Oleksandra V. Beznosova
____
Candidate

Department of Mathematics and Statistics
____
Department

This dissertation is approved, and it is acceptable in quality and form for publication on microfilm:

Approved by the Dissertation Committee:

____

____

____

____

____

____

____

____

Accepted:

____
Dean, Graduate School

____
Date

Bellman Functions, Paraproducts, Haar Multipliers and Weighted Inequalities

by

Oleksandra V. Beznosova

B.S., Applied Math, National Technical University of Ukraine, 2000
M.S., Applied Math, National Technical University of Ukraine, 2002
M.S., Pure Math, University of New Mexico, 2003

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy

Mathematics

The University of New Mexico
Albuquerque, New Mexico

May, 2008

ACKNOWLEDGMENTS

I heartily acknowledge Dr. Maria Cristina Pereyra, my graduate advisor, Teacher and dissertation chair, for her helpful advices and a lot of time spent with me. Her guidance and professional style will remain with me as I continue my career.

I also thank my committee members, Dr. Cornelis Onneweer, Dr. Wilfredo Urbina, Dr. Pedro Embid and Dr. Dimiter Vassilev, for their valuable recommendations pertaining to this study and assistance in my professional development.

Gratitude is extended to S. Treil and F. Nazarov for a useful conversations on subject of this work.

And finally to my husband, Boris, your love is the greatest gift of all.

Bellman Functions, Paraproducts, Haar Multipliers and Weighted Inequalities

by

Oleksandra V. Beznosova

B.S., Applied Math, National Technical University of Ukraine, 2000
M.S., Applied Math, National Technical University of Ukraine, 2002
M.S., Pure Math, University of New Mexico, 2003

ABSTRACT OF DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy

Mathematics

The University of New Mexico
Albuquerque, New Mexico

May, 2008

Bellman Functions, Paraproducts, Haar Multipliers and Weighted Inequalities

by

Oleksandra V. Beznosova

B.S., Applied Math, National Technical University of Ukraine, 2000
M.S., Applied Math, National Technical University of Ukraine, 2002
M.S., Pure Math, University of New Mexico, 2003
Ph.D., Mathematics, University of New Mexico, 2008

ABSTRACT

In this dissertation we use the method of Bellman functions in applications to the following problems: finding the sharp dependence of the norm of the paraproduct on weighted Lebesgue spaces $L_2(w)$ on the A_2 characteristic of the weight; finding the sharp dependence of the norm of the Haar multiplier operators on the Lebesgue spaces L_2 (unweighted, the weight is in the multiplier) on a suitable characteristic of the weight; finding the two weighted estimates for the weighed square function. We also provide alternative proofs of some known weighted inequalities and derive a few new ones using the Bellman function method. We extend some of the weighted inequalities to the two weighted case and to the case when the underlying measure is not Lebesgue, but a doubling measure $d\sigma$, absolutely continuous with respect to Lebesgue measure. These extensions allow us to use sharp extrapolation theorems to get bounds in weighted and unweighted L_p spaces.

Contents

1 Introduction

In this dissertation we are concerned with weighted norm inequalities. Weighted norm inequalities is an active research area in harmonic analysis. The boundedness of the Hilbert transform on weighted Lebesgue spaces was first shown by Helson and Szegö in [HeSz] in 1960. In 1973, Hunt, Muckenhoupt and Wheeden (see [HuMW]) presented a new necessary and sufficient condition for the boundedness of the Hilbert transform on the weighted spaces $L_p(w)$, the celebrated Muckenhoupt A_p-condition. A year later, in [CoFe], Coifman and Fefferman extended this result to a larger class of convolution singular integrals with standard kernels. The Calderón-Zygmund singular integrals are bounded on the weighted Lebesgue space $L_p(w)$ if and only if the weight belongs to the Muchenhoupt class A_p (see [Gr]).

The question that has been asked is:

How is the norm of a Calderón-Zygmund singular integral operator on weighted Lebesgue spaces $L_p(w)$ related to the Muckenhoupt (A_p) characteristic of the weight w, $\|w\|_{A_p}$. More precisely, we want to find a function $\varphi(x)$, sharp in terms of its growth, such that

$$\|Tf\|_{L_p(w)} \;\leq\; C\varphi\left(\|w\|_{A_p}\right)\|f\|_{L_p(w)}\,. \tag{1.1}$$

This kind of estimates for different singular integral operators is used often in the theory of partial differential equations, see [FeKPi], [AISa], [PetVo], [BaJa] and [DrPetVo].

In 1984 (see [Ru]) Rubio de Francia published his famous extrapolation theorem. Heuristically, it says that *there is no L_p but weighted L_2* (some authors attribute this statement to Rubio de Francia, see [Per1], others to A. Cordoba, see [CrMPe]). Rubio de Francia's extrapolation theorem attracted a lot of attention to weighted norm inequalities. In fact, see [Per2], when one is looking for the sharp bounds on the unweighted L_p norms of an operator it is useful to prove an estimate in the weighted case $L_2(w)$, extrapolate it to $L_p(w)$ and then use a particular choice of the weight w to conclude the unweighted L_p result.

The sharp version of Rubio de Francia's extrapolation theorem (see [DrGrPerPet]) allows one to trace, for each $1 < p < \infty$, the dependence of the $L_p(w)$ norm of an operator T on the A_p-characteristic of the weight w based on the function $\varphi(x)$ involved in the estimate (1.1) for $p = 2$. In the particular case when an operator T obeys linear bounds in $L_2(w)$, that is

$$\|Tf\|_{L_2(w)} \leq C\,\|w\|_{A_2}\,\|f\|_{L_2(w)}\,. \tag{1.2}$$

The sharp extrapolation theorem will return the following bounds in $L_p(w)$, for $w \in A_p$

$$\|Tf\|_{L_p(w)} \leq C\,\|w\|_{A_p}^{\max\left\{1,\frac{1}{p-1}\right\}}\,\|f\|_{L_p(w)}\,. \tag{1.3}$$

For a handful of operators the initial linear bound in $L_2(w)$, (1.2), is known to hold and known to be optimal. For some of them, but not all, the extrapolated bounds (1.3) are optimal as well. For others, (1.3) is optimal for $1 < p \leq 2$, but not for $p > 2$. It remains an open question to identify the optimal rate of growth for such operators for $p > 2$.

The extrapolation theorems will be discussed in Section 3.4.

Bounds on the operator norms of the type (1.2) is the main topic of our research in this dissertation. Inequalities of this type are not easy to derive. In chronological order of publication inequality (1.2) was shown to hold for the maximal function (Buckley 1993 [Buc1]), the martingale transforms (Wittwer 2000 [Wit1]), the square functions (dyadic case: Hukovic, Treil and Volberg 2000 [HTrVo], continuous case: Wittwer 2002 [Wit2]), the Beurling transform (Bañuelos, Janakiraman [BaJa] and independently by Dragicevic, Petermichl and Volberg 2006, [DrPetVo]), the Hilbert transform (Petermichl 2007 [Pet1]) and the Riesz transforms (Petermichl 2008, [Pet2]).

Each of the inequalities (1.2) for the Hilbert, Riesz, Beurling, martingale transforms, maximal and square functions is a result of a series of papers. For example, for the Hilbert transform on $\mathbb{R}$, Buckley showed that the inequality (1.2) holds with $\varphi(x) = x^2$ in [Buc1], Petermichl and Pott in [PetPo] improved the exponent of $\varphi(x)$ from 2 to 3/2 and in 2006 Petermichl obtained (1.2) for the Hilbert transform, see [Pet1]. Petermichl received the Salem Prize for this result in 2006, the most prestigious award for a young analyst.

It has been conjectured that the linear bound (1.2) is true for all Calderón-Zygmund singular integral operators.

The proofs of the results for particular operators, however, cannot be extended to the more general case because they heavily rely on the symmetries of the corresponding transform.

One of the main tools in the proofs of sharp bounds on the norms of the Hilbert, Riesz, Beurling (proof from [DrPetVo]) transforms on the weighted Lebesgue spaces $L_2(w)$, as well as the proof of square function result from [HTrVo], is the method of Bellman functions. First introduced by Burkholder in [Bur], it was developed as a method in control theory by Bellman. In his dissertation Buckley used a similar approach to prove, what is now known as, Buckley's inequality, see Section 3.3.

Recently the method of Bellman functions reappeared in harmonic analysis with the help of Nazarov, Treil and Volberg (see [NTr], [NTrVo2], [NTrVo3], [HTrVo], [DrVo]). Further modification of the method of Bellman functions, the method of sharp Bellman functions, can be found in [Va], [VaVo], [SlVo]. The method of Bellman functions turns out to be an extremely powerful tool and a very natural way to deal with weighted inequalities.

The main result of this dissertation is the linear bound on the norm of the dyadic paraproduct discussed in Section 6.5.

Paraproducts first appeared in the work of Bony on nonlinear partial differential equations (see [Bo]) and since then have played a central role in harmonic analysis. By the celebrated T(1) theorem of David and Journé [JoDa] a Calderón-Zygmund singular integral operator T can be written as $T = L + \pi_{b_1} + \pi^*_{b_2}$ where L is an almost translation invariant (convolution) operator, ($L1 = 0 = L^*1$), b_1 is the value of T at 1 and b_2 is the value of the adjoint of T at 1 (i.e. $b_2 = T^*(1)$), and π_b is the paraproduct associated to the function b. A dyadic version of this theorem can be found in [Per1]. So, if one wants to show that the norms of Calderón-Zygmund singular integral operators on the weighted Lebesgue spaces depend on the A_p-characteristic of the weight w at most linearly, it is natural to start with the paraproduct and with its simple dyadic version.

To obtain the linear bound on the norm of the dyadic paraproduct we use a number of tools, among them: the system of Weighted Haar functions from [CoJS], the bilinear embedding theorem from [NTrVo2] and five propositions developed in this dissertation. The propositions, as well as other tools used in the dissertation can be found in Section 3. Bellman function proofs of the propositions can be found in Section 7.

It turned out that some of the results developed for paraproducts allow us to obtain extensions of known results for Haar multiplier operators (see Section 5) and two weighted inequalities for the weighted square function (see Section 4).

Section 2 contains some basic facts and definitions.

2 Preliminaries

This introductory section closely follows the Lecture notes on dyadic harmonic analysis by M.C. Pereyra ([Per1]), where one can find much more detailed information on the discussed topics as well as proofs of the dyadic versions of the theorems stated here without a proof.

We will be working on the real line $\mathbb{R}$. All functions are real valued, $f : \mathbb{R} \to \mathbb{R}$. Unless specified, all integrals are taken with respect to the Lebesgue measure, over the whole line. We will be mostly emphasizing measures that are absolutely continuous with respect to the Lebesgue measure on the real line. For any set $S \subset \mathbb{R}$, $|S|$ is the Lebesgue measure of S. Unless specified, p and q are real numbers larger or equal than 1, $1 \leq p < \infty$, $1 \leq q < \infty$. When we will be dealing with real parameters that could be less than 1 or negative, we will try to use r or s.

The symbol $\langle \cdot\, ; \cdot \rangle$ stand for the scalar product in L_2

$$\langle f; g \rangle \; = \; \int fg.$$

Sometimes we will need the scalar product with respect to some measure σ, different from the Lebesgue measure, then we will denote it by $\langle \; ; \rangle_{d\sigma}$, i.e.

$$\langle f; g \rangle_{d\sigma} \; = \; \int fg \; d\sigma.$$

The L_p norms are defined in a regular way:

$$\|f\|_{L_p} \; := \; \left(\int |f|^p \right)^{1/p}$$

and

$$\|f\|_{L_p(d\sigma)} \; := \; \left(\int |f|^p \, d\sigma \right)^{1/p}.$$

2.1 The dyadic filtration

We are going to work in the dyadic world on the real line. Let D be the collection of all dyadic intervals,

$$D := \left\{ I \subset \mathbb{R} \ : \ I = \left[k2^{-j}, (k+1)2^{-j} \right), \ \ k, j \in \mathbb{Z} \right\}.$$

And let D_j denote the set of all dyadic intervals I of length 2^{-j},

$$D_j := \left\{ I \in D \ : \ |I| = 2^{-j} \right\},$$

it is also called the *j-th generation*.

Note that:

- Any two dyadic intervals $I, J \in D$ are either disjoint or one is contained in the other.
- Each dyadic interval I is in a unique generation D_j and there are exactly two subintervals of I in the next generation D_{j+1}, the *children* of I, which we will denote I_+ and I_-. Clearly, $I = I_+ \bigcup I_-$. The I_+ will always denote the right half of the dyadic interval I, while I_- will stand for the left half of I.
- For every interval $I \in D_j$ there is exactly one interval $\tilde{I} \in D_{j-1}$, such that $I \subset \tilde{I}$, $\tilde{I}$ will be called the *parent* of I. And there is exactly one interval $I^* \in D_j$, $I^* = \tilde{I} \setminus I$, called the *sibling* of I.

All the dyadic subintervals of a dyadic interval $J \in D$ will be denoted by $D(J) := \{I \in D : I \subset J\}$.

2.1.1 The Carleson sequences

A sequence of positive numbers $\{\lambda_I\}_{I\in D}$ is called a *Carleson sequence* if and only if there exists a constant $C > 0$, such that for every dyadic interval $J \in D$

$$\frac{1}{|J|}\sum_{I\in D(J)} \lambda_I \leq C.$$

By the Carleson constant of the sequence $\{\lambda_I\}_{I\in D}$ we will mean the smallest such constant. We will discuss Carleson sequences, as well as the famous Carleson Embedding Theorem, and give references in Section 3.1.

2.1.2 The Haar basis and averages

Let f be a locally integrable function, $f \in L_1^{loc}(\mathbb{R})$ and I be an interval. Denote the average of f over I by

$$m_I f \;:=\; \frac{1}{|I|}\int_I f.$$

Associated to any dyadic interval I, there is a *Haar function*

$$h_I(x) \;=\; \frac{1}{|I|^{1/2}}\left[\chi_{I_+}(x) - \chi_{I_-}(x)\right],$$

where χ_I is the *characteristic function* of the interval I:

$$\chi_I(x) \;=\; \begin{cases} 1, & x \in I \\ 0, & x \notin I. \end{cases}$$

The system of Haar functions $\{h_I\}_{I\in D}$ forms an orthonormal basis in $L_2(\mathbb{R})$. The Haar functions were first introduced by A. Haar in 1910, see [Ha]. They provide the oldest example of a wavelet basis.

In Section 2.2.4 we will introduce the weighted Haar functions.

2.1.3 The dyadic doubling measures

Let σ be a positive Borel measure on the real line, absolutely continuous with respect to the Lebesgue measure. We define $\sigma(I)$ to be

$$\sigma(I) := \int_I d\sigma.$$

We say that the measure σ is *dyadic doubling* if

$$D^d(d\sigma) := \sup_{I\in D} \frac{\sigma(\tilde{I})}{\sigma(I)} < \infty.$$

$D^d(d\sigma)$ is called the *dyadic doubling constant* of the measure σ. Note that the dyadic doubling constant of the measure σ is always greater than or equal to 2: $D^d(d\sigma) \geq 2$, and is equal to 2 in case of the Lebesgue measure.

We can consider the averages of a locally integrable function f with respect to measure σ:

$$m_I^\sigma f := \frac{1}{\sigma(I)} \int_I f d\sigma.$$

In this work we are only going to deal with doubling measures. The theory in the non-doubling case is much more complicated and only very few results are available, see [NTrVo1] and [To] for some results for non-doubling measures.

2.1.4 The dyadic BMO^d space

A locally integrable function b *has dyadically bounded mean oscillation*, $b \in BMO^d$ if and only if

$$\|b\|_{BMO_1^d} := \sup_{I\in D} \frac{1}{|I|} \int_I |b_I(x) - m_I b| \, dx < \infty.$$

When considered modulo constant function, BMO^d is a Banach space and $\|b\|_{BMO_1^d}$ is a norm. Function $b(x) = \log|x|$ is a typical example of a BMO^d function.

The continuous BMO space first appeared in [JohNi]. In the same paper the authors proved the famous John-Nirenberg inequality, which says that the distribution function of a function in the BMO space has to have exponential decay. In particular, it shows that $b(x) = \log|x|$ is a typical example of an unbounded BMO function. We state the dyadic version of the theorem, a proof based on a stopping time argument can be found, for example, in [Per1].

Theorem 2.1 (John-Nirenberg Inequality) *Given a function $b \in BMO^d$, any dyadic interval $I \in D$ and a positive number $\lambda > 0$, then there exist positive constants $C_1, C_2 > 0$ (independent of b, I, λ), such that*

$$|\{x \in I : \ |b(x) - m_I b| \ > \ \lambda\}| \ \leq \ C_1 \, |I| \, e^{-\frac{C_2 \lambda}{\|b\|_{BMO_1^d}}}.$$

The following remarkable self-improvement property can be deduced from the John-Nirenberg Inequality:

Corollary 2.2 (Self-improvement) *Given $b \in BMO^d$, then for all $p > 1$ there exists a constant $C_p > 0$ such that for all dyadic intervals $I \in D$*

$$\left(\frac{1}{|I|} \int_I |b(x) - m_I b|^p \, dx \right)^{1/p} \ \leq \ C_p \, \|b\|_{BMO_1^d}.$$

The reverse inequality holds by Hölder, therefore the left hand side provides an alternative definition for the BMO^d norm. We will use the alternative norm given by $p = 2$:

$$\|b\|^2_{BMO^d} \ := \ \sup_{I \in D} \frac{1}{|I|} \int_I |b(x) - m_I b|^2 \, dx.$$

Note that it follows from the fact, that $\{h_I\}_{I \in D(J)}$ is an orthonormal basis for $\{f \in L_2(J) : m_I f = 0\}$, that

$$\int_J |b(x) - m_J b|^2 \, dx = \sum_{I \in D(J)} |\langle b; h_I \rangle|^2. \tag{2.1}$$

Let us, however, show a different proof of (2.1), which, in our opinion, is more in the spirit of the method of Bellman functions. Let us first check that

$$\begin{aligned}\int_I |b(x) - m_I b|^2 \, dx &= \int_I \left(b^2(x) - 2b(x)m_I b + (m_I b)^2\right) dx \\ &= \int_I b^2(x)dx - 2m_I b \int_I b(x)dx + (m_I b)^2 \, |I| \\ &= \int_I b^2(x)dx - (m_I b)^2 \, |I| \, .\end{aligned}$$

Therefore

$$\begin{aligned}&\int_I |b(x) - m_I b|^2 \, dx - |\langle b; h_I \rangle|^2 \\ =& \int_I b^2(x)dx - |I| \, (m_I b)^2 - \frac{1}{|I|}\left(\int_{I^+} b(x)dx - \int_{I^-} b(x)dx\right)^2 \\ =& \int_I b^2(x)dx - |I| \left(\frac{m_{I^+}b + m_{I^-}b}{2}\right)^2 - \frac{(|I|/2)^2}{|I|}\left(m_{I^+}b - m_{I^-}b\right)^2 \\ =& \int_I b^2(x)dx - \frac{|I|}{4}\left[\left(m_{I^+}b + m_{I^-}b\right)^2 + \left(m_{I^+}b - m_{I^-}b\right)^2\right] \\ =& \int_I b^2(x)dx - \frac{|I|}{2} m_{I^+}^2 b - \frac{|I|}{2} m_{I^-}^2 b \\ =& \int_{I^+} b^2(x)dx - \left|I^+\right| m_{I^+}^2 b + \int_{I^-} b^2(x)dx - \left|I^-\right| m_{I^-}^2 b \\ =& \int_{I^+} |b(x) - m_{I^+}b|^2 \, dx + \int_{I^-} |b(x) - m_{I^-}b|^2 \, dx.\end{aligned}$$

Iterating this process we will get (2.1) and, hence,

$$\|b\|_{BMO^d}^2 = \sup_{J \in D} \frac{1}{|J|} \sum_{I \in D(J)} |\langle b; h_I \rangle|^2.$$

2.2 Weights

We say that w is a *weight* if it is a locally integrable (with respect to a given measure) function on the real line, positive almost everywhere (with respect to a given measure). The choice of measure should be clear from the context.

Since wdx is a measure, we could be interested in its dyadic doubling constant

$$D^d(w) \ := \ D^d(wdx) \ = \ \sup_{I \in D} \frac{\int_{\tilde{I}} wdx}{\int_I wdx}.$$

If the dyadic doubling constant $D^d(w)$ of the weight w is finite, we will say that w is a dyadic doubling weight.

Let us go over a few important classes of weights and their basic properties.

2.2.1 Reverse Hölder classes

A weight w belongs to the *dyadic Reverse Hölder class* RH_p^d $(1 < p < \infty)$ if there exists a constant $C > 0$ such that for all dyadic intervals $I \in D$

$$(m_I w^p)^{1/p} \ \leq \ C m_I w.$$

The smallest such constant C we will denote by $\|w\|_{RH_p^d}$. Note that by Hölder's inequality the dyadic Reverse Hölder classes satisfy:

$$if \quad 1 < p \leq q < \infty, \quad then \quad RH_q^d \ \subseteq \ RH_p^d \quad and \quad 1 \ \leq \ \|w\|_{RH_p^d} \ \leq \ \|w\|_{RH_q^d}.$$

Typical examples of RH_p^d weights are $w = |x|^\alpha$ for $\alpha > -1/p$. Note that in the dyadic case, unlike the continuous case, the RH_q^d condition for the weight w does not imply that w is a dyadic doubling weight. For example (see [Buc2]), one can see that $w = \chi_{\mathbb{R}\setminus[0,1]}$ is in RH_p^d, but is clearly not a dyadic doubling weight.

Reverse Hölder classes have a remarkable self-improvement property, discovered by Gehring in 1973, see [Ge].

Theorem 2.3 (Gehring's theorem) *Suppose $w \in RH_p^d$ for some $1 < p < \infty$. Then there exists $\epsilon > 0$, depending only on p and the RH_p^d constant of w, such that $w \in RH_{p+\epsilon}^d$.*

2.2.2 The dyadic Muckenhoupt classes

We say that the weight w belongs to the *dyadic Muckenhoupt class* A_p^d, $1 < p < \infty$, if

$$\|w\|_{A_p^d} := \sup_{I \in D} m_I w \left(m_I \left(w^{-\frac{1}{p-1}} \right) \right)^{p-1} < \infty.$$

Note that $\|w\|_{A_d^p}$ is not a norm, we will call it the A_p^d-*characteristic of* w. We will refer more often to A_2^d-weights:

$$\|w\|_{A_2^d} = \sup_{I \in D} m_I w \; m_I(w^{-1}).$$

Similarly to the dyadic reverse Hölder classes, by Hölder's inequality, $\|w\|_{A_p^d} \geq 1$ holds for all $1 < p < \infty$, as well as the following inclusion:

$$if \quad 1 < p \leq q < \infty \quad then \quad A_p^d \subseteq A_q^d, \quad \|w\|_{A_q^d} \leq \|w\|_{A_p^d}.$$

Also note another consequence of Hölder's inequality:

$$if \quad w \in A_p^d, \quad then \quad w^{\frac{1}{p}} \in A_{2-\frac{1}{p}}^d \quad and \quad \left\| w^{\frac{1}{p}} \right\|_{A_{2-\frac{1}{p}}^d} \leq \|w\|_{A_p^d}^{\frac{1}{p}}.$$

It is easy to see, let $w \in A_p^d$, then for all dyadic intervals $I \in D$

$$m_I w \left(m_I \left(w^{-\frac{1}{p-1}} \right) \right)^{p-1} \leq \|w\|_{A_p^d},$$

which we can rewrite as follows, using Hölder's inequality,

$$\begin{aligned}\|w\|_{A_p^d} &\geq m_I\left(\left(w^{\frac{1}{p}}\right)^p\right)\left(m_I\left(\left(w^{\frac{1}{p}}\right)^{-\frac{p}{p-1}}\right)\right)^{p-1} \\ &\geq \left(m_I\left(w^{\frac{1}{p}}\right)\right)^p\left(\left(m_I\left(\left(w^{\frac{1}{p}}\right)^{-\frac{p}{p-1}}\right)\right)^{\frac{p-1}{p}}\right)^p \\ &= \left[m_I\left(w^{\frac{1}{p}}\right)\left(m_I\left(\left(w^{\frac{1}{p}}\right)^{-\frac{1}{1-\frac{1}{p}}}\right)\right)^{1-\frac{1}{p}}\right]^p,\end{aligned}$$

which implies that $w^{\frac{1}{p}} \in A^d_{2-\frac{1}{p}}$ with $\left\|w^{\frac{1}{p}}\right\|^p_{A^d_{2-\frac{1}{p}}} \leq \|w\|_{A_p^d}$.

Note also that, unlike the dyadic reverse Hölder classes, $w \in A_p^d$ for some $1 < p < \infty$ implies that w is a dyadic doubling weight and, moreover, $D^d(w) \leq 2^p \|w\|_{A_p^d}$.

$$\begin{aligned}D^d(w) &= \sup_{I\in D}\frac{\int_{\tilde{I}} w dx}{\int_I w dx} = \sup_{I\in D}\frac{|\tilde{I}|\; m_{\tilde{I}}w}{|I|\; m_I w} \\ &\leq \sup_{I\in D}\frac{|\tilde{I}|}{|I|}\frac{\|w\|_{A_p^d}\left(m_{\tilde{I}}\left(w^{-\frac{1}{p-1}}\right)\right)^{-(p-1)}}{\left(m_I\left(w^{-\frac{1}{p-1}}\right)\right)^{-(p-1)}} \\ &= \sup_{I\in D}\frac{|\tilde{I}|}{|I|}\|w\|_{A_p^d}\left(\frac{m_I\left(w^{-\frac{1}{p-1}}\right)}{m_{\tilde{I}}\left(w^{-\frac{1}{p-1}}\right)}\right)^{p-1} \\ &= \|w\|_{A_p^d}\sup_{I\in D}\left(\frac{|\tilde{I}|}{|I|}\right)^p\left(\frac{\int_I w^{-\frac{1}{p-1}}dx}{\int_{\tilde{I}} w^{-\frac{1}{p-1}}dx}\right)^{p-1} \\ &\leq \textit{since } I \subset \tilde{I} \textit{ and } w \textit{ is positive} \\ &\leq \|w\|_{A_p^d}\sup_{I\in D}\left(\frac{|\tilde{I}|}{|I|}\right)^p \\ &= 2^p\|w\|_{A_p^d}.\end{aligned}$$

We define A^d_∞ to be:

$$A^d_\infty := \left\{w - \textit{weight s.t. } \|w\|_{A^d_\infty} := \sup_{I\in D} m_I w\, e^{-m_I(\ln w)} < \infty\right\}.$$

An alternative definition of A_∞^d is

$$A_\infty^d = \bigcup_{p>1} A_p^d,$$

moreover, if $1 < p < \infty$ then $\|w\|_{A_\infty^d} \le \|w\|_{A_p^d}$. The proof of the equivalence of these two definitions for the continuous case can be found in [GaRu], in the dyadic case we can conclude it in a similar way or see [KaPer].

Also note that $w \in A_\infty^d$ implies that w is a doubling weight $D^d(w) < \infty$ and moreover, $D^d(w) \le 2\|w\|_{A_\infty^d} e^{C\|w\|_{A_\infty^d}}$, by a similar calculation to the one above and the observation (see Theorem 2.6) that $\|w\|_{A_\infty^d} \sim \|\ln w\|_{BMO^d}$.

The other limiting case is A_1^d, which consists of all weights w, such that there exists a positive constant $C > 0$ that satisfies $m_I w \le Cw(x)$ for a.e. $x \in I$ for all dyadic intervals $I \in D$.

Typical examples of A_p^d weights are weights of the form $w = |x|^\alpha$ with $-1 < \alpha < p - 1$.

Note that both A_p^d and RH_p^d are dealing with the Reverse Hölder condition. Now let us discuss the connection between the dyadic Reverse Hölder and Muckenhoupt classes. In the continuous case every weight that satisfies Muckenhoupt condition of some order $p > 1$ (that is $w \in A_\infty$), satisfies a reverse Hölder condition of some order $q > 1$, that is, $A_\infty = \bigcup_{q>1} RH_q$ (see [CoFe]) and vice versa. However, this fails to be true in the dyadic case. As mentioned before, the dyadic Reverse Hölder $w \in RH_p^d$ condition does not imply that the weight w is a doubling weight, while A_p^d does. Also note that Buckley's example, weight $w = \chi_{\mathbb{R}\setminus[0,1]}$, is in the dyadic Reverse Hölder class RH_p^d for every p, but not in A_∞^d and, hence, not in any of the A_p^d.

