WEIGHTED NORM INEQUALITIES

TUOMAS HYTÖNEN

Lecture notes of a course at the University of Helsinki, Winter 2011.

0. Outline

These lectures are concerned with the following objects:

Spaces: A weight (function) $0 < w \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ is identified with the positive measure (denoted by the same symbol)

$$w(E) := \int_E w(x)\,\mathrm{d}x.$$

We consider the weighted L^p space

$$L^p(w) := \Big\{ f \text{ measurable} : \|f\|_{L^p(w)} := \Big(\int_{\mathbb{R}^n} |f(x)|^p w(x)\,\mathrm{d}x \Big)^{1/p} < \infty \Big\},$$

and its weak version

$$L^{p,\infty}(w) := \Big\{ f \text{ measurable} : \|f\|_{L^{p,\infty}(w)} := \sup_{t>0} t \cdot w(\{|f| > t\})^{1/p} < \infty \Big\},$$

where $\{|f| > t\} := \{x \in \mathbb{R}^n : |f(x)| > t\}$. Note that $\|f\|_{L^{p,\infty}(w)} \leq \|f\|_{L^p(w)}$.

Operators: The operators of interest are the usual ones from Real and Harmonic Analysis. The main examples are the Hardy–Littlewood maximal operator

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

and the Hilbert transform

$$Hf(x) := \lim_{\varepsilon \to 0} \frac{1}{\pi} \int_{|x-y|>\varepsilon} \frac{f(y)\,\mathrm{d}y}{x-y}.$$

Note that the integrations involved in these operators are with respect to the Lebesgue measure, not with respect to the weighted measure w, yet we want to consider the action of these operators on functions $f \in L^p(w)$. The presence of these integrations with respect to different measures is the main difficulty of the weighted theory.

Questions: The main questions of the weighted theory are of the following three types:

(1) Given an operator $T \in \{M, H, \ldots\}$ and an exponent p, under what conditions on the weight w is it true that

$$\|Tf\|_{L^p(w)} \leq C\|f\|_{L^p(w)} \qquad \text{or} \qquad \|Tf\|_{L^{p,\infty}(w)} \leq C\|f\|_{L^p(w)}$$

for some $C = C(T, p, w)$ which is independent of $f \in L^p(w)$? This is the qualitative one-weight question, which is fairly well understood after the contributions of Muckenhoupt, Hunt, Wheeden, Coifman, Fefferman. . . in the 1970s.

(2) Assuming we know the answer to the first question, how exactly does the constant C depend on w? This is the quantitative one-weight question, whose study was started by Buckley's investigation of M in the 1990s, and has been continued by Petermichl, Volberg, Lerner, Ombrosi, Pérez, Lacey. . . in the 2000s.

Version: March 3, 2011.

(3) Repeat the first question with different weights u and v on the left and right side of the inequality. For M, a characterization was obtained by Sawyer in the 1980s, but for singular integral operators only different sufficient conditions are known. Already for H, a complete understanding is still missing, despite a number of deep recent contributions to the problem.

The emphasis of these lectures is on the quantitative one-weight theory. Only restricted aspects of the two-weight theory will appear from time to time.

Framework: Instead of directly working with the classical operators $T \in \{M, H, \ldots\}$, we will most of the time consider certain simpler models, so-called dyadic operators. However, as it turns out, the model is sufficiently rich, so that we can recover the original classical operators by means of appropriate averages of the dyadic operators.

1. Maximal function estimates

1.1. **Abstract dyadic cubes.** Let $\mathscr{D}$ be a countable collection of measurable subsets of $\mathbb{R}^n$ (we could consider more general measure spaces) with the following property:

$$\forall Q, R \in \mathscr{D} : Q \cap R \in \{Q, R, \varnothing\}. \tag{1.1}$$

We will refer to the elements of $\mathscr{D}$ as the *dyadic cubes*. The main example to keep in mind is the standard dyadic cubes given by

$$\mathscr{D} = \{2^{-k}\big([0,1)^n + m\big) : k \in \mathbb{Z}, m \in \mathbb{Z}^n\}.$$

However, many of the basic results are valid assuming only the abstract dyadic structure as postulated by (1.1).

An important principle related to dyadic cubes is the following. It is the dyadic counterpart of the more complicated covering arguments of Classical Analysis.

Lemma 1.1 (Dyadic covering lemma)**.** *Let $\mathscr{Q} \subseteq \mathscr{D}$ be a finite collection. Let $\mathscr{Q}^*$ be the maximal cubes in $\mathscr{Q}$: all $R \in \mathscr{Q}$ which are not strictly contained in any bigger $Q \in \mathscr{Q}$. Then every $Q \in \mathscr{Q}$ is contained in some $Q \in \mathscr{Q}^*$ and*

$$\bigcup_{R \in \mathscr{Q}^*} R = \bigcup_{Q \in \mathscr{Q}} Q.$$

The elements of $\mathscr{Q}^*$ are pairwise disjoint.

(Often, the same maximality argument also works even when there are infinitely many cubes, but here it is particularly obvious.)

Proof. Let $Q \in \mathscr{Q}$; we want to prove that $Q \subseteq R$ for some $R \in \mathscr{Q}^*$. If Q is maximal, then $Q \subseteq Q \in \mathscr{Q}^*$. If not, then by definition $Q \subsetneq Q'$ for some $Q' \in \mathscr{Q}$. If Q' is maximal, then we are done. If not, then $Q \subsetneq Q' \subsetneq Q''$ for some $Q \in \mathscr{Q}$, and we check if Q'' is maximal or not. Since $\mathscr{Q}$ is finite, this chain of cubes must terminate after finitely many steps.

To see the equality of the unions, note that $\subseteq$ is clear, since $\mathscr{Q}_t^* \subseteq \mathscr{Q}_t$ by definition. And $\supseteq$ follows from the previous paragraph: any $Q \in \mathscr{Q}_t$ is contained in some $R \in \mathscr{Q}_t^*$.

For disjointness, consider two different $Q, Q' \in \mathscr{Q}$ with $Q \cap Q' \neq \varnothing$. By (1.1), this means that $Q \cap Q' \in \{Q, Q'\}$, that is, one of the cubes contains the other. Hence the contained cube is not maximal, thus not in $\mathscr{Q}^*$. □

1.2. **The dyadic maximal operator.** We consider the maximal operator given by

$$Mf(x) := \sup_{Q \in \mathscr{D}} 1_Q(x) \frac{1}{|Q|} \int_Q |f| \, \mathrm{d}y,$$

or with respect to another measure μ:

$$M^\mu f(x) := \sup_{Q \in \mathscr{D}} 1_Q(x) \frac{1}{\mu(Q)} \int_Q |f| \, \mathrm{d}\mu.$$

The integral average will be often abbreviated as

$$\langle f\rangle_Q := \fint_Q f\,\mathrm{d}y := \frac{1}{|Q|}\int_Q f\,\mathrm{d}y,$$

or with respect to other measures:

$$\langle f\rangle_Q^\mu := \fint_Q f\,\mathrm{d}\mu := \frac{1}{\mu(Q)}\int_Q f\,\mathrm{d}\mu.$$

1.3. **Universal maximal function estimates.** A fundamental property of the dyadic maximal function is the following set of estimates: the maximal function M^μ is always bounded on $L^p(\mu)$ for the same μ. The problems only arise when we insist on the boundedness of M ($\mathrm{d}\mu = \mathrm{d}x$) on $L^p(w)$ ($\mathrm{d}\mu' = w\,\mathrm{d}x$).

Theorem 1.1. *We have the estimates*

$$\|M^\mu f\|_{L^p(\mu)} \leq p'\|f\|_{L^p(\mu)}, \qquad p\in(1,\infty],$$

$$t\cdot\mu(\{M^\mu f > t\}) \leq \int_{\{M^\mu f>t\}} |f|\,\mathrm{d}\mu \leq \|f\|_{L^1(\mu)}.$$

Proof. We make the following approximation argument, which is often handy also later on: We may assume that the collection $\mathscr{D}$ is finite. Indeed, since $\mathscr{D}$ is countable, we have $\mathscr{D} = \{Q_i\}_{i=1}^\infty$. Let $\mathscr{D}_k := \{Q_i\}_{i=1}^k$, and let M_k^μ be the maximal function related to $\mathscr{D}_k$ in place of $\mathscr{D}$. Then $M_k^\mu f \uparrow M^\mu f$ as $k\to\infty$. Assuming we can prove the finite case, we have

$$t\cdot\mu(\{M_k^\mu f > t\}) \leq \int_{\{M_k^\mu f>t\}} |f|\,\mathrm{d}\mu \leq \int_{\{M^\mu f>t\}} |f|\,\mathrm{d}\mu$$

for all $k\in\mathbb{N}$. But the left side converges to $t\cdot\mu(\{M^\mu f>t\})$ as $k\to\infty$. The bound for $\|M^\mu f\|_{L^p(\mu)}$ also follows from that of $\|M_k^\mu f\|_{L^p(\mu)}$ by dominated convergence, for $p\in(1,\infty)$. (For $p=\infty$, the bound is immediate, since clearly $M^\mu f\leq \|f\|_{L^\infty(\mu)}$.)

With $\mathscr{D}$ finite, consider first the weak-type estimate. Let $f\geq 0$, without loss of generality (think why!). By definition, there holds $M^\mu f(x) > t$ if and only if $\langle f\rangle_Q^\mu > t$ for some $Q\ni x$, and in this case $M^\mu f > t$ at all points of Q. Let $\mathscr{Q}_t := \{Q\in\mathscr{D} : \langle f\rangle_Q^\mu > t\}$. Then

$$\{M^\mu f > t\} = \bigcup_{Q\in\mathscr{Q}_t} Q.$$

Let $\mathscr{Q}_t^*$ be the maximal cubes in $\mathscr{Q}_t$, thus they are pairwise disjoint. Hence

$$\begin{aligned}\mu(\{M^\mu f>t\}) &= \mu\Big(\bigcup_{Q\in\mathscr{Q}_t} Q\Big) = \mu\Big(\bigcup_{Q\in\mathscr{Q}_t^*} Q\Big) = \sum_{Q\in\mathscr{Q}_t^*}\mu(Q)\\ &\leq \sum_{Q\in\mathscr{Q}_t^*}\frac{1}{t}\int_Q f\,\mathrm{d}\mu = \frac{1}{t}\int_{\bigcup\mathscr{Q}_t^*} f\,\mathrm{d}\mu = \frac{1}{t}\int_{\{M^\mu f>t\}} f\,\mathrm{d}\mu.\end{aligned}$$

We turn to the L^p estimate for $p\in(1,\infty)$. Recall the useful formula

$$\int_\Omega |f|^p\,\mathrm{d}\mu = \int_0^\infty pt^{p-1}\mu(\{|f|>t\})\,\mathrm{d}t. \tag{1.2}$$

Then

$$\begin{aligned}\|M^\mu f\|_{L^p(\mu)}^p &= \int_0^\infty pt^{p-1}\mu(\{M^\mu f>t\})\,\mathrm{d}t \leq \int_0^\infty pt^{p-2}\int_{\{M^\mu f(x)>t\}} |f(x)|\,\mathrm{d}\mu(x)\,\mathrm{d}t\\ &= \int_{\mathbb{R}^n}\int_0^{M^\mu(x)} pt^{p-1}\,\mathrm{d}t|f(x)|\,\mathrm{d}\mu(x) = \int_{\mathbb{R}^n}\frac{p}{p-1}(M^\mu f(x))^{p-1}|f(x)|\,\mathrm{d}\mu(x).\end{aligned}$$

We apply Hölder's inequality with exponents p' and p, observing that $(p-1)p'=p$, to deduce

$$\|M^\mu f\|_{L^p(\mu)}^p \leq p'\|M^\mu f\|_{L^p(\mu)}^{p-1}\|f\|_{L^p(\mu)}.$$

The claim follows, in principle, after dividing by $\|M^\mu f\|_{L^p(\mu)}^{p-1}$. The only problem is to make sure that this factor is not infinite. But this is easy for a finite $\mathscr{D}$; we only need to check that each $1_Q \fint_Q f\,\mathrm{d}\mu$ belongs to $L^p(\mu)$ (as M^μ is the maximum of finitely many such functions). And we have

$$\Big\|1_Q \fint_Q f\,\mathrm{d}\mu\Big\|_{L^p(\mu)} = \mu(Q)^{1/p}\frac{1}{\mu(Q)}\int_Q f\,\mathrm{d}\mu \le \mu(Q)^{1/p}\frac{1}{\mu(Q)}\|f\|_{L^p(\mu)}\mu(Q)^{1/p'} = \|f\|_{L^p(\mu)}.$$

This completes the proof. □

1.4. **First weighted inequalities.** By a simple variant of the proof of the universal maximal function estimate, we obtain the following boundedness property of the unweighted M:

Proposition 1.1 (Fefferman–Stein 1971 [7])**.**

$$t\cdot w(\{Mf>t\}) \le \int_{\{Mf>t\}} |f|\,Mw \le \|f\|_{L^1(Mw)}.$$

Proof. We may again assume that $\mathscr{D}$ is finite to start with, and $f\ge 0$. Let $\mathscr{Q}_t := \{Q\in\mathscr{D} : \langle f\rangle_Q > t\}$, and $\mathscr{Q}_t^*$ consist of the maximal elements of $\mathscr{Q}_t$. Then

$$\begin{aligned} w(\{Mf>t\}) = \sum_{Q\in\mathscr{Q}_t^*} w(Q) = \sum_{Q\in\mathscr{Q}_t^*}\frac{w(Q)}{|Q|}|Q| \le \sum_{Q\in\mathscr{Q}_t^*}\inf_Q Mw\cdot\frac{1}{t}\int_Q f \le \frac{1}{t}\sum_{Q\in\mathscr{Q}_t^*}\int_Q f\,Mw \\ = \frac{1}{t}\int_{\bigcup\mathscr{Q}_t^*} f\,Mw = \frac{1}{t}\int_{\{Mf>t\}} f\,Mw. \end{aligned}$$ □

The previous result gave the boundedness of $M : L^1(Mw)\to L^{1,\infty}(w)$ between different weighted spaces. It also provides an immediate sufficient condition for $M : L^1(w)\to L^{1,\infty}(w)$:

1.5. **Muckenhoupt's class A_1.** A weight belongs to the class A_1 if

$$[w]_{A_1} := \operatorname{ess\,sup}\frac{Mw}{w} < \infty.$$

The condition can be formulated in various ways: Using the definition of the maximal function, it says that for a.e. x, for all dyadic $Q\ni x$, we have $\langle w\rangle_Q \le [w]_{A_1} w(x)$. Since there are only countably many $Q\in\mathscr{D}$, the union of the null sets in "a.e." above is also a null set, and we may permute "a.e. x" and "all Q" to the result that: for all dyadic Q, for a.e. $x\in Q$, there holds $\langle w\rangle_Q \le [w]_{A_1}w(x)$, and yet in other words:

$$\langle w\rangle_Q \le [w]_{A_1}\operatorname*{ess\,inf}_Q w \qquad \forall Q\in\mathscr{D}.$$

This is often applied in the inverted form

$$\operatorname*{ess\,sup}_Q\frac{1}{w} \le [w]_{A_1}\frac{1}{\langle w\rangle_Q} \qquad \forall Q\in\mathscr{D}.$$

Corollary 1.1.

$$t\cdot w(\{Mf>t\}) \le [w]_{A_1}\|f\|_{L^1(w)}.$$

Proof. This is immediate from Proposition 1.1 and the definition of $[w]_{A_1}$. □

Remark 1.1 (The Muckenhoupt–Wheeden conjecture)**.** Proposition (1.1) motivated the following analogous *Muckenhoupt–Wheeden conjecture* for the Hilbert transform:

$$\|Hf\|_{L^{1,\infty}(w)} \le C\|f\|_{L^1(Mw)},$$

with C independent of f and w. If true, this would imply the *weak Muckenhoupt–Wheeden conjecture*

$$\|Hf\|_{L^{1,\infty}(w)} \le C[w]_{A_1}\|f\|_{L^1(w)},$$

again with C independent of f and w.

These remained open for a long time. In 2010, counterexamples by Reguera, Thiele, Nazarov, Reznikov, Vasyunin and Volberg [20, 22, 23] show that both conjectures are *false.*

As it turns out, the A_1 condition is precisely what is needed for the weighted weak-type L^1 bounds of the maximal function:

Proposition 1.2.

$$\|M\|_{L^1(w)\to L^{1,\infty}(w)} = [w]_{A_1}.$$

Proof. We already proved $\leq$, so let us consider $\geq$. Denote $N := \|M\|_{L^1(w)\to L^{1,\infty}(w)}$, and note first that

$$w(\{Mf \geq t\}) = \lim_{\varepsilon\to 0} w(\{Mf > (1-\varepsilon)t\}) \leq \lim_{\varepsilon\to 0} \frac{N}{(1-\varepsilon)t}\|f\|_{L^1(w)} = \frac{N}{t}\|f\|_{L^1(w)}.$$

Fix $Q \in \mathscr{D}$ and consider an $f = f1_Q \geq 0$ in $L^1(w)$. Then $Mf \geq \langle f\rangle_Q$ on all of Q, and hence

$$w(Q) \leq w(\{Mf \geq \langle f\rangle_Q\}) \leq \frac{N}{\langle f\rangle_Q}\|f1_Q\|_{L^1(w)}.$$

The special case of $f = 1_E$ with $E \subseteq Q$ gives $w(Q) \leq Nw(E)/\langle 1_E\rangle_Q$, that is

$$\frac{w(Q)}{|Q|} \leq N\frac{w(E)}{|E|}.$$

We further specialize to $E := \{x \in Q : w(x) < \operatorname{ess\,inf}_Q w + \varepsilon\}$, which has positive measure by definition of $\operatorname{ess\,inf}$. Then

$$\frac{w(Q)}{|Q|} \leq N\frac{w(E)}{|E|} = \frac{N}{|E|}\int_E w\,\mathrm{d}x \leq \frac{N}{|E|}\int_E (\operatorname*{ess\,inf}_Q w + \varepsilon)\,\mathrm{d}x = N(\operatorname*{ess\,inf}_Q w + \varepsilon).$$

As $\varepsilon \to 0$, this gives by our reformulation of the A_1 condition that $[w]_{A_1} \leq N$. □

1.6. **The A_p classes.** With the boundedness of $M : L^1(w) \to L^{1,\infty}(w)$ for $w \in A_1$, it would be reasonably straightforward to see that the A_1 condition is also sufficient for the boundedness of $M : L^p(w) \to L^p(w)$ for $p > 1$. However, this condition is stronger than necessary for this estimate, and there is instead another condition A_p adapted to each p:

Proposition 1.3.

$$\|M\|^p_{L^p(w)\to L^{p,\infty}(w)} = [w]_{A_p} := \sup_{Q\in\mathscr{D}} \Big(\fint_Q w\,\mathrm{d}x\Big)\Big(\fint_Q w^{-1/(p-1)}\,\mathrm{d}x\Big)^{1/p}, \qquad p \in (1,\infty).$$

Proof. The estimate $\leq$ can be proven by an adaptation of the proof of Proposition 1.1, and is left as an exercise. Let us prove $\geq$. Let again $f = f1_Q \geq 0$ be in $L^p(w)$, and observe that $Mf \geq \langle f\rangle_Q$ on the cube Q. Let $N := \|M\|_{L^p(w)\to L^{p,\infty}(w)}$. Hence

$$w(Q) \leq w(\{Mf \geq \langle f\rangle_Q\}) \leq \frac{N^p}{\langle f\rangle_Q^p}\int_Q f^p w\,\mathrm{d}x,$$

and reorganizing,

$$\frac{w(Q)}{|Q|}\Big(\frac{1}{|Q|}\int_Q f\,\mathrm{d}x\Big)^p \leq N^p\frac{1}{|Q|}\int_Q f^p w\,\mathrm{d}x.$$

We would like to choose f so that $f = f^p$ on Q, i.e., $f = w^{-1/(p-1)}1_Q$. The problem is that, a priori, it is not clear that $f \in L^p(w)$. So we choose $f := (w+\varepsilon)^{-1/(p-1)}1_Q$ instead, which is a bounded function. Substituting back, and estimating $w \leq w + \varepsilon$ on the right,

$$\frac{w(Q)}{|Q|}\Big(\frac{1}{|Q|}\int_Q (w+\varepsilon)^{-1/(p-1)}\,\mathrm{d}x\Big)^p \leq N^p\frac{1}{|Q|}\int_Q (w+\varepsilon)^{-p/(p-1)}(w+\varepsilon)\,\mathrm{d}x,$$

where $(w+\varepsilon)^{-p/(p-1)}(w+\varepsilon) = (w+\varepsilon)^{-1/(p-1)}$. Dividing out the common factor, it follows that

$$\frac{w(Q)}{|Q|}\Big(\frac{1}{|Q|}\int_Q (w+\varepsilon)^{-1/(p-1)}\,\mathrm{d}x\Big)^{p-1} \leq N^p,$$

and letting $\varepsilon \to 0$ concludes the argument. □

1.7. **Comparison of the A_p classes; the class A_∞.** Let us introduce the local A_p characteristics

$$A_1(Q,w) := \langle w\rangle_Q \operatorname*{ess\,sup}_Q w^{-1}, \qquad A_p(Q,w) := \langle w\rangle_Q \langle w^{-1/(p-1)}\rangle_Q^{p-1} \quad (p\in(1,\infty)),$$

so that $[w]_{A_p} = \sup_{Q\in\mathscr{D}} A_p(Q,w)$ for $p\in[1,\infty)$.

It is a straightforward exercise to check that for $1<p<q<\infty$ and any $0<w\in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, we have

$$A_1(Q,w)\geq A_p(Q,w)\geq A_q(Q,w)\geq 1,$$

where these quantities may be finite or infinite. From these estimates it follows that if $A_p(Q,w)<\infty$ for some $p\in[1,\infty)$, then $A_q(Q,w)<\infty$ for all $q\in[p,\infty)$, and these local A_q characteristics are decreasing and bounded from below. By elementary analysis, there exists a limit

$$A_\infty(Q,w) := \lim_{q\to\infty} A_q(Q,w).$$

Let us find a more explicit expression for it. Since the factor $\langle w\rangle_Q$ is the same in all $A_q(Q,w)$, we only need to consider (let $\varepsilon_q := 1/(q-1)\to 0$ as $q\to\infty$)

$$\Big(\fint_Q w^{-1/(q-1)}\,\mathrm{d}x\Big)^{q-1} = \Big(\fint_Q \Big[\frac{w^{-\varepsilon_q}-1}{\varepsilon_q}\varepsilon_q+1\Big]\,\mathrm{d}x\Big)^{q-1} = \Big(1+\frac{1}{q-1}\fint_Q \frac{w^{-\varepsilon_q}-1}{\varepsilon_q}\,\mathrm{d}x\Big)^{q-1},$$

and recall that $(1+c_m/m)^m\to e^c$ as $m\to\infty$, if $c_m\to c$. Hence, assuming that the limit exists, we have

$$\lim_{q\to\infty}\Big(\fint_Q w^{-1/(q-1)}\,\mathrm{d}x\Big)^{q-1} = \exp\Big(\lim_{\varepsilon\to 0}\fint_Q \frac{w^{-\varepsilon}-1}{\varepsilon}\,\mathrm{d}x\Big).$$

It is immediate that we have the pointwise limit

$$\frac{w^{-\varepsilon}-1}{\varepsilon}\to -\log w,$$

and it remains to check the conditions for dominated convergence. Note that $w\in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and $A_p(Q,w)<\infty$ implies that $w,w^{-\delta}\in L^1(Q)$ for $\delta=1/(p-1)>0$. By the mean value theorem

$$\frac{w^{-\varepsilon}-1}{\varepsilon} = -\log w\cdot e^{-\varepsilon'\log w} = -\log w\cdot w^{-\varepsilon'}, \qquad \varepsilon'\in(0,\varepsilon).$$

Since $1+x\leq e^x$, we have $\log w\leq w-1\leq w$, and hence $|\log w| = \upsilon^{-1}\log w^\upsilon\leq \upsilon^{-1}w^\upsilon$ for $w\geq 1$. For $w<1$ we use $|\log w| = \log w^{-1}\leq \upsilon^{-1}w^{-\upsilon}$, so altogether, for $\varepsilon\leq\upsilon$

$$|\log w\cdot w^{-\varepsilon'}|\leq \max\{\upsilon^{-1}w^{-\upsilon},w\}\cdot\max\{1,w^{-\upsilon}\}\leq \upsilon^{-1}w^{-2\upsilon}+w.$$

With $\upsilon:=\delta/2$, both terms are integrable, so we have found an integrable majorant for the $(w^{-\varepsilon}-1)/\varepsilon$, uniformly in ε. Thus

$$A_\infty(Q,w) = \Big(\fint_Q w\,\mathrm{d}x\Big)\exp\Big(\fint_Q\lim_{\varepsilon\to 0}\frac{w^{-\varepsilon}-1}{\varepsilon}\,\mathrm{d}x\Big) = \Big(\fint_Q w\,\mathrm{d}x\Big)\exp\Big(-\fint_Q \log w\,\mathrm{d}x\Big),$$

and it is natural to define

$$[w]_{A_\infty} := \sup_{Q\in\mathscr{D}} A_\infty(Q,w).$$

By what we have shown above (it is also easy to check this directly from the definition), we have $A_p\subseteq A_\infty$ and $[w]_{A_\infty}\leq [w]_{A_p}$.

In fact, if $w\in A_\infty$, it can be shown (although we will not do so here) that $w\in A_p$ for some $p<\infty$.

1.8. **Weighted L^p norm estimates for the maximal function.** We return to the maximal function. Recall that $\|M\|_{L^p(w)\to L^{p,\infty}(w)} = [w]_{A_p}^{1/p}$. Also the boundedness on $L^p(w)$ depends on the same quantity, but we need a somewhat larger power:

Theorem 1.2 (Muckenhoupt 1972; Buckley 1993 [1, 19])**.** *The maximal operator M is bounded on $L^p(w)$ if and only if $w \in A_p$; more precisely*

$$\|M\|_{L^p(w)\to L^p(w)} \leq C_p[w]_{A_p}^{1/(p-1)}, \qquad p \in (1,\infty),$$

where C_p depends on p but not on w.

Here the qualitative statement is due to Muckenhoupt, the quantitative dependence due to Buckley. In fact, the Muckenhoupt–Buckley theorem will be deduced as a consequence of the following more precise version, which I proved with Carlos Pérez last week (10–14 January):

Theorem 1.3 (Hytönen–Pérez 2011 [12])**.**

$$\|M\|_{L^p(w)\to L^p(w)} \leq 4e \cdot p' \cdot \big([w]_{A_p}[w^{-1/(p-1)}]_{A_\infty}\big)^{1/p}.$$

To see that this implies the Muckenhoupt–Buckley theorem, observe the following immediate consequence of the definition:

$$[w]_{A_p} = [w^{-1/(p-1)}]_{A_{p'}}^{p-1}. \tag{1.3}$$

Since $[w^{-1/(p-1)}]_{A_{p'}} \geq [w^{-1/(p-1)}]_{A_\infty}$, it is easy to conclude after some algebra with the exponents.

1.9. **Sawyer's dual weight trick.** We want to prove an estimate of the form

$$\|Tf\|_{L^p(w)} \leq N\|f\|_{L^p(w)},$$

presently for $T = M$, but the trick given here is equally valid for other operators. Let us substitute $f = \phi\sigma$, where σ is going to be a new weight yet to be chosen. This leads to the equivalent formulation

$$\|M(\phi\sigma)\|_{L^p(w)} \leq N\|\phi\sigma\|_{L^p(w)} = N\Big(\int |\phi|^p\sigma^p w\,\mathrm{d}x\Big)^{1/p}.$$

We want to choose σ so that $\sigma^p w = \sigma$, i.e., $\sigma = w^{-1/(p-1)}$. Thus an equivalent problem is to prove that

$$\|T(\phi\sigma)\|_{L^p(w)} \leq N\|f\|_{L^p(\sigma)},$$

where $\sigma = w^{-1/(p-1)}$ is precisely the weight which appears in Theorem 1.3. The advantage of this reformulation is that the same weight σ appears inside the operator T and in the norm on the right side.

1.10. **The two-weight maximal inequality.** By Sawyer's trick, proving Theorem 1.3 is reduced to proving

$$\|M(\phi\sigma)\|_{L^p(w)} \leq N\|\phi\|_{L^p(\sigma)}$$

for certain N and σ depending on w. But let us for the moment consider this estimate on its own right, for general w and σ, which need not be related to each other. We would like to have a criterion for the pair of weights (w,σ) under which such an estimate holds. Let us again consider $\phi \geq 0$ and a finite $\mathscr{D}$ for simplicity, this restriction being easy to lift in the end.

Recall that

$$M(\phi\sigma)(x) = \sup_{Q\in\mathscr{D}} 1_Q(x)\langle f\sigma\rangle_Q,$$

where sup is actually max in the finite case. Let

$$E(Q) := \Big\{x \in Q : M(\phi\sigma)(x) = \langle\phi\sigma\rangle_Q > \langle\phi\sigma\rangle_{Q'} \text{ for all } Q' \supsetneq Q\Big\}$$

be the part of Q where the value of the maximal function is reached as the average on Q.

