ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 41, Number 3, 2011

SURVEY ARTICLE:
A USER'S GUIDE TO BELLMAN FUNCTIONS

JANINE WITTWER

CONTENTS

1. Introduction. The Bellman function method has been around for quite some time in mathematics, but not under that name, and not in the form in which I present it here. Its ideas are used in the theory of optimal control of stochastic processes, closely connected with Bellman's principle. For a detailed and interesting description of how the method I present here ties in with stochastic processes, see [**7**]. In Harmonic Analysis, the Bellman method probably made its first appearance in [**2**], in which Burkholder proves sharp estimates on martingale transforms. Then Nazarov, Treil and Volberg used it in [**6**], and gave it the name "Bellman function method" in honor of its use in control theory.

In this paper, I will describe how the Bellman function method can be used to prove bounds on sums indexed by dyadic intervals. I begin with

Received by the editors on October 15, 2007, and in revised form on September 28, 2010.
DOI:10.1216/RMJ-2011-41-3-631

a very simple example, then describe a step-by-step procedure. A word of warning to those who have been led to believe that this method is a panacea: In the Bellman function method, we find an upper bound on a dyadic sum by finding the so-called Bellman function corresponding to this sum. Once we have the Bellman function, the rest follows easily. However, finding the function can be tricky. While there are some tools we can use to help us, fundamentally we are forced to guess at the right function. This can be a time consuming and frustrating process. That being said, the numerous questions that have been answered with this method, many of which were open for many years, show that the method is certainly worth while.

Note. I recommend that after reading this introduction to Bellman functions, the reader also read [**7**] and possibly [**4**]. They are more advanced than this paper, and cover a wider variety of applications. [**7**], which explains the connection of the Bellman function method to stochastic control theory is written in a very accessible way, even for the reader with no prior knowledge of stochastic integrals. It works through many examples, often from both the stochastic control and the Bellman method points of view. In some examples, the Bellman function method is used to bound quantities where no sum seems to be involved, something not touched in this paper. [**4**] also contains an introduction to the Bellman method, but then moves quickly into new estimates of a variety of singular integral operators in scalar and matrix weighted L^p spaces.

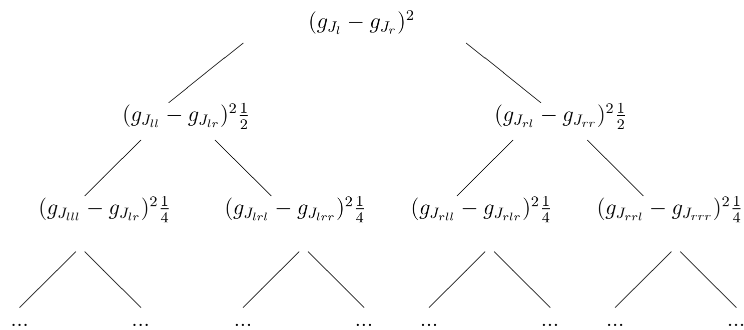
1.1. A simple example. The easiest way to get a feeling for how the method works is to look at an example. We will need some notation for this example: Let I denote a dyadic interval, i.e., an interval of the form $[k2^i, (k+1)2^i)$ where k and i are integers. I's halves, which are again dyadic intervals, are denoted by I_l (left half) and I_r (right half). g_I denotes the average of the function g over the interval I, ie $g_I = 1/|I| \int_I g(x)\, dx$.

Imagine that you wanted to prove the following theorem:

Theorem. *Let g be a positive function bounded above by 1, and J any dyadic interval. Then*

$$\sum_{I \subseteq J} (g_{I_l} - g_{I_r})^2 \frac{|I|}{|J|} \leq C.$$

One interesting aspect of this sum is that it is indexed by dyadic intervals, which have a binary structure. Thus, we can picture the terms of the sum in the following way:



Now, imagine that we have a tool that starts at the top level of this tree and bounds the term $(g_{J_l} - g_{J_r})^2$ by C. Then our tool traverses the tree to the next lower level, adds the terms there to the term we already have, and again bounds the new sum by C. So now we have $(g_{J_l} - g_{J_r})^2 + (1/2)(g_{J_{l_l}} - g_{J_{l_r}})^2 + (1/2)(g_{J_{r_l}} - g_{J_{r_r}})^2 \le C$. Assume that we can always traverse from one level to the next lower one, add the terms on that level, and not increase the bound. That means that we can bound all partial sums $\sum_{I \subseteq J, |I| > 2^n} (g_{I_l} - g_{I_r})^2 |I|/|J|$ by C, and thus the whole sum is bounded by C. This is exactly how the Bellman method works. It lets us transition from one level to the next, adding terms on the way, and never increasing the bound.

Let us see how this is done.

Proof. Let $B(x) = x - x^2$, $x \in [0,1]$. It is easy to verify that (for C=1)

1. $C \ge B(x) \ge 0$.

2. $B(x) \ge c(x_- - x_+)^2 + [B(x_-) + B(x_+)]/2$ for $x_-, x_+ \in [0,1]$ such that $x = (x_- + x_+)/2$.

This function B is the tool we are looking for. We refer to it as the Bellman function of the problem. Here is how we use it: Note that all averages of g are between 0 and 1, and that $g_J = (1/2)(g_{J_l} + g_{J_r})$. Thus we can let $x_- = g_{J_l}$, $x_+ = g_{J_r}$ and $x = g_J$. Then, by the above inequalities for $B(x)$,

$$C \geq B(g_J) \geq c\,(g_{J_l} - g_{J_r})^2 + \frac{B(g_{J_l}) + B(g_{J_r})}{2}$$

(Since $[B(g_{J_l}) + B(g_{J_r})]/2$ is positive, this shows that $C \geq c(g_{J_l} - g_{J_r})^2$.) Now we will apply the difference inequality again, this time to $B(g_{J_l})$ and $B(g_{J_r})$, with $x_- = g_{J_{l_l}}$, $x_+ = g_{J_{l_r}}$ and $x = g_{J_l}$, and $x_- = g_{J_{l_l}}$, $x_+ = g_{J_{r_r}}$ and $x = g_{J_r}$, respectively. Then

$$\begin{aligned} C &\geq c\,(g_{J_l} - g_{J_r})^2 + \frac{B(g_{J_l}) + B(g_{J_r})}{2} \\ &\geq c\,(g_{J_l} - g_{J_r})^2 + \frac{c}{2}\left(g_{(J_l)_r} - g_{(J_l)_l}\right)^2 + \frac{c}{2}\left(g_{(J_r)_l} - g_{(J_r)_r}\right)^2 \\ &\quad + \frac{B(g_{(J_l)_l}) + B(g_{(J_l)_r}) + B(g_{(J_r)_l}) + B(g_{(J_r)_r})}{4} \end{aligned}$$

and so

$$C \geq c\,(g_{J_l} - g_{J_r})^2 + \frac{c}{2}\left(g_{(J_l)_r} - g_{(J_l)_l}\right)^2 + \frac{c}{2}\left(g_{(J_r)_l} - g_{(J_r)_r}\right)^2.$$

Repeat this procedure n times, each time applying the inequality to the B's on the right, expressing $B(g_M)$ in terms of $B(g_{M_l})$, $B(g_{M_r})$ and $(g_{M_l} - g_{M_r})^2$.

We get

$$\begin{aligned} C \geq c &\sum_{I \subseteq J, |I| \geq 2^{-n}|J|} (g_{I_l} - g_{I_r})^2 \frac{|I|}{|J|} \\ &+ \sum_{\substack{K \subset J \\ |K| = 2^{-n-1}|J|}} B(g_K)\left(\frac{1}{2}\right)^{n+1}. \end{aligned}$$

The $|I|/|J|$ stems from the factors $1/2$ that we get with each repetition.

Since $B(g_K)$ is always positive, we can simply omit the second sum without changing the inequality, i.e.,

$$C \geq c \sum_{\substack{I \subseteq J \\ |I|/|J| \geq 2^{-n}}} (g_{I_l} - g_{I_r})^2 \frac{|I|}{|J|}.$$

Thus the partial sums are uniformly bounded. Letting $n \to \infty$ establishes

$$C' \geq \sum_{I \subseteq J} (g_{I_l} - g_{I_r})^2 \frac{|I|}{|J|},$$

which is what we set out to prove. $\square$

This simple example already captures the essence of the Bellman function method. We will analyze the steps involved, and discuss how to find the right Bellman function in the following sections.

2. Notation. All intervals in this paper will be understood to be dyadic, i.e. of the form $[k2^i, (k+1)2^i)$, with $i, k \in \mathbf{Z}$. The set of dyadic intervals will be denoted by D. If $I \in D$ is one such interval, then $I_l, I_r \in D$ denote the left and right half of I respectively. Let f_I denote $1/|I| \int_I f(x)\,dx$ for I a dyadic interval. Several of my examples use dyadic A_2 weights. A dyadic A_2 weight is a non-negative function w such that $w_I^{-1} w_I = 1/|I| \int_I 1/w(x)\,dx \; 1/|I| \int_I w(x)\,dx \leq C$ uniformly for all intervals $I \in D$. I will use $\mu(I) = w_I w_I^{-1}$ and $\|w\|_{A_2} = \sup_{I \in D} \mu(I)$. As is customary, I will refer to $\|w\|_{A_2}$ as the A_2 norm of w, even though it is not actually a norm. C will denote a constant, though not necessarily the same one at each occurrence.