The following properties in the continuous case were observed by Coifman and Fefferman in [CoFe] as well as García-Cuerva and Rubio de Francia in [GaRu]. In the dyadic case they are due to Buckley ([Buc2]); a complete proof can be found in

[KaPer].

Theorem 2.4 *Let $s > 1$.*

If $w \in A^d_\infty$ then $w^{1/s} \in RH^d_s$.

If $w \in RH^d_p$ and w is a doubling weight, then $w \in A^d_\infty$.

Sometimes it is more convenient to deal with A^d_p:

Lemma 2.5 *Let w be a weight, then for any two real numbers p and r, such that $1 < p < \infty$ and $1 \le r < \infty$,*

$$w^r \in A^d_p \quad \Leftrightarrow \quad w \in RH^d_r \bigcap A^d_{\frac{p+r-1}{r}}.$$

Moreover,

$$\|w^r\|_{A^d_p} \;\le\; \|w\|^r_{RH^d_r} \|w\|^r_{A^d_{\frac{p+r-1}{r}}}$$

and

$$\|w\|_{RH^d_r} \;\le\; \|w^r\|^{1/r}_{A^d_p}, \qquad \|w\|_{A^d_{\frac{p+r-1}{r}}} \;\le\; \|w^r\|^{1/r}_{A^d_p}.$$

In particular, when $r = p$,

$$w^p \in A^d_p \quad \Leftrightarrow \quad w \in RH^d_p \bigcap A^d_{2-\frac{1}{p}}$$

and

$$\|w^p\|_{A^d_p} \;\le\; \|w\|^p_{RH^d_p} \|w\|^p_{A^d_{2-\frac{1}{p}}},$$

$$\|w\|_{RH^d_p} \;\le\; \|w^p\|^{1/p}_{A^d_p},$$

$$\|w\|_{A^d_{2-\frac{1}{p}}} \;\le\; \|w^p\|^{1/p}_{A^d_p}.$$

This lemma follows from the following simple equation: for all dyadic intervals

$I \in D$ we can write

$$\left\{m_I(w^r)\left[m_I\left(w^{-\frac{r}{p-1}}\right)\right]^{p-1}\right\}^{1/r} = \frac{(m_I(w^r))^{1/r}}{m_I w}\left\{m_I w\left[m_I\left(w^{-\frac{1}{\frac{p+r-1}{r}-1}}\right)\right]^{\frac{p+r-1}{r}-1}\right\}. \tag{2.2}$$

If $w \in RH_r^d$, then

$$\sup_{I \in D} \frac{(m_I(w^r))^{1/r}}{m_I w} = \|w\|_{RH_r^d} < \infty$$

and, similarly, $w \in A^d_{\frac{p+r-1}{r}-1}$ implies that

$$\|w\|_{A^d_{\frac{p+r-1}{r}-1}} = \sup_{I \in D} m_I w\left[m_I\left(w^{-\frac{1}{\frac{p+r-1}{r}-1}}\right)\right]^{\frac{p+r-1}{r}-1} < \infty.$$

Then equation (2.2) implies that

$$\begin{aligned}
\|w^r\|_{A_p^d} &= \sup_{I \in D}\left\{m_I(w^r)\left[m_I\left(w^{-\frac{r}{p-1}}\right)\right]^{p-1}\right\} \\
&\leq \sup_{I \in D}\left[\frac{m_I(w^r)^{1/r}}{m_I w}\right]^r \sup_{I \in D}\left\{m_I w\left[m_I\left(w^{-\frac{1}{\frac{p+r-1}{r}-1}}\right)\right]^{\frac{p+r-1}{r}-1}\right\}^r \\
&\leq \|w\|^r_{RH_r^d}\,\|w\|^r_{A^d_{\frac{p+r-1}{r}-1}} < \infty.
\end{aligned}$$

The reverse implication and inequalities for norms hold because, by Hölder's inequality, for all dyadic intervals $I \in D$

$$\frac{m_I(w^r)^{1/r}}{m_I w} \geq 1$$

and

$$m_I w\left[m_I\left(w^{-\frac{1}{\frac{p+r-1}{r}-1}}\right)\right]^{\frac{p+r-1}{r}-1} \geq 1$$

hold for any $r \geq 1$ (inequality $\frac{p+r-1}{r} > 1$ trivially holds whenever $p > 1$).

The following theorem relates A^d_∞ and BMO^d weights. It follows from the integral form of the John-Nirenberg inequality, discussed in detail, for example, in [SlVa], where authors are concerned with sharp constants in it. The proof of it using the stopping time argument can be found in [Per1].

Theorem 2.6 *If a weight* $w \in A^d_\infty$ *then* $\ln w \in BMO^d$*. If* $b \in BMO^d$ *then* $w = e^{\delta b} \in A^d_\infty$ *for* δ *small enough. Moreover, if* $w \in A^d_\infty$*, then* $\|w\|_{A^d_\infty} \sim \|\ln w\|_{BMO^d}$*.*

We can introduce the dyadic Muckenhoupt classes with respect to a measure σ, other than the Lebesgue measure on the real line, as well. We will say that $w \in A^d_p(d\sigma)$ whenever

$$\|w\|_{A^d_p(d\sigma)} \;:=\; \sup_{I\in D} m^\sigma_I w \left(m^\sigma_I \left(w^{-\frac{1}{p-1}}\right)\right)^{p-1} \;<\; \infty$$

and similarly $w \in A^d_\infty(d\sigma)$ when

$$\|w\|_{A^d_\infty(d\sigma)} \;:=\; \sup_{I\in D} m^\sigma_I w \; e^{-m^\sigma_I(\ln w)} \;<\; \infty.$$

In some cases we will be dealing with the two-weighted Muckenhoupt condition. A pair of weights (v, w) that satisfies

$$[v, w]_{A^d_p} \;:=\; \sup_{I\in D} m_I v \;\left(m_I \left(w^{\frac{1}{p-1}}\right)\right)^{p-1} \;< \infty$$

is said to be of class A^d_p. The quantity $[v, w]_{A^d_p}$ is called the A^d_p-*characteristic of the pair.*

We will mostly be dealing with A^d_2: $(v, w) \in A^d_2$ if and only if

$$[v, w]_{A^d_2} \;:=\; \sup_{I\in D} m_I v \; m_I w \;< \infty.$$

And similarly we can define the $(d\sigma)$ version:

the pair of weights (v,w) belongs to the $A_p^d(d\sigma)$-class if and only if

$$[v,w]_{A_2^d(d\sigma)} \;:=\; \sup_{I\in D} m_I^\sigma v \;\left(m_I^\sigma\left(w^{\frac{1}{p-1}}\right)\right)^{p-1} \;<\infty.$$

Note that when $w=v^{-1}$, the condition $(v,w)\in A_p^d$ or $(v,w)\in A_p^d(d\sigma)$ is equivalent to $v\in A_p^d$ or $v\in A_p^d(d\sigma)$.

Also note that in the one weight condition, by Hölder's inequality, the inverse inequality holds, i.e. if $w=v^{-1}$ we know that for all dyadic intervals $I\in D$

$$1 \;\le\; m_I^\sigma v\left(m_I^\sigma\left(v^{-\frac{1}{p-1}}\right)\right)^{p-1}$$

and in particular for the Lebesgue measure

$$1 \;\le\; m_I v\left(m_I\left(v^{-\frac{1}{p-1}}\right)\right)^{p-1},$$

while in general, for $w\neq v^{-1}$, the above inequalities do not have to be true for all dyadic intervals I.

It might be natural to define the two weight Muckenhoupt condition $(v,w)\in A_p^d(d\sigma)$ to be

$$\begin{aligned}(v,w)\in A_p^d(d\sigma) \;\Leftrightarrow 0 \;&<\; \inf_{I\in D} m_I^\sigma v\left(m_I^\sigma\left(w^{\frac{1}{p-1}}\right)\right)^{p-1}\\ &\le\; \sup_{I\in D} m_I^\sigma v\left(m_I^\sigma\left(w^{\frac{1}{p-1}}\right)\right)^{p-1} \;<\; \infty.\end{aligned}$$

2.2.3 C_s^d condition

The C_s^d condition was first introduced in [KaPer] in relation to the Haar multiplier operators and is very convenient in this context.

We will say that that the weight w satisfies *condition* C_s^d, $s\in\mathbb{R}$, if there exists a constant $C > 0$, such that

$$m_I w^s \;\le\; C\,(m_I w)^s, \quad \forall I \in D.$$

The smallest such constant will be denoted by $\|w\|_{C_s^d}$. Note that when $s > 1$, the C_s^d condition coincides with the reverse Hölder condition, and we have

$$\|w\|_{C_s^d} \;=\; \|w\|_{RH_s^d}^s \quad (s > 1).$$

For $s < 0$, C_s^d condition is nothing but $A^d_{1-\frac{1}{s}}$:

$$w \in A^d_{1-\frac{1}{s}} \;\Leftrightarrow\; m_I w\,(m_I(w^s))^{-\frac{1}{s}} \;\le\; \|w\|_{A^d_{1-\frac{1}{s}}} \;<\; \infty,$$
$$w \in C_s^d \;\Leftrightarrow\; (m_I w)^{-s}\, m_I w^s \;\le\; \|w\|_{C_s^d} \;<\; \infty.$$

Since $-s > 0$, it is equivalent to:

$$m_I w\;(m_I w^s)^{-\frac{1}{s}} \;\le\; \|w\|_{C_s^d}^{-\frac{1}{s}}$$

and hence

$$\|w\|_{C_s^d} \;=\; \|w\|^{-s}_{A^d_{1-\frac{1}{s}}},$$

or, alternatively,

$$\|w\|^{\frac{1}{p-1}}_{A_p^d} \;=\; \|w\|_{C^d_{\frac{1}{1-p}}}.$$

For $0 \le s \le 1$ C_s condition coincides with Hölder's inequality and always holds with constant 1.

2.2.4 Weighted Haar basis

The weighted Haar functions were first introduced in [CoJS] in 1989 and are extremely useful in the weighted dyadic inequalities. In many cases discussed in this work, as well as in [Per2], [PetPo] weighted Haar functions allow us to split the sum we need to bound into two parts. One part is almost a weighted inequality on a weighted space and behaves similarly to the unweighted situations (Bessel's inequality is often useful for the first part). While the other part of the sum is simpler than the original expression, which allows us to obtain a desired bound in most cases.

Note that the weighted Haar functions we are introducing in this section are not normalized.

First define for a measure σ and dyadic interval $I \in D$

$$h_I^{d\sigma}(x) := \begin{cases} \sqrt{\frac{\sigma(I^-)}{\sigma(I^+)}} \quad , & x \in I^+ \\ -\sqrt{\frac{\sigma(I^+)}{\sigma(I^-)}} \quad , & x \in I^- \\ 0 \quad , & x \notin I \end{cases}$$

Then, if σ is the Lebesgue measure on $\mathbb{R}$

$$h_I^{dx}(x) := \begin{cases} 1 \quad , & x \in I^+ \\ -1 \quad , & x \in I^- \\ 0 \quad , & x \notin I \end{cases}$$

Note that the weighted Haar functions have zero mean,

$$\begin{aligned} \int_I h_I^{d\sigma}(x) d\sigma &= \int_{I^+} \sqrt{\frac{\sigma(I^-)}{\sigma(I^+)}} d\sigma - \int_{I^-} \sqrt{\frac{\sigma(I^+)}{\sigma(I^-)}} d\sigma \\ &= \sqrt{\sigma(I^+)\sigma(I^-)} - \sqrt{\sigma(I^+)\sigma(I^-)} = 0 \end{aligned}$$

and we can compute their $L_2(d\sigma)$ norms,

$$\begin{aligned}
\left\|h_I^{d\sigma}\right\|_{L_2(d\sigma)}^2 &= \int_I \left(h_I^{d\sigma}\right)^2 d\sigma \;=\; \int_{I^+} \frac{\sigma(I^-)}{\sigma(I^+)} d\sigma \;+\; \int_{I^-} \frac{\sigma(I^+)}{\sigma(I^-)} d\sigma \\
&= \sigma(I^-) \;+\; \sigma(I^+) \;=\; \sigma(I).
\end{aligned}$$

As a consequence of the zero mean dyadic filtration properties, we conclude that $\left\{h_I^{d\sigma}\right\}_{I\in D}$ are orthogonal in $L_2(d\sigma)$ with norms $\left\|h_I^{d\sigma}\right\|_{L_2(d\sigma)} = \sqrt{\sigma(I)}$. Then we can write

$$\begin{aligned}
h_I^{dx}(x) &= \frac{2}{\sqrt{\frac{\sigma(I^-)}{\sigma(I^+)}} + \sqrt{\frac{\sigma(I^+)}{\sigma(I^-)}}} \left[h_I^{d\sigma}(x) - \frac{\sqrt{\frac{\sigma(I^-)}{\sigma(I^+)}} - \sqrt{\frac{\sigma(I^+)}{\sigma(I^-)}}}{2} \chi_I \right] \\
&= \frac{2\sqrt{\sigma(I^+)\sigma(I^-)}}{\sigma(I^-) + \sigma(I^+)} \left[h_I^{d\sigma}(x) - \frac{\sigma(I^-) - \sigma(I^+)}{2\sqrt{\sigma(I^+)\sigma(I^-)}} \chi_I \right] \\
&= \frac{2\sqrt{\sigma(I^+)\sigma(I^-)}}{\sigma(I)} \left[h_I^{d\sigma}(x) + \frac{\sigma(I^+) - \sigma(I^-)}{2\sqrt{\sigma(I^+)\sigma(I^-)}} \chi_I \right].
\end{aligned}$$

Check: $x \in I^+$:

$$\begin{aligned}
1 &= \frac{2\sqrt{\sigma(I^+)\sigma(I^-)}}{\sigma(I)} \left[\sqrt{\frac{\sigma(I^-)}{\sigma(I^+)}} + \frac{\sigma(I^+) - \sigma(I^-)}{2\sqrt{\sigma(I^+)\sigma(I^-)}} \right] \\
&= \frac{2\sqrt{\sigma(I^+)\sigma(I^-)}}{\sigma(I)} \left[\frac{2\sigma(I^-) + \sigma(I^+) - \sigma(I^-)}{2\sqrt{\sigma(I^+)\sigma(I^-)}} \right] \;=\; 1.
\end{aligned}$$

$x \in I^-$:

$$\begin{aligned}
-1 &= \frac{2\sqrt{\sigma(I^+)\sigma(I^-)}}{\sigma(I)} \left[-\sqrt{\frac{\sigma(I^+)}{\sigma(I^-)}} + \frac{\sigma(I^+) - \sigma(I^-)}{2\sqrt{\sigma(I^+)\sigma(I^-)}} \right] \\
&= \frac{2\sqrt{\sigma(I^+)\sigma(I^-)}}{\sigma(I)} \left[\frac{-2\sigma(I^+) + \sigma(I^+) - \sigma(I^-)}{2\sqrt{\sigma(I^+)\sigma(I^-)}} \right] \;=\; -1.
\end{aligned}$$

Say we have two measures, $d\sigma$ and $d\mu$, $d\mu \ll d\sigma$ ($d\mu = wd\sigma$). Then we can write

$$\begin{aligned} h_I^{dx} &= \frac{2\sqrt{\sigma(I^+)\sigma(I^-)}}{\sigma(I)}\left[h_I^{d\sigma} + \frac{\sigma(I^+)-\sigma(I^-)}{2\sqrt{\sigma(I^+)\sigma(I^-)}}\chi_I\right] \\ &= \frac{2\sqrt{\mu(I^+)\mu(I^-)}}{\mu(I)}\left[h_I^{d\mu} + \frac{\mu(I^+)-\mu(I^-)}{2\sqrt{\mu(I^+)\mu(I^-)}}\chi_I\right]. \end{aligned}$$

Note that $\mu(I) = \int_I d\mu = \int_I wd\sigma = \sigma(I)\, m_I^\sigma w$.

$$h_I^{dx} = \frac{2\sqrt{\sigma(I^+)\sigma(I^-)}\sqrt{m_{I^+}^\sigma w\; m_{I^-}^\sigma w}}{\sigma(I)m_I^\sigma w}\left[h_I^{d\mu} + \frac{\sigma(I^+)m_{I^+}^\sigma w - \sigma(I^-)m_{I^-}^\sigma w}{2\sqrt{\sigma(I^+)\sigma(I^-)}\sqrt{m_{I^+}^\sigma w\; m_{I^-}^\sigma w}}\chi_I\right],$$

then

$$\begin{aligned} h_I^{d\sigma} &= \frac{\sqrt{m_{I^+}^\sigma w\; m_{I^-}^\sigma w}}{m_I^\sigma w}h_I^{d\mu} + \frac{\sigma(I^+)m_{I^+}^\sigma w - \sigma(I^-)m_{I^-}^\sigma w}{2m_I^\sigma w\sqrt{\sigma(I^+)\sigma(I^-)}}\chi_I - \frac{\sigma(I^+)-\sigma(I^-)}{2\sqrt{\sigma(I^+)\sigma(I^-)}}\chi_I \\ &= \frac{\sqrt{m_{I^+}^\sigma w\; m_{I^-}^\sigma w}}{m_I^\sigma w}h_I^{d\mu} + \frac{\sigma(I^+)m_{I^+}^\sigma w - \sigma(I^-)m_{I^-}^\sigma w - [\sigma(I^+)-\sigma(I^-)]\, m_I^\sigma w}{2m_I^\sigma w\sqrt{\sigma(I^+)\sigma(I^-)}}\chi_I. \end{aligned}$$

Let us simplify the coefficient in front of χ_I:

$$\frac{\sigma(I^+)m_{I^+}^\sigma w - \sigma(I^-)m_{I^-}^\sigma w - \frac{\sigma(I^+)-\sigma(I^-)}{\sigma(I)}\left(\sigma(I^+)m_{I^+}^\sigma w + \sigma(I^-)m_{I^-}^\sigma w\right)}{2m_I^\sigma w\sqrt{\sigma(I^+)\sigma(I^-)}}$$

$$= \frac{\sigma(I^+)\sigma(I)m_{I^+}^\sigma w - \sigma(I^-)\sigma(I)m_{I^-}^\sigma w - \sigma(I^+)^2 m_{I^+}^\sigma w}{2m_I^\sigma w\;\sigma(I)\sqrt{\sigma(I^+)\sigma(I^-)}}$$

$$+ \frac{\sigma(I^+)\sigma(I^-)m_{I^+}^\sigma w - \sigma(I^+)\sigma(I^-)m_{I^-}^\sigma w + \sigma(I^-)^2 m_{I^-}^\sigma w}{2m_I^\sigma w\;\sigma(I)\sqrt{\sigma(I^+)\sigma(I^-)}}$$

$$= \frac{m_{I^+}^\sigma w\left[\sigma(I^+)\sigma(I) - \sigma(I^+)^2 + \sigma(I^+)\sigma(I^-)\right] - m_{I^-}^\sigma w\left[\sigma(I^-)\sigma(I) + \sigma(I^+)\sigma(I^-) - \sigma(I^-)^2\right]}{2m_I^\sigma w\;\sigma(I)\sqrt{\sigma(I^+)\sigma(I^-)}}$$

$$= \frac{2\sigma(I^+)\sigma(I^-)m_{I^+}^\sigma w \;-\; 2\sigma(I^+)\sigma(I^-)m_{I^-}^\sigma w}{2m_I^\sigma w\;\sigma(I)\sqrt{\sigma(I^+)\sigma(I^-)}}$$

$$= \frac{\sqrt{\sigma(I^+)\sigma(I^-)}}{\sigma(I)}\frac{m_{I^+}^\sigma w - m_{I^-}^\sigma w}{m_I^\sigma w}.$$

So,

$$h_I^{d\sigma} = \frac{\sqrt{m_{I^+}^\sigma w \; m_{I^-}^\sigma w}}{m_I^\sigma w} h_I^{d\mu} + \frac{\sqrt{\sigma(I^+)\sigma(I^-)}}{\sigma(I)} \frac{m_{I^+}^\sigma w - m_{I^-}^\sigma w}{m_I^\sigma w} \chi_I. \tag{2.3}$$

Note that $\left\{ \frac{h_I^{d\sigma}}{\sqrt{\sigma(I)}} \right\}_{I \in D}$ is an orthonormal basis in $L_2(d\sigma)$, so for all functions f, square integrable with respect to measure σ, we have

$$\|f\|_{L_2(d\sigma)}^2 \;=\; \sum_{I \in D} \frac{\left\langle f ; h_I^{d\sigma} \right\rangle_{d\sigma}^2}{\sigma(I)}.$$

Sometimes it is more convenient to deal with simpler functions. Whenever possible, we will be using system of functions $\{H_I^w\}_{I \in D}$ defined by

$$H_I^w \;:=\; h_I^{dx} - A_I^w \chi_I, \quad where \quad A_I^w := \frac{m_{I^+}w - m_{I^-}w}{2 m_I w}. \tag{2.4}$$

The functions $\{H_I^w\}_{I \in D}$ form an orthogonal system in $L_2(w)$, suffices to verify that the functions have zero mean when integrated with respect to wdx,

$$\begin{aligned} \int H_I^w w \;&=\; \int_I h_I^{dx} w - \int_I \frac{m_{I^+}w - m_{I^-}w}{2 m_I w} w \\ &=\; \left|I^+\right| \; m_{I^+}w - \left|I^-\right| \; m_{I^-}w - \frac{|I|}{2} \left(m_{I^+}w - m_{I^-}w\right) \;=\; 0. \end{aligned}$$

Moreover, their $L_2(w)$ norms $\|H_I^w\|_{L_2(w)}$ satisfy the inequality $\|H_I^w\|_{L_2(w)} \leq \sqrt{|I| \; m_I w}$:

$$\begin{aligned} \|H_I^w\|_{L_2(w)}^2 \;&=\; \int_I \left(H_I^w\right)^2 w \\ &=\; \int_I \left(h_I^{dx}\right)^2 w - 2 \int_I h_I^{dx} \frac{m_{I^+}w - m_{I^-}w}{2 m_I w} w + \int_I \left(\frac{m_{I^+}w - m_{I^-}w}{2 m_I w} \right)^2 w \\ &=\; |I| \; m_I w - |I| \frac{\left(m_{I^+}w - m_{I^-}w\right)^2}{2 m_I w} + |I| \frac{\left(m_{I^+}w - m_{I^-}w\right)^2}{4 m_I w} \\ &=\; |I| \left(m_I w - \frac{\left(m_{I^+}w - m_{I^-}w\right)^2}{4 m_I w} \right) \\ &\leq\; |I| m_I w. \end{aligned}$$

Therefore, by Bessel's inequality we have:

$$\forall g \in L_2(w) \quad \sum_{I \in D} \frac{\langle g; H_I^w \rangle_{wdx}^2}{|I| m_I w} \ \le \ \|g\|_{L_2(w)}^2 ,$$

which we can also rewrite as:

$$\forall f \in L_2 \quad \sum_{I \in D} \frac{\left\langle f w^{1/2}; H_I^w \right\rangle^2}{|I| m_I w} \ \le \ \|f\|_{L_2}^2 . \tag{2.5}$$

We will refer to both systems $\left\{h_I^{d\sigma}\right\}_{I \in D}$ and $\left\{H_I^w\right\}_{I \in D}$ as to weighted Haar functions. It should be clear from the context which one of them we are using.

3 Main tools

We are ready to state the theorems we are going to use in the next sections.

3.1 The Carleson / Sawyer type theorems

The problem Carleson was working on was to classify positive measures $\mu(x,t)$ on the upper half plane $\mathbb{R}^2_+$, for which the mapping that takes square integrable functions on $\mathbb{R}$ to their harmonic extension to $\mathbb{R}^2_+$ is bounded from $L_2(\mathbb{R})$ into $L_2(\mathbb{R}^2_+, d\mu)$. More precisely, let $u(x+it) = f * P_t(x)$, $x, t \in \mathbb{R}$, $t > 0$ be harmonic extension of f, where $P_t = \frac{1}{\pi}\frac{t}{x^2+t^2}$ is the Poisson kernel. Then the question is: for which measures μ does there exist a constant $C > 0$, such that for all $f \in L_2(\mathbb{R})$

$$\int_{\mathbb{R}^2_+} |f * P_t(x)|^2 \, d\mu(x,t) \;\leq\; C \int_{\mathbb{R}} |f|^2 \quad ?$$

The answer is "if and only if the measure μ is a Carleson measure", i.e. there exists a constant $C > 0$, such that for all intervals $I \subset \mathbb{R}$

$$\mu(Q_I) \;\leq\; C\,|I|,$$

where $Q_I = \{(x,t): \;\; x \in I, \;\; 0 \leq t \leq |I|\}$ is the *Carleson box* corresponding to I. The Carleson Embedding Theorem (Theorem 3.1 below) is the dyadic analogue of this theorem. The Poisson averaging is replaced by averaging the function on the upper halves $T_I = \left\{z = x + it \in Q_I : \;\; t \geq \frac{|I|}{2}\right\}$ of the corresponding dyadic Carleson boxes, that is we define extension to $\mathbb{R}^2_+$ by:

$$u_d(x+it) \;:=\; \sum_{I \in D} (m_I f) \chi_{T_I}(x+it).$$

Note that the collection $\{T_I\}_{I\in D}$ provides a partition of $\mathbb{R}^2_+$. The problem is the same as before, except that now all the information we need from the measure is the mass of T_I, namely the positive sequence $\mu(T_I) =: \lambda_I$. Observe that $\mu(Q_J) = \sum_{I\in D(J)} \lambda_I$.

Theorem 3.1 (Carleson Embedding Theorem) *Given a Carleson sequence $\{\lambda_I\}_{I\in D}$, i.e.*

$$\forall J \in D \quad \frac{1}{|J|} \sum_{I\in D(J)} \lambda_I \leq Q,$$

then for all $f \in L_2(\mathbb{R})$

$$\sum_{I\in D} |m_I f|^2 \lambda_I \ \leq \ C\, Q \ \|f\|^2_{L_2}$$

holds with numerical constant $C > 0$, independent of Q, dyadic interval J, sequence $\{\lambda_I\}$ and function $f \in L_2$.

In [NTr] the authors presented a Bellman function proof of the above Carleson Embedding theorem. One can find both classical and Bellman function proofs in [Per1]. The Bellman function approach allowed to extend the Carleson Embedding theorem to the weighted case (see [NTrVo2], [Per1]).

Theorem 3.2 (Weighted Carleson Embedding theorem) *Let $\{\alpha_I\}_{I\in D}$ be a sequence of nonnegative numbers and v be a weight. If there exists a positive constant $Q > 0$ that for any dyadic interval $J \in D$*

$$\frac{1}{|J|} \sum_{I\in D(J)} \alpha_I m_I^2 v \ \leq \ Q m_J v \tag{3.1}$$

then

$$\sum_{I\in D} \alpha_I m_I^2 (f v^{\frac{1}{2}}) \ \leq \ CQ\, \|f\|^2_{L_2} \tag{3.2}$$

holds with some universal numerical constant $C > 0$ for any square integrable function f.

One can interpret the Weighted Carleson Embedding theorem as a Sawyer's type estimate (as discussed in the beginning of Section 3.2) as well. We can think of (3.1) as the inequality (3.2) on the test functions of the form $f^{(J)} = v^{1/2}\chi_J$, $J \in D$.

The following theorem from [Per2] extends the Carleson Embedding Theorem even further, to the $d\sigma$ case. We will need it in Section 4 where we discuss the square function, in all other results we use Theorem 3.2.

Theorem 3.3 ($d\sigma$-Sawyer's Estimate) *Given a dyadic doubling measure σ, $Q > 0$ and a sequence of positive numbers $\{\lambda_I\}_{I\in D}$. Suppose that for all dyadic intervals J*

$$\frac{1}{\sigma(J)} \sum_{I\in D(J)} |m_I^\sigma v|^2 \lambda_I \;\leq\; Q\, m_J^\sigma v, \tag{3.3}$$

then $\forall J \in D$

$$\frac{1}{\sigma(J)} \sum_{I\in D(J)} \left|m_I^\sigma (f v^{1/2})\right|^2 \lambda_I \;\leq\; 4Q\, m_J^\sigma f^2.$$

The following proposition is in the spirit of the Carleson Embedding Theorem as well.

