This is to certify that the
dissertation entitled

BELLMAN FUNCTION AND BMO

presented by

LEONID SLAVIN

has been accepted towards fulfillment
of the requirements for the

Doctoral degree in Mathematics

Major Professor's Signature

Aug. 11, 2004

Date

BELLMAN FUNCTION AND *BMO*

By

Leonid Slavin

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Mathematics

2004

ABSTRACT

BELLMAN FUNCTION AND BMO

By

Leonid Slavin

The Bellman function method is applied to three different problems in harmonic analysis. The first, introductory chapter outlines the specifics of the method, addresses its stochastic control origins, and gives a harmonic analysis perspective. A brief description of the main function space under consideration, BMO, is also provided. In the second chapter, the integral form of the John-Nirenberg inequality for BMO functions is examined, the corresponding Bellman function explicitly found, and the sharp constants in the inequality as well as the exact bounds on the region of its validity established. Two cases, those of the continuous and dyadic BMO, are treated and the results differ significantly between the cases. In the third chapter, the dyadic version of the Chang-Wilson-Wolff theorem for functions whose s-function is uniformly bounded is proved using a Bellman-type argument. Furthermore, a local version of the theorem is established, whereby the s-function is assumed to be bounded on a measurable subset E of $[0,1]$. Consequently, the exponential summability of the second order over E is derived. In the fourth chapter, the famous question of $H^1 - BMO$ duality is considered. Two cases, the continuous and dyadic ones, are treated and the same key lemma, based on a Bellman-type argument, is used in both to establish the embedding of BMO in the corresponding dual space. Moreover, in the dyadic case, an explicit estimate for the norm of the embedding is found.

To the memory of my father

ACKNOWLEDGMENTS

I would like to express my deepest and sincerest gratitude to my advisor, Professor Alexander Volberg. He has taught me all the skills of the trade I have had the ability to grasp and he is responsible for my reaching this stage in my career. I am very grateful for his patience and generosity with his time, ideas, and expertise. I am also very happy to be working in this field and credit him solely for the opportunity to be on the cutting edge of research. Whatever mathematical accomplishments I dare count to my name would not have existed, were it not for his help and encouragement. The time I have spent as his apprentice will forever be appreciated.

I would also like to thank my mentor and coauthor, Professor Vasily Vasyunin. He has opened my eyes to many a subtle mathematical phenomenon in the process of our joint work and taken time to educate me on not so subtle matters. His friendship, truly infinite patience, and preparedness to give the youngling an opportunity after opportunity to establish himself have helped me immensely.

I am very thankful to Professor Michael Frazier for the many useful discussions I have had the pleasure and advantage of having with him. He has helped me prepare for my talks and his insight has given me a broader perspective on certain areas of research. I am also very thankful for his help in preparation for my job interviews.

I would like to express my appreciation to Professors Alexander Volberg, Clifford Weil, Vladimir Peller, Joel Shapiro, and Kirill Vaninsky for agreeing to be on my committee, to have their patience tried with delays, and their kindness tested with this imperfect text.

Last but not least, I would like to thank my best friends Leonid Freidovich and Zeynep Altinsel for being there for me all these years.

Contents

List of Figures

Chapter 1

Preliminaries

1.1 Introduction

This work is a compilation of three results united by a common theme. In all three, a relatively new harmonic analysis technique is used to establish the crucial segments of the argument: the Bellman function method. With its origins in stochastic control, its timeline in analysis both short and rich, and its scope seemingly unlimited, the method is as novel as it is powerful. Every successful application, therefore, serves to further the method's legacy as well as explore and develop its subtle and intricate aspects that only come to light in a particular problem. That is why applying the Bellman function technique to problems that have been solved by other methods merits a researcher's time. Due to its nature, this technique often streamlines the proofs and/or reveals new properties of the function spaces under consideration. What is more, it has been applied to obtain sharp and dimensionless results in existing estimates and inequalities, establish their regions of validity, and, sometimes, disprove intriguing conjectures.

The other theme linking the results presented is, as the title suggests, the function space for whose elements they are valid. It is BMO, the space of functions of bounded mean oscillation (see below for detailed description). The first and the last results establish certain properties of BMO functions explicitly, whereas the second one deals with a space that is better (smaller) than BMO, and whose elements thus possess properties that are similar in nature to those of BMO functions, but stronger; that is to say it proves exponential integrability but of higher order than that of BMO.

Chapter 2 is on the John-Nirenberg inequality for BMO, perhaps the most fundamental property of BMO functions. It establishes sharp constants and bounds previously unknown. This is a joint result with Professor Vasily Vasyunin of St. Petersburg Department of Steklov Mathematical Institute and Mathematics Department of St. Petersburg State University. This work was inspired by that of Professor Alexander Volberg of Michigan State University, the mathematician who is, perhaps, the most responsible for establishing and promoting the Bellman function method as a tool in harmonic analysis. It is his presentation of the stochastic control Bellman set-up for the John-Nirenberg inequality in [15] and an earlier version of [24] that, at about the same time but at vastly distant venues, caught our attention and spurred further research. We then independently found the Bellman function for the problem, after which we joined the efforts to solve the problem completely. Many useful discussions with A. Volberg have allowed us to use his insight and expertise. Two formulations, the continuous and dyadic, are considered. Notably, the results are significantly different in the two cases.

Chapter 3 is on a joint result with A. Volberg. It gives a Bellman-function-type proof of a famous result of Chang, Wilson, and Wolff for functions whose square function is bounded and uses that result to establish that the square of such a function is exponentially summable. More precisely, for any φ from the unit ball in the corresponding space (with the usual factorization over constant functions and the s-function as the semi-norm), $e^{\alpha\varphi^2}$ is integrable. Furthermore, we give a local version of the theorem, whereas the s-function is assumed to be from $L^\infty(E)$, as opposed to $L^\infty([0,1])$, for a measurable subset E of $[0,1]$. The result is a weak-form estimate, similar to that in the Chang, Wilson, Wolff's original paper [2] and a uniform bound on the integral $\int_E e^{\alpha\varphi^2}$ for some $\alpha > 0$. The article [2] does not give the local version, although it seems that the authors' reasoning can be modified to make it work for an arbitrary E. Our proof is short, straightforward, and constructive — the usual advantages of the method used. The formulation of an extremal problem is discussed, although the "Bellman function" used in the proof is a supersolution of the optimization problem, i.e. a majorate of the "true" Bellman function. The Bellman approach to this problem has very intriguing implications, as discussed in the section on research prospects.

Chapter 4 is on a joint result with A.Volberg. It examines the old question of $H^1 - BMO$ duality, first addressed by Fefferman in [5]. We use a Bellman-type reasoning to establish a key lemma, which then works to prove two duality results: the continuous one and its dyadic analog. No optimization problem is set up; thus we use a hands-on, heuristic approach to find a "Bellman function," whose properties are dictated by the differential estimates in the lemma. As is often the case with this technique, we are able to obtain an explicit embedding constant in the dyadic case, although we do not discuss whether it is sharp. Moreover, the same proof seems to work in a multidimensional setting and, with the types of H^1 and BMO norms used, the constants of embedding are dimensionless.

One common feature, the hallmark of a Bellman function or Bellman-function-type proof, that is present in all three results is the unwrapping of a certain integral sum, undoubtedly the main connection between stochastic control and harmonic analysis. We expand on this connection in the next two sections.

1.2 Stochastic control Bellman function

We will formally derive the Bellman equation for a controlled stochastic process following the exposition in [15, 24, 10].

Let x^t be an d-dimensional stochastic process, satisfying the stochastic differential equation

$$x^t = x + \int_0^t \sigma(\alpha^s, x^s)\, dw^s + \int_0^t b(\alpha^s, x^s)\, ds. \tag{1.1}$$

Here t is the time, w^t is a d_1-dimensional Wiener process, $\sigma(a, y)$ is a $d \times d_1$ matrix, and b is a d-dimensional vector. Different choices of the control, α^s, which is a d-dimensional stochastic process, give us different trajectories, i.e. different solutions of (1.1). The derivation we give will be entirely formal; thus we do not address questions of existence or uniqueness of solutions.

Equation (1.1) is a part of an optimal control problem. Namely, given a profit function f^α, on the trajectory x^t, for the interval $[t, t + \Delta t]$, the profit is

$$f^{\alpha^t}(x^t)\Delta t + o(\Delta t).$$

Therefore, on the whole trajectory we earn

$$\int_0^\infty f^{\alpha^t}(x^t)\, dt.$$

We want to choose the control $\alpha = \{\alpha^s\}$ to maximize the average profit

$$v^\alpha(x) = \mathbb{E} \int_0^\infty f^{\alpha^t}(x^t)\, dt + \overline{\lim_{t\to\infty}}\, \mathbb{E}(F(x^t)), \tag{1.2}$$

for the process starting at x. Here $F \geq 0$ is the bonus function — one gets it when one retires. The Bellman function for the process (1.1) is the optimal average gain,

$$v(x) = \sup_{\alpha \in A} v^{\alpha}(x), \tag{1.3}$$

where A is the set of admissible controls; v satisfies the well-known Bellman (partial) differential equation, which is based on two ingredients: Bellman's principle and Ito's formula.

Bellman's principle states that

$$v(x) = \sup_{\alpha \in A} \mathbb{E}\left[\int_0^t f^{\alpha^s}(x^s)\,ds + v(x^t)\right]. \tag{1.4}$$

To explain it, we fix $t > 0$ and consider an individual trajectory. The profit for the interval $[0, t]$ is given by

$$\int_0^t f^{\alpha^s}(x^s)\,ds.$$

Suppose the trajectory has reached the point y at the moment t. The maximal average profit we can make starting at t and at the point y is precisely $v(y)$. Indeed, since the increments of w^s for $s \geq t$ do not depend on $w^\tau, \tau < t$, and equation (1.1) is time-invariant, there is no difference between starting at time 0 or at time t. Applying the full probability formula to take into account all possible endpoints $y = x^t$, we obtain (1.4).

Let us now explain the version of Ito's formula that we need. Fix a moment of time s and a small increment Δs. We want to estimate the difference $v(x^{s+\Delta s}) - v(x^s)$. Let $\Delta w^s = w^{s+\Delta s} - w^s$. Assuming enough smoothness, we can use Taylor's formula. Among others, we will have the term

$$\sum_{k=1}^{d} \frac{\partial v}{\partial x_k}(x^s) \sum_{j=1}^{d_1} \sigma_{kj}(\alpha^s, x^s)\Delta w_j^s + \sum_{k=1}^{d} \frac{\partial v}{\partial x_k}(x^s)\, b_k(\alpha^s, x^s)\Delta s.$$

After taking the expectation, the first term will vanish, since each Δw_k^s is independent of x^s and has zero mean. The second term can be rewritten as

$$\mathbb{E}\left(\mathcal{L}_1^{\alpha^s}(x^s)v\right)(x^s)\Delta s,$$

with the first order differential operator $\mathcal{L}_1^\alpha$ given by

$$\mathcal{L}_1^\alpha = \sum_{k=1}^d b_k(\alpha, s)\,\frac{\partial}{\partial x_k}.$$

The next term in the Taylor formula will be

$$\frac{1}{2}\sum_{k,j}\frac{\partial^2 v}{\partial x_j \partial x_k}\left(\sum_k \sigma_{jk}\Delta w_k^s + b_j(\alpha^s, x^s)\Delta s\right)\left(\sum_k \sigma_{ik}\Delta w_k^s + b_i(\alpha^s, x^s)\Delta s\right).$$

Averaging over probability, taking into account the fact that $\mathbb{E}\Delta w_k^s \Delta w_m^s = \Delta s$ if $k = m$ and 0 otherwise, and omitting the terms with $(\Delta s)^2$, we get

$$\mathbb{E}\left(\mathcal{L}_2^{\alpha^s}(x^s)v\right)(x^s)\Delta s,$$

where the second-order differential operator $\mathcal{L}_2^\alpha$ is given by

$$\mathcal{L}_2^\alpha = \sum_{i,j=1}^d a^{ij}(\alpha, x)\,\frac{\partial^2}{\partial x_i \partial x_j}; \qquad a^{ij}(\alpha, x) = \frac{1}{2}\sum_{k=1}^{d_1}\sigma_{ik}(\alpha, x)\,\sigma_{jk}(\alpha, x).$$

Gathering all the terms together and omitting the ones with powers of Δs greater than one, we obtain

$$\mathbb{E}(v(x^t)) = v(x) + \mathbb{E}\int_0^t \mathcal{L}^{\alpha^s}(x^s)\,v(x^s)\,ds, \tag{1.5}$$

where $\mathcal{L}^\alpha = \mathcal{L}_1^\alpha + \mathcal{L}_2^\alpha$. That is the application of Ito's formula we need. Putting (1.5) into Bellman's principle (1.4), we get

$$0 = \sup_{\alpha \in A} \left[\int_0^t f^{\alpha^s}(x^s) + \int_0^t \mathcal{L}^{\alpha^s}(x^s)\, v(x^s)\, ds \right].$$

Dividing by t an taking the limit as $t \to 0$ (assuming it is justified), we get Bellman's partial differential equation supplemented by the obstacle condition $v \geq F$

$$\begin{aligned} &\sup_{\alpha \in A} [\mathcal{L}^{\alpha}(x)v(x) + f^{\alpha}(x)] = 0, \ x \in \Omega \\ &v(x) \geq F(x), \ x \in \Omega. \end{aligned} \tag{1.6}$$

1.3 Harmonic analysis Bellman function

The scope of applications of the Bellman function method has been exceptionally broad. While the method seems well suited for proving weighted norm inequalities (see the early 1995 version of paper [16], the ground-breaking proof of the matrix Hunt-Muckenhoupt-Wheeden theorem in [13], or a more recent result in [20]; alternatively, see [12] for a two-weight negative result), it has found use in areas far from its origins. The (much needed) anthology of the existing Bellman function results is far beyond the scope of this chapter. An incomplete list of references, besides those already named, includes [7, 14, 17, 18, 19] and, perhaps the closest in spirit to the current work, [23].

This panoply of results seemed to necessitate the development of a uniform foundation. Such a foundation has been found in the very field from which the notion of a Bellman function first arose — that of stochastic control. The work on developing the stochastic control framework for harmonic analysis problems was begun in [15] and continued in [24]. It is in paper [15] that the famous result of Burkholder for martingale transforms [3] was interpreted from the stochastic perspective. We refer the reader to these articles for a thorough exposition and many interesting examples.

Here, we would like to indicate the general principles of the stochastic control framework. We will use them to develop the Bellman equation for the John-Nirenberg inequality after we introduce the space BMO in the next section.

The problems that can be treated using the Bellman function from stochastic control and thus the machinery of the previous section are always dyadic. It is often possible to pass from a dyadic problem to the corresponding problem with analytic or harmonic function using some kind of Green's formula. (That is exactly what is done in Chapter 4 in the continuous case.) It is, therefore, very interesting and unexpected that in the John-Nirenberg setting of Chapter 2, we, motivated by the stochastic control Bellman function formally derived for the inequality, first solve the continuous problem and only then use those results to return to the dyadic case.

First of all, we make a choice of variables so that the function space under consideration is mapped onto a Euclidean domain. Namely, assume that we want to prove a certain inequality for all functions from the δ-ball, F_δ, of a space F. Then, with every pair (φ, J), where $\varphi \in F_\delta$ and J is a dyadic interval, we associate a d-dimensional vector $x = (x_1, x_2, \ldots, x_d)$ in a domain Ω_δ whose geometry is determined by the space F. Very often, the coordinates x_i will have a martingale structure; we often see $x_1(\varphi, J) = \langle \varphi \rangle_J$. (We make a choice like this when dealing with the John-Nirenberg inequality below.) The vector x then is the state vector of our system, i.e. a solution of (1.1). Thus, given a function $\varphi \in F_\delta$, we know the state of the system on every dyadic level (for every generation of dyadic intervals), that is to say, at every moment t. Naturally, the time is now discrete and is equivalent to the order of the current generation. Instead of (1.1), we now have a discrete model

$$x^{n+1} = x^n + \sigma(\alpha^n, x^n)\Delta^n w + b(\alpha^n, x^n). \tag{1.7}$$

Most of the time, we have $\sigma = \mathrm{diag}\{\alpha_1, \alpha_2, \ldots, \alpha_d\}$. The exact form of the matrix

σ and the vector b is determined by the difference $x^{n+1} - x^n$. The Wiener process is very often just a series of coin tosses: we either move to the left or the right half of the interval I. Thus the equation (1.7) often becomes

$$x^{n+1} = x^n + \alpha^n \xi^n + b(\alpha^n, x^n),$$

where ξ^n is either 1 or -1. If we are maximizing a sum of the form

$$\frac{1}{|I|} \sum_{J \subseteq I} \dots,$$

that sum can be interpreted as the expectation of the cumulative gain $\int_0^\infty f^{\alpha^s}(x^s)\, ds$, allowing us to choose the profit function correctly. On the other hand, if we have a term of the form $\langle g(\varphi) \rangle_I$ to maximize, and $x_1 = \langle \varphi \rangle$, then the bonus (obstacle) function is $g(x_1)$.

We have described the method very empirically but just enough to be able to formulate the Bellman optimization problem for the John-Nirenberg inequality. First, however, we need to introduce the space in question, BMO, in order to understand the geometry of the state domain Ω.

1.4 The space BMO

1.4.1 John-Nirenberg inequality

The space of functions of bounded mean oscillation, or BMO, was introduced by John and Nirenberg in [8] in their work on partial differential equations and quickly began to play a very prominent role in harmonic analysis. An excellent reference on the complex-variable approach to BMO is [6], whereas [21] describes this space in the real-variable and multi-dimensional setting in great detail.

The original definition follows. Let I be an interval. Then

$$BMO(I) = \left\{\varphi \in L^1(I): \sup_{J \subseteq I} \frac{1}{|J|} \int_J |\varphi(t) - \langle\varphi\rangle_J| \, dt < \infty\right\}. \tag{1.8}$$

Here $\langle\varphi\rangle_I = \frac{1}{|I|}\int_I \varphi(t)\, dt$, and J is a subinterval of I. If we factorize the space (1.8) over constant functions, it becomes a Banach space with the norm

$$\|\varphi\|_{BMO(I)} = \sup_{J \subseteq I} \frac{1}{|J|} \int_J |\varphi(t) - \langle\varphi\rangle_J| \, dt. \tag{1.9}$$

The most fundamental result for BMO is the weak-form John-Nirenberg theorem, first proved in [8], that states that for every $\varphi \in BMO(I)$ one has

$$m\left(\{x \in I: \ |\varphi(x) - \langle\varphi\rangle_I| > \lambda\}\right) \le c_1 e^{-c_2 \lambda \|\varphi\|_{BMO(I)}}. \tag{1.10}$$

Finding the sharp constants in the inequality (1.10) is a natural goal. In an important development, Korenovskii found the sharp constant $c_2 = 2/e$ in [9].

A remarkable consequence of the John-Nirenberg inequality is that every BMO function φ is in $L^p(I), 1 \le p < \infty$ and

$$\left(\sup_{J \subseteq I} \frac{1}{|J|} \int_J |\varphi(t) - \langle\varphi\rangle_J|^p \, dt\right)^{1/p} \tag{1.11}$$

defines an equivalent norm on $BMO(I)$. We are particularly interested in using the L^2-based norm. Let

$$\|\varphi\|_{BMO(I),2} = \left(\sup_{J \subseteq I} \frac{1}{|J|} \int_J |\varphi(t) - \langle\varphi\rangle_J|^2 \, dt\right)^{1/2}. \tag{1.12}$$

We can — and this is the main reason for using $p = 2$ — rewrite (1.12) as

$$\|\varphi\|_{BMO(I),2} = \left(\sup_{J \subseteq I} \left\{\langle\varphi^2\rangle_J - \langle\varphi\rangle_I^2\right\}\right)^{1/2}. \tag{1.13}$$

Furthermore, we have the corresponding weak-form John-Nirenberg inequality

$$m\left(\{x\in I:\ |\varphi(x)-\langle\varphi\rangle_I|>\lambda\}\right)\le c_1e^{-c_2\lambda/\|\varphi\|_{BMO(I),2}}. \tag{1.14}$$

with constants c_1 and c_2 different from those in (1.10). One can rewrite (1.14) in the integral form, as follows. There exists $\varepsilon_0>0$ such that for every $\varepsilon<\varepsilon_0$ and every $\varphi\in BMO(I)$ such that $\|\varphi\|_{BMO(I),2}\le\varepsilon$, we have

$$\langle e^{\varphi}\rangle_I\le C(\varepsilon)e^{\langle\varphi\rangle_I}, \tag{1.15}$$

for some constant $C(\varepsilon)$. The inequality (1.15) is the reverse Jensen inequality. Finding the sharp value for ε_0 and the sharp expression for $C(\varepsilon)$ is highly desirable. Chapter 2 describes how it is done in both continuous and dyadic settings.

