Contemporary Mathematics

# Lecture Notes on Dyadic Harmonic Analysis

María Cristina Pereyra

Contents


| | |
|---|---|
| Introduction | 1 |
| 1. Main Characters | 3 |
| 2. Classical Tools | 16 |
| 3. $BMO$, $A_\infty$ and Stopping Times | 27 |
| 4. The $T(1)$ Theorem | 36 |
| 5. Carleson's Lemma and Bellman Functions | 43 |
| 6. Haar multipliers and Weighted Inequalities | 49 |
| References | 57 |


## Introduction

These Lecture Notes grew out of a series of lectures delivered by the author at the Analysis Summer School in the Instituto de Matemáticas of the Universidad Autónoma de México, Unidad Cuernavaca in June 2000. The lectures were intended for begining graduate students with a basic knowledge of real and complex analysis, measure theory and functional analysis. There were many exercises sprinkled throughout the lectures, which hopefully complemented and helped the reader test his/her understanding of the material presented. I have included those exercises and more in the lecture notes. Also, while I had hoped to cover topics related to weights, I did not have sufficient time to present them in Cuernavaca, but have included them in these notes.

The notes contain what I consider are the main actors and universal tools used in this area of mathematics. They also contain an overview of the classical problems that lead mathematicians to study these objects and to develop the tools that are now considered the *abc* of harmonic analysis. The modern twist is the connection to a parallel dyadic world where objects, statements and sometimes proofs are

1991 *Mathematics Subject Classification.* Primary 42B25, 42B20, 42A45; Secondary 47B38.

*Key words and phrases.* dyadic harmonic analysis, Hilbert transform, maximal function, square function, paraproducts, Haar multipliers, $BMO$, $A_p$ and $RH_p$ weights, singular integral operators, $T(1)$ Theorem, Carleson's Lemma, Bellman functions, stopping time, weighted inequalities.

simpler, but yet illuminating enough to guarantee that one can translate them into the non-dyadic world. This philosophy has been pushed to unexpected limits by Nazarov, Treil and Volberg, as well as by their students and collaborators. Most of the material related to Bellman functions I learned from them, either in preprints or in the Spring School on Analysis held in Paseky[1] a couple weeks before the School in Cuernavaca.

In these lectures we will concentrate on Haar analogues, this is just the tip of the time/frequency iceberg. A full dyadic model for the phase plane is given by the Walsh functions. Beautiful results are being obtained now with the more sophisticated time/frequency tools and very delicate combinatorial arguments. The pioneering work was done by C. Fefferman in 1972 [**Fef**]. A few years ago C. Thiele, relying heavily on Fefferman's ideas, presented a solution of a famous conjecture of Calderón for a Walsh model of the bilinear Hilbert transform in his PhD Thesis [**Th**]. Joining forces with M. Lacey they were able to prove the full conjecture in a work that earned them the 1997 Salem Price [**LT**]. There is lots of work in progress along these lines in connection to PDE's, for an overview see the lecture notes of a course taught by T. Tao during Spring 2001 at UCLA [**Tao**].

The main problem analyzed in the following pages is $L^p$ boundedness of operators. Namely, we want to know if a given linear (or sublinear) operator $T$ acts continuously from $L^p(X)$ into $L^q(Y)$, where $(X,\mu)$ and $(Y,\nu)$ are measure spaces, for some $1 \le p, q \le \infty$, i.e. is there a constant $C > 0$ such that for all $f \in L^p(X)$,

$$\|Tf\|_{L^q(Y)} = \left(\int_Y |Tf(y)|^q d\nu(y)\right)^{\frac{1}{q}} \le C \left(\int_X |f(x)|^p d\mu(x)\right)^{\frac{1}{p}} = C\|f\|_{L^p(X)} \ ?$$

We try to illustrate in the first lecture why people were interested in such inequalities. We do so by revisiting the most classical operators: the Hilbert transform, Hardy-Littlewood maximal function, square functions and paraproducts; noting their place in history as well as their dyadic counterparts.

In the second lecture we introduce the classical tools used to handle boundedness: Schur's Lemma, Cotlar's Lemma, interpolation and extrapolation, Calderón-Zygmund decomposition. We illustrate how to use these tools to prove boundedness estimates for the classic operators.

The third lecture introduces the space of bounded mean oscillation ($BMO$) and $A_\infty$ weights, as well as their dyadic counterparts; the "self-improvement" theorems of John-Nirenberg and Gehring are proved as the first examples of the power of stopping time techniques. An analogue of the John-Nirenberg Theorem for $RH_p$ weights is presented, the Weight Lemma, and its use is illustrated in proving characterizations of weights by summation conditions.

In the fourth lecture singular integral operators are introduced and the celebrated $T(1)$ theorem of David-Journé is proved following the dyadic proof of Coifman and Semmes. Some history on the Cauchy integral is provided as well as the $T(b)$ theorem. Finally we present a short survey on recent progress done in extending these tools to non-homogeneous (non-doubling measures) spaces.

In the fifth lecture the classical Carleson embedding theorem is presented and its dyadic counterparts. A stopping time proof is presented as well as Nazarov-Treil-Volberg's proof using Bellman functions. Either proof can be extended to handle weighted versions of the embedding theorem. We illustrate furthermore the

[1]School that I attented thanks to a Travel Grant from AWM/NSF, May 2000

Bellman function technique by proving Buckley's characterization of $A_\infty$ weights by summation conditions.

In the last lecture we use the weight lemma introduced in the third lecture to prove the boundedness of some non-constant Haar multipliers. From their boundedness one can deduce boundedness on weighted $L^p$ for our dyadic operators: constant Haar multipliers, paraproducts and square function. The last section is a survey on weighted inequalities. Much progress has been made in the last 5 years. Several longstanding problems have been solved like the single matrix-valued weight, and the two-weights problem for the Hilbert transform; as well as the study of sharp constants for the boundedness of the dyadic square function and the Hilbert transform on weighted spaces.

Some years ago we had a plan to write a book on this subject with Nets Katz, there is an unpublished manuscript by Katz that has been very inspiring, [**Ka2**]. There are many books in harmonic analysis that contain much more than what is here, including excellent expository books like [**Duo**] and [**Kr**], or concise and juicy surveys like [**Ch**] and [**Da2**], or the well known new and old testaments [**St1**], [**St2**], which are of an encyclopedic nature and are compulsory reading for anybody interested in modern harmonic analysis.

We were very lucky to have Steve Hofmann teaching simultaneously a beautiful course on his very recent proof of the Kato Problem (a 40 years old longstanding conjecture), see his Lecture Notes in this volume [**Hof**]. His proof is very classical and utilizes all these techniques: Littlewood-Paley analysis (square functions), maximal functions, Carleson's measures, sophisticated versions of the $T(b)$ theorem for corresponding singular integral operators adapted to the heat kernel. It was delightful to see all the classical techniques joining forces to produce such an astonishing and long overdue result. It does say something about the strength of the basic techniques.

Last but not least, I would like to warmly thank the organizers, Salvador Pérez-Esteva and Carlos Villegas, for inviting me to teach in the school. I would also like to thank all the participants, students, colleagues from throughout Mexico and abroad who attended the course and provided comments during and afterward. In particular Martha Guzmán-Partida, Lucero de Teresa, Magaly Folch, Stephan De Vievre and Steve Hofmann. You all made this experience a very rewarding one, socially and mathematically. Finally I would like to thank Kees Onneweer who volunteered to proofread the manuscript under a tight time schedule.

Disclaimer: All other mistakes are my son Nicolás' fault! He was in the making while these lectures were delivered, and he was born before I had time to finish them. I thought it was going to be easy to complete this project, little did I know!!

## 1. Main Characters

**1.1. The Hilbert Transform.** The Hilbert transform is the prototypical example of a singular integral operator. It is given formally by the principal value integral:

$$Hf(x) = \text{p.v.}\frac{1}{\pi}\int \frac{f(y)}{x-y}dy := \lim_{\epsilon\to 0}\frac{1}{\pi}\int_{|x-y|>\epsilon}\frac{f(y)}{x-y}dy.$$

Notice that it is also given by convolution with the distributional kernel $k(x) = \text{p.v.}\frac{1}{\pi x}$. If the kernel were an integrable function, then the integral operator: $Tf(x) = k * f(x) = \int k(x-y)f(y)dy$ would be automatically bounded in $L^p$ for all $1 \leq p \leq \infty$, by Young's inequality: $\|f * k\|_p \leq \|k\|_1 \|f\|_p$. Unfortunately the Hilbert transform kernel is not integrable, nevertheless the Hilbert transform is bounded in $L^p(\mathbb{R})$ for all $1 < p < \infty$; and although it is not bounded at the endpoints $p = 1$ and $p = \infty$; there are appropriate substitutes.

In particular one can compute the Fourier transform of the Hilbert transform, at least when applied to very smooth and compactly supported functions, and obtain:

$$(Hf)^\wedge(\xi) = \int_{\mathbf{R}} Hf(x)e^{-2\pi i x\xi}dx = -i\,\text{sgn}(\xi)\hat{f}(\xi);$$

here we define $\text{sgn}(\xi) = -1$ when $\xi < 0$, $\text{sgn}(\xi) = 1$ when $\xi > 0$, and $\text{sgn}(0) = 0$.

This automatically shows that the Hilbert transform is an isometry [2] on a dense subset of $L^2(\mathbb{R})$, and can then be extended by continuity as an isometry to $L^2(\mathbb{R})$. Notice also that from the above identity we conclude that $H^2 = -I$.

In the next lecture we will give alternative proofs, based on Cotlar's and Schur's lemmas, of the boundedness in $L^2$ of the Hilbert transform. We will also present the original proof by M. Riesz of the boundedness in $L^p$ for $1 < p < \infty$.

EXERCISE 1.1. Show that the Hilbert transform is not bounded in $L^1$ nor in $L^\infty$ by explicitly calculating its action on the characteristic function of the interval $[0, 1]$, which is a function in $L^1 \cap L^\infty$.

When $p = 1$, one can get away with a weaker notion of boundedness. Notice that if an operator is bounded in $L^p(X)$, $(X, \mu)$ a measure space, $p \geq 1$, then the following inequality is an immediate consequence of Tchebychev's inequality:

$$\mu(\{x \in X : |Tf(x) > \lambda\}) \leq C\left(\frac{\|f\|_{L^p(X)}}{\lambda}\right)^p, \quad C \geq 1. \tag{1.1}$$

EXERCISE 1.2. Check the above inequality for $C = 1$ and $T$ a bounded operator in $L^p(X)$.

An operator that satisfies (1.1) is said to be of *weak type* $(p, p)$. An operator that is bounded in $L^p$ is said to be of *strong type* $(p, p)$. We have just shown that strong $(p, p)$ implies weak $(p, p)$; but the converse, in general, is false.

We will show that the Hilbert transform is of weak type (1,1) in the next lecture. As for bounded functions they are mapped into a larger space *bounded mean oscillation*, $BMO$, to be defined later. This behaviour is shared by a large class of very important operators the so-called *Calderón-Zygmund singular integral operators*. The departure point of the Calderón-Zygmund theory is an a priori $L^2$ estimate; everything else unfolds from there. Having means other than Fourier analysis to obtain such $L^2$ estimate is crucial, that is the content of the celebrated $T(1)$ Theorem of David and Journé which we will discuss in our fourth lecture.

Why did mathematicians get interested in the Hilbert transform? Here are a few classical problems where the Hilbert transform appeared naturally.

[2] $\|Hf\|_2 = \|(Hf)^\wedge\|_2 = \|\hat{f}\| = \|f\|_2$, where Plancherel identity has been used twice

1.1.1. *Connection to complex analysis.* Consider a real valued function $f \in L^2(\mathbb{R})$ and let $F(z)$ be twice its analytic extension to the upper half plane $\mathbb{R}^2_+ = \{z = x + it : t > 0\}$, suitably normalized. $F(z)$ can be explicitly computed by means of the well known *Cauchy integral formula*:

$$F(z) = \frac{1}{2\pi i} \int_{\mathbb{R}} \frac{f(y)}{z - y} dy, \qquad z \in \mathbb{R}^2_+.$$

Notice the resemblance with the Hilbert transform. No principal value is needed here since the singularity is never achieved. By separating the real and imaginary parts of the kernel, one can obtain explicit formulae for the real and imaginary parts of $F(z) = u(z) + iv(z)$ in terms of convolutions with the so-called *Poisson* and *conjugate Poisson* kernels: $u(x + it) = f * P_t(x)$, $v(x + it) = f * Q_t(x)$. The function $u$ is the *harmonic extension* of $f$ to the upper-half plane, and the function $v$ is its *harmonic conjugate*.

EXERCISE 1.3. Show that the Poisson kernel is given by $P_t(x) = \frac{1}{\pi} \frac{t}{x^2+t^2}$, and the conjugate Poisson kernel by $Q_t(x) = \frac{1}{\pi} \frac{x}{x^2+t^2}$. Show that for each $t > 0$, $\hat{Q}_t(\xi) = -i \operatorname{sgn}(\xi) e^{-2\pi|t\xi|}$, therefore as $t \to 0$, $\hat{Q}_t(\xi)$ approaches $-i\operatorname{sgn}(\xi)$, the Fourier multiplier corresponding to the Hilbert transform.

The Poisson kernel is an example of an approximation of the identity that we will discuss more deeply in the next section. As such, the limit as $t \to 0$ of $u = P_t * f$ is $f$ in the $L^2$ sense and almost everywhere. On the other hand, as $t \to 0$, $v = Q_t * f$ approaches the Hilbert transform $Hf$ in $L^2$.

1.1.2. *Connection to Fourier series.* For functions integrable on $\mathbb{T} = [0, 1]$, the $n$-th Fourier coefficient is well defined by the formula

$$\hat{f}(n) = \int_0^1 f(x) e^{-2\pi i n x} dx.$$

Since $L^2(\mathbb{T}) \subset L^1(\mathbb{T})$, this is also well defined for square integrable functions. It is well known that the trigonometric system $\{e^{2\pi i n x}\}_{n \in \mathbb{Z}}$ is an orthonormal complete system in $L^2(\mathbb{T})$; therefore the following reconstruction and isometry formulae hold in $L^2$:

$$f(x) = \sum_{n \in \mathbb{Z}} \hat{f}(n) e^{2\pi i n x}, \qquad \|f\|^2_{L^2(\mathbb{T})} = \sum_{n \in \mathbb{Z}} |\hat{f}(n)|^2.$$

The $N$-th partial sum is given by

$$S_N f(x) = \sum_{|n| \le N} \hat{f}(n) e^{2\pi i n x}.$$

In the XIX century mathematicians asked for which $2\pi$-periodic functions $f$ would it be true that $\lim_{N \to \infty} S_N f(x) = f(x)$ at a given point $x \in \mathbb{T}$ ? Some partial answers were given, more than continuity at the point was always required (eg. Dini's condition). In 1889, Du Bois Raymond found a continuous function whose partial Fourier sum diverges at a point. With the advent of measure theory and $L^p$ spaces, new questions were formulated: Is there convergence a.e.? Is there convergence in the $L^p$ sense? The second question is answered positively for $1 < p < \infty$ and it is a consequence of the boundedness of the Hilbert transform in such $L^p$'s. The first question is much more difficult, for $p = 2$ the positive answer was given by L. Carleson in a celebrated paper published in 1965, see [**Car**] (settling the question for periodic continuous functions which had remained open until then); two years later, R. Hunt extended the result for the remaining $p$'s, $1 < p < \infty$, see

[**Hu**]. The case $p = 1$ had been ruled out by Kolmogorov's famous example of an integrable function whose Fourier series diverges everywhere, see [**Kol**].

By a limiting procedure on the unit disc, similar to the one described in the upper half plane, one can conclude that the boundary values of the harmonic conjugate of the harmonic extension of a periodic, real-valued, continuously differentiable, function $f$ on $\mathbb{T}$ are given by

$$\tilde{H}f(x) = \text{p.v.}\frac{1}{\pi}\int_0^1 \frac{f(t)}{\tan\big(\pi(x-t)\big)}\,dt; \tag{1.2}$$

here we are identifying $x$ with $z = e^{2\pi i x}$.

The singularity at the diagonal is comparable to that of the Hilbert transform; so this would be the appropriate analogue of the Hilbert transform on the unit circle. On Fourier side, one can check that a similar identity holds, namely: $(\tilde{H}f)^\wedge(n) = -i\text{sgn}(n)\hat{f}(n)$.

Note that the Fourier transform of the partial Fourier sum of a nice function is also given by a similar Fourier multiplier:

$$(S_N f)^\wedge(n) = \chi_{|k|\le N}(n)\hat{f}(n).$$

EXERCISE 1.4. Check that $\chi_{|k|\le N}(n) = \frac{1}{2}(\text{sgn}(n-N) - \text{sgn}(n+N))$. Furthermore remember that the Fourier transform maps modulations into translations, more precisely, check that if $M_N f(\theta) = f(x)e^{2\pi i\theta N}$, then $(M_N f)^\wedge(n) = \hat{f}(n-N)$. Finally check that $i(M_N\tilde{H}M_{-N})^\wedge(n) = \text{sgn}(n-N)\hat{f}(n)$. Similarly check that: $i(M_{-N}\tilde{H}M_N)^\wedge(n) = \text{sgn}(n+N)\hat{f}(n)$.

The exercise implies that $S_N = \frac{i}{2}(M_N\tilde{H}M_{-N} - M_{-N}\tilde{H}M_N)$.

EXERCISE 1.5. Show that the $S_N$'s are uniformly (in $N$) bounded in $L^p$, for each $1 < p < \infty$. Deduce, from the Uniform Boundedness Principle, that

$$\lim_{N\to\infty} \|S_N f - f\|_p = 0.$$

Therefore the convergence in $L^p$ of the partial Fourier sums is a consequence of the boundedness of the Hilbert transform in those spaces.

1.1.3. *Connection to stationary processes.* In the 50's Wiener and Massani studied stationary Gaussian processes, see [**MW**]. A *discrete stationary process* is a sequence $\{\xi_n\}_{n\in\mathbb{Z}}$ of random variables in the probability space $(\Omega, P)$ such that $E(\xi_n) = 0$ and $E(\xi_n^2) < \infty$[3]; and $E(\xi_n\xi_k) = \gamma(k-n)$ (stationary condition). This last condition implies that the correlation matrix has a Toeplitz structure. Notice that the sequence $\{\gamma(k)\}_{k\in\mathbb{Z}}$ is a positive definite sequence[4].

They were interested in the geometry of such a sequence of random variables in $L^2(\Omega, P)$. The inner product is naturally given by $\langle\xi,\eta\rangle = E(\xi\eta)$, and let $\mathcal{H}$ be the closure of the linear span of the sequence $\{\xi_n\}_{n\in\mathbb{Z}}$ in the norm induced by this inner product. The problem they had was to predict $\xi_1$ knowing the predecessors $\xi_k$, $k \le 0$. The best predictor in the Hilbert space context is the orthogonal projection of $\xi_1$ onto the span of $\{\xi_k : k \le 0\}$.

The celebrated Herglotz-Bochner-Schwartz theorem allows us to move into more familiar ground: given a positive definite sequence $\{\gamma(k)\}_{k\in\mathbb{Z}}$ there exists a unique positive measure $\mu \ge 0$ on the unit circle $\mathbb{T}$ (parameterized by $z =$

[3]as usual the expectation is given by $E(\xi) = \int_\Omega \xi dP$

[4]i.e. $\sum_{k,n}\gamma(k-n)x_k\overline{x_n} = E|\sum_k x_k\xi_k|^2 \ge 0$ for all $x_k$.

$e^{2\pi ix}$, $x \in [0,1)$), such that $\gamma(k) = \hat{\mu}(k) = \int_{\mathbb{T}} \overline{z}^k d\mu(z)$. Such measure $\mu$ is called the *spectral measure of the process*. Moreover,

$$E(\xi_k \xi_n) = \gamma(k-n) = \int_{\mathbb{T}} z^n \overline{z}^k d\mu(z) = \langle z^n, z^k \rangle_{L^2(\mu)}.$$

Instead of studying the geometry of $\{\xi_n\}_{n\in\mathbb{Z}}$ in $\mathcal{H}$ we study the geometry of $\{z^n\}_{n\in\mathbb{Z}}$ in $L^2(\mu)$.

Let the *past* be denoted by $\mathcal{P} = \overline{span\{z^n : n < 0\}}$ and the *future* by $\mathcal{F} = \overline{span\{z^n : n \geq 0\}}$. When is the angle between the past and the future positive? Remember that given two closed subspaces $E_1, E_2$ in a Hilbert space $H$ then $\cos(\angle E_1 E_2) = \sup\{|\langle e_1, e_2\rangle_H| : e_i \in E_i, \|e_i\| = 1\}$.

EXERCISE 1.6. Let $H, E_1, E_2$ be as above, assume that $E_1 \cap E_2 = \{0\}$, $E_1 + E_2$ is dense in $H$ then show that the following are equivalent: (I) $\angle E_1 E_2 > 0$, (II) $E_1 + E_2 = H$, (III) The projection onto $E_1$ parallel to $E_2$, $P_{E_1\|E_2}$[5] is bounded. Moreover if $\angle E_1 E_2 = \alpha$ then $\|P_{E_1\|E_2}\| = 1/\sin\alpha$. **Hint:** Use the Closed Graph Theorem for the equivalences.

Now our question can be rephrased as: When is the *Riesz projection* $P_+$ bounded? Here $P_+(\sum c_k z^k) = \sum_{k\geq 0} c_k z^k$ is the projection onto the future parallel to the past, or the projection onto the analytic part of the function. By the next exercises, this is equivalent to asking when is $\{z^n\}_{n\in\mathbb{Z}}$ a basis in $L^2(\mu)$.

EXERCISE 1.7. $\{x_n\}$ is a basis on a Banach space $X$ if and only if for all $x \in X$ there is a unique representation $x = \sum c_k x_k$ (this is equivalent to asking that the system $\{x_n\}$ is complete and linearly independent). Show that $\{x_n\}$ is a basis if and only if $\|P_n\| \leq C$ for all $n$, where $P_n(\sum c_k x_k) = \sum_{k=1}^n c_k x_k$. **Hint:** use Uniform Boundedness Principle.

EXERCISE 1.8. In our setting $P_n f = \sum_{k=0}^n \hat{f}(k) z^k$. Show that $P_n f = P_+ f - z^{n+1} P_+ \overline{z}^{n+1} f$. Show also that if $P_+$ is bounded on $L^2(\mu)$ then $\mu$ is absolutely continuous with respect to Lebesgue measure: $d\mu = w\, dm$, $w \geq 0$.

There is an explicit formula for $P_+ f(z)$ for $|z| < 1$, the Cauchy formula (the projection onto the analytic part of the function). We already discussed the boundary values when $|z| = 1$:

$$P_+ f(z) = \frac{1}{2} f(z) + i \frac{1}{2} \tilde{H} f(z), \qquad |z| = 1.$$

The behaviors of $H$ and $\tilde{H}$ are similar as we mentioned before. Therefore the question becomes: When is the Hilbert transform $H$ bounded in $L^2(w)$? Where $w$ is a *weight*, that is $0 \leq \omega \in L^1_{loc}$. This problem was solved in 1965 by Helson and Szegö, the necessary and sufficient conditions on the weight $w$ are that $w = e^{u+Hv}$, for $u, v \in L^\infty$, and $\|v\|_\infty < \pi/2$ [**HS**]. They used complex analysis methods. In 1973, Hunt, Muckenhoupt and Wheeden found an equivalent condition using purely real methods [**HMW**]: $H$ is bounded in $L^2(w)$ if and only if $w \in A_2$,

$$\sup_I \left(\frac{1}{|I|}\int_I w\right)\left(\frac{1}{|I|}\int_I w^{-1}\right) < \infty \quad A_2 - \text{condition}, \tag{1.3}$$

where the supremum is taken over all intervals $I$. We will say more about weights and weighted inequalities in later lectures.

[5] If $x \in E_1 + E_2$, then $x = x_1 + x_2$, $x_i \in E_i$, the decomposition is unique; define $P_{E_1\|E_2} x = x_1$.

We have borrowed the presentation in this section from lectures presented by S. Treil in the 2000 Spring Analysis School held in Paseky, Czeck Republic. For generalizations to the multivariate setting see [**TV2**].

**1.2. The Hardy-Littlewood Maximal Function.** A natural question for locally integrable functions $f$ in $\mathbb{R}$ is whether the averages on small intervals $I_x$ containing a point $x$ converge to the value of the function there, i.e.

$$\lim_{|I_x|\to 0} \frac{1}{|I_x|}\int_{I_x} f(t)\,dt = f(x).$$

It is clear that if the function $f$ is continuous this is true, by the Fundamental Theorem of Calculus. The Lebesgue Differentiation Theorem says that for locally integrable functions this is true almost everywhere (a.e.).

A natural object to study, instead of the limit, is the supremum. In this example it corresponds to the Hardy-Littlewod maximal function, a sublinear operator defined by

$$Mf(x) = \sup_{x\in I} \frac{1}{|I|}\int_I |f(t)|\,dt.;$$

here $I$ are intervals containing $x$.

It turns out that the boundedness properties of maximal operators imply convergence a.e. of corresponding limits. It should be clear that the maximal function is bounded in $L^\infty$. What is less obvious is that $M$ is of strong type $(p,p)$ for $p>1$ and of weak type $(1,1)$. We will show both facts in the next lecture.

EXERCISE 1.9. Show that the maximal function M is not of strong type $(1,1)$.

1.2.1. *Approximations of the identity.* The maximal function controls a large class of so-called *approximations of the identity.*

Given $\phi$ a real valued integrable function in $\mathbb{R}$, such that $\int \phi = 1$. Define for each $t>0$ $\phi_t(x) = \frac{1}{t}\phi(\frac{x}{t})$. We say that the family $\{\phi_t\}_{t>0}$ is an approximation of the identity. One can check that $\phi_t$ converges as $t\to 0$ to the Dirac delta function in the sense of distributions. In particular this implies that for nice functions (in the Schwartz class $\mathcal{S}$): $\lim_{t\to 0}\phi_t * g(x) = g(x)$ for all $x$. The question then becomes:

$$\text{When does} \qquad \lim_{t\to 0}\phi_t * f(x) = f(x) \quad a.e.?$$

EXERCISE 1.10. Given an approximation of the identity $\{\phi_t\}_{t>0}$ show that

$$\lim_{t\to 0}\|\phi_t * f - f\|_p = 0, \qquad \forall f\in L^p,\ \ 1\le p<\infty.$$

As a consequence there exists a subsequence $\phi_{t_k} * f(x)$ that converges a.e. to $f(x)$. Therefore if $\lim_{t\to 0}\phi_t * f(x)$ exists it must coincide with $f(x)$ a.e.

1.2.2. *More on Fourier series - Summability methods.* The classical examples of approximations of the identity arise in the study of Fourier integrals and convergence of truncated Fourier integrals. We will remind you the analogue problems in the setting of Fourier series and convergence of partial Fourier sums, where instead of approximations of the identity, as defined in the previous paragraph, we encounter summability kernels.

