Dyadic analysis and weights

Lecture notes of a course at the University of Helsinki,
Spring 2017

Tuomas P. Hytönen

Department of Mathematics and Statistics
P.O.B. 68 (Gustaf Hällströmin katu 2b)
FI-00014 University of Helsinki, Finland
tuomas.hytonen@helsinki.fi

May 5, 2017

Abstract

These are the lecture notes of a seven-week course at the University of Helsinki in Spring 2017. They deal with sharp weighted inequalities for Calderón–Zygmund operators using their domination by so-called sparse dyadic operators, incorporating several developments over the period 2015–2017. The first part of the lectures is concerned with the scalar-valued theory and the second part with its extension to vector-valued functions on matrix-weighted spaces. An important omission is the extension of this theory beyond Calderón–Zygmund operators, which has been under rapid development over the same period 2015–2017.

Contents

Chapter 1

Dyadic analysis of Calderón–Zygmund operators

1.1 Dyadic cubes and the maximal operator

We work in the Euclidean space $\mathbb{R}^d$. Its standard dyadic cubes are defined as

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

However, it is useful to adopt an axiomatic approach, where we assume that $\mathscr{D} = \bigcup_{k\in\mathbb{Z}^d} \mathscr{D}_k$ and, for some fixed $\lambda > 0$:

- each $\mathscr{D}_k$ is a partition of $\mathbb{R}^d$ consisting of cubes of sidelength $\lambda \cdot 2^k$ that are translates of each other by an integer-vector multiple of their sidelength;
- if $Q, R \in \mathscr{D}$, then $Q \cap R \in \{Q, R, \varnothing\}$.

The dyadic maximal operator is

$$M_{\mathscr{D}} f := \sup_{Q\in\mathscr{D}} 1_Q \langle |f| \rangle_Q, \qquad \text{where}$$
$$\langle f \rangle_Q := \fint_Q f := \frac{1}{|Q|} \int_Q f.$$

It is immediate that $M_{\mathscr{D}} f \leq Mf$, where M is the Hardy–Littlewood maximal operator,

$$Mf := \sup_{Q \text{ cube}} 1_Q \langle |f| \rangle_Q,$$

so that many boundedness properties of $M_{\mathscr{D}}$ could be deduced from the corresponding properties of M. However, $M_{\mathscr{D}}$ actually enjoys some better properties than M itself, so that it is beneficial to work with it directly. For instance, $M_{\mathscr{D}}$ is essentially insensitive to the underlying measure, and one can almost as easily study

$$M^{\mu}_{\mathscr{D}} f := \sup_{Q\in\mathscr{D}} 1_Q \langle |f| \rangle^{\mu}_Q,$$

where

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

for any locally finite (i.e. $\mu(Q) < \infty$ for all cubes Q) Borel measure μ.

1.1.1 Theorem. *$M_{\mathscr{D}}^\mu : L^1(\mu) \to L^{1,\infty}(\mu)$, where*

$$\|f\|_{L^{1,\infty}(\mu)} := \sup_{t>0} t\cdot\mu(|f|>t).$$

In fact,

$$t\cdot\mu(M_{\mathscr{D}}^\mu f > t) \le \int_{\{M_{\mathscr{D}}^\mu f>t\}} |f|\,\mathrm{d}y \le \|f\|_{L^1(\mu)}.$$

Proof. By definition, we have the equality of sets

$$\{M_{\mathscr{D}}^\mu f > t\} = \bigcup_{Q\in\mathscr{F}_\lambda} Q,$$

where $\mathscr{F}_\lambda := \{Q\in\mathscr{D} : \langle|f|\rangle_Q^\mu > \lambda\}$. Let $\mathscr{F}\subset\mathscr{F}_\lambda$ be any finite subcollection, and let $\mathscr{F}^*$ consist of the maximal cubes in $\mathscr{F}$: those that are not contained in any bigger element of $\mathscr{F}$.

We claim that every cube $Q\in\mathscr{F}$ is contained in some maximal $Q^*\in\mathscr{F}^*$. If Q itself is maximal, this is clear. If not, it means that $Q\subsetneq Q_1$ for some $Q_1\in\mathscr{F}$. If Q_1 is maximal, we are done; else we have $Q_1\subsetneq Q_2$ for some $Q_2\in\mathscr{F}$. Since $\mathscr{F}$ only contains finitely many cubes, after finitely many steps this process must terminate and we have $Q\subsetneq Q_1\subsetneq\ldots\subsetneq Q_n$, where Q_n is not contained in any bigger $Q'\in\mathscr{F}$. But this means that $Q_n\in\mathscr{F}^*$, and we are done.

From the previous claim it follows that

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

On the other hand, we claim that the cubes $Q\in\mathscr{F}^*$ are disjoint. Namely, any two dyadic cubes are either disjoint, or one is contained in the other. But if $Q\subsetneq Q'\in\mathscr{F}^*$, then Q is not maximal and hence not in $\mathscr{F}^*$. This only leaves the possibility of disjointness for two different cubes.

Thus we have

$$\begin{aligned}\mu\Big(\bigcup_{Q\in\mathscr{F}} Q\Big) &= \mu\Big(\bigcup_{Q\in\mathscr{F}^*} Q\Big) = \sum_{Q\in\mathscr{F}^*}\mu(Q) \le \sum_{Q\in\mathscr{F}^*}\frac{1}{\lambda}\int_Q |f|\,\mathrm{d}\mu \\ &= \frac{1}{\lambda}\int_{\bigcup_{Q\in\mathscr{F}^*} Q} |f|\,\mathrm{d}\mu \le \frac{1}{\lambda}\int_{\{M_{\mathscr{D}}^\mu f>t\}} |f|\,\mathrm{d}\mu.\end{aligned}$$

Let then $\mathscr{F}^n$ be an increasing sequence of finite collections such that $\bigcup_{n=1}^\infty \mathscr{F}^n = \mathscr{F}_\lambda$. (Note that the latter set, as a subset of $\mathscr{D}$, is countable.) Then it follows that

$$\mu(M_{\mathscr{D}}^\mu f > t) = \mu\Big(\bigcup_{Q\in\mathscr{F}_\lambda} Q\Big) = \lim_{n\to\infty}\mu\Big(\bigcup_{Q\in\mathscr{F}^n} Q\Big) \le \frac{1}{\lambda}\int_{\{M_{\mathscr{D}}^\mu f>t\}} |f|\,\mathrm{d}\mu,$$

and it is clear that the last integral is bounded by $\|f\|_{L^1(\mu)}$. □

1.1.2 Corollary. *For $p \in (1, \infty)$, we have $M^\mu_{\mathscr{D}} : L^p(\mu) \to L^p(\mu)$ with norm bounded by $p' = p/(p-1)$.*

Proof. We compute

$$\begin{aligned}
\|M^\mu_{\mathscr{D}} f\|^p_{L^p(\mu)} &= \int_0^\infty p t^{p-1} \mu(M^\mu_{\mathscr{D}} f > t)\,\mathrm{d}t \\
&\leq \int_0^\infty p t^{p-1} \frac{1}{t} \int_{\{M^\mu_{\mathscr{D}} f > t\}} |f|\,\mathrm{d}\mu\,\mathrm{d}t \\
&= \int_{\mathbb{R}^d} |f| \int_0^{M^\mu_{\mathscr{D}} f} p t^{p-2}\,\mathrm{d}t\,\mathrm{d}\mu \\
&= \int_{\mathbb{R}^d} |f| \frac{p}{p-1} (M^\mu_{\mathscr{D}} f)^{p-1}\,\mathrm{d}\mu \\
&= p' \|f\|_{L^p(\mu)} \Big(\int_{\mathbb{R}^d} (M^\mu_{\mathscr{D}} f)^{(p-1)p'}\,\mathrm{d}\mu \Big)^{1/p'} \\
&= p' \|f\|_{L^p(\mu)} \|M^\mu_{\mathscr{D}} f\|^{p-1}_{L^p(\mu)}.
\end{aligned}$$

From here the bound follows after dividing both sides by $\|M^\mu_{\mathscr{D}} f\|^{p-1}_{L^p(\mu)}$, provided that this number is finite. To guarantee the finiteness, we could first run the previous computation with a maximal operator defined by a finite collection $\mathscr{F} \subset \mathscr{D}$ instead of $\mathscr{D}$, to conclude that $\|M^\mu_{\mathscr{F}} f\|_{L^p(\mu)} \leq p' \|f\|_{L^p(\mu)}$, and then apply monotone convergence to an increasing family $\mathscr{F}^n$ such that $\mathscr{D} = \bigcup_{n=1}^\infty \mathscr{F}^n$ to get $\|M^\mu_{\mathscr{D}} f\|_{L^p(\mu)} = \lim_{n\to\infty} \|M^\mu_{\mathscr{F}^n} f\|_{L^p(\mu)}$. $\square$

1.2 Sparse collections and operators

A collection $\mathscr{S}$ of sets of finite measure (mostly, cubes or even dyadic cubes) is called γ-*sparse* if there are *disjoint major subsets* $E_S \subset S$ for each $S \in \mathscr{S}$, i.e., $E_S \cap E_{S'} = \varnothing$ if $S \neq S'$ (disjoint) and $|E_S| \geq \gamma |S|$ (major subset). More generally, for a general measure μ, we say that $\mathscr{S}$ is γ-*sparse* with respect to μ if there are disjoint $E_S \subset S$ for each $S \in \mathscr{S}$ such that $\mu(E_S) \geq \gamma \mu(S)$.

A *sparse operator* is an operator of the form

$$T^\mu_{\mathscr{S}} f = \sum_{S \in \mathscr{S}} 1_S \langle f \rangle^\mu_S,$$

where $\mathscr{S} \subset \mathscr{D}$ is a sparse collection of dyadic cubes.

1.2.1 Proposition. *A γ-sparse operator maps $T^\mu_{\mathscr{S}} : L^p(\mu) \to L^{p'}(\mu)$ with norm at most $\gamma^{-1} p p'$.*

Proof. We apply the dualisation

$$\|h\|_{L^p(\mu)} = \sup \Big\{ \int hg\,\mathrm{d}\mu : \|g\|_{L^{p'}(\mu)} \leq 1 \Big\}$$

to $h = T^\mu_{\mathscr{S}} f$ and estimate (assuming first that $f, g \geq 0$)

$$
\begin{aligned}
\int T^\mu_{\mathscr{S}} f \cdot g \,\mathrm{d}\mu &= \int_{\mathbb{R}^d} \sum_{S \in \mathscr{S}} \langle f \rangle^\mu_S 1_S g \,\mathrm{d}\mu \\
&= \sum_{S \in \mathscr{S}} \langle f \rangle^\mu_S \langle g \rangle^\mu_S \mu(S) \\
&\leq \sum_{S \in \mathscr{S}} \inf_{y \in S} M^\mu_{\mathscr{D}} f(y) \inf_{z \in S} M^\mu_{\mathscr{D}} g(z) \frac{\mu(E_S)}{\gamma} \\
&= \frac{1}{\gamma} \sum_{S \in \mathscr{S}} \inf_{y \in S} M^\mu_{\mathscr{D}} f(y) \inf_{z \in S} M^\mu_{\mathscr{D}} g(z) \int_{E_S} \mathrm{d}\mu(x) \\
&\leq \frac{1}{\gamma} \sum_{S \in \mathscr{S}} \int_{E_S} M^\mu_{\mathscr{D}} f(x) M^\mu_{\mathscr{D}} g(x) \,\mathrm{d}\mu(x) \\
&\leq \frac{1}{\gamma} \int_{\mathbb{R}^d} M^\mu_{\mathscr{D}} f(x) M^\mu_{\mathscr{D}} g(x) \,\mathrm{d}\mu(x) \\
&\leq \frac{1}{\gamma} \| M^\mu_{\mathscr{D}} f \|_{L^p(\mu)} \| M^\mu_{\mathscr{D}} g \|_{L^{p'}(\mu)} \\
&\leq \frac{1}{\gamma} \cdot p' \| f \|_{L^p(\mu)} \cdot p \| g \|_{L^{p'}(\mu)}.
\end{aligned}
$$

For general f, g, it is enough to observe that $|T^\mu_{\mathscr{S}} f \cdot g| \leq T^\mu_{\mathscr{S}} |f| \cdot |g|$ and apply the bound already proved for positive function to $|f|$ and $|g|$. $\square$

A variant of the same argument also gives a *weighted inequality* for $T_{\mathscr{S}}$. A function $w \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ with $w(x) \in (0, \infty)$ almost everywhere is called a weight. We usually identify the weight function $x \mapsto w(x)$ with the induced measure $E \mapsto \int_E w(x) \,\mathrm{d}x$, and denote the latter simply by $w(E)$. By a weighted inequality we understand something like

$$
\|Tf\|_{L^2(w)} \leq K \|f\|_{L^2(w)}, \qquad \|f\|_{L^2(w)} = \Big(\int_{\mathbb{R}^d} |f|^2 w \Big)^{1/2}.
$$

For $T = T_{\mathscr{S}}$, this is *not* a special case of the previous result, since we don't put the weight into the operator $T_{\mathscr{S}}$, but only in the space $L^2(w)$. There is a useful reformulation of the previous bound as follows: Given a function σ taking values in $(0, \infty)$, we replace f by $g\sigma$. Note that there is a bijective correspondence between f and g when σ is fixed. Thus the previous inequality is equivalent to

$$
\|T(f\sigma)\|_{L^2(w)} \leq K \|f\sigma\|_{L^2(w)} = K \|f\|_{L^2(\sigma^2 w)} = K \|f\|_{L^2(\sigma)},
$$

provided that we choose $\sigma := w^{-1}$ (the solution of $\sigma^2 w = \sigma$). This σ is called the (L^2-)*dual weight* of w.

We need the notation

$$
[w]_{A_2} := \sup_Q \langle w \rangle_Q \langle w^{-1} \rangle_Q,
$$

where the supremum is over all cubes, and we say that w is an A_2-weight and write $w \in A_2$ if $[w]_{A_2} < \infty$. The dyadic A_2 class $A_2^{\mathscr{D}}$ and the constant $[w]_{A_2^{\mathscr{D}}}$ are defined similarly by restricting the supremum to $Q \in \mathscr{D}$ only.

1.2.2 Theorem (Cruz-Uribe–Martell–Pérez 2010 [CUMP10])**.** *If $w \in A_2^{\mathscr{D}}$, every γ-sparse operator maps $T_{\mathscr{S}} : L^2(w) \to L^2(w)$ with norm at most $4\gamma^{-1}[w]_{A_2}^{\mathscr{D}}$.*

Proof. Let $\sigma := w^{-1}$ be the dual weight. We apply the dualisation

$$\|h\|_{L^2(w)} = \sup\Big\{ \int hg \,\mathrm{d}\mu : \|g\|_{L^2(w)} \leq 1 \Big\}$$

to $h = T_{\mathscr{S}}(f\sigma)$. It will be useful to observe that

$$\langle f\sigma\rangle_S = \frac{1}{|S|}\int_S f\sigma = \frac{\sigma(S)}{|S|}\frac{1}{\sigma(S)}\int_S f\sigma = \langle\sigma\rangle_S\langle f\rangle_S^\sigma.$$

We can then estimate (assuming again that $f, g \geq 0$)

$$\begin{aligned}
\int T_{\mathscr{S}}(f\sigma)\cdot g\cdot w &= \int_{\mathbb{R}^d}\sum_{S\in\mathscr{S}}\langle f\sigma\rangle_S 1_S g\cdot w \\
&= \sum_{S\in\mathscr{S}}\langle f\sigma\rangle_S\langle gw\rangle_S|S| \\
&= \sum_{S\in\mathscr{S}}\langle\sigma\rangle_S\langle w\rangle_S\langle f\rangle_S^\sigma\langle g\rangle_S^w|S| \\
&\leq \sum_{S\in\mathscr{S}}[w]_{A_2^{\mathscr{D}}}\langle f\rangle_S^\sigma\langle g\rangle_S^w\frac{|E_S|}{\gamma} \\
&\leq \frac{[w]_{A_2^{\mathscr{D}}}}{\gamma}\sum_{S\in\mathscr{S}}\inf_{y\in S}M_{\mathscr{D}}^\sigma f(y)\inf_{z\in S}M_{\mathscr{D}}^w g(z)\int_{E_S}\mathrm{d}x \\
&\leq \frac{[w]_{A_2^{\mathscr{D}}}}{\gamma}\sum_{S\in\mathscr{S}}\int_{E_S}M_{\mathscr{D}}^\sigma f(x)M_{\mathscr{D}}^w g(x)\,\mathrm{d}x \\
&\leq \frac{[w]_{A_2^{\mathscr{D}}}}{\gamma}\int_{\mathbb{R}^d}M_{\mathscr{D}}^\sigma f(x)M_{\mathscr{D}}^w g(x)\sigma(x)^{1/2}w(x)^{1/2}\,\mathrm{d}x \\
&\leq \frac{[w]_{A_2^{\mathscr{D}}}}{\gamma}\|M_{\mathscr{D}}^\sigma f\sigma^{1/2}\|_{L^2}\|M_{\mathscr{D}}^w g w^{1/2}\|_{L^2} \\
&= \frac{[w]_{A_2^{\mathscr{D}}}}{\gamma}\|M_{\mathscr{D}}^\sigma f\|_{L^2(\sigma)}\|M_{\mathscr{D}}^w g\|_{L^2(w)} \\
&\leq \frac{[w]_{A_2^{\mathscr{D}}}}{\gamma}\cdot 2\|f\|_{L^2(\sigma)}\cdot 2\|g\|_{L^2(w)}.
\end{aligned}$$

□

1.2.3 Exercise. For a given weight w and $p \in (1,\infty)$, find the L^p dual weight σ such that the inequalities $\|Tf\|_{L^p(w)} \leq K\|f\|_{L^p(w)}$ and $\|T(f\sigma)\|_{L^p(w)} \leq K\|f\|_{L^p(\sigma)}$ (for all f that make the respective right sides finite) are equivalent.

1.2.4 Exercise. Let w be a weight, and consider the operator $f \mapsto 1_Q \langle f \rangle_Q$. Show that the norm of this operator on $L^2(w)$ is $(\langle w \rangle_Q \langle w^{-1} \rangle_Q)^{1/2}$.

(Hint: The upper bound for the norm is basically Cauchy–Schwarz. For the lower bound, think of conditions for equality in Cauchy–Schwarz. To do the lower bound carefully, you may need to consider $(w+\varepsilon)^{-1}$ first and take the limit $\varepsilon \to 0$ in the end; note that w^{-1} is not assumed to be integrable over Q, but you can make the exercise slightly easier by adding this assumption.)

1.2.5 Exercise. For $p \in (1,\infty)$, find the norm of the operator $f \mapsto 1_Q \langle f \rangle_Q$ on $L^p(w)$.

(Hint: this has something in common with both previous exercises.)

1.3 Sharpness and extrapolation

We have proved the following bounds for sparse operators:

- $\|T_{\mathscr{S}}\|_{L^p \to L^p} \leq \gamma^{-1} p p' \leq c_\gamma p$ for $p \geq 2$.
- $\|T_{\mathscr{S}}\|_{L^2(w) \to L^2(w)} \leq 4\gamma^{-1} [w]_{A_2}^{\mathscr{D}} = c_\gamma [w]_{A_2}^{\mathscr{D}}$ for $w \in A_2^{\mathscr{D}}$.

We claim that both these bounds have the optimal rate of growth as a function of the parameters p and $[w]_{A_2}^{\mathscr{D}}$, in the following sense: If we have a bound of the form $\|T_{\mathscr{S}}\|_{L^p \to L^p} \leq \phi(p)$ or $\|T_{\mathscr{S}}\|_{L^2(w) \to L^2(w)} \leq \phi([w]_{A_2^{\mathscr{D}}})$, then $\phi(t) \geq ct$.

We begin with an example in the unweighted case:

1.3.1 Example. Consider the sparse collection $\mathscr{S} = \{[0, 2^{-k}) : k = 0, 1, 2, \ldots\}$ and $f = 1_{[0,1)}$. Then clearly $\|f\|_{L^p} = 1$ for all p. On the other hand, we have

$$\begin{aligned} T_{\mathscr{S}} f &= \sum_{k=0}^{\infty} 1_{[0,2^{-k})} \langle f \rangle_{[0,2^{-k})} = \sum_{k=0}^{\infty} 1_{[0,2^{-k})} \overset{\text{a.e.}}{=} \sum_{k=0}^{\infty} \sum_{j=k}^{\infty} 1_{[2^{-j-1}, 2^{-j})} \\ &= \sum_{j=0}^{\infty} 1_{[2^{-j-1}, 2^{-j})} \sum_{k=0}^{j} 1 = \sum_{j=0}^{\infty} 1_{[2^{-j-1}, 2^{-j})} (1+j), \end{aligned}$$

and hence

$$\|T_{\mathscr{S}} f\|_{L^p}^p = \sum_{j=0}^{\infty} 2^{-1-j} (1+j)^p.$$

To estimate this from below, let j_p be the unique integer such that $j_p \leq p < j_p + 1$. Then $2^{-j_p} \geq 2^{-p}$, and hence

$$\|T_{\mathscr{S}} f\|_{L^p}^p \geq 2^{-1-j_p} (1+j_p)^p \geq 2^{-1} 2^{-p} \cdot p^p = \Big(2^{-1/p} \frac{p}{2}\Big)^p,$$

and hence $\|T_{\mathscr{S}} f\|_{L^p} \geq 2^{-1/p} p/2 \geq p/4$ for $p \geq 1$. So if $\|T_{\mathscr{S}}\|_{L^p \to L^p} \leq \phi(p)$, then $\phi(p) \geq p/4$.

It is interesting that the sharpness of the weighted inequality can be deduced indirectly from the previous example, without the need to give any example in the weighted case. This is a consequence of the following extrapolation result. It is most naturally stated in terms of another weight constant:

$$[w]_{A_1} := \Big\| \frac{Mw}{w} \Big\|_\infty .$$

The dyadic version $[w]_{A_1}^{\mathscr{D}}$ is defined by using $M_{\mathscr{D}}$ in place of M.

1.3.2 Theorem (R. Fefferman–Pipher 1997 [FP97])**.** *Let T be an operator that satisfies*

$$\|T\|_{L^2(w)\to L^2(w)} \leq \phi([w]_{A_1})$$

for all $w \in A_1$. Then

$$\|T\|_{L^p\to L^p} \leq \sqrt{2}\phi(c_d p)$$

for $p \in [2,\infty)$. Here c_d is a dimensional constant, and may be taken as $c_d = 1$ if A_1 is replaced by $A_1^{\mathscr{D}}$ in the assumption.

Before proving the theorem, we note that it applies in particular to the situation where we have a similar bound in terms of A_2 rather than A_1:

1.3.3 Lemma. *$A_1 \subset A_2$ and $[w]_{A_2} \leq [w]_{A_1}$ (both dyadic and non-dyadic cases).*

Proof. Note that $\langle w\rangle_Q \leq \inf_{z\in Q} Mw(z) \leq [w]_{A_1} \inf_{z\in Q} w(z)$, and hence

$$\langle w\rangle_Q \langle w^{-1}\rangle_Q \leq [w]_{A_1} \fint_Q \inf_{z\in Q} w(z) w^{-1}(x)\,\mathrm{d}x \leq [w]_{A_1} \fint_Q w(x)w^{-1}(x)\,\mathrm{d}x = [w]_{A_1}.$$

□

In particular, we deduce the sharpness of our estimate for $\|T_{\mathscr{S}}\|_{L^2(w)\to L^2(w)}$: Suppose that $\|T_{\mathscr{S}}\|_{L^2(w)\to L^2(w)} \leq \phi([w]_{A_2^{\mathscr{D}}}) \leq \phi([w]_{A_1^{\mathscr{D}}})$. Then the theorem implies that $\|T_{\mathscr{S}}\|_{L^2(w)\to L^2(w)} \leq \sqrt{2}\phi(p)$. But our example shows that in this case $\sqrt{2}\phi(p) \geq p/4$, so that $\phi(p) \geq cp$ with $c = (\sqrt{2}\cdot 4)^{-1}$.

Proof of the Theorem. We prove both the dyadic and non-dyadic versions of the theorem simultaneously, understanding that all relevant objects (weights, maximal functions etc.) are taken to be dyadic in the dyadic version, and non-dyadic in the non-dyadic version.

We need to estimate

$$\|Tf\|_{L^p} = \||Tf|^2\|_{L^{p/2}}^{1/2} = \sup\Big\{ \Big(\int_{\mathbb{R}^d} |Tf|^2 g \Big)^{1/2} : \|g\|_{(p/2)'} \leq 1 \Big\}.$$

Note that the last expression is formally a weighted L^2 norm of Tf, but the function g need not be in the relevant weight class to apply the assumptions. This problem is fixed by applying to g the so-called *Rubio de Francia algorithm*

$$R_q g := \sum_{k=0}^\infty 2^{-k} \|M\|_{L^q\to L^q}^{-k} M^k g,$$

where

$$M^0 g := |g|, \qquad M^k g := M(M^{k-1}g),$$

and M is the maximal operator. The function $R_q g$ has three important properties:

1. $R_q g \geq |g|$. (Indeed, the zeroth term of the positive series is simply $|g|$.)
2. $\|R_q g\|_{L^q} \leq 2\|g\|_{L^q}$. (The L^q norm of the kth term is dominated by $2^{-k}\|g\|_{L^q}$ after using the boundedness of the maximal operator and cancelling the norms.)
3. $[R_q g]_{A_1} \leq 2\|M\|_{L^q \to L^q}$. (Applying M to the defining series, we obtain a similar series with $M^{k+1}g$ in place of $M^k g$. Making a change of variables, we find that $M(R_q g)$ is almost the same series as $R_q g$, only starting at $k = 1$ and multiplied by $2\|M\|_{L^q \to L^q}$.)