A little thought confirms that the sets $E(Q)$ are pairwise disjoint, and

$$M(\phi\sigma) = \sum_{Q\in\mathscr{D}} 1_{E(Q)}\langle\phi\sigma\rangle_Q =: \tilde{M}(\phi\sigma),$$

so it suffices to consider the $L^p(w)$ bound for this linearization $\tilde{M}$ of M. (Note that $M(\phi\sigma) = \tilde{M}(\phi\sigma)$ for the given function ϕ which was used to define the sets $E(Q)$, but once this definition is made, we may also consider the action of $\tilde{M}$, as defined above, on other functions. This $\tilde{M}$, unlike M, is a linear operator.) By disjointness,

$$\Big\| \sum_{Q\in\mathscr{D}} 1_{E(Q)}\langle\phi\sigma\rangle_Q \Big\|_{L^p(w)}^p = \sum_{Q\in\mathscr{D}} w(E(Q))(\langle\phi\sigma\rangle_Q)^p = \sum_{Q\in\mathscr{D}} w(E(Q))\Big(\frac{\sigma(Q)}{|Q|}\Big)^p(\langle\phi\rangle_Q^\sigma)^p,$$

and we want a condition for this to be bounded by $N^p\|\phi\|_{L^p(\sigma)}$. This is provided by the following:

Theorem 1.4 (Dyadic Carleson embedding theorem). *For $p\in(1,\infty)$, the estimate*

$$\Big(\sum_{Q\in\mathscr{D}} a_Q(\langle\phi\rangle_Q^\sigma)^p\Big)^{1/p} \leq N\|\phi\|_{L^p(\sigma)} \qquad \forall\phi\in L^p(\sigma)$$

holds if and only if

$$\Big(\sum_{Q\subseteq R} a_Q\Big)^{1/p} \leq \tilde{N}\sigma(R)^{1/p} \qquad \forall R\in\mathscr{D};$$

moreover, $\tilde{N}\leq N\leq p'\cdot\tilde{N}$.

Proof. The "only if" part is immediate by substituting $\phi = 1_R$. The "if" part is the main implication.

We use the identity (1.2) with the discrete set $\Omega = \mathscr{D}$ and measure $\mu(\{Q\}) = a_Q$, and $f(Q) = \langle\phi\rangle_Q^\sigma$. Writing $\mathscr{Q}_t := \{Q\in\mathscr{D} : \langle\phi\rangle_Q^\sigma > t\}$ and $\mathscr{Q}_t^*$ for its maximal cubes, this gives

$$\begin{aligned}
\sum_{Q\in\mathscr{D}} a_Q(\langle\phi\rangle_Q^\sigma)^p &= \int_0^\infty pt^{p-1}\sum_{Q\in\mathscr{Q}_t} a_Q\,\mathrm{d}t \leq \int_0^\infty pt^{p-1}\sum_{R\in\mathscr{Q}_t^*}\sum_{Q\subseteq R} a_Q\,\mathrm{d}t \\
&\leq \int_0^\infty pt^{p-1}\sum_{R\in\mathscr{Q}_t^*}\tilde{N}^p\sigma(R)\,\mathrm{d}t = \tilde{N}^p\int_0^\infty pt^{p-1}\sigma\Big(\bigcup_{R\in\mathscr{Q}_t^*} R\Big)\,\mathrm{d}t \\
&= \tilde{N}^p\int_0^\infty pt^{p-1}\sigma(\{M^\sigma\phi > t\})\,\mathrm{d}t = \tilde{N}^p\|M^\sigma\phi\|_{L^p(\sigma)}^p \leq \big(\tilde{N}\cdot p'\cdot\|\phi\|_{L^p(\sigma)}\big)^p,
\end{aligned}$$

where we used the usual properties of the maximal dyadic cubes, and the universal maximal function estimate in the last step. □

If we apply the Carleson embedding with $a_Q = w(E(Q))\big(\sigma(Q)/|Q|\big)^p$, we find that

$$\|\tilde{M}(\phi\sigma)\|_{L^p(w)} \leq N\|f\|_{L^p(\sigma)} \tag{1.4}$$

if and only if

$$\sum_{Q\subseteq R} w(E(Q))\Big(\frac{\sigma(Q)}{|Q|}\Big)^p \leq \tilde{N}^p\sigma(R) \qquad \forall R\in\mathscr{D}. \tag{1.5}$$

Note that on $E(Q)\subseteq Q\subseteq R$, we have $\sigma(Q)/|Q| \leq M(\sigma 1_R)$, and hence

$$\begin{aligned}
\sum_{Q\subseteq R} w(E(Q))\Big(\frac{\sigma(Q)}{|Q|}\Big)^p &= \int\sum_{Q\subseteq R} 1_{E(Q)}\Big(\frac{\sigma(Q)}{|Q|}\Big)^p w \\
&\leq \int\sum_{Q\subseteq R} 1_{E(Q)}M(1_R\sigma)^p w \leq \int_R M(1_R\sigma)^p w.
\end{aligned}$$

So if $\|1_R M(1_R\sigma)\|_{L^p(w)} \leq \tilde{N}\sigma(R)^{1/p}$, then (1.5) holds, hence by Carleson's embedding also (**??**), and therefore the original two-weight inequality

$$\|\tilde{M}(\phi\sigma)\|_{L^p(w)} \leq N\|f\|_{L^p(\sigma)}.$$

Conversely, if this estimate holds, then clearly also $\|1_R M(1_R\sigma)\|_{L^p(w)} \leq \tilde{N}\sigma(R)^{1/p}$; just substitute $f = 1_R$. So altogether we have proven:

Theorem 1.5 (Sawyer 1982 [25]). *The two-weight maximal function estimate*

$$\|M(\phi\sigma)\|_{L^p(w)} \leq N\|\phi\|_{L^p(\sigma)} \qquad \forall \phi \in L^p(\sigma)$$

holds if and only if

$$\|1_R M(1_R\sigma)\|_{L^p(w)} \leq \tilde{N}\sigma(R)^{1/p} \qquad \forall Q \in \mathscr{D};$$

moreover, $\tilde{N} \leq N \leq p' \cdot \tilde{N}$.

1.11. **Back to the Muckenhoupt–Buckley theorem; principal cubes.** Now we return to the case that $w \in A_p$ and $\sigma = w^{-1/(p-1)}$. We want to estimate the constant $\tilde{N}$ in Sawyer's two-weight theorem in this case. This is accomplished with the help of another linearization of M involving the following *principal cubes*: Let $\mathscr{S}_0 := \{R\}$ and recursively

$$\mathscr{S}_k := \bigcup_{S \in \mathscr{S}_{k-1}} \{Q \subset S : \langle\sigma\rangle_Q > 2\langle\sigma\rangle_S, Q \text{ is a maximal such cube}\},$$

and then $\mathscr{S} := \bigcup_{k=0}^{\infty} \mathscr{S}_k$. Let

$$E(S) := S \setminus \bigcup_{\substack{S' \in \mathscr{S} \\ S' \subsetneq S}} S'$$

be the part of S which is not contained in any smaller principal cube. The sets $E(S)$, $S \in \mathscr{S}$, are disjoint and partition R.

We also observe that $E(S)$ mush consist of a reasonable fraction of S. In fact, let $S \in \mathscr{S}$ and consider all maximal $S' \in \mathscr{S}$ strictly contained in S. They all satisfy $\sigma(S')/|S'| > 2\sigma(S)/|S|$, i.e., $|S'| < \frac{1}{2}\sigma(S')|S|/\sigma(S)$. Hence

$$\sum |S'| \leq \frac{1}{2}\frac{|S|}{\sigma(S)} \sum \sigma(S') \leq \frac{1}{2}\frac{|S|}{\sigma(S)}\sigma(S) = \frac{1}{2}|S|,$$

and thus $|E(S)| \geq \frac{1}{2}|S|$.

If $x \in E(S)$ and $Q \ni x$, then $\langle\sigma\rangle_Q \leq 2\langle\sigma\rangle_S$, and hence $1_R M(1_R\sigma) \leq 2\langle\sigma\rangle_S$ on $1_{E(S)}$. So altogether

$$\begin{aligned}
\|1_R M(1_R\sigma)\|_{L^p(w)}^p &\leq 2^p \Big\| \sum_{S \in \mathscr{S}} 1_{E(S)}\langle\sigma\rangle_S \Big\|_{L^p(w)}^p = 2^p \sum_{S \in \mathscr{S}} w(E(S)) \Big(\frac{\sigma(S)}{|S|}\Big)^p \\
&\leq 2^p \sum_{S \in \mathscr{S}} \frac{w(E(S))}{|S|} \Big(\frac{\sigma(S)}{|S|}\Big)^{p-1} \Big(\frac{\sigma(S)}{|S|}\Big)|S| \\
&\leq 2^{p+1} \sum_{S \in \mathscr{S}} A_p(S, w) A_\infty(S, \sigma) \exp\Big(\fint_S \log\sigma\Big)|E(S)| \\
&\leq 2^{p+1}[w]_{A_p}[\sigma]_{A_\infty} \int_R \sum_{S \in \mathscr{S}} \exp\Big(\fint_S \log\sigma\Big) 1_{E(S)}(x)\,\mathrm{d}x \\
&\leq 2^{p+1}[w]_{A_p}[\sigma]_{A_\infty} \int_R \sup_{Q \in \mathscr{D}} 1_Q(x) \exp\Big(\fint_Q \log\sigma 1_R\Big)\,\mathrm{d}x \\
&=: 2^{p+1}[w]_{A_p}[\sigma]_{A_\infty} \int_R M_0(1_R\sigma)(x)\,\mathrm{d}x,
\end{aligned}$$

where the maximal function M_0 is essentially defined by the last step above, formally

$$M_0 f(x) := \sup_{Q \in \mathscr{D}} 1_Q(x) \exp\Big(\fint_Q \log|f|\Big).$$

Taking for granted for the moment the following mapping property of M_0...

Lemma 1.2. *For all* $p \in (0, \infty]$,

$$\|M_0 f\|_{L^p} \leq e^{1/p}\|f\|_{L^p}.$$

...we may conclude that

$$\|1_R M(1_R\sigma)\|_{L^p(w)}^p \leq 2^{p+1}[w]_{A_p}[\sigma]_{A_\infty}\int_R M_0(1_R\sigma) \leq 2^{p+1}[w]_{A_p}[\sigma]_{A_\infty}\cdot e\cdot\sigma(R).$$

Hence we can take

$$\tilde{N} \leq \big(2^{p+1}e\cdot[w]_{A_p}[\sigma]_{A_\infty}\big)^{1/p} = 2^{1+1/p}e^{1/p}\big([w]_{A_p}[\sigma]_{A_\infty}\big)^{1/p} \leq 4e\big([w]_{A_p}[\sigma]_{A_\infty}\big)^{1/p}$$

in Sawyer's two-weight characterization for M, and the mentioned result finally tells us that

$$\|M(\sigma\phi)\|_{L^p(w)} \leq 4e\cdot p'\cdot([w]_{A_p}[\sigma]_{A_\infty}\big)^{1/p}\|\phi\|_{L^p(\sigma)}.$$

By Sawyer's two-weight trick, this concludes the proof of Theorem 1.3.

Remark 1.2. A careful reading of the proof shows that the product

$$[w]_{A_p}[\sigma]_{A_\infty} = \sup_Q A_p(Q,w)\cdot\sup_R A_\infty(R,\sigma),$$

where the supremum is taken independently for the two factors, could in fact have been replaced by the somewhat smaller quantity

$$\sup_Q A_p(Q,w)A_\infty(Q,\sigma).$$

1.12. **The logarithmic maximal operator.** Let us now prove Lemma 1.2 about the logarithmic maximal operator M_0.

Proof of Lemma 1.2. By Jensen's inequality and the basic properties of the logarithm, we have

$$M_0 f \leq Mf, \qquad M_0 f = (M_0|f|^{1/q})^q \leq (M|f|^{1/q})^q, \qquad q\in(0,\infty).$$

By the L^q boundedness of the usual maximal function for $q>1$, we have

$$\int[M_0 f]^p \leq \int[M|f|^{p/q}]^q \leq (q')^q\int(|f|^{p/q})^q = (q')^q\int|f|^p.$$

As $q\to\infty$, we have

$$(q')^q = \Big(\frac{q}{q-1}\Big)^q = \Big(1+\frac{1}{q-1}\Big)^q \to e,$$

and hence $\|M_0 f\|_{L^p}^p \leq e\|f\|_{L^p}^p$ for $p\in(0,\infty)$. □

The constant e is in fact optimal. One way of proving this is indicated in the exercises.

1.13. **Back to reality (from the dyadic world).** All the considerations above were about the dyadic maximal function. But in Classical Analysis, one is usually interested about the Hardy–Littlewood maximal function, where the supremum is taken over all cubes (or balls), not just the special dyadic ones. However, it turns out that it is reasonable straightforward to pass from one case to the other. Here it is important that our dyadic results are true for any dyadic system: not just the standard

$$\mathscr{D} := \{2^{-k}\big([0,1)^n+m\big) : k\in\mathbb{Z}, m\in\mathbb{Z}^n\},$$

but also the following perturbations:

$$\mathscr{D}^\alpha := \{2^{-k}\big([0,1)^n+m+(-1)^k\alpha\big) : k\in\mathbb{Z}, m\in\mathbb{Z}^n\}, \qquad \alpha\in\{0,\tfrac{1}{3}\}^n.$$

It is not difficult to convince oneself from a picture that each of these satisfies the dyadic property $Q\cap R\in\{Q,R,\varnothing\}$. And these 2^n dyadic systems $\mathscr{D}^\alpha$ are rich enough to approximate all cubes (with sides parallel to the axes) in $\mathbb{R}^n$:

Proposition 1.4. *If $Q\subset\mathbb{R}^n$ is any cube, there exists $\alpha\in\{0,\frac{1}{3}\}^n$ and $Q'\in\mathscr{D}^\alpha$ such that $Q\subset Q'$ and $\ell(Q')\leq 6\ell(Q)$.*

Proof. Case $n = 1$. Now cubes are just intervals. So let an interval I be given. There is exactly one integral power of 2 in the interval $(3\ell(I), 6\ell(I)]$, call it $2^{k(I)}$. Consider the set of all end-points of the intervals $I' \in \mathscr{D}^0 \cup \mathscr{D}^{1/3}$ of length $2^{k(I)}$. The distance of any two such points is at least $\frac{1}{3} \cdot 2^{k(I)} > \frac{1}{3} \cdot 3\ell(I) = \ell(I)$. But this means that I can contain at most one such end-point. If I does not contain an end-point of any $I' \in \mathscr{D}^0$ of length $2^{k(I)}$, then I is fully contained in one of these intervals (since they cover $\mathbb{R}$). And if I contains an end-point of some $I' \in \mathscr{D}^0$ of length $2^{k(I)}$, then I does not contain an end-point of any $I'' \in \mathscr{D}^{1/3}$ of length $2^{k(I)}$, and hence I must be contained in some $I'' \in \mathscr{D}^{1/3}$. So in any case $I \subset J \in \mathscr{D}^0 \cup \mathscr{D}^{1/3}$, where $\ell(J) = 2^{k(I)} \leq 6\ell(I)$.

Case $n > 1$. Now $Q = I_1 \times \cdots \times I_n$, where the I_j are intervals of the same length. By the one-dimensional case, for each I_j we find $\alpha_j \in \{0, \frac{1}{3}\}$ and $I'_j \in \mathscr{D}_1^{\alpha_j}$ (the subscript 1 refers to the one-dimensional dyadic system, to distinguish it from the n-dimensional version) such that $I_j \subset I'_j$, and $\ell(I'_j)$ is the unique power of two in the interval $(3\ell(I_j), 6\ell(I_j)] = (3\ell(Q), 6\ell(Q)]$ (hence it is the same number for each j). It follows that

$$Q = I_1 \times \cdots \times I_n \subset I'_1 \times \cdots \times I'_n =: Q' \in \mathscr{D}^\alpha, \qquad \alpha = (\alpha_1, \ldots, \alpha_n),$$

where $\ell(Q') \equiv \ell(I'_j) \leq 6\ell(I_j) \equiv 6\ell(Q)$. □

Let

$$M^\alpha f(x) := \sup_{Q \in \mathscr{D}^\alpha} 1_Q(x) \frac{1}{|Q|} \int_Q |f(y)| \, dy$$

be the dyadic maximal function related to $\mathscr{D}^\alpha$, and

$$M f(x) := \sup_{Q \text{ cube}} 1_Q(x) \frac{1}{|Q|} \int_Q |f(y)| \, dy$$

the Hardy–Littlewood maximal function.

Corollary 1.2. *We have the pointwise estimate.*

$$\max_{\alpha \in \{0, \frac{1}{3}\}} M^\alpha f(x) \leq M f(x) \leq 6^n \max_{\alpha \in \{0, \frac{1}{3}\}} M^\alpha f(x)$$

Proof. The first estimate is clear, since all $Q \in \mathscr{D}^\alpha$ are examples of all cubes. For the second estimate, consider $x \in \mathbb{R}^n$, a cube $Q \ni x$, and let $Q' \in \mathscr{D}^{\alpha(Q)}$ be a dyadic cube provided by the previous proposition with $\ell(Q') \leq 6\ell(Q)$ and $Q' \supset Q \ni x$. Then

$$\frac{1}{|Q|} \int_Q |f| \leq \frac{6^n}{|Q'|} \int_{Q'} |f| \leq 6^n M^{\alpha(Q)} f(x) \leq 6^n \max_{\alpha \in \{0, \frac{1}{3}\}^n} M^\alpha f(x).$$

Taking the supremum over all cubes $Q \ni x$ gives the claim. □

Corollary 1.3.

$$\|Mf\|_{L^p(w)} \leq 12^n \cdot N \cdot \|f\|_{L^p(w)}, \qquad N := 4e \cdot p' \cdot \big([w]_{A_p}[\sigma]_{A_\infty}\big)^{1/p}$$

Proof. Using the earlier results for each of the dyadic maximal functions, we have

$$\begin{aligned} \|Mf\|_{L^p(w)} &\leq 6^n \|\max_\alpha M^\alpha f\|_{L^p(w)} \leq 6^n \Big\| \sum_\alpha M^\alpha f \Big\|_{L^p(w)} \leq 6^n \sum_\alpha \|M^\alpha f\|_{L^p(w)} \\ &\leq 6^n \sum_\alpha N \|f\|_{L^p(w)} = 6^n \cdot 2^n \cdot N \|f\|_{L^p(w)}. \end{aligned}$$ □

The other results proven for the dyadic maximal function are easy to generalize to the Hardy–Littlewood maximal function by the same method. The previous proof should serve as sufficient illustration.

1.14. **Lerner's proof of the Muckenhoupt–Buckley theorem.** If we are interested just in the bound

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

instead of the more precise version with $\big([w]_{A_p}[\sigma]_{A_\infty}\big)^{1/p}$, an amazingly short argument due to Lerner is available. It is based on the following:

Proposition 1.5 (Lerner 2008 [15])**.**

$$(Mf)^{p-1} \leq \|w\|_{A_p} M_w[(M_\sigma(f\sigma^{-1}))^{p-1}w^{-1}].$$

Proof.

$$\begin{aligned}\Big(\frac{1}{|Q|}\int_Q f\Big)^{p-1} &= \frac{w(Q)}{|Q|}\Big(\frac{\sigma(Q)}{|Q|}\Big)^{p-1}\frac{|Q|}{w(Q)}\Big(\frac{1}{\sigma(Q)}\int_Q f\sigma^{-1}\sigma\Big)^{p-1} \\ &\leq \|w\|_{A_p}\frac{|Q|}{w(Q)}\inf_Q[M_\sigma(f\sigma^{-1})]^{p-1} \\ &\leq \|w\|_{A_p}\frac{1}{w(Q)}\int_Q[M_\sigma(f\sigma^{-1})]^{p-1}w^{-1}w.\end{aligned}$$

Taking the supremum over all dyadic $Q \ni x$ gives the assertion. □

After this pointwise bound, the norm inequality is just a question of applying the universal maximal function estimate:

Proof of Theorem 1.2 (Lerner 2008 [15]*).*

$$\begin{aligned}\|Mf\|_{L^p(w)} &= \|(Mf)^{p-1}\|_{L^{p'}(w)}^{1/(p-1)} \\ &\leq \|w\|_{A_p}^{1/(p-1)}\|M_w[(M_\sigma(f\sigma^{-1}))^{p-1}w^{-1}]\|_{L^{p'}(w)}^{1/(p-1)} \\ &\leq \|w\|_{A_p}^{1/(p-1)}\big(p\cdot\|(M_\sigma(f\sigma^{-1}))^{p-1}w^{-1}\|_{L^{p'}(w)}\big)^{1/(p-1)} \\ &= \|w\|_{A_p}^{1/(p-1)}p^{1/(p-1)}\|M_\sigma(f\sigma^{-1})\|_{L^p(\sigma)} \\ &\leq \|w\|_{A_p}^{1/(p-1)}p^{1/(p-1)}\cdot p'\cdot\|f\sigma^{-1}\|_{L^p(\sigma)} \\ &= p^{1/(p-1)}\cdot p'\cdot\|w\|_{A_p}^{1/(p-1)}\|f\|_{L^p(w)}.\end{aligned}$$

A standard calculus optimization shows that $p^{1/(p-1)} \leq e$, so altogether

$$\|Mf\|_{L^p(w)} \leq e\cdot p'\cdot\|w\|_{A_p}^{1/(p-1)}\|f\|_{L^p(w)};$$

in terms of the numerical constant, this is even slightly better than what we obtained by the previous method. □

It would be interesting to find a Lerner-type argument for the sharper Theorem 1.3.

2. Extrapolation theory

In this section, we deal with a surprising phenomenon in the weighted world: If an operator T is bounded on $L^r(w)$ for *some* $r \in (1,\infty)$ and for *all* $w \in A_r$, then it is in fact bounded on $L^p(w)$ for *all* $p \in (1,\infty)$ and for *all* $w \in A_p$. This is the content of the extrapolation theorem of Rubio de Francia (1984) [24], one of the highlights of the weighted theory. At the time it was proven, the interest in the quantitative weighted estimates had not yet started, but it turns out that a careful examination of the argument even provides a quantitative estimate. This was achieved by Dragičević, Grafakos, Pereyra and Petermichl (2005) [6]. This argument has been simplified by Duoandikoetxea (unpublished). The weighted boundedness of the maximal operator is a crucial ingredient of the proof. By using the sharp form of the Muckenhoupt–Buckley theorem, also the quantitative extrapolation can be made even slightly more precise: (This is again recent joint work of myself and Carlos Pérez [12].)

Consider weighted estimates of the form

$$\|Tf\|_{L^p(w)} \lesssim \sum [w]_{A_p}^{\alpha(p)} [w]_{A_\infty}^{\beta(p)} [w^{-1/(p-1)}]_{A_\infty}^{(p-1)\gamma(p)} \|f\|_{L^p(w)}, \tag{2.1}$$

where the sum is over a finite set of triples (α, β, γ). We now aim to extrapolate such bounds from one value of p to others.

Theorem 2.1. *Suppose that for some r and every $w \in A_r$, an operator T satisfies (2.1) for $p = r$. Then it satisfies the same bound for every $p \in (1, r)$ with*

$$\begin{aligned} \alpha(p) &= \alpha(r) + \tau(r)\frac{r-p}{p}, \qquad \tau := \alpha + \beta + \gamma \\ \beta(p) &= \beta(r), \\ \gamma(p) &= \gamma(r) + \tau(r)\frac{r-p}{p(p-1)}. \end{aligned}$$

In particular,

$$\tau(p) = \tau(r)\frac{r-1}{p-1}.$$

Lemma 2.1. *Let $g \in L^p(w)$ have norm 1. Let*

$$Rg := \sum_{k=0}^{\infty} \frac{2^{-k} M^k g}{\|M\|_{L^p(w) \to L^p(w)}^k}, \tag{2.2}$$

where $M^k g := M \circ \cdots \circ M(f)$ (k iterations of the maximal operator), and $M^0 g := g$. Then Rg satisfies the following properties:

$$|g| \le Rg, \qquad \|Rg\|_{L^p(w)} \le 2\|g\|_{L^p(w)} = 2, \qquad [Rg]_{A_1} \le 2\|M\|_{L^p(w) \to L^p(w)}.$$

Proof. The first one follows from the fact that all terms in the defining series are nonnegative, and the first one is g. For the second, by the triangle inequality in $L^p(w)$, we have

$$\|Rg\|_{L^p(w)} \le \sum_{k=0}^{\infty} 2^{-k} \frac{\|M^k g\|_{L^p(w)}}{\|M\|_{L^p(w) \to L^p(w)}^k} \le \sum_{k=0}^{\infty} 2^{-k} \|g\|_{L^p(w)} = 2\|g\|_{L^p(w)} = 2.$$

For the last property, use the triangle inequality (together with the sublinearity of the maximal operator) pointwise and change the summation variable to see that

$$\begin{aligned} M(Rg)(x) \le \sum_{k=0}^{\infty} \frac{2^{-k} M^{k+1} g(x)}{\|M\|_{L^p(w) \to L^p(w)}^k} &= 2\|M\|_{L^p(w) \to L^p(w)} \sum_{k=0}^{\infty} \frac{2^{-k-1} M^{k+1} g(x)}{\|M\|_{L^p(w) \to L^p(w)}^{k+1}} \\ 2\|M\|_{L^p(w) \to L^p(w)} \sum_{k=1}^{\infty} \frac{2^{-k} M^k g(x)}{\|M\|_{L^p(w) \to L^p(w)}^k} &\le 2\|M\|_{L^p(w) \to L^p(w)} Rg(x). \end{aligned}$$

Recalling the definition of A_1, this says exactly that $[Rg]_{A_1} \le 2\|M\|_{L^p(w) \to L^p(w)}^k$. □

Proof of Theorem 2.1. Fix some $p \in (1, r)$, $w \in A_p$, $f \in L^p(w)$ and $g := |f|/\|f\|_{L^p(w)}$. We will make extensive use of the auxiliary function $Rg \in A_1$ as provided by the previous lemma. Its A_1 property will in practice be applied via the following equivalent formulation:

$$\sup_Q (Rg)^{-1} \le [Rg]_{A_1} \langle Rg \rangle_Q^{-1}.$$

To prove the boundedness of T on $L^p(w)$, the quantity we need to estimate is

$$\begin{aligned} \|Tf\|_{L^p(w)} &= \Big(\int |Tf|^p (Rg)^{-(r-p)p/r} (Rg)^{(r-p)p/r} w \Big)^{1/p} \\ &\le \Big(\int |Tf|^r (Rg)^{-(r-p)} w \Big)^{1/r} \Big(\int (Rg)^p w \Big)^{1/p - 1/r} \\ &\le \|Tf\|_{L^r(W)} (2^p)^{1/p - 1/r}, \qquad W := (Rg)^{-(r-p)} w. \end{aligned}$$

By assumption, we have

$$\|Tf\|_{L^r(W)} \lesssim \sum [W]_{A_r}^{\alpha(r)} [W]_{A_\infty}^{\beta(r)} [W^{-1/(r-1)}]_{A_\infty}^{(r-1)\gamma(r)} \|f\|_{L^r(W)}, \tag{2.3}$$

where

$$\begin{aligned}\|f\|_{L^r(W)} &= \Big(\int |f|^r (Rf)^{-(r-p)} w\Big)^{1/r} \|f\|_{L^p(w)}^{(r-p)/r} \\ &\leq \Big(\int |f|^r |f|^{-(r-p)} w\Big)^{1/r} \|f\|_{L^p(w)}^{(r-p)/r} = \|f\|_{L^p(w)},\end{aligned}$$

so it remains to estimate the weight characteristics

$$[W]_{A_r}, \qquad [W]_{A_\infty}, \qquad [W^{-1/(r-1)}]_{A_\infty}.$$

So far, we do not even know if they are finite, so in principle (2.3) could be a useless statement saying only that $\|Tf\|_{L^r(W)} \leq \infty$.

