CONTINUITY OF WEIGHTED ESTIMATES IN HARMONIC ANALYSIS WITH RESPECT TO THE WEIGHT

By

Nikolaos Pattakos

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Mathematics

2012

ABSTRACT

CONTINUITY OF WEIGHTED ESTIMATES IN HARMONIC ANALYSIS WITH RESPECT TO THE WEIGHT

By

Nikolaos Pattakos

Given the class of A_p weights, $1 < p < \infty$, we are able to define a metric d_* on this set such that the operator norm of any Calderón-Zygmund operator T on $L^p(w)$, $w \in A_p$, is a continuous function with respect to w. Moreover, we find the "rate" of this continuity with respect to the weight and prove that it is sharp. This is done by finding the exact "rate" for the Hilbert transform H on the unit disk. We also study many properties of this new metric space (A_p, d_*) and identify its completion as a subset of $BMO(\mathbb{R}^d)$. In addition, we extend the continuity result to the case of matrix-valued A_2 weights W, for the Martingale transform M_σ^W and we show that it does not hold for the classical Martingale transform. The problem of continuity of weighted estimates with respect to the weight appears naturally in problems of PDE (Partial Differential Equations) with random coefficients, and can also be important to multivariate stationary processes.

To my mother Antigoni, my father Georgios and my brother Evangelos.

ACKNOWLEDGMENTS

It would be a shame to cut this section short, since so many people have helped me.

I would like to thank my dissertation advisor, Dr. Alexander Volberg for his support and guidance through these years. He has always been there answering my questions and helping me with many of the problems that I had to face during my Ph.D. study. I was not even done with my qualifying exams yet, but I already had a distinguished and very helpful advisor. Because of him and his research I started to get involved in Harmonic Analysis and I should confess that it is my favorite subject. I remember stepping into his office for the first time in the beginning of Fall semester 2008 and asking him to be my advisor for my Ph.D. study. He accepted right away without any hesitation, even though he did not know me, and I am grateful for that.

My deepest thanks also go to my defense committee members Dr. Vladimir Peller, Dr. Ignacio Uriarte-Tuero, Dr. Jeffrey Schenker and Dr. Shen, Chun-Yen for their expertise and precious time. Spending time with them and attending their classes has been very important for me. Every time I needed them they were there for me. I am also grateful to Ms. Barbara Miller, Graduate Secretary in the Department of Mathematics, for her generous help during my graduate study. Here is a good time to say thank you to my advisors from the Math department of the University of Crete in Greece. People such as Dr. Michalis Papadimitrakis, Dr. Themistoklis Mitsis, Dr. Souzana Papadopoulou, Dr. Konstantinos Skandalis, Dr. George Kostakis and Dr. Athanasios Feidas helped me enormously during my undergraduate years in Greece.

Moreover, I would like to express my gratitude to Dr. Nicholas Boros for his friendship and precious help throughout my studies here in MSU. We spent many hours discussing problems of Harmonic Analysis together. I would also like to thank my friends from the math department Mr. Manousos Maridakis, Mr. Ambar Rao, Mr. Alexander Reznikov, Mr. Michalis Orfanoudakis, Dr. Mathhew Bond and Dr. Diogo Oliveira e Silva for giving me more chances to discuss mathematics with them. In addition, friends like Mr. George Koutsimanis, Dr. George Perdikakis, Dr. Zacharias Fthenakis, Dr. Eleni Beli, Dr. Artemis Spyrou, Dr. Evangelos Milliordos and Dr. Georgia Mavrommati made my stay here to East Lansing a better place to be.

Finally, I would like to thank my mother Antigoni, my father George and my brother Evangelos for supporting and believing in me. I can not come up with the words to appropriately express my gratitude. I could not finish writing this part of my thesis without mentioning the encouragement and support that I received all these years from Father Methodios and Father Onisimos from Greece. They have always been there for me.

TABLE OF CONTENTS

Chapter 1

Introduction

Weighted inequalities have been studied extensively during the last thirty years and have found many applications in PDE, geometric measure theory and multivariate stationary processes. The first characterization of the famous Muckenhoupt A_p classes in dimension 1, was done in [18] by B. Muckenhoupt. The definition, which we will state in dimension d, is the following. For a positive $L^1_{loc}(\mathbb{R}^d)$ function w and $p \in (1, +\infty)$ we write that $w \in A_p$ if the quantity

$$[w]_{A_p} := \sup_Q \Big(\frac{1}{|Q|}\int_Q w(x)dx\Big)\Big(\frac{1}{|Q|}\int_Q w(x)^{-\frac{1}{p-1}}dx\Big)^{p-1},$$

is finite, where we consider the supremum over all cubes Q inside $\mathbb{R}^d$. The number $[w]_{A_p}$ is called the A_p characteristic of the weight w. The A_∞ class of weights is defined to be the union of all the A_p classes. That is

$$A_\infty = \bigcup_{p>1} A_p.$$

There is also the following useful characterization of A_∞. For a weight w we have that $w \in A_\infty$ if and only if the quantity

$$[w]_{A_\infty} := \sup_Q \Big(\frac{\frac{1}{|Q|}\int_Q w(x)dx}{\exp(\frac{1}{|Q|}\int_Q \log w(x)dx)}\Big),$$

is finite, where we consider the supremum over all cubes Q inside $\mathbb{R}^d$. This is called the A_∞ characteristic of the weight.

In [18], Muckenhoupt characterized all weights, w, in dimension 1, and in [4] Coifman and Fefferman all weights in dimension d, with the property that the Hardy-Littlewood maximal operator M defined as

$$Mf(x) := \sup_Q \frac{1}{|Q|} \int_Q |f(y)|dy,$$

where we consider the supremum over all cubes Q in $\mathbb{R}^d$ such that $x \in Q$, is bounded from $L^p(w)$ to $L^p(w)$, for $p \in (1, +\infty)$. That is, under what conditions on the weight w the inequality

$$\Big(\int_{\mathbb{R}^d} |Mf(x)|^p w(x)dx \Big)^{\frac{1}{p}} \leq C \Big(\int_{\mathbb{R}^d} |f(x)|^p w(x)dx \Big)^{\frac{1}{p}},$$

holds for all functions $f \in L^p(w)$ and a constant $C > 0$ independent of f. As it turns out, this is true if and only if $w \in A_p$. The classical proof of this fact starts by showing that the A_p condition on w is necessary and sufficient for the Hardy-Littlewood maximal operator to be of weak type (p,p) with respect to w. This means that M sends $L^{p,\infty}(w)$ to $L^{p,\infty}(w)$. Then assuming that the A_p condition is true one shows using Calderón-Zygmund decomposition, that w satisfies the Reverse Hölder inequality which implies that there is an $\epsilon > 0$ such that $w \in A_{p-\epsilon}$. Finally, by the use of the Marcinkiewicz interpolation theorem we are done. Observe that the A_p condition is equivalent to the requirement that all averaging operators

$$f \mapsto \frac{1}{|I|} \int_I f(t)dt \cdot \chi_I,$$

are uniformly bounded on $L^p(w)$ with respect to the finite interval I. The Hardy-Littlewood maximal operator is exactly an averaging type operator.

It is a very natural question to ask what happens when we choose p to be 1. In this case it does not make sense to require that the Maximal operator sends $L^1(\mathbb{R}^d)$ to $L^1(\mathbb{R}^d)$ since this makes no sense even in the unweighted case. It is very well known that M sends $L^1(\mathbb{R}^d)$ to $L^{1,\infty}(\mathbb{R}^d)$. Therefore, the correct question is under what conditions on the weight w, M is bounded from $L^1(w)$ to $L^{1,\infty}(w)$. The answer is exactly when $w \in A_1$. The precise definition of this class of weights is the following. If there is a positive constant c such that

$$Mw(x) \leq cw(x),$$

for almost every $x \in \mathbb{R}^d$, we say that $w \in A_1$. The smallest such c is called the A_1 characteristic of the weight and is denoted by $[w]_{A_1}$. The relation among the characteristics of a weight w, for $p \in [1, +\infty]$, is

$$[w]_{A_\infty} \leq [w]_{A_p} \leq [w]_{A_1}.$$

This implies that the A_p classes are nested. That is $A_1 \subset A_p \subset A_\infty$.

A very interesting fact is that the A_p condition is necessary and sufficient for many of the classical operators of Harmonic Analysis to be bounded from $L^p(w)$ to $L^p(w)$, for $p \in (1, +\infty)$. For instance, the Hilbert transform

$$Hf(x) := \frac{1}{\pi} p.v. \int_{\mathbb{R}} \frac{f(y)}{x-y} dy,$$

and the Riesz transforms

$$R_j f(x) := \frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}} p.v. \int_{\mathbb{R}^d} \frac{x_j - y_j}{|x-y|^{d+1}} f(y) dy,$$

$1 \leq j \leq d$, are examples of such operators. Additionally, in [12] it was proven that the Hilbert transform is of weak type (p,p) with respect to the weight w if and only if $w \in A_p$. Their argument can be adopted to the case of Riesz transforms. Furthermore, Stein showed in [28] that if any of the Riesz transforms is bounded on $L^p(w)$ then $w \in A_p$.

By the use of good-λ inequalities and the strong Maximal operator, one is able to prove that all Calderón-Zygmund operators are bounded on $L^p(w)$, provided that $w \in A_p$. Recently though, the focus has been on a problem, now known as the A_2 conjecture for Calderón-Zygmund operators, on how to find the sharp dependence of the operator norm and the A_2 characteristic of the weight. It states that for any singular integral operator T of Calderón-Zygmund type we have the estimate

$$\|T\|_{L^2(w) \to L^2(w)} \leq c[w]_{A_2}, \tag{1.1}$$

for all A_2 weights w, where c is a positive constant independent of the weight. This result turns out to be correct and was first proven in [13]. This linear estimate with respect to the A_2 characteristic of the weight, is sharp for many of the classical operators, such as the Hilbert and Riesz transforms. Such estimates play a very important role in PDE. In fact, Calderón-Zygmund operators naturally arise as fraction derivatives of solutions of PDE. If we recall, for example, in [26] the authors proved that the Ahlfors-Beurling operator in the complex plane $\mathbb{C}$ defined as

$$Tf(z) = \frac{1}{\pi} p.v. \int_{\mathbb{C}} \frac{f(\zeta)}{(z-\zeta)^2} d\zeta,$$

satisfies (1.1), and as a consequence they obtained borderline regularity properties for solutions of the Beltrami equation on $\mathbb{C}$ ($u_{\overline{z}} = \mu u_z$, where μ is a given function of $\|\mu\|_{L^\infty} < 1$). The main tool in their proof is the Bellman function technique which is very powerful to such kind of estimates. We are also going to use this technique to estimate the norm of a Riesz matrix operator.

Chapter 2

The Muckenhoupt A_∞ class as a metric space

The main purpose of this chapter is to define a natural metric structure on the classical Muckenhoupt A_p classes. As far as we know, this is the first time that such metric has been studied in the context of continuity of norms of Calderón-Zygmund operators. Classically, the A_p spaces have only been treated as sets with no additional structure on them.

Before we define the metric structure we need to state some useful and well known results about the A_p classes and their relation with the $BMO(\mathbb{R}^d)$ space. First of all, the space of BMO functions in $\mathbb{R}^d$, consists of locally integrable functions f such that the norm

$$\|f\|_* = \sup_Q \frac{1}{|Q|} \int_Q |f(x) - f_Q| dx$$

is finite. Notice that this quantity becomes a norm if we identify all functions inside $BMO(\mathbb{R}^d)$ that differ by a constant. If f is a BMO function then for any number $\lambda \in (0, \frac{c}{\|f\|_*}]$, the function $e^{\lambda f}$ is an A_p weight, $1 < p < \infty$, where the constant c depends on p and the dimension d. Secondly, for small BMO norm, the A_p characteristic of the weight $e^{\lambda f}$ is bounded by the number 2 for example (see e.g. [8]). A subset of $BMO(\mathbb{R}^d)$ that appears in many applications is $BLO(\mathbb{R}^d)$. It stands for the functions of bounded lower oscillation. A function $f \in L^1_{loc}(\mathbb{R}^d)$ is said to belong in $BLO(\mathbb{R}^d)$ if there is a positive constant c such that:

$$\frac{1}{|Q|}\int_Q f(y)dy - \inf_{x\in Q} f(x) \le c$$

for all cubes Q, where the infimum is to be understood as the essential infimum. It can be proved that for any $w \in A_1$, the function $\log w$ is in $BLO(\mathbb{R}^d)$. Also if a function $f \in BLO(\mathbb{R}^d)$ then for sufficiently small $\lambda > 0$ the function $e^{\lambda f} \in A_1$. The reference for all these results is [8].

Let us observe that if we have any weight w, any positive constant $c > 0$ and any $1 \le p \le \infty$, then $[w]_{A_p} = [cw]_{A_p}$. We define an equivalence relation in A_∞ in the following way: for $u, v \in A_\infty$ we will write $u \sim v$ if and only if there is a positive constant c such that $u = cv$ almost everywhere in $\mathbb{R}^d$. This allows us to define the quotient space:

$$\mathcal{A}_\infty = A_\infty \big/ \sim .$$

In the same way we define for $1 \le p < \infty$:

$$\mathcal{A}_p = A_p \big/ \sim .$$

For two elements $u, v \in \mathcal{A}_\infty$ we define the distance function d_* as:

$$d_*(u, v) = \|\log u - \log v\|_*.$$

It is obvious that all the requirements of a metric are satisfied and the reason for defining the equivalence relation is exactly because we need to have:

$$d_*(u,v) = 0 \Leftrightarrow u \sim v.$$

So we define a metric in $\mathcal{A}_\infty$, going through the $BMO(\mathbb{R}^d)$ space. Notice that the restriction of the d_* metric to $\mathcal{A}_p$, makes the class a metric space. The drawback of these "new" metric spaces is that none of them is complete. However, the following is an obvious remark that gives more informations about this "new" spaces. It states that small balls around the constant weight 1, are complete in the d_* metric.