1.4.2 H^1-BMO duality

A major reason BMO gained prominence is paper [5] in which Fefferman established it as dual to H^1. This is the result we take up in Chapter 4. Both spaces are considered on the unit circle $\mathbb{T}$. Again, we treat two cases, the continuous and dyadic. In the dyadic case, we use the following BMO norm

$$\|\varphi\|=\sup_{J\in D}\frac{1}{|J|}\sum_{I\subseteq J}\left(\langle\varphi\rangle_{I_+}-\langle\varphi\rangle_{I_-}\right)^2|I|, \tag{1.16}$$

where D is the dyadic lattice rooted in $\mathbb{T}$ and I_-,I_+ are the left and right halves of the dyadic arc I, correspondingly. Here, the advantage of the L^2-formulation for BMO is evident, since

$$\frac{1}{|J|}\sum_{I\subseteq J}\left(\langle\varphi\rangle_{I_+}-\langle\varphi\rangle_{I_-}\right)^2|I|=4\left(\langle\varphi^2\rangle_J-\langle\varphi\rangle_J^2\right).$$

(See Chapter 4 for a detailed explanation.) Thus we are using virtually the same BMO norm as in (1.12).

In the continuous case, the equivalence of the norm (1.12) and

$$\|\varphi\| = \sup_{\text{arc } I\subset\mathbb{T}} \frac{1}{|I|}\int_{Q_I} |\varphi'(\xi)|^2(1-|\xi|)\,dA(\xi), \tag{1.17}$$

where $\varphi(z)$ is the harmonic extension of φ into the disk and Q_I is the Carleson square based on the arc I, can be found, for instance, in [21].

1.5 The Bellman equation for the John-Nirenberg inequality

To preserve history, we will develop the equation in the way we first encountered it, although the attentive reader will notice that we use a slightly different choice of variables in Chapter 2. According to the previous section, we have the following underlying dyadic problem: Given that $\left\langle(\varphi-\langle\varphi\rangle_J)^2\right\rangle_J \le \delta^2$, for every dyadic subinterval J of I, prove that

$$\langle e^{\varphi}\rangle_I \le C(\delta)e^{\langle\varphi\rangle_I}.$$

For every J, let $x_1 = \langle\varphi\rangle_J$, $x_2 = \left\langle(\varphi-\langle\varphi\rangle_J)^2\right\rangle_J$. The Cauchy inequality and the assumption of the theorem give $\Omega_\delta = \{(x_1,x_2):\ x_1\in\mathbb{R},\ 0\le x_2\le\delta\}$.

If we take conditional expectations $\mathbb{E}(\,\cdot\,|x^n)$ of (1.7), we get

$$b_1(\alpha,x^n) = \mathbb{E}(x_1^{n+1}|x_1^n) - x_1^n = \frac{(x_1^n)^- + (x_1^n)^+}{2} - x_1^n = 0$$

and

$$b_2(\alpha, x^n) = \mathbb{E}(x_2^{n+1}|x_2^n) - x_2^n = \frac{(x_2^n)^- + (x_2^n)^+}{2} - x_2^n = \left(\frac{(x_2^n)^- - (x_2^n)^+}{2}\right)^2 = (\alpha_1^n)^2.$$

Furthermore, according to the reasoning of section 1.3, the profit function is 0 and the obstacle function is $F(x) = e^{x_1}$. We, therefore, solve the optimization problem

$$\sup_{\alpha=(\alpha_1,\alpha_2)} \left[\frac{1}{2}\langle d^2v, v\rangle - \frac{\partial v}{\partial x_2}\alpha_1^2\right] = 0$$
$$v(x) \geq e^{x_1}, x \in \Omega_\delta. \tag{1.18}$$

This is the formulation that led us (independently) to the corresponding family of solutions

$$v_\delta(x) = \frac{\sqrt{\delta^2 - x_2}}{1-\delta} e^{x_1 + \sqrt{\delta^2 - x_2} - \delta}.$$

As it turned out, this function was NOT the dyadic Bellman function for the domain Ω_δ. It was not until the crucial splitting tool (Lemma 4 in Chapter 2) was developed in [23] and the realization that the continuous case was the one to consider ensued, that the result started to develop further. The continuous Bellman function was first discovered and after much effort the dyadic one was found in the same family, although, somewhat bafflingly, not on the same level (to clarify these cryptic statements, the reader is encouraged to refer to Chapter 2 for a complete presentation).

1.6 Bellman-function-type proofs

We now attempt to address the type of Bellman function argument in which no explicit Bellman function is found. We can roughly consider two categories of proofs of this sort. First, we may have an extremal problem posed and, instead of finding the Bellman function itself, we find its majorate. Provided that is bounded, so is the

Bellman function and the result follows. The influential papers [15, 24] describe the procedure for finding such majorates, provided the corresponding Bellman equation has been obtained. Chapter 3 details a result of this kind. However, we do not even attempt to obtain the Bellman equation there. The reason is that the most important property of the formal stochastic-control-like development in the previous sections was that the Bellman function, defined as the supremum (taken over all elements of a function space) of an integral or a dyadic sum over an interval, would not depend on the interval. For instance, in the John-Nirenberg setting of Chapter 2 (slightly different from that of the previous section), the Bellman function, defined by

$$\mathbf{B}_\varepsilon(x) = \sup_{\|\varphi\|_{BMO(I)} \leq \varepsilon} \left\{ \langle e^{\varphi} \rangle_I : \ \langle \varphi \rangle_I = x_1, \langle \varphi^2 \rangle_I = x_2 \right\},$$

does not depend on the interval I. This allows for effective modeling using the stochastic control technique. In the case of the local Chang-Wilson-Wolff theorem, given a measurable subset E of $[0,1]$, the corresponding Bellman function

$$\mathbf{B}_t^\alpha(x) = \sup_{\varphi \in F_\alpha} \left\{ \int_E e^{t\varphi(s)}\, ds : \ \langle \varphi \rangle_{[0,1]} = x_1, S^{\varphi}_{[0,1]} = x_2 \right\}$$

is very E-specific. This is, of course, due to the fact that E does not scale as well as $[0,1]$ itself does, i.e. dividing an interval in half, we will not necessarily divide that interval's portion of E in half. Of equal importance is the fact that the variable S_I defined in the chapter does not have a martingale character to it; thus it does not scale correctly either.

A reasonable question then is whether one can obtain a more manageable Bellman function set-up in the case $E = [0,1]$ (the original Chang-Wilson-Wolff case), if, instead of S_I, one uses the s-function itself as a variable. This remains subject of further investigation.

The second category of Bellman-function-type arguments are the ones where the extremal problem does not even enter the consideration. Chapter 4 deals with a result of this kind, the H^1-BMO duality. Although it is possible to think of an underlying optimization problem, the Bellman function involved appears simply as a means to carry out a clever segment of the proof. It is a very utilitarian approach — one attempts to unwrap the corresponding sum over dyadic intervals with a nonexistent (as of yet) function B and imposes such differential and quantitative properties on the function as are required so that the estimates work out the right way, the constants are not too big, etc. After that, one attempts to find the function satisfying the properties so determined and, perhaps, optimize its parameters or tweak it otherwise. We daresay this is the way many explicit Bellman functions as well as their majorates are constructed. Admittedly, this is how the function in Chapter 3 was constructed (The function in Chapter 2, however, stands out as a pure product of stochastic control formalism.)

What, then, distinguishes a Bellman-type proof from any other? Precisely the unwrapping of an integral/dyadic sum, mentioned in the preceding paragraph. The three key lemmas of the three chapters that follow this introduction, the reader will notice, accomplish just that.

We are now in a position to present the main results of this thesis. It is our hope that the reader will appreciate the unifying ideas behind the proofs and the flexibility of the method in dealing with subtleties of each particular case. Some prospects for future research are detailed in the last section.

Chapter 2

Sharp constants and bounds in the John-Nirenberg inequality

2.1 Introduction

For any interval $I \subset \mathbb{R}$ and a function $\varphi \in L^1(I)$, we denote by $\langle\varphi\rangle_I$ the average of φ over I, $\langle\varphi\rangle_I = \frac{1}{|I|}\int_I \varphi(t)\,dt$. We define the space $BMO(I)$ as

$$BMO(I) = \left\{\varphi \in L^2(I) : \int_J |\varphi(t) - \langle\varphi\rangle_I|^2 dt \leq C^2|J|, \forall \text{ interval } J \subset I\right\} \tag{2.1}$$

with the best such C being the corresponding norm of φ. This definition can be rewritten in a more useful form:

$$BMO(I) = \{\varphi \in L^2(I) : \langle\varphi^2\rangle_J - \langle\varphi\rangle_J^2 \leq C^2, \forall J \subset I\} \tag{2.2}$$

with the norm

$$\|\varphi\|_{BMO(I)} = \left(\sup_{J\subset I}\{\langle\varphi^2\rangle_J - \langle\varphi\rangle_J^2\}\right)^{1/2}. \tag{2.3}$$

We also introduce $BMO^d(I)$, the dyadic analog of $BMO(I)$, whose definition is identical to that of $BMO(I)$, except that the intervals J over which the supremum is taken are members of the dyadic lattice based on I. Finally, by $BMO_\varepsilon(I)$ and $BMO_\varepsilon^d(I)$ we denote the ε-ball (the ball of radius ε centered at 0) in the corresponding space.

The following result is well known. (This is the integral form of the John-Nirenberg theorem.)

Theorem. *There exists* $\varepsilon_0 > 0$ *such that for every* $0 \le \varepsilon < \varepsilon_0$ *there is* $C(\varepsilon) > 0$ *such that for any function* $\varphi \in BMO_\varepsilon(I)$,

$$\langle e^\varphi \rangle_I \le C(\varepsilon) e^{\langle \varphi \rangle_I}. \tag{2.4}$$

We are interested in determining the sharp bound ε_0 and the exact expression for $C(\varepsilon)$. We will do that in the case of both, the conventional (continuous) BMO and the dyadic BMO.

In accordance with the ideology of the Bellman function method, with every ball $BMO_\varepsilon(I)$ $(BMO_\varepsilon^d(I))$ and the set of all subintervals $J \subset I$ we associate the domain $\Omega_\varepsilon = \{x = (x_1, x_2) : x_1 \in \mathbb{R}, x_1^2 \le x_2 \le x_1^2 + \varepsilon^2\}$ as follows

$$(\varphi, J) \longmapsto \left(\langle \varphi \rangle_J, \langle \varphi^2 \rangle_J\right). \tag{2.5}$$

This map is well-defined because $\langle \varphi \rangle_J^2 \le \langle \varphi^2 \rangle_J$ (Cauchy inequality) and $\varphi \in BMO_\varepsilon(I)$ $(BMO_\varepsilon^d(I))$. On Ω_ε, we define the following Bellman functions

$$\mathbf{B}_\varepsilon(x) = \sup_{\varphi \in BMO_\varepsilon(I)} \left\{ \langle e^\varphi \rangle_I : \ \langle \varphi \rangle_I = x_1, \langle \varphi^2 \rangle_I = x_2 \right\}, \tag{2.6}$$

$$\mathbf{B}_\varepsilon^d(x) = \sup_{\varphi \in BMO_\varepsilon^d(I)} \left\{ \langle e^\varphi \rangle_I : \ \langle \varphi \rangle_I = x_1, \langle \varphi^2 \rangle_I = x_2 \right\}. \tag{2.7}$$

Observe that these functions do not depend on I. Finding them explicitly would provide us with the complete solution of the John-Nirenberg problem. That is exactly what we are able to accomplish. Remarkably and unlike many other Bellman function problems, the results for the continuous and dyadic cases differ significantly. Which is more, we first solve the continuous problem and only then use the corresponding Bellman function family to solve the dyadic case.

2.2 Main results

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

$$C(\varepsilon) = \frac{e^{-\varepsilon}}{1-\varepsilon}. \tag{2.8}$$

Then, for any $\varphi \in BMO_\varepsilon(I)$,

$$\langle e^\varphi \rangle_I \le C(\varepsilon) e^{\langle \varphi \rangle_I}. \tag{2.9}$$

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

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

$$C^d(\varepsilon) = C(\delta), \tag{2.10}$$

where $C(\delta)$ *is defined by (2.8) and* $\delta = \delta(\varepsilon)$ *is the unique solution of the equation*

$$(1 - \sqrt{\delta^2 - \varepsilon^2}) e^{\sqrt{\delta^2 - \varepsilon^2}} \left[2 - e^{\varepsilon/\sqrt{2}} \right] - (1-\delta) e^{\delta - \varepsilon/\sqrt{2}} = 0. \tag{2.11}$$

Then, for any $\varphi \in BMO_\varepsilon^d(I)$,

$$\langle e^{\varphi}\rangle_I \le C^d(\varepsilon) e^{\langle\varphi\rangle_I}. \tag{2.12}$$

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

Theorems 1 and 2 are immediate consequences of the following results for the Bellman functions (2.6) and (2.7). Let

$$B_\delta(x) = \frac{1-\sqrt{\delta^2+x_1^2-x_2}}{1-\delta}\exp\left(x_1+\sqrt{\delta^2+x_1^2-x_2}-\delta\right). \tag{2.13}$$

Theorem 3. *If* $0 \le \varepsilon < 1$, *then*

$$\mathbf{B}_\varepsilon(x) = B_\varepsilon(x); \tag{2.14}$$

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

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

Theorem 4. *If* $0 \le \varepsilon < \sqrt{2}\log 2$, *then*

$$\mathbf{B}_\varepsilon^d(x) = B_{\delta(\varepsilon)}(x); \tag{2.15}$$

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

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

Indeed, since the function $t \mapsto (1-t)e^t$ is decreasing for positive t, $\mathbf{B}$ assumes its maximum when $x_2 = x_1^2 + \varepsilon^2$, i.e.

$$\mathbf{B}(x) \le \frac{e^{-\varepsilon}}{1-\varepsilon} e^{x_1},$$

giving (2.9) and (2.12) with the sharp constant (2.8).

2.3 The continuous case

2.3.1 The key lemmas

We first consider the continuous case and prove Theorem 3. One can observe that the proof does not work in the dyadic case. We split the proof of the identity (2.14) into two parts.

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

$$\mathbf{B}_\varepsilon(x) \ge B_\varepsilon(x).$$

Proof. We prove this inequality by explicitly finding a function φ for every point $x \in \Omega_\varepsilon$ such that $(\langle\varphi\rangle_I, \langle\varphi^2\rangle_I) = (x_1, x_2)$ and

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

Since $x_2 = x_1^2$ occurs if and only if $\varphi = x_1 = const$, it is clear that $\mathbf{B}_0(x) = B_0(x) = e^{x_1}$. 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

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

Direct calculation shows that $\varphi_{a,b,\gamma} \in BMO_{|\gamma|}(I)$, $\langle\varphi_{a,b,\gamma}\rangle_I = \gamma a + b$, $\langle\varphi^2_{a,b,\gamma}\rangle_I = 2\gamma^2 a + 2\gamma ab + b^2$, and

$$\langle e^{\varphi_{a,b,\gamma}}\rangle_I = \begin{cases} \frac{1-\gamma+a\gamma}{1-\gamma} & \text{if } \gamma < 1 \\ \infty & \text{if } \gamma \geq 1. \end{cases}$$

Since $\mathbf{B}_\varepsilon(x_1, x_1^2) = B_\varepsilon(x_1, x_1^2) = e^{x_1}$ for all ε, we only need to consider the points $x \in \Omega_\varepsilon$ with $x_2 > x_1^2$. Then we can set $a = 1 - \sqrt{\gamma^2 + x_1^2 - x_2}$ 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 $\mathbf{B}_\varepsilon(x) = \infty$. For $\gamma = \varepsilon \in (0,1)$, we get

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

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

$$\mathbf{B}_\varepsilon(x) \leq B_\varepsilon(x) \tag{2.16}$$

Proof. To establish (2.16), we first prove that $\mathbf{B}_\varepsilon(x) \leq B_{\varepsilon_1}(x)$, $\forall \varepsilon_1 > \varepsilon, \forall x \in \Omega_\varepsilon$, and take the limit as $\varepsilon_1 \to \varepsilon$. (Observe that B_ε is continuous in ε from above.) We need the following two results:

Lemma 3. *The function* B_ε *is concave in* Ω_ε, *i.e.*

$$B_\varepsilon(\alpha_- x^- + \alpha_+ x^+) \geq \alpha_- B_\varepsilon(x^-) + \alpha_+ B_\varepsilon(x^+) \tag{2.17}$$

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 4. *Fix* ε. *Take any* $\varepsilon_1 > \varepsilon$. *Then for every interval* I *and every* $\varphi \in BMO_\varepsilon(I)$, *there exists such a splitting* $I = I_- \cup I_+$ *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 $\varphi \in BMO_\varepsilon(I)$. Take any $\varepsilon_1 > \varepsilon$. Observe that $\varphi \in BMO_\varepsilon(J)$ for any subinterval J of I. Split I according to the rule from Lemma 4. Let $I^{0,0} = I$, $I^{1,0} = I_-$, $I^{1,1} = I_+$. Now split I_- and I_+ according to the rule of Lemma 4 and continue this splitting. 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 $2^n - 1$. The corresponding points given by (2.5) are indexed by the same pair of indices. Also, let $\alpha_{n,m} = |I^{n,m}|/|I|$. Since Lemma 4 provides for the value of α_+ uniformly separated from 0 and 1 on every step, we have

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

Then, using Lemma 3 repeatedly, we have

$$\begin{aligned}
B_{\varepsilon_1}(x^{0,0}) &\geq \frac{|I^{1,0}|}{|I^{0,0}|}B_{\varepsilon_1}(x^{1,0}) + \frac{|I^{1,1}|}{|I^{0,0}|}B_{\varepsilon_1}(x^{1,1}) \\
&\geq \frac{|I^{1,0}|}{|I^{0,0}|}\frac{|I^{2,0}|}{|I^{1,0}|}B_{\varepsilon_1}(x^{2,0}) + \frac{|I^{1,0}|}{|I^{0,0}|}\frac{|I^{2,1}|}{|I^{1,0}|}B_{\varepsilon_1}(x^{2,1}) \\
&\quad + \frac{|I^{1,1}|}{|I^{0,0}|}\frac{|I^{2,2}|}{|I^{1,1}|}B_{\varepsilon_1}(x^{2,2}) + \frac{|I^{1,1}|}{|I^{0,0}|}\frac{|I^{2,3}|}{|I^{1,1}|}B_{\varepsilon_1}(x^{2,3}) \qquad (2.18) \\
&= \frac{|I^{2,0}|}{|I^{0,0}|}B_{\varepsilon_1}(x^{2,0}) + \frac{|I^{2,1}|}{|I^{0,0}|}B_{\varepsilon_1}(x^{2,1}) + \frac{|I^{2,2}|}{|I^{0,0}|}B_{\varepsilon_1}(x^{2,2}) + \frac{|I^{2,3}|}{|I^{0,0}|}B_{\varepsilon_1}(x^{2,3}) \\
&\geq \frac{1}{|I^{0,0}|}\sum_{m=0}^{2^n-1}|I^{n,m}|B_{\varepsilon_1}(x^{n,m}) \geq \frac{1}{|I^{0,0}|}\sum_{m=0}^{2^n-1}|I^{n,m}|e^{x_1^{n,m}} = \frac{1}{|I|}\int_I e^{\varphi_n(s)}\,ds,
\end{aligned}$$

where we have used the fact that $B_{\varepsilon_1}(x) \geq e^{x_1}$ and φ_n is the step function,

$\varphi_n(s) = x_1^{n,m}$ for $s \in I^{n,m}$. Since φ_n converges almost everywhere to φ, Fatou's Lemma yields

$$B_{\varepsilon_1}(x) \geq \frac{1}{|I|} \limsup \int_I e^{\varphi_n(s)\,ds} \geq \frac{1}{|I|} \int_I \limsup e^{\varphi_n(s)}\,ds = \frac{1}{|I|} \int_I e^{\varphi(s)}\,ds = \langle e^{\varphi} \rangle_I .$$

Taking supremum over all φ with $\langle \varphi \rangle_I = x_1^{0,0} = x_1$ and $\langle \varphi^2 \rangle_I = x_2^{0,0} = x_2$, we obtain the inequality

$$B_{\varepsilon_1}(x) \geq \mathbf{B}_{\varepsilon}(x),$$

thus proving the lemma. □

To finish the proof of Theorem 3, we need to prove Lemmas 3 and 4. In the section "How to find the Bellman function" below, the function B_ε is explicitly constructed to be concave in Ω_ε, which is what Lemma 3 states. Thus Theorem 3 is contingent on Lemma 4.

