DYADIC ANALYSIS AND WEIGHTS

TUOMAS HYTÖNEN

Lecture notes of a course at the University of Helsinki, Spring 2014.

1. Dyadic cubes and the dyadic maximal operator

1.A. **Dyadic cubes: basics.** The standard dyadic cubes of $\mathbb{R}^d$ consist of the collection

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

It is the union over $k \in \mathbb{Z}$ of the collections

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

The main properties of dyadic cubes are:

- Each $\mathscr{D}_k$ is a *partition* (i.e., a pairwise disjoint cover) of $\mathbb{R}^d$.
- Each $\mathscr{D}_{k+1}$ is a *refinement* of the previous $\mathscr{D}_k$: i.e., we have
$$Q = \bigcup_{\substack{Q' \in \mathscr{D}_{k+1} \\ Q' \subseteq Q}} Q$$
for every $Q \in \mathscr{D}_k$.

We denote by $\ell(Q)$ the side-length of a cube, and by $|Q| = \ell(Q)^d$ its Lebesgue measure.

Definition 1.1. A collection of sets $\mathscr{Q}$ is said to

(1) be *nested* if
$$Q \cap R \in \{\varnothing, Q, R\} \qquad \text{for all } Q, R \in \mathscr{Q}, \tag{1.2}$$
(2) have an *infinite increasing chain* if there exist $Q_k \in \mathscr{Q}$, $k \in \mathbb{N}$, such that $Q_k \subsetneq Q_{k+1}$ for all $k \in \mathbb{N}$.

An element $Q \in \mathscr{Q}$ is said to be *maximal* if there does not exist any $R \in \mathscr{Q}$ with $R \supsetneq Q$. The collection of maximal elements in $\mathscr{Q}$ is denoted by $\mathscr{Q}^*$.

It is easy to check that any subcollection $\mathscr{Q} \subseteq \mathscr{D}$ of dyadic cubes is nested. On the other hand, this property fails for almost any other collection of sets. (Say, the intersection of balls is often not a ball at all, not to mention one of the original balls.)

The collection $\mathscr{D}$ has infinite increasing chains, for instance $Q_k := 2^k[0,1)^d$, $k \in \mathbb{Z}$, and it does not have any maximal elements (since every dyadic cube is always contained in a strictly bigger dyadic cube). On the other hand, if E is any bounded set, then $\mathscr{Q} = \{Q \in \mathscr{D} : Q \subseteq E\}$ does not have infinite increasing chains, since $\ell(Q_k) \geq 2\ell(Q_{k-1}) \geq \ldots \geq 2^k \ell(Q_0)$ in such a chain, and eventually the side-length would be too big to fit inside the bounded set E.

Lemma 1.3. *Let $\mathscr{Q}$ be a collection of sets.*

(1) *If $\mathscr{Q}$ does not have infinite increasing chains, then every $Q \in \mathscr{Q}$ is contained in a maximal $Q^* \in \mathscr{Q}^*$.*
(2) *If $\mathscr{Q}$ is nested, then any two $Q, R \in \mathscr{Q}^*$ are disjoint.*

Proof. (1): Let $Q_0 \in \mathscr{Q}$ be given; we need to find some maximal Q^* containing Q_0. If Q_0 itself is maximal, we take $Q^* = Q_0$. Otherwise, by definition, there exists some $Q_1 \in \mathscr{Q}$ with $Q_1 \supsetneq Q_0$. Suppose that we have already found $Q_k \supsetneq Q_{k-1} \supsetneq \ldots \supsetneq Q_1 \supsetneq Q_0$. If Q_k is maximal, then we are done with $Q^* := Q_k$. Otherwise, there exists some $Q_{k+1} \in \mathscr{Q}$ with $Q_{k+1} \supsetneq Q_k$. This process

Version: May 2, 2014.

must terminate after finitely many steps with some Q_n being maximal, since otherwise we could construct an infinite increasing chain, a contradiction.

(2): We understand of course that we mean two *different* $Q, R \in \mathscr{Q}^*$: $Q \neq R$. Since $\mathscr{Q}$ is nested, we have $Q \cap R \in \{\varnothing, Q, R\}$. If they are not disjoint, i.e., $Q \cap R \neq \varnothing$, this only leaves the options $Q \cap R \in \{Q, R\}$. Suppose for instance that $Q \cap R = Q$, the other case being symmetric. This means that $Q \subseteq R$, and since $Q \neq R$, that $Q \subsetneq R$. But this contradicts with the maximality of Q. Thus disjointness is the only possibility. $\square$