Proposition 3.4 *Let v be a weight, such that v^{-1} is a weight as well, and let $\{\lambda_I\}$ be a Carleson sequence of nonnegative numbers, that is, there exists a constant $Q > 0$ s.t.*

$$\forall J \in D \quad \frac{1}{|J|} \sum_{I\in D(J)} \lambda_I \;\leq\; Q,$$

then $\forall J \in D$

$$\frac{1}{|J|} \sum_{I\in D(J)} \frac{\lambda_I}{m_I v^{-1}} \;\leq\; 4Q\, m_J v \tag{3.4}$$

and therefore if $v \in A_2^d$ then for any $J \in D$ we have:

$$\frac{1}{|J|} \sum_{I\in D(J)} m_I v \lambda_I \;\leq\; 4Q\, \|v\|_{A_2^d}\, m_J v. \tag{3.5}$$

Proof of this proposition can be found in Section 7.2.

In fact we can prove the following stronger A_∞^d version of Proposition 3.4:

Proposition 3.5 *Let v be a weight, such that $\log v \in L_1^{loc}$, let $\{\lambda_I\}_{I\in D}$ be a Carleson sequence:*

$$\exists Q > 0 \quad s.t. \quad \forall J \in D \quad \frac{1}{|J|}\sum_{I\in D(J)} \lambda_I \;\leq\; Q$$

and v be a weight. Then $\forall J \in D$

$$\frac{1}{|J|}\sum_{I\in D(J)} e^{m_I(\log v)}\lambda_I \;\leq\; 4Q\; m_J v \tag{3.6}$$

and therefore, if $v \in A_\infty^d$ then for any dyadic interval $J \in D$ we have:

$$\frac{1}{|J|}\sum_{I\in D(J)} m_I v\; \lambda_I \;\leq\; 4Q\;\; \|v\|_{A_\infty^d}\, m_J v. \tag{3.7}$$

Proof of Proposition 3.5 can be found in Section 7.3.

3.2 The bilinear embedding theorems

In [Saw1] Sawyer showed that the maximal operator is bounded from $L_2(u)$ to $L_2(v)$ if and only if it is bounded on a very special class of "simple" functions of the form $u^{-1}\chi_I$, $I \subset \mathbb{R}$. Later in [Saw2] he extended the theorem to a larger class of integral operators. The difference from the case of maximal functions is that one should check the boundedness of the adjoint operator T^* on the functions of the form $v\chi_I$, $I \in D$ as well. The following theorem is a theorem of the Sawyer's type by Nazarov, Treil and Volberg and it was presented in [NTrVo2]. It is the key tool in our estimate on the norm of the dyadic paraproduct (Section 6).

Theorem 3.6 (Bilinear Embedding theorem by Nazarov, Treil, Volberg) *Let v and w be weights, $\{\alpha_I\}$ be a sequence of nonnegative numbers. Define a family* *of self adjoint operators*

$$T^{(J)} f \ := \sum_{I \in D(J)} \alpha_I \, m_I f \ \chi_I .$$

Assume that the weights v and w satisfy the following two conditions for all dyadic intervals J:

$$(1) \qquad \frac{1}{|J|} \int_J \left[T^{(J)} w \right]^2 v \ \leq \ Q^2 \, m_J w, \tag{3.8}$$

$$(2) \qquad \frac{1}{|J|} \int_J \left[T^{(J)} v \right]^2 w \ \leq \ Q^2 \, m_J v. \tag{3.9}$$

Then for any two non-negative functions $f, g \in L_2$,

$$\sum_{I \in D} \alpha_I \, m_I \left(f v^{1/2} \right) \ m_I \left(g w^{1/2} \right) |J| \ \leq \ CQ \, \|f\|_{L_2} \|g\|_{L_2} .$$

Moreover, condition (1) implies the inequality

$$\sum_{I \in D_1} \alpha_I \, m_I (f v^{1/2}) \, m_I (g w^{1/2}) |I| \ \leq \ CQ \, \|f\|_{L_2} \|g\|_{L_2} ,$$

where $D_1 := \left\{ I \in D \mid \frac{m_I^2 (f v^{1/2})}{m_I (f^2) \, m_I v} \geq \frac{m_I^2 (g w^{1/2})}{m_I (g^2) \, m_I w} \right\}$ and similarly condition (2) implies the inequality

$$\sum_{I \in D_2} \alpha_I \, m_I (f v^{1/2}) \, m_I (g w^{1/2}) |I| \ \leq \ CQ \, \|f\|_{L_2} \|g\|_{L_2} ,$$

where $D_2 = D \setminus D_1$.

We will only use the first part of the theorem. Let us rewrite it in a more convenient way. First, we write the left-hand side of (3.8) and (3.9) as sums:

$$
\begin{aligned}
\frac{1}{|J|}\int_J \left[T^{(J)}w\right]^2 v &= \frac{1}{|J|}\int_J \left(\sum_{I\in D(J)} \alpha_I\, m_I w\, \chi_I\right)^2 v \\
&\sim \frac{1}{|J|}\int_J \sum_{I\in D(J)} \alpha_I\, m_I w\, \chi_I \left(\sum_{K\in D(I)} \alpha_K\, m_K w\, \chi_K\right) v \\
&= \frac{1}{|J|}\sum_{I\in D(J)} \alpha_I\, m_I w \left(\sum_{K\in D(I)} \alpha_K\, m_K w \int_J \chi_K\, v\right) \\
&= \frac{1}{|J|}\sum_{I\in D(J)} \alpha_I\, m_I w \left(\sum_{K\in D(I)} \alpha_K\, m_K w\, m_K v\, |K|\right) \\
&= \frac{1}{|J|}\sum_{I\in D(J)} \alpha_I\, m_I w\, |I| \left(\frac{1}{|I|}\sum_{K\in D(I)} \alpha_K\, m_K w\, m_K v\, |K|\right).
\end{aligned}
$$

Similarly,

$$
\frac{1}{|J|}\int_J \left[T^{(J)}v\right]^2 w \sim \frac{1}{|J|}\sum_{I\in D(J)} \alpha_I\, m_I v\, |I| \left(\frac{1}{|I|}\sum_{K\in D(I)} \alpha_K\, m_K w\, m_K v\, |K|\right).
$$

Hence, we can state the Bilinear Embedding Theorem in the following way, similar to the one found in [Wit1].

Corollary 3.7 *Let v and w be weights. Let $\{\alpha_I\}$ be a sequence of nonnegative numbers, such that for all dyadic intervals $I \in D$ the following two inequalities hold with some constant $Q > 0$:*

$$
\frac{1}{|J|}\sum_{I\in D(J)} \alpha_I\, m_I w\, |I| \left(\frac{1}{|I|}\sum_{K\in D(I)} \alpha_K\, m_K w\, m_K v\, |K|\right) \le Q^2\, m_J w \tag{3.10}
$$

and

$$\frac{1}{|J|}\sum_{I\in D(J)}\alpha_I\; m_Iv\;|I|\left(\frac{1}{|I|}\sum_{K\in D(I)}\alpha_K\; m_Kw\; m_Kv\;|K|\right)\;\le\; Q^2\; m_Jv, \tag{3.11}$$

then for any two nonnegative functions $f, g \in L_2$

$$\sum_{I\in D}\alpha_I\; m_I\left(fv^{1/2}\right)\; m_I\left(gw^{1/2}\right)\;|I|\;\le\; CQ\;\|f\|_{L_2}\|g\|_{L_2} \tag{3.12}$$

holds with some numerical constant $C > 0$.

Now it is easy to see that when for all $I, J \in D$

$$\frac{1}{|I|}\sum_{K\in D(I)}\alpha_K\; m_Kw\; m_Kv\;|K|\;\le\; Q,$$

$$\frac{1}{|J|}\sum_{I\in D(J)}\alpha_I\; m_Iw\;|I|\;\le\; Q\; m_Jw$$

and

$$\frac{1}{|J|}\sum_{I\in D(J)}\alpha_I\; m_Iv\;|I|\;\le\; Q\; m_Jv,$$

all conditions of Theorem 3.7 will be satisfied. We can state it as a corollary.

Corollary 3.8 *Let* v *and* w *be weights. Let* $\{\alpha_I\}$ *be a sequence of nonnegative numbers such that for all dyadic intervals* $J \in D$ *the following three inequalities hold with some constant* $Q > 0$*:*

$$\frac{1}{|J|}\sum_{I\in D(I)}\alpha_I\; m_Iw\; m_Iv\;|I|\;\le\; Q,$$

$$\frac{1}{|J|}\sum_{I\in D(J)}\alpha_I\; m_Iw\;|I|\;\le\; Q\; m_Jw$$

and

$$\frac{1}{|J|}\sum_{I\in D(J)} \alpha_I\, m_I v\, |I| \;\le\; Q\, m_J v.$$

Then for any two nonnegative functions $f, g \in L_2$

$$\sum_{I\in D} \alpha_I\, m_I\left(f v^{1/2}\right)\; m_I\left(g w^{1/2}\right) |I| \;\le\; CQ\; \|f\|_{L_2}\, \|g\|_{L_2}$$

holds with some numerical constant $C > 0$.

A theorem, similar to Corollary 3.8 was used by Petermichl in the proof of the boundedness of the Hilbert transform on the weighted Lebesgue spaces, where it was proved directly, see [Pet1].

3.3 The weight lemmas / Buckley's type bounds

The embedding theorems from Sections 3.1 and 3.2 allow us to replace functions in the inequalities by weights. Bounds of the Buckley's type will allow us to handle this case.

In his thesis, as well as in [Buc2], Buckley proves the following dyadic theorem:

Theorem 3.9 (Buckley) *Suppose* w *is a weight.*

(i) For $s \neq 0$, $s \neq 1$, $w \in C_s^d$ *if and only if*

$$\forall J \in D \quad \frac{1}{|J|}\sum_{I\in D(J)} (m_I w)^s \left(\frac{m_{I^+}w - m_{I^-}w}{m_I w}\right)^2 |I| \;\le\; C\; \|w\|_{C_s^d} (m_J w)^s \qquad (3.13)$$

with positive constant C *that only depends on* p.

(ii) For $s = 0$, $w \in A_\infty^d$ *if and only if* w *is a doubling weight and*

$$\forall J \in D \quad \frac{1}{|J|}\sum_{I\in D(J)} \left(\frac{m_{I^+}w - m_{I^-}w}{m_I w}\right)^2 |I| \;\le\; C\; \log\left(\|w\|_{A_\infty^d}\right) \qquad (3.14)$$

(Fefferman-Kenig-Pipher inequality).

(iii) For $s = 1$, $w \in A^d_\infty$ if and only if w is a doubling weight and

$$\forall J \in D \quad \frac{1}{|J|} \sum_{I \in D(J)} m_I w \left(\frac{m_{I^+} w - m_{I^-} w}{m_I w} \right)^2 |I| \ \leq \ K \, m_J w \tag{3.15}$$

(Buckley's inequality),

with some positive constant K that may depend on the A^d_∞-characteristic of the weight w.

The Bellman function proof of (3.14), very similar to Buckley's original proof, can be found in [NTrVo2].

Note that (ii) and a continuous version of (i) are due to Fefferman, Kenig and Pipher, see [FeKPi]. Note also, that (3.13) and (3.14) trace the dependence on the C^d_s and A^d_∞ constants of the weight w, and the dependences are sharp. Note however, that (3.15) does not give the dependence of constant K on the A^d_∞-characteristics of the weight w.

Some partial results about the constant K in (3.15) are known. Wittwer traced the constant in Buckley's proof for the case when $w \in A^d_2$, see [Wit1]:

Theorem 3.10 (Wittwer's sharp version of Buckley's inequality) *There exists a positive constant $C > 0$ such that for any weight $w \in A^d_2$ and every dyadic interval $J \in D$:*

$$\frac{1}{|J|} \sum_{I \in D(J)} \left(\frac{m_{I^+} w - m_{I^-} w}{m_I w} \right)^2 |I| m_I w \ \leq \ C \, \|w\|_{A^d_2} m_J w.$$

In [Per2] Pereyra presented a Bellman function proof of a similar result, which allowed her to extend it to the $(d\sigma)$-case:

Theorem 3.11 ($A^d_2(d\sigma)$-Weight Lemma) *There exists a positive constant $C > 0$, such that if $v \in A^d_2(d\sigma)$ and σ is a doubling dyadic measure, then for every dyadic* *interval* $J \in D$

$$\frac{1}{\sigma(J)} \sum_{I \in D(J)} \frac{\left(m^{\sigma}_{I+}v - m^{\sigma}_{I-}v\right)^2}{m^{\sigma}_{I}v} \sigma(I) \ \le\ C\ \left(D^d(d\sigma)\right)^3 \|v\|_{A^d_2(d\sigma)}\, m^{\sigma}_J v.$$

Unfortunately, this $A^d_2(d\sigma)$ result is not enough for us. In order to extrapolate the Square function result, we needed an $A^d_p(d\sigma)$-bound for some $p > 2$. Fortunately, the proof from [Per2], with only minor changes in the Bellman function and the domain, extends to the following:

Proposition 3.12 ($A^d_p(d\sigma)$-Weight Lemma) *If* $v \in A^d_p(d\sigma)$ *for some* $p > 1$ *and* σ *is a dyadic doubling measure (i.e.* $D^d(d\sigma) := \sup_{I \in D} \frac{\sigma(\tilde{I})}{\sigma(I)} < \infty$, *where* $\tilde{I}$ *is a parent of* I*), then for every dyadic interval* $J \in D$

$$\frac{1}{\sigma(J)} \sum_{I \in D(J)} \frac{\left(m^{\sigma}_{I+}v - m^{\sigma}_{I-}v\right)^2}{m^{\sigma}_{I}v} \sigma(I) \ \le\ A\ m^{\sigma}_J v,$$

holds with constant $A = C^{\frac{2p-1}{p}} \left(D^d(d\sigma)\right)^{p+1} \|w\|_{A^d_p}$.

The proof of the A^d_p-Weight Lemma (Proposition 3.12), very similar to the one from [Per2], can be found in Section 7.6.

So, whenever $w \in A^d_p$ for some $1 < p < \infty$, the dependence of the constant K on the A^d_p-characteristic of the weight w in (3.15) is linear. We expect it to depend linearly on the A^d_∞-characteristic of the weight w as well, but as far as we know, it is an open problem.

Note that it follows from (3.14) and Proposition 3.5 that

$$\forall J \in D \quad \frac{1}{|J|} \sum_{I \in D(J)} m_I w \left(\frac{m_{I+}w - m_{I-}w}{m_I w}\right)^2 |I| \ \le\ C\ \|w\|_{A^d_\infty} \log \|w\|_{A^d_\infty} m_J w.$$

At the end of this section let us state a few more propositions, similar to Buckley's theorem, which we are going to use in Section 6 to obtain the bound on the norm of the dyadic paraproduct on weighted Lebesgue spaces.

Proposition 3.13 *There exists a positive constant $C > 0$, such that whenever w is a weight, such that w^{-1} is a weight as well, then for all dyadic intervals J:*

$$\frac{1}{|J|}\sum_{I\in D(J)}\frac{(m_{I^+}w - m_{I^-}w)^2}{(m_I w)^3}|I| \;\leq\; C\, m_J(w^{-1})$$

and therefore, if $w \in A_2^d$, the following inequality holds for all dyadic intervals $J \in D$:

$$\frac{1}{|J|}\sum_{I\in D(J)}\left(\frac{m_{I^+}w - m_{I^-}w}{m_I w}\right)^2 |I|\, m_I(w^{-1}) \;\leq\; C\,\|w\|_{A_2^d}\, m_J(w^{-1}). \tag{3.16}$$

The proof of Proposition 3.13 can be found in Section 7.4.

Proposition 3.14 *There exists a positive numerical constant $C > 0$, such that whenever w is a weight, such that w^{-1} is a weight as well, then $\forall J \in D$*

$$\frac{1}{|J|}\sum_{I\in D(J)}\left(\frac{m_{I^+}w - m_{I^-}w}{m_I w}\right)^2 |I| m_I^{1/4}w\; m_I^{1/4}(w^{-1}) \leq C m_J^{1/4}w\; m_J^{1/4}(w^{-1})$$

and therefore, if $w \in A_2^d$ then $\forall J \in D$

$$\frac{1}{|J|}\sum_{I\in D(J)}\left(\frac{m_{I^+}w - m_{I^-}w}{m_I w}\right)^2 |I|\, m_I w\, m_I(w^{-1}) \;\leq\; C\,\|w\|_{A_2^d}\,.$$

The proof of Proposition 3.14 can be found in Section 7.5.

3.4 Extrapolation theorems

In this dissertation we are focused on finding the sharp bounds on the norms of some dyadic operators and classes of dyadic operators (the dyadic paraproducts, the weighted dyadic square functions) on the weighted Lebesgue spaces $L_p(w)$. It turns out that in the case $p = 2$, finding the bounds on the norms on the space $L_2(w)$ is much easier due to the symmetry of $L_2(w)$, while for other values of p it would require a different approach. The essential tool that allows us to obtain the $L_p(w)$ estimates is the celebrated extrapolation theorem of Rubio de Francia, that first appeared in [Ru] in 1984. This beautiful theorem is one of the deepest results in the modern harmonic analysis and perhaps one of the most powerful tools in the study of weighted norm inequalities.

In this section we will discuss the original extrapolation theorem of Rubio de Francia, the sharp version of this theorem from [DrGrPerPet] and some further generalizations of it, in particular, to the cases of two weights. We will also point out a recent interesting result of A. Lerner, which in some cases allows one to obtain sharper bounds than the bounds produced by the sharp extrapolation theorem from [DrGrPerPet].

Let us state Rubio de Francia's extrapolation theorem now.

Theorem 3.15 (Rubio de Francia's extrapolation theorem) *Given an operator T, suppose that for some p_0, $1 \leq p_0 < \infty$ and every $w \in A_{p_0}$ there exists a constant C depending on $\|w\|_{A_{p_0}}$ only, such that*

$$\int_{\mathbb{R}^n} |Tf(x)|^{p_0}\, w(x)dx \ \leq\ C \int_{\mathbb{R}^n} |f(x)|^{p_0} w(x)dx.$$

Then for every p, $1 < p < \infty$ and every $w \in A_p$ there exists a constant depending only on $\|w\|_{A_p}$, such that

$$\int_{\mathbb{R}^n} |Tf(x)|^{p}\, w(x)dx \ \leq\ C \int_{\mathbb{R}^n} |f(x)|^{p} w(x)dx.$$

The original proof of this result, see [Ru], was very complicated. Garcia-Cuerva in [GaRu] presented an alternative proof, which is significantly simpler than the original one. Different proofs can also be found in [Gr] or [Du]. All the proofs mentioned above use Rubio de Francia iteration algorithm (see [Bl]) and require considering cases $1 < p < p_0$ and $p_0 < p < \infty$ separately. In [DrGrPerPet] the authors observed that one can trace dependence of the $L_p(w)$ norms on the $L_{p_0}(w)$ norm of the operator T in Rubio de Francia's extrapolation theorem using the proof from [Gr] and the sharp estimate on the $L_p(w)$ norm of Hardy-Littlewood's maximal function for $1 < p < \infty$:

$$\|M\|_{L_p(w)\to L_p(w)} \;\leq\; C(p)\,\|w\|_{A_p}^{\frac{p'}{p}}\,, \tag{3.17}$$

where p' is the dual exponent of p $\left(\frac{1}{p'}+\frac{1}{p}=1\right)$ and constant $C(p)$ only depends on p and the dimension. Estimate (3.17) for the centered maximal function was obtained by Buckley, see [Buc1]. But, since the centered maximal function, the uncentered maximal function and the dyadic maximal function are comparable modulo dimensional constant, (3.17) holds for the uncentered and the dyadic maximal functions as well.

Let us state the sharp version of Rubio de Francia extrapolation theorem from [DrGrPerPet].

Theorem 3.16 (The sharp extrapolation theorem) *Given an operator T, suppose that there is p_0, $1 \leq p_0 < \infty$, such that the operator T is bounded on $L_{p_0}(w)$ for all weights $w \in A_{p_0}$. Then the operator T is bounded on $L_p(w)$ for all $1 < p < \infty$ and weights $w \in A_p$. More precisely, suppose for each $B > 1$ there is a constant $N_{p_0}(B) > 0$, such that*

$$\|T\|_{L_{p_0}(w)\to L_{p_0}(w)} \;\leq\; N_{p_0}(B) \quad \textit{for all} \quad w \in A_{p_0} \quad \textit{with} \quad \|w\|_{A_{p_0}} \leq B, \tag{3.18}$$

then for any $1 < p < \infty$ and $B > 1$ there is $N_p(B) > 0$, such that for all weights $w \in A_p$ with $\|w\|_{A_p} \leq B$

$$\|T\|_{L_p(w)\to L_p(w)} \;\leq\; N_p(B). \tag{3.19}$$

Moreover,

$$N_p(B) \;\leq\; \begin{cases} 2^{\frac{1}{p_0}} N_{p_0}\left(2C(p')^{\frac{p-p_0}{p-1}} B\right) & \text{if} \quad p > p_0 \\ 2^{\frac{p_0-1}{p_0}} N_{p_0}\left(2^{p_0-1}\left(C(p)^{p-p_0} B\right)^{\frac{p_0-1}{p-1}}\right) & \text{if} \quad p < p_0. \end{cases} \tag{3.20}$$

$C(p)$ here is the constant appearing in (3.17), which depends on p and the dimension.

In the case when $p_0 = 2$ and $N_{p_0}(B)$ is a linear function of B, the above sharp extrapolation theorem implies that if

$$\|T\|_{L_2(w)\to L_2(w)} \;\leq\; C\;\|w\|_{A_2} \qquad \forall w \in A_2, \tag{3.21}$$

then for all $1 < p < \infty$ and $w \in A_p$

$$\|T\|_{L_p(w)\to L_p(w)} \;\leq\; C_p\;\|w\|_{A_p}^{\max\left\{1, \frac{1}{p-1}\right\}}, \tag{3.22}$$

with constant C_p that depends on p and the dimension.

Buckley (see [Buc1]) showed using power weights and power functions that for convolution operators with standard kernels the power is at least $\max\left\{1, \frac{1}{p-1}\right\}$. Hence, if one can show a linear bound (3.21) for a convolution operator with standard kernel, then the extrapolated result (inequality (3.22)) is sharp.

The above observation shows that one needs a good initial estimate of the form (3.18) or (3.21) to obtain good results by extrapolation. Linear estimates for $p_0 = 2$ are known for Hilbert, Riesz, Beurling, martingale transforms, square and maximal functions (see [Pet1], [Pet2], [PetVo], [Wit1], [HTrVo] and [Buc1]). From the above observation of Buckley, the sharp extrapolation theorem produces the sharp inequality (3.22) for Hilbert, Riesz, Beurling and martingale transforms. Comparison of the inequality (3.22) with the sharp result for maximal function (3.17) shows, however, that the sharp extrapolation theorem (Theorem 3.16) does not always produce sharp bounds for operators that are not self-adjoint at least for values $p > 2$.

Sometimes, using some additional information about operators that are not self-adjoint and more refined argument, one may be able to produce better $L_p(w)$-bounds than (3.22) or (3.19). In [Le] Lerner showed that for the square function estimate (3.22) can be improved for $p > 2$ at least to

$$\|S\|_{L_p(w)\to L_p(w)} \leq C \|w\|_{A_p}^{\max\left\{\frac{p}{2(p-1)},\frac{1}{p-1}\right\}}.$$

Lerner's extrapolation tool will be discussed later in this section.

Note that in Rubio de Francia extrapolation theorem we make no assumption on the operator T (originally T was required to be sublinear, but it turns out that the theorem holds if T is just well defined on its domain, see [Gr], [CrMPe]), so we can reinterpret Rubio de Francia extrapolation theorem as follows: if the inequality (3.18) holds for ordered pairs of functions of the form (g, f) (in the classical case ordered pairs correspond to $g = Tf$):

$$\|g\|_{L_{p_0}(w)} \leq N_{p_0}(B) \|f\|_{L_{p_0}(w)} \tag{3.23}$$

for all $w \in A_{p_0}$ with $\|w\|_{A_{p_0}} \leq B$, then the inequality (3.19) holds for such pairs as well:

$$\|g\|_{L_p(w)} \leq N_p(B) \|f\|_{L_p(w)}. \tag{3.24}$$

This property was first observed by [CrPe]. An extremely interesting discussion of the modern understanding of the Rubio de Francia extrapolation theory, various extensions of the extrapolation theorem to function spaces, more general than $L_p(w)$ and two weighted extrapolation, can be found in [CrMPe].

Another result closely related to the sharp extrapolation was obtained by Lerner in [Le]. As we discussed above, the sharp extrapolation does not always produce the sharp bound and knowing more than inequality (3.18) may allow one to extrapolate sharper than (3.19) and (3.20). Also, elimination of the operator, as in (3.23) and (3.24), significantly improves the range of possible applications of the theorem.

Theorem 3.17 (Lerner) *Suppose that for two measurable functions f and g*

$$\|g\|_{L_2(w)} \leq C\sqrt{\|v\|_{A_\infty}[w,v]_{A_2}}\,\|f\|_{L_2(v^{-1})} \tag{3.25}$$

holds for all weights v and w, where C is some absolute constant. Then for any $1 < p < \infty$ and any weight $w \in A_p$,

$$\|g\|_{L_p(w)} \leq C\,\|w\|_{A_p}^{\max\left\{1,\frac{p}{2}\right\}\frac{1}{p-1}}\,\|f\|_{L_p(w)}\,, \tag{3.26}$$

where the constant C depends only on p and the dimension.

In the case when $w = v^{-1}$ condition (3.25) can be written as:

$$\begin{aligned}\|g\|_{L_2(w)} &\leq C\sqrt{\|w^{-1}\|_{A_\infty}\|w\|_{A_2}}\,\|f\|_{L_2(w)}\\ &\leq C\,\|w\|_{A_2}\,\|f\|_{L_2(w)}\,,\end{aligned}$$

where the last inequality follows from the fact that $\|w^{-1}\|_{A_\infty} \leq \|w^{-1}\|_{A_2}$ and the symmetry of A_2: $\|w^{-1}\|_{A_2} = \|w\|_{A_2}$ (as discussed in Section 2.2.2). In this case the sharp extrapolation theorem implies that

$$\|g\|_{L_p(w)} \leq C\,\|w\|_{A_p}^{\max\left\{1,\frac{p'}{p}\right\}}\,\|f\|_{L_p(w)}\,.$$

For $1 < p < 2$ it coincides with the inequality obtained from Lerner's theorem, formula (3.26), while for $2 < p < \infty$ Lerner's result is sharper. However, Lerner's result is much harder to apply due to the form of condition (3.25). Lerner showed that condition (3.25) holds for the square function, hence for $p > 2$ the estimate is better, as discussed earlier in this section. It is not known if Lerner's estimate for the square function for $p > 2$ is sharp. We expect the dyadic paraproduct to behave like a square function.

My academic brother Dariusz Panek is working on extensions of the extrapolation theorems. In particular, he extended the sharp extrapolation theorem, as well as Lerner's theorem to the case of the doubling measure σ. Darek Panek also generalized Lerner's theorem to the cases when the condition (3.25) is replaced by

$$\|g\|_{L_2(w)} \;\leq\; C\,\|v\|_{A_\infty}^{\alpha}\,[w,v]_{A_2}^{\beta}\,\|f\|_{L_2(w)}$$

and when the A_∞-characteristic of the weight v is replaced by the A_p-characteristic of v for some $p > 2$:

$$\|g\|_{L_2(w)} \;\leq\; C\sqrt{\|v\|_{A_p(d\sigma)}\,[w,v]_{A_2(d\sigma)}}\,\|f\|_{L_2(v)}\,.$$

This can be applied to the estimates on the norms of the weighted square functions, that we derive in Section 4.

4 Square function

4.1 Introduction

Now we can take a look at the weighted dyadic square function. We define the dyadic square function with respect to the measure σ to be

$$S^d_\sigma f(x) \;:=\; \left(\sum_{I\in D} \left|m^\sigma_I f - m^\sigma_{\tilde{I}} f\right|^2 \chi_I(x)\right)^{1/2},$$

where $\tilde{I}$ is the parent of I.

