beta) \rangle_I \alpha_k + \mu(\gamma)(1 - \alpha_k) \right\} 2^{-k}$$

$$= \langle \mu \circ (\varphi_0 + \beta) \rangle_I \alpha + \mu(\gamma)(1 - \alpha).$$

This calculation, with the appropriate choice of μ and the reasoning used above for $\tilde{\varphi}_{\pm}$, gives $\langle \varphi_{\pm} \rangle_I = x_1$, $\langle \varphi_{\pm}^2 \rangle_I = x_2$, $\langle e^{\varphi_{\pm}} \rangle_I = B_{\delta\pm}^{\pm}(x)$. It remains to check that $\|\psi\|_{\mathrm{BMO}^d(I)} = \varepsilon$. This will immediately imply that $\|\varphi_{\pm}\|_{\mathrm{BMO}^d(I)} = \varepsilon$.

Take any (open) dyadic interval $J \subset I$. We have the following trichotomy:

(1) $J \subseteq (2^{-n}, 2^{-n+1})$ for a certain n and $\alpha_n = 0$. Then $\psi|_J = \gamma$ and $\langle \psi^2 \rangle_J - \langle \psi \rangle_J^2 = 0$.

(2) $J \subseteq (2^{-n}, 2^{-n+1})$ for a certain n and $\alpha_n = 1$. Then $\psi(t) = \varphi_0(2^n t - 1) + \beta$, $\forall t \in J$, and $\langle \psi^2 \rangle_J - \langle \psi \rangle_J^2 \leq \varepsilon^2$ (see the detailed consideration for φ_0 above). Also, if $J = (2^{-n}, 2^{-n+1})$, then $\langle \psi^2 \rangle_J - \langle \psi \rangle_J^2 = \varepsilon^2$.

(3) $J = (0, 2^{-n})$ for a certain n. Then

$$\psi(t) = \sum_{k=n+1}^{\infty} \left\{ \alpha_k \left[\varphi_0(2^k t - 1) + \beta \right] + (1 - \alpha_k)\gamma \right\} \chi_{(2^{-k}, 2^{-k+1})}(t)$$

for almost every $t \in J$. So,

$$\langle \psi \rangle_J = \frac{1}{|J|} \sum_{k=n+1}^{\infty} \left\{ \alpha_k \langle \varphi_0 + \beta \rangle_I + \gamma(1 - \alpha_k) \right\} 2^{-k} = \beta p + \gamma(1 - p)$$

and

$$\langle \psi^2 \rangle_J = (\varepsilon^2 + \beta^2)p + \gamma^2(1 - p),$$

where $p = 2^n \sum_{k=n+1}^{\infty} \alpha_k 2^{-k}$. We have

$$\langle \psi^2 \rangle_J - \langle \psi \rangle_J^2 = p \left[\varepsilon^2 + (\beta - \gamma)^2(1 - p) \right] \stackrel{\text{def}}{=} \eta(p).$$

We maximize η, subject to the constraint $0 \leq p \leq 1$. Since

$$\eta'(p) = \varepsilon^2 + (\beta - \gamma)^2(1 - 2p) \geq \varepsilon^2 - (\beta - \gamma)^2$$

$$= \varepsilon^2 - \left(\delta - \sqrt{\delta^2 - \varepsilon^2} \right)^2 = 2\sqrt{\delta^2 - \varepsilon^2} \left(\delta - \sqrt{\delta^2 - \varepsilon^2} \right) \geq 0,$$

we have $\langle \varphi^2 \rangle_J - \langle \varphi \rangle_J^2 \leq \eta(1) = \varepsilon^2$.

This completes the proof of the lemma. $\qquad\qquad\qquad\qquad\qquad\qquad \square$

Lemma 2d. *For every $x \in \Omega_\varepsilon$,*

$$\text{(5.4)} \qquad \boldsymbol{B}_\varepsilon^{d+}(x) \leq B_{\delta^+(\varepsilon)}^+(x); \ \ \boldsymbol{B}_\varepsilon^{d-}(x) \geq B_{\delta^-(\varepsilon)}^-(x).$$

Proof. We follow the template of Lemma 2c. As in the continuous case, we have a concavity-type result, Lemma 3d, allowing us to use the induction on the order of the dyadic generation to construct an integral sum for $\langle e^\varphi \rangle_I$. Lemma 4c, the splitting lemma, cannot have a dyadic analog, since in the dyadic setting an interval is always split in half. This lack of splitting flexibility forces us to use a Bellman function candidate satisfying a stronger concavity (convexity) condition. Namely, the following two inequalities are true.

Lemma 3d.

$$\text{(5.5)} \qquad B_{\delta^+(\varepsilon)}^+ \left(\frac{1}{2}x^- + \frac{1}{2}x^+ \right) \geq \frac{1}{2}B_{\delta^+(\varepsilon)}^+(x^-) + \frac{1}{2}B_{\delta^+(\varepsilon)}^+(x^+),$$

$$\text{(5.6)} \qquad B_{\delta^-(\varepsilon)}^- \left(\frac{1}{2}x^- + \frac{1}{2}x^+ \right) \leq \frac{1}{2}B_{\delta^-(\varepsilon)}^-(x^-) + \frac{1}{2}B_{\delta^-(\varepsilon)}^-(x^+)$$

for any straight-line segment with the endpoints $x^\pm \in \Omega_\varepsilon$ such that $(x^- + x^+)/2 \in \Omega_\varepsilon$.

Assuming this lemma for the time being, take $\varphi \in \mathrm{BMO}_\varepsilon^d(I)$. Observe that $\varphi \in \mathrm{BMO}_\varepsilon^d(J)$ for any dyadic subinterval J of I. Let $I^{0,0} = I$ and let $I^{n,m}$ be the m-th interval of the n-th generation in the dyadic lattice based on I. Let $x^{n,m} = \left(\langle \varphi \rangle_{I^{n,m}}, \langle \varphi^2 \rangle_{I^{n,m}} \right)$. The argument of Lemma 2c now translates verbatim to the dyadic case. For the sake of completeness we repeat its major points. Using (5.5) from Lemma 3d repeatedly, we get

$$B_{\delta^+(\varepsilon)}^+(x^{0,0}) \ \geq \ \frac{1}{2}B_{\delta^+(\varepsilon)}^+(x^{1,0}) + \frac{1}{2}B_{\delta^+(\varepsilon)}^+(x^{1,1})$$

$$\text{(5.7)} \qquad \geq \ \frac{1}{4}B_{\delta^+(\varepsilon)}^+(x^{2,0}) + \frac{1}{4}B_{\delta^+(\varepsilon)}^+(x^{2,1}) + \frac{1}{4}B_{\delta^+(\varepsilon)}^+(x^{2,2}) + \frac{1}{4}B_{\delta^+(\varepsilon)}^+(x^{2,3})$$

$$\geq \ \frac{1}{2^n} \sum_{m=0}^{2^n-1} B_{\delta^+(\varepsilon)}^+(x^{n,m}) = \frac{1}{|I|} \int_I e^{\varphi_n(s)} b_+(s_n(s)) \, ds,$$

where φ_n and s_n are the same step functions that appeared in the proof of Lemma 2c: $\varphi_n(s) = x_1^{n,k}$ and $s_n(s) = x_2^{n,k} - (x_1^{n,k})^2$ for $s \in I^{n,k}$. The function b_+ has a meaning similar to that in the proof of Lemma 2c:

$$b_+(t) = \frac{H\left(\sqrt{\delta^+(\varepsilon)^2 - t} \, \right)}{H\left(\delta^+(\varepsilon) \right)}.$$

The last equality is just the statement $B_{\delta^+(\varepsilon)}^+(x^{n,k}) = e^{\varphi_n(s)} b_+(s_n(s))$, $s \in I^{n,k}$.