A *summability kernel* is a sequence $\{K_N\}$ of continuous 1-periodic functions whose averages are 1, whose $L^1$ norms are uniformly bounded, and for all $0<\delta<\pi$,

$$\lim_{N\to\infty}\int_\delta^{1-\delta} |K_N(t)|dt = 0. \tag{1.4}$$

EXERCISE 1.11. Let $f \in L^1[\mathbb{T}]$, $\{K_N\}$ a summability kernel, show that

$$\lim_{N\to\infty} \|f * K_N\|_1 = 0.$$

We can consider families $\{K_r\}$ depending on a continuous parameter $r$ instead of the discrete $N$. For instance the Poisson kernel defined below, depends on the parameter $0 \le r < 1$, and we replace the limit "$\lim_{N\to\infty}$" by "$\lim_{r\to 1}$" wherever necessary.

Remember that for $f \in L^1[\mathbb{T}]$ the $N$-th partial sum is given by

$$S_N f(x) = \sum_{|n|\le N} \hat{f}(n) e^{2\pi i n x} = \int_0^1 f(t) D_N(x-t)\, dt = f * D_N(x).$$

Here $D_N$ is the Dirichlet kernel, and $\int_0^1 D_N = 1$ but the sequence $\{\|D_N\|_1\}_{N\in\mathbb{N}}$ is not uniformly bounded , this is the cause of many difficulties when trying to show a.e. convergence of partial Fourier sums.

Mathematicians used summability methods to overcome this difficulty. The Césaro method considered averages of partial sums:

$$\sigma_N f(x) = \frac{1}{N+1} \sum_{k=0}^{N} S_k f(x) = \int_0^1 f(t) F_N(x-t)\, dt = f * F_N(x),$$

where $F_N$ is the Fejér kernel which is positive and $\int_0^1 F_N = 1$, therefore $\|F_N\|_1 = 1$ for all $N$. The Poisson method considers an analytic extension to the unit disc, $F(z) = \sum_{n\ge 0} \hat{f}(n) z^n$, where $z = re^{2\pi i x}$. Notice that the real and imaginary parts of $F(z) = u(z) + iv(z)$ are given by:

$$u(re^{2\pi i x}) = \frac{1}{2} \sum_{n=-\infty}^{\infty} \hat{f}(n) r^{|n|} e^{2\pi i n x} = \frac{1}{2} \int_0^1 f(t) P_r(x-t)\, dt = \frac{1}{2} f * P_r(x),$$

$$v(re^{2\pi i x}) = \frac{-i}{2} \sum_{n=-\infty}^{\infty} \operatorname{sgn}(n) \hat{f}(n) r^{|n|} e^{2\pi i n x} = \frac{1}{2} \int_0^1 f(t) Q_r(x-t)\, dt = \frac{1}{2} f * Q_r(x),$$

where $P_r$ is the Poisson kernel on the disc, which is also positive and $\int_0^1 P_r = 1$; and $Q_r$ is the conjugate Poisson kernel on the disc.

EXERCISE 1.12. Find closed formulas for $D_N$, $F_N$, $P_r$ and $Q_r$. Check that $F_N$ and $P_r$ are summability kernels but not $D_N$. More precisely show that $\|D_N\|_1 \sim \ln N$. Compare $\lim_{r\to 1} Q_r(t)$ with (1.2) and observe that as $r \to 1$, $F(z)$ approaches $\frac{1}{2} f(x) + i \frac{1}{2} \tilde{H} f(x)$ (the limit should be taken in the sense of distributions).

More details can be found in [**Duo**], and [**Kat**].

1.2.3. *Convergence a.e.* The following theorem connects maximal operators and a.e. convergence.

THEOREM 1.13. *Given a family of linear operators, $\{T_t\}_{t\in\Lambda}$ in $L^p(X,\mu)$, indexed by a closed set of real numbers, $\Lambda$. Let $T^* f(x) = \sup_{t\in\Lambda} |T_t f(x)|$ be the maximal operator associated to $\{T_t\}_{t\in\Lambda}$. If $T^*$ is of weak type $(p,p)$, $p \ge 1$, and $t_o$ is in the closure of $\Lambda$, then the following sets are closed in $L^p$:*

$$\{f \in L^p(X,\mu) : \lim_{t\to t_o} T_t f(x) = f(x) \quad a.e.\} = A_{t_o}.$$

Its proof can be found in [**Duo**], p. 37.

In the case of the approximations of the identity, let $T_t f = \phi_t * f$. The Schwartz class $\mathcal{S}$ (or if you prefer $C_0^\infty$) is a subset of $A_0$. $\mathcal{S}$ is dense in $L^p$, therefore if $A_0$ is closed, we conclude that $A_0 = L^p$. The fact that $A_0$ is closed can be shown by the previous theorem, provided we can show that $T^*$ is of weak type $(p,p)$. One can control a large class of approximation of the identity kernels (in particular the Fejer and Poisson kernels) by the Hardy-Littlewood maximal function; therefore all we need to check is that M is of weak type $(p,p)$.

EXERCISE 1.14. Consider $\phi \in L^1$ such that $\phi$ is even and decreasing as a function of $t = |x| > 0$. Show that $T^* f(x) = \sup_{t>0} |\phi_t * f(x)| \leq \|\phi\|_1 M f(x)$. Show that if $M$ is of weak type $(p,p)$ so is $T^*$, $p \geq 1$.

1.2.4. *Lebesgue Differentiation Theorem.* The previous results can be used in particular to prove the well known Lebesgue Differentiation Theorem. Namely

$$f \in L^1(\mathbb{R}) \quad \Rightarrow \quad \lim_{t\to 0} \frac{1}{2t} \int_{x-t}^{x+t} f(s)\, ds = f(x) \quad \text{a.e..}$$

In higher dimensions one would like to consider similar problems. One could use instead of intervals, cubes or rectangles, or more general sets. If one uses cubes then the theory runs parallel to the 1-dim theory. If instead we use rectangles with sides parallel to the axis then the corresponding maximal function is of strong type $(p,p)$ for $p > 1$ but not of weak type $(1,1)$, moreover the Lebesgue differentiation theorem is false for functions $f \in L^1(\mathbb{R}^n)$, although it works for functions such that $f(1 + \log|f|)^{n-1} \in L^1_{loc}(\mathbb{R}^n)$, in particular if $f \in L^p_{loc}(\mathbb{R}^n)$. If we allow all possible rectangles then the corresponding maximal function is not even of strong type $(p,p)$ for any $p$. For this and much more on differentiability properties of basis of rectangles, see the classical book by Miguel de Guzmán [**Guz**].

**1.3. Square functions/Littlewood-Paley Theory.** The so-called *square functions* are ubiquitous objects in harmonic analysis. This is not just one object but several who share some properties. It is best to describe some of the most classical examples to give the flavor of the so-called Littlewood-Paley Theory. This theory has been a favorite tool for proving $L^p$ estimates.

1.3.1. *Trigonometric series and Littlewood-Paley square function.* Given a function $f \in L^2(\mathbb{T})$, denote the *$N$-th dyadic partial Fourier sum* by

$$P_N f(x) = \sum_{|n|\leq 2^N} \hat{f}(n) e^{2\pi i n x} \quad \big(= S_{2^N} f(x)\big).$$

$P_N f$ should be viewed as an approximation of $f$ which gets better as $N$ increases. Consider the difference operators

$$\Delta_N f(x) = P_{N+1} f(x) - P_N f(x) = \sum_{2^N < |n| \leq 2^{N+1}} \hat{f}(n) e^{2\pi i n x}, \;\; N \geq 1;$$

$$\Delta_0 f(x) = \sum_{|n|\leq 1} \hat{f}(n) e^{2\pi i n x}.$$

Because the trigonometric system is an orthonormal basis in $L^2(\mathbb{T})$ we have the reconstruction formula $f = \Delta_0 f + \sum_{N\geq 1} \Delta_N f$, and Plancherel's identity $\|f\|_2^2 = \sum_{n\in\mathbb{Z}} |\hat{f}(n)|^2$. It is clear that if we decide to change the sign of some of the Fourier coefficients and utilize them to crate a new function $Tf(x) = \sum \pm \hat{f}(n) e^{2\pi i n x}$,

then $Tf \in L^2(T)$, moreover, $T$ is an isometry: $\|Tf\|_2 = \|f\|_2$. This means that the trigonometric system is an *unconditional basis* in $L^2(\mathbb{T})$. The question then becomes: Is the trigonometric system and unconditional basis in $L^p(\mathbb{T})$ for $p \neq 2$? The answer is NO, there is no way we can decide if a function is in $L^p(\mathbb{T})$ just by analyzing the absolute value of its Fourier coefficients, unless $p = 2$. The closest substitute is obtained analyzing the *Littlewood-Paley square function*:

$$Sf(x) = \left( \sum_{n=0}^{\infty} |\Delta_n f(x)|^2 \right)^{\frac{1}{2}}.$$

THEOREM 1.15 (Littlewood-Paley). *Let $1 < p < \infty$, then $f \in L^p(\mathbb{T})$ if and only if $Sf \in L^p(\mathbb{T})$. Moreover $\|f\|_p \sim \|Sf\|_p$.*

Notice for $p = 2$ this is an immediate consequence of Plancherel's identity. For a proof for $p \neq 2$, see for example [**St2**] p.104, and Appendix D. As usual $A \sim B$ means that there exist constants $c, C > 0$ such that $cB \leq A \leq CB$.

1.3.2. *The g-function.* Let $\psi$ be a compactly supported $C^\infty(\mathbb{R})$ function with zero mean, $\int \psi = 0$. Let $\psi_t(x) = \frac{1}{t}\psi(\frac{x}{t})$. Define the family of operators $Q_t f = \psi_t * f$, which now play the role of the differences in the previous example. One obtains a *reproducing formula*: $f(x) = c^{-1}(\psi) \int_0^\infty Q_t^2 f(x) \frac{dt}{t}$.

EXERCISE 1.16. Show that $c(\psi) = \int_0^\infty |\hat{\psi}(t)|^2 \frac{dt}{t} < \infty$ is the right constant in the reproducing formula.

The analogue to the Littlewood-Paley square function is the *square function* defined by:

$$Sf(x) = \left( \int_0^\infty |Q_t f(x)|^2 \frac{dt}{t} \right)^{\frac{1}{2}}.$$

Same theorem holds,

THEOREM 1.17. *Let $1 < p < \infty$, then $f \in L^p(\mathbb{R})$ if and only if $Sf \in L^p(\mathbb{R})$. Moreover $\|f\|_p \sim \|Sf\|_p$.*

The case $p = 2$ is an immediate consequence of Plancherel. The inequalities for $p \neq 2$ are more subtle. One can view square functions as ordinary singular integrals, but now taking their values on a Hilbert space, a machinery similar to the one necessary to handle the Hilbert transform and its siblings can be used to prove this theorem, see [**St1**].

When the functions $\psi_t$ arise as derivatives of the Poisson kernel, more precisely, $\psi_t = t\frac{\partial}{\partial t} P_t$, then the square function is called the *g-function*:

$$g(f)(x) = \left( \int_0^\infty t|\nabla u(x,t)|^2 dt \right)^{\frac{1}{2}},$$

where $u(x,t) = P_t * f(x)$ is the harmonic extension of $f$. In this case one can use Green's formula to show that $\|f\|_2 = \|g(f)\|_2$.

There is a third very illustrative example, the *dyadic square function*. We have decided to present all dyadic analogues at the end of this first lecture. All these square functions share the property that we go from a function of $x$ to a function of $(x,t)$, $t > 0$, or of $(x,n)$, $n \geq 0$, or $n \in \mathbb{Z}$ ($Q_t f(x)$ or $\Delta_n f(x)$). There is always some reconstruction formula and the way the square function is constructed is by taking an $L^2$ (or $l^2$) norm on the new variable. The square function $Sf$ has now the same $L^p$ properties as the function $f$.

**1.4. Paraproducts and $BMO$.** The *paraproducts* are cousins of the square functions. As such, they represent a class of objects rather than one specific operator. Here we will discuss the continuous paraproduct, the dyadic paraproduct will be discussed in the next section.

Let $\phi, \psi$ be compactly supported $C^\infty$ functions (or functions in the Schwartz class), such that $\int \psi = 0$, and $\int \phi = 1$. Introduce the operators $Q_t f = \psi_t * f$ and $P_t f = \phi * f$. As before, the $Q_t$ operator represents differences and the $P_t$ is an approximation of the identity, and therefore represents averagings. The *paraproduct* is defined formally as a bilinear operator:

$$\pi(f, b) = \int_0^\infty Q_t(P_t f \, Q_t b)\frac{dt}{t}.$$

Heuristically the paraproduct can be thought as "half a product": $bf \sim \pi(f, b) + \pi(b, f)$. In our case one could do a formal Fourier analysis argument, see [**Ch**] p. 41. This will be more evident in the case of the dyadic paraproduct. For a fixed function $b$ consider the ordinary product as an operator in $L^p$, $M_b f = bf$.

EXERCISE 1.18. Show that $M_b$ is bounded in $L^p$ if and only if $b \in L^\infty$.

The paraproduct will behave better in the sense that for fixed $b$ we will have boundedness properties in a space larger than $L^\infty$. That space is the so-called space of *bounded mean oscillation* or $BMO$. A locally integrable function $b \in BMO$ if

$$\|b\|_{BMO} = \sup_I \frac{1}{|I|}\int_I |b(x) - m_I b|dx < \infty;$$

where $m_I b = \frac{1}{|I|}\int_I b$ denotes the mean value over $I$ of the function $b$. This means that the average oscillation of $b$ is uniformly bounded on every interval $I$.

The null elements in the $BMO$ norm are the constants, so a function in $BMO$ is defined only up to additive constants.

EXERCISE 1.19. Show that $BMO$ (modulo constant functions) is a Banach space. Show that $L^\infty \subset BMO$. Show that $\log|x| \in BMO$, hence $BMO$ is larger than $L^\infty$.

EXERCISE 1.20. Show that $\log|P(x)| \in BMO$ for any polynomial $P$ on $\mathbb{R}$.

In the third lecture we will show that the singularities allowed in $BMO$ are precisely like those of $\log|x|$. It is the content of the celebrated John-Nirenberg Inequality. A corollary of that inequality is that for each $1 < p < \infty$,

$$\|b\|_{BMO} \sim \left(\sup_I \frac{1}{|I|}\int_I |b(x) - m_I b|^p dx\right)^{\frac{1}{p}}.$$

THEOREM 1.21. *Given $b \in BMO$ then $\pi_b$ is bounded in $L^p$ for $1 < p < \infty$. Moreover $\|\pi_b f\|_p \leq \|b\|_{BMO}\|f\|_p$. (Here $\pi_b f = \pi(f, b)$.)*

The proof can be found in [**Ch**]. It uses square functions and Carleson's Lemma which we will introduce in its dyadic incarnation in the next section, and we will discuss more deeply in the fifth lecture.

The paraproduct appeared naturally in non-linear differential equations in the work of Bony, see [**Bo**]. It turns out that the paraproduct can be thought as a singular integral operator which is far from being translation invariant. Moreover, what the acclaimed $T(1)$ Theorem says is that a large class of singular integral

operators can be decomposed as a a piece which is close to a translation invariant (or convolution) operator plus some paraproducts:

$$T = S + \pi_{b_1} + \pi^*_{b_2};$$

where $\pi^*_b$ is the adjoint of $\pi_b$. We will state this more precisely in the fourth lecture.

**1.5. Dyadic analogues.** In this section we introduce *dyadic analogues* of each of the operators discussed above (not necessarily in the same order).

Intervals of the form $[k2^{-j}, (k+1)2^{-j})$ for integers $j, k$ are called *dyadic intervals*. The collection of all dyadic intervals is denoted by $\mathcal{D}$, and $\mathcal{D}_j$ denotes all dyadic intervals $I$, such that $|I| = 2^{-j}$, also called the $j$-th generation. It is clear that each $\mathcal{D}_j$ provides a partition of the real line, and that $\mathcal{D} = \cup_{j\in\mathbb{Z}}\mathcal{D}_j$.

EXERCISE 1.22. Show that given $I, J \in \mathcal{D}$, then either they are disjoint or one is contained in the other.

This "martingale" property is what makes the dyadic intervals so useful.

Each dyadic interval $I$ is in a unique generation $\mathcal{D}_j$, and there are exactly two subintervals of $I$ in the next generation $\mathcal{D}_{j+1}$, the *children* of $I$, which we will denote $I_r$, $I_l$, the *right* and *left* child respectively. Clearly, $I = I_l \cup I_r$.

Associated to any interval $I$ there is a *Haar function* defined by:

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

where $\chi_I(x) = 1$ if $x \in I$, $\chi_I(x) = 0$ otherwise. It is not hard to see that $\{h_I\}_{I\in\mathcal{D}}$ is an orthonormal basis in $L^2(\mathbb{R})$, that is the content of the next exercise.

EXERCISE 1.23. Show that $\int h_I = 0$, $\|h_I\|_2 = 1$ and that $\langle h_I, h_J\rangle = \delta_{I,J}$[6] for all $I, J \in \mathcal{D}$. Furthermore show that if $\langle f, h_I\rangle = 0$ for all $I \in \mathcal{D}$, then $f = 0$ in $L^2$.

We will introduce here two operators that will play the role of $Q_t$ and $P_t$ in the continuous case. Denote the average of a function $f$ on the interval $I$ by $m_I f = \frac{1}{|I|}\int_I f(t)\,dt$. Then the *expectation* operators are defined by

$$E_n f(x) = m_I f, \quad x \in I \in \mathcal{D}_n;$$

and the *difference* operators by

$$\Delta_n f(x) = E_{n+1} f(x) - E_n f(x).$$

EXERCISE 1.24. Show that $\Delta_n f(x) = \sum_{I\in\mathcal{D}_n}\langle f, h_I\rangle h_I(x)$. Show that for all $f \in L^2(\mathbb{R})$

$$E_{N+1} f(x) = \sum_{n\le N}\Delta_n f(x).$$

This provides another proof of the completeness of the Haar system, after observing that $\lim_{n\to\infty} E_n f = f$ in $L^2$; i.e.: $f = \sum_{n\in\mathbb{Z}}\Delta_n f$.

The Haar functions were introduced by A. Haar in 1909, see [**Ha**]. They provide the oldest example of a wavelet basis.

EXERCISE 1.25. Let $I = [k2^{-j}, (k+1)2^{-j})$, show that $h_I(x) = 2^{j/2}h(2^j x - k) = h_{j,k}(x)$, where $h = h_{[0,1]}$.

[6] As usual $\delta_{I,J} = 0$ if $I \neq J$, $\delta_{I,J} = 1$ if $I = J$.

1.5.1. *Dyadic maximal function.* The *dyadic maximal function* is defined as the ordinary maximal function, except that the supremum is taken over the dyadic intervals:

$$M^d f(x) = \sup_{x \in I \in \mathcal{D}} \frac{1}{|I|} \int_I |f(t)|\, dt = \sup_{n \in \mathbb{Z}} E_n f(x).$$

$M^d$ is bounded in $L^\infty$, we will show that it is of weak type $(1,1)$ and by interpolation it will be of strong type $(p,p)$ for all $1 < p < \infty$. The weak type property will be an immediate consequence of the Calderón-Zygmund decomposition to be discussed in the next lecture. The interpolation theorem is also discussed in the next lecture.

It is clear that $M$ is pointwise larger than $M^d$, $M^d f(x) \leq Mf(x)$ for all $x$. Therefore boundedness properties of $M^d$ are deduced from those of $M$. One can actually reverse the process. See [**Duo**] for such an approach.

1.5.2. *Dyadic square function.* The *dyadic square function* is defined formally by

$$S^d f(x) = \left( \sum_{n \in \mathbb{Z}} |\Delta_n f(x)|^2 \right)^{\frac{1}{2}}.$$

EXERCISE 1.26. Show that $\|S^d f\|_2 = \|f\|_2$ (this is a consequence of Plancherel).

THEOREM 1.27. *Let $1 < p < \infty$, then $f \in L^p(\mathbb{R})$ if and only if $S^d f \in L^p(\mathbb{R})$. Moreover $\|f\|_p \sim \|S^d f\|_p$.*

We will follow S. Buckley [**Bu1**] in his proof of this fact, by showing that $S^d f$ is bounded in $L^2(w)$ for all $w \in A_2$, see (1.3); and then a beautiful result of Rubio de Francia, the Extrapolation Theorem, will give boundedness in $L^p$ for all $1 < p < \infty$. We will discuss the extrapolation theorem as well as the proof of Theorem 1.27 in the next lecture.

By Exercise 1.24, $\Delta_n f(x) = \langle f, h_I \rangle h_I(x)$, where $x \in I \in \mathcal{D}_n$, therefore

$$(S^d f)^2(x) = \sum_{x \in I \in \mathcal{D}} \frac{|\langle f, h_I \rangle|^2}{|I|}.$$

From here it is now easy to see that,

COROLLARY 1.28. *$\{h_I\}_{I \in \mathcal{D}}$ is an unconditional basis in $L^p(\mathbb{R})$, for $1 < p < \infty$.*

1.5.3. *Dyadic paraproducts.* Formally the *dyadic paraproduct* is a bilinear operator $\pi^d(b,f) = \pi_b^d f$, given by

$$\pi_b^d f(x) = \sum_{n \in \mathbb{Z}} E_n f(x) \Delta_n b(x) = \sum_{I \in \mathcal{D}} m_I f \langle b, h_I \rangle h_I(x).$$

EXERCISE 1.29. Check formally that $bf = \sum_j \Delta_j b \big[E_j f + (I - E_j) f\big] = \pi_b^d f + \pi_f^d b + \sum_j \Delta_j b \Delta_j f$, so that the paraproduct can be thought as "half a product".

As mentioned before, the paraproduct will behave better than the ordinary product, in the sense that we do not need $b$ to be bounded to obtain boundedness in $L^p$. The substitute for $L^\infty$ in this dyadic world will be *dyadic* $BMO^d$. A locally integrable function $b$ is in $BMO^d$ if

$$\|b\|_{BMO^d} = \left( \sup_{I \in \mathcal{D}} \frac{1}{|I|} \int_I |b(x) - m_I b|^2 dx \right)^{\frac{1}{2}} < \infty.$$

EXERCISE 1.30. Show that for $b \in BMO^d$,

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

EXERCISE 1.31. Show that $BMO^d$ is strictly larger than $BMO$.

THEOREM 1.32. *Given $b \in BMO^d$, then $\pi_b$ is bounded in $L^p(\mathbb{R})$, for $1 < p < \infty$. Moreover, $\|\pi_b^d f\|_p \le C_p \|b\|_{BMO^d} \|f\|_p$.*

For $p = 2$ this theorem is an immediate consequence of Carleson's Lemma that we will prove in the fifth lecture.

A positive sequence $\{\lambda_I\}_{I\in\mathcal{D}}$ is a *Carleson sequence* if there exists a constant $C > 0$ such that for all $I \in \mathcal{D}$, $\sum_{J\in\mathcal{D}(I)} \lambda_J \le C|I|$.

LEMMA 1.33 (Carleson's Lemma). *Let $\{\lambda_I\}_{I\in\mathcal{D}}$, be a* Carleson sequence. *Given any positive sequence $\{a_I\}_{I\in\mathcal{D}}$, let $a^*(x) = \sup_{x\in I\in\mathcal{D}} a_I$; then*

$$\sum_{I\in\mathcal{D}} a_I \lambda_I \le C \int_{\mathbf{R}} a^*(x)\, dx, \tag{1.5}$$

Given $b \in BMO^d$, let $\lambda_I = |\langle b, h_I\rangle|^2$, and $a_I = m_I^2 f$, then the sequence $\{\lambda_I\}$ is Carleson with constant $C = \|b\|_{BMO^d}^2$, and $a^*(x) = (M^d f(x))^2$. By Carleson's Lemma and the boundedness in $L^2$ of the dyadic maximal function we obtain the boundedness in $L^2$ of the dyadic paraproduct,

$$\|\pi_b^d f\|_2^2 = \sum_{I\in\mathcal{D}} m_I^2 f\, b_I^2 \le \|b\|_{BMO^d}^2 \|M^d f\|_2^2 \le \|b\|_{BMO^d}^2 \|f\|_2^2.$$

For $p \ne 2$ one can use Littlewood-Paley theory (square functions) plus Carleson's Lemma. Instead we will show in the last lecture that $\pi_b^d$ is bounded in $L^2(w)$ for all $w \in A_2$ and invoke the extrapolation theorem, see Section 6.2.2. Alternatively we will prove that $\pi_b^d$ is of weak type $(1,1)$; by interpolation this time we can show that it is of strong type $(p,p)$ for $1 < p < 2$. If we could show the same for its adjoint, then a duality argument will give us the range $2 < p < \infty$.

EXERCISE 1.34. Show that the adjoint of $\pi_b$ is given by

$$\pi_b^* g(x) = \sum_{I\in\mathcal{D}} \langle g, h_I\rangle b_I \frac{\chi_I(x)}{|I|}.$$

Both the paraproduct and its adjoint are linear operators which are bounded in $L^2$, and such that for every dyadic interval $I$, the image under either of them is supported on $I$, we will see in the next lecture that this implies that they are of weak type $(1,1)$, see Lemma 2.10.

EXERCISE 1.35. Check formally that $\pi_b^d 1 = b$ and that $(\pi_b^d)^* 1 = 0$.

1.5.4. *Haar multipliers and the Hilbert transform.* The operators in this section do not have a priori a continuous analogue. We will see that averages over random dyadic grids of appropriate Haar multipliers will give back the Hilbert transform.

A *Haar multiplier* is an operator defined formally by

$$Tf(x) = \sum_{I\in\mathcal{D}} w_I(x) \langle f, h_I\rangle h_I(x);$$

where the symbol $w_I(x)$ is a function of both the space and the "frequency" variables $(x, I)$. This is completely analogous to pseudodifferential operators where the Haar system has been replaced by the trigonometric functions:

$$\Phi f(x) = \int_R a(x,\xi)\hat{f}(\xi)e^{2\pi i x\xi}d\xi,$$

the *symbol* here is $a(x,\xi)$. In both cases one would like to identify those symbols for which the corresponding operators are bounded in, for example, $L^p(\mathbb{R})$.

The simplest examples correspond to $w_I(x) = w_I$ and $w_I(x) = w(x)$. In the first case we encounter *constant Haar multipliers*; and in the second multiplication by $w$.

EXERCISE 1.36. Let $T_\alpha$ be a constant Haar multiplier,

$$T_\alpha f(x) = \sum_{I\in\mathcal{D}} \alpha_I \langle f, h_I\rangle h_I(x).$$

Show that $T$ is bounded in $L^p$ if and only if $\{\alpha_I\}_{I\in\mathcal{D}} \in l^\infty$. (Test action of $T$ on the Haar functions for the necessity. For the sufficiency use Plancherel for $p = 2$, and the dyadic square function theorem for $p \neq 2$.)

A class of Haar multipliers corresponding to the symbol $w_I(x) = \left(\frac{w(x)}{m_I w}\right)^{-\frac{1}{2}}$ is known to be bounded in $L^2$ if and only if the weight $w \in A_2^d$ (for the necessity check the action on Haar functions). These multipliers appeared in the study of weighted inequalities, see [**TV1**]. For this and other multipliers see [**KP1**]. We will say more about this type of multipliers in the last lecture.