Theorem 1. *Consider a closed ball $\bar{B}(1,r)$ of sufficiently small radius $r > 0$ and center at the weight 1, in the metric space $(\mathcal{A}_\infty, d_*)$, i.e. $\bar{B}(1,r) = \{w \in \mathcal{A}_\infty : d_*(w,1) \leq r\}$. Then $\bar{B}(1,r)$ is a complete metric space with respect to the metric d_*.*

Proof. :

Consider a Cauchy sequence $\{w_n\}_{n\in\mathbb{N}}$ in $(\bar{B}(1,r), d_*)$. This means that the sequence $\{\log w_n\}_{n\in\mathbb{N}}$ is Cauchy in the $BMO(\mathbb{R}^d)$ space. But $BMO(\mathbb{R}^d)$ is a Banach space and so there is a function $f \in BMO(\mathbb{R}^d)$ such that $\log w_n \to f$ in $BMO(\mathbb{R}^d)$ as $n \to \infty$. By the John-Nirenberg inequality we know that there is a dimensional constant $c > 0$ such that for all $\lambda \in (0, \frac{c}{\|f\|_*}]$ the function $e^{\lambda f} \in A_2$. But $|\|\log w_n\|_* - \|f\|_*| \leq \|\log w_n - f\|_* \to 0$ as $n \to \infty$. Here we use the fact that $w_n \in \bar{B}(1,r)$. This means that $\|\log w_n\|_* = \|\log w_n - \log 1\|_* \leq r$ and r is sufficiently small. Therefore, the number $\|f\|_*$ is small and so the number $\frac{c}{\|f\|_*}$ is really big. We are now allowed to choose for $\lambda = 1$ and we get that $e^f \in A_2$ or equivalently there is a weight $w \in A_2 \subset A_\infty$ with $f = \log w$. It is trivial now to see that $d_*(w_n, w) \to 0$ as $n \to \infty$. □

Of course in the previous Theorem, we can replace the $\mathcal{A}_\infty$ space by any of the other $\mathcal{A}_p$ spaces. We already mentioned that none of the $\mathcal{A}_p$ spaces is complete. The proof of this fact is very simple. Let us prove that $\mathcal{A}_1$ is not complete by finding a Cauchy sequence in the space that has no limit inside $\mathcal{A}_1$. It will follow that this example works for anyone of the $\mathcal{A}_p$ spaces. Consider a decreasing sequence $-1 < r_n < 0$ with $\lim_{n\to\infty} r_n = -1$. Define the A_1 weights $w_n = |x|^{r_n}$. Then:

$$d_*(w_{r_n}, w_{r_m}) = \|r_n \log|x| - r_m \log|x|\|_* = |r_n - r_m| \|\log|x|\|_*$$

and since $r_n \to 1$ we see that $\{w_n\}_{n\in\mathbb{N}}$ is Cauchy in $\mathcal{A}_1$, or equivalently the sequence $\{\log w_n\}_{n\in\mathbb{N}}$ is Cauchy in $BMO(\mathbb{R}^d)$. It's limit in the $BMO(\mathbb{R}^d)$ space is obviously the function $f(x) = -\log|x|$. This means that for $w(x) = \frac{1}{|x|}$ we have $d_*(w_n, w) \to 0$ as $n \to \infty$, but since w is not in $L^1_{loc}(\mathbb{R}^d)$ it can not be an A_1 weight. So the space $(\mathcal{A}_1, d_*)$ is not complete.

Let us also mention the following result in [9], by Garnett and Jones, that helps to understand better when a ball in $(\mathcal{A}_p, d_*)$ is complete. It states that for a function $f \in BMO(\mathbb{R}^d)$,

$$dist_{BMO}(f, L^\infty) := \inf\{\|f-g\|_* : g \in L^\infty\} \sim \frac{1}{\sup\{\lambda > 0 : e^{\lambda f} \in A_2\}}.$$

This means that if we have a Cauchy sequence in $\mathcal{A}_p$, the closer the sequence is to the $L^\infty(\mathbb{R}^d)$ space, the more chances it has to have a limit in $\mathcal{A}_p$.

So now we can try and find the completion of these spaces under the metric d_*. By definition the completion of $(\mathcal{A}_p, d_*)$ is the space $\bar{\mathcal{A}}_p$ that consists of the equivalence classes of all Cauchy sequences of $\mathcal{A}_p$. We can identify this space as a subspace of $BMO(\mathbb{R}^d)$. Indeed:

$$\bar{\mathcal{A}}_p = \{f \in BMO(\mathbb{R}^d) : \exists \{w_n\}_{n\in\mathbb{N}} \subset A_p : \lim_{n\to\infty} \| \log w_n - f\|_* = 0\},$$

and we can think of the $\mathcal{A}_p$ class as a subset of $\bar{\mathcal{A}}_p$, by identifying every weight w with it's logarithm, $\log w$, in $BMO(\mathbb{R}^d)$. Since the classical A_p spaces form an increasing "sequence" of the variable p (and of course the same is true for the $\mathcal{A}_p$ spaces), the same is true for this new subspaces of $BMO(\mathbb{R}^d)$, $\bar{\mathcal{A}}_1 \subset \bar{\mathcal{A}}_p \subset \bar{\mathcal{A}}_q \subset \bar{\mathcal{A}}_\infty \subset BMO(\mathbb{R}^d)$, for $1 \leq p \leq q \leq \infty$.

They are also convex subsets of $BMO(\mathbb{R}^d)$. Indeed, consider $1 < p < \infty$, and $f, g \in \bar{\mathcal{A}}_p$. This means that there are sequences $\{w_n\}_{n\in\mathbb{N}}, \{v_n\}_{n\in\mathbb{N}} \subset A_p$ such that: $f = \lim_{n\to\infty} w_n$, $g = \lim_{n\to\infty} v_n$, in $BMO(\mathbb{R}^d)$. Let $0 < t < 1$ be fixed. We will show that $tf + (1-t)g \in \bar{\mathcal{A}}_p$. For this, we only need to see that $tf + (1-t)g = \lim_{n\to\infty} \log(w_n^t v_n^{1-t})$, in $BMO(\mathbb{R}^d)$, and check using Hölder that the weight $w_n^t v_n^{1-t} \in A_p$, for all n, since:

$$[w^t v^{1-t}]_{A_p} \leq [w]_{A_p}^t [v]_{A_p}^{1-t},$$

for all $w, v \in A_p$. Thus, $tf + (1-t)g \in \bar{\mathcal{A}}_p$. It is trivial to see now that $\bar{\mathcal{A}}_\infty$ is also a convex subset of $BMO(\mathbb{R}^d)$. For $\bar{\mathcal{A}}_1$ the same holds, since if we have two A_1 weights, w, v, it is trivial to see that $w^t v^{1-t} \in A_1$ and actually that $[w^t v^{1-t}]_{A_1} \leq [w]_{A_1}^t [v]_{A_1}^{1-t}$.

Here, let us observe that for any $1 < p < \infty$, we have that $L^\infty(\mathbb{R}^d) \subset \bar{\mathcal{A}}_p$. There is a nice result of weighted theory (see [8]) that states the following (we will present the statement only for A_2): There are dimensional constants $c_1, c_2 > 0$, such that for a function ϕ in $\mathbb{R}^d$ we have:

a) $e^{\phi} \in A_2$ provided $\inf\{\|\phi - g\|_* : g \in L^\infty(\mathbb{R}^d)\} \le c_1$ and

b) $\inf\{\|\phi - g\|_* : g \in L^\infty(\mathbb{R}^d)\} \le c_2$ provided $e^{\phi} \in A_2$. This means that all functions $f \in BMO(\mathbb{R}^d)$ that satisfy the assumption a), belong to the $\bar{\mathcal{A}}_2$ space. Equivalently, there is a small neighborhood of $L^\infty(\mathbb{R}^d)$ inside $BMO(\mathbb{R}^d)$, that lies inside the $\bar{\mathcal{A}}_2$ space.

We should also mention that since:

$$BLO(\mathbb{R}^d) = \{\alpha \log w : \alpha \ge 0, w \in A_1\},$$

we can ask the question if the spaces $\bar{\mathcal{A}}_1$, $BLO(\mathbb{R}^d)$ are equal. Let us assume that they are. A classical result of weighted theory is that $BMO(\mathbb{R}^d) = BLO(\mathbb{R}^d) - BLO(\mathbb{R}^d)$. By our assumption we have that $BMO(\mathbb{R}^d) = \bar{\mathcal{A}}_1 - \bar{\mathcal{A}}_1$. Now consider a function $f \in BMO(\mathbb{R}^d)$. There are functions $\phi, \psi \in \bar{\mathcal{A}}_1$ such that $f = \phi - \psi$. We know that there are sequences of A_1 weights $\{\phi_n\}_{n\in\mathbb{N}}, \{\psi_n\}_{n\in\mathbb{N}}$ such that $f = \lim_{n\to\infty} \log\phi_n - \lim_{n\to\infty} \log\psi_n = \lim_{n\to\infty} \log\phi_n\psi_n^{-1}$, where the limit is in $BMO(\mathbb{R}^d)$. But $\phi_n\psi_n^{-1}$ is an A_2 weight for all n. So we get that $\bar{\mathcal{A}}_2 = BMO(\mathbb{R}^d)$. But this is obviously false.

Notice that from the argument follows the inclusion, $\bar{\mathcal{A}}_1 - \bar{\mathcal{A}}_1 \subset \bar{\mathcal{A}}_2$. Trivially, we have the more general fact, that for any $1 < p < \infty$, $\bar{\mathcal{A}}_1 + (1-p)\bar{\mathcal{A}}_1 \subset \bar{\mathcal{A}}_p$. Also, since we have that $w \in A_p \Leftrightarrow w^{1-p'} \in A_{p'}$, we get the equivalence $f \in \bar{\mathcal{A}}_p \Leftrightarrow (1-p')f \in \bar{\mathcal{A}}_{p'}$. For $p = 2$ we have $f \in \bar{\mathcal{A}}_2 \Leftrightarrow -f \in \bar{\mathcal{A}}_2$, which means that the $\bar{\mathcal{A}}_2$ class is symmetric with respect to the origin in the $BMO(\mathbb{R}^d)$ space. No other $\bar{\mathcal{A}}_p$ class has this property. Here we should remember the following about power weights. A function of the form $|x|^\alpha$ is an A_p weight in $\mathbb{R}^d$, if and only if $-d < \alpha < d(p-1)$. The interval for α is symmetric with respect to the origin, if and only if $p = 2$. Now we can see that there is a "correspondence" between the $\bar{\mathcal{A}}_2$ space and the interval $(-d, d)$.

Chapter 3

Continuity of weighted estimates and sharpness of result

3.1 The continuity on the weight

In this chapter we are going to study the behavior of the operator norm of a sub-linear operator T on $L^p(w)$ with respect to the weight w. Our goal is to show that if two weights w and w_0 are close in the metric d_*, defined in the previous chapter, then the numbers $\|T\|_{p,w}$ and $\|T\|_{p,w_0}$ are also close. The main Theorem of this chapter is the following.

Theorem 2. *Consider $1 < p < +\infty$ and $w_0 \in A_p$. Suppose that the sub-linear operator T on $\mathbb{R}^d$ satisfies the weighted estimate*

$$\|T\|_{L^p(w)\to L^p(w)} \leq F([w]_{A_p}),$$

for all $w \in A_p$, where F is a positive increasing function. Then there is a positive constant c that depends on p, the dimension d, $[w_0]_{A_p}$ and the function F such that for all weights w that are sufficiently close to w_0 in the metric d_*,

$$\|T\|_{L^p(w)\to L^p(w)} \leq \|T\|_{L^p(w_0)\to L^p(w_0)}(1 + cd_*(w, w_0)).$$

Moreover, we have

$$\lim_{d_*(w,w_0)\to 0} \|T\|_{L^p(w)\to L^p(w)} = \|T\|_{L^p(w_0)\to L^p(w_0)}.$$

Remark 3. In [2] Buckley showed that the Hardy-Littlewood maximal operator satisfies the estimate

$$\|M\|_{L^p(w)\to L^p(w)} \le c[w]_{A_p}^{\frac{1}{p-1}},$$

for $1 < p < +\infty$, and all weights $w \in A_p$, where the constant $c > 0$ is independent of the weight w. This means that the assumptions of Theorem 2 hold for M.

Remark 4. Consider any Calderón-Zygmund operator T. By [13] we know that:

$$\|T\|_{L^p(w)\to L^p(w)} \le c\,[w]_{A_p}^{max(1,\frac{1}{p-1})},$$

for any A_p weight w, where $c > 0$ is independent of the weight. This means that we can apply Theorem 2, for $1 < p < \infty$ and $F(x) = cx^{max(1,\frac{1}{p-1})}$.

Before moving on to the proof of Theorem 2 we need to see some preliminary results. For the proof of our theorem interpolation with change of measure is going to play an important role. In the following $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ will denote measure spaces. Suppose T is an operator of a class of functions on X into a class of functions on Y. T is called a sub-linear operator, if it satisfies the following properties:

i)If $f(x) = f_1(x) + f_2(x)$ and $Tf_1(x), Tf_2(x)$ are defined then $Tf(x)$ is defined,

ii)$|T(f_1(x) + f_2(x))| \le |Tf_1(x)| + |Tf_2(x)|$, μ almost everywhere,

iii)For any scalar k, we have $|T(kf(x))| = |k||Tf(x)|$, μ almost everywhere.

Let μ_0, μ_1 be two measures for $(X, \mathcal{M})$. If we define the measure $\mu = \mu_0 + \mu_1$, then μ_0, μ_1

are each absolutely continuous with respect to μ. Thus, by the Radon-Nikodym theorem, there exists two functions, α_0, α_1 such that for any $E \in \mathcal{M}$,

$$\mu_j(E) = \int_E \alpha_j(x) d\mu(x)$$

where $j = 0, 1$. In the following we will assume that α_0, α_1 are never zero. This is equivalent to asserting that the sets of measure zero with respect to μ_j, $j = 0, 1$, are the same as the sets of measure zero with respect to μ. Thus, in the various measure spaces that we will consider, the equivalence classes of functions will be the same. Let $0 \leq s \leq 1$, and define the measure μ_s on X by

$$\mu_s(E) = \int_E \alpha_0^{1-s}(x) \alpha_1^s(x) d\mu(x),$$

for each $E \in \mathcal{M}$. Also assume, that we have two measures ν_0, ν_1 on $\mathcal{N}$, and define the measures ν_r, for $0 \leq r \leq 1$, just as we did for μ_s above.

Given any real numbers $1 \leq p_0, p_1, q_0, q_1$ and any $0 \leq t \leq 1$, we define $p_t, q_t, s(t), r(t)$ as follows:

$$\frac{(1-t)p_t}{p_0} + \frac{tp_t}{p_1} = 1, \frac{(1-t)q_t}{q_0} + \frac{tq_t}{q_1} = 1$$

$$s(t) = \frac{(tp_t)}{p_1}, r(t) = \frac{(tq_t)}{q_1}.$$

We have the following Theorem by [29]:

Theorem 5. *Suppose that T is a sub-linear operator satisfying*

$$\|Tf\|_{q_j,\nu_j} \leq K_j \|f\|_{p_j,\mu_j}$$

for all $f \in L^{p_j}(X, \mathcal{M}, \mu_j)$, $j = 0, 1$. *Then, for* $0 \le t \le 1$, *we have*

$$\|Tf\|_{q_t, \nu_{r(t)}} \le K_0^{1-t} K_1^t \|f\|_{p_t, \mu_{s(t)}}$$

for all $f \in L^{p_t}(X, \mathcal{M}, \mu_{s(t)})$.