3. The general situation. In general, we use the method of Bellman functions to establish bounds on sums of the form

$$\frac{1}{|J|} \sum_{I \subseteq J} f(u_I, u_{I_l}, u_{I_r}) |I|,$$

where f is a positive function of $\mathbf{R}^{3n}$, $u_I \in \mathbf{R}^n$, and we have some control over $u_I - (u_{I_l} + u_{I_r})/2$. The components of u_I are typically averages of functions, or elements of sequences. In the example from the introduction, $f(u_I, u_{I_l}, u_{I_r}) = (u_{I_l} - u_{I_r})^2$, and u_I has just one component, namely, $u_I = 1/|I| \int_I g$.

We can think of the Bellman method as splitting into three components:

1. Given a particular sum to bound, specify what properties the Bellman function needs to satisfy.

2. Find a function which satisfies the specifications.
3. Use the function to run the Bellman function proof.

The third task is the easiest, as it always follows the same pattern as in the introduction. We will spend some time discussing tasks 1 and 2.

3.1. What does a Bellman function need to satisfy? A Bellman function needs to satisfy two inequalities. The first one, the boundedness condition, is typically simple, and takes the form

$$0 \leq B(v) \leq X,$$

where X varies according to the situation and represents the quantity with which we wish to bound our sum. The variable v represents the as of yet unspecified variable that the Bellman function depends on.

The second inequality, which we shall refer to as the difference inequality, takes the form

$$B(v) - \frac{B(v_-) + B(v_+)}{2} \geq f(v, v_-, v_+).$$

(Recall that our goal is to bound $1/|J| \sum_{I \subseteq J} f(u_I, u_{I_l}, u_{I_r})|I|$.) Later, when running the Bellman tool, we will plug u_I, u_{I_l} and u_{I_r} in for v, v_- and v_+, respectively. As we will see shortly, u_I, u_{I_l} and u_{I_r} cannot take on any old set of values. Any inequalities for $B(v)$ need only be satisfied when v, v_-, v_+ represent values that are possible for u_I, u_{I_l} and u_{I_r}. There are two particular instances of this:

Often, there are relationships between the individual components of u_I. For example, if $u_I = (f_I, (f^2)_I)$, then by the Cauchy-Schwarz inequality, $(f_I)^2 \leq (f^2)_I$. Thus, if $v = (v_1, v_2)$, it is enough for our purposes if $B(v)$ exists only when $v_1^2 \leq v_2$. We look for a Bellman function $B(v)$ with domain $\{v = (v_1, v_2) : (v_1)^2 \leq v_2\}$.

Furthermore, there is always a relationship between u_I , u_{I_l} and u_{I_r}. $B(v)$ will only need to satisfy the difference inequality when v, v_- and v_+ satisfy the same conditions. For example, in our introductory example, $u_I = g_I$ and $g_I = (g_{I_l} + g_{I_r})/2$. Thus the difference inequality, $B(v) \geq c(v_- - v_+)^2 + (B(v_-) + B(v_+))/2$, need only be true for v_-, v_+ such that $v = (v_- + v_+)/2$.

To get a feeling for what quantities we will include in v, let us take a look at the following examples.

3.1.1. Example 1. The following was a crucial lemma in [**11**]. We will reconstruct here how this lemma was proven. See the section on notation for a definition of A_2.

Lemma. *For w in A_2,*

$$\frac{1}{|J|}\sum_{I\subseteq J}\left|\frac{w_{I_l}-w_{I_r}}{w_I}\right|\left|\frac{w^{-1}_{I_l}-w^{-1}_{I_r}}{w^{-1}_I}\right|w_I|I|\le C\|w\|_{A_2}w_J.$$

Let us first decide on the variables for the Bellman function. In order to express the summand, we will need $v=(x,y)$: x to correspond to w_I and y to correspond to w_I^{-1}. Then $w_{I_l}, w_{I_r}, w^{-1}_{I_l}$ and $w^{-1}_{I_r}$ can then be expressed as x_-, x_+, y_- and y_+, respectively. In this example, $f(x,y,x_-,x_+,y_-,y_+)=|(x_- - x_+)/x||(y_- - y_+)/y|x$.

Since w_I and w_I^{-1} are not independent of each other, we will need to restrict the domain to reflect that. By the A_2 condition, $w_Iw_I^{-1}\le A$ for some constant A. In addition, the Cauchy-Schwarz inequality shows that $1\le 1/|I|\int_I w^{1/2}w^{-1/2}\le [1/|I|\int_I w1/|I|\int_I w^{-1}]^{1/2}$.

Thus,

$$1\le w_Iw_I^{-1}\le A.$$

Correspondingly, the domain for the Bellman function is $\{(x,y):1\le xy\le A\}$. When we plug in the variables, we also plug in $\|w\|_{A_2}$ for A.

Note that $w_J=(w_{J_l}+w_{J_r})/2$, $w_J^{-1}=(w^{-1}_{J_l}+w^{-1}_{J_r})/2$, and so the difference inequality for $B(x,y)$ must be fulfilled for (x,y), (x_-,y_-) and (x_+,y_+) which satisfy $x=(x_-+x_+)/2$, and $y=(y_-+y_+)/2$.

The upper bound for the Bellman function is CAx, since we wish to bound the sum by $C\|w\|_{A_2}w_J$. (The upper bound is applied only at the very first step, when the interval is J.)

In summary, $B(x,y)$ must satisfy (for some constants C_1, C_2)

1. $0\le B(x,y)\le C_1Ax$ on $\{x,y>0; 1\le xy\le A\}$
2. $$B(x,y)-\frac{B(x_-,y_-)+B(x_+,y_+)}{2}\ge C_2\left|\frac{(x_- - x_+)}{x}\frac{(y_- - y_+)}{y}x\right|$$

where $x = (x_- + x_+)/2$ and $y = (y_- + y_+)/2$, and $(x, y), (x_-, y_-)$, (x_+, y_+) are in the domain $\{x, y > 0; 1 \le xy \le A\}$.

This completes step 1 of this Bellman function proof. We will discuss step 2, tools for finding the function, in a later section. Assuming that we have found such a Bellman function, we can now run step 3 to prove the lemma. Note that the proof does not depend on the function, only on its existence.

Proof. Let $x = w_J, y = w_J,\ x_- = w_{J_l},\ x_+ = w_{J_r},\ y_- = w_{J_l}^{-1}$, $y_+ = w_{J_r}^{-1}$ and $A = \|w\|_{A_2}$.

Plug these quantities into the Bellman function, and use the two inequalities to get

$$\begin{aligned} &C_1\|w\|_{A_2} w_J \ge B(w_J, w_J^{-1}) \\ &\ge C_2\left|\frac{w_{J_l} - w_{J_r}}{w_J}\right|\left|\frac{w_{J_l}^{-1} - w_{J_r}^{-1}}{w_J^{-1}} w_J\right| + \frac{B(w_{J_l}, w_{J_l}^{-1}) + B(w_{J_r}, w_{J_r}^{-1})}{2}. \end{aligned}$$

Now $1 \le w_{I_l} w_{I_r}^{-1} \le \|w\|_{A_2}$ and $1 \le w_{I_r} w_{I_r}^{-1} \le \|w\|_{A_2}$ again, so we can use the difference inequality again, this time to the two Bellman functions on the right hand side.

$$\begin{aligned} C_1\|w\|_{A_2} w_J \ge C_2&\left|\frac{w_{J_l} - w_{J_r}}{w_J}\right|\left|\frac{w_{J_l}^{-1} - w_{J_r}^{-1}}{w_J^{-1}}\right| w_J \\ &+ \frac{1}{2}\left(C_2\left|\frac{w_{J_{l,l}} - w_{J_{l,r}}}{w_{J_l}}\right|\left|\frac{w_{J_{l,l}}^{-1} - w_{J_{l,r}}^{-1}}{w_{J_l}^{-1}}\right| w_{J_l}\right. \\ &+ \left. C_2\left|\frac{w_{J_{r,l}} - w_{J_{r,r}}}{w_{J_r}}\right|\left|\frac{w_{J_{r,l}}^{-1} - w_{J_{r,r}}^{-1}}{w_{J_r}^{-1}}\right| w_{J_r}\right) \\ &+ \frac{B(w_{J_{l,l}}, w_{J_{l,l}}^{-1}) + B(w_{J_{l,r}}, w_{J_{l,r}}^{-1})|}{4} \\ &+ \frac{B(w_{J_{r,l}}, w_{J_{r,l}}^{-1}) + B(w_{J_{r,r}}, w_{J_{r,r}}^{-1})}{4}. \end{aligned}$$

This process can be repeated as often as we want. After n iterations, we have the following formula:

$$C_1\|w\|_{A_2} w_J \geq C_2 \sum_{\substack{I\subseteq J \\ |I|\geq 2^{-n}|J|}} \left|\frac{w_{I_l}-w_{I_r}}{w_I}\right|\left|\frac{w_{I_l}^{-1}-w_{I_r}^{-1}}{w_I^{-1}}\right| w_I \frac{|I|}{|J|}.$$

Letting $n \to \infty$, we have the desired result. □

3.1.2. Example 2. We wish to prove

$$\frac{1}{|J|}\sum_{I\subseteq J}\left(\frac{w_{I_l}-w_{I_r}}{w_I}\right)^2 |I| \leq C = C(\|w\|_{A_2})$$

where $w \in A_2$.

