

# THE LINEAR BOUND IN $A_2$ FOR CALDERÓN–ZYGmund OPERATORS: A SURVEY

MICHAEL LACEY

*School of Mathematics  
Georgia Institute of Technology  
Atlanta GA 30332, USA  
E-mail: lacey@math.gatech.edu*

**Abstract.** For an  $L^2$ -bounded Calderón–Zygmund Operator  $T$  acting on  $L^2(\mathbb{R}^d)$ , and a weight  $w \in A_2$ , the norm of  $T$  on  $L^2(w)$  is dominated by  $C_T \|w\|_{A_2}$ . The recent theorem completes a line of investigation initiated by Hunt–Muckenhoupt–Wheeden in 1973 [12], has been established in different levels of generality by a number of authors over the last few years. It has a subtle proof, whose full implications will unfold over the next few years. This sharp estimate requires that the  $A_2$  character of the weight can be exactly once in the proof. Accordingly, a large part of the proof uses two-weight techniques, is based on novel decomposition methods for operators and weights, and yields new insights into the Calderón–Zygmund theory. We survey the proof of this Theorem in this paper.

**1. Introduction.** We survey recent developments on the norm behavior of classical Calderón–Zygmund operators on weighted spaces, with a special focus on the Muckenhoupt–Wheeden class of weights  $A_2$ . Indeed, after the 40 some-odd years since the class of  $A_p$  weights was introduced by Muckenhoupt and Wheeden, the theory has reached a natural milestone, with the sharp dependence of norm estimates being established. We concentrate on an exposition of the techniques behind this Theorem:

**THEOREM 1.** *Let  $T$  be an  $L^2$ -bounded Calderón–Zygmund operator acting on  $L^2(\mathbb{R}^d)$  (for precise definition see Definition 3). And, let  $w \in A_2$  (for precise definition see Definition 8). We then have the estimate*

$$\|Tf\|_{L^2(w)} \leq C_T \|w\|_{A_2} \|f\|_{L^2(w)}. \quad (2)$$

---

2010 *Mathematics Subject Classification:* Primary 42B20, 42B35; Secondary 47B38.

*Key words and phrases:* Calderón–Zygmund operators, weighted inequalities, maximal function, dyadic grids.

Research supported in part by grant NSF-DMS 0968499.

The paper is in final form and no version of it will be published elsewhere.

Here  $0 < C_T < \infty$  depends only on the operator  $T$  and dimension  $d$ .

The theory of weights came of age in 1973, with the result of Hunt–Muckenhoupt–Wheeden [12], which showed in one dimension that for  $w > 0$  a.e., the Hilbert transform is bounded on  $L^2(w)$  if and only if  $w \in A_2$ . This result was established for other suitable collections of singular integrals in higher dimensions. But, early proofs of this fact delivered a poor control on the norm, and indeed, the significance of the sharp dependence was a theme recognized over time.

The interest here is that the power of the  $A_2$  characteristic is in general sharp. Accordingly, the method of proof is delicate, and indeed sheds new light on methods and techniques appropriate for weighted spaces, as well as the structure of Calderón–Zygmund operators.

It is known that the estimate (2), together with sharp extrapolation [9], gives the sharp estimate on  $L^p(w)$ , for  $1 < p < \infty$ , accordingly, we concentrate on the  $L^2$  case. We recall definitions in the next two sections, and then recall different elements of one of the proofs known of this paper, the pleasingly direct proof of Hytönen–Pérez–Treil–Volberg [17]. The concluding section includes some historical remarks, and a variety of pointers to cognate results and approaches.

**Acknowledgments.** Due to my, and my father’s, personal connection to Polish mathematicians, it was my distinct pleasure to participate in the conference marking the centenary of birth of Józef Marcinkiewicz. It was a fitting testament to the life of Marcinkiewicz, of what was accomplished, what was lost, and finally, what the people of Poznań and Poland can now achieve, in their beautiful and prosperous city and country.

**2. Calderón–Zygmund Operators.** There are two canonical examples of Calderón–Zygmund operators that one can keep in mind. The first is Hilbert transform itself, defined by

$$Hf(x) := \lim_{\epsilon \rightarrow 0} \int f(x-y) \frac{dy}{y}.$$

Here, one should note that if  $f$  is Schwartz class, then the limit above exists for all  $x$ , and is referred to as the principal value of the integral. In brief,  $Hf(x) = \text{pv } f * \frac{1}{x}$ . But, the Hilbert transform is a convolution operator, which introduces a subtle simplification in its analysis in Lebesgue space. (There is no paraproduct to control.) Aside from the Hilbert transform, the other canonical convolution operators are the Beurling in the plane, and the vector of Riesz transforms.

A second example to keep in mind, one that motivated much of the development of the Theory in the 1980s, is the *Calderón Commutator* defined as follows. For Lipschitz function  $A$  on  $\mathbb{R}$ , let

$$C_A f(x) := \int f(y) \frac{A(x) - A(y)}{(x-y)^2} dy.$$

Note that we have  $C_A = [M_A, \frac{d}{dx}]H$ , where  $M_A$  is the operation of multiplication by  $A$ . We require  $A$  to be Lipschitz, as then we have, in some average sense,  $\frac{A(x)-A(y)}{(x-y)^2} \simeq (x-y)^{-1}$ . And the deep fact is that we have  $\|C_A\|_{2 \rightarrow 2} \lesssim \|A\|_{\text{Lip}}$ . But, this is not at all easy to prove!

A general definition of Calderón-Zygmund operators we will consider.

**CALDERÓN-ZYGMUND OPERATORS 3.** Let  $0 < \delta < 1$  and Let  $K(x, y) : \mathbb{R}^d \times \mathbb{R}^d \setminus \{(x, x) : x \in \mathbb{R}^d\} \rightarrow \mathbb{R}$  satisfy kernel estimates

$$\begin{aligned} |K(x, y)| &\leq C_T |x - y|^{-d-j}, \quad x \neq y \in \mathbb{R}^d, \\ |K(x, y) - K(x', y)| + |K(y, x) - K(y, x')| &\leq C_T \frac{|x - x'|^\delta}{|x - y|^{d+\delta}}, \end{aligned} \quad (4)$$

with the second condition holding provided  $|x - x'| < \frac{1}{2}|x - y|$ . Here,  $0 < C_T < \infty$ . Occasionally,  $K(x, y)$  will be referred to as a Calderón-Zygmund kernel.

Consider a linear operator  $T : L^2 \rightarrow L^2$  such that

$$Tf(x) = \int K(x, y)f(y) dy, \quad x \notin \text{supp } f,$$

for a fixed kernel  $K(x, y)$ .

We then say that  $T$  is Calderón-Zygmund Operator, and write  $T \in \text{CZO}_\delta$  and

$$\|T\|_{\text{CZO}_\delta} := \|T\|_{L^2(dx) \rightarrow L^2(dx)} + C_T < \infty.$$

One should note that in one dimension, that the kernel  $K(x, y) = \frac{1}{|x-y|}$  is a Calderón-Zygmund kernel, though the corresponding operator is not bounded. As well, it is hardly clear that the Calderón Commutator is a bounded operator. Thus, it is a natural question to find a simple characterization of the Calderón-Zygmund Operators. This class of operators was characterized by David and Journé [7], in the famous  $T1$  Theorem.