Likewise, applying (5.6) repeatedly, we obtain

$$B_{\delta^-(\varepsilon)}^-(x^{0,0}) \leq \frac{1}{|I^{0,0}|} \sum_{k=0}^{2^n-1} |I^{n,k}| B_{\delta^-(\varepsilon)}^-(x^{n,k}) = \frac{1}{|I|} \int_I e^{\varphi_n(s)} b_-(s_n(s)) \, ds.$$

Here

$$b_-(t) = \frac{H\left(-\sqrt{\delta^+(\varepsilon)^2 - t} \, \right)}{H\left(-\delta^+(\varepsilon) \right)}.$$

The technical convergence arguments of Lemma 2c completely carry over to the dyadic case (the quasi-Haar system in the proof of Lemma 2c now becomes the usual Haar system), and we obtain

$$B_{\delta^-(\varepsilon)}^-(\langle\psi\rangle_I, \langle\psi^2\rangle_I) \leq \frac{1}{|I|}\int_I e^{\psi(s)}\,ds \leq B_{\delta^+(\varepsilon)}^+(\langle\psi\rangle_I, \langle\psi^2\rangle_I).$$

Taking first supremum and then infimum over all $\psi \in \mathrm{BMO}_\varepsilon^d(I)$ with $\langle\psi\rangle_I = x_1$ and $\langle\psi^2\rangle_I = x_2$, we obtain the inequalities

$$B_{\delta^+(\varepsilon)}^+(x) \geq \boldsymbol{B}_\varepsilon^{d+}(x), \qquad B_{\delta^-(\varepsilon)}^-(x) \leq \boldsymbol{B}_\varepsilon^{d-}(x),$$

thus proving the lemma. $\qquad\square$

Proof of Lemma 3d. We prove the inequality in the most constructive manner: for every ε we will choose the smallest δ so that the statement of the lemma holds. From the proof of Lemma 1d, it is clear that $\delta(\varepsilon) > \varepsilon$.

One straightforward approach would be to choose $\delta(\varepsilon)$ large enough so that any straight-line segment $[x^-, x^+]$ with $x^-, x^+, x^0 \in \Omega_\varepsilon$ fits entirely inside $\Omega_{\delta(\varepsilon)}$. The statement of Lemma 3d would then follow from Lemma 3c. Let us investigate how large the $\delta(\varepsilon)$ so chosen would be with regard to ε. $\qquad\square$

Proposition 2. *If $\varepsilon \leq \frac{2\sqrt{2}}{3}\delta$, then the segment $[x^-, x^+]$ lies entirely in Ω_δ, for all $x^-, x^+ \in \Omega_\varepsilon$, such that $\frac{1}{2}x^- + \frac{1}{2}x^+ \in \Omega_\varepsilon$.*

Proof. We need to consider only those segments $[x^-, x^+]$ that have points outside Ω_ε, because otherwise $[x^-, x^+] \subset \Omega_\varepsilon \subset \Omega_\delta$. Fix such a segment and parameterize its points by

$$x(t) = (1-t)x^- + tx^+, \quad 0 \leq t \leq 1.$$

To prove the proposition, we need to check that for the function

$$\tau(t) = x_2(t) - x_1^2(t), \quad 0 \leq t \leq 1,$$

the inequality $\tau(t) \leq \delta^2$ holds.

Denote by a and b the points of intersection of $[x^-, x^+]$ with the upper boundary of Ω_ε, the parabola $x_2 = x_1^2 + \varepsilon^2$. Since $\frac{1}{2}x^- + \frac{1}{2}x^+ \in \Omega_\varepsilon$, the segment $[a, b]$ lies between this point and one of the endpoints $x^\pm$. Call this endpoint x^-. Since $\tau(t) \leq \varepsilon^2$ for $x(t) \in \Omega_\varepsilon$, we have

$$\max_{x(t)\in[x^-,x^+]} \tau(t) = \max_{x(t)\in[a,b]} \tau(t).$$

Therefore, instead of the initial segment $[x^-, x^+]$, it is sufficient to consider the shorter segment $[a, 2b - a]$. This means that without loss of generality we may assume the points x^- and $\frac{1}{2}x^- + \frac{1}{2}x^+$ to be on the upper boundary of Ω_ε, i.e.,

(5.8) $$x_2^- = (x_1^-)^2 + \varepsilon^2,$$

(5.9) $$\frac{1}{2}(x_2^- + x_2^+) = \frac{1}{4}(x_1^- + x_1^+)^2 + \varepsilon^2.$$

From (5.8) and (5.9) we get

$$x_2^+ = \frac{1}{2}\big((x_1^+)^2 - (x_1^-)^2\big) + x_1^- x_1^+ + \varepsilon^2.$$

Since $x^+ \in \Omega_\varepsilon$, we have the restriction $x_2^+ \geq (x_1^+)^2$, which is equivalent to the inequality

$$(5.10) \qquad (x_1^- - x_1^+)^2 \leq 2\varepsilon^2.$$

Now,

$$\begin{aligned} \tau(t) &= x_2(t) - x_1^2(t) \\ &= \left[(1-t)x_2^- + tx_2^+\right] - \left[(1-t)x_1^- + tx_1^+\right]^2 \\ &= \varepsilon^2 + \frac{1}{2}(x_1^- - x_1^+)^2(t - 2t^2). \end{aligned}$$

This function attains its maximum at $t = \frac{1}{4}$, so

$$\max \tau(t) = \varepsilon^2 + \frac{1}{16}(x_1^- - x_1^+)^2.$$

Taking into account inequality (5.10), we get

$$\max \tau(t) \leq \frac{9}{8}\varepsilon^2 \leq \delta^2.$$

This means that $[x^-, x^+] \subset \Omega_\delta$, as claimed. $\qquad \square$

Remark 5.2. We will make use of Proposition 2 throughout the proof of Lemma 3d; however, it has an immediate implication that is worth stating separately. Take $\varepsilon > 0$, $\delta \geq \frac{3}{2\sqrt{2}}\varepsilon$, and any two points $x^-, x^+ \in \Omega_\varepsilon$, such that $\frac{1}{2}x^- + \frac{1}{2}x^+ \in \Omega_\varepsilon$. Then the segment $[x^-, x^+] \subset \Omega_\delta$, and we can apply Lemma 3c to obtain

$$(5.11) \qquad \begin{aligned} B_\delta^+\left(\frac{1}{2}x^- + \frac{1}{2}x^+\right) &\geq \frac{1}{2}B_\delta^+(x^-) + \frac{1}{2}B_\delta^+(x^+), \\ B_\delta^-\left(\frac{1}{2}x^- + \frac{1}{2}x^+\right) &\leq \frac{1}{2}B_\delta^-(x^-) + \frac{1}{2}B_\delta^-(x^+). \end{aligned}$$

Thus, we can run the machine of Lemma 2d (with the restriction $\varepsilon < \frac{2\sqrt{2}}{3}$ for B^+) to establish that

$$B_{\frac{3}{2\sqrt{2}}\varepsilon}^+(x) \geq \boldsymbol{B}_\varepsilon^{d+}(x), \qquad B_{\frac{3}{2\sqrt{2}}\varepsilon}^-(x) \leq \boldsymbol{B}_\varepsilon^{d-}(x), \qquad \forall x \in \Omega_\varepsilon.$$

Together with Remark 5.1 in the proof of Lemma 1d, this gives us the two-sided estimate

$$(5.12) \qquad \frac{2\sqrt{2}}{3} \leq \varepsilon_0^d \leq \sqrt{2}\log 2.$$

At the end of the proof of Lemma 3d we will see that the correct value of ε_0^d is the right bound in (5.12).

So far, we have been trying to ensure that the segment $[x^-, x^+]$ lies inside the domain of concavity (convexity) of a certain function B_δ, so that we can then infer (5.11). Now, we try to enforce (5.11) directly instead.

Since we are searching for $\delta(\varepsilon)$ such that $\boldsymbol{B}_\varepsilon^d = B_{\delta(\varepsilon)}$, we attempt to solve the extremal problem

$$(5.13) \qquad \begin{aligned} \delta^\pm(\varepsilon) = \min_{\varepsilon < \delta < 1} \ &\left\{ \delta : \ \pm\left[B_\delta^\pm(x^0) - \frac{1}{2}\Big(B_\delta^\pm(x^-) + B_\delta^\pm(x^+)\Big)\right] \geq 0, \right. \\ &\left. \forall x^-, x^+ \in \Omega_\varepsilon \text{ such that } x^0 = \frac{1}{2}x^- + \frac{1}{2}x^+ \in \Omega_\varepsilon \right\}. \end{aligned}$$

Recalling representation (3.5),

$$B_\delta^\pm(x) = \frac{H\left(\pm\sqrt{\delta^2 + x_1^2 - x_2}\right)}{H(\pm\delta)}\, e^{x_1},$$

we are thus looking for the smallest possible δ such that, for every triple $x^\pm$, x^0 from Ω_ε, either inequality

(5.14)
$$2H\left(\sqrt{\delta^2 + (x_1^0)^2 - x_2^0}\right) - \exp\left(\frac{x_1^- - x_1^+}{2}\right) H\left(\sqrt{\delta^2 + (x_1^-)^2 - x_2^-}\right)$$
$$- \exp\left(\frac{x_1^+ - x_1^-}{2}\right) H\left(\sqrt{\delta^2 + (x_1^+)^2 - x_2^+}\right) \geq 0$$

or inequality
(5.15)
$$2H\left(-\sqrt{\delta^2 + (x_1^0)^2 - x_2^0}\right) - \exp\left(\frac{x_1^- - x_1^+}{2}\right) H\left(-\sqrt{\delta^2 + (x_1^-)^2 - x_2^-}\right)$$
$$- \exp\left(\frac{x_1^+ - x_1^-}{2}\right) H\left(-\sqrt{\delta^2 + (x_1^+)^2 - x_2^+}\right) \leq 0$$

holds, depending on whether we are considering the "+" or "−" in (5.13).