We need to express the fact that $w \in A_2$. Since the defining characteristic of A_2 weights is that $w_I w_I^{-1} \leq A$ for some constant A, we let x represent w_I, y represent w_I^{-1}, and we require that $1 \leq xy \leq A$. The summand is $f(x, x_-, x_+, y, y_-, y_+) = ((x_- - x_+)/x)^2$. Note that y is a hidden variable–it was not apparent from the sum that we would need it. The relationship between x and y is the same as in the previous example.

In summary, we need to find a function $B(x, y)$ in the domain $\{x, y > 0; 1 \leq xy \leq A\}$, such that the following two conditions hold:

1. $0 < B(x, y) \leq C$, (C may depend on A.)

2. $B(x, y) - [B(x_-, y_-) + B(x_+, y_+)]/2 \geq ((x_- - x_+)/x)^2$ whenever $x = (x_- + x_+)/2$, $y = (y_- + y_+)/2$ and $(x_-, y_-), (x_+, y_+), (x, y)$ in the domain.

We will demonstrate the existence of such a function in a later section. Given a function $B(x, y)$ as described above, substitute w_J for x, w_J^{-1} for y, and use the inequalities. Then

$$\begin{aligned}
C &\geq B(w_J, w_J^{-1}) \\
&\geq \left(\frac{w_{J_l}-w_{J_r}}{w_J}\right)^2 + \frac{B(w_{J_l}, w_{J_l}^{-1}) + B(w_{J_r}, w_{J_r}^{-1})}{2} \\
&\geq \left(\frac{w_{J_l}-w_{J_r}}{w_J}\right)^2 + \frac{1}{2}\left(\frac{w_{J_{l,l}}-w_{J_{l,r}}}{w_{J_l}}\right)^2
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{2}\left(\frac{w_{J_{r,l}}-w_{J_{r,r}}}{w_{J_r}}\right)^2 \\
&+\frac{B(w_{J_{l,l}},w_{J_{l,l}}^{-1})+B(w_{J_{l,r}},w_{J_{l,r}}^{-1})+B(w_{J_{r,l}},w_{J_{r,l}}^{-1})+B(w_{J_{r,r}},w_{J_{r,r}}^{-1})}{4} \\
\geq{}& \cdots \\
\geq{}& \frac{1}{|J|}\sum_{\substack{I\subseteq J\\ |I|/|J|\geq 2^{-n}}}\left(\frac{w_{I_l}-w_{I_r}}{w_I}\right)^2|I|.
\end{aligned}
$$

As before, the last step follows by using the difference inequality n times and dropping the Bellman function terms, which we can do since they are positive. Let $n\to\infty$. We have proven that

$$
\frac{1}{|J|}\sum_{I\subseteq J}\left(\frac{w_{I_l}-w_{I_r}}{w_I}\right)^2|I|\leq C
$$

3.1.3. Example 3. This example is more difficult, and stems from [**6**]. It is the proof of the Carleson embedding theorem using Bellman function methods.

The dyadic Carleson embedding theorem can be written as

$$
\frac{1}{|J|}\sum_{I\subseteq J}\mu_I(\phi_I)^2|I|\leq C\frac{1}{|J|}\int_J\phi^2,
$$

where $\{\mu_I\}$ is a normalized Carleson sequence, i.e., $\sum_{I\subseteq K}\mu_I|I|\leq|K|$ for any dyadic interval K, and $\mu_I\geq 0$. ϕ is a positive function in $L^2(dx)$.

Note. In the Bellman function method, we normally work with positive functions, and then extend the result, if possible and relevant, to all functions by linearity or sublinearity. This makes the variables positive, and thus easier to work with.

Our first step is to decide what variables B will depend on. In order to express the summand, we need to be able to express μ_I and $\phi_I=1/|I|\int_I\phi$.

Let x correspond to $\phi_I = 1/|I| \int_I \phi$. Then since $\phi_I = (\phi_{I_l} + \phi_{I_r})/2$, we will use $x = (x_- + x_+)/2$ in the difference equation.

μ_I is more tricky. The obvious choice would be to let y correspond to μ_I. However, if we did that, we would have no information about the relation between μ_I, μ_{I_l}, and μ_{I_r}. Furthermore, there would be no easy way of incorporating the Carleson condition. Thus the authors of [**6**] made y correspond to $M(I) = 1/|I| \sum_{K \subseteq I} \mu_K |K|$. Then

$$\begin{aligned} M(I) - \frac{M(I_l) + M(I_r)}{2} &= \frac{1}{|I|} \sum_{K \subseteq I} \mu_K |K| \\ &\quad - \frac{1/|I_l| \sum_{K \subseteq I_l} \mu_K |K| + 1/|I_r| \sum_{K \subseteq I_r} \mu_K |K|}{2} \\ &= \mu_I \geq 0. \end{aligned}$$

The Carleson condition is thus described by $y \leq 1$ and $y - (y_- + y_+)/2 \geq 0$.

Note that $1/|J| \int_J \phi^2$, while not appearing in the sum itself, does appear in the bound, so we also need a variable corresponding to it. We call this variable z. Again we choose $z = (z_- + z_+)/2$ to mimic the behavior of $1/|J| \int_J \phi^2$.

The Cauchy-Schwartz inequality implies that $((1/|J|) \int_J \phi)^2 \leq 1/|J| \int_J (\phi)^2$. This imposes the condition $x^2 \leq z$.

The summand is $f(x, x_-, x_+, y, y_-, y_+, z, z_-, z_+) = (y - (y_- + y_+)/2)x^2$.

In summary: We need to prove the existence of a function $B(x, y, z)$ on the domain $\{(x, y, z) : x, y, z > 0; y \leq 1, x^2 \leq z\}$ such that

1. $0 \leq B(x, y, z) \leq Cz$.
2. $B(x, y, z) - [B(x_-, y_-, z_-) + B(x_+, y_+, z_+)]/2 \geq (y - (y_- + y_+)/2)x^2$ for all (x, y, z), (x_-, y_-, z_-), (x_+, y_+, z_+) in the domain with $y - (y_- + y_+)/2 \geq 0$, $x = (x_- + x_+)/2$ and $z = (z_- + z_+)/2$.

Assume that we have found a $B(x, y, z)$ satisfying these conditions. Let $x = \phi_J$, $x_- = \phi_{J_l}$, $x_+ = \phi_{J_r}$, $y = M(J)$, $y_- = M(J_l)$, $y_+ = M(J_r)$, $z = (\phi^2)_J$, $z_- = (\phi^2)_{J_l}$, $z_+ = (\phi^2)_{J_r}$.

Using the inequalities,

$$\begin{aligned}
C(\phi^2)_J &\geq B(M(J),(\phi^2)_J,\phi_J)\\
&\geq \left(M(J)-\frac{M(J_l)+M(J_r)}{2}\right)(\phi_J)^2\\
&\quad+\frac{B(M(J_l),(\phi^2)_{J_l},\phi_{J_l})+B(M(J_r),(\phi^2)_{J_r},\phi_{J_r})}{2}\\
&=\mu_J(\phi_J)^2+\frac{B(M(J_l),(\phi^2)_{J_l},\phi_{J_l})+B(M(J_r),(\phi^2)_{J_r},\phi_{J_r})}{2}.
\end{aligned}$$

Repeat the inequality n times and drop the Bellman functions to get

$$C(\phi^2)_J \geq \sum_{\substack{I\subseteq J\\ |I|/|J|\geq 2^{-n}}} \mu_I(\phi_I)^2\frac{|I|}{|J|},$$

and let $n\to\infty$.

3.2. Step 2: Finding the right function. Now we turn to the trickiest step in a Bellman function proof, finding the Bellman function. In some cases it may be possible to prove the existence of B without ever explicitly finding it. In general, however, we will need to find a specific function and prove that it satisfies the inequalities. This will usually involve a fair amount of educated guessing. Bellman functions for related problems will be similar, so that a good starting point for finding a Bellman function is to look at previously used Bellman functions. Sometimes, all that is needed is a small modification. What we will discuss in this section is how to transform the difference condition into a differential condition which, while not necessarily easier to solve, is considerably easier to verify. Once we have a starting guess, we can use the differential condition to see if it is the right function, and, if not, get an idea of what modifications are needed.

