

Weighted Inequalities and Dyadic Harmonic Analysis

María Cristina Pereyra

Abstract We survey the recent solution of the so-called A_2 conjecture, that states: all Calderón–Zygmund singular integral operators are bounded on $L^2(w)$ with a bound that depends linearly on the A_2 characteristic of the weight w . We also survey corresponding results for commutators. We highlight the interplay of dyadic harmonic analysis in the solution of the A_2 conjecture, especially Hytönen’s representation theorem for Calderón–Zygmund singular integral operators in terms of Haar shift operators. We describe Chung’s dyadic proof of the corresponding quadratic bound on $L^2(w)$ for the commutator of the Hilbert transform with a BMO function, and we deduce *sharpness* of the bounds for the dyadic paraproduct on $L^p(w)$ that were obtained extrapolating Beznosova’s linear bound on $L^2(w)$. We show that if an operator T is bounded on the weighted Lebesgue space $L^r(w)$ and its operator norm is bounded by a power α of the A_r characteristic of the weight, then its commutator $[T, b]$ with a function b in BMO will be bounded on $L^r(w)$ with an operator norm bounded by the increased power $\alpha + \max\{1, \frac{1}{r-1}\}$ of the A_r characteristic of the weight. The results are sharp in terms of the growth of the operator norm with respect to the A_r characteristic of the weight for all $1 < r < \infty$.

Keywords Dyadic harmonic analysis • Weighted inequalities • A_2 conjecture • Haar shift operators • Commutator • Dyadic paraproduct • A_p weights • Random dyadic grids • Haar basis • Hilbert transform • Fourier multiplier

1 Introduction

The main problem of study in this chapter is the *weighted L^p inequalities*

M.C. Pereyra (✉)

Department of Mathematics and Statistics, University of New Mexico, MSC03 21501, Albuquerque, NM 87131-0001, USA

e-mail: crisp@math.unm.edu

$$\|Tf\|_{L^p(v)} \leq C(u, v) \|f\|_{L^p(u)}, \tag{1}$$

where $f \in L^p(u)$ iff $\|f\|_{L^p(u)} := (\int |f(x)|^p u(x) dx)^{1/p} < \infty$, and u, v are locally integrable positive a.e. functions defined on \mathbb{R}^n . When $u \equiv 1$ we denote $\|f\|_p := \|f\|_{L^p(u)}$.

Two-weight problem: *Find necessary and sufficient conditions on the weights so that above inequality holds for a given operator or class of operators T , and find the optimal rate of dependence of the constant $C(w)$ on the weight.*

In this survey we will concentrate on *one-weight inequalities*, $u = v$, for Calderón–Zygmund singular integral operators, more specifically for the Hilbert transform $T = H$ and for the commutator of the Hilbert transform with a function b in the space BMO of bounded mean oscillation, namely $T = [b, H] := bH - Hb$.

The Hilbert transform is bounded on $L^p(w)$ if and only if the weight w belongs to the Muckenhoupt A_p class [31]. This is also true for Calderón–Zygmund singular integral operators [11]. A weight w is in the *Muckenhoupt A_p class* if

$$[w]_{A_p} := \sup_I \left(\frac{1}{|I|} \int_I w \right) \left(\frac{1}{|I|} \int_I w^{-1/(p-1)} \right)^{p-1} < \infty. \tag{2}$$

In the last decade there has been a flurry of activity trying to identify the exact dependence of the operator bound on the A_p characteristic, $[w]_{A_p}$, of the weight. This dependence was first proved to be linear in A_2 for a few dyadic operators [30, 79, 80], then for the Beurling–Ahlfors [70], Hilbert [67], and Riesz transforms [68], and for the dyadic paraproduct [4]. Finally Tuomas Hytönen solved in the positive the A_2 conjecture [33]: If T is a Calderón–Zygmund singular integral operator, $w \in A_2$, then the dependence on the A_2 characteristic of the weight is linear, that is,

$$\|Tf\|_{L^2(w)} \leq C[w]_{A_2} \|f\|_{L^2(w)}. \tag{3}$$

Sharp extrapolation [20] then yields the correct L^p bounds for the class of Calderón–Zygmund singular integral operators:

$$\|Tf\|_{L^p(w)} \leq C_p [w]_{A_p}^{\max\{1, \frac{1}{p-1}\}} \|f\|_{L^p(w)}.$$

Remark. The long-standing two-weight problem for the Hilbert transform “à la Muckenhoupt” is an outstanding open problem: Characterize the pairs of weights (u, v) , in terms of conditions like the A_p condition in the one-weight problem, for which (1) holds. Recently there has been progress due to Lacey, Sawyer, Shen, and Uriarte-Tuero [45]. Note that Cotlar and Sadosky solved, years ago, the two-weight problem “à la Helson–Szëgo,” that is, using complex analysis techniques [13, 14].

In this chapter we want to highlight the interplay with dyadic harmonic analysis [60] in the solution of the A_2 conjecture. Initially the A_2 conjecture was shown to hold, one at a time, for dyadic operators and for operators such as the Hilbert trans-

form that have lots of symmetries. Stephanie Petermichl showed, in groundbreaking work in 2000, that the Hilbert transform can be written as an appropriate average of dyadic shift operators [32, 66], and later she showed, in a tour de force using Bellman function techniques, that for the dyadic shift operators the A_2 conjecture is true and therefore also for the Hilbert transform [67]. This work represented a quantum jump in our understanding of singular integral operators. Until then a simpler dyadic model, the *martingale transform*, was considered the toy model for singular integrals. One would first try to prove results for this model and then hope to prove them for a genuine singular integral operator, but the transition was by no means automatic [60]. Petermichl’s representation theorem made this transition trivial for the Hilbert transform. For a while it seemed that the miracle of this representation theorem was a consequence of the symmetries of the operator. Similar constructions were found for other symmetric operators: the Riesz transform (n -dimensional analogue of the Hilbert transform) [68], the Beurling–Ahlfors transform [70], and for sufficiently smooth convolution Calderón–Zygmund singular integral operators [76]. The fact that for the Beurling–Ahlfors transform the A_2 conjecture holds for $p > 2$ (linear estimate in A_p characteristic in the range of $p > 2$) had important implications in the theory of quasiconformal mappings [2].

All these operators have a representation as averages of dyadic Haar shift operators of bounded complexity. In 2008, Oleksandra Beznosova showed that the linear bound on $L^2(w)$ also holds for the dyadic paraproduct, an operator not in the above class [4]. Hytönen was able to prove a representation theorem valid for *all* Calderón–Zygmund singular integral operators (not only convolution) in terms of dyadic Haar shift operators of arbitrary complexity, paraproducts, and adjoints of the paraproducts. Different groups of researchers had already shown that the A_2 conjecture was true for all these Haar shift operators [17, 18, 44], using techniques other than Bellman function which had dominated the scene until then. However, the dependence of the operator bound on the complexity was exponential and prevented one from deducing the A_2 conjecture for general Calderón–Zygmund singular operators. Only for those operators that were averages of dyadic shift operators of bounded complexity one could deduce the A_2 conjecture. Hytönen was able to overcome this obstacle as well, proving a polynomial dependence on the complexity and the linear dependence on the A_2 characteristic of the weight for Haar shift operators, therefore proving the A_2 conjecture [33]. Precursors to Petermichl’s and Hytönen’s results can be found in Figiel’s work [24]. Nowadays some of the simpler arguments yielding polynomial and even linear dependence on the complexity use minimally Bellman functions [54, 74], or do not use them at all [37, 42].

The commutator $[b, H]$ is more singular than the operator H , and this is reflected on the nature of its bounds on weighted L^p spaces. Daewon Chung showed in [8] that

$$\|[b, H]f\|_{L^2(w)} \leq C [w]_{A_2}^2 \|f\|_{L^2(w)}. \tag{4}$$

That is, the dependence on the A_2 characteristic of the operator bound is now *quadratic* as opposed to the *linear* bound enjoyed by the Hilbert transform. Chung’s proof can be labeled as a dyadic proof. It suffices to consider the commutator

with Petermichl's Haar shift operator [69]. Then known linear bounds for the shift operator [67] and for the dyadic paraproduct [4] can be used, and Bellman function arguments can be invoked as did all of Chung's predecessors until then. We observe that the sharp bounds for the commutator of the Hilbert transform imply that Beznosova's bounds [4] are the sharp bounds for the dyadic paraproduct in $L^p(w)$, which was not known until now. The author in collaboration with Chung and Carlos Pérez established a transference theorem that states that if a linear operator T obeys a linear bound on $L^2(w)$ then its commutator with a BMO function obeys a quadratic bound [10]. In light of Hytönen's theorem this means that *all* commutators of Calderón–Zygmund singular operators with BMO functions obey a quadratic bound as in inequality (4). The argument follows the classical Coifman, Rochberg, and Weiss argument [12] exploiting the Cauchy integral formula and some very precise quantitative results in the theory of A_2 weights and BMO functions. Generalizations of these results to commutators with fractional integrals and to the two-weight setting appear in [16], and weak-type estimates and strong estimates involving instead the A_1 characteristic of the weight appear in [59]. In this note we present the simple modifications necessary to state a transference theorem that provides bounds on $L^r(w)$, $r \neq 2$, for the commutator given corresponding bounds on $L^r(w)$ for the initial operator.

The author strongly believes that Petermichl and Hytönen's representation theorem in terms of dyadic operators could have important consequences in applications, in the same way that the $T(1)$ theorem [19] had repercussions in computational harmonic analysis via the Beylkin, Coifman, and Rokhlin algorithm to decompose singular integral operators [3].