Using $\sup_Q (Rg)^{-1} \leq [Rg]_{A_1} \langle Rg\rangle_Q^{-1}$ or Hölder's or Jensen's inequality where appropriate, we compute

$$\begin{aligned}\langle W\rangle_Q &= \langle (Rg)^{-(r-p)} w\rangle_Q \\ &\leq [Rg]_{A_1}^{r-p} \langle Rg\rangle_Q^{-(r-p)} \langle w\rangle_Q, \\ \langle W^{-1/(r-1)}\rangle_Q^{r-1} &= \langle (Rg)^{(r-p)/(r-1)} w^{-1/(r-1)}\rangle_Q^{r-1} \\ &\leq \langle Rg\rangle_Q^{r-p} \langle w^{-1/(p-1)}\rangle_Q^{p-1}, \\ \exp\langle -\log W\rangle_Q &= \big(\exp\langle \log(Rg)\rangle_Q\big)^{r-p} \exp\langle -\log w\rangle_Q \\ &\leq \langle Rg\rangle_Q^{r-p} \exp\langle -\log w\rangle_Q,\end{aligned}$$

and

$$\begin{aligned}&\big(\exp\langle -\log W^{-1/(r-1)}\rangle_Q\big)^{r-1} \\ &\quad = \big(\exp\langle \log(Rg)^{-1}\rangle_Q\big)^{r-p} \big(\exp\langle -\log w^{-1/(r-1)}\rangle_Q\big)^{r-1} \\ &\quad \leq [Rg]_{A_1}^{r-p} \langle Rg\rangle^{-(r-p)} \big(\exp\langle -\log w^{-1/(p-1)}\rangle_Q\big)^{p-1}.\end{aligned}$$

Multiplying the appropriate estimates and using the definition, we then have

$$\begin{aligned}&[W]_{A_r} \leq [Rg]_{A_1}^{r-p} [w]_{A_p}, \qquad [W]_{A_\infty} \leq [Rg]_{A_1}^{r-p} [w]_{A_\infty}, \\ &[W^{-1/(r-1)}]_{A_\infty}^{r-1} \leq [Rg]_{A_1}^{r-p} [w^{-1/(p-1)}]_{A_\infty}^{p-1}.\end{aligned}$$

Also recall that

$$[Rg]_{A_1} \leq 2\|M\|_{L^p(w)\to L^p(w)} \lesssim [w]_{A_p}^{1/p} [w^{-1/(p-1)}]_{A_\infty}^{1/p}.$$

Thus, with $\tau(r) := \alpha(r) + \beta(r) + \gamma(r)$, we obtain that

$$\begin{aligned}\|T\|_{L^p(w)\to L^p(w)} &\lesssim \sum [W]_{A_r}^{\alpha(r)} [W]_{A_\infty}^{\beta(r)} [W^{-1/(r-1)}]_{A_\infty}^{(r-1)\gamma(r)} \\ &\lesssim \sum [Rg]_{A_1}^{\tau(r)(r-p)} [w]_{A_p}^{\alpha(r)} [w]_{A_\infty}^{\beta(r)} [w^{-1/(p-1)}]_{A_\infty}^{(p-1)\gamma(r)} \\ &\lesssim \sum [w]_{A_p}^{\alpha(r)+\tau(r)(r-p)/p} [w]_{A_\infty}^{\beta(r)} [w^{-1/(p-1)}]_{A_\infty}^{(p-1)\gamma(r)+\tau(r)(r-p)/p},\end{aligned}$$

from which the asserted exponents can be easily read. □

A similar result also holds for exponents bigger than the original one:

Theorem 2.2. *Suppose that for some r and every $w \in A_r$, an operator T satisfies* (2.1) *for $p = r$. Then it satisfies the same bound for every $p \in (r,\infty)$ with*

$$\begin{aligned}\alpha(p) &= \frac{r-1}{p-1}\alpha(r) + \frac{p-r}{(p-1)p}\tau(r), \qquad \tau := \alpha+\beta+\gamma\\ \beta(p) &= \frac{r-1}{p-1}\beta(r) + \frac{p-r}{p}\tau(r),\\ \gamma(p) &= \frac{r-1}{p-1}\gamma(r).\end{aligned}$$

In particular,

$$\tau(p) = \tau(r).$$

Sketch of proof. The proof involves very similar ideas as the previous one, the main difference being the use of duality. Fix some $p \in (r,\infty)$, $w \in A_p$, $f \in L^p(w)$. By duality, we have

$$\|Tf\|_{L^p(w)} = \sup_{\substack{h\geq 0\\ \|h\|_{L^{p'}(w)=1}}} \int |Tf|hw.$$

We fix one such h, and try to bound the expression on the right.

Observe that the pointwise multiplication operators

$$h \mapsto wh : L^{p'}(w) \to L^{p'}(w^{1-p'}), \qquad g \mapsto \frac{1}{w}g : L^{p'}(w^{1-p'}) \to L^{p'}(w)$$

are isometric. Let R be as in the previous proof, except with p' and $\sigma = w^{1-p'}$ in place of p and w:

$$Rg := \sum_{k=0}^{\infty} \frac{2^{-k}M^k g}{\|M\|^k_{\mathscr{B}(L^{p'}(\sigma))}},$$

and $R'h := w^{-1}R(wh)$. Then one checks that

$$h \leq R'h, \qquad \|R'h\|_{L^{p'}(w)} \leq 2\|h\|_{L^{p'}(w)} = 2, \qquad [wR'h]_{A_1} \leq 2\|M\|_{\mathscr{B}(L^{p'}(\sigma))}.$$

The estimation then starts from

$$\begin{aligned}\int |Tf|hw \leq \int |Tf|(R'h)w &= \int |Tf|(R'h)^{(p-r)/[r(p-1)]}(R'h)^{(r-1)p/[r(p-1)]}w\\ &\leq \Big(\int |Tf|^r(R'h)^{(p-r)/(p-1)}w\Big)^{1/r}\Big(\int (R'h)^{p/(p-1)}w\Big)^{1/r'}\\ &\leq \|Tf\|_{L^r(W)}2^{p'/r'}, \qquad W := (R'h)^{(p-r)/(p-1)}w.\end{aligned}$$

The assumption on the boundedness of T can now be applied, and it remains to estimate the weight characteristics of W, in analogy with the previous proof. □

3. Lerner's "magic formula"

The topic of this section is a certain formula, discovered by Lerner (2010) [16], which provides very useful and precise information about a measurable function in terms of its "local oscillations". As such, this formula has nothing to do with weights, and could be part of a general course in Real Analysis. However, the formula was developed with applications to weighted norm inequalities in mind, and it has proven to be very powerful in this context. At the time of writing these lectures, it is not known whether the most general weighted inequalities can be proven by the use of Lerner's formula; but some important special cases can be derived from this formula in an elegant way, much simpler than any other known method.

Before stating and proving the actual formula, we need some preparations.

3.1. **The median of a function.** Let $f : Q \to \mathbb{R}$ be a measurable function. Here Q could be any set of finite positive measure, but later on it will mostly be a cube; hence the choice of the letter. The *median* of f on Q is any real number $m_f(Q)$ with the following two properties:

$$|Q \cap \{f > m_f(Q)\}| \leq \tfrac{1}{2}|Q|, \qquad |Q \cap \{f < m_f(Q)\}| \leq \tfrac{1}{2}|Q|.$$

It is left as an exercise to show that a median always exists; it need not be unique, but the set of all medians is always a closed interval.

The median can be thought of as a substitute for the average of the function on Q. An advantage is the fact that the median exists for any measurable function, whereas the average $\langle f \rangle_Q = |Q|^{-1} \int_Q f \, \mathrm{d}x$ requires f to be integrable. (The median is also more stable in the sense that it does not "see" singularities of a function which appear in sets of small measure, and it is often preferred in applied statistics: on the economy pages of a newspaper one can often read about the median prediction for the profit of a company.) A disadvantage is the possible non-uniqueness. Because of this, one needs to be somewhat careful when working with the median.

The following simple observation is handy for estimating the median:

Lemma 3.1. *The following claims hold for all medians $m_f(Q)$ and real numbers α:*

- *If $|Q \cap \{f \geq \alpha\}| > \frac{1}{2}|Q|$, then $m_f(Q) \geq \alpha$.*
- *If $|Q \cap \{f \leq \alpha\}| > \frac{1}{2}|Q|$, then $m_f(Q) \leq \alpha$.*
- *If $|Q \cap \{f = \alpha\}| > \frac{1}{2}|Q|$, then $m_f(Q) = \alpha$.*

Proof. Consider the first case, and suppose for contradiction that $m_f(Q) < \alpha$ is a median. So in particular $|Q \cap \{f > m_f(Q)\}| \leq \frac{1}{2}|Q|$, hence

$$\tfrac{1}{2}|Q| \leq |Q \cap \{f \leq m_f(Q)\}| \leq |Q \cap \{f < \alpha\}| < \tfrac{1}{2}|Q|,$$

a contradiction.

The second claim can be proven similarly, or by reduction to the first claim by considering $(-f, -\alpha)$ in place of (f, α). The third claim follows at once from the first and second. □

There is a median analogue of Lebesgue's differentiation theorem (which deals with averages):

Proposition 3.1 (Fujii 1991 [9])**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be measurable. Then for almost every $x \in \mathbb{R}^n$, we have*

$$\lim_{\substack{Q \ni x \\ |Q| \to 0}} m_f(Q) = f(x),$$

where the limit is along all cubes containing x, whose volume goes to zero, and along all medians of f on these cubes.

Proof. We introduce the auxiliary functions

$$s_k := \sum_{j \in \mathbb{Z}} \frac{j}{2^k} 1_{\{j2^{-k} \leq f < (j+1)2^{-k}\}} =: \sum_{j \in \mathbb{Z}} \frac{j}{2^k} 1_{E_{kj}}.$$

Then $s_k \leq f < s_k + 2^{-k}$ at every point. Observe that $\{E_{kj}\}_{j \in \mathbb{Z}}$ is a partition of $\mathbb{R}^n$ for every k. Now every $1_{E_{kj}} \in L^1_{\mathrm{loc}}$, so we may apply the usual Lebesgue differentiation theorem, to the result that

$$\frac{|Q \cap E_{kj}|}{|Q|} = \frac{1}{|Q|} \int_Q 1_{E_{kj}} \, \mathrm{d}x \underset{\substack{Q \ni x \\ |Q| \to 0}}{\longrightarrow} 1_{E_{kj}}(x) \tag{3.1}$$

for almost every $x \in \mathbb{R}^n$. Let us explicitly denote the exceptional null set by N_{kj}, so the above convergence holds for every $x \in N_{kj}^c$. Let $N := \bigcup_{k,j} N_{k,j}$. This is another null set, and the convergence (3.1) holds for every $x \in N^c$ and every $k, j \in \mathbb{Z}$.

We turn to the actual claim of the lemma. Written out in terms of the definition of the limit, it says that for almost every $x \in \mathbb{R}^n$,

$$\forall \varepsilon > 0 \ \exists \delta > 0 : \forall Q \ni x, |Q| < \delta \Rightarrow |m_f(Q) - f(x)| < \varepsilon, \tag{3.2}$$

where $m_f(Q)$ is any median of f on Q.

Let $x \in N^c$ and $\varepsilon > 0$ be given. We choose k so that $2^{-k} < \varepsilon$. There is a unique j (determined by x and k) such that $x \in E_{kj}$. By (3.1), we have the existence of a $\delta > 0$ such that

$$\frac{|Q \cap E_{kj}|}{|Q|} > \frac{2}{3} \qquad \forall Q \ni x \text{ such that } |Q| < \delta. \tag{3.3}$$

We now check that this same δ is also good for (3.2). So fix any cube Q as in (3.2). Recalling that $E_{kj} \subseteq \{f \geq j2^{-k}\} \cap \{f \leq (j+1)2^{-k}\}$, Lemma 3.1 and (3.3) imply that

$$j2^{-k} \leq m_f(Q) \leq (j+1)2^{-k}.$$

But we also have $j2^{-k} \leq f(x) < (j+1)2^{-k}$ since $x \in E_{kj}$, and thus $|m_f(Q) - f(x)| \leq 2^{-k} < \varepsilon$, and this is what we wanted to prove. □

3.2. **The decreasing rearrangement.** This is another concept, which can be defined for any measurable function f. We denote

$$f^*(t) := \inf\{\alpha \geq 0 : |\{|f| > \alpha\}| \leq t\} \qquad (\inf \varnothing := \infty).$$

We make the following observations:

- f^* is non-increasing.
  Indeed, if $s > t$, the condition $|\{|f| > \alpha\}| \leq s$ is easier to satisfy than $|\{|f| > \alpha\}| \leq t$. So the set of admissible α's is bigger for s, and the infimum of a bigger set is smaller.
- The set inside the infimum is of the form $[\alpha_0, \infty)$ (or $\varnothing$). Hence the infimum is reached as a minimum; in particular, $f^*(t)$ itself is an admissible value of α, so that
  $$|\{|f| > f^*(t)\}| \leq t. \tag{3.4}$$
  Indeed, if the set is nonempty and α belongs to this set, then every $\alpha' > \alpha$ satisfies
  $$|\{|f| > \alpha'\}| \leq |\{|f| > \alpha\}| \leq t,$$
  so also α' belongs to the set. So it remains to show that the infimum α_0 also belongs to the set. This follows from $\{|f| > \alpha_0\} = \bigcup_{j=1}^{\infty}\{|f| > \alpha_0 + j^{-1}\}$ and the monotonicity of the measure,
  $$|\{|f| > \alpha_0\}| = \lim_{j\to\infty} |\{|f| > \alpha_0 + j^{-1}\}| \leq t.$$
- We have $(f1_Q)^*(t) = \inf\{\alpha \geq 0 : |Q \cap \{|f| > \alpha\}| \leq t\}$.
  It suffices to check that $Q \cap \{|f| > \alpha\} = \{|1_Q f| > \alpha\}$. But this is easy to see.

A very useful connection between the median and the decreasing rearrangement is the following:

Lemma 3.2. *The following estimate holds for all* $\lambda \in (0, \frac{1}{2})$ *and all medians* $m_f(Q)$*:*

$$|m_f(Q)| \leq (f1_Q)^*(\lambda|Q|).$$

Proof. The right side is the infimum of $\{\alpha \geq 0 : |Q \cap \{|f| > \alpha\}| \leq \lambda|Q|\}$. It suffices to prove that if $\alpha < |m_f(Q)|$, then it is not in this set, for this implies that the infimum of the set is at least $|m_f(Q)|$.

So let $0 \leq \alpha < |m_f(Q)|$, where $m_f(Q)$ is a median. We prove that $|Q \cap \{|f| > \alpha\}| \geq \frac{1}{2}|Q| > \lambda|Q|$. Suppose first that $m_f(Q) > 0$. Then

$$|Q \cap \{|f| > \alpha\}| \geq |Q \cap \{f > \alpha\}| \geq |Q \cap \{f \geq m_f(Q)\}| = |Q| - |Q \cap \{f < m_f(Q)\}| \geq \tfrac{1}{2}|Q|.$$

If $m_f(Q) < 0$, then $\alpha < |m_f(Q)| = -m_f(Q)$ implies $-\alpha > m_f(Q)$, and hence

$$|Q \cap \{|f| > \alpha\}| \geq |Q \cap \{f < -\alpha\}| \geq |Q \cap \{f \leq m_f(Q)\}| = |Q| - |Q \cap \{f > m_f(Q)\}| \geq \tfrac{1}{2}|Q|.$$

So we are done; of course the case $m_f(Q) = 0$ is trivial from the beginning. □

Remark 3.1. The limiting case $\lambda = \frac{1}{2}$ of the previous estimate is more tricky. It is only true that *there exists* a median $m_f(Q)$ such that $|m_f(Q)| \leq (f1_Q)^*(\frac{1}{2}|Q|)$, but this need not be the case for all medians. Therefore we prefer to work with the more flexible estimate with $\lambda < \frac{1}{2}$, where we do not need to specify the choice of the median which we work with.

Related to this point, there is a slightly careless claim in Lerner's original paper that "it is easy to see that $|m_f(Q)| \leq (f1_Q)^*(\frac{1}{2}|Q|)$", and this is also used in the proof of his formula. We will need to slightly modify the proof to avoid the problems related to this estimate.

3.3. **Local oscillations of a function.** The following quantity should be understood as a measure of how well the function f can be approximated by a constant in the cube (or another set of finite positive measure) Q:

$$\omega_\lambda(f;Q) := \inf_c \big((f-c)1_Q\big)^*(\lambda|Q|).$$

We also define an associated maximal function

$$M^{\#}_{\lambda,Q}f(x) := \sup_{Q'\in\mathscr{D}(Q)} 1_{Q'}(x)\omega_\lambda(f;Q'),$$

where $\mathscr{D}(Q)$ is the collection of dyadic subcubes of Q, obtained by repeatedly dividing into 2^n equal cubes. (In this section, we always consider this usual dyadic structure of $\mathbb{R}^n$, no longer an abstract dyadic structure as before. In particular, we will exploit the fact the *dyadic parent* $\hat{Q}$ of Q [the minimal dyadic cube, which strictly contains Q] has measure $|\hat{Q}| = 2^n|Q|$.)

Finally, it will be convenient to have a variant of $\omega_\lambda(f;Q)$ involving the median rather than an arbitrary constant. We let

$$\tilde{\omega}_\lambda(f;Q) := \sup_{m_f(Q)} \big((f-m_f(Q))1_Q\big)^*(\lambda|Q|),$$

where the supremum (on purpose, not the infimum here!) is taken over all medians $m_f(Q)$ of f on Q. Then

Lemma 3.3. *For $\lambda \in (0,\frac{1}{2})$, we have*

$$\omega_\lambda(f;Q) \leq \tilde{\omega}_\lambda(f;Q) \leq 2\omega_\lambda(f;Q).$$

Before turning to the proof, we record a useful observation for later purposes as well. For a function f and a constant c, there holds

$$m_{f-c}(Q) = m_f(Q) - c \tag{3.5}$$

as an *equality of sets*: the set of all medians of $f-c$ is obtained by translating the set of all medians of f, as stated. This follows immediately from the definition.

Proof of Lemma 3.3. The first estimate is obvious. For the second, recalling the definitions, we need to show the following: for any median $m_f(Q)$ and any constant c, we have

$$\inf\{\alpha \geq 0 : |Q\cap\{|f-m_f(Q)|>\alpha\}| \leq \lambda|Q|\} \leq 2\big((f-c)1_Q\big)^*(\lambda|Q|).$$

And this means that $\alpha = 2\big((f-c)1_Q\big)^*(\lambda|Q|)$ should be an admissible value in the set inside the infimum, i.e., that

$$\big|Q\cap\big\{|f-m_f(Q)| > 2\big((f-c)1_Q\big)^*(\lambda|Q|)\big\}\big| \leq \lambda|Q|. \tag{3.6}$$

Let us prove this. By triangle inequality, (3.5) and Lemma 3.2, we have

$$\begin{aligned} |f-m_f(Q)| &\leq |f-c| + |m_f(Q)-c| \\ &= |f-c| + |m_{f-c}(Q)| \leq |f-c| + \big((f-c)1_Q\big)^*(\lambda|Q|). \end{aligned}$$

(Here, given a median $m_f(Q)$ of f, we have that $m_f(Q)-c$ is a median of $f-c$, and it is important that the bound of Lemma 3.2 holds for all these medians; this is ensured by $\lambda < \frac{1}{2}$.) Hence

$$\begin{aligned} &\big|Q\cap\big\{|f-m_f(Q)| > 2\big((f-c)1_Q\big)^*(\lambda|Q|)\big\}\big| \\ &\leq \big|Q\cap\big\{|f-c| > \big((f-c)1_Q\big)^*(\lambda|Q|)\big\}\big| \leq \lambda|Q| \end{aligned}$$

by (3.4) in the last step. This proves (3.6), and hence the Lemma. □

3.4. **Lerner's formula.** Now we are fully prepared for the main result of this section:

Theorem 3.1 (Lerner 2010 [16])**.** *Let $Q_0 \subset \mathbb{R}^n$ be a cube, and $f : Q_0 \to \mathbb{R}$ a measurable function. Then there exist dyadic subcubes Q_j^k of Q_0 such that for almost every $x \in Q_0$,*

$$|f(x) - m_f(Q_0)| \leq 4M^{\#}_{1/4,Q_0} f(x) + 4\sum_{k=1}^{\infty}\sum_j \omega_{2^{-n-2}}(f;\hat{Q}_j^k)\cdot 1_{Q_j^k}(x),$$

where $m_f(Q_0)$ is any median of f. Moreover,

- *$\{Q_j^k\}_j$ is a disjoint collection for any fixed k,*
- *the sets $\Omega_k := \bigcup_j Q_j^k$ satisfy $\Omega_{k+1} \subset \Omega_k$, and*
- *$|Q_j^k \cap \Omega_{k+1}| \leq \frac{1}{2}|Q_j^k|$.*

Proof. Fix a median $m_f(Q_0)$, and let $f_1 := f - m_f(Q_0)$. Fujii's lemma ensures that

$$|f_1(x)| = \lim_{\substack{Q\in\mathscr{D}(Q_0)\\ Q\ni x, |Q|\to 0}} |m_{f_1}(Q)| \leq \sup_{Q\in\mathscr{D}(Q_0)} 1_Q(x) \sup_{m_{f_1}(Q)} |m_{f_1}(Q)| =: m^{\Delta}_{Q_0} f_1(x), \tag{3.7}$$

where the last equality simply defines yet another maximal operator $m^{\Delta}_{Q_0}$, and $\sup_{m_{f_1}(Q)}$ is the supremum over all medians. We let

$$\Omega_1 := Q_0 \cap \{m^{\Delta}_{Q_0} f_1 > \tilde{\omega}_{1/4}(f;Q_0)\} = \bigcup_j Q_j^1,$$

where Q_j^1 are the maximal cubes in $\mathscr{D}(Q_0)$ such that $\sup_{m_{f_1}(Q)} |m_{f_1}(Q)| > \tilde{\omega}_{1/4}(f;Q_0)$. By Lemma 3.2, we have $|m_{f_1}(Q_0)| \leq \omega_{1/4}(f;Q_0) \leq \tilde{\omega}_{1/4}(f;Q_0)$, so Q_0 itself cannot be among these cubes; thus all Q_j^1 are strict subcubes of Q_0, so that also $\hat{Q}_j^1 \in \mathscr{D}(Q_0)$.

We claim that

$$|\Omega_1| = \sum_j |Q_j^1| \leq \tfrac{1}{2}|Q_0|. \tag{3.8}$$

To see this, we make the following auxiliary considerations:

$$\begin{aligned}
(f_1 1_{Q_0})^*(\tfrac{1}{4}|Q_0|) &= \big((f - m_f(Q_0))1_{Q_0}\big)^*(\tfrac{1}{4}|Q_0|) \qquad \text{(definition of } f_1) \\
&\leq \tilde{\omega}_{1/4}(f;Q_0) \qquad \text{(definition of } \tilde{\omega}_{1/4}) \\
&< \sup_{m_{f_1}(Q_j^1)} |m_{f_1}(Q_j^1)| \qquad \text{(definition of } Q_j^1) \\
&\leq (f_1 1_{Q_j^1})^*(\lambda|Q_j^1|) \qquad \text{(Lemma 3.2, for any } \lambda \in (0,\tfrac{1}{2})) \\
&= \inf\big\{\alpha \geq 0 : |Q_j^1 \cap \{|f_1| > \alpha\}| \leq \lambda|Q_j^1|\big\} \qquad \text{(definition of } f^*).
\end{aligned}$$

That the infimum is strictly bigger than the left side means that $\alpha = (f_1 1_{Q_0})^*(\frac{1}{4}|Q_0|)$ is not an admissible value; hence we have the opposite inequality

$$\big|Q_j^1 \cap \big\{|f_1| > (f_1 1_{Q_0})^*(\tfrac{1}{4}|Q_0|)\big\}\big| > \lambda|Q_j^1|$$

We sum this over all Q_j^1, recalling that these are disjoint (being maximal), and all contained in Q_0:

$$\begin{aligned}
\lambda\sum_j |Q_j^1| &< \sum_j \big|Q_j^1 \cap \big\{|f_1| > (f_1 1_{Q_0})^*(\tfrac{1}{4}|Q_0|)\big\}\big| \\
&\leq \big|Q_0 \cap \big\{|f_1| > (f_1 1_{Q_0})^*(\tfrac{1}{4}|Q_0|)\big\}\big| \leq \tfrac{1}{4}|Q_0| \qquad \text{(by (3.4))}.
\end{aligned}$$

Hence $|\Omega_1| \leq (4\lambda)^{-1}|Q_0|$ for any $\lambda \in (0,\frac{1}{2})$, and letting $\lambda \to \frac{1}{2}$ gives the claimed estimate (3.8).

Choosing any medians $m_{f_1}(Q_j^1)$, we can now write the identity

$$\begin{aligned}
(f - m_f(Q_0))1_{Q_0} = f_1 1_{Q_0} &= f_1 1_{Q_0\setminus\Omega_1} + \sum_j m_{f_1}(Q_j^1)\cdot 1_{Q_j^1} + \sum_j (f_1 - m_{f_1}(Q_j^1))\cdot 1_{Q_j^1} \\
&= f_1 1_{Q_0\setminus\Omega_1} + \sum_j m_{f_1}(Q_j^1)\cdot 1_{Q_j^1} + \sum_j (f - m_f(Q_j^1))\cdot 1_{Q_j^1},
\end{aligned} \tag{3.9}$$

observing in the last step that

$$\begin{aligned} f_1 - m_{f_1}(Q_j^1) &= (f - m_f(Q_0)) - m_{f-m_f(Q_0)}(Q_j^1) \\ &= (f - m_f(Q_0)) - (m_f(Q_j^1) - m_f(Q_0)) = f - m_f(Q_j^1) \end{aligned}$$

for some median $m_f(Q_j^1)$.

For the first two terms on the right of (3.9), we have the following estimates:

$$\begin{aligned} |f_1| 1_{Q_0 \setminus \Omega_1} &\leq m^{\Delta}_{Q_0} f_1 \cdot 1_{Q_0\setminus\Omega_1} && \text{(by (3.7))} \\ &\leq \tilde{\omega}_{1/4}(f; Q_0) \cdot 1_{Q_0\setminus\Omega_1} && \text{(definition of } \Omega_1) \\ &\leq 2\omega_{1/4}(f; Q_0) \cdot 1_{Q_0\setminus\Omega_1} && \text{(Lemma 3.3)} \\ &\leq 2M^{\#}_{1/4,Q_0} f \cdot 1_{Q_0\setminus\Omega_1} && \text{(definition of } M^{\#}_{1/4,Q_0}), \end{aligned}$$

and

$$\begin{aligned} &|m_{f_1}(Q_j^1)| \\ &\leq |m_{f_1}(Q_j^1) - m_{f_1}(\hat{Q}_j^1)| + |m_{f_1}(\hat{Q}_j^1)| \\ &= |m_{f - m_f(\hat{Q}_j^1)}(Q_j^1)| + |m_{f_1}(\hat{Q}_j^1)| \qquad \text{(by (3.5), applied twice)} \\ &\leq \big((f - m_f(\hat{Q}_j^1))1_{Q_j^1}\big)^*(\tfrac{1}{4}|Q_j^1|) + \tilde{\omega}_{1/4}(f; Q_0) \qquad \text{(Lemma (3.2); definition of } Q_j^1) \\ &\leq \big((f - m_f(\hat{Q}_j^1))1_{\hat{Q}_j^1}\big)^*(2^{-n-2}|\hat{Q}_j^1|) + \tilde{\omega}_{1/4}(f; Q_0) \qquad (g^* \leq h^* \text{ if } g \leq h;\ |Q_j^1| = 2^{-n}|\hat{Q}_j^1|) \\ &\leq \tilde{\omega}_{2^{-n-2}}(f; \hat{Q}_j^1) + \tilde{\omega}_{1/4}(f; Q_0) \qquad \text{(definition of } \tilde{\omega}) \\ &\leq 2\omega_{2^{-n-2}}(f; \hat{Q}_j^1) + 2\omega_{1/4}(f; Q_0) \qquad \text{(Lemma 3.3)} \end{aligned}$$

And the summands in the third term on the right of (3.9) are of the same form as the left side $(f - m_f(Q_0))1_{Q_0}$ which we started from. So the point is now that we can iterate the same algorithm on these smaller cubes.

Let us check that such iteration gives the following identity:

$$(f - m_f(Q_0))1_{Q_0} = \sum_{k=1}^K f_k 1_{\Omega_{k-1}\setminus\Omega_k} + \sum_{k=1}^K \sum_j m_{f_k}(Q_j^k) \cdot 1_{Q_j^k} + \sum_j (f - m_f(Q_j^K)) \cdot 1_{Q_j^K}, \quad (3.10)$$

where $\Omega_0 := Q_0$, $\Omega_k := \bigcup_j Q_j^k$, and each f_k is of the form

$$f_k = \sum_i (f - m_f(Q_i^{k-1}))1_{Q_i^{k-1}}$$

(where the summation contains just one term for $k - 1 = 0$, with $Q_i^0 = Q_0$), and all Q_j^k are contained in some Q_i^{k-1}, in such a way that $|Q_i^{k-1} \cap \Omega_k| \leq \frac{1}{2}|Q_i^{k-1}|$. Moreover,

$$\begin{aligned} |f_k| \cdot 1_{\Omega_{k-1}\setminus\Omega_k} &\leq 2M^{\#}_{1/4;Q_0} f \cdot 1_{\Omega_{k-1}\setminus\Omega_k}, \\ |m_{f_k}(Q_j^k)| &\leq 2\omega_{2^{-n-2}}(f; \hat{Q}_j^k) + 2\omega_{1/4}(f; Q_i^{k-1}) \qquad \forall Q_j^k \subset Q_i^{k-1}. \end{aligned} \quad (3.11)$$

Indeed, for $K = 1$, these are just the identity (3.9) and the subsequent estimates for $|f_1|$ and $m_{f_1}(Q_j^1)$. Assuming (3.10) for some K, we check it for $K + 1$, thereby proving this identity by induction. To this end, we only need to apply (3.9) to each Q_j^K in place of Q_0, producing cubes $Q_i^{K+1} \subset Q_j^K$ (in place of Q_j^1) whose union $\Omega_{K+1} = \bigcup_i Q_i^{K+1}$ satisfies $|Q_i^K \cap \Omega_{K+1}| \leq \frac{1}{2}|Q_j^K|$ and

$$\begin{aligned} &(f - m_f(Q_j^K)) \cdot 1_{Q_j^K} =: f_{K+1} \cdot 1_{Q_j^K} \\ &= f_{K+1} \cdot 1_{Q_j^K\setminus\Omega_{K+1}} + \sum_{i:Q_i^{K+1}\subset Q_j^K} m_{f_{K+1}}(Q_i^{K+1}) + \sum_{i:Q_i^{K+1}\subset Q_j^K} (f - m_f(Q_i^{K+1})) \cdot 1_{Q_i^{K+1}}, \end{aligned} \quad (3.12)$$

where

$$\begin{aligned} |f_{K+1}| \cdot 1_{Q_j^K\setminus\Omega_{K+1}} &\leq 2M^{\#}_{1/4;Q_j^K} f \cdot 1_{Q_j^K\setminus\Omega_{K+1}} \leq 2M^{\#}_{1/4;Q_0} f \cdot 1_{Q_j^K\setminus\Omega_{K+1}} \\ |m_{f_{K+1}}(Q_i^{K+1})| &\leq 2\omega_{2^{-n-2}}(f; \hat{Q}_i^{K+1}) + 2\omega_{1/4}(f; Q_j^K). \end{aligned}$$

Summing (3.12) over j and reorganizing $\sum_j \sum_{i:Q_i^{K+1}\subset Q_j^K} = \sum_i$, we have

$$\begin{aligned}
&\sum_j (f - m_f(Q_j^K)) \cdot 1_{Q_j^K} \\
&\qquad = f_{K+1} \cdot 1_{\Omega_K \setminus \Omega_{K+1}} + \sum_i m_{f_{K+1}}(Q_i^{K+1}) + \sum_i (f - m_f(Q_i^{K+1})) \cdot 1_{Q_i^{K+1}},
\end{aligned}$$

which may be substituted to (3.10) to produce a similar formula with $K+1$ in place of K. This proves (3.10) for all K by induction.