Let $v\in\mathbf{R}^n$. Assuming for the moment that B is C^2, we write

$$B(v)-\frac{B(v_-)+B(v_+)}{2}=-\frac{1}{2}\left(B(v_-)-B(v)+B(v_+)-B(v)\right)$$

and expand $B(v_-)$ and $B(v_+)$ in power series around v with second order remainders.

$$B(v_-)=B(v)+DB(v)(v_--v)+(v_--v)\frac{D^2B(\xi_1)}{2!}(v_--v)^t,$$

and similarly for $B(v_+)$. Here ξ_1, ξ_2 are points on the line segments between v_- and v, and v_+ and v, respectively.

Putting this all together we have

$$\begin{aligned} &B(v) - \frac{B(v_-)+B(v_+)}{2} \\ &= -1/2\Big[DB(v)(v_- - v) + DB(v)(v_+ - v) \\ &\quad + (v_- - v)\frac{\big(D^2B(\xi_1)\big)}{2!}(v_- - v)^t + (v_+ - v)\frac{\big(D^2B(\xi_2)\big)}{2!}(v_+ - v)^t\Big] \\ &= DB(v)\left(v - \frac{v_- + v_+}{2}\right) \quad -\frac{1}{2}\Big[(v_- - v)\frac{\big(D^2B(\xi_1)\big)}{2!}(v_- - v)^t \\ &\qquad\qquad + (v_+ - v)\frac{\big(D^2B(\xi_2)\big)}{2!}(v_+ - v)^t\Big]. \end{aligned}$$

Recall that we assumed that v was such that we had information about $v - (v_- + v_+)/2$. This is where we need to use it. If $v - (v_- + v_+)/2 = 0$, as in many of the examples we are working with, the first order terms cancel, and we are left with

$$\begin{aligned} &B(v) - \frac{B(v_-)+B(v_+)}{2} \\ &= -\frac{1}{2}(v_- - v)\frac{\big(D^2B(\xi_1)\big)}{2!}(v_- - v)^t + (v_+ - v)\frac{\big(D^2B(\xi_2)\big)}{2!}(v_+ - v)^t. \end{aligned}$$

Since ξ_1 and ξ_2 are not known, we typically seek B with conditions on the matrix of second derivatives which hold for any possible ξ_1 and ξ_2.

Given a candidate for B, it is normally considerably more straightforward to estimate the differential expression for any possible ξ_1 and ξ_2 than the difference equation for any v_- and v_+. Even in cases where not all the first order terms cancel, the differential equation is much easier to work with than the difference inequality.

Note. To formulate the differential condition, we assumed that B is C^2. This is not a real restriction - if there is a B satisfying the difference condition, we can usually mollify it to get a B in C^2 which satisfies the differential condition.

3.2.1. Example 1 continued. We now turn to finding the Bellman function for our first example.

In Step 1, we determined that what we need is a Bellman function that satisfies:

1. $0 \le B(x,y) \le C_1 Ax$ on $\{x,y>0; 1\le xy\le A\}$.
2.
$$\begin{aligned} B(x,y) - \frac{B(x_-,y_-)+B(x_+,y_+)}{2} &\ge C_2 \left| \frac{(x_- - x_+)}{x}\frac{(y_- - y_+)}{y} x \right| \\ &= C_2 \left| (x_- - x_+)(y_- - y_+)\frac{1}{y} \right|, \end{aligned}$$

where $x = (x_- + x_+)/2$ and $y = (y_- + y_+)/2$, and (x,y), (x_-,y_-), (x_+,y_+) are in the domain $\{x,y>0; 1\le xy\le A\}$.

In this example, $v = (x,y)$ with $x = (x_- + x_+)/2$, and $y = (y_- + y_+)/2$. Thus the first derivative terms in the power series expansion cancel, and, since $x_- - x = (x_- - x_+)/2$, etc., we are left with

$$\begin{aligned} &\left(B(x,y) - \frac{B(x_-,y_-)+B(x_+,y_+)}{2} \right) \\ &\qquad = -\frac{1}{2}\left(\frac{x_- - x_+}{2}, \frac{y_- - y_+}{2} \right) \\ &\qquad\quad \times \frac{\left(D^2B(\xi_1,\zeta_2) + D^2B(\xi_2,\zeta_2)\right)}{2!} \left(\frac{x_- - x_+}{2}, \frac{y_- - y_+}{2} \right)^t . \end{aligned}$$

Thus we need a Bellman function that satisfies

$$\begin{aligned} (1) \qquad & -\frac{1}{2}\left(\frac{x_- - x_+}{2}, \frac{y_- - y_+}{2} \right) \frac{\left(D^2B(\xi_1,\zeta_1) + D^2B(\xi_2,\zeta_2)\right)}{2!} \\ &\qquad \times \left(\frac{x_- - x_+}{2}, \frac{y_- - y_+}{2} \right)^t \\ &\qquad\qquad \ge C \left| (x_- - x_+)(y_- - y_+)\frac{1}{y} \right| . \end{aligned}$$

Since we do not know the actual values of (ξ_1, ζ_1) or (ξ_2, ζ_2), we need to make sure that the differential condition is satisfied if at all possible $(\xi_1, \zeta_1), (\xi_2, \zeta_2)$. We have $1 \le x_- y_- \le A$, $1 \le x_+ y_+ \le A$ and $1 \le xy \le A$. Also, $x_- = 2x - x_+ \le 2x$, and similar for x_+, y_-, y_+. Without loss of generality, $x_+ \le \xi \le x_-$ and either $y_+ \le \zeta \le y_-$, or $y_- \le \zeta \le y_+$. In the first case, $1 \le x_+ y_+ \le \xi\zeta \le x_- y_- \le A$. In the second case, $\xi\zeta \le x_- y_+ \le 2x2y \le 4A$ and $\xi\zeta \ge x_+ y_- \ge x_+ y_+ \ge 1$. Thus it is always true that $1 \le \xi\zeta \le 4A$, and we will estimate the matrix of second derivatives in that domain. (A more careful analysis would show that in fact $1 \le \xi\zeta \le 2A$, but we do not need that here.)

Note that 1 can be restated as

$$
\begin{aligned}
&(x_- - x_+, y_- - y_+)\left(D^2 B(\xi_1, \zeta_1) + D^2 B(\xi_2, \zeta_2)\right)(x_- - x_+, y_- - y_+)^t \\
&\quad \ge \left|(x_- - x_+, y_- - y_+)\begin{pmatrix} 0 & C/y \\ C/y & 0 \end{pmatrix}(x_- - x_+, y_- - y_+)^t\right| \\
&\quad \sim \left|(x_- - x_+, y_- - y_+)\begin{pmatrix} 0 & C'/\zeta_i \\ C'/\zeta_i & 0 \end{pmatrix}(x_- - x_+, y_- - y_+)^t\right|,
\end{aligned}
$$

where the last line follows from the fact that $C/y \sim C/\zeta_i$, since $y/2 \le y_- \le 2y$ and $y/2 \le y_+ \le 2y$, and so $y/2 \le \zeta_i \le 2y$.

Using matrix notation, and splitting the absolute value inequality into two inequalities, we see that it suffices to find a function such that the following two matrixes are positive semi-definite on $\{\xi, \zeta > 0; 1 \le \xi\zeta \le 4A\}$:

(2)

$$
\begin{pmatrix} -B_{xx} & -B_{xy} - C'/\zeta \\ -B_{xy} - C'\zeta & -B_{yy} \end{pmatrix} \text{ and } \begin{pmatrix} -B_{xx} & -B_{xy} + C'/\zeta \\ -B_{xy} + C'/\zeta & -B_{yy} \end{pmatrix}.
$$

The second derivatives here are evaluated at (ξ, ζ).

To find such a function, it is often helpful to look at Bellman functions that have been used to prove similar estimates. Hukovikc, Treil and Volberg used the function $B(x, y) = x(-(4A/xy) - xy + 4A + 1)$ to prove a sharp weighted bound on the dyadic square function ([**3**]). I adjusted this function via educated guesses to

$$
B(x, y) = x\left(\frac{-4A}{xy} - \frac{xy}{4A} + 4A + 1\right),
$$

which satisfies our differential condition as well as the desired upper bound.

3.2.2. Example 2 continued. In Example 2 we have the same variables as in Example 1, so the same terms cancel in the power series. We thus need to find a function that satisfies

1. $0 < B(x,y) \leq C$ whenever (x,y) in the domain $1 \leq xy \leq A$. (C will depend on A.)

2. $-(1/2)((x_- - x_+)/2), ((y_- - y_+)/2))\big((D^2B(\xi_1,\xi_2) + D^2B(\xi_2,\zeta_2)\big)/2!((x_- - x_+)/2), (y_- - y_+)/2)^t \geq ((x_- - x_+)/x)^2$.

In matrix notation, this leads us to ask for B which satisfies that

$$-D^2B - \begin{pmatrix} c/x^2 & 0 \\ 0 & 0 \end{pmatrix}$$

be positive semi definite in the domain $\{1 \leq xy \leq 4A\}$. (Note that I wrote c/x^2 rather than $1/x^2$ in the second matrix. This is because, in the previous equation, derivatives were evaluated in terms of (ξ,ζ), while the right hand side was expressed in terms of x. If we wish to have the same variable on the right hand side, we need to use the fact that $x \sim \xi_i$, which introduces the constant.)