In addition to the previous theorem we need also the following proved in [15]:

Theorem 6. *If the* A_∞ *characteristic of a weight* w *is small, i.e.* $[w]_{A_\infty} \le 1 + \delta < 2$, *then the function* $f = \log w$, *and any cube* Q *satisfy*

$$\frac{1}{|Q|} \int_Q |f(x) - f_Q| dx \le 32\sqrt{\delta}.$$

We will give a rough idea for A_2, since for A_∞ is similar. We will show that for any A_2 weight w:

$$\|\log w\|_* \le 2\sqrt{[w]_{A_2} - 1}.$$

Indeed, for any real number x we have that: $2 + x^2 \le e^x + e^{-x}$. Now apply it with $x = \log(\frac{w}{w_Q})$ and get:

$$\frac{1}{|Q|} \int_Q \Big(2 + \Big(\log \frac{w}{w_Q}\Big)^2\Big) \le \frac{1}{|Q|} \int_Q \frac{w}{w_Q} + \frac{1}{|Q|} \int_Q \frac{w_Q}{w}.$$

Hence,

$$\frac{1}{|Q|} \int_Q |\log w - \log(w_Q)|^2 \le 1 + w_Q (w^{-1})_Q - 2 \le [w]_{A_2} - 1.$$

By Hölder's inequality, we have f: $\frac{1}{|Q|}\int_Q f \leq \sqrt{\frac{1}{|Q|}\int_Q f^2}$, for any positive function f. Thus,

$$\frac{1}{|Q|}\int_Q |\log w - \log(w_Q)| \leq \sqrt{[w]_{A_2} - 1}.$$

Now using the well known inequality

$$\frac{1}{|Q|}\int_Q |\log w - (\log w)_Q| \leq 2 \inf_{r\in\mathbb{R}} \frac{1}{|Q|}\int_Q |\log w - r|,$$

we get exactly what we want. Now for a general A_∞ weight the proof follows the same lines. See [15] for more details and for the proof that the square root is sharp.

Now we are ready to present the proof of Theorem 2.

Proof. First we will show that for any sub-linear operator T that satisfies the assumptions of our theorem we have:

$$\|T\|_{L^p(w)\to L^p(w)} \leq \|T\|_{L^p(w_0)\to L^p(w_0)}(1 + c\delta),$$

for all weights $w \in A_p$ with $d_*(w, w_0) \leq \delta$. Let $\delta > 0$ be a small number that we consider to be fixed. Fix also an A_p weight w, with $d_*(w, w_0) < \delta$. This means that $\|\log \frac{w}{w_0}\|_* \leq \delta$. We would like to write our weight w as $w = w_0^{1-t}W^t$, for some small and positive number t (which is going to be about δ), and some weight $W \in A_p$. From the expression we can see that

$$W = \frac{w^{\frac{1}{t}}}{w_0^{\frac{1}{t}}} w_0.$$

For this, let us consider only the case $p = 2$, but the general case is identical to this one. Since $w_0 \in A_2$ we know that there is a small $\epsilon > 0$ such that $w_1 := w_0^{1+\epsilon} \in A_2$. Then obviously $w_0 = w_1^{1-s}$ for small $s > 0$. To continue, consider the function $f = \log \left(\frac{w}{w_0}\right)^{\frac{1}{s}}$. The BMO norm of f is really small since:

$$\|f\|_* = \frac{1}{s} d_*(w, w_0) \leq \frac{1}{s}\delta,$$

and so by the John-Nirenberg inequality we have that for all $\lambda \in (0, \frac{c}{\|f\|_*}]$ the function $e^{\lambda f} = \left(\frac{w}{w_0}\right)^{\frac{\lambda}{s}} \in A_2$, where c is a positive constant that depends only on the dimension. If we choose $\lambda = \frac{c_0}{\delta}$, $c_0 > 0$ is any constant less than or equal to sc, we see that $w_2 := \left(\frac{w}{w_0}\right)^{\frac{c_0}{\delta s}} \in A_2$, which implies that the function $w_1^{1-s} w_2^s \in A_2$. Then:

$$W := \frac{w^{\frac{1}{t}}}{w_0^{\frac{1}{t}}} w_0 = w_1^{1-s} w_2^s \in A_2,$$

where we put $t = \frac{\delta}{c_0}$. Here we should mention that the A_2 norm of W can be chosen to be bounded above by a constant that depends only on the A_2 norm of w_1. On the other hand, $[w_1]_{A_2}$ depends only on the A_2 norm of w_0, and this is fixed. With this in mind, let us assume that the A_2 characteristic of W is bounded above by c. The important thing here is that it does not depend on δ. Write $\gamma = \|T\|_{L^p(w_0) \to L^p(w_0)}$. By the interpolation result of Stein and Weiss, Theorem 5, for $X = Y = \mathbb{R}^d$, $\mathcal{M} = \mathcal{N} = \mathcal{L}$, where by $\mathcal{L}$ we denote the σ-algebra of Lebesgue measurable sets in $\mathbb{R}^d$, and $\mu_0 = \nu_0 = w_0 dx, \mu_1 = \nu_1 = W dx$, we get

$$\begin{aligned} \|T\|_{L^p(w) \to L^p(w)} &\leq \gamma^{1-t} \|T\|^t_{L^p(W) \to L^p(W)} \\ &\leq \gamma^{1-t} c^t F\Big([W]_{A_p}\Big)^t \\ &\leq \gamma^{1-t} c^t F(c)^t \end{aligned}$$

and the right-hand side goes to γ as $t \to 0^+$ or equivalently as $\delta \to 0^+$. In other words:

$$\limsup_{d_*(w,w_0)\to 0} \|T\|_{L^p(w)\to L^p(w)} \leq \|T\|_{L^p(w_0)\to L^p(w_0)}$$

and in addition we have the desired estimate:

$$\|T\|_{L^p(w)\to L^p(w)} \leq \|T\|_{L^p(w_0)\to L^p(w_0)}(1 + c\delta),$$

where c is a constant depending on n, p, $[w_0]_{A_p}$ and the function F, for all weights w in A_p that are δ close to w_0 in the metric d_*.

We can also conclude the following new result:

Proposition 7. *The set*

$$\{\log w : w \in A_p\}$$

is open in $BMO(\mathbb{R}^d)$ for all $1 < p < +\infty$.

Proof. To see this fix $w_0 \in A_p$ and choose sufficiently small $\delta > 0$. For $f \in BMO(\mathbb{R}^d)$ with $\|f - \log w_0\|_* \leq \delta$, write $f = \log u$, where u is a positive function. Then follow the previous reasoning in the beginning of the proof, with $w = u$ and write $u = w_0^{1-t}W^t$, for $0 < t < 1$. It follows that $W \in A_p$, if $\delta > 0$ is small depending only on the A_p norm of w_0, and so $u = w_0^{1-t}W^t$ is an A_p weight, by Hölder's inequality. As we can see, this is exactly the same argument as before. $\square$

There is only one thing remaining to finish the proof of Theorem 2. We need to show that

$$\|T\|_{L^p(w_0)\to L^p(w_0)} \leq \liminf_{d_*(w,w_0)\to 0} \|T\|_{L^p(w)\to L^p(w)}.$$

We are going to resent two different proofs. The first appeared in [24] and the second in [22]. Both approaches give different information about the weights involved in the calculations.

For the first proof we assume that our operator T is linear. Let us also assume for simplicity that $p = 2$ and that $\|T\|_{L^2(w_0)\to L^2(w_0)} = 1$. Note that other p's can be treated similarly. Let M_ϕ denote the operation of multiplication by ϕ. To finish the proof of the continuity at $w = w_0$ we are going to assume that the quantity

$$\liminf_{d_*(w,w_0)\to 0} \|T\|_{L^2(w)\to L^2(w)}$$

which is equal to

$$\liminf_{d_*(w,w_0)\to 0} \left\| M_{w_0^{-\frac{1}{2}} w^{\frac{1}{2}}} T M_{w_0^{\frac{1}{2}} w^{-\frac{1}{2}}} \right\|_{L^2(w_0)\to L^2(w_0)}$$

is strictly less than 1 and get a contradiction. This means that there is $\tau > 0$ small, and a sequence of A_2 weights w_n such that $d_*(w_n, w_0) \to 0$ as $n \to \infty$ and in addition:

$$\|w_0^{-\frac{1}{2}} w_n^{\frac{1}{2}} T w_0^{\frac{1}{2}} w_n^{-\frac{1}{2}} g\|_{L^2(w_0)} \le (1-\tau)\|g\|_{L^2(w_0)} \tag{3.1}$$

for all functions $g \in L^2(w_0)$. Fix now any cube Q in $\mathbb{R}^d$. Here we can make the normalization assumption $\frac{1}{|Q|}\int_Q \frac{w_n}{w_0} dx = 1$ for all $n \in \mathbb{N}$. We claim two things::

$1^*)\|w_n^{-\frac{1}{2}} - w_0^{-\frac{1}{2}}\|_{L^2(w_0,Q)} \to 0$ as $n \to \infty$ where by $L^2(w_0, Q)$ we mean the $L^2(w_0)$ norm over Q, and

$2^*)$ there exists a subsequence k_n such that $w_{k_n} \to w_0$ almost everywhere in the cube Q.

Obviously 2^* follows from 1^*. For a proof of 1^*, see Lemma after the end of this proof.

Now without loss of generality we can assume that the subsequence is the original sequence w_n. Notice that 1^* implies $\|w_n^{-\frac{1}{2}} f - w_0^{-\frac{1}{2}} f\|_{L^2(w_0,Q)} \to 0$ as $n \to \infty$ for all bounded f, and so for $g = f w_0^{-\frac{1}{2}}$, we get $\|T(w_0^{\frac{1}{2}} w_n^{-\frac{1}{2}} g) - Tg\|_{L^2(w_0,Q)} \to 0$ as $n \to \infty$ and this implies that for a subsequence of w_n (which again we assume that is the whole sequence), $w_0^{-\frac{1}{2}} w_n^{\frac{1}{2}} T w_0^{\frac{1}{2}} w_n^{-\frac{1}{2}} g \to Tg$ almost everywhere in the cube Q. It is time to apply Fatou's Lemma in inequality (3.1) and get:

$$\begin{aligned}
\left\| \liminf_{n\to\infty} w_0^{-\frac{1}{2}} w_n^{\frac{1}{2}} T w_0^{\frac{1}{2}} w_n^{-\frac{1}{2}} g \right\|_{L^2(w_0,Q)} &\leq \liminf_{n\to\infty} \left\| w_0^{-\frac{1}{2}} w_n^{\frac{1}{2}} T w_0^{\frac{1}{2}} w_n^{-\frac{1}{2}} g \right\|_{L^2(w_0,Q)} \\
&\leq (1-\tau)\|g\|_{L^2(w_0,Q)}.
\end{aligned}$$

Here $g = f w_0^{-\frac{1}{2}}$ with bounded f form a dense family in $L^2(w_0, Q)$. For g from this dense family it follows:

$$\|Tg\|_{L^2(w_0)} \leq (1-\tau)\|g\|_{L^2(w_0)}$$

by letting the cube Q expand to infinity, for g in some dense subclass of $L^2(w_0)$. By assumption $\|T\|_{L^2(w_0)\to L^2(w_0)} = 1$ and this is how we have our contradiction. All that remains is the following Lemma:

Lemma 8. *Let $w_0, w \in A_2$ such that $d_*(w, w_0) \leq \epsilon$, where ϵ is sufficiently small. Let us have a normalization assumption $\frac{1}{|Q|}\int_Q \frac{w}{w_0} dx = 1$. Then $\|w^{-\frac{1}{2}} - w_0^{-\frac{1}{2}}\|_{L^2(w_0,Q)} \leq |Q|^{\frac{1}{2}} c(\epsilon)^{\frac{1}{2}}$, where $c(\epsilon)$ is a positive constant that goes to 0 as ϵ goes to 0.*

Notice that this Lemma states that the weight $w^{-\frac{1}{2}}$ is close to $w_0^{-\frac{1}{2}}$ in the $L^2(w_0)$ norm of the cube Q.

Proof. : We want to estimate the expression:

$$\frac{1}{|Q|}\left\|w^{-\frac{1}{2}} - w_0^{-\frac{1}{2}}\right\|^2_{L^2(w_0,Q)} = \frac{1}{|Q|}\int_Q \frac{w_0}{w} + 1 - \frac{2}{|Q|}\int_Q \left(\frac{w_0}{w}\right)^{\frac{1}{2}}.$$

The last integral can be taken care of really easy, since by our normalization assumption and Cauchy-Schwartz we get the following:

$$\frac{1}{|Q|}\int_Q \left(\frac{w_0}{w}\right)^{\frac{1}{2}} = \frac{1}{|Q|}\int_Q \left(\frac{w}{w_0}\right)^{-\frac{1}{2}} \geq \left(\frac{1}{|Q|}\int_Q \left(\frac{w}{w_0}\right)^{\frac{1}{2}}\right)^{-1} \geq \left(\frac{1}{|Q|}\int_Q \frac{w}{w_0}\right)^{-\frac{1}{2}} = 1.$$

Therefore, the quantity that we need to estimate is bounded above by:

$$\frac{1}{|Q|}\left\|w^{-\frac{1}{2}} - w_0^{-\frac{1}{2}}\right\|^2_{L^2(w_0,Q)} \leq \frac{1}{|Q|}\int_Q \frac{w_0}{w} - 1.$$

It is time to use the fact that $d_*(w, w_0) \leq \epsilon$. We get that the weight $\frac{w}{w_0}$ is in the A_2 class and actually because the BMO norm of $\log\left(\frac{w}{w_0}\right)$ is really small, the A_2 characteristic is bounded by $1 + c(\epsilon)$, where $c(\epsilon)$ is a positive constant that goes to 0 as ϵ goes to 0. So we have the desired inequality:

$$\frac{1}{|Q|}\left\|w^{-\frac{1}{2}} - w_0^{-\frac{1}{2}}\right\|^2_{L^2(w_0,Q)} \leq \left[\frac{w}{w_0}\right]_{A_2} - 1 \leq c(\epsilon).$$

$\square$

Observe that the proof just presented can not be used to the case when the operator T is sub-linear. The linearity assumption for this proof is really essential.

The second proof covers all cases. Here we need not make any linearity assumptions for T. Our operator is going to be sub-linear. The main tool for the proof is the inequality

(proved earlier in this chapter)

$$\|T\|_{L^p(u)\to L^p(u)} \leq \|T\|_{L^p(v)\to L^p(v)}(1 + c_{[v]_{A_p}} d_*(u,v)), \tag{3.2}$$

that holds for all A_p weights $u, v \in A_p$ that are sufficiently close in the d_* metric, and for sublinear operators T that satisfy the assumptions of our Theorem. The positive constant $c_{[v]_{A_p}}$ that appears in the inequality depends on the dimension n, p, the function F and the A_p characteristic of the weight v. Since the quantities n, p, F are fixed we only write the subscript $c_{[v]_{A_p}}$ to emphasize this dependence on the characteristic.