With these properties, it is easy to conclude (let $q := (p/2)'$)

$$\begin{aligned}
\Big(\int_{\mathbb{R}^d} |Tf|^2 g\Big)^{1/2} &\leq \Big(\int_{\mathbb{R}^d} |Tf|^2 R_q g\Big)^{1/2} \\
&\leq \phi([R_q g]_{A_1})\Big(\int_{\mathbb{R}^d} |f|^2 R_q g\Big)^{1/2} \\
&\leq \phi([R_q g]_{A_1})\Big(\int_{\mathbb{R}^d} (|f|^2)^{p/2}\Big)^{1/p}\Big(\int_{\mathbb{R}^d} (R_q g)^q\Big)^{1/2q} \\
&= \phi([R_q g]_{A_1})\|f\|_{L^p}\|R_q g\|_{L^q}^{1/2},
\end{aligned}$$

where $\|R_q g\|_{L^q}^{1/2} \leq (2\|g\|_{L^q})^{1/2} \leq \sqrt{2}$ and, recalling that $q = (p/2)'$ so that $q' = p/2$,

$$[R_q g]_{A_1} \leq 2\|M\|_{L^q \to L^q} \leq \begin{cases} 2q' = p, & \text{in the dyadic case,} \\ 2c_d q' = c_d p, & \text{in the non-dyadic case.} \end{cases}$$

In the dyadic case we used the bound for the dyadic maximal operator that we proved earlier. We take for granted the above bound for the Hardy–Littlewood maximal operator from Real Analysis. (This can also be deduced from the dyadic version by the method of 'parallel dyadic cubes' that we might discuss later.) Putting together these estimates, are proof is complete. □

The next two exercises are preparations for continuous (non-dyadic) versions of the sharp inequalities for sparse operators studied above.

1.3.4 Exercise. Check that the function $f = 1_{(0,1)}(x)\log x$ satisfies $\|f\|_{L^p} \geq cp$ for $p \in [1, \infty)$.

1.3.5 Exercise. Compute the Hilbert transform

$$Hf(x) := \lim_{\varepsilon \to 0}\Big(\int_{-\infty}^{x-\varepsilon} + \int_{x+\varepsilon}^{\infty}\Big)\frac{f(y)}{x-y}\,\mathrm{d}y$$

of $f = 1_{(0,1)}$, and deduce that $\|H\|_{L^p \to L^p} \geq cp$ for $p \in [2, \infty)$.

1.3.6 Exercise. Prove an analogue of Theorem 1.3.2 starting from the assumption that $\|T\|_{L^s(w)\to L^s(w)} \leq \phi([w]_{A_1})$ for a fixed $s \neq 2$.

1.3.7 Remark. R. Fefferman and Pipher [FP97] used their Theorem 1.3.2 "directly", to deduce L^p bounds from $L^2(w)$ bounds. The idea of using it "backwards", to deduce the sharpness of weighted estimates from the sharpness of unweighted ones, was introduced by Luque, Pérez and Rela 2013 [LPR15].[1]

1.4 Lerner's abstract domination theorem

Besides linear operators, we frequently encounter *positive sublinear* operators. By this we mean that for all functions f and g we have that $Tf \geq 0$ is a non-negative function, $T(\alpha f) = |\alpha| Tf$ for constants α, and $T(f+g) \leq Tf + Tg$. A prime example is the maximal operator M and its dyadic version. We will see other examples later. Let us record a useful observation:

1.4.1 Lemma. *If T is linear or positive sublinear, then*

$$|Tf - Tg| \leq |T(f-g)|.$$

Proof. For linear operators this is clear, since $Tf - Tg = T(f-g)$. In the positive sublinear case, let us fix a point x and assume first that $Tf(x) \geq Tg(x)$. Then

$$Tf(x) = T(f-g+g)(x) \leq T(f-g)(x) + Tg(x)$$

so that

$$|Tf(x) - Tg(x)| = Tf(x) - Tg(x) \leq T(f-g)(x).$$

But if $Tg(x) > Tf(x)$, the same argument show that

$$|Tg(x) - Tf(x)| \leq T(g-f)(x) = T((-1)(f-g))(x) = T(f-g)(x),$$

giving the same bound in both cases. □

1.4.2 Theorem (Lerner 2015 [Ler16])**.** *Let T be linear or positive sublinear, and consider the associated Lerner's maximal operator*

$$M_T f(x) := \sup_{Q \ni x} \sup_{y \in Q} |T(1_{(3Q)^c} f)(y)|.$$

Suppose that both T and M_T are bounded from L^1 to $L^{1,\infty}$. Then for every boundedly supported $f \in L^1(\mathbb{R}^d)$ and $\varepsilon \in (0,1)$, there is a $(1-\varepsilon)$-sparse family $\mathscr{S}$ of dyadic cubes such that

$$|Tf| \leq \frac{c_d c_T}{\varepsilon} \sum_{S \in \mathscr{S}} 1_S \fint_{3S} |f|,$$

where c_d depends only on dimension and

$$c_T := \|T\|_{1 \to 1,\infty} + \|M_T\|_{1\to 1,\infty}.$$

[1] When possible, we have dated results according to their preprint year (here: 2013), which is usually strictly earlier than the publication year (here: 2015).

1.4.3 Remark. Despite the "maximal" character of M_T, it is not necessarily bigger than T itself. E.g., if $T = I$ is the identity operator, then $M_T = 0$.

The heart of the Theorem is contained in the following lemma:

1.4.4 Lemma. *Under the assumptions of Theorem 1.4.2, for any cube Q_0 and $f \in L^1(3Q_0)$ and $\varepsilon \in (0,1)$, there are disjoint subcubes $Q'_j \in \mathscr{D}(Q_0)$ such that*

$$\sum_j |Q'_j| \leq \varepsilon |Q_0|.$$

and, if $Q_j \in \mathscr{D}(Q_0)$ are (possibly bigger) disjoint cubes such that $\bigcup_j Q_j \supset \bigcup_j Q'_j$, then

$$\Big|1_{Q_0} T(1_{3Q_0} f) - \sum_j 1_{Q_j} T(1_{3Q_j} f)\Big| \leq 1_{Q_0} \frac{c_d c_T}{\varepsilon} \fint_{3Q_0} |f|.$$

1.4.5 Remark. For the purposes of Theorem 1.4.2, the case $Q_j = Q'_j$ of the Lemma suffices. The possibility of allowing bigger cubes is relevant for a vector-valued generalisation of Theorem 1.4.2.

Proof. Given a cube Q_0, consider any disjoint family of its subcubes $Q_j \in \mathscr{D}(Q_0)$ and write $\Omega := \bigcup Q_j$. Then we have the obvious identity

$$1_{Q_0} T(1_{3Q_0} f) = 1_{Q_0 \setminus \Omega} T(1_{3Q_0} f) + \sum_j 1_{Q_j} T(1_{3Q_0} f).$$

and hence

$$1_{Q_0} T(1_{3Q_0} f) - \sum_j 1_{Q_j} T(1_{3Q_j} f) = 1_{Q_0 \setminus \Omega} T(1_{3Q_0} f) + \sum_j 1_{Q_j} (T(1_{3Q_0} f) - T(1_{3Q_j} f))$$

Taking absolute values and using Lemma 1.4.1, we have

$$\begin{aligned}
&\Big|1_{Q_0} T(1_{3Q_0} f) - \sum_j 1_{Q_j} T(1_{3Q_j} f)\Big| \\
&= 1_{Q_0 \setminus \Omega} |T(1_{3Q_0} f)| + \sum_j 1_{Q_j} |T(1_{3Q_0} f) - T(1_{3Q_j} f)| \\
&\leq 1_{Q_0 \setminus \Omega} |T(1_{3Q_0} f)| + \sum_j 1_{Q_j} |T(1_{3Q_0} f - 1_{3Q_j} f)| \\
&= 1_{Q_0 \setminus \Omega} |T(1_{3Q_0} f)| + \sum_j 1_{Q_j} |T(1_{3Q_0 \setminus 3Q_j} f)|
\end{aligned}$$

Hence, we need to prove that

$$1_{Q_0 \setminus \Omega} |T(1_{3Q_0} f)| + \sum_j 1_{Q_j} |T(1_{3Q_0 \setminus 3Q_j} f)| \leq 1_{Q_0} \frac{c_d c_T}{\varepsilon} \fint_{3Q_0} |f|. \tag{1.4.6}$$

For a $\lambda > 0$ to be chosen, let us define a preliminary candidate for the set Ω by

$$\Omega'' := Q_0 \cap \{|T(1_{3Q_0} f)| > \lambda \text{ or } M_T(1_{3Q_0} f) > \lambda\}.$$

Thus, by the assumed L^1 to $L^{1,\infty}$ bounds,

$$\begin{aligned}|\Omega''| &\leq |\{|T(1_{3Q_0}f)| > \lambda\}| + |\{M_T(1_{3Q_0}f) > \lambda\}| \\ &\leq \frac{1}{\lambda}\|T\|_{1\to 1,\infty}\|1_{3Q_0}f\|_1 + \frac{1}{\lambda}\|M_T\|_{1\to 1,\infty}\|1_{3Q_0}f\|_1 \\ &= \frac{3^d}{\lambda}c_T \fint_{3Q_0}|f|\cdot|Q_0|.\end{aligned} \tag{1.4.7}$$

Let then $Q'_j \in \mathscr{D}(Q_0)$ be the maximal dyadic subcubes such that

$$\frac{|Q'_j \cap \Omega''|}{|Q'_j|} > 2^{-d-1}.$$

Then the cubes Q'_j are disjoint, and $\Omega' := \bigcup_j Q'_j = \{M_d(1_{\Omega'}) > 2^{-d-1}\}$, so that

$$|\Omega'| \leq 2^{d+1}\|1_{\Omega''}\|_1 = 2^{d+1}|\Omega''| \leq \frac{2\cdot 6^d}{\lambda}c_T \fint_{3Q_0}|f|\cdot|Q_0| = \varepsilon|Q_0| \tag{1.4.8}$$

if we choose

$$\lambda := \frac{2\cdot 6^d}{\varepsilon}c_T \fint_{3Q_0}|f|.$$

Since $1_{\Omega''} \leq M_d(1_{\Omega''})$ almost everywhere, we see that Ω'' is contained in Ω', except perhaps for a subset of measure zero. In particular, if $\Omega := \bigcup_j Q_j \supset \bigcup_j Q'_j = \Omega' \supset \Omega''$, then we have

$$1_{Q_0\setminus\Omega}|T(1_{3Q_0}f)| \leq 1_{Q_0\setminus\Omega}\lambda. \tag{1.4.9}$$

On the other hand, the maximality of Q'_j implies that its dyadic parent $\hat{Q}'_j$ satisfies the opposite inequality, and hence

$$\frac{|Q'_j \cap \Omega''|}{|Q'_j|} \leq \frac{|\hat{Q}'_j \cap \Omega'|}{2^{-d}|\hat{Q}'_j|} \leq \frac{2^{-d-1}}{2^{-d}} = \frac{1}{2}.$$

Thus $|Q'_j\setminus\Omega''| \geq \frac{1}{2}|Q'_j| > 0$, so in particular Q'_j intersects $(\Omega'')^c$. When $\bigcup_j Q'_j \subset \bigcup Q_i$, then also any $Q_i \supset Q'_j$ intersects $(\Omega'')^c$. This means that there exists some $z \in Q_i$ such that

$$\lambda \geq M_T(1_{3Q_0}f)(z) = \sup_{Q\ni z}\sup_{y\in Q}|T(1_{3Q_0\setminus 3Q}f)(y)| \geq \sup_{y\in Q_i}|T(1_{3Q_0\setminus 3Q_i}f)(y)|,$$

and hence

$$1_{Q_i}|T(1_{3Q_0\setminus 3Q_i}f)| \leq 1_{Q_i}\lambda.$$

A combination of (1.4.9) and the previous bound gives the required estimate (1.4.6), recalling the definition of λ, and this completes the proof of the Lemma. □

Proof of Theorem 1.4.2. We first consider $1_{Q_0}T(1_{3Q_0}f)$ for a fixed cube Q_0. By Lemma 1.4.4, we have

$$1_{Q_0}|T(1_{3Q_0}f)| \leq 1_{Q_0}\frac{c_d c_T}{\varepsilon}\fint_{3Q_0}|f| + \sum_j 1_{Q_j^1}|T(1_{3Q_j^1}f)|, \qquad \sum_j |Q_j^1| \leq \varepsilon|Q_0|.$$

Applying the same estimate to each Q_j^1 in place of Q_0, and continuing by induction, we obtain

$$1_{Q_0}|T(1_{3Q_0}f)| \leq \frac{c_d c_T}{\varepsilon}\sum_{n=0}^{N-1}\sum_j 1_{Q_j^n}\fint_{3Q_j^n}|f| + \sum_k 1_{Q_k^N}|T(1_{3Q_k^N}f)|, \tag{1.4.10}$$

where Q_0 is the unique cube of the form Q_j^0, and the cubes of the form Q_j^n are subcubes of some Q_i^{n-1} in such a way that

$$\sum_{j:Q_j^n\subset Q_i^{n-1}}|Q_j^n| \leq \varepsilon|Q_i^{n-1}|.$$

In particular,

$$\sum_j |Q_j^n| \leq \varepsilon\sum_i |Q_i^{n-1}| \leq \ldots \leq \varepsilon^n|Q_0|,$$

so that the support of the last term in (1.4.10) becomes negligible in the limit $N\to\infty$. Thus, almost everywhere, we have

$$1_{Q_0}|T(1_{3Q_0}f)| \leq \frac{c_d c_T}{\varepsilon}\sum_{n=0}^{\infty}\sum_j 1_{Q_j^n}\fint_{3Q_j^n}|f|, \tag{1.4.11}$$

where the pairwise disjoint subsets

$$E_j^n := Q_j^n \setminus \bigcup_k Q_k^{n+1}$$

have measure $|E_j^n| \geq (1-\varepsilon)|Q_j^n|$.

Let us finally use this local estimate to deduce a global bound. Fix a boundedly supported $f\in L^1(\mathbb{R}^d)$ and consider the maximal dyadic cubes Q with the property that Q does not contain the support of f. We claim that these cubes form a partition of $\mathbb{R}^d$, *provided* that we use a dyadic system with the following additional property:

(1.4.12) Every bounded set is contained in some dyadic cube $Q\in\mathscr{D}$.

Note that this is *not* satisfied by the standard dyadic system: e.g., all balls centred at the origin are not contained in any dyadic cube. However, it is not difficult to construct dyadic systems with this special property: In one dimension, starting from $I_0 := [0,1)$, let $I_{k+1} := I_k + (-1)^k|I_k|$ (so that we extend the previous interval to the left and to the right alternatingly). We can then define a dyadic system $\mathscr{D}$ by taking all intervals obtained from the intervals I_k by either shifting them by an integer multiple of their side-length, or dividing them into halves arbitrarily many times. Then it is easy to check that any bounded set is contained in some interval of $\mathscr{D}$; in fact, in some I_k. A construction in $\mathbb{R}^d$ can be achieved by considering cubes of the form $Q = J_1 \times \cdots \times J_d$, where the intervals have equal length and belong to the one-dimensional system just described.

Returning to the proof, if f is nontrivial, then sufficiently small cubes cannot contain its support, and hence every point x is contained in some Q with $Q \not\supset \operatorname{supp} f$. On the other hand, we assumed that every bounded set is contained in some dyadic cube Q. Applying this to the bounded set $\operatorname{supp} f \cup Q$, where Q is any given dyadic cube, we we see that Q is contained in a dyadic cube Q' that also contains $\operatorname{supp} f$, and hence all sufficiently large dyadic $Q' \supset Q$ violate the condition that $Q' \not\supset \operatorname{supp} f$. Thus, every Q with $Q \not\supset \operatorname{supp} f$ is contained in a maximal cube with this property.

Let us denote the disjoint collection of cubes just discussed by $\mathscr{S}_0$. For every $S \in \mathscr{S}_0$, maximality implies that $3S \supset \hat{S} \supset \operatorname{supp} f$, and hence $f = 1_{3S} f$. Then the partition property implies that

$$|Tf| = \sum_{S \in \mathscr{S}_0} 1_S |Tf| = \sum_{S \in \mathscr{S}_0} 1_S |T(1_{3S} f)|.$$

Applying (1.4.11) to each S in place of Q_0, we obtain

$$|Tf| \leq \frac{c_d c_T}{\varepsilon} \sum_{S \in \mathscr{S}} 1_S \fint_{3S} |f|,$$

where $\mathscr{S}$ consists of all cubes of the form Q_j^n, starting from some $Q_0 = S \in \mathscr{S}_0$. This collection is $(1-\varepsilon)$-sparse, by the observations above. □

1.5 Triples of dyadic cubes

1.5.1 Proposition. *Suppose that $\mathscr{D}$ is a system of dyadic cubes on $\mathbb{R}^d$. Then $\{3Q : Q \in \mathscr{D}\}$ can be divided into 3^d subcollections, each of which has the same covering and nestedness properties as $\mathscr{D}$.*

Proof. Let us first consider $d = 1$.

We introduce the notation

$$\tilde{I} := \text{the neighbour of } \hat{I} \text{ that touches } I.$$

The relevance of this definition comes from the basic observation (best seen by drawing a picture) that

$$3I \subset 3\tilde{I}.$$

Now suppose that $\mathscr{C}$ is a subcollection of $\mathscr{D}$ such that

- it contains every third interval of each fixed length scale, and

- if $I \in \mathscr{C}$, then $\tilde{I} \in \mathscr{C}$.

It is clear from the first property that the triples $3I$ of $I \in \mathscr{C}$ of a fixed length form a partition of $\mathbb{R}$. Suppose then that $I, J \in \mathscr{C}$ and $3I \cap 3J \neq \varnothing$. Suppose for instance that $\ell(I) \leq \ell(J)$, let $\ell(I) = 2^{-k}\ell(J)$, and let $\tilde{I}^{(k)} := (\tilde{I}^{(k-1)})^{\sim}$ (with $\tilde{I}^{(0)} := I$) be the k-fold version of $I \mapsto \tilde{I}$. Then by induction $3I \subset 3\tilde{I}^{(k)}$, where $I^{(k)} \in \mathscr{C}$ (by the second property) is an interval of the same length as J. Now $\varnothing \neq 3I \cap 3J \subset 3\tilde{I}^{(k)} \cap 3J$, and hence the partition property implies that $J = I^{(k)}$. But then $3I \subset 3\tilde{I}^{(k)} = 3J$, proving the nestedness.

So it remains to check that we can divide $\mathscr{D}$ into three collections like $\mathscr{C}$.

We introduce a relation on $\mathscr{D}$ as follows. For intervals I, J of equal length $\ell(I) = \ell(J)$, we say that $I \sim J$ if and only if $I = J \dot{+} 3m$ for some $m \in \mathbb{Z}$.

Consider two consecutive intervals $I, J = I \dot{+} 3$ of some equivalence class. If $I = \hat{I}_\ell$, then $J = \hat{J}_r$ and $\hat{J} = \hat{I} \dot{+} 1$. (It is instructive to draw a picture.) In this case $\tilde{I} = \hat{I} \dot{-} 1$, and $\tilde{J} = \hat{J} \dot{+} 1 = \hat{I} \dot{+} 2 = \tilde{I} \dot{+} 3$. If $I = \hat{I}_r$, then $J = \hat{J}_\ell$ and $\hat{J} = \hat{I} \dot{+} 2$. In this case $\tilde{I} = \hat{I} \dot{+} 1$ and $\tilde{J} = \hat{J} \dot{-} 1 = \hat{I} \dot{+} 1 = \tilde{I}$. Thus, $\tilde{J}$ is either $\tilde{I}$ or $\tilde{I} \dot{+} 3$ depending on the relative position of I and its parent interval, but in either case we have $\tilde{J} \sim \tilde{I}$. This argument for consecutive intervals easily extends by induction to any equal-length intervals $I \sim J$, showing that we always have $\tilde{I} \sim \tilde{J}$ as well. Similarly, it is easy to check that if $\tilde{I} \sim \tilde{J}$, then also $I \sim J$.

Let us fix a unit interval I_0 and for $k \geq 1$, define inductively $I_k := \tilde{I}_{k-1}$ and choose I_{-k} (among two possibilities) such that $\tilde{I}_{-k} = I_{-(k-1)}$. We say that $J \in \mathscr{D}$ is in the class of I_0 if $J \sim I_k$ when $\ell(I_k) = \ell(J)$. By the observations above, it follows that J is in the class of I_0 if and only if $\tilde{J}$ is in the class of I_0. Moreover, at every length scale, exactly every third interval is in the class of I_0. If we choose two other intervals I_0', I_0'' such that $I_0 \not\sim I_0' \not\sim I_0'' \not\sim I_0$ and define the classes of I_0' and I_0'' similarly, then every dyadic interval is in the class of exactly one of I_0, I_0', I_0'', and each of these classes has the properties of the collection $\mathscr{C}$ above.

Thus we have constructed the one-dimensional classes $\mathscr{C}_1^\alpha$, $\alpha = 0, 1, 2$. In d-dimensions, we can simply define

$$\mathscr{C}_d^\alpha = \mathscr{C}_d^{(\alpha_1, \ldots, \alpha_d)} := \{Q = I_1 \times \cdots \times I_d \in \mathscr{D} : I_i \in \mathscr{C}_1^{\alpha_i}\ \forall i = 1, \ldots, d\}.$$

The required properties are easily verified from the one-dimensional versions observing that $3Q = 3I_1 \times \cdots \times 3I_d$ and $Q \cap R = (I_1 \cap J_1) \times \cdots \times (I_d \cap J_d)$ if Q is as above and $R = J_1 \times \cdots \times J_d$. $\square$

1.5.2 Corollary. *Under the assumptions of Theorem 1.4.2, for every compactly supported $f \in L^1(\mathbb{R}^d)$ and $\varepsilon > 0$, there are $3^{-d}(1-\varepsilon)$-sparse subcollections $\mathscr{S}_i \subset \mathscr{D}_i$ of 3^d different dyadic-type collections $\mathscr{D}_i$ such that*

$$|Tf| \leq \frac{c_d c_T}{\varepsilon} \sum_{i=1}^{3^d} T_{\mathscr{S}_i}|f|. \tag{1.5.3}$$

Proof. By Theorem 1.4.2 and the trivial pointwise bound $1_S \leq 1_{3S}$, we have

$$|Tf| \leq \frac{c_d c_T}{\varepsilon} \sum_{S \in \mathscr{S}} 1_S \fint_{3S} |f| \leq \frac{c_d c_T}{\varepsilon} \sum_{S \in \mathscr{S}} 1_{3S} \fint_{3S} |f|$$

Letting $\{3Q : Q \in \mathscr{D}\} = \bigcup_{i=1}^{3^d} \mathscr{D}_i$ be the splitting provided by Proposition 1.5.1, we let $\mathscr{S}_i := \{3S \in \mathscr{D}_i : S \in \mathscr{S}\}$. Then clearly

$$\sum_{S\in\mathscr{S}} 1_{3S} \fint_{3S} |f| = \sum_{i=1}^{3^d} \sum_{R\in\mathscr{S}_i} 1_R \fint_R |f| = \sum_{i=1}^{3^d} T_{\mathscr{S}_i}|f|$$

and the pairwise disjoint sets $E_S \subset S \in \mathscr{S}$ are also good for checking the sparseness of $3S \in \mathscr{S}_i$, since $|E_S| \geq (1-\varepsilon)|S| = 3^{-d}(1-\varepsilon)|3S|$. □

1.5.4 Corollary. *Let $X, Y \subset L^1_{\mathrm{loc}}(\mathbb{R}^d)$ be Banach spaces such that*

- *compactly supported functions $f \in X$ are dense in X, and*
- *if $f \in X$, then $|f| \in X$ and $\||f|\|_X \leq c_X \|f\|_X$.*

If T is an operator that satisfies the assumptions of Theorem 1.4.2, then

$$\|T\|_{X\to Y} \leq c_d c_T c_X \sup_{\mathscr{S}} \|T_{\mathscr{S}}\|_{X\to Y}, \tag{1.5.5}$$

where the supremum is over all 3^{-d-1}-sparse subcollections $\mathscr{S}$ of any system of dyadic cubes on $\mathbb{R}^d$. In particular, such operators satisfy

$$\begin{aligned} \|T\|_{L^p\to L^p} &\leq c_d c_T c_X pp', \qquad p \in (1,\infty), \\ \|T\|_{L^2(w)\to L^2(w)} &\leq c_d c_T c_X [w]_{A_2}, \qquad w \in A_2. \end{aligned}$$

Proof. Let $f \in X$ be compactly supported. Since $X \subset L^1_{\mathrm{loc}}(\mathbb{R}^d)$? this means that $f \in L^1(\mathbb{R}^d)$. By the previous corollary, fixing $\varepsilon = \frac{2}{3}$, there are 3^{-d-1}-sparse subcollection $\mathscr{S}_i \subset \mathscr{D}_i$, where $\mathscr{D}_i$, $i = 1, \ldots, 3^d$, are systems of dyadic cubes on $\mathbb{R}^d$, such that

$$|Tf| \leq c_d c_T \sum_{i=1}^{3^d} T_{\mathscr{S}_i}|f|.$$

Thus

$$\begin{aligned} \|Tf\|_Y &\leq c_d c_T \sum_{i=1}^{3^d} \|T_{\mathscr{S}_i}|f|\|_Y \\ &\leq c_d c_T \sum_{i=1}^{3^d} \|T_{\mathscr{S}_i}\|_{X\to T} \||f|\|_X \\ &\leq c_d c_T 3^d \sup_{\mathscr{S}} \|T_{\mathscr{S}}\|_{X\to T} c_X \|f\|_X. \end{aligned}$$

Since compactly supported functions are dense in X, the operator T may be extended by continuity to all $f \in X$, and the same bound above holds in general. (Note that in the two intermediate steps the collections $\mathscr{S}_i$ depend on f; however, in the right, where we have taken the supremum over all sparse collections, the only dependence on f is via the explicit expression $\|f\|_X$.) □

We conclude this section with applications of Proposition 1.5.1 to the maximal operator.