Constant Haar multipliers are considered models for singular integral operators. In particular, the family of constant Haar multipliers given by choices of sign $\sigma$:

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

where here $\sigma_I = \pm 1$, has proved to be a very good model. Heuristically it is expected that if certain estimates can be done uniformly on $\sigma$ for this family, then the same estimates will hold for the Hilbert transform. That has been the driving force behind the work of Nazarov, Treil and Volberg, see [**NTV2**]. But the passage from the multipliers to the Hilbert transform was not at all obvious. Very recently, Stephanie Petermichl, a student of Volberg, showed in her PhD Thesis [**Pet1**] that certain averages over translated and dilated dyadic grids of the operator

$$H_{\mathcal{D}} f(x) = \sum_{I\in\mathcal{D}} \langle f, h_I\rangle \big[h_{I_r}(x) - h_{I_l}(x)\big],$$

produce a non-zero operator which has the following properties: (1) It commutes with all translations and dilations, and (2) It is antisymmetric. It turns out that the only operators with such properties are constant multiples of the Hilbert transform, rendering now the heuristic very precise, see [**Pet2**]. Similar results have shown to hold for the Riesz transforms, the higher dimensional analogues of the Hilbert transform, see [**PV**].

## 2. Classical Tools

In this lecture we will introduce some classical tools to prove $L^p$ or weak boundedness of a given linear or sublinear operator. These are well known results and

can be found in most harmonic analysis books. Staying faithful to the lectures, we have decided to present the proofs of some of these results because they are very elegant, short, and anybody interested in the subject should see them in detail at least once. We illustrate the use of these results proving boundedness results for the Hilbert transform, the dyadic maximal function, the dyadic square function, the dyadic paraproduct and some Haar multipliers.

**2.1. Schur's Lemma.** Consider the integral operator $T$ with kernel $k(x,y)$, i.e. formally

$$Tf(x) = \int k(x,y)f(y)\,dy.$$

We seek conditions on the kernel for boundedness in $L^p$. That is, under which conditions there exists a constant $C > 0$ such that for all $f \in L^p$

$$\|Tf\|_p \leq C\|f\|_p.$$

We denote by $\|T\| = \|T\|_{L^p \to L^p}$ the operator norm, that is the infimum of the constants $C$ in the above inequality. Let $p'$ denote the dual index to $p$, $\frac{1}{p} + \frac{1}{p'} = 1$.

LEMMA 2.1 ( Schur's Lemma). *Suppose there exist measurable positive functions $w_1, w_2$, and positive constants $C_1, C_2$ such that for a.e. $x, y$*

$$\begin{aligned} \int w_1(x)|k(x,y)|\,dx &\leq C_1 w_2(y), \\ \int w_2^{\frac{p'}{p}}(y)|k(x,y)|\,dy &\leq C_2 w_1^{\frac{p'}{p}}(x). \end{aligned}$$

*Then $T$ is bounded in $L^p$. Moreover $\|Tf\|_p \leq C_1^{\frac{1}{p}} C_2^{\frac{1}{p'}} \|f\|_p$.*

PROOF. By duality it is enough to check that

$$\int |g(x)||Tf(x)|\,dx \leq C_1^{\frac{1}{p}} C_2^{\frac{1}{p'}} \|f\|_p \|g\|_{p'}.$$

Indeed, multiplying by one and using Hölder's inequality:

$$\begin{aligned} &\int |g(x)||Tf(x)|\,dx \\ &= \int\int |g(x)| w_1^{-\frac{1}{p}}(x) w_2^{\frac{1}{p}}(y)|k(x,y)||f(y)| w_1^{\frac{1}{p}}(x) w_2^{-\frac{1}{p}}(y)\,dx\,dy \\ &\leq \left(\int |g(x)|^{p'} w_1^{-\frac{p'}{p}}(x) \int w_2^{\frac{p'}{p}}(y)|k(x,y)|\,dy\,dx\right)^{\frac{1}{p'}} \times \\ &\qquad \left(\int |f(y)|^p w_2^{-1}(y) \int w_1(x)|k(x,y)|\,dx\,dy\right)^{\frac{1}{p}} \\ &\leq C_1^{\frac{1}{p}} C_2^{\frac{1}{p'}} \|f\|_p \|g\|_{p'}. \end{aligned}$$

□

This can be thought as a weighted version of a familiar linear algebra result: if the sums along the rows and along the columns of a matrix are uniformly bounded ($w_i = 1$) then we get boundedness in $L^p$. We will state Schur's lemmma for an operator given by its infinite matrix in the Haar basis in $L^2$ without weights; its

proof and the statement of its generalizations to $L^p$ and weights is left as an exercise for the reader.

LEMMA 2.2 (Dyadic Schur's Lemma). *Suppose that*

$$\sup_J \sum_{I\in\mathcal{D}} |\langle Th_I, h_J\rangle| + \sup_I \sum_{J\in\mathcal{D}} |\langle Th_I, h_J\rangle| \leq \infty.$$

*Then $T$ is bounded in $L^2$.*

EXERCISE 2.3. Prove the Dyadic Schur's Lemma. State and prove the analogue to Lemma 2.1 for the dyadic case.

When the kernel is a convolution kernel, $k(x,y) = k(x-y)$, and the weights are equal to one, both conditions reduce to integrability of $k(t)$, and the lemma is nothing more than Young's inequality $\|k * f\|_p \leq \|k\|_1 \|f\|_p$.

We could say that an integral operator is "trivially" bounded if the weights $w_1, w_2$ are easy to guess. Some analysts become experts on finding very complicated weights. Another approach is to break the operator into a sum of "trivial" pieces that can be handled with Schur's Lemma and hope that the interactions are small enough so that one can superimpose the estimates. That is the content of the next lemma.

**2.2. Cotlar's Lemma.** The notion of *almost-orthogonality* and the Almost-Orthogonal Lemma for self-adjoint operators were introduced by M. Cotlar in 1955 in a hard to find argentinean journal [**Co**]. It has become a tool of wide use in analysis. The statement we present is a generalization due to E. Stein.

LEMMA 2.4 (Cotlar-Stein Lemma ). *Let $\mathcal{H}$ be a Hilbert space, $\{T_j\}$ a sequence of bounded operators in $\mathcal{H}$, $\{T_j^*\}$ their adjoints* [7]. *Let $\{a(j)\}$ be a sequence of positive numbers such that $A = \sum_{j\in\mathbb{Z}} \sqrt{a(j)} < \infty$, and*

$$\|T_i T_j^*\| + \|T_i^* T_j\| \leq a(i-j).$$

*Then* $\|\sum_{j=-\infty}^{\infty} T_j\| \leq A.$

PROOF. It is enough to show that for all integers $n < m$ is true that

$$\|\sum_{j=n}^{m} T_j\| \leq A.$$

Denote by $S = \sum_{j=n}^m T_j$. Remember that $\|S\| = \|SS^*\|^{\frac{1}{2}}$, where now $SS^*$ is selfadjoint. Applying the same norm identity $N$ times we obtain $\|S\| = \|(SS^*)^k\|^{\frac{1}{2k}}$, for $k = 2^N$. But

$$(SS^*)^k = \sum_{j_1,\ldots,j_{2k}=n}^{m} T_{j_1} T_{j_2}^* \ldots T_{j_{2k-1}} T_{j_{2k}}^*.$$

We are going to estimate the norm of each summand in two different ways. Remember that from the hypothesis all the operators are uniformly bounded $\|T_j\| = \|T_j T_j^*\|^{\frac{1}{2}} \leq \sqrt{a(0)}$; also

$$\begin{aligned} \|(T_{j_1} T_{j_2}^*) \ldots (T_{j_{2k-1}} T_{j_{2k}}^*)\| &\leq a(j_1 - j_2) \ldots a(j_{2k-1} - j_{2k}), \\ \|T_{j_1}(T_{j_2}^* T_{j_3}) \ldots (T_{j_{2k-2}}^* T_{j_{2k-1}}) T_{j_{2k}}^*\| &\leq \sqrt{a(0)} a(j_2 - j_3) \ldots a(j_{2k-2} - j_{2k-1}) \sqrt{a(0)} \end{aligned}$$

[7] As usual $\langle T_j x, y\rangle_{\mathcal{H}} = \langle x, T_j^* y\rangle_{\mathcal{H}}$.

The left-hand sides are the same, so we can estimate it by the geometric mean of the right-hand sides:

$$\|T_{j_1}T_{j_2}^*\dots T_{j_{2k-1}}T_{j_{2k}}^*\| \le \sqrt{a(0)a(j_1-j_2)a(j_2-j_3)\dots a(j_{2k-1}-j_{2k})}.$$

Remember that for all $n<m$ and for any $i$, $\sum_{j=n}^m \sqrt{a(i-j)} \le A$. Therefore

$$\begin{aligned}
\|S\|^{2k} &= \|(SS^*)^k\| \\
&\le \sum_{j_1,\dots,j_{2k-1}=n}^m \sqrt{a(0)\dots a(j_{2k-2}-j_{2k-1})}\left(\sum_{j_{2k}=n}^m \sqrt{a(j_{2k-1}-j_{2k})}\right) \\
&\le \dots \le \sqrt{a(0)}\sum_{j_1=n}^m\left(\sum_{j_2=n}^m \sqrt{a(j_1-j_2)}\right)A^{2k-2} \\
&\le \sqrt{a(0)}A^{2k-1}(m-n+1)
\end{aligned}$$

Taking $2k$-th root on both sides and letting $k\to\infty$, we see that $\|S\|\le A$. □

It should be clear that if the operators $T_j$ are orthogonal to each other and uniformly bounded by $M$, the conditions of the lemma are immediately satisfied with $a(0)=M$, and $a(j)=0$ for all $j\neq 0$. That explains why this is called the Almost Orthogonal Lemma. As an example we will show how it applies to the Hilbert transform.

2.2.1. *$L^2$ boundedness of the Hilbert transform.*

PROOF. Let $\mathcal{H}=L^2(\mathbb{R})$. For $f\in L^2(\mathbb{R})$, and any integer $j$, let

$$H_jf(x)=\int_{2^j<|t|\le 2^{j+1}}\frac{f(x-t)}{t}dt = f*k_j(x),$$

where the convolution kernel is given by

$$k_j(t)=\frac{\chi_{\Delta_j}(t)}{t},\quad \text{and}\quad \Delta_j=\{t\in\mathbb{R}: 2^j<|t|\le 2^{j+1}\}.$$

EXERCISE 2.5. Show that $k_j\in L^1(\mathbb{R})$. Moreover $\|k_j\|_1<M$ for all $j\in\mathbb{Z}$. Therefore $H_j$'s are uniformly bounded in $L^2$.

Next observe that each $H_j$ is almost selfadjoint, $H_j^*=-H_j$. Therefore to check the hypothesis of Cotlar's Lemma all we need to do is to estimate $\|H_iH_j\|$. But $H_iH_jf=(k_i*k_j)*f$, so there is always the trivial bound given by Young's inequality: $\|H_iH_jf\|_2\le\|k_i*k_j\|_1\|f\|_2$.

EXERCISE 2.6. Check that $\|k_i*k_j\|_1\le C2^{-|i-j|}$. (This is done carefully estimating $|k_i*k_j(x)|$ pointwise and then integrating.)

Now Cotlar's Lemma can be invoked to conclude that $H$ is bounded in $L^2$.

□

One can actually prove the boundedness of the Hilbert transform just using the Dyadic Schur's Lemma. This was observed by Coifman and Semmes in [**CJS**]. The same argument can be used to show boundedness of the Cauchy integral on Lipschitz curves and the $T(b)$ theorem; provided one shows that a certain Haar system adapted to the geometry of the Lipschitz curve is an unconditional basis in $L^2$, you can consult the original sources, or for a very fresh discussion of this

and many other topics, see [**Tao**]. We will come back to these points in the fourth lecture.

For a multilinear version of Schur's Lemma and an $L^p$ version of Cotlar's Lemma see [**KP1**]. See also the work by L. Grafakos and R. Torres [**GT**].

**2.3. Calderón-Zygmund decomposition.** In this section we show a decomposition theorem for functions instead of operators. It is an invaluable tool to show weak-type inequalities. Furthermore, it provides the first example of an *stopping time* argument.

LEMMA 2.7 (Calderón-Zygmund Decomposition). *Given a function in $L^1(\mathbb{R})$, $\lambda$ a positive real number. There exist disjoint dyadic intervals $\{J_i\}$ such that:*

(i) $|f(x)| \leq \lambda$ *for a.e* $x \notin \cup J_i$,

(ii) $\lambda < \dfrac{1}{|J_i|}\displaystyle\int_{J_i} |f| \leq 2\lambda$,

(iii) $\displaystyle\sum |J_i| \leq \frac{\|f\|_1}{\lambda}$.

COROLLARY 2.8. *Given $f \in L^1$ and $\lambda > 0$, we can decompose $f = g + b$ where*

(i) *The "good" function $g$ is bounded, $\|g\|_\infty \leq 2\lambda$,*

(ii) *The "bad" function $b = \sum b_i$ is such that $supp(b_i) \subset J_i$, $\int b_i = 0$ and $\int |b_i| \leq 4\lambda |J_i|$.*

EXERCISE 2.9. Check that if we define

$$g(x) = \begin{cases} f(x) & x \notin \cup J_i \\ m_{J_i} f & x \in J_i \end{cases}, \quad b_i = (f - m_{J_i} f)\chi_{J_i};$$

the decomposition of the corollary is fulfilled. Check also that since $f \in L^1$ then $g \in L^2$, and $\|g\|_2^2 \leq 2\lambda \|f\|_1$.

PROOF OF THE CALDERÓN-ZYGMUND DECOMPOSITION. Since $f \in L^1$ there exists some integer $N$ such that $m_I|f| \leq \lambda$ for all $I \in \mathcal{D}_N$ (clearly on larger intervals the same will hold!). Consider now the following *stopping time*: for each $I \in D_N$, look at all dyadic subintervals $J \subset I$, considering first the kids, then the grandchildren, etc., ask whether $m_J|f| > \lambda$? If the answer is YES, stop and let $J = J_i$ (don't look into the subintervals of $J_i$). If the answer is NO, continue asking to the subintervals of $J$. By construction the intervals $\{J_i\}$ are the *maximal dyadic intervals* for which $m_J|f| > \lambda$ and therefore they must be *disjoint*. To reach $J_i$ we must have answered NO when we checked all its ancestors, that is $m_J|f| \leq \lambda$ for all $J_i \subset J$. In particular for $\tilde{J}_i$ the parent of $J_i$, and we prove (ii):

$$\lambda < \frac{1}{|J_i|}\int_{J_i} |f| \leq \frac{2}{|\tilde{J}_i|}\int_{\tilde{J}_i} |f| \leq 2\lambda.$$

Also observe that (iii) is also an immediate consequence of our choice of intervals,

$$\|f\|_1 \geq \sum_i \int_{J_i} |f| > \sum_I |J_i|\lambda.$$

Finally for points $x \notin \cup J_i$ it holds that $m_J|f| \leq \lambda$ for all dyadic interval $J \ni x$, therefore by the Lebesgue Differentiation Theorem $|f(x)| \leq \lambda$ for a.e. $x \notin \cup J_I$.

□

In the first lecture we stated that the Lebesgue Differentiation Theorem could be deduced from the weak $(1,1)$ boundedness of the Hardy-Littlewood maximal function, see Section 1.2.4. We will present in the next section the classical proof relying on a covering lemma instead of on the Calderón-Zygmund decomposition.

In the previous Lecture we mentioned the following result, here is the proof.

LEMMA 2.10. *Let $T$ be a linear or sublinear operator which is of strong type $(2,2)$. Suppose that for every dyadic interval $I$, the function $Th_I$ is supported only on $I$. Then $T$ is of weak type $(1,1)$.*

PROOF. Suppose we have $f \in L^1(\mathbb{R})$. We pick $\lambda > 0$. We now apply the Calderón-Zygmund decomposition. We write $f = g + b$ with $\|g\|_\infty \leq \lambda$, $\|g\|_1 \leq 2\|f\|_1$, and $b$ supported on a disjoint sequence of dyadic intervals $\{I_j\}$, having mean zero on those intervals, and so that $\sum_j |I_j| \leq \lambda^{-1}||f||_1$. Now observe that

$$|\{x : |(Tf)(x)| \geq \lambda\}| \leq |\{x : |(Tg)(x)| \geq \frac{\lambda}{2}\}| + |\{x : |(Tb)(x)| \geq \frac{\lambda}{2}\}|.$$

For the first term on the right, we apply the fact that $T$ is of strong type (2,2) and hence of weak type (2,2) and that by Hölder's inequality

$$\|g\|_2 \leq \|g\|_1^{\frac{1}{2}} \|g\|_\infty^{\frac{1}{2}} \leq \sqrt{2\lambda \|f\|_1}.$$

Thus,

$$|\{x : |(Tg)(x)| \geq \frac{\lambda}{2}\}| \leq \frac{4C||g||_2^2}{\lambda^2} \leq \frac{8C||f||_1}{\lambda}.$$

On the other hand $\langle b, h_J \rangle = 0$ for $J$ dyadic unless $J \subset I_j$ for some $j$. Thus by assumption,

$$|\{x : |(Tb)(x)| \geq \frac{\lambda}{2}\}| \leq |\cup_j I_j| \leq \frac{||f||_1}{\lambda}.$$

All these estimates show that $T$ is of weak type (1,1) which was to be shown. □

Clearly, the constant Haar multipliers, the dyadic paraproduct and its adjoint satisfy the hypotheses of Lemma 2.10 Therefore they are of weak type (1,1).

2.3.1. *Maximal and dyadic maximal function are of weak type* $(1,1)$. Remember the definitions of the maximal and dyadic maximal functions:

$$Mf(x) = \sup_{x \in I} \frac{1}{I} \int_I |f|, \qquad M^d f(x) = \sup_{x \in I \in \mathcal{D}} \frac{1}{I} \int_I |f|.$$

As an immediate corollary of the Calderón-Zygmund decomposition we obtain the following:

LEMMA 2.11. *$M^d$ is of weak-type $(1,1)$.*

PROOF. Given a function $f \in L^1$, we want to estimate the size of the set $E_\lambda^d = \{x \in \mathbb{R} : M^d f(x) > \lambda\}$. Consider the maximal dyadic intervals for which $m_I|f| > \lambda$; this are exactly the intervals $\{J_i\}$ given by the Calderón-Zygmund decomposition. Clearly $x \in J_i$ for some $i$ if and only if $M^d f(x) > \lambda$, hence $E_\lambda^d = \cup J_i$ and $|E_\lambda^d| = \sum |J_i| \leq \frac{\|f\|_1}{\lambda}$. Hence $M^d$ is of weak type $(1,1)$. □

THEOREM 2.12 (Hardy-Littlewood [1930]). *$M$ is of weak-type $(1,1)$.*

PROOF. Consider now the set $E_\lambda = \{x \in \mathbb{R} : Mf(x) > \lambda\}$. $E_\lambda$ is an open set[8]. Therefore we can approximate from inside with compact sets $K \subset E_\lambda$. Clearly $K \subset \cup_{x \in E_\lambda} I_x = E_\lambda$, and by compactness there is a finite subcovering $\{I_1, \dots, I_n\}$ of $K$.

EXERCISE 2.13. Let $\{I_1, \dots, I_n\}$ be a finite family of open intervals in $\mathbb{R}$. Show that there is a subfamily $\{J_1, \dots, J_k\}$ such that the $J_i$ are pairwise disjoint and such that $\sum_{i=1}^k |J_i| \geq \frac{1}{2} \left| \bigcup_{j=1}^n I_j \right|$. (see [**Gar**] p.25).

Choose the pairwise disjoint intervals $\{J_i\}$ given by the exercise, notice that each of them satisfies the inequality $m_{J_i}|f| > \lambda$. Now we can estimate the size of $K$, for all $K \subset E_\lambda$

$$|K| \leq \left| \bigcup_{j=1}^n I_j \right| \leq 2 \sum_{i=1}^k |J_i| \leq 2 \sum_{i=1}^k \frac{1}{\lambda} \int_{J_i} |f| \leq 2 \frac{\|f\|_1}{\lambda}.$$

Which shows that $|E_\lambda| \leq 2\frac{\|f\|_1}{\lambda}$, hence the Hardy-Littlewood maximal function is of weak-type $(1,1)$. □

The original proof uses a beautiful argument based on the so-called *Rising Sun Lemma*, it can be found in [**Koo**] p.234.

2.3.2. *The Hilbert transform is of weak type* $(1,1)$.

THEOREM 2.14 (Kolmogorov). *The Hilbert transform is of weak-type* $(1,1)$.

PROOF. Fix $\lambda > 0$, assume $f \geq 0$, and $f \in L^1$. By the corollary to the Calderón-Zygmund Decomposition there exists $\{J_1, J_2, \dots\}$ disjoint dyadic intervals such that $f = g + b$ with the properties listed there. In particular, $g \in L^2$ and $\|g\|_2^2 \leq 2\lambda\|f\|_1$. It should be clear that

$$|\{x : |Hf(x)| > \lambda\}| \leq |\{x : |Hg(x)| > \frac{\lambda}{2}\}| + |\{x : |Hb(x)| > \frac{\lambda}{2}\}|.$$

Since $g \in L^2$ and we just showed that $H$ is bounded in $L^2$ then by Tchebychev's inequality (if you prefer strong $(2,2)$ implies weak $(2,2)$)

$$|\{x : |Hg(x)| > \frac{\lambda}{2}\}| \leq \frac{4}{\lambda^2}\|g\|_2^2 \leq \frac{8}{\lambda}\|f\|_1.$$

One can also show that

$$|\{x : |Hb(x)| > \frac{\lambda}{2}\}| \leq \frac{c}{\lambda}\|f\|_1;$$

but this is more delicate. We will prove a stronger result in the fourth lecture, see Lemma 4.1. Together these two estimates show that $H$ is of weak-type $(1,1)$. □

[8] If $x \in E_\lambda$ then there is an open interval $I_x \ni x$ such that $m_{I_x}|f| > \lambda$ but this means that $I_x \in E_\lambda$.

**2.4. Interpolation.** This tool has already been mentioned. It is extremely useful and there are books dedicated to the subject. We present here the classical Marcinkiewicz Interpolation Theorem.

THEOREM 2.15 (Marcinkiewicz Interpolation Theorem [1939]). *Let $1 \le p < q$. Let $T$ be a sublinear operator defined on $L^p + L^q$. Suppose that $T$ is of weak-type $(p,p)$ and $(q,q)$ (when $q = \infty$ we replace weak by strong-type). Then $T$ is of strong-type $(r,r)$ for all $p < r < q$.*

Immediate consequences of this theorem are the boundedness in $L^p$ of the Hardy-Littlewood maximal function and the Hilbert transform for all $1 < p < \infty$. We have shown that $M$ is bounded in $L^\infty$ and is of weak-type $(1,1)$, therefore $M$ is bounded in $L^p$ for $1 < p < \infty$. We have also shown that $H$ is bounded in $L^2$ and is of weak-type $(1,1)$, therefore $H$ is bounded in $L^p$ for $1 < p < 2$, but this implies that its adjoint is bounded on the dual spaces that is in $L^q$ for $2 < q < \infty$, but $H^* = -H$; therefore $H$ is bounded on the full range $1 < p < \infty$.

PROOF. We will prove the interpolation theorem assuming that the operator $T$ is of strong-type $(p,p)$ and $(q,q)$ (this result is the so-called Riesz-Thorin Interpolation Theorem, proved several years before Marcinkiewicz theorem). It will be obvious from the proof that all that is required is the weak-type condition.

Given $p < r < q$, there is a $0 < t < 1$ such that

$$\frac{1}{r} = \frac{1-t}{p} + \frac{t}{q}.$$

By hypothesis $\|Tf\|_p \le M_p \|f\|_p$ and $\|Tf\|_q \le M_q \|f\|_q$. We will show that $\|Tf\|_r \le C M_p^{1-t} M_q^t \|f\|_r$. A first attempt to use Hölder's inequality shows that $\|Tf\|_r \le C M_p^{1-t} M_q^t \|f\|_p^{1-t} \|f\|_q^t$... a more subtle idea is needed.

Let $E_\lambda(f) = \{x \in \mathbb{R} : |f(x)| > \lambda\}$, and let $d_f(\lambda) = |E_\lambda(f)|$ be the *distribution function* of the function $f$.

EXERCISE 2.16. Show that for $1 \le p < \infty$, $\|f\|_p^p = p \displaystyle\int_0^\infty \lambda^{p-1} d_f(\lambda)\, d\lambda$.

EXERCISE 2.17. Given $\phi > 0$, an increasing and differentiable function, show that $\displaystyle\int_{\mathbf{R}} \phi(f)(x)\, dx = \int_0^\infty \phi'(\lambda) d_f(\lambda)\, d\lambda$.

For each $\lambda$ decompose $f = f_p^\lambda + f_q^\lambda$, where $f_p^\lambda(x) = f(x)\chi_{\{x : |f(x)| > c\lambda\}}$ hence $f_q^\lambda = f(x)\chi_{\{x : |f(x)| \le c\lambda\}}$, for $c$ a positive constant to be chosen later.

EXERCISE 2.18. Show that if $f \in L^r$, $p < r < q$, then $f_p^\lambda \in L^p$ and $f_q^\lambda \in L^q$. Moreover show that $\|f_p^\lambda\|_p^p \le (c\lambda)^{p-r} \|f\|_r^r$, and $\|f_q^\lambda\|_q^q \le (c\lambda)^{q-r} \|f\|_r^r$.

Also $|T(f_p^\lambda + f_q^\lambda)| \le |T(f_p^\lambda)| + |T(f_q^\lambda)|$, therefore $d_{Tf}(\lambda) \le d_{Tf_p^\lambda}(\frac{\lambda}{2}) + d_{Tf_q^\lambda}(\frac{\lambda}{2})$.

Case $q = \infty$: Let $c = (2M_\infty)^{-1}$, then clearly $d_{Tf_q^\lambda}(\frac{\lambda}{2}) = 0$. By Tchebychev's inequality and the boundedness in $L^p$ of $T$, $d_{Tf_p^\lambda}(\frac{\lambda}{2}) \le \big(2M_p \lambda^{-1} \|f_p^\lambda\|_p\big)^p$, which is just a weak-type $(p,p)$ inequality! Therefore by Exercise 2.16,

$$\|Tf\|_r^r = r \int_0^\infty \lambda^{r-1} d_{Tf}(\lambda)\, d\lambda \le r \int_0^\infty \lambda^{r-1} d_{Tf_p^\lambda}\big(\frac{\lambda}{2}\big)\, d\lambda.$$

EXERCISE 2.19. Show that the right hand side of the last inequality is *equal* to $r(r-p)^{-1}(2M_p)^p (2M_\infty)^{r-p} \|f\|_r^r$.

Therefore $\|Tf\|_r \leq C M_p^{1-t} M_\infty^t \|f\|_r$, for all $p < r$.

EXERCISE 2.20. Do the case $q < \infty$. Observe that the proof works if instead of assuming strong-type inequalities we assume weak-type inequalities.