Proof of Lemma 4. We fix an interval I and a function $\varphi \in BMO_\varepsilon(I)$. 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 $\alpha_- = \alpha_+ = \frac{1}{2}$. 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 by: $\rho(\alpha_+) = \max_{x \in [\xi, x^0]} \{x_2 - x_1^2\}$. 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 $\xi = x^-$. We stop when $\rho(\alpha_+) = \varepsilon_1^2$ and fix that splitting. It remains to check that

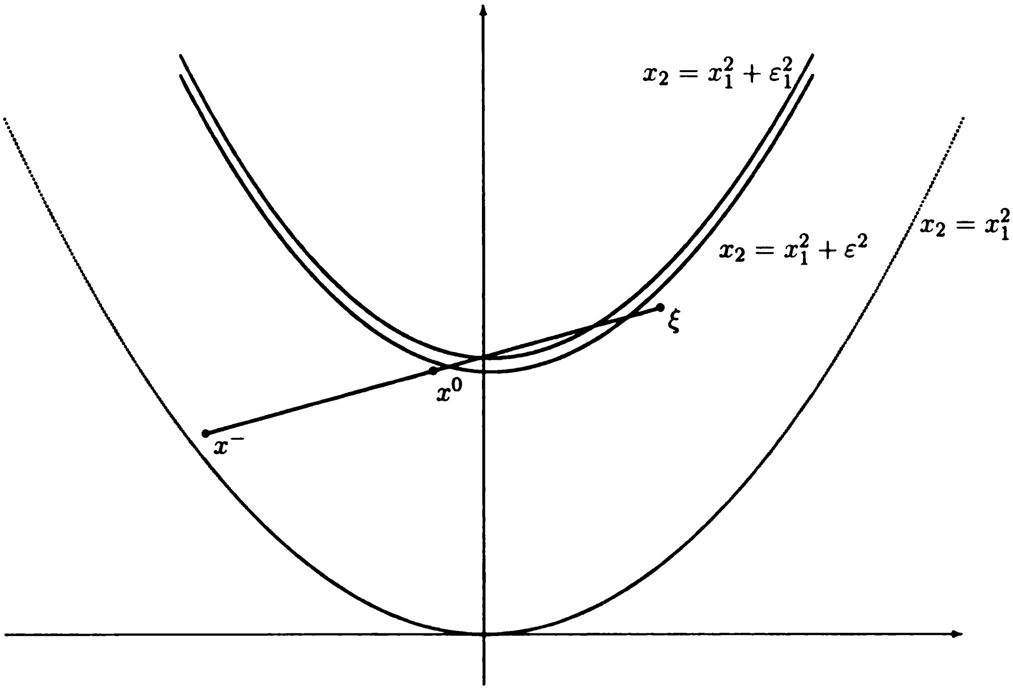


Figure 2.1: The initial splitting: $\alpha_- = \alpha_+ = \frac{1}{2}$, $\xi = x^+$.

such a moment occurs at all 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, $x^0 \in \Omega_\varepsilon$), we conclude that there is a point $\alpha_+ \in \left[\frac{1}{2}, 1\right]$ with $\rho(\alpha_+) = \varepsilon_1^2$.

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 $\xi_1 - x_1^0 = x_1^+ - x_1^0 = \alpha_-(x_1^+ - x_1^-)$. Analogously, 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

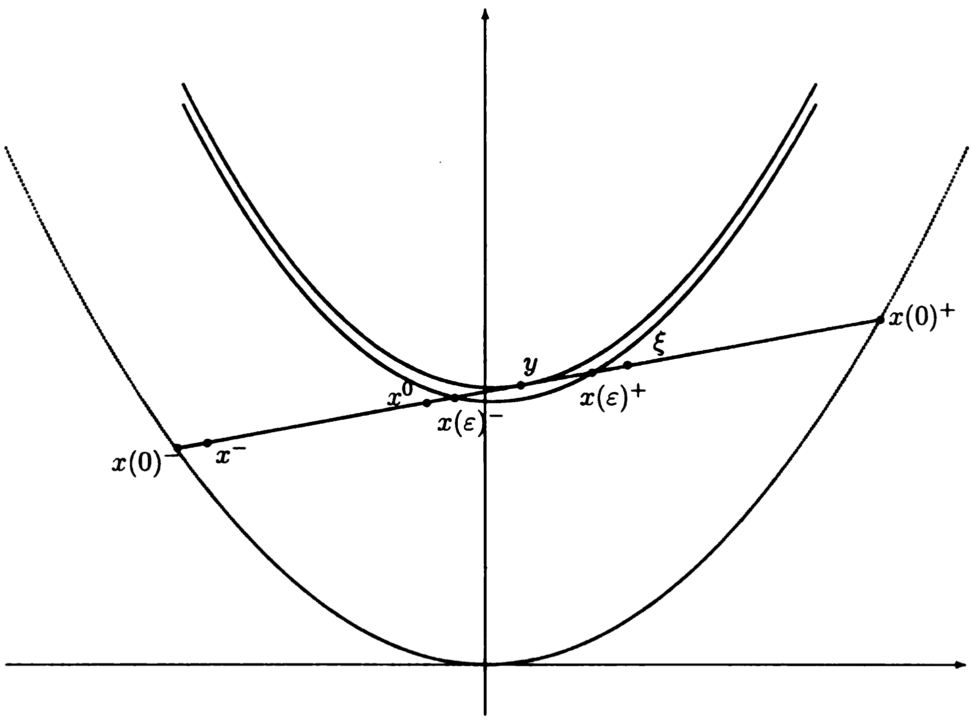


Figure 2.2: The stopping time: $[x^-, \xi]$ is tangent to the parabola $x_2 = x_1^2 + \varepsilon^2$.

is, therefore, $x_2 = 2x_1y_1 - y_1^2 + \varepsilon_1^2$. 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,

$$2\sqrt{\varepsilon_1^2-\varepsilon^2} \;=\; |x(\varepsilon)_1^+ - x(\varepsilon)_1^-| \le |x_1^0-\xi_1| = \min\{\alpha_\pm\}|x_1^+ - x_1^-|$$

$$\le\; \min\{\alpha_\pm\}|x(0)_1^+ - x(0)_1^-| = \min\{\alpha_\pm\}2\varepsilon_1,$$

which implies

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

As promised, this estimate does not depend on φ or I. □

2.3.2 How to find the Bellman function

We first observe that the Bellman function $\mathbf{B}$ must be of the form

$$\mathbf{B}_\varepsilon(x) = \exp\left\{x_1 + w_\varepsilon(x_2 - x_1^2)\right\} \tag{2.19}$$

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

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

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

or

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

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

or

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

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

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

By the Jensen inequality ($\langle e^\varphi \rangle \geq e^{\langle \varphi \rangle}$), we get that $\mathbf{B}(0, x_2 - x_1^2) \geq 1$. Hence, there exists a positive function $w = \log \mathbf{B}(0, \cdot)$ defined on the interval $[0, \varepsilon^2]$ such that (2.19) holds. Furthermore, $x_2 = x_1 = 0$ if and only if $\varphi = 0$. Thus $\mathbf{B}(0,0) = 1$ and $w(0) = 0$.

To use the machinery of Lemma 2, we need the Bellman function candidate B to be a concave function. We thus want

$$-\frac{\partial^2 B}{\partial x_i \partial x_j} \tag{2.20}$$

to be a nonnegative matrix.

Using (2.19), we get

$$\begin{aligned}
\frac{\partial B}{\partial x_1} &= (1 - 2x_1 w')B, \\
\frac{\partial B}{\partial x_2} &= w' B, \\
\frac{\partial^2 B}{\partial x_1^2} &= \left((1 - 2x_1 w')^2 - 2w' + 4x_1^2 w''\right) B, \\
\frac{\partial^2 B}{\partial x_1 \partial x_2} &= \left(w'(1 - 2x_1 w') - 2x_1 w''\right) B, \\
\frac{\partial^2 B}{\partial x_2^2} &= \left((w')^2 + w''\right) B.
\end{aligned}$$

The matrix (2.20) turns into

$$-\begin{bmatrix} \dfrac{\partial^2 B}{\partial x_1^2} & \dfrac{\partial^2 B}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 B}{\partial x_1 \partial x_2} & \dfrac{\partial^2 B}{\partial x_2^2} \end{bmatrix} = -B \begin{bmatrix} 1 & -2x_1 \\ 0 & 1 \end{bmatrix} R \begin{bmatrix} 1 & 0 \\ -2x_1 & 1 \end{bmatrix}, \tag{2.21}$$

where

$$R = \begin{bmatrix} 1-2u' & u' \\ u' & (u')^2 + u'' \end{bmatrix}. \tag{2.22}$$

For the extremal function (if any) we must have equality at every step in (2.18) from Lemma 2, so the matrix (2.20) has to be degenerate. Because of the representation (2.21) and (2.22), this translates into

$$(1-2u')\left((u')^2 + u''\right) = (u')^2, \tag{2.23}$$

$$2u' - 1 \geq 0. \tag{2.24}$$

We solve equation (2.23) by

$$\begin{aligned} (1-2u')u'' &= 2(u')^3 \\ \left(\frac{1}{2(u')^3} - \frac{1}{(u')^2}\right) u'' &= 1 \\ \left(\frac{1}{u'} - \frac{1}{4(u')^2}\right)' &= 1 \\ \frac{1}{u'} - \frac{1}{4(u')^2} &= t + const \\ -\left(1 - \frac{1}{2u'}\right)^2 &= t + const. \end{aligned}$$

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 we continue

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

and

$$1 - \frac{1}{2w'} = \pm\sqrt{\delta^2 - t}. \tag{2.25}$$

Condition (2.24) requires that the square root be strictly less than 1. Therefore, the only feasible solutions for w are those for $\delta < 1$ and such a solution is defined on the interval $[0, \delta^2]$. Equation (2.25) gives

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

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

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

which, together with (2.19), gives formula (2.13)

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

2.3.3 How to find the extremal function

We now show how to find the extremal function that appeared without any explanation in the proof of Lemma 1. As mentioned in the previous section, for the extremal function there is equality at every step in the chain of inequalities (2.18). Thus in the splitting process we only proceed along the vector field defined by the kernel vectors of the matrix (2.20). Using (2.13), we obtain the quadratic form of that matrix

$$\sum_{i,j=1}^{2} \frac{\partial^2 B}{\partial x_i \partial x_j} \Delta_I \Delta_j = \frac{\left(\left(x_1 + \sqrt{\delta^2 + x_1^2 - x_2}\right)\Delta_1 - \frac{1}{2}\Delta_2\right)^2}{\sqrt{\delta^2 + x_1^2 - x_2}\,(1-\delta)} \exp\left\{x_1 + \sqrt{\delta^2 + x_1^2 - x_2} - \delta\right\}. \tag{2.26}$$

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

$$\left(x_1 + \sqrt{\delta^2 + x_1^2 - x_2}\right) dx_1 = \frac{1}{2} dx_2. \tag{2.27}$$

Introducing the variable $t = \sqrt{\delta^2 + x_1^2 - x_2}$, 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 (2.27) by $x_1\,dx_1 - t\,dt$, we get $t\,dx_1 = -t\,dt$, i.e. $t = c - x_1$ and

$$x_2 = \delta^2 + x_1^2 - t^2 = 2cx_1 + \delta^2 - c^2.$$

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

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

It covers the whole domain, i.e.

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

Furthermore, 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

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

Therefore, if the points $x^{\pm}$ are on such a segment, we have equality in (2.17) (with $\delta = \varepsilon$).

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

We now write what it means to be on one of our trajectories in terms of the function φ. First, consider the case when the point x^0 is on the boundary, i.e. x_1^0 is arbitrary and $x_2^0 = (x_1^0)^2 + \delta^2$. Let $I = [0,1]$. Split it at point a : $I_- = [0,a]$; $I_+ = [a,1]$. We can choose which point is to the right of x^0. Assume it is x^-. We try to place it on the trajectory $x_2 = x_1^2 + \delta^2$. Then for all $a \in [0,1]$ we have

$$\delta^2 = x_2^- - (x_1^-)^2 = \frac{1}{a}\int_0^a \varphi^2(t)\,dt - \left(\frac{1}{a}\int_0^a \varphi(t)\,dt\right)^2. \tag{2.28}$$

Introducing the function $\psi(a) = \frac{1}{\delta}\int_0^a \varphi(t)\,dt$, we have $\varphi = \delta\psi'$ and, after multiplication by a, (2.28) turns into

$$\begin{aligned}
\int_0^a \psi'(t)^2\,dt &= a + \frac{1}{a}\psi^2(a) \\
\psi'(a)^2 = \frac{d}{da}\left(a + \frac{1}{a}\psi^2(a)\right) &= 1 - \frac{\psi^2(a)}{a^2} + \frac{2\psi(a)\psi'(a)}{a} \\
\left(\psi'(a) - \frac{\psi(a)}{a}\right)^2 &= 1 \\
\psi'(a) - \frac{\psi(a)}{a} &= \pm 1 \\
\left(\frac{\psi(a)}{a}\right)' &= \pm\frac{1}{a} \\
\psi(a) &= \pm a\left(\log\frac{1}{a} + const\right).
\end{aligned}$$

Finally, we have (since we seek to maximize e^{φ}, we choose the plus sign)

$$\varphi(t) = \delta\left(\log\frac{1}{t} - 1\right) + x_1^0, \tag{2.29}$$

where the constant has been chosen so that x_1^0 is the mean value of φ.

Now, 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 $\omega(x_1^-)$ 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, $x_1^+ = x_1^- - \delta$. Then for the splitting parameter $a = \alpha_-$ we have the equation

$$x_1^0 = \alpha_- x_1^- + \alpha_+ x_1^+ = a x_1^- + (1-a)x_1^+ = x_1^+ + a\delta,$$

giving us

$$a = \frac{x_1^0 - x_1^+}{\delta} = 1 - \frac{x_1^- - x_1^0}{\delta} = 1 - \frac{1}{\delta}\sqrt{\delta^2 + (x_1^0)^2 - x_2^0}.$$

Since $x_2^+ = (x_1^+)^2$, φ is constant on I_+, $\varphi|_{I_+} = x_1^+$, and for the interval I_- we use the procedure outlined above, since the point x^- is on the top boundary curve. We only need to renormalize the function from (2.29) to this interval: $\varphi|_{I_-} = \delta\left(\log \frac{a}{t} - 1\right) + x_1^-$. This yields the function we used to prove Lemma 1.

2.4 The dyadic case

2.4.1 Preliminary considerations

We now explicitly construct $\mathbf{B}_\varepsilon^d$, proving Theorem 4 and thus Theorem 2. It is apparent that the proofs given above for the continuous case do not go through in the dyadic case. In particular, Lemma 4 becomes irrelevant, since in the dyadic case one cannot choose the splitting of an interval because it is always split in half, i.e., in the language of Lemma 3 and Lemma 4, $\alpha_- = \alpha_+ = \frac{1}{2}$. This seemingly makes the proof of Lemma 2 impossible to use. We, however, are still able to use the chain of inequalities in Lemma 2 to establish the results in the dyadic case.

A fundamental question that comes to mind first is whether the dyadic Bellman function is different from the continuous one and if so, whether the size of the BMO^d ball in which the John-Nirenberg inequality is valid is the same in both cases (namely, if it is still the unit ball). The following (counter)example answers the former question in the positive and the latter, in the negative.

Fix $\varepsilon > 0$. Let the function φ be defined on $I = (0, 1]$ by

$$\varphi|_{\left(2^{-(k+1)}, 2^{-k}\right]} = (k-1)a, \quad k = 0, 1, \ldots, \tag{2.30}$$

with the constant a to be determined later. We have the following picture for φ

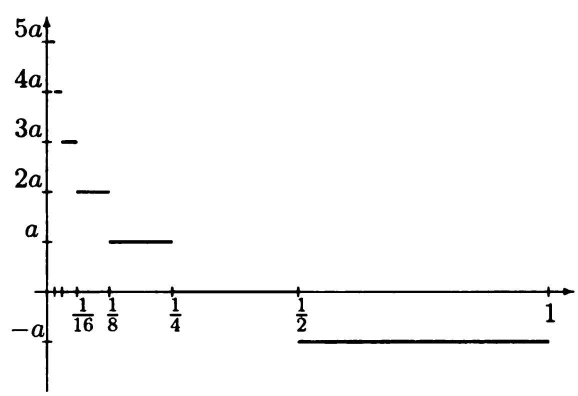


Figure 2.3: The counterexample to the conjecture $\varepsilon_0^d = 1$.

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

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

and

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

where we have used the identities

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

Then

$$\begin{aligned}\|\varphi\|^2_{BMO^d} &= \sup_{J \text{ dyad} \subset I} \left\{\langle\varphi^2\rangle_J - \langle\varphi\rangle_J^2\right\} \\ &= \sup_n \left\{\langle\varphi^2\rangle_{I_n} - \langle\varphi\rangle_{I_n}^2\right\} = \sup_n \left\{a^2(n^2+2) - a^2n^2\right\} = 2a^2.\end{aligned}$$

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

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

The latter sum converges if and only if $e^a < 2$, i.e. $a < \log 2$. In terms of ε_0^d from Theorem 2, we obtain the crucial estimate

$$\varepsilon_0^d \le \sqrt{2}\log 2. \tag{2.31}$$

Informally, we have just shown that the space BMO^d is substantially worse than BMO. However, having developed the Bellman function family (2.13) for the continuous BMO, we would like to look for the dyadic Bellman function within that family. From the example (2.30) above, it is clear that, should the dyadic Bellman function be in fact found in that family, we would have to have

$$\mathbf{B}_\varepsilon^d = B_{\delta(\varepsilon)} \tag{2.32}$$

for some $\delta(\varepsilon) > \varepsilon$. One straightforward approach would be to choose $\delta(\varepsilon)$ large enough so that any straight-line segment $[x^-, x^+]$ with $x^-, x^+ \in \Omega_\varepsilon$ would fit entirely inside $\Omega_{\delta(\varepsilon)}$. Then we would be able to use the chain of inequalities in Lemma 2 without the help of Lemma 4. Let us investigate how large the $\delta(\varepsilon)$ so chosen would be with regard to ε.

Consider the situation when the segment $[x^-, x^+]$ "sticks out" the most. This happens when the middle point $x^0 = \frac{1}{2}(x^- + x^+)$ as well as one of the endpoints, say x^+, are on the top boundary, $x_2 = x_1^2 + \varepsilon^2$, and the other endpoint, x^-, is on the bottom boundary, $x_2 = x_1^2$.

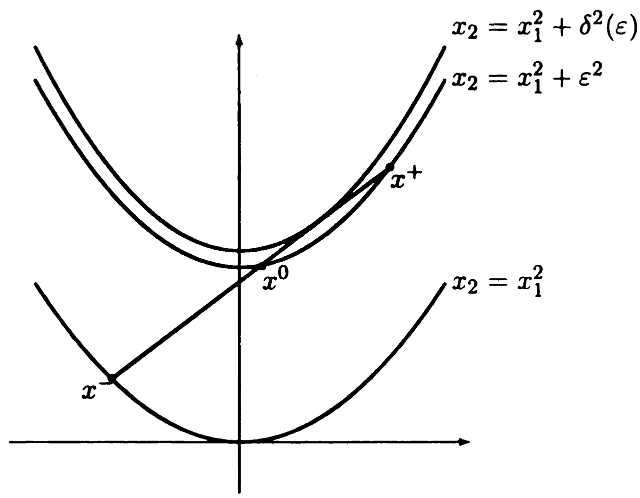


Figure 2.4: The worst case scenario: the largest portion of $[x^-, x^+]$ is outside Ω_ε.