Summarizing the above, we see that the $B(x,y)$ that we are searching for satisfies

1. $\begin{pmatrix} -B_{xx} - c/x^2 & -B_{xy} \\ -B_{xy} & -B_{yy} \end{pmatrix}$ is positive semi definite in $1 \leq xy \leq 4A$.
2. $0 < B(x,y) \leq C$ in $1 \leq xy \leq A$.

We know that the upper bound should involve A, so a good guess would be to look for functions in the variable $u = xy$. Say $B(x,y) = g(u)$. What would g need to satisfy? A calculation shows that

$$-D^2g - \begin{pmatrix} c/x^2 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} -y^2g'' - c/x^2 & -ug'' + g' \\ -ug'' + g' & -x^2g'' \end{pmatrix}.$$

This will be positive semi definite if the sub-determinants are positive, i.e.

1. $-y^2g'' - \frac{c}{x^2} \geq 0$.
2. $cg'' + 2ug''g' + (g')^2 \geq 0$.

The first condition can be written as

$$u^2 g'' + c \leq 0.$$

A natural function to try would be g such that this is 0, i.e., $g''(u) = -c/u^2$, so $g(u) = c \ln(u)$. We can verify that this function satisfies the second condition, too, and thus is the function we were looking for.

3.2.3. Example 3 continued. Recall that, to finish Example 3, we need $B(x,y,z)$ on the domain $\{(x,y,z) : x,y,z > 0; y \leq 1, x^2 \leq z\}$ such that

1. $0 \leq B(x,y,z) \leq Cz$,
2. $B(x,y,z) - [B(x_-,y_-,z_-) + B(x_+,y_+,z_+)]/2 \geq ((y - (y_- + y_+)/2))x^2$ for all $y - ((y_- + y_+)/2) \geq 0$, $x = ((x_- + x_+)/2)$ and $z = ((z_- + z_+)/2)$, when (x,y,z), (x_-,y_-,z_-), (x_+,y_+,z_+) are in the domain.

We will again expand in a second degree power series. This time $(2y - y_- - y_+)$ is non-negative, but not necessarily 0, so the first derivative terms in y will not cancel each other. The terms in x and z do. We have

$$\begin{aligned}
&2\left(B(x,y,z) - \frac{B(x_-,y_-,z_-) + B(x_+,y_+,z_+)}{2}\right)\\
&= \frac{\partial B(x,y,z)}{\partial y}(2y - y_- - y_+)\\
&\quad - (x_- - x, y_- - y, z_- - z) D^2 B(\xi_1, \zeta_1, \eta_1)(x_- - x, y_- - y, z_- - z)^t\\
&\quad - (x_+ - x, y_+ - y, z_+ - z) D^2 B(\xi_2, \zeta_2, \eta_2)(x_+ - x, y_+ - y, z_+ - z)^t,
\end{aligned}$$

where (ξ_1, ζ_1, η_1) and (ξ_2, ζ_2, η_2) are again points on the lines between (x,y,z) and (x_-,y_-,z_-), (x_+,y_+,z_+), respectively.

Since we want this to be $\geq 2(y - ((y_- + y_+)/2))x^2 = (2y - y_- - y_+)x^2$, the conditions become

1. $\partial B/\partial y \geq x^2$,
2. $-D^2 B$ positive semi definite for $(x,y,z) \in \{x,y,z \geq 0; y \leq 1; x \leq z^{1/2}\}$,
3. $0 \leq B(x,y,z) \leq cz$ for $(x,y,z) \in \{x,y,z \geq 0; y \leq 1; x \leq z^{1/2}\}$.

Note that the domain of the Bellman function is convex this time, so we know that (ξ_i, ζ_i, η_i) are in the domain whenever (x, y, z), (x_-, y_-, z_-), (x_+, y_+, z_+) are.

Again, finding the right Bellman function will most likely require many educated guesses. A first step might be to try $z - x^2$, since it has the correct bounds on the domain and is concave. However, it does not satisfy the first derivative condition. Multiplying the $-x^2$ by $-y$ would solve that, but would destroy the concavity of the function. So we might try to divide $-x^2$ by y instead, or rather by $1 + y$, since the function could become unbounded otherwise. So our next guess might be $z - (x^2/(1 + y))$. This leads to the function used in [**6**], $4(z - (x^2/(1 + y)))$.

Note that we needed two differential conditions in this case. This is because the condition on $\partial B/\partial y$ does not control the entire first derivative, and is thus not enough to control the remainder.

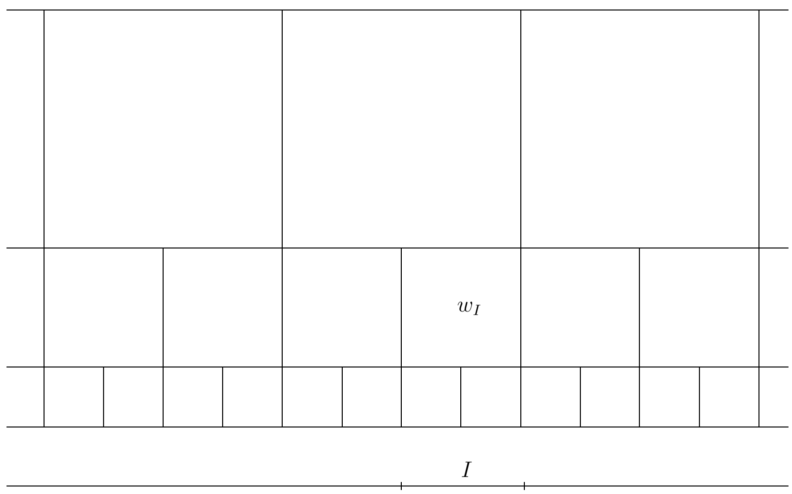
3.3. The continuous case. Frequently, we look at dyadic operators because they are easier to work with than their continuous counterparts–but once we've bounded the dyadic version, how do we get back to the continuous operator? In this section, I would like to give the reader just a hint of how we can sometimes transition from the dyadic to the continuous world. Example 1 above was a lemma in [**11**], where it was used to control the martingale transform, a dyadic analog of the Hilbert transform. In [**9**], Petermichl and I transformed this dyadic lemma into a continuous one in order to control the Hilbert transform. As you will see below, the proof of the continuous case is very similar to the dyadic one. We restate the difference condition on B as a differential condition. The transition from the coarsest level (corresponding to w_J) to the finest levels (corresponding to w_I for very small I's.) is done by using Green's theorem.

In our continuous world, our weight function ω is defined on the Taurus, $\mathbf{T}$ and extended to the disk by harmonic extension, i.e., $\omega(z) = \int \omega(t) P_z(t)\, dm(t)$, where $P_z(t) = 1 - |z|^2/|1 - \overline{z}t|^2$.

We consider ω to be an $\mathbf{A}_2$ weight if

$$\sup_{z\in\mathbf{D}} \omega(z)\omega^{-1}(z) = Q_2^{\mathrm{inv}}(\omega)$$

is finite.



Note that since the harmonic extension is an average of the values on the boundary, $\omega(z)$ is an average just like our w_I in the dyadic case. When z is the center of the disk, the average is over the whole taurus, (corresponding to averaging over a very large dyadic interval I), while if is z close to the boundary, the average is more localized, (w_I over a very small interval). In that light, you can see that the above definition of continuous A_2 is analog to our discrete one.

To provide some more intuition for why we can link dyadic averages to harmonic extensions, I would like to point out that the collection of dyadic averages of w can be thought of as a discrete extension of w to the upper half plane: Divide the upper half plane into squares as in the picture below. Each dyadic interval I is associated with a square of size $I \times I$ which starts at height $|I|$ above it. I have only drawn three sizes of squares–picture increasingly small squares below my lowest row, becoming infinitely small as they approach the x axis, and increasingly large squares above my highest row. We think of the extension of w being the function that is assigned the value w_I on the square associated with the interval I.

You can see that this extension is not so far off from forming a discrete harmonic extension of w to the upper half plane, in that the values of the extension are averages of the closest boundary values. Also, as we approach a particular point on the boundary, the values of the extension approach the value of w at that point.

Keeping in mind that the upper half plane is just a Möbius transform away from the disk, you can see that the harmonic extension of ω on the taurus really has a lot in common with our dyadic averages.

The analogue to Example 1 in the continuous case is:

Lemma 3.1. *If $\omega \in \mathbf{A}_2$*

$$\int_{\mathbf{D}} \frac{|\omega(\xi)'||\omega^{-1}(\xi)'|}{\omega(\xi)\omega^{-1}(\xi)}\omega(\xi)\log\frac{1}{|S_z(\xi)|}\,dA(\xi) \leq cQ\omega(z)$$

where $S_z(\xi) = (\xi - z)/(1 - \overline{z}\xi)$, a Möbius transform.