This chapter is organized as follows. In Sect. 2 we define the Hilbert transform and the dyadic Haar shift operators, recall some of their basic properties, state Petermichl's representation theorem, and show how it provides a straightforward proof of the boundedness of the Hilbert transform on $L^p(\mathbb{R})$ (Riesz's theorem). In Sect. 3 we discuss weighted inequalities for the Hilbert transform and recount the prehistory of linear estimates for dyadic operators on $L^2(w)$. We state the sharp extrapolation theorem and deduce $L^p(w)$ bounds from linear bounds and observe that these bounds are sometimes sharp, but not always, as Buckley's estimates for the maximal function show. We then define the Haar shift operators of complexity (m, n) , discuss their boundedness properties, and state Hytönen's theorem (the A_2 conjecture), as well as his representation theorem. In Sect. 4 we define the commutator, state its boundedness properties, and sketch Chung's dyadic proof of the quadratic estimate on $L^2(w)$. We note that this quadratic estimate is sharp, and we show that Chung's dyadic method of proof implies that Beznosova's bound for the dyadic paraproduct is sharp as well. Finally we state a variation of the transference theorem for commutators on $L^r(w)$ with $r \neq 2$ and present its proof in the Appendix.

2 Hilbert Transform Versus Dyadic Shift Operators

We define the Hilbert transform both on Fourier and space domains, and we describe its boundedness and symmetry properties. We introduce the dyadic intervals, the random dyadic grids, and corresponding Haar bases, and we emphasize some of the properties these bases share with wavelets such as being an unconditional basis on L^p spaces. We define Petermichl’s Haar shift operators and describe their symmetry properties; we state Petermichl’s representation theorem and show how it provides a straightforward proof of the boundedness of the Hilbert transform on $L^p(\mathbb{R})$.

2.1 Hilbert Transform

In this section we recall the definition of the Hilbert transform on Fourier domain as a Fourier multiplier and on space domain as a convolution with a singular kernel. We also recall how symmetry properties completely characterize the Hilbert transform. These are well-known facts that can be found in any Fourier analysis book such as [21, 26, 73]. You will also find here the definition of *BMO*, the space of functions of bounded mean oscillation.

2.1.1 Fourier Multiplier

The *Fourier transform* of a Schwartz function is defined by

$$\widehat{f}(\xi) := \int_{\mathbb{R}} f(x) e^{-2\pi i \xi x} dx.$$

With some work one can define the Fourier transform on $L^2(\mathbb{R})$ and show that it is an isometry, that is, $\|\widehat{f}\|_2 = \|f\|_2$ (Plancherel’s identity).

On Fourier side the *Hilbert transform* can be defined as a *Fourier multiplier*:

$$\widehat{Hf}(\xi) = -i \operatorname{sgn}(\xi) \widehat{f}(\xi), \tag{5}$$

where $\operatorname{sgn}(\xi) = 1$ if $\xi > 0$, $\operatorname{sgn}(\xi) = -1$ if $\xi < 0$, and is zero at $\xi = 0$.

The absolute value of the symbol $m_H(\xi) := -i \operatorname{sgn}(\xi)$ is 1 a.e., and Plancherel’s identity used twice implies that $H : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$ and that it is an isometry:

$$\|Hf\|_2 = \|\widehat{Hf}\|_2 = \|\widehat{f}\|_2 = \|f\|_2.$$

2.1.2 Singular Integral Operator

Since the Hilbert transform is given on Fourier side by

$$\widehat{Hf}(\xi) = m_H(\xi) \widehat{f}(\xi),$$

multiplication on Fourier side comes from convolution on space with the distributional kernel k_H which is the inverse Fourier transform of the multiplier m_H . A calculation yields

$$k_H(x) := (m_H)^\vee(x) = \text{p.v.} \frac{1}{\pi x}.$$

For a distributional kernel, the integration must be done in the principal value sense:

$$Hf(x) = k_H * f(x) = \text{p.v.} \frac{1}{\pi} \int \frac{f(x-y)}{y} dy := \lim_{\epsilon \rightarrow 0} \frac{1}{\pi} \int_{|x-y|>\epsilon} \frac{f(y)}{x-y} dy. \quad (6)$$

Had the kernel k_H been integrable, boundedness on $L^p(\mathbb{R})$ would be a consequence of the Hausdorff–Young’s inequality for $p \geq 1$: if $g \in L^1(\mathbb{R})$, $f \in L^p(\mathbb{R})$, then $\|g * f\|_p \leq \|g\|_1 \|f\|_p$. But k_H is not in $L^1(\mathbb{R})$; despite this fact, H is bounded on $L^p(\mathbb{R})$ for all $1 < p < \infty$, as Marcel Riesz proved in 1927:

$$\|Hf\|_p \leq C_p \|f\|_p.$$

However, H is not bounded on $L^1(\mathbb{R})$ nor on $L^\infty(\mathbb{R})$, but there are appropriate substitutes: H is of weak type (1,1) and is bounded on BMO [21, 26, 73]. Recall that a function $b : \mathbb{R} \rightarrow \mathbb{R}$ belongs to BMO , the *space of bounded mean oscillation*, if and only if

$$\|b\|_{BMO} := \sup_I \frac{1}{|I|} \int_I |b(x) - m_I b| dx < \infty, \quad (7)$$

where $m_I b$ denotes the integral average of b on the interval I , $m_I b = \frac{1}{|I|} \int_I b(x) dx$. This space was introduced by John and Nirenberg in the 1960s [39]. The space of bounded functions $L^\infty(\mathbb{R})$ is a proper subset of BMO ; the canonical example of a function that is not bounded but it is in BMO is $\log|x|$ [26].

2.1.3 Symmetries

The Hilbert transform commutes with translations and dilations and anticommutes with reflections, and it is essentially the only bounded linear operator in $L^2(\mathbb{R})$ that has those properties. In what follows $h \in \mathbb{R}$ and $\delta > 0$.

- Convolution $\Leftrightarrow H$ commutes with *translations* $\tau_h f(x) := f(x - h)$

$$\tau_h(Hf) = H(\tau_h f).$$

- Homogeneity of kernel $\Leftrightarrow H$ commutes with dilations $D_\delta f(x) = f(\delta x)$

$$D_\delta(Hf) = H(D_\delta f).$$

- Kernel odd $\Leftrightarrow H$ anticommutes with reflections $\tilde{f}(x) := f(-x)$

$$(H\tilde{f}) = -H(\tilde{f}).$$

Theorem 1 ([26, 73]). *Let T be a linear and bounded operator in $L^2(\mathbb{R})$ that commutes with translations and dilations and anticommutes with reflections, then T must be a constant multiple of the Hilbert transform: $T = cH$.*

Using this principle, Petermichl [66] showed that we can write H as a suitable “average of dyadic operators”; see also [32].

2.2 Dyadic Shift Operators

We first introduce the dyadic intervals and associated Haar basis, as well as random dyadic grids. We recall some important properties of the Haar basis shared with wavelet bases such as being an unconditional system in L^p spaces and weighted $L^p(w)$ whenever $w \in A_p$. We then describe Petermichl’s averaging theorem and give some intuition why this should work. We deduce Riesz’s theorem from this representation, that is, the boundedness on $L^p(\mathbb{R})$ of the Hilbert transform.

2.2.1 Dyadic Intervals

The *standard dyadic grid* \mathcal{D} is the collection of intervals of the form $[k2^{-j}, (k + 1)2^{-j})$, for all integers $k, j \in \mathbb{Z}$. They are organized by generations: $\mathcal{D} = \cup_{j \in \mathbb{Z}} \mathcal{D}_j$, and our labeling is such that $I \in \mathcal{D}_j$ iff $|I| = 2^{-j}$. They satisfy:

- *Trichotomy or nestedness:* $I, J \in \mathcal{D}$ then $I \cap J = \emptyset$, $I \subseteq J$, or $J \subset I$.
- *One parent, two children:* If $I \in \mathcal{D}_j$, then there is a unique interval $\tilde{I} \in \mathcal{D}_{j-1}$ such that $I \subset \tilde{I}$ and $|\tilde{I}| = 2|I|$. There are exactly two disjoint intervals, the right and left children $I_r, I_l \in \mathcal{D}_{j+1}$, such that $I = I_r \cup I_l$ and $|I| = 2|I_r| = 2|I_l|$.

2.2.2 Random Dyadic Grids

A dyadic grid in \mathbb{R} is a collection of intervals, organized in generations, each of them being a partition of \mathbb{R} , that have the trichotomy and two children per interval property. For example, the shifted and rescaled regular dyadic grid will be a dyadic grid. However, these are *not* all possible dyadic grids.

The following parametrization will capture *all* dyadic grids. Consider the *scaling or dilation parameter* r with $1 \leq r < 2$ and the *random parameter* β with $\beta = \{\beta_i\}_{i \in \mathbb{Z}}$, $\beta_i = 0, 1$; let $x_j = \sum_{i < -j} \beta_i 2^i$ and then define

$$\mathcal{D}_j^\beta := x_j + \mathcal{D}_j, \quad \text{and} \quad \mathcal{D}_j^{r,\beta} := r \mathcal{D}_j^\beta.$$

The family of intervals $\mathcal{D}^{r,\beta}$ so defined is a dyadic grid. Here r is a dilation parameter, and β a random parameter that encode all possible dyadic grids. Notice that for the standard dyadic grid zero is never an interior point of a dyadic interval, and it is always on the right side of any dyadic interval it belongs to. If we translate \mathcal{D} by a fixed number it will simply shift zero, and it will still have this singular property. The translated grids correspond to parameters β such that β_j is constant for all sufficiently large j . But these are not all the possible grids. Once we have an interval in a dyadic grid its descendants are completely determined, simply subdivide; however, there are two possible choices for the parent, four possible choices for the grandparent, and 2^n choices for the n th-parent. The parameter β captures all of these possibilities. Those β 's that do not become eventually constant eliminate the presence of a singular point such as zero in the standard grid.