We have

$$x_2^0 = (x_1^0)^2 + \varepsilon^2, \qquad x_2^+ = (x_1^+)^2 + \varepsilon^2, \qquad x_2^0 = (x_1^0)^2 + \varepsilon^2,$$

and

$$x_1^0 = \frac{x_1^- + x_1^+}{2}; \qquad x_2^0 = \frac{x_2^- + x_2^+}{2}.$$

Eliminating x^- and parameterizing everything by x_1^0, we get

$$x_1^+ = x_1^0 + \frac{\varepsilon}{\sqrt{2}}.$$

The equation of the line segment $[x^0, x^+]$ is then

$$x_2 = \left(\frac{\varepsilon}{\sqrt{2}} + 2x_1^0\right) x_1 - \frac{\varepsilon}{\sqrt{2}} x_1^- - (x_1^0)^2 + \varepsilon^2, \quad x_1^0 \le x_1 \le x_1^0 + \frac{\varepsilon}{\sqrt{2}}$$

and the distance between this segment and the top boundary curve is

$$d(x_1) = -x_1^2 + \left(\frac{\varepsilon}{\sqrt{2}} + 2x_1^0\right) x_1 - \frac{\varepsilon}{\sqrt{2}} x_1^0 - (x_1^0)^2, \quad x_1^0 \le x_1 \le x_1^0 + \frac{\varepsilon}{\sqrt{2}},$$

with its maximum $d\left(x_1^0 + \frac{\varepsilon}{2\sqrt{2}}\right) = \frac{\varepsilon^2}{8}$. Therefore, any segment $[x^-, x^+]$ with $x^-, x^+ \in \Omega_\varepsilon$ will lie inside $\Omega_{\frac{3}{2\sqrt{2}}\varepsilon}$. Thus, for any $\varepsilon < \frac{2\sqrt{2}}{3}$ we can run the machine of Lemma 2 to establish that

$$B_{\frac{3}{2\sqrt{2}}\varepsilon}(x) \ge \mathbf{B}_\varepsilon^d(x), \quad \forall x \in \Omega_\varepsilon. \tag{2.33}$$

This gives us the following estimates:

$$\frac{2\sqrt{2}}{3} \le \varepsilon_0^d \le \sqrt{2}\log 2 \tag{2.34}$$

and, provided the dyadic Bellman function can indeed be found inside the family B_δ,

$$\delta(\varepsilon) \le \frac{3}{2\sqrt{2}}\varepsilon. \tag{2.35}$$

In what follows, we are able to prove that the example (2.30) does indeed produce the sharp constant and that the dyadic Bellman function is a member of the family B_δ. Clearly, we need to employ more subtle reasoning than the one above.

2.4.2 Detailed considerations

The following simple lemma shows that the dyadic Bellman function is concave, something that could not be shown directly in the continuous case.

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

$$\mathbf{B}^d_\varepsilon(x) \geq \frac{1}{2}\mathbf{B}^d_\varepsilon(x^-) + \frac{1}{2}\mathbf{B}^d_\varepsilon(x^+). \tag{2.36}$$

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

$$\langle e^{\varphi_n}\rangle_{I_\pm} \longrightarrow \mathbf{B}^d_\varepsilon(x^\pm) \quad \text{as } n \to \infty.$$

We need to check that $\varphi_n \in BMO^d_\varepsilon(I)$. But $BMO^d_\varepsilon(I) = BMO^d_\varepsilon(I_-) \cup BMO^d_\varepsilon(I_+) \cup \{\varphi : \langle\varphi^2\rangle_I - \langle\varphi\rangle_I^2 \leq \varepsilon^2\}$. Since, by assumption, $x \in \Omega_\varepsilon$, we have $\langle\varphi^2\rangle_I - \langle\varphi\rangle_I^2 \leq \varepsilon^2$. 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

$$\mathbf{B}^d_\varepsilon(x) \geq \lim \langle e^{\varphi_n}\rangle_I = \frac{1}{2}\mathbf{B}^d_\varepsilon(x^-) + \frac{1}{2}\mathbf{B}^d_\varepsilon(x^+),$$

which completes the proof. □

Observe that in the continuous case $BMO_\varepsilon(I) \neq BMO_\varepsilon(I_-) \cup BMO_\varepsilon(I_+) \cup \{\varphi : \langle\varphi^2\rangle_I - \langle\varphi\rangle_I^2 \leq \varepsilon^2\}$, 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 $\mathbf{B}_\varepsilon^d$ is concave in Ω_ε. Furthermore, the reasoning of (2.19) still works and we conclude that

$$\mathbf{B}_\varepsilon^d(x) = \exp\left\{x_1 + w(x_2 - x_1^2)\right\} \tag{2.37}$$

for a nonnegative function w such that $w(0) = 0$. What is more, we expect the corresponding matrix $-d^2\mathbf{B}_\varepsilon^d$ (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 (2.13) with $\delta \geq \varepsilon$. This, somewhat heuristic, argument supports our believe that the dyadic Bellman function is a member of that family.

In our desire to use Lemma 2, we have been trying to ensure that the segment $[x^-, x^+]$ lies inside the domain of concavity of a certain function B, so that we can conclude that

$$B(x) \geq \frac{1}{2}B(x^-) + \frac{1}{2}B(x^+) \tag{2.38}$$

and proceed with the unwrapping of the integral sum (2.18). Now, we try to enforce the condition (2.38) directly instead. Since we are searching for $\delta(\varepsilon)$ such that $\mathbf{B}_\varepsilon^d = B_{\delta(\varepsilon)}$, we attempt to solve the extremal problem

$$\delta(\varepsilon) = \min_{\varepsilon<\delta<\min\left\{\frac{3}{2\sqrt{2}}\varepsilon,1\right\}} \left\{\delta:\ B_\delta(x) \geq \frac{1}{2}B_\delta(x^-) + \frac{1}{2}B_\delta(x^+),\right.$$
$$\left. \forall x^-, x^+ \in \Omega_\varepsilon \text{ such that } x = \frac{x^- + x^+}{2} \in \Omega_\varepsilon\right\}, \tag{2.39}$$

where we have used (2.35). Recalling the definition (2.13) of B_δ, we write the "concavity" condition as

$$\frac{1-\sqrt{\delta^2+x_1^2-x_2}}{1-\delta}\exp\left(x_1+\sqrt{\delta^2+x_1^2-x_2}-\delta\right)$$

$$\geq \frac{1}{2}\frac{1-\sqrt{\delta^2+(x_1^-)^2-x_2^-}}{1-\delta}\exp\left(x_1^-+\sqrt{\delta^2+(x_1^-)^2-x_2^-}-\delta\right) \tag{2.40}$$

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

Let

$$\alpha=\sqrt{\delta^2+x_1^2-x_2},\ \alpha_\pm=\sqrt{\delta^2+\left(x_1^\pm\right)^2-x_2^\pm},\ \theta=\frac{1}{2}(x_1^--x_1^+).$$

Using this notation and multiplying (2.40) by $2(1-\delta)e^{-x_1+\delta}$, we can rewrite it as

$$2(1-\alpha)e^\alpha \geq (1-\alpha_-)e^{\theta+\alpha_-}+(1-\alpha_+)e^{-\theta+\alpha_+}.$$

A straightforward calculation shows that $\alpha_-^2+\alpha_+^2=2\alpha^2+2\theta^2$. The condition $x, x_-, x_+ \in \Omega_\varepsilon$ can be rewritten as $\alpha,\alpha_-,\alpha_+\in[\sqrt{\delta^2-\varepsilon^2},\delta]$. Finally, let

$$f(\alpha,\alpha_-,\alpha_+,\theta)=2(1-\alpha)e^\alpha-(1-\alpha_-)e^{\theta+\alpha_-}-(1-\alpha_+)e^{-\theta+\alpha_+} \tag{2.41}$$

We will solve the extremal problem (2.39) as the following two-stage problem:

For $0<\varepsilon\leq\delta<\min\left\{\frac{3}{2\sqrt{2}}\varepsilon,1\right\}$, let

$$\begin{aligned} S_{\delta,\varepsilon} &= \left\{(a,b,c,d)\in\mathbb{R}^4:\ a,b,c\in[\sqrt{\delta^2-\varepsilon^2},\delta];\ b^2+c^2=2a^2+2d^2\right\} \\ g(\delta,\varepsilon) &= \min\left\{f(\alpha,\alpha_-,\alpha_+,\theta):(\alpha,\alpha_-,\alpha_+,\theta)\in S_{\delta,\varepsilon}\right\}, \end{aligned} \tag{2.42}$$

$$\delta(\varepsilon)=\min\{\delta: g(\delta,\varepsilon)\geq 0\}. \tag{2.43}$$

While it is not a priori obvious that the set $\{\delta : g(\delta,\varepsilon)\geq 0\}$ is not empty, in what follows we show that the problem (2.42), (2.43) has a solution (unique, by our formulation).

2.4.3 Stage 1

We will often refer to finding the solution of (2.42) as optimization over the cube $[\sqrt{\delta^2-\varepsilon^2},\delta]^3$, even though $S_{\delta,\varepsilon}$ is in 4-space. We will use Lagrange multipliers in the interior and on the faces of the cube and straightforward one-variable optimization on the edges. We make extensive use of the inherent symmetry of the problem. In addition, since we are only interested in the possible negative values of f, we will disregard any and all other possible minimums.

Interior of the cube

Let

$$H(\alpha,\alpha_-,\alpha_+,\theta,\lambda) = 2(1-\alpha)e^{\alpha}-(1-\alpha_-)e^{\theta+\alpha_-}-(1-\alpha_+)e^{-\theta+\alpha_+}-\lambda(\alpha_-^2+\alpha_+^2-2\alpha^2-2\theta^2).$$

From $\nabla H = 0$ we get

$$\begin{aligned}
-2\alpha e^{\alpha}+4\alpha\lambda &= 0\\
\alpha_- e^{\theta+\alpha_-}-2\alpha_-\lambda &= 0\\
\alpha_+ e^{-\theta+\alpha_+}-2\alpha_+\lambda &= 0\\
-(1-\alpha_-)e^{\theta+\alpha_-}+(1-\alpha_+)e^{-\theta+\alpha_+}+4\theta\lambda &= 0\\
\alpha_-^2+\alpha_+^2-2\alpha^2-2\theta^2 &= 0.
\end{aligned}$$

Since no α can be zero, we get

$$e^{\alpha}=e^{\alpha_-}=e^{\alpha_+}=2\lambda, \quad \text{so} \quad \alpha_-+\alpha_+=2\alpha,$$

and so at any possible point of minimum in the interior

$$f = 2(1-\alpha)e^{\alpha} - (1-\alpha_-)e^{\alpha} - (1-\alpha_+)e^{\alpha} = e^{\alpha}[-2\alpha + \alpha_- + \alpha_+] = 0.$$

Faces

There are six faces. We make the following observations:

- if $\alpha = \delta$, then $\alpha_- = \alpha_+ = \delta$ and $\theta = 0$. Hence, the intersection of the domain $S_{\delta,\varepsilon}$ and the interior of this face is empty;
- because of the symmetry in α_-, α_+ (θ can be ≤ 0 or ≥ 0), it suffices to consider only the faces $\alpha_+ = \delta$ and $\alpha_+ = \sqrt{\delta^2 - \varepsilon^2}$. The extremum points (if any) will have their "twins" on the faces $\alpha_- = \delta$ and $\alpha_- = \sqrt{\delta^2 - \varepsilon^2}$.

Thus the only faces we need to consider are $\alpha = \sqrt{\delta^2 - \varepsilon^2}$, $\alpha_+ = \delta$, and $\alpha_+ = \sqrt{\delta^2 - \varepsilon^2}$.

1. **Face** $\alpha = \sqrt{\delta^2 - \varepsilon^2}$

Let

$$h_1(\alpha_-, \alpha_+, \theta, \lambda) = H(\sqrt{\delta^2 - \varepsilon^2}, \alpha_-, \alpha_+, \theta, \lambda) = 2(1 - \sqrt{\delta^2 - \varepsilon^2})e^{\sqrt{\delta^2 - \varepsilon^2}}$$
$$-(1-\alpha_-)e^{\theta+\alpha_-} - (1-\alpha_+)e^{-\theta+\alpha_+} - \lambda(\alpha_-^2 + \alpha_+^2 - 2(\delta^2 - \varepsilon^2) - 2\theta^2).$$

From $\nabla h_1 = 0$ we get

$$\alpha_- e^{\theta+\alpha_-} - 2\alpha_-\lambda = 0$$
$$\alpha_+ e^{-\theta+\alpha_+} - 2\alpha_+\lambda = 0$$
$$-(1-\alpha_-)e^{\theta+\alpha_-} + (1-\alpha_+)e^{-\theta+\alpha_+} + 4\theta\lambda = 0$$
$$\alpha_-^2 + \alpha_+^2 = 2(\delta^2 - \varepsilon^2) + 2\theta^2.$$

Therefore, $e^{\theta+\alpha_-} = e^{-\theta+\alpha_+} = 2\lambda$ and $\alpha_+ = \alpha_- + 2\theta$. Plugging this into the last equation, we get

$$(\alpha_- + \theta)^2 = \delta^2 - \varepsilon^2, \quad \alpha_- = -\theta + \sqrt{\delta^2 - \varepsilon^2}, \quad \alpha_+ = \theta + \sqrt{\delta^2 - \varepsilon^2},$$

where the plus sign is chosen in the square root to ensure that each α is positive. Lastly, we obtain at any possible extremum point,

$$h_1 = 2(1 - \sqrt{\delta^2 - \varepsilon^2})e^{\sqrt{\delta^2-\varepsilon^2}} - 2(1 - \sqrt{\delta^2 - \varepsilon^2})e^{\sqrt{\delta^2-\varepsilon^2}} = 0.$$

2. **Face** $\alpha_+ = \delta$

Let

$$\begin{aligned} &h_2(\alpha, \alpha_-, \theta, \lambda) = H(\alpha, \alpha_-, \delta, \theta, \lambda) \\ &= 2(1-\alpha)e^{\alpha} - (1-\alpha_-)e^{\theta+\alpha_-} - (1-\delta)e^{-\theta+\delta} - \lambda(\alpha_-^2 + \delta^2 - 2\alpha^2 - 2\theta^2). \end{aligned}$$

From $\nabla h_2 = 0$ we get

$$\begin{aligned} e^{\alpha} = e^{\theta+\alpha_-} &= 2\lambda \\ -(1-\alpha_-)e^{\theta+\alpha_-} + (1-\delta)e^{-\theta+\delta} + 4\theta\lambda &= 0 \\ \alpha_-^2 + \delta^2 &= 2\alpha^2 + 2\theta^2, \end{aligned}$$

from which we get $\alpha = \theta + \alpha_-$, $(1-\delta)e^{-\theta+\delta} = (1 - \alpha_- - 2\theta)e^{\theta+\alpha_-}$, and $\delta^2 = (\alpha_- + 2\theta)^2$. If $\alpha_- = -2\theta + \delta$, then we get $(1-\delta)e^{\delta} = (1+\delta)e^{-\delta}$, which only has the trivial solution $\delta = 0$. So $\alpha_- = 2\theta + \delta$, $\alpha = -\theta + \delta$ and, at any possible extremum point,

$$h_2 = 2(1 + \theta - \delta)e^{-\theta+\delta} - (1 + 2\theta - \delta)e^{-\theta+\delta} - (1-\delta)e^{-\theta+\delta} = 0.$$

3. **Face** $\alpha_+ = \sqrt{\delta^2 - \varepsilon^2}$

Let

$$h_3(\alpha, \alpha_-, \theta, \lambda) = H(\alpha, \alpha_-, \sqrt{\delta^2 - \varepsilon^2}, \theta, \lambda) = 2(1-\alpha)e^{\alpha} - (1-\alpha_-)e^{\theta+\alpha_-}$$
$$- (1 - \sqrt{\delta^2 - \varepsilon^2})e^{-\theta+\sqrt{\delta^2-\varepsilon^2}} - \lambda(\alpha_-^2 + (\delta^2 - \varepsilon^2) - 2\alpha^2 - 2\theta^2).$$

We observe that this case is identical with the previous one if $\sqrt{\delta^2 - \varepsilon^2}$ is substituted for δ. Therefore, at any possible extremum point, $h_3 = 0$.

Edges

We have a total of 12 edges:

(1) $\alpha = \sqrt{\delta^2 - \varepsilon^2},\ \alpha_+ = \delta$

(2) $\alpha = \sqrt{\delta^2 - \varepsilon^2},\ \alpha_+ = \sqrt{\delta^2 - \varepsilon^2}$

(3) $\alpha = \sqrt{\delta^2 - \varepsilon^2},\ \alpha_- = \delta$

(4) $\alpha = \sqrt{\delta^2 - \varepsilon^2},\ \alpha_- = \sqrt{\delta^2 - \varepsilon^2}$

(5) $\alpha = \delta,\ \alpha_+ = \delta$

(6) $\alpha = \delta,\ \alpha_+ = \sqrt{\delta^2 - \varepsilon^2}$

(7) $\alpha = \delta,\ \alpha_- = \delta$

(8) $\alpha = \delta,\ \alpha_- = \sqrt{\delta^2 - \varepsilon^2}$

(9) $\alpha_+ = \delta,\ \alpha_- = \delta$

(10) $\alpha_+ = \delta,\ \alpha_- = \sqrt{\delta^2 - \varepsilon^2}$

(11) $\alpha_+ = \sqrt{\delta^2 - \varepsilon^2},\ \alpha_- = \delta$

(12) $\alpha_+ = \sqrt{\delta^2 - \varepsilon^2},\ \alpha_- = \sqrt{\delta^2 - \varepsilon^2}$

Edges (5) through (8) have $\alpha = \delta$, which implies $\alpha_- = \alpha_+ = \delta$, $\theta = 0$. The pairs (1)-(3), (2)-(4), and (10)-(11) are symmetric. On edge (12), using the fact that $\alpha_-^2 + \alpha_+^2 = 2\alpha^2 + 2\theta^2$, we get $\alpha^2 = \delta^2 - \varepsilon^2 - \theta^2$, giving $\theta = 0$, $\alpha = \alpha_- = \alpha_+$ and $f = 0$. This leaves us with the following four (renumbered) edges to consider (out of the symmetric pairs we chose the edges sharing a vertex):

$$\langle 1\rangle\ \alpha_+ = \delta,\ \alpha_- = \delta$$
$$\langle 2\rangle\ \alpha_+ = \sqrt{\delta^2-\varepsilon^2},\ \alpha_- = \delta$$
$$\langle 3\rangle\ \alpha = \sqrt{\delta^2-\varepsilon^2},\ \alpha_- = \delta$$
$$\langle 4\rangle\ \alpha = \sqrt{\delta^2-\varepsilon^2},\ \alpha_+ = \sqrt{\delta^2-\varepsilon^2}$$

1. **Edge** $\alpha_+ = \delta,\ \alpha_- = \delta$

Since $\alpha_-^2 + \alpha_+^2 = 2\alpha^2 + 2\theta^2$, we have $\alpha = \sqrt{\delta^2-\theta^2}$, $-\varepsilon \le \theta \le \varepsilon$. The function to minimize is, therefore,

$$F(\theta) = 2(1-\sqrt{\delta^2-\varepsilon^2})e^{\sqrt{\delta^2-\varepsilon^2}} - (1-\delta)e^{\delta}(e^{\theta}+e^{-\theta}).$$

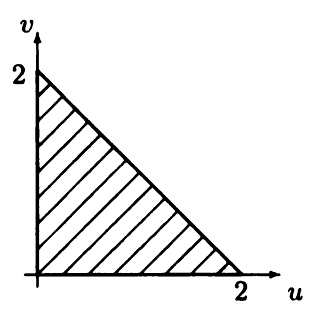
We only need to show that $F(\theta_0) \ge 0$ at any possible extremum point $\theta_0 \in (-\varepsilon,\varepsilon)$. We will, however, show more; namely, that

$$F(\theta) \ge 0,\ \forall\theta \in [-\delta,\delta],\ \forall\delta \in [0,1].$$

Let $u = \delta+\theta$, $v = \delta-\theta$. Then, slightly abusing notation, we have

$$F(u,v) = 2(1-\sqrt{uv})e^{\sqrt{uv}} - \left(1-\frac{u+v}{2}\right)(e^u+e^v)\,,\ u+v = 2\delta,\ 0 \le u,v \le 2\delta.$$

We will accomplish our goal if we show that $F(u,v) \ge 0$, $0 \le u \le 2$, $0 \le v \le 2-u$.



Every point (u, v) of minimum inside the region in the picture would have to satisfy the equations

$$F_u = -ve^{\sqrt{uv}} + \frac{1}{2}(e^u + e^v) - \left(1 - \frac{u+v}{2}\right)e^u = 0$$
$$F_v = -ue^{\sqrt{uv}} + \frac{1}{2}(e^u + e^v) - \left(1 - \frac{u+v}{2}\right)e^v = 0.$$

Adding the two equations, we get

$$\frac{e^u + e^v}{2} = e^{\sqrt{uv}}. \tag{2.44}$$

Subtracting the second equation from the first and using (2.44), we get

$$e^v(1-v) = e^u(1-u),$$

which implies (since $u, v \geq 0$) that $u = v$. The value of F on any possible extremum point in the interior is then $F(u, u) = 0$.