Use inequality (3.2) with $u = w_0$ and $v = w$

$$\|T\|_{L^p(w_0)\to L^p(w_0)} \leq \|T\|_{L^p(w)\to L^p(w)}(1 + c_{[w]_{A_p}} d_*(w, w_0)).$$

At this point if we know that the constant $c_{[w]_{A_p}}$ remains bounded as the distance $d_*(w, w_0)$ goes to 0 we are done.

For this reason we assume that $d_*(w, w_0) = \delta$ is very close to 0. Then the function $\frac{w}{w_0}$ is an A_p weight with A_p characteristic very close to 1 (see [8]). How close depends only on δ, not on w. Thus, if R is large enough, the weight $(\frac{w}{w_0})^R \in A_p$, with A_p characteristic independent of w (again see [8]). Note that from the classical A_p theory, for sufficiently small $\epsilon > 0$, we have $w_0^{1+\epsilon} \in A_p$. Choose the numbers R, ϵ such that we have the relation $\frac{1}{R} + \frac{1}{1+\epsilon} = 1$, i.e. such that R and $R' = 1 + \epsilon$ are conjugate numbers. Then

$$< w >_Q < w^{-\frac{1}{p-1}} >_Q^{p-1} = \Big\langle \frac{w}{w_0} w_0 \Big\rangle_Q \Big\langle \Big(\frac{w}{w_0}\Big)^{-\frac{1}{p-1}} w_0^{-\frac{1}{p-1}} \Big\rangle_Q^{p-1}$$

and by Hölder's inequality it is less than or equal to

$$\Big\langle \Big(\frac{w}{w_0}\Big)^R \Big\rangle_Q^{\frac{1}{R}} \Big\langle w_0^{R'} \Big\rangle_Q^{\frac{1}{R'}} \Big\langle \Big(\frac{w}{w_0}\Big)^{-\frac{1}{p-1}\cdot R} \Big\rangle_Q^{\frac{p-1}{R}} \Big\langle w_0^{-\frac{1}{p-1}\cdot R'} \Big\rangle_Q^{\frac{p-1}{R'}}.$$

Separating the R-terms from the R'-terms and applying Hölder's inequality one more time we obtain that this is at most

$$\Big[\Big(\frac{w}{w_0}\Big)^R\Big]_{A_p}^{\frac{1}{R}} [w_0^{1+\epsilon}]_{A_p}^{\frac{1}{R'}} \le C,$$

where C is a constant independent of the weight w. Therefore, $[w]_{A_p} \le C$.

The last step is to remember how we obtained the constant $c_{[w]_{A_p}}$ that appears in inequality (3.2). We used the Riesz-Thorin interpolation theorem with change in measure and then expressed one of the terms that appears in our calculations as a Taylor series. The constant $c_{[w]_{A_p}}$ appears at exactly this point and it is not difficult to see that it depends continuously on $[w]_{A_p}$. Since this characteristic is bounded for w close to w_0 in the metric d_* we have that $c_{[w]_{A_p}}$ is bounded as well. This completes the proof. □

A consequence of the proof is the following remark.

Remark 9. Fix a weight $w_0 \in A_p$ and a positive number δ sufficiently small. There is a positive constant C that depends on $[w_0]_{A_p}$ and δ such that for all weights w with $d_*(w, w_0) < \delta$ we have $[w]_{A_p} \le C$. In addition, from the inequality (see the proof of Theorem 2)

$$[w]_{A_p} \le \Big[\Big(\frac{w}{w_0}\Big)^R\Big]_{A_p}^{\frac{1}{R}} [w_0^{1+\epsilon}]_{A_p}^{\frac{1}{R'}}, \tag{3.3}$$

and Lebesgue dominated convergence theorem (by letting $R \to +\infty$ and remembering that the A_p constant of the weight $(\frac{w}{w_0})^R$ is independent of R) we obtain

$$\limsup_{d_*(w,w_0)\to 0} [w]_{A_p} \leq [w_0]_{A_p}.$$

In order to get the remaining inequality

$$[w_0]_{A_p} \leq \liminf_{d_*(w,w_0)\to 0} [w]_{A_p},$$

we rewrite (3.3) as

$$[w_0]_{A_p} \leq \left[\left(\frac{w_0}{w}\right)^R\right]_{A_p}^{\frac{1}{R}} [w^{1+\epsilon}]_{A_p}^{\frac{1}{R'}},$$

and we proceed in the same way as before. In this case the number ϵ depends on $[w]_{A_p}$. But we already know that for w close to w_0 in the d_* metric the A_p characteristic of w is bounded from above. This means that we are allowed to choose the same number ϵ for all weights w that are sufficiently close to w_0 and we are done. Therefore, the A_p characteristic of a weight $w \in A_p$ is a continuous function of the weight with respect to the metric d_*, i.e. the following equality is true

$$\lim_{d_*(w,w_0)\to 0} [w]_{A_p} = [w_0]_{A_p}.$$

3.2 The sharp rate of convergence

In the following we are going to consider the Hilbert transform, H, the Riesz projection, P_+ and weights in A_2 on the circle. We are going to show that Theorem 2 is sharp for the Hilbert transform and that it is not sharp for the Riesz projection. This result is interesting because these two operators, H and P_+, are very closely related, i.e.

$$H = -iP_+ + i(I - P_+).$$

But as we shall see they do not behave in the same way.

We start with a weight $w \in A_2$, such that $[w]_{A_2} = 1 + \delta$, where $\delta > 0$ is really close to 0. We know that there exists an outer function h such that $w = |h|^2$. Outer means that $h = e^{u+i\widetilde{u}}$, where $\widetilde{u}$ denotes the harmonic conjugate of the function u. As we already have mentioned, $\log w = 2u$ is in $BMO(\mathbb{T})$ with norm $\|\log w\|_* \leq c\sqrt{\delta}$. This means that the conjugate function of u has small BMO norm, i.e. $\|\widetilde{u}\|_* \leq c\sqrt{\delta}$. From [15] the square root of δ is sharp. So we can choose our function u such that $c_1\sqrt{\delta} \leq \|u\|_*$. Observe also that $\frac{\overline{h}}{h} = e^{-2i\widetilde{u}}$. Let us now look at the operator $f \mapsto e^{if}$, that maps the space $BMO(\mathbb{T})$ continuously into itself (this is clear since if the oscillation of f is bounded then the same should be true for the function e^{if}). Of course, this is not a linear operator but it has some nice properties. For example, for $\epsilon > 0$ small, it maps the ball $B(0, \epsilon) = \{f \in BMO(\mathbb{T}) : \|f\|_* \leq \epsilon\}$ into another ball of center 0 and radius say $c\epsilon$. We claim that the ball $B(0, \epsilon)$ is mapped homeomorphically onto it's image, and that the image contains a ball $B(0, c\epsilon)$, for some c. For this it suffices to see that the derivative of this map at the point 0, is exactly the linear map $f \mapsto if$ which is a continuous surjection from $BMO(\mathbb{T})$ onto itself. Then make use of the inverse function theorem for Banach spaces. We did all this in order to be able to claim that we can choose our function h satisfying:

$$c_2\sqrt{\delta} \leq \left\|\frac{\overline{h}}{h}\right\|_* \leq c_1\sqrt{\delta}.$$

Let $f_{\pm}$ denote the analytic and anti-analytic parts of a bounded function f on the circle. Now the space $BMO(\mathbb{T})$ can be written as the direct sum of the $BMOA$ and $\overline{BMOA}$ spaces, the BMO analytic and the BMO anti-analytic spaces respectively. Without loss of generality we can assume that $c_1\sqrt{\delta} \geq \left\|\left(\frac{\overline{h}}{h}\right)_{-}\right\|_{\overline{BMOA}} \geq \frac{c_2\sqrt{\delta}}{2}$. But,

$$\left\|\left(\frac{\overline{h}}{h}\right)_{-}\right\|_{\overline{BMOA}} = dist\Big(\frac{\overline{h}}{h}, H^{\infty}\Big) = \sup_{\|\phi\|_1 \leq 1, \phi \in H_0^1} \Big| \int \frac{\overline{h}}{h}\phi \Big|$$

$$= \sup_{\|\phi_1\|_2, \|\phi_2\|_2 \leq 1, \phi_2(0)=0} \Big| \int \frac{\overline{h}}{h}\phi_1\phi_2 \Big|,$$

where $\|.\|_2$ is the norm in the Hardy space H^2. This last supremum is exactly equal to

$$\sup_{\|\phi_1\|_2, \|\phi_2\|_2 \leq 1, \phi_2(0)=0} (H_{\frac{\overline{h}}{h}}\phi_1, \overline{\phi_2}) = \|H_{\frac{\overline{h}}{h}}\|,$$

where $H_{\frac{\overline{h}}{h}} : H^2 \to H^2_{-}$ is the Hankel operator of symbol $\frac{\overline{h}}{h}$. Now consider the spaces:

$$H_{+} = clos_{L^2(w)}\{1, z, z^2, ...\}, H_{-} = clos_{L^2(w)}\{\overline{z}, \overline{z}^2, ...\}.$$

These spaces are called the future and the past spaces (the terminology comes from the probability, where w plays the role of the spectral density of a stationary stochastic process, see e.g. [31] and the literature cited therein).

The next step is to find the angle θ of these two spaces in $L^2(w)$. This is exactly

$$\sup_{\|\phi_-\|_{L^2(w)}=1, \|\phi_+\|_{L^2(w)}=1} \Big|\Big(\phi_+, \phi_-\Big)_{L^2(w)}\Big|.$$

If we write down just one of these inner products we see the following

$$\Big|\int \phi_+\overline{\phi_-}|h|^2\Big| = \Big|\int (\phi_+ h)(\overline{\phi_-} h)\frac{\overline{h}}{h}\Big|.$$

The first two functions that appear in the integrand are analytic since they are products of analytic functions. Note that since the function ϕ_- is anti-analytic, the function $\overline{\phi_-}$ is analytic. Also their H^2 norm is ≤ 1. This means that the supremum is exactly equal to

$$\frac{c_2\sqrt{\delta}}{2} \leq \|H_{\frac{\overline{h}}{h}}\| = \sup_{\|\phi_-\|_{L^2(w)}=1, \|\phi_+\|_{L^2(w)}=1} \Big|\Big(\phi_+, \phi_-\Big)_{L^2(w)}\Big| \leq c_1\sqrt{\delta}.$$

Therefore, the $\cos\theta$ is exactly of the order $\sqrt{\delta}$. This means that $\sin\theta - 1$ is of the order δ. Now, all that remains is an easy problem. We are given that the cosine of the angle of two directions is of the order $\sqrt{\delta}$ and we would like to find the order of $\sup \frac{\|u+v\|}{\|u-v\|}$ over all vectors u that have the first direction and v that have the second direction. Using the theorem of cosines we can see that the order of this supremum must be $1 + c\sqrt{\delta}$. Thus

$$\|H\|_{L^2(w)\to L^2(w)} \geq \sup \frac{\|u+v\|}{\|u-v\|} \asymp 1 + c\sqrt{\delta}$$

and

$$\|P_+\|_{L^2(w)\to L^2(w)} = \frac{1}{\sin\theta} \asymp 1 + c\delta.$$

This means that P_+ converges faster to its L^2 norm, as $[w]_{A_2} \to 1$, than the Hilbert transform. This should not be a surprise, since the multiplier that corresponds to P_+ takes only the values $\{0, 1\}$ and the multiplier for the Hilbert transform attains the values $\{-1, 1\}$. So the jump for P_+ is only 1 and for H is 2.

Chapter 4

Bellman functions and an application to Littlewood-Paley estimates

In this chapter we construct a new Bellman function based on the results of the previous chapter. It can be used to estimate the norms of second order Riesz transforms and to give a better understanding of some Littlwood-Paley estimates that first appeared in [26]. For instance, using our main Theorem 2 and techniques from [7], [20], [26], we can prove that for any $f, g \in C_c^\infty(\mathbb{R}^2)$ the quantity

$$2\int_0^{+\infty}\int_{\mathbb{R}^2}\Big(\Big|\frac{\partial f^h}{\partial x_1}\Big|+\Big|\frac{\partial f^h}{\partial x_2}\Big|\Big)^{\frac{1}{2}}\Big(\Big|\frac{\partial g^h}{\partial x_1}\Big|+\Big|\frac{\partial g^h}{\partial x_2}\Big|\Big)^{\frac{1}{2}}dydt$$

is bounded from above by

$$(p^*-1)(1+c\sqrt{\delta})\|f\|_{L^p(w)}\|g\|_{L^{p'}(w^{1-p'})},$$

for any A_p weight w on $\mathbb{R}^2$ with $[w]_{A_p} \le 1+\delta < 2$. The functions on the left hand side are the heat extensions of f, g respectively, and c is a constant that depends on $p \in (1, +\infty)$ and $p^* - 1 = \max\{p-1, \frac{1}{p-1}\}$. For example, for an f in $\mathbb{R}^d$ say, the heat extension to $\mathbb{R}^{d+1}$ is the convolution of f with the heat kernel, that is

$$f^h(x,t) = c_d \int_{\mathbb{R}^d} f(y) \exp\Big(-\frac{|x-y|^2}{4t}\Big) dy.$$

The main Theorem of this chapter is the following.

Theorem 10. *For any $1 < Q < 2$, $1 < p < +\infty$ define the domain $D_Q^p = \{0 < (X, Y, x, y, r, s) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R} : |x|^p < Xs^{p-1}, |y|^{p'} < Yr^{p'-1}, 1 < rs^{p-1} < Q\}$. Let K be any compact subset of D_Q^p. Then there exists a function $B = B_{Q,K}^{(p)}(X, Y, x, y, r, s)$ infinitely differentiable in a small neighborhood of K, and at the same time for any $\epsilon > 0$, $B_{Q,K}$ can be chosen in such a way that*

(1) $0 \le B \le (p^* - 1)(1+\epsilon)(1 + c\sqrt{\delta})X^{1/p}Y^{1/p'}$

(2) $-d^2B \ge 2|dx||dy|$,

where $Q = 1 + \delta$ and c is a constant that depends on p and the dimension d.

Proof. By Theorem 2 we know that for the martingale transform T_r and an A_p weight w, on $\mathbb{R}$, of characteristic $[w]_{A_p} < 1 + \delta < 2$,

$$\|T_r\|_{L^p(w) \to L^p(w)} \le \|T_r\|_{L^p \to L^p}(1 + c\sqrt{\delta}),$$

where c is a constant that depends on p. It is really easy to see that the interpolation in [29] works for the vectorized martingale transform T_r. This means that using the techniques from the previous chapter, the above inequality is also true for the vectorized martingale transform (acting on functions with values in a separable Hilbert space). But a famous result by Burkholder (see [3], and an extension of it in [7]), states that $\|T_r\|_{L^p \to L^p} = p^* - 1$.