The random dyadic grids were introduced by Nazarov, Treil, and Volberg in their study of Calderón–Zygmund singular integrals on nonhomogeneous spaces [56] and are utilized by Hytönen in his representation theorem [32, 33]. The advantage of this parametrization is that there is a very natural probability space, say (Ω, P) associated to the parameters, and averaging here means calculating the expectation in this probability space, that is, $\mathbb{E}f = \int_\Omega f \, dP$.

2.2.3 Haar Basis

Given an interval I , its associated *Haar function* is defined to be

$$h_I(x) := |I|^{-1/2}(\chi_{I_r}(x) - \chi_I(x)),$$

where $\chi_I(x) = 1$ if $x \in I$, zero otherwise. Note that $\|h_I\|_2 = 1$, and it has zero integral $\int h_I = 0$. One can check, from these integral properties and the nestedness properties of the dyadic intervals, that $\{h_I\}_{I \in \mathcal{D}}$ is an orthonormal system in $L^2(\mathbb{R})$. Furthermore, the system is complete, that is, it is an orthonormal basis in $L^2(\mathbb{R})$.

Alfred Haar introduced in 1910 the Haar basis in $L^2([0, 1])$ and showed that for continuous functions their Haar expansions converge uniformly [28], unlike their expansions in the trigonometric (Fourier) basis [21, 26, 73].

A basis is *unconditional* in $L^p(\mathbb{R})$ if and only if changes in the signs of the coefficients of a function keep it in the same space with comparable norms [82]. The trigonometric system $\{e^{2\pi i n x}\}_{n \in \mathbb{Z}}$ does not form an unconditional basis in $L^p([0, 1])$ for $p \neq 2$ [26, 82]. On the other hand, the Haar basis $\{h_I\}_{I \in \mathcal{D}}$ is an unconditional basis in $L^p(\mathbb{R})$. More precisely we can define an operator, the *martingale transform*,

given by

$$T_\sigma f(x) = \sum_{I \in \mathcal{D}} \sigma_I \langle f, h_I \rangle h_I, \quad \text{where } \sigma_I = \pm 1. \tag{8}$$

Unconditionality of the Haar basis in $L^p(\mathbb{R})$ reduces then to show that the martingale transform is bounded in $L^p(\mathbb{R})$ with norm independent of the choice of signs:

$$\sup_\sigma \|T_\sigma f\|_p \leq C_p \|f\|_p.$$

This was proved by Burkholder who also found the optimal constant C_p [7].

The Haar system $\{h_I\}_{I \in \mathcal{D}}$ is an unconditional basis in $L^p(w)$ if and only if $w \in A_p$. This fact is deduced from the boundedness of the martingale transform on $L^p(w)$ [75]. For sharp linear bounds in $L^2(w)$ for the martingale transform see [79].

The Haar basis is the first example of a wavelet basis, that is, a basis $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$, that is found by translating and dilating appropriately a fixed function ψ , the wavelet, more precisely, $\psi_{j,k}(x) := 2^{-j/2} \psi(2^j x + k)$. The Haar functions are translates and dyadic dilates of the function $h(x) := \chi_{[0,1/2)}(x) - \chi_{[1/2,1)}(x)$. These unconditionality properties are shared by a large class of wavelets [29, 75, 82].

2.2.4 Petermichl’s Dyadic Shift Operator

Petermichl’s dyadic shift operator S associated to the standard dyadic grid \mathcal{D} is defined for function $f \in L^2(\mathbb{R})$ by

$$Sf(x) := \sum_{I \in \mathcal{D}} \langle f, h_I \rangle H_I(x), \quad \text{where } H_I := 2^{-1/2}(h_{I_r} - h_{I_l}).$$

Petermichl’s shift operator is an isometry in $L^2(\mathbb{R})$, that is, it preserves L^2 -norms, $\|Sf\|_2 = \|f\|_2$. Notice that if $I \in \mathcal{D}$, $Sh_I(x) = H_I(x)$. A periodic version of the Hilbert transform that we denote by H_p , has the property that it maps cosines into sines, $H_p \cos(x) = \sin(x)$. Draw the profiles of h_I and H_I and you can view them as a localized sine and cosine. This indicates that this shift operator may be a good dyadic model for the Hilbert transform. More evidence comes from the way it interacts with translations, dilations, and reflections.

Denote by $S_{r,\beta}$ Petermichl’s shift operator associated to the dyadic grid $\mathcal{D}_{r,\beta}$. Each shift operator $S_{r,\beta}$ does not commute with translations and dilations, nor does it anticommute with reflections; however, one can verify that the following symmetries for the family of shift operators $\{S_{r,\beta}\}_{(r,\beta) \in \Omega}$ hold:

- Translation: $\tau_h(S_{r,\beta} f) = S_{\tau_h(r,\beta)}(\tau_h f)$, where $\tau_h(r, \beta) \in \Omega$.
- Dilation: $D_\delta(S_{r,\beta} f) = S_{D_\delta(r,\beta)}(D_\delta f)$, where $D_\delta(r, \beta) \in \Omega$.
- Reflection: $\widetilde{S_{r,\beta} f} = S_{r,\tilde{\beta}}(\tilde{f})$, where $\tilde{\beta}_i = 1 - \beta_i$.

Where the maps $\tau_h, D_\delta : \Omega \rightarrow \Omega$ are bijections. Each shift dyadic operator does not have the symmetries that characterize the Hilbert transform, but the average over all dyadic grids will, therefore,

Theorem 2 (Petermichl’s [32, 66]).

$$\mathbb{E}S_{r,\beta} = \int_{\Omega} S_{r,\beta} dP(r, \beta) = cH.$$

Petermichl’s result then follows once one verifies that $c \neq 0$ (which she did!). Similar trick works for the *Beurling–Ahlfors* [70] and the *Riesz transforms* [68]. Vagharshakyan showed that sufficiently smooth one-dimensional Calderón–Zygmund convolution operators are averages of Haar shift operators of bounded complexity [76].

2.2.5 L^p Boundedness of the Hilbert Transform: A Dyadic Proof

Estimates for the Hilbert transform H follow from uniform estimates for Petermichl’s shift operators.

Lemma 1 (Riesz [17]). *The Hilbert transform is bounded on L^p for $1 < p < \infty$.*

$$\|Hf\|_p \leq C_p \|f\|_p.$$

Proof. Suffices to check that

$$\sup_{(r,\beta) \in \Omega} \|S_{r,\beta} f\|_p \leq C_p \|f\|_p.$$

Case $p = 2$ follows from orthonormality of the Haar basis. First rewrite Petermichl’s shift operator in the following manner, where \tilde{I} is the parent of I in the dyadic grid $\mathcal{D}^{r,\beta}$:

$$S_{r,\beta} f = \sum_{I \in \mathcal{D}_{r,\beta}} \frac{1}{\sqrt{2}} \operatorname{sgn}(I, \tilde{I}) \langle f, h_{\tilde{I}} \rangle h_I,$$

where $\operatorname{sgn}(I, \tilde{I}) = 1$ if I is the right child of \tilde{I} and -1 if I is the left child. We can now use Plancherel to compute the L^2 norm, and noticing that each parent has two children,

$$\|S_{r,\beta} f\|_2^2 = \sum_{I \in \mathcal{D}_{r,\beta}} \frac{|\langle f, h_{\tilde{I}} \rangle|^2}{2} = \|f\|_2^2.$$

Minkowski integral inequality then shows that

$$\|\mathbb{E}S_{r,\beta} f\|_2 \leq \mathbb{E}\|S_{r,\beta} f\|_2 \leq \|f\|_2.$$

Case $p \neq 2$ follows from the unconditionality of the Haar basis on $L^p(\mathbb{R})$. \square

3 Weighted Inequalities and the A_2 Conjecture

In this section we discuss weighted inequalities for the Hilbert transform and recount the prehistory of linear estimates for dyadic operators on $L^2(w)$. We state the sharp extrapolation theorem and deduce $L^p(w)$ bounds from linear bounds and observe that these bounds are sometimes sharp, but not always, as Buckley's estimates for the maximal function show. We then define the Haar shift operators of complexity (m, n) , discuss their boundedness properties, and finally state Hytönen's theorem (A_2 conjecture).

3.1 Boundedness on Weighted L^p

The Hilbert transform is bounded on weighted $L^p(w)$; the celebrated 1973 Hunt–Muckenhoupt–Wheeden theorem says:

Theorem 3 (Hunt–Muckenhoupt–Wheeden [31]).

$$w \in A_p \Leftrightarrow \|Hf\|_{L^p(w)} \leq C_p(w) \|f\|_{L^p(w)}.$$

Dependence of the constant on the A_p characteristic was found 30 years later.