We next want to pass to the limit $K \to \infty$ in (3.10). Observe that the last term is supported on $\bigcup_j Q_j^K = \Omega_K$, where

$$\Omega_k \subset \Omega_{k-1} \subset \cdots \subset Q_0, \qquad |\Omega_k| \leq 2^{-1}|\Omega_{k-1}| \leq \cdots \leq 2^{-k}|Q_0|.$$

Hence $\Omega_\infty := \bigcap_{k=0}^\infty \Omega_k$ has measure zero. If $x \notin \Omega_\infty$, then $x \notin \Omega_K$ for all K bigger than some $K(x)$, and hence the last term in (3.10) vanishes for all these K. That is, for almost every $x \in Q_0$ (namely, $x \in Q_0 \setminus \Omega_\infty$), we have

$$(f(x) - m_f(Q_0)) \cdot 1_{Q_0}(x) = \sum_{k=1}^\infty \Big(f_k(x) \cdot 1_{\Omega_{k-1}\setminus\Omega_k}(x) + \sum_j m_{f_k}(Q_j^k) \cdot 1_{Q_j^k}(x) \Big), \tag{3.13}$$

where the existence of this $\sum_{k=0}^\infty = \lim_{K\to\infty} \sum_{k=0}^K$ follows from the identity (3.10) and the just established convergence of the last term to zero.

We now estimate by absolute values, using the bounds (3.11). Here it is convenient to reorganize the latter sum as $\sum_j = \sum_i \sum_{j:Q_j^k \subset Q_i^{k-1}}$ again. This gives

$$\begin{aligned}
&|(f - m_f(Q_0)) \cdot 1_{Q_0}| \\
&\leq \sum_{k=1}^\infty \Big(2M^\#_{1/4;Q_0} f \cdot 1_{\Omega_{k-1}\setminus\Omega_k} + \sum_i \sum_{j:Q_j^k\subset Q_i^{k-1}} 2\big(\omega_{2^{-n-2}}(f;\hat{Q}_j^k) + \omega_{1/4}(f;Q_i^{k-1})\big)1_{Q_j^k} \Big) \\
&\leq 2M^\#_{1/4;Q_0} f \cdot 1_{Q_0} + 2\sum_{k,j} \omega_{2^{-n-2}}(f;\hat{Q}_j^k) \cdot 1_{Q_j^k} + 2\sum_{k=1}^\infty \sum_i \omega_{1/4}(f;Q_i^{k-1}) \cdot 1_{Q_i^{k-1}},
\end{aligned}$$

where we used $\sum_{k=1}^\infty 1_{\Omega_{k-1}\setminus\Omega_k} \leq 1_{Q_0}$ and $\sum_{i:Q_j^k\subset Q_i^{k-1}} 1_{Q_j^k} \leq 1_{Q_i^{k-1}}$ by the disjointness of these sets. The first two terms are of the asserted form, and it remains to investigate the last one.

For $k = 1$, the i-sum contains just the single term

$$\omega_{1/4}(f;Q_0) \cdot 1_{Q_0} \leq M^\#_{1/4,Q_0} f \cdot 1_{Q_0}.$$

And for $k \geq 2$, we have

$$\begin{aligned}
\omega_{1/4}(f;Q_i^{k-1}) &= \inf_c (1_{Q_i^{k-1}}(f-c))^*(\tfrac{1}{4}|Q_i^{k-1}|) \\
&\leq \inf_c (1_{\hat{Q}_i^{k-1}}(f-c))^*(2^{-n-2}|\hat{Q}_i^{k-1}|) = \omega_{2^{-n-2}}(f;\hat{Q}_i^{k-1});
\end{aligned}$$

hence, by this estimate and relabeling the summation variables,

$$\sum_{k=2}^\infty \sum_i \omega_{1/4}(f;Q_i^{k-1}) \cdot 1_{Q_i^{k-1}} \leq \sum_{k=1}^\infty \sum_j \omega_{2^{-n-2}}(f;\hat{Q}_j^k) \cdot 1_{Q_j^k}.$$

Putting these estimates together, we have proven Lerner's formula exactly as stated. □

4. A dyadic model of singular integrals: the dyadic shifts

We now turn to the investigation of weighted inequalities for the second class of operators of interest, the singular integral operators. Just like with maximal operators, our approach will proceed via a dyadic model first. Unlike with the maximal operators, the connection of this dyadic model to the "real world" of classical operators is not so obvious at the first sight, and establishing this connection will depend on nontrivial representation theorems, which we take up at a later point. For the moment, we just introduce and study the dyadic model operators on their own right.

4.1. **Dyadic cubes, Haar functions.** Again, we will need more precisely specified dyadic cubes than just the general property $Q \cap R \in \{Q, R, \varnothing\}$. The exact coordinate representation is still not important, but we will impose the following structural conditions:

$$\mathscr{D} = \bigcup_{k\in\mathbb{Z}} \mathscr{D}_k,$$

where each $\mathscr{D}_k$ is a partition of $\mathbb{R}^d$ (we now use d for dimension, liberating n for other purposes) into axes-parallel cubes of sidelength $\ell(Q) = 2^{-k}$, and each $Q \in \mathscr{D}_k$ is the disjoint union of 2^d cubes $Q' \in \mathscr{D}_{k+1}$, called its children. Likewise, each cubes has a unique parent, which we denote by $Q^{(1)} = \hat{Q}$, and the jth generation ancestor is defined inductively by $Q^{(j)} := \widehat{Q^{(j-1)}}$.

We have the averaging (or in probabilistic language: conditional expectation) operators

$$\mathbb{E}_k f := \sum_{Q\in\mathscr{D}_k} 1_Q \langle f\rangle_Q,$$

which are pointwise dominated by the maximal function Mf. If $f \in L^p(\mathbb{R}^d)$ for some $p \in [1,\infty)$, one easily checks that $\langle f\rangle_Q \to 0$ as $\ell(Q) \to \infty$. Lebesgue's differentiation theorem tells that $\langle f\rangle_Q \to f(x)$ when $\ell(Q) \to 0$ and $Q \ni x$, for all $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ almost every x. Hence, for $f \in L^p(\mathbb{R}^d)$, $p \in [1,\infty)$, we have the pointwise convergence

$$\mathbb{E}_k f \to \begin{cases} 0, & k \to -\infty \\ f, & k \to +\infty. \end{cases}$$

For $p \in (1,\infty)$, the domination by $Mf \in L^p(\mathbb{R}^d)$ yields the corresponding norm convergence in $L^p(\mathbb{R}^d)$. Finally, this can be expressed differently as

$$f = f - 0 = \lim_{\substack{a\to-\infty\\ b\to+\infty}} (\mathbb{E}_b f - \mathbb{E}_a f) = \lim_{\substack{a\to-\infty\\ b\to+\infty}} \sum_{k=a}^{b-1} (\mathbb{E}_{k+1} f - \mathbb{E}_k f) =: \sum_{k=-\infty}^{\infty} \mathbb{D}_k f.$$

We further expand

$$\mathbb{E}_k f = \sum_{Q\in\mathscr{D}_k} 1_Q \mathbb{E}_k f =: \sum_{Q\in\mathscr{D}_k} \mathbb{E}_Q f, \qquad \mathbb{D}_k f = \sum_{Q\in\mathscr{D}_k} 1_Q \mathbb{D}_k f =: \sum_{Q\in\mathscr{D}_k} \mathbb{D}_Q f.$$

It is straightforward to check that all operators $\mathbb{E}_k, \mathbb{D}_k, \mathbb{E}_Q, \mathbb{D}_Q$ are projections, and also that $\mathbb{D}_Q \mathbb{D}_R = 0$ for $Q \neq R$.

For $Q \in \mathscr{D}_k$, we have

$$\mathbb{D}_Q f = 1_Q(\mathbb{E}_{k+1} f - \mathbb{E}_k f) = \sum_{Q' \text{ child of } Q} 1_{Q'} \langle f\rangle_{Q'} - 1_Q \langle f\rangle_Q.$$

We can then identify the range $\mathsf{R}(\mathbb{D}_Q)$ as

$$\begin{aligned} \mathsf{R}(\mathbb{D}_Q) &= \Big\{ f : \operatorname{supp} f \subseteq Q, f \text{ constant of children of } Q, \int_Q f = 0 \Big\} \\ &= \Big\{ f = \sum_{Q' \text{ child of } Q} 1_{Q'} a_{Q'} : a_{Q'} \text{ constant }, \sum_{Q' \text{ child of } Q} a_Q = 0 \Big\}. \end{aligned}$$

Indeed, it is immediate from the formula of $\mathbb{D}_Q f$, that functions in the range of $\mathbb{D}_Q f$ have the asserted form. Conversely, it is easy to check that if f is of the form on the right of the previous equality, then $\mathbb{D}_Q f = f$, so that $f \in \mathsf{R}(\mathbb{D}_Q)$.

From the second formula for $\mathsf{R}(\mathbb{D}_Q)$, it is immediate that $\dim \mathsf{R}(\mathbb{D}_Q) = 2^d - 1$, as there are 2^d variables $a_{Q'}$ with one linear constraint. We now describe a convenient basis for this linear space. In dimension $d = 1$, consider the Haar functions

$$h_I^0 := |I|^{-1/2} 1_I, \qquad h_I^1 := |I|^{-1/2}\big(1_{I_{\text{left}}} - 1_{I_{\text{right}}}\big),$$

where I_{left} and I_{right} are the left and right halves of the interval I. Then, in d dimensions, let

$$h_Q^\varepsilon(x) := \prod_{j=1}^d h_{I_j}^{\varepsilon_j}(x_j), \qquad \varepsilon = (\varepsilon_1, \ldots, \varepsilon_d) \in \{0,1\}^d, \quad x = (x_1, \ldots, x_d) \in \mathbb{R}^d, \quad Q = I_1 \times \cdots I_d.$$

By Fubini's theorem, we have

$$\int h_Q^\varepsilon h_Q^{\varepsilon'} = \prod_{j=1}^d \int h_{I_j}^{\varepsilon_j} h_{I_j}^{\varepsilon_j'}.$$

If $\varepsilon \neq \varepsilon'$, then $\varepsilon_j \neq \varepsilon'$ for some j, and hence $\{\varepsilon_j, \varepsilon_j'\} = \{0,1\}$. Hence $h_{I_j}^{\varepsilon_j} h_{I_j}^{\varepsilon_j'} = |I_j|^{-1/2} h_{I_j}^1$, which clearly integrates to zero. It is also immediate that $(h_{I_j}^{\varepsilon_j})^2 = |I_j|^{-1} \cdot 1_{I_j}$, which integrates to 1. So we have

$$\int h_I^\varepsilon h_I^{\varepsilon'} = \delta_{\varepsilon,\varepsilon'}.$$

With $\varepsilon' = 0$, this shows that all h_I^ε, $\varepsilon \neq 0$, have zero integral, and for two different $\varepsilon, \varepsilon' \in \{0,1\}^d \setminus \{0\}$, we see that we have produced $2^d - 1$ orthonormal, hence linearly independent, functions in $\mathsf{R}(\mathbb{D}_Q)$.

4.2. **Definition and basic properties of dyadic shifts.** A dyadic shift with parameters (m,n) is an operator of the following form:

$$\text{Ш} f = \sum_{K \in \mathscr{D}} A_K f, \qquad A_K f = \sum_{\substack{I,J \in \mathscr{D},\ I,J \subseteq K \\ \ell(I) = 2^{-m}\ell(K) \\ \ell(J) = 2^{-n}\ell(K)}} a_{IJK} \langle f, h_I \rangle h_J,$$

where the a_{IJK} are constants with the normalization

$$|a_{IJK}| \leq \frac{\sqrt{|I||J|}}{|K|},$$

and each h_I is one of the Haar functions h_I^ε, $\varepsilon \in \{0,1\}^d \setminus \{0\}$, as just defined.

The symbol Ш is the Cyrillic capital letter 'sha' (for 'shift'). As an operator, Ш is a sum of component A_K associated to all dyadic cubes K (intuitively, to all positions and length scales). And each a_K has an expression in terms of the Haar functions h_I and h_J, where $I, J \subseteq K$ are dyadic subintervals, smaller than K by a fixed factor determined by the shift parameters. We record some basic observations:

Lemma 4.1.

$$|A_K f| \leq 1_K \frac{1}{|K|} \int_K |f|.$$

Proof. Using the bound for a_{IJK} and $|h_I| \leq 1_I / \sqrt{|I|}$, we have

$$\begin{aligned}|A_K f| &\leq \sum_{\substack{I,J \in \mathscr{D},\ I,J \subseteq K \\ \ell(I) = 2^{-m}\ell(K) \\ \ell(J) = 2^{-n}\ell(K)}} \frac{\sqrt{|I||J|}}{|K|} \cdot \int_I |f| \frac{1}{\sqrt{|I|}} \cdot \frac{1_J}{\sqrt{|J|}} \\ &= \frac{1}{|K|} \sum_{\substack{I \in \mathscr{D},\ I \subseteq K \\ \ell(I) = 2^{-m}\ell(K)}} \int_I |f| \sum_{\substack{J \in \mathscr{D},\ J \subseteq K \\ \ell(J) = 2^{-n}\ell(K)}} 1_J = \frac{1}{|K|} \int_K |f| \cdot 1_K,\end{aligned}$$

since dyadic cubes of fixed sidelength are disjoint. □

Corollary 4.1.

$$\|A_K f\|_p \leq \|f\|_p, \qquad \forall p \in [1,\infty].$$

Proof.

$$\|A_K f\|_p \leq \Big\|1_K \frac{1}{|K|}\int_K |f|\Big\|_p = |K|^{1/p}\frac{1}{|K|}\|1_K f\|_1 \leq |K|^{1/p}\frac{1}{|K|}|K|^{1/p'}\|f\|_p = \|f\|_p.$$ □

Lemma 4.2.

$$\|Ш f\|_2 \leq \|f\|_2.$$

Proof. We use the orthonormality of the Haar functions. Let

$$\mathscr{H}_K^m := \operatorname{span}\{h_I : I \subseteq K, \ell(I) = 2^{-m}\ell(K)\},$$

and let $\mathbb{P}_K^m$ be the orthogonal projection of L^2 onto this subspace. For a fixed m, these spaces are orthogonal, as K ranges over $\mathscr{D}$.

We have $\langle f, h_I\rangle = \langle \mathbb{P}_K^m f, h_I\rangle$ for all I appearing in A_K, and hence $A_K f = A_K\mathbb{P}_K^m f$. Also, $h_J = \mathbb{P}_K^n h_J$ for all J appearing in A_K, and hence $A_K f = \mathbb{P}_K^n A_K f$. We can apply these identities and Pythagoras' theorem to the result that

$$\begin{aligned}\|Ш f\|_2 = \Big\|\sum_{K\in\mathscr{D}}\mathbb{P}_K^n A_K \mathbb{P}_K^m f\Big\|_2 &= \Big(\sum_{K\in\mathscr{D}}\|\mathbb{P}_K^n A_K\mathbb{P}_K^m f\|_2^2\Big)^{1/2} \\ &\leq \Big(\sum_{K\in\mathscr{D}}\|\mathbb{P}_K^m f\|_2^2\Big)^{1/2} \leq \|f\|_2,\end{aligned}$$

where we used the L^2 boundedness of $A_K$in the second-to-last step. □

Remark 4.1. For the representation of general singular integrals, it is also necessary to consider generalized dyadic shifts, where it is allowed that the h_I or h_J may be of the non-cancellative type, i.e., $h_I^0 = |I|^{-1/2}\cdot 1_I$ or $h_J^0 = |J|^{-1/2}\cdot 1_J$. In this case, the result of the previous lemma is not automatically true, unless stronger conditions on the coefficients a_{IJK} are imposed. On the other hand, many results about the shifts may be proven by allowing the non-cancellative Haar functions, but requiring $\|Ш f\|_2 \leq \|f\|_2$ as an assumption.

4.3. **The weak-type** $(1,1)$ **estimate for dyadic shifts.** In practice, the weighted norm estimates for singular-integral type operators always rely on some information about their behaviour in the unweighted L^1 space.

Proposition 4.1. *Let* Ш *be a dyadic shift of parameters* (m,n)*. Then*

$$\|Ш f\|_{L^{1,\infty}} \lesssim (1+m)\|f\|_{L^1}.$$

Proof. We need to prove that $|\{|f| > \lambda\}| \lesssim (1+m)\|f\|_1/\lambda$. This is a classical-style argument based on the Calderón–Zygmund decomposition. Let $\mathscr{B}$ be the collection of maximal dyadic cubes $L \in \mathscr{D}$ with the property that $\langle |f|\rangle_L > \lambda$. Note that $\langle |f|\rangle_L \leq |L|^{-1}\|f\|_1 \to 0$ as $|L| \to \infty$, so this property cannot hold for arbitrarily large cubes L, and the set $\mathscr{B}$ of maximal cubes is well defined. We write the Calderón–Zygmund decomposition (observe the similarity to Lerner's decomposition, which used the median $m_f(L)$ instead of the mean $\langle f\rangle_L$):

$$f = \Big(f\cdot 1_{\mathbb{R}^d\setminus\bigcup\mathscr{B}} + \sum_{L\in\mathscr{B}} 1_L\langle f\rangle_L\Big) + \sum_{L\in\mathscr{B}} 1_L(f - \langle f\rangle_L) =: g + \sum_{L\in\mathscr{B}} b_L =: g + b.$$

If $x \in \mathbb{R}^d \setminus \bigcup\mathscr{B}$, it means that all dyadic cubes Q containing x satisfy $\langle |f|\rangle_Q \leq \lambda$, hence by Lebesgue's theorem, also $|g(x)| = |f(x)| \leq \lambda$ almost everywhere on this set. If $x \in L \in \mathscr{B}$, then

$$|g(x)| = |\langle f\rangle_L| \leq \frac{1}{|L|}\int_L |f| \leq \frac{2^d}{|\hat{L}|}\int_{\hat{L}} |f| \leq 2^d\lambda$$

by the maximality of L. So altogether $\|g\|_\infty \leq 2^d\lambda$, and then

$$\|g\|_2^2 = \int |g|^2 \leq \|g\|_\infty \int |g| \leq 2^d\lambda\Big(\int_{\mathbb{R}^d\setminus\bigcup\mathscr{B}} |f| + \sum_{L\in\mathscr{B}} |B||\langle f\rangle_L|\Big)$$

$$\leq 2^d\lambda\Big(\int_{\mathbb{R}^d\setminus\bigcup\mathscr{B}} |f| + \sum_{L\in\mathscr{B}} \int_L |f|\Big) = 2^d\lambda\|f\|_1.$$

We can then estimate

$$|\{|Ш f| > \lambda\}| \leq |\{|Ш g| > \tfrac{1}{2}\lambda\}| + |\{|Ш b| > \tfrac{1}{2}\lambda\}|,$$

and

$$|\{|Ш g| > \tfrac{1}{2}\lambda\}| \leq \frac{4}{\lambda^2}\|Ш g\|_2^2 \leq \frac{4}{\lambda^2}\|g\|_2^2 \leq \frac{4}{\lambda^2}2^d\lambda\|f\|_1 = 2^{d+2}\frac{1}{\lambda}\|f\|_1,$$

which is a bound of the required form. Note that this only used the L^2 boundedness of the operator Ш, not its structure, which is typical for the estimation of the 'good' part in the Calderón–Zygmund decomposition.

For the 'bad' part b, we argue as follows. Observe first that

$$Ш b = \sum_{K\in\mathscr{D}} A_K \sum_{L\in\mathscr{B}} b_L = \sum_{L\in\mathscr{B}}\sum_{K\in\mathscr{D}} A_K b_L.$$

Since $A_K b_L = A_K(1_K b_L)$ and b_L is supported on L, we only need to consider cubes K with $K\cap L \neq \varnothing$, hence either $K\subseteq L$ or $L\subset K$. But there is yet another reduction coming from the fact that $b_L = 1_L(f - \langle f\rangle_L)$ has integral zero. Namely, suppose that $K\supset L$ and $\ell(K)\geq 2^{m+1}\ell(L)$. Then all the h_I appearing in A_K have $\ell(I) = 2^{-n}\ell(K)\geq 2\ell(L)$. The Haar functions h_I are constant on all cubes of sidelength $2^{-1}\ell(I)\geq\ell(L)$, so in particular on L. Denoting by c_{IL} this constant valued, we have $\langle b_L, h_I\rangle = \int b_L c_{IL} = c_{IL}\int b_L = 0$, and hence $A_K b_L = 0$. So among the big cubes $K\supset L$, only those with $K\subseteq L^{(m)}$ may give a nonzero contribution $A_K b_L$. Hence

$$Ш b = \sum_{L\in\mathscr{B}}\Big(\sum_{K\subseteq L} A_K b_L + \sum_{K:L\subset K\subseteq L^{(m)}} A_K b_L\Big).$$

Now a weaker version of the assertion could be obtained as follows: Note that all the terms $A_K b_L$ appearing above are supported on $L^{(m)}$, and hence

$$|\{|Ш b| > \tfrac{1}{2}\lambda\}| \leq |\{|Ш b| > 0\}| \leq \Big|\bigcup_{L\in\mathscr{B}} L^{(m)}\Big| \leq \sum_{L\in\mathscr{B}} |L^{(m)}|$$

$$= 2^{dm}\sum_{L\in\mathscr{B}}|L| = \frac{2^{dm}}{\lambda}\sum_{L\in\mathscr{B}}\int_L |f| \leq 2^{dm}\frac{1}{\lambda}\|f\|_1.$$

So this would give the exponential factor 2^{md} in place of the claimed m.

We argue slightly more carefully

$$|\{|Ш b| > \tfrac{1}{2}\lambda\}| \leq \Big|\Big\{\Big|\sum_{L\in\mathscr{B}}\sum_{K\subseteq L} A_K b_L\Big| > 0\Big\}\Big| + \Big|\Big\{\sum_{L\in\mathscr{B}}\Big|\sum_{K:L\subset K\subseteq L^{(m)}} A_K b_L\Big| > \tfrac{1}{2}\lambda\Big\}\Big|.$$

The first set on the right is contained in $\bigcup_{L\in\mathscr{B}} L$, so its measure is dominated by $\|f\|_1/\lambda$, by the argument just given, but without the expansions $L^{(m)}$. So it remains to bound

$$\Big|\Big\{\Big|\sum_{L\in\mathscr{B}}\sum_{K:L\subset K\subseteq L^{(m)}} A_K b_L\Big| > \tfrac{1}{2}\lambda\Big\}\Big| \leq \frac{2}{\lambda}\Big\|\sum_{L\in\mathscr{B}}\sum_{K:L\subset K\subseteq L^{(m)}} A_K b_L\Big\|_1$$

$$\leq \frac{2}{\lambda}\sum_{L\in\mathscr{B}}\sum_{K:L\subset K\subseteq L^{(m)}}\|A_K b_L\|_1$$

$$\leq \frac{2}{\lambda}\sum_{L\in\mathscr{B}}\sum_{K:L\subset K\subseteq L^{(m)}}\|b_L\|_1 \leq \frac{2}{\lambda}\sum_{L\in\mathscr{B}} m\|b_L\|_1$$

$$\leq \frac{4m}{\lambda}\sum_{L\in\mathscr{B}}\|1_L f\|_1 \leq 4m\frac{1}{\lambda}\|f\|_1.$$

This completes the proof. □

4.4. **Local oscillations of a shift.** We are finally in a position to tie together some of the concepts developed in this section and the previous one.

Proposition 4.2 (Cruz-Uribe–Martell–Pérez 2010 [5])**.** *Let* Ш *be a dyadic shift with parameters* (m,n)*. Then for* $\lambda\in(0,1)$*,*

$$\omega_\lambda(Ш f;Q)\lesssim\frac{(1+m)2^{dn}}{\lambda}\fint_{Q^{(n)}}|f|.$$

Note that here we only get an exponential estimate in terms of the shift parameter n, and for this reason, the result is not very helpful in estimating series of shifts with arbitrarily large shift parameters (as they appear in the general representation formula for singular integral operators). On the other hand, this result provides a rather direct and elegant route to some applications involving shifts with small parameters only, in which case the exponential dependence causes no trouble. In practice, this proposition is applied in companion with Lerner's formula, where we take $\lambda\in\{2^{-2},2^{-d-2}\}$; hence the factor λ^{-1} may be absorbed in the implicit dimensional constant.

To streamline the proof, we begin with a lemma, which will have another application later on as well:

Lemma 4.3. *For all* $p\in(1,\infty)$*, we have*

$$g^*(t)\leq\|g\|_{L^{p,\infty}}t^{-1/p}\leq\|g\|_{L^p}t^{-1/p}.$$

Proof. By definition,

$$g^*(t)=\inf\{\alpha\geq 0:|\{|g|>\alpha\}|\leq t\}.$$

Now if $g\in L^{p,\infty}$, this means that $\alpha\cdot|\{|g|>\alpha\}|^{1/p}\leq\|g\|_{L^{p,\infty}}$; hence

$$|\{|g|>\alpha\}|\leq\alpha^{-p}\|g\|^p_{L^{p,\infty}}\leq t$$

at least for all $\alpha\geq\|g\|_{L^{p,\infty}}t^{-1/p}$. So the infimum of the admissible values of α for $g^*(t)$ is at most $\|g\|_{L^{p,\infty}}t^{-1/p}$, exactly as claimed. □

Proof of Proposition 4.2. We write

$$1_Q Ш f=1_Q Ш(1_{Q^{(n)}}f)+1_Q Ш(1_{\mathbb{R}^d\setminus Q^{(n)}}f),\tag{4.1}$$

where the latter term is

$$1_Q Ш(1_{\mathbb{R}^d\setminus Q^{(n)}}f)=\sum_{K\in\mathscr{D}}1_Q A_K(1_{\mathbb{R}^d\setminus Q^{(n)}}f),$$

and only cubes K with $Q\cap K\neq\varnothing\neq K\setminus Q^{(n)}$ need to be considered. This means that $K\supseteq Q^{(n+1)}$. But then all the h_J appearing in A_K are constant on cubes of sidelength $2^{-n-1}\ell(K)\geq\ell(Q)$, so in particular on Q. Thus the latter term on the right of (4.1) is actually a constant times 1_Q. Writing c_Q for this constant, we conclude that

$$\begin{aligned}\omega_\lambda(Ш f;Q)&=\inf_c\big(1_Q(Ш f-c)\big)^*(\lambda|Q|)\\&\leq\big(1_Q(Ш f-c_Q)\big)^*(\lambda|Q|)=\big(1_Q Ш(1_{Q^{(n)}}f)\big)^*(\lambda|Q|)\\&\leq\frac{1}{\lambda|Q|}\|Ш(1_{Q^{(n)}}f)\|_{L^{1,\infty}}\qquad\text{(by Lemma 4.3)}\\&\lesssim\frac{2^{dn}}{\lambda|Q^{(n)}|}(1+m)\|1_{Q^{(n)}}f\|_{L^1}\qquad\text{(by Proposition 4.1).}\end{aligned}$$

□

4.5. **The sharp weighted estimate for dyadic shifts.** We are ready for the main result of this section:

Theorem 4.1 (Lacey–Petermichl–Reguera 2010 [14])**.** *Let* Ш *be a dyadic shift with parameters* (m,n)*. Then*

$$\|Ш f\|_{L^2(w)} \leq C_d(m,n)[w]_{A_2}\|f\|_{L^2(w)}.$$

The original proof of Lacey–Petermichl–Reguera was relatively complicated, but we will present a simpler approach by Cruz-Uribe–Martell–Pérez, based on their Proposition stated above. Both these proofs give a constant $C_d(m,n)$, which is exponential in the shift parameters, and we will stop paying attention to its precise value here. However, the more complicated Lacey–Petermicl–Reguera approach has the advantage that it can be modified to yield a much better polynomial dependence, which is necessary for applications to general singular integrals. At the time of writing, it is not known if such an improvement is possible for the Cruz-Uribe–Martell–Pérez approach.