1.5.6 Lemma. *Let $R \subset \mathbb{R}^d$ be a cube. Then there is a dyadic cube $Q \in \mathscr{D}$ such that $R \subset 3Q$ and $\ell(Q) \leq \ell(R)$.*

Proof. We consider $d = 1$ first. Let J be an interval, and consider dyadic intervals I of length $\ell(I) \leq \ell(J) < 2\ell(I)$. Then J is contained in the union of two such consecutive intervals. If I is either of these intervals, then $J \subset 3I$.

In general, if $R = J_1 \times \cdots \times J_d$, we find dyadic I_i such that $J_i \subset 3I_i$ and $\ell(I_i) \leq \ell(J_i) < 2\ell(I_i)$. Then all I_i have equal length, and therefore $Q := I_1 \times \cdots \times I_d$ is a dyadic cube such that $R \subset 3Q$ and $\ell(Q) \leq \ell(R)$. □

1.5.7 Proposition. *There are 3^d dyadic maximal operators $M_{\mathscr{D}_i}$ such that the Hardy–Littlewood maximal operator M satisfies*

$$Mf \leq 3^d \sup_{1 \leq i \leq 3^d} M_{\mathscr{D}_i} f.$$

Proof. Given $x \in \mathbb{R}^d$ and a cube $R \ni$, let $Q \in \mathscr{D}$ be a such that $R \subset 3Q$ and $\ell(Q) \leq \ell(R)$. Then

$$\fint_R |f| \leq \frac{|3Q|}{|R|} \fint_{3Q} |f| \leq 3^d \fint_{3Q} |f| \leq 3^d M_{\mathscr{D}_i} f(x),$$

where $\mathscr{D}_i \ni 3Q$ is one of the dyadic systems provided by Proposition 1.5.1. □

1.5.8 Corollary.

$$\|M\|_{L^p \to L^p} \leq 9^d p'.$$

Proof.

$$\|Mf\|_{L^p} \leq 3^d \sum_{i=1}^{3^d} \|M_{\mathscr{D}_i} f\|_{L^p} \leq 3^d \sum_{i=1}^{3^d} p' \|f\|_{L^p} = 9^d p' \|f\|_{L^p}.$$

□

1.6 Domination of Calderón–Zygmund operators

The main concrete application of Lerner's abstract domination theorem is to the following class of operators:

1.6.1 Definition. We say that T is a Calderón–Zygmund operator, if T is a bounded linear operator on $L^2(\mathbb{R}^d)$, and it has a representation

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

where the kernel K satisfies

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

and

$$\begin{aligned}&|K(x,y)-K(x',y)|+|K(y,x)-K(y,x')| \\ &\qquad\qquad \leq \omega\Big(\frac{|x-x'|}{|x-y|}\Big)\frac{1}{|x-y|^d}, \qquad |x-y|>2|x-x'|\end{aligned}$$

for some *modulus of continuity* ω, by which we mean an increasing and subadditive (i.e., $\omega(a+b) \leq \omega(a)+\omega(b)$) function $\omega : [0,\infty) \to [0,\infty)$ with $\omega(0)=0$. We say that the kernel or the modulus satisfies the Dini condition if

$$\|\omega\|_{Dini} := \int_0^1 \omega(t)\frac{\mathrm{d}t}{t} < \infty.$$

Recall that Lerner's Theorem 1.4.2 requires that $T, M_T : L^1 \to L^{1,\infty}$. This is a relatively straightforward consequence of the following results in the classical Calderón–Zygmund theory (to which we return in the following section):

- The Hardy–Littlewood maximal operator maps $M : L^1 \to L^{1,\infty}$.
- Each Calderón–Zygmund operator maps $T : L^1 \to L^{1,\infty}$.
- The *maximal truncated Calderón–Zygmund operators*

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

  also map $T_\# : L^1 \to L^{1,\infty}$.

Note that the classical property $T : L^1 \to L^{1,\infty}$ is already "one half" of the requirements of Theorem 1.4.2. The other half concerning M_T is a consequence of $M, T_\# : L^1 \to L^{1,\infty}$ and the following:

1.6.2 Lemma.

$$M_T f(x) \leq T_\# f(x) + c_d(c_K + \|\omega\|_{Dini}) M f(x).$$

Proof. Let $Q \ni x$ and $z \in Q$ be fixed for the moment. We need to estimate

$$T(1_{(3Q)^c} f)(z) = T(1_{(3Q)^c} f)(z) - T_{2\sqrt{d}\ell(Q)} f(x) + T_{2\sqrt{d}\ell(Q)} f(x),$$

where the last term is bounded by $T_\# f(x)$ by definition. On the other hand,

$$T_{2\sqrt{d}\ell(Q)} f(x) = T(1_{B(x,2\sqrt{d}\ell(Q))^c} f)(x),$$

and $3Q \subset B(x, 2\sqrt{d}\ell(Q))$ so that

$$\begin{aligned}
&T(1_{(3Q)^c} f)(z) - T_{2\sqrt{d}\ell(Q)} f(x) \\
&\qquad = \int_{(3Q)^c} K(z,y) f(y)\,\mathrm{d}y - \int_{|y-x|>2\sqrt{d}\ell(Q)} K(x,y) f(y)\,\mathrm{d}y \\
&\qquad = \int_{|y-x|>2\sqrt{d}\ell(Q)} (K(z,y) - K(x,y)) f(y)\,\mathrm{d}y \\
&\qquad\quad + \int_{(3Q)^c \cap B(x,2\sqrt{d}\ell(Q))} K(z,y) f(y)\,\mathrm{d}y =: I + II.
\end{aligned}$$

In term I, we have $z, x \in Q$, thus $|z - x| \leq \sqrt{d}\ell(Q) < \frac{1}{2}|x - y|$, so that

$$\begin{aligned}
|I| &\leq \int_{|y-x|>2\sqrt{d}\ell(Q)} \omega\Big(\frac{\sqrt{d}\ell(Q)}{|x-y|}\Big) \frac{1}{|x-y|^d} |f(y)|\,\mathrm{d}y \\
&\leq \sum_{k=1}^{\infty} \int_{2^k\sqrt{d}\ell(Q)<|x-y|\leq 2^{k+1}\sqrt{d}\ell(Q)} \omega(2^{-k}) \frac{1}{(2^k\sqrt{d}\ell(Q))^d} |f(y)|\,\mathrm{d}y \\
&\leq \sum_{k=1}^{\infty} \omega(2^{-k}) c_d \fint_{B(x,2^{k+1}\sqrt{d}\ell(Q)} |f(y)|\,\mathrm{d}y \\
&\leq \sum_{k=1}^{\infty} \omega(2^{-k}) c_d Mf(x) \leq \|\omega\|_{Dini} c_d Mf(x),
\end{aligned}$$

where the last bound used Exercise 1.6.6 below.

In term II, we have $z \in Q$, $y \in (3Q)^c$, thus $|y - z| \geq \ell(Q)$, so that

$$\begin{aligned}
|II| &\leq \int_{(3Q)^c \cap B(x,2\sqrt{d}\ell(Q))} \frac{c_K}{|y-z|^d} |f(y)|\,\mathrm{d}y \\
&\leq \int_{B(x,2\sqrt{d}\ell(Q))} \frac{c_K}{\ell(Q)^d} |f(y)|\,\mathrm{d}y \\
&\leq c_K c_d \fint_{B(x,2\sqrt{d}\ell(Q))} |f(y)|\,\mathrm{d}y \leq c_K c_d Mf(x).
\end{aligned}$$

Combining the estimates, and taking the supremum over $z \in Q$ and $Q \ni x$, we arrive at the claim of the lemma. □

Thus, taking for granted the above-listed results of classical Calderón–Zygmund theory, we have:

1.6.3 Theorem. *Every Calderón–Zygmund operator T satisfies the assumptions, and hence the conclusions, of Theorem 1.4.2. In particular, every Calderón–Zygmund operator T satisfies*

$$\|T\|_{L^2(w)\to L^2(w)} \leq c_d c_T [w]_{A_2}, \qquad w \in A_2. \tag{1.6.4}$$

1.6.5 Remark. The bound (1.6.4) is known as the A_2 theorem. It was first proved by Hytönen in July 2010 [Hyt12]), although only for the Hölder moduli of continuity $\omega(t) = t^\delta$, $\delta \in (0,1]$; note that Calderón–Zygmund operators are often defined using these Hölder moduli only. In the stated generality of Dini moduli, the theorem was first proved by Lacey in January 2015 [Lac17] and simplified by Lerner in December 2015 [Ler16].

The intermediate step of domination of T by $T_{\mathscr{S}}$ has some history of its own. The sparse operators $T_{\mathscr{S}}$ were used in the proof of some special cases of the A_2 theorem by Cruz-Uribe, Martell and Pérez in January 2010 [CUMP10], but not in the first proof of the full A_2 theorem [Hyt12]. They reappeared in the simple proof of the A_2 theorem by Lerner in 2012 [Ler13], where the norm domination (1.5.5) first appeared. The stronger pointwise domination (1.5.3) was an open question for a while, and was settled independently in 2014 by Conde-Alonso and Rey [CAR16] and by Lerner and Nazarov [LN15]. All these results still made slightly stronger assumptions on the modulus of continuity than the Dini condition.

1.6.6 Exercise. Check that there are constants c, c' such that every modulus of continuity satisfies $c\|\omega\|_{Dini} \leq \sum_{k=1}^{\infty} \omega(2^{-k}) \leq c'\|\omega\|_{Dini}$.

1.6.7 Exercise. Consider Lerner's maximal operator M_T, when $T = M$, the Hardy–Littlewood maximal operator, and show that $M_M f \leq c_d M f$.

1.6.8 Exercise. Prove the analogue of Lemma 1.6.2 for the maximal truncated Calderón–Zygmund operator $T_\#$ in place of the linear Calderón–Zygmund operator T, i.e., prove a pointwise bound for $M_{T_\#}$ which allows to conclude the $L^1 \to L^{1,\infty}$ boundedness of this operator, and hence the A_2 theorem for $T_\#$. (This extension of the A_2 theorem to $T_\#$ was first obtained by Hytönen, Lacey, Martikainen, Orponen, Reguera, Sawyer and Uriarte-Tuero in 2011 [HLM$^+$12] by a rather difficult argument; now it is only a slight variation of the result for T.) Hint: It suffices (why?) to consider T_ε with an arbitrary but fixed $\varepsilon > 0$.

1.6.9 Exercise. Consider again the Hilbert transform H from Exercise 1.3.5. Taking for granted that $H : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is bounded (This can be proved in various ways, but it is not completely trivial.), check that H is a Calderón–Zygmund operator with a modulus of continuity of the form $\omega(t) = ct$ for some constant c. Conclude from the previous results that

$$\|H\|_{L^2(w)\to L^2(w)} \leq c[w]_{A_2}, \qquad w \in A_2, \tag{1.6.10}$$

and argue by extrapolation (without a concrete example) that this dependence on $[w]_{A_2}$ is optimal. (This special A_2 theorem for the Hilbert transform was first proved by Petermichl [Pet07] and encouraged the quest for the general result.)

1.6.11 Exercise. Show the optimality of (1.6.10) by working out the following concrete example (without using extrapolation): Consider the weight $w(x) = |x|^\alpha$, and the function $f(x) = |x|^{-\alpha} 1_{(-1,0)}(x)$ and estimate the quantities $[w]_{A_2}$, $\|f\|_{L^2(w)}$ and $\|Hf\|_{L^2(w)}$. (Hint: For the last one, you only need a lower bound, so it is enough to consider $\|1_{(0,1)} Hf\|_{L^2(w)}$.)

1.7 Some classical Calderón–Zygmund theory

We now review the aspect of the classical Calderón–Zygmund theory that were needed in the previous section.

1.7.1 Proposition (Calderón–Zygmund decomposition)**.** *Given $f \in L^1(\mathbb{R}^d)$ and $\lambda > 0$, there exists a decomposition $f = g + b$, where*

$$\|g\|_\infty \leq 2^d\lambda, \quad \|g\|_1 \leq \|f\|_1, \quad \|g\|_2^2 \leq 2^d\lambda\|f\|_1$$

and $b = \sum_i b_i$, where

$$\operatorname{supp} b_i \subseteq Q_i, \quad \int b_i = 0, \quad \sum_i |Q_i| \leq \frac{1}{\lambda}\|f\|_1, \quad \sum_i \|b_i\|_1 \leq 2\|f\|_1$$

for some dyadic cubes Q_i.

Proof. Let $Q_i \in \mathscr{D}$ be the maximal dyadic cubes such that $\fint_{Q_i} |f| > \lambda$. Then they are pairwise disjoint, and

$$\sum_i |Q_i| = |\{M_{\mathscr{D}} f > \lambda\}| \leq \frac{1}{\lambda}\|f\|_1.$$

We define $b_i := 1_{Q_i}(f - \langle f\rangle_{Q_i})$, whence the first two properties of b_i are clear, and it remains to estimate

$$\sum_i \|b_i\|_1 \leq \sum_i (\|1_{Q_i} f\|_1 + |Q_i||\langle f\rangle_{Q_i}|) \leq \sum_i 2\int_{Q_i} |f| \leq 2\|f\|_1$$

by the disjointness of the cubes. To ensure that $f = g + b$, we must then define

$$g := 1_{(\bigcup_i Q_i)^c} f + \sum_i 1_{Q_i}\langle f\rangle_{Q_i},$$

where the terms are disjointly supported. If $x \in (\bigcup_i Q_i)^c$, then all dyadic cubes $Q \ni x$ satisfy $\fint_Q |f| \leq \lambda$, and thus

$$|g(x)| = |f(x)| = \lim_{\substack{Q \ni x \\ \ell(Q) \to 0}} \fint_Q |f| \leq \lambda$$

at almost every such x by Lebesgue's differentiation theorem. On the other hand, the maximality of Q_i implies that its dyadic parent $\hat{Q}_i$ satisfies the opposite inequality, $\fint_{\hat{Q}_i} |f| \leq \lambda$. Thus

$$|g(x)| = |\langle f\rangle_{Q_i}| \leq \frac{1}{|Q_i|}\int_{Q_i} |f| \leq \frac{|\hat{Q}_i|}{|Q_i|} \cdot \frac{1}{|\hat{Q}_i|}\int_{\hat{Q}_i} |f| \leq 2^d \cdot \lambda$$

for $x \in Q_i$, and we see that $|g(x)| \leq 2^d\lambda$ in both cases. Moreover,

$$\|g\|_1 = \int_{(\bigcup_i Q_i)^c} |f| + \sum_i |Q_i||\langle f\rangle_{Q_i}| \leq \int_{(\bigcup_i Q_i)^c} |f| + \sum_i \int_{Q_i} |f| = \|f\|_1$$

by the disjointness of the cubes. Finally,

$$\|g\|_2^2 = \int |g|^2 \le \|g\|_\infty \|g\|_1 \le 2^d \lambda \|f\|_1. \qquad \square$$

1.7.2 Theorem (Calderón–Zygmund)**.** *If T is a Calderón–Zygmund operator, then*

$$\|T\|_{L^1 \to L^{1,\infty}} \le c_d \Big(\|T\|_{L^2 \to L^2} + \|\omega\|_{\mathrm{Dini}} \Big).$$

Proof. Fix $\lambda > 0$; we need to estimate $\lambda |\{|Tf| > \lambda\}|$.

Let $f = g + b$ the the Calderón–Zygmund decomposition of f at level $\alpha\lambda$ (instead of λ), where α is to be determined. If Q_i are the corresponding cubes, let B_i be the concentric ball of twice the diameter and $\Omega^* := \bigcup_i B_i$. Then

$$|\{|Tf| > \lambda\}| \le |\{|Tg| > \lambda/2\}| + |\{|Tb| > \lambda/2\} \setminus \Omega^*| + |\Omega^*|,$$

where the last term satisfies

$$|\Omega^*| \le \sum_i |B_i| = \sum_i c_d |Q_i| \le \frac{c_d}{\alpha\lambda} \|f\|_1.$$

Moreover,

$$|\{|Tg| > \lambda/2\}| \le \frac{1}{(\lambda/2)^2} \|Tg\|_2^2 \le \frac{4}{\lambda^2} \|T\|_{L^2 \to L^2}^2 \|g\|_2^2 \le \frac{4}{\lambda^2} \|T\|_{L^2 \to L^2}^2 \cdot 2^d \alpha\lambda \|f\|_1.$$

Finally, we estimate the bad part:

$$|\{|Tb| > \lambda/2\} \setminus \Omega^*| \le \int_{(\Omega^*)^c} \frac{|Tb|}{\lambda/2} \le \frac{2}{\lambda} \sum_i \int_{(\Omega^*)^c} |Tb_i| \le \frac{2}{\lambda} \sum_i \int_{(B_i)^c} |Tb_i|.$$

The ith term here is

$$\begin{aligned}\int_{(B_i)^c} |Tb_i(x)| \,\mathrm{d}x &= \int_{(B_i)^c} \Big| \int_{Q_i} K(x,y) b_i(y) \,\mathrm{d}y \Big| \,\mathrm{d}x \\ &= \int_{(B_i)^c} \Big| \int_{Q_i} [K(x,y) - K(x,z_i)] b_i(y) \,\mathrm{d}y \Big| \,\mathrm{d}x,\end{aligned}$$

where z_i is the common centre of the cube Q_i and the ball B_i, and we used the fact that $\int b_i(y)\,\mathrm{d}y = 0$ for the last identity. For $y \in Q_i$ and $x \in (B_i)^c$, we have $|y - z_i| \le \frac{1}{2}\operatorname{diam}(Q_i)$ and $|x - z_i| \ge \frac{1}{2}\operatorname{diam}(B_i) = \operatorname{diam}(Q_i)$. Thus we can use continuity of the kernel to conclude that

$$\begin{aligned}\int_{(B_i)^c} |Tb_i(x)| \,\mathrm{d}x &\le \int_{(B_i)^c} \int_{Q_i} \omega\Big(\frac{\frac{1}{2}\operatorname{diam}(Q_i)}{|x - z_i|} \Big) \frac{1}{|x - z_i|^d} |b_i(y)| \,\mathrm{d}y \,\mathrm{d}x \\ &= \int_{(B_i)^c} \omega\Big(\frac{\frac{1}{2}\operatorname{diam}(Q_i)}{|x - z_i|} \Big) \frac{1}{|x - z_i|^d} \,\mathrm{d}x \cdot \|b_i\|_1 \\ &= c_d \int_{\operatorname{diam}(Q_i)}^\infty \omega\Big(\frac{\frac{1}{2}\operatorname{diam}(Q_i)}{r} \Big) \frac{1}{r^d} r^{d-1} \,\mathrm{d}r \cdot \|b_i\|_1 \\ &= c_d \int_0^{1/2} \omega(t) \frac{\mathrm{d}t}{t} \cdot \|b_i\|_1 \le c_d \|\omega\|_{\mathrm{Dini}} \|b_i\|_1.\end{aligned}$$

Summing over i, we find that

$$\frac{2}{\lambda}\sum_i \int_{(B_i)^c} |Tb_i| \leq \frac{2}{\lambda} c_d \|\omega\|_{\mathrm{Dini}} \sum_i \|b_i\|_1 \leq \frac{2}{\lambda} c_d \|\omega\|_{\mathrm{Dini}} \cdot 2\|f\|_1$$

Altogether, we have

$$|\{|Tf| > \lambda\}| \leq \frac{c_d}{\lambda}\|f\|_1 \Big(\alpha \|T\|^2_{L^2 \to L^2} + \frac{1}{\alpha} + \|\omega\|_{\mathrm{Dini}}\Big), \tag{1.7.3}$$

and choosing $\alpha = 1/\|T\|_{L^2 \to L^2}$ provides the claimed bound. $\square$

1.7.4 Exercise. Suppose that we did the previous proof only with $\alpha = 1$, leading to the bound

$$\|T\|_{L^1 \to L^{1,\infty}} \leq c_d \Big(\|T\|^2_{L^2 \to L^2} + 1 + \|\omega\|_{\mathrm{Dini}}\Big).$$

Apply this to the operator αT in place of T, where $\alpha > 0$ is a constant, and see how the different quantities depend on α to deduce (1.7.3) and thus the statement of Theorem 1.7.2 by this alternative route. (The trick of this exercise is an example of a "scaling argument", which is useful in many contexts.)

We next study the truncated singular integrals

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

and the maximal truncation

$$T_\sharp f(x) := \sup_{\varepsilon > 0} |T_\varepsilon f(x)|.$$

We also need a variant of the maximal operator,

$$M_\delta f := \big(M(|f|^\delta)\big)^{1/\delta}.$$

1.7.5 Theorem (Cotlar's inequality)**.** *For $\delta \in (0,1)$, we have*

$$\begin{aligned} T_\sharp f &\leq c_{d,\delta}\big(\|\omega\|_{\mathrm{Dini}} + \|T\|_{L^1 \to L^{1,\infty}}\big) Mf + c_{d,\delta} M_\delta(Tf) \\ &\leq c_{d,\delta}\big(\|\omega\|_{\mathrm{Dini}} + \|T\|_{L^2 \to L^2}\big) Mf + c_{d,\delta} M_\delta(Tf). \end{aligned}$$

Proof. For a fixed x, we need to estimate $T_\varepsilon f(x)$ uniformly in $\varepsilon > 0$. We first observe the identity, for every $x' \in B(x, \frac{1}{2}\varepsilon)$,

$$\begin{aligned} T_\varepsilon f(x) &= T(1_{B(x,\varepsilon)^c} f)(x) \\ &= [T(1_{B(x,\varepsilon)^c} f)(x) - T(1_{B(x,\varepsilon)^c} f)(x')] + Tf(x') - T(1_{B(x,\varepsilon)} f)(x'). \end{aligned}$$

We take the L^δ average of this over $x' \in B(x, \frac{1}{2}\varepsilon)$. Since L^δ is only a quasi-Banach space, this introduces a constant c_δ.

For the first term, we have

$$\begin{aligned}
&|T(1_{B(x,\varepsilon)^c}f)(x) - T(1_{B(x,\varepsilon)^c}f)(x')| \\
&= \Big| \int_{|x-y|>\varepsilon} [K(x,y) - K(x',y)] f(y)\,\mathrm{d}y \Big| \\
&\le \int_{|x-y|>\varepsilon} \omega\Big(\frac{\varepsilon/2}{|x-y|}\Big) \frac{1}{|x-y|^d} |f(y)|\,\mathrm{d}y \\
&\le \sum_{k=0}^{\infty} \int_{2^k\varepsilon<|x-y|<2^{k+1}} \omega(2^{-1-k}\varepsilon) \frac{1}{(2^k\varepsilon)^d} |f(y)|\,\mathrm{d}y \\
&\le 2^d \sum_{k=0}^{\infty} \omega(2^{-k-1}) \fint_{B(x,2^{k+1}\varepsilon)} |f(y)|\,\mathrm{d}y \\
&\le 2^d Mf(x) \sum_{k=0}^{\infty} \omega(2^{-k-1}) \le c 2^d Mf(x) \|\omega\|_{\mathrm{Dini}}.
\end{aligned}$$

For $Tf(x')$, it is immediate that

$$\Big(\fint_{B(x,\frac{1}{2}\varepsilon)} |Tf(x')|^\delta \,\mathrm{d}x' \Big)^{1/\delta} \le M_\delta(Tf)(x),$$

by definition of M_δ.

Concerning the final term, we note that

$$\begin{aligned}
\Big(\fint_{B(x,\frac{1}{2}\varepsilon)} |T(1_{B(x,\varepsilon)}f)|^\delta \Big)^{1/\delta} &\le \frac{c_\delta}{|B(x,\frac{1}{2}\varepsilon)|} \|T(1_{B(x,\varepsilon)}f)\|_{L^{1,\infty}} \\
&\le \frac{c_\delta \|T\|_{L^1\to L^{1,\infty}}}{|B(x,\frac{1}{2}\varepsilon)|} \|1_{B(x,\varepsilon)}f\|_{L^1} \\
&\le c_{\delta,d} \|T\|_{L^1\to L^{1,\infty}} Mf(x),
\end{aligned}$$

where the first bound follows from the general estimate

$$\begin{aligned}
\fint_F |g|^\delta &= \frac{1}{|F|} \int_0^\infty \delta\lambda^{\delta-1} |F \cap \{|g|>\lambda\}|\,\mathrm{d}\lambda \\
&\le \int_0^A \delta\lambda^{\delta-1}\,\mathrm{d}\lambda + \frac{1}{|F|}\int_A^\infty \delta\lambda^{\delta-2}\|g\|_{L^{1,\infty}}\,\mathrm{d}\lambda \\
&= A^\delta + \frac{\delta}{1-\delta} A^{\delta-1} \frac{\|g\|_{L^{1,\infty}}}{|F|} = \frac{1}{1-\delta}\Big(\frac{\|g\|_{L^{1,\infty}}}{|F|}\Big)^\delta
\end{aligned} \tag{1.7.6}$$

by choosing $A = \|g\|_{L^{1,\infty}}/|F|$. □

1.7.7 Corollary.

$$\|T_\sharp\|_{L^1\to L^{1,\infty}} \le c_d\big(\|T\|_{L^2\to L^2} + \|\omega\|_{\mathrm{Dini}}\big)$$

Proof. From Cotlar's inequality, it follows that

$$\begin{aligned}\|T_\sharp\|_{L^1\to L^{1,\infty}} \le c_{d,\delta}\big(\|\omega\|_{\mathrm{Dini}} + \|T\|_{L^2\to L^2}\big)\|M\|_{L^1\to L^{1,\infty}} \\ + c_{d,\delta}\|M_\delta\|_{L^{1,\infty}\to L^{1,\infty}}\|T\|_{L^1\to L^{1,\infty}}.\end{aligned}$$

Since

$$\|M\|_{L^1\to L^{1,\infty}} \le c_d, \qquad \|T\|_{L^1\to L^{1,\infty}} \le c_d\big(\|T\|_{L^2\to L^2} + \|\omega\|_{\mathrm{Dini}}\big),$$

it remains to check that

$$\|M_\delta\|_{L^{1,\infty}\to L^{1,\infty}} \le c_{d,\delta}, \qquad \forall \delta\in(0,1),$$

and fix some $\delta\in(0,1)$, say $\delta=\frac{1}{2}$. We prove the final bound in the dyadic case, as the general case follows by the method of adjacent dyadic systems.