**THEOREM 5 ( $T1$  Theorem).** *An operator  $T$  with Calderón-Zygmund kernel, is  $L^2$ -bounded if and only if for  $\mathbf{T} > 0$ , we have the two uniform estimates over all cubes  $I \subset \mathbb{R}^d$ .*

$$\int_I |T\chi_I|^2 dx \leq \mathbf{T}^2 |I|, \quad (6)$$

$$\int_I |T^*\chi_I|^2 dx \leq \mathbf{T}^2 |I|. \quad (7)$$

In the second line,  $T^*$  is the adjoint of  $T$ , namely it has the kernel  $K(y, x)$ .

The import of this result is that the full  $L^2$ -inequality already follows from the boundedness of the operator on a very small set of functions, namely the indicators of cubes. We should note that this is not the formulation of the Theorem as in [7], but the version found in [53, Chapter V]. Clearly, we prefer the form above, over its more familiar formulation, as it does not require the supplemental space  $BMO$ . We refer to the two conditions (6) and (7) as *Sawyer testing conditions*, as their use in characterizing the bounded of operators first appeared in his two-weight Theorems on the Maximal Function [49] and the Fractional Integrals [50].

Let us use this Theorem to see that the Calderón Commutator is bounded. Let us take the interval  $I = [a, b]$ , and then using integration by parts,

$$\begin{aligned} C_A(\chi_I)(x) &= \int_a^b \frac{A(x) - A(y)}{(x - y)^2} dy \\ &= \frac{A(x) - A(b)}{(x - b)} - \frac{A(x) - A(a)}{(x - a)} + \int_a^b \frac{A'(y)}{x - y} dy. \end{aligned}$$

The first two terms are bounded by  $\|A\|_{\text{Lip}}$ , and the third is the Hilbert transform applied to  $A' \in L^\infty$ . Hence the testing condition for  $C_A$  follows from the  $L^2$ -boundedness of the Hilbert transform.

**3. The  $A_p$  Weights.** The  $A_p$  weights have the definition

DEFINITION 8. For  $w$  an a. e. positive function (a weight) on  $\mathbb{R}^d$ , we define the  $A_p$  characteristic of  $w$  to be

$$\|w\|_{A_p} := \sup_I |I|^{-1} \int_I w \, dx \cdot \left[ |I|^{-1} \int_I w^{-1/(p-1)} \, dx \right]^{p-1}, \quad 1 < p < \infty,$$

where the supremum is over all cubes in  $\mathbb{R}^d$ . In the case of  $p = 1$ , we set

$$\|w\|_{A_1} := \left\| \frac{Mw}{w} \right\|_\infty,$$

where  $M$  denotes the Maximal Function.

We note that  $\|w\|_{A_p}$  is not a norm, but continue to use the familiar notation. Below, we will also write  $w(I) = \int_I w \, dx$  for the (non-negative) measure with density  $w$ . It is a critical property, one that is key to the many beautiful properties of the  $A_p$  theory, that we have  $w > 0$  a.e. In particular, this means that  $\sigma := w^{-1/(p-1)}$  is unambiguously defined. Also, note that we have  $w\sigma^{p-1} \equiv 1$ , which casts the definition of  $A_p$  in a clear light: It requires that this pointwise equality continue to hold in an average sense, uniformly over all locations and scales.

We will refer to  $\sigma$  as the *dual measure*. This language is justified by a useful observation from [49]. The inequalities below are all equivalent for a linear operator  $T$ :

$$\begin{aligned} \|Tf\|_{L^p(w)} &\leq C \|f\|_{L^p(w)}, \\ \|T(\sigma f)\|_{L^p(w)} &\leq C \|Tf\|_{L^p(\sigma)}, \\ \|T^*(w\phi)\|_{L^{p'}(\sigma)} &\leq C \|Tf\|_{L^{p'}(w)}. \end{aligned} \tag{9}$$

To pass from the first line to the second, use the change of variables  $f \mapsto \sigma \cdot f$ . There is a routine calculation, which is based on the basic identity of the weighted theory that  $p(p' - 1) = p'$ . And, note that the last line is the formal dual inequality to the second. Thus, the inequality (9) expresses duality in a natural way: Interchange the roles of  $w$  and  $\sigma$ , and replace  $p$  by dual index  $p'$ .

Of course we are primarily interested in the case of  $p = 2$ . Two examples of  $A_2$  weights to keep in mind, in dimension 1, are as follows. First, for an arbitrary measurable set  $E \subset \mathbb{R}$ , and  $N > 0$ , the weight is  $w = N\chi_E + \chi_{\mathbb{R}-E}$ . As long as  $|E| > 0$ , one has  $\|w\|_{A_2} \leq \max(N, 1/N)$ . Indeed, we can assume  $N > 1$ . For an interval  $I$  we have

$$\frac{w(I)}{|I|} \cdot \frac{\sigma(I)}{|I|} \leq \frac{N|E \cap I| + |E^c \cap I|}{|I|} \leq N.$$

This shows that an  $A_2$  weight need not have any smoothness associated with it.

A second example is the borderline case of  $w(x) = |x|$ . This is *not* an  $A_2$  weight as the dual measure  $\sigma(x) = |x|^{-1}$  is not locally integrable. But if we mollify the zero, setting for  $0 < \alpha < 1$ ,  $w_\alpha(x) = |x|^\alpha$ , then we have  $\|w_\alpha\|_{A_2} \simeq (1 - \alpha)^{-1}$ . (Test on indicators of cubes.) It is for such examples that one can verify that  $\|H\|_{L^2(w_\alpha)} \simeq (1 - \alpha)^{-1}$ . But,

these examples are somewhat misleading, in that the simple behavior of their zeros is not at all indicative of intricacy of the general  $A_2$  measure.

We comment on classical Theorem of Muckenhoupt [32] concerning the  $A_p$  weights and the Maximal Function, defined by

$$Mf(x) := \sup_{t>0} (2t)^{-1} \int_{[-t,t]^d} |f(x-u)| dt.$$

**THEOREM 10.** *For  $w > 0$ , we have the following equivalences:*

1.  $w \in A_p$ ;
2.  $M$  is bounded as a map from  $L^p(w)$  to  $L^{p,\infty}(w)$ ;
3.  $M$  is bounded as a map from  $L^p(w)$  to  $L^p(w)$ .

*Note that the weak and strong type norms are equivalent.*

Clearly, the strong-type inequality implies the weak-type. Using the formulation (9), and applying the maximal function to a the indicator of a cube directly proves that  $w \in A_p$ . So, the content of the result is that the  $A_p$  property implies the strong type inequality. Here, the fact that  $w > 0$  a.e. is decisive, and the shortest – six lines – proof of this is due to Lerner [30]. Nevertheless, it seems confusing that the weak and strong types should be equivalent. The sharp dependence of the Maximal Function on the  $A_p$  characteristic is helpful here. For  $w \in A_p$ , how does the norm depend upon  $\|w\|_{A_p}$ ? Buckley [3] studied the question and proved

**THEOREM 11.** *For  $w \in A_p$ , we have*

$$\begin{aligned} \|M\|_{L^p(w) \rightarrow L^{p,\infty}(w)} &\lesssim \|w\|_{A_p}^{1/p}, \\ \|M\|_{L^p(w) \rightarrow L^p(w)} &\lesssim \|w\|_{A_p}^{1/(p-1)} \end{aligned} \quad (12)$$

Thus, the norm dependence is rather different. This is a basic set of inequalities, due to the notion of Rubio de Francia extrapolation [48].

A final, critical property for us is the so-called  $A_\infty$ -property. It states that an  $A_p$  weight cannot be too concentrated in any one cube. Indeed, as we will illustrate in the context of the Maximal Function, this is the single property of  $A_p$  weights that can be used to prove sharp results, and it can only be used once.