We can simplify this formulation by introducing the variables

$$z = x_2^0 - (x_1^0)^2,$$
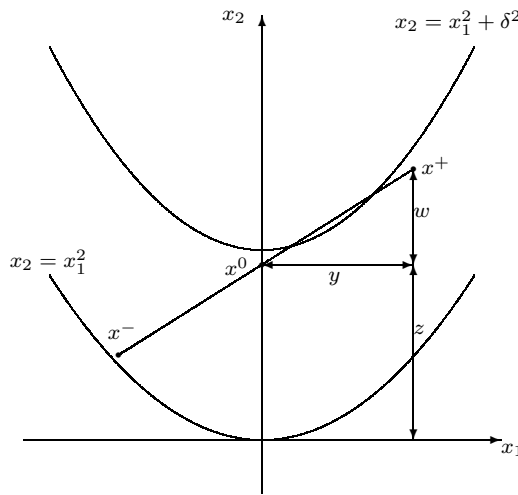$$y = x_1^+ - x_1^0,$$
$$w = x_2^+ - x_2^0 - 2x_1^0 y.$$

FIGURE 4. Geometrical meaning of w, y, and z for the case $x_1^0 = 0$.

Let

$$F^{\pm}(w,y,z) \overset{\text{def}}{=} 2H\left(\pm\sqrt{\delta^2 - z}\,\right) - e^{-y}H\left(\pm\sqrt{\delta^2 + y^2 - z + w}\,\right)$$
$$- e^y H\left(\pm\sqrt{\delta^2 + y^2 - z - w}\,\right).$$
(5.16)

In this notation, inequalities (5.14) and (5.15) become

$$(5.17) \qquad\qquad F^+(w,y,z) \geq 0 \quad\text{and}\quad F^-(w,y,z) \leq 0\,.$$

Because of the symmetry $F(w,y,z) = F(-w,-y,z)$, we can fix the sign of either w or y. Let us assume $w \geq 0$. Since H is strictly decreasing on the positive semi-axis and strictly increasing on the negative semi-axis (see Proposition 1), we have

$$H\left(\pm\sqrt{\delta^2 + y^2 - z + w}\,\right) \leq H\left(\pm\sqrt{\delta^2 + y^2 - z - w}\,\right),$$

and, therefore,

$$F^{\pm}(w,y,z) - F^{\pm}(w,-y,z)$$
$$(5.18) \qquad = (e^y - e^{-y})\left[H\left(\pm\sqrt{\delta^2 + y^2 - z + w}\,\right) - H\left(\pm\sqrt{\delta^2 + y^2 - z - w}\,\right)\right] \leq 0$$

for $y \geq 0$, i.e. each inequality $F^+(w,y,z) \geq 0$, $F^-(w,-y,z) \leq 0$ for nonnegative y implies the same inequality for all y. Therefore, it is sufficient to only consider $y \geq 0$. Note also that it is enough to consider only the special case $x^{\pm} = (\pm y, z \pm w)$, $x^0 = (0,z)$. Indeed, plugging these points into (5.14), (5.15), we get (5.17). Figure 4 gives the picture for such a line segment, for the case $w > 0$, $y > 0$. The domain where we need to investigate the function F^- is symmetrical to that of F^+: we have to consider the positively slanted segments $[x^-, x^+]$ for F^+ and negatively slanted ones for F^-. However, it is more convenient to consider the same domain in both cases and investigate $F^-(w,-y,z)$ rather than $F^-(w,y,z)$. Therefore, not only w but also y will be positive in what follows.

Now we determine the appropriate domain on which to investigate the functions $F^{\pm}$. The fact that the point x^- is above the parabola $x_2 = x_1^2$ is equivalent to the assertion that x^+ is under the parabola $x_2 = 2z - x_1^2$. We are only interested in those segments $[x^-, x^+]$ intersecting the parabola $x_2 = x_1^2 + \varepsilon^2$ (otherwise, as has been already pointed out, by Lemma 3c we know that $\pm F^{\pm}(w,y,z) \geq 0$). Therefore, the endpoint x^+ must be above the line $x_2 = 2x_1\sqrt{\delta^2 - z} + z$, passing through the point $x^0 = (0,z)$ and tangent to the parabola $x_2 = x_1^2 + \delta^2$, and to the right of the point of tangency. To have at least one such point in Ω_{ε}, the point $(\sqrt{z - \varepsilon^2/2}, z + \varepsilon^2/2)$ of intersection of the parabolas $x_2 = 2z - x_1^2$ and $x_2 = x_1^2 + \varepsilon^2$ must be above the tangent line $x_2 = 2x_1\sqrt{\delta^2 - z} + z$ (see Figure 5), which means

$$(5.19) \qquad\qquad z + \frac{\varepsilon^2}{2} \geq 2\sqrt{z - \varepsilon^2/2}\,\sqrt{\delta^2 - z} + z\,.$$

Solving this inequality for z, we get

$$z \geq \frac{2\delta^2 + \varepsilon^2 + 2\delta\sqrt{\delta^2 - \varepsilon^2}}{4}\,.$$

On the other hand, the point x^0 is in Ω_{ε}, hence $z \leq \varepsilon^2$. These two inequalities give a nonempty domain if and only if $\varepsilon^2 \geq \frac{8}{9}\delta^2$ (in accordance with Proposition 2). Therefore, the range of z that interests us is

$$(5.20) \qquad\qquad \frac{2\delta^2 + \varepsilon^2 + 2\delta\sqrt{\delta^2 - \varepsilon^2}}{4} \leq z \leq \varepsilon^2\,.$$
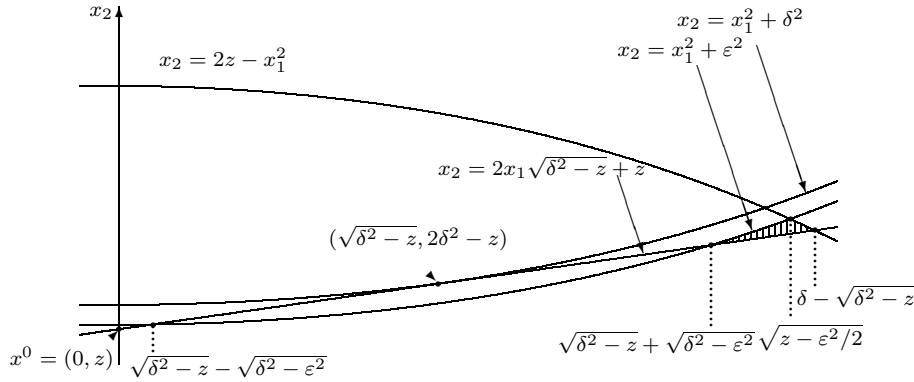
FIGURE 5. The triangle of the possible position of x^+ in which we are interested.