Proof of Theorem 4.1 by Cruz-Uribe–Martell–Pérez [5]. We consider a function $f \in L^2(w) \cap L^2$ (the intersection of weighted and unweighted L^2 spaces). This intersection includes in particular all bounded compactly supported functions, so such f are dense in $L^2(w)$. By $f \in L^2$, we know that $Ш f \in L^2$ is a well-defined function, and we can compute the norm

$$\|Ш f\|^2_{L^2(w)} = \int_{\mathbb{R}^d} |Ш f|^2 w = \sum_{\varepsilon} \int_{\mathbb{R}^d_\varepsilon} |Ш f|^2 w,$$

where $\mathbb{R}^d_\varepsilon$ are the 2^d quadrants of $\mathbb{R}^d$. Fix one such quadrant, and consider an increasing sequence of dyadic cubes Q_N which exhausts all $\mathbb{R}^d_\varepsilon$, i.e., $Q_{N-1} \subset Q_N$ and $\bigcup_{N=1}^\infty Q_N = \mathbb{R}^d_\varepsilon$. By monotone convergence, it would suffice to bound the $L^2(w)$ norm of $1_{Q_N} Ш f$, uniformly in N. In order to make use of Lerner's formula, we still want to subtract the median of $Ш f$ on this cube. And we observe that for $g = Ш f \in L^2$, we have

$$\begin{aligned} |m_g(Q_N)| &\leq (1_{Q_N} g)^*(\lambda |Q_N|) \qquad \text{(by Lemma 3.2, for } \lambda \in (0, \tfrac{1}{2})) \\ &\leq \frac{1}{\sqrt{\lambda |Q_N|}} \|g\|_{L^2} \qquad \text{(by Lemma 4.3)} \\ &\to 0 \qquad \text{as } |Q_N| \to \infty. \end{aligned}$$

Hence we conclude that, on $\mathbb{R}^d_\varepsilon$, we have

$$1_{Q_N}\big(Ш f - m_{Ш f}(Q_N)\big) \to Ш f,$$

and hence

$$\begin{aligned} \int_{\mathbb{R}^d_\varepsilon} |Ш f|^2 w &= \int_{\mathbb{R}^d_\varepsilon} \liminf_{N\to\infty} 1_{Q_N} |Ш f - m_{Ш f}(Q_N)|^2 w \\ &\leq \liminf_{N\to\infty} \int_{\mathbb{R}^d_\varepsilon} 1_{Q_N} |Ш f - m_{Ш f}(Q_N)|^2 w \qquad \text{(by Fatou's lemma)} \\ &\leq \liminf_{N\to\infty} \int \Big(4M^\#_{1/4,Q_N}(Ш f) + 4\sum_{k,j} \omega_{2^{-d-2}}(Ш f; \hat{Q}^k_j) \cdot 1_{Q^k_j}\Big)^2 w \end{aligned} \tag{4.2}$$

by Lerner's formula in the last step.

The first term is straightforward by what we already know:

$$\begin{aligned} M^\#_{1/4,Q_N}(Ш f) &= \sup_{Q \in \mathscr{D}(Q_N)} 1_Q \cdot \omega_{1/4}(Ш f; Q) \qquad \text{(definition of } M^\#_{1/4,Q_N}) \\ &\lesssim \sup_{Q \in \mathscr{D}(Q_N)} 1_Q \cdot \fint_{Q_N^{(n)}} |f| \qquad \text{(Proposition 4.2)} \\ &\leq Mf \qquad \text{(definition of the usual maximal function } M); \end{aligned}$$

note that we absorbed the dependence both the dimension and the shift parameters into the implicit constant in the second step. By the previous pointwise estimate and the Muckenhoupt–Buckley theorem, we now conclude that

$$\|M^{\#}_{1/4,Q_N}(\text{Ш}f)\|_{L^2(w)} \lesssim \|Mf\|_{L^2(w)} \lesssim [w]_{A_2}\|f\|_{L^2(w)}.$$

For the second term on the right of (4.2), we also get a pointwise bound from Proposition 4.2:

$$\sum_{k,j}\omega_{2^{-d-2}}(\text{Ш}f;\hat{Q}^k_j)\cdot 1_{Q^k_j} \lesssim \sum_{k,j}\fint_{(\hat{Q}^k_j)^{(n)}}|f|\cdot 1_{Q^k_j} =: \sum_{k,j}\fint_{P^k_j}|f|\cdot 1_{Q^k_j} =: F,$$

where $P^k_j := (\hat{Q}^k_j)^{(n)} = (Q^k_j)^{(n+1)}$, and it remains to estimate $\|F\|_{L^2(w)}$. We do this with the help of the duality

$$\|F\|_{L^2(w)} = \sup\Big\{\int Fhw : \|h\|_{L^2(w)} = 1\Big\}.$$

We manipulate the integral somewhat in the spirit of Lerner's proof of the Muckenhoupt–Buckley theorem (denote $\sigma := w^{-1}$):

$$\begin{aligned}\int Fhw &= \sum_{k,j}\frac{1}{|P^k_j|}\int_{P^k_j}|f|\cdot\int_{Q^k_j}hw \\ &= \sum_{k,j}\frac{1}{\sigma(P^k_j)}\int_{P^k_j}(|f|w)\sigma\cdot\frac{1}{w(Q^k_j)}\int_{Q^k_j}hw\cdot\frac{\sigma(P^k_j)}{|P^k_j|}\frac{w(Q^k_j)}{|Q^k_j|}\cdot|Q^k_j|.\end{aligned} \tag{4.3}$$

And we make the following estimates:

$$\frac{\sigma(P^k_j)}{|P^k_j|}\frac{w(Q^k_j)}{|Q^k_j|} \leq 2^{d(n+1)}\frac{\sigma(P^k_j)}{|P^k_j|}\frac{w(P^k_j)}{|P^k_j|} \leq 2^{d(n+1)}[w]_{A_2}$$

and, by the properties of the cubes Q^k_j provided by Lerner's theorem,

$$|Q^k_j| \leq 2|Q^k_j\setminus\Omega_{k+1}| =: 2|E^k_j|, \qquad \Omega_{k+1} = \bigcup_i Q^{k+1}_i.$$

Note that the sets E^k_j are pairwise disjoint with respect to both k and j. Finally, for any $x \in E^k_j \subseteq Q^k_j$, we have

$$\frac{1}{\sigma(P^k_j)}\int_{P^k_j}(|f|w)\sigma \leq M_\sigma(|f|w)(x), \qquad \frac{1}{w(Q^k_j)}\int_{Q^k_j}hw \leq M_w(h)(x)$$

by definition. Substituting these into (4.3) and absorbing the dependence on dimension and shift parameters, we have

$$\begin{aligned}\int Fhw &\lesssim \sum_{k,j}\inf_{E^k_j}M_\sigma(|f|w)\cdot\inf_{E^k_j}M_w(h)\cdot|E^k_j| \\ &\leq \sum_{k,j}\int_{E^k_j}M_\sigma(|f|w)M_w(h) \\ &\leq \int_{\mathbb{R}^d}M_\sigma(|f|w)M_w(h)\sigma^{1/2}w^{1/2} \qquad \text{(by disjointness of } E^k_j \text{ and } \sigma w = 1) \\ &\leq \Big(\int_{\mathbb{R}^d}M_\sigma(|f|w)^2\sigma\Big)^{1/2}\Big(\int_{\mathbb{R}^d}M_w(h)^2w\Big)^{1/2} \qquad \text{(Cauchy–Schwarz)} \\ &\lesssim \Big(\int_{\mathbb{R}^d}(|f|w)^2\sigma\Big)^{1/2}\Big(\int_{\mathbb{R}^d}h^2w\Big)^{1/2} \qquad \text{(universal maximal inequality)} \\ &= \Big(\int_{\mathbb{R}^d}|f|^2w\Big)^{1/2} \qquad (\text{by } w\sigma = 1 \text{ and } \|h\|_{L^2(w)} = 1).\end{aligned}$$

This completes the proof. □

By an application of the extrapolation theory, we obtain the corresponding bounds in $L^p(w)$, $p \in (1,\infty)$. We recall the extrapolation theorem (both upper and lower extrapolation in the same statement) in the following simpler form, which only takes into account the dependence of the estimates on $[w]_{A_p}$ not on $[w]_{A_\infty}$. Consider an estimate of the form

$$\|Tf\|_{L^p(w)} \lesssim [w]_{A_p}^{\tau(p)} \|f\|_{L^p(w)}.$$

If T satisfies this for some $p = r \in (1,\infty)$ and all $w \in A_r$, then it satisfies this for all $p \in (1,\infty)$ and all $w \in A_p$, with

$$\tau(p) = \max\Big\{\frac{r-1}{p-1}, 1\Big\}\tau(r).$$

An application of this result with $r = 2$ and $\tau(2) = 1$, in combination with Theorem 4.1 gives:

Corollary 4.2. *Let* Ш *be a dyadic shift with parameters* (m,n)*. Then*

$$\|\text{Ш} f\|_{L^p(w)} \le C_{d,p}(m,n)[w]_{A_p}^{\max\{1,1/(p-1)\}}\|f\|_{L^p(w)}.$$

By the representation theorem which we prove in the next section, this implies the same type of estimate for the Hilbert transform H in place of Ш. And one can check for H, that the dependence $[w]_{A_p}^{\max\{1,1/(p-1)\}}$ is the best possible; hence it must also be the best possible in the above statement.

5. Representation of singular integrals by dyadic shifts

We now take up the question of representing classical singular integral operators with the help of the dyadic shifts, in such a way that we can obtain sharp norm inequalities for these classical operators from the work already done for shifts. A number of representation theorems of this type are available. They all rely on the notion of random dyadic systems.

5.1. **Random dyadic cubes.** Let $\omega = (\omega_k)_{k\in\mathbb{Z}} \in \Omega := (\{0,1\}^d)^{\mathbb{Z}}$, and consider dyadic systems of the form

$$\mathscr{D}^\omega = \bigcup_{k\in\mathbb{Z}} \mathscr{D}^\omega_k, \qquad \mathscr{D}^\omega_k = \Big\{2^{-k}([0,1)^d + m) + \sum_{j>k} 2^{-j}\omega_j : m \in \mathbb{Z}\Big\}$$
$$= \Big\{I + \sum_{j>k} 2^{-j}\omega_j : I \in \mathscr{D}^0_k\Big\},$$

i.e., $\mathscr{D}^\omega_k$ is obtained by translating the cubes in the standard system $\mathscr{D}^\omega_k$ by the truncated binary expansion $\sum_{j>k} 2^{-j}\omega_j \in 2^{-k}[0,1]^d$.

There is a natural way to introduce a probability measure of Ω: we require that the coordinates ω_k are independent, and $\mathbb{P}(\omega_k = \eta) = 2^{-d}$ for each $\eta \in \{0,1\}^d$. A random choice of $\omega \in \Omega$ then induces a random choice of a dyadic system $\mathscr{D}^\omega$. We will be mostly concerned about averages (expectations) of some functions of ω, taken with respect to the measure $\mathbb{P}$ over the whole probability space Ω.

We will also be interested in dilated dyadic systems $\mathscr{D}^{\omega,r} := \{rI : I \in \mathscr{D}^\omega\}$, where

$$r\prod_{j=1}^d [a_j, b_j) := \prod_{j=1}^d [ra_j, rb_j),$$

for $r \in [1,2)$. We also take averages with respect to this parameter r, but with respect to the measure $\mathrm{d}r/r$. The average of $f(r)$ would then be $(\log 2)^{-1}\int_1^2 f(r)\,\mathrm{d}r$, but we will drop the factor $(\log 2)^{-1}$, as this is just a universal numerical constant, which is of no concern to us.

5.2. **Special shifts and their averages.** We restrict our considerations to shifts of a particularly simple structure, which permits an easy evaluation of the required averages. For a dyadic cube K, write $\inf K$ for its "lower left" corner (i.e., the pointwise infimum separately in all coordinates). Let

$$\text{Ш} = \sum_{K\in\mathscr{D}} \gamma(\ell(K))A_K,$$

where the $\gamma(\ell(K))$ are bounded coefficients, and all operators A_K are obtained from

$$A_{[0,1)^d}f(x) =: A_0 f(x) = \int a(x,y)f(y)\,\mathrm{d}y, \qquad \operatorname{supp} a \subseteq [0,1]^d\times[0,1]^d,$$

by

$$A_K f(x) = \int_{[0,1)^d} \frac{1}{|K|} a\Big(\frac{x-\inf K}{\ell(K)}, \frac{y-\inf K}{\ell(K)}\Big) f(y)\,\mathrm{d}y.$$

It is easy to see that if A_0 is of the form required by the definition of a shift, then so are all A_K.

We consider shifts of this type associated to different dyadic systems $\mathscr{D} = \mathscr{D}^{\omega,r}$. Let us write explicitly

$$\text{Ш}^{\omega,r} = \sum_{K\in\mathscr{D}^{\omega,r}} \gamma(\ell(K))A_K.$$

Since all $\text{Ш}^{\omega,r}$ are uniformly bounded operators on L^2 (and by the results of the previous section, on all $L^p(w)$ for $p\in(1,\infty)$ and $w\in A_p$) it is formally clear that so is their average $\int_1^2\int_\Omega \text{Ш}^{\omega,r}\,\mathrm{d}\mathbb{P}(\omega)\,\mathrm{d}r/r$. Let us be a bit more precise about the existence of these integrals.

Let us define the averaging operators

$$\mathbb{E}_k^{\omega,r} f := \sum_{Q\in\mathscr{D}_k^{\omega,r}} 1_Q \langle f\rangle_Q.$$

Then one readily observes the following identity for a Haar function h_J, $J\in\mathscr{D}^{\omega,r}$:

$$\mathbb{E}_k^{\omega,r} h_J = \begin{cases} h_J, & \ell(J) > r2^{-k} \\ 0, & \ell(J)\leq r2^{-k}. \end{cases}$$

The reason is simple: h_J is constant on cubes smaller than J, so the averaging has no effect, and $\int h_J = 0$, so averaging over cubes equal to or larger than its support results in zero. If $\text{Ш}^{\omega,r}$ has shift parameters (m,n), it follows that (recall that the h_J appearing in A_K have $\ell(J) = 2^{-n}\ell(K)$)

$$\sum_{\substack{K\in\mathscr{D}^{\omega,r} \\ 2^{-s}<\ell(K)/r\leq 2^{-t}}} \gamma(\ell(K))A_K f = (\mathbb{E}_{s+n}^{\omega,r} - \mathbb{E}_{t+n}^{\omega,r})\text{Ш}^{\omega,r} f \underset{\substack{s\to+\infty \\ t\to-\infty}}{\longrightarrow} \text{Ш}^{\omega,r} f, \tag{5.1}$$

where the convergence takes place pointwise for $\text{Ш}^{\omega,r} f\in L^1_{\text{loc}}$, in particular, when $f\in L^p(w)$ for $p\in(1,\infty)$ and $w\in A_p$. Since $\mathbb{E}_s^{\omega,r} g\leq Mg$ and the maximal operator is bounded on $L^p(w)$, the convergence also takes place in $L^p(w)$. Parameterizing the cubes $K\in\mathscr{D}^{\omega,r}$ on the left of (5.1) as

$$K = r(L+\omega(\ell(L))), \qquad \omega(\ell(L)) := \sum_{j:2^{-j}<\ell(L)} 2^{-j}\omega_j, \qquad L\in\mathscr{D}^0,$$

it can be written as

$$\begin{aligned} \sum_{\substack{K\in\mathscr{D}^{\omega,r} \\ 2^{-s}<\ell(K)/r\leq 2^{-t}}} \gamma(\ell(K))A_K f(x) &= \sum_{\substack{L\in\mathscr{D}^0 \\ 2^{-s}<\ell(L)\leq 2^{-t}}} \gamma(r\ell(L))A_{r(L+\omega(\ell(L)))} f(x), \\ = \sum_{\substack{K\in\mathscr{D}^0 \\ 2^{-s}<\ell(L)\leq 2^{-t}}} &\frac{\gamma(r\ell(K))}{r^d|L|}\int a\Big(\frac{x/r-\inf L-\omega(\ell(L))}{\ell(L)}, \frac{y/r-\inf L-\omega(\ell(L))}{\ell(L)}\Big) f(y)\,\mathrm{d}y. \end{aligned}$$

Every term here is jointly measurable with respect to $(\omega,r,x)\in\Omega\times[1,2)\times\mathbb{R}^d$; hence so is their sum as well the limit in (5.1).

So $Ш^{\omega,r}f(x)$ is jointly measurable, and we have

$$
\begin{aligned}
&\int_{\mathbb{R}^d}\Big(\int_\Omega\int_1^2 |Ш^{\omega,r}f(x)|\frac{\mathrm{d}r}{r}\,\mathrm{d}\mathbb{P}(\omega)\Big)^p w(x)\,\mathrm{d}x \\
&\quad\lesssim \int_{\mathbb{R}^d}\int_\Omega\int_1^2 |Ш^{\omega,r}f(x)|^p\frac{\mathrm{d}r}{r}\,\mathrm{d}\mathbb{P}(\omega)w(x)\,\mathrm{d}x \qquad \text{(by Jensen)} \\
&\quad= \int_\Omega\int_1^2\int_{\mathbb{R}^d} |Ш^{\omega,r}f(x)|^p w(x)\,\mathrm{d}x\frac{\mathrm{d}r}{r}\,\mathrm{d}\mathbb{P}(\omega) \qquad \text{(by Fubini)} \\
&\quad\lesssim \int_\Omega\int_1^2 [w]_{A_p}^{\max\{1,1/(p-1)\}p}\int_{\mathbb{R}^d}|f(x)|^p w(x)\,\mathrm{d}x\frac{\mathrm{d}r}{r}\,\mathrm{d}\mathbb{P}(\omega) \qquad \text{(by Corollary 4.2)} \\
&\quad\lesssim [w]_{A_p}^{\max\{1,1/(p-1)\}p}\int_{\mathbb{R}^d}|f(x)|^p w(x)\,\mathrm{d}x \qquad \text{(integrating out } \omega \text{ and } r).
\end{aligned}
\tag{5.2}
$$

So in particular the integral

$$
Tf(x) := \int_\Omega\int_1^2 Ш^{\omega,r}f(x)\frac{\mathrm{d}r}{r}\,\mathrm{d}\mathbb{P}(\omega)
$$

exists for almost every $x\in\mathbb{R}^d$, and defines a bounded operator on $L^p(w)$, of norm at most a constant times $[w]_{A_p}^{\max\{1,1/(p-1)\}}$. It remains to see which classical operators can be obtained in this way.

5.3. **Alternative representation of the average.** Recall (5.1), and the subsequent observation that these truncated shifts are pointwise dominated by $M(Ш^{\omega,r}f)(x)$. By essentially the same computation as in (5.2), using the boundedness on $L^p(w)$ of both M and $Ш^{\omega,r}$, we find that

$$
\int_\Omega\int_1^2 M(Ш^{\omega,r}f)(x)\frac{\mathrm{d}r}{r}\,\mathrm{d}\mathbb{P}(\omega)<\infty
$$

at almost every $x\in\mathbb{R}^d$. This justifies the use of dominated convergence to deduce from (5.1) that

$$
\begin{aligned}
Tf(x) &= \lim_{\substack{b\to+\infty\\ a\to-\infty}}\int_\Omega\int_1^2 \sum_{\substack{K\in\mathscr{D}^{\omega,r}\\ 2^{-s}<\ell(K)/r\leq 2^{-t}}}\gamma(\ell(K))A_K f(x)\frac{\mathrm{d}r}{r}\,\mathrm{d}\mathbb{P}(\omega) \\
&= \lim_{\substack{b\to+\infty\\ a\to-\infty}}\sum_{k=a}^{b-1}\int_1^2\int_\Omega\sum_{L\in\mathscr{D}^0_k}\gamma(r2^{-k})A_{r(2^{-k}L+\omega(\ell(L)))}f(x)\,\mathrm{d}\mathbb{P}(\omega)\frac{\mathrm{d}r}{r}.
\end{aligned}
\tag{5.3}
$$

To proceed with the evaluation, observe that $\omega(\ell(L))=\omega(2^{-k})=\sum_{j>k}2^{-j}\omega_j$ ranges (uniformly) over $2^{-k}[0,1]^d$ as $\omega\in\Omega$, and also write $L\in\mathscr{D}^0_k$ as $L=2^{-k}[0,1)^d+m$, with $m\in\mathbb{Z}^d$. Then

$$
\begin{aligned}
&\int_\Omega\sum_{L\in\mathscr{D}^0_k}\gamma(r2^{-k})A_{r(2^{-k}L+\omega(\ell(L)))}f(x)\,\mathrm{d}\mathbb{P}(\omega) \\
&\quad= \int_{[0,1)^d}\sum_{m\in\mathbb{Z}^d}\gamma(t)A_{t([0,1)^d+m+u)}f(x)\,\mathrm{d}u, \qquad t:=2^{-k}r \\
&\quad= \int_{[0,1)^d}\sum_{m\in\mathbb{Z}^d}\frac{\gamma(t)}{t^d}\int a\Big(\frac{x}{t}-m-u,\frac{y}{t}-m-u\Big)f(y)\,\mathrm{d}y\,\mathrm{d}u.
\end{aligned}
$$

Observing that $u+m$ ranges over $\mathbb{R}^d$ as $u\in[0,1)^d$ and $m\in\mathbb{Z}^d$, we can continue

$$
\begin{aligned}
&= \int_{\mathbb{R}^d}\frac{\gamma(t)}{t^d}\int a\Big(\frac{x}{t}-u,\frac{y}{t}-u\Big)f(y)\,\mathrm{d}y\,\mathrm{d}u = \frac{\gamma(t)}{t^d}\int\int_{\mathbb{R}^d}a\Big(\frac{x}{t}-u,\frac{y}{t}-u\Big)\,\mathrm{d}uf(y)\,\mathrm{d}y \\
&= \frac{\gamma(t)}{t^d}\int\int_{\mathbb{R}^d}a\Big(\frac{x-y}{t}+z,z\Big)\,\mathrm{d}zf(y)\,\mathrm{d}y =: \frac{\gamma(t)}{t^d}\int\tilde{a}((x-y)/t)f(y)\,\mathrm{d}y.
\end{aligned}
$$

It is easy to justify the changes in the integration order, since $|a(x,y)| \leq 1_{[0,1]^d}(x)1_{[0,1]^d}(y)$, so that

$$|\tilde{a}(x)| \leq \int |a(x+z,z)|\,\mathrm{d}z \leq \int_{[0,1)^d} 1_{[0,1)^d}(x+z)\,\mathrm{d}z \leq 1_{[-1,1]^d}(x).$$

Substituting back to (5.3) and using

$$\sum_{k=a}^{b-1}\int_1^2 F(r2^{-k})\frac{\mathrm{d}r}{r} = \sum_{k=a}^{b-1}\int_{2^{-k}}^{2^{1-k}} F(t)\frac{\mathrm{d}t}{t} = \int_{2^{1-b}}^{2^{1-a}} F(t)\frac{\mathrm{d}t}{t},$$

we have

$$Tf(x) = \int_0^\infty \frac{\gamma(t)}{t^d}\int_{\mathbb{R}^d}\tilde{a}\Big(\frac{x-y}{t}\Big)f(y)\,\mathrm{d}y\frac{\mathrm{d}t}{t},$$

where the outer integral should be understood as the limit of $\int_{2^{-b}}^{2^{-a}}$ as $a \to +\infty$ and $b \to -\infty$.

Finally, suppose that $x \notin \operatorname{supp} f$, which means that $|x-y| \geq \delta > 0$ for all y appearing in the above integral. We claim that in this case the integral converges absolutely. Indeed,

$$\int_0^\infty \Big|\frac{\gamma(t)}{t^d}\tilde{a}\Big(\frac{x-y}{t}\Big)\Big|\frac{\mathrm{d}t}{t} \leq \int_{c|x-y|}^\infty \|\gamma\|_\infty \frac{\mathrm{d}t}{t^{d+1}} \lesssim \frac{\|\gamma\|_\infty}{|x-y|^d},$$

and $y \mapsto |x-y|^{-d} \cdot 1_{|x-y|\geq\delta}$ is in $L^{p'}(\sigma)$, so that this function can be integrated against $f \in L^p(w)$ (exercise). We summarize our observations in the following:

Proposition 5.1. *Let*

$$Tf(x) := \int_\Omega \int_1^2 Ш^{\omega,r} f(x)\frac{\mathrm{d}r}{r}\,\mathrm{d}\mathbb{P}(\omega), \qquad Ш^{\omega,r} = \sum_{K\in\mathscr{D}^{\omega,r}} \gamma(\ell(K))A_K,$$

where

$$A_K f(x) = \int_{[0,1)^d} \frac{1}{|K|} a\Big(\frac{x-\inf K}{\ell(K)}, \frac{y-\inf K}{\ell(K)}\Big)f(y)\,\mathrm{d}y.$$

Then the integral defining T exists for all $f \in L^p(w)$ and almost every $x \in \mathbb{R}^d$, and defines a bounded linear operator on $L^p(w)$ with

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

For $x \notin \operatorname{supp} f$, we have the formula

$$Tf(x) = \int_{\mathbb{R}^d} K(x-y)f(y)\,\mathrm{d}y, \quad K(x) := \int_0^\infty \gamma(t)\tilde{a}\Big(\frac{x}{t}\Big)\frac{\mathrm{d}t}{t^{d+1}}, \quad \tilde{a}(x) := \int_{\mathbb{R}^d} a(x+z,z)\,\mathrm{d}z.$$

So we have a weighted estimate for some convolution-type operators, but they are rather implicitly described. It remains to determine, what type of kernels K can arise from the above representation with different choices of a and γ.

5.4. **Choosing a particular shift.** We now specialize to dimension $d = 1$. Our approach is to take one suitable fixed choice of the function a, but then make full use of the freedom to pick an arbitrary $\gamma \in L^\infty(0,\infty)$. Recall that $a(x,y)$ should be the kernel of $A_{[0,1)}$, thus of the form

$$a(x,y) = \sum_{\substack{I,J\in\mathscr{D},\ I,J\subseteq[0,1)\\ \ell(I)=2^{-m},\ \ell(J)=2^{-n}}} a_{IJ}h_I(y)h_J(x) \tag{5.4}$$

for some $m, n \in \mathbb{N}$. And the choice we make will be as follows:

$$a(x,y) := h(x)g(y),$$

where (denoting by $\hat{h}_I := 1_{I_{\text{left}}} - 1_{I_{\text{right}}}$ the L^∞-normalized Haar function on I)

$$\begin{aligned} h &= 7\cdot 1_{[0,1/4)} - 1\cdot 1_{[1/4,1/2)} + 1\cdot 1_{[1/2,3/4)} - 7\cdot 1_{[3/4,1)} \\ &= 3\cdot \hat{h}_{[0,1)} + 4\cdot(\hat{h}_{[0,1/2)} + \hat{h}_{[1/2,1)}) \end{aligned}$$

and

$$g = -1\cdot 1_{[0,1/4)} + 1\cdot 1_{[1/4,3/4)} - 1\cdot 1_{[3/4,1)} = -\hat{h}_{[0,1/2)} + \hat{h}_{[1/2,1)}.$$

Thus

$$\begin{aligned}a(x,y) &= 3\big(-\hat{h}_{[0,1/2)}+\hat{h}_{[1/2,1)}\big)(y)\hat{h}_{[0,1)}(x)\\ &\quad + 4\big(-\hat{h}_{[0,1/2)}+\hat{h}_{[1/2,1)}\big)(y)(\hat{h}_{[0,1/2)}+\hat{h}_{[1/2,1)}) =: a_1(x,y)+a_2(x,y)\end{aligned}$$

is actually not precisely of the form (5.4), but it is a sum of two such kernels with parameters (m,n) equal to $(1,0)$ and $(1,1)$. But it is clear that Proposition 5.1 also applies to such an a; the convolution operator T is then simply the sum of averages of shifts of type $(1,0)$ and $(1,1)$, but both of them satisfy the required bounds, and hence so does the sum.

5.5. **Some computations.** In the unit square $[0,1)\times[0,1)$, the function $a(x,y)$ looks as follows, where each small square in the figure has sidelength $1/4$:

-7	1	-1	7
7	-1	1	-7
7	-1	1	-7
-7	1	-1	7

To compute $\tilde{a}(x)=\int a(x+z,z)\,\mathrm{d}z$, we start from the point $(x,0)$ on the x-axis, and integrate the values of a along the line of slope 1 starting from this point. In order that this line meets the unit square, it is necessary that $x\in[-1,1]$. For $x=j\frac{1}{4}$, $j\in\mathbb{Z}$, the line crosses only full $\frac{1}{4}\times\frac{1}{4}$-squares in the above figure, and we easily find the following values:

x	-1	$-3/4$	$-1/2$	$-1/4$	0	$1/4$	$1/2$	$3/4$	1
$\tilde{a}(x)$	0	$-7/4$	2	$5/4$	0	$-5/4$	-2	$7/4$	0

It is also not difficult to see that $\tilde{a}$ is piecewise linear between these values. We observe that $\tilde{a}$ is an odd function, $\tilde{a}(-x)=-\tilde{a}(x)$. Its derivative is the piecewise constant even function (the constants being the slopes of $\tilde{a}$) given in a neighbourhood of the positive axis by

$$\tilde{a}'=-5\cdot 1_{(0,1/4)}-3\cdot 1_{(1/4,1/2)}+15\cdot 1_{(1/2,3/4)}-7\cdot 1_{(3/4,1)}\qquad\text{on }\mathbb{R}_+,$$

and its second derivative, in the distributional sense, it the a combination of Dirac masses at the discontinuities of $\tilde{a}'$, the coefficients being equal to the size of the jumps at these points:

$$\tilde{a}''=2\delta_{1/4}+18\delta_{1/2}-22\delta_{3/4}+7\delta_1\qquad\text{on }\mathbb{R}_+.$$

With $d=1$, our formula for the kernel K form Proposition 5.1 reads as

$$K(x)=\int_0^\infty \gamma(t)\tilde{a}(\frac{x}{t})\frac{\mathrm{d}t}{t^2}.$$

Obviously K is also odd when $\tilde{a}$ is, so we only need to consider $x>0$. Then, writing $\tilde{a}$ as an integral of its derivative, we have

$$\tilde{a}(\frac{x}{t})=-\int_x^\infty \tilde{a}'(\frac{y}{t})\frac{\mathrm{d}y}{t}=\int_x^\infty\int_y^\infty \tilde{a}''(\frac{z}{t})\frac{\mathrm{d}z}{t}\frac{\mathrm{d}y}{t},$$

we have

$$\begin{aligned}K(x)&=\int_x^\infty\int_y^\infty\Big(\int_0^\infty\gamma(t)\tilde{a}''(\frac{z}{t})\frac{\mathrm{d}t}{t^4}\Big)\,\mathrm{d}z\,\mathrm{d}y=\int_x^\infty\int_y^\infty K''(z)\,\mathrm{d}z\,\mathrm{d}y,\\ K''(x)&=\int_0^\infty\gamma(t)\tilde{a}''(\frac{x}{t})\frac{\mathrm{d}t}{t^4}=\int_0^\infty\gamma(\frac{x}{s})\tilde{a}''(s)\frac{s^2\,\mathrm{d}s}{x^3},\\ x^3K''(x)&=\int_0^\infty\gamma(\frac{x}{s})\big(2\delta_{1/4}(s)+18\delta_{1/2}(s)-22\delta_{3/4}(s)+7\delta_1(s)\big)s^2\,\mathrm{d}s\\ &=2\gamma(4x)(\frac{1}{4})^2+18\gamma(2x)(\frac{1}{2})^2-22\gamma(\frac{4}{3}x)(\frac{3}{4})^2+7\gamma(x)\\ &=\frac{1}{8}\gamma(4x)+\frac{9}{2}\gamma(2x)-\frac{99}{8}\gamma(\frac{4}{3}x)+7\gamma(x).\end{aligned}\tag{5.5}$$

Since $\gamma\in L^\infty(0,\infty)$, it is clear that $x\mapsto x^3K''(x)$ needs to be in this same space. A key point is that we can choose it to be *any* function in $L^\infty(0,\infty)$, by a suitable choice of γ.