Theorem 4 (Petermichl [67]).

$$\|Hf\|_{L^p(w)} \leq C [w]_{A_p}^{\max\{1, \frac{1}{p-1}\}} \|f\|_{L^p(w)}.$$

Proof (Sketch of the proof). For $p = 2$ suffices to find uniform (on the grids) linear estimates for Petermichl's shift operator (this was the hard part which she did using Bellman functions and a bilinear Carleson embedding theorem due to Nazarov, Treil, and Volberg [55]). For $p \neq 2$ a sharp extrapolation theorem [20] that we will discuss in Sect. 3.1.2 automatically gives the result from the *linear estimate* in $L^2(w)$. \square

3.1.1 Chronology of First Linear Estimates on $L^2(w)$

In 1993, Steve Buckley showed that the maximal function obeys a linear bound in $L^2(w)$ [6]. Starting in 2000, one at a time over a span of 10 years, a handful of dyadic operators or operators with enough symmetries that could be written as averages of dyadic operators were shown to obey a linear bound in $L^2(w)$; see (3):

- *Martingale transform* (Janine Wittwer [79] in 2000)
- *Dyadic square function* (Sanja Hukovic, Treil, Volberg [30], Wittwer [80] in 2000)

- *Beurling transform* (Petermichl, Volberg [70] in 2002)
- *Hilbert transform* (Stephanie Petermichl [67] in 2003, published 2007)
- *Riesz transforms* (Stephanie Petermichl [68] in 2008)
- *Dyadic paraproduct in \mathbb{R}* (Oleksandra Beznosova [4] in 2008)

These estimates were based on Bellman functions and bilinear Carleson estimates by Nazarov, Treil, and Volberg [55]. See [61] for Bellman function extensions of the results for dyadic square functions to homogeneous spaces. See [9] for a neat Bellman function transference lemma that allows to use Bellman functions in \mathbb{R} to deduce results in \mathbb{R}^n with no sweat, similar considerations are used in [74]. There are now simpler Bellman function proofs that recover the estimates for the dyadic shift operators [54,74] and for the dyadic paraproduct [52]. The Bellman function method was introduced in harmonic analysis by Nazarov, Treil, and Volberg, and with their students and collaborators, they have been able to use this method to obtain a number of astonishing results not only in this area; see [77,78] and references.

3.1.2 Estimates in $L^p(w)$ via Sharp Extrapolation

The $L^p(w)$ inequalities can be deduced from the linear bounds on $L^2(w)$, thanks to a sharp version of Rubio de Francia’s extrapolation theorem [25].

Theorem 5 (Sharp Extrapolation Theorem [20]). *If for all $w \in A_r$ there is $\alpha > 0$, and $C > 0$ such that*

$$\|Tf\|_{L^r(w)} \leq C[w]_{A_r}^\alpha \|f\|_{L^r(w)},$$

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

$$\|Tf\|_{L^p(w)} \leq C_{p,r}[w]_{A_p}^{\alpha \max\{1, \frac{r-1}{p-1}\}} \|f\|_{L^p(w)}.$$

Duoandikoetxea found recently a shorter proof of this theorem [22]. Sharp extrapolation from $r = 2$ is sharp for the martingale, Hilbert, Beurling–Ahlfors, and Riesz transforms for all $1 < p < \infty$ [20]. Therefore the theorem cannot be improved in terms of the power on the A_p characteristic of the weight. However, it is not necessarily sharp for each individual operator. The theorem is sharp for the dyadic square function and $1 < p \leq 2$, see [20], but it is not sharp for $p > 2$, see [46]. The optimal power for the square function is $\max\{\frac{1}{2}, \frac{1}{p-1}\}$ (see [18]), which corresponds to sharp extrapolation starting at $r = 3$ with square root power instead of starting at $r = 2$ with linear power; see also [50]. We conclude that *sharp extrapolation is not always sharp*. Buckley’s estimates for the maximal function are a more dramatic example of the above statement.

Remember the Hardy–Littlewood maximal function is defined as

$$Mf(x) = \sup_{I \ni x} \frac{1}{|I|} \int_I |f(y)| \, dy.$$

The maximal function is known to be bounded on $L^p(\mathbb{R})$ for $1 < p$; it is not bounded on $L^1(\mathbb{R})$, but it is of weak type $(1, 1)$ [21, 26, 73]. Muckenhoupt showed in 1972 [53] that the maximal function is bounded on $L^p(w)$ if and only if $w \in A_p$. The optimal dependence on the A_p characteristic of the weight was discovered by Buckley 20 years later.

Theorem 6 (Buckley [6]). *Let $w \in A_p$ and $1 < p$, then*

$$\|Mf\|_{L^p(w)} \leq C_p [w]_{A_p}^{\frac{1}{p-1}} \|f\|_{L^p(w)}.$$

This estimate is key in the proof of the sharp extrapolation theorem. Observe that if we start with Buckley’s estimate on $L^r(w)$, then sharp extrapolation will give the right power for all $1 < p \leq r$; however, for $p > r$, it will simply give $\frac{1}{r-1}$ which is bigger than the correct power $\frac{1}{p-1}$.

3.1.3 Estimates for Larger Classes of Operators

Petermichl’s shift operator and the martingale transform are the simplest among a larger class of Haar shift operators that we now define.

A Haar shift operator of complexity (m, n) , $S_{m,n}$, is defined as follows:

$$S_{m,n}f(x) := \sum_{L \in \mathcal{D}} \sum_{I \in \mathcal{D}_m(L), J \in \mathcal{D}_n(L)} c_{I,J}^L \langle f, h_I \rangle h_J(x), \tag{9}$$

where the coefficients $|c_{I,J}^L| \leq \frac{\sqrt{|I||J|}}{|L|}$ and $\mathcal{D}_m(L)$ denote the dyadic subintervals of L with length $2^{-m}|L|$.

The normalization of the coefficients ensures that $\|S_{m,n}f\|_2 \leq \|f\|_2$. The reader can now check that the martingale transform is a Haar shift operator of complexity $(0, 0)$ and Petermichl’s shift operator is a Haar shift operator of complexity $(0, 1)$. However, the dyadic paraproduct π_b , which is defined for a function $b \in BMO$ as

$$\pi_b f(x) := \sum_{I \in \mathcal{D}} m_I f \langle b, h_I \rangle h_I(x), \quad \text{where } m_I f = \frac{1}{|I|} \int_I f(x) dx,$$

is not a Haar shift operator. The Haar shift operators were introduced in [44] and used in [17, 18]. Later, a larger class, the generalized dyadic shift operators, that included the paraproducts was defined [33, 37], where the Haar functions in (9) were replaced by $|I|^{-1/2} \chi_I(x)$ and boundedness on $L^2(\mathbb{R})$ is now part of the definition since it will not follow from the normalization of the coefficients. In this setting the dyadic paraproduct, the martingale transform, and Petermichl’s Haar shift operator are generalized dyadic shift operators of complexity $(0, 1)$, $(1, 1)$, and $(1, 2)$, respectively. The adjoint of the dyadic paraproduct, defined by

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

is a generalized dyadic shift operator of complexity $(1, 0)$, and the composition $\pi_b^* \pi_b$ is of complexity $(0, 0)$. On the other hand the composition $\pi_b \pi_b^*$ is not a generalized dyadic shift operator; localization has been lost.

The following authors either extend to other settings or recover most of the previous known results (the linear bounds on $L^2(w)$) and can extend them to the larger class of Haar shift operators, and in particular averaging appropriately, they can get Hilbert, Riesz, and Beurling–Ahlfors transforms:

- Lacey, Moen, Pérez, and Torres [43] obtain sharp bound on weighted L^p spaces for fractional integral operators.
- Lacey, Petermichl, and Reguera [44] use a *corona decomposition* and a *two-weight theorem for “well-localized operators”* of Nazarov, Treil, and Volberg, to recover linear bounds for Haar shifts operators on $L^2(w)$; they do not use Bellman functions. Dependence on the complexity is exponential. This result does not include dyadic paraproducts.
- Cruz-Uribe, Martell, and Pérez [17, 18] recover all results for Haar shift operators. No Bellman functions, no two-weight results. Instead they use a local median oscillation introduced by Lerner [47, 48]. The method is very flexible, they can get new results such as the sharp bounds for the square function for $p > 2$, they can recover also the result for the dyadic paraproduct, they can get results for vector-valued maximal operators and two-weight results as well. Dependence on complexity is exponential.

After these results were posted a lot of activity followed and results covering larger classes of operators appeared:

- Lerner [48, 50] showed that all standard convolution-type operators in arbitrary dimension gave the expected result for $p \in (1, 3/2] \cap [3, \infty)$. He also showed sharp estimates on $L^p(w)$ for all $p > 1$ and for all sort of square functions. This is based on controlling them with Wilson’s intrinsic square function [81].
- Hytönen, Lacey, Reguera, Sawyer, Uriarte-Tuero, and Vagharshakyan posted a preprint in 2010 which was then replaced by a 2011 preprint with more authors [37]. They obtain the desired result for a general class of Calderón–Zygmund non-convolution operators, still requiring smoothness of the kernels.
- Pérez, Treil, and Volberg [65] showed that all Calderón–Zygmund operators obey an almost linear estimate on $L^2(w)$: $[w]_{A_2} \log(1 + [w]_{A_2})$. They identified the obstacle that would remove the log term.

3.1.4 The A_2 Conjecture (Now Theorem)

The A_2 conjecture said that all Calderón–Zygmund singular integral operators should obey a linear bound on $L^2(w)$. This was finally proved by Tuomas Hytönen in 2010.

Theorem 7 (Hytönen [33]). *Let $1 < p < \infty$ and let T be any Calderón–Zygmund singular integral operator in \mathbb{R}^n ; then there is a constant $c_{T,n,p} > 0$ such that*

$$\|Tf\|_{L^p(w)} \leq c_{T,n,p} [w]_{A_p}^{\max\{1, \frac{1}{p-1}\}} \|f\|_{L^p(w)}.$$

It is enough to consider the case $p = 2$ thanks to sharp extrapolation. Hytönen proves the representation theorem, gets linear estimates on $L^2(w)$ with respect to the A_2 characteristic for Haar shift operators, and gets polynomial dependence in the complexity. Together these imply the theorem for $p = 2$. We consider the representation theorem to be of independent interest, and we state it here.