Therefore,

$$\|T_r\|_{L^p(w)\to L^p(w)} \le (p^*-1)(1+c\sqrt{\delta}).$$

Now, by using duality we arrive to the point (we denote by $|\,.\,|$ the norm in our Hilbert space) that the expression

$$\frac{1}{4|J|}\sum_{I\in\mathcal{D}(J)} |<f>_{I_+} - <f>_{I_-}||<g>_{I_+} - <g>_{I_-}||I|$$

is bounded from above by

$$(p^*-1)(1+c\sqrt{\delta}) <|f|^p w>_J^{1/p} <|g|^{p'} w^{1-p'}>_J^{1/p'}$$

for any $J\in\mathcal{D}$, any vector functions $f\in L^p(w)$ and $g\in L^{p'}(w^{1-p'})$. The definition of the Bellman function is the following.

$$B(X,Y,x,y,r,s) = \sup\Big\{\frac{1}{4|J|}\sum_{I\in\mathcal{D}(J)} |<f>_{I_+} - <f>_{I_-}||<g>_{I_+} - <g>_{I_-}||I| :$$

$$<f>_J = x, <g>_J = y, <w>_J = r, <w^{1-p'}>_J = s,$$

$$<|f|^p w>_J = X, <|g|^{p'} w^{1-p'}>_J = Y\Big\}.$$

Obviously, this function satisfies inequality (1) in the statement of our Theorem and it does not depend on the choice of the interval J since averages of functions are translation invariant. We claim that for all 6-tuples $a^+ = (X^+, Y^+, x^+, y^+, r^+, s^+), a^- =$

$(X^-, Y^-, x^-, y^-, r^-, s^-) \in D_Q^p$, such that $\frac{a^+ + a^-}{2} \in D_Q^p$, the inequality is true

$$B\Big(\frac{a^+ + a^-}{2}\Big) - \frac{B(a^+) + B(a^-)}{2} \geq \frac{1}{4}|x^+ - x^-||y^+ - y^-|.$$

To prove this let us consider a positive ϵ. Find functions f^+, g^+, w^+ on J_+ such that they satisfy the conditions in the supremum of the function B for the vector a^+ and

$$B(a^+) - \epsilon \leq \frac{1}{|J_+|} \sum_{I \in \mathcal{D}(J_+)} | < f >_{I+}^{+} - < f >_{I-}^{+} || < g >_{I+}^{+} - < g >_{I-}^{+} ||I|.$$

Do the same for the vector a^- in the interval J_-. Define the functions F, G, W on the interval J as: $F = \begin{cases} f_+ & on\ J_+ \\ f_- & on\ J_- \end{cases}$ $G = \begin{cases} g_+ & on\ J_+ \\ g_- & on\ J_- \end{cases}$ and $W = \begin{cases} w_+ & on\ J_+ \\ w_- & on\ J_-. \end{cases}$

Observe that they satisfy the required equalities in order to be acceptable for the supremum that defines the Bellman function for the vector $\frac{a^+ + a^-}{2}$ and therefore,

$$B(\frac{a^+ + a^-}{2}) \geq \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} |F_{I+} - F_{I-}||G_{I+} - G_{I-}||I| =$$

$$\frac{1}{2|J_+|} \sum_{I \in \mathcal{D}(J_+)} |f_{I+}^+ - f_{I-}^+||g_{I+}^+ - g_{I-}^+||I| +$$

$$\frac{1}{2|J_-|} \sum_{I \in \mathcal{D}(J_-)} |f_{I+}^- - f_{I-}^-||g_{I+}^- - g_{I-}^-||I| + \frac{1}{|J|} term(I = J)$$

$$\geq \frac{1}{2}(B(a^+) - \epsilon) + \frac{1}{2}(B(a^-) - \epsilon) + \frac{1}{4}|x_+ - x_-||y_+ - y_-|$$

Now we need to mollify this function B, in order to take the smooth version of it. This can be done in exactly the same way as in [20]. The concavity inequality remains the same after the mollification and the size condition can become $1+C_K\epsilon$ times worse, where C_K is just a constant that depends on the compact set K. $\square$

For a nice application of Theorem 10, we can formulate the following result.

Theorem 11. *Let $1<p<+\infty$, and any scalar A_p weight w on $\mathbb{R}^d$ of $[w]_{A_p}<1+\delta<2$. Then*

$$\|\mathcal{R}\|_{L^p(\mathbb{R}^d,\mathbb{R}^d,wdx)\to L^p(\mathbb{R}^d,\mathbb{R}^d,wdx)} \le (p^*-1)(1+c\sqrt{\delta})\,,$$

where $\mathcal{R}=(R_iR_j)_{i,j=1}^d$, is a matrix with each entry a product of two Riesz transforms.

Observe that if we let δ go to 0, which means w becomes a constant weight

$$\|\mathcal{R}\|_{L^p(\mathbb{R}^d,\mathbb{R}^d)\to L^p(\mathbb{R}^d,\mathbb{R}^d)} \le (p^*-1).$$

Proof. We can show that for any A_p weight, w, on $\mathbb{R}^d$ of $[w]_{A_p}\le 1+\delta<2$, and any vector functions $\Phi=(\phi_1,...,\phi_d), \Psi=(\psi_1,...,\psi_d)\in C_c^\infty(\mathbb{R}^d)$ the quantity

$$2\int_{\mathbb{R}^{d+1}_+}\Big(\sum_{i,j=1}^d\Big|\frac{\partial\phi_j^h(x,t)}{\partial x_i}\Big|^2\Big)^{\frac12}\Big(\sum_{i,j=1}^d\Big|\frac{\partial\psi_j^h(x,t)}{\partial x_i}\Big|^2\Big)^{\frac12}dxdt$$

is bounded from above by

$$(p^*-1)(1+c\sqrt{\delta})\|\Phi\|_{L^p(w)}\|\Psi\|_{L^{p'}(w^{1-p'})}.$$

The proof of this inequality, follows the standard techniques appearing in [7], [20], [26], in which the existence of the Bellman function implies a Littlewood-Paley type estimate and it, in its turn, implies the desired estimate.

In addition, expressing the norm of $\mathcal{R}$ by duality we obtain (here $\Phi = (\phi_j)_{j=1}^d$, $\Psi = (\psi_i)_{i=1}^d$ are vector functions on $\mathbb{R}^d$):

$$< \mathcal{R}\Phi, \Psi >= 2\int_{\mathbb{R}_+^{d+1}} \sum_{i,j=1}^d \frac{\partial^2 \phi_j^h(x,t)}{\partial x_i \partial x_j} \psi_i^h(x,t) dx dt$$
$$= 2\int_{\mathbb{R}_+^{d+1}} \sum_{i,j=1}^d \frac{\partial \phi_j^h(x,t)}{\partial x_j} \frac{\partial \psi_i^h(x,t)}{\partial x_i} dx dt,$$

where we get the second equality because ϕ_j, ψ_i are smooth with compact support, and hence ϕ_j^h, ψ_i^h are Schwarz functions. Now, we only need to observe that:

$$\int_{\mathbb{R}_+^{d+1}} \sum_{i,j=1}^d \frac{\partial \phi_j^h(x,t)}{\partial x_j} \frac{\partial \psi_i^h(x,t)}{\partial x_i} dx dt = \int_{\mathbb{R}_+^{d+1}} \sum_{i=1}^d \Big(\sum_{j=1}^d \frac{\partial \phi_j^h(x,t)}{\partial x_i} \frac{\partial \psi_i^h(x,t)}{\partial x_j} \Big) dx dt$$

which in its turn is equal to

$$\int_{\mathbb{R}_+^{d+1}} trace \Big[\Big(\frac{\partial \phi_j^h(x,t)}{\partial x_i} \Big)_{i,j=1}^d \Big(\frac{\partial \psi_i^h(x,t)}{\partial x_j} \Big)_{i,j=1}^d \Big] dx dt,$$

and that on the other hand, point-wisely

$$\Big| trace \Big[\Big(\frac{\partial \phi_j^h(x,t)}{\partial x_i} \Big)_{i,j=1}^d \Big(\frac{\partial \psi_i^h(x,t)}{\partial x_j} \Big)_{i,j=1}^d \Big] \Big| \le \Big\| \Big(\frac{\partial \phi_j^h(x,t)}{\partial x_i} \Big)_{i,j=1}^d \Big\|_2 \Big\| \Big(\frac{\partial \psi_i^h(x,t)}{\partial x_j} \Big)_{i,j=1}^d \Big\|_2.$$

This means we are done. $\square$

Chapter 5

Matrix weights and the A_2 condition

In the previous chapters we considered only scalar weights w and studied the behavior of the operator norm of an operator T on $L^p(w)$, $1 < p < \infty$, with respect to the weight w. We treated all operators at the same time meaning that the exact same proof works for all of them. The only property that we required from the operator was that it is strongly bounded on $L^p(w)$ and that its operator norm depends only on the A_p characteristic of the weight. Weighted estimates for matrix valued weights W have also been studied in the literature and one of the main references is the paper [30] by Dr. Treil and Dr. Volberg. They considered L^1_{loc} matrices $W \in \mathbb{C}^{d\times d}$ that are invertible, self-adjoint and positive almost everywhere with respect to Lebesgue measure. One of their main results is the characterization of all matrices W with the property that the inequality

$$\Big(\int_{\mathbb{R}}(W(x)Hf(x), Hf(x))_{\mathbb{C}^d}dx\Big)^{\frac{1}{2}} \leq C\Big(\int_{\mathbb{R}}(W(x)f(x), f(x))_{\mathbb{C}^d}dx\Big)^{\frac{1}{2}},$$

holds for all $f \in L^2(W)$, for some positive constant C that depends on the dimension d and the weight W, and H is the Hilbert transform that acts coordinate-wise on the vector function f. The space $L^2(W)$ consists of all measurable functions $f : \mathbb{R} \to \mathbb{C}^d$ such that

$$\|f\|^2_{L^2(W)} = \|f\|^2_{2,W} = \int_{\mathbb{R}}(W(x)f(x), f(x))_{\mathbb{C}^d}dx < \infty.$$

Their Theorem states that the class of such weights is the matrix A_2 class that consists of W that satisfy

$$[W]_{A_2} = \sup_I \Big\| < W >_I^{\frac{1}{2}} < W^{-1} >_I^{\frac{1}{2}} \Big\| < \infty,$$

where the supremum is taken over all finite intervals I of the real line $\mathbb{R}$, and the quantities $< W >_I, < W^{-1} >_I$ are used to denote the averages of W and W^{-1} over the interval I respectively. Notice that this characteristic for the matrix weight W is a generalization of the scalar one and that the former is the square root of the latter. From now on in this paper, wherever we write the symbol for the A_2 characteristic we mean the one given for matrix weights. Throughout the paper we always assume that the weight W is non-degenerate in the sense that there is no vector $e \in \mathbb{C}^d$ such that $W(t)e = 0$ almost everywhere, because otherwise we can always restrict ourselves to the orthogonal complement of such e. We have to point out that the A_2 condition just stated is equivalent to the requirement that

$$< W^{-1} >_I \leq [W]_{A_2}^2 < W >_I^{-1},$$

in the sense of quadratic forms. In addition, it is equivalent to the statement that all averaging operators

$$f \mapsto \Big(\frac{1}{|I|} \int_I f(x) dx \Big) \chi_I,$$

are uniformly bounded in $L^2(W)$ with respect to the finite interval I. For this reason if we consider any direction $e \in \mathbb{C}^d$, $\|e\| = 1$, we see that the scalar weight $w_e(x) = (W(x)e, e)_{\mathbb{C}^d}$ is an A_2 weight of characteristic at most $[W]_{A_2}$. Immediately we obtain that the diagonal elements of W are scalar A_2 weights and that the weight $trace(W)$ is also an A_2 weight of characteristic at most $d \cdot [W]_{A_2}$.

The motivation of studying estimates of this type comes from stochastic processes and operator theory. Let us consider a multivariate random stationary process. For simplicity we consider the case of discrete time i.e. a sequence of d-tuples $x(n) = (x_1(n), ..., x_d(n))$, $n \in \mathbb{Z}$, of scalar random variables such that $\mathbb{E}|x_j(n)|^2 < +\infty$ and the correlation matrix

$$Q(n,k) = \{Q(n,k)_{i,j}\}_{1\le i,j\le d} := \{\mathbb{E}x_i(n)\overline{x_j(n)}\}_{1\le i,j\le d},$$

depends only on the difference $n - k$ (we use the symbol $\mathbb{E}$ to denote the expectation). Without loss of generality we can assume that the process is complex valued. It is well known (see [27]) that there exists a matrix valued non-negative measure M on the unit circle $\mathbb{T}$ whose Fourier coefficients coincide with the entries of the correlation matrix

$$Q(n,k) = \widehat{M}(n,k),$$

$n, k \in \mathbb{Z}$ and that if the process is completely regular then its spectral measure, M, is absolutely continuous with respect to the normalized Lebesgue measure m on the unit circle, i.e. $dM = W dm$. The past of the process is defined as

$$\mathcal{X}_n = span\{x_j(k) : 1 \le j \le d, k < n\}$$

and the future as

$$\mathcal{X}^n = span\{x_j(k) : 1 \le j \le d, k \ge n\}.$$

By writing span we mean the closed linear span in the complex Hilbert space $L^2(\Omega, dP)$. If we consider the mapping

$$x_j(k) \mapsto z^k e_j,$$

where $\{e_j\}_{1 \le j \le d}$ is the standard orthonormal basis of $\mathbb{C}^d$, then we obtain an isometric isomorphism between $span\{x_j(k) : 1 \le j \le d, k \in \mathbb{Z}\}$ and $L^2(W)$. The past and the future of the process are mapped to the subspaces of $L^2(W)$

$$X_n = span\{z^k \mathbb{C}^d : k < n\}$$

and

$$X^n = span\{z^k \mathbb{C}^d : k \ge n\},$$

respectively. In this representation the angle between past and future is nonzero if and only if the Riesz projection P_+ is bounded in the weighted space $L^2(W)$. All these applications are thoroughly discussed in the introduction of [30] and the references therein.

In this chapter we are going to study the behavior of the operator norm of some dyadic operators on $L^2(W)$ with respect to a matrix weight W. As we shall see there are many important differences between the scalar and the matrix cases. We will prove that for a dimensional analogue of the Martingale transform the operator norm on $L^2(W)$ does not approach the unweighted norm as the matrix weight W "approaches" the identity matrix Id. This already is in contrast with the scalar case where such thing can not happen as we showed in the previous chapters. It seems that as we consider more than one dimensions the flatness (meaning closeness to 1) of the A_2 characteristic does not suffice for continuity results of the kind of Theorem 2. It is also interesting and surprising that trivial dyadic operators are examples of such not well behaved operators. Here we should mention that none of the techniques we used in the scalar case work for the matrix case. Firstly, the Riesz-Thorin interpolation theorem with change in measure does not work and secondly, there is no useful BMO theory for the case of matrices. Useful in the sense that there is a nice interplay between BMO and the A_2 class.