In [Buc2] Buckley showed that the dyadic square function S^d is bounded in $L_2(w)$ for all A^d_2-weights with an operator bound of the order $\|w\|^{3/2}_{A^d_2}$. Hukovic, Treil and Volberg in [HTrVo] showed that the norm of the dyadic square function depends linearly on the A^d_2-characteristic of the weight w. Using the sharp version of Rubio de Francia extrapolation theorem from [DrGrPerPet] one can conclude that for $w \in A^d_p$, $\left\|S^d\right\|_{L_p(w)\to L_p(w)} \le C\,\|w\|^{\max\left\{\frac{1}{p-1},1\right\}}_{A^d_p}$ and we can not extrapolate from the above A^d_2 bound better than this since the extrapolation theorem is sharp (for some operators it produces the sharp results). As shown in [DrGrPerPet], for $1 < p \le 2$, the rate of growth $\|w\|^{\frac{1}{p-1}}_{A^d_p}$ is sharp. However, Lerner in [Le] was able to improve the exponent of $\|w\|_{A^d_p}$ for $p > 2$ to

$$\left\|S^d\right\|_{L_p(w)\to L_p(w)} \;\le\; C\;\|w\|^{\max\left\{1,\frac{p}{2}\right\}\frac{1}{p-1}}_{A^d_p}. \tag{4.1}$$

He first observed that the norm of the square function is bounded from $L_2(v^{-1})$ to $L_2(w)$ by

$$\left\|S^d\right\|_{L_2(v^{-1})\to L_2(w)} \;\le\; C\sqrt{\|v\|_{A^d_\infty}\,[v,w]_{A^d_2}}$$

and his extrapolation tool allows to extrapolate this kind of two-weight bounds to (4.1).

My academic brother Darek Panek showed that Lerner's extrapolation theorem can be generalized to the $d\sigma$-case and $\|v\|_{A^d_\infty}$ can be replaced by $\|v\|_{A^d_q(d\sigma)}$ for some $q > 2$.

4.2 The two-weight L_2 estimate for square function

To jumpstart the extrapolation, we prove the following theorem:

Theorem 4.1 *Let v and w be weights, $v \in A^d_q(d\sigma)$ for some $q > 1$ and the couple $(v, w) \in A^d_2(d\sigma)$. Then the dyadic square function S^d_σ is bounded from $L_2(v^{-1}d\sigma)$ to $L_2(wd\sigma)$, moreover for all $f \in L_2(v^{-1}d\sigma)$*

$$\left\|S^d_\sigma f\right\|^2_{L_2(wd\sigma)} \le C\frac{2q-1}{q}(D^d(d\sigma))^{q+1}\|v\|_{A^d_q(d\sigma)}[v,w]_{A^d_2(d\sigma)}\|f\|^2_{L_2(v^{-1}d\sigma)}$$

holds with some numerical constant $C > 0$.

Proof. First note that $\left\|S^d_\sigma f\right\|^2_{L_2(d\mu)}$ can be written as:

$$\left\|S^d_\sigma f\right\|^2_{L_2(d\mu)} = \sum_{I \in D}\left|m^\sigma_I f - m^\sigma_{\tilde{I}} f\right|^2 \mu(I) = \sum_{\tilde{I} \in D}(\alpha_{\tilde{I}^+} + \alpha_{\tilde{I}^-}),$$

where $\alpha_I = \left|m_I^\sigma f - m_{\tilde I}^\sigma f\right|^2 \mu(I)$. Let I^* be the sibling of I, both I and I^* are kids of $\tilde I$. Then

$$\begin{aligned}
\left|m_I^\sigma f - m_{\tilde I}^\sigma f\right| &= \left|\frac{1}{\sigma(I)}\int_I f d\sigma - \frac{1}{\sigma(\tilde I)}\int_{\tilde I} f d\sigma\right| = \frac{\sigma(\tilde I)\int_I f d\sigma - \sigma(I)\int_{\tilde I} f d\sigma}{\sigma(I)\sigma(\tilde I)} \\
&= \frac{\left|(\sigma(I)+\sigma(I^*))\int_I f d\sigma - \sigma(I)\left(\int_I f d\sigma + \int_{I^*} f d\sigma\right)\right|}{\sigma(I)\sigma(\tilde I)} \\
&= \frac{\left|\sigma(I)\int_I f d\sigma + \sigma(I^*)\int_I f d\sigma - \sigma(I)\int_I f d\sigma - \sigma(I)\int_{I^*} f d\sigma\right|}{\sigma(I)\sigma(\tilde I)} \\
&= \frac{1}{\sigma(\tilde I)} \frac{\left|\sigma(I^*)\int_I f d\sigma - \sigma(I)\int_{I^*} f d\sigma\right|}{\sigma(I)} \\
&= \frac{1}{\sigma(\tilde I)}\sqrt{\frac{\sigma(I^*)}{\sigma(I)}}\left|\sqrt{\frac{\sigma(I^*)}{\sigma(I)}}\int_I f d\sigma - \sqrt{\frac{\sigma(I)}{\sigma(I^*)}}\int_{I^*} f d\sigma\right| \\
&= \frac{1}{\sigma(\tilde I)}\sqrt{\frac{\sigma(I^*)}{\sigma(I)}}\left|\langle f; h_{\tilde I}^{d\sigma}\rangle_{d\sigma}\right|.
\end{aligned}$$

Hence,

$$\begin{aligned}
\alpha_{\tilde I^+} &= \frac{\mu(\tilde I^+)}{\sigma(\tilde I)^2}\frac{\sigma(\tilde I^-)}{\sigma(\tilde I^+)}\left|\langle f; h_{\tilde I}^{d\sigma}\rangle_{d\sigma}\right|^2, \\
\alpha_{\tilde I^-} &= \frac{\mu(\tilde I^-)}{\sigma(\tilde I)^2}\frac{\sigma(\tilde I^+)}{\sigma(\tilde I^-)}\left|\langle f; h_{\tilde I}^{d\sigma}\rangle_{d\sigma}\right|^2
\end{aligned}$$

and

$$\left\|S_\sigma^d f\right\|_{L_2(d\mu)}^2 = \sum_{I\in D}\left[\frac{\mu(I^+)\sigma(I^-)}{\sigma(I^+)} + \frac{\mu(I^-)\sigma(I^+)}{\sigma(I^-)}\right]\frac{1}{\sigma(I)^2}\left|\langle f; h_I^{d\sigma}\rangle_{d\sigma}\right|^2.$$

Note, that in the case when $d\mu = w d\sigma$, $\mu(I) = \int_I w d\sigma = \sigma(I)\, m_I^\sigma w$, we can write:

$$\left\|S_\sigma^d f\right\|_{L_2(d\mu)}^2 = \sum_{I\in D}\left(\frac{\sigma(I^-)m_{I^+}^\sigma w + \sigma(I^+)m_{I^-}^\sigma w}{\sigma(I)}\right)\frac{\left|\langle f; h_I^{d\sigma}\rangle_{d\sigma}\right|^2}{\sigma(I)}.$$

We want to bound the norm $\left\|S_\sigma^d\right\|_{L_2(v^{-1}d\sigma)\to L_2(wd\sigma)}$, hence we are looking at the inequalities of the following type:

$$\forall f\in L_2(v^{-1}d\sigma):\quad \sum_{I\in D}\left(\frac{\sigma(I^-)m_{I^+}^\sigma w+\sigma(I^+)m_{I^-}^\sigma w}{\sigma(I)}\right)\frac{\left|\left\langle f;h_I^{d\sigma}\right\rangle_{d\sigma}\right|^2}{\sigma(I)}\ \le\ A_1\left\|f\right\|_{L_2(v^{-1}d\sigma)}^2$$

(then we can claim that $\left\|S_\sigma^d\right\|_{L_2(v^{-1}d\sigma)\to L_2(wd\sigma)}\le\sqrt{A_1}$) or, equivalently,

$$\forall f\in L_2(d\sigma):\quad \sum_{I\in D}\left(\frac{\sigma(I^-)m_{I^+}^\sigma w+\sigma(I^+)m_{I^-}^\sigma w}{\sigma(I)}\right)\frac{\left|\left\langle fv^{1/2};h_I^{d\sigma}\right\rangle_{d\sigma}\right|^2}{\sigma(I)}\ \le\ A_1\left\|f\right\|_{L_2(d\sigma)}^2.$$

Let us denote expression $\frac{\sigma(I^-)m_{I^+}^\sigma w+\sigma(I^+)m_{I^-}^\sigma w}{\sigma(I)^2}$ by $B_{I,\sigma}^w$:

$$B_{I,\sigma}^w\ =\ \frac{\sigma(I^-)m_{I^+}^\sigma w+\sigma(I^+)m_{I^-}^\sigma w}{\sigma(I)^2}. \tag{4.2}$$

In order to handle the scalar product $\left|\left\langle fv^{1/2};h_I^{d\sigma}\right\rangle_{d\sigma}\right|^2$, we switch to the weighted Haar system in $L_2(vd\sigma)$, $\left\{h_I^{vd\sigma}\right\}_{I\in D}$, a system of functions orthogonal in $L_2(vd\sigma)$ with norms $\left\|h_I^{vd\sigma}\right\|_{L_2(vd\sigma)}=\sigma(I)m_I^\sigma v$, related to the $\left\{h_I^{d\sigma}\right\}_{I\in D}$ via

$$h_I^{d\sigma}\ =\ \frac{\sqrt{m_{I^+}^\sigma v\ m_{I^-}^\sigma v}}{m_I^\sigma v}h_I^{vd\sigma}+\frac{\sqrt{\sigma(I^+)\sigma(I^-)}}{\sigma(I)}\frac{m_{I^+}^\sigma v-m_{I^-}^\sigma v}{m_I^\sigma v}\chi_I,$$

(see (2.3) with $d\mu=vd\sigma$). Then expanding the square yields three summonds,

$$\begin{aligned}\left|\left\langle fv^{1/2};h_I^{d\sigma}\right\rangle_{d\sigma}\right|^2\ &=\ \left|\left\langle fv^{1/2};\frac{\sqrt{m_{I^+}^\sigma v\ m_{I^-}^\sigma v}}{m_I^\sigma v}h_I^{vd\sigma}\right\rangle_{d\sigma}\right.\\ &\qquad\left.+\left\langle fv^{1/2};\frac{\sqrt{\sigma(I^+)\sigma(I^-)}}{\sigma(I)}\frac{m_{I^+}^\sigma v-m_{I^-}^\sigma v}{m_I^\sigma v}\chi_I\right\rangle_{d\sigma}\right|^2\\ &=\ I_1+I_2+I_3,\end{aligned}$$

where

$$
\begin{aligned}
I_1 &= \frac{m^\sigma_{I^+}v\; m^\sigma_{I^-}v}{\left(m^\sigma_I v\right)^2}\left\langle fv^{1/2}; h^{vd\sigma}_I\right\rangle^2_{d\sigma},\\
I_2 &= 2\frac{\sqrt{m^\sigma_{I^+}v\, m^\sigma_{I^-}v}}{m^\sigma_I v}\left\langle fv^{1/2}; h^{vd\sigma}_I\right\rangle_{d\sigma}\frac{\sqrt{\sigma(I^+)\sigma(I^-)}}{\sigma(I)}\frac{m^\sigma_{I^+}v - m^\sigma_{I^-}v}{m^\sigma_I v}\left\langle fv^{1/2}; \chi_I\right\rangle_{d\sigma}\\
&= 2\frac{\sqrt{m^\sigma_{I^+}v\; m^\sigma_{I^-}v}}{m^\sigma_I v}\sqrt{\sigma(I^+)\sigma(I^-)}\frac{m^\sigma_{I^+}v - m^\sigma_{I^-}v}{m^\sigma_I v}\left\langle fv^{-1/2}; h^{vd\sigma}_I\right\rangle_{vd\sigma} m^\sigma_I(fv^{1/2}),\\
I_3 &= \frac{\sigma(I^+)\sigma(I^-)}{\sigma(I)^2}\frac{\left(m^\sigma_{I^+}v - m^\sigma_{I^-}v\right)^2}{\left(m^\sigma_I v\right)^2}\left|\left\langle fv^{1/2}; \chi_I\right\rangle_{d\sigma}\right|^2\\
&= \sigma(I^+)\sigma(I^-)\frac{\left(m^\sigma_{I^+}v - m^\sigma_{I^-}v\right)^2}{\left(m^\sigma_I v\right)^2}\left(m^\sigma_I(fv^{1/2})\right)^2.
\end{aligned}
$$

Thus we can decompose $\left\|S^d_\sigma f\right\|^2_{L_2(w)}$ into three sums:

$$
\begin{aligned}
\left\|S^d_\sigma f\right\|^2_{L_2(w)} &= \sum_{I\in D} B^w_{I,\sigma}\left|\left\langle fv^{1/2}; h^{d\sigma}_I\right\rangle_{d\sigma}\right|^2\\
\left(\textstyle\sum_1 :=\right) &= \sum_{I\in D} B^w_{I,\sigma}\frac{m^\sigma_{I^+}v\; m^\sigma_{I^-}v}{\left(m^\sigma_I v\right)^2}\left\langle fv^{-1/2}; h^{vd\sigma}_I\right\rangle^2_{vd\sigma}\\
\left(\textstyle\sum_2 :=\right) &\quad + 2\sum_{I\in D} B^w_{I,\sigma}\frac{\sqrt{m^\sigma_{I^+}v\; m^\sigma_{I^-}v}}{m^\sigma_I v}\sqrt{\sigma(I^+)\sigma(I^-)}\\
&\qquad \times\frac{m^\sigma_{I^+}v - m^\sigma_{I^-}v}{m^\sigma_I v}\left\langle fv^{-1/2}; h^{vd\sigma}_I\right\rangle_{vd\sigma} m^\sigma_I(fv^{1/2})\\
\left(\textstyle\sum_3 :=\right) &\quad + \sum_{I\in D} B^w_{I,\sigma}\sigma(I^+)\sigma(I^-)\frac{\left(m^\sigma_{I^+}v - m^\sigma_{I^-}v\right)^2}{\left(m^\sigma_I v\right)^2}\left(m^\sigma_I(fv^{1/2})\right)^2,
\end{aligned}
$$

where $B^w_{I,\sigma}$ is defined in (4.2).

Note that by Cauchy-Schwarz $\sum_2 \le \left(\sum_1\right)^{1/2}\left(\sum_3\right)^{1/2}$ and hence it is enough to bound $\sum_1$ and $\sum_3$. Let us take a look at $\sum_1$ first.

$$
\begin{aligned}
\sum\nolimits_1 &= \sum_{I\in D}\frac{\left(\sigma(I^-)m^\sigma_{I^+}w + \sigma(I^+)m^\sigma_{I^-}w\right)}{\sigma(I)^2}\frac{m^\sigma_{I^+}v\; m^\sigma_{I^-}v}{\left(m^\sigma_I v\right)^2}\left\langle fv^{-1/2}; h^{vd\sigma}_I\right\rangle^2_{vd\sigma}\\
&= \sum_{I\in D}\frac{\sigma(I^-)m^\sigma_{I^-}v\; m^\sigma_{I^+}w\; m^\sigma_{I^+}v + \sigma(I^+)m^\sigma_{I^+}v\; m^\sigma_{I^-}w\; m^\sigma_{I^-}v}{\sigma(I)^2\;\left(m^\sigma_I v\right)^2}\left\langle fv^{-1/2}; h^{vd\sigma}_I\right\rangle^2_{vd\sigma}.
\end{aligned}
$$

Note that by assumption $(v, w) \in A_2^d(d\sigma)$, i.e. $[v, w]_{A_2^d(d\sigma)} := \sup_{I\in D} m_I v \, m_I w < \infty$, so for $\sum_1$ we can write an inequality

$$\sum\nolimits_1 \; \leq \; [v, w]_{A_2^d(d\sigma)} \sum_{I\in D} \frac{\sigma(I^-)m_{I^-}^\sigma v + \sigma(I^+)m_{I^+}^\sigma v}{(\sigma(I)m_I^\sigma v)^2} \left\langle fv^{-1/2}; h_I^{vd\sigma} \right\rangle_{vd\sigma}^2$$

and using the fact that $\sigma(I^-)m_{I^-}^\sigma v + \sigma(I^+)m_{I^+}^\sigma v = \sigma(I)m_I^\sigma v$, we have

$$\sum\nolimits_1 \; \leq \; [v, w]_{A_2^d(d\sigma)} \sum_{I\in D} \frac{\left\langle fv^{-1/2}; h_I^{vd\sigma} \right\rangle_{vd\sigma}^2}{\sigma(I)m_I^\sigma v}.$$

Note that, as discussed in Section 2.2.4, we can write Bessel's inequality for the orthogonal system of functions $\left\{h_I^{vd\sigma}\right\}_{I\in D}$ with norms $\left\|h_I^{vd\sigma}\right\|_{L_2(vd\sigma)} = \sqrt{\sigma(I)m_I^\sigma v}$ in $L_2(vd\sigma)$ as

$$\forall g \in L_2(vd\sigma) \qquad \sum_{I\in D} \frac{\left\langle g; h_I^{vd\sigma} \right\rangle_{vd\sigma}^2}{\sigma(I)m_I^\sigma v} \; \leq \; \|g\|_{L_2(vd\sigma)}^2 .$$

And hence $\sum_1$ is bounded from above in the following way:

$$\sum\nolimits_1 \; \leq \; [v, w]_{A_2^d(d\sigma)} \left\|fv^{-1/2}\right\|_{L_2(vd\sigma)}^2 \; = \; [v, w]_{A_2^d(d\sigma)} \|f\|_{L_2(d\sigma)}^2 .$$

Now we want to show that for

$$A_1 \; = \; C\frac{2q-1}{q}(D(d\sigma))^{q+1} \|v\|_{A_q^d(d\sigma)} (v, w)_{A_2^d(d\sigma)}$$

we have:

$$\begin{aligned}
\sum\nolimits_3 \; &= \; \sum_{I\in D} B_{I,\sigma}^w \, \sigma(I^+)\sigma(I^-)\frac{\left(m_{I^+}^\sigma v - m_{I^-}^\sigma v\right)^2}{(m_I^\sigma v)^2} \left(m_I^\sigma(fv^{1/2})\right)^2 \\
&\leq \; A_1 \|f\|_{L_2(d\sigma)}^2 .
\end{aligned}$$

We localize the last inequality in the following way: $\forall J \in D$

$$\frac{1}{\sigma(J)} \sum_{I \in D(J)} B^w_{I,\sigma}\, \sigma(I^+)\sigma(I^-) \frac{\left(m^\sigma_{I^+}v - m^\sigma_{I^-}v\right)^2}{\left(m^\sigma_I v\right)^2} \left(m^\sigma_I (f v^{1/2})\right)^2 \leq A_1 m^\sigma_J f^2$$

and use the $(d\sigma)$-version of the Sawyer's Estimate (Theorem 3.3).

So, it is enough for us to prove the following inequality for all dyadic intervals $J \in D$

$$\frac{1}{\sigma(J)} \sum_{I \in D(J)} \left(\sigma(I^-) m^\sigma_{I^+} w + \sigma(I^+) m^\sigma_{I^-} w\right) \left(m^\sigma_{I^+}v - m^\sigma_{I^-}v\right)^2 \frac{\sigma(I^+)\sigma(I^-)}{\sigma(I)^2} \;\leq\; A_1 m^\sigma_J v. \tag{4.3}$$

Let us look at the left-hand side of (4.3).

$$\begin{aligned}
&\frac{1}{\sigma(J)} \sum_{I \in D(J)} \frac{m^\sigma_I v \left(\sigma(I^-)^2 \int_{I^+} w d\sigma + \sigma(I^+)^2 \int_{I^-} w d\sigma\right)}{\sigma(I)^3} \frac{\left(m^\sigma_{I^+}v - m^\sigma_{I^-}v\right)^2}{m^\sigma_I v} \sigma(I) \\
\leq\; & \frac{1}{\sigma(J)} \sum_{I \in D(J)} m^\sigma_I v\; m^\sigma_I w \left[\frac{\sigma(I^+)^2 + \sigma(I^-)^2}{\sigma(I)^2}\right] \frac{\left(m^\sigma_{I^+}v - m^\sigma_{I^-}v\right)^2}{m^\sigma_I v} \sigma(I) \\
\leq\; & [v,w]_{A^d_2(d\sigma)} \frac{1}{\sigma(J)} \sum_{I \in D(J)} \frac{\left(m^\sigma_{I^+}v - m^\sigma_{I^-}v\right)^2}{m^\sigma_I v} \sigma(I).
\end{aligned}$$

Note here that $\frac{1}{2} \leq \frac{\sigma(I^+)^2 + \sigma(I^-)^2}{\sigma(I)^2} \leq 1$, but we might lose sharpness in the dependence on the doubling constant of measure σ, $D^d(d\sigma)$. By the $A^d_p(d\sigma)$-Weight Lemma (Proposition 3.12), applied with $p = q$, (4.3) holds with constant

$$A_1 \;=\; \frac{4(2q-1)}{q} (D^d(d\sigma))^{q+1} \|v\|_{A^d_q(d\sigma)} [v,w]_{A^d_2(d\sigma)},$$

hence

$$\sum\nolimits_3 \;\leq\; C \frac{2q-1}{q} (D^d(d\sigma))^{q+1} \|v\|_{A^d_q(d\sigma)} [v,w]_{A^d_2(d\sigma)} \|f\|^2_{L_2(d\sigma)}.$$

Which completes the proof of Theorem 4.1. ∎

5 Haar multipliers

5.1 Introduction

This section closely follows [Per2]. Using a similar approach, we generalize results of Pereyra, but lose sharpness in one of the cases considered in [Per2].

The Haar multipliers are operators of the form

$$T_w^t f(x) \;=\; \sum_{I\in D} \left(\frac{w(x)}{m_I w}\right)^t \langle f; h_I\rangle h_I(x),$$

where t is any real number, $-\infty < t < \infty$, f is a locally integrable function on $\mathbb{R}$ and w is a weight.

Haar multiplier operators do not have continuous analogues. However, see [PetPo], the averages over random dyadic grids of the Martingale transforms (which are Haar multipliers with constant symbol) give back the Hilbert transform. This averaging process allowed Petermichl to obtain in [Pet1] the sharp linear bound on the norm of the Hilbert transform on the weighted Lebesgue spaces.

In [KaPer] Katz and Pereyra studied boundedness properties of the Haar multiplier operators and proved (in the case when $w \in A_\infty^d$) that the Haar multiplier operator T_w^t is bounded in $L_p(w)$ if and only if w satisfies the C_s^d condition for $s = tp$, which we introduced in Section 2.2.3. The results of Nazarov, Treil and Volberg for martingale transform from [NTrVo2] inspired Pereyra to approach the problem of finding the sharp bounds on the norms of the Haar multiplier operators on the weighted Lebesgue spaces. In [Per2] sharp bounds on the norm of the Haar multipliers can be found for cases $t = 1$, $1/2$ and $-1/2$.

In this section we present bounds for T_w^t in $L_p(\mathbb{R})$ in terms of the corresponding C_{tp}-characteristic of the weight w for all $t \in \mathbb{R}$.

5.2 L_2-bounds for Haar multipliers

The following theorem generalizes the results of Pereyra to the arbitrary values of the parameter t, but does not recover the sharp bound for $t = 1$.

Theorem 5.1 *Let t be a real number, $-\infty < t < \infty$. Let w be a weight in C^d_{2t}, such that $w^{2t} \in A^d_p$ for some $p > 1$ and that satisfies the C^d_{2t} condition with constant $\|w\|_{C^d_{2t}}$. Then the Haar multiplier operator T^t_w is bounded on L_2. Moreover*

$$\left\|T^t_w\right\|_{L_2 \to L_2} \;\le\; C_p \, \|w\|^{1/2}_{C^d_{2t}} \left\|w^{2t}\right\|^{1/2}_{A^d_p}$$

holds with some positive constant $C_p > 0$ that only depends on p.

Proof. First let us compute the formal adjoint $(T^t_w)^*$:

$$\begin{aligned}
\left\langle T^t_w f; g\right\rangle &= \left\langle \sum_{I\in D} \left(\frac{w(x)}{m_I w}\right)^t \langle f; h_I\rangle h_I; g \right\rangle \\
&= \sum_{I\in D} \langle f; h_I\rangle \left\langle g \left(\frac{w(x)}{m_I w}\right)^t ; h_I \right\rangle \\
&= \left\langle f; \sum_{I\in D} \frac{1}{(m_I w)^t} \left\langle g w^t; h_I\right\rangle h_I \right\rangle \\
&= \left\langle f; (T^t_w)^* g\right\rangle .
\end{aligned}$$

So, $(T^t_w)^* f(x) = \sum_{I\in D} \frac{\langle f w^t; h_I\rangle}{(m_I w)^t} h_I(x)$. And by Plancherel

$$\left\|T^t_w\right\|_{L_2\to L_2} \;=\; \left\|(T^t_w)^*\right\|_{L_2\to L_2} \;=\; \left(\sup_{f\in L_2,\, \|f\|_{L_2}=1} \sum_{I\in D} \frac{\langle f w^t; h_I\rangle^2}{(m_I w)^{2t}} \right)^{1/2} . \tag{5.1}$$

So, in order to find a bound on $\|T^t_w\|_{L_2\to L_2}$, we will establish the following inequality:

$$\sum_{I\in D} \frac{\left|\langle f w^t; h_I\rangle\right|^2}{(m_I w)^{2t}} \;\le\; A \, \|f\|^2_{L_2} \qquad \forall f \in L_2, \tag{5.2}$$

from which we can deduce that $\|T^t_w\|_{L_2\to L_2} \le A$.

We will first use the system of weighted Haar functions we discussed in Section 2.2.4, see (2.4),

$$H_I^{(w^{2t})} := h_I\sqrt{|I|} - A_I^{(w^{2t})}\chi_I, \quad A_I^{(w^{2t})} := \frac{m_{I^+}(w^{2t}) - m_{I^-}(w^{2t})}{2m_I(w^{2t})}.$$

Note that $\left\{H_I^{(w^{2t})}\right\}_{I\in D}$ are orthogonal in $L_2(w^{2t})$, with norms $\left\|H_I^{(w^{2t})}\right\|_{L_2(w^{2t})} \leq \sqrt{|I|\, m_I(w^{2t})}$. So, according to (2.5) we have:

$$\sum_{I\in D} \frac{\left|\left\langle fw^t; H_I^{(w^{2t})}\right\rangle\right|^2}{|I|\, m_I(w^{2t})} \leq \|f\|_{L_2}^2 .$$

We can decompose the sum (5.2) into

$$\begin{aligned}
\sum_{I\in D} \frac{|\langle fw^t; h_I\rangle|^2}{(m_I w)^{2t}} &= \\
\left(\sum\nolimits_1 :=\right) &= \sum_{I\in D} \frac{\left|\left\langle fw^t; H_I^{(w^{2t})}\right\rangle\right|^2}{|I|\,(m_I w)^{2t}} \\
\left(\sum\nolimits_2 :=\right) &\quad + 2\sum_{I\in D} \frac{\left|\left\langle fw^t; H_I^{(w^{2t})}\right\rangle\right|}{(m_I w)^{2t}} A_I^{(w^{2t})} m_I(fw^t) \\
\left(\sum\nolimits_3 :=\right) &\quad + \sum_{I\in D} \left(A_I^{(w^{2t})}\right)^2 m_I^2(fw^t)\,|I|\,\frac{1}{(m_I w)^{2t}}.
\end{aligned}$$

First note that by Cauchy-Schwarz $\sum_2 \leq \left(\sum_1\right)^{1/2}\left(\sum_3\right)^{1/2}$, so it is enough to bound the first and the third sums. We apply Bessel's inequality to the first sum:

$$\begin{aligned}
\sum\nolimits_1 &= \sum_{I\in D} \frac{\left|\left\langle fw^t; H_I^{(w^{2t})}\right\rangle\right|^2}{|I|\,(m_I w)^{2t}} \\
&= \sum_{I\in D} \frac{m_I(w^{2t})}{(m_I w)^{2t}} \frac{\left|\left\langle fw^t; H_I^{(w^{2t})}\right\rangle\right|^2}{|I|\, m_I(w^{2t})} \leq \|w\|_{C_{2t}} \|f\|_{L_2}^2 .
\end{aligned}$$

Let us find a bound on the third sum now. First, by the weighted Carleson Lemma (Theorem 3.2), it is enough to prove the following inequality

$$\forall J \in D \quad \frac{1}{|J|} \sum_{I \in D(J)} \left(A_I^{(w^{2t})} \right)^2 m_I^2(w^{2t}) \, |I| \, \frac{1}{(m_I w)^{2t}} \;\leq\; A \, m_J(w^{2t}).$$

We have

$$\begin{aligned} \sum\nolimits_4 \;&:=\; \frac{1}{|J|} \sum_{I \in D(J)} \left(A_I^{(w^{2t})} \right)^2 m_I^2(w^{2t}) \, |I| \, \frac{1}{(m_I w)^{2t}} \\ &=\; \frac{1}{|J|} \sum_{I \in D(J)} \frac{m_I(w^{2t})}{(m_I w)^{2t}} \frac{\left(m_{I^+}(w^{2t}) - m_{I^-}(w^{2t})\right)^2}{4 m_I(w^{2t})} \, |I| \\ &\leq\; \|w\|_{C_{2t}} \frac{1}{|J|} \sum_{I \in D(J)} \frac{\left(m_{I^+}(w^{2t}) - m_{I^-}(w^{2t})\right)^2}{4 m_I(w^{2t})} \, |I| \end{aligned}$$

which, by the A_p^d-Weight Lemma (Proposition 3.12), is bounded by

$$C_p \, \|w\|_{C_{2p}^d} \left\| w^{2t} \right\|_{A_p^d} m_J(w^{2t})$$

for all $p > 1$ whenever $w^{2t} \in A_p^d$, which completes the proof of Theorem 5.1. ■

Thus, combining the estimates for $\sum_1$, $\sum_2$ and $\sum_3$, we conclude that for all $p > 1$ and $-\infty < t < \infty$, the following inequality

$$\left\| T_w^t \right\|_{L_2 \to L_2} \;\leq\; C(p) \;\; \|w\|_{C_{2t}^d}^{1/2} \left\| w^{2t} \right\|_{A_p^d}^{1/2}, \tag{5.3}$$

holds whenever $w^{2t} \in A_p^d$ with a positive constant $C(p)$ that only depends on p.