On the boundary $v = 0$, $0 \leq u \leq 2$, we have $F(u, 0) = 2 - \left(1 - \frac{u}{2}\right)(e^u + 1)$. The equation $[F(u, 0)]' = 0$ gives $e^u(1-u) = 1$. Since $u \geq 0$, we conclude that $u = 0$. Then $F(0, 0) = 0$. The case $u = 0$ is identical.

If $u + v = 2$, we have $F(u, v) = F(u, 2-u) = 2(1 - \sqrt{u(2-u)})e^{\sqrt{u(2-u)}}$. Since $0 \leq u(2-u) = 1 - (u-1)^2 \leq 1$, and the function $s \mapsto e^s(1-s)$ is nonnegative for $s \in [0, 1]$, we conclude that $F \geq 0$ on this piece of the boundary. This completes the consideration of this edge.

2. **Edge** $\alpha_+ = \sqrt{\delta^2 - \varepsilon^2}$, $\alpha_- = \delta$

Since $\alpha_-^2 + \alpha_+^2 = 2\alpha^2 + 2\theta^2$, we have $\alpha = \sqrt{\delta^2 - \frac{\varepsilon^2}{2} - \theta^2}$, $-\frac{\varepsilon}{\sqrt{2}} \leq \theta \leq \frac{\varepsilon}{\sqrt{2}}$. The function to minimize is, therefore,

$$F(\theta) = 2\left(1 - \sqrt{\delta^2 - \frac{\varepsilon^2}{2} - \theta^2}\right) e^{\sqrt{\delta^2 - \frac{\varepsilon^2}{2} - \theta^2}} - (1-\delta)e^{\delta+\theta} - \left(1 - \sqrt{\delta^2 - \varepsilon^2}\right) e^{\sqrt{\delta^2-\varepsilon^2} - \theta}.$$

Assume that F has a local minimum for some $\theta_0 \in \left(-\frac{\varepsilon}{\sqrt{2}}, \frac{\varepsilon}{\sqrt{2}}\right)$. Then $F'(\theta_0) = 0$, i.e.

$$2\theta_0 e^{\sqrt{\delta^2 - \frac{\varepsilon^2}{2} - \theta_0^2}} - (1-\delta)e^{\delta+\theta_0} + \left(1 - \sqrt{\delta^2 - \varepsilon^2}\right) e^{\sqrt{\delta^2-\varepsilon^2} - \theta_0} = 0. \tag{2.45}$$

Expressing $(1 - \sqrt{\delta^2 - \varepsilon^2})\, e^{\sqrt{\delta^2-\varepsilon^2} - \theta_0}$ from (2.45), we get

$$F(\theta_0) = 2e^{\theta_0}\left[\left(1 - \left(\sqrt{\delta^2 - \frac{\varepsilon^2}{2} - \theta_0^2} - \theta_0\right)\right) e^{\sqrt{\delta^2 - \frac{\varepsilon^2}{2} - \theta_0^2} - \theta_0} - (1-\delta)e^{\delta}\right]. \tag{2.46}$$

Observe that $(1-x)e^x - (1-y)e^y \geq 0$ if $y \geq 0$ and $-y \leq x \leq y$. Together with (2.46), this consideration implies that if

$$-\delta \leq \sqrt{\delta^2 - \frac{\varepsilon^2}{2} - \theta_0^2} - \theta_0 \leq \delta, \tag{2.47}$$

then $F(\theta_0) \geq 0$. We now solve the inequality (2.47).

First, since $|\theta_0| \leq \varepsilon/\sqrt{2} \leq \delta$, we have that

$$\sqrt{\delta^2 - \frac{\varepsilon^2}{2} - \theta_0^2} - \theta_0 \geq -\delta, \ \forall \theta_0 \in \left[-\frac{\varepsilon}{\sqrt{2}}, \frac{\varepsilon}{\sqrt{2}}\right].$$

Secondly, solving the (right-hand) inequality (2.47), we obtain that

$$\sqrt{\delta^2 - \frac{\varepsilon^2}{2} - \theta_0^2} - \theta_0 \leq \delta$$

if and only if

$$\theta_0 \geq \frac{-\delta + \sqrt{\delta^2 - \varepsilon^2}}{2} \quad \text{or} \quad \theta_0 \leq \frac{-\delta - \sqrt{\delta^2 - \varepsilon^2}}{2}.$$

We must check whether $-\frac{\varepsilon}{\sqrt{2}} \leq \frac{-\delta-\sqrt{\delta^2-\varepsilon^2}}{2}$. This inequality is equivalent to

$$\sqrt{\delta^2-\varepsilon^2} \leq \sqrt{2}\varepsilon - \delta.$$

Since $\varepsilon \geq \frac{2\sqrt{2}}{3}\delta \geq \frac{1}{\sqrt{2}}\delta$, we can square the last inequality, getting as a result

$$\varepsilon \geq \frac{2\sqrt{2}}{3}\delta$$

again, which is true by the formulation of our extremal problem (2.42)–(2.43). Finally, we conclude that (2.47) holds if and only if

$$\frac{-\delta+\sqrt{\delta^2-\varepsilon^2}}{2} \leq \theta_0 \leq \frac{\varepsilon}{\sqrt{2}} \quad \text{or} \quad -\frac{\varepsilon}{\sqrt{2}} \leq \theta_0 \leq \frac{-\delta-\sqrt{\delta^2-\varepsilon^2}}{2}. \tag{2.48}$$

We now try and fill the gap in the condition (2.48). Expressing $(1-\delta)e^{\delta+\theta_0}$ from (2.45), we get

$$F(\theta_0) = 2e^{-\theta_0}\left[\left(1-\left(\sqrt{\delta^2-\tfrac{\varepsilon^2}{2}-\theta_0^2}+\theta_0\right)\right)e^{\sqrt{\delta^2-\frac{\varepsilon^2}{2}-\theta_0^2}+\theta_0} - \left(1-\sqrt{\delta^2-\varepsilon^2}\right)e^{\sqrt{\delta^2-\varepsilon^2}}\right]. \tag{2.49}$$

Similarly to (2.47), if

$$-\sqrt{\delta^2-\varepsilon^2} \leq \sqrt{\delta^2-\frac{\varepsilon^2}{2}-\theta_0^2}+\theta_0 \leq \sqrt{\delta^2-\varepsilon^2}, \tag{2.50}$$

then $F(\theta_0) \geq 0$. Carefully solving this inequality, we obtain that (2.50) holds if and only if

$$\frac{-\delta-\sqrt{\delta^2-\varepsilon^2}}{2} \leq \theta_0 \leq \frac{-\delta+\sqrt{\delta^2-\varepsilon^2}}{2}. \tag{2.51}$$

To summarize, the fact that $\theta_0 \in \left(-\frac{\varepsilon}{\sqrt{2}}, \frac{\varepsilon}{\sqrt{2}}\right)$ implies either (2.48) or (2.51), which, in turn, imply (2.47) or (2.50), respectively. Either one of the latter inequalities implies that $F(\theta_0) \geq 0$. This completes the consideration of this edge.

3. **Edge** $\alpha = \sqrt{\delta^2 - \varepsilon^2}$, $\alpha_- = \delta$

Since $\alpha_-^2 + \alpha_+^2 = 2\alpha^2 + 2\theta^2$, we have $\alpha_+ = \sqrt{\delta^2 - 2\varepsilon^2 + 2\theta^2}$, $\frac{\varepsilon}{\sqrt{2}} \leq |\theta| \leq \varepsilon$. The function to minimize is, therefore,

$$F(\theta) = 2\left(1 - \sqrt{\delta^2 - \varepsilon^2}\right) e^{\sqrt{\delta^2 - \varepsilon^2}} - (1-\delta)e^{\delta+\theta} - \left(1 - \sqrt{\delta^2 - 2\varepsilon^2 + 2\theta^2}\right) e^{\sqrt{\delta^2 - 2\varepsilon^2 + 2\theta^2} - \theta}.$$

We seek the absolute minimum of F. First, we observe that we only need to consider $\theta < 0$. Indeed, since $0 \leq \sqrt{\delta^2 - 2\varepsilon^2 + 2\theta^2} \leq \delta$, we have

$$(1-\delta)e^{\delta} \leq \left(1 - \sqrt{\delta^2 - 2\varepsilon^2 + 2\theta^2}\right) e^{\sqrt{\delta^2 - 2\varepsilon^2 + 2\theta^2}}.$$

So, if $\theta \geq 0$, then

$$\begin{aligned}(1-\delta)e^{\delta+\theta} &+ \left(1 - \sqrt{\delta^2 - 2\varepsilon^2 + 2\theta^2}\right) e^{\sqrt{\delta^2 - 2\varepsilon^2 + 2\theta^2} - \theta} \\ &\leq (1-\delta)e^{\delta-\theta} + \left(1 - \sqrt{\delta^2 - 2\varepsilon^2 + 2\theta^2}\right) e^{\sqrt{\delta^2 - 2\varepsilon^2 + 2\theta^2} + \theta}.\end{aligned}$$

$$\begin{aligned}&2\left(1 - \sqrt{\delta^2 - \varepsilon^2}\right) e^{\sqrt{\delta^2 - \varepsilon^2}} - (1-\delta)e^{\delta+\theta} - \left(1 - \sqrt{\delta^2 - 2\varepsilon^2 + 2\theta^2}\right) e^{\sqrt{\delta^2 - 2\varepsilon^2 + 2\theta^2} - \theta} \\ &\geq 2\left(1 - \sqrt{\delta^2 - \varepsilon^2}\right) e^{\sqrt{\delta^2 - \varepsilon^2}} - (1-\delta)e^{\delta-\theta} - \left(1 - \sqrt{\delta^2 - 2\varepsilon^2 + 2\theta^2}\right) e^{\sqrt{\delta^2 - 2\varepsilon^2 + 2\theta^2} + \theta},\end{aligned}$$

i.e. $F(\theta) \geq F(-\theta)$. Therefore,

$$\min_{\frac{\varepsilon}{\sqrt{2}} \leq |\theta| \leq \varepsilon} F(\theta) = \min_{-\varepsilon \leq \theta \leq -\frac{\varepsilon}{\sqrt{2}}} F(\theta).$$

We now look for the minimum of F over the interval $-\varepsilon \leq \theta \leq -\frac{\varepsilon}{\sqrt{2}}$. We have

$$F'(\theta) = e^{\theta}\left[\left(1 - \left(\sqrt{\delta^2 - 2\varepsilon^2 + 2\theta^2} - 2\theta\right)\right) e^{\sqrt{\delta^2 - 2\varepsilon^2 + 2\theta^2} - 2\theta} - (1-\delta)e^{\delta}\right].$$

We claim that $F'(\theta) \leq 0, \ -\varepsilon \leq \theta \leq -\frac{\varepsilon}{\sqrt{2}}$. Indeed, since $\theta \leq \frac{\varepsilon}{\sqrt{2}}$ and $\varepsilon \geq \frac{2\sqrt{2}}{3}\delta$, we have $2\theta \leq -\sqrt{2}\varepsilon \leq -\frac{4}{3}\delta$. Thus $\delta + 2\theta \leq 0$ and $\sqrt{\delta^2 - 2\varepsilon^2 + 2\theta^2} - 2\theta \geq \delta$. Since the function $s \mapsto (1-s)e^s$ is decreasing for $s \geq 0$, we conclude that $F'(0) \leq 0$.

Finally, since there are no extremum points in the interior of the interval, and $F'(\theta) \leq 0$, we conclude that

$$\min_{\frac{\varepsilon}{\sqrt{2}} \leq |\theta| \leq \varepsilon} F(\theta) = \min_{-\varepsilon \leq \theta \leq -\frac{\varepsilon}{\sqrt{2}}} F(\theta) = F\left(-\frac{\varepsilon}{\sqrt{2}}\right),$$

i.e. the minimum of f over this edge is attained at a vertex.

4. **Edge** $\alpha = \sqrt{\delta^2 - \varepsilon^2}, \ \alpha_+ = \sqrt{\delta^2 - \varepsilon^2}$

Since $\alpha_-^2 + \alpha_+^2 = 2\alpha^2 + 2\theta^2$, we have $\alpha_- = \sqrt{\delta^2 - \varepsilon^2 + 2\theta^2}, \ -\frac{\varepsilon}{\sqrt{2}} \leq |\theta| \leq \frac{\varepsilon}{\sqrt{2}}$. The function to minimize is, therefore,

$$\begin{aligned} F(\theta) = 2\left(1 - \sqrt{\delta^2 - \varepsilon^2}\right) e^{\sqrt{\delta^2 - \varepsilon^2}} - \left(1 - \sqrt{\delta^2 - \varepsilon^2 + 2\theta^2}\right) e^{\sqrt{\delta^2 - \varepsilon^2 + 2\theta^2} + \theta} \\ - \left(1 - \sqrt{\delta^2 - \varepsilon^2}\right) e^{\sqrt{\delta^2 - \varepsilon^2} - \theta}. \end{aligned}$$

As in the previous case, we only need to consider $\theta < 0$. Indeed, assume that $\theta \geq 0$. Since

$$\left(1 - \sqrt{\delta^2 - \varepsilon^2 + 2\theta^2}\right) e^{\sqrt{\delta^2 - \varepsilon^2 + 2\theta^2}} \leq \left(1 - \sqrt{\delta^2 - \varepsilon^2}\right) e^{\sqrt{\delta^2 - \varepsilon^2}},$$

we have

$$\begin{aligned} &\left(1 - \sqrt{\delta^2 - \varepsilon^2 + 2\theta^2}\right) e^{\sqrt{\delta^2 - \varepsilon^2 + 2\theta^2} + \theta} + \left(1 - \sqrt{\delta^2 - \varepsilon^2}\right) e^{\sqrt{\delta^2 - \varepsilon^2} - \theta} \\ &\quad \leq \left(1 - \sqrt{\delta^2 - \varepsilon^2 + 2\theta^2}\right) e^{\sqrt{\delta^2 - \varepsilon^2 + 2\theta^2} - \theta} + \left(1 - \sqrt{\delta^2 - \varepsilon^2}\right) e^{\sqrt{\delta^2 - \varepsilon^2} + \theta}. \end{aligned}$$

Therefore, as above $F(\theta) \geq F(-\theta)$ and we have

$$\min_{-\frac{\varepsilon}{\sqrt{2}} \leq \theta \leq \frac{\varepsilon}{\sqrt{2}}} F(\theta) = \min_{-\frac{\varepsilon}{\sqrt{2}} \leq \theta \leq 0} F(\theta).$$

Assume now that there exists a local extremum θ_0 in this interval. Then $F'(\theta_0) = 0$. This gives

$$\left(1 - \sqrt{\delta^2 - \varepsilon^2}\right) e^{\sqrt{\delta^2 - \varepsilon^2} - \theta_0} = \left(1 - \sqrt{\delta^2 - \varepsilon^2 + 2\theta_0^2} - 2\theta_0\right) e^{\sqrt{\delta^2 - \varepsilon^2 + 2\theta_0^2} + \theta_0}$$

and

$$\begin{aligned} F(\theta_0) = 2e^{\theta_0} \Bigg[& \left(1 - \sqrt{\delta^2 - \varepsilon^2}\right) e^{\sqrt{\delta^2 - \varepsilon^2}} \\ & - \left(1 - \left(\sqrt{\delta^2 - \varepsilon^2 + 2\theta_0^2} + \theta_0\right)\right) e^{\sqrt{\delta^2 - \varepsilon^2 + 2\theta_0^2} + \theta_0} \Bigg]. \end{aligned}$$

We know that if $\sqrt{\delta^2 - \varepsilon^2 + 2\theta_0^2} + \theta_0 \geq \sqrt{\delta^2 - \varepsilon^2}$, then $F(\theta_0) \geq 0$. Solving this inequality we obtain

$$|\theta_0| \geq 2\sqrt{\delta^2 - \varepsilon^2}.$$

What to do if $|\theta_0| < 2\sqrt{\delta^2 - \varepsilon^2}$? Assume that is the case. Here we have to consider the behavior of the derivative. We have

$$F'(\theta) = e^{-\theta} \left[\left(1 - \sqrt{\delta^2 - \varepsilon^2}\right) e^{\sqrt{\delta^2 - \varepsilon^2}} - \left(1 - \left(\sqrt{\delta^2 - \varepsilon^2 + 2\theta^2} + 2\theta\right)\right) e^{\sqrt{\delta^2 - \varepsilon^2 + 2\theta^2} + 2\theta} \right].$$

Recall that $(1 - x)e^x - (1 - y)e^y \geq 0$ if $y \geq 0$ and $-y \leq x \leq y$. Thus if

$$-\sqrt{\delta^2 - \varepsilon^2} < \sqrt{\delta^2 - \varepsilon^2 + 2\theta^2} + 2\theta < \sqrt{\delta^2 - \varepsilon^2},$$

then $F'(\theta) < 0$. The left-hand inequality translates into $0 < |\theta| < 2\sqrt{\delta^2 - \varepsilon^2}$, while the right-hand one, into $|\theta| > 0$.

To summarize: we have $F'(\theta) < 0$ if $-2\sqrt{\delta^2 - \varepsilon^2} < \theta < 0$, thus there are no points of local extremum on this subinterval; if there is an extremum point $\theta_0 \leq -2\sqrt{\delta^2 - \varepsilon^2}$, then $F(\theta_0) \geq 0$. Altogether, the two "interesting" points where the absolute minimum may be attained are $\theta = 0$ and $\theta = \frac{\varepsilon}{\sqrt{2}}$, both corresponding to some vertices of the cube, to be considered in the next subsection. This completes the consideration of the last edge.

Vertices

We formally have a total of eight vertices, but some are non-existent and only two give us nontrivial results.

1. As mentioned before, the set $S_{\delta,\varepsilon}$ intersects the face $\alpha = \delta$ at one point only, the vertex $\alpha = \alpha_- = \alpha_+ = \delta$ where we have $f = 0$.

2. The vertex $\alpha = \alpha_- = \alpha_+ = \sqrt{\delta^2 - \varepsilon^2}$ also gives $\theta = 0$ and $f = 0$.

3. The vertex $\alpha = \sqrt{\delta^2 - \varepsilon^2}$, $\alpha_- = \alpha_+ = \delta$ was considered above as a part of the edge $\alpha_- = \alpha_+ = \delta$ on which we have $f \geq 0$.

4. The vertex $\alpha = \alpha_- = \sqrt{\delta^2 - \varepsilon^2}$, $\alpha_+ = \delta$ gives $\theta = \pm\frac{\varepsilon}{\sqrt{2}}$. Then

$$f|_{\theta=\frac{\varepsilon}{\sqrt{2}}} = 2\left(1 - \sqrt{\delta^2 - \varepsilon^2}\right) e^{\sqrt{\delta^2-\varepsilon^2}} - \left(1 - \sqrt{\delta^2 - \varepsilon^2}\right) e^{\sqrt{\delta^2-\varepsilon^2}+\varepsilon/\sqrt{2}} - (1-\delta)e^{\delta-\varepsilon/\sqrt{2}}$$

and

$$f|_{\theta=-\frac{\varepsilon}{\sqrt{2}}} = 2\left(1 - \sqrt{\delta^2 - \varepsilon^2}\right) e^{\sqrt{\delta^2-\varepsilon^2}} - \left(1 - \sqrt{\delta^2 - \varepsilon^2}\right) e^{\sqrt{\delta^2-\varepsilon^2}-\varepsilon/\sqrt{2}} - (1-\delta)e^{\delta+\varepsilon/\sqrt{2}}.$$

Since $\left(1 - \sqrt{\delta^2 - \varepsilon^2}\right) e^{\sqrt{\delta^2-\varepsilon^2}} \geq (1 - \delta)e^{\delta}$, we have

$$f|_{\theta=\frac{\varepsilon}{\sqrt{2}}} \leq f|_{\theta=-\frac{\varepsilon}{\sqrt{2}}}.$$

5. The vertex $\alpha = \alpha_+ = \sqrt{\delta^2 - \varepsilon^2}$, $\alpha_- = \delta$ is symmetric to the previous one and gives the same result.