5.6. **Solving a functional equation.** To put the problem of finding a suitable γ into an appropriate framework, we introduce a bit of notation. On the space $L^\infty(0,\infty)$, let Δ_a be the operator given by

$$\Delta_a f(x) := f(ax), \qquad a, x > 0.$$

It is immediate that Δ_a is a bounded linear operator on $L^\infty(0,\infty)$, of norm $\|\Delta_a\| = 1$, and invertible with $\Delta_a^{-1} = \Delta_{a^{-1}}$. Also, $\Delta_a\Delta_b = \Delta_{ab}$ for any $a, b > 0$.

Consider a functional equation of the form

$$\sum_{i=0}^{k} b_i \Delta_{a_i} \gamma = m, \tag{5.6}$$

where $m \in L^\infty(0,\infty)$ is a given function, $a_i > 0$ and $b_i \in \mathbb{C}$ are given coefficients, and we should solve this for the unknown function γ. Here is a sufficient condition for the existence of a solution:

Lemma 5.1. *Suppose that the coefficients satisfy*

$$|b_0| > \sum_{i=1}^{k} |b_i|.$$

Then (5.6) *has a unique solution* $\gamma \in L^\infty(0,\infty)$*, for any given* $m \in L^\infty(0,\infty)$ *and any* $a_i > 0$*.*

Proof. We can rewrite (5.6) as

$$\begin{aligned} m &= b_0 \Delta_{a_0} \gamma + \sum_{i=1}^{k} b_i \Delta_{a_i} \gamma \\ &= b_0 \Delta_{a_0} \Big(I + \sum_{i=1}^{k} \frac{b_i}{b_0} \Delta_{a_i/a_0} \Big) \gamma =: b_0 \Delta_{a_0} (I + T)\gamma, \end{aligned} \tag{5.7}$$

where I is the identity operator, and T is defined by the last equality. For the operator norm of T, we can estimate

$$\|T\| \leq \sum_{i=1}^{k} \Big| \frac{b_i}{b_0} \Big| \|\Delta_{a_i/a_0}\| = \frac{1}{|b_0|} \sum_{i=1}^{k} |b_i| < 1.$$

It is then a general fact from Functional Analysis that the operator $I + T$ is invertible. Indeed, the series

$$\sum_{j=0}^{\infty} (-1)^j T^j$$

converges to an operator, which is easily checked to be the inverse of $I + T$. Hence (5.7) has the unique solution

$$\gamma = b_0^{-1} (I + T)^{-1} \Delta_{1/a_0} m.$$

□

In the particular case of solving

$$m(x) = \frac{1}{8}\gamma(4x) + \frac{9}{2}\gamma(2x) - \frac{99}{8}\gamma(\frac{4}{3}x) + 7\gamma(x),$$

we have

$$\frac{99}{8} - \Big(\frac{1}{8} + \frac{9}{2} + 7 \Big) = 12 + \frac{3}{8} - \Big(\frac{1}{8} + 4 + \frac{1}{2} + 7 \Big) = \frac{3}{4} > 0,$$

so this particular equation can be solved for every $m \in L^\infty(0,\infty)$.

5.7. **Vagharshakyan's representation theorem.** Combining the results of this section, we have now proven the following:

Theorem 5.1 (Vagharshakyan 2010 [26]). *Let $K:\mathbb{R}\setminus\{0\}\to\mathbb{C}$ be a twice differentiable kernel with*

$$K(-x)=-K(x),\qquad |x\cdot K(x)|+|x^2\cdot K'(x)|+|x^3\cdot K''(x)|\leq C. \tag{5.8}$$

Then there exists an operator T, bounded on all $L^p(w)$ for $p\in(1,\infty)$ and $w\in A_p$, which has the representations

$$\begin{aligned} Tf(x) &= \sum_{n=0}^{1} c_n\int_1^2\int_\Omega Ш_n^{\omega,r}f(x)\,\mathrm{d}\mathbb{P}(\omega)\frac{\mathrm{d}r}{r} \qquad \textit{for a.e. } x\in\mathbb{R},\\ &= \int_{\mathbb{R}} K(x-y)f(y)\,\mathrm{d}y \qquad \textit{for a.e. } x\notin\operatorname{supp} f, \end{aligned}$$

where $Ш_n^{\omega,r}$ is a dyadic shift of parameters $(1,n)$, associated with the dyadic system $\mathscr{D}^{\omega,r}$. In particular, this T satisfies

$$\|Tf\|_{L^p(w)}\lesssim [w]_{A_p}^{\max\{1,1/(p-1)\}}\|f\|_{L^p(w)}\qquad \forall p\in(1,\infty),\ \forall w\in A_p. \tag{5.9}$$

Indeed, if K is as in the assumption (5.8), so in particular $\mapsto x^3\cdot K''(x)\in L^\infty(0,\infty)$, and we have shown the existence of $\gamma\in L^\infty(0,\infty)$, which solves (5.5). The assumption (5.8) also implies that $K(x),K'(x)\to 0$ as $x\to\infty$, and hence K itself can be recovered as

$$K(x)=\int_x^\infty\int_y^\infty K''(z)\,\mathrm{d}z\,\mathrm{d}y=\int_0^\infty \gamma(t)\tilde{a}(\frac{x}{t})\frac{\mathrm{d}t}{t^2},$$

and this suffices for the existence of T as claimed in the Theorem, as shown in Proposition 5.1.

Corollary 5.1 (Vagharshakyan 2010 [26]). *Let T be any bounded linear operator on $L^2(\mathbb{R})$ such that*

$$Tf(x)=\int_{\mathbb{R}}K(x-y)f(y)\,\mathrm{d}y\qquad \textit{for a.e. } x\notin\operatorname{supp} f$$

for a twice differentiable kernel satisfying (5.8). *Then* (5.9) *holds.*

The difference compared to Theorem 5.1 is that we claim the weighted bound (5.9) for *all* operators with kernel K, not just one such operator. However, it turns out that this difference is not very big, for the kernel *almost* uniquely specifies the operator. We state without proof the following result from the general theory of singular integral operators. It is not particularly difficult, but would take us a little away from the main line of these lectures:

Proposition 5.2. *Suppose that T is a bounded linear operator on some L^p (or even just from L^p to $L^{p,\infty}$) with kernel* 0*, i.e.,*

$$Tf(x)=0\qquad \textit{for a.e. } x\notin\operatorname{supp} f.$$

Then T is a multiplication operator given by $Tf(x)=b(x)f(x)$ for some $b\in L^\infty$.

Proof of Corollary 5.1. Let T be as in Corollary 5.1, and let T_0 be the operator with the same kernel provided by Theorem 5.1. Then $T-T_0$ is bounded on L^2 and has kernel 0, thus by Proposition 5.2, we have $(T-T_0)f=bf$ for some $b\in L^\infty$, and hence

$$\begin{aligned} \|Tf\|_{L^p(w)}\leq \|T_0f\|_{L^p(w)}+\|bf\|_{L^p(w)} &\lesssim [w]_{A_p}^{\max\{1,1/(p-1)\}}\|f\|_{L^p(w)}+\|b\|_{L^\infty}\|f\|_{L^p(w)}\\ &\lesssim [w]_{A_p}^{\max\{1,1/(p-1)\}}\|f\|_{L^p(w)}, \end{aligned}$$

using $\|b\|_{L^\infty}\lesssim 1\leq [w]_{A_p}^{\max\{1,1/(p-1)\}}$ in the last step. □

Corollary 5.2 (Petermichl 2007 [21]). *Let H be the Hilbert transform given by*

$$Hf(x)=\lim_{\varepsilon\to 0}\int_{|x-y|>\varepsilon}\frac{f(y)\,\mathrm{d}y}{x-y}.$$

Then

$$\|Hf\|_{L^p(w)} \lesssim [w]_{A_p}^{\max\{1,1/(p-1)\}} \|f\|_{L^p(w)} \qquad \forall p \in (1,\infty),\ \forall w \in A_p.$$

Proof. We use the known fact that H is bounded on L^2. From the definition, it is clear that it has kernel $K(x) = 1/x$ which is easily checked to satisfy (5.8). Hence the result follows immediately from Corollary 5.1. □

In fact, this earlier result of Petermichl can be obtained slightly more easily than as a corollary of the more general result of Vagharshakyan. Namely, in this case $x^3 \cdot K''(x) = 2$, and it is immediate to check that a suitable constant function γ solves (5.5), so the part of the argument involving the solution of the general functional equation becomes unnecessary. Also it should be observed that the rest of our proof of Theorem 5.1 and Corollary 5.1 goes back to Petermichl's original proof of Corollary 5.2.

For the Hilbert transform, it can in fact be shown that the average of the dyadic shifts provided by Theorem 5.1 is precisely H, so the additional multiplication operator does not appear. This observation is due to myself, and it was proven in the course "Martingales and Harmonic Analysis" in Spring 2008, and published in [11].

6. Implications of the A_∞ condition

Recall that A_∞ is the largest of the A_p classes, so that $A_p \subseteq A_\infty$ for all $p \in [1,\infty]$. In this section, we investigate some consequences of the A_∞ condition, in other words, properties common to all A_p weights, irrespective of the value p.

6.1. **The reverse Hölder inequality.** It follows immediately from Hölder's inequality that $\fint_Q w \leq \big(\fint_Q w^r\big)^{1/r}$ for all $r > 1$. A remarkable consequence of the A_∞ condition is that this inequality can be reversed for small enough values of r. The qualitative result of this type is classical, and goes back to Coifman and Fefferman (1974) [3].

Theorem 6.1. *For $w \in A_\infty$, let $r(w) := 1 + 2^{-d-3}/[w]_{A_\infty}$. Then*

$$\Big(\fint_Q w^{r(w)}\Big)^{1/r(w)} \leq 2 \fint_Q w.$$

Note that $r(w)' = 1 + 2^{d+3}[w]_{A_\infty}$.

Proof by A. de la Torre (unpublished). We use the dyadic maximal function on the dyadic subcubes of a given Q_0:

$$\int_{Q_0} w^{1+\varepsilon} \leq \int_{Q_0} M_d(w1_{Q_0})^\varepsilon w = \int_0^\infty \varepsilon t^{\varepsilon-1} w(Q_0 \cap \{M_d(w1_{Q_0}) > t\})\,\mathrm{d}t.$$

Let Q_i be the maximal dyadic subcubes of Q_0 with $\langle w\rangle_{Q_i} > t$. For $t \geq \langle w\rangle_{Q_0}$, these are necessarily strict subcubes of Q_0. Denoting by $\hat{Q}_i$ the dyadic parent of Q_i, we then have

$$\begin{aligned} w(Q_0 \cap \{M_d(w1_{Q_0}) > t\}) &= \sum_i w(Q_i) \leq \sum_i w(\hat{Q}_i) \leq \sum_i t|\hat{Q}_i| = 2^d t \sum_i |Q_i| \\ &= 2^d t |\{Q_0 \cap M_d(w1_{Q_0}) > t\}|, \end{aligned}$$

where $w(\hat{Q}_i)/|\hat{Q}_i| \leq t$ be the maximality of Q_i. Hence

$$\begin{aligned} \int_{Q_0} w^{1+\varepsilon} &\leq \int_0^{\langle w\rangle_{Q_0}} \varepsilon t^{\varepsilon-1} w(Q_0)\,\mathrm{d}t + \int_{\langle w\rangle_{Q_0}}^\infty \varepsilon t^\varepsilon 2^d |\{Q_0 \cap M_d(w1_{Q_0}) > t\}|\,\mathrm{d}t \\ &\leq \langle w\rangle_{Q_0}^\varepsilon w(Q_0) + \frac{\varepsilon 2^d}{1+\varepsilon} \int_{Q_0} M_d(w1_{Q_0})^{1+\varepsilon}. \end{aligned}$$

So far we have not used that $w \in A_\infty$; we do it now. This condition says that

$$\fint_Q w \leq [w]_{A_\infty} \exp\Big(\fint_Q \log w\Big)$$

for all cubes, and taking the supremum over $Q \ni x$, $Q \subseteq Q_0$, that $M_d(w1_{Q_0})(x) \leq M_0(w1_{Q_0})(x)$, where M_0 is the dyadic logarithmic maximal function. Hence

$$\int_{Q_0} M_d(w1_{Q_0})^{1+\varepsilon} \leq [w]_{A_\infty}^{1+\varepsilon} \int_{Q_0} M_0(w1_{Q_0})^{1+\varepsilon} \leq [w]_{A_\infty}^{1+\varepsilon} \cdot e \int_{Q_0} w^{1+\varepsilon},$$

so altogether

$$\fint_{Q_0} w^{1+\varepsilon} \leq \Big(\fint_{Q_0} w\Big)^{1+\varepsilon} + 2^d [w]_{A_\infty}^{1+\varepsilon} \frac{e\cdot\varepsilon}{1+\varepsilon} \fint_{Q_0} w^{1+\varepsilon}.$$

Under the a priori assumption that $\fint_{Q_0} w^{1+\varepsilon} < \infty$, it suffices to see that the last term can be absorbed for $\varepsilon = \delta_d/[w]_{A_\infty}$ and $\delta_d = 1/c_d$ sufficiently small. Indeed, with this choice,

$$2^d[w]_{A_\infty}^{1+\varepsilon} \frac{e\cdot\varepsilon}{1+\varepsilon} \leq 2^d \delta_d([w]_{A_\infty})^{\delta_d/[w]_{A_\infty}} \cdot e \leq 2^d\delta_d e^{\delta_d/e} \cdot e,$$

where we used the elementary calculus fact that $t^{1/t} \leq e^{1/e}$ for $t \geq 1$. Since $e^{1/e}\cdot e < 4$, the choice of $\delta_d = 2^{-d-3}$ yields $2^d\delta_d e^{\delta_d/e}\cdot e < 2^{-1}$, hence

$$\fint_{Q_0} w^{1+\varepsilon} \leq \Big(\fint_{Q_0} w\Big)^{1+\varepsilon} + \frac{1}{2}\fint_{Q_0} w^{1+\varepsilon}, \qquad \varepsilon = 2^{-d-3}/[w]_{A_\infty},$$

and the claim follows under the a priori higher integrability assumption.

The extra assumption may be lifted in various ways. For example, consider the piecewise constant w_k defined by $w_k(x) := \langle w\rangle_R$ whenever $x \in R$ and R is a dyadic subcube of Q_0 with $\ell(R) = 2^{-k}\ell(Q_0)$. If Q is a dyadic subcube of Q_0 with $\ell(Q) > 2^{-k}\ell(Q_0)$, let R_i be the subcubes of sidelength $2^{-k}\ell(Q_0)$ contained in it. Then, since $t \mapsto -\log t$ is a convex function,

$$\fint_Q (-\log w_k) = \fint_Q \sum_i 1_{R_i}(-\log\langle w\rangle_{R_i}) \leq \fint_Q \sum_i 1_{R_i}\langle -\log w\rangle_{R_i} = \fint_Q(-\log w).$$

Since also $\fint_Q w_k = \fint_Q w$, we find that all the relevant A_∞ constants of w_k appearing in the previous argument are dominated by those of w. Thus the proof under a priori higher integrability shows that

$$\Big(\fint_{Q_0} w_k^{r(w)}\Big)^{1/r(w)} \leq 2\fint_{Q_0} w_k = 2\fint_{Q_0} w,$$

and the pointwise convergence $w_k \to w$ (Lebesgue's differentiation theorem) and Fatou's lemma

$$\fint_{Q_0} w^{r(w)} = \fint_{Q_0} \lim_{k\to\infty} w_k^{r(w)} \leq \liminf_{k\to\infty} \fint_{Q_0} w_k^{r(w)}$$

complete the argument. □

Corollary 6.1 (Fefferman–Pipher 1997 [8]). *Let $w \in A_\infty$. If Q is a cube and $E \subseteq Q$ a measurable subset, then*

$$\frac{w(E)}{w(Q)} \leq C\Big(\frac{|E|}{|Q|}\Big)^\delta, \tag{6.1}$$

where

$$C = 2, \qquad \delta = \frac{1}{1+2^{d+3}[w]_{A_\infty}}.$$

Proof.

$$\frac{w(E)}{w(Q)} = \frac{1}{\langle w\rangle_Q}\fint_Q 1_E w \leq \frac{1}{\langle w\rangle_Q}\Big(\fint_Q 1_E^{r(w)'}\Big)^{1/r(w)'}\Big(\fint_Q w^{r(w)}\Big)^{1/r(w)}$$
$$\leq \frac{1}{\langle w\rangle_Q}\Big(\frac{|E|}{|Q|}\Big)^{1/r(w)'} 2\fint_Q w = 2\Big(\frac{|E|}{|Q|}\Big)^\delta. \qquad \square$$

Remark 6.1. The condition that (6.1) hold for some $C, \delta > 0$, is one of the classical definitions of $w \in A_\infty$. Indeed, we have just shown that (6.1) follows from our definition, and the converse is also true (although we will not prove it here). The advantage of the definition that we use, in view of quantitative estimates, is that it carries the unique A_∞ constant $[w]_{A_\infty}$, whereas (6.1) in the classical formulation involves two parameters C and δ.

6.2. **Comparison of singular integrals and maximal functions.** The main topic of the previous sections has been the domination of $\|Tf\|_{L^p(w)}$ by $\|f\|_{L^p(w)}$, where T is either a maximal or a singular integral operator. For such questions, it was necessary that $w \in A_p$, for the same exponent p for which we investigate this inequality. We now turn to a different type of problem of comparing the norms $\|Tf\|_{L^p(w)}$ and $\|Mf\|_{L^p(w)}$. It turns out that only $w \in A_\infty$ is needed here, irrespective of the value of p.

To be more precise, let us say that T is a *Calderón–Zygmund operator* if T is a bounded linear operator on the unweighted L^2 space, having the kernel representation

$$Tf(x) = \int_{\mathbb{R}^d} K(x,y)f(y)\,\mathrm{d}y, \qquad x \notin \operatorname{supp} f,$$

where K satisfies the *standard estimates*

$$|K(x,y)| \leq \frac{C}{|x-y|^d}$$

and, for some $\theta \in (0,1]$,

$$|K(x+h,y) - K(x,y)| + |K(x,y+h) - K(x,y)| \leq \frac{C|h|^\theta}{|x-y|^{d+\theta}}, \qquad \text{if } |x-y| > 2|h|.$$

We fill further be interested in the *truncated Calderón–Zygmund operators*

$$T_\varepsilon f(x) := \int_{|x-y|>\varepsilon} K(x,y)f(y)\,\mathrm{d}y$$

and the *maximal truncations*

$$T_\natural f(x) := \sup_{\varepsilon>0} \Big| \int_{|x-y|>\varepsilon} K(x,y)f(y)\,\mathrm{d}y \Big|.$$

We agree to use the ℓ^∞-metric $|x-y| = \max_i |x_i - y_i|$ here; this is convenient, since the "balls" of this metric are the usual cubes, and it is easy to go back to the results for the Euclidean metric, when needed. (E.g., one can check that the difference of $T_\natural f$ for the ℓ^∞ metric and $T_\natural f$ for the Euclidean metric is pointwise dominated by the Mf.)

One of the main results of this section is the following theorem which, in its qualitative form, again goes back to Coifman and Fefferman (1974) [3]:

Theorem 6.2. *The following estimate holds for all $p \in (0,\infty)$, $w \in A_\infty$, and all bounded compactly supported functions f on $\mathbb{R}^d$, provided that the left side is finite:*

$$\|T_\natural f\|_{L^p(w)} \leq C_d \cdot 2^{1/p}(1+p) \cdot [w]_{A_\infty} \|Mf\|_{L^p(w)}$$

6.3. **Whitney decomposition, good-λ inequality.** The proof of Theorem 6.2 will consist of several steps. The main intermediate goal is to prove an estimate of the type

$$|\{T_\natural f > 2\lambda, Mf \leq \gamma\lambda\}| \leq Ce^{-c/\gamma}|\{T_\natural f > \lambda\}|, \tag{6.2}$$

as well as a similar weighted estimate for w in place of the Lebesgue measure on both sides. Note that $e^{-c/\gamma} \to 0$ rapidly as $\gamma \to 0$; thus, the set $\{T_\natural f > 2\lambda, Mf \leq \gamma\lambda\}$, which is obviously a subset of $\{T_\natural f > \lambda\}$, is only a very small fraction of it for small γ. Estimates of this type are commonly referred to as "good-λ inequalities".

It is not difficult to check that $\Omega := \{T_\natural f > \lambda\}$ is an open set, and it is also a bounded set when f is bounded and compactly supported. Thus we may make the following *Whitney decomposition* of Ω: Let Q_j be the maximal dyadic cubes with $3Q_j \subseteq \Omega$. These are disjoint and cover Ω. By maximality, it follows that $7Q_j \supset 3\hat{Q}_j$ intersects Ω^c. To prove (6.2), it suffices to prove[1]

$$|Q \cap \{T_\natural f > 2\lambda, Mf \leq \gamma\lambda\}| \leq Ce^{-c/\gamma}|Q| \tag{6.3}$$

for all $Q \in Q_j$. And here we may assume that there exists some $\tilde{x} \in Q$ with $Mf(\tilde{x}) \leq \gamma\lambda$, for otherwise the set on the left is empty and the inequality is trivial. Also, we may pick some $\bar{x} \in 7Q_j \cap \Omega^c$ so that $T_\natural f(\bar{x}) \leq \lambda$ by definition of Ω.

---

[1]With γ on the right in place of $e^{-c/\gamma}$, this estimate goes back to Coifman and Fefferman (1974) [3]. In the stated form, it is due to Buckley (1993) [1]. Note that $e^{-c/\gamma}$ tends to 0 much faster than γ as $\gamma \to 0$.

Let

$$f = f \cdot 1_{20Q} + f \cdot 1_{(20Q)^c} =: f_1 + f_2.$$

Lemma 6.1. *If $\gamma \leq \gamma_d$, then*

$$T_\natural f_2(x) \leq \tfrac{3}{2}\lambda, \qquad x \in Q.$$

Proof. We want to exploit the knowledge that $T_\natural f(\bar{x}) \leq \lambda$ for some $\bar{x} \in 7Q$. So let us write

$$\begin{aligned}
\int_{|y-x|>\varepsilon} K(x,y) f_2(y)\,\mathrm{d}y &= \int_{\substack{|y-x|>\varepsilon \\ |y-x_Q|>10\ell(Q)}} K(x,y) f(y)\,\mathrm{d}y \\
&= \int_{|y-\bar{x}|>20\ell(Q)+\varepsilon} K(\bar{x},y) f(y)\,\mathrm{d}y \\
&\quad + \int_{|y-\bar{x}|>20\ell(Q)+\varepsilon} [K(x,y) - K(\bar{x},y)] f(y)\,\mathrm{d}y \\
&\quad + \int_{\substack{|y-x|>\varepsilon \\ |y-x_Q|>10\ell(Q) \\ |y-\bar{x}|\leq 20\ell(Q)+\varepsilon}} K(x,y) f(y)\,\mathrm{d}y =: I + II + III.
\end{aligned}$$

It is immediate that $|I| \leq T_\natural f(\bar{x}) \leq \lambda$. Next,

$$|II| \leq \int_{|y-\bar{x}|>20\ell(Q)} \frac{C|x-\bar{x}|^\theta}{|y-\bar{x}|^{d+\theta}} |f(y)|\,\mathrm{d}y \leq \int_{|y-\bar{x}|>20\ell(Q)} \frac{C\ell(Q)^\theta}{|y-\bar{x}|^{d+\theta}} |f(y)|\,\mathrm{d}y$$

Splitting the integration over annuli $20 \cdot 2^j \cdot \ell(Q) < |y - \bar{x}| \leq 40 \cdot 2^j \cdot \ell(Q)$, we can easily dominate the right side by $C \cdot Mf(z)$ for any z with $|z - \bar{x}| < 20\ell(Q)$, and choosing $z = \tilde{x}$ we get

$$|II| \leq C \cdot Mf(\tilde{x}) \leq C\gamma\lambda.$$

Finally, in term III, we have both $|y - x| > \varepsilon$ and $|y - x| > |y - x_Q| - |x - x_Q| > 9\ell(Q)$, so that $|y - x| \geq (\varepsilon + \ell(Q))/2$. Hence

$$|III| \leq \int_{|y-\bar{x}|\leq 20(\ell(Q)+\varepsilon)} \frac{C}{(\ell(Q)+\varepsilon)^d} |f(y)|\,\mathrm{d}y \leq CMf(z)$$

for any z as before, and we get again $|III| \leq C\gamma\lambda$ by choosing $z = \tilde{x}$. Altogether, we have shown that

$$T_\natural f_2(x) \leq \lambda + C\gamma\lambda \leq \tfrac{3}{2}\lambda$$

for any small enough γ. □

From the previous lemma and the subadditivity of $T_\natural$, it follows that if $T_\natural f(x) > 2\lambda$ for some $x \in Q$, then also $T_\natural f_1(x) > \frac{1}{2}\lambda$. Thus we can estimate

$$|Q \cap \{T_\natural f > 2\lambda, Mf \leq \gamma\lambda\}| \leq |Q \cap \{T_\natural f_1 > \tfrac{1}{2}\lambda, Mf \leq \gamma\lambda\}|, \qquad f_1 = f \cdot 1_{20Q}.$$

In other words, we have localized our problem into a neighbourhood of Q also inside the operator $T_\natural$.

To proceed further, we resort to the Calderón–Zygmund decomposition of the function f_1. For some dimensional constant C_1, let $P \in \mathscr{P}$ be the maximal dyadic cubes with $\langle |f_1| \rangle_P > C_1\gamma\lambda$, and write

$$f_1 = \Big(f_1 \cdot 1_{\mathbb{R}^d \setminus \bigcup_{P\in\mathscr{P}} P} + \sum_{P\in\mathscr{P}} \langle f_1 \rangle_P \cdot 1_P \Big) + \sum_{P\in\mathscr{P}} (f_1 - \langle f_1 \rangle_P) 1_P = g + b.$$

We have, as usual,

$$|Q \cap \{T_\natural f_1 > \tfrac{1}{2}\lambda, Mf \leq \gamma\lambda\}| \leq |Q \cap \{T_\natural g > \tfrac{1}{4}\lambda, Mf \leq \gamma\lambda\}| + |Q \cap \{T_\natural b > \tfrac{1}{4}\lambda, Mf \leq \gamma\lambda\}|.$$

6.4. **The good part** g. As usual, we have $|f_1| \leq C_1\gamma\alpha$ on $(\bigcup_{P\in\mathscr{P}} P)^c$ by Lebesgue's differentiation theorem, and $\langle |f_1|\rangle_P \leq 2^d C_1\gamma\lambda$ by maximality of P; hence $\|g\|_\infty \leq C\gamma\lambda$. In our specific situation, we can also make an observation concerning the support of g. Note that $P \subset \{Mf > C_1\gamma\lambda\}$, and consider $x \in (40Q)^c$. Then

$$Mf_1(x) = \sup_{R\ni x} \frac{1}{|R|}\int_R |f_1| \leq \frac{1}{|10Q|}\int_{20Q} |f| \leq 2^d Mf(z), \qquad z \in 20Q,$$

since if $\int_R |f_1| \neq 0$, then R must intersect $\operatorname{supp} f_1 \subseteq 20Q$, and in this case $\ell(R) \geq 10\ell(Q)$. Recall that we had some $\tilde{x} \in Q$ with $Mf(\tilde{x}) \leq \lambda\gamma$. Taking $z = \tilde{x}$, we find that $Mf_1(x) \leq 2^d\gamma\lambda$ for $x \in (40Q)^c$. So if $C_1 \geq 2^d$, we find that all $x \in (40Q)^c$ implies $x \notin P$ for any P, and hence

$$P \subseteq 40Q \qquad \forall P \in \mathscr{P}.$$

Thus also $\operatorname{supp} g \subseteq 40Q$.

For a sharp estimation of $T_\natural g$, we will require some relatively precise information about the action of Calderón–Zygmund operators on bounded functions:

Lemma 6.2. *Suppose that S is an operator such that $\|S\|_{L^p\to L^{p,\infty}} \leq Cp$ for $p \in [p_0,\infty)$. Then, if h is a bounded function supported on a cube Q, then*

$$|\{|Sh| > \alpha\}| \leq e^{-\alpha/(Ce\|h\|_\infty)}|Q|, \qquad \alpha \geq Ce\|h\|_\infty p_0.$$

Remark 6.2. The operator $S = T_\natural$ satisfies the norm growth assumption of this lemma with $p_0 = 2$. We take this (nontrivial) fact for granted for the moment. It is interesting that this sharp unweighted estimate (with precise dependence on the exponent p) plays a role in obtaining sharp weighted estimates (with precise dependence on $[w]_{A_p}$ or another related quantity).