□

COROLLARY 2.21. *Let $\{\alpha_I\}$ be a sequence satisfying $|\alpha_I| \leq A$, for every $I$. Then the Haar multiplier with coefficients given by $\alpha$,*

$$T_\alpha f = \sum_{I \in \mathcal{D}} \alpha_I \langle f, h_I \rangle h_I$$

*is bounded in $L^p(\mathbb{R})$ for all $1 < p < \infty$.*

PROOF. By Lemma 2.10, $T_\alpha$ is of weak-type $(1,1)$, and is bounded in $L^2$ by hypothesis. Thus by Marcinkiewicz Interpolation Theorem $T_\alpha$ is bounded on $L^p(\mathbb{R})$ for all $1 < p \leq 2$. But $T_\alpha$ is selfadjoint, hence $T_\alpha$ is bounded on $L^p(\mathbb{R})$ for all $1 < p < \infty$. □

It is plain that the boundedness of $\{\alpha_I\}$ is also necessary for the boundedness of $T_\alpha$ on any $L^p$ space as may be seen by applying $T_\alpha$ to the Haar basis.

2.4.1. *Riesz's proof of $L^p$ boundedness of the Hilbert transform.* Riesz proved in 1927 the $L^p$ boundedness of the Hilbert transform for $p > 1$ using only the Riesz-Thorin Interpolation Theorem and the boundedness in $L^2$ of $H$. The proof is so beautiful that we want to sketch it here. This proof is based on the following lemma:

LEMMA 2.22. *If $H$ is bounded in $L^p$ then it is bounded in $L^{2p}$.*

RIESZ'S PROOF OF $L^p$ BOUNDEDNESS OF THE HILBERT TRANSFORM. We know $H$ is bounded in $L^2$ then is bounded in $L^4$, $L^8$, ..., $L^{2^n}$ for all $n$. Then by strong interpolation $H$ is bounded for $2 < p < 2^n$ for all $n$, hence for all $p > 2$ and by duality its adjoint $H^* = -H$ is bounded for all $1 < p < 2$. □

PROOF OF THE CASE $p = 2$. Given $f$ in the Schwartz class (or if you prefer a compactly supported $C^\infty$ function), so that $f \in L^p$ for all $p > 1$ and the class is dense in each space. Let $u$ be its harmonic extension to the upper half plane and $v$ its harmonic conjugate. Remember that as $t \to 0$ for $z = x + it$ then $u \to f$ and $v \to Hf$. Then the complex valued function $F(z) = u(z) + iv(z)$ is analytic in $R^2_+$. But so is its square: $F^2 = u^2 - v^2 + 2iuv$. Taking limit as $t \to 0$ of the real and imaginary parts of $F^2$ we obtain,

$$\lim_{t \to 0} \operatorname{Re} F^2(z) = \lim_{t \to 0} (u^2 - v^2) = f^2 - (Hf)^2,$$

$$\lim_{t \to 0} \operatorname{Im} F^2(z) = \lim_{t \to 0} 2uv = 2f\,Hf.$$

On the other hand it is also true that

$$\lim_{t \to 0} \operatorname{Im} F^2(z) = H(\lim_{t \to 0} \operatorname{Re} F^2(z)) = H(f^2 - (Hf)^2).$$

Therefore $2f\,Hf = H(f^2 - (Hf)^2)$. Remembering that $H^2 = -I$ we conclude that $(Hf)^2 = 2H(f\,Hf) + f^2$. Therefore, by the Cauchy-Schwartz inequality and the

fact that $H$ is an isometry,

$$\begin{aligned}\|Hf\|_4^4 &\leq 2(4\int |H(f\,Hf)|^2 + \int |f|^4) = 8\int |f\,Hf|^2 + 2\|f\|_4^4 \\ &\leq 8\left(\int |f|^4\right)^{\frac{1}{2}}\left(\int |Hf|^4\right)^{\frac{1}{2}} + 2\|f\|_4^4 \;=\; 8\|Hf\|_4^2\|f\|_4^2 + 2\|f\|_4^4.\end{aligned}$$

EXERCISE 2.23. If $A, B < \infty$, $c, d > 0$ and $A \leq c\sqrt{AB} + dB$ then $A \leq kB$ for some other positive constant $k$.

Therefore setting $A = \|Hf\|_4^4 < \infty$ and $B = \|f\|_4^4 < \infty$ and applying the exercise we conclude that $H$ is bounded in $L^4$ (modulo the observation that for functions in the Schwartz class we get again functions in the same class after applying the Hilbert transform). □

EXERCISE 2.24. Show the Lemma for any $p > 1$.

**2.5. Extrapolation.** An alternative to the interpolation theorems, that require two end-points results, is the *extrapolation theorem* of Rubio de Francia. His moto was *there is no $L^p$ but weighted $L^2$.*

Remember that a weight is in $A_2$ if there is a constant $C > 0$ such that for all intervals $I$

$$\left(\frac{1}{|I|}\int_I w\right)\left(\frac{1}{|I|}\int_I w^{-1}\right) < C.$$

THEOREM 2.25 (Rubio de Francia's Extrapolation Theorem). *Assume $T$ is a bounded linear operator in $L^2(w)$ for all weights $w \in A_2$, that is*

$$\int |Tf|^2 w \leq C\int |f|^2 w.$$

*Then $T$ is bounded in $L^p$ for all $1 < p < \infty$.*

For a proof and much more about weighted inequalities, see the classical book [**GC-RF**]. One can actually replace $q = 2$ by any $1 < q < \infty$ provided one replaces the $A_2$-condition by a corresponding $A_q$-condition:

$$\sup_I \left(\frac{1}{|I|}\int_I w\right)\left(\frac{1}{|I|}\int_I w^{\frac{-1}{q-1}}\right)^{q-1} < \infty; \tag{2.1}$$

and one assumes boundedness in $L^q(w)$ for all $A_q$ weights $w$. Actually the conclusion is stronger, it can be shown that $T$ will be bounded in $L^p(w)$ for all $w \in A_p$, and for all $1 < p < \infty$.

We already mentioned in section 1.1.3 that $w \in A_2$ is a necessary and sufficient condition for the boundedness of the Hilbert transform in $L^2(w)$. The same is true for the maximal function. Similarly the necessary and sufficient condition for boundedness in $L^p(w)$ of either $H$ or $M$ is that $w \in A_p$. Remember that the dyadic maximal function $M^d$ is pointwise bounded by the $M$, therefore it will also be bounded in $L^p(w)$ for weights in $A_p$.

Unfortunately the extrapolation theorem says nothing about the endpoints $p = 1$ and $p = \infty$.

2.5.1. *Boundedness of the dyadic square function in $L^p$.* As an illustration of the power of the extrapolation theorem we will present S. Buckley's proof of the boundedness in $L^p$, $1 < p < \infty$ of the dyadic square function

$$S^d f = \left( \sum_{n \in \mathbb{Z}} |\Delta_n f(x)|^2 \right)^{\frac{1}{2}} = \left( \sum_{x \in I \in \mathcal{D}} \frac{|\langle f, h_I \rangle|^2}{|I|} \right)^{\frac{1}{2}};$$

see Section 1.5.2 for the notation, for the original paper see [**Bu1**], and for variations see [**Per**].

PROOF. According to the extrapolation theorem, all we have to do is check that $S^d$ is bounded in $L^2(w)$ for all $w \in A_2$.

EXERCISE 2.26. Check that $\|S^d f\|^2_{L^2(w)} = \sum_{I \in \mathcal{D}} |\langle f, h_I \rangle|^2 m_I w$

Notice that $|\langle f, h_I \rangle|^2 = |m_{I_r} f - m_I f|^2 |I| = |m_{I_l} f - m_I f|^2 |I|$. Also $w \in A_2$ implies that $w$ is a *doubling* weight, that is, we can compare the mass of an interval $I$ and its double $2I$ (or its parent $\tilde{I}$ )[9].

What we need to check is that for all $w \in A_2$ and $f \in L^2(w)$ then

$$2W = \sum_{I \in \mathcal{D}} |m_I f - m_{\tilde{I}} f|^2 w(\tilde{I}) \leq C \int |f|^2 w. \tag{2.2}$$

EXERCISE 2.27. Pairing the terms that have a common parent, show that

$$W = \sum_{I \in \mathcal{D}} (m_I^2 f - m_{\tilde{I}}^2 f) w(\tilde{I}).$$

We are using the fact that for $I, I^*$ children of $\tilde{I}$ then $m_{\tilde{I}} f = \frac{m_I f + m_{I^*} f}{2}$.

Therefore, adding and subtracting $2w(I) m_I^2 f$, we get

$$W = \sum_{I \in \mathcal{D}} \left[ 2w(I) m_I^2 f - w(\tilde{I}) m_{\tilde{I}}^2 f \right] + \sum_{I \in \mathcal{D}} \left[ w(\tilde{I}) - 2w(I) \right] m_I^2 f = W_1 + W_2.$$

EXERCISE 2.28. Check that $W_1 = \sum_{M=-\infty}^{\infty} (a_m - a_{m+1})$ is a telescoping sum, where $a_m = \sum_{I \in \mathcal{D}_m} 2w(I) m_I^2 f = 2 \int (E_m f)^2 w$.

Check also that $a_m \leq 2 \int |M^d f|^2 w \leq C \int |f|^2 w$. (The last inequality is the fact that the maximal function is bounded in $L^2(w)$ if $w \in A_2$, see [**CF**], [**Duo**]).

By the exercise $W_1 \leq C \int |f|^2 w$.

EXERCISE 2.29. Check that $W_2 = \sum_{I \in \mathcal{D}} (w(\tilde{I}) - 2w(I))(m_I^2 f - m_{\tilde{I}}^2 f)$.

Hence, by the Cauchy-Schwartz inequality,

$$\begin{aligned} W_2 &\leq \left( \sum_{I \in \mathcal{D}} \frac{(w(\tilde{I}) - 2w(I))^2}{w(\tilde{I})} (m_I f + m_{\tilde{I}} f)^2 \right)^{\frac{1}{2}} \left( \sum_{I \in \mathcal{D}} w(\tilde{I}) (m_I f - m_{\tilde{I}} f)^2 \right)^{\frac{1}{2}} \\ &= W_3^{\frac{1}{2}} W^{\frac{1}{2}} \leq \frac{W_3 + W}{2}. \end{aligned}$$

[9] more precisely, there is a constant $C > 0$ such that for all intervals: $w(\tilde{I}) = \int_{\tilde{I}} w \leq C w(I)$

Therefore $W \leq C\int |f|^2 w + W_3$. So all it remains to show is that $W_3 \leq C\int |f|^2 w$. Notice that $m_I|f| \leq 2m_{\tilde{I}}|f|$ and let $b_{\tilde{I}}^2 = \frac{(w(\tilde{I})-2w(I))^2}{w(\tilde{I})}$, then

$$W_3 \leq \sum_{I\in\mathcal{D}} w(I) b_I^2 m_I^2 |f|.$$

This sum looks very close to the $L^2$ norm of the dyadic paraproduct introduced in Section 1.5.3. There we showed that boundedness in $L^2$ was a consequence of Carleson's Lemma. In our case we will need a weighted version of this lemma whose proof is similar to the proof of the ordinary Carleson's Lemma that we will discuss in the fifth lecture.

The class $A_\infty$ is defined as the union of the $A_p$ classes for $p > 1$. We will give an alternative definition in the next lecture.

Given a weight $w$, an *$w$-Carleson sequence* $\{\lambda_I\}_{I\in\mathcal{D}}$ is a positive sequence such that there is a constant $C > 0$ such that for all dyadic intervals $I$: $\sum_{J\in\mathcal{D}(I)} \lambda_J \leq Cw(I)$.

LEMMA 2.30 (Weighted Carleson's Embedding Theorem). *Given a weight $w \in A_\infty$ and a $w$-*Carleson sequence $\{\lambda_I\}_{I\in\mathcal{D}}$ *then for all $f \in L^2(w)$ holds*

$$\sum_{I\in\mathcal{D}} \lambda_I m_I^2 |f| \leq C\int |f|^2 w.$$

The weight $w$ in $A_2$ is in $A_\infty$ by definition. Assuming the lemma, all we need to check is that the sequence $\frac{(w(\tilde{I})-2w(I))^2}{w(\tilde{I})}$ is $w$-Carleson. This is S. Buckley characterization of classes of weights by summation conditions [**Bu1**], which we will prove in the last lecture, see p. 48.

□

We will prove again this result in the last lecture using Haar multipliers. There is also a Bellman function proof which yields the sharp constant, see [**Huk**], [**HTV**].

## 3. $BMO$, $A_\infty$ and Stopping Times

**3.1. Dyadic $BMO$ and self-improvement.** Remember that *dyadic $BMO^d$* is the collection of locally integrable functions $b$ such that

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

It is clear that bounded functions are in ordinary $BMO$ (as above without restricting to dyadic intervals), and that $BMO$ is contained in $BMO^d$ (these are Banach spaces when considered modulo the constant functions). The function $b(x) = \log|x|$ is a prototype of a $BMO$ function. It can be checked that the function $b(x) = \log|x|\chi_{(0,\infty)}$ is in $BMO^d$ but not in $BMO$. The celebrated *John-Nirenberg Inequality* says that the function $\log|x|$ is, in terms of distribution functions, typical of unbounded $BMO$ functions.

THEOREM 3.1 (John-Nirenberg Inequality). *Given a function $b \in BMO$, any interval $I$, and a positive number $\lambda > 0$, then there exist positive constants $C_1, C_2 > 0$ (independent of $b, I, \lambda$) such that*

$$|\{x \in I : |b(x) - m_I b| > \lambda\}| \leq C_1 |I| e^{-\frac{C_2\lambda}{\|b\|_{BMO}}}. \tag{3.1}$$

EXERCISE 3.2. Show that if a function $b$ satisfies the John-Nirenberg Inequality (3.1) then $b \in BMO$.

We will prove this classical theorem in the next section.

COROLLARY 3.3 (Self-improvement). *Given $b \in BMO$ then for all $p > 1$ there exists a constant $C_p > 0$ such that for all intervals $I$*

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

PROOF. By Exercise 2.16,

$$\begin{aligned}\int_I |b(x) - m_I b|^p \, dx &= \int_0^\infty p\lambda^{p-1} |\{x \in I : |b(x) - m_I b| > \lambda\}| \, d\lambda \\ &\leq |I| \int_0^\infty p\lambda^{p-1} C_1 e^{-\frac{C_2 \lambda}{\|b\|_{BMO}}} d\lambda \\ &= |I| \|b\|_{BMO}^p C_1 C_2^{-p} p\Gamma(p).\end{aligned}$$

The last equality after performing the change of variable $s = \frac{C_2 \lambda}{\|b\|_{BMO}}$. □

Notice that the reverse inequality is nothing more than Hölder's inequality. Therefore the left hand side provides an alternative definition for the $BMO$ norm. These *reverse Hölder inequalities* can be thought as *self-improvement* inequalities: being in $L^1(I)$ we conclude that we are in $L^p(I)$ for all $p > 1$, this is usually false (the reverse is always true).

The John-Nirenberg Theorem holds in $BMO^d$, provided we restrict our attention to dyadic intervals $I$. Clearly the corollary would still hold for dyadic intervals. Therefore we will often use the alternative norm given by $p = 2$:

$$\|b\|_{BMO^d}^2 \sim \sup_{I \in \mathcal{D}} \frac{1}{|I|} \int_I |b(x) - m_I b|^2 dx = \sup_{I \in \mathcal{D}} \frac{1}{|I|} \sum_{J \in \mathcal{D}} |\langle b, h_J \rangle|^2.$$

The last equality is Exercise 1.30.

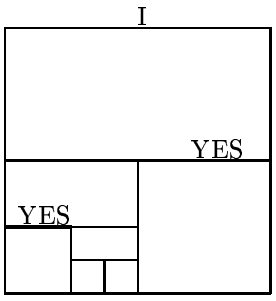
STOPPING TIME PROOF OF THE JOHN-NIRENBERG INEQUALITY. The classical proof iterates the Calderón-Zygmund decomposition ad infinitum. The present proof gives our second example of a *stopping time* argument. We will do the proof for dyadic intervals $I$, in this case all we need is $b \in BMO^d$, but it clearly applies to all intervals when $b \in BMO$.

Fix $I \in \mathcal{D}$, we will define recursively a sequence of mutually disjoint collections of disjoint dyadic subintervals of $I$:

$$\mathcal{J}_0(I) = \{I\}, \ \mathcal{J}_1(I), \ \mathcal{J}_2(I), \ldots$$

The *stopping time* question to select the intervals in $\mathcal{J}_1(I)$ is:

$$\text{Is } \frac{1}{|J|}\int_J |b - m_I b| > 2\|b\|_{BMO^d} \ ?$$




Start at $J = I$, the answer is NO. Ask the question now to the children $J$ of $I$. If the answer is YES, put that interval in $\mathcal{J}_1(I)$ and STOP. If the answer is NO, continue asking to the children of $J$, and repeat. Notice that each $J \in \mathcal{J}_1(I)$ is maximal with respect to the stopping time question, i.e. for all $K \in \mathcal{D}(I)$ such that $J \subset K$, $\frac{1}{|K|}\int_K |b - m_I b| \leq 2\|b\|_{BMO^d}$. By construction the intervals in $\mathcal{J}_1(I)$ are disjoint intervals.

Given $\mathcal{J}_{N-1}(I)$ define

$$\mathcal{J}_N(I) = \bigcup_{J \in \mathcal{J}_{N-1}(I)} \mathcal{J}_1(J).$$

These are disjoint families of dyadic subintervals of $I$. Notice that for each $J \in \mathcal{J}_{N-1}(I)$ we are doing a Calderón-Zygmund decomposition with $f = (b - m_J b)\chi_J$ and $\lambda = 2\|b\|_{BMO^d}$. Therefore the following hold:

(i) $|b(x) - m_J b| \leq 2\|b\|_{BMO^d}$ for a.e. $x \in J \setminus \bigcup_{K \in \mathcal{J}_1(J)} K$,

(ii) $|m_K b - m_J b| \leq \dfrac{1}{|K|}\displaystyle\int_K |b - m_J b| \leq 4\|b\|_{BMO^d}$ for all $K \in \mathcal{J}_1(J)$,

(iii) $\displaystyle\sum_{K \in \mathcal{J}_1(J)} |K| \leq \frac{\|(b - m_J b)\chi_J\|_1}{2\|b\|_{BMO^d}} \leq \frac{|J|}{2}$.

EXERCISE 3.4. Show that for a.e. $x \in I \setminus \bigcup_{K \in \mathcal{J}_N(I)} K$ then $|b(x) - m_I b| \leq 4N\|b\|_{BMO^d}$. Show also that $\displaystyle\sum_{K \in \mathcal{J}_N(I)} |K| \leq \frac{|I|}{2^N}$.

If $\lambda \geq 4\|b\|_{BMO^d}$, then choose an integer $N \geq 1$ such that $4N\|b\|_{BMO^d} \leq \lambda < 4(N+1)\|b\|_{BMO^d}$. Then, by the previous exercise, the subset of $x \in I$ such that $|b(x) - m_I b| > 4N\|b\|_{BMO^d}$ is contained in the $\bigcup_{K \in \mathcal{J}_N(I)} K$ modulo a set of measure zero, therefore,

$$|\{x \in I : |b(x) - m_I b| > \lambda\}| \leq \sum_{K \in \mathcal{J}_N(I)} |K| \leq \frac{|I|}{2^N}.$$

But $2^N \leq e^{\frac{\lambda \ln 2}{4\|b\|_{BMO^d}}} < 2^{N+1}$, thus

$$|\{x \in I : |b(x) - m_I b| > \lambda\}| \leq 2|I| e^{-\frac{\lambda \ln 2}{4\|b\|_{BMO^d}}}.$$

If $\lambda < 4\|b\|_{BMO^d}$ then

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

provided we choose $C_1 \geq e^{4C_2}$. In particular if we choose $C_2 = \frac{\ln 2}{4}$ and $C_1 \geq 2$ we obtain the desired inequality for all $\lambda > 0$. □

**3.2. $A_\infty$ and $A_\infty^d$ weights.** Given a non-negative, locally integrable function $w$, i.e. a *weight*, we say that it is in $A_\infty$ if there is a constant $C > 0$ such that for all intervals $I$ the following *reverse Jensen inequality* holds:

$$\frac{1}{|I|}\int_I w \leq C e^{\frac{1}{|I|}\int_I \ln w}. \tag{3.2}$$

A weight $w$ is in *dyadic* $A_\infty^d$ if the same inequality holds for dyadic intervals.

3.2.1. *$A_\infty$ vs $BMO$.* A typical example of an $A_\infty$ weight is $w = |x|^\alpha$ for $\alpha > -1$. In this case $\ln w = \alpha \ln |x| \in BMO$. This is not just a coincidence.

THEOREM 3.5. *If a weight $w \in A_\infty^d$ then $\ln w \in BMO^d$. If $b \in BMO^d$ then $w = e^{\delta b} \in A_\infty^d$ for $\delta$ small enough.*

PROOF OF THE SECOND PART. We will prove the first part of this theorem using a stopping time argument in Section 3.4. There is an alternative proof using Bellman functions, see [NTV4] (we learned about this proof from Volberg in the Spring School on Analysis held in Paseky, Czech Republic, Spring 2000).

As for the second part it can be seen that it is a consequence of the John-Nirenberg Inequality. Given $b \in BMO^d$, and $I \in \mathcal{D}$, then

$$\frac{1}{|I|}\int_I e^{\delta b(x)}dx = e^{\delta m_I b}\frac{1}{|I|}\int_I e^{\delta b(x)-m_I(\delta b)}dx.$$

If we can show that $\frac{1}{|I|}\int_I e^{\delta b - m_I(\delta b)} < C$ for all dyadic intervals $I$, then we would have shown that $w = e^{\delta b} \in A_\infty^d$. But, by Exercise 2.17 and the John-Nirenberg Inequality applied to $\delta b \in BMO^d$,

$$\begin{aligned}\frac{1}{|I|}\int_I e^{\delta b(x)-m_I(\delta b)}dx &\leq \frac{1}{|I|}\int_0^\infty e^\lambda |\{x \in I : |\delta b(x) - m_I(\delta b)| > \lambda\}| d\lambda + 1 \\ &\leq C_1 \int_0^\infty e^{\lambda\left(1-\frac{C_2}{\delta \|b\|_{BMO^d}}\right)} d\lambda + 1.\end{aligned}$$

The right hand side will be finite as long as $\delta\|b\|_{BMO^d} < C_2$, that is for $\delta$ small enough, $w = e^{\delta b} \in A_\infty^d$. □

3.2.2. *$A_\infty$ vs $A_p$.* Remember that we said that for $1 < p < \infty$, $w \in A_p$ if there is $C > 0$ such that for all intervals $I$

$$\left(\frac{1}{|I|}\int_I w\right)\left(\frac{1}{|I|}\int_I w^{\frac{-1}{p-1}}\right)^{p-1} < C.$$

Typical examples of $A_p$ weights are $w = |x|^\alpha$ for $-1 < \alpha < p-1$.

The dyadic $A_p^d$ classes are defined similarly restricting to dyadic intervals.

By Hölder's inequality $w \in A_p$ implies $w \in A_r$ for all $r > p$ with the same or smaller constant. It was observed by García-Cuerva, see [**GC-RF**], that $A_\infty$ is the natural limit as $p \to \infty$ of $A_p$ (hence the name $A_\infty$).

EXERCISE 3.6. Show that $\lim_{p\to\infty}\left(\frac{1}{|I|}\int_I w^{\frac{-1}{p-1}}\right)^{p-1} = e^{-m_I(\ln w)}$. Show that this implies that if $w \in A_p$ then $w \in A_\infty$.

We can also consider the limiting case when $p \to 1$.

EXERCISE 3.7. Show that $\lim_{p\to 1}\left(\frac{1}{|I|}\int_I w^{\frac{-1}{p-1}}\right)^{p-1} = \|w^{-1}\|_{L^\infty(I)}$.

This leds us to consider the space $A_1$ of non-negative functions $w$ so that for all intervals $I$,

$$\frac{1}{|I|}\int_I w\,dx \leq Cw(x), \qquad \text{for a.e. } x \in I.$$

EXERCISE 3.8. Show that $A_1 \subset A_p$ for all $p > 1$.

The converse to the second part of Exercise 3.6 is also true, therefore $A_\infty = \bigcup_{p>1} A_p$ (same is true for the dyadic counterparts).

EXERCISE 3.9. Show that if $w \in A_\infty$ then there is a $p > 1$ such that $w \in A_p$. This can be deduced from the John-Nirenberg Inequality similarly to the proof of the second part of Theorem 3.5.

3.2.3. *$A_\infty$ vs $RH_p$.* In the non-dyadic world there is another characterization of $A_\infty$ due to Coifman and C. Fefferman [**CF**]. We need to introduce the *reverse Hölder classes* of weights, $RH_p$.

A weight $w \in RH_p$, $1 < p < \infty$, if there is a constant $C > 0$ such that for all intervals $I$

$$\left(\frac{1}{|I|}\int_I w^p\right)^{\frac{1}{p}} \leq C\frac{1}{|I|}\int_I w.$$

Notice that the reverse inequality holds trivially with constant one, by Hölder's inequality. Also it is immediate that if $w \in RH_p$ then $w \in RH_r$ for all $1 < r < p$. Typical examples of $RH_p$ weights are $w = |x|^\alpha$ for $\alpha > -1/p$. It is clear how to define the larger class dyadic $RH_p^d$.

THEOREM 3.10 (Coifman-Fefferman). $A_\infty = \bigcup_{p>1} RH_p$.

For a proof see [**GC-RF**]. Half of this theorem is not true in the dyadic case, unless we assume that the weight is *dyadic doubling*. One can show that $RH_p^d$ does not imply dyadic doubling, on the other hand $w \in A_\infty^d$ does, see [**Bu1**]. In the non-dyadic world $A_\infty$ implies doubling, therefore so do $A_p$ and $RH_p$. One can give an alternative definition of $A_\infty$ based on a more quantitative doubling property, these weights assign to a subset $E$ of an interval $I$ a fair share of $I$'s weight, when compared to the ratio of the Lebesgue measures of $E$ and $I$. More precisely, $w \in A_\infty$ if and only if for any $\alpha$, $0 < \alpha < 1$, there exists a $\beta$, $0 < \beta < 1$, so that, for all intervals $I$ and all subsets $E \subset I$, $|E| \geq \alpha|I| \Rightarrow w(E) \geq \beta w(I)$, see [**GC-RF**] or [**Duo**] .

EXERCISE 3.11. Show that if $w \in A_\infty$ then there is a $q > 1$ such that $w \in RH_q$. This can be deduced from the John-Nirenberg Inequality similarly to the proof of the second part of Theorem 3.5. (Use the fact that $\ln w \in BMO$.)

In the dyadic case it is true that $A_\infty^d \subset \bigcup_{q>1} RH_q^d$. See [**Bu1**] for a proof. There is an alternative proof using Bellman functions, see [**NTV4**].

EXERCISE 3.12. Show that the weight $w = \chi_{\mathbf{R}\setminus[0,1]}$ is in $RH_p^d$ but is not in $A_\infty^d$ (because it is not dyadic doubling).