5.1 The not well behaved dyadic operators

Before we study such examples we need to establish some notation. For a function f (scalar or matrix valued) and a finite interval I we denote by $< f >_I$ the average of f over the interval I, that is the number $\frac{1}{|I|}\int_I f(x)dx$. For a given interval I we denote the right half as I_+ and the left half as I_-. For such interval there is a Haar function associated to it, which we call h_I, defined in the following way

$$h_I(x) = \frac{1}{\sqrt{|I|}}(\chi_{I_+}(x) - \chi_{I_-}(x)),$$

where $\chi_A(x)$ represents the characteristic function of the set A. It is obvious that if we restrict ourselves to dyadic subintervals, $I, J \in \mathcal{D}$, of the real line we have

$$(h_I, h_J)_{L^2} := \int_{\mathbb{R}} h_I(x)h_J(x)dx = \delta_{IJ},$$

where δ_{IJ} is equal to 1 if $I = J$ and equal to 0 if $I \neq J$. By $\mathcal{D}$ we denote the set $\mathcal{D} = \cup_{k=-\infty}^{+\infty}\mathcal{D}_k$, where $\mathcal{D}_k = \{[\frac{j}{2^k}, \frac{j+1}{2^k}) : j \in \mathbb{Z}\}$. In addition, for an interval $J \in \mathcal{D}$ we are going to denote by $\mathcal{D}(J)$ the set of all dyadic subintervals of J, including J itself.

Given a sequence of signs enumerated by the dyadic intervals $\sigma = \{\sigma_I\}_{I\in\mathcal{D}}$, $\sigma_I \in$

$\{-1, +1\}$, we define the martingale transform T_σ of a function f to be

$$T_\sigma f(x) = \sum_{I \in \mathcal{D}} \sigma_I (f, h_I)_{L^2} h_I(x).$$

The boundedness properties of this operator on $L^2(w)$ for a scalar A_2 weight w were studied in [33], where the sharp dependence of the operator norm and the A_2 characteristic was found for the first time. For a function $f : \mathbb{R} \to \mathbb{C}^d$ we can define the Martingale transform $T_\sigma f$ to be the vector in $\mathbb{C}^d$ with coordinates the numbers $(T_\sigma f_1, T_\sigma f_2, ..., T_\sigma f_d)$. For our purposes we will define a more general operator than this which we still call the Martingale transform of the function f defined as

$$M_\sigma f(x) = \sum_{\substack{I \in \mathcal{D} \\ 1 \le j \le d}} \sigma_I^j (f, h_I^j)_{L^2} h_I^j(x),$$

where the function $h_I^j(x)$ is the vector $h_I e_j$ and $\{e_j\}_{1 \le j \le d}$ is a fixed orthonormal basis of $\mathbb{C}^d$. Here the sequence $\sigma = \{\sigma_I^j\}$, $I \in \mathcal{D}$ and $1 \le j \le d$, is again a sequence of signs.

Let us also define the projections $P_{I,j}$ (that are orthogonal in the unweighted L^2 space) in $L^2(W)$ by the formula

$$P_{I,j} f = h_I (\int_I f(t) h_I(t) dt, e_j)_{\mathbb{C}^d} e_j.$$

Notice that $P_{I,j} f = (f, W^{-1} h_I^j)_{2,W} h_I^j$, where we denote by $(,)_{2,W}$ the inner product in $L^2(W)$, from which it follows

$$\|P_{I,j}\|_{2,W}^2 = \|W^{-1} h_I^j\|_{2,W}^2 \|h_I^j\|_{2,W}^2.$$

Observe that after some easy calculations we obtain

$$\|P_{I,j}\|_{2,W}^2 = (< W >_I e_j, e_j)_{\mathbb{C}^d}(< W^{-1} >_I e_j, e_j)_{\mathbb{C}^d}.$$

The claim is that such quantity can not be controlled by the A_2 characteristic. Let us choose a matrix weight $W \in \mathbb{C}^{2\times 2}$ and an orthonormal basis $\{e_j\}_{1\leq j\leq d}$ in $\mathbb{C}^2$ such that the operator norm of $P_{I,j}$ is not close to its unweighted norm, which is 1, no matter how "close" W is to the identity matrix Id. Assume that one of the bases vectors is $e_1 = \frac{1}{\sqrt{2}}(1,1)$ and that W is a diagonal 2×2 matrix with two scalar A_2 weights w and v for diagonal elements. Let it be that w is in the $1,1$ spot and v in the $2,2$ spot. Then

$$\begin{aligned} \|P_{I,1}\|_{2,W}^2 &= \frac{1}{4}(< w >_I + < v >_I)(< w^{-1} >_I + < v^{-1} >_I) \\ &\geq \frac{1}{4}(2+ < w >_I < v^{-1} >_I + < w^{-1} >_I < v >_I) \\ &\geq \frac{1}{4}\Big(2 + \frac{< w >_I}{< v >_I} + \frac{< v >_I}{< w >_I}\Big), \end{aligned}$$

and the weights w, v have no relation with each other. Both of them can have A_2 characteristic close to 1, as close as we like, but the quotients $\frac{<w>_I}{<v>_I}$ and $\frac{<v>_I}{<w>_I}$ can not be controlled in general. This means that even though the matrix weight W has A_2 characteristic close to 1 the norm $\|P_{I,1}\|_{2,W}$ is not close to $\|P_{I,1}\|_{2,Id} = 1$. Here notice that the Martingale transform M_σ can be written in the form

$$M_\sigma f = \sum_{\substack{I\in\mathcal{D}\\ 1\leq j\leq d}} \sigma_I^j (f, W^{-1}h_I^j)_{L^2(W)} h_I^j(x) = \sum_{\substack{I\in\mathcal{D}\\ 1\leq j\leq d}} \sigma_I^j P_{I,j} f.$$

Since the operator norms of the projections $P_{I,j}$ in $L^2(W)$ are not continuous with respect to the weight W we can not expect this Martingale transform to have an operator norm in $L^2(W)$ that is continuous with respect to W. Now we define the projection

$$P_I f = \sum_{1 \leq j \leq d} P_{I,j} f = h_I \Big(\int_I f(t) h_I(t) dt \Big).$$

In [30] it has been proved that the operator norm in $L^2(W)$ is exactly equal to

$$\|P_I\|_{2,W} = \Big\| < W >_I^{\frac{1}{2}} < W^{-1} >_I^{\frac{1}{2}} \Big\|.$$

This expression is obviously less than or equal to $[W]_{A_2}$ which immediately shows that P_I behaves nicely compared to its component operators that are the ones who do not. Here we have a collection of projections $\{P_{I,j}\}_{1 \leq j \leq d}$ such that some of them do not become "flat" as the weight W becomes "flat" but their sum is "flat". Therefore, we already see that strange things can occur in more than one dimensions.

5.2 Some results about "flatness" and a Riesz basis

In this section we will discuss some results which give us hope that there are important quantities of dyadic harmonic analysis that obey the same rules as their scalar dimensional analogues. We also present an important example of a Riesz basis for $L^2(W)$. All of them were proved in [30] but we present them here to show that the dependence on the A_2 characteristic is the "correct" one and because we need them for our calculations.

We start with a Lemma.

Lemma 12. *Let A and B be nonsingular positive $d \times d$ matrices. Then:*

$$\sqrt{\det A \det B} \leq \det\Big(\frac{A+B}{2}\Big).$$

Proof. It suffices to prove the Lemma in the special case when $\frac{A+B}{2} = I$, since we can always consider the matrices C^*AC, C^*BC where the matrix $C = \left(\frac{A+B}{2}\right)^{-\frac{1}{2}}$. Write $A = I + D$, $B = I - D$, $D = D^*$, and let $\lambda_1, ...\lambda_d$ be the eigenvalues of D. Then the eigenvalues of A, B are $1 + \lambda_1, ..., 1 + \lambda_d$ and $1 - \lambda_1, ..., 1 - \lambda_d$, respectively. It follows that:

$$\det(AB) = \det A \det B = \prod_{i=1}^{d}(1+\lambda_i)(1-\lambda_i) \leq 1 = \Big[\det\Big(\frac{A+B}{2}\Big)\Big]^{\frac{1}{2}},$$

which is exactly what we need. $\square$

Lemma 13. *Let W be a matrix weight such that W and W^{-1} are summable on a measurable set I. Then for any vector $e \in \mathbb{C}^d$*

$$\frac{(< W >_I e, e)_{\mathbb{C}^d}}{([< W^{-1} >_I]^{-1} e, e)_{\mathbb{C}^d}} \geq 1.$$

Moreover, the operators $< W >_I^{\frac{1}{2}} < W^{-1} >_I^{\frac{1}{2}}$ are expanding in the sense that they satisfy $\| < W >_I^{\frac{1}{2}} < W^{-1} >_I^{\frac{1}{2}} e\| \geq \|e\|$ for all vectors $e \in \mathbb{C}^d$.

Proof. Fix a vector e and define $f = [W^{-1}(I)]^{-1}e$. Then

$$\begin{aligned}
|I|([W^{-1}(I)]^{-1}e, e)_{\mathbb{C}^d} &= |I|(e, f)_{\mathbb{C}^d} = \int_I (W^{\frac{1}{2}}(t)e, W^{-\frac{1}{2}}(t)f)_{\mathbb{C}^d} dt \\
&\leq \Big(\int_I (W(t)e, e)_{\mathbb{C}^d} dt\Big)^{\frac{1}{2}} \Big(\int_I (W^{-1}(t)f, f)_{\mathbb{C}^d} dt\Big)^{\frac{1}{2}} \\
&= (W(I)e, e)_{\mathbb{C}^d}^{\frac{1}{2}} (W^{-1}(I)f, f)_{\mathbb{C}^d}^{\frac{1}{2}}
\end{aligned}$$

$$= (W(I)e,e)^{\frac{1}{2}}_{\mathbb{C}^d}([W^{-1}(I)]^{-1}e,e)^{\frac{1}{2}}_{\mathbb{C}^d}.$$

Thus,

$$|I|^2 \leq \frac{(W(I),e,e)_{\mathbb{C}^d}}{([W^{-1}(I)]^{-1}e,e)_{\mathbb{C}^d}},$$

which is exactly what we wanted to prove. □

With the help of these two Lemmas we are able to show the following.

Lemma 14. *Let us consider a matrix weight $W \in A_2$. There is a constant c independent of the weight W such that for all $J \in \mathcal{D}$:*

$$\frac{1}{|J|}\sum_{I\in\mathcal{D}(J)} trace\Big((W_I)^{-\frac{1}{2}}(W_{I_+} - W_{I_-})(W_I)^{-\frac{1}{2}}\Big)^2|I| \leq c\log[W]_{A_2},$$

and

$$\frac{1}{|J|}\sum_{I\in\mathcal{D}(J)} \Big\|[(W_I)^{-\frac{1}{2}}(W_{I_+} - W_{I_-})(W_I)^{-\frac{1}{2}}]\Big\|^2|I| \leq c\log[W]_{A_2}.$$

Proof. For a dyadic interval I let us denote by $\mu(I) = \det W_I, \nu(I) = \det(W^{-1})_I$ and $m(I) = \mu(I)\nu(I)$. Since $W_I = \frac{W_{I_+}+W_{I_-}}{2}$ Lemma 12 implies that $\mu(I)^2 \geq \mu(I_+)\mu(I_-)$ and similarly $\nu(I)^2 \geq \nu(I_+)\nu(I_-)$. Also, we define the matrices:

$$A = (W_I)^{-\frac{1}{2}}W_{I_+}(W_I)^{-\frac{1}{2}}, B = (W_I)^{-\frac{1}{2}}W_{I_-}(W_I)^{-\frac{1}{2}}.$$

Observe that $\frac{A+B}{2} = I$ and as it was done in the proof of the previous Lemma, we write $A = I + D$, $B = I - D$ and let $\lambda_1, ..., \lambda_d$ be the eigenvalues of D. Then:

$$\det A \det B = \prod_{i=1}^{d}(1-\lambda_i^2) = \exp\Big(\sum_{i=1}^{d}\log(1-\lambda_i^2)\Big)$$

$$
\begin{aligned}
&\leq \exp\Big(-\sum_{i=1}^{d}\lambda_i\Big)\\
&= \exp(-trace(D^2))\\
&= \exp\Big(-\frac{1}{4}trace((A-B)^2)\Big).
\end{aligned}
$$

So we have proved that

$$\mu(I) \geq (\mu(I_+)\mu(I_-))^{\frac{1}{2}} \exp\Big(\frac{1}{2}\cdot\frac{1}{4}trace([(W_I)^{-\frac{1}{2}}(W_{I_+}-W_{I_-})(W_I)^{-\frac{1}{2}}]^2)\Big).$$

But $\nu(I)^2 \geq \nu(I_+)\nu(I_-)$ and this implies

$$m(I) \geq (m(I_+)m(I_-))^{\frac{1}{2}} \exp\Big(\frac{1}{2}\cdot\frac{1}{4}trace([W_I)^{-\frac{1}{2}}(W_{I_+}-W_{I_-})(W_I)^{-\frac{1}{2}}]^2)\Big).$$

Applying this last inequality to I_+, I_- and then to the halves of these intervals we get on the nth step

$$m(I) \geq \Big[\prod m(J)\Big]^{\frac{1}{2^n}} \exp\Big(\frac{1}{8}\sum\frac{|J|}{|I|}\cdot trace([(W_I)^{-\frac{1}{2}}(W_{I_+}-W_{I_-})(W_I)^{-\frac{1}{2}}]^2)\Big),$$

where the product is over all subintervals, I, of J of length $|I| = |J|2^{-n}$ and the summation is over all subintervals, I, of J of length $|I| > |J|2^{-n}$. We know that the operators $(W_I)^{\frac{1}{2}}((W^{-1})_I)^{\frac{1}{2}}$ are expanding and this implies that $m(I) \geq 1$. We take logarithms, let n go to infinity and obtain the inequality

$$\frac{1}{|J|}\sum_{I\in\mathcal{D}(J)} trace([(W_I)^{-\frac{1}{2}}(W_{I_+}-W_{I_-})(W_I)^{-\frac{1}{2}}]^2)|I| \leq 8\log\sup_{I_0}[\det(W_{I_0})\det((W^{-1})_{I_0})].$$

The right hand side of this inequality is less than or equal to the quantity

$$8\log([W]_{A_2}^{2d}) = 16d\log[W]_{A_2}.$$

Therefore, we have shown that

$$\frac{1}{|J|}\sum_{I\in\mathcal{D}(J)} trace([(W_I)^{-\frac{1}{2}}(W_{I_+} - W_{I_-})(W_I)^{-\frac{1}{2}}]^2)|I| \le 16d\log[W]_{A_2}.$$

Notice that the matrix $(W_I)^{-\frac{1}{2}}(W_{I_+} - W_{I_-})(W_I)^{-\frac{1}{2}}$ is self-adjoint which implies that we have the same estimate with the square of the operator norm of the matrix in the place of trace. □

The following Lemma is a type of a weighted Carleson embedding theorem.