The derivatives in the integral are to be understood using the notation $f(z)' = \partial f/\partial z = (f_x - if_y)/2$. We also will use $\partial f/\partial\overline{z} = (f_x + if_y)/2$. Note that with this notation $\Delta f = 4(\partial^2 f/\partial z\partial\overline{z})$.

For comparisons's sake, here is the result from Example 1 again:

Lemma. *If w in A_2, then*

$$\frac{1}{|J|}\sum_{I\subseteq J}\left|\frac{w_{I_l} - w_{I_r}}{w_I}\right|\left|\frac{w_{I_l}^{-1} - w_{I_r}^{-1}}{w_I^{-1}}\right| w_I|I| \leq C\|w\|_{A^2}w_J.$$

To start with, think of the $\log 1/|S_z(\xi)|\,dA(\xi)$ as simply a volume element belonging to the measure. It is the price we pay for working on the disk. Similarly, the $1/|J|$ and $|I|$ in the dyadic sum can be thought of as a kind of measure. What is left are the terms in w and ω. They correspond as follows:

$$\begin{aligned}
\omega(\xi) &\iff w_I\\
\omega(z) &\iff w_J\\
\omega'(\xi) &\iff w_{I_l} - w_{I_r}\\
\omega^{-1\prime}(\xi) &\iff w_{I_l}^{-1} - w_{I_r}^{-1}\\
Q &\iff \|w\|_{A^2}.
\end{aligned}$$

Thus, the above two lemmas are really a continuous and dyadic analog of each other.

Therefore, we can hope to use the same Bellman function as we used for Example 1 for the continuous case. I have largely copied the proof from [**9**] below, and I'll intersperse explanatory remarks in parenthesis. So here is the proof of Lemma 3.1.

Let

$$B(x,y) = x\Big(-\frac{4Q}{xy} - \frac{xy}{4Q} + 4Q + 1\Big).$$

(This is the same equation as we used in Example 1. Recall that we have

$$1 \le xy \le Q \Longrightarrow 0 \le B(x,y) \le cQx$$

and also that the two matrixes

$$-d^2B \pm \begin{pmatrix} 0 & C'/\zeta \\ C'/\zeta & 0 \end{pmatrix}$$

are positive semidefinite, so

$$-\overline{\left(\frac{\partial\omega}{\partial z}, \frac{\partial\omega^{-1}}{\partial z}\right)}\, d^2B(\omega,\omega^{-1}) \begin{pmatrix} \partial\omega/\partial z \\ \partial\omega^{-1}/\partial z \end{pmatrix}$$
$$\ge \left| \overline{\left(\frac{\partial\omega}{\partial z}, \frac{\partial\omega^{-1}}{\partial z}\right)} \begin{pmatrix} 0 & C'/\omega^{-1} \\ C'\omega^{-1} & 0 \end{pmatrix} \begin{pmatrix} \partial\omega/\partial z \\ \partial\omega^{-1}/\partial z \end{pmatrix} \right| .)$$

Let us also consider the function $b : \mathbf{C} \to \mathbf{R}$

$$b(z) = B(h(z)) = B(\omega(z), \omega^{-1}(z)).$$

(By elementary but tedious calculations,

$$-\Delta b(z) = -\overline{\left(\frac{\partial\omega}{\partial z}, \frac{\partial\omega^{-1}}{\partial z}\right)}\, d^2B \begin{pmatrix} \partial\omega/\partial z \\ \partial\omega^{-1}\partial z \end{pmatrix}),$$

which, using the above inequality and some algebra, is

$$\ge \frac{C'}{\omega^{-1}(z)} |\omega(z)'||\omega^{-1}(z)'|.$$

Thus,

$$0 \le b(z) \le cQ\omega(z)$$

and

$$-\Delta b(z) \ge C\omega(z)\frac{|\omega(z)'||\omega^{-1}(z)'|}{\omega(z)\omega^{-1}(z)}.$$

Now we can run the continuous Bellman process

$$\int_{\mathbf{D}} \frac{|\omega(\xi)'||\omega^{-1}(\xi)'|}{\omega(\xi)\omega^{-1}(\xi)}\omega(\xi)\log\frac{1}{|S_z(\xi)|}\,dA(\xi) \le c\int_{\mathbf{D}} -\Delta b(\xi)\log\frac{1}{|S_z(\xi)|}\,dA(\xi)$$

(using the second condition on $b(z)$)

$$= c\int_{\mathbf{D}} -\Delta b(S_{-z}(\xi))\log\frac{1}{|\xi|}\,dA(\xi) = c\left(b(z) - \int_{\mathbf{T}} b\,dm\right)$$

(using Green's theorem)

$$\le cQ\omega(z)$$

(by the first condition on $b(z)$).

This ends the proof of Lemma 3.1.

Note how we, just as in the discrete case, can drop the second integral merely because we are subtracting something positive. Remember that for z on the boundary, $b(z)$ is equivalent to $B(w_I, w_I^{-1})$ for infinitesimally small I.

Note. This is just one example of the many uses of the Bellman function method in the continuous case, and only shows the tip of the iceberg. There are many different ways in which the method has been used, including using extensions other than harmonic ones (e.g., heat extensions, see, for example, [**8**]).

4. The converse. In previous sections we saw that the existence of a Bellman function implies the desired bound on the corresponding sum. A converse is also true. Under reasonably general conditions, we can show that whenever we have a bounded dyadic sum, there is a corresponding Bellman function.

Again, this is best demonstrated by an example.

It is well known that the dyadic square function

$$S_d f = \left[\sum_{I \in D} (f_{I_r} - f_{I_l})^2 \chi_I \right]^{1/2}$$

is bounded in L^2. Thus,

$$\sum_{I \in D} (f_{I_l} - f_{I_r})^2 |I| \leq \|f\|_{L^2}^2$$

for $f \in L^2$, or, equivalently (use $f\chi_J$),

$$(*) \qquad \frac{1}{|J|} \sum_{I \subset J} (f_{I_l} - f_{I_r})^2 |I| \leq \frac{1}{|J|} \int_J f^2$$

for any dyadic interval J, $f \in L^2$.

If we were to prove this bound using the method of Bellman functions, we would search for a function $B(x, y)$ on the domain $\{x, y > 0, x^2 \leq y\}$ such that

1. $0 \leq B(x, y) \leq y$
2. $B(x, y) - [B(x_-, y_-) + B(x_+, y_+)]/2 \geq (x_- - x_+)^2$ whenever $x = (x_- + x_+)/2$, and $y = (y_- + y_+)/2$.

Here x corresponds to $1/|I| \int_I f$ and y corresponds to $1/|I| \int_I f^2$.

Finding this Bellman function would prove the result for non-negative L^2 functions. We would extend to all L^2 functions by sublinearity of Sf. However, our goal here is not to prove $(*)$. Rather, we will assume that equation $(*)$ holds, and prove from there that such a $B(x, y)$ exists.

For positive x, y such that $x^2 \leq y$, let

$$B(x, y) = \sup_{\{f \in L^2 : f \geq 0, f_J = x, (f^2)_J = y\}} \frac{1}{|J|} \sum_{I \subseteq J} (f_{I_l} - f_{I_r})^2 |I|.$$

Claim. *$B(x, y)$ is well defined on the domain $\{x, y > 0, x^2 \leq y\}$ and satisfies*

1. $0 \leq B(x,y) \leq y$,

2. $B(x,y) - [B(x_-,y_-) + B(x_+,y_+)]/2 \geq (x_- - x_+)^2$ whenever $x = (x_- + x_+)/2$, and $y = (y_- + y_+)/2$.

on this domain.

Proof. We need to show that, for fixed positive x_0, y_0 with $x_0^2 \leq y_0$, the set $\{f \in L^2, f \geq 0 : f_J = x_0, (f^2)_J = y_0\}$ is not empty to guarantee that the supremum is well defined.

Assume for a moment that $J = [0,1)$. Let $f(t) = x_0(\alpha+1)t^\alpha \chi_{[0,1)}(t)$ for a fixed constant α. A calculation shows that $\int_0^1 f(t)\,dt = x_0$ and $\int_0^1 f^2(t)\,dt = x_0^2(\alpha+1)^2/(2\alpha+1)$. Since for $\alpha \in (-(1/2), 0]$, $(\alpha+1)^2/(2\alpha+1)$ takes all values in $[1,\infty)$, we can choose α such that $x_0^2(\alpha+1)^2/(2\alpha+1) = y_0$. (Recall that $y_0 \geq x_0^2$.) Then for that α, $f(t)$ is non-negative, in L^2, and satisfies $\int_0^1 f = x_0$, $\int_0^1 f^2 = y_0$. Thus the supremum is well defined if $J = [0,1)$. If J is any other interval, we simply rescale the above example. If $J = [k2^j, (k+1)2^j)$, then $\tilde{f}(t) = |f(((t - k2^j)/2^j)|$ will have the desired averages on J.

Next, we prove the inequalities. Clearly, $B(x,y)$ is positive. By assumption $(*)$,

$$B(x,y) \leq y.$$

Thus it remains to prove that $B(x,y)$ also satisfies the difference inequality.