**LEMMA 13.** *Let  $w \in A_p$ ,  $I$  is a cube and  $E \subset I$ . We then have*

$$\frac{|E|}{|I|} \leq \|w\|_{A_p}^{1/p} \left[ \frac{w(E)}{w(I)} \right]^{1/p}. \quad (14)$$

*Proof.* The property that  $w > 0$  a.e. allows us to write

$$\begin{aligned} \frac{|E|}{|I|} &= \frac{\int_E w^{1/p}(x) w(x)^{-1/p} dx}{|I|} \\ &\leq \frac{w(E)^{1/p} \sigma(I)^{1/p'}}{|I|} \\ &= \left[ \frac{w(E)}{w(I)} \right]^{1/p} \frac{w(I)^{1/p} \sigma(I)^{1/p'}}{|I|} \end{aligned}$$

which proves the Lemma. ■

**4. Dyadic Grids.** Combinatorial arguments, stopping time arguments or decompositions of functions and operators, will frequently be done with the help of *dyadic grids*. In this section, we collect a number of elementary facts that we will need from time to time. At different moments, the methods and constructions of this section will in fact be decisive for us.

By a *grid* we mean a collection  $\mathcal{G}$  of cubes in  $\mathbb{R}^d$  with  $I \cap I' \in \{\emptyset, I, I'\}$  for all  $I, I' \in \mathcal{I}$ . The cubes can be taken to be a product of clopen intervals, although the behavior of functions or weights on the boundary of cubes in a grid will never be a concern for us. If  $G, G' \in \mathcal{G}$ , with  $G'$  the smallest element of  $\mathcal{G}$  that strictly contains  $G$ , we refer to  $G'$  as the  $\mathcal{G}$ -parent of  $G$ , and  $G'$  is a  $\mathcal{G}$ -child of  $G'$ . Let  $\text{Child}_{\mathcal{G}}(G')$  denote the collection of all  $\mathcal{G}$ -children of  $G'$ . If the grid is understood, the  $\mathcal{G}$  is suppressed in the notation.

We will say that  $\mathcal{G}$  is a *dyadic grid* if each cube  $I \in \mathcal{G}$  these two properties hold. (1)  $I$  is the union of  $2^d$ -subcubes of equal volume (the children of  $I$ ), and (2) the set of cubes  $\{I' \in \mathcal{G} : |I'| = |I|\}$  partition  $\mathbb{R}^d$ .

Associated to any dyadic grid  $\mathcal{D}$  are the usual conditional expectations and martingale differences are given by

$$\mathbb{E}_I f := \chi_I \cdot |I|^{-1} \int_I f \, dx, \quad \Delta_I f := \sum_{I' \in \text{Child}(I)} \mathbb{E}_{I'} f - \mathbb{E}_I f. \quad (15)$$

And, we also set

$$\mathbb{E}_k f := \sum_{\substack{I \in \mathcal{D} \\ \ell(I) = 2^{-k}}} \mathbb{E}_I f, \quad \Delta_k f := \sum_{\substack{I \in \mathcal{D} \\ \ell(I) = 2^{-k}}} \Delta_I f.$$

Then, by the Martingale Convergence Theorem, for  $f \in L^1(dx)$ ,  $\mathbb{E}_k f \rightarrow f$  a.e. And, by the Muckenhoupt Theorem for the Maximal Function, for  $w \in A_2$ , and  $f \in L^2(w)$ , the same conclusion holds.

**4.1. Proof of Buckley's Maximal Function Inequality.** As an illustration of the use the  $A_\infty$  condition, let us return to Buckley's estimate (12), and prove it in the dyadic case. Namely, for choice of dyadic grid  $\mathcal{D}$  in  $\mathbb{R}^d$ , we define the associated Maximal Function

$$Mf(x) := \sup_{I \in \mathcal{D}} \chi_I(x) \mathbb{E}_I |f|$$

where here we have introduced the notation  $\mathbb{E}_I \phi := |I|^{-1} \int_I \phi$ . Also, we are continuing with the same notation for the Maximal Function, suppressing its dependence on the choice of grid. For this operator, we will prove (12).

We make the definition of the stopping cubes.

**DEFINITION 16.** Let  $\mathcal{G}$  be a grid,  $\sigma$  a weight. Given cube  $I \in \mathcal{G}$ , we set the *stopping children of  $I_0$* , written  $\mathcal{C}(I)$ , to be the maximal dyadic cubes  $I' \subset I$  for which  $\mathbb{E}_{I'} \sigma > 4\mathbb{E}_I \sigma$ . A basic property of this collection is that

$$\sum_{I' \in \mathcal{C}(I)} |I'| < \frac{1}{4} |I|. \quad (17)$$

We set the *stopping cubes of  $I$*  to be the collection  $\mathcal{S}(I) = \bigcup_{j \geq 0} \mathcal{S}_j(I)$ , where we inductively define  $\mathcal{S}_0(I) := \{I\}$ , and  $\mathcal{S}_{j+1}(I) = \bigcup_{I' \in \mathcal{S}_j(I)} \mathcal{C}(I')$ . Thus, these are the max-

imal dyadic cubes, so that passing from parent to child in  $\mathcal{S}$ , the average value of  $\sigma$  is increasing by at least factor 4.

*Proof of (12).* It is the fundamental Theorem of Eric Sawyer [49] that for the Maximal Function, we have a powerful variant of the David Journé  $T1$  Theorem. Namely, for *any* pairs of weights  $(w, \sigma)$ , we have the equivalence between these two inequalities

$$\begin{aligned} \|M(\sigma f)\|_{L^p(w)} &\leq C_1 \|f\|_{L^p(\sigma)}, \\ \int_I M(\sigma \chi_I)^p w \, dx &\leq C_2^p \sigma(I), \quad I \in \mathcal{D}. \end{aligned} \quad (18)$$

Moreover, letting  $C_1$  and  $C_2$  be the optimal constants in these two inequalities, we have  $C_1 \simeq C_2$ . Notice that this shows that the Maximal Function bound reduces to a testing condition.

And so, in the special case that  $w \in A_p$ , and  $\sigma = w^{1-p'}$ , we estimate the constant  $C_2$ .

$$\int_I M(\sigma \chi_I)^p w(dx) \lesssim \|w\|_{A_p}^{p'} \sigma(I).$$

We do so by passing to the stopping cubes  $\mathcal{S}(I)$ , and estimating as below, where we will use some common manipulations in the  $A_p$  theory.

$$\begin{aligned} \int_I M(\sigma \chi_I)^p w(dx) &\leq \int_I \left[ \sum_{S \in \mathcal{S}(I)} \frac{\sigma(S)}{|S|} \cdot \chi_S \right]^p w(dx) \\ &\lesssim \sum_{S \in \mathcal{S}(I)} \int_I \left[ \frac{\sigma(S)}{|S|} \cdot \chi_S \right]^p w(I) \end{aligned} \quad (19)$$

$$\leq \|w\|_{A_p} \sum_{S \in \mathcal{S}(I)} \sigma(S) \quad (20)$$

$$\lesssim \|w\|_{A_p} \|\sigma\|_{A_{p'}} \sigma(I) \quad (21)$$

$$= \|w\|_{A_p}^{p'} \sigma(I). \quad (22)$$

This proves our estimate. Here, we have taken these steps.