Proof of Lemma 6.2. By definition,

$$\alpha|\{|Sh| > \alpha\}|^{1/p} \leq \|Sh\|_{L^{p,\infty}} \leq Cp\|h\|_{L^p} \leq Cp\|h\|_{L^\infty}|Q|^{1/p};$$

thus

$$|\{|Sh| > \alpha\}| \leq \Big(\frac{Cp\|h\|_\infty}{\alpha}\Big)^p |Q|.$$

Choosing $p = \alpha/(Ce\|h\|_\infty) \geq p_0$ completes the proof. □

We apply Lemma 6.2 to $h = g$ and $40Q$ in place of Q. Recall that $\|g\|_\infty \leq C\gamma\lambda$. With $\alpha = \frac{1}{4}\lambda \geq Ce\|g\|_\infty$ in Lemma 6.2, we obtain

$$|\{T_\natural g > \tfrac{1}{4}\lambda\}| \leq e^{-c/\gamma}|40Q| = Ce^{-c/\gamma}|Q|,$$

as required.

6.5. **The bad part** b. We turn our attention to

$$b = \sum_{P\in\mathscr{P}} b_P = \sum_{P\in\mathscr{P}} (f_1 - \langle f_1\rangle_P)1_P,$$

where we recall that $\langle |f_1|\rangle_P > C_1\gamma\lambda$. The claim is that

$$|Q \cap \{T_\natural b > \tfrac{1}{4}\lambda, Mf \leq \gamma\lambda\}| \leq Ce^{-c/\gamma}|Q|. \tag{6.4}$$

We first observe that

$$\{Mf \leq \gamma\lambda\} \subseteq \Big(\bigcup_{P\in\mathscr{P}} 2P\Big)^c. \tag{6.5}$$

Indeed, if $x \in 2P$, then

$$Mf(x) \geq \frac{1}{|2P|}\int_{2P} |f| \geq \frac{2^{-d}}{|P|}\int_P |f_1| > 2^{-d}C_1\gamma\lambda \geq \gamma\lambda$$

if $C_1 \geq 2^d$. This proves (6.5), and hence proving (6.4) is reduced to showing

$$\Big|Q \cap \Big(\bigcup_{P\in\mathscr{P}} 2P\Big)^c \cap \{T_\natural b > \tfrac{1}{4}\lambda\}\Big| \leq Ce^{-c/\gamma}|Q|. \tag{6.6}$$

This is the final part in the proof of (6.3).

Lemma 6.3. *For $x \in \big(\bigcup_{P\in\mathscr{P}} 2P\big)^c$, we have*

$$T_\natural b(x) \leq C\gamma\lambda\Big(1 + \sum_{P\in\mathscr{P}}\sum_{j=0}^{\infty} 1_{3^jP}(x)\cdot 3^{-j(d+\theta)}\Big) =: C\gamma\lambda(1+\phi).$$

Proof. For given x and $\varepsilon > 0$, we first investigate

$$\begin{aligned} T_\varepsilon b(x) &= \sum_{P\in\mathscr{P}} T_\varepsilon b_P(x) \\ &= \sum_{\substack{P\in\mathscr{P}\\ \varepsilon<d(x,P)}} T_\varepsilon b_P(x) + \sum_{\substack{P\in\mathscr{P}\\ d(x,P)\leq\varepsilon\leq d(x,P)+\ell(P)}} T_\varepsilon b_P(x) + \sum_{\substack{P\in\mathscr{P}\\ \varepsilon>d(x,P)+\ell(P)}} T_\varepsilon b_P(x). \end{aligned} \tag{6.7}$$

In the first sum on the right, we have

$$\begin{aligned} T_\varepsilon b_P(x) &= \int_{|x-y|>\varepsilon} K(x,y)b_P(y)\,\mathrm{d}y = \int K(x,y)b_P(y)\,\mathrm{d}y \\ &= \int [K(x,y)-K(x,x_P)]b_P(y)\,\mathrm{d}y, \end{aligned}$$

where we dropped the constraint $|x-y|>\varepsilon$, since this is satisfied for every $y\in\operatorname{supp} b_P\subseteq P$, and we used the fact that $\int b_P = 0$ in the last step to introduce the constant $K(x,x_P)$, where x_P is the centre of P. This leads to the estimate

$$|T_\varepsilon b_P(x)| \leq \int \frac{C|y-x_P|^\theta}{|x-x_P|^{d+\theta}}|b_P(y)|\,\mathrm{d}y \leq \frac{C\ell(P)^\theta}{|x-x_P|^{d+\theta}}\|b_P\|_1 \leq C\gamma\lambda\frac{\ell(P)^\theta}{|x-x_P|^{d+\theta}}|P|,$$

where the last step follows readily from the definition of the component b_P in the Calderón–Zygmund decomposition.

In the second sum on the right of (6.7), we make the simpler estimate (recall that $x\notin 2P$, so that $|x-y|\eqsim|x-x_P|$ for all $y\in P$)

$$|T_\varepsilon b_P(x)| \leq \int |K(x,y)b_P(y)|\,\mathrm{d}y \leq \frac{C}{|x-x_P|^d}\|b_P\|_1 \leq \frac{C\gamma\lambda|P|}{(d(x,P)+\ell(P))^d} \leq \frac{C}{\varepsilon^d}|P|.$$

Moreover, in this term we have $\frac{1}{2}\ell(P)\leq d(x,P)\leq\varepsilon$, so P is contained in a cube of centre x and sidelength 6ε. Since all the cubes $P\in\mathscr{P}$ are also pairwise disjoint, we conclude that

$$\sum_{\substack{P\in\mathscr{P}\\ d(x,P)\leq\varepsilon\leq d(x,P)+\ell(P)}} |T_\varepsilon b_P(x)| \leq \sum_{\substack{P\in\mathscr{P}\\ d(x,P)\leq\varepsilon\leq d(x,P)+\ell(P)}} \frac{C}{\varepsilon^d}|P| \leq \frac{C}{\varepsilon^d}(6\varepsilon)^d \leq C.$$

Finally, in the third sum on the right of (6.7), the sets $\{y:|y-x|>\varepsilon\}$ and P do not intersect (since $|x-y|\leq d(x,P)+\ell(P)\leq\varepsilon$ for all $y\in P$), and hence $T_\varepsilon b_P(x)=0$.

Altogether then we have shown that

$$|T_\varepsilon b(x)| \leq C\gamma\lambda\Big(\sum_{P\in\mathscr{P}}\frac{\ell(P)^\theta|P|}{|x-x_P|^{d+\theta}}+1\Big), \qquad x\notin\bigcup_{P\in\mathscr{P}}2P,$$

and we may take the supremum over $\varepsilon>0$ of the left side to get the same upper bound for $T_\natural b(x)$.

It only remains to express the function on the right in somewhat different terms. Note that $\ell(P)^\theta|P| = \ell(P)^{d+\theta}$. We have

$$1_{(2P)^c}(x)\Big(\frac{\ell(P)}{|x-x_P|}\Big)^{d+\theta} \leq \sum_{j=1}^{\infty} 1_{3^jP\setminus 3^{j-1}P}(x)\Big(\frac{\ell(P)}{3^{j-1}\ell(P)}\Big)^{-d-\theta} \leq C\sum_{j=0}^{\infty}1_{3^jP}(x)\cdot 3^{-j(d+\theta)},$$

and this clearly completes the proof of the lemma. □

With the previous Lemma, the proof of (6.6) is again reduced further, as we find that

$$\Big|Q\cap\Big(\bigcup_{P\in\mathscr{P}}2P\Big)^c\cap\{T_\natural b>\tfrac{1}{4}\lambda\}\Big| \leq |Q\cap\{C\gamma\lambda(1+\phi)>\tfrac{1}{4}\lambda\}| \leq |Q\cap\{\phi>c/\gamma\}|$$

when γ is small enough. And we would like to prove that

$$|Q \cap \{\phi > c/\gamma\}| \leq Ce^{-c/\gamma}|Q|. \tag{6.8}$$

Such local exponential integrability is typical of BMO functions, which motivates the following lemma:

Lemma 6.4. *Let $\mathscr{P} \subset \mathscr{D}$ be any disjoint collection of dyadic cubes in $\mathbb{R}^d$. Then*

$$\phi := \sum_{P\in\mathscr{P}} \sum_{j=0}^{\infty} 1_{3^j P} \cdot 3^{-j(d+\theta)} \in \mathrm{BMO}_d \qquad (\textit{the dyadic } \mathrm{BMO})$$

with $\|\phi\|_{\mathrm{BMO}_d} \lesssim 1$, independent of $\mathscr{P}$.

Proof. Let R be a dyadic cube. Note that $3^j P$ is a union of 3^{jd} dyadic cubes of the same size as P. For $\ell(P) \geq \ell(R)$, it follows that either $R \subset 3^j P$ or $R \cap 3^j P = \varnothing$; in either case, $1_{3^j P}$ is constant on R.

Consider then $\ell(P) < \ell(R)$ with $P \cap 3R = \varnothing$. Let j be the smallest integer such that $3^j P \cap R \neq \varnothing$. Since $3^j P$ intersects both R and $P \subset (3R)^c$, we have (measuring distance in the ℓ^∞-sense, so that the "diameter" of a cube is equal to its sidelength) $\ell(3^j P) \geq \mathrm{dist}(P, R) \geq \ell(R)$. Since $3^j P$ intersects R and $\ell(3^j P) \geq \ell(R)$, it follows that $3^k P \supset R$ for all $k > j$. So the smallest integer j, say j_P, with $3^j P \cap R \neq \varnothing$ is the only j (if any) for which $1_{3^j P}$ is not constant on R. Altogether, we find that

$$1_R \phi = 1_R \sum_{\substack{P\in\mathscr{P} \\ \ell(P)<\ell(R) \\ P\subset 3R}} \sum_j 1_{3^j P} 3^{-j(d+\theta)} + 1_R \sum_{\substack{P\in\mathscr{P} \\ \ell(P)<\ell(R) \\ P\cap 3R=\varnothing}} 1_{3^{j_P} P} 3^{-j_P(d+\theta)} + 1_R c_R,$$

and hence, using $3^{j_P}\ell(P) \geq \mathrm{dist}(P,R) \geq \frac{1}{2}(\ell(R) + \mathrm{dist}(P,R))$ in the second sum,

$$\begin{aligned}
\int_R |\phi - c_R| &\leq \sum_{\substack{P\in\mathscr{P} \\ \ell(P)<\ell(R) \\ P\subset 3R}} \sum_j |3^j P| 3^{-j(d+\theta)} + |R| \sum_{\substack{P\in\mathscr{P} \\ \ell(P)<\ell(R) \\ P\cap 3R=\varnothing}} 3^{-j_P(d+\theta)} \\
&\leq \sum_{\substack{P\in\mathscr{P} \\ \ell(P)<\ell(R) \\ P\subset 3R}} \sum_j |P| 3^{-j\theta} + |R| \sum_{\substack{P\in\mathscr{P} \\ \ell(P)<\ell(R) \\ P\cap 3R=\varnothing}} \Big(\frac{2\ell(P)}{\ell(R)+\mathrm{dist}(P,R)}\Big)^{d+\theta} \\
&\lesssim \sum_{\substack{P\in\mathscr{P} \\ \ell(P)<\ell(R) \\ P\subset 3R}} |P| + |R| \sum_{\substack{P\in\mathscr{P} \\ \ell(P)<\ell(R) \\ P\cap 3R=\varnothing}} \frac{|P|\cdot\ell(R)^\theta}{(\ell(R)+\mathrm{dist}(P,R))^{d+\theta}} \\
&\lesssim |R| + |R| \sum_{\substack{P\in\mathscr{P} \\ \ell(P)<\ell(R) \\ P\cap 3R=\varnothing}} |P| \inf_{x\in P} \frac{\ell(R)^\theta}{(\ell(R)+\mathrm{dist}(x,R))^{d+\theta}} \\
&\lesssim |R| + |R| \int_{\mathbb{R}^d} \frac{\ell(R)^\theta}{(\ell(R)+\mathrm{dist}(x,R))^{d+\theta}} \,\mathrm{d}x \lesssim |R| + |R| \int_0^\infty \frac{\ell(R)^\theta t^{d-1}\,\mathrm{d}t}{(\ell(R)+t)^{d+\theta}} \lesssim |R|. \qquad \square
\end{aligned}$$

We are ready to prove (6.8). From the previous Lemma and the John–Nirenberg inequality, it follows that

$$|Q \cap \{|\phi - \langle\phi\rangle_Q| > \lambda\}| \leq Ce^{-c\lambda}|Q|,$$

and of course we have

$$|Q \cap \{\phi > c/\gamma\}| \leq |Q \cap \{|\phi - \langle\phi\rangle_Q| > c/\gamma - |\langle\phi\rangle_Q|\}| \leq Ce^{-c/\gamma + c|\langle\phi\rangle_Q|}|Q|.$$

It only remains to check that $|\langle\phi\rangle_Q| \leq C$: We have

$$\langle\phi\rangle_Q \leq \frac{1}{|Q|}\|\phi\|_1 = \frac{1}{|Q|}\sum_{P\in\mathscr{P}}\sum_{j=0}^{\infty} 3^{-j(d+\theta)}|3^j P| = \frac{1}{|Q|}\sum_{P\in\mathscr{P}} |P| \sum_{j=0}^{\infty} 3^{-j\theta} \leq \frac{C}{|Q|}\sum_{P\in\mathscr{P}} |P|,$$

and recalling that the cubes P are disjoint and contained in $40Q$ completes the estimate.

Remark 6.3. The exponential integrability estimate for the function ϕ goes back to Carleson's (1966) [2] proof of the almost-everywhere convergence of Fourier series.

So we have now completed the proof of (6.6), thus of (6.4), and this in turn was the last missing part of the proof of (6.3), which we repeat here:

$$|Q \cap \{T_\natural f > 2\lambda, Mf \leq \gamma\lambda\}| \leq Ce^{-c/\gamma}|Q|. \tag{6.9}$$

6.6. **Completion of the proof of Theorem 6.2.** We combine Corollary 6.1 with (6.9): for $E := Q \cap \{T_\natural f > 2\lambda, Mf \leq \gamma\lambda\}$, where we choose $\gamma := \gamma_d/[w]_{A_\infty}$, we have

$$\begin{aligned}\frac{w(E)}{w(Q)} \leq 2\Big(\frac{|E|}{|Q|}\Big)^{1/(c_d[w]_{A_\infty})} &\leq 2(Ce^{-c/(\gamma_d[w]_{A_\infty})})^{1/(c_d[w]_{A_\infty})} \\ &= 2C^{1/(c_d[w]_{A_\infty})}e^{-c/(c_d\gamma_d)} \leq 2C^{1/c_d}e^{-c/(c_d\gamma_d)},\end{aligned}$$

since $[w]_{A_\infty} \geq 1$. The right side approaches zero as $\gamma_d \to 0$, and hence we conclude that

$$w\Big(Q \cap \Big\{T_\natural f > 2\lambda, Mf \leq \frac{\gamma_{d,\delta}}{[w]_{A_\infty}}\lambda\Big\}\Big) \leq \delta w(Q) \tag{6.10}$$

as soon as $\gamma_{d,\delta}$ is chosen small enough, depending only on the dimension d and the parameter δ. In fact, one can readily check that

$$2C^{1/c_d}e^{-c/(c_d\gamma_{d,\delta})} = \delta \qquad \Leftrightarrow \qquad \gamma_{d,\delta} = \frac{c}{c_d \log(2C^{1/c_d}\delta^{-1})} \tag{6.11}$$

Summing (6.10) over all cubes Q in the Whitney decomposition of $\{T_\natural f > 2\lambda\}$), we obtain

$$w\Big(T_\natural f > 2\lambda, Mf \leq \frac{\gamma_{d,\delta}}{[w]_{A_\infty}}\lambda\Big) \leq \delta w(T_\natural f > \lambda). \tag{6.12}$$

This is a powerful *good λ inequality*, which encodes a lot of information about the size of $T_\natural f$. Its typical application is to the estimation of $L^p(w)$ norms, as in Theorem 6.2, as follows:

$$\begin{aligned}\|T_\natural f\|^p_{L^p(w)} &= \int_0^\infty p\lambda^{p-1}w(T_\natural f > \lambda)\,\mathrm{d}\lambda \\ &= 2^p\int_0^\infty p\lambda^{p-1}w(T_\natural f > 2\lambda)\,\mathrm{d}\lambda \\ &\leq 2^p\int_0^\infty p\lambda^{p-1}\Big[w\Big(T_\natural f > 2\lambda, Mf \leq \frac{\gamma_{d,\delta}}{[w]_{A_\infty}}\lambda\Big) + w\Big(Mf > \frac{\gamma_{d,\delta}}{[w]_{A_\infty}}\lambda\Big)\Big]\,\mathrm{d}\lambda \\ &\leq 2^p\delta\int_0^\infty p\lambda^{p-1}w(T_\natural f > \lambda) + 2^p\Big(\frac{[w]_{A_\infty}}{\gamma_{d,\delta}}\Big)^p\int_0^\infty p\lambda^{p-1}w(Mf > \lambda)\,\mathrm{d}\lambda \\ &\leq 2^p\delta\|T_\natural f\|^p_{L^p(w)} + 2^p\Big(\frac{[w]_{A_\infty}}{\gamma_{d,\delta}}\Big)^p\|Mf\|^p_{L^p(w)}.\end{aligned}$$

Hence, under the a priori assumption that $\|T_\natural f\|_{L^p(w)} < \infty$, we have

$$(1 - 2^p\delta)\|T_\natural f\|^p_{L^p(w)} \leq 2^p\Big(\frac{[w]_{A_\infty}}{\gamma_{d,\delta}}\Big)^p\|Mf\|^p_{L^p(w)}.$$

Let us explicitly choose $\delta := 2^{-1-p}$, and $\gamma_{d,\delta}$ as in (6.11). Then

$$\begin{aligned}\|T_\natural f\|_{L^p(w)} &\leq \frac{2}{(1-2^p\delta)^{1/p}}\frac{[w]_{A_\infty}}{\gamma_{d,\delta}}\|Mf\|_{L^p(w)} \\ &\leq 2\cdot 2^{1/p}\cdot[w]_{A_\infty}\frac{c_d}{c}\log(2C^{1/c_d}2^{1+p})\|Mf\|_{L^p(w)} \\ &\leq C_d\cdot 2^{1/p}(1+p)[w]_{A_\infty}\|Mf\|_{L^p(w)}.\end{aligned}$$

which is exactly the claim of Theorem 6.2.

Remark 6.4. Our primary interest in Theorem 6.2 is when $p \in [1, \infty)$ (or more generally, say, $p \in [2^{-1}, \infty)$), in which case the factor $2^{1/p}$ is of no concern to us. In fact, a better estimate for $p \in (0, 2^{-1})$ may be obtained by the following modification of the last computation. Take instead $\delta = p2^{-p}$ and $\gamma_{d,\delta}$ as in (6.11). Then

$$\begin{aligned}\|T_\natural f\|_{L^p(w)} &\leq 2 \cdot (1-p)^{-1/p} \cdot [w]_{A_\infty} \frac{c_d}{c} \log(2C^{1/c_d} p^{-1} 2^p) \|Mf\|_{L^p(w)} \\ &\leq C_d \cdot \big(1 + \log p^{-1}\big) [w]_{A_\infty} \|Mf\|_{L^p(w)},\end{aligned}$$

using in particular that $(1-p)^{-1/p} \leq (1-2^{-1})^{-1/2^{-1}} = 2^2 = 4$ for $p \in (0, 2^{-1}]$.

6.7. **Some facts from the classical theory of singular integrals.** In the (lengthy) proof of Theorem 6.2 above, we required at one point (Remark 6.2) the following sharp growth bound from the classical theory of singular integrals:

$$\|T_\natural\|_{L^p \to L^p} \leq Cp, \qquad p \to \infty. \tag{6.13}$$

Let us briefly indicate why this is the case. Recall that the operators under consideration here are bounded $T : L^2 \to L^2$ with the kernel representation

$$Tf(x) = \int_{\mathbb{R}^d} K(x,y) f(y) \, \mathrm{d}y, \qquad x \notin \operatorname{supp} f,$$

where K satisfied the *standard estimates.*

Proposition 6.1. *We have $T : L^1 \to L^{1,\infty}$ boundedly.*

This is proven in a similar way as the corresponding result for the dyadic shifts, using the Calderón–Zygmund decomposition $f = g + b$, the L^2 boundedness for Tg and the kernel bounds for Tb.

Proposition 6.2. *We have $T : L^p \to L^p$ boundedly for $p \in (1,2)$, and in particular $\|T\|_{L^p \to L^p} \leq C/(p-1) \leq Cp'$ for $p \in (1, 3/2)$.*

The qualitative statement is a consequence of the boundedness of $T : L^2 \to L^2$ and $T : L^1 \to L^{1,\infty}$ by the usual *Marcinkiewicz interpolation theorem.* The quantitative statement follows in the same way, by keeping careful track of the constants in the usual proof of the Marcinkiewicz theorem.

Proposition 6.3. *We have $T : L^p \to L^p$ boundedly for $p \in (2, \infty)$, and in particular $\|T\|_{L^p \to L^p} \leq Cp$ for $p \in (3, \infty)$.*

Note that the adjoint operator T^* satisfies exactly the same assumptions, and thus conclusions, as T. By the previous proposition, we have $T^* : L^q \to L^q$ for $q \in (1,2)$ and $\|T^*\|_{L^q \to L^q} \leq Cq'$ for $q \in (1, 3/2)$. But it is a general fact that an operator has the same norm as its adjoint in the dual space, thus

$$\|T\|_{L^p \to L^p} = \|T^*\|_{(L^p)^* \to (L^p)^*} = \|T^*\|_{L^{p'} \to L^{p'}} \leq Cp'' = Cp$$

for $p \in (3, \infty)$, whence $p' \in (1, 3/2)$.

The proof of (6.13) is the concluded by the following:

Proposition 6.4 (Cotlar's inequality)**.**

$$T_\natural f(x) \lesssim M(Tf)(x) + Mf(x).$$

Indeed, recalling that the norm of the maximal operator in L^p is at most $C_d p' \leq C_d$ for $p \geq 2$, Cotlar's inequality and the previous proposition implies that

$$\|T_\natural f\|_p \lesssim \|M(Tf)\|_p + \|Mf\|_p \lesssim \|Tf\|_p + \|f\|_p \lesssim (p+1)\|f\|_p \lesssim p\|f\|_p$$

for $p \in [3, \infty)$. We conclude this section with:

Proof of Cotlar's inequality. By definition of $T_\natural$, it suffices to prove the pointwise estimate for T_ε, with an arbitrary $\varepsilon > 0$. Consider a point x_0. Then

$$\begin{aligned} T_\varepsilon f(x_0) &= \int_{|y-x_0|>\varepsilon} K(x_0,y)f(y)\,\mathrm{d}y =: \int_{\mathbb{R}^d} K(x_0,y)f_2(y)\,\mathrm{d}y, \qquad f_2(y) := 1_{B(x_0,\varepsilon)}(y)f(y), \\ &= \int_{\mathbb{R}^d} [K(x_0,y)-K(x,y)]f_2(y)\,\mathrm{d}y + \int_{\mathbb{R}^d} K(x,y)f_2(y)\,\mathrm{d}y, \qquad x \in B(x_0,\tfrac{1}{2}\varepsilon) =: B_\varepsilon. \end{aligned}$$

In the first term, we can estimate

$$|K(x_0,y)-K(x,y)| \leq C\frac{|x_0-x|^\theta}{|x_0-y|^{d+\theta}} \leq C\frac{\varepsilon^\theta}{|x_0-y|^{d+\theta}},$$

and splitting the integration region $|y-x_0| > \varepsilon$ into the annuli $2^j\varepsilon > |y-x_0| \leq 2^{j+1}\varepsilon$, $j = 0,1,2,\ldots$, it is easy to dominate

$$\int_{\mathbb{R}^d} |[K(x_0,y)-K(x,y)]f_2(y)|\,\mathrm{d}y \lesssim Mf(x_0).$$

We turn to the second term

$$\int_{\mathbb{R}^d} K(x,y)f_2(y)\,\mathrm{d}y = Tf_2(x) = Tf(x) - Tf_1(x), \qquad f_1 := f - f_2 = 1_{B(x_0,\varepsilon)}f = 1_{2B_\varepsilon}f.$$

So altogether we have shown that

$$|T_\varepsilon f(x_0)| \lesssim Mf(x_0) + |Tf(x)| + |Tf_1(x)|, \qquad x \in B_\varepsilon, \qquad f_1 = 1_{2B_\varepsilon}f.$$

We take this to the power $\delta \in (0,1)$, and take the average over B_ε. Observing that the first two terms are constants, this gives

$$|T_\varepsilon f(x_0)| \lesssim Mf(x_0) + \Big(\fint_{B_\varepsilon} |Tf|^\delta\Big)^{1/\delta} + \Big(\fint_{B_\varepsilon} |Tf_1|^\delta\Big)^{1/\delta}. \tag{6.14}$$

By Hölder's inequality, we may replace the δ in the first average by 1, and then dominate this average by $M(Tf)(x_0)$.

It remains to deal with the last term. Its estimation is based on the $L^1 \to L^{1,\infty}$ boundedness of T and the fact that $\delta \in (0,1)$:

$$\begin{aligned} \int_{B_\varepsilon} |Tf_1|^\delta &= \int_0^\infty \delta\lambda^{\delta-1}|B_\varepsilon \cap \{|Tf_1| > \lambda\}|\,\mathrm{d}\lambda \\ &\leq \int_0^A \delta\lambda^{\delta-1}|B_\varepsilon|\,\mathrm{d}\lambda + \int_A^\infty \delta\lambda^{\delta-1}|\{|Tf_1|>\lambda\}|\,\mathrm{d}\lambda \\ &\leq A^\delta|B_\varepsilon| + \int_A^\infty \delta\lambda^{\delta-2}C\|f_1\|_1\,\mathrm{d}\lambda \\ &= A^\delta|B_\varepsilon| + \frac{\delta}{1-\delta}A^{\delta-1}C\|f_1\|_1 = \frac{|B_\varepsilon|}{1-\delta}\Big(\frac{C\|f_1\|_1}{|B_\varepsilon|}\Big)^\delta, \qquad A := \frac{C\|f_1\|_1}{|B_\varepsilon|}. \end{aligned}$$

Thus

$$\Big(\fint_{B_\varepsilon} |Tf_1|^\delta\Big)^{1/\delta} \leq (1-\delta)^{-1/\delta}\frac{C\|f_1\|_1}{|B_\varepsilon|} \leq C_\delta Mf(x_0),$$

where the last step follows readily by recalling that $\|f_1\|_1 = \int_{2B_\varepsilon}|f|$.

Thus the three terms in (6.14) are estimated as

$$|T_\varepsilon f(x_0)| \lesssim Mf(x_0) + M(Tf)(x_0) + Mf(x_0) \lesssim M(Tf)(x_0) + Mf(x_0),$$

and the proof is completed by taking supremum over $\varepsilon > 0$. □

Remark 6.5. In fact, our proof showed that $T_\natural f(x) \lesssim M_\delta(Tf)(x) + Mf(x)$, where $M_\delta g := \big(M(|g|^\delta)\big)^{1/\delta}$ and $\delta \in (0,1)$. This is not needed for our present applications, but can be used to check that $T_\natural : L^1 \to L^{1,\infty}$ from the fact that $T, M : L^1 \to L^{1,\infty}$ and $M_\delta : L^{1,\infty} \to L^{1,\infty}$. It is not true that $M : L^{1,\infty} \to L^{1,\infty}$.

7. Estimates for singular integrals involving A_1 weights

In this final section, we explore some results available in the most restricted Muckenhoupt class of A_1 weights. The goal is to prove the following:

Theorem 7.1 (Lerner–Ombrosi–Pérez 2009 [18])**.** *Let T be a Calderón–Zygmund operator and $w \in A_1$. Then*

$$\|Tf\|_{L^p(w)} \leq Cpp'[w]_{A_1}\|f\|_{L^p(w)}, \qquad p \in (1,\infty), \tag{7.1}$$

and

$$\|Tf\|_{L^{1,\infty}(w)} \leq C[w]_{A_1}(1+\log[w]_{A_1})\|f\|_{L^1(w)}, \tag{7.2}$$

where the constant C depends only on the dimension d, and parameters of the operator T.

In fact, we will obtain a slightly sharper version with part of the A_1 control replaced by A_∞, as we did with the maximal function.

Some observations are in order. First, the A_1 condition is qualitatively stronger than necessary for (7.1); just A_p would suffice, as we have shown in the special case of one-dimensional convolution operators earlier. However, quantitatively, as $p \to 1$, the dependence on the A_p constant is of the form $[w]_{A_p}^{1/(p-1)}$, while we have a linear dependence on the A_1 constant for all $p \in (1,\infty)$.

Concerning (7.2), such a logarithmic estimate seems to be the best available positive result on the end-point $p = 1$. The A_1 *conjecture*, claiming that

$$\|Tf\|_{L^{1,\infty}(w)} \leq C[w]_{A_1}\|f\|_{L^1(w)},$$

has been recently *disproven* (Nazarov–Reznikov–Vasyunin–Volberg 2010 [20]).