Theorem 8 (Hytönen [33]). *Let T be a Calderón–Zygmund singular integral operator; then*

$$Tf = \mathbb{E} \left(\sum_{(m,n) \in \mathbb{N}^2} a_{m,n} S_{m,n}^{r,\beta} f \right),$$

where the coefficients in the series are of the form $a_{m,n} = e^{-(m+n)\alpha/2}$, α is the smoothness parameter of T , and $S_{m,n}^{r,\beta}$ are Haar shift operators of complexity (m, n) when $(m, n) \neq (0, 0)$, and when $(m, n) = (0, 0)$ they are a linear combination of a Haar shift of complexity $(0, 0)$, a dyadic paraproduct, and the adjoint of the dyadic paraproduct, all based on the dyadic grid $\mathcal{D}_{r,\beta}$, and \mathbb{E} is the expectation in the probability space (Ω, P) associated to the random dyadic grids $\mathcal{D}_{r,\beta}$.

Leading to the solution of the A_2 conjecture were the results of Pérez, Treil, and Volberg [65]. Since the appearance of Hytönen’s theorem several simplifications of the argument have appeared [34, 38, 42, 54, 74], as well as an extension to metric spaces with geometric doubling condition [58]. There is also a very nice survey of the A_2 conjecture [41].

Can we expect more singular operators to have worst estimates? Yes, for example, the commutators of $b \in BMO$ with T a Calderón–Zygmund singular integral operator.

4 Sharp Weighted Inequalities for the Commutator

In this section we define the commutator, state its boundedness properties, and sketch Chung’s dyadic proof of the quadratic estimate on $L^2(w)$. We note that this quadratic estimate is sharp, and we show that Chung’s dyadic method of proof implies that Beznosova’s bounds for the dyadic paraproduct are sharp as well.

Finally we state a variation of the transference theorem for commutators on $L^r(w)$ with $r \neq 2$ and present its proof in Appendix.

4.1 The Commutator

The commutator $[b, H]$ of $b \in BMO$ and H the Hilbert transform is defined:

$$[b, H]f = b(Hf) - H(bf).$$

It is well known that the commutator $[b, H]$ is bounded on $L^p(\mathbb{R})$.

Theorem 9 (Coifman et al. [12]). *Let $b \in BMO$ and $1 < p < \infty$, then*

$$\|[H, b]f\|_p \leq C_p \|b\|_{BMO} \|f\|_p.$$

However, the commutator is not of weak type $(1, 1)$ as Carlos Pérez showed [62]. The commutator $[b, H]$ is more singular than H . Another way to quantify this roughness is to observe that the maximal function M controls H ; however, to control the commutator we need M^2 [63].

Observe that separately bH and Hb are *not* bounded on $L^p(\mathbb{R})$ when $b \in BMO$, simply because multiplication by a BMO function does not preserve $L^p(\mathbb{R})$ (one needs the multiplier to be bounded and $L^\infty(\mathbb{R}) \subsetneq BMO$). *The commutator introduces some key cancellation.* This is very much connected to the celebrated $H^1 - BMO$ duality by Fefferman and Stein [23] (H^1 denotes the Hardy space on the line).

Coifman, Rochberg, and Weiss have a beautiful argument in [12] to prove boundedness on $L^p(\mathbb{R})$ of the commutator based on the boundedness of the Hilbert transform on $L^p(v)$ for $v \in A_2$; it is this argument that was exploited to obtain the following weighted inequalities for the commutator in quite a general framework; here we state the estimate for the Hilbert transform.

Theorem 10 (Alvarez et al. [1]). *If $w \in A_p$ and $b \in BMO$, then*

$$\|[H, b]f\|_{L^p(w)} \leq C_p(w) \|b\|_{BMO} \|f\|_{L^p(w)}.$$

4.2 Chung’s Dyadic Argument

Daewon Chung proved the following sharp bound on $L^p(w)$ for the commutator of the Hilbert transform and a BMO function:

Theorem 11 (Chung [8]).

$$\|[H, b]f\|_{L^p(w)} \leq C_p \|b\|_{BMO[w]_{A_p}}^{2\max\{1, \frac{1}{p-1}\}} \|f\|_{L^p(w)}.$$

The result is sharp in $L^2(w)$, meaning that in that case the quadratic power cannot be improved. Similar examples show extrapolated bounds are sharp in $L^p(w)$; see [8].

Chung’s proof is based on a decomposition of the product $b f$ using the dyadic paraproduct $\pi_b f$, its adjoint $\pi_b^* f$, and a related operator $\pi_f b$; this line of argument was suggested in [69]. He works with Petermichl’s dyadic shift operator S instead of H , and Bellman functions. This argument works for dyadic shift operators (hence for Riesz and Beurling transforms, and it is sharp for them as well). We will sketch Chung’s proof after some preliminaries on paraproducts.

4.2.1 Dyadic Paraproduct

Recall that dyadic paraproduct associated to the function $b \in BMO$ is defined by

$$\pi_b f(x) = \sum_{I \in \mathcal{D}} m_I f \langle b, h_I \rangle h_I(x), \quad \text{where } m_I f = \frac{1}{|I|} \int_I f(x) dx.$$

The dyadic paraproduct is bounded on $L^p(\mathbb{R})$ for $1 < p < \infty$ and is of weak type $(1, 1)$ [60]. Paraproducts appeared in the work of Bony [5] on paradifferential equations; they also appeared in the proof of the $T(1)$ theorem [19].

Theorem 12 (Beznosova [4]). *Let $b \in BMO$, $w \in A_2$, then for all $f \in L^2(w)$*

$$\|\pi_b f\|_{L^2(w)} + \|\pi_b^* f\|_{L^2(w)} \leq C \|b\|_{BMO[w]_{A_2}} \|f\|_{L^2(w)}.$$

Ordinary multiplication $M_b f = b f$ is not bounded on $L^p(\mathbb{R})$ unless $b \in L^\infty(\mathbb{R})$. The space BMO includes unbounded functions. Hence the boundedness properties of the paraproduct are better than those of the ordinary product. It is well known that the following decomposition holds:

$$b f = \pi_b f + \pi_b^* f + \pi_f b. \tag{10}$$

The first two terms are not only bounded on $L^p(\mathbb{R})$ but are also bounded on $L^p(w)$ (follows by extrapolation from boundedness on $L^2(w)$) when $b \in BMO$ and $w \in A_p$; the enemy in this decomposition is the third term $\pi_f b$. It is because of this relation with the ordinary product that the name “paraproduct” was coined.

Proof (Sketch of Chung’s proof of Theorem 11). Apply the decomposition (10) to the commutator with Petermichl’s shift operator S :

$$[S, b]f = [S, \pi_b]f + [S, \pi_b^*]f + [S(\pi_f b) - \pi_{Sf}(b)]. \tag{11}$$

The first two terms give quadratic bounds from the linear bounds for S , π_b , and π_b^* . Boundedness of the commutator on $L^p(w)$ will be recovered from the uniform boundedness of the third commutator. Surprisingly (at the time this was discovered) the third term is better; it obeys a linear bound, and so do halves of the other two commutators:

$$\begin{aligned} & \|S(\pi_f b) - \pi_{Sf}(b)\|_{L^2(w)} + \|S\pi_b f\|_{L^2(w)} + \|\pi_b^* S\|_{L^2(w)} \\ & \leq C \|b\|_{BMO[w]_{A_2}} \|f\|_{L^2(w)}. \end{aligned}$$

Providing uniform *quadratic bounds* for the commutator $[S, b]$, hence

$$\|[H, b]\|_{L^2(w)} \leq C \|b\|_{BMO[w]_{A_2}^2} \|f\|_{L^2(w)}.$$

□

Chung proved his linear estimates using Bellman functions. A posteriori one realizes that the operators $[S(\pi_{\{\cdot\}} b) - \pi_{S\{\cdot\}}(b)]$, $S\pi_b$, and $\pi_b^* S$ are generalized Haar shift operators; hence, the linear bound is a particular case of the results in [33, 34, 37, 38]. For the commutator the bad terms are the nonlocal operators $\pi_b S$ and $S\pi_b^*$.

4.2.2 Commutators Versus Paraproducts

Beznosova proved the linear bound for the dyadic paraproduct, and then sharp extrapolation shows that the following bounds hold in $L^p(w)$ for $w \in A_p$:

$$\|\pi_b f\|_{L^p(w)} \leq C_p \|b\|_{BMO[w]_{A_p}^{\max\{1, \frac{1}{p-1}\}}} \|f\|_{L^p(w)}.$$

It was not known whether these were sharp for some or all $1 < p < \infty$.

Theorem 13. *The above estimate is optimal in the power $\max\{1, \frac{1}{p-1}\}$.*

Proof. Suppose there is an $\alpha < 1$ and a $p > 1$ such that for all $b \in BMO$ weights $w \in A_p$ and for all $f \in L^p(w)$ the following estimate holds:

$$\|\pi_b f\|_{L^p(w)} \leq C_p \|b\|_{BMO[w]_{A_p}^{\alpha \max\{1, \frac{1}{p-1}\}}} \|f\|_{L^p(w)}.$$

One can verify that the same estimate holds for π_b^* . Then we will obtain the following bound for the commutator of the Hilbert transform and b :

$$\|[b, H]f\|_{L^p(w)} \leq C_p \|b\|_{BMO[w]_{A_p}^{(1+\alpha) \max\{1, \frac{1}{p-1}\}}} \|f\|_{L^p(w)}.$$

And this is a contradiction because the power $2 \max\{1, \frac{1}{p-1}\}$ is optimal for $[b, H]$.