- (19) Pointwise, the sum  $\sum_{S \in \mathcal{S}(I)} \frac{\sigma(S)}{|S|} \cdot \chi_S(x)$  is super-geometric, so comparable to its maximal term in the summand. This allows us to move the  $p$ th power inside the sum.
- (20) We are using the definition of  $A_p$  here.
- (21) The  $A_\infty$  property is decisive. By (17) and (14), we have, using the notation for the stopping children from Definition 16,  $\sum_{I' \in \mathcal{C}(I)} \sigma(I') \leq (1-c\|\sigma\|_{A_{p'}}^{-1})\sigma(I)$ , permitting us to sum a geometric series to get this estimate.
- (22) By inspection,  $\|\sigma\|_{A_{p'}} = \|w\|_{A_p}^{p'-1}$ .

■

**4.2. Random Dyadic Grids, Good and Bad Cubes.** We are used to thinking of a dyadic grid as being canonical, namely the cubes

$$\mathcal{D} := \{2^k(n + [0, 1)^d) : k \in \mathbb{Z}, n \in \mathbb{Z}^d\}.$$

This choice has a strong *edge effect*, it for instance distinguishes the origin, in that it is the vertex of infinitely many cubes. This sort of anomaly on the other hand should be typically rare. Quantifying this is achieved by a random grid. To present one typical example, if  $\mathcal{G}$  is a dyadic grid, and the interval  $[0, 1)^d$  is in  $\mathcal{G}$ , it has one  $2^d$  possible parents, found by taking the cube to be the product of one of the two intervals  $[0, 2)$  and  $[-1, 1)$  in each coordinate separately. To randomize  $\mathcal{G}$ , these possible choices of grids should be equally likely.

For any  $\beta = \{\beta_l\} \in \mathcal{B} := \{\{0, 1\}^d\}^{\mathbb{Z}}$ , and cube  $I$ , set

$$I \dot{+} \beta = I + \sum_{l < \ell(I)} \beta_l 2^{-l}, \quad (23)$$

where  $\ell(I) := |I|^{1/d}$  is the side length of the cube. Then, define the dyadic grid  $\mathbb{D}_\beta$  to be the collection of cubes  $\mathbb{D}_\beta = \{I \dot{+} \beta : I \in \mathcal{D}\}$ . This parametrization of dyadic grids appears explicitly in [13], and implicitly in [35, section 9.1].

Place the uniform probability measure  $\mathbb{P}$  on the space  $\mathcal{B}$ . Namely, the probability that any coordinate  $\beta_j$  takes any one value in  $\{0, 1\}^d$  is  $2^{-d}$ , and the coordinates  $\beta_j$  are independent of one another.

Let us see how the randomization affects the edge effect mentioned above. Let  $0 < \gamma < 1$  be a fixed parameter, and  $r \in \mathbb{Z}_+$  is a fixed integer. We say that a pair of intervals  $(I, J) \in \mathcal{D}_\beta$  are *good* if the smaller interval, say  $I$ , satisfies  $2^r \ell(I) < \ell(J)$ , and

$$\text{dist}(I, \partial J) \geq \ell(I)^\gamma \ell(J)^{1-\gamma}.$$

And an interval  $I$  is said to be *good*, if for all intervals  $J$  with  $\ell(J) > 2^r \ell(I)$ , we have that the pair  $(I, J)$  is good. Otherwise, we say that the cube is *bad*.

An important property of goodness is the independence of the location or scale of a cube  $I$  and its goodness. Take  $I \dot{+} \beta \in \mathcal{D}_\beta$ . The spatial position of  $I$  is given by the formula (23), which only depends upon  $\beta_j$  for  $2^{-j} < \ell(I)$ . And, for a larger cube  $J$ , the position of  $J$  can be written as

$$J + \sum_{j : 2^{-j} < \ell(I)} 2^{-j} \beta_j + \sum_{j : \ell(I) \leq 2^{-j} < \ell(J)} 2^{-j} \beta_j.$$

And hence, the position of  $J$  *relative to*  $I$  depends only on the coordinates  $\beta_j$  for  $\ell(I) \leq 2^{-j} < \ell(J)$ , and hence is independent of the location of  $I$ .

As a consequence the probability of a given cube is bad is independent of the location or scale of  $I$ . Denoting this probability by  $\pi_{r, \gamma}$ , it is an elementary exercise to see that  $\pi_{r, \gamma} \lesssim 2^{-r\gamma}$ . As it will turn out, it will be sufficient to have this probability less than one, for a choice of  $\gamma$  that depends upon the Calderón–Zygmund Operator  $T$ , and can be taken to be a small multiple of the constant  $\delta$  in the Definition 3.

**4.3. Haar Shifts, Dyadic Calderón–Zygmund Operators.** In one dimension, the Martingale Difference in (15) is given by the rank-one projection  $\Delta_I f = \langle f, h_I \rangle \cdot h_I$  where  $h_I$  is the Haar function, given by  $h_I := (-\chi_{I_-} + \chi_{I_+})|I|^{-1/2}$ , where  $I_\pm$  denotes the two children of  $I$ . And then, the simplest possible dyadic Calderón–Zygmund operator would be a *martingale transform*

$$Tf := \sum_I \varepsilon_I \langle f, h_I \rangle \cdot h_I.$$

The amenability of these operators to issues of measurability, and stopping time arguments has long been exploited, leading to a remarkable set of properties that are known for these objects.

Below, we will say that martingale transforms have *complexity* 1. To motivate this upcoming definition, let us recall the remarkable result of Stephanie Petermichl, concerning the Hilbert transform. In one dimension, consider the *dual* to the classical Haar function given by  $g_I = (-h_{I_-} + h_{I_+})/\sqrt{2}$ , and the special operator given by

$$Uf = U_\beta f := \sum_{I \in \mathcal{D}_\beta} \langle f, g_I \rangle \cdot h_I.$$

The Hilbert transform can be recovered from the operators  $U_\beta$ , namely the result below holds.

**THEOREM 24.** *Let  $\text{Dil}_\delta f(x) = f(x/\delta)$ . For non-zero constant  $c$ , we have*

$$\mathbb{E}_\beta \int_1^2 \text{Dil}_\delta U_\beta \text{Dil}_{1/\delta} \frac{d\delta}{\delta} = cH$$

Here, the expectation is taken over  $\beta \in \beta$ .

The Hilbert transform is distinguished by different properties, including being  $L^2$ -bounded, translation and dilation invariant, and (formally) satisfying  $H(\cos) = c \cdot \sin$ . By inspection,  $U_\beta$  is  $L^2$ -bounded. The averaging procedure above provides translation invariance, and dilation invariance, as we have used the Haar measure for the dilation group in the average. For the last property, note that  $g_I$  is a localized cosine, while  $h$  is a localized sinus. We refer the reader to [41, 13] for a precise proof of this Theorem.

The import of this result is that in situations where there is a translational and dilational invariance, one can prove results about the Hilbert transform by considering the much simpler operators  $U$  – where *tail behavior is no longer an issue*. Similar representations are available for other distinguished convolution kernels. For instance, the Beurling operator [8] can be recovered from martingale transforms, while the Riesz transforms are closer to the Hilbert transform [42]. The most general result known in this direction is [54], which shows that all smooth, odd one dimensional Calderón–Zygmund kernels can be obtained by a variant of Stephanie Petermichl’s method.