Let us analyze this result now.

1. For $t = -1/2$, since $\|w\|_{C_{-1}^d} = \|w\|_{A_2^d}$ (as discussed in Section 2.2.3), (5.3) becomes

$$\left\| T_w^{-1/2} \right\|_{L_2 \to L_2} \;\leq\; C(p) \, \|w\|_{C_{-1}^d}^{1/2} \left\| w^{-1} \right\|_{A_p^d}^{1/2} \;=\; C(p) \, \|w\|_{A_2^d}^{1/2} \left\| w^{-1} \right\|_{A_p^d}^{1/2}$$

and if we chose $p = 2$, since $w^{-1} \in A_2^d$ whenever $w \in A_2^d$, $\|w\|_{A_2^d} = \|w^{-1}\|_{A_2^d}$ (as discussed in Section 2.2.2)

$$\left\|T_w^{-1/2}\right\|_{L_2 \to L_2} \le C \; \|w\|_{A_2^d},$$

which recovers the bound from [Per2] that is known to be sharp.

2. For $t = 1/2$ (5.3) becomes

$$\left\|T_w^{1/2}\right\|_{L_2 \to L_2} \le C(p) \, \|w\|_{C_1^d}^{1/2} \|w\|_{A_p^d}^{1/2} = C(p) \, \|w\|_{A_p^d}^{1/2} \tag{5.4}$$

and when $p = 2$

$$\left\|T_w^{1/2}\right\|_{L_2 \to L_2} \le C \, \|w\|_{A_2^d}^{1/2},$$

which is known to be a sharp bound from [Per2] as well.

3. For $t = 1$ we can write:

$$\|T_w\|_{L_2 \to L_2} \le C(p) \, \|w\|_{C_2^d}^{1/2} \left\|w^2\right\|_{A_p^d}^{1/2},$$

where $\|w\|_{C_2^d}$ is such that $\forall I \in D \; m_I w^2 \le \|w\|_{C_2^d} (m_I w)^2$. This coincides with the dyadic Reverse Hölder RH_2^d-condition:

$$\forall I \in D \qquad m_I(w^2) \le \|w\|_{RH_2^d}^2 (m_I w)^2.$$

So, we get

$$\|T_w\|_{L_2 \to L_2} \le C(p) \, \|w\|_{RH_2^d} \left\|w^2\right\|_{A_p^d}^{1/2}.$$

By Theorem 2.4

$$w^2 \in A_\infty^d \quad \Rightarrow \quad w \in RH_2^d,$$

in particular, by Lemma 2.5

$$w^2 \in A_p^d \quad \Leftrightarrow \quad w \in RH_2^d \bigcap A^d_{\frac{p+1}{2}}$$

and we know that

$$\left\|w^2\right\|_{A_p^d}^{1/2} \;\leq\; \|w\|_{RH_2^d} \, \|w\|_{A^d_{\frac{p+1}{2}}} \;,$$

hence

$$\|T_w\|_{L_2 \to L_2} \;\leq\; C(p)\, \|w\|_{RH_2^d}^2 \, \|w\|_{A^d_{\frac{p+1}{2}}} \;,$$

which is slightly worse than the sharp bound from [Per2] $\|T_w\|_{L_2 \to L_2} \;\leq\; C\, \|w\|_{RH_2^d}^2$, with constant C that possibly depends linearly on the dyadic doubling constant of the weight w, since, as discussed in Section 2.2.2, the dyadic doubling constant of w is bounded from above by the $A^d_{\frac{p+1}{2}}$-characteristic of the weight w.

6 The dyadic paraproduct

6.1 Introduction

The dyadic paraproduct is defined as

$$\pi_b f := \sum_{I \in D} m_I f \; b_I \; h_I,$$

where $b_I := \langle b; h_I \rangle$ and b and f are locally integrable functions on the real line.

In order for the dyadic paraproduct to be bounded on L_p we need b to be in BMO^d, see, for example, [Per1].

Paraproducts first appeared in the work of Bony on nonlinear partial differential equations (see [Bo]) and since then have played a central role in harmonic analysis. The general paraproduct holds the key to the class of singular integral operators with standard kernels. By the celebrated T(1) theorem of David and Journé [JoDa] a Calderón-Zygmund singular integral operator T can be written as $T = L + \pi_{b_1} + \pi^*_{b_2}$ where L is an almost translation invariant (convolution) operator, ($L1 = 0 = L^*1$), b_1 is the value of T at 1 and $b_2 = T^*(1)$. A dyadic version of this theorem can be found in [Per1]. So, if one is looking for a bound on the norms of some reasonably large class of singular integral operators it is natural to start with the paraproduct and with its simple dyadic version.

We are going to prove that the bound on the norm of the dyadic paraproduct on the weighted spaces $L_2(w)$ depends linearly on the A_2^d characteristic of the weight w. Therefore, in order to prove the linear bounds on the norms of singular integral operators, defined on $L_2(w)$ and with standard kernels in the dyadic case one has to concentrate on the operator L.

A typical representative of such operators is the Hilbert transform defined by

$$Hf(x) = P.V. \frac{1}{\pi} \int \frac{f(x)}{x-y} dy.$$

Helson & Szegö in [HeSz] gave a necessary and sufficient condition for a weight w so that H maps $L^2(w)$ into itself continuously.

In 1973, Hunt, Muckenhoupt and Wheeden (see [HuMW]) presented a new condition, where for the first time the A_p-condition for weights appeared as a necessary and sufficient condition for the boundedness of the Hilbert transform in $L_p(w)$.

A year later, in [CoFe], Coifman and Fefferman extended this result to a larger class of operators.

The question that has been asked is:

How is the norm of a Calderón-Zygmund singular integral operator on the weighted Lebesgue spaces $L_p(w)$ related to the Muckenhoupt (A_p) characteristic of the weight w, $\|w\|_{A_p}$. More precisely, what we want to find is a function $\varphi(x)$, sharp in terms of its growth, such that

$$\|Tf\|_{L_p(w)} \leq C\varphi\left(\|w\|_{A_p}\right) \|f\|_{L_p(w)}. \tag{6.1}$$

This kind of estimates for different singular integral operators is used often in the theory of partial differential equations, see [FeKPi], [AISa], [PetVo], [BaJa] and [DrPetVo]. Some partial answers have been given to this question.

For the Hilbert transform on $\mathbb{R}$, Buckley showed $\varphi(x) = x^2$ in [Buc1]. Petermichl and Pott in [PetPo] improved the exponent of $\varphi(x)$ from 2 to $\frac{3}{2}$ and in 2006 Petermichl obtained (6.1) with $\varphi(x) = x$ for the Hilbert transform, see [Pet1].

Later in [Pet2] Petermichl used similar ideas to show linear bounds for the norm of the Riesz transforms in $\mathbb{R}^n$. It was also shown that the norm of the dyadic Martingale transform on the weighted space $L_2(w)$ depends linearly on $\|w\|_{A_2^d}$, see [Wit1].

Boundedness of the paraproduct operator on the weighted $L_p(w)$ has been known for a long time, a direct proof of it can be found, for example, in [KaPer].

The best known bound on the norm of the dyadic paraproduct until recently, was

$$\|\pi_b\|_{L_2(w)\to L_2(w)} \leq C\phi(\|w\|_{A_2})\,\|b\|_{BMO^d} \tag{6.2}$$

with $\phi(x) = x^2$, a statement can be found in [DrGrPerPet].

First we were able to improve the above result from $\phi(x) = x^2$ to $\phi(x) = x^{3/2}$ without making any significant changes to the structure of the proof. Then using a suggestion of F. Nazarov we tried the duality approach which allowed us to recover 3/2 in multiple ways and using the version of the bilinear embedding theorem (Lemma 3.8), similar to the one from [Pet1], we were able to improve to $\phi(x) = x(1+\log^{1/2} x)$. Using the sharp version of the bilinear embedding theorem from [NTrVo2] (Theorem 3.7) slightly improved the power of the logarithm in the bound ($\phi(x) = x(1+\log^{1/4} x)$). And finally, the theorem presented in this dissertation shows the linear bound and in fact the proof can rely on either one of the bilinear embedding theorems.

6.2 First approach, the $\varphi(x) = x^{3/2}$ result

In this section we will take a first look at the paraproduct, and using the approach from [DrGrPerPet] we will prove that the norm of the dyadic paraproduct on the weighted space $L_2(w)$ depends on the A_2^d-characteristic of the weight w at most as $\|w\|_{A_2^d}^{3/2}$, i.e. (6.2) holds with $\varphi(x) = x^{3/2}$:

$$\|\pi_b\|_{L_2(w)\to L_2(w)} \;\leq\; C\;\|w\|_{A_2^d}^{3/2}\,\|b\|_{BMO^d}\,. \tag{6.3}$$

First we approached the problem the same way as the authors did in [KaPer].

Observe that

$$\forall f \in L_2 \quad \|Tf\|_{L_2(w)} \leq C \|f\|_{L_2(w)} \iff \forall g \in L_2 \quad \left\|w^{1/2} T w^{-1/2} g\right\|_{L_2} \leq C \|g\|_{L_2}.$$

Then we can rewrite (6.3) as:

$$\forall f \in L_2 \quad \left\|(w^{1/2}\pi_b w^{-1/2})(f)\right\|_{L_2} \ \leq \ C \|b\|_{BMO^d} \|w\|_{A_2^d}^{3/2} \|f\|_{L_2}, \tag{6.4}$$

with $b \in BMO^d$, where

$$\pi_b f = \sum_{I \in D} m_I f \langle b; h_I \rangle h_I.$$

Let M_w^α be a constant Haar multiplier, sometimes called dyadic multiplication, defined by

$$M_w^\alpha \ := \ \sum_{I \in D} \left(\frac{1}{m_I w}\right)^\alpha \langle f; h_I \rangle h_I.$$

It is easy to see that $M_w^\alpha M_w^{-\alpha} = Id$, where Id stands for the identity operator. Note the the Haar multiplier operators, encountered in Section 5, can be written as a composition of a dyadic multiplication followed by multiplication by the weight w to the appropriate power, more precisely,

$$T_w^\alpha f = w^\alpha M_w^\alpha f = \sum_{I \in D} \left(\frac{w}{m_I w}\right)^\alpha \langle f; h_I \rangle h_I.$$

We can rewrite the left-hand side of (6.4) as follows:

$$\begin{aligned}
\left\|(w^{1/2}\pi_b w^{-1/2})(f)\right\|_{L_2} &= \left\|(w^{1/2} M_w^{1/2} M_w^{-1/2} \pi_b w^{-1/2})(f)\right\|_{L_2} \\
&\leq \left\|w^{1/2} M_w^{1/2}\right\|_{(L_2 \to L_2)} \left\|(M_w^{-1/2}\pi_b w^{-1/2})(f)\right\|_{L_2} \\
&= \left\|T_w^{1/2}\right\|_{(L_2 \to L_2)} \left\|(M_w^{-1/2}\pi_b w^{-1/2})(f)\right\|_{L_2} \\
&\leq C \|w\|_{A_2^d}^{1/2} \left\|M_w^{-1/2}\pi_b w^{-1/2} f\right\|_{L_2}.
\end{aligned}$$

The last inequality ($\left\|T_w^{1/2}\right\|_{(L_2\to L_2)} \le C\,\|w\|_{A_2^d}^{1/2}$) follows from (5.4), which we discussed earlier in Section 5.

Note that when we split the norm $\left\|\left(w^{1/2}M_w^{1/2}M_w^{-1/2}\pi_b w^{-1/2}\right)(f)\right\|_{L_2}$ into the product of the operator norm $\left\|w^{1/2}M_w^{1/2}\right\|_{L_2\to L_2}$ and the L_2 norm $\left\|\left(M_w^{-1/2}\pi_b w^{-1/2}\right)(f)\right\|_{L_2}$ we may lose sharpness.

Our goal now is to bound $\left\|\left(M_w^{-1/2}\pi_b w^{-1/2}\right)(f)\right\|_{L_2}$ by

$$\left\|\left(M_w^{-1/2}\pi_b w^{-1/2}\right)(f)\right\|_{L_2} \le C\,\|b\|_{BMO^d}\,\|w\|_{A_2^d}\,\|f\|_{L_2}, \tag{6.5}$$

from which we can conclude that (6.4) holds. Let us take a look at the left hand side of (6.5) now.

$$\begin{aligned}
\left\|(M_w^{-1/2}\pi_b w^{-1/2})(f)\right\|_{L_2}^2 &= \left\|M_w^{-1/2}\left[\pi_b[w^{-1/2}f]\right]\right\|_{L_2}^2\\
&= \left\|M_w^{-1/2}\left[\sum_{I\in D} m_I(w^{-1/2}f)\langle b;h_I\rangle h_I\right]\right\|_{L_2}^2\\
&= \left\|\sum_{J\in D}\left(\frac{1}{m_J w}\right)^{-\frac{1}{2}}\left\langle\sum_{I\in D} m_I(w^{-1/2}f)\langle b;h_I\rangle h_I;h_J\right\rangle h_J\right\|_{L_2}^2\\
&= \left\|\sum_{J\in D}(m_J w)^{1/2}\sum_{I\in D} m_I(w^{-1/2}f)\langle b;h_I\rangle\langle h_I;h_J\rangle h_J\right\|_{L_2}^2\\
&= \left\|\sum_{I\in D}(m_I w)^{1/2} m_I(w^{-1/2}f)\langle b;h_I\rangle h_I\right\|_{L_2}^2\\
&= \sum_{I\in D}\left[(m_I w)^{1/2} m_I(w^{-1/2}f)\langle b;h_I\rangle\right]^2\\
&= \sum_{I\in D} m_I w\, m_I^2(w^{-1/2}f)\langle b;h_I\rangle^2,
\end{aligned}$$

i.e., taking advantage of the orthonormality of the Haar system, we can write:

$$\left\|(M_w^{-1/2}\pi_b w^{-1/2})(f)\right\|_{L_2}^2 = \sum_{I\in D} m_I w\, m_I^2(w^{-1/2}f) b_I^2.$$

So, the square of (6.5) can be written as:

$$\sum_{I\in D} m_I w \, m_I^2(w^{-1/2}f) b_I^2 \;\le\; C\,\|b\|_{BMO^d}^2 \,\|w\|_{A_2^d}^2 \,\|f\|_{L_2}^2 \,, \tag{6.6}$$

and this is the inequality we will now prove.

Since w is an almost everywhere positive function, $m_I w$ is positive and we can apply the weighted Carleson embedding theorem (Theorem 3.2) to (6.6) with $\alpha_I = m_I w \, b_I^2$ and $v = w^{-1}$. Proving (6.6) becomes equivalent to proving:

$$\forall J \in D \quad \frac{1}{|J|} \sum_{I\in D(J)} m_I w \, m_I^2(w^{-1}) b_I^2 \;\le\; C\,\|b\|_{BMO^d}^2 \,\|w\|_{A_2^d}^2 \, m_J(w^{-1}). \tag{6.7}$$

By the definition of A_2^d, $m_I w m_I(w^{-1}) \le \|w\|_{A_2^d}$ for all dyadic intervals $I \in D$, hence

$$\frac{1}{|J|} \sum_{I\in D(J)} m_I w \, m_I^2(w^{-1}) b_I^2 \;\le \|w\|_{A_2^d} \frac{1}{|J|} \sum_{I\in D(J)} m_I(w^{-1}) b_I^2$$

and (6.7) will follow from:

$$\forall J \in D \quad \frac{1}{|J|} \sum_{I\in D(J)} m_I(w^{-1}) b_I^2 \;\le\; C\,\|b\|_{BMO^d}^2 \,\|w\|_{A_2^d} \, m_J(w^{-1}). \tag{6.8}$$

We already mentioned in Section 2.1.4 that since $b \in BMO^d$, $\{b_I^2\}_{I\in D}$ is a Carleson sequence with constant $\|b\|_{BMO^d}^2$, i.e.

$$\forall J \in D \quad \frac{1}{|J|} \sum_{I\in D(J)} b_I^2 \le \|b\|_{BMO^d}^2 \tag{6.9}$$

Applying Proposition 3.4 with $\lambda_I = b_I^2$ and $v = w^{-1}$ (note here that $\|w^{-1}\|_{A_2} = \|w\|_{A_2}$), we get the following bound:

$$\forall J \in D \quad \frac{1}{|J|} \sum_{I\in D(J)} m_I(w^{-1}) b_I^2 \;\le\; C\,\|b\|_{BMO^d}^2 \,\|w\|_{A_2} \, m_J(w^{-1}),$$

i.e. (6.8) holds, so we have shown that

$$\|\pi_b\|_{L_2(w)\to L_2(w)} \;\le\; C\;\|b\|_{BMO^d}\,\|w\|_{A_2^d}^{3/2}.$$

6.3 Duality approach, the $\varphi(x) = x\left(1+\log^{1/2} x\right)$ result

In order to improve the estimate on the norm of the dyadic paraproduct we obtained in Section 6.2, we need to refine our argument. In this section we will use a duality approach, which allows to recover the bound (6.4), $\varphi(x) = x^{3/2}$, of the previous section and improve it to $\varphi(x) = x\left(1+\log^{1/2} x\right)$, i.e. at the end of this section we will see that

$$\|\pi_b\|_{L_2(w)\to L_2(w)} \;\le\; C\;\|b\|_{BMO^d}\,\|w\|_{A_2^d}\left(1+\log^{1/2}\|w\|_{A_2^d}\right). \tag{6.10}$$

In personal communication F. Nazarov suggested us to try bounding the norm of the dyadic paraproduct by duality, i.e. show that for all $f \in L_2(w)$ and $g \in L_2(w^{-1})$

$$\langle \pi_b f; g\rangle \le C\varphi(\|w\|_{A_2})\,\|b\|_{BMO^d}\,\|f\|_{L_2(w)}\,\|g\|_{L_2(w^{-1})} \tag{6.11}$$

Or, alternatively, for all $f, g \in L_2$

$$\left\langle \pi_b(fw^{-1/2}); gw^{1/2}\right\rangle \le C\varphi(\|w\|_{A_2})\,\|b\|_{BMO^d}\,\|f\|_{L_2}\,\|g\|_{L_2}.$$

In this section we will show that (6.11) holds with $\varphi(x) = x\left(1+\log^{1/2} x\right)$.

We can write $\left\langle \pi_b(fw^{-1/2}); gw^{1/2}\right\rangle$ as

$$\begin{aligned}\left\langle \pi_b(fw^{-1/2}); gw^{1/2}\right\rangle &= \left\langle \sum_{I\in D} m_I(fw^{-1/2})b_I h_I; gw^{1/2}\right\rangle \\ &= \sum_{I\in D} m_I(fw^{-1/2})b_I\langle gw^{1/2}; h_I\rangle \;=:\; \sum\nolimits_1.\end{aligned}$$

Thus, we need to show that

$$\sum\nolimits_1 \;=\; \sum_{I\in D} m_I(fw^{-1/2}) b_I \langle gw^{1/2}; h_I\rangle \;\leq\; C\,\|b\|_{BMO^d}\,\varphi(\|w\|_{A_2^d})\,\|f\|_{L_2}\,\|g\|_{L_2}$$

with $\varphi(x) = x\left(1+\log^{1/2} x\right)$. Let us first show an extremely easy alternative way to get $\varphi(x) = x^{\frac{3}{2}}$. By Cauchy-Schwarz inequality

$$\sum_{I\in D} m_I(fw^{-1/2}) b_I\langle gw^{1/2}; h_I\rangle \leq \left(\sum_{I\in D} m_I w\; m_I^2(fw^{-1/2})\; b_I^2\right)^{\frac{1}{2}} \left(\sum_{I\in D} \frac{\langle gw^{1/2}; h_I\rangle^2}{m_I w}\right)^{\frac{1}{2}}. \tag{6.12}$$

From (6.6) of the previous Section 6.2, we know that the first sum is bounded by

$$\sum_{I\in D} m_I^2(fw^{-1/2}) m_I w b_I^2 \leq C\,\|w\|^2_{A_2^d}\,\|b\|^2_{BMO^d}\,\|f\|^2_{L_2}. \tag{6.13}$$

Note that

$$\sum_{I\in D} \frac{\langle gw^{1/2}; h_I\rangle^2}{m_I w} \;=\; \left\|\left(T_w^{1/2}\right)^* g\right\|^2_{L_2}.$$

By (5.4) we know that $\left\|\left(T_w^{1/2}\right)^*\right\|_{L_2\to L_2} = \left\|T_w^{1/2}\right\|_{L_2\to L_2} \leq C\,\|w\|^{1/2}_{A_2^d}$ and hence

$$\sum_{I\in D} \frac{\langle gw^{1/2}; h_I\rangle^2}{m_I w} \;\leq\; C\;\;\|w\|_{A_2^d}\,\|g\|^2_{L_2}. \tag{6.14}$$

Combining (6.13) and (6.14) in (6.12), we get

$$\sum_{I\in D} m_I\left(fw^{-1/2}\right) b_I \left\langle gw^{1/2}; h_I\right\rangle \;\leq\; C\;\;\|w\|^{3/2}_{A_2^d}\,\|b\|_{BMO^d}\,\|f\|_{L_2}\,\|g\|_{L_2}.$$

So, (6.11) holds with $\varphi(x) = x^{3/2}$, i.e.

$$\|\pi_b\|_{L_2(w)\to L_2(w)} \;\leq\; C\;\;\|b\|_{BMO^d}\,\|w\|^{3/2}_{A_2^d},$$

which recovers the result of Section 6.2.

Note that (6.12) is not the only possible partition of the sum we are interested in. Clearly, there are a few other different ways to split it up and get the same exponent $\varphi(x) = x^{\frac{3}{2}}$.

In order to improve the estimate we will need another tool. We are going to use the weighted Haar basis, introduced in Section 2.2.4.

We will decompose the sum from (6.12) using a weighted Haar system of functions, introduced in Section 2.2.4. Let H_I^w be defined in the following way:

$$H_I^w \;:=\; h_I\sqrt{|I|} - A_I^w \chi_I \;\; and \;\; A_I^w \;:=\; \frac{m_{I^+}w - m_{I^-}w}{2m_I w}.$$

Then the system of functions $\{w^{1/2}H_I^w\}$ is orthogonal in L_2 with norms bounded from above by $\left\|w^{1/2}H_I^w\right\|_{L_2} \leq \sqrt{|I|m_I w}$. So, by Bessel's inequality we have:

$$\forall g \in L_2 \qquad \sum_{I\in D} \frac{1}{|I|m_I w}\left\langle g; w^{1/2}H_I^w\right\rangle^2 \;\leq\; \|g\|_{L_2}^2 . \tag{6.15}$$

Then we can break $\sum_1$ into two sums:

$$\begin{aligned}
\sum\nolimits_1 \;&=\; \sum_{I\in D} m_I\left(fw^{-1/2}\right) b_I \left\langle gw^{1/2}; h_I\right\rangle \\
&=\; \sum_{I\in D} m_I\left(fw^{-1/2}\right) b_I \frac{1}{\sqrt{|I|}}\left\langle g; w^{1/2}H_I^w\right\rangle \\
&\quad + \sum_{I\in D} m_I\left(fw^{-1/2}\right) b_I \frac{1}{\sqrt{|I|}}\left\langle gw^{1/2}; A_I^w\chi_I\right\rangle \\
&=: \; \sum\nolimits_2 + \sum\nolimits_3 .
\end{aligned}$$

Our goal now is to bound $\sum_2$ and $\sum_3$. When we are splitting the sum into two pieces and bounding them separately, we can lose sharpness because we are not taking into account cancelations between them. However, here it is likely that this step is sharp because of the following heuristic reasons. The sum $\sum_2$ is close to the "weighted" version of a paraproduct over a weighted space $L_2(w)$, which behaves similar to the unweighted situation, while $\sum_3$ takes into account the difference between the norm of the paraproduct on weighted and unweighted L_2. In the simplest case $w = const$, $\|w\|_{A_2^d} = 1$, $\sum_1 = \sum_2$ and $\sum_3 = 0$ we recover classical results. Note also, that for weights with small A_2^d-characteristics $\sum_2$ will be dominating and $\sum_3$ will be close to 0, while for $\|w\|_{A_2^d}$ large $\sum_3$ becomes more important.

The linear bound on $\sum_2$ is very straight-forward. We decompose $\sum_2$ into the product of two sums using Cauchy-Schwarz:

$$\begin{aligned}\sum\nolimits_2 &= \sum_{I\in D} m_I\left(fw^{-1/2}\right) b_I \frac{1}{\sqrt{|I|}}\left\langle g; w^{1/2}H_I^w\right\rangle \\ &\leq \left(\sum_{I\in D} m_I^2\left(fw^{-1/2}\right) b_I^2 m_I w\right)^{1/2}\left(\sum_{I\in D}\frac{1}{|I|m_I w}\left\langle g; w^{1/2}H_I^w\right\rangle^2\right)^{1/2}.\end{aligned}$$

By (6.6)

$$\sum_{I\in D} m_I^2(fw^{-1/2})m_I w b_I^2 \leq C\|w\|_{A_2^d}^2\|b\|_{BMO^d}^2\|f\|_2^2.$$

By Bessel's inequality, as discussed in Section 2.2.4, we have

$$\sum_{I\in D}\frac{1}{|I|m_I w}\left\langle g; w^{1/2}H_I^w\right\rangle^2 \leq \|g\|_{L_2}^2.$$

Therefore we get the desired bound for $\sum_2$,

$$\sum\nolimits_2 = \sum_{I\in D} m_I\left(fw^{-1/2}\right) b_I \frac{1}{\sqrt{|I|}}\left\langle g; w^{1/2}H_I^w\right\rangle \leq C\|w\|_{A_2^d}\|b\|_{BMO^d}\|f\|_2\|g\|_2.$$

Note here that $\sum_1 = \sum_2 + \sum_3$ and $\sum_2$ admits a bound linear with respect to the A_2^d-characteristic of the weight w. The logarithmic factor in (6.11) comes from $\sum_3$. In the next two sections we will show how to improve the bound on $\sum_3$ to $\varphi(x) = x\left(1+\log^{1/4} x\right)$ (in Section 6.4) and to the linear dependence (in Section 6.5).

Now we need to bound $\sum_3$.

$$\sum\nolimits_3 = \sum_{I\in D} b_I A_I^w \sqrt{|I|}\, m_I \left(fw^{-1/2}\right)\ m_I \left(gw^{1/2}\right).$$

By Corollary 3.8, applied to $v = w^{-1}$ and $\alpha_I = \frac{|b_I A_I^w|}{\sqrt{|I|}}$, it is enough to establish the following three inequalities:

$$\sum\nolimits_4 := \frac{1}{|J|} \sum_{I\in D(J)} |b_I A_I^w| \sqrt{|I|}\, m_I w\ m_I w^{-1}\ \le\ C\, \|b\|_{BMO^d}\, \varphi\left(\|w\|_{A_2^d}\right), \tag{6.16}$$

$$\sum\nolimits_5 := \frac{1}{|J|} \sum_{I\in D(J)} |b_I A_I^w| \sqrt{|I|}\, m_I w\ \le\ C\, \|b\|_{BMO^d}\, \varphi\left(\|w\|_{A_2^d}\right) m_J w, \tag{6.17}$$

$$\sum\nolimits_6 := \frac{1}{|J|} \sum_{I\in D(J)} |b_I A_I^w| \sqrt{|I|}\, m_I w^{-1}\ \le\ C\, \|b\|_{BMO^d}\, \varphi\left(\|w\|_{A_2^d}\right) m_J(w^{-1}), \tag{6.18}$$

with $\varphi(x) = x \log^{1/2} x$.