The $RH_p$ classes are *self-improving*, this is a beautiful theorem of Gehring, discovered while studying quasiconformal mappings, see [**Ge**]. The same holds for the dyadic classes for which we will state it. This is another example of a self-improvement inequality.

THEOREM 3.13 (Gehring's Theorem). *If $w \in RH_p^d$ then $w \in RH_{p+\epsilon}^d$ for some $\epsilon > 0$.*

We will prove this theorem in a couple of sections using a stopping time argument. Here is a short proof based on a lemma, the catch is that the lemma is a corollary of Gehring's Theorem...

LEMMA 3.14. *If $w \in RH_p^d$ then $w^p \in RH_q^d$ for some $q > 1$.*

CHEAP PROOF OF GEHRING'S THEOREM. $w \in RH_p^d$ implies $w^p \in RH_q^d$ for some $q > 1$, then both statements imply that $w \in RH_{pq}^d$. Set $pq = p + \epsilon$. □

EXERCISE 3.15. Prove the Lemma using Gehring's Theorem.

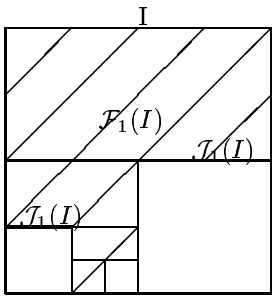
EXERCISE 3.16. Show, using the Coifman-Fefferman Theorem, that $w \in RH_p$ if and only if $w^p \in A_\infty$. Therefore the lemma holds in the non-dyadic case.

In the non-dyadic case one could consider that the important step in the proof of Gehring's Theorem is the Coifman-Fefferman characterization of $A_\infty$. This can be shown using Bellman functions as well, see [**NTV4**].

**3.3. Decaying Stopping Times.** This section we are borrowing from [**KP1**].

Let us first define a *stopping time* $\mathcal{J}$ for an interval $I$, and a given property.

For a given interval $I$, let $\mathcal{J}(I)$ be the collection of dyadic intervals contained in $I$ which are maximal with respect to a given property. Let $\mathcal{F}(I)$ be the collection of dyadic intervals contained in $I$ but not contained in any interval $J \in \mathcal{J}(I)$.
We say the property is *admissible* if there exists an integer $j$ such that for all $I \in \mathcal{D}_j$, $I \in \mathcal{F}(I)$ and hence $\mathcal{F}(I)$ is not empty.
Given an admissible property, let $\mathcal{J}_0(I) = \{I\}$. For $n > 0$ define now the collections $\mathcal{J}_n(I)$ and $\mathcal{F}_n(I)$ inductively. $\mathcal{J}_n(I)$ is the collection of intervals belonging to $\mathcal{J}(J)$ for some $J$ in $\mathcal{J}_{n-1}(I)$. Similarly, $\mathcal{F}_n(I)$ is the collection of intervals belonging to $\mathcal{F}(J)$ for some $J$ in $\mathcal{J}_{n-1}(I)$.




The family of collections of intervals $(\mathcal{J}_n, \mathcal{F}_n)$ is the *stopping time* $\mathcal{J}$ for the interval $I$ corresponding to the given admissible property. The intervals in $\mathcal{F}_n$ are "good", those in $\mathcal{J}_n$ are "bad" but not too bad because their parents are "good". Clearly for each $n > 0$ the intervals in $\mathcal{J}_n(I)$ are pairwise disjoint. By definition the elements of $\mathcal{J}_n(I)$ and $\mathcal{F}_n(I)$ are subintervals of the elements of $\mathcal{J}^{n-1}(I)$. Also $\mathcal{D}(I) = \bigcup_{j=0}^{\infty} \mathcal{F}^j(I)$, and the $\mathcal{F}^j$'s are disjoint collections of dyadic subintervals.

We say that $\mathcal{J}$ is a *decaying stopping time* if there exists $0 < c < 1$ so that for every $I \in \mathcal{D}$, one has

$$\sum_{J \in \mathcal{J}_1(I)} |J| \leq c|I|.$$

Iterating this property we conclude that for decaying stopping times,

$$\sum_{J \in \mathcal{J}_k(I)} |J| \leq c^k |I|. \tag{3.3}$$

Notice that this was exactly the case in the proof of the John-Nirenberg Inequality.

We now prove a basic lemma in the theory of weights. It should be viewed as an analogue of the John-Nirenberg Inequality.

Given a weight $w$, we define the stopping time $\mathcal{J}^w$ where $\mathcal{J}^w(I)$ denotes the set of pairwise disjoint dyadic subintervals $J$ of $I$ which are maximal with respect to the property that $m_J w \geq \lambda m_I w$ or $m_J w \leq \frac{1}{\lambda} m_I w$, where $\lambda \geq 1$ is to be specified in the proof of the following Lemma. It depends only on the $RH_p^d$ constant of $w$.

LEMMA 3.17 (Weight Lemma). *Let $w \in RH_p^d$. Then, for $\lambda$ sufficiently large, $\mathcal{J}^w$ is a decaying stopping time.*

PROOF. First let $\lambda > 3$. We may divide $I$ into three disjoint subsets,

$$I = \bigcup_j I_j^\lambda \bigcup_j I_j^{\frac{1}{\lambda}} \bigcup G,$$

where the intervals $I_j^\lambda, I_j^{\frac{1}{\lambda}}$ are an enumeration of all the different elements of $\mathcal{J}_1^w(I)$ such that for each $j$

$$\lambda\, m_I w \le m_{I_j^\lambda} w, \qquad m_{I_j^{\frac{1}{\lambda}}} w \le \frac{1}{\lambda} m_I w \tag{3.4}$$

and where $\frac{1}{\lambda} m_I w \le w(x) \le \lambda\, m_I w$ a.e. $x$ on $G$.

Suppose the lemma is false. Then $G$ can be arbitrarily small. Suppose $|G| \le \frac{|I|}{3\lambda}$. Thus $\int_G w \le \frac{1}{3}\int_I w$, and since $\lambda > 3$, $\sum_j \int_{I_j^{1/\lambda}} w \le \frac{1}{3}\int_I w$, so that $\sum_j \int_{I_j^\lambda} w \ge \frac{1}{3}\int_I w$. The maximality assumption in the definition of $\mathcal{J}^w(I)$ implies that $m_{I_j^\lambda} w \le 2\lambda m_I w$, therefore,

$$\sum_j |I_j^\lambda| \ge \frac{1}{6\lambda}|I|. \tag{3.5}$$

We will now use (3.5) and (3.4) to contradict $w \in RH_p^d$,

$$\int_I w^p \ge \sum_j \int_{I_j^\lambda} w^p \ge \sum_j \frac{1}{|I_j^\lambda|^{p-1}} \left( \int_{I_j^\lambda} w \right)^p = \sum_j |I_j^\lambda| (m_{I_j^\lambda} w)^p,$$

with the second inequality being just an application of Hölder's inequality. But

$$\sum_j |I_j^\lambda| (m_{I_j^\lambda} w)^p \ge \lambda^p \sum_j |I_j^\lambda| (m_I w)^p \ge \frac{1}{6}\lambda^{p-1} |I| (m_I w)^p.$$

This contradicts $RH_p^d$ provided we chose $\lambda \ge (6C)^{\frac{1}{p-1}}$, where $C$ is the $RH_p^d$ constant of $w$. If this is the case, then $|G| \ge \frac{1}{3\lambda}|I|$ and hence, we have proved the lemma with $c = 1 - \frac{1}{3\lambda}$. □

We will use this lemma to show the classical Gehring's Theorem in the next section. We will also use it to prove the fact that for weights $w \in A_\infty^d$, then $\ln w$ is in dyadic $BMO^d$.

3.3.1. *Proof of Gehring's Theorem.*

THEOREM 3.18 (Gehring's theorem). *Suppose $w \in RH_p^d$ for some $1 < p < \infty$. Then there exists an $\epsilon > 0$, depending only on $p$ and the $RH_p^d$ constant of $w$, so that $w \in RH_{p+\epsilon}^d$.*

PROOF. For any interval $J$, define $\mathcal{G}_1(I) = \left(\bigcup \mathcal{J}_1^w(I)\right)^c$, where we are denoting $\bigcup \mathcal{J}_1^w(I) = \bigcup_{J \in \mathcal{J}_1^w(I)} J$. Observe that for almost every $x \in \mathcal{G}_1(I)$, one has $\lambda^{-1} m_I w \le w(x)$, and one has, for $0 < c < 1$, that $|\mathcal{G}_1(I)| \ge (1-c)|I|$, by Lemma 3.17. Thus $\int_{\mathcal{G}_1(I)} w^p \ge (1-c)|I|\lambda^{-p}(m_I w)^p$. Since $w \in RH_p^d$ this means

there exists $a > 0$ depending only on $p$ and the $RH_p^d$ constant for $w$ (since $\lambda$ and $c$ depend only on these) so that for every $I$,

$$\int_{\mathcal{G}_1(I)} w^p \geq a \int_I w^p. \tag{3.6}$$

We define $\mathcal{G}_j(I) = \bigcup \mathcal{J}_{j-1}^w(I) \setminus \bigcup \mathcal{J}_j^w(I)$. Clearly for $x \in \mathcal{G}_j(I)$ we have that

$$w(x) \leq (2\lambda)^j m_I w. \tag{3.7}$$

EXERCISE 3.19. Show that for every $j$, $\displaystyle\int_{\bigcup \mathcal{J}_j^w(I)} w^p \leq (1-a) \int_{\bigcup \mathcal{J}_{j-1}^w(I)} w^p$.

Using the exercise $j-1$ times we obtain

$$\int_{\mathcal{G}_j(I)} w^p \leq (1-a)^{j-1} \int_I w^p. \tag{3.8}$$

Now we estimate

$$\begin{aligned}
\int_I w^{p+\epsilon} &\leq \sum_{j=1}^{\infty} \int_{\mathcal{G}_j(I)} w^{p+\epsilon} \leq (m_I w)^\epsilon \sum_{j=1}^{\infty} (2\lambda)^{j\epsilon} \int_{\mathcal{G}_j(I)} w^p \\
&\leq (m_I w)^\epsilon \sum_{j=1}^{\infty} (2\lambda)^{j\epsilon} \int_{\mathcal{G}_j(I)} w^p \\
&\leq (m_I w)^\epsilon \sum_{j=1}^{\infty} (2\lambda)^{j\epsilon} (1-a)^{j-1} \int_I w^p. \qquad (3.9)
\end{aligned}$$

Here we have obtained the second inequality using (3.7) and the third using (3.8). Now we choose $\epsilon$ sufficiently small so that $(1-a)(2\lambda)^\epsilon < 1$, and the sum in (3.9) converges. There is a $C > 0$ so that for every $I$,

$$\frac{1}{|I|} \int_I w^{p+\epsilon} \leq C (m_I w)^\epsilon \frac{1}{|I|} \int_I w^p \leq C (m_I w)^{p+\epsilon}.$$

The last inequality by the $RH_p^d$ condition. Therefore $w \in RH_{p+\epsilon}^d$, which was to be shown. □

**3.4. $RH_p^d$ vs $BMO^d$.** We now highlight some connections between $RH_p^d$ and $BMO^d$. We say a weight $w$ is *dyadic doubling* if there is a constant $C$ so that for every dyadic interval $I$ one has

$$m_{\tilde{I}} w \leq C m_I w, \quad \tilde{I} \text{ parent of } I. \tag{3.10}$$

We call the smallest such $C$, the doubling constant for $w$. (We note that if $w$ satisfies $RH_p$ for every interval instead of just the dyadic ones, then it is automatically doubling.) If $w$ is dyadic doubling then for every $J \in \mathcal{J}^w(I)$ one has $m_J w \geq \frac{1}{C\lambda} m_I w$.

LEMMA 3.20 (Buckley's Lemma). *Given a weight $w$, then $w$ is dyadic doubling and is in $RH_p^d$ for some $1 < p < \infty$ if and only if $w \in A_\infty^d$.*

For a proof see [**Bu1**].

THEOREM 3.21. *If $w$ is a dyadic doubling weight which is in $RH_p^d$ for some $1 < p < \infty$, then $\ln w \in BMO^d$.*

By Buckley's Lemma, this is equivalent to Theorem 3.5.

PROOF. First for $\mu > 0$, we estimate for every $I \in \mathcal{D}$,

$$|\{x \in I : |\ln w(x) - \ln m_I w| > \mu\}| \leq C_1 e^{-c_2 \mu}|I|, \tag{3.11}$$

for $C_1, c_2 > 0$ constants independent of $\mu$ and $I$. To see this, we observe that $\{x \in I : |\ln w(x) - \ln m_I w| > \ln \lambda\} \subset \bigcup \mathcal{J}_j^w(I)$. Iterating and using the doubling condition, we see that with $B$ the doubling constant for $w$,

$$\{x \in I : |\ln w(x) - \ln m_I w| > j \ln \lambda + (j-1)\ln B\} \subset \bigcup \mathcal{J}_j^w(I).$$

Applying the Weight Lemma 3.17, we see that there is $0 < c < 1$ such that

$$|\{x \in I : |\ln w(x) - \ln m_I w| > \mu\}| \leq c^{\frac{\mu}{\ln(B\lambda)}}|I| = e^{-C_2\mu}|I|,$$

which is the John-Nirenberg Inequality (3.1). We see that there is a constant $C$ so that for every $I$,

$$\int_I |\ln w(x) - \ln m_I w| dx \leq C|I|.$$

Thus the function $\ln w(x)$ is in $BMO^d$. □

**3.5. $A_\infty^d$ and summation conditions.** We prove a more "dyadic" relation between $A_\infty^d$ weights and $BMO^d$ first discovered by R. Fefferman, Kenig, and Pipher, see [**FKP**]. The relation holds for weights in $RH_p^d$ (without assuming doubling), therefore by Buckley's Lemma it holds for $A_\infty^d$ weights.

THEOREM 3.22. *Let $w \in RH_p^d$ for some $1 < p < \infty$. We define the function*

$$b(x) = \sum_{I \in \mathcal{D}} \frac{\langle w, h_I\rangle}{m_I w} h_I.$$

*Then $b(x)$ is in dyadic $BMO$.*

PROOF. Remember that a function $b$ is in dyadic $BMO$ if and only if the sequence $\lambda_I = |\langle b, h_I\rangle|^2$ satisfies the Carleson condition: $\sum_{J \in \mathcal{D}(I)} \lambda_J \leq C|I|$, for all dyadic intervals $I$. Thus it suffices to show that there exist a constant $C > 0$, so that

$$\sum_{J \in \mathcal{D}(I)} \left| \frac{\langle w, h_J\rangle}{m_J w} \right|^2 \leq C|I| \qquad \forall I \in \mathcal{D}.$$

We define $\mathcal{F}_j^w(I)$ to be the set of dyadic intervals contained in some interval of $\mathcal{J}_{j-1}^w(I)$ but not contained in any interval of $\mathcal{J}_j^w(I)$. This is the decaying stopping time $\mathcal{J}^w$ introduced in Section 3.3. For any interval $J$, we define the function $w_J(x)$ supported on $J$ to be equal to $w(x)$ when $x$ is not contained in any interval of $\mathcal{J}_1^w(J)$ and to be equal to $m_K w$ when $x \in K \in \mathcal{J}_1^w(J)$. Then

$$|J| m_J^2 w \;\; + \sum_{K \in \mathcal{F}^w(J)} |\langle w, h_K\rangle|^2 = \int_J w_J^2(x).$$

However, by definition of $\mathcal{J}^w(J)$, in particular by (3.4),

$$\sum_{K \in \mathcal{F}^w(J)} |\langle w, h_K\rangle|^2 \leq \int_J w_J^2(x) \leq 2\lambda m_J w \int_J w_J(x) = 2\lambda |J| m_J^2 w.$$

Now for any $K \in \mathcal{F}_j^w(I)$, there is a unique interval $J$ of $\mathcal{J}_{j-1}^w(I)$ containing $K$. Then $K \in \mathcal{F}_1^w(J)$ and $m_K w > \frac{1}{\lambda} m_J w$. Thus

$$\begin{aligned}\sum_{K\in\mathcal{F}_j^w(I)} \left|\frac{\langle w, h_K\rangle}{m_K w}\right|^2 &\leq \lambda^2 \sum_{J\in\mathcal{J}_{j-1}^w(I)} \sum_{K\in\mathcal{F}_1^w(J)} \left|\frac{\langle w, h_K\rangle}{m_J w}\right|^2 \\ &\leq \lambda^2 \sum_{J\in\mathcal{J}_{j-1}^w(I)} \frac{1}{m_J^2 w} 2\lambda |J| m_J^2 w \\ &\leq 2\lambda^3 \sum_{J\in\mathcal{J}_{j-1}^w(I)} |J| \leq 2\lambda^3 c^{j-1}|I|,\end{aligned}$$

where we have used (3.3) to get the last inequality. Thus

$$\sum_{J\in\mathcal{D}(I)} \left|\frac{\langle w, h_J\rangle}{m_J w}\right|^2 \leq 2\lambda^3(1 + c + c^2 + \dots)|I| \leq \frac{2\lambda^3|I|}{1-c},$$

which was to be shown. □

COROLLARY 3.23 (Fefferman-Kenig-Pipher). *Let $w \in A_\infty^d$ then there is a constant $C > 0$ such that for all dyadic intervals $I$,*

$$\sum_{J\in\mathcal{D}(I)} \left|\frac{\langle w, h_J\rangle}{m_J w}\right|^2 \leq C|I|. \tag{3.12}$$

There is an alternative characterization of $A_\infty^d$ due to Buckley.

THEOREM 3.24 (Buckley). *Let $w \in A_\infty^d$ then there is a constant $C > 0$ such that for all dyadic intervals $I$,*

$$\sum_{J\in\mathcal{D}(I)} m_J w \left|\frac{\langle w, h_J\rangle}{m_J w}\right|^2 \leq C|I| m_I w. \tag{3.13}$$

We will prove this result using Bellman functions in the fifth lecture.

EXERCISE 3.25. Prove Buckley's Theorem using the Weight Lemma 3.17.

## 4. The $T(1)$ Theorem

**4.1. Singular Integral Operators - Standard kernels.** The prototypical example of a singular integral operator is the Hilbert transform, the operator introduced in the first lecture,

$$Hf(x) = \text{p.v.} \int_{-\infty}^{\infty} \frac{f(y)}{x-y}\, dx = \lim_{\epsilon\to 0} \int_{|x-y|>\epsilon} \frac{f(y)}{x-y}\, dx.$$

We do not expect the integral to converge. Thus $H$ is the first operator we have met for which we see explicitly a distributional kernel which is not an integrable function. However the singular support of the kernel is contained only in the diagonal. Off the diagonal the kernel is given simply by $K(x,y) = (x-y)^{-1}$. This property of having the singular support on the diagonal is called pseudo-locality, see [**St2**]. Pseudo-locality is somewhat related to the hypotheses of Lemma 2.10 which is a sort of Haar locality.

We abstract from the Hilbert transform to a more general class of kernels. A kernel $K(x,y)$ is said to be *standard* if it is a measurable function away from the diagonal and there are constants $C > 0$ and $\delta > 0$ so that

$$|K(x,y)| \leq \frac{C}{|x-y|} \qquad \text{(size condition)}$$

$$\begin{aligned} |K(x,y) - K(x',y)| &\leq C\frac{|x-x'|^\delta}{|x-y|^{1+\delta}} \\ |K(x,y) - K(x,y')| &\leq C\frac{|y-y'|^\delta}{|x-y|^{1+\delta}} \end{aligned} \qquad \text{(cancellation conditions)}$$

for all $x, y, x', y' \in \mathbb{R}$ with $x \neq y$ and with $2|x-x'| \leq |x-y|$ and $2|y-y'| \leq |x-y|$.

Observe that the kernel of $H$ is standard with $\delta = 1$.

LEMMA 4.1. *Suppose $T$ is a bounded operator on $L^2(\mathbb{R})$ whose kernel is standard. Then $T$ is of weak type $(1,1)$.*

PROOF. Given $f \in L^1(\mathbb{R})$, we proceed as in the proof of Corollary 2.8 writing $f = g + b$ with $\|g\|_\infty \leq 2\lambda$ and with $b = \sum_j b_j$ with each $b_j$ supported in an interval $I_j$, having mean zero on $I_j$, and with $\{I_j\}$ pairwise disjoint and so that $|\cup I_j| \leq \frac{\|f\|_1}{\lambda}$. Notice that by definition of $g$, $\|g\|_1 \leq \|f\|_1$ (see exercise 2.9).

As before, from the $L^2$ boundedness of $T$, it follows that

$$\left|\left\{x : Tg(x) > \frac{\lambda}{2}\right\}\right| \leq \frac{C\|f\|_1}{\lambda}.$$

We need only show the same for the set where $|Tb|$ is large. We define $3I_j$ to be the (non-dyadic) interval having the same center as $I_j$ but having triple the length. We let $x_j$ be the center of $I_j$. We need only show that there is a constant $C$ so that for every $j$, we have

$$\|\chi_{(3I_j)^c} Tb_j\|_1 \leq C\|b_j\|_1, . \tag{4.1}$$

If we have done this, we may write $Tb = e_1 + e_2$, where

$$e_1 = \sum_j \chi_{(3I_j)^c} Tb_j, \qquad e_2 = \sum_j \chi_{3I_j} Tb_j.$$

EXERCISE 4.2. Check $\|e_1\|_1 \leq 2C\|f\|_1$, while $e_2$ is supported only on $\cup_j 3I_j$.

We necessarily have

$$|\{x : |e_2(x)| > \lambda\}| \leq |\cup_j 3I_j| \leq 3\sum_j |I_j| \leq \frac{3\|f\|_1}{\lambda},$$

and also that $$|\{x : |e_1(x)| > \lambda|\}| \leq \frac{\|e\|_1}{\lambda} \leq \frac{C\|f\|_1}{\lambda}.$$

Thus, we need only prove (4.1).

Suppose $x \notin 3I_j$. Then

$$(Tb_j)(x) = \int_{I_j} K(x,y)b_j(y)\,dy,$$

where $K$ is the kernel of $T$. Recalling that $b$ has mean zero, we compute

$$\begin{aligned}
\int_{(3I_j)^c} |(Tb_j)(x)|\,dx &= \int_{(3I_j)^c} \left| \int_{I_j} K(x,y) b_j(y)\,dy \right| dx \\
&\leq \int_{(3I_j)^c} \left| \int_{I_j} (K(x,y) - K(x,y')) b_j(y)\,dy \right| dx \\
&\leq C \int_{(3I_j)^c} \int_{I_j} \frac{|y-y'|^\delta}{|y'-x|^{1+\delta}} |b_j(y)|\,dy\,dx \\
&\leq C \int_{I_j} |y-y'|^\delta |b_j(y)| \int_{(3I_j)^c} \frac{1}{|x-y'|^{1+\delta}}\,dx\,dy \\
&\leq C \int_{I_j} \frac{|y-y'|^\delta}{|I_j|^\delta} |b_j(y)| dy \\
&\leq C\|b_j\|_1.
\end{aligned}$$

Therefore the Lemma is proven. □

In what preceded, we used the following inequality: let $\delta > 0$ be given, let $I$ be any interval and $x_I$ its center, then there exists a constant $C$ depending only on $\delta$ so that $\int_{I^c} |y - x_I|^{-1-\delta} dy \leq C|I|^{-\delta}$.

From the symmetry between $x$ and $y$ in the definition of standard kernel, and the weak boundedness just shown, it follows by interpolation that if an operator $T$ having a standard kernel is of strong type (2,2) then it is of strong type $(p,p)$ for all $1 < p < \infty$. More can be said about the limiting case when $p \to \infty$.

LEMMA 4.3. *Suppose $T$ is a bounded operator in $L^2(\mathbb{R})$ whose kernel is standard, then it extends to a bounded operator from $L^\infty(\mathbb{R})$ into $BMO$.*

The result is due independently to Peetre, Spanne and Stein, [**FS**]. A proof for bounded functions with compact support can be found in [**Duo**] p. 121. Those functions are not dense in $L^\infty$ therefore we can not use a continuity argument to extend to all the space. The advantage of considering those functions is that they are in $L^2$ therefore the action of the operator on them is well defined at a.e. point. The proof in general follows the same lines as the proof for compactly supported bounded functions, provided we have been able to define $Tf$ for $f \in L^\infty$. The problem being that even the truncated integrals $\int_{|x-y|>\epsilon} K(x,y) f(y)\,dy$ might not be defined for a bounded function, since $K(x,y)$ is only $O(|x-y|^{-1})$ at infinity. We can define $Tf$, for all $f \in L^\infty$, as a distribution that acts on $C^\infty$, compactly supported functions $\phi$ such that $\int \phi = 0$. Let $I$ be any interval that contains the support of $\phi$. Let $f_1 = f \cdot \chi_{3I}$, and $f_2 = f - f_1$. Then $Tf_1$ is defined since $f_1$ is bounded and has compact support. Define $\langle Tf_2, \phi\rangle = \langle f_2, T^*\phi\rangle$. By the standard estimates on $K$ and the assumption that $\phi$ has mean value zero, one concludes that for $x \notin 3I$, $|T^*\phi(x)| \leq C(1+|x|)^{-1-\delta}$, so $\langle f_2, T^*\phi\rangle$ is well defined.

**4.2. The $T(1)$ Theorem.** Integral operators given by a standard kernel that are bounded in $L^2$ have many other boundedness properties. We will give sufficient conditions for the boundedness of such operators in $L^2$. We will do this by studying the interaction of such operators with the Haar basis. Some of the conditions are going to involve the image of a bounded function, a procedure similar to the one described in the preceding paragraph will do the job.

Given any point $x$ and any dyadic intervals $I$ and $J$, we define $\rho_{xI}$ to be the largest of $|I|$ and the euclidean distance from $x$ to $I$. Similarly, we define $\rho_{IJ}$ to be the largest of $|I|$, $|J|$, and the Euclidean distance between $I$ and $J$.

We let $T$ be an operator with standard kernel. And we consider $Th_I(x)$ for $x \notin 3I$. We let $x_I$ be the center of $I$. We observe

$$Th_I(x) = \int_I K(x,y)h_I(y)\,dy = \int_I (K(x,y) - K(x,x_I))h_I(y)\,dy.$$

Using the cancellation conditions of the kernel and the fact that $|h_I(x)| = \frac{1}{\sqrt{|I|}}$, we obtain that for $x \notin 3I$,

$$|Th_I(x)| + |T^*h_I(x)| \leq \frac{C|I|^{\frac{1}{2}+\delta}}{\rho_{xI}^{1+\delta}}, \tag{4.2}$$

(See the proof of the boundedness of the Hilbert transform in Section 2.3.2.)

None of the following is precisely standard terminology. We say that any linear operator $T$ (whether it has a standard kernel or not) is a singular integral operator if it satisfies (4.2). In particular $Th_I$ must be well defined a.e. and for all $I \in \mathcal{D}$. And its adjoint is defined at least on the Haar basis, meaning that: $\langle Th_I, h_J\rangle = \langle h_I, T^*h_J\rangle$. Clearly for an operator $T$ to be bounded on $L^2(\mathbb{R})$, it must be bounded restricted to the Haar basis and so must its adjoint. That is, there must be a constant $C$ so that for every $I$,

$$\|Th_I\|_2 + \|T^*h_I\|_2 \leq C. \tag{4.3}$$

We refer to (4.3) as the *weak boundedness property*. At last, we are ready to state a version of the famous $T(1)$ theorem of David and Journé.