A more general definition is as follows. In higher dimensions, we mention that the martingale differences are finite rank projections, but there is no canonical choice of the Haar functions in this case. Below, by *Haar function* we will a function  $h_I$ , supported on  $I$ , constant on its children, and orthogonal to  $\chi_I$  (and no assumption on normalizations). And, by a *generalized Haar function* as a function  $h_I$  which is a linear combination of  $\chi_I$ , and  $\{\chi_{I'} : I' \in \text{Child}(I)\}$ . Such a function supported on  $I$  but need not be orthogonal to constants.

**DEFINITION 25.** For integers  $(m, n) \in \mathbb{Z}_+^2$ , we say that linear operator  $S$  is a (*generalized*) *Haar shift operator of parameters*  $(m, n)$  if

$$Sf(x) = \sum_{I \in \mathcal{D}_{I'}} \sum_{\substack{J' \in \mathcal{D} \\ I', J' \subset I}}^{(m, n)} \frac{\langle f, h_{J'}^{I'} \rangle}{|I|} h_{I'}^{J'}$$

where (1) in the second sum, the superscript  $(m,n)$  on the sum means that in addition we require  $\ell(I') = 2^{-m}\ell(I)$  and  $\ell(J') = 2^{-n}\ell(I)$ , and (2) the function  $h_{J'}^{I'}$  is a (generalized) Haar function on  $J'$ , and  $h_{I'}^{J'}$  is one on  $I'$ , with the joint normalization that

$$\|h_{J'}^{I'}\|_\infty \|h_{I'}^{J'}\|_\infty \leq 1. \quad (26)$$

In particular, this means that we have the representation

$$Sf(x) = \sum_{I \in \mathcal{D}} |I|^{-1} \int_I f(y) s_I(x, y) dy \quad (27)$$

where  $s_I(x, y)$  is supported on  $I \times I$ , with  $L^\infty$  norm at most one. We say that the *complexity* of  $S$  is  $\max(m, n)$ .

These are dyadic variants of Calderón–Zygmund operators. Note in particular that (26) is analogous to (4), while the ‘smoothness’ criteria is replaced by the parameters  $(m, n)$ . Consider a Haar shift operator. It is an  $L^2$ -bounded operator, in particular its norm is at most one. The situation for generalized shifts is far more subtle, and here, we should single out the following definition, for its distinguished role in the theory, though not necessarily this paper. We call an operator  $S$  a paraproduct if it is a generalized Haar shift of parameters  $(0, 1)$  or  $(1, 0)$ . To be specific, it, or its dual, is of the form

$$Sf = \sum_{I \in \mathcal{D}} \mathbb{E}_I f \cdot h_I \quad (28)$$

where  $h_I$  is a Haar function. A fundamental fact here is the following special case of the  $T1$  Theorem, in the dyadic case.

**THEOREM 29.** *Let  $S$  be as in (28). Then,  $S$  is  $L^2$ -bounded if and only if we have*

$$\|S\chi_I\|_2 \lesssim |I|^{1/2}.$$

This is a particular variant of the famous Carleson Embedding Theorem, and the main step in extending the David Journé  $T1$  Theorem to the dyadic setting.

More generally, we have the following quantitative form of the Dyadic  $T1$  Theorem.

**THEOREM 30.** *Let  $S$  be a generalized Haar shift operator of complexity  $\mu$ . Then  $S$  extends to a bounded operator on  $L^2(\mathbb{R}^d)$  if and only if we have, uniformly over cubes  $I$ ,*

$$\begin{aligned} \int_I |S\chi_I|^2 dx &\leq \mathbf{S}^2 |I|, \\ \int_I |S^* \chi_I|^2 dx &\leq \mathbf{S}^2 |I|, \end{aligned}$$

Moreover, we have  $\|S\|_{2 \rightarrow 2} \lesssim \mu \mathbf{S} + \mu^2$ .

There are two points to make here. The first is that there is a weak dependence of the norm of the operator as a function of the complexity  $\mu$ . The second, is the familiar, but not mentioned to this point, feature of the Calderón–Zygmund theory, that the operators have strong features. If  $S$  is a bounded operator, then, the sum in (27) is unconditional in  $I$ . The import of this feature, important for proof of the main Theorem, is that decompositions of dyadic cubes lead immediately to decompositions of operators.

In the second, an  $L^2$ -bounded Calderón–Zygmund operator is necessarily bounded on many other spaces. Of particular interest for us is the endpoint estimate for  $L^1$ :

**THEOREM 31.** *Let  $S$  be a dyadic shift operator of complexity  $\mu$ , which is bounded on  $L^2(\mathbb{R}^d)$ . Then, we have the estimate*

$$\sup_{\lambda>0} \lambda |\{Sf > \lambda\}| \lesssim \{(1 + \|S\|_{2 \rightarrow 2})^2 + \mu\} \|f\|_1. \quad (32)$$

This is a well-known principle, but the weak-dependence on the complexity is a new feature, see [17, Theorem 5.2].

**4.4. A Weighted Version of the Dyadic  $T1$  Theorem.** A crucial step is to prove a weighted version of the  $T1$  Theorem, one that holds for general weights. To emphasize this point, for a pair of weights  $(w, \sigma)$ , which are not necessarily related, we set the *two weight  $A_2$*  condition to be

$$\|w, \sigma\|_{A_2} := \sup_{I \in \mathcal{D}} \frac{w(I) \sigma(I)}{|I| |I|}.$$

We have this variant of the  $T1$  Theorem, for generalized Haar shift operators, in the weighted setting.

**THEOREM 33.** *Let  $S$  be a generalized Haar shift operator of complexity  $\mu$ , and  $(w, \sigma)$  a pair of weights. We have  $\|S(\sigma f)\|_{L^2(w)} \leq C \|f\|_\sigma$ , where*

$$\begin{aligned} C &\lesssim_d \mu \mathbf{S} + \mu^2 \|w, \sigma\|_{A_2}^{1/2} \\ &\int_I S(\sigma \chi_I)^2 w(dx) \leq \mathbf{S}^2 \sigma(I), \\ &\int_I S^*(w \chi_I)^2 \sigma(dx) \leq \mathbf{S}^2 w(I), \end{aligned}$$

Here, we are considering the weighted inequality in its natural form, see (9). And we are bounding the weighted norm of the Haar shift in terms of the two-weight  $A_2$  condition, as well as the testing condition. Of particular importance for the proof of the linear bound is the *very weak dependence* of the constants on the  $A_2$  characteristic. For the proof, see [37] and for the quantitative estimate above [17, Theorem 3.4]. In particular, the proof is a weighted variant of the usual proof of the dyadic  $T1$  Theorem, with an important point being that one should use *weighted Haar functions* to give the proof.

**5. The Random BCR Algorithm.** A proof of the  $T1$  Theorem must, implicitly, or explicitly, decompose the Calderón–Zygmund operator into appropriate components. In the language of random dyadic shifts, the remarkable result of [17, Theorem 4.1] is

**THEOREM 34.** *Let  $T$  be a Calderón–Zygmund Operator  $T$  with smoothness parameter  $\delta$ . Then, we can write*

$$T = C \mathbb{E}_\beta \sum_{(m,n) \in \mathbb{Z}_+^2} 2^{-(m+n)\delta/2} S_{m,n}^\beta \quad (35)$$

where (a) the expectation is taken over the space of random dyadic grid;  $S_{m,n}$  is a (random) dyadic shift; (c) the shifts of parameters  $(0, 1)$  and  $(1, 0)$  are generalized shifts; (d)

all other shifts need not be generalized; (e) the constant  $C$  is a function of  $T$ , and the smoothness parameter  $\delta$ . In particular, we will have, uniformly over the probability space,

$$\|S_{m,n}^\beta\|_{2 \rightarrow 2} \lesssim 1.$$

The focus with Theorem 24 is noteworthy. The prior result obtains the Hilbert transform as a convex combination of Haar shifts of *bounded* complexity. The Theorem above obtains it as a sum of Haar shifts, but one that is rapidly converging in complexity.