Note that $B(x,y)$ is actually independent of the choice of J: Let $J = [k2^j, (k+1)2^j)$, and $I = [s2^i, (s+1)2^i)$. We can rescale any $f \in L^2$ with averages $f_J = x, (f^2)_J = y$ to $\tilde{f}(t) = f(2^i(((t - k2^j)/2^j) + s2^i)$ to get a function with exactly the same averages on I and children as f had on J and its children. Thus the supremum does not change with J. In what follows, we will set $J = [0,1)$ to simplify the presentation.

Let $x = (x_- + x_+)/2$, and $y = (y_- + y_+)/2$ be given. Fix $\varepsilon > 0$. By the definition of a supremum, since $B(x_-,y_-) = \sup_{\{f \in L^2 : f \geq 0, f_J = x_-, (f^2)_J = y_-\}} 1/|J| \sum_{I \subseteq J} (f_{I_l} - f_{I_r})^2 |I|$, there exists a $u(t)$ that almost attains the supremum, i.e., $u_{[0,1)} = x_-$, $(u^2)_{[0,1)} = y_-$, and

$$B(x_-,y_-) - \sum_{I \subseteq [0,1)} (u_{I_l} - u_{I_r})^2 |I| \leq \varepsilon.$$

Similarly, there exists a $v(t)$ such that $v_{[0,1)} = x_+, (v^2)_{[0,1)} = y_+$, and

$$B(x_+, y_+) - \sum_{I \subseteq [0,1)} (v_{I_l} - v_{I_r})^2 \, |I| \leq \varepsilon.$$

We will define a new function $f(t)$ on $[0,1)$ by simply concatenating u and v.

Let $f(t) = u(2t)\chi_{[0,1)}(t) + v(2t-1)\chi_{[0,1)}(t)$.

Note that

$$\sum_{I \subseteq [0,1/2)} (f_{I_l} - f_{I_r})^2 |I| = \sum_{K \subseteq [0,1)} (u_{K_l} - u_{K_r})^2 \, \frac{|K|}{2}$$

and

$$\sum_{I \subseteq [1/2,1)} (f_{I_l} - f_{I_r})^2 |I| = \sum_{K \subseteq [0,1)} (v_{K_l} - v_{K_r})^2 \frac{|K|}{2}.$$

Furthermore, $\int_0^1 f = \int_0^{1/2} u(2t) + \int_{1/2}^1 v(2t-1) = (u_{[0,1)} + v_{[0,1)})/2 = (x_- + x_+)/2 = x$, and similarly, $\int_0^1 f^2 = y$. Thus f satisfies all the properties to be included in the supremum that defines $B(x,y)$.

$$\begin{aligned}
B(x,y) &= \sup_{\{h \in L^2 : h \geq 0, \int_0^1 h = x, \int_0^1 h^2 = y\}} \sum_{I \subseteq [0,1)} (h_{I_l} - h_{I_r})^2 \, |I| \\
&\geq \sum_{I \subseteq [0,1)} (f_{I_l} - f_{I_r})^2 |I| \\
&= \sum_{I \subseteq [0,1/2)} (f_{I_l} - f_{I_r})^2 |I| + \sum_{I \subseteq [1/2,1)} (f_{I_l} - f_{I_r})^2 |I| \\
&\quad + (f_{[0,1/2)} - f_{[1/2,1)})^2 = \frac{1}{2} \sum_{I \subseteq [0,1)} (u_{I_l} - u_{I_r})^2 |I| \\
&\quad + \frac{1}{2} \sum_{I \subseteq [0,1)} (v_{I_l} - v_{I_r})^2 |I| + (x_- - x_+)^2 \\
&\geq \frac{1}{2}(B(x_-, y_-) + \varepsilon) + \frac{1}{2}(B(x_+, y_+) + \varepsilon) + (x_- - x_+)^2.
\end{aligned}$$

So

$$B(x,y) - \frac{B(x_-, y_-) + B(x_+, y_+)}{2} \geq (x_- - x_+)^2 + \varepsilon.$$

Since ε was arbitrary, we may let it go to 0. Thus the inequality is proven. $\square$

Thus, given the bounded sum, we were able to produce a Bellman function corresponding to it. The above argument can be used in many cases. The fundamental requirements are that we able to combine v and u into a new function, a requirement that is normally satisfied. So in most situations, a Bellman function exists if, and only if, the corresponding result is true.

In the next section, we will use this if and only if nature of Bellman functions to disprove an operator bound.

4.1. Using Bellman functions to disprove a bound. In this section we will use the Bellman function method to disprove a bound. We will proceed by proving that if the estimate were true, a Bellman function with certain properties would need to exist. We then prove that no such function exists.

It is known that the dyadic square function $S_d f = [\sum_{I\in D}(f_{I_r} - f_{I_l})^2\chi_I]^{1/2}$ is bounded in $L^2(w)$, where w is an A_2 weight (for example, [**2, 12**].) It is natural to ask if a mixed type square function, formed using both dx and w, might be bounded in $L^2(w)$. The answer follows quite easily from Bellman function techniques.

Define $f_{w_I} = 1/(\int_I w(x)\,dx)\int_I f(x)w(x)$.

Theorem 4.1. *Fix $1 \leq A < \infty$. Let*

$$\mathcal{S}f(y) = \left[\sum_{I\in D}\left(f_{I_r} - f_{w_{I_l}}\right)^2 \chi_I(y)\right]^{1/2}.$$

Then

$$\sup_{\{w\in A_2:\|w\|_{A_2}\leq A,\|w\chi^{-1}_{[0,1)}\|_{L^2(w)}=1\}} \|\mathcal{S}\left(w^{-1}\chi_{[0,1)}\right)\|_{L^2(w)} = \infty.$$

Thus, the mixed square function is not bounded in $L^2(w)$.

Proof. Let

$$\mathcal{T} f(y)=\left[\sum_{I \in D}\left(f_{I_l}-f_{w I_l}\right)^2 \chi_I(y)\right]^{1/2}.$$

(Here the first term is averaged over I_l instead of I_r.)

$$\mathcal{T} f(y)=\left[\sum_{I \in D}\left(f_{I_l}-f_{I_r}+f_{I_r}-f_{w I_l}\right)^2 \chi_I(y)\right]^{1/2} \leq S_d f(y)+\mathcal{S} f(y).$$

Since $S_d f$ is positive, it is enough to show that $\mathcal{T} f$ is unbounded.

Apply $\mathcal{T}$ to $f=w^{-1}\chi_{[0,1)}$. Since $\|w^{-1}\chi_{[0,1)}\|_{L^2(w)}=(\int_{[0,1)} w^{-1})^{1/2}$, we will estimate

$$\sup_{\{w \in A_2 : \|w\|_{A_2} \leq A, \int_0^1 w^{-1}=1\}} \|\mathcal{T} w^{-1}\chi_{[0,1)}\|_{L^2(w)}^2$$

from below.

$$\begin{aligned}
\|\mathcal{T} w^{-1}\chi_{[0,1)}\|_{L^2(w)}^2 & \geq \sum_{I_l \subseteq[0,1)}\left(\frac{1}{|I_l|} \int_{I_l} w^{-1}-\frac{|I_l|}{\int_{I_l} w}\right)^2 \int_I w \\
& =\sum_{I_l \subseteq[0,1)}\left(w_{I_l}^{-1}-\frac{1}{w_{I_l}}\right)^2\left(\int_{I_l} w+\int_{I_r} w\right) \\
& \geq \sum_{I_l \subseteq[0,1)}\left(w_{I_l}^{-1}-\frac{1}{w_{I_l}}\right)^2 w_{I_l}|I_l|.
\end{aligned}$$

So our task becomes to estimate

$$\sup_{\{w : \|w\|_{A^2} \leq A, \int_0^1 w^{-1}=1\}} \sum_{I_l \subseteq[0,1)}\left(w_{I_l}^{-1}-\frac{1}{w_{I_l}}\right)^2 w_{I_l}|I_l|.$$

We will switch from a sum over only left sided intervals, $I_l \subseteq [0,1)$, to a sum over all intervals $I \subseteq [0,1)$. To do that, first notice that

$$\begin{aligned}
& \sup_{\{w : \|w\|_{A_2} \leq A, \int_0^1 w^{-1}=1\}} \sum_{I_l \subseteq[0,1)}\left(w_{I_l}^{-1}-\frac{1}{w_{I_l}}\right)^2 w_{I_l}|I_l| \\
& \quad=\sup_{\{w : \|w\|_{A_2} \leq A, \int_0^1 w^{-1}=1\}} \sum_{I_r \subseteq[0,1]}\left(w_{I_r}^{-1}-\frac{1}{w_{I_r}}\right)^2 w_{I_r}|I_r|,
\end{aligned}$$

since for every $w(t)$ in the set we are taking the supremum over, $w(1-t)$ is also in the set. The sum of these two is at least as large as

$$\sup_{\{w:\|w\|_{A_2}\leq A,\int_0^1 w^{-1}=1\}} \sum_{I\subseteq[0,1]} \left(w_I^{-1} - \frac{1}{w_I}\right)^2 w_I|I|.$$