□

4.3 Transference Theorem in $L^r(w)$ for Commutators

The following transference theorem holds:

Theorem 14 (Chung et al. [10]). *If a linear operator T obeys linear bounds in $L^2(w)$ for all $w \in A_2$*

$$\|Tf\|_{L^2(w)} \leq C[w]_{A_2} \|f\|_{L^2(w)},$$

then its commutator with $b \in BMO$ obeys quadratic bounds for all $w \in A_2$,

$$\|[T, b]f\|_{L^2(w)} \leq C[w]_{A_2}^2 \|b\|_{BMO} \|f\|_{L^2(w)}. \tag{12}$$

Proof follows the beautiful Coifman–Rochberg–Weiss classical argument using the Cauchy integral formula and immediately generalizes to higher-order commutators $T_b^k := [b, T_b^{k-1}]$. Under the same assumptions of Theorem 14,

$$\|T_b^k f\|_{L^2(w)} \leq Ck! [w]_{A_2}^{k+1} \|b\|_{BMO}^k \|f\|_{L^2(w)}. \tag{13}$$

Extrapolation gives bounds on $L^p(w)$; they are sharp for all $1 < p < \infty$, all $k \geq 1$, and all dimensions, as examples involving the Riesz transforms show [10].

As a corollary of these and Hytönen’s theorem we conclude that for each Calderón–Zygmund singular integral operators T there is a constant $C > 0$ such that for all BMO functions b and for all A_2 weights w , (12) holds. Sharp extrapolation then shows that for all Calderón–Zygmund singular operators T ,

$$\|[T, b]f\|_{L^p(w)} \leq C_p [w]_{A_p}^{2 \max\{1, \frac{1}{p-1}\}} \|b\|_{BMO} \|f\|_{L^p(w)}. \tag{14}$$

A refinement of the argument in [10] shows that

Theorem 15. *If a linear operator T obeys a power bound in $L^r(w)$ for all $w \in A_r$,*

$$\|Tf\|_{L^r(w)} \leq C[w]_{A_r}^\alpha \|f\|_{L^r(w)},$$

then its commutator with $b \in BMO$ obeys the following bounds for all $w \in A_r$:

$$\|[T, b]f\|_{L^r(w)} \leq C_{n,r} [w]_{A_r}^{\alpha + \max\{1, \frac{1}{r-1}\}} \|b\|_{BMO} \|f\|_{L^r(w)}.$$

Notice that in the case of T a Calderón–Zygmund singular integral operator, we recover the $L^p(w)$ norm obtained from sharp extrapolation in [10], because the initial estimate on $L^p(w)$ corresponds to $\alpha = \max\{1, \frac{1}{p-1}\}$; hence in this case Theorem 15 gives (14). Because this bound is known to be sharp for the Hilbert and Riesz transforms, we deduce that the power obtained in Theorem 15 cannot be improved.

We present the proof of this result in the Appendix.

Generalizations and variations of these results have already appeared. Cruz-Uribe and Moen [16] prove corresponding estimates for commutators with fractional integrals (they also use the classical Coifman–Rochberg–Weiss argument). They use the machinery developed by Cruz-Uribe, Martell, and Pérez [18] and Lerner’s local mean oscillation [48] to obtain two-weight estimates for the commutators with Calderón–Zygmund singular integral operators and fractional integrals. Carmen Ortiz-Caraballo [59] shows the following quadratic estimate for $b \in BMO$, and any Calderón–Zygmund operator T , on $L^p(w)$ where the weight is in $A_1 \subset \cap_{p>1} A_p$, the following estimate was obtained before Hytönen proved the A_2 conjecture, so it was the first nontrivial bound valid for all commutators of Calderón–Zygmund singular integral operators:

$$\|[T, b]\|_{L^p(w)} \leq C_n \|b\|_{BMO} p (p')^2 [w]_{A_1}^2.$$

There are now mixed A_p - A_∞ estimates that hold for all Calderón–Zygmund singular integral operators [34–36, 49]; inequality (15) is an example of such an estimate when $p = 2$. These estimates can be transferred to the commutators [36].

Theorem 16 (Hytönen and Pérez [36]). *If a linear operator T obeys the following bounds in $L^2(w)$ for all $w \in A_2$:*

$$\|Tf\|_{L^2(w)} \leq C [w]_{A_2}^{\frac{1}{2}} ([w]_{A_\infty} + [w^{-1}]_{A_\infty})^{\frac{1}{2}} \|f\|_{L^2(w)}, \tag{15}$$

then its commutator of order $k \geq 1$ with $b \in BMO$ obeys the following bounds for all $w \in A_2$:

$$\|T_b^k f\|_{L^2(w)} \leq C [w]_{A_2}^{\frac{1}{2}} ([w]_{A_\infty} + [w^{-1}]_{A_\infty})^{k+\frac{1}{2}} \|b\|_{BMO} \|f\|_{L^2(w)}.$$

The two-weight problem is still an outstanding open problem for most operators. Necessary and sufficient conditions are known for the maximal function via Sawyer-type conditions [51, 72], for the martingale transform and other dyadic operators [55] (these are of Sawyer type as well with respect to the dyadic operators), and for the dyadic square function (Beznosova, O., personal communication); compare to [81]. As for sufficient conditions many different sets are known, including several sets for the Hilbert transform [15, 40, 45, 57]. In all these cases the conditions are somehow inherent to the operator studied: “Sawyer-type conditions.” An exception being sufficient conditions in terms of “bump conditions” in Orlicz spaces [16, 18]. Lacking are theorems of the nature; operator A is bounded from $L^p(u)$ into $L^p(v)$ if and only if operator B is bounded from $L^p(u)$ into $L^p(v)$.

Appendix

Proof (Sketch the proof of Theorem 15).

We “conjugate” the operator as follows: if z is any complex number we define

$$T_z(f) = e^{zb} T(e^{-zb} f).$$

Then, a computation gives (for instance for “nice” functions)

$$[b, T](f) = \frac{d}{dz} T_z(f)|_{z=0} = \frac{1}{2\pi i} \int_{|z|=\epsilon} \frac{T_z(f)}{z^2} dz, \quad \epsilon > 0$$

by the Cauchy integral theorem; see [1, 12].

Now, by Minkowski’s inequality,

$$\|[b, T](f)\|_{L^r(w)} \leq \frac{1}{2\pi \epsilon^2} \int_{|z|=\epsilon} \|T_z(f)\|_{L^r(w)} |dz|, \quad \epsilon > 0.$$

The key point is to find the appropriate radius ϵ . First we look at the inner norm,

$$\|T_z(f)\|_{L^r(w)} = \|T(e^{-zb} f)\|_{L^r(w e^{rRezb})},$$

and try to find appropriate bounds on z . To do this we use the main hypothesis, namely that T is bounded on $L^r(v)$ if $v \in A_r$ with

$$\|T\|_{L^r(v)} \leq C[v]_{A_r}^\alpha.$$

Let $v = w e^{rRezb}$. We must check that if $w \in A_r$ then $v \in A_r$ for $|z|$ sufficiently small:

$$[v]_{A_r} = \sup_Q \left(\frac{1}{|Q|} \int_Q w e^{rRezb(x)} dx \right) \left(\frac{1}{|Q|} \int_Q w^{-\frac{1}{r-1}}(x) e^{-\frac{r}{r-1} Rezb(x)} dx \right)^{r-1}.$$

Now, since $w \in A_r$, then $w \in RH_q$ for some $q > 1$ [11]. Recall that $w \in RH_q$ if and only if there is a constant $C > 0$ such that for all cubes Q ,

$$\left(\frac{1}{|Q|} \int_Q w^q dx \right)^{\frac{1}{q}} \leq \frac{C}{|Q|} \int_Q w.$$

The following precise reverse Hölder condition for A_r weights holds [64]:

Lemma 2. *If $w \in A_r$ and $q = 1 + \frac{1}{2^{2r+n+1}[w]_{A_r}} (< 2)$, then $w \in RH_q$ and*

$$\left(\frac{1}{|Q|} \int_Q w^q dx \right)^{\frac{1}{q}} \leq \frac{2}{|Q|} \int_Q w. \tag{16}$$

It is well known that if $w \in A_r$ then $\sigma := w^{-\frac{1}{r-1}} \in A_{r'}$ with $r' = \frac{r}{r-1}$ the dual exponent of r , and $[w]_{A_r}^{r'} = [\sigma]_{A_{r'}}^r$. Applying Lemma 2 to σ and r' we conclude then that if $s = 1 + \frac{1}{2^{2r'+n+1}[\sigma]_{A_{r'}}} < 2$ then $\sigma \in RH_s$ and

$$\left(\frac{1}{|Q|} \int_Q \sigma^s dx \right)^{\frac{1}{s}} \leq \frac{2}{|Q|} \int_Q \sigma. \tag{17}$$

Let $t = \min\{q, s\}$, where q and s are as above, then $t \leq q$ and $t \leq s$. Holder's inequality with $p = q/t > 1$ and $p = s/t > 1$, respectively, implies that

$$\left(\frac{1}{|Q|} \int_Q w^t dx \right)^{\frac{1}{t}} \leq \frac{2}{|Q|} \int_Q w, \quad \left(\frac{1}{|Q|} \int_Q \sigma^t dx \right)^{\frac{1}{t}} \leq \frac{2}{|Q|} \int_Q \sigma.$$

Using these and Holder's inequality twice with $p = t$, we have for an arbitrary Q

$$\left(\frac{1}{|Q|} \int_Q w(x)e^{rRezb(x)} dx \right) \left(\frac{1}{|Q|} \int_Q \sigma(x)e^{-r'Rezb(x)} dx \right)^{r-1} \leq 4[w]_{A_r} [e^{t'rRezb}]_{A_r}^{\frac{1}{t}}.$$

Now, since $b \in BMO$, it is well known that $e^{\eta b} \in A_r$ for η small enough [21,27]. We need a quantitative version of this result.