Let Q_i be the maximal dyadic cubes such that $\fint_{Q_i}|f|^\delta > \lambda^\delta$; hence $\{M_\delta f > \lambda\} = \bigcup_i Q_i$. Let us first consider a union of finitely many of these cubes only, say $F := \bigcup_{i=1}^N Q_i$. Then $|F|<\infty$, and

$$|F| = \sum_{i=1}^N |Q_i| \le \frac{1}{\lambda^\delta}\sum_{i=1}^N \int_{Q_i}|f|^\delta = \frac{1}{\lambda^\delta}\int_F |f|^\delta \le \frac{1}{\lambda^\delta}\frac{|F|^{1-\delta}\|f\|_{L^{1,\infty}}^\delta}{(1-\delta)}$$

by (1.7.6) in the last step. Simplifying, this shows that $|F| \le c_\delta\|f\|_{L^{1,\infty}}/\lambda$, and since this is true for any finite union $F = \bigcup_{i=1}^N Q_i \subset \{M_\delta f > \lambda\}$, letting $N\to\infty$ it follows that $|\{M_\delta f>\lambda\}| \le c_\delta\|f\|_{L^{1,\infty}}/\lambda$. □

Now we have proved the estimates from classical Calderón–Zygmund theory needed for Lerner's dyadic domination of Calderón–Zygmund operators.

1.8 Variational estimates*

In this section we show the applicability of Lerner's domination to so-called variational Calderón–Zygmund operators. However, this section will not be self-contained, but we will instead borrow and use some results from the theory of variational operators as black boxes. In a way, the possibility of using such black box input is also an illustration of the power of Lerner's abstract domination theorem.

Given a family of linear operators $(S_\varepsilon)_{\varepsilon\in(0,\infty)}$, we define its r-variation operator

$$V_\varepsilon^r Sf(x) := \sup\Big(\sum_{j=1}^N |S_{\varepsilon_j}f(x) - S_{\varepsilon_{j+1}}f(x)|^r\Big)^{1/r},$$

where the supremum is over all increasing sequences $\varepsilon\le\varepsilon_0\le\ldots\le\varepsilon_N$ (with the additional requirement that $0<\varepsilon_0$ if $\varepsilon=0$), where the length N is finite but arbitrary. We also denote $V^rS := V_0^rS$.

Two main examples of S_ε are the averaging operators

$$A_\varepsilon f(x) := \fint_{B(x,\varepsilon)} f(y)\,\mathrm{d}y$$

and the truncated Calderón–Zygmund operators

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

1.8.1 Remark (Variational vs. maximal operators). Interpreting $r=\infty$ in the "usual way", $V^\infty S$ is essentially a maximal operator. Indeed

$$V^\infty Sf(x) = \sup_{\varepsilon<\delta} |S_\varepsilon f(x) - S_\delta f(x)| \le 2\sup_{\varepsilon>0} |S_\varepsilon f(x)|,$$

and if $\inf_{\delta>0}|Sf(x)| = 0$ (e.g., if $\lim_{\delta\to\infty}|S_\delta f(x)| = 0$), then $V^\infty Sf(x) \ge \sup_{\varepsilon>0}|S_\varepsilon f(x)|$.

Since $\|\ \|_{\ell^r} \ge \|\ \|_{\ell^s} \ge \|\ \|_{\ell^\infty}$ for $r<s<\infty$, the variational operators V^rS are bigger than the maximal operator for finite r, and they increase in size as r decreases.

1.8.2 Remark. A prime application of maximal operators is the study of pointwise convergence of $S_\varepsilon f(x)$ as $\varepsilon\to 0$: Suppose that this convergence holds (for a.e. x) for every $f\in F$, where $F\subset L^p$ is dense, and suppose that $V^\infty : L^p \to L^{p,\infty}$ is bounded. Then the convergence holds for all $f\in L^p$.

To see this, define the sublinear operator

$$\Lambda f(x) := \limsup_{\varepsilon,\delta\to 0} |S_\varepsilon f(x) - S_\delta f(x)| \le V^\infty Sf(x).$$

By assumption, $\Lambda f = 0$ (a.e.) if $f\in F$. If $f\in L^p$ and $g\in F$, we have $\Lambda f = \Lambda(f-g+g) \le \Lambda(f-g)+\Lambda g = \Lambda(f-g)$, thus

$$\begin{aligned}|\{\Lambda f>\varepsilon\}| &\le |\{\Lambda(f-g)>\varepsilon\}| \le |\{V^\infty S(f-g)>\varepsilon\}|\\ &\le \varepsilon^{-p}\|V^\infty S(f-g)\|^p_{L^{p,\infty}} \le \varepsilon^{-p}\|V^\infty S\|^p_{L^p\to L^{p,\infty}}\|f-g\|^p_{L^p}.\end{aligned}$$

Taking the limit $g\to f$ in L^p, we see that $|\{\Lambda f>\varepsilon\}| = 0$, hence $|\{\Lambda>0\}| = \bigcup_{n=1}^\infty |\{\Lambda f > n^{-1}\}| = 0$, and thus $\Lambda f = 0$ almost everywhere.

On the other hand, the knowledge that $V^rS : L^p\to L^{p,\infty}$ (or the stronger $V^rS : L^p\to L^p$) allows to conclude the convergence of $S_\varepsilon f(x)$ without any a priori knowledge of this convergence on a dense subspace, which is useful in some applications. Equally importantly, the variational bounds give quantitative information on the rate of convergence.

We state without proof two theorems about variational operators:

1.8.3 Theorem. *For $r>2$, the operator*

$$\tilde V^r Af(x) := \sup_{z\in\mathbb{R}^d} V^r_{|z-x|}Af(z)$$

is bounded from $L^1(\mathbb{R}^d)$ to $L^{1,\infty}(\mathbb{R}^d)$.

1.8.4 Theorem. *Let T be a Calderón–Zygmund operator of convolution-type, meaning that its kernel has the form $K(x,y) = k(x-y)$. Suppose also that*

$$\int_{\partial B(0,t)} k(u)\,\mathrm{d}\sigma(u) = 0$$

for all $t > 0$, where σ is the $(d-1)$-dimensional surface measure, and moreover that at least one of the following additional conditions holds:

1. *$k(x) = \dfrac{1}{|x|^d} k\Big(\dfrac{x}{|x|}\Big)$ for all $x \in \mathbb{R}^d \setminus \{0\}$ (k is homogeneous), or*
2. *$|\nabla k(x)| \leq c'_K |x|^{-d-1}$ (k is smooth).*

Then, for every $r > 2$, the operator $V^r T$ is bounded from $L^1(\mathbb{R}^d)$ to $L^{1,\infty}(\mathbb{R}^d)$.

1.8.5 Remark. Theorem 1.8.3 as stated is from de França Silva and Zorin-Kranich 2016 [FZ16], but a similar bound for $V^r A$ in place of $\tilde{V}^r A$ is much older. Case (1) of Theorem 1.8.4 is from Campbell, Jones, Reinhold and Wierdl [CJRW03]. In case (2), an $L^p \to L^p$ bound for $V^r T$ is from Mirek, Stein and Trojan [MST15], and the $L^1 \to L^{1,\infty}$ bound follows from this by the results of [CJRW03], as noted in [FZ16].

Our present goal is to study the dyadic domination of $V^r T$. In order to deduce this by Lerner's method, we need the $L^1 \to L^{1,\infty}$ bounds of this operator itself (which is given by Theorem 1.8.4), as well as of $M_{V^r T}$. A key to the latter is the following bound:

1.8.6 Lemma (de França Silva & Zorin-Kranich 2016 [FZ16])**.** *Let T be a Calderón–Zygmund operator. If $|x - x'| \leq \varepsilon/2$, then*

$$|V^r_\varepsilon T f(x) - V^r_\varepsilon T f(x')| \leq c_d(\|\omega\|_{Dini} + c_K) M f(x) + c_d c_K \tilde{V}^r A|f|(x).$$

Proof. By the triangle inequality, we have

$$\begin{aligned} &|V^r_\varepsilon T f(x) - V^r_\varepsilon T f(x')| \\ &\leq \sup \Big(\sum_j |(T_{\varepsilon_j} f(x) - T_{\varepsilon_{j+1}} f(x)) - (T_{\varepsilon_j} f(x') - T_{\varepsilon_{j+1}} f(x'))|^r \Big)^{1/r}, \end{aligned}$$

where the supremum is over all increasing sequences $\varepsilon \leq \varepsilon_0 \leq \varepsilon_1 \leq \ldots \leq \varepsilon_N$. Moreover,

$$\begin{aligned} &(T_{\varepsilon_j} f(x) - T_{\varepsilon_{j+1}} f(x)) - (T_{\varepsilon_j} f(x') - T_{\varepsilon_{j+1}} f(x')) \\ &= \int_{\varepsilon_j < |y-x| < \varepsilon_{j+1}} K(x,y) f(y)\,\mathrm{d}y - \int_{\varepsilon_j < |y-x'| < \varepsilon_{j+1}} K(x',y) f(y)\,\mathrm{d}y \\ &= \Big(\int_{\varepsilon_j < |y-x| < \varepsilon_{j+1}} - \int_{\varepsilon_j < |y-x'| < \varepsilon_{j+1}} \Big) K(x,y) f(y)\,\mathrm{d}y \\ &\quad + \int_{\varepsilon_j < |y-x'| < \varepsilon_{j+1}} (K(x,y) - K(x',y)) f(y)\,\mathrm{d}y = I_j + II_j, \end{aligned} \tag{1.8.7}$$

where

$$\Big(\sum_j |II_j|^r\Big)^{1/r} \leq \sum_j \int_{\varepsilon_j<|y-x'|<\varepsilon_{j+1}} |K(x,y)-K(x',y)||f(y)|\,\mathrm{d}y$$
$$\leq \int_{|y-x'|>\varepsilon} \omega\Big(\frac{|x-x'|}{|x'-y|}\Big)\frac{1}{|x'-y|^d}|f(y)|\,\mathrm{d}y$$
$$\leq \sum_{k=0}^{\infty}\int_{2^k\varepsilon<|y-x'|<2^{k+1}} \omega(2^{-k-1})\frac{1}{(2^k\varepsilon)^d}|f(y)|\,\mathrm{d}y$$
$$\leq c_d\sum_{k=0}^{\infty}\omega(2^{-k-1}) \fint_{B(x',2^{k+1}\varepsilon)} |f(y)|\,\mathrm{d}y \leq c_d\|\omega\|_{Dini} Mf(x),$$

observing that $|x-x'| \leq \varepsilon/2$ implies that the balls $B(x', 2^{k+1}\varepsilon)$ also contain x.

For the first term on the left of (1.8.7), we write it as

$$I_j = \int_{\mathbb{R}^d} \chi_j(y) K(x,y) f(y)\,\mathrm{d}y,$$

where (suppressing the dependence on the fixed x, x')

$$\chi_j(y) := 1_{\varepsilon_j<|y-x|<\varepsilon_{j+1}} - 1_{\varepsilon_j<|y-x'|<\varepsilon_{j+1}}$$

We consider two cases. If $\varepsilon_{j+1} \leq \varepsilon_j + \varepsilon$, then we simply estimate

$$|\chi_j(y)| \leq 1_{\varepsilon_j<|y-x|<\varepsilon_{j+1}} + 1_{\varepsilon_j<|y-x'|<\varepsilon_{j+1}}$$

If $\varepsilon_{j+1} > \varepsilon_j + \varepsilon$, then we write

$$\begin{aligned}\chi_j(y) &= 1_{|y-x|>\varepsilon_j} - 1_{|y-x'|>\varepsilon_j} - \Big(1_{|y-x|>\varepsilon_{j+1}} - 1_{|y-x'|>\varepsilon_{j+1}}\Big)\\ &= 1_{|y-x|>\varepsilon_j>|y-x'|} - 1_{|y-x'|>\varepsilon_j>|y-x|}\\ &\quad - \Big(1_{|y-x|>\varepsilon_{j+1}>|y-x'|} - 1_{|y-x'|>\varepsilon_{j+1}>|y-x|}\Big).\end{aligned}$$

Since $\Big||y-x| - |y-x'|\Big| \leq |x-x'| \leq \varepsilon/2$, we deduce that

$$\begin{aligned}|\chi_j(y)| \leq\ & 1_{\varepsilon_j<|y-x|<\varepsilon_j+\varepsilon/2} + 1_{\varepsilon_j<|y-x'|<\varepsilon_j+\varepsilon/2}\\ &+ \Big(1_{\varepsilon_{j+1}-\varepsilon/2<|y-x'|<\varepsilon_{j+1}} + 1_{\varepsilon_{j+1}-\varepsilon/2<|y-x|<\varepsilon_{j+1}}\Big).\end{aligned}$$

Let us define $\eta_{3j} := \varepsilon_j$. In the first case that $\varepsilon_{j+1} \leq \varepsilon_j + \varepsilon$, we define $\eta_{3j+1} := \eta_{3j+2} := \varepsilon_{j+1}$. In the second case that $\varepsilon_{j+1} > \varepsilon_j + \varepsilon$, we define $\eta_{3j+1} := \varepsilon_j + \varepsilon/2$ and $\eta_{3j+2} := \varepsilon_{j+1} - \varepsilon/2$. In both cases, we have checked that

$$|\chi_j(y)| \leq \sum_{\hat{x}=x,x'}\sum_{i=0,2} 1_{\eta_{3j+i}<|y-\hat{x}|<\eta_{3j+i+1}},$$

where η_j is an increasing sequence with $\eta_{3j+i+1} \le \eta_{3j+i} + \varepsilon$ for $i = 0, 2$. Hence, using

$$|K(x,y)| \le \frac{c_K}{|x-y|^d} \le \frac{2^d c_K}{|x'-y|^d} \tag{1.8.8}$$

we have

$$\begin{aligned}
|I_j| &= \Big| \int \chi_j(y) K(x,y) f(y) \, \mathrm{d}y \Big| \\
&\le \sum_{\hat{x},i} \int_{\eta_{3j+i} < |y-\hat{x}| < \eta_{3j+i+1}} \frac{2^d c_K}{|\hat{x}-y|^d} |f(y)| \, \mathrm{d}y \\
&\le c_d c_K \sum_{\hat{x},i} \frac{1}{\eta_{3j+i+1}^d} \int_{\eta_{3j+i} < |y-\hat{x}| < \eta_{3j+i+1}} |f(y)| \, \mathrm{d}y,
\end{aligned}$$

where

$$\begin{aligned}
&\frac{1}{\eta_{3j+i+1}^d} \int_{\eta_{3j+i} < |y-\hat{x}| < \eta_{3j+i+1}} |f(y)| \, \mathrm{d}y \\
&= \Big(\frac{1}{\eta_{3j+i+1}^d} \int_{|y-\hat{x}| < \eta_{3j+i+1}} - \frac{1}{\eta_{3j+i}^d} \int_{|y-\hat{x}| < \eta_{3j+i}} |f(y)| \, \mathrm{d}y \\
&\qquad + \Big(\frac{1}{\eta_{3j+i}^d} - \frac{1}{\eta_{3j+i+1}^d} \Big) \int_{|y-\hat{x}| < \eta_{3j+i}} \Big) |f(y)| \, \mathrm{d}y \\
&= c_d \Big(\fint_{B(\hat{x},\eta_{3j+i+1})} - \fint_{B(\hat{x},\eta_{3j+i})} \Big) |f(y)| \, \mathrm{d}y \\
&\qquad + c_d \Big(1 - \frac{\eta_{3j+i}^d}{\eta_{3j+i+1}^d} \Big) \fint_{B(\hat{x},\eta_{3j+i})} |f(y)| \, \mathrm{d}y = III_j + IV_j.
\end{aligned}$$

Here, by definition,

$$\Big(\sum_j |III_j|^r \Big)^{1/r} \le c_d V_\varepsilon^r A|f|(\hat{x}) \le c_d \tilde{V}^r A|f|(x);$$

the case $\hat{x} = x$ of the last inequality is clear, and the case $\hat{x} = x'$ also follows from the definition of $\tilde{V}^r A$, recalling that $|x'-x| < \varepsilon/2$.

On the other hand,

$$|IV_j| \le c_d \Big| 1 - \frac{\eta_{3j+i}^d}{\eta_{3j+i+1}^d} \Big| Mf(x),$$

and, using the mean value estimate $1 - \zeta^d = d\xi^{d-1}(1-\zeta) \le d(1-\zeta)$ with

$\zeta = \eta_{3j+i}/\eta_{3j+i+1}$, we may continue with

$$\begin{aligned}
\Big(\sum_j \Big|1 - \frac{\eta_{3j+i}^d}{\eta_{3j+i+1}^d}\Big|^r\Big)^{1/r} &\leq d\Big(\sum_j \Big|1 - \frac{\eta_{3j+i}}{\eta_{3j+i+1}}\Big|^r\Big)^{1/r} \\
&= d\Big(\sum_{k=1}^{\infty} \sum_{j:k\varepsilon\leq\eta_{3j+i}<(k+1)\varepsilon} \Big|\frac{\eta_{3j+i+1}-\eta_{3j+i}}{\eta_{3j+i+1}}\Big|^r\Big)^{1/r} \\
&\leq d\Big(\sum_{k=1}^{\infty} \Big|\sum_{j:k\varepsilon\leq\eta_{3j+i}<(k+1)\varepsilon} \frac{\eta_{3j+i+1}-\eta_{3j+i}}{k\varepsilon}\Big|^r\Big)^{1/r}.
\end{aligned}$$

In the inner sum, we have $\eta_{3j+i+1} \leq \eta_{3j+i} + \varepsilon < (k+2)\varepsilon$. On the other hand, since $3j+i+1 \leq 3(j+1)+i$, the intervals $[\eta_{3j+i}, \eta_{3j+i+1}) \subset [k\varepsilon, (k+2)\varepsilon)$ are pairwise disjoint, so their lengths sum up to at most 2ε. Thus we continue the estimate with

$$\leq d\Big(\sum_{k=1}^{\infty}\Big|\frac{2\varepsilon}{k\varepsilon}\Big|^r\Big)^{1/r} = 2d\Big(\sum_{k=1}^{\infty} k^{-r}\Big)^{1/r} \leq cd,$$

since $r > 2$. □

With the hard work done in Lemma 1.8.6, it is now reasonably straightforward to obtain the following:

1.8.9 Lemma. *Let T be a Calderón–Zygmund operator. Then*

$$M_{V^rT}f(x) \leq V^rTf(x) + c_d(\|\omega\|_{Dini} + c_K)Mf(x) + c_dc_K\tilde{V}^rA|f|(x).$$

Proof. Recall that

$$M_{V^rT}f(x) = \sup_{Q\ni x}\sup_{z\in Q} V^rT(1_{(3Q)^c}f)(z).$$

Fix some Q and z as here, and abbreviate $\tilde{f} := 1_{(3Q)^c}f$

With

$$v_i := v_{\varepsilon_i,\varepsilon_{i+1}} := |T_{\varepsilon_i}\tilde{f}(z) - T_{\varepsilon_{i+1}}\tilde{f}(z)| = \Big|\int_{\varepsilon_j<|y-z|<\varepsilon_{j+1}} K(z,y)\tilde{f}(y)\,\mathrm{d}y\Big|$$

we have

$$V^rT\tilde{f}(z) = \sup\Big(\sum_i v_i^r\Big)^{1/r},$$

where the supremum is over all increasing sequences $0 < \varepsilon_0 < \varepsilon_1 < \ldots < \varepsilon_N$.

We want to break the sum at the thresholds $\ell(Q)$ and $2\sqrt{d}\ell(Q)$. If these numbers are not already part of the sequences of ε_i's, we can add them by using the triangle inequality

$$\begin{aligned}
&v_i \leq v_{\varepsilon_i,\ell(Q)} + v_{\ell(Q),\varepsilon_{i+1}} \quad \text{if} \quad \varepsilon_i < \ell(Q) < \varepsilon_{i+1} \leq 2\sqrt{d}\ell(Q), \\
&v_i \leq v_{\varepsilon_i,2\sqrt{d}\ell(Q)} + v_{2\sqrt{d}\ell(Q),\varepsilon_{i+1}} \quad \text{if} \quad \ell(Q) \leq \varepsilon_i < 2\sqrt{d}\ell(Q) < \varepsilon_{i+1}, \\
&v_i \leq v_{\varepsilon_i,\ell(Q)} + v_{\ell(Q),2\sqrt{d}\ell(Q)} + v_{2\sqrt{d}\ell(Q),\varepsilon_{i+1}} \quad \text{if} \quad \varepsilon_i < \ell(Q) < 2\sqrt{d}\ell(Q) < \varepsilon_{i+1},
\end{aligned}$$

leading to

$$\sup\Big(\sum_i v_i^r\Big)^{1/r} \leq \Big(\sup_{\varepsilon_N \leq \ell(Q)} + \sup_{\substack{\ell(Q)\leq\varepsilon_0 \\ \varepsilon_N \leq 2\sqrt{d}\ell(Q)}} + \sup_{2\sqrt{d}\ell(Q)\leq\varepsilon_0}\Big)\Big(\sum_i v_i^r\Big)^{1/r},$$

where e.g. the first supremum on the right is over all increasing sequences of $0 < \varepsilon_0 \leq \varepsilon_1 \leq \ldots \leq \varepsilon_N \leq \ell(Q)$, and the other two suprema have an analogous meaning.

In the first supremum, the integration domain in v_i is always contained in $B(z, \ell(Q))$, which does not meet $\operatorname{supp} \tilde{f} \subset (3Q)^c$, since $z \in Q$ and $\operatorname{dist}(Q, (3Q)^c) = \ell(Q)$. In the second supremum, we simply dominate $\| \ \|_{\ell^r} \leq \| \ \|_{\ell^1}$, leading to

$$\begin{aligned}\sum_{\substack{\ell(Q)\leq\varepsilon_i \\ \varepsilon_{i+1}\leq 2\sqrt{d}\ell(Q)}} &\int_{\varepsilon_i<|z-y|<\varepsilon_{i+1}} |K(z,y)||f(y)|\,\mathrm{d}y \\ &\leq \int_{\ell(Q)<|z-y|<2\sqrt{d}\ell(Q)} \frac{c_K}{|z-y|^d}|f(y)|\,\mathrm{d}y \\ &\leq c_d c_K \fint_{B(z,2\sqrt{d}\ell(Q))} |f(y)|\,\mathrm{d}y \leq c_d c_K Mf(x),\end{aligned}$$

since $x \in Q$ also belongs to the same ball.

Finally, the third supremum is by definition $V^r_{2\sqrt{d}\ell(Q)}\tilde{f}(z) = V^r_{2\sqrt{d}\ell(Q)} f(z)$, where we could replace $\tilde{f} = 1_{(3Q)^c} f$ by f, since the truncation parameter in the variation operator already ensures that we only integrate over $B(2\sqrt{d}\ell(Q))^c \subset (3Q)^c$ anyway.

Since $|x - z| < \sqrt{d}\ell(Q) = \frac{1}{2} \cdot 2\sqrt{d}\ell(Q)$, the previous Lemma 1.8.6 ensures that

$$V^r_{2\sqrt{d}\ell(Q)} f(z) \leq V^r_{2\sqrt{d}\ell(Q)} Tf(x) + c_d(\|\omega\|_{Dini} + c_K)Mf(x) + c_d c_K \tilde{V}^r A|f|(x),$$

and the first term on the right is obviously dominated by $V^r Tf(x)$. In combination with the bound for the second supremum above, this concludes the proof. □

1.8.10 Theorem (de França Silva & Zorin-Kranich 2016)**.** *Let $r > 2$ and T be a Calderón–Zygmund operator for which $V^r T : L^1 \to L^{1,\infty}$. Then for every compactly supported $f \in L^1(\mathbb{R}^d)$, there are sparse collection $\mathscr{S}_i \subset \mathscr{D}_i$ such that*

$$V^r T \leq c_{d,r,T} \sum_{i=1}^{3^d} T_{\mathscr{S}_i}|f|$$

and consequently $\|V^r T\|_{L^2(w)\to L^2(w)} \leq c_{d,r,T}[w]_{A_2} \quad \forall w \in A_2$.

Proof. We of course want to apply Lerner's Theorem 1.4.2, and the condition that $V^r T : L^1 \to L^{1,\infty}$ is part of the assumption. The other condition of

Theorem 1.4.2, that $M_{V^rT} : L^1 \to L^{1,\infty}$ follows from Lemma 1.8.9 and the $L^1 \to L^{1,\infty}$ bounds of the operators V^rT, M and $\tilde{V}^rA$, where the last result is contained in Theorem 1.8.3. □

1.8.11 Remark. Theorem 1.8.4 provides concrete conditions under which the property that $V^rT : L^1 \to L^{1,\infty}$ is valid. However, in the above theorem we don't need to make any reference to these conditions, only to the boundedness $V^rT : L^1 \to L^{1,\infty}$. If, in the future, this boundedness is verified for other classes of Calderón–Zygmund operators, Theorem 1.8.10 automatically applies to these operators as well.