Lemma 15. *Let $W \in \mathbb{C}^{d\times d}$ be an A_2 matrix weight and consider the quantity*

$$\mu_I = |I|\Big\| < W >_I^{-\frac{1}{2}} (< W >_{I_+} - < W >_{I_-}) < W >_I^{-\frac{1}{2}} \Big\|^2.$$

Then there is a positive dimensional constant c such that

$$\sum_{I\in\mathcal{D}} \mu_I \| < W >_I^{-\frac{1}{2}} < W^{\frac{1}{2}} f >_I \|^2 \le c[W]_{A_2}^2 \log[W]_{A_2} \|f\|_2^2,$$

holds for all $f \in L^2(\mathbb{R} \to \mathbb{C}^d)$.

This was proved in [30] but the authors were not interested in the dependence of the inequality with respect to the A_2 characteristic of the weight. If we just simply follow their proof we are able to obtain the square of $[W]_{A_2}$ and the logarithm with the use of Lemma

14.

Suppose now that we have a collection of subspaces E_n of a Hilbert space H. We assume that the only vector perpendicular to every E_n is the zero vector. We will call such collections complete.

The collection is called minimal if there is a family of bounded projections (not necessarily orthogonal) $\mathcal{E}_n$

$$\mathcal{E}_n = Id \cdot \chi_{E_n},$$

and is called uniformly minimal if

$$\sup_{n\in\mathbb{N}} \|\mathcal{E}_n\|_{H\to H} < \infty.$$

For a minimal collection of subspaces E_n we can define the bi-orthogonal or dual system by

$$E'_n = (\mathcal{E}_n)^*(H) = span\{E_k : k \neq n\}^{\perp},$$

where $(\mathcal{E}_n)^*$ denotes the dual operator of $\mathcal{E}_n$.

A complete system of subspaces E_n is called an unconditional basis if there exists an isomorphism U from H onto another Hilbert space H' that maps the collection E_n into an orthogonal system. Such an isomorphism is called the orthogonalizer of the collection. An equivalent statement is that there exists a constant $C > 0$ such that for any finite collection of vectors $f_n \in E_n$

$$\frac{1}{C}\sum \|f_k\|_H^2 \leq \Big\| \sum f_k \Big\|_H^2 \leq C \sum \|f_k\|_H^2.$$

Notice that if the subspaces E_n were orthogonal then we would have that these quantities are equal by the Pythagorean theorem. Instead of that we have that they are comparable.

A collection of vectors f_n is called an unconditional basis if the corresponding system of one dimensional spaces is an unconditional basis and we call the collection of vectors a Riesz basis if it is almost normalized, that is

$$0 < \inf_{n \in \mathbb{N}} \|f_n\|_H \leq \sup_{n \in \mathbb{N}} \|f_n\|_H < \infty.$$

Notice that we do not allow the vectors f_n to be arbitrarily large or arbitrarily small inside the Hilbert space H.

The following is a very important result of [21].

Theorem 16. *A complete collection of subspaces E_n of a Hilbert space H is an unconditional basis if and only if it is uniformly minimal and the following two conditions hold for some positive constant C independent of f*

$$\sum_n \|P_{E_n} f\|_H^2 \leq C\|f\|_H^2$$

and

$$\sum_n \|P_{E_n'} f\|_H^2 \leq C\|f\|_H^2,$$

for all $f \in H$, where by P_{E_n} and $P_{E_n'}$ we denote the orthogonal projections onto E_n and E_n' respectively. .

Our Hilbert space now is going to be $L^2(W)$ where $W \in \mathbb{C}^{d \times d}$ is an A_2 matrix weight. A construction of a Riesz basis in $L^2(W)$ was done in [30]. We need this basis to define a Martingale transform that its operator norm on $L^2(W)$ is going to be a continuous function of the weight W. For this reason let us see how this construction was done.

Denote by e_I^k, $1 \leq k \leq d$, an orthonormal basis of $\mathbb{C}^d$, consisting of eigenvectors of the positive self-adjoint matrix $< W >_I$ and let

$$\begin{aligned} w_I^k &= \Big(\frac{1}{|I|}\int_I (W(t)e_I^k, e_I^k)_{\mathbb{C}^d} dt\Big)^{-\frac{1}{2}} = (< W >_I e_I^k, e_I^k)_{\mathbb{C}^d}^{-\frac{1}{2}} \\ &= ([W_I]^{-1} e_I^k, e_I^k)_{\mathbb{C}^d}^{\frac{1}{2}} = \| < W >_I^{-\frac{1}{2}} e_I^k \|. \end{aligned}$$

Define the vectors

$$f_I^k(x) = w_I^k h_I(x) e_I^k.$$

Observe that

$$(f_I^k, f_J^{k'})_{L^2} = w_I^k w_J^{k'} (e_I^k, e_J^{k'})_{\mathbb{C}^d} \int_{\mathbb{R}} h_I(x) h_J(x) dx$$

which is equal to zero if $I \neq J$ since $h_I \perp h_J$ in L^2 and if $k \neq k'$ and $I = J$ it is again equal to zero since the vectors $e_I^k, e_I^{k'}$ are orthogonal in $\mathbb{C}^d$. This means that the vectors $\{f_I^k\}$ are orthogonal in the unweighted L^2 space.

In the $L^2(W)$ space similar things happen but the situation is slightly different. For instance,

$$\begin{aligned} (f_I^k, f_I^k)_{L^2(W)} = (w_I^k)^2 \int_{\mathbb{R}} h_I^2(x)(W(x)e_I^k, e_I^k)_{\mathbb{C}^d} dx &= (w_I^k)^2 (< W >_I e_I^k, e_I^k)_{\mathbb{C}^d} \\ &= 1 \end{aligned}$$

and for $k \neq k'$ we have that $(f_I^k, f_I^{k'})_{L^2(W)} = 0$ since the vectors $e_I^k, e_I^{k'}$ are orthogonal in $\mathbb{C}^d$. But for $I \neq J$ we do not have orthogonality in general. The reason for that is that the vectors e_I^k and $e_J^{k'}$ for $I \neq J$ have no relation for an arbitrary matrix weight W. This is a difficulty that we are able to overcome easily since we can prove that for W with "flat" A_2 characteristic these vectors are almost orthogonal. We will make this more precise later.

Now let us define a collection of spaces E_I as

$$E_I = span\{f_I^k : 1 \leq k \leq d\} = h_I \mathbb{C}^d,$$

$I \in \mathcal{D}$. The vectors $\{f_I^k\}_{1 \leq k \leq d}$ constitute an orthonormal basis of E_I inside $L^2(W)$. Here notice that from our previous considerations it follows that the subspaces E_I and E_J are orthogonal in the unweighted L^2 space for $I \neq J$. The E_I's are d-dimensional subspaces of $L^2(W) \cap L^2$. It is east to prove that if a vector function f is orthogonal to all E_I's then $f = 0$ almost everywhere. This means that our collection $\{E_I\}_{I \in \mathcal{D}}$ is a complete system of subspaces.

Let us define the projections

$$P_I f(x) = h_I(x) \Big(\int_I f(t) h_I(t) dt \Big).$$

Notice that we considered these projections before in the example of the not well behaved Martingale transform. Also, $\|P_I\|_2 = 1$ and

$$\|P_I\|_{2,W} = \Big\| < W >_I^{\frac{1}{2}} < W^{-1} >_I^{\frac{1}{2}} \Big\|.$$

That is, these projections are orthogonal in L^2 but they are almost orthogonal in $L^2(W)$ for "flat" $W \in A_2$. In addition, inside $L^2(W)$ (and L^2) we have the equality

$$P_I = Id \cdot \chi_{E_I},$$

which implies that our collection $\{E_I\}_{I\in\mathcal{D}}$ is minimal. Actually,

$$\sup_{I\in\mathcal{D}} \|P_I\|_{2,W} \leq [W]_{A_2} < \infty,$$

and so the collection is uniformly minimal. Let us denote by E'_I the bi-orthogonal system and by P_{E_I} the orthogonal projection onto E_I and by $P_{E'_I}$ the orthogonal projection onto E'_I . Using techniques from [30] and Lemma 15 we can show that the following is true.

Theorem 17. *There is a positive dimensional constant C with the property*

$$\sum_{I\in\mathcal{D}} \|P_{E_I} f\|_{2,W}^2 \leq (1 + C\sqrt{\log[W]_{A_2}})\|f\|_{2,W}^2,$$

for all $f \in L^2(W)$ and

$$\sum_{I\in\mathcal{D}} \|P_{E'_I} g\|_{2,W}^2 \leq (1 + C\sqrt{\log[W]_{A_2}})\|g\|_{2,W}^2,$$

for all $g \in L^2(W)$.

Notice that this proves that the uniformly minimal collection $\{E_I\}_{I\in\mathcal{D}}$ is actually an unconditional basis in $L^2(W)$ and that the vectors f_I^k are a Riesz basis in $L^2(W)$. Here is a short outline of the proof of Theorem 17. What we want to show is that

$$\sum_{\substack{I\in\mathcal{D}\\ 1\leq k\leq d}} |(f, f_I^k)_{2,W}|^2 \leq (1 + C\sqrt{\log[W]_{A_2}})\|f\|_{2,W}^2.$$

For this reason we define the vectors $g_I^k = f_I^k + \chi_I A_I e_I^k$, where $A_I = \frac{1}{2}|I|^{-\frac{1}{2}} < W >_I^{-1} (< W >_{I_+} - < W >_{I_-}) < W >_I^{-\frac{1}{2}}$. Since the collection $\{g_I^k\}$ is orthogonal in $L^2(W)$ and since by Bessel's inequality (the norms $\|g_I^k\|_{2,W}$ are uniformly bounded because $\sup_{I,k}\|g_I^k\|_{L^2(W)} \leq 1 + c\sqrt{\log[W]_{A_2}}$)

$$\sum_{\substack{I\in\mathcal{D}\\1\leq k\leq d}} \frac{1}{\|g_I^k\|_{2,W}^2}|(f,g_I^k)_{2,W}|^2 \leq \|f\|_{2,W}^2,$$

it suffices to prove

$$\sum_{\substack{I\in\mathcal{D}\\1\leq k\leq d}} |(f,\chi_I A_I e_I^k)_{2,W}|^2 \leq (1+C\sqrt{\log[W]_{A_2}})\|f\|_{2,W}^2,$$

for all $f \in L^2(W)$. This statement is equivalent to

$$\sum_{\substack{I\in\mathcal{D}\\1\leq k\leq d}} |I|^2\|A_I^*(W^{\frac{1}{2}}f)_I\|^2 \leq (1+C\sqrt{\log[W]_{A_2}})\|f\|_2^2,$$

for all $f \in L^2$. But $\|A_I^* e\| \leq \frac{1}{2}(\mu_I)^{\frac{1}{2}}|I|^{-1}\|(W_I)^{-\frac{1}{2}}e\|$, for all vectors e, where μ_I is the quantity that appears in Lemma 15. This means we are done. Before we go further and study the Martingale transform in the next section, note that the system of subspaces E_I in $L^2(W)$ has the same geometry as the system $W^{\frac{1}{2}}E_I$ in the unweighted L^2 space. The bi-orthogonal to the latter system is $W^{-\frac{1}{2}}E_I$.

5.3 The Martingale transform revisited: M_σ and M_σ^W

In this section we will try to use the construction of the Riesz basis which was presented before to prove that a new Martingale transform, M_σ^W, which depends on the matrix weight W is "flat" for "flat" A_2 weight W. We already know that we can not expect the usual Martingale transform M_σ to behave nicely for a general A_2 weight. This is because the

$L^2(W)$ operator norms of the projections $P_{I,j}$ that we considered in section 5.1 are not "flat" in general for "flat" A_2 matrix weight W. Let us fix such an A_2 weight W and for each dyadic interval I we consider the eigenvectors e_I^k of the matrix $< W >_I$ as we did in the previous section, their eigenvalues λ_I^k, and the vectors $h_I e_I^k$ (see section 5.2). For a vector function f we define the operator

$$M_\sigma^W f = \sum_{\substack{I \in \mathcal{D} \\ 1 \leq k \leq d}} \sigma_I^k (f, h_I e_I^k)_2 \cdot h_I e_I^k$$

and the projections

$$P_{I,k}^W f = (f, W^{-1} h_I e_I^k)_{2,W} h_I e_I^k = h_I \Big(\int_I f(t) h_I(t) dt, e_I^k \Big)_{\mathbb{C}^d} e_I^k.$$

We write the super-index W because they depend on the matrix weight. The claim is the following Theorem which is a substitute for the not well behaved Martingale transform M_σ (see section 5.1).

Theorem 18. *Let W be a matrix weight with $[W]_{A_2} = 1 + \delta$ where $\delta > 0$. There is a dimensional constant $c > 0$ such that for all δ sufficiently close to 0 we have the estimate*

$$\|M_\sigma^W\|_{2,W} \leq 1 + c\sqrt{[W]_{A_2} - 1}.$$

Proof. We claim that the projections $P_{I,k}^W$ behave nice in $L^2(W)$. The square of their norm $\|P_{I,k}^W\|_{2,W}^2$ is equal to $\|h_I e_I^k\|_{2,W}^2 \|h_I e_I^k\|_{2,W^{-1}}^2$ and in its turn this is equal to the product (the brackets $(,)$ mean inner product in $\mathbb{C}^d$)

$$\begin{aligned}
(< W >_I e_I^k, e_I^k)(< W^{-1} >_I e_I^k, e_I^k) &= \lambda_I^k (< W^{-1} >_I e_I^k, e_I^k) \\
&= (< W^{-1} >_I^{\frac{1}{2}} \sqrt{\lambda_I^k} e_I^k, < W^{-1} >_I^{\frac{1}{2}} \sqrt{\lambda_I^k} e_I^k) \\
&= (< W^{-1} >_I^{\frac{1}{2}} < W >_I^{\frac{1}{2}} e_I^k, < W^{-1} >_I^{\frac{1}{2}} < W >_I^{\frac{1}{2}} e_I^k) \\
&\leq [W]_{A_2}^2
\end{aligned}$$

which is equal to $1 + c\delta$ for $[W]_{A_2} = 1 + \delta$, $\delta \approx 0$. To continue we denote by $(.,.)_H$ the usual inner product in the Hilbert space $H = L^2(W)$ (we denote the dual space $L^2(W^{-1}) = H^*$) and we use the notation $x_{I,k} = h_I e_I^k$. In order to estimate the operator norm of M_σ^W on $L^2(W)$ we only need to estimate the expression

$$\sum_{\substack{I \in \mathcal{D} \\ 1 \leq k \leq d}} |(x_{I,k}, W^{-1} f)_H| |(x_{I,k}, W\psi)_{H^*}|,$$

for $f \in L^2(W^{-1})$ and $\psi \in L^2(W)$. This sum is equal to

$$\sum_{\substack{I \in \mathcal{D} \\ 1 \leq k \leq d}} \|x_{I,k}\|_H \|x_{I,k}\|_{H^*} |(\frac{x_{I,k}}{\|x_{I,k}\|_H}, F)_H| |(\frac{x_{I,k}}{\|x_{I,k}\|_{H^*}}, \Psi)_{H^*}|, \tag{5.1}$$

where we denote by $F = W^{-1} f$ and $\Psi = W\psi$. We bound $\|x_{I,k}\|_H \|x_{I,k}\|_{H^*}$ by the norm of the projection $P_{I,k}^W$ on $L^2(W)$ which is less than or equal to $[W]_{A_2}^2$ and then by the use of Cauchy-Schwartz inequality, expression (5.1) is bounded above by

$$[W]_{A_2}^2 \Big(\sum_{\substack{I \in \mathcal{D} \\ 1 \leq k \leq d}} |(\frac{x_{I,k}}{\|x_{I,k}\|_H}, F)_H|^2 \Big)^{\frac{1}{2}} \Big(\sum_{\substack{I \in \mathcal{D} \\ 1 \leq k \leq d}} |(\frac{x_{I,k}}{\|x_{I,k}\|_{H^*}}, \Psi)_{H^*}|^2 \Big)^{\frac{1}{2}}. \tag{5.2}$$