The tangent line $x_2 = 2x_1\sqrt{\delta^2 - z} + z$ intersects the parabola $x_2 = x_1^2 + \varepsilon^2$ at the points with $x_1 = \sqrt{\delta^2 - z} \pm \sqrt{\delta^2 - \varepsilon^2}$ and the parabola $x_2 = 2z - x_1^2$ at the point with $x_1 = \delta - \sqrt{\delta^2 - z}$. Since the point x^+ must be above the tangent line $x_2 = 2x_1\sqrt{\delta^2 - z} + z$ and under the parabolas $x_2 = x_1^2 + \varepsilon^2$ and $x_2 = 2z - x_1^2$, we get the restriction for the first coordinate of x^+ (recall that $x_1^+ = y$)

$$(5.21) \qquad \sqrt{\delta^2 - z} + \sqrt{\delta^2 - \varepsilon^2} \le y \le \delta - \sqrt{\delta^2 - z},$$

and for the second coordinate $(x_2^+ = z + w)$

$$2y\sqrt{\delta^2 - z} + z \le x_2 \le \min\{y^2 + \varepsilon^2, 2z - y^2\}.$$

Putting these together, we get the following range of w:
$$(5.22)$$
$$2y\sqrt{\delta^2 - z} \le w \le \begin{cases} y^2 - z + \varepsilon^2 & \text{for} & \sqrt{\delta^2 - z} - \sqrt{\delta^2 - \varepsilon^2} \le y \le \sqrt{z - \varepsilon^2/2} \\ z - y^2 & \text{for} & \sqrt{z - \varepsilon^2/2} \le y \le \delta - \sqrt{\delta^2 - z} \end{cases}.$$

We thus need to investigate the functions $F^\pm$ on the domain given by (5.20), (5.21), and (5.22).

Proposition 3. $F^+(w, y, z) \ge 0$ and $F^-(w, -y, z) \le 0$ on the domain (5.20), (5.21), and (5.22) if and only if these inequalities hold for all y satisfying (5.21), z satisfying (5.20), and $w = \min\{y^2 - z + \varepsilon^2, z - y^2\}$.

Proof. Using the identity

$$\frac{d}{dt}H(\pm\sqrt{t}) = -\frac{1}{2}e^{\pm\sqrt{t}},$$

we can easily calculate the partial derivative of $F^\pm$ with respect to w:

$$\frac{\partial F^\pm}{\partial w}(w, \pm y, z) = \frac{1}{2}\exp\{\mp y \pm \sqrt{\delta^2 + y^2 - z + w}\} - \frac{1}{2}\exp\{\pm y \pm \sqrt{\delta^2 + y^2 - z - w}\}.$$

To prove the proposition, it is sufficient to show that this derivative is negative for F^+ and positive for F^-. In both cases we arrive at the same inequality:

$$(5.23) \qquad -y + \sqrt{\delta^2 + y^2 - z + w} < y + \sqrt{\delta^2 + y^2 - z - w}\,.$$

We have the following chain of equivalent inequalities:

$$\sqrt{\delta^2 + y^2 - z + w} < \sqrt{\delta^2 + y^2 - z - w} + 2y\,,$$
$$w < 2y^2 + 2y\sqrt{\delta^2 + y^2 - z - w}\,,$$
$$\delta^2 - z < y^2 + 2y\sqrt{\delta^2 + y^2 - z - w} + (\delta^2 + y^2 - z - w)\,,$$
$$\sqrt{\delta^2 - z} < y + \sqrt{\delta^2 + y^2 - z - w}\,.$$

The last inequality is evidently true by (5.21). $\qquad\square$

Now, we shall investigate the functions

$$F_1^{\pm}(y, z) \stackrel{\text{def}}{=} F^{\pm}(\min\{y^2 - z + \varepsilon^2, z - y^2\}, \pm y, z)\,,$$

or, equivalently,

$$F_1^{\pm}(y,z) = \begin{cases} 2H(\pm\sqrt{\delta^2 - z}) - e^{\mp y}H(\pm\sqrt{\delta^2 + 2y^2 - 2z + \varepsilon^2}) - e^{\pm y}H(\pm\sqrt{\delta^2 - \varepsilon^2}) \\ \qquad\qquad \text{for } \sqrt{\delta^2 - z} + \sqrt{\delta^2 - \varepsilon^2} \le y \le \sqrt{z - \varepsilon^2/2} \\ \\ 2H(\pm\sqrt{\delta^2 - z}) - e^{\mp y}H(\pm\delta) - e^{\pm y}H(\pm\sqrt{\delta^2 + 2y^2 - 2z}) \\ \qquad\qquad \text{for } \sqrt{z - \varepsilon^2/2} \le y \le \delta - \sqrt{\delta^2 - z} \end{cases}\,.$$

Proposition 4. $F_1^+(y, z) \ge 0$ and $F_1^-(y, z) \le 0$ on the domain given by (5.20) and (5.21) if and only if these inequalities hold for all z satisfying (5.20) and $y = \sqrt{\delta^2 - \varepsilon^2/2}$.

Proof. Checking whether $\pm F_1^{\pm}(y, z) \ge 0$ is equivalent to checking whether $\pm e^{\mp y}F_1^{\pm}(y, z) \ge 0$ (or $\pm e^{\pm y}F_1^{\pm}(y, z) \ge 0$). However, the latter functions are much better behaved, and so we work with them instead. (We are grateful to the referee for suggesting this trick; it has significantly shortened our calculations.) Specifically, to prove the proposition we check that the functions $\pm e^{\mp y}F_1^{\pm}(y, z)$ are decreasing on $y \in \left(\sqrt{\delta^2 - z} + \sqrt{\delta^2 - \varepsilon^2}, \sqrt{z - \varepsilon^2/2}\right)$ and the functions $\pm e^{\pm y}F_1^{\pm}(y, z)$ are increasing on $y \in \left(\sqrt{z - \varepsilon^2/2}, \delta - \sqrt{\delta^2 - z}\right)$.

Since

$$\pm\frac{\partial}{\partial y}\left(e^{\mp y}F_1^{\pm}(y, z)\right) = -2e^{\mp y}H\left(\pm\sqrt{\delta^2 - z}\right) + 2e^{\mp 2y}H\left(\pm\sqrt{\delta^2 + 2y^2 - 2z + \varepsilon^2}\right)$$

$$\mp \exp\left\{\mp 2y \pm \sqrt{\delta^2 + 2y^2 - 2z + \varepsilon^2}\right\}\left(-\frac{1}{2}\right)4y$$

$$= 2e^{\mp y}\left[H\left(\pm\sqrt{\delta^2 + 2y^2 - 2z + \varepsilon^2} \mp y\right) - H\left(\pm\sqrt{\delta^2 - z}\right)\right]\,,$$

we will have proved that this expression is not positive if we check that

$$\sqrt{\delta^2 + 2y^2 - 2z + \varepsilon^2} - y \ge \sqrt{\delta^2 - z}\,.$$

Rewriting this inequality in its equivalent forms,

$$\sqrt{\delta^2 + 2y^2 - 2z + \varepsilon^2} \geq y + \sqrt{\delta^2 - z},$$
$$y^2 - z + \varepsilon^2 \geq 2y\sqrt{\delta^2 - z},$$
$$y^2 - 2y\sqrt{\delta^2 - z} + (\delta^2 - z) \geq \delta^2 - \varepsilon^2,$$

we see that the last inequality is true as the squared left inequality in (5.21).

On the second part of the interval, we have

$$\pm\frac{\partial}{\partial y}\Big(e^{\pm y}F_1^{\pm}(y,z)\Big) = 2e^{\pm y}H\big(\pm\sqrt{\delta^2 - z}\,\big) - 2e^{\pm 2y}H\big(\pm\sqrt{\pm\delta^2 + 2y^2 - 2z}\,\big)$$
$$\mp \exp\Big\{\pm 2y \pm \sqrt{\delta^2 + 2y^2 - 2z}\Big\}\Big(-\frac{1}{2}\Big)4y$$
$$= 2e^{\pm y}\Big[H\big(\pm\sqrt{\delta^2 - z}\,\big) - H\big(\pm\sqrt{\delta^2 + 2y^2 - 2z} \pm y\big)\Big],$$

and here the situation is even simpler: this expression is clearly positive because $y > \sqrt{\delta^2 - z}$. $\square$

Finally, we need to investigate the functions

$$F_2^{\pm}(z) \overset{\text{def}}{=} F_1^{\pm}(\sqrt{z - \varepsilon^2/2}\,, z)$$
$$= 2H\big(\pm\sqrt{\delta^2 - z}\,\big) - e^{\mp\sqrt{z - \varepsilon^2/2}}H(\pm\delta) - e^{\pm\sqrt{z - \varepsilon^2/2}}H\big(\pm\sqrt{\delta^2 - \varepsilon^2}\,\big).$$

Proposition 5. $F_2^{+}(z) \geq 0$ and $F_2^{-}(z) \leq 0$ on the interval (5.20) if and only if these inequalities hold at the right endpoint $z = \varepsilon^2$.