1.8.12 Remark. The key Lemma 1.8.6 is taken from the original paper of de França Silva and Zorin-Kranich [FZ16] almost verbatim, but after this we have deviated from their original presentation in completing the proof of Theorem 1.8.10. The original paper used Lemma 1.8.6 as an input for an earlier approach to dyadic domination by Lacey from January 2015 [Lac17], whereas we have used the simplified approach of Lerner from December 2015 [Ler16].

1.8.13 Exercise. Suppose that $V^rSf(x) < \infty$ at some point x. Show that $\lim_{\varepsilon\to 0} S_\varepsilon f(x)$ exists at this point. (Hint: Given $\delta > 0$, check first that there are at most $N_\delta < \infty$ disjoint intervals $[a_i, b_i)$ with $|S_{a_i}f(x) - S_{b_i}f(x)| \geq \delta$.)

1.8.14 Exercise. Check that if $f \in \bigcup_{p\in[1,\infty)} L^p(\mathbb{R}^d)$, then both $A_\varepsilon f(x)$ and $T_\varepsilon f(x)$ tend to zero as $\varepsilon \to \infty$.

1.8.15 Exercise. For $0 < a < b$, prove that

$$V_a^rTf(x) \leq V_b^rTf(x) + c_d c_K(1 + \log(b/a))Mf(x).$$

1.8.16 Exercise. Define $\tilde{V}^rT$ in a way analogous to $\tilde{V}^rA$. Prove a pointwise bound for $\tilde{V}^rTf$, which allows to conclude that $\tilde{V}^rT : L^1 \to L^{1,\infty}$.

1.8.17 Exercise. Prove a pointwise bound for $M_{V^rA}f$, which allows to conclude that $M_{V^rA} : L^1 \to L^{1,\infty}$ (and hence to apply Lerner's theorem to V^rA.)

1.8.18 Exercise (Motivation for "$r > 2$" in the variational estimates)**.** Consider the standard dyadic intervals $\mathscr{D}$ of $\mathbb{R}$, and define the dyadic analogue of the averaging operators A_ε by $E_jf(x) := \langle f\rangle_{Q_j(x)}$, where $Q_j(x)$ is the unique dyadic cube of side-length 2^{-j} that contains x. The corresponding variation operator is $V^rEf := \sup(\sum_i |E_{j_i}f - E_{j_{i+1}}f|^r)^{1/r}$, where the supremum is over all increasing integer sequences j_i.

Define the L^∞-normalised Haar functions $h_I^\infty := 1_{I_\ell} - 1_{I_r}$, where $I_{\ell/r}$ is the left/right half of I, and the Rademacher functions $r_j := \sum_{I\in\mathscr{D}_j[0,1)} h_I^\infty$, where $\mathscr{D}_j[0,1) = \{I \in \mathscr{D} : I \subseteq [0,1), \ell(I) = 2^{-j}\}$. Check that the functions $(r_i)_{i=0}^\infty$ are orthonormal: $\int r_ir_j = \delta_{ij}$ ($:= 1$ if $i = j$, and $:= 0$ else). Check that $E_jr_i = r_i$ if $j > i$ and $E_jr_i = 0$ if $j \leq i$. Then consider a function of the form $f = \sum_{i=0}^\infty a_ir_i$. Check that, pointwise on $[0,1)$, we have $V^rEf \geq (\sum_{i=0}^\infty |a_i|^r)^{1/r}$ (hint: a very easy choice of j_i works), while $\|f\|_{L^1} \leq \|f\|_{L^2} = (\sum_{i=0}^\infty |a_i|^2)^{1/2}$. Conclude with a suitable choice of $(a_i)_{i=0}^\infty$ that $V^rE : L^1 \not\to L^{1,\infty}$ if $r < 2$.

1.9 A_∞ weights and the reverse Hölder inequality

1.9.1 Definition. We say that $w \in A_\infty^{\mathscr{D}}$ with constant $[w]_{A_\infty}^{\mathscr{D}}$ if

$$\int_Q M_Q w \leq [w]_{A_\infty}^{\mathscr{D}} \int_Q w$$

for all $Q \in \mathscr{D}$, where $M_Q w := \sup_{Q' \in \mathscr{D}, Q' \subseteq Q} 1_{Q'} \langle w \rangle_{Q'}$.

1.9.2 Proposition. *We have $A_2^{\mathscr{D}} \subset A_\infty^{\mathscr{D}}$ and $[w]_{A_\infty}^{\mathscr{D}} \leq e[w]_{A_2}^{\mathscr{D}}$.*

Proof. Since $x \mapsto \exp x$ is convex, we deduce from Jensen's inequality that

$$\langle w^{-1} \rangle_Q = \langle \exp \log w^{-1} \rangle_Q \geq \exp \langle \log w^{-1} \rangle_Q = \exp \langle -\log w \rangle_Q = \frac{1}{\exp \langle \log w \rangle_Q}$$

and hence

$$\langle w \rangle_Q \leq [w]_{A_2}^{\mathscr{D}} \langle w^{-1} \rangle_Q^{-1} \leq [w]_{A_2}^{\mathscr{D}} \exp \langle \log w \rangle_Q.$$

On the other hand, for $p \in (1, \infty)$,

$$\exp \langle \log w \rangle_Q = (\exp \langle \log w^{1/p} \rangle_Q)^p \leq (\langle w^{1/p} \rangle_Q)^p,$$

and hence

$$M_Q w \leq [w]_{A_2}^{\mathscr{D}} (M_Q (w^{1/p}))^p \leq [w]_{A_2}^{\mathscr{D}} (M_{\mathscr{D}} (1_Q w^{1/p}))^p.$$

Thus

$$\int_Q M_Q w \leq [w]_{A_2}^{\mathscr{D}} \int (M_{\mathscr{D}} (1_Q w^{1/p}))^p \leq [w]_{A_2}^{\mathscr{D}} (p')^p w(Q).$$

Being valid for every $p \in (1, \infty)$, this implies that

$$[w]_{A_\infty}^{\mathscr{D}} \leq [w]_{A_2}^{\mathscr{D}} \lim_{p \to \infty} (p')^p,$$

where

$$(p')^p = \Big(\frac{p}{p-1} \Big)^p = \frac{1}{(1 - \frac{1}{p})^p} \to \frac{1}{e^{-1}} = e$$

as $p \to \infty$. □

1.9.3 Lemma. *Let Q_0 be a cube, and w and u be weights related by the following condition: Whenever $Q = Q_0$, or $Q \subset Q_0$ is a maximal cube with $\langle w \rangle_Q > \lambda$ for some $\lambda > \langle w \rangle_{Q_0}$, then $u(Q) \leq K w(Q)$ with some fixed constant K. Then*

$$\fint_{Q_0} (M_{Q_0} w)_N^{\varepsilon} u \leq K \langle w \rangle_{Q_0}^{1+\varepsilon} + \frac{\varepsilon}{1+\varepsilon} K 2^d \fint_{Q_0} (M_{Q_0} w)_N^{1+\varepsilon},$$

where $x_N := \min(x, N)$.

Proof. If $N \le \langle w\rangle_{Q_0}$, then the estimate holds even without the second term on the right:

$$\int_{Q_0} (M_{Q_0}w)_N^{\varepsilon} u \le N^{\varepsilon} u(Q_0) \le \langle w\rangle_{Q_0}^{\varepsilon} K w(Q_0) = K\langle w\rangle_{Q_0}^{1+\varepsilon}|Q_0|.$$

We then assume that $N > \langle w\rangle_{Q_0}$ and write

$$\begin{aligned}\int_{Q_0} (M_{Q_0}w)_N^{\varepsilon} u &= \int_0^{\infty} \varepsilon\lambda^{\varepsilon-1} u(\{(M_{Q_0}w)_N > \lambda\})\,\mathrm{d}\lambda \\ &= \int_0^{N} \varepsilon\lambda^{\varepsilon-1} u(\{M_{Q_0}w > \lambda\})\,\mathrm{d}\lambda \\ &= \int_0^{\langle w\rangle_{Q_0}} \varepsilon\lambda^{\varepsilon-1} u(Q_0)\,\mathrm{d}\lambda + \int_{\langle w\rangle_{Q_0}}^{N} \varepsilon\lambda^{\varepsilon-1} \sum_{Q\in\mathscr{Q}_\lambda} u(Q)\,\mathrm{d}\lambda,\end{aligned}$$

where $\mathscr{Q}_\lambda$ is the collection of maximal cubes $Q \subset Q_0$ with $\langle w\rangle_Q > \lambda$. By assumption, we have $u(Q_0) \le Kw(Q_0)$; hence

$$\begin{aligned}\int_0^{\langle w\rangle_{Q_0}} \varepsilon\lambda^{\varepsilon-1} u(Q_0)\,\mathrm{d}\lambda &\le K\int_0^{\langle w\rangle_{Q_0}} \varepsilon\lambda^{\varepsilon-1} w(Q_0)\,\mathrm{d}\lambda \\ &= K\langle w\rangle_{Q_0}^{\varepsilon} w(Q_0) = K\langle w\rangle_{Q_0}^{1+\varepsilon}|Q_0|.\end{aligned}$$

For $Q \in \mathscr{Q}_\lambda$, we have

$$u(Q) \le Kw(Q) \le Kw(\hat{Q}) = K\langle w\rangle_{\hat{Q}}|\hat{Q}| \le K\lambda\cdot 2^d|Q|,$$

and hence

$$\begin{aligned}\int_{\langle w\rangle_{Q_0}}^{N} \varepsilon\lambda^{\varepsilon-1} \sum_{Q\in\mathscr{Q}_\lambda} u(Q)\,\mathrm{d}\lambda &\le 2^d K\int_{\langle w\rangle_{Q_0}}^{N} \varepsilon\lambda^{\varepsilon} \sum_{Q\in\mathscr{Q}_\lambda} |Q|\,\mathrm{d}\lambda \\ &\le 2^d K\int_0^{N} \varepsilon\lambda^{\varepsilon}|\{M_{Q_0}w > \lambda\}|\,\mathrm{d}\lambda \\ &= 2^d K\int_0^{\infty} \varepsilon\lambda^{\varepsilon}|\{(M_{Q_0}w)_N > \lambda\}|\,\mathrm{d}\lambda \\ &= 2^d K\frac{\varepsilon}{1+\varepsilon}\int_{Q_0} (M_{Q_0}w)_N^{1+\varepsilon}.\end{aligned}$$

□

1.9.4 Theorem (Reverse Hölder inequality)**.** *If $Q_0 \in \mathscr{D}$ and $w \in A_\infty^{\mathscr{D}}$, then*

$$\fint_{Q_0} w^{1+\varepsilon} \le 2\Big(\fint_{Q_0} w\Big)^{1+\varepsilon}, \qquad \varepsilon = \frac{1}{2^{d+1}[w]_{A_\infty}^{\mathscr{D}} - 1}.$$

Proof. We apply Lemma 1.9.3 with two choices of u and K: the trivial case $u = w$, $K = 1$; as well as $u = (M_{Q_0}w)_N$ and $K = [w]_{A_\infty}^{\mathscr{D}}$. For the latter case, it suffices to check the case $N=\infty$, i.e., $u=M_{Q_0}w$, since clearly $u_N(Q)\leq u(Q)$ for every cube Q. The case

$$u(Q_0)=\int_{Q_0} M_{Q_0}w \leq [w]^{\mathscr{D}}_{A_\infty} w(Q_0)$$

is immediate from the definition of $A^{\mathscr{D}}_\infty$. Concerning maximal $Q\subset Q_0$ with $\langle w\rangle_Q>\lambda$, we note that

$$\sup_{\substack{Q'\supsetneq Q\\ Q'\subseteq Q_0}} \langle w\rangle_{Q'} \leq \lambda < \langle w\rangle_Q,$$

and hence, for all $x\in Q$,

$$M_{Q_0}w(x)=\sup_{\substack{Q'\ni x\\ Q'\subseteq Q_0}} \langle w\rangle_{Q'} = \sup_{\substack{Q'\ni x\\ Q'\subseteq Q}} \langle w\rangle_{Q'} = M_Q w(x), \tag{1.9.5}$$

and thus

$$u(Q)=\int_Q M_{Q_0}w=\int_Q M_Q w\leq [w]^{\mathscr{D}}_{A_\infty} w(Q)$$

also in this case.

We first apply Lemma 1.9.3 to $(u,K)=((M_{Q_0}w)_N,[w]^{\mathscr{D}}_{A_\infty})$ to deduce that

$$\fint_{Q_0}(M_{Q_0}w)_N^{1+\varepsilon}\leq K\langle w\rangle_{Q_0}^{1+\varepsilon}+\frac{\varepsilon}{1+\varepsilon}[w]^{\mathscr{D}}_{A_\infty}2^d\fint_{Q_0}(M_{Q_0}w)_N^{1+\varepsilon},$$

and hence

$$\fint_{Q_0}(M_{Q_0}w)_N^{1+\varepsilon}\leq \frac{[w]^{\mathscr{D}}_{A_\infty}}{1-\frac{\varepsilon}{1+\varepsilon}2^d[w]^{\mathscr{D}}_{A_\infty}}\langle w\rangle_{Q_0}^{1+\varepsilon} \tag{1.9.6}$$

provided that $\frac{\varepsilon}{1+\varepsilon}K2^d<1$. Note that subtracting a multiple of the integral $\fint_{Q_0}(M_{Q_0}w)_N^{1+\varepsilon}$ from both sides is legitimate, since $\fint_{Q_0}(M_{Q_0}w)_N^{1+\varepsilon}\leq\fint_{Q_0}N^{1+\varepsilon}=N^{1+\varepsilon}<\infty$.

We then apply Lemma 1.9.3 to $(u,K)=(w,1)$ to deduce that

$$\begin{aligned}\fint_{Q_0}(M_{Q_0}w)_N^{\varepsilon}w &\leq \langle w\rangle_{Q_0}^{1+\varepsilon}+\frac{\varepsilon}{1+\varepsilon}2^d\fint_{Q_0}(M_{Q_0}w)_N^{1+\varepsilon}\\ &\leq \langle w\rangle_{Q_0}^{1+\varepsilon}+\frac{\frac{\varepsilon}{1+\varepsilon}2^d[w]^{\mathscr{D}}_{A_\infty}}{1-\frac{\varepsilon}{1+\varepsilon}2^d[w]^{\mathscr{D}}_{A_\infty}}\langle w\rangle_{Q_0}^{1+\varepsilon}=\frac{1}{1-\frac{\varepsilon}{1+\varepsilon}2^d[w]^{\mathscr{D}}_{A_\infty}}\langle w\rangle_{Q_0}^{1+\varepsilon},\end{aligned}$$

substituting the bound from (1.9.6) in the second step.

Observing that $w\leq M_{Q_0}w=\lim_{N\to\infty}M_{Q_0}w$, we deduce from monotone convergence that

$$\fint_{Q_0}w^{1+\varepsilon}\leq\frac{1}{1-\frac{\varepsilon}{1+\varepsilon}2^d[w]^{\mathscr{D}}_{A_\infty}}\langle w\rangle_{Q_0}^{1+\varepsilon},$$

and choosing ε so that $\frac{\varepsilon}{1+\varepsilon}2^d[w]^{\mathscr{D}}_{A_\infty}=\frac12$, namely $\varepsilon=(2^{d+1}[w]^{\mathscr{D}}_{A_\infty}-1)^{-1}$, we obtain the Theorem. □

1.9.7 Remark. The reverse Hölder inequality of A_∞ weights is a classical result of Coifman and C. Fefferman [CF74]; its sharp form stated in Theorem 1.9.4 is from [HPR12], which also contains an extension to spaces of homogeneous type.

1.9.8 Exercise. Show the following converse of Theorem 1.9.4: If a weight w satisfies the reverse Hölder inequality

$$\Big(\fint_Q w^{1+\varepsilon}\Big)^{1/(1+\varepsilon)} \leq K \fint_Q w$$

for all $Q \in \mathscr{D}$, then $w \in A_\infty^{\mathscr{D}}$. Estimate $[w]_{A_\infty}^{\mathscr{D}}$ in terms of K and ε. (Hint: Use the boundedness of the maximal operator in $L^{1+\varepsilon}$.)

1.9.9 Exercise. Consider the following truncated version of M_Q:

$$M_Q^N f(x) := \sup_{\substack{Q' \in \mathscr{D}, Q' \subseteq Q \\ \ell(Q') \geq 2^{-N}\ell(Q)}} 1_{Q'}(x)\langle |f| \rangle_{Q'},$$

and define the truncated dyadic A_∞ constant as the smallest constant in the following inequality:

$$\int_Q M_Q^N w \leq [w]_{A_\infty}^{\mathscr{D},N} \int_Q w \qquad \forall Q \in \mathscr{D}.$$

Show that $[w]_{A_\infty}^{\mathscr{D},N} < \infty$ for any weight w, and that $[w]_{A_\infty}^{\mathscr{D},N} \to [w]_{A_\infty}^{\mathscr{D}}$ as $N \to \infty$.

1.9.10 Exercise. The following condition is often used as the definition of the (dyadic) A_∞: There are constants $\delta, \eta \in (0,1)$ such that for all (dyadic) cubes Q and all measurable subsets $E \subset Q$, if $|E| \leq \delta |Q|$, then $w(E) \leq \eta w(Q)$. Prove that this condition implies the dyadic A_∞ condition as we have defined it.

Hint: Prove that $[w]_{A_\infty}^{\mathscr{D},N} \leq \frac{1}{\delta} + \eta [w]_{A_\infty}^{\mathscr{D},N}$ by splitting the integral $\int_{Q_0} M_{Q_0}^N w$ over $Q_0 \setminus E$ and E, where $E = \{x \in Q_0 : M_{Q_0}^N w > \lambda\}$, where λ is chosen appropriately to ensure that $|E| \leq \delta |Q_0|$. Check that (1.9.5) also works for $M_{Q_0}^N$.

Chapter 2

Introduction to the theory of matrix weights

2.1 The matrix-weighted L^2 space

We are now going to study matrix-valued weights $W : \mathbb{R}^d \to \mathscr{L}(\mathbb{C}^n) \approx \mathbb{C}^{n\times n}$.

A scalar-valued weight is usually assumed to be positive almost everywhere. We assume that a matrix-valued weight is also positive almost everywhere, *in the sense of self-adjoint operators*. Namely, we demand that W is self-adjoint, and

$$(Wx|x) > 0 \qquad \forall\, x \in \mathbb{C}^n.$$

Recall from Linear Algebra that a self-adjoint matrix can always be diagonalised; namely, there is an orthonormal basis $(e_i)_{i=1}^n$ of $\mathbb{C}^n$ (consisting of eigenvectors of W) and real eigenvalues $(\lambda_i)_{i=1}^n$ such that

$$W = \sum_{i=1}^n \lambda_i e_i \otimes e_i, \qquad e_i \otimes e_i := e_i(e_i|\ )$$

Testing the previous inequality with $x = e_i$, we find that $\lambda_i > 0$ if W is positive.

The eigenvalue expansion leads to the *functional calculus* of W, i.e., a way of defining functions of W via

$$\phi(W) := \sum_{i=1}^n \phi(\lambda_i) e_i \otimes e_i.$$

Important particular cases are the square root $\phi(t) = \sqrt{t}$ and the inverse $\phi(t) = t^{-1}$; it is easy to see that W^{-1} so defined in the special case of a positive matrix coincides with the usual matrix inverse.

Note that all functions $\phi(W)$ are also self-adjoint matrices. In particular, we have

$$(Wx|y) = (W^{1/2}W^{1/2}x|y) = (W^{1/2}x|W^{1/2}y), \quad (Wx|x) = \|W^{1/2}x\|^2.$$

For scalar-valued functions f and w, the quantity $\|f\|_{L^2(w)}$ can be expressed in different equivalent ways:

$$\|f\|_{L^2(w)}^2 = \int |f|^2 w = \int |w^{1/2} f|^2 = \int w f \bar{f}.$$

For a vector-valued function $f : \mathbb{R}^d \to \mathbb{R}^n$ and a matrix-valued weight $W : \mathbb{R}^d \to \mathscr{L}(\mathbb{R}^n) \approx \mathbb{R}^{n\times n}$, we take the last two expressions as the basis of the definition:

$$\|f\|_{L^2(W)}^2 := \int \|W^{1/2} f\|^2 = \int (Wf|f).$$

2.2 The matrix A_2 condition

The following proposition motivates the definition

$$[W]_{A_2} := \sup_Q \|\langle W\rangle_Q^{1/2} \langle W^{-1}\rangle_Q^{1/2}\|_{op}^2,$$

and $W \in A_2$ if and only if $[W]_{A_2} < \infty$.

2.2.1 Proposition. *The norm of the operator $f \mapsto 1_Q\langle f\rangle_Q$ on $L^2(W)$ is equal to $\|\langle W\rangle_Q^{1/2}\langle W^{-1}\rangle_Q^{1/2}\|_{op}$.*

Proof. I. Upper bound for the norm:

$$\begin{aligned}
\|1_Q\langle f\rangle_Q\|_{L^2(W)} &= \Big(\int_Q (W\langle f\rangle_Q|\langle f\rangle_Q)\Big)^{1/2} = \Big(|Q|(\langle W\rangle_Q\langle f\rangle_Q|\langle f\rangle_Q)\Big)^{1/2} \\
&= \Big(|Q|\|\langle W\rangle_Q^{1/2}\langle f\rangle_Q\|^2\Big)^{1/2} = |Q|^{1/2} \sup_{\|x\|\leq 1} |(\langle W\rangle_Q^{1/2}\langle f\rangle_Q|x)|
\end{aligned}$$

where

$$\begin{aligned}
|(\langle W\rangle_Q^{1/2}\langle f\rangle_Q|x)| &= \Big|\Big(\langle f\rangle_Q\Big|\langle W\rangle_Q^{1/2}x\Big)\Big| = \Big|\Big\langle (f|\langle W\rangle_Q^{1/2}x)\Big\rangle_Q\Big| \\
&= \Big|\Big\langle \Big(W^{1/2}f\Big|W^{-1/2}\langle W\rangle_Q^{1/2}x\Big)\Big\rangle_Q\Big| \\
&\leq \Big\langle \Big|(W^{1/2}f|W^{-1/2}\langle W\rangle_Q^{1/2}x)\Big|\Big\rangle_Q \\
&\leq \Big\langle \|W^{1/2}f\|\|W^{-1/2}\langle W\rangle_Q^{1/2}x\|\Big\rangle_Q \\
&\leq \Big\langle \|W^{1/2}f\|^2\Big\rangle_Q^{1/2}\Big\langle \|W^{-1/2}\langle W\rangle_Q^{1/2}x\|^2\Big\rangle_Q^{1/2},
\end{aligned}$$

where $\left\langle \|W^{1/2}f\|^2 \right\rangle_Q^{1/2} = |Q|^{-1/2}\|1_Q f\|_{L^2(W)}$ and

$$\begin{aligned}
&\left\langle \|W^{-1/2}\langle W\rangle_Q^{1/2}x\|^2 \right\rangle_Q \\
&= \left\langle \left(W^{-1/2}\langle W\rangle_Q^{1/2}x \middle| W^{-1/2}\langle W\rangle_Q^{1/2}x \right) \right\rangle_Q \\
&= \left\langle \left(W^{-1}\langle W\rangle_Q^{1/2}x \middle| \langle W\rangle_Q^{1/2}x \right) \right\rangle_Q \\
&= \left(\langle W^{-1}\rangle_Q \langle W\rangle_Q^{1/2}x \middle| \langle W\rangle_Q^{1/2}x \right) \\
&= \left(\langle W^{-1}\rangle_Q^{1/2} \langle W\rangle_Q^{1/2}x \middle| \langle W^{-1}\rangle_Q^{1/2}\langle W\rangle_Q^{1/2}x \right) \\
&= \|\langle W^{-1}\rangle_Q^{1/2}\langle W\rangle_Q^{1/2}x\|^2 \le \|\langle W^{-1}\rangle_Q^{1/2}\langle W\rangle_Q^{1/2}\|_{op}^2
\end{aligned}$$

Combining the estimates, we have shown that

$$\|1_Q\langle f\rangle_Q\|_{L^2(W)} \le \|\langle W^{-1}\rangle_Q^{1/2}\langle W\rangle_Q^{1/2}\|_{op}\|f\|_{L^2(W)}.$$

II. Lower bound for the norm: Let us then assume that

$$\|1_Q\langle f\rangle_Q\|_{L^2(W)} \le K\|f\|_{L^2(W)}.$$

We test this estimate with the function

$$f := \Sigma_\varepsilon x := (W+\varepsilon)^{-1}x.$$

Thus

$$\|\langle W\rangle_Q^{1/2}\langle \Sigma_\varepsilon\rangle_Q x\|^2 = (\langle W\rangle_Q\langle\Sigma_\varepsilon\rangle_Q x|\langle\Sigma_\varepsilon\rangle_Q x) \le K^2 \fint_Q (W\Sigma_\varepsilon x|\Sigma_\varepsilon x),$$

where

$$\begin{aligned}
(W\Sigma_\varepsilon x|\Sigma_\varepsilon x) = ((W+\varepsilon-\varepsilon)\Sigma_\varepsilon x|\Sigma_\varepsilon x) &= (x|\Sigma_\varepsilon x) - \varepsilon(\Sigma_\varepsilon x|\Sigma_\varepsilon x) \\
&\le (x|\Sigma_\varepsilon x) = (\Sigma_\varepsilon x|x),
\end{aligned}$$

and hence

$$\|\langle W\rangle_Q^{1/2}\langle\Sigma_\varepsilon\rangle_Q x\|^2 \le K^2\left(\langle\Sigma_\varepsilon\rangle_Q x \middle| x\right) = K^2\|\langle\Sigma_\varepsilon\rangle_Q^{1/2}x\|^2.$$

The matrix Σ_ε is positive and integrable, hence its average $\langle\Sigma_\varepsilon\rangle_Q$ is positive and hence invertible. Choosing $x = \langle\Sigma_\varepsilon\rangle_Q^{-1}y$ we deduce that

$$\|\langle W\rangle_Q^{1/2}\langle\Sigma_\varepsilon\rangle_Q^{1/2}y\| \le K\|y\|,$$

hence

$$\|\langle\Sigma_\varepsilon\rangle_Q^{1/2}\langle W\rangle_Q^{1/2}\|_{op} = \|(\langle\Sigma_\varepsilon\rangle_Q^{1/2}\langle W\rangle_Q^{1/2})^*\|_{op} = \|\langle W\rangle_Q^{1/2}\langle\Sigma_\varepsilon\rangle_Q^{1/2}\|_{op} \le K,$$

which can be further written as

$$\|\langle\Sigma_\varepsilon\rangle_Q^{1/2}\langle W\rangle_Q^{1/2}x\| \le K\|x\|$$

or (since $\langle W\rangle_Q$ is positive and hence invertible)

$$\fint_Q(\Sigma_\varepsilon y|y) = (\langle\Sigma_\varepsilon\rangle_Q y|y) = \|\langle\Sigma_\varepsilon\rangle_Q^{1/2}y\|^2 \le K^2\|\langle W\rangle_Q^{-1/2}y\|^2.$$

Here

$$(\Sigma_\varepsilon y|y) = \sum_i(\lambda_i+\varepsilon)^{-1}|(e_i|y)|^2 \ \uparrow\ \sum_i \lambda_i^{-1}|(e_i|y)|^2 = (W^{-1}y|y),$$

so monotone convergence implies the integrability of the latter, and we finally deduce that

$$\|\langle W^{-1}\rangle_Q^{1/2}y\| \le K\|\langle W\rangle_Q^{-1/2}y\|,$$

or

$$\|\langle W^{-1}\rangle_Q^{1/2}\langle W\rangle_Q^{1/2}x\| \le K\|x\|,$$

so that

$$\|\langle W\rangle_Q^{1/2}\langle W^{-1}\rangle_Q^{1/2}\|_{op} = \|\langle W^{-1}\rangle_Q^{1/2}\langle W\rangle_Q^{1/2}\|_{op} \le K.$$

□

2.2.2 Exercise. For self-adjoint matrices A, B, we introduce the partial order $\le$ as follows:

$$A \le B \qquad \overset{def}{\Leftrightarrow} \qquad (Ax|x) \le (Bx|x) \quad \forall x \in \mathbb{C}^n.$$

For positive matrices A, B, show that

$$A \le B \ \Leftrightarrow\ \|A^{1/2}B^{-1/2}\|_{op} \le 1 \ \Leftrightarrow\ \|B^{-1/2}A^{1/2}\|_{op} \le 1 \ \Leftrightarrow\ B^{-1} \le A^{-1},$$

i.e., all four listed conditions are equivalent.