For the sequence $\{(A_I^w)^2 |I|\}_{I\in D}$ note that

$$\frac{1}{|J|} \sum_{I\in D(J)} (A_I^w)^2 |I| = \frac{1}{|J|} \sum_{I\in D(J)} \left(\frac{m_{I^+} w - m_{I^-} w}{2 m_I w} \right)^2 |I|.$$

By the sharp version of Fefferman-Kenig-Pipher inequality (Theorem 3.9, inequality (3.14)), since $\log \|w\|_{A_\infty^d} \le \log \|w\|_{A_2^d}$, we have that

$$\frac{1}{|J|} \sum_{I\in D} (A_I^w)^2\, |I|\ =\ \frac{1}{|J|} \sum_{I\in D} \left(\frac{m_{I^+} w - m_{I^-} w}{2 m_I w} \right)^2 |I|\ \le\ C\ \log\left(\|w\|_{A_2^d}\right).$$

Together with (6.9) this shows that the sequence $\{|b_I A_I^w| \sqrt{|I|}\}_{I\in D}$ is a Carleson sequence with Carleson constant $C\, \|b\|_{BMO^d} \log^{1/2} \|w\|_{A_2^d}$, i.e. for all dyadic intervals

$J \in D$

$$\frac{1}{|J|}\sum_{I\in D(J)} |b_I A_I^w|\sqrt{|I|} \;\leq\; \left(\frac{1}{|J|}\sum_{I\in D(J)} b_I^2\right)^{\frac{1}{2}} \left(\frac{1}{|J|}\sum_{I\in D(J)} |A_I^w|^2|I|\right)^{\frac{1}{2}}$$
$$\leq\; C\; \|b\|_{BMO^d} \log^{\frac{1}{2}} \|w\|_{A_2^d}\,.$$

Therefore,

$$\sum\nolimits_4 := \frac{1}{|J|}\sum_{I\in D(J)} |b_I A_I^w|\sqrt{|I|}\; m_I w\; m_I w^{-1} \leq C\,\|b\|_{BMO^d}\,\|w\|_{A_2^d}\log^{\frac{1}{2}}\|w\|_{A_2^d}\,. \quad (6.19)$$

Both $\sum_5$ and $\sum_6$ can also be bounded using inequality (3.5) in Proposition 3.4 with $v = w$ and $v = w^{-1}$ respectively (we are using the fact that $\|w^{-1}\|_{A_2^d} = \|w\|_{A_2^d}$ here):

$$\sum\nolimits_5 := \frac{1}{|J|}\sum_{I\in D(J)} |b_I A_I^w|\sqrt{|I|}\; m_I w \leq C\,\|b\|_{BMO^d}\,\|w\|_{A_2^d}\log^{\frac{1}{2}}\|w\|_{A_2^d}\, m_J w, \quad (6.20)$$

$$\sum\nolimits_6 := \frac{1}{|J|}\sum_{I\in D(J)} |b_I A_I^w|\sqrt{|I|}\; m_I(w^{-1}) \leq C\,\|b\|_{BMO^d}\,\|w\|_{A_2^d}\log^{\frac{1}{2}}\|w\|_{A_2^d}\, m_J(w^{-1}), \quad (6.21)$$

Hence (6.11) holds with $\varphi(x) = x\left(1+\log^{1/2} x\right)$, which implies inequality (6.10), i.e.

$$\|\pi_b\|_{L_2(w)\to L_2(w)} \;\leq\; C\,\|b\|_{BMO^d}\,\|w\|_{A_2^d}\left(1+\log^{1/2}\|w\|_{A_2^d}\right).$$

6.4 Further refinement. The $\varphi(x) = x\left(1+\log^{1/4} x\right)$ result

In the previous section in formula (6.20) we bounded $\sum_5$ by
$C\,\|b\|_{BMO^d}\,\|w\|_{A_2^d}\log^{\frac{1}{2}}\|w\|_{A_2^d}\, m_J w$. We also know that since b_I^2 is a Carleson sequence with Carleson constant $\|b\|_{BMO^d}^2$, by Proposition 3.4, inequality (3.5), we have

$$\forall J \in D \quad \frac{1}{|J|}\sum_{I\in D(J)} b_I^2\; m_I w \;\leq\; C\,\|b\|_{BMO^d}^2\,\|w\|_{A_2^d}\, m_J w.$$

Furthermore, by Wittwer's sharp version of Buckley's inequality (Theorem 3.10) we have

$$\begin{aligned}\forall J \in D \quad \frac{1}{|J|} \sum_{I \in D(J)} (A_I^w)^2 \, |I| \, m_I w \;&=\; \frac{1}{|J|} \sum_{I \in D(J)} \left(\frac{m_{I^+} w - m_{I^-} w}{2 m_I w} \right)^2 |I| \, m_I w \\ &\leq\; C \; \|w\|_{A_2^d} \, m_J w.\end{aligned}$$

So, by Cauchy-Schwarz inequality we have

$$\begin{aligned}\sum\nolimits_5 \;&\leq\; \left(\frac{1}{|J|} \sum_{I \in D(J)} (A_I^w)^2 \, |I| \, m_I w \right)^{1/2} \left(\frac{1}{|J|} \sum_{I \in D(J)} b_I^2 |I| \, m_I w \right)^{1/2} \\ &\leq\; C \; \|b\|_{BMO^d} \, \|w\|_{A_2^d} \, m_J w.\end{aligned}$$

Also, by Proposition 3.13:

$$\forall J \in D \quad \frac{1}{|J|} \sum_{I \in D(J)} \left(\frac{m_{I^+} w - m_{I^-} w}{m_I w} \right)^2 |I| \, m_I(w^{-1}) \;\leq\; C \, \|w\|_{A_2^d} \, m_J(w^{-1}),$$

and by Cauchy-Schwarz, $\sum_6$ can also be bounded by

$$\begin{aligned}\sum\nolimits_6 \;&=\; \frac{1}{|J|} \sum_{I \in D(J)} |b_I A_I^w| \sqrt{|I|} \, m_I(w^{-1}) \\ &\leq\; \left(\frac{1}{|J|} \sum_{I \in D(J)} b_I^2 \, m_I(w^{-1}) \right)^{1/2} \left(\frac{1}{|J|} \sum_{I \in D(J)} (A_I^w)^2 \, |I| \, m_I(w^{-1}) \right)^{1/2} \\ &\leq\; C \; \|b\|_{BMO^d} \, \|w\|_{A_2^d} \, m_J(w^{-1}).\end{aligned}$$

Corollary 3.8, however, does not allow us to improve the bound on

$$\sum\nolimits_3 \;=\; \sum_{I \in D} b_I A_I^w \sqrt{|I|} \, m_I \left(f w^{-1/2} \right) \; m_I \left(g w^{1/2} \right)$$

since we are taking the supremum of the bounds on $\sum_4$, $\sum_5$ and $\sum_6$, and we only know that $\sum_5$ and $\sum_6$ depend on the A_2^d-characteristic of the weight w linearly, while $\sum_4$ is still bounded by $C\left\|b\right\|_{BMO^d}\left\|w\right\|_{A_2^d}\log^{1/2}\left(\left\|w\right\|_{A_2^d}\right)$.

Using the bilinear embedding theorem (Corollary 3.7), we will be able to take into account the linear bounds on $\sum_5$ and $\sum_6$ and decrease the power of the logarithm to 1/4: $\varphi(x) = x\left(1+\log^{1/4}x\right)$.

By Corollary 3.7 we need to prove the following two inequalities for the double sums:

$$\frac{1}{|J|}\sum_{I\in D(J)}|b_I A_I^w|\sqrt{|I|}m_I w^{-1}\frac{1}{|I|}\sum_{K\in D(I)}|b_K A_K^w|\sqrt{|K|}m_K w\, m_K w^{-1} \ \le\ Q^2 m_J w^{-1} \tag{6.22}$$

and

$$\frac{1}{|J|}\sum_{I\in D(J)}|b_I A_I^w|\sqrt{|I|}m_I w\frac{1}{|I|}\sum_{K\in D(I)}|b_K A_K^w|\sqrt{|K|}m_K w\, m_K w^{-1} \ \le\ Q^2 m_J w \tag{6.23}$$

to deduce that

$$\sum_{I\in D}|b_I A_I^w|\sqrt{|I|}\, m_I(fw^{-1/2})\, m_I(gw^{1/2}) \ \le\ CQ\left\|f\right\|_{L_2}\left\|g\right\|_{L_2}. \tag{6.24}$$

Let us take a close look at (6.22) and (6.23) now. In Section 6.3 (see (6.19)) we showed that $\forall I\in D$, $\forall J\in D$

$$\sum\nolimits_4 \ =\ \frac{1}{|I|}\sum_{K\in D(I)}|b_K A_K^w|\sqrt{|K|}m_K w\, m_K w^{-1} \ \le\ C\left\|w\right\|_{A_2^d}\log^{1/2}\left\|w\right\|_{A_2^d}\left\|b\right\|_{BMO^d} \tag{6.25}$$

and we observed at the beginning of this section that for $\sum_5$ and $\sum_6$ we had linear bounds:

$$\sum\nolimits_5 = \frac{1}{|J|} \sum_{I \in D(J)} |b_I A_I^w| \sqrt{|I|} m_I w \leq C \|w\|_{A_2^d} \|b\|_{BMO^d} \, m_J w \tag{6.26}$$

and

$$\sum\nolimits_6 = \frac{1}{|J|} \sum_{I \in D(J)} |b_I A_I^w| \sqrt{|I|} m_I w^{-1} \leq C \|w\|_{A_2^d} \|b\|_{BMO^d} \, m_J w^{-1}. \tag{6.27}$$

From which we can conclude that (6.24) is true with

$$\begin{aligned} Q &= \sqrt{C \|w\|_{A_2^d}^2 \log^{1/2}\left(\|w\|_{A_2^d}\right) \|b\|_{BMO^d}^2} \\ &= C \|w\|_{A_2^d} \log^{1/4}\left(\|w\|_{A_2^d}\right) \|b\|_{BMO^d} , \end{aligned}$$

which shows that

$$\begin{aligned} \sum\nolimits_3 &= \sum_{I \in D} m_I(f w^{-1/2}) \, |b_I A_I^w| \sqrt{|I|} \, m_I(g w^{1/2}) \\ &\leq C \, \|b\|_{BMO^d} \|w\|_{A_2^d} \log^{1/4}\left(\|w\|_{A_2^d}\right) \|f\|_{L_2} \|g\|_{L_2} , \end{aligned}$$

i.e. we have proved the following inequality for the dyadic paraproduct operator:

$$\|\pi_b\|_{L_2(w) \to L_2(w)} \leq C \, \|b\|_{BMO^d} \|w\|_{A_2^d} \left(1 + \log^{1/4}\left(\|w\|_{A_2^d}\right)\right).$$

6.5 Main result (linear bound for the dyadic paraproduct)

The reader probably noticed that the only obstacle on the way to the linear bound for the norm of the dyadic paraproduct is the logarithmic term in the inequality for

$\sum_4$:

$$\begin{aligned}\sum_4 &= \frac{1}{|J|}\sum_{I\in D(J)} |b_I A_I^w|\,\sqrt{|I|}\, m_I w\, m_I(w^{-1}) \\ &\le C\ \|b\|_{BMO^d}\,\|w\|_{A_2^d}\log^{1/2}\left(\|w\|_{A_2^d}\right).\end{aligned}$$

Proposition 3.14 will let us fix this problem. We shall state the result as a theorem and give a complete proof, most of which will repeat parts of the previous sections.

Theorem 6.1 (Main result) *The norm of dyadic paraproduct on the weighted Lebesgue space $L_2(w)$ is bounded from above by a constant multiple of the product of the A_2^d characteristic of the weight w and the BMO^d norm of b, i.e. there exists a positive constant $C>0$, such that for all $b\in BMO^d$ and $w\in A_2^d$, for all $f\in L_2(w)$ and all $g\in L_2(w^{-1})$*

$$\langle \pi_b f; g\rangle \ \le\ C\ \|w\|_{A_2^d}\ \|b\|_{BMO^d}\ \|f\|_{L_2(w)}\ \|g\|_{L_2(w^{-1})}\,. \tag{6.28}$$

This theorem, together with the sharp version of the Rubio De Francia's extrapolation theorem from [DrGrPerPet], produces $L_p(w)$ bounds of the following type:

Theorem 6.2 *Let $w\in A_p^d$ and $b\in BMO^d$. Then the norm of the dyadic paraproduct π_b on the weighted $L_p(w)$ spaces satisfies the following inequalities*

$$\|\pi_b\|_{L_p(w)\to L_p(w)} \ \le\ C_1(p)\,\|w\|_{A_p^d}\,\|b\|_{BMO^d} \quad \text{when } p\ge 2$$

and

$$\|\pi_b\|_{L_p(w)\to L_p(w)} \ \le\ C_2(p)\,\|w\|_{A_p^d}^{\frac{1}{p-1}}\,\|b\|_{BMO^d} \quad \text{when } p<2,$$

where $C_1(p)$ and $C_2(p)$ are constants that only depend on p.

Proof. [Theorem 6.1] In order to prove Theorem 6.1 it is enough to show that $\forall\ f, g \in L_2$

$$\sum\nolimits_1 := \left\langle \pi_b \left(f w^{-1/2} \right); g w^{1/2} \right\rangle \ \leq \ C \left\| w \right\|_{A_2^d} \left\| b \right\|_{BMO^d} \left\| f \right\|_2 \left\| g \right\|_2 .$$

We are going to decompose this sum using the weighted Haar system of functions (see [CoJS]), introduced in Section 2.2.4. Let H_I^w be defined in the following way:

$$H_I^w \ := \ h_I \sqrt{|I|} - A_I^w \chi_I \ \ and \ \ A_I^w \ := \ \frac{m_{I^+} w - m_{I^-} w}{2 m_I w},$$

then $\{w^{1/2} H_I^w\}$ is orthogonal in L_2 with norms satisfying the inequality $\left\| w^{1/2} H_I^w \right\|_{L_2} \leq \sqrt{|I| m_I w}$. By Bessel's inequality, as discussed in Section 2.2.4, we have:

$$\forall g \in L_2 \qquad \sum_{I \in D} \frac{1}{|I| m_I w} \left\langle g; w^{1/2} H_I^w \right\rangle_{L_2}^2 \ \leq \ \|g\|_{L_2}^2 . \tag{6.29}$$

We can break $\sum_1$ into two sums:

$$\begin{aligned}
\sum\nolimits_1 \ &= \ \sum_{I \in D} m_I \left(f w^{-1/2} \right) b_I \left\langle g w^{1/2}; h_I \right\rangle \\
&= \ \sum_{I \in D} m_I \left(f w^{-1/2} \right) b_I \frac{1}{\sqrt{|I|}} \left\langle g; w^{1/2} H_I^w \right\rangle \\
&\quad + \sum_{I \in D} m_I \left(f w^{-1/2} \right) b_I \frac{1}{\sqrt{|I|}} \left\langle g w^{1/2}; A_I^w \chi_I \right\rangle \\
&=: \ \sum\nolimits_2 + \sum\nolimits_3 .
\end{aligned}$$

We claim that both sums, $\sum_2$ and $\sum_3$, can be bounded with a bound that depends on $\|w\|_{A_2^d}$ at most linearly:

$$\sum_{I \in D} m_I \left(f w^{-1/2} \right) b_I \frac{1}{\sqrt{|I|}} \left\langle g; w^{1/2} H_I^w \right\rangle \ \leq \ C \left\| w \right\|_{A_2^d} \left\| b \right\|_{BMO^d} \left\| f \right\|_{L_2} \left\| g \right\|_{L_2} \tag{6.30}$$

and

$$\sum_{I\in D} m_I\left(fw^{-1/2}\right) b_I A_I^w \sqrt{|I|} m_I\left(gw^{1/2}\right) \leq C \|w\|_{A_2^d} \|b\|_{BMO^d} \|f\|_{L_2} \|g\|_{L_2}. \qquad (6.31)$$

The bound on $\sum_2$ is very straight-forward. We decompose $\sum_2$ into the product of two sums using Cauchy-Schwarz:

$$\begin{aligned} \sum\nolimits_2 &= \sum_{I\in D} m_I\left(fw^{-1/2}\right) b_I \frac{1}{\sqrt{|I|}} \left\langle g; w^{1/2} H_I^w \right\rangle \\ &\leq \left(\sum_{I\in D} m_I^2\left(fw^{-1/2}\right) b_I^2 m_I w\right)^{1/2} \left(\sum_{I\in D} \frac{1}{|I| m_I w} \left\langle g; w^{1/2} H_I^w \right\rangle^2\right)^{1/2}. \end{aligned}$$

By (6.29)

$$\sum_{I\in D} \frac{1}{|I| m_I w} \left\langle g; w^{1/2} H_I^w \right\rangle^2 \ \leq \ \|g\|_{L_2}^2.$$

So, for (6.30) to hold it is enough to show that

$$\sum_{I\in D} m_I^2\left(fw^{-1/2}\right) b_I^2 m_I w \ \leq \ C \|w\|_{A_2^d}^2 \|b\|_{BMO^d}^2 \|f\|_{L_2}^2. \qquad (6.32)$$

By the weighted Carleson embedding theorem (Theorem 3.2), (6.32) holds if and only if

$$\forall J \in D \quad \frac{1}{|J|} \sum_{I\in D(J)} m_I^2 w^{-1}\, m_I w\, b_I^2 \ \leq \ C \|w\|_{A_2^d}^2 \|b\|_{BMO^d}^2 \ m_J w^{-1}.$$

Since $\forall I \in D \quad m_I w\, m_I w^{-1} \leq \|w\|_{A_2^d}$, it is enough to verify that

$$\forall J \in D \quad \frac{1}{|J|} \sum_{I\in D(J)} m_I w^{-1}\, b_I^2 \ \leq \ C \|w\|_{A_2^d} \|b\|_{BMO^d}^2 \ m_J w^{-1}. \qquad (6.33)$$

Inequality (6.33) follows from the fact that $b \in BMO^d$ and hence the sequence $\{b_I^2\}_{I\in D}$ is a Carleson sequence with Carleson constant $\|b\|^2_{BMO^d}$:

$$\forall J \in D \quad \frac{1}{|J|}\sum_{I\in D(J)} b_I^2 \;\le\; \|b\|^2_{BMO^d}\,, \tag{6.34}$$

and Proposition 3.4. Estimate (3.5) applied to $\lambda_I = b_I^2$ and w^{-1} ($w^{-1} \in A_2^d$ and $\|w^{-1}\|_{A_2^d} = \|w\|_{A_2^d}$) provides (6.33), so the inequality (6.30) holds.

Now we need to prove the inequality (6.31). It is a little bit more involved. We want to show that

$$\sum\nolimits_3 = \sum_{I\in D} b_I A_I^w \sqrt{|I|}\; m_I\left(fw^{-1/2}\right)\; m_I\left(gw^{1/2}\right) \le C\,\|w\|_{A_2^d}\,\|b\|_{BMO^d}\,\|f\|_2\,\|g\|_2\,.$$

We are going to use the bilinear embedding theorem (Corollary 3.8) here. Note, however, that all other versions of it would work as well. So, in order to complete the proof it is enough to show that the following three inequalities hold:

$$\forall J \in D \quad \sum\nolimits_4 := \frac{1}{|J|}\sum_{I\in D(J)} |b_I A_I^w|\sqrt{|I|}\; m_I w\; m_I w^{-1} \;\le\; C\,\|b\|_{BMO^d}\,\|w\|_{A_2^d}\,, \tag{6.35}$$

$$\forall J \in D \quad \sum\nolimits_5 := \frac{1}{|J|}\sum_{I\in D(J)} |b_I A_I^w|\sqrt{|I|}\; m_I w \;\le\; C\,\|b\|_{BMO^d}\,\|w\|_{A_2^d}\, m_J w, \tag{6.36}$$

$$\forall J \in D \quad \sum\nolimits_6 := \frac{1}{|J|}\sum_{I\in D(J)} |b_I A_I^w|\sqrt{|I|}\; m_I w^{-1} \;\le\; C\,\|b\|_{BMO^d}\,\|w\|_{A_2^d}\, m_J w^{-1}, \tag{6.37}$$

Proposition 3.14 helps us handle the first sum (6.35). Note that by Cauchy-Schwarz

$$\frac{1}{|J|}\sum_{I\in D(J)} |b_I A_I^w|\sqrt{|I|} m_I w\, m_I w^{-1}$$

$$\le \left(\frac{1}{|J|}\sum_{I\in D(J)} b_I^2 m_I w\, m_I w^{-1}\right)^{\frac{1}{2}} \left(\frac{1}{|J|}\sum_{I\in D(J)} (A_I^w)^2 |I| m_I w\, m_I w^{-1}\right)^{\frac{1}{2}}.$$

It follows from (6.34) that

$$\frac{1}{|J|}\sum_{I\in D(J)} b_I^2\, m_I w\; m_I w^{-1} \;\le\; \|w\|_{A_2^d}\,\frac{1}{|J|}\sum_{I\in D(J)} b_I^2 \;\le\; \|w\|_{A_2^d}\;\|b\|^2_{BMO^d}\,,$$

and by Proposition 3.14

$$\frac{1}{|J|}\sum_{I\in D(J)} (A_I^w)^2|I|\, m_I w\, m_I w^{-1} \;=\frac{1}{|J|}\sum_{I\in D(J)}\left(\frac{m_{I^+}w - m_{I^-}w}{m_I w}\right)^2 |I| m_I w\, m_I w^{-1}$$

$$\le\; C\,\|w\|_{A_2^d}\,.$$

Inequality (6.36) follows by Cauchy-Schwarz as described in the beginning of Section 6.4. The tools required there are: sharp Buckley's inequality (Theorem 3.10) and Proposition 3.4 applied to $\lambda_I = b_I^2$.

Proposition 3.13, inequality (3.16), together with (6.33) allow us to establish inequality (6.37) in a similar way.

This completes the proof of the Theorem 6.1. ■

7 Bellman function proofs

7.1 Introduction

In this section we present the proofs of the propositions and lemmas used to obtain the results for the dyadic square function, the Haar multiplier operators and the dyadic paraproduct described in previous sections.

All the proofs are based on the technique of Bellman functions. The Bellman function method is a relatively new yet powerful tool in modern harmonic analysis. This technique was first introduced by Burkholder in his paper of 1984 [Bur], where he uses it to prove Paley's inequality. Later on, Bellman used Burkholder's ideas to produce an extremely efficient method for stochastic control, and nowadays it reappeared in harmonic analysis with the help of Nazarov, Treil, Volberg, Pereyra, Petermichl, Dragicevic, Vasyunin and others. They've provided an elegant Bellman function proof of known results in harmonic analysis and delivered new interesting results using this powerful method. The Bellman function technique is an extremely powerful tool and is a very natural approach to weighted norm inequalities.

In many cases Bellman functions allow one to get optimal results, see, for example [Va],[VaVo], [Per2], [Pet1], [Pet2]. In some cases Bellman function results can be extended, for example to L_p results, without using any additional tools like extrapolation as in [DrGrPerPet], or to the cases of doubling measures, rather than just Lebesgue measure on real line, see [Per2], and give multi-dimensional analogues, see [DrVo] and [SlVo].

7.2 Proof of Proposition 3.4

We are going to show that for any Carleson sequence $\{\lambda_I\}_{I\in D}$ with constant Q, $\lambda_I \geq 0$ and

$$\forall J \in D \quad \frac{1}{|J|}\sum_{I\in D(J)} \lambda_I \ \leq \ Q,$$

the inequality (3.4) holds for any dyadic interval J, that is,

$$\frac{1}{|J|}\sum_{I\in D(J)} \frac{\lambda_I}{m_I w^{-1}} \ \leq \ 4Q m_J w.$$

Note that inequality (3.5) follows from inequality (3.4).

Lemma 7.1 *Suppose there exists a real valued function of 3 variables $B(x) = B(u,v,l)$, whose domain $\mathfrak{D}$ is given by those $x = (u,v,l) \in \mathbb{R}^3$ such that*

$$u, v > 0, \quad uv \geq 1 \quad \textit{and} \quad 0 \leq l \leq 1,$$

whose range is given by

$$0 \ \leq \ B(x) \ \leq \ u,$$

and such that the following convexity property holds:

$$\forall x,\ x_\pm \in \mathfrak{D} \quad \textit{such that} \quad x - \frac{x_+ + x_-}{2} = (0,0,\alpha)$$

$$B(x) - \frac{B(x_+) + B(x_-)}{2} \ \geq \ \frac{1}{4v}\alpha. \tag{7.1}$$

Then Proposition 3.4 holds.

Proof.

Fix a dyadic interval J. Let $x_J = (u_J, v_J, l_J)$ where $u_J = m_J w$, $v_J = m_J(w^{-1})$ and $l_J = \frac{1}{|J|Q}\sum_{I\in D(J)} \lambda_I$. Clearly for each dyadic J, x_J belongs to the domain $\mathfrak{D}$.

Let $x_{\pm} := x_{J^{\pm}} \in \mathfrak{D}$. By definition,

$$x_J - \frac{x_{J^+} + x_{J^-}}{2} \;=\; (0, 0, \alpha_J),$$

where $\alpha_J := \frac{1}{|J|Q}\lambda_J$. Then, by the convexity condition (7.1) and since $2|J^+| = 2|J^-| = |J|$,

$$\begin{aligned} |J|\, m_J w \;\geq\; |J|\; B(x_J) \;&\geq\; |J|\frac{B(x_{J^+})}{2} + |J|\frac{B(x_{J^-})}{2} + \frac{1}{4Qm_J(w^{-1})}\lambda_J \\ &=\; |J^+|B(x_{J^+}) + |J^-|B(x_{J^-}) + \frac{1}{4Qm_J(w^{-1})}\lambda_J. \end{aligned}$$

We can now use the same lower bound estimate for $|J^+|B(x_{J^+})$ and $|J^-|B(x_{J^-})$. Iterating this procedure and using the assumption that $B \geq 0$ on $\mathfrak{D}$ we get:

$$m_J w \geq \frac{1}{4|J|Q} \sum_{I \in D(J)} \frac{\lambda_I}{m_I(w^{-1})}$$

which implies Proposition 3.4. ■

So, Proposition 3.4 will hold if we can show the existence of the function B of Bellman type, satisfying the conditions of Lemma 7.1.

Lemma 7.2 *The following function*

$$B(u, v, l) := u - \frac{1}{v(1 + l)}$$

is defined on $\mathfrak{D}$*,* $0 \leq B(x) \leq u$ *for all* $x = (u, v, l) \in \mathfrak{D}$ *and satisfies the following differential inequalities on* $\mathfrak{D}$*:*

$$\frac{\partial B}{\partial l}(u, v, l) \;\geq\; \frac{1}{4v} \tag{7.2}$$

and

$$-(du,dv,dl)\,d^2B(u,v,l)\begin{pmatrix} du \\ dv \\ dl \end{pmatrix} \geq 0, \tag{7.3}$$

where $d^2B(u,v,l)$ denotes the Hessian matrix of the function B evaluated at (u,v,l). Moreover, conditions (7.2) and (7.3) imply the convexity condition (7.1).