Proof. We use a trick similar to the one from the proof of Proposition 4. Namely, to prove the assertion, it is sufficient to check that the functions $\pm e^{\mp\sqrt{z - \varepsilon^2/2}}F_2^{\pm}(z)$ are strictly decreasing, i.e. to check that the expressions

$$\pm\frac{d}{dz}\Big[e^{\mp\sqrt{z - \varepsilon^2/2}}F_2^{\pm}(z)\Big]$$
$$= -\frac{1}{\sqrt{z - \varepsilon^2/2}}e^{\mp\sqrt{z - \varepsilon^2/2}}H(\pm\sqrt{\delta^2 - z}\,) \pm e^{\mp\sqrt{z - \varepsilon^2/2}\pm\sqrt{\delta^2 - z}}$$
$$+ \frac{1}{\sqrt{z - \varepsilon^2/2}}e^{\mp 2\sqrt{z - \varepsilon^2/2}}H(\pm\delta)$$
$$= \frac{e^{\mp 2\sqrt{z - \varepsilon^2/2}}}{\sqrt{z - \varepsilon^2/2}}\Big[H(\pm\delta) - H\big(\pm\sqrt{\delta^2 - z} \pm \sqrt{z - \varepsilon^2/2}\,\big)\Big]$$

are not positive. This is the same as checking that $\delta \geq \sqrt{\delta^2 - z} + \sqrt{z - \varepsilon^2/2}$. Squaring both sides, we get

$$\frac{\varepsilon^2}{2} \geq 2\sqrt{z - \varepsilon^2/2}\,\sqrt{\delta^2 - z},$$

which is the same as (5.19) and so is true by (5.20), as we have already seen. $\square$

Define two functions of δ and ε by putting

$$g^{\pm}(\delta, \varepsilon) \overset{\text{def}}{=} F_2^{\pm}(\varepsilon^2)$$

or, equivalently,

$$g^{\pm}(\delta, \varepsilon) = \left(2 - e^{\pm \varepsilon/\sqrt{2}}\right) H(\pm\sqrt{\delta^2 - \varepsilon^2}) - e^{\mp \varepsilon/\sqrt{2}} H(\pm\delta)$$

$$= (1 \mp \sqrt{\delta^2 - \varepsilon^2}) e^{\pm\sqrt{\delta^2 - \varepsilon^2}} \left(2 - e^{\pm \varepsilon/\sqrt{2}}\right) - (1 \mp \delta) e^{\pm\delta \mp \varepsilon/\sqrt{2}}.$$

Propositions 3–5 can now be summarized as follows:

$$B_{\delta}^{+}\left(\frac{1}{2}x^- + \frac{1}{2}x^+\right) \geq \frac{1}{2}B_{\delta}^{+}(x^-) + \frac{1}{2}B_{\delta}^{+}(x^+) \Longleftrightarrow g^+(\delta, \varepsilon) \geq 0,$$

$$B_{\delta}^{-}\left(\frac{1}{2}x^- + \frac{1}{2}x^+\right) \leq \frac{1}{2}B_{\delta}^{-}(x^-) + \frac{1}{2}B_{\delta}^{-}(x^+) \Longleftrightarrow g^-(\delta, \varepsilon) \leq 0,$$

where the inequalities on the left are meant to hold for all $x^-, x^+ \in \Omega_{\varepsilon}$ such that $\frac{1}{2}(x^- + x^+) \in \Omega_{\varepsilon}$.

Accordingly, to find the Bellman functions we are looking for the smallest such δ for a given ε, i.e.,

$$(5.24) \qquad \delta^+(\varepsilon) = \min_{\varepsilon < \delta < 1}\{\delta : g^+(\delta, \varepsilon) \geq 0\},$$

$$(5.25) \qquad \delta^-(\varepsilon) = \min_{\varepsilon < \delta}\{\delta : g^-(\delta, \varepsilon) \leq 0\}.$$

The following simple result will complete the proof of Lemma 3d.

Proposition 6. *For any ε, $0 \leq \varepsilon < \sqrt{2}\log 2$, the equation $g^+(\delta, \varepsilon) = 0$ has a unique solution in the interval $(\varepsilon, 1)$ and it is $\delta^+(\varepsilon)$ from (5.24).*

For any $\varepsilon \geq 0$ the equation $g^-(\delta, \varepsilon) = 0$ has a unique solution in the interval $\left(\varepsilon, \frac{3}{2\sqrt{2}}\varepsilon\right)$ and it is $\delta^-(\varepsilon)$ from (5.25).

Proof. At the left endpoint, $\delta = \varepsilon$, we have

$$g^{\pm}(\varepsilon, \varepsilon) = 2 - e^{\pm \varepsilon/\sqrt{2}} - (1 \mp \varepsilon)e^{\pm\varepsilon \mp \varepsilon/\sqrt{2}}$$

$$= e^{\pm(1 - \frac{1}{\sqrt{2}})\varepsilon}\left(2e^{\mp(1 - \frac{1}{\sqrt{2}})\varepsilon} - e^{\pm(\sqrt{2} - 1)\varepsilon} - 1 \pm \varepsilon\right) \stackrel{\text{def}}{=} e^{\pm(1 - \frac{1}{\sqrt{2}})\varepsilon}h(\pm\varepsilon).$$

The function h satisfies $h(0) = h'(0) = 0$, $h''(s) = (\sqrt{2} - 1)^2(e^{(\frac{1}{\sqrt{2}} - 1)s} - e^{(\sqrt{2} - 1)s})$, and $\operatorname{sgn} h''(s) = -\operatorname{sgn} s$. Therefore $\operatorname{sgn} h(s) = -\operatorname{sgn} s$, i.e. $g^+(\varepsilon, \varepsilon) < 0$ and $g^-(\varepsilon, \varepsilon) > 0$.

At the point $\delta = 1$, we have

$$g^+(1, \varepsilon) = H\left(\sqrt{1 - \varepsilon^2}\right)\left(2 - e^{\varepsilon/\sqrt{2}}\right).$$

Since $\varepsilon < \sqrt{2}\log 2$, we conclude that $g^+(1, \varepsilon) > 0$. Therefore, the equation $g^+(\delta, \varepsilon) = 0$ has a solution in the interval $(\varepsilon, 1)$.

At the point $\delta = \frac{3}{2\sqrt{2}}\varepsilon$, we have

$$g^-\left(\frac{3}{2\sqrt{2}}\varepsilon, \varepsilon\right) = \left(1 - \frac{\varepsilon}{2\sqrt{2}}\right)e^{-\frac{\varepsilon}{2\sqrt{2}}} - \left(1 + \frac{\varepsilon}{2\sqrt{2}}\right)e^{-\frac{3\varepsilon}{2\sqrt{2}}}$$

$$= e^{-\frac{\varepsilon}{\sqrt{2}}}\left[H\left(\frac{\varepsilon}{2\sqrt{2}}\right) - H\left(-\frac{\varepsilon}{2\sqrt{2}}\right)\right] < 0$$

by statement (b) of Proposition 1. This proves the existence of a root of g^- in the interval $\left(\varepsilon, \frac{3}{2\sqrt{2}}\varepsilon\right)$.

To check the uniqueness of these roots, we differentiate $g^{\pm}$ with respect to δ:

$$\frac{\partial g^{\pm}}{\partial \delta}(\delta, \varepsilon) = \delta \left[e^{\pm(\delta - \varepsilon/\sqrt{2})} - e^{\pm\sqrt{\delta^2 - \varepsilon^2}} \left(2 - e^{\pm\varepsilon/\sqrt{2}} \right) \right] .$$

Since $\delta > \sqrt{\delta^2 - \varepsilon^2}$, we have

$$\frac{\partial g^+}{\partial \delta}(\delta, \varepsilon) > \delta e^{\pm\sqrt{\delta^2 - \varepsilon^2}} \left[e^{-\varepsilon/\sqrt{2}} - 2 + e^{\varepsilon/\sqrt{2}} \right] > 0 ,$$

i.e. the function g^+ is increasing in δ and the root of the equation $g^+(\delta, \varepsilon) = 0$ is unique and satisfies (5.24).