We have thus solved the first part of the extremal problem (2.42)–(2.43) in the sense that the only possible nontrivial (meaning negative) minimum the function f can have on $S_{\delta,\varepsilon}$ is given by

$$g(\delta,\varepsilon)=\left(1-\sqrt{\delta^2-\varepsilon^2}\right)e^{\sqrt{\delta^2-\varepsilon^2}}\left(2-e^{\varepsilon/\sqrt{2}}\right)-(1-\delta)e^{\delta-\varepsilon/\sqrt{2}}. \tag{2.52}$$

2.4.4 Stage 2

We are now in a position to solve the minimization problem (2.43)

$$\delta(\varepsilon)=\min\left\{\delta\in\left(\varepsilon,\min\left\{\frac{3}{2\sqrt{2}}\varepsilon,1\right\}\right)\;:\;g(\delta,\varepsilon)\geq 0\right\}. \tag{2.53}$$

Differentiating g with respect to δ, we get

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

We observe that $g_\delta>0$. Indeed, checking if $\delta-\frac{\varepsilon}{\sqrt{2}}>\sqrt{\delta^2-\varepsilon^2}$, we obtain that this condition is equivalent to $\delta<\frac{3}{2\sqrt{2}}\varepsilon$, hence it is satisfied. Then

$$g_\delta(\delta,\varepsilon)\geq\delta e^{\sqrt{\delta^2-\varepsilon^2}}\left[-1+e^{\varepsilon/\sqrt{2}}\right]>0.$$

Therefore, if the equation $g(\delta,\varepsilon)=0$ has a solution, then it is unique and solves our extremal problem. We thus look for a solution of the equation

$$g(\delta,\varepsilon)=0 \tag{2.54}$$

in the interval $\left[\varepsilon,\min\left\{\frac{3}{2\sqrt{2}}\varepsilon,1\right\}\right)$. We have two cases

1. $\varepsilon<\dfrac{2\sqrt{2}}{3}$

   We seek the solution $\delta\in\left[\varepsilon,\frac{3}{2\sqrt{2}}\varepsilon\right)$. We have

$$g(\varepsilon,\varepsilon)=2-e^{\varepsilon/\sqrt{2}}-(1-\varepsilon)e^{\varepsilon-\varepsilon/\sqrt{2}}<0$$

and

$$g\left(\frac{3\varepsilon}{2\sqrt{2}},\varepsilon\right)=\left[\left(1-\frac{\varepsilon}{2\sqrt{2}}\right)\left(2-e^{\varepsilon/\sqrt{2}}\right)-\left(1-\frac{3\varepsilon}{2\sqrt{2}}\right)\right]e^{\frac{\varepsilon}{2\sqrt{2}}}>0,$$

which implies that there is solution δ of (2.54) inside the interval $\left[\varepsilon,\frac{3}{2\sqrt{2}}\varepsilon\right)$.

2. $\dfrac{2\sqrt{2}}{3}\leq\varepsilon<1$

We seek the solution $\delta\in[\varepsilon,1)$. As above, $g(\varepsilon,\varepsilon)<0$. On the other hand,

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

If $\varepsilon<\sqrt{2}\log 2$, we have $g(1,\varepsilon)>0$ and, hence, there is a unique solution of (2.54) in the interval $\left[\frac{2\sqrt{2}}{3},\sqrt{2}\log 2\right)$.

Putting the two cases together, we see that, provided $\varepsilon\in(0,\sqrt{2}\log 2)$ there is a unique solution $\delta(\varepsilon)$ of equation (2.54), which also solves our extremal problem (2.53). On the other hand, in light of example (2.30) this is the best we can hope for, since the example implies that $\mathbf{B}^d_\varepsilon(x)=\infty$ if $\varepsilon\geq\sqrt{2}\log 2$. Therefore, we have proved that

$$\varepsilon^d_0=\sqrt{2}\log 2,$$

as Theorem 2 asserts. □

Some of the prospects and future directions of research on this topic are discussed at the end. We are now turning to the result that deals with a property not unlike the one this chapter has been devoted to. The space considered there, the Chang-Wilson-Wolff space, is better than BMO. Thus the result is stronger: instead of summability of the exponent e^{φ} we get summability of $e^{\alpha\varphi^2}$. We deviate from the formalism of the Bellman function method, so the proofs are Bellman-function-type proofs.

Chapter 3

Bellman-function-type proof of a local Chang-Wilson-Wolff theorem and related results

3.1 Introduction

Let D be the dyadic lattice on $[0,1]$. For each $I \in D$ and every $x \in [0,1]$, define the dyadic cone $\Gamma_I(x)$ to be

$$\Gamma_I(x) = \{J \in D : \; J \subseteq I, J \ni x\}. \tag{3.1}$$

Let $\varphi \in L^1([0,1])$. For every $I \in D$, let $\langle\varphi\rangle_I = \frac{1}{|I|}\int_I \varphi(s)\,ds$. Also, let I_- and I_+ be the left and right halves of I, respectively. Let

$$S\varphi(x) = \left(\sum_{I \in \Gamma_{[0,1]}(x)} \left(\langle\varphi\rangle_{I_+} - \langle\varphi\rangle_{I_-}\right)^2\right)^{1/2}. \tag{3.2}$$

be the s-function of φ. Let E be a measurable subset of $[0,1]$. We introduce the Chang-Wilson-Wolff space

$$F(E) = \left\{\varphi \in L^1 : \ \|S\varphi(x)\|_{L^\infty(E)} < \infty\right\}. \tag{3.3}$$

Then $\|S\varphi(x)\|_{L^\infty(E)}$ defines a semi-norm in this space. (In the case $E = [0,1]$, it is actually a norm, with the usual factorization over constant functions).

Fix $\varphi \in F(E)$. For every $I \in D$ such that $m(I \cap E) \neq 0$, let

$$S_I = \left\| \sum_{J \in \Gamma_I(x)} \left(\langle\varphi\rangle_{J_+} - \langle\varphi\rangle_{J_-} \right)^2 \right\|_{L^\infty(E)} \tag{3.4}$$

and $S_I = 0$ if $m(I \cap E) = 0$. Obviously, S_I depends on φ but we will not indicate this dependence when the context is unambiguous. We prove that if $S_{[0,1]} = \|S\varphi(x)\|^2_{L^\infty(E)} \leq 1$, then there exist absolute constants $\alpha > 0$ and $C > 0$ such that

$$\int_E e^{\alpha\left(\varphi - \langle\varphi\rangle_{[0,1]}\right)^2} \leq C. \tag{3.5}$$

The integrability of $e^{\alpha u^2}$ over the disk $\mathbb{D}$ under the assumption $\int_{\mathbb{D}} |\nabla u(x)|^2 dx \leq 1$ has been studied in [22] and [11]. In [1], the authors study a question like ours but for functions analytic in $\mathbb{D}$. Namely, they answer in the affirmative the question of Beurling and Moser as to whether the fact that $\int_{\mathbb{D}} |f'(z)|^2\, dz \leq \pi$ implies $\int_{\mathbb{T}} \exp |f(e^{i\theta})|^2\, d\theta \leq C$ for some absolute constant C. Considering for an instant $\mathbb{T}$ instead of $[0,1]$, we note that even in the case $E = \mathbb{T}$ (3.5) holds under weaker conditions, since if $\varphi \in L^2(\mathbb{T})$ and $\varphi(z)$ is its harmonic extension into the disk, $\int_{\mathbb{D}} |\nabla\varphi(z)|^2\, dz \leq C$ implies $\|S\varphi\|_{L^\infty(\mathbb{T})} \leq C$ with a different constant (see, for instance, [4]).

Our result follows immediately from a local version of the famous Chang-Wilson-Wolff theorem, which we prove first using a Bellman-type argument. In [2], the authors treat the case $E = [0,1]$, while we prove the theorem for arbitrary E. It has to be said that the authors' reasoning would seem to work for arbitrary E as well. However, the Bellman-function-type proof we give incorporates any E effortlessly.

3.2 Main results

We split the proof into several lemmas. Lemma 1 sets up the stage for the unwrapping of the typical Bellman-function integral sum and Lemma 2 uses it to prove a certain integral estimate, not unlike those in Chapter 1.

Lemma 1. *Assume there exists a C^2-function $B = B(x,L) : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}_+$ such that $\dfrac{\partial B}{\partial L} > 0$ and there exists $\theta \in (0,1)$ such that*

$$\frac{\dfrac{\partial^2 B}{\partial x^2}(a,b)}{\dfrac{\partial B}{\partial L}(c,d)} \leq 4(1-\theta) \tag{3.6}$$

for any $(a,b),(c,d) \in \mathbb{R} \times \mathbb{R}_+$ such that $b - d \leq -\theta\delta^2$ and $|a-c| \leq \frac{1}{2}\delta$ for some $\delta \geq 0$. Then, for every $\varphi \in F(E)$ and every $I \in D$ such that $m(I \cap E) \neq 0$,

$$B\left(\langle\varphi\rangle_I, S_I\right) \geq \frac{1}{2}B\left(\langle\varphi\rangle_{I_-}, S_{I_-}\right) + \frac{1}{2}B\left(\langle\varphi\rangle_{I_+}, S_{I_+}\right). \tag{3.7}$$

Proof. Assume the existence of such a function B. By definition (3.4),

$$S_I = \max\left\{S_{I_-}, S_{I_+}\right\} + \left(\langle\varphi\rangle_{I_+} - \langle\varphi\rangle_{I_-}\right)^2.$$

Let $L_0 = \max\left\{S_{I_-}, S_{I_+}\right\}$, $L = S_I = L_0+\delta^2$, $L^- = S_{I_-}$, $L^+ = S_{I_+}$, $x = \langle\varphi\rangle_I$, $x^- = \langle\varphi\rangle_{I_-}$, $x^+ = \langle\varphi\rangle_{I_+}$, $\delta = |x^+ - x^-|$. Then

$$\begin{aligned}
&B\left(\langle\varphi\rangle_I, S_I\right) - \frac{1}{2}B\left(\langle\varphi\rangle_{I_-}, S_{I_-}\right) - \frac{1}{2}B\left(\langle\varphi\rangle_{I_+}, S_{I_+}\right)\\
&= B(x,L) - \frac{1}{2}\left(B(x^+,L^+) + B(x^-,L^-)\right)\\
&\geq B(x, L_0+\delta^2) - \frac{1}{2}\left(B(x^+,L_0) + B(x^-,L_0)\right)
\end{aligned}$$

since $\frac{\partial B}{\partial L} > 0$. We continue

$$\begin{aligned}
&= \left[B(x,L_0+\delta^2) - B(x,L_0+\theta\delta^2)\right]\\
&\quad + \left[B(x,L_0+\theta\delta^2) - B(x,L_0)\right]\\
&\quad + \left[B(x,L_0) - \frac{1}{2}\left(B(x^+,L_0) + B(x^-,L_0)\right)\right]\\
&\geq \frac{\partial B}{\partial L}(x,\beta)\delta^2(1-\theta) - \frac{1}{4}\frac{\partial^2 B}{\partial x^2}(\gamma, L_0)\delta^2
\end{aligned}$$

for some $\beta \in [L_0+\theta\delta^2, L_0+\delta^2]$ and some γ between x^- and x^+.

$$= \frac{1}{4}\frac{\partial B}{\partial L}(x,\beta)\left[4(1-\theta) - \frac{\frac{\partial^2 B}{\partial x^2}(\gamma, L_0)}{\frac{\partial B}{\partial L}(x,\beta)}\right] \geq 0,$$

since $\gamma \leq |x - x^-| = |x - x^+| = \frac{1}{2}\delta$ and $L_0 - \beta \leq -\theta\delta^2$.

This completes the proof. □

The choice of B

We introduce the family of functions B parametrized by $t \geq 0$. Let

$$B_t(x,L) = e^{tx+At^2L} \tag{3.8}$$

for some $A \geq 0$. We find the best suitable A below. For now, we check that the functions B_t satisfy the conditions of Lemma 1. Fix any $\theta \in (0,1)$. Take a, b, c, d as in the formulation of Lemma 1. Then

$$\frac{\frac{\partial^2 B_t}{\partial x^2}(a,b)}{\frac{\partial B_t}{\partial L}(c,d)} = \frac{t^2 \exp(at + At^2 b)}{At^2 \exp(ct + At^2 d)} = \frac{1}{A}\exp((a-c)t + At^2(b-d))$$
$$\leq \frac{1}{A}\exp\left(\frac{1}{2}\delta t - At^2\theta\delta^2\right) = \frac{1}{A}\exp\left(\frac{1}{2}\left((\delta t) - 2A\theta(\delta t)^2\right)\right)$$
$$\leq \frac{1}{A}\exp\left(\frac{1}{2}\left(\frac{1}{8A\theta}\right)\right) = 4(1-\theta),$$

where we have used the fact that $y - 2A\theta y^2 \leq \frac{1}{8A\theta}$ and the last equality is guaranteed by a suitable choice of A. In Lemma 3, we use Lemma 1 with the function B defined by (3.8) to prove the estimate

$$m\left(\left\{\varphi - \langle\varphi\rangle_{[0,1]} > \lambda\right\}\right) \leq e^{-\lambda^2/(2AS_{[0,1]})}.$$

This suggests that we need to choose θ so that A is the smallest possible. Let $D = \frac{1}{4A}$. Then the equation relating θ and A becomes

$$De^{D/(4\theta)} = 1 - \theta, \tag{3.9}$$

which defines D as a function of θ in the interval $(0,1)$. We are seeking the maximum of D. Differentiating (3.9) with respect to θ, we get

$$e^{D/(4\theta)}\left[D'\left(1 + \frac{D}{4\theta} - \frac{D^2}{4\theta^2}\right)\right] + 1 = 0.$$

Setting $D' = 0$ and solving for D gives

$$D = \frac{4\theta^2}{1-\theta}. \tag{3.10}$$

Solving (3.9) and (3.10) simultaneously yields

$$\theta = \frac{2w\left(\frac{1}{4}\right)}{1+2w\left(\frac{1}{4}\right)}; \qquad D = \frac{16w^2\left(\frac{1}{4}\right)}{1+2w\left(\frac{1}{4}\right)}, \tag{3.11}$$

where w is Lambert's w-function, i.e. $w(x)$ is the solution of the equation $we^w = x$. Therefore, for the best A we obtain

$$A = \frac{1+2w\left(\frac{1}{4}\right)}{64w^2\left(\frac{1}{4}\right)} \approx 0.5291. \tag{3.12}$$

We are now in a position to prove the following lemma.

Lemma 2. *For every* $\varphi \in F(E)$ *and every* $t \geq 0$,

$$\int_E e^{t\left(\varphi(s)-\langle\varphi\rangle_{[0,1]}\right)} ds \leq e^{At^2 S_{[0,1]}}. \tag{3.13}$$

Proof. Let B_t be given by (3.8) with A given by (3.12). We will apply Lemma 1 repeatedly, at every step omitting the terms corresponding to dyadic intervals whose intersection with E has zero measure, i.e at every step we use the simple fact that

$$\sum_{|J|=2^{-j}} |J| B\left(\langle\varphi\rangle_J, S_J\right) \geq \sum_{\substack{|J|=2^{-j}\\ m(J\cap E)\neq 0}} |J| B\left(\langle\varphi\rangle_J, S_J\right),$$

which allows us to apply Lemma 1 to every term in the latter sum. Thus we have

$$\begin{aligned} e^{t\langle\varphi\rangle_{[0,1]}+At^2 S_{[0,1]}} &= B\left(\langle\varphi\rangle_{[0,1]}, S_{[0,1]}\right) \\ &\geq \frac{1}{2}B\left(\langle\varphi\rangle_{[0,1/2]}, S_{[0,1/2]}\right) + \frac{1}{2}B\left(\langle\varphi\rangle_{[1/2,1]}, S_{[1/2,1]}\right) \\ &\geq \sum_{\substack{|I|=2^{-1}\\ m(I\cup E)\neq 0}} |I|\, B\left(\langle\varphi\rangle_I, S_I\right) \geq \sum_{\substack{|I|=2^{-n}\\ m(I\cup E)\neq 0}} |I|\, B\left(\langle\varphi\rangle_I, S_I\right) \end{aligned}$$

$$= \sum_{\substack{|I|=2^{-n} \\ m(I\cup E)\neq 0}} |I|\, e^{t\langle\varphi\rangle_I + At^2 S_I}.$$

Now let $n \to \infty$ and observe that, since $S_{[0,1]} < \infty$, we have $S_I \to 0$ as $|I| \to 0$. On the other hand,

$$\sum_{\substack{|I|=2^{-n} \\ m(I\cup E)\neq 0}} |I|\, e^{t\langle\varphi\rangle_I} \longrightarrow \int_E e^{t\varphi(s)} ds \quad \text{as } n \to \infty,$$

which completes the proof. □

Lemma 3. *Let* $\varphi \in F(E)$. *Let* $E_\lambda = \left\{x \in E : \varphi(x) - \langle\varphi\rangle_{[0,1]} > \lambda\right\}$. *Let* $S = S_{[0,1]}$. *Then*

$$m(E_\lambda) \le e^{-\lambda^2/(2AS)}, \tag{3.14}$$

where A *is given by (3.12).*

Proof. We apply Chebyshev's inequality to (3.13) with $t = \lambda/(AS)$.

$$m(E_\lambda)\, e^{\lambda^2/(AS)} \le \int_E e^{\lambda\left(\varphi(s) - \langle\varphi\rangle_{[0,1]}\right)/(AS)} ds \le e^{\lambda^2/(2AS)},$$

concluding the proof. □

We are now in a position to prove the main result.

Theorem. *If* $\|\varphi\|_{F(E)} \le 1$, *then*

$$\int_E e^{\alpha\left(\varphi - \langle\varphi\rangle_{[0,1]}\right)^2} \le C(\alpha) \tag{3.15}$$

for any $\alpha < \frac{1}{2A}$, *with* A *given by (3.12).*

Proof. Lemma 3 implies that

$$m\left(\left\{\left|\varphi - \langle\varphi\rangle_{[0,1]}\right| \geq \lambda\right\}\right) \leq 2e^{-\lambda^2/(2A)}.$$

Fix any $r > 1$ such that $\alpha r^2 < \frac{1}{2A}$. Let

$$F_k = \left\{\frac{r^k}{A} < \left|\varphi - \langle\varphi\rangle_{[0,1]}\right| \leq \frac{r^{k+1}}{A}\right\}, \quad k = 0, 2, \dots \tag{3.16}$$

Take any $\alpha < \frac{1}{2A}$. Then, using (3.16),

$$\int_{F_k} e^{\alpha\left(\varphi - \langle\varphi\rangle_{[0,1]}\right)^2} \leq m(F_k)e^{\alpha r^{2k+2}/A^2} \leq 2e^{-r^{2k}/(2A^3)}e^{\alpha r^{2k+2}/A^2} = 2e^{-r^{2k}\left(1/(2A)-\alpha r^2\right)/A^2}.$$

Therefore,

$$\int_E e^{\alpha\left(\varphi - \langle\varphi\rangle_{[0,1]}\right)^2} = \int_{\left\{\left|\varphi - \langle\varphi\rangle_{[0,1]}\right| \leq \frac{1}{A}\right\}} e^{\alpha\left(\varphi - \langle\varphi\rangle_{[0,1]}\right)^2} + \sum_{k=0}^{\infty}\int_{F_k} e^{\alpha\left(\varphi - \langle\varphi\rangle_{[0,1]}\right)^2}$$

$$\leq m\left(\left\{\left|\varphi - \langle\varphi\rangle_{[0,1]}\right| \leq \frac{1}{A}\right\}\right) e^{\frac{\alpha}{A^2}} + \sum_{k=0}^{\infty} 2e^{-r^{2k}\left(1/(2A)-\alpha r^2\right)/A^2},$$

where the last sum converges and depends only on α (provided the best choice of r has been made). This completes the proof. □

3.3 Bellman function considerations

Lemma 1 and Lemma 2 are where the Bellman function technique is used, the rest of the proof follows using standard arguments. One can observe that no extremal problem as such has been posed. However, in the spirit of Chapter 1, we can try and associate with every dyadic interval I such that $m(I \cap E) \neq 0$ and every function $\varphi \in F(E)$ a point in a certain two-dimensional domain. First, we let F_α be the

"α-ball" in $F(E)$, $F_\alpha = \left\{\varphi \in F(E): \; \|S\varphi(x)\|_{L^\infty(E)} \le \alpha\right\}$. What is the domain in this case? This depends on the choice of the variables. Lemma 1 suggests the choice $(x_1, x_2) = (\langle\varphi\rangle_I, S_I)$. Then the domain associated with F_α is $\Omega_\alpha = \mathbb{R} \times [0, \alpha]$. We can define the Bellman function

$$\mathbf{B}_t^\alpha(x) = \sup_{\varphi \in F_\alpha} \left\{ \int_E e^{t\varphi(s)}\, ds : \; \langle\varphi\rangle_{[0,1]} = x_1, S^\varphi_{[0,1]} = x_2 \right\}. \tag{3.17}$$

If we let $\tilde{\varphi} = \varphi + c$, then $\langle\tilde{\varphi}\rangle = \langle\varphi\rangle + c$ and $S^{\tilde{\varphi}} = S^\varphi$. Furthermore,

$$\int_E e^{t\tilde{\varphi}(s)}\, ds = e^{tc} \int_E e^{t\varphi(s)}\, ds.$$

Taking the supremum in the last identity, we get

$$\mathbf{B}_t^\alpha(x_1 + c, x_2) = e^{tc}\, \mathbf{B}_t^\alpha(x_1, x_2).$$

If we set $c = -x_1$, we obtain that

$$\mathbf{B}_t^\alpha(x_1, x_2) = e^{tx_1} f_t(x_2) \tag{3.18}$$

for some $f_t \ge 0$. This is one of the considerations that led us to choose the family (3.8) of Bellman function majorates. We observe that the inequality (3.7)

$$B\left(\langle\varphi\rangle_I, S_I\right) \ge \frac{1}{2} B\left(\langle\varphi\rangle_{I_-}, S_{I_-}\right) + \frac{1}{2} B\left(\langle\varphi\rangle_{I_+}, S_{I_+}\right)$$

implies some sort of concave behavior about B, although the functions (3.8) are decidedly not concave and (3.7) holds because of the subtle interaction between $\frac{\partial^2 B}{\partial x^2}$ and $\frac{\partial B}{\partial L}$. If one were to produce a concave function U of the form (3.18), which would also be increasing with respect to the second argument, one could conclude that

$$\frac{1}{2}U\left(\langle\varphi\rangle_{I_-}, S_{I_-}\right) + \frac{1}{2}U\left(\langle\varphi\rangle_{I_+}, S_{I_+}\right) \le U\left(\langle\varphi\rangle_I, \frac{1}{2}\left(S_{I_-} + S_{I_+}\right)\right) \le U\left(\langle\varphi\rangle_I, S_I\right),$$

where the last inequality is due to the definition of S_I and the fact that $\frac{\partial U}{\partial L} \ge 0$. However, a concave function of the form (3.18) does not exist, signaling the need for more delicate considerations.