Proof. Range conditions are obvious. It is nothing but a calculus exercise to check the differential conditions as well:

$$\frac{\partial B}{\partial l}(u,v,l) = \frac{1}{v(1+l)^2} \geq \frac{1}{4v}$$

since $0 \leq l \leq 1$ and

$$\begin{aligned} -(du,dv,dl)\,d^2B(u,v,l)\begin{pmatrix} du \\ dv \\ dl \end{pmatrix} &= (du,dv,dl)\begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{2}{v^3(1+l)} & \frac{1}{v^2(1+l)^2} \\ 0 & \frac{1}{v^2(1+l)^2} & \frac{2}{v(1+l)^3} \end{pmatrix}\begin{pmatrix} du \\ dv \\ dl \end{pmatrix} \\ &\geq 0. \end{aligned}$$

Finally, let us see how differential conditions (7.2) and (7.3) imply the convexity condition (7.1):

$$\begin{aligned} B(x) - \frac{B(x_+)+B(x_-)}{2} &= \left[B(x) - B\left(\frac{x_+ + x_-}{2}\right)\right] \\ &\quad + \left[B\left(\frac{x_+ + x_-}{2}\right) - \frac{B(x_+)+B(x_-)}{2}\right] \\ &= \frac{\partial B}{\partial l}(u,v,l')\alpha - \frac{1}{2}\int_{-1}^{1}(1-|t|)b''(t)dt, \end{aligned} \tag{7.4}$$

where $b(t) := B(x(t))$, $x(t) := \frac{1+t}{2}x_+ + \frac{1-t}{2}x_-$, $-1 \leq t \leq 1$, note that $x(t) \in \mathfrak{D}$ whenever x_+ and x_- do, since $\mathfrak{D}$ is a convex domain and $x(t)$ is a point on the line segment between x_+ and x_-, and l' is a point between l and $\frac{l_+ + l_-}{2}$. The first summand in (7.4) appears as an application of the Mean Value Theorem. The second is an exercise in calculus, which we describe now.

It is easy to see that

$$\begin{aligned}
-\frac{1}{2}\int_{-1}^{1}(1-|t|)b''(t)dt &= -\frac{1}{2}\int_{-1}^{0}(1+t)b''(t)dt-\frac{1}{2}\int_{0}^{1}(1-t)b''(t)dt \\
&= -\frac{1}{2}(1+t)b'(t)\mid_{-1}^{0}+\frac{1}{2}\int_{-1}^{0}b'(t)dt-\frac{1}{2}(1-t)b'(t)\mid_{0}^{1} \\
&\quad -\frac{1}{2}\int_{0}^{1}b'(t)dt \\
&= -\frac{1}{2}b'(0)+\frac{1}{2}b(t)\mid_{-1}^{0}+\frac{1}{2}b'(0)-\frac{1}{2}b(t)\mid_{0}^{1} \\
&= b(0)-\frac{b(1)-b(-1)}{2}.
\end{aligned}$$

Note also that $b(0)=B(x(0))=B(x)$, similarly $b(-1)=B(x_-)$ and $b(1)=B(x_+)$.

The differential inequalities trivially imply that $-b''(t)\geq 0$ and

$$B(x)-\frac{B(x_+)+B(x_-)}{2}=\frac{\partial B}{\partial l}(u,v,l')\alpha-\frac{1}{2}\int_{-1}^{1}(1-|t|)b''(t)dt\geq\frac{1}{4v}\alpha.$$

This completes the proofs of both Lemma 7.2 and Proposition 3.4. ■

7.3 Proof of Proposition 3.5

We are going to show that for any Carleson sequence of nonnegative numbers $\{\lambda_I\}_{I\in D}$ with constant Q, $\lambda_I\geq 0$

$$\forall J\in D\quad \frac{1}{|J|}\sum_{I\in D}\lambda_I\ \leq\ Q$$

inequality (3.6) holds for every dyadic interval J:

$$\frac{1}{|J|}\sum_{I\in D(J)}e^{m_I(\log v)}\lambda_I\ \leq\ 4Q\,m_J v.$$

Note that inequality (3.7) trivially follows from inequality (3.6).

Lemma 7.3 *Suppose there exists a real valued function of 3 variables $B(x) = B(u, v, l)$, whose domain $\mathfrak{D}$ is given by those $x = (u, v, l) \in \mathbb{R}^3$, such that*

$$u > 0, \quad ue^{-v} \geq 1 \quad \textit{and} \quad 0 \leq l \leq 1,$$

whose range is given by

$$0 \;\leq\; B(x) \;\leq\; u,$$

and such that the following convexity property holds:

$\forall x, x_\pm \in \mathfrak{D}$ *such that* $x - \frac{x_+ + x_-}{2} = (0, 0, \alpha)$ *with some* $\alpha \geq 0$

$$B(x) - \frac{B(x_+) + B(x_-)}{2} \;\geq\; \frac{e^v}{4}\alpha. \tag{7.5}$$

Then Proposition 3.5 holds.

Proof. Fix a dyadic interval J. Let $x_J = (u_J, v_J, l_J)$, where $u_J = m_J v$, $v_J = m_J(\log v)$ and $l_J = \frac{1}{Q|J|}\sum_{I \in D(J)} \lambda_I$. Clearly for each dyadic interval I, x_I belongs to the domain $\mathfrak{D}$. Let $x_\pm := x_{J^\pm} \in \mathfrak{D}$.

By the above definition,

$$x_J - \frac{x_{J^+} + x_{J^-}}{2} \;=\; (0, 0, \alpha_J)$$

with $\alpha_J = \frac{1}{Q|J|}\lambda_J \geq 0$.

Then, by convexity condition (7.5)

$$|J| m_J v \;\geq\; |J| B(x_J) \;\geq\; |J|\frac{B(x_{J^+})}{2} + |J|\frac{B(x_{J^-})}{2} + \frac{e^{v_J}}{4Q}\lambda_J.$$

Iterating this procedure on $B(x_{J^+})$ and $B(x_{J^-})$ and using the assumption that B is nonnegative, we get

$$\begin{aligned}
|J|m_J v &\geq |J^+|B(x_{J^+}) + |J^-|B(x_{J^-}) + \frac{e^{v_J}}{4Q}\lambda_J \\
&\geq \frac{1}{4Q}\sum_{I\in D(J)} e^{v_I}\lambda_I \\
&= \frac{1}{4Q}\sum_{I\in D(J)} e^{m_I(\log v)}\lambda_I,
\end{aligned}$$

which implies inequality (3.6) and Proposition 3.5. ■

So, Proposition 3.5 will hold if we can show the existence of the function B of Bellman type, satisfying the conditions of Lemma 7.3.

Lemma 7.4 *The following function*

$$B(u,v,l) \;:=\; u - \frac{e^v}{1+l}$$

is defined on the domain $\mathfrak{D}$, $0 \leq B(x) \leq u$ *for all* $x = (u,v,l) \in \mathfrak{D}$ *and satisfies the following differential inequalities on* $\mathfrak{D}$*:*

$$\forall x = (u,v,l) \in \mathfrak{D} \qquad \frac{\partial B}{\partial l}(x) \;\geq\; \frac{e^v}{4} \tag{7.6}$$

and $\forall x \in \mathfrak{D}$ *and* $\forall (du,dv,dl) \in \mathbb{R}^3$

$$-(du,dv,dl)d^2B(x)\begin{pmatrix} du \\ dv \\ dl \end{pmatrix} \;\geq\; 0, \tag{7.7}$$

i.e. the Hessian matrix of B, d^2B *is non-positive definite on* $\mathfrak{D}$*. Moreover, conditions (7.6) and (7.7) imply convexity condition (7.5).*

Proof. Range conditions are obvious. The positivity of B follows from the positivity of l and inequality $ue^{-v} \geq 1$ ($u \geq e^v$) on the domain $\mathfrak{D}$. The bound from above on B follows from the positivity of the fraction $\frac{e^v}{1+l}$.

It is easy to see that differential conditions hold as well:

$$\frac{\partial B}{\partial l} = \frac{e^v}{(1+l)^2} \geq \frac{e^v}{4}$$

since $0 \leq l \leq 1$, and the Hessian

$$-d^2B(x) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{e^v}{1+l} & -\frac{e^v}{(1+l)^2} \\ 0 & -\frac{e^v}{(1+l)^2} & \frac{2e^v}{(1+l)^3} \end{pmatrix}$$

is non-negative definite.

Finally, let us see how differential conditions (7.6) and (7.7) imply convexity condition (7.5).

We introduce the function $b(t) := B(x_t)$, $-1 \leq t \leq 1$, where $x_t = \frac{1+t}{2}x_{J^+} + \frac{1-t}{2}x_{J^-}$. Note that since domain $\mathfrak{D}$ is convex, $x_t \in \mathfrak{D}$ whenever x_{J^+} and x_{J^-} do. Also note that $x_{-1} = x_{J^-}$, $x_1 = x_{J^+}$ and $x_0 = x_J - (0, 0, \alpha)$, where $\alpha = \frac{1}{|J|Q}\lambda_J$.

If $x_t = (u_t, v_t, l_t)$, then

$$\begin{aligned} b''(t) &= \frac{d^2}{dt^2}B(u_t, v_t, l_t) \\ &= \frac{d}{dt}\left(\frac{\partial B}{\partial u}(u_t, v_t, l_t)\frac{du_t}{dt} + \frac{\partial B}{\partial v}(u_t, v_t, l_t)\frac{dv_t}{dt} + \frac{\partial B}{\partial l}(u_t, v_t, l_t)\frac{dl_t}{dt}\right) \\ &= \left(\frac{du_t}{dt}, \frac{dv_t}{dt}, \frac{dl_t}{dt}\right) d^2B(u_t, v_t, l_t) \begin{pmatrix} \frac{du_t}{dt} \\ \frac{dv_t}{dt} \\ \frac{dl_t}{dt} \end{pmatrix}. \end{aligned}$$

So, since $-d^2B$ is non-negative definite,

$$-b''(t) \;\geq\; 0.$$

On the other hand, see (7.4),

$$\begin{aligned} B(x) - \frac{B(x_+) + B(x_-)}{2} &= \left[B(x) - B\left(\frac{x_+ + x_-}{2}\right)\right] \\ &\quad + \left[B\left(\frac{x_+ + x_-}{2}\right) - \frac{B(x_+) + B(x_-)}{2}\right] \\ &= \frac{\partial B}{\partial l}(u_J, v_J, l'_J)\alpha - \frac{1}{2}\int_{-1}^{1}(1 - |t|)b''(t)dt, \end{aligned}$$

where l'_J is a point between l_J and $\frac{l_{J+} + l_{J-}}{2}$.

Using the fact that $-b''(t)$ is positive and $\frac{\partial B}{\partial l}(u, v, l) \geq \frac{e^v}{4}$,

$$\begin{aligned} B(x) - \frac{B(x_+) + B(x_-)}{2} &= \frac{\partial B}{\partial l}(u_J, v_J, l'_J)\alpha - \frac{1}{2}\int_{-1}^{1}(1 - |t|)b''(t)dt \\ &\geq \frac{\partial B}{\partial l}(u_J, v_J, l'_J) \\ &\geq \frac{e^{v_J}}{4}. \end{aligned}$$

This completes the proof of Lemma 7.4 and Proposition 3.5. ■

7.4 Proof of Proposition 3.13

We are going to show that

$$\forall J \in D \quad \frac{1}{|J|}\sum_{I \in D(J)} \frac{(m_{I^+}w - m_{I^-}w)^2}{m_I^3 w}|I| \;\leq\; C m_J w^{-1} \tag{7.8}$$

holds with some numerical constant C for any weight w, such that expressions $m_I w$ and $m_I(w^{-1})$ are meaningful for all dyadic intervals I (for example, both functions w and w^{-1} are locally integrable).

Lemma 7.5 *Suppose there exists a real-valued function of two variables $B(x) = B(u,v)$, whose domain $\mathfrak{D}$ is given by those $x = (u,v) \in \mathbb{R}^2$ such that*

$$u, v \; > \; 0, \tag{7.9}$$

$$uv \; \geq \; 1, \tag{7.10}$$

whose range is given by

$$0 \; \leq \; B(x) \; \leq \; v$$

and such that the following convexity property holds for all $x, x_\pm \in \mathfrak{D}$:

$$if \quad x = \frac{x_+ + x_-}{2} \quad then \quad B(x) - \frac{B(x_+) + B(x_-)}{2} \; \geq \; C\frac{1}{u^3}(u_+ - u_-)^2 \tag{7.11}$$

with some numerical constant C. Then Proposition 3.13 will be proved (inequality (7.8) holds for all dyadic intervals J).

Proof. Let $u_I := m_I w$, $v_I := m_I w^{-1}$, $v_+ = v_{I^+}$, $v_- = v_{I^-}$ and similarly for $u_\pm$. Then by Hölder's inequality (u,v) and $(u_\pm, v_\pm)$ belong to the $\mathfrak{D}$.

Fix $J \in D$, by the convexity property and range conditions

$$|J| m_J w^{-1} \; \geq \; |J| B(u_J, v_J)$$

$$\geq \; |J^+| B(u_+, v_+) + |J^-| B(u_-, v_-) + C|J| \frac{1}{m_J^3 w}(m_{J^+} w - m_{J^-} w)^2.$$

Iterating this process and using positivity of function B, we get

$$|J| m_J w^{-1} \; \geq \; C \sum_{I \in D(J)} |I| \frac{1}{m_I^3 w}(m_{I^+} w - m_{I^-} w)^2,$$

which completes the proof of Lemma 7.5. ∎

To prove inequality (7.8) and Proposition 3.13 we need to show the existence of the function B of Bellman type satisfying the conditions of Lemma 7.5.

Lemma 7.6 *The following function*

$$B(u,v) \;=\; v - \frac{1}{u}$$

is defined on domain $\mathfrak{D}$, $0 \le B(u,v) \le v$ *for all* $(u,v) \in \mathfrak{D}$ *and satisfies the following differential inequality in* $\mathfrak{D}$*:*

$$-\,(du,dv)\,d^2B(u,v)\begin{pmatrix} du \\ dv \end{pmatrix} \;=\; \frac{2}{u^3}|du|^2.$$

Moreover, it implies the convexity condition (7.11) with some numerical constant C *independent of everything.*

Proof. First note that since $uv \ge 1$ and u and v are both positive in the domain $\mathfrak{D}$, B is well-defined and

$$0 \;\le\; B(u,v) \;=\; \frac{uv-1}{u} \;=\; v - \frac{1}{u} \;\le\; v$$

on $\mathfrak{D}$ and $-\,(du,dv)\,d^2B(u,v)\begin{pmatrix} du \\ dv \end{pmatrix} \;=\; 2u^{-3}|du|^2.$

It is just a calculus exercise to derive condition (7.11):

$$B(u,v) - \frac{B(u_+,v_+) + B(u_-,v_-)}{2} \;\ge\; C\frac{1}{u^3}\,(u_+ - u_-)^2 \tag{7.12}$$

from the differential condition

$$-\,(du,dv)\,d^2B(u,v)\begin{pmatrix} du \\ dv \end{pmatrix} \;=\; 2u^{-3}\,|du|^2\,.$$

For $t \in [-1,1]$ define (u_t, v_t) to be a point on the line segment from (u_-, v_-) to (u_+, v_+):

$$u_t \ := \ \frac{1}{2}(1+t)u_+ + \frac{1}{2}(1-t)u_-$$

and

$$v_t \ := \ \frac{1}{2}(1+t)v_+ + \frac{1}{2}(1-t)v_-$$

Note that since domain $\mathfrak{D}$ is convex, $(u_t, v_t) \in \mathfrak{D}$. Also note that $u_0 = u$, $u_1 = u_+$, $u_{-1} = u_-$ and, similarly, $v_0 = v$, $v_1 = v_+$, $v_{-1} = v_-$;

$$\frac{du_t}{dt} \ = \ \frac{u_+ - u_-}{2} \quad and \quad \frac{dv_t}{dt} \ = \ \frac{v_+ - v_-}{2}.$$

Define $b(t)$ to be $b(t) := B(u_t, v_t)$, then (7.12) is equivalent to proving

$$b(0) - \frac{b(1) + b(-1)}{2} \ \geq \ C\frac{1}{u_0^3}\left(u_1 - u_{-1}\right)^2.$$

It is easy to see that (as observed in Section 7.2)

$$b(0) - \frac{b(1) + b(-1)}{2} \ = \ -\frac{1}{2}\int_{-1}^{1}(1-|t|)b''(t)dt.$$

Note that

$$\begin{aligned}
-b''(t) &= -\frac{d^2}{dt^2}B(u_t, v_t) \\
&= -\frac{d}{dt}\left(\frac{\partial B}{\partial u}(u_t, v_t)\frac{du_t}{dt}(t) + \frac{\partial B}{\partial v}(u_t, v_t)\frac{dv_t}{dt}(t)\right) \\
&= -\left(\frac{\partial^2 B}{\partial u^2}(u_t, v_t)\left(\frac{du_t}{dt}\right)^2 + 2\frac{\partial^2 B}{\partial u \partial v}(u_t, v_t)\left(\frac{du_t}{dt}\frac{dv_t}{dt}\right) + \frac{\partial^2 B}{\partial v^2}(u_t, v_t)\left(\frac{dv_t}{dt}\right)^2\right) \\
&= -\left(\frac{du_t}{dt}, \frac{dv_t}{dt}\right) d^2 B(u_t, v_t) \begin{pmatrix} \frac{du_t}{dt} \\ \frac{dv_t}{dt} \end{pmatrix} \\
&\geq 2u_t^{-3}\left|\frac{du_t}{dt}\right|^2 \\
&= 2u_t^{-3}\left|\frac{u_+ - u_-}{2}\right|^2 \\
&= \frac{1}{2}u_t^{-3}\left|u_+ - u_-\right|^2.
\end{aligned}$$

In order to make a lower bound on $-b''(t)$ independent of t and match inequality (7.12), we need to bound u_t from above by a constant multiple of $u_0 = u$

$$u_t = u_0 + t\frac{u_1 - u_{-1}}{2}.$$

Note that

$$\frac{|u_1 - u_{-1}|}{2} \leq \frac{|u_1| + |u_{-1}|}{2} = \frac{u_1 + u_{-1}}{2} = u_0$$

and hence

$$u_t \leq u_0 + tu_0 = u_0(1+t) \leq 2u_0,$$

so

$$-b''(t) \geq \frac{1}{16}u_0^{-3}\left|u_1 - u_{-1}\right|^2.$$

Since $\frac{1}{2}\int_{-1}^{1}(1-|t|)dt = \frac{1}{2}$, we conclude that

$$\begin{aligned} b(0) - \frac{b(1)+b(-1)}{2} &= -\frac{1}{2}\int_{-1}^{1}(1-|t|)b''(t)dt \\ &\geq \frac{1}{16}u_0^{-3}\left|u_1 - u_{-1}\right|^2 \frac{1}{2}\int_{-1}^{1}(1-|t|)dt \\ &= \frac{1}{32}u_0^{-3}\left|u_1 - u_{-1}\right|^2 . \end{aligned}$$

This completes the proof of Lemma 7.6 and Proposition 3.13. ∎

7.5 Proof of Proposition 3.14

We are going to prove that there is a numerical constant $C > 0$, such that for all dyadic intervals $J \in D$

$$\frac{1}{|J|}\sum_{I\in D(J)}\left(\frac{m_{I^+}w - m_{I^-}w}{m_I w}\right)^2 |I| m_I^{1/4}w\; m_I^{1/4}w^{-1} \;\leq\; C m_J^{1/4}w\; m_J^{1/4}w^{-1}, \tag{7.13}$$

using the Bellman function technique.

Lemma 7.7 *Suppose there exists a real-valued function of two variables* $B(x) = B(u,v)$*, whose domain* $\mathfrak{D}$ *is given by those* $x = (u,v) \in \mathbb{R}^2$ *such that*

$$u, v \;>\; 0 \quad and \quad uv \;\geq\; 1, \tag{7.14}$$

whose range is given by

$$0 \;\leq\; B(x) \;\leq\; \sqrt[4]{uv}, \qquad x \in \mathfrak{D},$$

and such that the following convexity property holds for all $x, x_\pm \in \mathfrak{D}$*:*

$$if \;\; x = \frac{x_+ + x_-}{2} \quad then \quad B(x) - \frac{B(x_+)+B(x_-)}{2} \;\geq\; C\frac{v^{1/4}}{u^{7/4}}(u_+ - u_-)^2 \tag{7.15}$$

with a numerical constant C. Then Proposition 3.14 will be proved.

Proof. Let $u_I := m_I w$, $v_I := m_I w^{-1}$, $v_+ = v_{I^+}$, $v_- = v_{I^-}$ and similarly for $u_\pm$. Then by Hölder's inequality (u, v) and $(u_\pm, v_\pm)$ belong to $\mathfrak{D}$.

Fix $J \in D$, by the convexity and range conditions

$$|J| \sqrt[4]{m_J w \, m_J w^{-1}} \;\geq\; |J| B(u_J, v_J)$$

$$\geq\;\; |J^+| B(u_+, v_+) + |J^-| B(u_-, v_-) + |J| C \frac{m_J^{1/4} w^{-1}}{m_J^{7/4} w} \left(m_{J^+} w - m_{J^-} w\right)^2 .$$

Iterating this process and using the fact that $B(u, v) \geq 0$ we get:

$$|J| \sqrt[4]{m_J w \, m_J w^{-1}} \;\geq\; C \sum_{I \in D(J)} |I| \frac{m_I^{1/4} w^{-1}}{m_I^{7/4} w} \left(m_{J^+} w - m_{J^-} w\right)^2,$$

which completes the proof of Lemma 7.7. ■

Now, in order to complete the proof of (7.13) we need to show the existence of the Bellman type function B which satisfies the conditions of Lemma 7.7.

Lemma 7.8 *The following function*

$$B(u, v) := \sqrt[4]{uv}$$

is defined on $\mathfrak{D}$, $0 \leq B(u, v) \leq \sqrt[4]{uv}$ for all $(u, v) \in \mathfrak{D}$, and satisfies the following differential inequality in $\mathfrak{D}$:

$$-\left(du, dv\right) d^2 B(u, v) \begin{pmatrix} du \\ dv \end{pmatrix} \;\geq\; \frac{1}{8} \frac{v^{1/4}}{u^{7/4}} |du|^2. \tag{7.16}$$

Furthermore, this implies the convexity condition (7.15) of Lemma 7.7.

Proof. Since u and v are positive in the domain $\mathfrak{D}$, the function $B = \sqrt[4]{uv}$ is well defined on $\mathfrak{D}$ and condition $0 \leq B(u, v) \leq \sqrt[4]{uv}$ is trivially satisfied.

Let us prove the differential inequality (7.16) now:

$$\begin{aligned}
-(du,dv)\,d^2B(u,v)\begin{pmatrix} du \\ dv \end{pmatrix} &= \frac{1}{16}(du,dv)\begin{pmatrix} 3v^{\frac{1}{4}}u^{\frac{-7}{4}} & -v^{\frac{-3}{4}}u^{\frac{-3}{4}} \\ -v^{\frac{-3}{4}}u^{\frac{-3}{4}} & 3v^{\frac{-7}{4}}u^{\frac{1}{4}} \end{pmatrix}\begin{pmatrix} du \\ dv \end{pmatrix} \\
&= \frac{1}{8}(du,dv)\begin{pmatrix} v^{\frac{1}{4}}u^{\frac{-7}{4}} & 0 \\ 0 & v^{\frac{-7}{4}}u^{\frac{1}{4}} \end{pmatrix}\begin{pmatrix} du \\ dv \end{pmatrix} \\
&\quad + \frac{1}{16}(du,dv)\begin{pmatrix} v^{\frac{1}{4}}u^{\frac{-7}{4}} & -v^{\frac{-3}{4}}u^{\frac{-3}{4}} \\ -v^{\frac{-3}{4}}u^{\frac{-3}{4}} & v^{\frac{-7}{4}}u^{\frac{1}{4}} \end{pmatrix}\begin{pmatrix} du \\ dv \end{pmatrix} \\
&\geq \frac{1}{8}v^{\frac{1}{4}}u^{\frac{-7}{4}}|du|^2,
\end{aligned}$$

as we wanted to show.

Now we only need to check the convexity condition (7.15). We fix an interval I and let

$$b(t) \;:=\; B(u_t, v_t),$$

$$u_t \;:=\; \frac{1}{2}(t+1)u_+ + \frac{1}{2}(1-t)u_-, \quad v_t \;:=\; \frac{1}{2}(t+1)v_+ + \frac{1}{2}(1-t)v_-,$$

$$-1 \;\leq\; t \;\leq\; 1.$$

In order to prove inequality (7.15) it is enough to establish

$$b(0) - \frac{b(1)+b(-1)}{2} \;\geq\; C\,\frac{v^{1/4}}{u^{7/4}}\,|u_+ - u_-|^2\,.$$

It is easy to see that (as we did in Section 7.2)

$$b(0) - \frac{1}{2}\left(b(-1)+b(1)\right) \;=\; -\frac{1}{2}\int_{-1}^{1}(1-|t|)b''(t)dt.$$

Note that

$$-b''(t) \;\geq\; \frac{1}{32}v_t^{1/4}u_t^{-7/4}\left(u_1 - u_{-1}\right)^2 \;\geq\; 0 \tag{7.17}$$

and that $\forall t \in [-1/2; 1/2] \quad u_t = u_0 + \frac{1}{2}t(u_1 - u_{-1})$, since domain $\mathfrak{D}$ is convex $u_t \in \mathfrak{D}$, and

$$|u_1 - u_{-1}| \ \leq \ |u_1| + |u_{-1}|, \quad |t| \leq 1/2, \ \ u_1, u_{-1} \geq 0,$$

$$-\frac{u_0}{2} \ = \ -\frac{1}{4}(u_1 + u_{-1}) \ \leq \ \frac{1}{2}t(u_1 - u_{-1}) \ \leq \ \frac{1}{4}(u_1 + u_{-1}) \ = \ \frac{u_0}{2},$$

and similarly for v:

$$-\frac{v_0}{2} \ = \ -\frac{1}{4}(v_1 + v_{-1}) \ \leq \ \frac{1}{2}t(v_1 - v_{-1}) \ \leq \ \frac{1}{4}(v_1 + v_{-1}) \ = \ \frac{v_0}{2},$$

so $u_t \ \leq \ \frac{3}{2}u_0$ and $v_t \ \geq \ \frac{1}{2}v_0$ for $t \in [-1/2; 1/2]$. Together with (7.17) it makes

$$-b''(t) \ \geq \ C v_0^{1/4} u_0^{-7/4} (u_1 - u_{-1})^2,$$

for $|t| < 1/2$.

For $1/2 < |t| \leq 1$ by (7.17), $-b''(t) \geq 0$, therefore

$$\begin{aligned} b(0) - \frac{b(-1) + b(1)}{2} \ &\geq \ -\frac{1}{2}\int_{-1/2}^{1/2} (1 - |t|)\, b''(t)dt \\ &\geq \ \frac{1}{2}\int_{-1/2}^{1/2} (1 - |t|)\, C v_0^{1/4} u_0^{-7/4} (u_1 - u_{-1})^2 dt. \end{aligned}$$

So,

$$B(u,v) - \frac{1}{2}(B(u_+, v_+) - B(u_-, v_-)) = b(0) - \frac{1}{2}(b(1) + b(-1)) \geq C\frac{v^{1/4}}{u^{7/4}}|du|^2$$

with numerical constant C. This completes the proof of Lemma 7.8 and Proposition 3.14. ∎

7.6 Proof of the $A_p(d\sigma)$-Weight Lemma (Proposition 3.12)

Proof. We are going to prove that for any weight $v \in A_p^d(d\sigma)$, for all dyadic intervals $J \in D$ inequality

$$\frac{1}{\sigma(J)} \sum_{I \in D(J)} \frac{\left(m_{I^+}^\sigma v - m_{I^-}^\sigma v\right)^2}{m_I^\sigma v} \sigma(I) \;\leq\; A\, m_J^\sigma v$$

holds with constant $A = C^{\frac{2p-1}{p}} \left(D^d(d\sigma)\right)^{p+1} \|w\|_{A_p^d}$.

Fix $0 < \epsilon \leq 1/2$ and define

$$\mathfrak{D}_{p;Q} \;:=\; \left\{(v,u) \mid\; 1 \leq vu^{p-1} \leq Q;\;\; u,v > 0\right\}$$

and

$$\mathfrak{D}_{p;2\epsilon^{1-p}Q} \;:=\; \left\{(v,u) \mid\; 1 \leq vu^{p-1} \leq 2\epsilon^{1-p}Q;\;\; u,v > 0\right\}.$$

Suppose we can find a function $B(v,u)$, defined on the domain $\mathfrak{D}_{p;Q}$ with the following properties:

(i) $0 \leq B(v,u) \leq 2Qv$ for all $(v,u) \in \mathfrak{D}_{p;Q}$,

(ii) for all triplets $(v,u), (v_\pm, u_\pm) \in \mathfrak{D}_{p;Q}$, such that $v = sv_+ + (1-s)v_-$ and $u = su_+ + (1-s)u_-$, where $\epsilon \leq s \leq 1-\epsilon$ ($\epsilon \leq 1-s \leq 1-\epsilon$), the following convexity condition holds for all such s:

$$\Delta_s B(v,u) \;=\; B(v,u) - sB(v_+,u_+) - (1-s)B(v_-,u_-) \;\geq\; C\frac{|v_+ - v_-|^2}{v}$$

with some constant C. Then Proposition 3.12 will be proved with constant $A \leq \frac{2Q}{C}$.