2.2.3 Exercise. Show that $W \in A_2$ if and only if $\langle W\rangle_Q \le C\langle W^{-1}\rangle_Q^{-1}$, if and only if $W^{-1} \in A_2$, and the optimal constant satisfies $C = [W]_{A_2} = [W^{-1}]_{A_2}$.

2.2.4 Exercise. Show that any matrix weight W satisfies the estimate

$$\langle W^{-1}\rangle_Q^{-1} \le \langle W\rangle_Q.$$

Hint: $(\langle W^{-1}\rangle_Q^{-1}x|x) = \langle(W^{-1/2}\langle W^{-1}\rangle_Q^{-1}x|W^{1/2}x)\rangle_Q$.

2.2.5 Remark. The set-up and the results about matrix weights are from Treil and Volberg [TV97].

2.3 Convex body domination

Given a linear operator T acting on scalar-valued functions, we may define its extension on vector-valued functions component-wise by

$$Tf := \sum_{i=1}^{n} e_i T(e_i|f), \tag{2.3.1}$$

where $(e_i)_{i=1}^n$ is an orthonormal basis on $\mathbb{R}^n$. For a linear operator, this definition is independent of the chosen orthonormal basis (Exercise 2.3.19). One can still make the same definition (2.3.1) for a non-linear operator, and this is often done, but it is not as "canonical", since it depends on the chosen basis. Our analysis below will exploit changes of basis, and therefore we stick to linear operators only.

We now state a version of the dyadic domination theorem for vector-valued functions that is suitable for studying their estimation in the matrix-weighted norms.

2.3.2 Theorem (Nazarov, Petermichl, Treil, Volberg 2017 [NPTV17])**.** *Let $T : L^1(\mathbb{R}^d) \to L^{1,\infty}(\mathbb{R}^d)$ be a linear operator such that also $M_T : L^1(\mathbb{R}^d) \to L^{1,\infty}(\mathbb{R}^d)$. We define the action of T on $L^1(\mathbb{R}^d;\mathbb{R}^n)$ component-wise. For compactly supported $f \in L^1(\mathbb{R}^d;\mathbb{R}^n)$, there is a $(1-\varepsilon)$-sparse collection $\mathscr{S}$ of dyadic cubes such that*

$$Tf(x) \in \frac{c_{d,n}c_T}{\varepsilon} \sum_{S\in\mathscr{S}} 1_S(x)\langle\!\langle f\rangle\!\rangle_{3S}, \tag{2.3.3}$$

where

$$\begin{aligned}\langle\!\langle f\rangle\!\rangle_Q &:= \Big\{ \fint_Q \phi f : \phi \in B_{L^\infty(Q)} \Big\} \subset \mathbb{R}^n,\\ B_{L^\infty(Q)} &:= \{\phi \in L^\infty(Q) : \|\phi\|_\infty \leq 1\}.\end{aligned}$$

More precisely, there exist functions $k_S \in B_{L^\infty(S\times 3S)}$ such that

$$Tf(x) = \frac{c_{d,n}c_T}{\varepsilon} \sum_{S\in\mathscr{S}} 1_S(x) \fint_{3S} k_S(x,y) f(y)\,\mathrm{d}y. \tag{2.3.4}$$

Note that $\langle\!\langle f\rangle\!\rangle_Q$ is a subset, not an element, of $\mathbb{R}^n$. Accordingly, the formula (2.3.3) has an inclusion "$\in$" instead of "$=$".

Theorem 2.3.2 will be used to prove the current record $\|T\|_{L^2(W)\to L^2(W)} \leq c_{d,n}c_T[W]_{A_2}^{3/2}$ (also from [NPTV17]) for Calderón–Zygmund operators on matrix weighted spaces.

Before starting the proof of the theorem, we provide various preparations.

2.3.5 Lemma. *For $f \in L^1(Q;\mathbb{R}^n)$, the set $\langle\!\langle f\rangle\!\rangle_Q \subset \mathbb{R}^n$ is symmetric, convex and compact.*

Proof. Recall that a set E is symmetric if $x \in E$ implies $-x \in E$ and convex if $x_i \in E$ for $i = 1, 2$ and $\lambda_i \in [0, 1]$ with $\lambda_1 + \lambda_2 = 1$ imply that $\lambda_1 x_1 + \lambda_2 x_2 \in E$. So let $x = \fint_Q \phi f$, $x_i = \fint_Q \phi_i f \in \langle\!\langle f \rangle\!\rangle_Q$, where $\phi, \phi_i \in B_{L^\infty(Q)}$. Then $-x_1 = \fint_Q (-\phi_1) f \in \langle\!\langle f \rangle\!\rangle_Q$, since also $-\phi_i \in B_{L^\infty(Q)}$, and $\lambda_1 x_1 + \lambda_2 x_2 = \fint_Q (\lambda_1 \phi_1 + \lambda_2 \phi_2) f \in \langle\!\langle f \rangle\!\rangle_Q$, since $\lambda_1 \phi_1 + \lambda_2 \phi_2 \in B_{L^\infty}$.

A subset of $\mathbb{R}^n$ is compact if and only if it is bounded and closed. For boundedness, it is immediate that $|x_1| \leq \fint_Q |\phi_1||f| \leq \fint |f|$, so that $\langle\!\langle f \rangle\!\rangle_Q \subset B(0, \langle |f| \rangle_Q)$.

The compactness follows easily from the functional analytic Proposition 2.3.8 below. Namely, let $x_k = \fint_Q \phi_k f \in \langle f \rangle_Q$ converge to some $x \in \mathbb{R}^n$. Now there is a subsequence $(\phi_{k_j})_{j=1}^\infty$ and a $\phi \in L^\infty(Q)$ such that $\int_Q \phi_{k_j} h \to \int_Q \phi h$ for every $h \in L^1$, in particular for each component function $h = f_i$ of $f = (f_i)_{i=1}^n$, and hence $x_{k_j} = \fint_Q \phi_{k_j} f \to \fint_Q \phi f$. On the other hand, we also have $x_{k_j} \to x$, and thus $x = \fint_Q \phi f \in \langle\!\langle f \rangle\!\rangle_Q$. $\square$

We record the following facts from Functional Analysis. Let $S \subset \mathbb{R}^d$ be a measurable subset. (The results would be valid much more abstractly, but we will not need this here.)

2.3.6 Proposition. $(L^1(S))^* = L^\infty(S)$, *i.e., every* $g \in L^\infty(S)$ *defines a bounded linear functional* $\Lambda : L^1(S) \to \mathbb{R}$ *via* $\Lambda f := \int_S gf$, *and every bounded bounded linear functional* $\Lambda : L^1(S) \to \mathbb{R}$ *has this form for some* $g \in L^\infty(S)$ *such that* $\|g\|_S = \|\Lambda\|_{L^1(S) \to \mathbb{R}}$.

2.3.7 Proposition. $L^1(S)$ *is separable, i.e., there is a countable dense subset.*

Sketch of proof. All rational linear combinations of indicators of dyadic cubes is an example of a countable dense subset of $L^1(\mathbb{R}^d)$; their restrictions to S gives a similar subset of $L^1(S)$. $\square$

2.3.8 Proposition (Weak* sequential compactness of the unit ball)**.** *For every sequence* $(\phi_k)_{k=1}^\infty$ *in* $B_{L^\infty(S)}$, *there is a subsequence* $(\phi_{k_j})_{j=1}^\infty$ *and a further* $\phi \in B_{L^\infty(S)}$ *such that* $\int \phi_{k_j} h \to \int \phi h$ *for every* $h \in L^1(S)$.

Proof. Let $(\psi_i)_{i=1}^\infty \subset L^1(S)$ be a countable dense subset. Since the sequence of real numbers $\int \phi_k \psi_1$ is bounded, there is an infinite subsequence K_1 such that $\int \phi_k \psi_1$ converges when $K_1 \ni k \to \infty$. By induction if an infinite sequence K_{j-1} is already chosen, since the sequence of real numbers $\int \phi_k \psi_j$ is bounded, there is a further infinite subsequence $K_j \subset K_{j-1}$ such that $\int \phi_k \psi_j$ converges when $K_j \ni k \to \infty$. Let us form the sequence $K = \{k_1, k_2, \ldots\}$ so that $k_1 \in K_1$ and, if k_{j-1} is already chosen, we pick $k_j \in K_j$ so that $k_j > k_{j-1}$. Now $\int \phi_k \psi_j$ converges for every j as $K \ni k \to \infty$, since the convergence only depends on the tail $(k_j, k_{j+1}, \ldots) \subset K_j$, and we know that convergence happens along the sequence K_j.

We then show that $\int \phi_k h$ converges as $K \in k \to \infty$ for every $h \in L^1(S)$ by verifying Cauchy's criterion. Let $\varepsilon > 0$ be given. Choose k so that $\|h - \psi_k\|_{L^1} <$

$\varepsilon/3$. Then

$$\Big|\int \phi_m h - \int \phi_\ell h\Big| = \Big|\int (\phi_m - \phi_\ell)(h - \psi_k) + \int (\phi_m - \phi_\ell)\psi_k\Big|$$
$$\leq \|\phi_m - \phi_\ell\|_\infty \|h - \psi_k\|_1 + \Big|\int \phi_m \psi_k - \int \phi_\ell \psi_k\Big| =: I + II.$$

Here $I \leq 2\varepsilon/3$ for all m, ℓ, and $II < \varepsilon/3$ as soon as $K \ni m, \ell \geq N_\varepsilon$, since $\int \phi_m \psi_k$ converges as $m \to \infty$.

Thus we have shown the existence of

$$\Lambda h := \lim_{K \ni k \to \infty} \int \phi_k h$$

for every $h \in L^1(S)$, and clearly this defines a linear mapping $\Lambda : L^1(S) \to \mathbb{R}$ of norm at most $\sup \|\phi_k\|_\infty \leq 1$. By Proposition 2.3.6, there exists $\phi \in B_{L^\infty(S)}$ such that $\Lambda h = \int \phi h$. $\square$

2.3.9 Lemma. *Let $f \in L^1(Q; \mathbb{R}^n)$ and let $g : \mathbb{R}^d \to \mathbb{R}^n$ be a measurable function such that $g(x) \in \langle\!\langle f \rangle\!\rangle_Q$ for almost every $x \in \mathbb{R}^d$. Then there is a function $k \in B_{L^\infty(\mathbb{R}^d \times Q)}$ such, at almost every $x \in \mathbb{R}^d$, we have*

$$g(x) = \fint_Q k(x, y) f(y)\,\mathrm{d}y. \tag{2.3.10}$$

On the formal level, the conclusion (2.3.10) is obvious, since $g(x) \in \langle\!\langle f \rangle\!\rangle_Q$ means the existence of some $k(x, \cdot) \in B_{L^\infty(Q)}$ such that (2.3.10) holds; the point of the lemma is that the function $k(x, y)$ can be chosen to be jointly measurable.

Proof. I: Case of simple g. Suppose first that $g = \sum_{j=1}^J a_j 1_{A_j}$ is a simple function, where the sets A_k is measurable and disjoint. Since $a_j \in \langle\!\langle f \rangle\!\rangle_Q$, we have $a_j = \fint_Q \phi_j(y) f(y)\,\mathrm{d}y$ for some $\phi_j \in B_{L^\infty}(Q)$. But then

$$g(x) = \fint_Q \sum_{j=1}^J 1_{A_j}(x)\phi_j(y) f(y)\,\mathrm{d}y,$$

where $\sum_{j=1}^J 1_{A_j}(x)\phi_j(y) \in B_{L^\infty(\mathbb{R}^d \times Q)}$.

II: Approximation by simple functions. Let g be a general function as in the assumptions. and $\varepsilon > 0$. Since $\langle\!\langle f \rangle\!\rangle_Q$ is compact, we can find a finite ε-net $(y_j)_{j=1}^J$ in $\langle\!\langle f \rangle\!\rangle_Q$, namely, for every $y \in \langle\!\langle f \rangle\!\rangle_Q$ there is some y_j such that $|y - y_j| \leq \varepsilon$. We then define the measurable sets

$$A_j := \{x \in \mathbb{R}^d : |g(x) - y_j| \leq \varepsilon \text{ and } |g(x) - y_i| > \varepsilon\ \forall i = 1, \ldots, j-1\}$$

and let

$$g_\varepsilon := \sum_{j=1}^J y_j 1_{A_j}.$$

Then $\|g - g_\varepsilon\|_\infty \le \varepsilon$ and g_ε is a simple function with $g_\varepsilon(x) \in \langle\!\langle f \rangle\!\rangle_Q$ for all $x \in \mathbb{R}^d$.

III: Conclusion. By the previous parts, for a general g as in the assumptions, we can find simple functions $g_j \to g$ in $L^\infty(\mathbb{R}^d)$ and kernels $k_j(x,y) \in B_{L^\infty(\mathbb{R}^d \times Q)}$ such that

$$g_j(x) = \fint_Q k_j(x,y) f(y)\,\mathrm{d}y. \tag{2.3.11}$$

By Proposition 2.3.8, there is a subsequence J and a function $k \in B_{L^\infty(\mathbb{R}^d \times Q)}$ such that $\int_{\mathbb{R}^d \times Q} k_j h \to \int_{\mathbb{R}^d \times Q} kh$ for all $h \in L^1(\mathbb{R}^d \times Q)$. Integrating (2.3.11) against any $\psi \in L^1(\mathbb{R}^d)$, we thus find that

$$\begin{aligned}
\int_{\mathbb{R}^d} g(x)\psi(x)\,\mathrm{d}x &= \lim_{J \ni j \to \infty} \int_{\mathbb{R}^d} g_j(x)\psi(x)\,\mathrm{d}x \\
&= \lim_{J \ni j \to \infty} \int_{\mathbb{R}^d} \fint_Q k_j(x,y) f(y)\,\mathrm{d}y \psi(x)\,\mathrm{d}x \\
&= \int_{\mathbb{R}^d} \fint_Q k(x,y) f(y)\,\mathrm{d}y \psi(x)\,\mathrm{d}x =: \int_{\mathbb{R}^d} G(x)\psi(x)\,\mathrm{d}x.
\end{aligned}$$

When such an inequality holds for all $\psi \in L^1(\mathbb{R}^d)$, it follows that $g = G$ almost everywhere. (For instance, given $x \in \mathbb{R}^d$, consider $\psi := |B(x,r)|^{-1} 1_{B(x,r)}$ and let $r \to 0$; then the left side converges to $g(x)$ and the right side to $G(x)$ at almost every x.) But, looking at the formula of G, the identity $g = G$ is precisely what the lemma claimed. □

In order to prove Theorem 2.3.2, we need a workable criterion to check the membership of a vector in $\langle\!\langle f \rangle\!\rangle_Q$. A key tool in this respect is provided by the following fundamental result about the shape of convex sets, which we prove in the following section:

2.3.12 Theorem (John ellipsoid theorem [Joh48]). *Let $K \subset \mathbb{R}^n$ be a compact convex symmetric set. Then there is a closed ellipsoid E centred at the origin such that $E \subset K \subset \sqrt{n}E$.*

Such an ellipsoid is referred to as the John ellipsoid of K.

By definition, a closed ellipsoid E centred at the origin is the image of the unit ball $B = \{x \in \mathbb{R}^n : |x| \le 1\}$ under the action of a matrix $A \in \mathscr{L}(\mathbb{R}^n)$, i.e. $E = AB$. (An ellipsoid centred at $x_0 \in \mathbb{R}^n$ is then a set of the form $x_0 + AB$.) To express this in another form, recall from Linear Algebra that every matrix has a *singular value decomposition*

$$A = \sum_{i=1}^n \sigma_i e_i \otimes f_i, \qquad e_i \otimes f_i(x) := e_i(f_i|x),$$

where $(e_i)_{i=1}^n$ and $(f_i)_{i=1}^n$ are two orthonormal bases of $\mathbb{R}^n$, and $\sigma_1 \ge \sigma_2 \ge \ldots \ge \sigma_n \ge 0$ are the singular values of A. The singular value decomposition leads to the *polar decomposition*

$$A = \Big(\sum_{i=1}^{n} \sigma_i e_i \otimes e_i\Big)\Big(\sum_{j=1}^{n} e_j \otimes f_j\Big) =: RU,$$

where the first identity follows easily from the result of Exercise 2.3.20 below. Since the matrix U maps B bijectively into itself, we find that

$$\begin{aligned} AB = RB &= \Big\{\sum_{i=1}^{n} \sigma_i e_i (e_i|x) : x \in B\Big\} = \Big\{\sum_{i=1}^{n} \sigma_i x_i e_i : \sum_{i=1}^{n} |x_i|^2 \leq 1\Big\} \\ &= \Big\{\sum_{i=1}^{n} y_i e_i : \sum_{i=1}^{n} \Big(\frac{|y_i|}{\sigma_i}\Big)^2 \leq 1\Big\}. \end{aligned}$$

With the change of variable $y_i = \sigma_i x_i$, the last formula is clearly valid in the non-degenerate case when all $\sigma_i > 0$; in the degenerate case, it is still valid provided that we interpret

$$\frac{|y_i|}{0} := \begin{cases} 0, & \text{if } |y_i| = 0, \\ \infty, & \text{if } |y_i| > 0, \end{cases}$$

so that

$$\sum_{i=1}^{n} \Big(\frac{|y_i|}{\sigma_i}\Big)^2 \leq 1 \qquad \Leftrightarrow \qquad \sum_{i:\sigma_i>0} \Big(\frac{|y_i|}{\sigma_i}\Big)^2 \leq 1 \text{ and } y_i = 0 \text{ if } \sigma_i = 0.$$

The orthonormal vectors $(e_i)_{i=1}^n$ are called the *principal axes* of the ellipsoid $E = AB = RB$. They lead to a useful sufficient condition for membership in $\langle\langle f\rangle\rangle_Q$ as follows:

2.3.13 Lemma. *Let $(e_i)_{i=1}^n$ be the principal axes of the John ellipsoid of $\langle\langle f\rangle\rangle_Q$. If $|(e_i|x)| \leq \frac{1}{n} \fint_Q |(e_i|f)|$ for every $i = 1, \ldots, n$, then $x \in \langle\langle f\rangle\rangle_Q$.*

Proof. For some numbers $\sigma_1 \geq \ldots \geq \sigma_n \geq 0$, the John ellipsoid E consists of all $y \in \mathbb{R}^n$ such that

$$\sum_{i=1}^{n} \Big(\frac{|(e_i|y)|}{\sigma_i}\Big)^2 \leq 1.$$

Since

$$y := \fint_Q \operatorname{sgn}((e_j|f)) f \in \langle\langle f\rangle\rangle_Q \subset \sqrt{n} E,$$

and hence $\frac{1}{\sqrt{n}} y \in E$

$$\begin{aligned} 1 \geq \sum_{i=1}^{n} \Big(\frac{|(e_i|y)|}{\sqrt{n}\sigma_i}\Big)^2 \geq \Big(\frac{|(e_j|y)|}{\sqrt{n}\sigma_j}\Big)^2 &= \Big(\frac{1}{\sqrt{n}\sigma_j} \fint_Q \operatorname{sgn}((e_j|f))(e_j|f)\Big)^2 \\ &= \Big(\frac{1}{\sqrt{n}\sigma_j} \fint_Q |(e_j|f)|\Big)^2. \end{aligned}$$

Thus

$$\fint_Q |(e_j|f)| \leq \sqrt{n}\sigma_j$$

for each $j = 1, \ldots, n$.

Now, if $x \in \mathbb{R}^n$ satisfies $|(e_i|x)| \leq \frac{1}{n} \fint_Q |(e_i|f)|$ for each i, then

$$\sum_{i=1}^n \Big(\frac{|(e_i|x)|}{\sigma_i}\Big)^2 \leq \sum_{i=1}^n \Big(\frac{1}{n\sigma_i} \fint_Q |(e_i|f)|\Big)^2 \leq \sum_{i=1}^n \Big(\frac{1}{\sqrt{n}}\Big)^2 = \sum_{i=1}^n \frac{1}{n} = 1,$$

and hence $x \in E \subset \langle\!\langle f \rangle\!\rangle_Q$. □

2.3.14 Lemma. *Under the assumptions of Theorem 2.3.2, for any cube Q_0 and $f \in L^1(3Q_0; \mathbb{R}^n)$ and $\varepsilon \in (0,1)$, there are disjoint subcubes $Q_j \in \mathscr{D}(Q_0)$ such that*

$$\sum_j |Q_j| \leq \varepsilon |Q_0|.$$

and

$$1_{Q_0} T(1_{3Q_0} f) - \sum_j 1_{Q_j} T(1_{3Q_j} f) \in 1_{Q_0} \frac{c_{d,n} c_T}{\varepsilon} \langle\!\langle f \rangle\!\rangle_{3Q_0}.$$

More precisely, there is a function $k_{Q_0} \in B_{L^\infty(Q_0 \times 3Q_0)}$ such that

$$\begin{aligned} &1_{Q_0}(x) T(1_{3Q_0} f)(x) - \sum_j 1_{Q_j}(x) T(1_{3Q_j} f)(x) \\ &\qquad = 1_{Q_0}(x) \frac{c_{d,n} c_T}{\varepsilon} \fint_{3Q_0} k_{Q_0}(x,y) f(y) \,\mathrm{d}y. \end{aligned} \tag{2.3.15}$$

Proof. We apply the scalar-valued version of this result, Lemma 1.4.4, to each of the n functions $(e_i|f)$, where $(e_i)_{i=1}^n$ are the principal axes of the John ellipsoid of $\langle\!\langle f \rangle\!\rangle_Q$. Thus, for each $i = 1, \ldots, n$, we find disjoint cubes $Q_j^i \in \mathscr{D}(Q_0)$ such that

$$\sum_j |Q_j^i| \leq \varepsilon |Q_0|$$

and, if $Q_j \in \mathscr{D}(Q_0)$ are (possibly bigger) disjoint cubes such that $\bigcup_j Q_j \supset \bigcup_j Q_j^i$, then

$$\Big| 1_{Q_0} T(1_{3Q_0}(e_i|f)) - \sum_j 1_{Q_j} T(1_{3Q_j}(e_i|f)) \Big| \leq 1_{Q_0} \frac{c_d c_T}{\varepsilon} \fint_{3Q_0} |(e_i|f)| \tag{2.3.16}$$

Let $\{Q_j\}_j$ be the maximal cubes among all $\{Q_j^i\}_{i,j}$. Then, on the one hand,

$$\sum_j |Q_j'| \leq \sum_{i,j} |Q_j^i| \leq n\varepsilon |Q_0|.$$

On the other hand, we clearly have $\bigcup_j Q_j \supset \bigcup_j Q_j^i$ for each i, and hence (2.3.16) holds for this particular choice of the cubes Q_j.