Thus

$$\begin{aligned}\sup_{\{w:\|w\|_{A_2}\leq A,\int_0^1 w^{-1}=1\}} &\sum_{I_l\subseteq[0,1]} \left(w_{I_l}^{-1} - \frac{1}{w_{I_l}}\right)^2 w_{I_l}|I_l| \\ &\geq \frac{1}{2} \sup_{\{w:\|w\|_{A_2}\leq A,\int_0^1 w^{-1}=1\}} \sum_{I\subseteq[0,1]} \left(w_I^{-1} - c(1/w_I)\right)^2 w_I|I|.\end{aligned}$$

To prove that this is unbounded, we use the following lemma, from which the theorem follows as a special case:

Lemma 4.2. *Fix $A \geq 1$. Let $R_A = \{(x,y) : 1 \leq xy \leq A; x, y > 0\}$ and $D(x,y) = \{w \in A_2$ such that $\|w\|_{A_2} \leq A, \int_{[0,1]} w = x, \int_{[0,1]} w^{-1} = y\}$.*

Say $f(x,y) \geq 0$ for all $(x,y) \in R_A$. Then for all $(x,y) \in R_A$ such that $f(x,y) > 0$,

$$B(x,y) = \sup_{D(x,y)} \sum_{I\subseteq[0,1]} f(w_I, w_I^{-1})|I| = \infty.$$

The theorem follows from the lemma for $f(x,y) = (y-(1/x))^2 x$ by considering $B(x,1)$ for any $x \neq 1$, since $f(x,y) = 0$ on $(x, 1/x)$ only. □

Proof of Lemma 4.2. This proof will proceed very similarly to the previous section where we showed that a converse of the Bellman method holds. First we will show that the supremum is well defined, and then we will show that that $B(x,y)$ satisfies a difference inequality wherever it is finite. This will lead to a contradiction.

For $B(x,y)$ to be well defined, the supremum needs to be taken over a non-empty set. We need to show that for any such (x,y) there is a $w \in A_2$ such that $w_J = x$, $w_J^{-1} = y$.

Let $w(t) = ct^\alpha$. It is well known, and can be verified by a simple calculation, that for $\alpha \in (-1,1)$ $w(t)$ is an A_2 weight. Assume for a moment that $J = [0,1)$. Then $w_J = c/(1+\alpha)$, $w_J^{-1} = 1/c(1-\alpha)$. A calculation shows that for $\alpha = \sqrt{1-(1/xy)}$ and $c = x(\alpha+1)$, we have the required weight. The case for general J follows by scaling, which does not change averages: If $J = [k2^j, (k+1)2^j)$, $w((t-k2^j)/2^j)$ is the required weight. Thus the supremum is never taken over an empty set, and is well defined.

Now we will prove that $B(x,y)$ satisfies the usual type of difference inequality.

Fix $(x,y), (x_-,y_-), (x_+,y_+) \in R_A$ with $x = (x_-+x_+)/2$ and $y = (y_-+y_+)/2$. We will show that $B(x,y)$ satisfies the following difference inequality whenever $B(x_-,y_-), B(x_+,y_+) < \infty$:

$$B(x,y) - \frac{B(x_-,y_-)+B(x_+,y_+)}{2} \geq f(x,y).$$

If $B(x_-,y_-) < \infty$, there is a $u(t) \in D(x_-,y_-)$ which essentially achieves the supremum for $B(x_-,y_-)$, i.e.,

$$B(x_-,y_-) - \sum_{I\subseteq[0,1]} f(u_I, u_I^{-1})|I| \leq \varepsilon,$$

and similarly, there is a $v(t) \in D(x_+,y_+)$ such that

$$B(x_+,y_+) - \sum_{I\subseteq[0,1]} f(v_I, v_I^{-1})|I| \leq \varepsilon.$$

Let

$$w(t) = u(2t)\chi_{[0,1/2)} + v(1-2t)\chi_{[1/2,1)}, \quad t \in [0,1),$$

and extend it to $\mathbf{R}$ by

$$w(t) = w|(\text{fractionalpart}\,(t)|), \quad t \notin [0,1).$$

Let us verify that w in $D(x,y)$.

$$\begin{aligned}\int_{[0,1]} w &= \frac{1}{2}\left(\frac{1}{1/2}\int_{[0,1/2)} u(2t) + \frac{1}{1/2}\int_{[1/2,1)} v(1-2t)\right)\\ &= \frac{x_+ + x_-}{2} = x,\end{aligned}$$

and similarly, $\int_{[0,1)} w^{-1} = y$.

If $I \subset [0,1/2)$, $w_I w_I^{-1} = u_J u^{-1}{}_J \leq A$, (where J is a larger dyadic interval) and when $I \subset [1/2,1)$, $w_I w_I^{-1} = v_J v^{-1}{}_J \leq A$. When $|I| \geq 1$, $I = [k,m)$ some integers k,m. Since the graph of w repeats itself $m-k$ times on that interval, $w_I = 1/(m-k)(m-k)\int_{[0,1)} w = x$, and similarly $w_I^{-1} = y$. So we have $w_I w_I^{-1} = xy \leq A$. Thus for all dyadic I, $w_I w_I^{-1} \leq A$. Thus $\|w\|_{A^2} \leq A$, and $w \in D(x,y)$.

Now we will prove the difference inequality.

Note that

$$\sum_{I \subseteq [0,1/2)} f(w_I, w_I^{-1})|I| = \sum_{K \subseteq [0,1)} f(u_K, u_K^{-1})\frac{|K|}{2} \geq \frac{1}{2}\left(B(x_-,y_-) - \varepsilon\right)$$

and

$$\sum_{I \subseteq [1/2,1)} f(w_I, w_I^{-1})|I| = \sum_{K \subseteq [0,1)} f(v_K, v_K^{-1})\frac{|K|}{2} \geq \frac{1}{2}\left(B(x_+,y_+) - \varepsilon\right).$$

Thus,

$$\begin{aligned} B(x,y) &\geq \sum_{I \subseteq [0,1]} f(w_I, w_I^{-1})|I| = f(w_{[0,1]}, w_{[0,1]}^{-1}) \\ &\quad + \sum_{I \subseteq [0,1/2]} f(w_I, w^{-1}{}_I)|I| + \sum_{I \subseteq [1/2,1]} f(w_I, w^{-1}{}_I)|I| \\ &\geq f(x,y) + \frac{B(x_-,y_-) + B(x_+,y_+)}{2} - \varepsilon. \end{aligned}$$

And so, letting $\varepsilon \to 0$, the difference inequality is proven.

We can derive a contradiction immediately: If $B(x,y) < \infty$, let $x_- = x, x_+ = x, y_- = y, y_+ = y$. Then all the Bellman functions in the difference inequality cancel each other, leaving us with $f(x,y) \leq 0$ But $f(x,y) > 0$. Therefore $B(x,y)$ must be ∞. This proves the lemma. $\square$

REFERENCES

1. S. Buckley *Estimates for operator norms on weighted spaces and reverse Jensen inequalities*, Trans. Amer. Math. Soc. **340** (1993), 253–272.

2. D.L. Burkholder, *Explorations in Martingale theory and its applications*, Ecole Probab. Saint-Flour **19** (1989), 1–66; Lect. Notes Math. **1464**, Springer, Berlin, 1991.

3. S. Hukovic, S. Treil and A. Volberg, *The Bellman functions and sharp weighted inequalities for square functions*, in *Complex analysis, operators, and related topics*, Birkhäuser, Basel, 2000.

4. F. Nazarov and S. Treil, *The hunt for a Bellman function: applications to estimates for singular integral operators and to other classical problems of harmonic analysis*, Algebra Anal., **8** (1996), 32–162.

5. ———, *The weighted norm inequalities for Hilbert transform are now trivial*, C.R. Acad. Sci. Paris **323** (1996), 717–722.

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

7. ———, *Bellman function in stochastic control and harmonic analysis* (*how our Bellman function got its name*), Oper. Theory: Advances Appl. **129** (2001), 393–423.

8. S. Petermichl and A. Volberg, *Heating of the Ahlfors-Beurling operator*: *Weakly quasiregular maps on the plane are quasiregular*, Duke Math. J. **112** (2002), 281–305.

9. S. Petermichl and J. Wittwer, *A sharp weighted estimate on the norm of Hilbert transform via invariant A_2 characteristic of the weight*, Michigan Math. J. **50** (2002), 71–87.

10. S. Treil and A. Volberg, *Wavelets and the angle between past and future*, J. Funct. Anal. **143** (1997), 269–308

11. J. Wittwer, *A sharp estimate on the norm of the martingale transform*, Math. Res. Lett. **7**, (2000), 1–12.

12. ———, *A sharp estimate on the norm of the continuous square function*, Proc. Amer. Math. Soc. **130** (2002), 2335–2342.

Westminster College, 1840 S. 1300 E, Salt Lake City, UT 84105
Email address: jwittwer1@gmail.com