Since $-\delta + \varepsilon/\sqrt{2} < -\sqrt{\delta^2 - \varepsilon^2}$ for $\varepsilon < \delta < \frac{3}{2\sqrt{2}}\varepsilon$, we have

$$\frac{\partial g^-}{\partial \delta}(\delta, \varepsilon) \le \delta e^{-\sqrt{\delta^2 - \varepsilon^2}} \left(-1 + e^{-\varepsilon/\sqrt{2}} \right) < 0 ,$$

i.e. the function g^- is decreasing in δ on $(\varepsilon, \frac{3\varepsilon}{2\sqrt{2}})$ and the root of the equation $g^+(\delta, \varepsilon) = 0$ is unique in this interval and satisfies (5.25). $\qquad\square$

The proof of Lemma 3d is now complete. $\qquad\square$

5.1. **How to find the dyadic Bellman function.** For simplicity, we only consider the case of $\boldsymbol{B}^d = \boldsymbol{B}^{d+}$. What prompted us to look for the dyadic Bellman function in the family B_δ from (3.5)? First, this family was first developed when solving the formal optimal control problem from [7, 16], where the space under consideration was the dyadic BMO. Second, and more important, the following simple proposition shows that the dyadic Bellman function is locally concave, something that could not be shown directly in the continuous case.

Proposition 7. *For any three points* $x^-, x^+, x \in \Omega_\varepsilon$ *such that* $x = \frac{1}{2}(x^- + x^+)$ *we have*

$$\boldsymbol{B}_\varepsilon^d(x) \ge \frac{1}{2}\boldsymbol{B}_\varepsilon^d(x^-) + \frac{1}{2}\boldsymbol{B}_\varepsilon^d(x^+).$$

Proof. Take a sequence $\{\varphi_n\} \in \mathrm{BMO}_\varepsilon^d(I_-) \cup \mathrm{BMO}_\varepsilon^d(I_+)$ such that

$$\langle e^{\varphi_n} \rangle_{I_\pm} \longrightarrow \boldsymbol{B}_\varepsilon^d(x^\pm) \text{ as } n \to \infty.$$

We need to check that $\varphi_n \in \mathrm{BMO}_\varepsilon^d(I)$. However,

$$\mathrm{BMO}_\varepsilon^d(I) = \left\{ \varphi : \ \varphi|_{I_-} \in \mathrm{BMO}_\varepsilon^d(I_-), \ \varphi|_{I_+} \in \mathrm{BMO}_\varepsilon^d(I_+), \ \langle\varphi^2\rangle_I - \langle\varphi\rangle_I^2 \le \varepsilon^2 \right\}.$$

Since, by assumption, $x \in \Omega_\varepsilon$, we have $\langle\varphi^2\rangle_I - \langle\varphi\rangle_I^2 \le \varepsilon^2$. Then we can pass to the limit in the identity

$$\langle e^{\varphi_n} \rangle_I = \frac{1}{2}\langle e^{\varphi_n} \rangle_{I_-} + \frac{1}{2}\langle e^{\varphi_n} \rangle_{I_+}$$

to get

$$\boldsymbol{B}_\varepsilon^d(x) \ge \lim\langle e^{\varphi_n} \rangle_I = \frac{1}{2}\boldsymbol{B}_\varepsilon^d(x^-) + \frac{1}{2}\boldsymbol{B}_\varepsilon^d(x^+),$$

which completes the proof. $\qquad\square$

Observe that the statement of the proposition does not hold in the continuous case. In that case, we have

$$\mathrm{BMO}_\varepsilon(I) \ne \left\{ \varphi : \ \varphi|_{I_-} \in \mathrm{BMO}_\varepsilon(I_-), \ \varphi|_{I_+} \in \mathrm{BMO}_\varepsilon(I_+), \ \langle\varphi^2\rangle_I - \langle\varphi\rangle_I^2 \le \varepsilon^2 \right\},$$

since there are other intervals to consider: those with the left endpoint in I_- and the right one in I_+.

We have just proved that $\boldsymbol{B}_\varepsilon^d$ is locally concave in Ω_ε. Furthermore, the reasoning of (4.9) still works, and we conclude that

$$\boldsymbol{B}_\varepsilon^d(x) = \exp\left\{x_1 + w(x_2 - x_1^2)\right\}$$

for a nonnegative function w such that $w(0) = 0$. What is more, we expect the corresponding matrix $-d^2\boldsymbol{B}_\varepsilon^d$ (assuming sufficient smoothness) to be degenerate in order for the supremum to be attained for an extremal function. But we have already described all functions with these properties. They are the functions B_δ from (3.5). The condition $\delta \geq \varepsilon$ appears because the function $\boldsymbol{B}_\varepsilon^d$ has to be defined on Ω_ε, $\Omega_\varepsilon \subset \Omega_\delta$ for $\delta \geq \varepsilon$, and Ω_δ is exactly the domain of B_δ. Thus we look for $\boldsymbol{B}^d$ within that family.

5.2. How to find the dyadic extremal function.

Again, we consider only the "+" case. Recall that in the continuous case we were looking for a function that would produce equality on every step of (5.7), i.e. in the Bellman induction of Lemma 2d. Thus, such a function was found by analyzing what it took to make B_δ behave as a linear function, that is to have

$$B_\delta(\alpha_- x^- + \alpha_+ x^+) = \alpha_- B_\delta(x^-) + \alpha_+ B_\delta(x^+).$$

We now employ similar reasoning. Namely, we construct the dyadic extremal function for a point on the top boundary so that we have equality in Lemma 3d, i.e.

$$(5.26) \qquad B_{\delta(\varepsilon)}\left(\frac{1}{2}x^- + \frac{1}{2}x^+\right) = \frac{1}{2}B_{\delta(\varepsilon)}(x^-) + \frac{1}{2}B_{\delta(\varepsilon)}(x^+)$$

at every dyadic split $I = I_- \cup I_+$. We construct a function φ_0 on $I = [0,1]$ for the point $x = (0, \varepsilon^2)$. Then the function φ_a, $\varphi_a(t) = \varphi_0(t) + a$, is an extremal function for the point $(a, a^2 + \varepsilon^2)$. The proof of Lemma 3d gives us a hint for our construction: the extremum in (5.13) was realized by a line segment whose center and one of the endpoints (say x^-) lay on the top boundary curve of Ω_ε, $x_2 = x_1^2 + \varepsilon^2$, i.e. $x = (0, \varepsilon^2)$ and $x^- = (a, a^2 + \varepsilon^2)$, while the other endpoint, x^+, lay on the bottom boundary curve $x_2 = x_1^2$, i.e. $x^+ = (-a, a^2)$. From the condition $x = \frac{1}{2}x_- + \frac{1}{2}x_+$ we get $a = \varepsilon/\sqrt{2}$. Only constant functions correspond to the points of the bottom boundary, so we have to put $\varphi_0(t) = x_1^+ = -a$ for $\frac{1}{2} < t < 1$, and on I^- we have to take the scaled function φ_a: $\varphi_0(t) = \varphi_a(2t) = \varphi_0(2t) + a$ for $0 < t < \frac{1}{2}$. This recursive relation yields $\varphi_0(t) = (n-1)a$ for $2^{-n-1} < t < 2^{-n}$. This is the function on Figure 3.

We now describe how to construct an extremal function φ when $(x_1, x_2) \neq (0, \varepsilon^2)$. If $x_2 = x_1^2 + \varepsilon^2$, i.e. x is on the top boundary, we simply let $\varphi = \varphi_0 + x_1$ to get the desired result. Likewise, if x is on the bottom boundary, we let $\varphi = x_1$, i.e. set the function to be constant on the whole interval. What should we do if x is in the interior of Ω_ε? We present two different perspectives on how this situation can be dealt with. Both lead to the same expression for the extremal function φ.