7.1. **Technical lemmas.** The proof of Theorem 7.1 depends on a number of technical estimates. We first record the following two classical facts:

Lemma 7.1. *Let $w_1, w_2 \in A_1$ and $w := w_1 w_2^{1-p}$. Then $w \in A_p$ and $[w]_{A_p} \leq [w_1]_{A_1}[w_2]_{A_1}^{p-1}$.*

Lemma 7.2 (Coifman–Rochberg 1980 [4])**.** *Let f be a function with $Mf < \infty$ almost everywhere, and $\delta \in (0,1)$. Then $(Mf)^\delta \in A_1$, and*

$$[(Mf)^\delta]_{A_1} \leq \frac{C_d}{1-\delta}.$$

Proofs of Lemmas 7.1 and 7.2. Left as exercises. □

The next lemma is more recent. Only a year earlier (2008), it was proven by the same authors [17] with the constant $Cp(1+\log p)$ in place of Cp.

Lemma 7.3 (Lerner–Ombrosi–Pérez 2009 [18])**.** *Let T be a Calderón–Zygmund operator, w be any weight, and $p, r \in (1,\infty)$. Then for any bounded compactly supported f, we have*

$$\Big\|\frac{Tf}{M_r w}\Big\|_{L^p(M_r w)} \leq Cp\Big\|\frac{Mf}{M_r w}\Big\|_{L^p(M_r w)}.$$

Proof. By duality with respect to the weighted measure $M_r w$, we have

$$\Big\|\frac{Tf}{M_r w}\Big\|_{L^p(M_r w)} = \sup\Big\{\int \frac{|Tf|}{M_r w} h\, M_r w = \int |Tf| h : \|h\|_{L^{p'}(M_r w)} \leq 1\Big\}.$$

Hence we estimate

$$\begin{aligned}\int |Tf|h &= \int |Tf|(M_r w)^{-1/p'}(M_r w)^{1/p'} h \\ &\leq \int |Tf|\big((M_r w)^{1/(2p')}\big)^{-2} R\big((M_r w)^{1/p'} h\big) =: \int |Tf| W,\end{aligned}$$

where R is Rubio de Francia's operator used in proving the extrapolation theorems,

$$R\phi := \sum_{k=0}^\infty \frac{2^{-k} M^k \phi}{\|M\|_{L^{p'} \to L^{p'}}^k}.$$

It satisfies

$$\phi \leq R\phi, \qquad \|R\phi\|_{L^{p'}} \leq 2\|\phi\|_{L^{p'}}, \qquad M(R\phi) \leq 2\|M\|_{L^{p'}\to L^{p'}} R\phi \leq Cp \cdot R\phi,$$

where the last condition says that $[R\phi]_{A_1} \leq Cp$.

By Lemma 7.2, we also have that

$$[(M_r w)^{1/(2p')}]_{A_1} = [(Mw^r)^{1/(2rp')}]_{A_1} \leq \frac{C_d}{1 - 1/(2rp')} \leq 2C_d,$$

and hence by Lemma 7.1 that

$$[W]_{A_\infty} \leq [W]_{A_3} \leq [R\big((M_r w)^{1/p'} h\big)]_{A_1} [(M_r w)^{1/(2p')}]^2_{A_1} \leq Cp.$$

Now we have by Theorem 6.2 that

$$\int |Tf| W \leq C[W]_{A_\infty} \int Mf\, W \leq Cp \int Mf\, W,$$

provided that $\int |Tf| W < \infty$. Let us first finish the estimate under this a priori assumption.

We have

$$\begin{aligned}
\int Mf\, W &= \int Mf (M_r w)^{-1/p'} \cdot R\big((M_r w)^{1/p'} h\big) \\
&\leq \Big(\int (Mf)^p (M_r w)^{-p/p'}\Big)^{1/p} \times \Big(\int \big[R\big((M_r w)^{1/p'} h\big)\big]^{p'}\Big)^{1/p'} \\
&\leq \Big\|\frac{Mf}{M_r w}\Big\|_{L^p(M_r w)} \times 2\Big(\int h^{p'}\, M_r w\Big)^{1/p'} \leq 2\Big\|\frac{Mf}{M_r w}\Big\|_{L^p(M_r w)},
\end{aligned}$$

where we used that $-p/p' = 1 - p$ and the boundedness of R on $L^{p'}$.

It remains to verify the finiteness of $\int |Tf|\, W$, which is almost the same computation as above:

$$\int |Tf|\, W \leq \Big(\int |Tf|^p (M_r w)^{-p/p'}\Big)^{1/p} \times \Big(\int \big[R\big((M_r w)^{1/p'} h\big)\big]^{p'}\Big)^{1/p'},$$

where the boundedness of the second factor was already checked above. If f is bounded with compact support, then $f \in L^q$ for all q, and hence $Tf \in L^q$ for all $q \in (1,\infty)$. By Lemmas 7.1 and 7.2, we have

$$(M_r w)^{-p/p'} = 1 \cdot \big((Mw^r)^{1/r}\big)^{1-p} \in A_1 \cdot (A_1)^{1-p} \subset A_p \subset A_\infty \subset L^s_{\text{loc}}$$

for some $s > 1$, where the last step follows from the reverse Hölder inequality.

If K is a compact set which contains the support of f, then

$$\int_K |Tf|^p (M_r w)^{1-p} \leq \Big(\int |Tf|^{ps'}\Big)^{1/s'} \Big(\int_K (M_r w)^{(1-p)s}\Big)^{1/s} < \infty.$$

On the other hand, for x outside the support of f (say, at a distance at least 1 from it), we have

$$|Tf(x)| \leq \int |K(x,y)||f(y)|\,\mathrm{d}y \leq \int \frac{C}{|x-y|^d}|f(y)|\,\mathrm{d}y \leq \frac{C}{\operatorname{dist}(x, \operatorname{supp} f)^d}\|f\|_1,$$

and $\max\{1, \operatorname{dist}(x, \operatorname{supp} f)\}^{-d} \leq C_f(1+|x|)^{-d} \in L^p((M_r w)^{-p/p'})$ for $(M_r w)^{-p/p'} \in A_p$, as we have observed earlier. This completes the proof. □

Our final technical result is the following. It will lead to an almost immediate proof of (7.1).

Lemma 7.4 (Lerner–Ombrosi–Pérez 2009 [18])**.** *Let T be a Calderón–Zygmund operator, w an arbitrary weight, and $p, r \in (1,\infty)$. Then for bounded and compactly supported f, we have*

$$\|Tf\|_{L^p(w)} \leq Cpp'(r')^{1/p'}\|f\|_{L^p(M_r w)}.$$

Proof. We first apply the substitution $f = gv$, where the new weight v is to be chosen. This leads to the equivalent claims that

$$\|T(gv)\|_{L^p(w)} \leq Cpp'(r')^{1/p'}\|gv\|_{L^p(M_r w)} = Cpp'(r')^{1/p'}\|g\|_{L^p(v^p M_r w)},$$

where we choose v so that $v^p M_r w = v$, i.e., $v = (M_r w)^{-1/(p-1)} = (M_r w)^{1-p'}$.

Next, we apply duality. Note that the dual of $L^p(w)$ with respect to the unweighted duality $\langle g, f\rangle = \int g\cdot f$ is $L^{p'}(\sigma)$, where $\sigma := 1_{\{w>0\}}w^{1-p'}$. This is the familiar dual weight, except that we need to take into account the possible zeros of w, now that w is allowed to be arbitrary. Then

$$\|T(gv)\|_{L^p(w)} = \sup\Big\{\Big|\int h\cdot T(gv)\Big| = \Big|\int T^*h\cdot gv\Big| : \|h\|_{L^{p'}(\sigma)} \leq 1\Big\}$$

where, by density, we may restrict the supremum to bounded and compactly supported functions h. By Hölder's inequality,

$$\Big|\int T^*h\cdot gv\Big| \leq \|T^*h\|_{L^{p'}(v)}\|g\|_{L^p(v)},$$

so that we are reduced to proving that

$$\|T^*h\|_{L^{p'}(v)} \leq Cpp'(r')^{1/p'}\|h\|_{L^{p'}(\sigma)}.$$

Here, it is important that the adjoint T^* is also a Calderón–Zygmund operator, so that we can apply Lemma 7.3. We may assume that h is supported in $\{w > 0\}$, so that we can freely write negative powers of w when they are multiplied by h.

Indeed, recalling that $v = (M_r w)^{1-p'}$, we have

$$\|T^*h\|_{L^{p'}(v)} = \Big\|\frac{T^*h}{M_r w}\Big\|_{L^{p'}(M_r w)} \leq Cp\Big\|\frac{Mh}{M_r w}\Big\|_{L^{p'}(M_r w)} = Cp\|Mh\|_{L^{p'}((M_r w)^{1-p'})}$$

and $Mh = \sup_Q\langle|h|\rangle_Q$ can be pointwise estimated by

$$\langle|h|\rangle_Q = \langle|h|w^{-1/p}w^{1/p}\rangle_Q \leq \langle w^r\rangle_Q^{1/(pr)}\langle(|h|w^{-1/p})^{(pr)'}\rangle_Q^{1/(pr)'},$$

where we used Hölder's inequality with exponents pr and $(pr)'$. Taking supremum over Q gives

$$Mh \leq (M_r w)^{1/p}\big[M\big((|h|w^{-1/p})^{(pr)'}\big)\big]^{1/(pr)'},$$

and hence

$$\begin{aligned}
\|Mh\|_{L^{p'}((M_r w)^{1-p'})} &\leq \Big(\int (M_r w)^{p'/p}\big[M\big((|h|w^{-1/p})^{(pr)'}\big)\big]^{p'/(pr)'}(M_r w)^{1-p'}\Big)^{1/p'} \\
&\leq \Big\|M\big((|h|w^{-1/p})^{(pr)'}\big)\Big\|_{L^{p'/(pr)'}}^{1/(pr)'} \qquad (\text{using } p'/p + 1 - p' = 0) \\
&\leq \Big(C\big(p'/(pr)'\big)'\Big\|(|h|w^{-1/p})^{(pr)'}\Big\|_{L^{p'/(pr)'}}\Big)^{1/(pr)'} \\
&= \Big(C\big(p'/(pr)'\big)'\Big)^{1/(pr)'}\|h\|_{L^{p'}(w^{1-p'})}
\end{aligned}$$

The second factor is $\|h\|_{L^{p'}(\sigma)}$, as we wanted, so it only remains to estimate the factor in front.

We have

$$\begin{aligned}
\big(p'/(pr)'\big)' &= \frac{p'/(pr)'}{p'/(pr)' - 1} = \frac{p'}{p' - (pr)'} = \frac{p/(p-1)}{p/(p-1) - pr/(pr-1)} \\
&= \frac{pr-1}{(pr-1) - r(p-1)} = \frac{pr-1}{r-1} \leq \frac{pr}{r-1} = pr'
\end{aligned}$$

and thus

$$\Big(C\big(p'/(pr)'\big)'\Big)^{1/(pr)'} \leq \big(Cpr'\big)^{1-1/(pr)} \leq C\cdot p\cdot(r')^{1-1/p+(1-1/r)/p} = Cp(r')^{1/p'+1/(r'p)},$$

where $(r')^{1/r'} \leq e^{1/e}$, and this completes the estimate. □

7.2. **Proof of Theorem 7.1.** In fact, we prove the following slightly sharper version, again implementing the philosophy that part of the A_p control may be replaced by the weaker A_∞ control in the quantitative bounds:

Theorem 7.2 (Hytönen–Pérez 2011 [12]). *Let T be a Calderón–Zygmund operator and $w \in A_1$. Then*

$$\|Tf\|_{L^p(w)} \leq Cpp'[w]_{A_1}^{1/p}[w]_{A_\infty}^{1/p'}\|f\|_{L^p(w)}, \qquad p \in (1,\infty), \tag{7.3}$$

and

$$\|Tf\|_{L^{1,\infty}(w)} \leq C[w]_{A_1}(1+\log[w]_{A_\infty})\|f\|_{L^1(w)}, \tag{7.4}$$

where the constant C depends only on the dimension d, and parameters of the operator T.

Proof of (7.3). This follows almost instantly from Lemma 7.4, which says that

$$\|Tf\|_{L^p(w)} \leq Cpp'(r')^{1/p'}\|f\|_{L^p(M_r w)}$$

for $p, r \in (1,\infty)$. We choose $r = r(w) = 1 + 2^{-d-3}/[w]_{A_\infty}$ as in the reverse Hölder inequality. Then $r' = 1 + 2^{d+3}[w]_{A_\infty} \leq C[w]_{A_\infty}$ and

$$M_r w \leq 2Mw \leq 2[w]_{A_1} w;$$

hence

$$(r')^{1/p'}\|f\|_{L^p(M_r w)} \leq C[w]_{A_\infty}^{1/p'}\|f\|_{L^p(2[w]_{A_1}w)} \leq C[w]_{A_\infty}^{1/p'}[w]_{A_1}^{1/p}\|f\|_{L^p(w)}. \qquad \square$$

Proof of (7.4). As is typical for weak-type $(1,1)$ proofs, this is based on the Calderón–Zygmund decomposition: For $\lambda > 0$, let Q_j be the maximal dyadic cubes with $\fint_{Q_j}|f| > \lambda$, let $\Omega := \bigcup_j Q_j$, and write

$$f = \Big(f\cdot 1_{\Omega^c} + \sum_j \langle f\rangle_{Q_j}\cdot 1_{Q_j}\Big) + \sum_j (f - \langle f\rangle_{Q_j})\cdot 1_{Q_j} =: g + \sum_j b_j =: g + b.$$

We also let $\tilde\Omega := \bigcup_j 2Q_j$. Then

$$w(|Tf| > \lambda) \leq w(\tilde\Omega) + w(\tilde\Omega^c \cap \{|Tb| > \tfrac12\lambda\}) + w(\tilde\Omega^c \cap \{|Tg| > \tfrac12\lambda\}).$$

The part $\tilde\Omega$. From the definition of Q_j, we have $|Q_j| < \lambda^{-1}\int_{Q_j}|f|$; thus

$$w(2Q_j) = \frac{w(2Q_j)}{|Q_j|}2^d|Q_j| \leq [w]_{A_1}\inf_{Q_j} w \cdot 2^d \cdot \frac{1}{\lambda}\int_{Q_j}|f| \leq \frac{2^d[w]_{A_1}}{\lambda}\int_{Q_j}|f|w,$$

and hence

$$w(\tilde\Omega) \leq \sum_j w(2Q_j) \leq \frac{2^d[w]_{A_1}}{\lambda}\sum_j\int_{Q_j}|f|w \leq \frac{2^d[w]_{A_1}}{\lambda}\|f\|_{L^1(w)},$$

since the cubes Q_j are disjoint.

The "bad" part b. This is in fact the easier part in the present case. We have

$$w(\tilde\Omega^c \cap \{|Tb| > \tfrac12\lambda\}) \leq \frac{2}{\lambda}\int_{\tilde\Omega^c}|Tb|w \leq \frac{2}{\lambda}\int_{\tilde\Omega^c}\sum_j|Tb_j|w \leq \frac{2}{\lambda}\sum_j\int_{(2Q_j)^c}|Tb_j|w.$$

We first estimate pointwise, for $x \in (2Q_j)^c$:

$$\begin{aligned}|Tb_j(x)| &= \Big|\int_{Q_j}K(x,y)b_j(y)\,\mathrm{d}y\Big| = \Big|\int_{Q_j}[K(x,y)-K(x,y_j)]b_j(y)\,\mathrm{d}y\Big| \\ &\leq \int_{Q_j}\frac{C|y-y_j|^\theta}{|x-y_j|^{d+\theta}}|b_j(y)|\,\mathrm{d}y \leq \frac{C\ell(Q_j)^\theta}{|x-y_j|^{d+\theta}}\int_{Q_j}|b_j(y)|\,\mathrm{d}y \leq \frac{C\ell(Q_j)^\theta}{|x-y_j|^{d+\theta}}\|1_{Q_j}f\|_{L^1},\end{aligned}$$

where y_j is the centre of Q_j, and we used the fact that $\int b_j = 0$. By estimating the value of $|x-y_j|^{-d-\theta}$ in the annuli $2^{k+1}Q_j \setminus 2^kQ_j$, we easily find that

$$1_{(2Q_j)^c}(x)\frac{\ell(Q_j)^\theta}{|x-y_j|^{d+\theta}} \leq C\sum_{k=0}^\infty 2^{-k\theta}\frac{1_{2^kQ_j}}{|2^kQ_j|}.$$

Hence

$$\begin{aligned}\|1_{(2Q_j)^c}Tb_j\|_{L^1(w)} &\leq C\|1_{Q_j}f_j\|_{L^1}\sum_{k=0}^{\infty}2^{-k\theta}\frac{w(2^kQ_j)}{|2^kQ_j|}\\ &\leq C\|1_{Q_j}f_j\|_{L^1}\sum_{k=0}^{\infty}2^{-k\theta}[w]_{A_1}\inf_{Q_j}w\\ &\leq C[w]_{A_1}\|1_{Q_j}f_j\|_{L^1(w)}\sum_{k=0}^{\infty}2^{-k\theta}\leq C[w]_{A_1}\|1_{Q_j}f_j\|_{L^1(w)},\end{aligned}$$

and finally

$$\begin{aligned}w(\tilde{\Omega}^c\cap\{|Tb|>\tfrac{1}{2}\lambda\}) &\leq \frac{2}{\lambda}\sum_j\|1_{(2Q_j)^c}Tb_j\|_{L^1(w)}\\ &\leq \frac{C[w]_{A_1}}{\lambda}\sum_j\|1_{Q_j}f\|_{L^1(w)}\leq\frac{C[w]_{A_1}}{\lambda}\|f\|_{L^1(w)}\end{aligned}$$

by the disjointness of the Q_j.

The "good" part g. This is the harder part in the present case, which relies on the technical Lemma 7.4. With some $p,r\in(1,\infty)$ to be chosen, we have

$$\begin{aligned}w(\tilde{\Omega}^c\cap\{|Tg|>\tfrac{1}{2}\lambda\}) &\leq \Big(\frac{2}{\lambda}\Big)^p\int_{\tilde{\Omega}^c}|Tg|^pw=\Big(\frac{2}{\lambda}\|Tg\|_{L^p(1_{\tilde{\Omega}^c}w)}\Big)^p\\ &\leq\Big(\frac{2}{\lambda}Cpp'(r')^{1/p'}\|g\|_{L^p\big(M_r(1_{\tilde{\Omega}^c}w)\big)}\Big)^p\\ &\leq (Cpp')^p(r')^{p-1}\frac{1}{\lambda^p}\int|g|^pM_r(1_{\tilde{\Omega}^c}w)\\ &\leq (Cpp')^p(r')^{p-1}\frac{1}{\lambda}\int|g|M_r(1_{\tilde{\Omega}^c}w)\qquad(\text{since }|g|\leq 2^d\lambda).\end{aligned}$$

By definition of g, we have

$$\int|g|M_r(1_{\tilde{\Omega}^c}w)=\int_{\Omega^c}|f|M_r(1_{\tilde{\Omega}^c}w)+\sum_j|\langle f\rangle_{Q_j}|\int_{Q_j}M_r(1_{\tilde{\Omega}^c}w)$$

The first term we simply estimate by

$$\int_{\Omega^c}|f|M_r(1_{\tilde{\Omega}^c}w)\leq\int_{\mathbb{R}^d}|f|M_rw.$$

For the second term, let us investigate $M_w(1_{\tilde{\Omega}^c}w)(x)$ for $x\in Q_j$. By definition

$$M_r(1_{\tilde{\Omega}^c}w)(x)=\sup_{Q\ni x}\Big(\frac{1}{|Q|}\int_Q1_{\tilde{\Omega}^c}w^r\Big)^{1/r}.$$

If $Q\ni x$ is a cube for which the above integral is nonzero, then Q much intersect both $x\in Q_j\ni x$ and $\tilde{\Omega}^c\subset(2Q_j)^c$. Hence $\ell(Q)\geq\frac{1}{2}\ell(Q_j)$. But then $3Q\supset Q_j$, and we have

$$\Big(\frac{1}{|Q|}\int_Q1_{\tilde{\Omega}^c}w^r\Big)^{1/r}\leq\Big(\frac{3^d}{|3Q|}\int_{3Q}1_{\tilde{\Omega}^c}w^r\Big)^{1/r}\leq C\inf_{Q_j}M_rw.$$

Thus

$$\begin{aligned}\sum_j|\langle f\rangle_{Q_j}|\int_{Q_j}M_r(1_{\tilde{\Omega}^c}w) &\leq C\sum_j|\langle f\rangle_{Q_j}|\cdot|Q_j|\inf_{Q_j}M_rw\\ &\leq C\sum_j\|1_{Q_j}f\|_{L^1}\cdot\inf_{Q_j}M_rw\leq C\sum_j\|1_{Q_j}f\|_{L^1(M_rw)}\leq\|f\|_{L^1(M_rw)}.\end{aligned}$$

Altogether, we have now shown that

$$w(\tilde{\Omega}^c \cap \{|Tg| > \tfrac{1}{2}\lambda\}) \leq (Cpp')^p (r')^{p-1} \frac{1}{\lambda} \|g\|_{L^1\left(M_r(1_{\tilde{\Omega}^c} w)\right)} \leq (Cpp')^p (r')^{p-1} \frac{1}{\lambda} \|f\|_{L^1(M_r w)},$$

and it remains to choose suitable p and r. First, let again $r := r(w) = 1 + 2^{-d-3}/[w]_{A_\infty}$, so that

$$\|f\|_{L^1(M_r w)} \leq 2\|f\|_{L^1(Mw)} \leq 2[w]_{A_1} \|f\|_{L^1(Mw)},$$

as before. Second, let $p := 1 + 1/\log(1 + [w]_{A_\infty}) \leq c$, so that $p' = 1 + \log(1 + [w]_{A_\infty})$ and

$$\begin{aligned}(Cpp')^p (r')^{p-1} &\leq C\big(1 + \log(1 + [w]_{A_\infty})\big)^{1+1/\log(1+[w]_{A_\infty})} (1 + 2^{d+3}[w]_{A_\infty})^{1/\log(1+[w]_{A_\infty})} \\ &\leq C(1 + \log[w]_{A_\infty}),\end{aligned}$$

by using the fact that $t^{1/t} \leq e^{1/e}$. These estimates combined, we have

$$w(\tilde{\Omega}^c \cap \{|Tg| > \tfrac{1}{2}\lambda\}) \leq (Cpp')^p (r')^{p-1} \frac{1}{\lambda} \cdot \|f\|_{L^1(M_r w)} \leq \frac{C}{\lambda}(1 + \log[w]_{A_\infty}) \cdot [w]_{A_1} \|f\|_{L^1(w)},$$

and this completes the proof. Note that the "good" part was the only one which produced the additional factor $1 + \log[w]_{A_\infty}$. □

8. Conclusion

These lectures have given an overview, by no means exhaustive, of a number of aspects in the recently active area of quantitative weighted norm inequalities. As mentioned in the beginning, we have concentrated on the one-weight theory, barely touching the two-weight theory, where many basic questions still remain open.

In terms of one-weight theory, the aim has been to survey a number of different types of questions, and not necessarily the most general results available. Thus, for the quantitative A_p bounds for singular integrals, we proved the dyadic representation theorem of Vagharshakyan, and the resulting sharp A_p bounds for one-dimensional convolution operators. The most general results in this direction are the following ones essentially due to myself (2010) [10]; the following formulation of Theorem 8.1 being from Hytönen–Pérez–Treil–Volberg (2010) [13]:

Theorem 8.1 (General dyadic representation theorem [10, 13])**.** *Let T be any Calderón–Zygmund operator. Then there exist dyadic shift of $Ш^\omega_{m,n}$ of parameters (m,n) for all $m,n \in \mathbb{N}$ such that*

$$\langle g, Tf\rangle = c \sum_{m,n=0}^{\infty} 2^{-(m+n)\theta/2} \int_\Omega \langle g, Ш^\omega_{m,n} f\rangle \,\mathrm{d}\mathbb{P}(\omega)$$

for all bounded, compactly supported functions f and g. Here all dyadic shifts with $(m,n) \in \mathbb{N}^2 \setminus \{(0,0)\}$ are of the cancellative type, but the shifts with parameters $(0,0)$ can be of the non-cancellative type; cf. Remark 4.1.

Theorem 8.2 (The A_2 theorem [10])**.** *Let T be any Calderón–Zygmund operator. Then it satisfies*

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

By extrapolation, we also have

$$\|Tf\|_{L^p(w)} \leq C[w]_{A_p}^{\max\{1,1/(p-1)\}} \|f\|_{L^p(w)} \qquad \forall p \in (1,\infty),\ \forall w \in A_p.$$

Before being proven in July 2010, this was known as the A_2 *conjecture*. Theorem 8.2 is established with the help of Theorem 8.1, and this is the only known way at the time of writing. Notice that Theorem 8.2 is not a direct corollary of Theorem 8.1 by the results we have proven in these lectures, because these estimates,

$$\|Ш^\omega_{m,n} f\|_{L^2(w)} \leq C_d(m,n)[w]_{A_2} \|f\|_{L^2(w)},$$

where $C_d(m,n)$ is exponential in (m,n), are too weak to give convergence of the infinite series above. For the proof of Theorem 8.2, it was necessary to improve this estimate so as to have good dependence not only on $[w]_{A_2}$ but also on (m,n).

The End

References

[1] Stephen M. Buckley. Estimates for operator norms on weighted spaces and reverse Jensen inequalities. *Trans. Amer. Math. Soc.*, 340(1):253–272, 1993.

[2] Lennart Carleson. On convergence and growth of partial sums of Fourier series. *Acta Math.*, 116:135–157, 1966.

[3] R. R. Coifman and C. Fefferman. Weighted norm inequalities for maximal functions and singular integrals. *Studia Math.*, 51:241–250, 1974.

[4] R. R. Coifman and R. Rochberg. Another characterization of BMO. *Proc. Amer. Math. Soc.*, 79(2):249–254, 1980.

[5] David Cruz-Uribe, José María Martell, and Carlos Pérez. Sharp weighted estimates for approximating dyadic operators. *Electron. Res. Announc. Math. Sci.*, 17:12–19, 2010.

[6] Oliver Dragičević, Loukas Grafakos, María Cristina Pereyra, and Stefanie Petermichl. Extrapolation and sharp norm estimates for classical operators on weighted Lebesgue spaces. *Publ. Mat.*, 49(1):73–91, 2005.

[7] C. Fefferman and E. M. Stein. Some maximal inequalities. *Amer. J. Math.*, 93:107–115, 1971.

[8] R. Fefferman and J. Pipher. Multiparameter operators and sharp weighted inequalities. *Amer. J. Math.*, 119(2):337–369, 1997.

[9] Nobuhiko Fujii. A condition for a two-weight norm inequality for singular integral operators. *Studia Math.*, 98(3):175–190, 1991.

[10] T. Hytönen. The sharp weighted bound for general Calderón–Zygmund operators. Preprint, arXiv:1010.0755 (2010).

[11] T. Hytönen. On Petermichl's dyadic shift and the Hilbert transform. *C. R. Math. Acad. Sci. Paris*, 346(21-22):1133–1136, 2008.

[12] T. Hytönen and C. Pérez. Sharp weighted bounds involving A_∞. Manuscript in preparation (2011).

[13] T. Hytönen, C. Pérez, S. Treil, and A. Volberg. Sharp weighted estimates for dyadic shifts and the A_2 conjecture. Preprint, arXiv:1010.0755 (2010).

[14] Michael T. Lacey, Stefanie Petermichl, and Maria Carmen Reguera. Sharp A_2 inequality for Haar shift operators. *Math. Ann.*, 348(1):127–141, 2010.

[15] Andrei K. Lerner. An elementary approach to several results on the Hardy-Littlewood maximal operator. *Proc. Amer. Math. Soc.*, 136(8):2829–2833, 2008.

[16] Andrei K. Lerner. A pointwise estimate for the local sharp maximal function with applications to singular integrals. *Bull. Lond. Math. Soc.*, 42(5):843–856, 2010.

[17] Andrei K. Lerner, Sheldy Ombrosi, and Carlos Pérez. Sharp A_1 bounds for Calderón-Zygmund operators and the relationship with a problem of Muckenhoupt and Wheeden. *Int. Math. Res. Not. IMRN*, (6):Art. ID rnm161, 11, 2008.

[18] Andrei K. Lerner, Sheldy Ombrosi, and Carlos Pérez. A_1 bounds for Calderón-Zygmund operators related to a problem of Muckenhoupt and Wheeden. *Math. Res. Lett.*, 16(1):149–156, 2009.

[19] Benjamin Muckenhoupt. Weighted norm inequalities for the Hardy maximal function. *Trans. Amer. Math. Soc.*, 165:207–226, 1972.

[20] F. Nazarov, A. Reznikov, V. Vasyunin, and A. Volberg. Weak norm estimates of weighted singular operators and Bellman functions. Manuscript (2010).

[21] 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(5):1355–1375, 2007.

[22] M. C. Reguera. On Muckenhoupt-Wheeden conjecture. Preprint, arXiv:1008.3943 (2010).

[23] M. C. Reguera and C. Thiele. The Hilbert transform does not map $L^1(Mw)$ to $L^{1,\infty}(w)$. Preprint, arXiv:1011.1767 (2010).

[24] José L. Rubio de Francia. Factorization theory and A_p weights. *Amer. J. Math.*, 106(3):533–547, 1984.

[25] Eric T. Sawyer. A characterization of a two-weight norm inequality for maximal operators. *Studia Math.*, 75(1):1–11, 1982.

[26] Armen Vagharshakyan. Recovering singular integrals from Haar shifts. *Proc. Amer. Math. Soc.*, 138(12):4303–4309, 2010.