Lemma 3. *Given $b \in BMO$ then there are $0 < \alpha_n < 1$ and $\beta_n > 1$ such that if $\eta \leq \min\{1, r - 1\} \frac{\alpha_n}{\|b\|_{BMO}}$, then $[e^{\eta b}]_{A_r} \leq \beta_n^r$.*

This follows from a similar computation to the one done for $r = 2$ in [10]. In our case, we need to ensure that $|t'rRezb| \leq \min\{1, r - 1\} \frac{\alpha_n}{\|b\|_{BMO}}$ to deduce that $[e^{t'rRezb}]_{A_r} \leq \beta_n^r$. That is, we are constrained to consider complex numbers z such that $|Rezb| \leq \min\{\frac{1}{r}, \frac{r-1}{r}\} \frac{\alpha_n}{t'\|b\|_{BMO}}$.

Recall that $t = \min\{1 + \frac{1}{2^{2r+n+1}[w]_{A_r}}, 1 + \frac{1}{2^{2r'+n+1}[w]_{A_r}^{\frac{1}{r-1}}}\}$; a calculation now shows that

$$t' = \begin{cases} 1 + 2^{2r+n+1}[w]_{A_r} & p \geq 2 \\ 1 + 2^{2r'+n+1}[w]_{A_r}^{\frac{1}{r-1}} & p < 2 \end{cases}.$$

Furthermore $t' \sim [w]_{A_r}^{\max\{1, \frac{1}{r-1}\}}$ with comparability constant depending exponentially in the dimension n and $\max\{r, r'\}$.

For $|z| \leq \epsilon$, with $\epsilon^{-1} \sim \|b\|_{BMO} [w]_{A_r}^{\max\{1, \frac{1}{r-1}\}}$, and since $1 < t < 2$, thus $t' > 2$, we have that

$$[v]_{A_r} = [we^{rRezb}]_{A_r} \leq 4[w]_{A_r} \beta_n^{\frac{r}{t'}} \leq 4[w]_{A_r} \beta_n^{r/2}.$$

Observe that $\|e^{-zb} f\|_{L^r(w)} = \|e^{-zb} f\|_{L^r(we^{rRez})} = \|f\|_{L^r(w)}$, and if $|z| \leq \frac{\alpha_n}{rt' \|b\|_{BMO}}$,

$$\|T_z(f)\|_{L^r(w)} = \|T(e^{-zb} f)\|_{L^r(w)} \leq [v]_{A_r}^\alpha \|f\|_{L^r(w)} \leq 4[w]_{A_r}^\alpha \beta_n^{r/2} \|f\|_{L^r(w)}.$$

Choose the radius $\epsilon = \frac{\alpha_n}{rt' \|b\|_{BMO}}$, and we can continue estimating the norm of the commutator

$$\|[b, T](f)\|_{L^r(w)} \leq \frac{1}{2\pi \epsilon^2} \int_{|z|=\epsilon} 4[w]_{A_r}^\alpha \beta_n^{r/2} \|f\|_{L^r(w)} |dz| = \frac{1}{\epsilon} 4[w]_{A_r}^\alpha \beta_n^{r/2} \|f\|_{L^r(w)}.$$

Finally, observe that ϵ^{-1} is essentially $[w]_{A_r}^{\max\{1, \frac{1}{r-1}\}}$ $\|b\|_{BMO}$, so we conclude that

$$\|[b, T](f)\|_{L^r(w)} \leq C_{n,r} [w]_{A_r}^{\alpha + \max\{1, \frac{1}{r-1}\}} \|b\|_{BMO}.$$

This finishes the proof of the theorem. □

Acknowledgments The author would like to thank the organizers of the February Fourier Talks, at The Norbert Wiener Center for Harmonic Analysis and Applications, University of Maryland, for inviting her to deliver a talk in the fifth edition of the FFTs on February 18–19, 2010, that was the seed of this chapter. The author dedicates this chapter to the memory of her friend and mentor Cora Sadosky [1940–2010].

References

1. Alvarez, J., Bagby, R., Kurtz, D., Perez, C.: Weighted estimates for commutators of linear operators. *Studia Mat.* **104**(2), 195–209 (1994)
2. Astala, K., Iwaniec, T., Saksman, E.: Beltrami operators in the plane. *Duke Math. J.* **107**(1), 27–56 (2001)
3. Beylkin, G., Coifman, R., Rokhlyn, V., Fast Wavelet Transforms and Numerical Algorithms I. *Comm. Pure Appl. Math.* **44**(2), 141–183 (1991)
4. Beznosova, O., Linear bound for dyadic paraproduct on weighted Lebesgue space $L^2(w)$. *J. Func. Anal.* **255**(4), 994–1007 (2008)
5. Bony, J.: Calcul symbolique et propagation des singularités pour les equations aux dérivées non-linéaires. *Ann. Sci. Ecole Norm. Sup.* **14**, 209–246 (1981)
6. Buckley, S.: Estimates for operator norms and reverse Jensen’s inequalities. *Trans. Amer. Math. Soc.* **340**(1), 253–272 (1993)
7. Burkholder, D.: Boundary value problems and sharp inequalities for martingale transforms. *Ann. Probab.* **12**(3), 647–702 (1984)
8. Chung, D.: Sharp estimates for the commutators of the Hilbert, Riesz and Beurling transforms on weighted Lebesgue spaces. To appear in *Indiana U. Math. J.* **60**(5), (2011)
9. Chung, D.: Weighted inequalities for multivariable dyadic paraproducts. *Pub. Mat.* **55**(2), 475–499 (2011)
10. Chung, D., Pereyra, M.C., Pérez, C.: Sharp bounds for general commutators on weighted Lebesgue spaces. *Trans. Amer. Math. Soc.* **364**(3), 1163–1177 (2012)
11. Coifman, R., Fefferman, C.: Weighted norm inequalities for maximal fuctions and singular integrals. *Studia Math.* **51**, 241–250 (1974)

12. Coifman, R., Rochberg, R., Weiss, G.: Factorization theorems for Hardy spaces in several variables. *Ann. Math.* **103**, 611–635 (1976)
13. Cotlar, M., Sadosky, C.: On the Helson–Szegő theorem and a related class of modified Toeplitz kernels. Harmonic analysis in Euclidean spaces (Proceedings of Symposium on Pure Mathematics, Williams College, Williamstown, MA, 1978), Part 1, pp. 383–407. Proceedings of Symposium on Pure Mathematics, vol. XXXV, Part, American Mathematical Society, Providence, RI (1979)
14. Cotlar, M., Sadosky, C.: On some L^p versions of the Helson–Szegő theorem. In: Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II (Chicago, IL, 1981), pp. 306–317. Wadsworth Mathematical Series, Wadsworth, Belmont, CA (1983)
15. Cruz-Uribe, D., Pérez, C.: On the two-weight problem for singular integral operators. *Ann. Scuola Norm. Super. Pisa Cl. Sci. (5)* **I**, 821–849 (2002)
16. Cruz-Uribe, D., SFO, Moen, K., Sharp norm inequalities for commutators of classical operators. *Pub. Mat.* **56**(1), 147–190 (2012)
17. Cruz-Uribe, D., SFO, Martell, J.M., Pérez, C.: Sharp weighted estimates for approximating dyadic operators. *Electron. Res. Announc. Math. Sci.* **17**, 12–19 (2010)
18. Cruz-Uribe, D., SFO, Martell, J., Pérez, C.: Sharp weighted estimates for classical operators. *Adv. Math.* **229**(1), 408–441 (2012)
19. David, G., Journé, J.-L.: A boundedness criterion for generalized Calderón–Zygmund operators. *Ann. Math.* **120**, 371–397 (1984)
20. Dragičević, O., Grafakos, L., Pereyra, M.C., Petermichl, S.: Extrapolation and sharp norm estimates for classical operators on weighted Lebesgue spaces. *Publ. Math* **49**(1), 73–91 (2005)
21. Duoandikoetxea, J., Fourier Analysis. Graduate Studies in Mathematics, vol. 29. American Mathematical Society, Providence (2001)
22. Duoandikoetxea, J.: Extrapolation of weights revisited: new proofs and sharp bounds. *J. Funct. Anal.* **260**(6), 1886–1901 (2011)
23. Fefferman, C., Stein, E.: H^p spaces of several variables. *Acta Math.* **129**, 137–193 (1972)
24. Figiel, T.: Singular integral operators: a martingale approach. In: Geometry of Banach Spaces (Strbl, 1989), London Mathematical Society, vol. 158. Lecture Notes Series, pp. 95–110. Cambridge University Press, Cambridge (1990)
25. Garcia-Cuerva, J., Rubio de Francia, J.L.: Weighted norm inequalities and related topics. North-Holland Mathematics Studies, vol. 116. North-Holland, Amsterdam, Holland (1981)
26. Grafakos, L.: Classical Fourier Analysis. 2nd edn. Graduate Texts in Mathematics, vol. 249. Springer, New York (2008)
27. Grafakos, L.: Modern Fourier Analysis. Graduate Texts in Mathematics, vol. 250, 2nd edn. Springer, New York (2009)
28. Haar, A.: Zur Theorie der orthogonalen Funktionen systeme. *Math. Ann.* **69**, 331–371 (1910)
29. Hernandez, E., Weiss, G.: A First Course on Wavelets. 3rd edn. Academic, New York (2008)
30. Hukovic, S., Treil, S., Volberg, A.: The Bellman function and sharp weighted inequalities for square functions. In: Complex Analysis, Operators and Related Topics. Operator Theory: Advances and Applications, vol. 113, pp. 97–113. Birkhäuser, Basel (2000)
31. Hunt, R., Muckenhoupt, B., Wheeden, R.: Weighted norm inequalities for the conjugate function and the Hilbert transform. *Trans. Amer. Math. Soc.* **176**, 227–252 (1973)
32. Hytönen, T.P.: On Petermichl’s dyadic shift and the Hilbert transform. *C. R. Acad. de Sci. Paris, Ser. I* **346**(21–22), 1133–1136 (2008)
33. Hytönen, T.P.: The sharp weighted bound for general Calderón–Zygmund operators. *Ann. Math. (2)* **175**(3), 1473–1506 (2012)
34. Hytönen, T.P.: Representation of singular integrals by dyadic operators, and the A_2 theorem. Preprint (2011) available at arXiv:1108.5119
35. Hytönen, T.P., Lacey, M.T.: The $A_p - A_\infty$ inequality for general Calderón–Zygmund operators. Preprint (2011) available at arXiv:1106.4797
36. Hytönen, T.P., Pérez, C.: Sharp weighted bounds involving A_∞ . To appear Journal of Analysis and P.D.E., available at arXiv:1103.5562