Let us abbreviate

$$F := 1_{Q_0} T(1_{3Q_0} f) - \sum_j 1_{Q_j} T(1_{3Q_j} f).$$

Then the left side of (2.3.16) is simply $|(e_i|F)|$, and we see that (2.3.16) holding for all $i = 1, \ldots, n$ is exactly the sufficient condition, provided by Lemma 2.3.13, for

$$F \in 1_{Q_0} n \frac{c_d c_T}{\varepsilon} \langle\!\langle f \rangle\!\rangle_{3Q_0}.$$

Replacing ε by ε/n we obtain the first claim of the lemma with $c_{d,n} = c_d n^2$. The second claim is then a direct consequence of this via Lemma 2.3.9. □

Proof of the Convex Body Domination Theorem 2.3.2. (This is almost the same as the proof of Lerner's Theorem 1.4.2, replacing the use of Lemma 1.4.4 by Lemma 2.3.14.) Iterating the conclusion of Lemma 2.3.14, we obtain disjoint families $\{Q_j^k\}_j$ of dyadic subcubes of Q_0 such that $\bigcup_j Q_j^{k+1} \subset \bigcup_i Q_i^k$,

$$\sum_{j: Q_j^{k+1} \subset Q_i^k} |Q_j^{k+1}| \leq \varepsilon |Q_i^k|$$

and, with some $k_{Q_j^k} \in B_{L^\infty(Q_j^k \times 3Q_j^k)}$,

$$\begin{aligned} 1_{Q_0}(x) T(1_{3Q_0} f)(x) = \sum_{k=0}^{K-1} \sum_j 1_{Q_j^k}(x) \frac{c_{d,n} c_T}{\varepsilon} \fint_{3Q_j^k} k_{Q_j^k}(x,y) f(y) \, \mathrm{d}y \\ + \sum_j 1_{Q_j^K}(x) T(1_{3Q_j^K} f)(x). \end{aligned}$$

As $K \to \infty$, we conclude that (a.e.)

$$1_{Q_0}(x) T(1_{3Q_0} f)(x) = \sum_{k=0}^{\infty} \sum_j 1_{Q_j^k}(x) \frac{c_{d,n} c_T}{\varepsilon} \fint_{3Q_j^k} k_{Q_j^k}(x,y) f(y) \, \mathrm{d}y, \tag{2.3.17}$$

where $\{Q_j^k\}_{k,j}$ is $(1-\varepsilon)$-sparse, since the sets $E(Q_j^k) := Q_j^k \setminus \bigcup_i Q_i^{k+1}$ are disjoint and $|E(Q_j^k)| \geq (1-\varepsilon)|Q_j^k|$.

As in the proof of Theorem 1.4.2, we pick a partition $\mathscr{S}_0$ of $\mathbb{R}^d$ by dyadic cubes Q_0 such that $3Q_0 \supset \operatorname{supp} f$. Then

$$Tf(x) = \sum_{Q_0 \in \mathscr{S}_0} 1_{Q_0}(x) Tf(x) = \sum_{Q_0 \in \mathscr{S}_0} 1_{Q_0}(x) T(1_{3Q_0} f)(x),$$

and we apply (2.3.17) to each $Q_0 \in \mathscr{S}_0$. The resulting collection $\mathscr{S}$ that consists of all Q_j^k related to each $Q_0 \in \mathscr{S}_0$ is still $(1-\varepsilon)$-sparse. □

2.3.18 Corollary (to Theorem 2.3.2)**.** *Under the assumptions of Theorem 2.3.2, we can also write*

$$\begin{aligned} Tf(x) &= \frac{c_{d,n}c_T}{\varepsilon} \sum_{j=1}^{3^d} \sum_{S \in \mathscr{S}_j} 1_S(x) \fint_S k_S(x,y) f(y) \, \mathrm{d}y \\ &\in \frac{c_{d,n}c_T}{\varepsilon} \sum_{j=1}^{3^d} \sum_{S \in \mathscr{S}_j} 1_S(x) \langle\!\langle f \rangle\!\rangle_S \end{aligned}$$

for $3^{-d}(1-\varepsilon)$*-sparse collections* $\mathscr{S}_j \subset \mathscr{D}_j$ *of dyadic-type collections* $\mathscr{D}_j$ *and functions* $k_S \in B_{L^\infty(S \times S)}$.

Proof. Recall from Proposition 1.5.1 that $\{3Q : Q \in \mathscr{D}\}$ can be divided into 3^d dyadic-type subcollections $\mathscr{D}_j$. Thus, defining $k_{3S}(x,y) := 1_S(x) k_S(x,y) \in B_{L^\infty(3S \times 3S)}$ (i.e., just extending k_S by zero for $x \in 3S \setminus S$),

$$\begin{aligned} \sum_{S \in \mathscr{S}} 1_S(x) \fint_{3S} k_S(x,y) f(y) \, \mathrm{d}y &= \sum_{S \in \mathscr{S}} 1_{3S}(x) \fint_{3S} k_{3S}(x,y) f(y) \, \mathrm{d}y \\ &= \sum_{j=1}^{3^d} \sum_{\substack{S \in \mathscr{S} \\ 3S \in \mathscr{D}_j}} 1_{3S}(x) \fint_{3S} k_{3S}(x,y) f(y) \, \mathrm{d}y \\ &= \sum_{j=1}^{3^d} \sum_{S' \in \mathscr{S}_j} 1_{S'}(x) \fint_{S'} k_{S'}(x,y) f(y) \, \mathrm{d}y, \end{aligned}$$

where each $\mathscr{S}_j := \{S' = 3S \in \mathscr{D}_j : S \in \mathscr{S}\}$ is $3^{-d}(1-\varepsilon)$-sparse, since the sets $E(3S) := E(S)$ are still disjoint and $|E(3S)| = |E(S)| \geq (1-\varepsilon)|S| = (1-\varepsilon)3^{-d}|3S|$. □

2.3.19 Exercise. Show that definition (2.3.1) is independent of the chosen orthonormal basis $(e_i)_{i=1}^n$ on the right, i.e., with another orthonormal basis $(e_i')_{i=1}^n$, we get the same result. Check also that

$$(v|Tf(x)) = T(v|f)(x)$$

for any vector $v \in \mathbb{R}^n$

2.3.20 Exercise. For any vectors $e, f, g, h \in \mathbb{R}^n$, prove the operator identity

$$(e \otimes f)(g \otimes h) = (f|g) e \otimes h.$$

(Warning: We defined $e \otimes f(x) := e(f|x)$, but some other texts may also use the 'wrapped' definition "$e \otimes f(x) := f(e|x)$".)

2.4 The John ellipsoid

We now turn to the proof of the key geometric tool behind the Convex Body Domination Theorem 2.3.2, which we restate for convenience:

2.4.1 Theorem (John ellipsoid theorem [Joh48])**.** *Let $K \subset \mathbb{R}^n$ be a compact convex symmetric set. Then there is a closed ellipsoid E centred at the origin such that $E \subset K \subset \sqrt{n}E$.*

In fact, we will show that, in the non-degenerate case, the ellipsoid of maximal measure contained in K satisfies the requirements of the theorem. We begin with the existence of such a maximiser:

2.4.2 Lemma. *Let $K \subset \mathbb{R}^n$ be a compact non-empty set. Among all closed origin-centred ellipsoids $E \subset K$, there is one of maximal measure.*

Note that it may happen that K only contains degenerate ellipsoids of measure zero. In this case any of them has 'maximal measure', and the conclusion is trivially true.

Proof. The ellipsoids in question are given by $E = AB$, where B is the closed unit ball and the matrix $A \in \mathscr{L}(\mathbb{R}^n) \eqsim \mathbb{R}^{n\times n}$ is subject to the condition $AB \subset K$. We claim that the set of such A is a compact subset of $\mathscr{L}(\mathbb{R}^n)$. Boundedness follows from the boundedness of K as follows: $Ax \in K \subset B(0,R)$ for all $x \in B$ means that

$$\|A\|_{op} = \sup\{|Ax| : x \in B\} \leq R.$$

To check closedness, suppose that $A_nB \subset K$ and $\|A_n - A\|_{op} \to 0$ for some matrix A. Then for all $x \in B$, we have $|A_nx - Ax| \leq \|A_n - A\|_{op} \to 0$, and thus $K \ni A_nx \to Ax$. Since K is closed, we find that $Ax \in K$. Since this holds for allă$x \in B$, we have $AB \subset K$. Thus the set of relevant A is a closed and bounded subset of the finite-dimensional space $\mathscr{L}(\mathbb{R}^n)$, and therefore compact. On the other hand, the measure of the ellipsoid $E = AB$ is $|E| = |\det A|$ is a continuous function of A (as a polynomial in the coefficients of A with respect to a fixed basis) and therefore reaches its maximum on a compact set. □

The technical core of the proof is contained in the following lemma:

2.4.3 Lemma. *Let $B \subset \mathbb{R}^n$ be the closed unit ball and $p = (p_1, 0)$ with $p_1 > 1$. Then the ellipsoid E consisting of all $(x_1, x_\perp) \in \mathbb{R}\times\mathbb{R}^{n-1}$ such that*

$$\Big(\frac{|x_1|}{a}\Big)^2 + \Big(\frac{|x_\perp|}{b}\Big)^2 \leq 1$$

is contained in the convex hull K of $B \cup \{p, -p\}$ for any $b \in (0,1)$ and $a > 1$ given by

$$a^2 = p_1^2 - (p_1^2 - 1)b^2 = p_1^2(1-b^2) + b^2$$

Taking this for granted for a moment (we return to it soon), we prove the following special case of Theorem 2.4.1

2.4.4 Lemma. *Let $K \subset \mathbb{R}^n$ be a convex symmetric body, and suppose that the closed unit ball B has maximal volume among all origin-centred ellipsoids $E \subset K$. Then $K \subset \sqrt{n}B$.*

Proof. Consider a point $p \in K$. Since K is symmetric, also $-p \in K$, and since K is convex and $B \subset K$, we know that the convex hull of $B \cup \{-p, p\}$ is contained in K. By Lemma 2.4.3, this convex hull, and hence K itself, contains for every $t \in (0,1)$ the ellipsoid E_t of all $(x_1, x_\perp) \in \mathbb{R} \times \mathbb{R}^{n-1}$ such that

$$\frac{|x_1|^2}{|p|^2 - (|p|^2-1)t} + \frac{|x_\perp|^2}{t} \leq 1.$$

By scaling properties of the Lebesgue measure,

$$|E_t|^2 = |B|^2(|p|^2 - (|p|^2-1)t)t^{n-1} = |B|^2(|p|^2 t^{n-1} - (|p|^2-1)t^n).$$

Since $B = E_1 \subset K$ has maximal volume among all ellipsoids $E \subset K$, we must have $|E_t| \leq |E_1|$ for $t \in (0,1)$, so in particular

$$\begin{aligned} 0 \leq \lim_{t \to 1-} \frac{|E_1|^2 - |E_t|^2}{1-t} &= \frac{d}{dt}|E_t|^2\Big|_{t=1} \\ &= |B|^2(|p|^2(n-1)t^{n-2} - (|p|^2-1)nt^{n-1})\Big|_{t=1} \\ &= |B|^2(|p|^2(n-1) - (|p|^2-1)n) = |B|^2(n - |p|^2), \end{aligned}$$

hence $|p| \leq \sqrt{n}$, and thus $p \in \sqrt{n}B$. $\square$

From the special case, the general theorem follows quite easily:

Proof of the John Ellipsoid Theorem 2.4.1. Case I: K has non-empty interior. Let $E \subset K$ be an origin-centred ellipsoid of maximal measure, whose existence is guaranteed by Lemma 2.4.2. Then $E = AB$ for some non-degenerate matrix A, and we find that $B \subset A^{-1}K$ is an origin-centred ellipsoid of maximal measure in the convex symmetric set $A^{-1}K$. By Lemma 2.4.4, we have $A^{-1}K \subset \sqrt{n}B$, and thus $K \subset \sqrt{n}AB = \sqrt{n}E$.

Case II: K has empty interior. Let $v_1, \dots, v_k \in K$ be a maximal collection of linearly independent vectors. Then $K \subset \operatorname{span}(v_1, \dots, v_k) =: V \approx \mathbb{R}^k$, and K has a nonempty interior viewed as a subset of V, since it contains (by convexity and symmetry) all vectors of the form $\sum_{i=1}^k \lambda_i v_i$ with $\sum_{i=1}^k |v_i| \leq 1$. By Case I, we find an ellipsoid $E \subset K \subset \sqrt{k}E \subset \sqrt{n}E$; this E is a proper ellipsoid in the k-dimensional subspace V, and hence a degenerate ellipsoid in $\mathbb{R}^n$. $\square$

To complete the proof, we still need to provide:

Proof of Lemma 2.4.3. Let us start by finding the tangent lines from $p = (p_1, 0_\perp)$ to B. Let $z = (z_1, z_\perp) \in S = \partial B$ be a point where such a tangent touches B. Thus $z \perp (z - p)$, i.e., $|z_1|^2 + |z_\perp|^2 = |z|^2 = 1 = z \cdot p = z_1 p_1$. Thus $z_1 = 1/p_1$ and $|z_\perp| = \sqrt{1 - 1/p_1^2} = p_1^{-1}\sqrt{p_1^2 - 1}$.

Since $z, p \in K$, the line-segments connecting p and any such z are contained in K. These lines consist of points of the form $(x_1, x_\perp)$ with $p_1^{-1} \le x_1 \le p_1$ and $|x_\perp| = (p_1 - x_1)/\sqrt{p_1^2 - 1}$. Since also the points $(x_1, 0_\perp)$ (being on the line from the origin to p) lie in K, we conclude that any point of the form $(x_1, x_\perp)$ with $p_1^{-1} \le x_1 \le p_1$ and $|x_\perp| \le (p_1 - x_1)/\sqrt{p_1^2 - 1}$ lies in K.

Let now $x = (x_1, x_\perp)$ be a point of the ellipsoid, and assume by symmetry that $x_1 \ge 0$. (Else, we would do similar considerations with $-p$ in place of p; note that otherwise the point $-p$ is nowhere used in this proof.)

Case $0 \le x_1 \le p_1^{-1}$*:* We want to check that in this case $x \in B$. We have

$$|x_\perp|^2 \le b^2(1 - x_1^2/a^2),$$

and this is $\le 1 - x_1^2$ provided that $(1 - b^2/a^2)x_1^2 \le 1 - b^2$. For $b < 1 < a$, the left side is largest at $x_1 = p_1^{-1}$, leading to the constraint

$$a^2 \le \frac{b^2}{1 - p_1^2(1 - b^2)} \tag{2.4.5}$$

if $p_1^2(1 - b^2) < 1$, while any a is good in the complementary case.

Case $p_1^{-1} \le x_1 \le p_1$*:* We want to check that $|x_\perp| \le (p_1 - x_1)/\sqrt{p_1^2 - 1}$, which we already checked to imply that $x \in K$ for this range of x_1. Letting $|x_\perp|$ be as large as it can be within the ellipsoid, we are led to the constraint

$$b^2(1 - x_1^2/a^2) \le \frac{(p_1 - x_1)^2}{p_1^2 - 1},$$

which can be written as

$$P(x_1) := [(p_1^2 - 1)b^2/a^2 + 1]x_1^2 - 2p_1x_1 + [p_1^2 - (p_1^2 - 1)b^2] \ge 0.$$

For a polynomial $P(x_1) = Ax_1^2 + 2Bx_1 + C$ with $A > 0$, the condition $P(x_1) \ge 0$ is equivalent to $AC - B^2 \ge 0$, which in this case reads as

$$\begin{aligned} 0 &\le [(p_1^2 - 1)b^2/a^2 + 1][p_1^2 - (p_1^2 - 1)b^2] - p_1^2 \\ &= (p_1^2 - 1)p_1^2b^2/a^2 - (p_1^2 - 1)^2b^4/a^2 - (p_1^2 - 1)b^2 \\ &= (p_1^2 - 1)b^2/a^2 \cdot [p_1^2 - (p_1^2 - 1)b^2 - a^2], \end{aligned}$$

which is satisfies provided that

$$a^2 \le p_1^2 - (p_1^2 - 1)b^2 = p_1^2(1 - b^2) + b^2. \tag{2.4.6}$$

We check that the second constraint (2.4.6) is always stronger than the first constraint (2.4.5). Namely,

$$\begin{aligned} [p_1^2(1 - b^2) + b^2][1 - p_1^2(1 - b^2)] - b^2 &= p_1^2(1 - b^2) - p_1^4(1 - b^2)^2 - b^2p_1^2(1 - b^2) \\ &= p_1^2(1 - b^2)[1 - p_1^2(1 - b^2) - b^2] \\ &= p_1^2(1 - b^2)(1 - b^2)(1 - p_1^2) < 0. \end{aligned}$$

□

2.4.7 Remark. The John ellipsoid theorem is originally from [Joh48]. Our presentation using elementary geometry has been adapted from [How97]. There is also a version of John's theorem for compact convex sets that are not necessarily symmetric. In this case, the factor $\sqrt{n}$ must be replaced by n. The ellipsoid theorem can be restated in terms of norms on an n-dimensional vector space as follows: For any norm $\| \, \|_X$, there is a Hilbertian norm $\| \, \|$ (i.e., one induced by an inner product) such that $\|x\| \leq \|x\|_X \leq \sqrt{n}\|x\|$ for all x. (This restatement is based on the fact that there is a bijective correspondence between norms and their closed unit balls $\{x : \|x\|_X \leq 1\}$, and closed unit balls are precisely the compact, convex and symmetric sets; an ellipsoidal unit ball corresponds to a Hilbertian norm.) For a more functional analytic approach from this perspective, see e.g. [Pis86].

2.4.8 Remark. The constant $\sqrt{n}$ in the John ellipsoid theorem is optimal. In fact, if Q is a cube of sidelength 2 centred at the origin, and E is a closed ellipsoid centred at the origin such that $E \subset Q \subset tE$, then $t \geq \sqrt{n}$, where the equality is reached if and only if E is the unit ball B. While it might be intuitively "clear" that B is the "best" ellipsoid inside Q, this is perhaps not entirely trivial to justify rigorously, and hints to do this are provided in the following exercises:

2.4.9 Exercise. Let $(e_i)_{i=1}^n$ and $(f_i)_{i=1}^n$ be two orthonormal bases of $\mathbb{R}^n$. Consider the cube

$$Q := \Big\{x = \sum_{i=1}^n x_i e_i : \max_i |x_i| \leq 1\Big\}$$

and the ellipsoid

$$E := \Big\{y = \sum_{i=1}^n y_i f_i : \sum_{i=1}^n \Big(\frac{y_i}{\sigma_i}\Big)^2 \leq 1\Big\}.$$

Prove that $E \subset Q$ *if and only if*

$$\max_{1 \leq i \leq n} \sum_{j=1}^n (e_i|f_j)^2 \sigma_j^2 \leq 1,$$

and deduce that if $E \subset Q$, then

$$\sum_{j=1}^n \sigma_j^2 \leq n.$$

2.4.10 Exercise. Let E and Q be as in the previous exercise. Prove that if $Q \subset E$, then

$$\sum_{i=1}^n \frac{1}{\sigma_i^2} \leq 1.$$

Hint: Write the condition that $x \in E$ for each of the 2^n corners x of Q, and take the *average* of the obtained bound over all corners. You may find it helpful to observe (and verify) that

$$\mathbb{E}\Big(\sum_{j=1}^n \varepsilon_j t_j\Big)^2 = \sum_{j=1}^n t_j^2,$$

where ε_j are independent random variables taking the values ± 1 with equal probability, and $\mathbb{E}$ is the expectation (the average) over all possible values of these random variables.

2.4.11 Exercise. Let E and Q be as above, and $t \geq 1$. Prove that if $E \subset Q \subset tE$, then $t \geq \sqrt{n}$, and the equality holds if and only if $E = B$ is the unit ball. Hint: Use the conclusions of the previous exercises, the latter suitably scaled for tE in place of E. Observe that $E = B$ is equivalent to $\sigma_i \equiv 1$ for all $i = 1, \ldots, n$.

2.4.12 Exercise. Let E and Q be as above. Prove that if $E \subset Q$, then $|E| \leq |B|$, and the equality holds if and only if $E = B$ is the unit ball. Hint: Express $|E|$ in terms of the number $\sigma_1, \ldots, \sigma_n$, and use the inequality between geometric and arithmetic means.

2.5 Calderón–Zygmund operators on $L^2(W)$

Our goal is now to apply the Convex Body Domination Theorem 2.3.2 to estimate the norm of Calderón–Zygmund operators on the matrix-weighted $L^2(W)$ space. That is, we want to prove an inequality of the type

$$\|Tf\|_{L^2(W)} \leq K\|f\|_{L^2(W)},$$

for $W \in A_2$, with an estimate on K in terms of $[W]_{A_2}$.

Let us observe some reformulations. Since $\|f\|_{L^2(W)} = \|W^{1/2}f\|_{L^2}$, substituting $h = W^{1/2}f$ and solving $f = W^{-1/2}h = \Sigma^{1/2}h$, where $\Sigma := W^{-1}$ (note that $W \in A_2$ is almost everywhere invertible), we find that it is equivalent to prove that

$$\|W^{1/2}T(\Sigma^{1/2}h)\|_{L^2} \leq K\|h\|_{L^2},$$

which is the boundedness of the operator $h \mapsto W^{1/2}T(\Sigma^{1/2}h)$ on the *unweighted* space $L^2 = L^2(\mathbb{R}^d; \mathbb{R}^n)$.

We also recall that $W \in A_2$ implies $\Sigma \in L^1_{\text{loc}}$, and hence $\Sigma^{1/2} \in L^2_{\text{loc}}$. If $h \in L^2_c$ (compactly supported L^2 functions), then $\Sigma^{1/2}h \in L^1_c$ is a function to which the Convex Body Domination Theorem 2.3.2 applies, and we deduce (more precisely from Corollary 2.3.18) that

$$T(\Sigma^{1/2}h)(x) = c_{d,n}c_T \sum_{j=1}^{3^d} \sum_{S \in \mathscr{S}_j} 1_S(x) \fint_S k_S(x,y)\Sigma(y)^{1/2}h(y)\,\mathrm{d}y$$

for some $k_S \in B_{L^\infty}(S \times S)$. Hence

$$\begin{aligned} &W(x)^{1/2} T(\Sigma^{1/2} h)(x) \\ &= c_{d,n} c_T \sum_{j=1}^{3^d} \sum_{S \in \mathscr{S}_j} 1_S(x) \fint_S W(x)^{1/2} k_S(x,y) \Sigma(y)^{1/2} h(y) \,\mathrm{d}y, \end{aligned}$$

and thus

$$\begin{aligned} &|W(x)^{1/2} T(\Sigma^{1/2} h)(x)| \\ &\leq c_{d,n} c_T \sum_{j=1}^{3^d} \sum_{S \in \mathscr{S}_j} 1_S(x) \fint_S |W(x)^{1/2} k_S(x,y) \Sigma(y)^{1/2} h(y)| \,\mathrm{d}y \\ &\leq c_{d,n} c_T \sum_{j=1}^{3^d} \sum_{S \in \mathscr{S}_j} 1_S(x) \fint_S \|W(x)^{1/2} \Sigma(y)^{1/2}\|_{op} |h(y)| \,\mathrm{d}y, \end{aligned} \tag{2.5.1}$$

where in the last step we used the fact that the $k_S(x,y)$ is a scalar of modulus at most one, so it commutes with matrix multiplication and

$$\begin{aligned} |W(x)^{1/2} k_S(x,y) \Sigma(y)^{1/2} h(y)| &= |k_S(x,y) W(x)^{1/2} \Sigma(y)^{1/2} h(y)| \\ &\leq \|W(x)^{1/2} \Sigma(y)^{1/2}\|_{op} |h(y)|. \end{aligned}$$

Notice that the vector-valued function $h \in L^2(\mathbb{R}^d; \mathbb{R}^n)$ and the scalar-valued function $|h| \in L^2(\mathbb{R}^d)$ have the same L^2 norm. Hence, from the pointwise inequality (2.5.1) we deduce:

2.5.2 Lemma. *For any Calderón–Zygmund operator T, or more generally any operator that satisfies the assumptions of Lerner's Theorem 1.4.2, and any matrix-weights W and Σ, we have*

$$\|W^{1/2} T \Sigma^{1/2}\|_{L^2(\mathbb{R}^d;\mathbb{R}^n) \to L^2(\mathbb{R}^d;\mathbb{R}^n)} \leq c_{d,n} c_T \sup_{\mathscr{S}} \|T_{\mathscr{S}}^{W,\Sigma}\|_{L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)},$$

where the supremum is over all η-sparse collections $\mathscr{S}$ of dyadic cubes for a fixed number $\eta = \eta_d \in (0,1)$, and

$$T_{\mathscr{S}}^{W,\Sigma} \phi(x) := \sum_{S \in \mathscr{S}} 1_S(x) \fint_S \|W(x)^{1/2} \Sigma(y)^{1/2}\|_{op} \phi(y) \,\mathrm{d}y.$$

Note that $T_{\mathscr{S}}^{W,\Sigma}$ is a scalar-valued operator, but its kernel is defined in terms of norms of the matrix-valued functions W and Σ. We now turn to the analysis of these operators.

We use the standard duality

$$\|T_{\mathscr{S}}^{W,\Sigma} \phi\|_{L^2} = \sup \Big\{ \int \psi(x) T_{\mathscr{S}}^{W,\Sigma} \phi(x) \,\mathrm{d}x : \|\psi\|_{L^2} \leq 1 \Big\},$$

where

$$\int \psi(x) T_{\mathscr{S}}^{W,\Sigma} \phi(x)\,\mathrm{d}x = \sum_{S\in\mathscr{S}} \frac{1}{|S|} \iint_{S\times S} \psi(x) \|W(x)^{1/2}\Sigma(y)^{1/2}\|_{op} \phi(y)\,\mathrm{d}x\,\mathrm{d}y,$$

and it suffices to consider the case $\phi, \psi \geq 0$, since replacing both functions by their absolute values only increases the above expression.