Perspective 1. Let us forget for a moment that we are to construct a *dyadic* extremal function; then we can split I so that x^+ is on the bottom boundary and x^- is on the top one. Let α be the splitting parameter, i.e. we have $I_- = (0, \alpha)$,

$I_+ = (\alpha, 1)$, and $x = \alpha x^- + (1 - \alpha)x^+$. We would like to choose the splitting so that

$$B_{\delta(\varepsilon)}(x) = \alpha B_{\delta(\varepsilon)}(x^-) + (1 - \alpha)B_{\delta(\varepsilon)}(x^+).$$

Then we can set φ to be constant on the right subinterval and the appropriately scaled function φ_0 on the left one, and apply (5.7) from Lemma 3d to I_- and I_+ separately. To do this, we place x^-, x, and x^+ on a line ω_δ^+ tangent to the curve $x_2 = x_1^2 + \delta^2$, since, according to section 4.2, B_δ^+ is a linear function along any such segment. More precisely, we consider the line through x that is tangent to $x_2 = x_1^2 + \delta^2$, and we set x^- to be the point of intersection of the line and the curve $x_2 = x_1^2 + \varepsilon^2$ and x^+ to be the point of intersection of the line and the curve $x_2 = x_1^2$. Let us calculate α. To avoid confusion, we will temporarily use x^0 when referring to the "midpoint" of our segment. Let us recall the notation of Lemma 1d:
(5.27)
$$r_1 = \sqrt{\delta^2 - \varepsilon^2}, \ r_2 = \sqrt{\delta^2 - x_2^0 + (x_1^0)^2}, \ \beta = r_2 - r_1, \ \gamma = r_2 - \delta, \ \alpha = \frac{\delta - r_2}{\delta - r_1}.$$

Also let

(5.28) $$\beta_1 = \beta + x_1^0; \ \gamma_1 = \gamma + x_1^0.$$

According to (4.18), the line $\omega_\delta^+(c)$ tangent to $x_2 = x_1^2 + \delta^2$ at the point $(c, c^2 + \delta^2)$ has the equation

(5.29) $$x_2 = 2cx_1 + \delta^2 - c^2.$$

We calculate c using the fact that this line passes through x^0. Since in our geometry $c \geq x_1^0$, we have $c = x_1^0 + r_2$. Then (5.29) becomes

$$x_2 = 2(x_1^0 + r_2)x_1 + \delta^2 - (x_1^0 + r_2)^2$$

or, equivalently,

$$(x_1 - (x_1^0 + r_2))^2 = \delta^2 + x_1^2 - x_2.$$

This line intersects the top boundary curve at the point $x^- = (\beta_1, \beta_1^2 + \varepsilon^2)$ (where we have used the fact that $x_1^0 \leq x_1^- \leq c$); the intersection with the bottom curve is at $x^+ = (\gamma_1, \gamma_1^2)$. The (horizontal) length of the segment $[x^+, x^-]$ is $\delta - r_1$, while that of the segment $[x^+, x^0]$ is $\delta - r_2$; so we get $x^0 = \alpha x^- + (1 - \alpha)x^+$. Putting everything together, we obtain the function $\tilde{\varphi}_+$ from the proof of Lemma 1d:

$$\tilde{\varphi}_+(t) = x_1^0 + \begin{cases} \varphi_0\left(\frac{t}{\alpha}\right) + \beta & \text{for } 0 < t < \alpha, \\ \gamma & \text{for } \alpha < t < 1. \end{cases}$$

We must pay the price for ignoring the fact that $(0, \alpha)$ is not, in general, a dyadic interval and, therefore, $\tilde{\varphi}_+$ is not in $\text{BMO}_\varepsilon^d(I)$. How to construct an appropriate rearrangement φ_+ of $\tilde{\varphi}_+$ is detailed in the proof of Lemma 1d.

Perspective 2. It is useful to consider another perspective on constructing an extremal function. We will start with the function φ_0 built for the point $x = (0, \varepsilon^2)$ and arrive at the same function φ_+, for an arbitrary point x^0, as in Lemma 1d, but using a different reasoning and skipping the $\tilde{\varphi}_+$ phase altogether. The main feature of this construction is that on every step we define our function on a dyadic subinterval of $(0, 1)$, as opposed to choosing an α and then approximating it dyadically as in Perspective 1.

Here is the simple logic: starting with $I = (0, 1)$, we will define our function on the right half of I, then redefine I to be the other half, and then repeat the

procedure. Consider, as before, the line through x^0 tangent to $x_2 = x_1^2 + \delta^2$. Let x^t and x^b be the points of intersection of the tangent with the top and bottom boundary of Ω_ε, respectively. If x^0 is closer to x^t than to x^b, set φ to be the appropriately scaled (and adjusted to have the prescribed average) function φ_0 on I_+ and replace x_0 with $2x_0 - x^t$. If, on the contrary, x^0 is closer to x^b than to x^t, set φ to be the appropriately chosen constant on I_+ and replace x_0 with $2x_0 - x^b$. In either case, replace I with I_- and repeat. If x^0 is exactly in the middle between x^b and x^t, let φ be the scaled φ_0 on I_+ and constant on I_-; then stop.

We will now make this procedure more precise and show why the function so obtained is the same as the one used to prove Lemma 1d.

Start with a point $x^0 \in \Omega_\varepsilon$. Let $x^* = x^0$, $I = (0,1)$ (the initial settings; x^* and I will be redefined in the procedure). Let r_1, β_1, and γ_1 be defined by (5.27) and (5.28) (these will not be redefined). Then:

1. Let $r_2 = \sqrt{\delta^2 - x_2^* + (x_1^*)^2}$
   – if $\delta + r_1 < 2r_2$, go to step 2;
   – if $\delta + r_1 > 2r_2$, go to step 4;
   – if $\delta + r_1 = 2r_2$, go to step 6.
2. Let $\varphi|_{I_+} = \gamma_1$.
3. Let $x^b = (\gamma_1, \gamma_1^2)$, $x^* := 2x^* - x^b$, $I := I_-$. Go to step 1.
4. Let $\varphi|_{I_+} = \varphi_0(2^k t + 1) + \beta_1$.
5. Let $x^t = (\beta_1, \beta_1^2 + \varepsilon^2)$, $x^* := 2x^* - x^t$, $I := I_-$. Go to step 1.
6. Let $\varphi|_{I_+} = \varphi_0(2^k t + 1) + \beta_1$, $\varphi|_{I_-} = \gamma_1$. Stop.

Since on every run of the loop we define φ on half of the current interval I and then rename the other half I, in the end we have defined φ almost everywhere on $(0,1)$. Furthermore, since every interval in the process is dyadic and $\varphi \in \mathrm{BMO}_\varepsilon^d(J)$ for every interval J that turns up on step 2, 4, or 6, we conclude that $\varphi \in \mathrm{BMO}_\varepsilon^d([0,1])$. All the action happens on the same line tangent to the parabola $x_2 = x_1^2 + \delta^2$, guaranteeing equality in (5.7) of Lemma 3d.

The inequality $\delta + r_1 < 2r_2$ (or $>, =$) is equivalent to the inequality $\delta - r_2 < r_2 - r_1$ (or $>, =$), i.e. the statement that the distance from x^0 to the bottom boundary curve is smaller than that to the top one. Alternatively, this inequality is equivalent to $\frac{\delta - r_2}{\delta - r_1} < \frac{1}{2}$, i.e., in the language of Perspective 1, $\alpha < \frac{1}{2}$. But comparing the current α to $1/2$ is the same as determining whether the current dyadic digit of the original α is 0 or 1. Indeed, if the current x^* is closer to the top boundary, its next iteration will be twice as far from it; the same holds for the bottom boundary. Let us quantify this.

Let $z_0 = \alpha$, $z_k = \frac{\delta - (r_2)_k}{\delta - r_1}$, the value on the k-th step of our procedure. By construction, if $z_{k-1} > 1/2$, then $z_k = 2z_{k-1} - 1$, and if $z_{k-1} < 1/2$, then $z_k = 2z_{k-1}$. Thus $z_k = \{2z_{k-1}\}$, the fractional part of $2z_{k-1}$. Then $\alpha_k \stackrel{\text{def}}{=} [2z_{k-1}]$ (the integer part) is the k-th dyadic digit of α. Recalling definition (5.3), we see that the function φ so obtained is indeed the same as φ_+ in Perspective 1.

6. Conclusion

In this section, we summarize what has been achieved, specify which obstacles need to be overcome on the way to generalizing the results, and outline immediate and long-term prospects.

From a purely practical viewpoint, we have obtained sharp new results in a widely-used inequality. In addition, the dyadic BMO formulation is common in applications, therefore exploring the problem in this setting — and showing that the results differ significantly from the continuous setting — is important.

Equally important is the methodological aspect of this work. We have added another nontrivial example to the short list of explicit Bellman functions. This application of the method can be viewed as an excellent case study, following every step in the recent explicit Bellman template. As far as we know, our transition to the dyadic case from the continuous one is unique in literature; as mentioned in the introduction, the usual way is the opposite. The dyadic setting has been prevalent in Bellman function studies, thus our obtaining of an explicit continuous-case Bellman function is noteworthy in itself.