THEOREM 4.4 ($T(1)$ Theorem ). *Let $T$ be a singular integral operator satisfying the weak boundedness property. Suppose that $T(1)$ and $T^*(1)$ are in $BMO$. Then $T$ is bounded on $L^2(\mathbb{R})$.*

If $T$ is an integral operator with a standard kernel, then $T$ is defined on $C_0^\infty$. Had $T$ been bounded in $L^2$ then it would have had a natural extension to $L^\infty$, as discussed in the previous section. A similar procedure allows us to define $T(1)$, even if $T$ is not a priori bounded, see [**Ch**] p.52. In our case, we assume that $T(1)$ and $T^*(1)$ are well defined and are functions in $BMO$, which is an infinite number of assumptions! In particular we can pair functions in $BMO$ with functions in its dual $H^1$, see [**FS**]. The Haar functions are prototypical atoms in $H^1$, therefore $\langle T(1), h_I\rangle = \int T^*h_I$; and $\langle h_I, T^*1\rangle = \int Th_I$.

PROOF. Let $T(1) = b_1$ and $T^*(1) = b_2$. Consider $L = T - \pi_{b_1} - \pi_{b_2}^*$, where $\pi_b$ is the dyadic paraproduct defined in Section 1.5.3. First observe that $L(1) = L^*(1) = 0$ (by Exercise 1.35). Secondly, observe that $L$ is a singular integral since $\pi_{b_1}$ and $\pi_{b_2}$ are (in fact $\pi_{b_1}h_I(x) = \pi_{b_1}^*h_I(x) = \pi_{b_2}h_I(x) = \pi_{b_2}^*h_I(x) = 0$, when $x \notin I$). Finally, since $\pi_{b_1}$ and $\pi_{b_2}$ are bounded in $L^2(\mathbb{R})$, the operator $L$ satisfies the weak boundedness property. If $L$ is bounded, then $T$ is. Thus, without loss of generality, we may assume $T(1) = T^*(1) = 0$.

We bound $T$ by decomposing it as

$$T = \sum_{j\in\mathbb{Z}} \left(E_jT\Delta_j + \Delta_jT\Delta_j + \Delta_jTE_j\right) \quad \left(= \sum_{j\in\mathbb{Z}} \left(E_{j+1}TE_{j+1} - E_jTE_j\right)\right).$$

We will apply Cotlar's Lemma to the sums of each of the three terms. First we will show that $\sum_{j\in\mathbb{Z}} \Delta_j T \Delta_j$, is bounded. Observe that

$$(\Delta_j T \Delta_j)^*(\Delta_k T \Delta_k) = (\Delta_j T \Delta_j)(\Delta_k T \Delta_k)^* = 0,$$

whenever $j \neq k$. Then fixing $j$, it suffices to bound $\Delta_j T \Delta_j$. We observe that its matrix is indexed by pairs of intervals $(I, J)$ each of which is in $\mathcal{D}_j$ and is given by $A_{IJ} = \langle h_J, T h_I \rangle$. Applying (4.2), we conclude that $|A_{IJ}| \leq \left(2^j \rho_{IJ}\right)^{-1-\delta}$. when $\rho_{IJ} > 2^{-j}$, and is bounded by $C$ in any case by the weak boundedness property. Now observe that for any $(I, J) \in \mathcal{D}_j \times \mathcal{D}_j$, we have that $\rho_{IJ} = k2^{-j}$ for some integer $k$. What we have shown is that for some constant $C$, $|A_{IJ}| \leq C(1+k)^{-1-\delta}$. Thus to bound $\Delta_j T \Delta_j$, it suffices to bound the integer indexed matrix given by $m_{jk} = (|j-k|+1)^{-1-\delta}$. This is bounded by Schur's lemma since, whenever $\delta > 0$,

$$\sum_{j\in\mathbb{Z}} m_{jk} = \sum_{j\in\mathbb{Z}} m_{kj} = \sum_{j\in\mathbb{Z}} \left(\frac{1}{1+|j|}\right)^{1+\delta} < \infty.$$

Thus to bound $T$, it suffices to prove that $\sum_j E_j T \Delta_j$ is bounded (since by the symmetry between the hypotheses on $T$ and $T^*$ this will be the same as bounding $\sum_j \Delta_j T E_j$). We will apply Cotlar's Lemma together with Schur's Lemma. Clearly $(E_j T \Delta_j)(E_k T \Delta_k)^* = 0$, when $j \neq k$. When $I \in \mathcal{D}_j$ and $J \in \mathcal{D}_k$, we let $A_{IJ} = \langle E_j T h_I, E_k T h_J \rangle$. We define

$$C_1^{j,k} = \sup_{J\in\mathcal{D}_k} \sum_{I\in\mathcal{D}_j} A_{IJ}, \quad \text{and} \quad C_2^{j,k} = \sup_{I\in\mathcal{D}_j} \sum_{J\in\mathcal{D}_k} A_{IJ}.$$

It suffices to show that there exists $C$ and $0 < c < 1$ so that

$$\sqrt{C_1^{j,k} C_2^{j,k}} \leq C c^{|j-k|}.$$

It also suffices to do this with $k \geq j$ by symmetry.

We define $\mathcal{B}_{jk}$ to be the set of intervals $J$ in $\mathcal{D}_k$ so that the distance from $J$ to an endpoint of any interval in $\mathcal{D}_j$ is bigger than $2^{-\frac{j+k}{2}}$. Since $k > j$ both $E_j T h_I$ and $E_k T h_J$ with $I \in \mathcal{D}_j$ and $J \in \mathcal{D}_k$ are constant on intervals in $\mathcal{D}_k$. For any $\hat{J} \in \mathcal{D}_k$, we denote by $(E_j T h_I)(\hat{J})$ and $(E_k T h_J)(\hat{J})$, the value of the functions $E_j T h_I$ and $E_k T h_J$ respectively on $\hat{J}$. We define for any $J \in \mathcal{D}_k$ that $K(J)$ is the unique element of $\mathcal{D}_j$ which contains $J$. Note that since $T^*(1) = 0$, we have that $\int T h_J = 0$, and hence by definition

$$\int E_k T h_J = \int T h_J = 0. \tag{4.4}$$

We estimate $A_{IJ}$ when $I \in \mathcal{D}_j$ and $J \in \mathcal{B}_{jk}$. We have that

$$\begin{aligned}
&\left| \int E_j T h_I(x) E_k T h_J(x)\, dx \right| \\
&= \left| \int_{K(J)} E_j T h_I(x) E_k T h_J(x)\, dx + \int_{K(J)^c} E_j T h_I(x) E_k T h_J(x)\, dx \right| \\
&= \left| -\int_{K(J)^c} E_j T h_I(J) E_k T h_J(x)\, dx + \int_{K(J)^c} E_j T h_I(x) E_k T h_J(x)\, dx \right| \\
&\leq C \sum_{K \in \mathcal{D}_k; K \subset K(J)^c} |I|^{\frac{1}{2}+\delta} |J|^{\frac{1}{2}+\delta} |K| \left( \frac{1}{\rho_{IJ}^{1+\delta} \rho_{JK}^{1+\delta}} + \frac{1}{\rho_{IK}^{1+\delta} \rho_{JK}^{1+\delta}} \right).
\end{aligned}$$

The second inequality by (4.4).

When $J \notin B_{jk}$, we simply use the size estimates to obtain

$$\left| \int E_j T h_I(x) E_k T h_J(x)\, dx \right| \leq C \sum_{K \in \mathcal{D}_k} \frac{|K| |I|^{\frac{1}{2}+\delta} |J|^{\frac{1}{2}+\delta}}{\rho_{IK}^{1+\delta} \rho_{JK}^{1+\delta}}. \tag{4.5}$$

In fact, this last inequality is true for all $I \in \mathcal{D}_j$ and $J \in \mathcal{D}_k$. Now we compute

$$\begin{aligned}
\sum_{J \in \mathcal{D}_k} \sum_{K \in \mathcal{D}_k; K \subset K(J)^c} \frac{|I|^{\frac{1}{2}+\delta} |J|^{\frac{1}{2}+\delta} |K|}{\rho_{IJ}^{1+\delta} \rho_{JK}^{1+\delta}} &\leq C \sum_{J \in \mathcal{D}_k} \frac{|I|^{\frac{1+\delta}{2}} |J|^{\frac{1+\delta}{2}}}{\rho_{IJ}^{1+\delta}} \\
&\leq C (2^{k-j})^{\frac{1-\delta}{2}},
\end{aligned}$$

while on the other hand

$$\begin{aligned}
\sum_{J \in \mathcal{D}_k} \sum_{K \in \mathcal{D}_k; K \subset K(J)^c} \frac{|I|^{\frac{1}{2}+\delta} |J|^{\frac{1}{2}+\delta} |K|}{\rho_{IK}^{1+\delta} \rho_{JK}^{1+\delta}} &\leq C \sum_{K \in \mathcal{D}_k} \frac{|I|^{\frac{1+\delta}{2}} |K|^{\frac{1+\delta}{2}}}{\rho_{IK}^{1+\delta}} \\
&\leq C (2^{k-j})^{\frac{1-\delta}{2}},
\end{aligned}$$

and

$$\sum_{J \in \mathcal{D}_K \setminus \mathcal{B}_{jk}} \sum_{K \in \mathcal{D}_k} \frac{|K| |I|^{\frac{1}{2}+\delta} |J|^{\frac{1}{2}+\delta}}{\rho_{IK}^{1+\delta} \rho_{JK}^{1+\delta}} \leq C \sum_{K \in \mathcal{D}_k} \frac{|K| |I|^{\delta}}{\rho_{IK}^{1+\delta}} \leq C.$$

Combining the last inequalities, we obtain that

$$C_2^{j,k} \leq C(1 + (2^{k-j})^{\frac{1-\delta}{2}}).$$

On the other hand, applying (4.5), we see that

$$\sum_{I \in \mathcal{D}_j} |A_{IJ}| \leq C \sum_{K \in \mathcal{D}_k} \sum_{I \in \mathcal{D}_j} \frac{|K| |I|^{\frac{1}{2}+\delta} |J|^{\frac{1}{2}+\delta}}{\rho_{IK}^{1+\delta} \rho_{JK}^{1+\delta}} \leq C (2^{k-j})^{-\frac{1}{2}}.$$

Thus we conclude

$$C_1^{j,k} C_2^{j,k} \leq C\big( (2^{k-j})^{-\frac{\delta}{2}} + (2^{k-j})^{-\frac{1}{2}} \big).$$

Recalling our restriction to $k > j$, we see that this is precisely what we needed to show. □

This and the previous section follow very closely N. Katz unpublished manuscript [**Ka1**]. Despite the fact that I know he now prefers the proof of this theorem that only uses Schur's Lemma, I have chosen to present this proof that also uses Cotlar's Lemma. The idea behind these proofs is due to Coifman and Semmes, see

[**CJS**]. Analyzing the decay of the coefficients of the matrix induced by the Haar basis (or a wavelet basis) is a powerful technique. It can be used in other contexts like matrix-valued weighted spaces and non-homogenous spaces. We will say more about the first topic in the last lecture and about the second, at the end of the next section.

**4.3. Cauchy Integral, $T(b)$ Theorem.** Let $\Gamma = \{x + iy : y = A(x), x \in \mathbb{R}\}$, $A : \mathbb{R} \to \mathbb{R}$ is a Lipschitz function[10]. The Cauchy integral of a function $f : \Gamma \to \mathbb{C}$ is given by:

$$C_\Gamma f(z) = \frac{1}{2\pi i} \int_\Gamma \frac{f(w)}{z - w}\, dw,$$

where $z = x + i(A(x) + t)$, $t > 0$; that is $z \in \Omega^+$ the upper domain determined by the curve $\Gamma$: When $\Gamma = \mathbb{R}$ this is the familiar Cauchy integral discussed in the first lecture. For suitable curves $\Gamma$ and functions $f$, the function $C_\Gamma f(z)$ is an analytic function on $\Omega^+$.

It is not hard to check that for smooth curves $\Gamma$ (that is smooth mappings $A$) $C_\Gamma$ is a bounded operator in $L^2$ (a perturbation argument plus boundedness of the Hilbert transform). Calderón showed in 1977 that boundedness still holds for Lipschitz mappings $A$ with small Lipschitz constant [**Cal**]. In 1982 Coifman, McIntosh and Meyer proved boundedness of the Cauchy integral for all Lipschitz mappings [**CMM**]. Guy David found in 1986 necessary and sufficient conditions on the curve $\Gamma$ for $C_\Gamma$ to be bounded in $L^2(\Gamma)$, namely the curve must be *Ahlfors-David regular*[11] [**Da2**]. There are many proofs of the boundedness of the Cauchy Integral on Lipschitz curves, in Murai's book there are listed at least twelve proofs [**Mu**], and since then a few new proofs have been found, including Verdera and Melnikov's beautiful geometric proof in 1995 using the Menger curvature [**MV**].

One of the most popular proofs uses a variant of the $T(1)$ Theorem, the so-called $T(b)$ Theorem. The function 1 is replaced by a *para-acreetive* function $b$, this means that the function is bounded and essentially bounded away from zero in a measure theoretic sense, see [**Ch**], [**Da1**], [**Da2**] for more details. Setting $b(x) = x + iA'(x)$, one can check that $b$ is para-acreetive and $C_\Gamma b = 0$; then one invokes the $T(b)$ Theorem to get the desired boundedness.

The Coifman-Semmes proof of the $T(1)$ Theorem that we presented in the previous section can be extended to *homogeneous spaces* (this is a code word for the space $L^2(d\mu)$, where $\mu$ is a doubling measure). The homogeneous spaces are close to Euclidean space and the Calderón-Zygmund theory still holds there (that is why it was important to have an analogue to the $T(1)$ theorem that permitted to show $L^2$ boundedness of singular integral operators). The Cauchy integral operator is defined by

$$C_\mu f(z) = \int \frac{f(\xi)}{z - \xi} d\mu(\xi).$$

The argument is similar to the one presented before, one must analize the matrix of the operator in a "Haar basis", the decay of the coefficients must be carefully determined and then the boundedness can be deduced from Schur's and Cotlar's Lemmas. The new ingredient is that this time the basis is a weighted Haar system and one must check that this is an unconditional basis, see [**Ch**], [**Da2**].

[10] there exists a constant $M > 0$ such that $|A(x_1) - A(x_2)| \leq M|x_1 - x_2|$ for all $x_i \in \mathbb{R}$

[11] This means that $|\Gamma \cap B_r(x)| \leq Cr$ for all balls centered at $x$ and of radius $r$.

**4.4. Non-homogeneous Spaces.** Until recently, people did not venture into the land of *non-homogeneous spaces* (spaces defined by *non-doubling* measures). Most of the theory can be developed in that context as well. This is very much work in progress. Particularly beautiful is the proof of the boundedness of the Cauchy integral on non-homogeneous spaces given in 1997 by Nazarov, Treil and Volberg using random dyadic grids and weighted Haar functions, see [**NTV3**]. Simultaneously, X. Tolsa proved the same result, more along the lines of Melnikov-Verdera's result, see [**To1**]. Their conclusion was that the Cauchy integral operator is bounded in $L^2(\mu)$ for *regular measures* $\mu$[12] on the plane if and only if it is bounded on characteristic functions of squares (this is the analogue to the condition $T(1) \in BMO$). The random dyadic grids can be used to prove the $T(b)$ Theorem on non-homogeneous spaces, see [**NTV5**]. For a more traditional approach, see [**To2**], where a Littlewood-Paley Theory for non-homogeneous spaces is developed as well.

## 5. Carleson's Lemma and Bellman Functions

In this lecture we discuss the celebrated Carleson Embedding Theorem. We recall the classical problem that gave rise to this theorem. As mentioned in the first lecture, its dyadic analogue is equivalent to the boundedness of the dyadic paraproduct. It can be deduced from Carleson's Lemma, whose classical proof we recall. We present a proof by Nazarov, Treil and Volberg of the embedding theorem using only Bellman functions (no maximal functions or stopping times). Weighted versions of the theorem and the lemma are stated. We illustrate furthermore the Bellman function technique by proving Buckley's characterization of $A_\infty$ weights by summation conditions.

**5.1. Carleson Embedding Theorem.** Carleson's problem was to classify those positive measures $\mu(x,t)$ on the upper half plane $\mathbb{R}^2_+$ for which the mapping that takes square integrable functions on $\mathbb{R}$ to their harmonic extension to $\mathbb{R}^2_+$ is bounded from $L^2(\mathbb{R})$ into $L^2(\mathbb{R}^2_+, d\mu)$. More precisely, let $u(x+it) = f * P_t(x)$, $x \in \mathbb{R}$, $t > 0$, $P_t$ the Poisson kernel, be the harmonic extension of $f$, then the question is: For which measures $\mu$ does there exist a constant $C > 0$ such that for all $f \in L^2(\mathbb{R})$

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

The answer is if and only if the measure $\mu$ is a *Carleson measure*; i.e. if and only if there exists a constant $C > 0$ such that for all intervals $I$ in $\mathbb{R}$

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

where $Q_I = \{(x,t) : x \in I, 0 \leq t \leq |I|\}$ is the *Carleson box* corresponding to $I$.

We will not prove this result here, you can find the proof in many books, for example [**Ni**], [**St2**], [**Gar**]. Maybe even in the lecture notes by S. Hofmann in this volume [**Hof**].

The dyadic analogue of this result is what we will call Carleson's Embedding Theorem. The Poisson averaging is replaced by averaging the function on the upper

[12] This is a necessary condition, and it means that $\mu(B_r) \leq Cr$.

halves $T_I = \{z = x + it \in Q_I : t \geq \frac{|I|}{2}\}$ of the corresponding dyadic Carleson boxes, that is we define an extension to $\mathbb{R}^2_+$ by:

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

Notice that the collection $\{T_I\}_{I\in\mathcal{D}}$ provides a partition of $\mathbb{R}^2_+$. The problem is the same as before, except that now all the information we need from the measure is the mass of $T_I$, namely the positive sequence $\mu(T_I) = \lambda_I$. Observe that $\mu(Q_I) = \sum_{J\in\mathcal{D}(I)} \lambda_J$. Thus we will say that a sequence $\{\lambda_I\}_{I\in\mathcal{D}}$ of positive numbers is a *Carleson sequence* if there exists a constant $C > 0$ such that for all $I \in \mathcal{D}$

$$\sum_{J\in\mathcal{D}(I)} \lambda_J \leq C|I|.$$

THEOREM 5.1 (Carleson's Embedding Theorem). *Given a Carleson sequence $\{\lambda_I\}_{I\in\mathcal{D}}$, then there exists a constant $C > 0$ such that for all $f \in L^2(\mathbb{R})$*

$$\sum_{I\in\mathcal{D}} |m_I f|^2 \lambda_I \leq C\|f\|_2^2.$$

EXERCISE 5.2. The converse is immediate. Test on Haar functions the above inequality and you will obtain the Carleson sequence condition.

As mentioned in the first lecture, this implies the boundedness in $L^2$ of the dyadic paraproduct $\Pi_b f$, where the Carleson sequence is $\lambda_I = |\langle b, h_I\rangle|^2$ (the Carleson condition on the sequence is equivalent to $b \in BMO^d$!).

We will prove something slightly different, namely:

LEMMA 5.3 (Carleson's Lemma). *Given a Carleson sequence $\{\lambda_I\}_{I\in\mathcal{D}}$ and a positive sequence $\{a_I\}_{I\in\mathcal{D}}$, let $a^*(x) = \sup_{I\ni x} a_I$, then*

$$\sum_{I\in\mathcal{D}} a_I \lambda_I \leq \int_{\mathbb{R}} a^*(x)\, dx. \tag{5.1}$$

This is precisely the Lemma we stated in the first lecture, see (1.5).

PROOF OF CARLESON'S EMBEDDING THEOREM. In our case, let $a_I = m_I^2 f$, then $a^*(x) = M_d^2 f(x)$, where $M_d$ is the dyadic maximal function. Assuming (5.1), and remembering that the dyadic maximal function is bounded in $L^2$, we conclude that

$$\sum_{I\in\mathcal{D}} m_I^2 f\, \lambda_I \leq C\|f\|_2^2,$$

this proves Carleson's Embedding Theorem. □

The previous proof invokes the maximal function, and that is a problem when one is in a context where there is no such obvious analogue of that object. For example when dealing with matrix or operator valued Hilbert transforms. A different proof is desirable and that is what we will present in the next section.

PROOF OF CARLESON'S LEMMA. To prove (5.1), define the characteristic function $\chi(I,t)$ to be one if $0 < t < a_I$, zero otherwise. Clearly,

$$\sum_{I\in\mathcal{D}} a_I \lambda_I = \int_0^\infty \sum_{I\in\mathcal{D}} \chi(I,t)\lambda_I\, dt.$$

For all $t > 0$, let $E_t = \{x \in \mathbb{R} : a^*(x) > t\}$. Clearly $E_t$ is the union of all the intervals $I$ such that $a_I > t$, hence by Tchebychev's inequality, $|E_t| \leq t^{-1} \int a^*(x)\, dx$ (where the integral on the right hand side can be assumed finite, otherwise there is nothing to prove).

Denote by $I_k$ the maximal dyadic intervals contained in $E_t$; then $E_t$ is the disjoint union of the $I_k$'s, and for all $t > 0$,

$$\sum_{I \in \mathcal{D}} a_I \lambda_I \leq \sum_{I \subset E_t} \lambda_I = \sum_k \sum_{I \subset I_k} \lambda_I \leq C \sum_k |I_k| = C|E_t|;$$

We used the fact that $\{\lambda_I\}$ is a Carleson sequence for the last inequality,

To finish the proof, just observe that by Fubini,

$$\int_0^\infty |E_t|\, dt = \int_{\mathbb{R}} a^*(x)\, dx.$$

□

**5.2. Bellman function proof.** The Bellman functions appeared in control theory and have been succesfully applied to a number of classical and not so classical problems in harmonic analysis by Nazarov, Treil, Volberg, their students and collaborators, see [**HTV**], [**NT**], [**NTV4**], [**Pet1**], [**Pet2**], [**PW**], [**Wi**]. The ingenuity of their arguments is striking. No big machinery is used, basically some convexity arguments and, of course, a great deal of imagination to produce an appropiate Bellman function (these beasts could be functions of 11 variables). As the authors themselves describe it, the method of Bellman functions is more a craftmanship, one learns by working examples and seeing how it works case by case. There is not a single Bellman function but a whole zoo, each problem will have its own Bellman function. The number of variables depends also on each particular case, and the first challenge is to single out all important properties that the function should have, among them some differential inequality must be satisfied; the final challenge is to find a function satisfying all desired properties. Sometimes one can write down "the Bellman function" and read the properties from it, but such function can be quite difficult to handle, the charm is that most of the times one is satisfied with any function that has the given properties... still, finding such functions requires lots of practice. If one is interested in sharp constants then one wants the "best" Bellman function, and finding such thing can be quite challenging, see [**HTV**], [**PW**]. Here we will present the Bellman function proof of the Carleson's embedding Theorem. We are borrowing this proof from [**NT**].

Bellman Function Proof of Carleson's Embedding Theorem. Fix a dyadic interval $I$, and real numbers $F$, $f$, $M$. Consider all sequences $\{\lambda_I\}_{I \in \mathcal{D}}$ satisfying the Carleson condition with the same constant $C$ as in the theorem (without loss of generality we can assume $C = 1$), namely: $\sum_{J' \in \mathcal{D}(J)} \lambda_{J'} \leq |J|$, for all $J \in \mathcal{D}$, and such that $\sum_{J \in \mathcal{D}(I)} \lambda_J = M|I|$. In that case we will say that the sequence belongs to $\mathcal{C}_M(I)$.

If the theorem were true then the quantity (for functions $\phi$ in $L^2$)

$$\mathcal{B}(F, f, M) = \frac{1}{|I|} \sup \left\{ \sum_{J \subset I} \lambda_J m_J^2 \phi : m_J \phi^2 = F,\ m_J \phi = f,\ \{\lambda_J\} \in \mathcal{C}_M(I) \right\}$$

is finite, moreover it would satisfy the inequality $\mathcal{B}(F, f, M) \leq CF$.

Notice that $\mathcal{B}(F,f,M)$ does not depend on the interval $I$ (although at a first glance it seems to depend on $I$! A simple rescaling argument shows that this is not true). Hence we get a function $\mathcal{B}(F,f,M)$ of three variables, the *Bellman function* associated to the Carleson embedding theorem.

(1) What is the domain of definition of $\mathcal{B}(F,f,M)$? Since we know where the variables come from, then they must satisfy the following:

$$f^2 \le F \qquad \text{(Hölder's inequality)},$$
$$0 < M < 1 \qquad \text{(normalized Carleson condition)}.$$

(2) What is the range of $\mathcal{B}(F,f,M)$? Since we believe the theorem is true, then there should be some constant $C$ such that for all triples $(F,f,M)$ in the domain of $\mathcal{B}$: $\quad 0 \le \mathcal{B}(F,f,M) \le CF$.

The most important property that this Bellman function has is the following:

(3) Consider all triples $(F,f,M)$, $(F_r,f_r,M_r)$, $(F_l,f_l,M_l)$ in the domain of $\mathcal{B}$ such that $F = \frac{F_r+F_l}{2}$, $f = \frac{f_r+f_l}{2}$, and $M = \frac{M_r+M_l}{2} + \Delta M$. For all such triples, the following concavity condition holds:

$$\mathcal{B}(F,f,M) - \frac{1}{2}\{\mathcal{B}(F_r,f_r,M_r) + \mathcal{B}(F_l,f_l,M_l)\} \ge \Delta M f^2,$$

To prove this last condition, consider all functions $\phi$ such that

$$m_{I_r}\phi = f_r, \quad m_{I_l}\phi = f_l, \quad m_{I_r}\phi^2 = F_r, \quad m_{I_l}\phi^2 = F_l. \tag{5.2}$$

Clearly for any such $\phi$: $m_I\phi = \frac{f_r+f_l}{2} = f$, and $m_I\phi^2 = \frac{F_r+F_l}{2} = F$.

Among all normalized Carleson sequences $\{\lambda_I\} \in \mathcal{C}_M(I)$, let

$$M_r = \frac{1}{|I_r|}\sum_{J:J\subset I_r}\lambda_J, \quad M_l = \frac{1}{|I_l|}\sum_{J:J\subset I_l}\lambda_J, \quad \Delta M = M - \frac{M_r+M_l}{2} = \frac{\lambda_I}{|I|}.$$

The expression in the definition of the $\mathcal{B}(F,f,M)$, before taking the supremum, can be split into the average of the corresponding expressions for $\mathcal{B}(F_r,f_r,M_r)$ and $\mathcal{B}(F_l,f_l,M_l)$ plus the term $\frac{\lambda_I}{|I|}f^2$. Now taking the supremum over those functions $\phi$ satisfying the above average conditions (5.2) we conclude that

$$\frac{1}{2}\left(\mathcal{B}(F_r,f_r,M_r) + \mathcal{B}(F_l,f_l,M_l)\right) + \frac{\lambda_I}{|I|}f^2 \le \mathcal{B}(F,f,M).$$

The last inequality because the set of functions over which we are taking the supremum is smaller than the one corresponding to $\mathcal{B}(F,f,M)$ (we are excluding all those functions whose averages on the right and left halves of $I$ are not the prescribed values in (5.2)).