Let $v = v_I := m_I^\sigma v$, $u = u_I := m_I^\sigma v^{-\frac{1}{p-1}}$, $v_+ := v_{I^+}$, $v_- := v_{I^-}$, $u_+ := u_{I^+}$, $u_- := u_{I^-}$. Let $Q := \|v\|_{A_p^d(d\sigma)} = \sup_{I \in D} m_I^\sigma v \left(m_I^\sigma v^{-\frac{1}{p-1}}\right)^{p-1} = \sup_{I \in D} v_I u_I^{p-1}$. Clearly $(v,u), (v_+,u_+), (v_-,u_-) \in \mathfrak{D}_{p;Q}$. For each $I \in D$ let $s = s_I = \frac{\sigma(I^+)}{\sigma(I)}$ (note that $1 - s_I = 1 - \frac{\sigma(I^+)}{\sigma(I)} = \frac{\sigma(I^-)}{\sigma(I)}$). Then the definition of the dyadic doubling constant of σ,

$D^d(d\sigma) := \sup_{I\in D} \frac{\sigma(\tilde{I})}{\sigma(I)} < \infty$, where $\tilde{I}$ is a parent of I, implies that

$$\frac{\sigma(I^{\pm})}{\sigma(I)} \;\geq\; \frac{1}{D^d(d\sigma)} \;>\; 0, \quad s,\; 1-s \;\geq\; \frac{1}{D^d(d\sigma)} \;>\; 0,$$

$$0 \;<\; \frac{1}{D^d(d\sigma)} \;=\; \epsilon \;\leq\; s \;\leq\; 1 - \frac{1}{D^d(d\sigma)},$$

note also that $\epsilon := \frac{1}{D^d(d\sigma)} \leq \frac{1}{2}$ since $D^d(d\sigma) \geq 2$, as discussed in Section 2.2.

Now we can fix a dyadic interval $J \in D$. By the convexity and range conditions (i) and (ii), we conclude that

$$\begin{aligned} \sigma(J)Q\, m_J^\sigma v \;&\geq\; \sigma(J)B(v_J, u_J) \\ &\geq\; \sigma(J)s_J B(v_{J^+}, u_{J^+}) + \sigma(J)(1-s_J)B(v_{J^-}, u_{J^-}) + C\frac{|v_{J^+} - v_{J^-}|^2}{v_J}\sigma(J) \\ &=\; \sigma(J^+)B(v_{J^+}, u_{J^+}) + \sigma(J^-)B(v_{J^-}, u_{J^-}) + C\frac{|v_{J^+} - v_{J^-}|^2}{v_J}\sigma(J). \end{aligned}$$

Iterating this process, since $B \geq 0$ on $\mathfrak{D}_{p;Q}$, we get

$$\sigma(J)Q\, m_J^\sigma v \;\geq\; C \sum_{I\in D(J)} \frac{|v_{I^+} - v_{I^-}|^2}{v_I}\sigma(I),$$

which implies Proposition 3.12 with constant $A = \frac{1}{C}$.

Now we want to show that such function B with the above properties (i) and (ii) exists. We claim that the function

$$\begin{aligned} B(v,u) \;&:=\; v\left(2Q - \frac{2Q}{vu^{p-1}} - \frac{p\epsilon^{p-1}}{2p-1}\ln\left(vu^{p-1}\right)\right) \\ &=\; 2Qv - 2Qu^{1-p} - \frac{p\epsilon^{p-1}}{2p-1}v\ln v - \frac{p(p-1)\epsilon^{p-1}}{2p-1}v\ln u \end{aligned}$$

satisfies the above properties with constant $C = \frac{3p\epsilon^p}{4(2p-1)}$.

Let us check range conditions first: $0 \leq B(v,u) \leq 2Qv$ for all $(v,u) \in \mathfrak{D}_{p;Q}$.

Inequality

$$B(v,u) \;=\; v\left(2Q - \frac{2Q}{vu^{p-1}} - \frac{p\epsilon^{p-1}}{2p-1}\ln\left(vu^{p-1}\right)\right) \leq Qv$$

is trivial since $v,u \geq 0$, $vu^{p-1} \geq 1$ and $\frac{p\epsilon^{p-1}}{2p-1} \geq 0$. In order to show that $B(v,u)$ is positive on the domain $\mathfrak{D}_{p;Q}$, first note that $v > 0$ and $1 \leq vu^{p-1} \leq Q$ on $\mathfrak{D}_{p;Q}$. So, it is enough to show that $f(x) := 2Q - \frac{2Q}{x} - \frac{p\epsilon^{p-1}}{2p-1}\ln x \geq 0$ on the interval $1 \leq x \leq Q$. Indeed, $f'(x) = \frac{2Q}{x^2} - \frac{p\epsilon^{p-1}}{2p-1}\frac{1}{x}$, which is positive when $x \leq 2\frac{2p-1}{p\epsilon^{p-1}}Q$, and negative otherwise. Note also that when $p > 1$, $0 < \epsilon \leq 1/2$ then $2\frac{2p-1}{p\epsilon^{p-1}}Q \geq 2Q$. So, it is enough to check the end point $x = 1$: $f(1) = 0$, which proves the positivity of $B(v,u) \geq 0$.

We are going to show that the convexity condition holds for our function $B(v,u)$ with constant $C = \frac{3p\epsilon^p}{4(2p-1)}$. In order to do so, let us first check that

$$-\left(du, dv\right) d^2B(u,v)\begin{pmatrix} du \\ dv \end{pmatrix} \geq C\frac{|dv|^2}{v}$$

holds on the extended domain $\mathfrak{D}_{p;2\epsilon^{1-p}Q}$ with constant $C = \frac{p\epsilon^{p-1}}{2(2p-1)}$. To save space we might write $-d^2B$ instead of $-\left(du, dv\right) d^2B(u,v)\begin{pmatrix} du \\ dv \end{pmatrix}$.

$$\frac{\partial B}{\partial v} = 2Q - \frac{p\epsilon^{p-1}}{2p-1}\ln v - \frac{p\epsilon^{p-1}}{2p-1} - \frac{p(p-1)\epsilon^{p-1}}{2p-1}\ln u,$$

$$\frac{\partial B}{\partial u} = -2Q(1-p)u^{-p} - \frac{p(p-1)\epsilon^{p-1}}{2p-1}\frac{v}{u},$$

$$\frac{\partial^2 B}{\partial v^2} = -\frac{p\epsilon^{p-1}}{2p-1}\frac{1}{v}, \qquad \frac{\partial^2 B}{\partial v \partial u} = -\frac{p(p-1)\epsilon^{p-1}}{2p-1}\frac{1}{u},$$

$$\frac{\partial^2 B}{\partial u^2} = -2Qp(p-1)u^{-p-1} + \frac{p(p-1)\epsilon^{p-1}}{2p-1}\frac{v}{u^2},$$

then

$$
\begin{aligned}
-d^2B &= -\frac{\partial^2 B}{\partial v^2}(dv)^2 - 2\frac{\partial^2 B}{\partial v \partial u} dv du - \frac{\partial^2 B}{\partial u^2}(du)^2 \\
&= \frac{p\epsilon^{p-1}}{2p-1}\frac{1}{v}(dv)^2 + 2\frac{p(p-1)\epsilon^{p-1}}{2p-1}\frac{1}{u} du dv \\
&\quad + \left[2Qu^{-p-1} - \frac{\epsilon^{p-1}}{2p-1}\frac{v}{u^2}\right] p(p-1)(du)^2 \\
&= \frac{p\epsilon^{p-1}}{2(2p-1)}\frac{(dv)^2}{v} + \frac{p\epsilon^{p-1}}{2p-1}\left[\frac{(dv)^2}{2v} + 2\frac{(p-1)\sqrt{2v}}{\sqrt{2v}u} du dv + \frac{(p-1)^2 2v}{u^2}(du)^2\right] \\
&\quad + \left[2Qp(p-1)u^{-p-1} - \frac{p(p-1)\epsilon^{p-1}}{2p-1}\frac{v}{u^2} - \frac{p(p-1)^2\epsilon^{p-1}}{2p-1}\frac{2v}{u^2}\right](du)^2 \\
&= \frac{p\epsilon^{p-1}}{2(2p-1)}\frac{(dv)^2}{v} + \frac{p\epsilon^{p-1}}{2p-1}\left[\frac{dv}{\sqrt{2v}} + \frac{(p-1)\sqrt{2v}}{u} du\right]^2 + Kp(p-1)\epsilon^{p-1}(du)^2,
\end{aligned}
$$

where

$$
K = 2Q\epsilon^{1-p}u^{-p-1} - \frac{1}{2p-1}\frac{v}{u^2} - \frac{p-1}{2p-1}\frac{2v}{u^2}.
$$

Since $(v, u) \in \mathfrak{D}_{p;2\epsilon^{1-p}Q}$, $1 \leq vu^{p-1} \leq 2\epsilon^{1-p}Q$, $2Q\epsilon^{1-p} \geq vu^{p-1}$, we have

$$
K \;\geq\; vu^{p-1}u^{-p-1} - \frac{1}{2p-1}\frac{v}{u^2}(1 + 2(p-1)) \;=\; \frac{v}{u^2} - \frac{2p-1}{2p-1}\frac{v}{u^2} \;=\; 0,
$$

so $K \geq 0$, which shows that

$$
-(du, dv)\, d^2B(u, v) \begin{pmatrix} du \\ dv \end{pmatrix} \;\geq\; \frac{p\epsilon^{p-1}}{2(2p-1)}\frac{(dv)^2}{v}. \tag{7.18}
$$

Now we are ready to prove that the convexity condition (ii) holds for our function B.

Let $b(t) := B(v_t, u_t)$, $v_t := (s-t)v_+ + (t + (1-s))v_-$ and $u_t := (s-t)u_+ + (t + (1-s))u_-$, where $-1+s \leq t \leq s$. Note that $v_0 = v$, $v_{-1+s} = v_+$, $v_s = v_-$ and similarly for u_t, i.e. (v_t, u_t) is a point on the line segment from (v_+, u_+) to (v_-, u_-).

So, $v_t \le \max\{v_+, v_-\}$ and $u_t \le \max\{u_+, u_-\}$. Since

$$v \;=\; sv_+ + (1-s)v_-, \quad v_+ \;\le\; \frac{1}{s}v, \quad v_- \;\le\; \frac{1}{1-s}v,$$

and, similarly, for u, we have $\forall -1+s \le t \le s$

$$v_t \;\le\; \max\left\{\frac{1}{s}, \frac{1}{1-s}\right\} v \quad \textit{and} \quad u_t \;\le\; \max\left\{\frac{1}{s}, \frac{1}{1-s}\right\} u.$$

Note also that $\frac{dv}{dt} = v_- - v_+$ and $\frac{du}{dt} = u_- - u_+$, which implies that

$$\begin{aligned} -b''(t) &= -\left(\frac{du_t}{dt}, \frac{dv_t}{dt}\right) d^2B(u_t, v_t) \begin{pmatrix} \frac{du_t}{dt} \\ \frac{dv_t}{dt} \end{pmatrix} \\ &\ge \frac{p\epsilon^{p-1}}{2(2p-1)} \frac{|v_- - v_+|^2}{v(t)} \\ &\ge \frac{p\epsilon^{p-1}}{2(2p-1)} \max{}^{-1}\left\{\frac{1}{s}, \frac{1}{1-s}\right\} \frac{|v_- - v_+|^2}{v}. \end{aligned} \tag{7.19}$$

Next, we will show that if (v,u), $(v_\pm, u_\pm) \in \mathfrak{D}_{p;Q}$, then for all t, such that $\forall s-1 \le t \le s$ $(v_t, u_t) \in \mathfrak{D}_{p;2\epsilon^{1-p}Q}$.

We have the following Lemma:

Lemma 7.9 *Let* $\mathfrak{D}_{p;Q} := \{(x,y) \mid x, y > 0,\ \ 1 \le xy^{p-1} \le Q\}$, $0 < \epsilon \le \frac{1}{2}$. *Then for any triplet* (x,y), (x_+, y_+), (x_-, y_-), *such that* $x = sx_+ + (1-s)x_-$ *and* $y = sy_+ + (1-s)y_-$ *with some* $\epsilon \le s \le 1-\epsilon$, *and any* $t \in [0,1]$, $x_t := tx_+ + (1-t)x_-$ *and* $y_t := ty_+ + (1-t)y_-$ *the couple* (x_t, y_t) *belongs to the extended domain* $\mathfrak{D}_{p;2\epsilon^{-(p-1)}Q}$.

A proof of this Lemma can be found at the end of this section.

Now consider

$$\begin{aligned} -\Delta_s B(v,u) &= sB(v_+,u_+) + (1-s)B(v_-,u_-) - B(v,u) \\ &= sb(-1+s) + (1-s)b(s) - b(0) \\ &= \int_{s-1}^{0} (s(1-s)+st)b''(t)dt + \int_0^s (s(1-s)-(1-s)t)b''(t)dt. \end{aligned}$$

This is a calculus exercise similar to the one presented in Section 7.2.

By (7.19) and observing that

$$A = \int_{s-1}^{0} (s(1-s)+st)dt + \int_0^s (s(1-s)-(1-s)t)dt$$

is the area of the triangle with base on the interval $[s-1; s]$ and height $s(1-s)$, hence $A = \frac{s(1-s)}{2}$, we have

$$\begin{aligned} \Delta_s B(v,u) &\geq \frac{p\epsilon^{p-1}}{2(2p-1)} \max{}^{-1}\left\{\frac{1}{s}, \frac{1}{1-s}\right\} \frac{|v_- - v_+|^2}{v} \left[\int_{s-1}^{0} (s(1-s)+st)dt \right. \\ &\quad \left. + \int_0^s (s(1-s)-(1-s)t)dt\right] \\ &= \frac{p\epsilon^{p-1}}{4(2p-1)} s(1-s)\min\{s,1-s\} \frac{|v_- - v_+|^2}{v}, \end{aligned}$$

and, since $\epsilon \leq s \leq 1-\epsilon$ and $\epsilon \leq 1-s \leq 1-\epsilon$, then $\min\{s,1-s\} \geq \epsilon$ and $s(1-s) \geq \epsilon(1-\epsilon) \geq \epsilon/2$,

$$\Delta_s B(v,u) \geq \frac{p\epsilon^{p+1}}{8(2p-1)} \frac{|v_- - v_+|^2}{v}.$$

This completes the proof of Proposition 3.12 $\left(A = C\frac{(2p-1)}{p}\left(D^d(d\sigma)\right)^{p+1} Q\right)$. ∎

Proof. [Lemma 7.9] Let $0 < \epsilon \leq \frac{1}{2}$,

$$\mathfrak{D}_{p;Q} := \left\{(x,y) \mid x,y > 0,\ \ 1 \leq xy^{p-1} \leq Q\right\},$$

$$\mathfrak{D}_{p;2\epsilon^{-(p-1)}Q} := \left\{(x,y) \mid x,y>0,\ \ 1 \le xy^{p-1} \le 2\epsilon^{-(p-1)}Q\right\},$$

(x,y), $(x_\pm, y_\pm) \in \mathfrak{D}_{p;Q}$ be such that $x = sx_+ + (1-s)x_-$, $y = sy_+ + (1-s)y_-$ with some $\epsilon \le s \le 1-\epsilon$ and $(x_t, y_t) := (tx_+ + (1-t)x_-, ty_+ + (1-t)y_-)$, $0 \le t \le 1$. We want to show that $(x_t, y_t) \in \mathfrak{D}_{p;2\epsilon^{-(p-1)}Q}$.

Obviously both x_t and y_t are positive and by convexity of the lower boundary, if $xy^{p-1} \ge 1$ then $x_t y_t^{p-1} \ge 1$. So we only need to show that $x_t y_t^{p-1} \le 2\epsilon^{-(p-1)}Q$.

First note that since $(x,y) \in \mathfrak{D}_{p;Q}$, we can write

$$\begin{aligned} Q \ \ge\ xy^{p-1} &= (sx_+ + (1-s)x_-)\,(sy_+ + (1-s)y_-)^{p-1} \\ &= sx_+y_+^{p-1}\left(s + (1-s)\frac{y_-}{y_+}\right)^{p-1} + (1-s)x_-y_-^{p-1}\left(s\frac{y_+}{y_-} + (1-s)\right)^{p-1} \end{aligned}$$

and since every term is positive,

$$sx_+y_+^{p-1}\left(s + (1-s)\frac{y_-}{y_+}\right)^{p-1} \ \le\ Q$$

and

$$(1-s)x_-y_-^{p-1}\left(s\frac{y_+}{y_-} + (1-s)\right)^{p-1} \ \le\ Q.$$

So,

$$\left(s + (1-s)\frac{y_-}{y_+}\right)^{p-1} \ \le\ Qs^{-1}x_+^{-1}y_+^{1-p}$$

and

$$\left(s\frac{y_+}{y_-} + (1-s)\right)^{p-1} \ \le\ Q(1-s)^{-1}x_-^{-1}y_-^{1-p}$$

and, since both s and $1-s$ are nonnegative, we have

$$\left(\frac{y_-}{y_+}\right)^{p-1} \ \le\ Qs^{-1}(1-s)^{1-p}x_+^{-1}y_+^{1-p}$$

and

$$\left(\frac{y_+}{y_-}\right)^{p-1} \leq Q(1-s)^{-1}s^{1-p}x_-^{-1}y_-^{1-p}.$$

Now we take a look at $x_t y_t^{p-1}$:

$$\begin{aligned} x_t y_t^{p-1} &= (tx_+ + (1-t)x_-)(ty_+ + (1-t)y_-)^{p-1} \\ &= tx_+ y_+^{p-1}\left(t + (1-t)\frac{y_-}{y_+}\right)^{p-1} + (1-t)x_- y_-^{p-1}\left(t\frac{y_+}{y_-} + (1-t)\right)^{p-1}. \end{aligned}$$

Note that $\left(t + (1-t)\frac{y_-}{y_+}\right)$ is a weighted average between 1 and $\frac{y_-}{y_+}$, so $\left(t + (1-t)\frac{y_-}{y_+}\right) \leq \max\left(1, \frac{y_-}{y_+}\right)$, similarly $\left(t\frac{y_+}{y_-} + (1-t)\right) \leq \max\left(1, \frac{y_+}{y_-}\right)$. Let us first assume that $\frac{y_-}{y_+} \geq 1$, then $\frac{y_+}{y_-} \leq 1$,

$$t + (1-t)\frac{y_-}{y_+} \leq \frac{y_-}{y_+}$$

and

$$t\frac{y_+}{y_-} + (1-t) \leq 1$$

and we can bound $x_t y_t^{p-1}$ from above by:

$$\begin{aligned} x_t y_t^{p-1} &= tx_+ y_+^{p-1}\left(t + (1-t)\frac{y_-}{y_+}\right)^{p-1} + (1-t)x_- y_-^{p-1}\left(t\frac{y_+}{y_-} + (1-t)\right)^{p-1} \\ &\leq tx_+ y_+^{p-1}\left(\frac{y_-}{y_+}\right)^{p-1} + (1-t)x_- y_-^{p-1} \\ &\leq tx_+ y_+^{p-1} Q s^{-1}(1-s)^{1-p}x_+^{-1}y_+^{1-p} + (1-t)x_- y_-^{p-1} \\ &\leq tQs^{-1}(1-s)^{1-p} + (1-t)Q. \end{aligned}$$

Now note that $\epsilon \leq s \leq 1-\epsilon < 1$ and $\epsilon \leq (1-s) \leq 1-\epsilon < 1$, however only one of them, either s or $1-s$ can be less than 1/2. So, $s(1-s)^{p-1} \geq \frac{1}{2}\epsilon^{p-1}$ and we can write

$$x_t y_t^{p-1} \leq 2t\epsilon^{1-p}Q + (1-t)Q \leq 2\epsilon^{1-p}Q.$$

In the case when $\frac{y_-}{y_+} \leq 1$ we can argue in the same way.

The proof of Lemma 7.9 is complete. ■

References

[AISa] K. Astala, T. Ivanec, E. Saksman. Beltrami operators in the plane. *Duke Math. J.*, 107:27-56, 2001.

[BaJa] R. Bañuelos, P. Janakiraman. L^p-bounds for the Beurling-Ahlfors transform. *Trans. Amer. Math. Soc.*, to appear.

[Bl] S. Bloom. Solving weighted norm inequalities using the Rubio de Francia algorythm. *Proc. AMS*, 101:306-312, 1987.

[Bo] J. Bony. Calcul symbolique et propagation des singularités pour les équations aux derivées non-lineaires. *Ann. Sci. École Norm. Sup.*, 14:209-246, 1981.

[Buc1] S. Buckley. Estimates for operator norms on weighted spaces and reverse Jensen inequalities. *Trans. Amer. Math. Soc.*, 340:253-272, 1993.

[Buc2] S. Buckley. Summation conditions on weights. *Michigan Math. J.*, 40:153-170, 1993.

[Bur] D.I.Burkholder. Boundary value problems and sharp inequalities for martingale transforms. *The Annals of Probability*, 12:647–702, 1984.

[CoFe] R. Coifman, C. Fefferman. Weighted norm inequalities for maximal functions and singular integrals. *Studia Math.*, 51:241-250, 1974.

[CoJS] R. Coifman, P. Jones, S. Semmes. Two elementary proofs of the L^2 boundedness of Cauchy integrals on Lipschitz curves. *Journal of the AMS*, 2:553-564, 1989.

[CrMPe] D. Cruz-Uribe, J.-M. Martell, C. Pérez. *Weights, extrapolation and the theory of Rubio de Francia.* Book in preparation, 2007.

[CrPe] D. Cruz-Uribe, C. Pérez. Two weight extrapolation via the maximal operator. *J. Func. Analysis*, 174:1-17, 2000.

[JoDa] G. David, J.-L. Journé. A boundedness criterion for generalized Calderón-Zygmund operators. *Ann. of Math.*, 20:371-397, 1984.

[DrGrPerPet] O. Dragicevic, L. Grafakos, M.C. Pereyra and S. Petermichl. Extrapolation and sharp norm estimates for classical operators on weighted Lebesgue spaces. *Publ. Mat.*, 49:73-91, 2005.

[DrPetVo] O. Dragicevic, S. Petermichl, A. Volberg. A rotation method which gives linear L^p estimates for powers of the Ahlfors-Beurling operator. *J. Math. Pures Appl. (9)*, 6:492-509, 2006.

[DrVo] O. Dragicevic, A. Volberg. Bellman function and the dimensionless estimates of the Littlewood-Paley type. *J. of Operator Theory*, 56:167-198, 2006.

[Du] J. Duoandikoetxea. *Fourier analysis.* vol. 29 of Grad. Stud. in Math., AMS, Providence, RI, 2001.

[FeKPi] C. Fefferman, C. Kenig, J. Pipher. The theory of weights and the Dirichlet problem for elliptic equations. *Annals of Math.*, 134:65-124, 1991.

[GaRu] J. García-Cuerva, J. Rubio de Francia. *Weighted norm inequalities and related topics.* Math. Studies, North-Holland, 1985.

[Ge] F.W. Gehring. The L_p-integrability of the partial derivatives of a quasiconformal mapping. *Acta Math.*, 130:266-277, 1973.

[Gr] L. Grafakos. *Classical and modern Fourier analysis.* Prentice Hall, NJ, 2003.

[Ha] A. Haar. Zur Theorie der orthogonalen Funktion Systeme. *Math. Ann.*, 69:331-371, 1910.

[HeSz] H. Helson, G. Szegö. A problem in prediction theory. *Am. Math. Pura. Appl.*, 51:107-138, 1960.

[HTrVo] S. Hukovic, S. Treil, A. Volberg. The Bellman function and sharp weighted inequalities for square functions. *Oper. Theory Adv. Appl.*, 113:97-113, 2000.

[HuMW] R. Hunt, B. Muckenhoupt, R. Wheeden. Weighted norm inequalities for the conjugate function and the Hilbert transform. *Trans. AMS*, 176:227-252, 1973.

[JohNi] F. John, L. Nirenberg. On functions of bounded mean oscillation. *Comm. Pure Appl. Math.*, 14:415-426, 1961.

[KaPer] N.H. Katz, M.C. Pereyra. Haar multipliers, paraproducts, and weighted inequalities. *Analysis of Divergence*, Chapter 10:145-170, 1999.

[Le] A.K. Lerner. On some weighted norm inequalities for Littlewood-Paley operators. *Illinois J. Math.*, to appear.

[NTr] F. Nazarov, S. Treil. The hunt for a Bellman function: applications to estimates of singular integral operators and to other classical problems in harmonic analysis. *Algebra and Analysis*, 5:32-162, 1996. *Translation in St. Petersburg Math. J.*, 8:383-388, 1997.

[NTrVo1] F. Nazarov, S. Treil, A. Volberg. Cauchy integral and Calderón-Zygmund operators on nonhomogeneous spaces. *Internat. Math. Res. Notices*, 15:703-726, 1997.

[NTrVo2] F. Nazarov, S. Treil, A. Volberg. The Bellman functions and two-weight inequalities for Haar multipliers. *Journal of the AMS*, 12:909-928, 1999.

[NTrVo3] F. Nazarov, S. Treil, A. Volberg. Bellman function in stochastic control and harmonic analysis. *Oper. Theory Adv. Appl.*, 129:393-423, 2002.

[Per0] M. C. Pereyra. On the resolvent of dyadic paraproducts. *Rev. Mat. Ibéroamericana*, 10:627-664, 1994.

[Per1] M.C. Pereyra. Lecture notes on dyadic harmonic analysis. *Contemporary Mathematics*, 289:1-60, 2001.

[Per2] M.C. Pereyra. Haar multipliers meet Bellman functions. Preprint 2005.

[Pet1] S. Petermichl. The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of classical A_p characteristic. *Amer. J. Math.* 129:1355-1375, 2007.

[Pet2] S. Petermichl. The sharp bound for the Riesz transforms. *Proc. AMS*, 136:1237-1249, 2008.

[PetPo] S. Petermichl, S. Pott. An estimate for weighted Hilbert transform via square functions. *Trans. Amer. Math. Soc.*, 354:281-305, 2002.

[PetVo] S. Petermichl, A. Volberg. Heating of the Beurling operator: weakly quasiregular maps on the plane are quasilinear. *Duke Math. J.*, 112:281-305, 2002.

[Ru] J. Rubio de Francia. Factorization theory and A_p weights. *Amer. J. Math.*, 106:533-547, 1984.

[Saw1] E.T. Sawyer. A characterization of a two weight norm inequality for maximal operators. *Studia Math.*, 75:I-II, 1982.

[Saw2] E.T. Sawyer. A characterization of a two weight norm inequality for fractional and Poisson integrals. *Trans. Am. Math. Soc.*, 308:533-545, 1988.

[SlVa] L. Slavin, V. Vasyunin. Sharp results in the integral form of John-Nirenberg inequality. Preprint 2007.

[SlVo] L. Slavin, A. Volberg. Bellman function and the $H^1 - BMO$ duality. *Harm. An., PDE's and Rel. Topics*, 428:113-126, 2007.

[To] X. Tolsa. L^2-bounddness of the Cauchy integral operator for continuous measures. *Duke Math. J.*, 98:269-304, 1999.

[TrVo] S.Treil, A.Volberg. Wavelets and the angle between past and future. *J. Funct. Anal.*, 143:269–308, 1997.

[Va] V. Vasyunin. The exact constant in the inverse Hölder inequality for Muckenhoupt weights. *St.Petersburg Mathematical Journal*, 15:49–79, 2004.

[VaVo] V. Vasyunin, A. Volberg. The bellman functions for the simplest two weight inequality: an investigation of a particular case. *Alg. i An.*, 18:24-56, 2006. *Trans. in St.Petersburg Math. J.*, 18:201-222, 2007.

[Wit1] J. Wittwer. A sharp estimate on the norm of the martingale transform. *Math. Res. Lett.*, 7:1-12, 2000.

[Wit2] J. Wittwer. A sharp estimate on the norm of the continuous square function. *Proc. AMS*, 130:2335-2342, 2002.