The approach used to obtain the results of this chapter may benefit from being put on a more formal Bellman function method footing. However, as the preceding discussion demonstrates, the questions of formulating the extremal problem exactly, the choice of the variables, and, of course, finding the Bellman function explicitly can be very complicated in this case. Some intriguing perspectives are discussed at the end.

We now turn to a result, in which the Bellman function does not appear in connection with any extremal problem at all, although, doubtless, the corresponding formalism may be developed. This result brings us back to the space BMO, the main space of interest in this work. The old and famous question of $H^1 - BMO$ duality is examined using the Bellman-function-type approach, which proves powerful; one Bellman-function lemma yields concise proofs in both, the dyadic and continuous settings.

Chapter 4

Bellman function and the $H^1 - BMO$ duality

4.1 Introduction

In this chapter, we aim to demonstrate technique rather than new results. Specifically, with the aid of the Bellman function method we prove one, the more technically involved, direction of the famous Fefferman duality theorem by elementary means. Namely, we establish the fact that $BMO_0(\mathbb{T}) \subset H^1(\mathbb{T})^*$ ($BMO_0(\mathbb{T}) = \{\varphi \in BMO(\mathbb{T}), \varphi(0) = 0\}$, and as usual, $\varphi(z)$ is the harmonic continuation of φ into $\mathbb{D}$). The most complex technical tool we use is Green's formula; thus the proof serves to show further the power of the Bellman function method in harmonic analysis. An application of the same method in the dyadic case serves to establish that $BMO^d \subset \left(\overset{d}{F}{}_1^{02}\right)^*$ (with an explicit estimate for the constant of embedding), with the Triebel-Lizorkin space $\overset{d}{F}{}_1^{02}$ giving a convenient characterization for $H^1_d(\mathbb{T})$, the dyadic version of $H^1(\mathbb{T})$. A simple argument demonstrates the converse inclusion. One technical "trick" allows us to deal with both, the dyadic and continuous cases. The key to the proof is a lemma whose hypotheses include the existence of a certain function. We will call it *the* Bellman function slightly abusing the language, since we make no claim as to its uniqueness.

The formulation of the key lemma. Let $D = D_{I_0}$ be the dyadic lattice rooted in an interval I_0. For an interval $I \in D$, let I_- and I_+ be its left and right halves, respectively. Consider two functions, $S : D \to (0, \infty)$ and $M : D \to [0, \bar{M}]$, such that

$$S_{I_-} = S_{I_+} \geq S_I \quad \text{and} \quad M_I \geq \frac{1}{2}(M_{I_-} + M_{I_+}), \forall I \in D. \tag{4.1}$$

Lemma. *Let S and M be as above. Assume there exists a C^2-function B : $(0, \infty) \times [0, \bar{M}] \to \mathbb{R}$ satisfying*

$$0 \leq B(x, y) \leq 2\bar{M}\sqrt{x}, \qquad -\frac{\partial B}{\partial x}\frac{\partial B}{\partial y} \geq \frac{\bar{M}}{2}, \quad \frac{\partial^2 B}{\partial x^2} \leq 0, \quad \frac{\partial^2 B}{\partial y^2} \geq 0. \tag{4.2}$$

Then, for any positive integer n,

$$\sum_{\substack{J \in D \\ |J| \geq 2^{-n+1}}} |J| \sqrt{(S_{J_+} - S_J)\left(M_J - \frac{1}{2}(M_{J_-} + M_{J_+})\right)} \leq \sqrt{\frac{\bar{M}}{2}}\, 2^{-n} \sum_{\substack{J \in D \\ |J| = 2^{-n}}} \sqrt{S_J}. \tag{4.3}$$

We will prove the lemma and demonstrate our Bellman function later. For now, we will establish the main results.

4.2 The dyadic case

Consider the dyadic lattice $D = D_{\mathbb{T}}$ on $\mathbb{T}$. Let $\overset{d}{F}{}^{02}_{1}$ be the dyadic Triebel-Lizorkin space,

$$\overset{d}{F}{}^{02}_{1} = \left\{ f \in L^1 : \int_{\mathbb{T}} \left(\sum_{I \ni \theta; I \in D} \left(\langle f \rangle_{I_+} - \langle f \rangle_{I_-} \right)^2 \right)^{1/2} d\theta < \infty \right\} \tag{4.4}$$

with the norm

$$\|f\|_{F_1^{d02}} = \int_{\mathbb{T}} \left(\sum_{I \ni \theta; I \in D} \left(\langle f \rangle_{I_+} - \langle f \rangle_{I_-} \right)^2 \right)^{1/2} d\theta.$$

The definition of the dyadic BMO is the same (up to a constant multiple) as the one we used in Chapter 2:

$$BMO^d = \left\{ \varphi \in L^1 : \sup_{J \in D} \frac{1}{|J|} \sum_{I \subset J} \left(\langle \varphi \rangle_{I_+} - \langle \varphi \rangle_{I_-} \right)^2 |I| < \infty \right\} \tag{4.5}$$

with the norm

$$\|\varphi\|_{BMO^d} = \sup_{J \in D} \left(\frac{1}{|J|} \sum_{I \subset J} \left(\langle \varphi \rangle_{I_+} - \langle \varphi \rangle_{I_-} \right)^2 |I| \right)^{1/2}.$$

To see the equivalence of the definitions (2.3) and (4.5), recall the Haar system: for every dyadic arc J, let

$$h_J = \begin{cases} \dfrac{1}{\sqrt{|J|}} & \text{on } J_- \\ -\dfrac{1}{\sqrt{|J|}} & \text{on } J_+ \\ 0 & \text{elsewhere} \end{cases} . \tag{4.6}$$

It is easy to check that $\{h_J\}_{J \in D}$ form an orthonormal system in $L_0^2 = \{f \in L^2(\mathbb{T}) : \int_{\mathbb{T}} f(\theta)\, d\theta = 0\}$, what is more (and well-known) is the fact that the Haar system actually is a basis for L_0^2. For any function $f \in L^1$ and every $J \in D$ one can compute the corresponding Haar coefficient,

$$(f, h_J) = \frac{\sqrt{|J|}}{2} \left(\langle f \rangle_{J_-} - \langle f \rangle_{J_+} \right). \tag{4.7}$$

For $f \in L_0^2$ we then have $f = \sum_{J \in D} (f, h_J) h_J$ and

$$\|f\|_{L^2} = \sum_{J\in D}(f, h_J)^2 = \sum_{J\in D} \frac{|J|}{4}\left(\langle f\rangle_{J_-} - \langle f\rangle_{J_+}\right)^2.$$

We state our main result.

Theorem 1. $BMO^d = \left(\overset{d}{F}{}^{02}_{1}\right)^*$.

Proof. The more difficult inclusion is handled using the Bellman-function lemma stated above.

Lemma 1. $BMO^d \subset \left(\overset{d}{F}{}^{02}_{1}\right)^*$. *More precisely, in terms of the Haar coefficients, for every* $\varphi \in BMO^d$ *and* $f \in \overset{d}{F}{}^{02}_{1}$,

$$\begin{aligned}\sum_{J\in D} |(f, h_J)|\,|(\varphi, h_J)| &= \frac{1}{4}\sum_{J\in D} |J|\,|\,\langle f\rangle_{J_+} - \langle f\rangle_{J_-}\,|\,|\,\langle \varphi\rangle_{J_+} - \langle \varphi\rangle_{J_-}\,| \qquad (4.8)\\ &\le \frac{1}{4\sqrt{2}}\|\varphi\|_{BMO^d}\,\|f\|_{\overset{d}{F}{}^{02}_{1}}.\end{aligned}$$

Proof. Fix $\varphi \in BMO^d$, $f \in \overset{d}{F}{}^{02}_{1}$. For every $J \in D$ define

$$M_J = \frac{1}{|J|}\sum_{I\subset J}\left(\langle\varphi\rangle_{I_+} - \langle\varphi\rangle_{I_-}\right)^2 |I|.$$

Then $0 \le M_J \le \bar{M} \overset{def}{=} \|\varphi\|^2_{BMO^d}$ and $M_J - \frac{1}{2}\left(M_{J_+} + M_{J_-}\right) = \left(\langle\varphi\rangle_{J_+} - \langle\varphi\rangle_{J_-}\right)^2$. Define

$$S_J = \sum_{I\supsetneq J}\left(\langle f\rangle_{I_+} - \langle f\rangle_{I_-}\right)^2.$$

Then $S_{J_+} = S_{J_-} = \sum_{I\supset J}\left(\langle f\rangle_{I_+} - \langle f\rangle_{I_-}\right)^2$ and $S_{J_+} - S_J = S_{J_-} - S_J = \left(\langle f\rangle_{J_+} - \langle f\rangle_{J_-}\right)^2$. We thus see that the conditions (4.1) of the lemma are satisfied.

Assuming the existence of the function B in the lemma and using (4.3), we obtain

$$\sum_{\substack{J\in D\\|J|\geq 2^{-n+1}}} |J|\,|\langle f\rangle_{J_+}-\langle f\rangle_{J_-}|\,|\langle\varphi\rangle_{J_+}-\langle\varphi\rangle_{J_-}|\leq\sqrt{\frac{\bar{M}}{2}}\,2^{-n}\sum_{\substack{J\in D\\|J|=2^{-n}}}\sqrt{\sum_{I\supsetneq J}\left(\langle f\rangle_{J_+}-\langle f\rangle_{J_-}\right)^2}.$$

Letting $n\to\infty$, we get the statement (4.8) of the theorem. □

The converse inclusion follows along more conventional lines.

Lemma 2. $\left(\overset{d}{F}{}_1^{02}\right)^*\subset BMO^d$.

Proof. We want to show that for every continuous linear functional l on $\overset{d}{F}{}_1^{02}$ there exists $\varphi\in BMO^d$ such that

$$\|\varphi\|_{BMO^d}\leq c\|l\| \tag{4.9}$$

and

$$l(f)=\int_{\mathbf{T}}\varphi(\theta)f(\theta)\,d\theta,\quad\forall f\in\overset{d}{F}{}_1^{02}. \tag{4.10}$$

First, we observe that $L_0^2\in\overset{d}{F}{}_1^{02}$. Indeed, for $f\in L_0^2$,

$$\begin{aligned}\|f\|^2_{\overset{d}{F}{}_1^{02}}&=\left(\int_{\mathbf{T}}\left(\sum_{I\ni\theta;I\in D}\left(\langle f\rangle_{I_+}-\langle f\rangle_{I_-}\right)^2\right)^{\frac{1}{2}}d\theta\right)^2\leq\int_{\mathbf{T}}\sum_{I\ni\theta;I\in D}\left(\langle f\rangle_{I_+}-\langle f\rangle_{I_-}\right)^2d\theta\\&=\sum_{I\in D}\int_{\mathbf{T}}\chi_I(\theta)\left(\langle f\rangle_{I_+}-\langle f\rangle_{I_-}\right)^2d\theta=\sum_{I\in D}|I|\left(\langle f\rangle_{I_+}-\langle f\rangle_{I_-}\right)^2=4\|f\|^2_{L_0^2}.\end{aligned}$$

Let $l\in\left(\overset{d}{F}{}_1^{02}\right)^*$. We can apply the Riesz representation theorem to $l|_{L_0^2}$ and conclude that there exists a function $\varphi\in L_0^2$ such that

$$l(f)=\int_{\mathbf{T}}\varphi(\theta)f(\theta)\,d\theta,\quad\forall f\in L_0^2. \tag{4.11}$$

We test l on appropriate elements of L_0^2 to see that $\varphi \in BMO^d$. Let a_I be an atom associated with a dyadic arc I, i.e. be supported on I with $|a_I| \leq \frac{1}{|I|}$, *a.e.* and $\int_I a(\theta)\, d\theta = 0$. We have

$$\begin{aligned}
\|a_I\|_{\overset{d}{F}{}_1^{02}} &= \int_{\mathbf{T}} \left(\sum_{J \ni \theta; J \in D} \left(\langle a_I \rangle_{J_+} - \langle a_I \rangle_{J_-} \right)^2 \right)^{1/2} d\theta \\
&= \int_I \left(\sum_{J \ni \theta; J \subset I} \left(\langle a_I \rangle_{J_+} - \langle a_I \rangle_{J_-} \right)^2 \right)^{1/2} d\theta \\
&\leq \left(\int_I \sum_{J \ni \theta; J \subset I} \left(\langle a_I \rangle_{J_+} - \langle a_I \rangle_{J_-} \right)^2 d\theta \right)^{1/2} \left(\int_I 1\, d\theta \right)^{1/2} \\
&= 2\|a_I\|_{L_0^2} \sqrt{|I|} \leq \frac{2}{\sqrt{|I|}} \sqrt{|I|} = 2,
\end{aligned}$$

and hence

$$\left| \int_I (\varphi - \langle \varphi \rangle_I)\, a_I \right| = \left| \int_{\mathbf{T}} \varphi\, a_I \right| = |l(a_I)| \leq \|l\| \|a_I\|_{\overset{d}{F}{}_1^{02}} \leq 2\|l\|.$$

Since this is true for any atom a_I, we conclude that $\int_I |\varphi - \langle \varphi \rangle_I| \leq 2\|l\||I|$ and thus that $\varphi \in BMO^d$ with the norm estimate (4.9). Here we have used the equivalence of the L^1- and L^2-based BMO norms, which is due to the John-Nirenberg inequality (see section 1.4.1). The proof of Lemma 2 (and hence Theorem 1) thus depends on proving that L_0^2 is dense in $\overset{d}{F}{}_1^{02}$. Together with (4.11) this will yield the result.

Take $f \in \overset{d}{F}{}_1^{02}$. Let f_n be the truncation of its Haar expansion at the n-th generation of the dyadic lattice,

$$f_n = \sum_{\substack{J \in D \\ |J| \geq 2^{-n}}} (f, h_J) h_J.$$

While $\{f_n\}$ may not converge in the L_0^2-norm, we show that it does converge (to f) in the $\overset{d}{F}{}_1^{02}$-norm. We have

$$\|f - f_n\|_{\overset{d}{F}{}_1^{02}} = \int_{\mathbf{T}} \left(\sum_{J \ni \theta} \frac{4}{|J|} (f - f_n, h_J)^2 \right)^{1/2} d\theta = \int_{\mathbf{T}} \left(\sum_{\substack{J \ni \theta \\ |J| < 2^{-n}}} \frac{4}{|J|} (f, h_J)^2 \right)^{1/2} d\theta.$$

Since $f \in \overset{d}{F}{}_1^{02}$, the dominated convergence theorem applies, so $\|f - f_n\|_{\overset{d}{F}{}_1^{02}} \to 0$ as $n \to \infty$. This concludes the proof of Lemma 2 and Theorem 1. □

4.3 The continuous case

Following [21], we define $H^1 = H^1(\mathbb{T})$ using the area integral; specifically

$$H^1 = \left\{ f \in L^1 : \int_{\mathbf{T}} \left(\int_{\Gamma_\alpha(e^{i\theta})} |f'(\xi)|^2 dA(\xi) \right)^{1/2} d\theta < \infty \right\} \tag{4.12}$$

with the corresponding norm

$$\|f\|_{H^1} = \int_{\mathbf{T}} \left(\int_{\Gamma_\alpha(e^{i\theta})} |f'(\xi)|^2 dA(\xi) \right)^{1/2} d\theta. \tag{4.13}$$

Here $f(z)$ is the harmonic extension of f into $\mathbb{D}$. $\Gamma_\alpha(e^{i\theta})$ is the cone-like region with vertex $e^{i\theta}$: $\Gamma_\alpha(e^{i\theta}) = \left\{ z \in \mathbb{D} : \dfrac{|e^{i\theta} - z|}{1 - |z|} < \dfrac{1}{\sin \alpha} \right\}$. We will specify the angle α a little later.

The corresponding definition of $BMO_0 = BMO_0(\mathbb{T})$ is

$$BMO_0 = \left\{ \varphi \in L^1 : \sup_{\text{arc } I \subset \mathbf{T}} \frac{1}{|I|} \int_{Q_I} |\varphi'(\xi)|^2 (1 - |\xi|) \, dA(\xi) < \infty, \ \varphi(0) = 0 \right\}, \tag{4.14}$$

where $\varphi(z)$ is the harmonic extension of φ into $\mathbb{D}$ and Q_I is the Carleson square corresponding to the arc I, $Q_I = \{z \in \mathbb{D} : z/|z| \in I, 1 - |I| \leq |z| < 1\}$. The norm

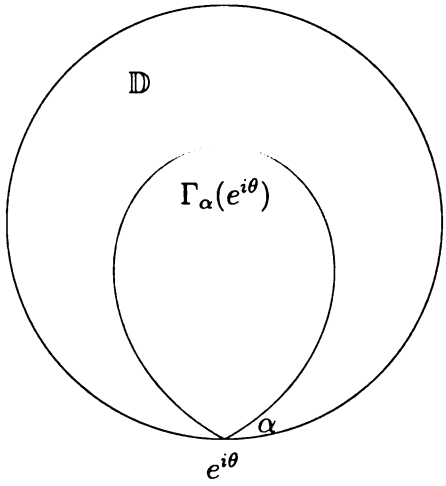


Figure 4.1: The region $\Gamma_\alpha(e^{i\theta})$.

in this space is then

$$\|\varphi\|_{BMO} = \sup_{\text{arc } I\subset\mathbf{T}} \left(\frac{1}{|I|}\int_{Q_I} |\varphi'(\xi)|^2(1-|\xi|)dA(\xi)\right)^{1/2}.$$

We are now in a position to state the main result.

Theorem 2. $BMO_0 \subset (H^1)^*$. *More precisely,*

$$\left|\int_{\mathbf{T}} \varphi(e^{i\theta})\bar{f}(e^{i\theta})\, d\theta\right| \le C\|\varphi\|_{BMO_0}\|f\|_{H^1},\ \forall\varphi\in BMO_0, \forall f\in H^1. \tag{4.15}$$

Proof. Not surprisingly, the proof starts off dyadic. For every $J \in D = D_{\mathbf{T}}$ define

$$M_J = \frac{1}{|J|}\int_{Q_J} |\varphi'(\xi)|^2(1-|\xi|)\, dA(\xi).$$

Clearly, $M_J \le \bar{M} \stackrel{def}{=} \|\varphi\|_{BMO_0}$. We have

$$M_J - \frac{1}{2}\left(M_{J_+} - M_{J_-}\right) = \frac{1}{|J|}\int_{TQ_J} |\varphi'(\xi)|^2(1-|\xi|), dA(\xi).$$

Here TQ_J is the top half of the (dyadic) square Q_J, $TQ_J = Q_J \backslash (Q_{J_+} \cup Q_{J_-})$.