37. Hytönen, T.P., Lacey, M.T., Martikainen, H., Orponen, T., Reguera, M.C., Sawyer, E.T., Uriarte-Tuero, I.: Weak and strong type estimates for maximal truncations of Calderón–Zygmund operators on A_p weighted spaces. Preprint (2011) available at arXiv:1103.5229
38. Hytönen, T.P., Pérez, C., Treil, S., Volberg, A.: Sharp weighted estimates for dyadic shifts and the A_2 conjecture. To appear in *Journal für die Reine und Angewandte Mathematik*, available at arXiv:1010.0755
39. John, F., Nirenberg, L.: On functions of bounded mean oscillation. *Comm. Pure Appl. Math.* **14**, 415–426 (1961)
40. Katz, N.H., Pereyra, M.C.: On the two weights problem for the Hilbert transform. *Rev. Mat. Iberoamericana* **13**(1), 211–243 (1997)
41. Lacey, M.T.: The linear bound in A_2 for Calderón–Zygmund operators: a survey. Marcinkiewicz centenary volume, *Banach Center Publ.*, 95, Polish Acad. Sci. Inst. Math., Warsaw, 97–114 (2011)
42. Lacey, M.T.: On the A_2 inequality for Calderón–Zygmund operators. Preprint (2011) available at arXiv:1106.4802
43. Lacey, M.T., Moen, K., Pérez, C., Torres, R.: The sharp bound the fractional operators on weighted L^p spaces and related Sobolev inequalities. *J. Funct. Anal.* **259**(5), 1073–1097 (2010)
44. Lacey, M.T., Petermichl, S., Reguera, M.C.: Sharp A_2 inequality for Haar shift operators. *Math. Ann.* **348**(1), 127–141 (2010)
45. Lacey, M.T., Sawyer, E., Shen, C.-Y., Uriarte-Tuero, I.: The two weight inequality for the Hilbert transform, coronas and energy conditions. Preprint (2011) available at arXiv:1108.2319
46. Lerner, A.K.: On some weighted norm inequalities for Littlewood–Paley operators. *Illinois J. Math.* **52**(2), 653–666 (2008)
47. Lerner, A.K.: A pointwise estimate for the local sharp maximal function with applications to singular integrals. *Bull. London Math. Soc.* **42**(5), 843–856 (2010)
48. Lerner, A.K.: A “local mean oscillation” and some of its applications. *Function spaces, Approximation, Inequalities and Lineability, Lectures of the Spring School in Analysis*, pp. 71–106. Matfyzpress, Prague (2011)
49. Lerner, A.K.: Mixed A_p - A_r inequalities for classical singular integrals and Littlewood–Paley operators. To appear in *J. Geom. Anal.*, available at arXiv:1105.5735
50. Lerner, A.K.: Sharp weighted norm inequalities for Littlewood–Paley operators and singular integrals. *Adv. Math.* **226**, 3912–3926 (2011)
51. Moen, K.: Sharp one-weight and two-weight bounds for maximal operators. *Studia Math.* **194**, 163–180 (2009)
52. Moraes, J.C., Pereyra, M.C.: Weighted estimates for dyadic paraproducts and t -Haar multipliers with complexity (m, n) . Submitted to *Pub. Mat.*, available at arXiv:1108.3109
53. Muckenhoupt, B.: Weighted norm inequalities for the Hardy maximal function. *Trans. Amer. Math. Soc.* **165**, 207–226 (1972)
54. Nazarov, F., Volberg, A.: A simple sharp weighted estimate of the dyadic shifts on metric spaces with geometric doubling. Preprint (2011) available at arXiv:1104.4893
55. Nazarov, F., Treil, S., Volberg, A.: The Bellman functions and two-weight inequalities for Haar multipliers. *J. Amer. Math. Soc.* **12**(4), 909–928 (1999)
56. Nazarov, F., Treil, S., Volberg, A.: The Tb -theorem on non-homogeneous spaces. *Acta Math.* **190**(2), 151–239 (2003)
57. Nazarov, F., Treil, S., Volberg, A.: Two weight estimate for the Hilbert transform and corona decomposition for non-doubling measures. Preprint (2005) available at arXiv:1003.1596
58. Nazarov, F., Reznikov, A., Volberg, A.: The proof of A_2 conjecture in a geometrically doubling metric space. Preprint (2011), available at arXiv:1106.1342
59. Ortiz, C.: Quadratic A_1 bounds for commutators of singular integrals with BMO functions. *Indiana U. Math. J.*, available at arXiv:1104.1069
60. Pereyra, M.C.: *Lecture Notes on Dyadic Harmonic Analysis*. Contemporary Mathematics, vol. 289. American Mathematical Society, Providence, RI, Chapter 1, pp. 1–61 (2001)
61. Pereyra, M.C.: Haar multipliers meet Bellman functions. *Rev. Mat. Iberoamericana* **25**(3), 799–840 (2009)

62. Pérez, C.: Endpoint Estimates for Commutators of Singular Integral Operators. *J. Funct. Anal.* **128**(1), 163–185 (1995)
63. Pérez, C.: Sharp estimates for commutators of singular integrals via iterations of the Hardy–Littlewood maximal function. *J. Fourier Anal. Appl.* **3**, 743–756 (1997)
64. Pérez, C.: A course on Singular Integrals and weights. To appear in Birkhäuser as part of the series “Advanced courses in Mathematics at the C.R.M., Barcelona”.
65. Pérez, C., Treil, S., Volberg, A.: On A_2 conjecture and corona decomposition of weights. Preprint (2010) available at arXiv:1006.2630
66. Petermichl, S., Dyadic shift and a logarithmic estimate for Hankel operators with matrix symbol. *C. R. Acad. Sci. Paris Sér. I Math.* **330**(6), 455–460 (2000)
67. Petermichl, S.: The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p characteristic. *Amer. J. Math.* **129**, 1355–1375 (2007)
68. Petermichl, S.: The sharp weighted bound for the Riesz transforms. *Proc. Amer. Math. Soc.* **136**, 1237–1249 (2008)
69. Petermichl, S., Pott, S.: An estimate for weighted Hilbert transform via square functions. *Trans. Amer. Math. Soc.* **354**(4), 1699–1703 (2002)
70. Petermichl, S., Volberg, A.: Heating of the Ahlfors-Beurling operator: Weakly quasiregular maps on the plane are quasiregular. *Duke Math J.* **112**, 281–305 (2002)
71. Riesz, M.: *Sur le fonctions conjuguées*. *Math. Zeit.* **27**, 218–244 (1927)
72. Sawyer, E.T.: A characterization of a two weight norm inequality for maximal operators. *Studia Math.* **75**(1), 1–11 (1982)
73. Stein, E.: Harmonic Analysis: real-variable methods, orthogonality, and oscillatory integrals. With the assistance of Timothy S. Murphy. Princeton Mathematical Series, vol. 43. Monographs in Harmonic Analysis, vol. III. Princeton University Press, Princeton (1993)
74. Treil, S.: Sharp A_2 estimates of Haar shifts via Bellman function. Preprint (2011) available at arXiv:1105.2252
75. Treil, S., Volberg, A.: Wavelets and the angle between past and future. *J. Funct. Anal.* **143**(2), 269–308 (1997)
76. Vagharyakhan, A.: Recovering singular integrals from the Haar shifts. *Proc. Amer. Math. Soc.* **138**(12), 4303–4309 (2010)
77. Vasuynin, V., Volberg, A.: Notes on Bellman functions in harmonic analysis. Preprint (2010), available at <http://sashavolberg.wordpress.com/>
78. Volberg, A.: Bellman function technique in harmonic analysis. Lectures of INRIA Summer School in Antibes. France (2011), available at arXiv:1106.3899
79. Wittwer, J.: A sharp estimate on the norm of the martingale transform. *Math. Res. Lett.* **7**(1), 1–12 (2000)
80. Wittwer, J.: A sharp estimate on the norm of the continuous square function. *Proc. Amer. Math. Soc.* **130**(8), 2335–2342 (2002) (electronic)
81. Wilson, M.: Weighted Littlewood–Paley theory and exponential-square integrability. *Lecture Notes in Mathematics*, vol. 1924. Springer, Berlin (2008)
82. Wojtaszczyk, P.: A mathematical introduction to wavelets. *London Mathematical Society Student Texts*, vol. 37. Cambridge University Press, Cambridge (2003)