There are several natural questions one may ask:

1. Can the results be extended to the L^p-based BMO?

The choice of variables (2.1) (and so the associated Bellman function definitions) depends heavily on the L^2-structure of our BMO. For $p > 1$ it is possible to consider the choice $x_2 = \langle |\varphi|^p \rangle$, although the associated norms are not the usual L^p-based BMO norms. It appears that an altogether different Bellman setup may be needed for the L^1 case.

2. Can the results be extended to higher dimensions?

Once we move to higher dimensions, there is the question of how one defines BMO. Typical definitions use cubes or balls, although others are possible. Since our technique depends critically on one's ability to split a body in $\mathbb{R}^n$ into bodies of the same type, it seems that the dyadic case is more amenable to higher-dimensional considerations because in the dyadic situation we have no problem splitting a cube into a union of smaller cubes. In the continuous case, however, the crucial splitting tool we have used, Lemma 4c, is pointedly one-dimensional. We could easily generalize our results to the n-parameter BMO on rectangles, but this appears to be of little interest.

Often in Bellman proofs one relies on a certain dyadic Bellman function to handle all dimensions. Naturally, our continuous-to-dyadic way of solving the problem does not go through in that sense. In addition, the continuous and dyadic results are expected to be increasingly different as dimension grows. Overall, new techniques are needed (work is underway) to deal with the higher-dimensional case.

Despite our present inability to handle the multi-dimensional case, we would like to put forth two related conjectures for the BMO defined on cubes.

Conjecture 1. *Theorems* 1c *and* 2c *remain true in the multi-dimensional case, i.e. in the nondyadic case the Bellman function does not depend on the dimension.*

Conjecture 2. *In the dyadic n-dimensional case the Bellman functions are $B_{\delta_n^\pm}^\pm$, where the parameters $\delta_n^\pm = \delta_n^\pm(\varepsilon)$ are the solutions of the equations*

$$(1 \mp \sqrt{\delta^2 - \varepsilon^2}) \exp(\pm\sqrt{\delta^2 - \varepsilon^2} \mp \delta) \left(2^n - e^{\pm(2^{n/2} - 2^{-n/2})\varepsilon}\right) = (1 \mp \delta)(2^n - 1)e^{\mp\varepsilon 2^{-n/2}},$$

and, therefore, the corresponding constants $C_n^d(\varepsilon)$ and $\varepsilon_0^d(n)$ are

$$C_n^d(\varepsilon) = \frac{(2^n - 1)e^{-\varepsilon 2^{-n/2}}}{2^n - e^{(2^{n/2} - 2^{-n/2})\varepsilon}}$$

and

$$\varepsilon_0^d(n) = \frac{n \log 2}{2^{n/2} - 2^{-n/2}}.$$

These conjectures are true if it is true that the extremal function corresponding to the point $(0, \varepsilon^2)$ *is*

$$\varphi_0(t_1, \ldots, t_n) = \varepsilon \left(n \log \frac{1}{\max t_k} - 1 \right)$$

in the nondyadic case and

$$\varphi_0^d(t_1, \ldots, t_n) = -\varepsilon 2^{-n/2} + \sum_{k=1}^{\infty} (2^{n/2} - 2^{-n/2}) \varepsilon \chi_{_{[0, 2^{-k}]}}(\max t_k)$$

in the dyadic one.

3. Can the classical weak-form John–Nirenberg inequality be handled by the methods of this paper?

At the moment, this appears to be the most promising of all directions of further research on the topic. By design, the Bellman function for a distributional inequality will have one more variable (at least, another parameter), compared to the integral case. This implies that the order of the Bellman PDE in the weak-form case will be higher.

On the other hand, we have a ready choice of variables by just reusing the ones from this paper. The usual logic that allows one to establish a finite-difference inequality for the Bellman function still applies. In [12], a Bellman-type function satisfying this inequality (a so-called supersolution) was found for the dyadic BMO. This showed that the Bellman function method works for the weak form of the John–Nirenberg inequality. However, not being the true Bellman function, that supersolution only gave suboptimal (not sharp) constants in the inequality. It is our hope to be able to find the true Bellman function for this inequality as well. Being the averages of functions, our variables have a clear martingale structure, thus we expect to be able to rewrite that inequality as a homogeneous Monge-Ampère equation, just as we have done here. Though that equation will not reduce to an ODE, there has been a recent surge (and success) in in-depth studies of the connection of such PDEs with the Bellman function method. All of this gives the problem a very promising outlook.

Acknowledgment

References

[1] D. L. Burkholder. Boundary value problems and sharp inequalities for martingale transforms. *Annals of Probability*, Vol. 12 (1984), no. 3, pp. 647–702. MR744226 (86b:60080)

[2] C. Fefferman. Characterizations of bounded mean oscillations. *Bull. Amer. Math. Soc.*, Vol. 77 (1971), pp. 587–588. MR0280994 (43:6713)

[3] F. John, L. Nirenberg. On functions of bounded mean oscillation. *Comm. Pure Appl. Math.*, Vol. 14 (1961), pp. 415–426. MR0131498 (24:A1348)

[4] A. Korenovskii. The connection between mean oscillations and exact exponents of summability of functions. *Math. USSR-Sb.*, Vol. 71 (1992), no. 2, pp. 561–567. MR1099524 (92b:26019)

[5] A. Melas. The Bellman functions of dyadic-like maximal operators and related inequalities. Advances in Mathematics, Vol. 192 (2005), no. 2, pp. 310–340. MR2128702 (2005k:42052)

[6] F. Nazarov, S. Treil. The hunt for Bellman function: Applications to estimates of singular integral operators and to other classical problems in harmonic analysis. *Algebra i Analiz*, Vol. 8 (1997), no. 5, pp. 32–162. MR1428988 (99d:42026)

[7] F. Nazarov, S. Treil, A. Volberg. Bellman function in stochastic control and harmonic analysis in *Systems, Approximation, Singular integral operators, and related topics*, ed. A Borichev, N. Nikolski, *Operator Theory: Advances and Applications*, Vol. 129, 2001, pp. 393–424, Birkhauser Verlag. MR1882704 (2003b:49024)

[8] F. Nazarov, S. Treil, A. Volberg. The Bellman functions and two-weight inequalities for Haar multipliers. 1995, Preprint, MSU, pp. 1–25.

[9] F. Nazarov, S. Treil, A. Volberg. The Bellman functions and two-weight inequalities for Haar multipliers. *Journal of the American Mathematical Society*, Vol. 12 (1999), no. 4, pp. 909–928. MR1685781 (2000k:42009)

[10] L. Slavin. Bellman function and BMO. Ph.D. thesis, Michigan State University, ProQuest LLC, Ann Arbor, MI, 2004. MR2706427

[11] L. Slavin, A. Volberg. The *s*-function and the exponential integral. Topics in harmonic analysis and ergodic theory, pp. 215–228, *Contemp. Math.*, Vol. 444, Amer. Math. Soc., Providence, RI, 2007. MR2423630 (2009c:26048)

[12] T. Tao. Bellman function and the John–Nirenberg inequality, Preprint, http://www.math.ucla.edu/~tao/~preprints/harmonic.html.

[13] V. Vasyunin. The sharp constant in the reverse Hölder inequality for Muckenhoupt weights. *Algebra i Analiz*, Vol. 15 (2003), no. 1, pp. 73–117. MR1979718 (2004h:42017)

[14] V. Vasyunin. The sharp constant in the John–Nirenberg inequality, Preprint POMI no. 20, 2003. http://www.pdmi.ras.ru/preprint/index.html

[15] V. Vasyunin, A. Volberg. The Bellman functions for a certain two weight inequality: The case study. *Algebra i Analiz*, Vol. 18 (2006), No. 2. MR2244935 (2007k:47053)

[16] A. Volberg. Bellman approach to some problems in harmonic analysis. *Équations aux Dérivées Partielles*, Exposé n. XX, 2002.

Department of Mathematical Sciences, University of Cincinnati, Cincinnati, Ohio 45221-0025

E-mail address: `leonid.slavin@uc.edu`

St. Petersburg Department of the V. A. Steklov Mathematical Institute, Russian Academy of Sciences, St. Petersburg, Russia

E-mail address: `vasyunin@pdmi.ras.ru`