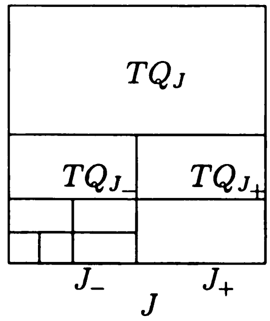


Figure 4.2: The decomposition $Q_J = \bigcup_{I\in D, I\subset J} TQ_I$.

Define

$$S_J = \int_{\Gamma_J^d} |f'(\xi)|^2 dA(\xi) = \sum_{I\supsetneq J; I\in D} \int_{TQ_I} |f'(\xi)|^2 dA(\xi).$$

Here Γ_J^d is the dyadic cone, $\Gamma_J^d = \bigcup_{I\supsetneq J} TQ_J$. *Observation:* For some fixed α, $\Gamma_J^d \subset \Gamma_\alpha(e^{i\theta})$, $\forall\theta\in J$.

We have $S_{J_-} = S_{J_+} = \sum_{I\supseteq J; I\in D} \int_{TQ_I} |f'(\xi)|^2 dA(\xi)$ and thus, $S_{J_-} - S_J = S_{J_+} - S_J = \int_{TQ_J} |f'(\xi)|^2 dA(\xi)$. Therefore, the conditions (4.1) of the key lemma are satisfied. Assuming the existence of the function B in the lemma and using (4.3), we have

$$\sum_{|J|\geq 2^{-n+1}} |J| \left(\frac{1}{|J|}\int_{TQ_J} |\varphi'(\xi)|^2 (1-|\xi|)\, dA(\xi)\right)^{1/2} \left(\int_{TQ_J} |f'(\xi)|^2 dA(\xi)\right)^{1/2}$$

$$\leq \sqrt{\frac{M}{2}}\, 2^{-n} \sum_{|J|=2^{-n}} \left(\int_{\Gamma_J^d} |f'(\xi)|^2 dA(\xi)\right)^{1/2}.$$

Let us estimate the left-hand side as $n\to\infty$.

$$\lim_{n\to\infty} \sum_{|J|\geq 2^{-n+1}} |J| \left(\frac{1}{|J|}\int_{TQ_J} |\varphi'(\xi)|^2 (1-|\xi|)\, dA(\xi)\right)^{1/2} \left(\int_{TQ_J} |f'(\xi)|^2 dA(\xi)\right)^{1/2}$$

$$= \sum_{J \in D} |J|^{1/2} \left(\int_{TQ_J} |\varphi'(\xi)|^2 (1 - |\xi|)\, dA(\xi) \right)^{1/2} \left(\int_{TQ_J} |f'(\xi)|^2 dA(\xi) \right)^{1/2}$$

$$\geq \sum_{J \in D} |J|^{1/2} \int_{TQ_J} |\varphi'(\xi)|\, |f'(\xi)| (1 - |\xi|)^{1/2} dA(\xi)$$

$$\geq C' \sum_{J \in D} |J| \int_{TQ_J} |\varphi'(\xi)|\, |f'(\xi)|\, dA(\xi)$$

$$\geq C' \int_{\mathbb{D}} |\varphi'(\xi)|\, |f'(\xi)| \log \frac{1}{|\xi|}\, dA(\xi)$$

(Here we have used the fact that $(1 - |\xi|)^{1/2} \sim |J|^{1/2}$ and $|J| \sim \log \frac{1}{|\xi|}$ if $\xi \in TQ_J$. In addition, $\bigcup_{J \in D} TQ_J = \mathbb{D}$.)

$$\geq C' \left| \int_{\mathbb{D}} \partial\varphi \bar{\partial} \bar{f} \log \frac{1}{|\xi|} dA(\xi) \right|$$

$$= C'' \left| \int_{\mathbb{D}} \Delta(\varphi \bar{f}) \log \frac{1}{|\xi|} dA(\xi) \right|,$$

where we have used the fact that $\partial\varphi\bar{\partial}\bar{f} = \partial\bar{\partial}(\varphi\bar{f}) = \frac{1}{4}\Delta(\varphi\bar{f})$, since φ and f are analytic.

Recall Green's formula.

$$\frac{1}{2\pi} \int_{\mathbf{T}} F(e^{i\theta}) d\theta - F(0) = \frac{1}{2\pi} \int_{\mathbb{D}} \Delta F(\xi) \log \frac{1}{|\xi|} dA(\xi).$$

Since $\varphi(0) = 0$, we get $\lim_{n \to \infty} (LHS) \geq C \left| \int_{\mathbf{T}} \varphi(e^{i\theta}) \bar{f}(e^{i\theta}) d\theta \right|$.

On the right-hand side we obtain, as $n \to \infty$,

$$\frac{1}{\sqrt{2}}\|\varphi\|_{BMO_0}\int_{\mathbf{T}}\left(\int_{\Gamma^d_\alpha(e^{i\theta})}|f'(\xi)|^2dA(\xi)\right)^{1/2}d\theta,$$

where $\Gamma^d_\alpha(e^{i\theta}) = \bigcup_{J\ni e^{i\theta}}\Gamma^d_J$. Since each $\Gamma^d_J \subset \Gamma_\alpha(e^{i\theta})$, we have $\Gamma^d_\alpha(e^{i\theta}) \subset \Gamma_\alpha(e^{i\theta})$, and thus

$$\int_{\mathbf{T}}\left(\int_{\Gamma^d_\alpha(e^{i\theta})}|f'(\xi)|^2dA(\xi)\right)^{1/2}d\theta \le \int_{\mathbf{T}}\left(\int_{\Gamma_\alpha(e^{i\theta})}|f'(\xi)|^2dA(\xi)\right)^{1/2}d\theta = \|f\|_{H^1}.$$

Putting together the estimates for the right- and left-hand sides, we obtain the statement (4.15). □

4.3.1 Multi-dimensional setting

Without going into detail, we note that an almost identical proof allows us to extend the results to the multi-dimensional setting. Fix a dyadic lattice D on $\mathbb{R}^n$. As before, for $I \in D$ define Q_I and TQ_I. Given $x \in \mathbb{R}^n$, introduce the "strange" cones

$$\Gamma^d_x = \bigcup_{I\ni x, I\in D} TQ_I.$$

For $f \in L^1(\mathbb{R}^n)$, let $f(y,t)$ be its harmonic extension into $\mathbb{R}^{n+1}_+$. Let

$$\mathcal{H}^1(\mathbb{R}^n) = \left\{f \in L^1(\mathbb{R}^n) : \int_{\mathbb{R}^n}\left(\int_{\Gamma^d_x}|\nabla f(y,t)|^2\,dy\,dt\right)^{1/2}dx < \infty\right\}.$$

with the corresponding norm. We use a quite natural $BMO(\mathbb{R}^n)$.

$$BMO(\mathbb{R}^n) = \left\{\varphi \in L^1(\mathbb{R}^n) : \int_I |\nabla\varphi(y,t)|^2 t\,dy\,dt \le C^2|I|, \forall I, \text{ a cube in } \mathbb{R}^n\right\}.$$

The best such C is the corresponding BMO norm. In this notation, we can state the following theorem

Theorem. *$BMO(\mathbb{R}^n) \subset (\mathcal{H}^1(\mathbb{R}^n))^*$ and the constant of embedding does not depend on dimension.*

4.4 The proof of the key lemma

Proof. Fix $J \in D$. Let $S = S_J$; $S_0 = S_{J_-} = S_{J_+}$; $M = M_J$; $M_- = M_{J_-}$; $M_+ = M_{J_+}$. Then

$$\frac{1}{2}B(S_{J_-}, M_{J_-}) + \frac{1}{2}B(S_{J_+}, M_{J_+})$$

$$= \frac{1}{2}B(S_0, M_-) + \frac{1}{2}B(S_0, M_+) - B(S_0, M) + B(S_0, M) - B(S, M) + B(S, M)$$

$$\geq B\left(S_0, \frac{1}{2}(M_- + M_+)\right) - B(S_0, M) + \frac{\partial B}{\partial S}(\hat{S}, M)(S_0 - S) + B(S, M)$$

$$= -\frac{\partial B}{\partial M}(S_0, \dot{M})\left(M - \frac{1}{2}(M_- + M_+)\right) + \frac{\partial B}{\partial S}(\hat{S}, M)(S_0 - S) + B(S, M),$$

for some $\hat{S} \in [S, S_0]$ and $\dot{M} \in \left[\frac{1}{2}(M_- + M_+), M\right]$. Since $\frac{\partial^2 B}{\partial S^2} \leq 0$, we have $\frac{\partial B}{\partial S}(\hat{S}, M) \geq \frac{\partial B}{\partial S}(S_0, M)$ and since $\frac{\partial^2 B}{\partial M^2} \leq 0$, we have $-\frac{\partial B}{\partial M}(S_0, \hat{M}) \geq -\frac{\partial B}{\partial M}(S_0, M)$. Therefore,

$$\frac{1}{2}B(S_{J_-}, M_{J_-}) + \frac{1}{2}B(S_{J_+}, M_{J_+})$$

$$\geq -\frac{\partial B}{\partial M}(S_0, M)\left(M - \frac{1}{2}(M_- + M_+)\right) + \frac{\partial B}{\partial S}(S_0, M)(S_0 - S) + B(S, M)$$

$$\geq 2\sqrt{-\frac{\partial B}{\partial M}(S_0, M)\frac{\partial B}{\partial S}(S_0, M)}\sqrt{\left(M - \frac{1}{2}(M_- + M_+)\right)(S_0 - S) + B(S, M)}$$

$$\geq \sqrt{2\bar{M}}\sqrt{\left(M - \frac{1}{2}(M_- + M_+)\right)(S_0 - S) + B(S, M)}.$$

Now,

$$2^{-n}\sum_{\substack{J\in D\\ |J|=2^{-n}}}\frac{1}{2}B(S_J, M_J) = 2^{-n}\sum_{\substack{J\in D\\ |J|=2^{-n+1}}}\left[\frac{1}{2}B(S_{J_-}, M_{J_-}) + \frac{1}{2}B(S_{J_+}, M_{J_+})\right]$$

$$\geq 2^{-n}\sum_{\substack{J\in D\\ |J|=2^{-n+1}}}\left[\sqrt{2\bar{M}}\sqrt{(S_{J_+} - S_J)\left(M_J - \frac{1}{2}(M_{J_-} + M_{J_+})\right)}\right.$$

$$\left. + B(S_J, M_J)\right]$$

$$= 2^{-n}F_n + 2^{-n+1}\sum_{\substack{J\in D\\ |J|=2^{-n+2}}}\left[\frac{1}{2}B(S_{J_-}, M_{J_-}) + \frac{1}{2}B(S_{J_+}, M_{J_+})\right]$$

$$\geq 2^{-n}F_n + 2^{-n+1}F_{n-1} + 2^{-n+1}\sum_{\substack{J\in D\\ |J|=2^{-n+2}}} B(S_J, M_J)$$

$$\geq \cdots \geq \sum_{k=1}^{n} 2^{-k}F_k + B(S_{I_0}, M_{I_0}),$$

where we have set

$$F_k = \sum_{\substack{J\in D\\ |J|=2^{-k+1}}}\sqrt{2\bar{M}}\sqrt{(S_{J_+} - S_J)\left(M_J - \frac{1}{2}(M_{J_-} + M_{J_+})\right)}.$$

Using the fact that $B(S_J, M_J) \leq 2\bar{M}\sqrt{S_J}, \forall J \in D$, $B(S_{I_0}, M_{I_0}) \geq 0$ and the definition of F_k, we obtain the statement of the lemma. □

4.5 A sample Bellman function

We are looking for a function B such that

$$0 \le B \le C_1\sqrt{x}, \quad \frac{\partial^2 B}{\partial x^2} \le 0, \quad \frac{\partial^2 B}{\partial y^2} \ge 0, \quad -\frac{\partial B}{\partial y}\frac{\partial B}{\partial x} \ge C_2,$$

for some positive constants C_1, C_2. The proof of the key lemma using B suggests that we want to choose these constants in order to minimize the ratio $C_1/\sqrt{C_2}$. To make the estimates "sharper," we require that $\frac{\partial^2 B}{\partial y^2} = 0$. (We cannot require equality for $\frac{\partial^2 B}{\partial x^2} \le 0$.) This means that B is a linear function of y. Furthermore, $\frac{\partial B}{\partial y}$ must be negative. Because of the homogeneity in the way B is used in the lemma, we only need to choose one coefficient in that linear function. Thus we seek B in the form

$$B(x,y) = \sqrt{x}(A - y).$$

Therefore,

$$-\frac{\partial B}{\partial x}\frac{\partial B}{\partial y} = \frac{A-y}{2} \ge \frac{A-\bar{M}}{2},$$

since $y \le \bar{M}$. We have $C_1 = A$ and for the ratio to be minimized

$$\frac{C_1}{\sqrt{C_2}} = \frac{\sqrt{2}A}{\sqrt{A-\bar{M}}}.$$

The minimum of this ratio is attained at $A = 2\bar{M}$, thus producing the function

$$B(x,y) = \sqrt{x}(2\bar{M} - y),$$

satisfying conditions (4.3).

Some possible directions of future research on the Bellman function and $H^1 - BMO$ duality are discussed in the next section.

Research prospects

In this section, we discuss some possible directions for future research on the topics presented throughout the thesis. We do not attempt to embrace all possible research prospects but rather concentrate on those of immediate importance and the greatest promise.

John-Nirenberg inequality

The most natural continuation of the research on this topic would be to find the Bellman function for the weak-form John-Nirenberg inequality (1.14)

$$m\left(\{x \in I:\ |\varphi(x) - \langle\varphi\rangle_I| > \lambda\}\right) \le c_1 e^{-c_2\lambda/\|\varphi\|_{BMO(I)}},$$

where the L^2-based BMO norm is used. While it is entirely possible that the extremal functions are different for the weak and integral form of the inequality, it would be very interesting to find out how they are related. The corresponding formulation would look something like this

$$B(s,t) = \sup\Big\{|\{x \in [0,1]:\ |\varphi(x) - <\varphi>_{[0,1]}| > t\}|,\ \varphi \in BMO_p,\ \|\varphi\|_{BMO_p} \le s\Big\}.$$

Finding B exactly would give sharp constants for the traditional John-Nirenberg inequality. This problem is open even in the case $p = 2$.

The next order of business would be to explore the possibility of finding the Bellman function for the integral form in the L^p-based formulations with $p \neq 2$. We observe that the ease with which we managed to associate a plane domain with the ε-ball in BMO in the L^2 setting will not be there. Somewhat similar formulations are possible in the case of even, positive p, however the number of variables involved would grow very fast, as would the number of constraints relating those variables. The essential feature of the formulation we have used is that, after considerations of homogeneity, the Bellman function can be sought as a function of one variable. That is unlikely to happen if the number of the variables in the set-up increases.

Perhaps a more perspective direction of research (and, possibly, of more interest to the general mathematical audience) would be to try to use the results obtained to deal with $BMO(Q)$ with Q being a cube in $\mathbb{R}^n$. The John-Nirenberg inequality holds in higher-dimensional BMO; hence the question of best constants.

Chang-Wilson-Wolff theorem

Having successfully treated the local Chang-Wilson-Wolff theorem in the dyadic case using the Bellman function technique, we would like to apply the method to the continuous version of the theorem. Namely, for $f \in L^1(\mathbb{R}^n)$, let $\Gamma_\gamma(x) = \{(y,t) \in \mathbb{R}^{n+1}_+ : \ |x-y| < \gamma t\}$. Let

$$A_\gamma f(x) = \left(\int_{\Gamma_\gamma(x)} |\nabla_y f(y,t)|^2 t^{1-n} \, dy \, dt \right)^{1/2}.$$

Assume $A_\gamma f \in L^\infty(\mathbb{R}^n)$. Then, as shown in [2],

$$\sup_{I:\ \text{cube}} \frac{1}{|I|} \int_I \exp\left(c_1 \frac{|f - \langle f \rangle_I|^2}{\|A_\gamma f\|_\infty^2} \right) < c_2,$$

where c_1, c_2 depend on γ and the dimension. A successful application of the Bellman function technique (how to do it is far from obvious in this case) may yield sharp constants. If the constants are independent of γ, one may replace A_γ by the g-function, $g(f)(x) = \left(\int_0^\infty |\nabla_x f(x,t)|^2 t\,dt\right)^{1/2}$, thus solving a famous open problem.

On the other hand, it would be instrumental to try and pose the extremal problem even in the dyadic case, so that the Bellman function could be found explicitly. The difficulties with the choice of variables and the right scaling (an impossibility in the case $E \neq [0,1]$) need to be addressed.

$H^1 - BMO$ duality

The successful formulation of an optimization problem in this case would be a significant accomplishment, since it would not only be the first of its kind, but also would shed some light on how to proceed to find the best function exactly. If the optimization problem is explicitly solved, the sharp constant of embedding would be produced, in the one- and multi-dimensional cases, which, to the best of our knowledge, is an open problem.

Bibliography

[1] S.-Y. A. Chang, D. E. Marshall. On a sharp inequality concerning the Dirichlet integral. *American Journal of Mathematics*, Vol. 107 (1985), no. 5, pp. 1015-1033.

[2] S.-Y. A. Chang, J. M. Wilson, T. H. Wolff. Some weighted norm inequalities concerning the Schrödinger operators. *Comment. Math. Helvetici*, Vol. 60 (1985), no. 2, pp. 217-246.

[3] D. L. Burkholder. Boundary value problems and sharp inequalities for martingale transforms. *Annals of Probability*, Vol. 12 (1984), no. 3, pp. 647-702.

[4] D. Donoho, M. Frazier. Matrix boundedness on discrete Littlewood-Paley Spaces with applications to Haar and interpolating wavelets. Preprint, MSU, 2002.

[5] C. Fefferman. Characterizations of bounded mean oscillations. *Bull. Amer. Math. Soc.*, Vol. 77 (1971), pp. 587-588.

[6] J. B. Garnett, *Bounded Analytic functions*. Pure and Applied Math. #96, Academic Press, 1981.

[7] S. Hucovic, S. Treil, A. Volberg. The Bellman functions and sharp weighted inequalities for square functions. *Operator Theory: Advances and Applications*, the volume in memory of S. A. Vinogradov, Vol. 113, Birkhauser Verlag, 2000, pp. 97-113.

[8] F. John, L. Nirenberg. On functions of bounded mean oscillation. *Comm. Pure Appl. Math.*, Vol. 14 (1961), pp. 415-426.

[9] 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.

[10] N. Krylov, *Optimal control of diffusion processes*. Springer-Verlag, New York, 1980, 308 p.

[11] J. Moser. A sharp form of an inequality by N. Trudinger. *Indiana University Mathematical Journal*, Vol. 20 (1971), no. 11, pp. 1077-1092.

[12] F. Nazarov. A counterexample to Sarason's conjecture. Preprint, MSU, 1997.

[13] 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 Analis*, Vol. 8 (1997), No. 5, pp. 32-162.

[14] F. Nazarov, S. Treil. The weighted norm inequalities for Hilbert transform are now trivial. *C. R. Acad. Sci. Paris Sér. I Math.*, Vol. 323 (1996), no. 7, pp. 717-722.

[15] 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.

[16] 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.

[17] F. Nazarov, A. Volberg. Heating of the Ahlfors-Beurling operator and estimates of its norm. *St. Petersburg Math. Journal*, Vol. 14 (2003), no. 3.

[18] F. Nazarov, A. Volberg. The Bellman function, the two-weight Hilbert transform, and embeddings of the model spaces K_θ. Dedicated to the memory of Thomas H. Wolff. *J. Anal. Math.*, Vol. 87 (2002), pp. 385-414.

[19] St. Petermichl, A. Volberg. Heating of the Ahlfors-Beurling operator: weakly quasiregular maps on the plane are quasiregular. *Duke Math. Journal*, Vol. 112 (2002), no. 2, pp. 281-305.

[20] St. Petermichl, J. Wittwer. A sharp estimate for the weighted Hilbert transform via Bellman functions. *Michigan Math. Journal*, Vol. 50 (2002), no. 1, pp. 71-87.

[21] E. Stein, *Harmonic Analysis: Real-variable Methods, Orthogonality, and Oscillatory Integrals.* Princeton University Press, Princeton, New Jersey, 1993.

[22] N. Trudinger. On imbeddings into Orlicz spaces and some applications. *Journal of Mathematics and Mechanics*, Vol. 17 (1967), no. 5, pp. 473-483.

[23] V. Vasyunin. The sharp constant in the reverse Holder inequality for Muckenhoupt weights. *Algebra i Analis*, Vol. 15 (2003), no. 1, pp. 73-117.

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