In the dyadic setting, similar results were proved by Figiel [11], and independently by [1], with the latter article being broadly influential. The method of expanding Calderón–Zygmund operators using this method reveals subtle approximation theory properties of these operators. This method is not random, but has the disadvantage of using operators which are not purely dyadic.

Indeed, the Theorem above looks wrong. Using standard Haar basis in one dimension, the inner product  $\langle Hh_{[0,1]}, h_{[0,2^k]} \rangle$  does not have the good decay properties in terms of complexity claimed above. Instead, one needs a concept like the goodness property of §4.2. And indeed, this is the main point, the inner product  $\langle Hh_I, h_J \rangle$  will be small, if the pair of intervals  $(I, J)$  are good.

**6. The Corona and the Linear Bound.** In the prior proofs of the linear bound for operators, one used the averaging technique of Petermichl, see Theorem 24, to represent the Calderón–Zygmund operator as an average of Haar shifts of *bounded* complexity. And then, verified the linear bound for such shifts. But, the representation (35) gives one another option. For an  $A_2$  weight, and an arbitrary Haar shift operator  $S$ , verify the linear bound, with only moderate growth in the complexity  $\mu$  of the Haar shift. Here, we can allow any polynomial dependence on the complexity. We have already described this in two different places, the first is the dyadic two-weight  $T1$  Theorem, Theorem 33, and the second is the weak- $L^1$  inequality, (32).

The relevant result is [17, Equation (5.5)].

**THEOREM 36.** *Let  $S$  be a generalized Haar shift operator of complexity  $\mu$ , and  $\mathbf{S} = \|S\|_{2 \rightarrow 2}$ . For  $w \in A_2$  and  $\sigma = w^{-1}$ , and cube  $I$ , we have*

$$\left[ \int_I |S\sigma\chi_I|^2 w(dx) \right]^{1/2} \lesssim (\mu + 1)(\mathbf{S} + \mu + 1) \|w\|_{A_2} \sigma(I)^{1/2}.$$

The method of proof here, aside from the dependence on the complexity, is derived from [22], and is a subtle extension of the method used to prove (12), the sharp dependence on the  $A_p$  characteristic for the Maximal Function. Indeed, the interested reader should first consult [22], which does not seek to track dependence of the bound on the constants. This argument uses the stopping cubes, as given in Definition 16. And, this decomposition is then used to decompose the operator. Then, the main step is to identify a notable extension of the John–Nirenberg inequalities that holds in the two-weight setting, for the decomposed operator. With this, we conclude our discussion of the proof of the linear bound for Calderón–Zygmund operators.

## 7. History

**7.1.** The weighted theory came of age with the paper [12] of Hunt–Muckenhoupt–Wheeden, showing that for non-negative weight  $w$ , the Hilbert transform is bounded on  $L^2(w)$  if and only if  $w \in A_2$ . Still, early proofs combined properties of the weight, including the  $A_\infty$  property we have used, with the Reverse Hölder inequality, and the good-lambda technique, to deliver estimates for the norm of the Hilbert transform of the order of  $\|w\|_{A_2}^2$ . These and the other comments about history reflect the authors' knowledge, but as he was not a participant in the development of the subject, they will certainly be incomplete. Apologies for omissions and gaps are extended in advance.

**7.2.** The rapid development of the  $A_p$  theory in the 1970's lent some credence to the thought that similar variants of the  $A_p$  condition could be used to characterize the two-weight inequalities as well. The characterization for the Hardy operator [33] confirmed this. It was a surprise when Sawyer [49] showed that such conditions cannot be used for the Maximal Function, instead one must use the testing conditions in (18). (For a little more detail, consult the counterexample discussed in Sawyer's paper.)

**7.3.** In the two-weight setting, the Hardy operator is somehow the easiest to study, the Maximal Function is the next step harder, then the fractional integrals, and finally the singular integrals. It took several years for the proof of the two-weight inequalities for the fractional integrals to be characterized. Sawyer gave the characterization in the  $T1$  language in [50]. This was contemporaneous with the David–Journé  $T1$  Theorem, but the connection was not widely appreciated until much later, especially by the work of Nazarov–Treil–Volberg. For history on this last point, see [55].

**7.4.** In the two-weight setting, one can have the fractional integral operators mapping  $L^p$  into  $L^p$ , indeed this is the hard case. In the case of  $L^p$  being mapped into  $L^q$ , there is a second characterization due to [18], also see [19, Chapter 3], and [51]. This characterization can be used to prove the sharp  $A_{p,q}$  bound for the fractional integrals on  $\mathbb{R}^d$ , see [21].

**7.5.** The paper of Sawyer–Wheeden [51] extends the two-weight inequality for the fractional integrals to homogeneous spaces; this is an interesting direction, which has been, and will be, explored in many different directions.

**7.6.** The question of the sharp dependence of the norm estimates of different operators, in terms of the  $A_p$  characteristic was specifically raised by Buckley [3], where the estimate (12) was proved. These bounds for the Maximal Function, together with the Rubio de Francia extrapolation technique leads to an important simplification of the analysis of many of the weighted inequalities. Namely, as is demonstrated in [9], identifying a sharp exponent in  $A_p$  characteristic for a single distinguished choice of  $p$  can prove the entire range of inequalities. For the Calderón–Zygmund operators, this index is  $p = 2$ .

**7.7.** In a different direction, Fefferman and Pipher [10] recognized the interest of this question, for singular integrals, with the weights  $w \in A_1$ . Wittwer [56] proved the linear bound for  $A_2$  weights, for martingale transforms. Petermichl and Volberg [45] showed the same for the Beurling operator, proving a conjecture of Astala on quasi-conformal maps as a consequence. Much later, a certain two-weight inequality for the Beurling operator

was proved [23] as a crucial step in proving another conjecture of Astala. These examples motivate in part the interest in such questions. Other motivations are derived from considerations in spectral theory [20], operator theory [39], and orthogonal polynomials [40].

**7.8.** It was an important breakthrough when Stefanie Petermichl proved the linear bound for the Hilbert transform [43]. This technique was based on the one hand, the representation of the Hilbert transform as an operator of complexity one, and on the other on the Bellman function method. The latter, deep, technique could require substantive modification if the Haar shift changes; these modifications were spelled out for the Riesz transforms in [44], and dyadic paraproducts [2].

**7.9.** An inequality used in some of these developments was the so-called bilinear embedding inequality of Nazarov–Treil–Volberg, [38]. The latter is a deep extension of the (weighted) Carleson embedding inequality to a two-weight setting. This inequality can also be interpreted in the language of fractional integrals, and the Sawyer method can be used to prove it, and extend it to other  $L^p$  settings [24], as well as vector-valued settings [52].

**7.10.** Andrei Lerner [28] devised a remarkable inequality, giving pointwise control of a function in terms of a sum of local oscillations. This inequality can be used to provide equally remarkable proofs of the sharp  $A_p$  inequalities for dyadic Calderón–Zygmund operators [5, 6], even in certain vector-valued situations. As of yet, it is not understood how to use this method on continuous Calderón–Zygmund operators.

**7.11.** Commutators of the form  $[T, M_b]$  are of interest, for instance, the Calderón Commutator can be written in this form. And the paper of Chung–Pereyra–Pérez [4] gives a complete discussion of this question in the setting of  $A_p$  weights. The two-weight variants appear to be largely open.