We bound each one of the two sums separately. Notice that for the first sum we project the vector F orthogonally onto the unit vectors $\frac{x_{I,k}}{\|x_{I,k}\|_H}$ which constitute a bases of the vector space E_I introduced in section 5.2. This means that $\sum_{1\le k\le d}|(\frac{x_{I,k}}{\|x_{I,k}\|_H},F)_H|^2 = \|P_{E_I}F\|_{2,W}^2$ and therefore,

$$\Big(\sum_{\substack{I\in\mathcal{D}\\1\le k\le d}}|(\frac{x_{I,k}}{\|x_{I,k}\|_H},F)_H|^2\Big)^{\frac{1}{2}} \le \Big(\sum_{I\in\mathcal{D}}\|P_{E_I}F\|_{2,W}^2\Big)^{\frac{1}{2}} \le \Big((1+c\sqrt{\log[W]_{A_2}})\|F\|_{2,W}^2\Big)^{\frac{1}{2}}$$

by Theorem 17, which is exactly what we want. For the second sum of (5.2) similar considerations apply. Namely, we project the vector Ψ onto the unit vectors $\frac{x_{I,k}}{\|x_{I,k}\|_{H^*}}$ which are almost orthogonal in H^* (the cosine of the angle between the vectors $x_{I,k}$ and $x_{I,k'}$, for $k\neq k'$, is of the order $\sqrt{\delta}$) and they constitute a bases of E_I'. So by the use of the law of cosines we can bound the sum $\sum_{1\le k\le d}|(\frac{x_{I,k}}{\|x_{I,k}\|_{H^*}},\Psi)_{H^*}|^2$ from above by $(1+c\sqrt{\delta})\|P_{E_I'}\Psi\|_{2,W^{-1}}^2$. Since

$$\sum_{I\in\mathcal{D}}\|P_{E_I'}\Psi\|_{2,W^{-1}}^2 \le (1+c\sqrt{\log[W]_{A_2}})\|\Psi\|_{2,W^{-1}}^2$$

by Theorem 17, we are done. In summary we have the desired estimate

$$\|M_\sigma^W\|_{2,W} \le 1+c\sqrt{[W]_{A_2}-1}.$$

$\square$

5.4 Open problems about matrix weights

As we showed in the beginning of this chapter for any A_2 weight $W\in\mathbb{C}^{d\times d}$ and any direction $e\in\mathbb{C}^d$ the scalar weight $w_e(x)=(W(x)e,e)_{\mathbb{C}^d}$ is A_2 with characteristic at most

$[W]_{A_2}$. Since the A_2 characteristic does not change when we multiply the weight by a positive constant number it is not hard to see that for any vector $y \in \mathbb{C}^d$, $y \neq 0$, the weight $w_y(x) = (W(x)y, y)_{\mathbb{C}^d}$ is an A_2 scalar weight and actually,

$$\sup_{y \in \mathbb{C}^d \setminus \{0\}} [w_y]_{A_2} \leq [W]_{A_2}.$$

The question that rises is the following. Suppose that for a matrix weight W and any non-zero vector y the scalar weights $w_y := (W(x)y, y)_{\mathbb{C}^d}$ and $(w^{-1})_y = (W^{-1}y, y)_{\mathbb{C}^d}$ are in the A_2 class with uniform bound for the A_2 characteristic. Do we necessarily have that $W \in A_2$? In [16] Dr. Lauzon and Dr. Treil proved that for $W \in \mathbb{R}^{2\times 2}$ the answer to this question is positive. That is if the weights $w_y, (w^{-1})_y$ are uniformly A_2 over all unit vectors $y \in \mathbb{R}^2$, then W satisfies the matrix A_2 condition. In addition, they proved that for $n \geq 6$ there are $W \in \mathbb{R}^{n\times n}$ such that for all directions $y \in \mathbb{R}^n$ the scalar weights $w_y, (w^{-1})_y$ are uniformly A_2 but $W \notin A_2$. For dimensions $n = 3, 4$ and 5 the answer is not known.

We should mention that the problem of characterizing matrix weights $W \in \mathbb{C}^{d\times d}$ that satisfy

$$\Big(\int_{\mathbb{R}} (W(x)^{\frac{2}{p}} Hf(x), Hf(x))_{\mathbb{C}^d}^{\frac{p}{2}} dx\Big)^{\frac{1}{p}} \leq C\Big(\int_{\mathbb{R}} (W(x)^{\frac{2}{p}} f(x), f(x))_{\mathbb{C}^d}^{\frac{p}{2}} dx\Big)^{\frac{1}{p}},$$

for $1 < p < \infty$, where $C > 0$ is a constant, has been solved in [19] and [32] with different methods. It states that this holds if and only if $W \in A_{p,q}$ where $\frac{1}{p} + \frac{1}{q} = 1$. To introduce the $A_{p,q}$ condition we need some preliminary definitions. Let $t \mapsto \rho_t$, $t \in \mathbb{R}$ be a function whose values are norms (or even semi-norms) on $\mathbb{R}^n$ (or $\mathbb{C}^n$). We assume this function to be measurable in the sense that for any vector $x \in \mathbb{R}^n$ the function $t \mapsto \rho_t(x)$ is measurable.

The $L^p(\rho)$ space consists of measurable vector functions f such that

$$\|f\|^p_{L^p(\rho)} := \int_{\mathbb{R}} \rho_t(f(t))^p dt < \infty.$$

The weighted $L^p(W)$ space with a matrix weight W is a special case of the space $L^p(\rho)$ where $\rho_t(x) = \|W(t)^{\frac{1}{p}} x\|$. For a norm ν on $\mathbb{R}^n$ (or $\mathbb{C}^n$) we denote by ν^* the dual norm defined as

$$\nu^*(x) = \sup_{y \neq 0} \frac{|(x,y)|}{\nu(y)}.$$

The notation $(,)$ denotes the inner product on $\mathbb{R}^n$ (or $\mathbb{C}^n$ depending on which case we are considering). For a normed valued function ρ we define the dual function $\rho^*_t = (\rho_t)^*$. We also denote by $< \rho >_{I,p}$ the p-average

$$< \rho >_{I,p} (x) = \Big(\frac{1}{|I|} \int_I \rho_t(x)^p dt \Big)^{\frac{1}{p}},$$

where $x \in \mathbb{R}^n$. We say that a normed valued function ρ satisfies the $A_{p,q}$ condition for $1 < p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$ if there is a positive constant C such that

$$< \rho^* >_{I,q} \leq C < \rho >^*_{I,p}, \tag{5.3}$$

for all finite intervals I in $\mathbb{R}$. Notice that the opposite inequality always holds with $C = 1$. This follows by Hölder's inequality and the fact that for a norm ν and all $x, y \in \mathbb{R}^n$ we have $|(x,y)| \leq \nu(x)\nu^*(y)$. We can denote the smallest C that satisfies inequality (5.3) by $[\rho]_{A_{p,q}}$ (notice that it is always bigger or equal to 1) and try to study the relation between operator norms on $L^p(\rho)$ with respect to the quantity $[\rho]_{A_{p,q}}$. A problem like this seems more difficult than the $p = 2$ case.

The $A_{p,q}$ classes do not have the analogous properties of their scalar dimensional analogues namely the A_p classes. In [1] the author showed that there is an A_2 matrix weight $W \in \mathbb{C}^{2\times 2}$ such that $W \notin A_{p,q}$ for all $p < 2$, $\frac{1}{p} + \frac{1}{q} = 1$. This is a counter example to the open ended property of the A_p classes. He also presented an example of a 2×2 A_2 weight W such that $W^r \notin A_2$ for all $r > 1$.

The behavior of the dimensional constants c that appear in our calculations when d grows is unclear and it can be an interesting problem. It is quite natural to understand the infinite-dimensional stationary stochastic processes, and, thus, to understand operator-valued weights. However, this is an open question. In [10] it was shown that for the Martingale transform

$$T_\sigma f = \sum_{I \in \mathcal{D}} \sigma_I (f, h_I) h_I,$$

the following is true. There are constants $0 < a, A < \infty$, such that for every d there is an A_2 weight $W_d \in \mathbb{C}^{d\times d}$ with $[W_d]_{A_2} \leq A$ and

$$\sup_\sigma \|T_\sigma\|_{2,W_d} \geq a(\log d)^{\frac{1}{2}}.$$

In [11] the same authors proved that a completely analogous result holds for the Hilbert transform.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] M. Bownik, *Inverse volume inequalities for matrix weights* , Indiana Univ. Math. J. 50 (2001), no. 1, 383-410.

[2] Stephen M. Buckley, *Estimates for operator norms on weighted spaces and reverse Jensen inequalities*, Trans. Amer. Math. Soc., volume 340, number 1, November 1993.

[3] D. Burkholder, *Boundary value problems and sharp estimates for the martingale transforms*, Ann. of Prob. 12 (1984), 647-702.

[4] R. Coifman and C. Fefferman, *Weighted norm inequalities for maximal functions and singular integrals*, Studia Math. 51 (1974), 241-250.

[5] R. Coifman, P. Jones and J. Rubio de Francia, *Constructive decomposition of B.M.O. functions and factorization of A_p weights*, Proc. Amer. Math. Soc. 87 (1983), 675-676.

[6] J.G. Conlon and T. Spencer, *A strong central limit theorem for a class of random surfaces*, arXiv:1105.2814v2.

[7] O. Dragicevic and A. Volberg, *Bellman function, Littlewood-Paley estimates and asymptotics for the Ahlfors-Beurling operator in $L^p(\mathbb{C})$*, Indiana Univ. Math. J. 54 (2005), no. 4, 971-995.

[8] J. Garcia-Cuerva and J. Rubio De Francia, *Weighted norm inequalities and related topics*, North Holland Math. Stud. 116, North Holland, Amsterdam 1985.

[9] J. Garnett and P.W. Jones, *The distance in BMO to L^∞*, Ann. of Math. 108 (1978), 373-393.

[10] A. Gillespie, S. Pott, S. Treil and A. Volberg, *Logarithmic growth for matrix Martingale transform*, J. London Math. Soc. (2) 64 (2001), no. 3, 624-636.

[11] A. Gillespie, S. Pott, S. Treil and A. Volberg, *Logarithmic growth for weighted Hilbert transform and vector Hankel operators*, J. Operator Theory 52 (2004), no. 1, 103-112.

[12] R. Hunt, B. Muckenhoupt and R. Wheeden, *Weighted norm inequalities for the conjugate function and the Hilbert transform*, Trans. Amer. Math. Soc. 176 (1973), 227-251.

[13] T. Hytonen, *The sharp weighted bound for general Calderón-Zygmund operators*, arXiv:1007.4330.

[14] P. Jones, *Factorization of A_p weights*, Ann. of Math. 111 (1980), 511-530.

[15] M. B. Korey, *Ideal weights: Asymptotically optimal versions of doubling, absolute continuity, and bounded mean oscillation*, The Journal of Fourier Analysis and Applications, Volume 4, Issues 4 and 5, 1998.

[16] M. Lauzon and S. Treil, *Scalar and vector Muckenhoupt weights*, Indiana Univ. Math. J. 56 (4) (2007) 1989-2015.

[17] N. G. Meyers, *An L^p estimate for the gradient of solutions of second order elliptic divergence equations*, Ann. Scuola Norm. Sup. Pisa (3) 17 189-206, (1963).

[18] B. Muckenhoupt, *Weighted norm inequalities for the Hardy maximal function*, Trans. Amer. Math. Soc. 165 (1972), 207-226.

[19] F. Nazarov and S. Treil, *The hunt for a Bellman function: applications to estimates for singular integral operators and to other classical problems of harmonic analysis*, Algebra i Analiz, 8:5 (1996), 32162.

[20] F. Nazarov and A. Volberg, *Heating of the Ahlfors-Beurling operator and estimates of it's norms*, St. Petersburg Math. J. Vol. 15 (2004), No. 4, Pages 563-573.

[21] N. K. Nikolskii, *"Treatise on the Shift Operator"*, Springer-Verlag, New York, 1985.

[22] M. Papadimitrakis and N. Pattakos, *Continuity of weighted estimates for sublinear operators*, arXiv:1206.4580v1.

[23] N. Pattakos and A. Volberg, *Continuity of weighted estimates in A_p norm*, Proc. Amer. Math. Soc. 140 (2012), 2783-2790.

[24] N. Pattakos and A. Volberg, *The Muckenhoupt A_∞ class as a metric space and continuity of weighted estimates*, Math. Res. Lett. 19 (2012), no. 02, 499-510.

[25] N. Pattakos and A. Volberg, *A new weighted Bellman function*, C. R. Acad. Sci. Paris, Ser. I 349 (2011) 1151-1154.

[26] S. Petermichl and A. Volberg, *Heating of the Ahlfors-Beurling operator: Weakly quasiregular maps on the plane are quasiregular*, Duke Math. J. 112 (2002), no. 2, 281-305.

[27] Yu. A. Rozanov, *Stationary stochastic processes*, Holden-Day, SF, 1967.

[28] E. M. Stein, *Harmonic Analysis: Real Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton Univ. Press, Princeton, 1993.

[29] E. Stein and G. Weiss, *Interpolation of operators with change of measures*, Transactions of the AMS, Volume 87, Number 1, Pages 159-172 (1958).

[30] S. Treil and A. Volberg, *Wavelets and the angle between the Past and the Future*,Journal of Functional Analysis 143, 269-308 (1997).

[31] S. Treil and A. Volberg, *Completely regular multivariate stationary processes and the Muckenhoupt condition*, Pacific J. Math. Vol. 190 (1999), No. 2, 361382.

[32] A. Volberg, *Matrix A_p weights via S-functions*, Journal of the American Mathematical Society Volume 10, Number 2, April 1997, Pages 445-466.

[33] J. Wittwer, *A sharp estimate on the norm of the Martingale transform*, Math. Res. Lett. 7, 112 (2000).