Let us recapitulate: if the theorem is true, then the Bellman function has the properties described above.

LEMMA 5.4. *If one can find a function of three variables satisfying* (1), (2), *and* (3), *then the theorem holds.*

We are done if a Bellman function can be found. Condition number (3) is some sort of restricted concavity (since it is a condition on dyadic differences). It is often easier to work with differential conditions. In this case, if $\delta M = 0$, then (3) implies concavity of $\mathcal{B}$. This means that

(3′) the second differential $d^2\mathcal{B} \le 0$.

Also, setting $F = F_r = F_l$, $f = f_r = f_l$, $M_r = M_l = M - \Delta M$ in (3), we obtain

$$\mathcal{B}(F, f, M) - \mathcal{B}(F, F, M - \delta M) \geq \Delta M f^2,$$

which means that

$$(3'') \quad \frac{\partial \mathcal{B}}{\partial M} \geq f^2.$$

EXERCISE 5.5. Show that (3′) and (3″) imply (3) (infinitesimal conditions imply finite difference condition because the domain is convex!).

EXERCISE 5.6. Check that $\mathcal{B}(F, f, M) = 4\left(F - \frac{f^2}{1+M}\right)$ satisfies conditions (1), (2), (3′) and (3″).

PROOF OF THE LEMMA. Given a function $\phi \in L^2$, and a normalized Carleson sequence $\{\lambda_I\}$, define for each dyadic interval $I$

$$F_I = m_I \phi^2, \quad f_I = m_I \phi, \quad M_I = \frac{1}{|I|} \sum_{J \subset I} \lambda_J.$$

Property (3) implies (as we already explained while deducing it) that:

$$|I|\mathcal{B}(F_I, f_I, M_I) - |I_r|\mathcal{B}(F_{I_r}, f_{I_r}, M_{I_r}) - |I_l|\mathcal{B}(F_{I_l}, f_{I_l}, M_{I_l}) \geq \lambda_I f_I^2,$$

because clearly the assumptions are satisfied, namely $F_I = \frac{F_{I_r} + F_{I_l}}{2}$, $f_I = \frac{f_{I_r} + f_{I_l}}{2}$, and $M_I = \frac{M_{I_r} + M_{I_l}}{2} + \frac{\lambda_I}{|I|}$. Iterating n times we obtain

$$\sum_{J : J \subset I, |J| \geq 2^{-n+1}|I|} f_J^2 \lambda_J \leq |I|\mathcal{B}(F_I, f_I, M_I) \quad - \sum_{J : J \subset I, |J| = 2^{-n}|I|} |J|\mathcal{B}(F_J, f_J, M_J).$$

Since the Bellman function is positive, we can discard the negative sum on the right hand side, and taking the limit as $n$ tends to infinity we obtain

$$\sum_{J \subset I} \lambda_J m_J^2 \phi = \sum_{J \subset I} \lambda_J f_J^2 \leq C|I|F_I = C \int_I \phi^2,$$

for all dyadic intervals $I$. In particular for $I = [-2^k, 0]$ and $I = [0, 2^k]$, any $k > 0$. Now let $k$ go to infinity and we obtain the theorem. □

□

**5.3. Weighted Carleson's Lemma. Buckley's Theorem.** One can extend either of the previous proofs to prove a weighted version of Carleson's lemma (more precisely of the embedding theorem).

Remember that a positive sequence $\{\lambda_I\}_{I \in \mathcal{D}}$ is a *w-Carleson sequence* if there is a constant $C > 0$ such that for all dyadic intervals $I$:

$$\sum_{J \in \mathcal{D}(I)} \lambda_J \leq C w(I).$$

LEMMA 5.7 (Weighted Carleson's Embedding Theorem). *Given a weight $w \in A_\infty$ and a $w$-*Carleson sequence $\{\lambda_I\}_{I \in \mathcal{D}}$, *then there exists a constant $C > 0$ such that for all $f \in L^2(w)$*

$$\sum_{I \in \mathcal{D}} \lambda_I m_I^2 |f| \leq C \int |f|^2 w.$$

EXERCISE 5.8. Check that the converse always holds by testing on characteristic functions of intervals.

EXERCISE 5.9. Provide a Bellman function proof of the theorem.

We used this theorem to prove the boundedness of the square function in $L^2(w)$ at the end of the second lecture. Provided we were able to show that the sequence

$$\lambda_{\tilde{I}} = \frac{(w(\tilde{I}) - 2w(I))^2}{w(\tilde{I})} = m_{\tilde{I}} w \left| \frac{\langle w, h_{\tilde{I}} \rangle}{m_{\tilde{I}} w} \right|^2$$

was a $w$-Carleson sequence. That is the content of Buckley's theorem see p.36. Here we present Nazarov, Treil and Volberg's proof based on Bellman funcions that we learned in the Spring School on Analysis in Paseky, May 2000, see [**NTV4**].

THEOREM 5.10 (Buckley's Theorem). *Let $w \in A_\infty^d$ then there is a constant $C > 0$ such that for all dyadic intervals $I$,*

$$\sum_{J \in \mathcal{D}(I)} m_J w \left| \frac{\langle w, h_J \rangle}{m_J w} \right|^2 \leq C|I| m_I w. \tag{5.3}$$

PROOF. Let us consider a function $B(W, L)$ defined in the domain $D = \{1 \leq We^{-L} < \infty, W > 0, w \in \mathbb{R}\}$ (this is a convex domain in $\mathbb{R}$). Later we will think of $W = m_I w$, $L = m_I \log w$, and the domain simply emphasizes the $A_\infty$ condition of the weight, and Jensen's inequality, namely that: $e^{m_I(\log w)} \leq m_I w \leq C e^{m_I(\log w)}$.

Suppose the function $B$ has the following properties:

(1) $0 \leq B \leq W$ (range).

(2) $-d^2 B \geq \dfrac{2e^L}{2W} \left( \dfrac{dW}{W} \right)^2$ (concavity condition).

Where $d^2 B$ is given by $(dW, dL) \begin{pmatrix} B_{WW} & B_{WL} \\ B_{LW} & B_{LL} \end{pmatrix} \begin{pmatrix} dW \\ dL \end{pmatrix}$, and $B_{WW} = \frac{\partial^2 B}{\partial W \partial W}$, etc.

The infinitesimal concavity condition implies the following finite difference condition (because the domain of $B$ is convex), which will be useful to us.

EXERCISE 5.11. Show that (1), (2) imply that for all pairs $(W, L)$, $(W_r, L_r)$, $(W_l, L_l)$ in the domain $D$ such that $(W, L) = \frac{(W_r, L_r) + (W_l, L_l)}{2}$; then there is some constant $c$ such that

$$B(W, L) - \frac{1}{2}\{B(W_r, L_r) + B(W_l, L_l)\} \geq c \frac{e^L}{W} \left( \frac{W_r - W_l}{W} \right)^2.$$

Fix a dyadic interval $I$, let $W = m_I w$, $L = m_I \log w$, where $w \in A_\infty$. Let $W_r = m_{I_r} w$, $L_r = m_{I_r} \log w$, similarly for $W_l$ and $L_l$. Then clearly all three pairs are in the domain (since $w \in A_\infty$), and they satisfy the averaging property in the

exercise. Then by the range property and iterating the exercise we conclude that:

$$\begin{aligned} m_I w &\geq B(m_I w, m_I \log w) - \frac{1}{2}\{B(m_{I_r} w, m_{I_r}(\log w)) + B(m_{I_l} w, m_{I_l}(\log w))\} \\ &\quad + \frac{1}{2}\{B(m_{I_r} w, m_{I_r}(\log w)) + B(m_{I_l} w, m_{I_l}(\log w))\} \\ &\geq c \sum_{J \in \mathcal{D}(I)} \frac{e^{2m_J(\log w)}}{m_J w} \left(\frac{m_{J_r} w - m_{J_l} w}{m_J w}\right)^2 \frac{|J|}{|I|} \\ &\geq \frac{c}{C^2} \frac{1}{|I|} \sum_{J \in \mathcal{D}(I)} m_J w \left|\frac{\langle w, h_J\rangle}{m_J w}\right|^2. \end{aligned}$$

In the last inequality we used the $A_\infty$ condition. This ends the proof as long as we can find a function that satisfies the given properties.

EXERCISE 5.12. Show that $B(W, L) = W - \dfrac{e^{2L}}{W}$ satisfies conditions (1), (2).

□

EXERCISE 5.13. Provide a Bellman function proof of the Fefferman, Kenig, Pipher characterization of $A_\infty$, see p.36.

## 6. Haar multipliers and Weighted Inequalities

In this lecture we introduce some non constant Haar multipliers and prove their boundedness in $L^2$. As a corollary we can prove boundedness in weighted spaces for the basic dyadic operators: constant Haar multipliers, dyadic paraproduct, and dyadic square function. The same idea can be used for more complicated singular integral operators, like the Hilbert Transform. In the last section we present a short survey of some more serious problems like the two-weights problem for the Hilbert transform; matrix-valued weights; and sharp constants for the dyadic square function and the Hilbert transform.

**6.1. Boundedness of Haar Multipliers.** Remember from Section 1.5.4, that a *Haar multiplier* is an operator of the form:

$$Tf(x) = \sum_{I \in \mathcal{D}} w_I(x) \langle f, h_I\rangle h_I(x),$$

where the *symbol* $w_I(x)$ is a function of both the variables $x \in \mathbf{R}$ and $I \in \mathcal{D}$. These operators are formally similar to pseudodifferential operators, but the trigonometric functions have been replaced by step functions. When the symbol is a function independent of $x$, $w_I(x) = w_I$, the corresponding operators are the *constant Haar multipliers*, known to be bounded in $L^p$ if and only if the sequence $\{w_I\}_{I \in \mathcal{D}}$ is bounded. We will be concerned with symbols of the form

$$w_I^t(x) = \left(\frac{w(x)}{m_I w}\right)^t,$$

where $t$ is a real number, $w$ is a weight, and $m_I w$ denotes the mean value of $w$ over $I$. The corresponding multipliers will be denoted $T_w^t$. We prove that under certain conditions on the weight these operators are bounded in $L^2(\mathbb{R})$.

The proof presented here is based on a stopping time argument suggested by P. W. Jones that utilizes the Weight Lemma 3.17. The argument can be adapted to the

case $p \neq 2$ using a version of Cotlar's Lemma in $L^p$, see [**KP1**]. A proof very much in the spirit of Buckley's proof of the boundedness of the square function in Section 2.5.1 can be found in [**Per**]; the $L^p$ result is obtained there by an extrapolation argument.

The non constant Haar multipliers corresponding to $t = 1$ appeared in [**Per**], in connection with the existence and boundedness of *the resolvent of the dyadic paraproduct*. The ones corresponding to $t = \pm 1/2$ appeared in the work of Treil and Volberg concerning matrix valued weighted inequalities for the Hilbert transform [**TV1**], in particular they were used to prove that the Haar system is an unconditional basis in $L^2(w)$ if and only if $w \in A_2$.

THEOREM 6.1. *Let $w$ be a weight. Define the operator*

$$T_w f(x) = \sum_{I \in \mathcal{D}} \frac{w(x)}{m_I w} \langle f, h_I \rangle h_I(x) = w(x)(M_w f)(x),$$

*where $M_w$ is the (possibly unbounded) Haar multiplier with coefficients $\frac{1}{m_I w}$. Then $T_w$ is bounded on $L^p$ if and only if $w \in RH_p^d$.*

REMARK 6.2. If $\frac{1}{\lambda} < w(x) < \lambda$ for a.e. $x$, then there is a trivial bound, far from being sharp. Just observe,

$$\|T_w f\|_p^p = \int w^p |M_w f|^p \leq \lambda^p \|M_w f\|_p^p \leq C\lambda^{2p} \|f\|_p^p,$$

where the last inequality uses the fact that $M_w$ is a constant Haar multiplier with symbol bounded by $\lambda$.

PROOF. The necessity of $w \in RH_p^d$ follows immediately from applying $T_w$ to the Haar functions. For the proof of sufficiency, it suffices to prove the theorem for

$$Tf(x) = \sum_{I \in \mathcal{D}([0,1])} \frac{w(x)}{m_I w} \langle f, h_I \rangle h_I(x).$$

We will only prove in these notes the sufficiency for $p = 2$. For $p \neq 2$, see [**KP1**].

If $w \in RH_2^d$, by the Weight Lemma 3.17, the stopping time $\mathcal{J} = \mathcal{J}^w([0,1])$ is decaying[13]. We will abuse our notation denoting by $\bigcup \mathcal{J}_n(I)$ the set $\bigcup_{J \in \mathcal{J}_n(I)} J$, and $\mathcal{F}_j = \mathcal{F}_j([0,1])$, similarly for $\mathcal{J}_j$.

We define $T_j = T\Delta_{\mathcal{F}_j} = wM_j$, where

$$M_j f(x) = \sum_{J \in \mathcal{F}_j} \frac{1}{m_J w} \langle f, h_J \rangle h_J(x) = \sum_{I \in \mathcal{J}_{j-1}} M_I f(x),$$

and we define for every dyadic $I$, the multiplier

$$M_I f(x) = \sum_{J \in \mathcal{F}^w(I)} \frac{1}{m_J w} \langle f, h_J \rangle h_J(x).$$

[13]Remember that $\mathcal{J}^w(I)$ denotes the set of pairwise disjoint dyadic subintervals of $I$ which are maximal with respect to the property that $m_J w \geq \lambda m_I w$ or $m_J w \leq \frac{1}{\lambda} m_I w$ for some $\lambda \geq (6C)^{1/(p-1)}$, $C$ the $RH_2$ constant of $w$. $\mathcal{F}^w(I)$ denotes the collection of those dyadic intervals in $I$ but not contained in any interval $J \in \mathcal{J}^w(I)$

Each $M_I$ is a bounded constant Haar multiplier since by the definition of the stopping time, $(m_J w)^{-1} \leq \lambda (m_I w)^{-1}$ for all $I \in \mathcal{D}$ and for all $J \in \mathcal{F}^w(I)$. Therefore, for any $f \in L^2$ and for each $I$, $\|M_I f\|_2 \leq \frac{\lambda}{m_I w} \|f\|_2$. Here $\lambda$ is fixed, so that $\mathcal{J}$ is decaying. In fact, defining $\Delta_I f = \sum_{K \in \mathcal{F}^w(I)} \langle f, h_K \rangle h_K$, we have $M_I \Delta_I f = M_I f$ hence,

$$\|M_I f\|_2 \leq \frac{\lambda}{m_I w} \|\Delta_I f\|_2. \tag{6.1}$$

Also notice that $f_j = \Delta_{\mathcal{F}_j} f = \sum_{I \in \mathcal{J}_{j-1}([0,1])} \Delta_I f$.

We shall prove that the $T_j$'s satisfy the conditions in Cotlar's Lemma 2.4. We will begin by proving that each is bounded on $L^2$. We write

$$\int |T_j f|^2 = \int_{\bigcup \mathcal{J}_{j-1}([0,1]) \setminus \bigcup \mathcal{J}_j([0,1])} |T_j f|^2 + \int_{\bigcup \mathcal{J}_j([0,1])} |T_j f|^2, \tag{6.2}$$

and we will estimate each term separately.

Observe that for every $I$, on $I \setminus \bigcup \mathcal{J}(I)$ one has almost everywhere that $w \leq \lambda m_I w$. Thus by the remark right before this proof and (6.1) we conclude that,

$$\begin{aligned} \int_{\bigcup \mathcal{J}_{j-1}([0,1]) \setminus \bigcup \mathcal{J}_j([0,1])} |T_j f|^2 &= \sum_{J \in \mathcal{J}_{j-1}([0,1])} \int_{J \setminus \bigcup \mathcal{J}(J)} |T_j f|^2 \\ &\leq \sum_{J \in \mathcal{J}_{j-1}([0,1])} (\lambda m_J w)^2 \int_{J \setminus \bigcup \mathcal{J}(J)} |M_J f|^2 \\ &\leq \sum_{J \in \mathcal{J}_{j-1}([0,1])} \lambda^4 \|\Delta_J f\|_2^2 \;=\; \lambda^4 \|f_j\|_2^2. \end{aligned}$$

Here the last line comes from the disjointness of the elements of $\mathcal{J}_{j-1}$. We have completed our estimate of the first term of (6.2).

Observe that for every $I$, we have that $M_I f$ is constant on all $J \in \mathcal{J}(I)$. We denote its value on $J$ by $(M_I f)(J)$. Also observe that for such $J$, we have by the definition of $\mathcal{J}$ that $m_J w \leq 2\lambda m_I w$. Now we estimate using the $RH_2^d$ condition, and (6.1),

$$\begin{aligned} \int_{\bigcup \mathcal{J}_j([0,1])} |T_j f|^2 &= \sum_{J \in \mathcal{J}_{j-1}([0,1])} \int_{\bigcup \mathcal{J}(J)} |T_j f|^2 \\ &= \sum_{J \in \mathcal{J}_{j-1}([0,1])} \sum_{K \in \mathcal{J}(J)} \big(M_J f(K)\big)^2 \int_K w^2 \\ &\leq C \sum_{J \in \mathcal{J}_{j-1}([0,1])} \sum_{K \in \mathcal{J}(J)} \big(M_J f(K)\big)^2 |K| m_K^2 w \\ &\leq C \sum_{J \in \mathcal{J}_{j-1}([0,1])} 2^2 \lambda^2 m_J^2 w \int_J |M_J f|^2 \;\leq\; 2^2 \lambda^4 \|f_j\|_2^2. \end{aligned}$$

Thus we have shown that there is a constant $C > 0$ so that $\int |T_j f|^2 \leq C \|f_j\|^2$. We claim further that there exists $0 < d < 1$ so that for any $k > j$, one has

$$\int_{\bigcup \mathcal{J}_{k-1}([0,1])} |T_j f|^2 \leq C d^{k-j} \|f_j\|_2^2. \tag{6.3}$$

We will use Hölder's inequality, the decaying[14] of $\mathcal{J}$, and the fact that $w \in RH^d_{2+\epsilon}$ to compute

$$\begin{aligned}
\int_{\bigcup \mathcal{J}_{k-1}([0,1])} |T_j f|^2 &= \sum_{J\in\mathcal{J}_j([0,1])} \int_{\bigcup \mathcal{J}_{k-j-1}(J)} |T_j f|^2 \\
&= \sum_{J\in\mathcal{J}_j([0,1])} |(M_j f)(J)|^2 \int_{\bigcup \mathcal{J}_{k-j-1}(J)} w^2 \\
&\le \sum_{J\in\mathcal{J}_j([0,1])} |(M_j f)(J)|^2 \left(\int_{\bigcup \mathcal{J}_{k-j-1}(J)} w^{2+\epsilon}\right)^{\frac{2}{2+\epsilon}} \left|\bigcup \mathcal{J}_{k-j-1}(J)\right|^{\frac{\epsilon}{2+\epsilon}} \\
&\le \sum_{J\in\mathcal{J}_j([0,1])} c^{\frac{\epsilon(k-j-1)}{2+\epsilon}} |(M_j f)(J)|^2 \left(\int_{\bigcup \mathcal{J}_{k-j-1}(J)} w^{2+\epsilon}\right)^{\frac{2}{2+\epsilon}} |J|^{\frac{\epsilon}{2+\epsilon}} \\
&\le \sum_{J\in\mathcal{J}_j([0,1])} c^{\frac{\epsilon(k-j-1)}{2+\epsilon}} |(M_j f)(J)|^2 |J| \left(\frac{1}{|J|}\int_J w^{2+\epsilon}\right)^{\frac{2}{2+\epsilon}} \\
&\le \sum_{J\in\mathcal{J}_j([0,1])} c^{\frac{\epsilon(k-j-1)}{2+\epsilon}} |(M_j f)(J)|^2 |J|\, C\,(m_J w)^2 \\
&\le C \sum_{J\in\mathcal{J}_j([0,1])} c^{\frac{\epsilon(k-j-1)}{2+\epsilon}} |J| \left(\frac{1}{|J|}\int_J w\,(M_j f)(J)\right)^2 \\
&\le \sum_{J\in\mathcal{J}_j([0,1])} Cc^{\frac{\epsilon(k-j-1)}{2+\epsilon}} \int_J w^2 |M_j f|^2 = Cc^{\frac{\epsilon(k-j-1)}{2+\epsilon}} \int |T_j f|^2 \\
&\le Cc^{\frac{\epsilon(k-j-1)}{2+\epsilon}} \|f_j\|_2^2.
\end{aligned}$$

This proves (6.3). But (6.3) implies the condition in Cotlar's Lemma because $T_k f$ is supported on $\bigcup \mathcal{J}_{k-1}([0,1])$. Thus the theorem is proven for $p = 2$. □

REMARK 6.3. Notice that $T_w h_J = \frac{w(x)}{m_J w} h_J(x)$ is supported on $J$, and the previous theorem for $p = 2$ will give the strong type (2,2) for the operator, hence by Lemma 2.10, $T_w$ is of weak type (1,1). Thus by interpolation it is of strong type $(p,p)$ for $1 < p < 2$. It is not true that the adjoint $T_w^*$ is localized when acting on Haar functions, hence we cannot repeat the argument and then use duality to get $2 < p < \infty$.

EXERCISE 6.4. Give a Bellman function proof of this result.

6.1.1. *Some Corollaries.* We can consider the following one parameter family of Haar multipliers:

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

where $t \in \mathbb{R}$, and $M_w^t$ is the constant Haar multiplier with symbol $(m_I w)^{-t}$.

Checking the action on the Haar functions one obtains a necessary condition for the boundedness of $T_w^t$ in $L^p(\mathbb{R})$, namely *condition* $C_{tp}$:

$$m_I w^{tp} \le (m_I w)^{tp}, \quad \forall I \in \mathcal{D}.$$

[14]Remember that $\mathcal{J}$ decaying implies that there exists $0 < c < 1$ such that $\left|\bigcup \mathcal{J}_k(I)\right| \le c^k |I|$.

Clearly condition $C_s$ coincides with $RH_s^d$ when $s > 1$, with $A^d_{1-1/s}$ when $s < 0$, and it always holds when $0 \leq s \leq 1$. The following exercise will allow us to deduce the boundedness of $T_w^t$ as a simple corollary of Theorem 6.1.

EXERCISE 6.5. If $w \in C_{tp} \cap A^d_\infty$ then (i) $w^t \in RH_p^d$ and (ii) $m_I w^t \leq C(m_I w)^t$ for all $I \in \mathcal{D}$.

COROLLARY 6.6. *$T_w^t$ is bounded in $L^p(\mathbb{R})$ if $w \in C_{tp} \cap A^d_\infty$, for all $t \in \mathbb{R}$, and $1 < p < \infty$.*

PROOF. Notice that we can factorize $T_w^t$ as a composition of two bounded operators. In fact: $T_w^t = T_{w^t} S_{w,t}$, where $S_{w,t}$ is the constant Haar multiplier with symbol $c_I = \frac{m_I w^t}{(m_I w)^t}$ which is a bounded sequence by Exercise 6.5, therefore $S_{w,t}$ is a bounded operator in $L^p$. Since $w \in A^d_\infty \cap C_{tp}$ also by Exercise 6.5, $w^t \in RH_p^d$, therefore by Theorem 6.1, $T_{w^t}$ is bounded in $L^p$. □

EXERCISE 6.7. Assume $w \in A^d_\infty \cap C_{tp}$ for some $1 < p < \infty$ then $T_w^t$ is of weak type $(1,1)$, for all $t \in \mathbb{R}$.

**6.2. Weighted inequalities.** As an example on how to prove weighted inequalities with the aid of the Haar multipliers $T_w^t$, we show that $A_p^d$ is a sufficient condition for the boundedness in $L^p(w)$ of constant Haar multipliers, the dyadic square function, and dyadic paraproducts.

Recall from the previous section that, for $w$ some weight, and for $t$ any real number, we denote by $M_w^t$, the multiplier given by $M_w^t f = \sum_{I \in \mathcal{D}} (m_I w)^{-t} \langle f, h_I \rangle h_I$.

COROLLARY 6.8. *Let $w \in A_p^d$. Then for any $1 < p < \infty$, the operators $w^{\frac{1}{p}} M_w^{\frac{1}{p}}$ and $M_w^{-\frac{1}{p}} w^{-\frac{1}{p}}$ are bounded on $L^p(\mathbb{R})$.*

PROOF. Since, $A_p^d \subset A^d_\infty$, setting $t = 1/p$, $t = -1/p$ in Corollary 6.6 we get, respectively that $T_w^{\frac{1}{p}} = w^{\frac{1}{p}} M_w^{\frac{1}{p}}$ is bounded in $L^p$ and $w^{-\frac{1}{p}} M_w^{-\frac{1}{p}}$ is bounded on $L^{\frac{p}{p-1}}(\mathbb{R})$. Since $\frac{p}{p-1}$ is the dual index of $p$ and $M_w^{-\frac{1}{p}} w^{-\frac{1}{p}}$ is the dual of $w^{-\frac{1}{p}} M_w^{-\frac{1}{p}}$, we have shown that $w^{\frac{1}{p}} M_w^{\frac{1}{p}}$ is bounded in $L^p$. □

This allows us to interchange the weight $w$ and the multiplier $M_w^{-1}$ when proving weighted norm inequalities. We will apply this philosophy to our basic dyadic operators.