**7.12.** Lerner conjectured that the Littlewood–Paley Square function would have a different behavior in terms of its  $A_p$  characteristic. Namely, the case of  $p = 3$  was the critical index, and the power on the  $A_p$ -characteristic was  $1/2$ . He used his ‘local oscillation’ inequality, as well as other considerations, to prove this inequality in full generality [29].

**7.13.** The paper [22] proved the  $A_2$  linear bound for all Haar shifts, using a Corona decomposition that has been useful to the complete resolution of the Conjecture. The technique is obtaining a natural Corona decomposition in order to verify the testing conditions. This paper gave a rather poor dependence in terms of the complexity of the Haar shift parameter, but the role of complexity was only brought to the fore in [14].

**7.14.** Pérez–Treil–Volberg used the full strength of the non-homogeneous Harmonic analysis, and in particular the innovative paper [36], to prove a remarkable extension of the  $T1$  Theorem to the  $A_2$  setting. Loosely, an operator  $T$  with a Calderón–Zygmund kernel, then  $T$  extends to a bounded operator on  $L^2(w)$ ,  $w \in A_2$ , if and only if the testing conditions of Theorem 33 hold. Then, it was shown [16] that the linear bound holds

for Calderón–Zygmund operators with sufficiently smooth kernels. This proof used the Belykin–Coifman–Roklin [1] decomposition, and the method of [22] to verify the testing conditions. A short time later, Hytönen [14], used a random variant of the Belykin–Coifman–Roklin method to give a proof of the linear bound for arbitrary smoothness, again using the  $A_2$   $T1$  Theorem of [36]. This proof of the full conjecture was then streamlined in [17], giving the line of argument we have followed in this survey.

**7.15.** Lerner has conjectured that the weak-type bound on Calderón–Zygmund operators should obey the linear bound in  $A_p$  for all  $1 < p < \infty$ . This has been verified for dyadic Calderón–Zygmund operators, without careful attention to behavior of the exponents in terms of complexity [15], and for the smooth case, with enough derivatives, in [16]. The principal technique is again derived from [22], as well as a (simple) testing condition for the weak-type inequality for singular integrals given in [25], also see [26]. Indeed, this argument proves the linear bound in  $A_p$  for the *maximal truncations* of singular integrals, as this is the kind of operator that we have the testing conditions for. It seems likely that this conjecture would follow from Theorem 34, if one tracks complexity constants.

**7.16.** The endpoint case of these estimates is also of interest, namely, for  $p = 1$ . It is an elementary consequence of a covering lemma argument that for an arbitrary weight  $w$ , the Maximal Function  $M$  maps  $L^1(Mw)$  into  $L^{1,\infty}(w)$ . It was then the subject of conjecture if the same inequality holds for singular integrals. This was disproved for Haar multipliers by Maria Reguera [46], and then for the Hilbert transform by Reguera–Thiele [47].

**7.17.** With the failure of the most optimistic form of the conjecture above, one can then ask if its natural variant for  $w \in A_1$  holds. Namely, does the Hilbert transform map  $L^1(w)$  into  $L^{1,\infty}(w)$  for  $w \in A_1$ , with norm estimate dominated by a constant times  $\|w\|_{A_1}$ ? This also fails in the dyadic case [34]. On the other hand, the Hilbert transform does map  $L^1(w)$  into  $L^{1,\infty}(w)$ , and the best known upper bound on the norm is  $\|w\|_{A_1} \log_+ \|w\|_{A_1}$ . See [31] for more information on these last two points.

**7.18.** A interesting part of the linear bound in  $A_2$  is that one needs a substantive portion of two-weight theory to address it. This is Theorem 33 above. The general two-weight question is a rather intricate one, with a full discussion carrying us beyond the scope of this text. The interested reader should consult [55] for a general introduction, and the more recent papers [26, 25, 27, 36].

## References

- [1] G. Belykin, R. Coifman, V. Rokhlin, *Fast wavelet transforms and numerical algorithms*. I, Comm. Pure Appl. Math. 44(2) (1991), 141–183.
- [2] O. V. Beznosova, *Linear bound for the dyadic paraproduct on weighted Lebesgue space  $L_2(w)$* , J. Funct. Anal. 255(4), (2008), 994–1007.
- [3] S. M. Buckley, *Estimates for operator norms on weighted spaces and reverse Jensen inequalities*, Trans. Amer. Math. Soc. 340 (1993), 253–272.

- [4] D. Chung, C. Pérez, M. C. Pereyra, *Sharp bounds for general commutators on weighted Lebesgue spaces*, <http://arxiv.org/abs/1002.2396> (2010).
- [5] D. Cruz-Uribe, J. M. Martell, C. Pérez, *Sharp weighted estimates for approximating dyadic operators*, *Electron. Res. Announc. Math. Sci.* 17 (2010), 12–19.
- [6] D. Cruz-Uribe, J. M. Martell, C. Pérez, *Sharp weighted estimates for classical operators*, eprint=<http://arxiv.org/abs/1001.4254> (2010).
- [7] G. David, J.-L. Journé, *A boundedness criterion for generalized Calderón–Zygmund operators*, *Ann. of Math.* 120 (1984), 371–397.
- [8] O. Dragičević, A. Volberg, *Sharp estimate of the Ahlfors–Beurling operator via averaging martingale transforms*, *Michigan Math. J.* 51 (2003), 415–435.
- [9] O. Dragičević, L. Grafakos, M. C. Pereyra, S. Petermichl, *Extrapolation and sharp norm estimates for classical operators on weighted Lebesgue spaces*, *Publ. Mat.* 49 (2005), 73–91.
- [10] R. Fefferman, J. Pipher, *Multiparameter operators and sharp weighted inequalities*, *Amer. J. Math.* 119 (1997), 337–369.
- [11] T. Figiel, *Singular integral operators: a martingale approach*, [in:] *Geometry of Banach spaces* (Strobl, 1989), *London Math. Soc. Lecture Note Ser.* vol. 158, Cambridge Univ. Press, Cambridge 1990, 95–110.
- [12] R. Hunt, B. Muckenhoupt, R. Wheeden, *Weighted norm inequalities for the conjugate function and Hilbert transform*, *Trans. Amer. Math. Soc.* 176 (1973), 227–251.
- [13] T. Hytönen, *On Petermichl’s dyadic shift and the Hilbert transform*, *C. R. Math. Acad. Sci. Paris* 346 (2008), 1133–1136.
- [14] T. Hytönen, *The sharp weighted bound for general Calderon–Zygmund operators*, <http://arxiv.org/abs/1007.4330> (2010).
- [15] T. Hytönen, M. T. Lacey, M. C. Reguera, A. Vagharshakyan, *Weak and Strong-type estimates for Haar Shift Operators: Sharp power on the  $A_p$  characteristic*, <http://www.arxiv.org/abs/0911.0713> (2009).
- [16] T. Hytönen, M. T. Lacey, M. C. Reguera, E. T. Sawyer, I. Uriarte-Tuero, A. Vagharshakyan, *Weak and Strong type  $A_p$  Estimates for Calderon–Zygmund Operators*, <http://www.arxiv.org/abs/1006.2530>, (2010).
- [17] T. Hytönen, C. Pérez, S. Treil, A. Volberg, *Sharp weighted estimates of the dyadic shifts and  $A_2$  conjecture*, eprint=<http://arxiv.org/abs/1010.0755>.
- [18] M. A. Gabidzashvili, V. Kokilashvili, *Two weight weak type inequalities for fractional type integrals*, *Ceskoslovenska Akademie Ved.* 45 (1989), 1–11.
- [19] I. Genebashvili, A. Gogatishvili, V. Kokilashvili, M. Krbeč, *Weight theory for integral transforms on spaces of homogeneous type*, *Pitman Monographs and Surveys in Pure and Applied Mathematics* vol. 92, Longman, Harlow, 1998.
- [20] H.-K. Kwon, S. Treil, *Similarity of operators and geometry of eigenvector bundles*, <http://www.arxiv.org/abs/0712.0114> (2007).
- [21] M. T. Lacey, K. Moen, C. Pérez, R. H. Torres, *Sharp weighted bounds for fractional integral operators*, *J. Funct. Anal.* 259 (2010), 1073–1097.
- [22] M. T. Lacey, S. Petermichl, M. C. Reguera, *Sharp  $A_2$  inequality for Haar shift operators*, *Math. Ann.* 348 (2010), 127–141.
- [23] M. T. Lacey, E. T. Sawyer, I. Uriarte-Tuero, *Astala’s Conjecture on Distortion of Hausdorff Measures under Quasiconformal Maps in the Plane*, <http://www.arxiv.org/abs/0805.4711>, *Acta Math.*, to appear (2008).
- [24] M. T. Lacey, E. T. Sawyer, I. Uriarte-Tuero, *Two Weight Inequalities for Discrete Positive Operators*, eprint=<http://www.arxiv.org/abs/0911.3437>, submitted, (2009).