We expand the matrix product by the identity

$$W(x)^{1/2}\Sigma(y)^{1/2} = W(x)^{1/2}\langle W\rangle_S^{-1/2}\langle W\rangle_S^{1/2}\langle \Sigma\rangle_S^{1/2}\langle \Sigma\rangle_S^{-1/2}\Sigma(y)^{1/2},$$

and hence

$$\begin{aligned}
&\|W(x)^{1/2}\Sigma(y)^{1/2}\|_{op} \\
&\leq \|W(x)^{1/2}\langle W\rangle_S^{-1/2}\|_{op} \cdot \|\langle W\rangle_S^{1/2}\langle \Sigma\rangle_S^{1/2}\|_{op} \cdot \|\langle \Sigma\rangle_S^{-1/2}\Sigma(y)^{1/2}\|_{op} \\
&\leq \|W(x)^{1/2}\langle W\rangle_S^{-1/2}\|_{op} \cdot [W]_{A_2}^{1/2} \cdot \|\langle \Sigma\rangle_S^{-1/2}\Sigma(y)^{1/2}\|_{op}.
\end{aligned}$$

This leads to

$$\begin{aligned}
&\int \psi(x) T_{\mathscr{S}}^{W,\Sigma}\phi(x)\,\mathrm{d}x \\
&\leq [W]_{A_2}^{1/2} \sum_{S\in\mathscr{S}} \frac{1}{|S|} \int_S \|W(x)^{1/2}\langle W\rangle_S^{-1/2}\|_{op}\psi(x)\,\mathrm{d}x \\
&\qquad\qquad \times \int_S \|\langle \Sigma\rangle_S^{-1/2}\Sigma(y)^{1/2}\|_{op}\phi(y)\,\mathrm{d}y \\
&\leq [W]_{A_2}^{1/2} \Big(\sum_{S\in\mathscr{S}} \frac{1}{|S|} \Big[\int_S \|W(x)^{1/2}\langle W\rangle_S^{-1/2}\|_{op}\psi(x)\,\mathrm{d}x \Big]^2 \Big)^{1/2} \\
&\qquad\qquad \times \Big(\sum_{S\in\mathscr{S}} \frac{1}{|S|} \Big[\int_S \|\langle \Sigma\rangle_S^{-1/2}\Sigma(y)^{1/2}\|_{op}\phi(y)\,\mathrm{d}y \Big]^2 \Big)^{1/2}
\end{aligned} \tag{2.5.3}$$

by Cauchy–Schwarz in the last step.

Noting that self-adjoint matrices A, B satisfy

$$\|AB\|_{op} = \|(AB)^*\|_{op} = \|B^*A^*\|_{op} = \|BA\|_{op},$$

the last two quadratic sums have exactly the same form. In the next section we prove the following bound:

2.5.4 Lemma. *For any $W \in A_2$ and any γ-sparse collection $\mathscr{S}$ of dyadic cubes,*

$$\Big(\sum_{S\in\mathscr{S}} \frac{1}{|S|} \Big[\int_S \|W(x)^{1/2}\langle W\rangle_S^{-1/2}\|_{op}\psi(x)\,\mathrm{d}x \Big]^2 \Big)^{1/2} \leq c_{d,n,\gamma}[W]_{A_2}^{1/2}\|\psi\|_{L^2(\mathbb{R}^d)}.$$

Taking this for granted for the moment, and applying this to the two quadratic sums on the right of (2.5.3), both as written and with Σ and ϕ in place of W and ψ, we deduce that

$$\int \psi(x) T_{\mathscr{S}}^{W,\Sigma}\phi(x)\,\mathrm{d}x \leq c_{d,n}[W]_{A_2}^{1/2} \cdot [W]_{A_2}^{1/2} \cdot [\Sigma]_{A_2}^{1/2} = c_{d,n}[W]_{A_2}^{3/2},$$

using $[\Sigma]_{A_2} = [W]_{A_2}$ in the last step. (We also fix the sparseness parameter γ.)

In view of Lemma 2.5.2 and the reductions in the beginning of the section, this implies:

2.5.5 Theorem (Nazarov, Petermichl, Treil, Volberg 2017 [NPTV17])**.** *For any Calderón–Zygmund operator T, or more generally any operator that satisfies the assumptions of Lerner's Theorem 1.4.2, we have*

$$\|T\|_{L^2(W)\to L^2(W)} \leq c_{d,n} c_T [W]_{A_2}^{3/2}$$

for any matrix-weight $W \in A_2$.

2.5.6 Remark. This is the best available bound at the time of writing, but it is not known whether it is optimal. Rather, it is conjectured that a linear bound $c_{d,n} c_T [W]_{A_2}$, as in the scalar-valued case, should remain valid even for matrix weights W.

A qualitative form of the theorem, that $T : L^2(W) \to L^2(W)$ is bounded for $W \in A_2$, is essentially contained in the paper of Treil and Volberg [TV97], who explicitly considered the Hilbert transform, but pointed out the applicability of the argument to more general Calderón–Zygmund operators. A different proof giving the slightly weaker quantitative bound than Theorem 2.5.5,

$$\|T\|_{L^2(W)\to L^2(W)} \leq c_{d,n} c_T [W]_{A_2}^{3/2} (1 + \log[W]_{A_2}),$$

is due to Bickel, Petermichl and Wick 2014 [BPW16].

The following exercises introduce some tools that will be needed in the proof of Lemma 2.5.4:

2.5.7 Exercise. The *trace* of a matrix A is defined as

$$\operatorname{tr} A := \sum_{i=1}^{n} (Ae_i | e_i),$$

where $(e_i)_{i=1}^n$ is any orthonormal basis. Show that this is well-defined, i.e., the result is independent of the chosen orthonormal basis. If G is a matrix-valued function, check that $\int_S \operatorname{tr} G(x)\,\mathrm{d}x = \operatorname{tr} \int_S G(x)\,\mathrm{d}x$.

2.5.8 Exercise. Let A be a positive self-adjoint matrix. Show that $\|A\|_{op} \leq \operatorname{tr} A \leq n\|A\|_{op}$. Hint: Write $\|A\|_{op}$ and $\operatorname{tr} A$ in terms of the eigenvalues of A.

2.5.9 Exercise. Prove that

$$\fint_S \|W(x)^{1/2} \langle W \rangle_S^{-1/2}\|_{op}^2 \,\mathrm{d}x \leq c_n.$$

Hint: Recall that $\|A\|_{op}^2 = \|A^* A\|_{op}$ and use the previous exercises.

2.5.10 Exercise. Suppose that $\mathscr{S}$ is *disjoint* (a stronger condition than sparse!) and W is a general matrix weight (not necessarily in A_2). Show that in this case Lemma 2.5.4 holds with just a dimensional constant in place of $c_{d,n}[W]_{A_2}^{1/2}$.

2.5.11 Exercise. Let $W \in A_2$ be a $\mathscr{L}(\mathbb{R}^n)$-valued matrix weight, and $x \in \mathbb{R}^n$ a non-zero vector. Prove that $(Wx|x) \in A_2$ is a scalar-valued A_2 weight and

$$[(Wx|x)]_{A_2} \leq [W]_{A_2}.$$

Hint: Consider functions of the form $f(t) = \phi(t)x$ in Proposition 2.2.1, where x is the given vector and ϕ is a scalar-valued function.

2.5.12 Exercise. Let $w \in A_2$ and $\sigma = w^{-1}$ be scalar-valued weights. Prove the linear bound $\|T^{w,\sigma}_{\mathscr{S}}\|_{L^2 \to L^2} \leq c_n [w]_{A_2}$ in this case. Hint: With suitable changes of variables, this can be reduced to Theorem 1.2.2. Why does your argument not work for matrix weights? (Or, if it does, you have proved the matrix A_2 conjecture!)

2.6 Sparse bounds with matrix A_∞ weights

In order to complete the proof of Theorem 2.5.5, we still need to prove Lemma 2.5.4. The proof will exploit the following notion of matrix A_∞ weights.

2.6.1 Definition. For a matrix weight $W : \mathbb{R}^d \to \mathscr{L}(\mathbb{R}^n)$, we set

$$[W]^{\mathscr{D}}_{A_\infty} := \sup_{e \in \mathbb{R}^n \setminus \{0\}} [(We|e)]^{\mathscr{D}}_{A_\infty},$$

and we say that $W \in A^{\mathscr{D}}_\infty$ if this number is finite.

2.6.2 Lemma. *We have $A^{\mathscr{D}}_2 \subset A^{\mathscr{D}}_\infty$ and $[W]^{\mathscr{D}}_{A_\infty} \leq e[W]^{\mathscr{D}}_{A_\infty}$.*

Proof. This is immediate by

$$[(We|e)]^{\mathscr{D}}_{A_\infty} \leq e[(We|e)]^{\mathscr{D}}_{A_2} \leq e[W]^{\mathscr{D}}_{A_2},$$

where the first step is the embedding $A_2 \subset A_\infty$ for scalar weights from Proposition 1.9.2, and the second step is Exercise 2.5.11. (The dyadic and the non-dyadic versions are proved in exactly the same way.) □

With Lemma 2.6.2 at hand, Lemma 2.5.4 will be a consequence of the following slightly sharper statement:

2.6.3 Proposition. *For any γ-sparse collection $\mathscr{S} \subset \mathscr{D}$ and any $W \in A^{\mathscr{D}}_\infty$, we have*

$$\Big(\sum_{S \in \mathscr{S}} |S| \Big[\fint_S \|W(x)^{1/2} \langle W \rangle_S^{-1/2}\|_{op} \psi(x) \, \mathrm{d}x \Big]^2 \Big)^{1/2} \leq c_{d,n,\gamma} [W]^{1/2}_{A^{\mathscr{D}}_\infty} \|\psi\|_{L^2(\mathbb{R}^d)}.$$

Note that we have also rewritten the left side of the estimate in a different but obviously equivalent way compared to Lemma 2.5.4. Using Proposition 2.6.3 instead of Lemma 2.5.4 in the proof of Theorem 2.5.5, we also deduce the following sharper version:

2.6.4 Corollary (Nazarov, Petermichl, Treil, Volberg 2017 [NPTV17])**.** *For any Calderón–Zygmund operator T, or more generally any operator that satisfies the assumptions of Lerner's Theorem 1.4.2, we have*

$$\|T\|_{L^2(W)\to L^2(W)} \leq c_{d,n} c_T [W]_{A_2}^{1/2} [W]_{A_\infty}^{1/2} [W^{-1}]_{A_\infty}^{1/2}$$

for any matrix-weight $W \in A_2$, where $[W]_{A_\infty} := \sup_{\mathscr{D}} [W]_{A_\infty}^{\mathscr{D}}$.

Proof of Proposition 2.6.3. Let us first consider the function

$$\begin{aligned}
\|W(x)^{1/2}\langle W\rangle_S^{-1/2}\|_{op}^2 &= \|\langle W\rangle_S^{-1/2} W(x) \langle W\rangle_S^{-1/2}\|_{op} \\
&\leq \operatorname{tr}(\langle W\rangle_S^{-1/2} W(x) \langle W\rangle_S^{-1/2}) \qquad \text{by Exercise 2.5.8} \\
&= \sum_{i=1}^n (\langle W\rangle_S^{-1/2} W(x) \langle W\rangle_S^{-1/2} e_i | e_i) \\
&= \sum_{i=1}^n (W(x)\langle W\rangle_S^{-1/2} e_i | \langle W\rangle_S^{-1/2} e_i) =: \sum_{i=1}^n w_{S,i}(x),
\end{aligned}$$

where $(e_i)_{i=1}^n$ is any fixed orthonormal basis of $\mathbb{R}^n$. By definition of matrix A_∞, we have $[w_{S,i}]_{A_\infty}^{\mathscr{D}} \leq [W]_{A_\infty}^{\mathscr{D}}$.

By the subadditivity of the maximal function and the definition of scalar A_∞, we have

$$\begin{aligned}
\int_Q M_Q\Big(\sum_{i=1}^n w_{S,i}\Big) &\leq \sum_{i=1}^n \int_Q M_Q(w_{S,i}) \\
&\leq \sum_{i=1}^n [w_{S,i}]_{A_\infty}^{\mathscr{D}} \int_Q w_{S,i} \leq [W]_{A_\infty}^{\mathscr{D}} \int_Q \sum_{i=1}^n w_{S,i},
\end{aligned}$$

and hence

$$[w_S]_{A_\infty}^{\mathscr{D}} := \Big[\sum_{i=1}^n w_{S,i}\Big]_{A_\infty^{\mathscr{D}}} \leq [W]_{A_\infty}^{\mathscr{D}}.$$

By Theorem 1.9.4, w_S satisfies the reverse Hölder inequality

$$\Big(\fint_S w_S^{1+\delta}\Big)^{1/(1+\delta)} \leq 2 \fint_S w_S, \qquad \delta := \frac{\varepsilon_d}{[W]_{A_\infty}^{\mathscr{D}}}. \tag{2.6.5}$$

Using Hölder's inequality with exponents $r = 2(1+\delta) > 2$ and $r' < 2$, we can now estimate

$$\begin{aligned}
\Big(\fint_S \|W(x)^{1/2}\langle W\rangle_S^{-1/2}\|_{op}\psi(x)\,\mathrm{d}x\Big)^2 &\leq \Big(\fint_S w_S^{1/2}\psi\Big)^2 \\
&\leq \Big(\fint_S w_S^{r/2}\Big)^{2/r} \Big(\fint_S \psi^{r'}\Big)^{2/r'},
\end{aligned}$$

where, by (2.6.5)

$$\Big(\fint_S w_S^{r/2}\Big)^{2/r} = \Big(\fint_S w_S^{1+\delta}\Big)^{1/(1+\delta)} \leq 2 \fint_S w_S,$$

and moreover, substituting the definition of w_S and using the linearity of the trace,

$$\begin{aligned}\fint_S w_S &= \fint_S \operatorname{tr}(\langle W\rangle_S^{-1/2} W(x) \langle W\rangle_S^{-1/2}) \, \mathrm{d}x \\ &= \operatorname{tr}\Big(\fint_S \langle W\rangle_S^{-1/2} W(x) \langle W\rangle_S^{-1/2} \, \mathrm{d}x\Big) \\ &= \operatorname{tr}\Big(\langle W\rangle_S^{-1/2} \fint_S W(x) \, \mathrm{d}x \langle W\rangle_S^{-1/2}\Big) \\ &= \operatorname{tr}\Big(\langle W\rangle_S^{-1/2} \langle W\rangle_S \langle W\rangle_S^{-1/2}\Big) = \operatorname{tr}(I) = n,\end{aligned}$$

where I is the $n \times n$ identity matrix.

Combining the previous estimates, we have checked that

$$\Big(\fint_S \|W(x)^{1/2} \langle W\rangle_S^{-1/2}\|_{op} \psi(x) \, \mathrm{d}x\Big)^2 \leq 2n \Big(\fint_S \psi^{r'}\Big)^{2/r'} \leq 2n \inf_S (M_{\mathscr{D}} \psi^{r'})^{2/r'},$$

and hence, using also the sparseness,

$$\begin{aligned}&\Big(\sum_{S \in \mathscr{S}} |S| \Big[\fint_S \|W(x)^{1/2} \langle W\rangle_S^{-1/2}\|_{op} \psi(x) \, \mathrm{d}x\Big]^2\Big)^{1/2} \\ &\leq \Big(\sum_{S \in \mathscr{S}} \frac{|E(S)|}{\gamma} 2n \inf_S (M_{\mathscr{D}} \psi^{r'})^{2/r'}\Big)^{1/2} \leq c_{n,\gamma} \Big(\int_{\mathbb{R}^d} (M_{\mathscr{D}} \psi^{r'})^{2/r'}\Big)^{1/2}.\end{aligned}$$

Since $r' < 2$, the maximal operator is bounded on $L^{2/r'}$ with norm $(2/r')'$, so that

$$\Big(\int_{\mathbb{R}^d} (M_{\mathscr{D}} \psi^{r'})^{2/r'}\Big)^{1/2} \leq \Big(\big(\frac{2}{r'}\big)'\Big)^{1/r'} \Big(\int_{\mathbb{R}^d} (\psi^{r'})^{2/r'}\Big)^{1/2},$$

and the last factor is the desired $\|\psi\|_{L^2}$.

It only remains to see the dependence of the factor $((2/r')')^{1/r'}$ on $[W]_{A_\infty}^{\mathscr{D}}$. Recall that $r = 2(1+\delta)$, hence $r' = r/(r-1) = 2(1+\delta)/(1+2\delta)$ and $2/r' = (1+2\delta)/(1+\delta) = 1 + \delta/(1+\delta) \leq 1+\delta$. Then $1/r' \leq (1+\delta)/2$ and $(2/r')' = 1 + (1+\delta)/\delta = 2 + 1/\delta \leq 3/\delta$. Thus

$$\Big(\big(\frac{2}{r'}\big)'\Big)^{1/r'} \leq \Big(\frac{3}{\delta}\Big)^{(1+\delta)/2} \leq 3\delta^{-1/2} \delta^{-\delta/2} = 3\delta^{-1/2} e^{\delta \log(1/\delta)/2} \leq c\delta^{-1/2},$$

since the function $\delta \log(1/\delta)$ is continuous on $(0,1]$ and has a finite limit as $\delta \to 0$. Recalling that $\delta = \varepsilon_d / [W]_{A_\infty}^{\mathscr{D}}$, we conclude that

$$\delta^{-1/2} \leq c_d [W]_{A_\infty^{\mathscr{D}}}^{1/2},$$

completing the proof. □

2.6.6 Remark. While we have reproduced Proposition 2.6.3 from [NPTV17], a closely related statement and proof already appears in Isralowitz, Kwon and

Pott 2015 [IKP15]. These authors also prove a predecessor of Theorem 2.5.5 for the usual sparse operators $T_{\mathscr{S}} f = \sum_{S\in\mathscr{S}} 1_S \fint_S f$; however, before the Convex Body Domination Theorem 2.3.2, its relevance to Calderón–Zygmund operators was unclear.

2.6.7 Remark. While the sharp matrix-weighted estimate for Calderón–Zygmund operators (even individual special cases like the Hilbert transform, or its dyadic models) remains open, the full matrix analogues of the sharp A_2 theorems from the scalar case have been achieved for the maximal operator M and the *dyadic square function*

$$Sf(x) := \Big(\sum_{I\in\mathscr{D}} \frac{1_I(x)}{|I|}|\langle h_I, f\rangle|^2\Big)^{1/2},$$

where $h_I := |I|^{-1/2}(1_{I_\ell} - 1_{I_r})$ (with $I_{\ell/r}$ the left/right half of I) are the L^2-normalised Haar functions. Since M and S are nonlinear operators, their extension to the matrix-valued setting is not canonical, but requires a choice of an interpretation. The following definitions incorporate the weight into the operator itself:

$$M_W f(x) := \sup_{Q\ni x} \fint_Q |W(x)^{1/2}W(y)^{-1/2}f(y)|\,\mathrm{d}y, \tag{2.6.8}$$

$$S_W f(x) := \Big(\sum_{I\in\mathscr{D}} \frac{1_I(x)}{|I|}|\langle W\rangle_I^{1/2}\langle h_I, f\rangle|^2\Big)^{1/2}, \tag{2.6.9}$$

where definition (2.6.8) was introduced by Christ and Goldberg [CG01] and definition (2.6.9) by Petermichl and Pott [PP03].

Both these operators satisfy the linear bounds

$$\|M_W\|_{L^2\to L^2} \le c_{d,n}[W]_{A_2}, \tag{2.6.10}$$

$$\|S_W\|_{L^2(W)\to L^2} \le c_n[W]_{A_2}, \tag{2.6.11}$$

where (2.6.10) was proved by Isralowitz, Kwon and Pott in 2015 [IKP15] and (2.6.11) by Hytönen, Petermichl and Volberg in February 2017 [HPV17].

For scalar weights, it is easy to check that $\|M_w\|_{L^2\to L^2} = \|M\|_{L^2(w)\to L^2(w)}$ and $\|S_w\|_{L^2(w)\to L^2} = \|S\|_{L^2(w)\to L^2(w)}$, so that (2.6.10) and (2.6.11) correspond to the scalar-weighted results of Buckley [Buc93] for M and Hukovic, Treil and Volberg [HTV00] for S. These are already known to be sharp, and hence so are the matrix-weighted extensions (2.6.10) and (2.6.11).

The proofs of both (2.6.10) and (2.6.11) make use of Proposition 2.6.3, or essentially equivalent considerations.

Bibliography

[BPW16] Kelly Bickel, Stefanie Petermichl, and Brett D. Wick. Bounds for the Hilbert transform with matrix A_2 weights. *J. Funct. Anal.*, 270(5):1719–1743, 2016. arXiv:1402.3886.

[Buc93] Stephen M. Buckley. Estimates for operator norms on weighted spaces and reverse Jensen inequalities. *Trans. Amer. Math. Soc.*, 340(1):253–272, 1993.

[CAR16] José M. Conde-Alonso and Guillermo Rey. A pointwise estimate for positive dyadic shifts and some applications. *Math. Ann.*, 365(3-4):1111–1135, 2016. arXiv:1409.4351.

[CF74] R. R. Coifman and C. Fefferman. Weighted norm inequalities for maximal functions and singular integrals. *Studia Math.*, 51:241–250, 1974.

[CG01] Michael Christ and Michael Goldberg. Vector A_2 weights and a Hardy-Littlewood maximal function. *Trans. Amer. Math. Soc.*, 353(5):1995–2002, 2001.

[CJRW03] James T. Campbell, Roger L. Jones, Karin Reinhold, and Máté Wierdl. Oscillation and variation for singular integrals in higher dimensions. *Trans. Amer. Math. Soc.*, 355(5):2115–2137, 2003.

[CUMP10] 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.

[FP97] R. Fefferman and J. Pipher. Multiparameter operators and sharp weighted inequalities. *Amer. J. Math.*, 119(2):337–369, 1997.

[FZ16] F. C. de França Silva and P. Zorin-Kranich. Sparse domination of sharp variational truncations. *ArXiv e-prints*, April 2016.

[HLM+12] Tuomas P. Hytönen, Michael T. Lacey, Henri Martikainen, Tuomas Orponen, Maria Carmen Reguera, Eric T. Sawyer, and Ignacio Uriarte-Tuero. Weak and strong type estimates for maximal truncations of Calderón-Zygmund operators on A_p weighted spaces. *J. Anal. Math.*, 118(1):177–220, 2012. arXiv:1103.5229.

[How97] Ralph Howard. The John ellipsoid theorem. Seminar notes, University of South Carolina, http://people.math.sc.edu/howard/Notes/john.pdf, 1997.

[HPR12] Tuomas Hytönen, Carlos Pérez, and Ezequiel Rela. Sharp reverse Hölder property for A_∞ weights on spaces of homogeneous type. *J. Funct. Anal.*, 263(12):3883–3899, 2012.

[HPV17] T. Hytönen, S. Petermichl, and A. Volberg. The sharp square function estimate with matrix weight. *ArXiv e-prints*, February 2017.

[HTV00] S. Hukovic, S. Treil, and A. Volberg. The Bellman functions and sharp weighted inequalities for square functions. In *Complex analysis, operators, and related topics*, volume 113 of *Oper. Theory Adv. Appl.*, pages 97–113. Birkhäuser, Basel, 2000.

[Hyt12] Tuomas P. Hytönen. The sharp weighted bound for general Calderón-Zygmund operators. *Ann. of Math. (2)*, 175(3):1473–1506, 2012. arXiv:1007.4330.

[IKP15] J. Isralowitz, H.-K. Kwon, and S. Pott. Matrix weighted norm inequalities for commutators and paraproducts with matrix symbols. *ArXiv e-prints*, July 2015.

[Joh48] Fritz John. Extremum problems with inequalities as subsidiary conditions. In *Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, 1948*, pages 187–204. Interscience Publishers, Inc., New York, N. Y., 1948.

[Lac17] Michael T. Lacey. An elementary proof of the A_2 bound. *Israel J. Math.*, 217(1):181–195, 2017. arXiv:1501.05818.

[Ler13] Andrei K. Lerner. A simple proof of the A_2 conjecture. *Int. Math. Res. Not. IMRN*, (14):3159–3170, 2013. arXiv:1202.2824.

[Ler16] Andrei K. Lerner. On pointwise estimates involving sparse operators. *New York J. Math.*, 22:341–349, 2016. arXiv:1512.07247.

[LN15] A. K. Lerner and F. Nazarov. Intuitive dyadic calculus: the basics. *ArXiv e-prints*, August 2015.

[LPR15] Teresa Luque, Carlos Pérez, and Ezequiel Rela. Optimal exponents in weighted estimates without examples. *Math. Res. Lett.*, 22(1):183–201, 2015. arXiv:1307.5642.

[MST15] M. Mirek, E. M. Stein, and B. Trojan. $L^p(\mathbb{Z}^d)$-estimates for discrete operators of Radon type: Variational estimates. *ArXiv e-prints*, December 2015.

[NPTV17] F. Nazarov, S. Petermichl, S. Treil, and A. Volberg. Convex body domination and weighted estimates with matrix weights. *ArXiv e-prints*, January 2017.

[Pet07] 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.

[Pis86] Gilles Pisier. *Factorization of linear operators and geometry of Banach spaces*, volume 60 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1986.

[PP03] S. Petermichl and S. Pott. A version of Burkholder's theorem for operator-weighted spaces. *Proc. Amer. Math. Soc.*, 131(11):3457–3461, 2003.

[TV97] S. Treil and A. Volberg. Wavelets and the angle between past and future. *J. Funct. Anal.*, 143(2):269–308, 1997.