6.2.1. *Constant Haar Multipliers.*

COROLLARY 6.9. *Let $\{\alpha_I\}_{I \in \mathcal{D}}$ be a bounded set of numbers and $T_\alpha$ be the associated Haar multiplier, i.e $T_\alpha f = \sum_{I \in \mathcal{D}} \alpha_I \langle f, h_I \rangle h_I$. Let $1 < p < \infty$ be given and let $w \in A_p^d$. Then $T_\alpha$ is bounded from $L^p(w)$ to $L^p(w)$.*

PROOF. It suffices to show that $w^{\frac{1}{p}} T_\alpha w^{-\frac{1}{p}}$ is bounded on $L^p(\mathbb{R})$. Since constant Haar multipliers commute, then $T_\alpha = M_w^{-\frac{1}{p}} T_\alpha M_w^{\frac{1}{p}}$, and is bounded on $L^p(\mathbb{R})$. Then $w^{\frac{1}{p}} T_\alpha w^{-\frac{1}{p}} = (w^{\frac{1}{p}} M_w^{\frac{1}{p}})(M_w^{-\frac{1}{p}} T_\alpha M_w^{\frac{1}{p}})(M_w^{-\frac{1}{p}} w^{-\frac{1}{p}})$. All three factor on the right hand side are bounded, thus so is the left hand side. □

6.2.2. *Paraproducts.*

COROLLARY 6.10. *Let $w \in A_p^d$ and $b$ in dyadic BMO. Then $\pi_b$, the dyadic paraproduct defined by $\pi_b f = \sum_{I\in\mathcal{D}} b_I m_I f\, h_I$, is bounded on $L^p(w)$.*

PROOF. We will prove the result for $p = 2$ and then an extrapolation argument proves it for $p \neq 2$. It suffices to bound $w^{\frac{1}{2}}\pi_b w^{-\frac{1}{2}}$ on $L^2$. We may write $w^{\frac{1}{2}}\pi_b w^{-\frac{1}{2}} = (w^{\frac{1}{2}}M_w^{\frac{1}{2}})(M_w^{-\frac{1}{2}}\pi_b w^{-\frac{1}{2}})$. Thus it suffices to bound $M_w^{-\frac{1}{2}}\pi_b w^{-\frac{1}{2}}$ on $L^2$. However for any $f \in L^2$,

$$\|M_w^{-\frac{1}{2}}\pi_b w^{-\frac{1}{2}} f\|_2^2 = \sum_{I\in\mathcal{D}} m_I w\, b_I^2\, m_I^2(w^{-\frac{1}{2}} f).$$

Remember that $w$, $w^{-1} \in A_2^d$, which implies that $w^{-1} \in RH^d_{\frac{2+\epsilon}{2}}$ for some $\epsilon > 0$. This fact plus Carleson's Lemma 5.3, boundedness of the maximal function on $L^{\frac{2+2\epsilon}{2+\epsilon}}$, and Hölder's inequality are used to estimate

$$\begin{aligned}
\|M_w^{-\frac{1}{2}}\pi_b w^{-\frac{1}{2}} f\|_2^2 &\leq \sum_I m_I w\, b_I^2 (m_I w^{-\frac{2+\epsilon}{2}})^{\frac{2}{2+\epsilon}} (m_I f^{\frac{2+\epsilon}{1+\epsilon}})^{\frac{2+2\epsilon}{2+\epsilon}} \\
&\leq C\sum_I m_I w\, b_I^2\, m_I(w^{-1})(m_I f^{\frac{2+\epsilon}{1+\epsilon}})^{\frac{2+2\epsilon}{2+\epsilon}} \leq C\sum_I b_I^2 (m_I f^{\frac{2+\epsilon}{1+\epsilon}})^{\frac{2+2\epsilon}{2+\epsilon}} \\
&\leq C\int (\sup_{I\ni x} m_I f^{\frac{2+\epsilon}{1+\epsilon}})^{\frac{2+2\epsilon}{2+\epsilon}}\, dx \leq C\|f^{\frac{2+\epsilon}{1+\epsilon}}\|_{\frac{2+2\epsilon}{2+\epsilon}}^{\frac{2+2\epsilon}{2+\epsilon}} = C\|f\|_2^2.
\end{aligned}$$

Which was to be shown. □

6.2.3. *Dyadic Square Function.* We can also prove weighted inequalities for the dyadic square function. We already presented Buckley's proof of this result in Section 2.5.1 In particular the classical dyadic Littlewood-Paley Theory can be deduced setting $w = 1$. The next result can be viewed as weighted Littlewood-Paley Theory.

COROLLARY 6.11. *Let $w \in A_p^d$ then the dyadic square function $S^d$ is bounded in $L^p(w)$.*

PROOF. We prove it for $p = 2$ and use extrapolation for $p \neq 2$. It suffices to show that $w^{\frac{1}{2}}S^d w^{-\frac{1}{2}}$ is bounded on $L^2$.

Remember that $S^d f(x) = \left(\sum_j |\Delta_j f|^2\right)^{\frac{1}{2}}$. Computing we get

$$\begin{aligned}
\|w^{\frac{1}{2}}S^d w^{-\frac{1}{2}} f\|_2^2 &= \sum_j \int w(x)|\Delta_j(f w^{-\frac{1}{2}})|^2 \\
&= \sum_{I\in\mathcal{D}} |\langle f w^{-\frac{1}{2}}, h_I\rangle|^2 m_I w = \|M_w^{-\frac{1}{2}} w^{-\frac{1}{2}} f\|_2^2 \leq C\|f\|_2^2.
\end{aligned}$$

Remember that $\Delta_j f = \sum_{I\in\mathcal{D}_j}\langle f, h_I\rangle h_I$. The last inequality by Corollary 6.8. □

For operators like the Hilbert transform it is more complicated than these dyadic examples, we will say more in the next section, see [**TV1**], [**KP2**]. But the same ideas are used. The following tautology holds: *an operator $T$ is bounded from $L^2(u)$ into $L^2(v)$ if and only if the operator $v^{\frac{1}{2}}Tu^{-\frac{1}{2}}$ is bounded from $L^2(\mathbb{R})$ into itself.* As we mentioned before, and we hope has been highlighted by the examples, the Haar multipliers will allow us to replace multiplication on space side by multiplication on frequency side, whenever it is convenient. For example to

show the boundedness of the Hilbert transform from $L^2(w)$ into itself it is enough to check that $M_w^{\frac{1}{2}} H M_{w^2}^{-1}$ is bounded in $L^2(\mathbb{R})$. In this case the estimates are more laborious to obtain than in the examples presented here. The strategy followed in [**TV1**], [**KP2**] follows the method introduced by Coifman and Semmes [**CJS**], and widely used in Wavelet Theory. One studies the decay of the matrix of the operator in the Haar basis. Some pieces are analized like if they were constant Haar multipliers, others as if they were paraproducts, very much in the spirit of the $T(1)$ Theorem of David and Journé [**DJ**].

**6.3. Hilbert transform survey.** The *Hilbert transform* is after the Fourier transform probably the most important operator in analysis. We have tried to convey that message throughout these notes. It is deeply connected to complex and Fourier analysis as well as to PDE's, see Section 1.1. It is the prototype of a large class of singular integral operators, the Calderón-Zygmund class, see Section 4.1; it can also be viewed as a Fourier multiplier whose generalizations include pseudo-differential operators and Fourier integral operators. One can perturb the geometry and consider the Cauchy integral along curves or sets, see Section 4.3. One can also change the measures and consider the problem of boundedness on weighted spaces. All these variants are intimately connected and in all cases the problem is to find necessary and sufficient conditions on the symbol of the pseudo-differential operator, on the kernel of the singular integral, on the curve or set, or on the measures for the boundedness of the corresponding operator in some function spaces. In the case of the measures, the problem is the so-called *two-weights problem* which is described in detail in the next section.

Another important related operator is the *bilinear Hilbert transform*

$$B(f,g) = \text{p.v.} \int_{-\infty}^{\infty} \frac{f(x-t)g(x+t)}{t}\, dt.$$

Its analysis requires the use of very fine time-frequency analysis involving deep combinatorial arguments [**LT**]. For an excellent presentation and generalizations see [**GN1**], [**GN2**]. One could attempt the same variants discussed for the Hilbert transform, namely change the geometry, the symbol or the measures, in the bilinear case. In particular the *three-weights problem* for the bilinear Hilbert transform, a very deep problem, would be of enormous interest to tackle.

6.3.1. *Two-weight problem, Matrix-valued weights.* The *two-weights problem* consists in finding necessary and sufficient conditions on a pair of weights $(u,v)$ such that a given operator $T$ is bounded from $L^p(u)$ into $L^p(v)$ for $1 < p < \infty$, i.e.

$$\int |Tf(x)|^p v(x)\, dx \leq C \int |f(x)|^p u(x)\, dx.$$

This problem is particularly interesting and difficult when the operator is the Hilbert transform. For equal weights $u = v = w$ and $p = 2$, this problem was posed and solved by Helson and Szëgo in the 60's. It arose in the context of prediction theory and stationary processes, see Section 1.1.3, and the techniques used involved complex analysis, see [**HS**]. In 1973 Hunt, Muckenhoupt and Wheeden found an alternative characterization of the weights $w$ using purely real analysis techniques, see [**HMW**]. The necessary and sufficient condition is the celebrated $A_p$ condition

$$\|w\|_{A_p} = \sup_I \left(\frac{1}{|I|}\int_I w\right)\left(\frac{1}{|I|}\int_I \frac{1}{w^{p-1}}\right)^{p-1} < \infty.$$

Soon after, Coifman and C. Fefferman extended this to a general class of singular integral operators as well as maximal operators, see [**CF**]. In the early 70's Mischa Cotlar and Cora Sadosky generalized à la Helson-Szegö for $u \neq v$, and provided extensions to the operator valued context, and the bidisc, see [**CS1**], [**CS2**]. No substantial progress was made until 1996, when Treil and Volberg began developing a new machinery to handle these problems.

To finish this notes I would like to describe progress that has been made in the last five years.

The real analysis techniques did not allow to consider matrix or operator valued weights (important in characterizing regularity properties for stationary processes in terms of their spectral measure.) In this context we let $\mathcal{H}$ be a Hilbert space, and $f : \mathbb{R} \to \mathcal{H}$, $W(x) : \mathcal{H} \to \mathcal{H}$ for $x \in \mathbb{R}$, and we can introduce $L^2(W)$ as the space of those functions $f$ such that

$$\|f\|^2_{L^2(W)} = \int \langle W(x)f(x), f(x)\rangle dx < \infty.$$

In 1996, Treil and Volberg considered $U = V = W$, $n \times n$ positive definite matrix-valued weights. They proved in [**TV1**] that the necessary and sufficient condition for the boundedness of the Hilbert transform in $L^2(W)$ is the analogue of the scalar $A_2$ condition, namely:

$$\|W\|_{A_2} = \sup_I \left\| \left(\frac{1}{|I|}\int_I W\right)\left(\frac{1}{|I|}\int_I W^{-1}\right)\right\| < C.$$

The tools used include Carleson embedding theorems and weighted Haar basis.

In 1997, with N. Katz we considered the case where the weights were different $U \neq V$ and operator valued, see [**KP2**]. We found *sufficient* conditions, including $A_2$ conditions on the weights and some conditions which amounted to boundedness of some paraproducts and Haar multipliers (in the scalar case the boundedness of those operators was obtained from Gehring's lemma, unavailable in the matrix or operator case). The tools used involved: Haar multipliers, Schur's lemma, Cotlar's lemma, and dyadic paraproducts.

In 1998, Nazarov, Treil and Volberg considered a *dyadic toy model* in dimension $n = 1$, for different scalar valued weights $u \neq v$. The *toy model* is given by a collection of *signed Haar multipliers*

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

where $\sigma = \{\sigma_I\}_{I\in\mathcal{D}}$ is a sequence of $\pm 1$. They found necessary and sufficient conditions for the two-weights problem to be solved uniformly for all choices of signs. The techniques used involve: Bellman functions, Carleson embedding theorems and more, see [**NTV2**].

In 1999, Gillespie, Nazarov, Pott, Treil and Volberg found an infinite dimensional counterexample, for the case of equal operator valued weights $U = V = W$, see [**GNPTV**], [**NTV2**]. Independently N. Katz obtained the same result [**Ka2**]. This ruled out the possibility of having a Hunt-Muckenhoupt-Wheeden type result for the Hilbert transform in infinite dimensions, unless further conditions are imposed on the weights.

Very recently, Nazarov, Treil and Volberg have found a characterization for the two-weights problem [**NTV6**]. The conditions are of Sawyer-type (This type of conditions had appeared first in Sawyer's work on two-weights problems for maximal

functions [**Saw**]). They are in direct analogy to the ones found for the toy model. It had been accepted that whatever was true for the toy model should eventually be true for the Hilbert transform and hence for a larger class of singular integral operators, but the two-weights problem seemed to defeat, for a while, this folklore knowledge... In any case, the transition from the toy model to the Hilbert transform had not been, by any means, simple. As we mentioned in our first lecture, in a surprising development, Stephanie Petermichl showed in her PhD Thesis [**Pet1**] that one can write the Hilbert transform as an average over scaled dyadic grids $\mathcal{D}_s$ of the following multipliers [**Pet2**]:

$$\sum_{I\in\mathcal{D}_s} \langle f, h_I\rangle (h_{I_l} - h_{I_r}).$$

This opens the door for much more straightforward arguments from the dyadic world to the Hilbert transform.

6.3.2. *Sharp Constants.* In the one weight problem it becomes important to understand the dependence of the operator norm on the $A_p$ constant of the weight. More precisely, what can be said about the power $r$ in the inequality

$$\int |Tf(x)|^p w(x)dx \leq C\|w\|_{A_p}^r \int |f(x)|^p w(x)dx.$$

This problem was first considered by Stein in his thesis [**St3**] who proved that for the Hilbert transform and power weights in $A_2$ then $r = 2$. Buckley considered all $A_2$ weights for singular integrals, and the maximal and square functions. He gave sharp estimates for the maximal function, showing that $r = p'$ where $\frac{1}{p} + \frac{1}{p'} = 1$, see [**Bu2**]. In her PhD thesis Sanja Hukovic improved the best known constant for the square function, she used Bellman functions, see [**Hu**]. In joint work with Treil and Volberg they found the sharp constant for the square function, see [**HTV**]. In a another recent development Stefanie Petermichl and Janine Wittwer have found the sharp constant for the Hilbert transform for the so-called invariant $A_2$ weights, see [**PW**], [**Wi**]. The Bellman functions techniques seem to be ideally suited to understanding sharp constants.

## References


[Bu1] S. Buckley, *Summation conditions on weights.* Michigan Math. J. **40** (1993), 153–170. MR **94d:**42021

[Bu2] S. Buckley, *Estimates for operator norms and reverse Jensen's inequalities.* Trans. Amer. Math. Soc. **340** (1993), p. 253–272. MR **94a:**42011

[Bo] J. Bony, *Calcul symbolique et propagation des singularités pour les equations aux dérivées non-linéaires.* Ann. Sci. Ecole Norm. Sup. **14** (1981), 209–246. MR **84h:**35177

[Cal] A. Calderón, *Cauchy integrals on Lipschitz curves and related operators.* Proc. Nat. Acad. Sci. USA 74 (1977), 1324–1327. MR **57** # 6445

[Car] L. Carleson, *On convergence and growth of partial sums of Fourier series.* Acta Math. **116** (1966), 135–157. MR **33**:774

[Ch] M. Christ, *Lectures on singular integral operators.* Regional Conferences Series in Math. AMS. **77** (1990). MR **92f:**42021

[CF] R. Coifman, C. Fefferman, *Weighted norm inequalities for maximal fuctions and singular integrals.* Studia Math. **51** (1974), 241–250. MR **50** # 10670

[CJS] R. Coifman, P. W. Jones, S. Semmes, *Two elementary proofs of the $L^2$ boundedness of Cauchy integrals on Lipschitz curves.* J. of the AMS **2** (1989), 553–564. MR **90k:**42017

[CMM] R. Coifman, A. McIntosh, Y. Meyer, *L'intégrale de Cauchy définit un opérateur borné sur $L^2$ pour les courbes lipschitziennes.* Ann. of Math. **116** (1982) 361–387. MR **84m:**42027

[Co] M. Cotlar, *A combinatorial inequality and its applications to $L^2$ spaces.* Rev. Mat. Cuyana **1**(1955), 41–55. MR 18,219a

[CS1] M. Cotlar, C. Sadosky, *On the Helson-Szegö theorem and a related class of modified Toeplitz kernels.* Harmonic analysis in Euclidean spaces (Proc. Sympos. Pure Math., Williams Coll., Williamstown, Mass., 1978), Part 1, pp. 383–407, Proc. Sympos. Pure Math., XXXV, Part, Amer. Math. Soc., Providence, R.I., 1979 MR **81j:**42022

[CS2] M. Cotlar, C. Sadosky, *On some $L^p$ versions of the Helson-Szegö theorem.* Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II (Chicago, Ill., 1981), 306–317, Wadsworth Math. Ser., Wadsworth, Belmont, CA, 1983. MR **85i:**42015

[Da1] G. David, *Opérateurs intégraux singuliers sur certaines courbes du plan complexe.* Ann. Sci. Ec. Norm. Sup. **17** (1984) 157–189. MR **85k:**42026

[Da2] G. David, *Wavelets and singular integrals on curves and surfaces.* Lecture Notes in Mathematics, **1465**. Springer-Verlag, Berlin, 1991. MR **92k:**42021

[DJ] G. David, J.-L. Journé, *A boundedness criterion for generalized Calderón-Zygmund operators.* Ann. of Math. **20** (1984), 371–397. MR **85k:**42041

[DJS] G. David, J.-L. Journé, S. Semmes, *Opérateurs de Calderón-Zygmund, fonctions paraaccrétives et interpolation.* Rev. Mat. Iberoamericana **1** (1985) n$^o$ 4, 1–56. MR **88f:**47024

[Duo] J. Duoandikoetxea, *Análisis de Fourier.* Addison-Wesley/Universidad Autónoma de Madrid (1995). *Fourier analysis.* Translated and revised from the 1995 Spanish original by David Cruz-Uribe. Graduate Studies in Mathematics, 29. AMS, Providence, RI, 2001. MR **1 800 316**

[Fef] C. Fefferman, *Pointwise convergence of Fourier series.* Ann. of Math. **98** (1973), 551–571. MR **49:**5676

[FS] C. Fefferman, E. Stein, *$H^p$ spaces of several variables.* Acta Math **129** (1972), 137–193. MR 56# 6263

[FKP] R. Fefferman, C. Kenig, J. Pipher, *The theory of weights and the Dirichlet problem for elliptic equations.* Ann. of Math **134** (1991), no 1, 65–124. MR **93h:**31010

[GC-RF] J. García-Cuerva, J. L. Rubio de Francia, *Weighted norm inequalities and related topics.* North-Holland Mathematics Studies **116**, Amsterdam, Holland (1981). MR **87d:**42023

[Gar] J. Garnett, Bounded Analytic Functions. Pure and Applied Math. # 96, Academic Press, NY (1981) MR **83g:**30037

[Ge] F. W. Gehring, *The $L^p$-integrability of the partial derivatives of a quasiconformal mapping.* Acta Math. **130** (1973), 265–277. MR **53** # 5861

[GN1] J. Gilbert, A. Nahmod, *Hardy spaces and a Walsh model for bilinear cone operators.* Trans. Amer. Math. Soc. **351** (1999), no 8, 3267–3300. MR **99m**:42022

[GN2] J. Gilbert, A. Nahmod, *Boundedness of bilinear operators with non-smooth symbols.* Math. Res. Lett. **7** (2000), no 5-6, 767-778. MR 1-809 300

[GNPTV] A. Gillespie, F. Nazarov, S. Pott, S. Treil, A. Volberg, *Logarithmic growth for weighted Hilbert transform and vector Hankel operators.* Preprint at: http://www.math.msu.edu/~volberg/

[GT] L. Grafakos, R. Torres, *A multilinear Schur test and multiplier operators.* J. of Functional Analysis, to appear.

[Guz] M. de Guzmán, *Differentiation of Integrals in $\mathbb{R}^n$.* Lectures Notes in Math. **481**, Springer-Verlag (1975). MR **56** # 10031

[Ha] A. Haar, *Zur Theorie der orthogonalen Funktionen systeme.* Math. Ann. **69** (1910), 331–371.

[HS] H. Helson, G. Szegö, *A problem in prediction theory.* Ann. Math. Pura Appl. **51** (1960), 107–138. MR **22** # 12343

[Hof] S. Hoffman, *On the square root problem of Kato.* Lecture Notes in this volume.

[Huk] S. Hukovic, *Singular integral operators in weighted spaces and Bellman functions.* PhD Thesis, Brown University (1998).

[HTV] S. Hukovic, S. Treil, A. Volberg, *The Bellman function and sharp weighted inequalities for square functions.* In "Complex analysis, operators and related topics", Oper. Theory Adv. Appl. **113** (2000), p. 97–113, Birkhäuser Basel. MR 1 771 755

[Hu] R. Hunt, *On the convergence of Fourier series.* in "Orthogonal expansions and their continuous analogues", Ed. D. T. Haimo, Southern Illinois Press (1968), 235–255. MR **38** # 6296

[HMW] R. Hunt, B. Muckenhoupt, R. Wheeden, *Weighted norm inequalities for the conjugate function and the Hilbert transform.* Trans. AMS **176** (1973), 227–252. MR **47** # 701

[Ka1] N. H. Katz, *Lecture notes.* Preprint 1996.

[Ka2] N. H. Katz, *Matrix valued Paraproducts.* J. of Fourier Anal. and Appl. **300** (1997), 913–921. MR **99f:**42046

[KP1] N. H. Katz, M. C. Pereyra, *Haar multipliers, paraproducts and weighted inequalities.* Analysis of divergence (Orono, ME, 1997), Eds. Bray and Stanojević, 145–170, Appl. Numer. Harmon. Anal., Birkhäuser Boston, Boston, MA, 1999. MR **2001d:**42008

[KP2] N. H. Katz, M. C. Pereyra, *On the two weights problem for the Hilbert transform.* Rev. Mat. Iberoamericana **13** (1997) # 1, 211–243. MR **99c:**47036

[Kat] Y. Katznelson, *An introduction to harmonic analysis.* Dover Publications, Inc., New York, 1976. MR **54** # 10976

[Kol] A. N. Kolmogorov, *Une série de Fourier-Lebesgue divergente partout.* C. R. Acad. Sci. Paris Sér. A-B **183** (1926), 1387–1328.

[Koo] Koosis, *Introduction to $H_p$ spaces.* London Math. Society Lecture Series **40**, Cambridge University Press (1980). MR **81c:**30062

[Kr] S. Krantz, *A panorama of harmonic analysis.* Carus Mathematical Monographs **27**, MAA, Washington 1999. MR **2000m:42001**

[LT] M. Lacey, C. Thiele, *$L^p$ estimates on the bilinear Hilbert transform for $2 < p < \infty$.* Ann. of Math. (2) **146** (1997) # 3, 693–724. MR **99b:**42014

[MW] P. Massani, N. Wiener, *The prediction theory of multivariate stochastic processes. I. The regularity condition.* Acta Mat. **98** (1957), 111–150. MR **20** # 4323

[MV] M. Melnikov, J. Verdera, *A geometric proof of the $L^2$ boundedness of the Cauchy integral on Lipschitz graphs.* Internat. Math. Res. Notices 1995, no. 7, 325–331. MR **96f:**45011

[Me] Y. Meyer, *Ondelettes et Opérateurs II.* Herman, Paris, 1990. MR **93i:**42003

[Mu] T, Murai, *A real variable method for the Cauchy transform and analytic capacity.* Lecture Notes in Mathematics, **1307**, Springer-Verlag, Berlin, 1988. MR **89k:**30022

[NT] F. Nazarov, S. Treil, *The Hunt for Bellman function: applications to estimate of singular integral operators and to other classical problems in harmonic analysis.* (Russian) Algebra i Analiz **8** (1996), no. 5, 32-162; translation in St. Petersburg Math. J. **8** (1997), no. 5, 721–824. MR **99d:**42026

[NTV1] F. Nazarov, S. Treil, A. Volberg *Counterexample to the infinite-dimensional Carleson embedding theorem.* C. R. Acad. Sci. Paris Sr. I Math. **325** (1997), no. 4, 383–388. MR **98d:**46039

[NTV2] F. Nazarov, S. Treil, A. Volberg *The Bellman functions and two-weight inequalities for Haar multipliers.* J. Amer. Math. Soc. **12** (1999), no. 4, 909–928. MR **2000k:**42009

[NTV3] F. Nazarov, S. Treil, A. Volberg, *Cauchy integral and Calderón-Zygmund operators on non-homogeneous spaces.* Int. Math. Res. Not. **15** (1997), 703–726. MR **99e:**42028

[NTV4] F. Nazarov, S. Treil, A. Volberg, *How our Bellman function got its name.* Preprint (2001).

[NTV5] F. Nazarov, S. Treil, A. Volberg, *$T(b)$ theorem on non-homogeneous spaces.* Preprint (1999).

[NTV6] F. Nazarov, S. Treil, A. Volberg, Personal communication.

[Ni] N. K. Nikolskii, *Treatise on the shift operator.* Springer-Verlag (1985). MR **87i:**47042

[Per] M. C. Pereyra, *On the resolvent of dyadic paraproducts.* Rev. Mat. Iberoamericana **10** (1994) # 3, 627–664. MR **96e:**42021

[Pet1] S. Petermichl, *Dyadic shift and a logarithmic estimate for Hankel operators with matrix symbol.* C. R. Acad. Sci. Paris Sr. I Math. **330** (2000) # 6, 455–460. MR **2000m:**42016

[Pet2] S. Petermichl, *Some sharp estimates involving Hilbert transform.* PhD Thesis, Michigan State University (2000).

[PV] S. Petermichl, A. Volberg, *Why Riesz transforms are averages of dyadic shifts.* Preprint (2001).

[PW] S. Petermichl, J. Wittwer, *A sharp estimate for weighted Hilbert transform via Bellman functions.* Preprint (2000).

[Saw] E. T. Sawyer, *A characterization of a two weight norm inequality for maximal operators.* Studia Math. **75** (1982), p. 1–11. MR **84i:**42032

[St1] E. Stein, *Singular integrals and differentiability properties of functions.* Princeton Mathematical Series, No. 30 Princeton University Press, Princeton, N.J. 1970. MR **44** # 7280

[St2] E. Stein, *Harmonic Analysis: real-variable methods, orthogonality, and oscillatory integrals.* With the assistance of Timothy S. Murphy. Princeton Mathematical Series, 43. Monographs in Harmonic Analysis, III. Princeton University Press, Princeton, NJ, 1993. MR **95c:**42002

[St3] E. Stein, *On certain operators on $L^p$ spaces.* PhD Thesis, Chicago University 1956.

[Tao] T. Tao, *Harmonic analysis on the phase plane.* Lecture Notes for Math 254, Winter 2001, available at http://www.math.ucla.edu/ tao/254a.1.01w/

[Th] C. Thiele, *The quartile operator and pointwise convergence of Walsh series.* Trans. of AMS **352** (2000), no 12, 5745–5766. MR **2001b:**42007

[To1] X. Tolsa, *$L^2$-boundedness of the Cauchy integral operator for continuous measures.* Duke Math. J. 98:2 (1999), 269–304. MR **2000d:**31001

[To2] X. Tolsa, *Littlewood-Paley theory and the $T(1)$ theorem with non-doubling measures.* Preprint (2000).

[TV1] S. Treil, A. Volberg, *Wavelets and the angle between past and future.* J. Funct. Anal. **143** (1997), no 2, 269–308. MR **99k**:42073

[TV2] S. Treil, A. Volberg, *Completely regular multivariate stationary processes and the Muckenhoupt condition.* Pacific J. of Math. **190** (1999) no 2, 361–382. MR **2001a:**42008

[Vo] A. Volberg, *Matrix $A_p$ weights via $S$-functions.* J. Amer. Math. Soc. **10** (1997), no. 2, 445–466. MR **98a:**42013

[Wi] J. Wittwer, *A sharp estimate on the norm of the martingale transform.* Math. Res. Lett. **7** (2000), no 1, 1–12. MR **2001e:**42022

[Zy] A. Zygmund, *Trigonometric Series I, II.* Cambridge University Press, 2nd Edition (1959). MR **21** # 6498

Department of Mathematics and Statistics, University of New Mexico, Albuquerque, NM 87131

*E-mail address*: crisp@math.unm.edu