- [25] M. T. Lacey, E. T. Sawyer, I. Uriarte-Tuero, *A characterization of two weight norm inequalities for maximal singular integrals with one doubling measure*, <http://arxiv.org/abs/0805.0246>, submitted to Analysis and PDE, (2008).
- [26] M. T. Lacey, E. T. Sawyer, I. Uriarte-Tuero, *Two Weight Inequalities for Maximal Truncations of Dyadic Calderón-Zygmund Operators*, Submitted, <http://www.arxiv.org/abs/0911.3920>, submitted, (2009).
- [27] M. T. Lacey, E. T. Sawyer, I. Uriarte-Tuero, *A Two Weight Inequality for the Hilbert transform Assuming an Energy Hypothesis*, <http://www.arXiv.org/abs/1001.4043> (2010).
- [28] A. K. Lerner, *A pointwise estimate for local sharp maximal function with applications to singular integrals*, Bull. LMS (2009), to appear.
- [29] A. K. Lerner, *Sharp weighted norm inequalities for Littlewood-Paley operators and singular integrals*, eprint=<http://arxiv.org/abs/1005.1422> (2010).
- [30] A. K. Lerner, *An elementary approach to several results on the Hardy-Littlewood maximal operator*, Proc. Amer. Math. Soc. 136 (2008), 2829–2833.
- [31] A. K. Lerner, S. Ombrosi, C. Pérez,  *$A_1$  bounds for Calderón-Zygmund operators related to a problem of Muckenhoupt and Wheeden*, Math. Res. Lett. 16 (2009), 149–156.
- [32] B. Muckenhoupt, *Weighted norm inequalities for the Hardy maximal function*, Trans. Amer. Math. Soc. 165 (1972), 207–226.
- [33] B. Muckenhoupt, *Hardy's inequality with weights*, Studia Math. 44 (1972), 31–38.
- [34] F. Nazarov, A. Reznikov, V. Vasyunin, A. Volberg, *Personal Communication*, (2010).
- [35] F. Nazarov, S. Treil, A. Volberg, *The Tb-theorem on non-homogeneous spaces*, Acta Math. 190 (2003), 151–239.
- [36] F. Nazarov, S. Treil, A. Volberg, *Two weight estimate for the Hilbert transform and Corona decomposition for non-doubling measures*, <http://arxiv.org/abs/1003.1596> (2004).
- [37] F. Nazarov, S. Treil, A. Volberg, *Two weight inequalities for individual Haar multipliers and other well localized operators*, Math. Res. Lett. 15 (2008), 583–597.
- [38] F. Nazarov, S. Treil, A. Volberg, *The Bellman functions and two-weight inequalities for Haar multipliers*, J. Amer. Math. Soc. 12 (1999), 909–928.
- [39] F. Nazarov, A. Volberg, *The Bellman function, the two-weight Hilbert transform, and embeddings of the model spaces  $K_\theta$* , J. Anal. Math. 87 (2002), 385–414.
- [40] F. Nazarov, F. Peherstorfer, A. Volberg, P. Yuditskii, *Asymptotics of the best polynomial approximation of  $|x|^p$  and of the best Laurent polynomial approximation of  $\operatorname{sgn}(x)$  on two symmetric intervals*, Constr. Approx. 29 (2009), 23–39.
- [41] S. Petermichl, *Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol*, language=English, with English and French summaries, C. R. Acad. Sci. Paris Sér. I Math. 330 (2000), 455–460.
- [42] S. Petermichl, S. Treil, A. Volberg, *Why the Riesz transforms are averages of the dyadic shifts?*, [in:] Proceedings of the 6th International Conference on Harmonic Analysis and Partial Differential Equations (El Escorial, 2000), Publ. Mat. (2002), 209–228.
- [43] S. Petermichl, *The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical  $A_p$  characteristic*, Amer. J. Math. 129 (2007), 1355–1375.
- [44] S. Petermichl, *The sharp weighted bound for the Riesz transforms*, Proc. Amer. Math. Soc. 136 (2008), 1237–1249.
- [45] S. Petermichl, A. Volberg, *Heating of the Ahlfors-Beurling operator: weakly quasiregular maps on the plane are quasiregular*, Duke Math. J. 112 (2002), 281–305.
- [46] M. C. Reguera, *On Muckenhoupt-Wheeden Conjecture*, <http://www.arxiv.org/abs/1008.3943>, (2010).

- [47] M. C. Reguera, C. Thiele, *The Hilbert transform does not map  $L^1(Mw)$  to  $L^{1,\infty}(w)$* , <http://www.arxiv.org/abs/1011.1767>, (2010).
- [48] J. L. Rubio de Francia, *Factorization theory and  $A_p$  weights*, Amer. J. Math. 106 (1984), 533–547.
- [49] E. T. Sawyer, *A characterization of a two-weight norm inequality for maximal operators*, Studia Math. 75 (1982), 1–11.
- [50] E. T. Sawyer, *A characterization of two weight norm inequalities for fractional and Poisson integrals*, Trans. Amer. Math. Soc. 308 (1988), 533–545.
- [51] E. Sawyer, R. L. Wheeden, *Weighted inequalities for fractional integrals on Euclidean and homogeneous spaces*, Amer. J. Math., 114 (1992), 813–874.
- [52] J. Scurry, *A Characterization of Two-Weight Inequalities for a Vector-Valued Operator*, <http://arxiv.org/abs/1007.3089>, (2010).
- [53] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series vol. 43, Princeton University Press, Princeton, NJ, 1993.
- [54] A. Vagharshakyan *Recovering Singular Integrals from Haar Shifts*, <http://arxiv.org/abs/1007.2994>, (2009).
- [55] A. Volberg, *Calderón-Zygmund capacities and operators on nonhomogeneous spaces*, CBMS Regional Conference Series in Mathematics vol. 100, Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2003.
- [56] J. Wittwer, *A sharp estimate on the norm of the martingale transform*, Math. Res. Lett. 7 (2000), 1–12.