1.B. **The dyadic maximal operator.** The dyadic maximal operator M_d is defined by

$$M_d f(x) := \sup_{\substack{Q \in \mathscr{D} \\ Q \ni x}} \frac{1}{|Q|} \int_Q |f| \, \mathrm{d}y, \tag{1.4}$$

where the supremum is over all dyadic cubes that contain the point $x \in \mathbb{R}^d$. It satisfies the following fundamental estimate, whose proof is a good illustration of the basic properties of dyadic cubes:

Proposition 1.5. *For every $f \in L^1(\mathbb{R}^d)$ and $\lambda > 0$, we have*

$$|\{x \in \mathbb{R}^d : M_d f(x) > \lambda\}| \leq \frac{1}{\lambda} \|f\|_{L^1} := \frac{1}{\lambda} \int_{\mathbb{R}^d} |f| \, \mathrm{d}y.$$

The set appearing in the estimate is a particular case of a *level set* of a function. We will often use the shorter notation

$$\{g > \lambda\} := \{x \in \mathbb{R}^d : g(x) > \lambda\}.$$

Proof. Let

$$\mathscr{Q}_\lambda := \{Q \in \mathscr{D} : \frac{1}{|Q|} \int_Q |f| \, \mathrm{d}y > \lambda\}.$$

From the definition of the dyadic maximal operator, we have $M_d f(x) > \lambda$ if and only if there exists a $Q \in \mathscr{Q}_\lambda$ with $Q \ni x$. Thus

$$\{M_d f > \lambda\} = \bigcup_{Q \in \mathscr{Q}_\lambda} Q. \tag{1.6}$$

Since $\mathscr{Q}_\lambda \subseteq \mathscr{D}$, it is nested. Let us check that it does not have infinite increasing chains. This follows by observing that for every $Q \in \mathscr{Q}_\lambda$, we have

$$|Q| \leq \frac{1}{\lambda} \int_Q |f| \, \mathrm{d}y \leq \frac{1}{\lambda} \|f\|_1, \tag{1.7}$$

so that the measure of the cubes $Q \in \mathscr{Q}_\lambda$ is bounded from above. This is not the case for dyadic cubes in an infinite increasing chain, since $|Q_{k+1}| \geq 2^d |Q_k|$ for $Q_{k+1} \supsetneq Q_k$. We conclude from Lemma 1.3(1) that

$$\bigcup_{Q \in \mathscr{Q}_\lambda} Q = \bigcup_{Q^* \in \mathscr{Q}_\lambda^*} Q^* : \tag{1.8}$$

here $\supseteq$ is clear since $\mathscr{Q}_\lambda \supseteq \mathscr{Q}_\lambda^*$, and $\subseteq$ follows from Lemma 1.3(1), since every Q on the left is contained in some Q^* on the right by the lemma. The union on the right of (1.8) is disjoint by Lemma 1.3(2).

Now we just put the pieces together:

$$\begin{aligned} |\{M_d f > \lambda\}| &= \Big| \bigcup_{Q \in \mathscr{Q}_\lambda} Q \Big| = \Big| \bigcup_{Q \in \mathscr{Q}_\lambda^*} Q \Big| \leq \sum_{Q \in \mathscr{Q}_\lambda^*} |Q| \leq \frac{1}{\lambda} \sum_{Q \in \mathscr{Q}_\lambda^*} \int_Q |f| \, \mathrm{d}y \\ &= \frac{1}{\lambda} \int_{\mathbb{R}^d} \Big(\sum_{Q \in \mathscr{Q}_\lambda^*} 1_Q(y) \Big) |f(y)| \, \mathrm{d}y \overset{(\star)}{\leq} \frac{1}{\lambda} \int_{\{M_d f > \lambda\}} |f(y)| \, \mathrm{d}y \leq \frac{1}{\lambda} \|f\|_{L^1}. \end{aligned}$$

Here $(\star)$ was based on the fact that the cubes $Q \in \mathscr{Q}_\lambda$ are pairwise disjoint, so at most one Q contains any given $y \in \mathbb{R}^d$; on the other hand, if $y \in \{M_d f > \lambda\} = \bigcup_{Q \in \mathscr{Q}_\lambda^*} Q$, then there exists exactly one such cube. Note that the second-to-last estimate above gave a slightly more precise bound than claimed in the proposition. $\square$

1.C. **The dyadic maximal operator for general measures.** The definition (1.4) immediately generalizes to other measures μ in place of the Lebesgue measure: we just replace every occurrence of the Lebesgue measure, both in integrations and in taking the measures of sets, by the measure μ. For the expression to be meaningful, we only ask that μ is a locally finite Borel measure (so that cubes are measurable, and $\mu(Q) < \infty$). Thus we define

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

This also satisfies the analogue of Proposition 1.5, again replacing every occurrence of the Lebesgue measure by μ:

Proposition 1.9. *For every $f \in L^1(\mu)$ and $\lambda > 0$, we have*

$$\mu(\{M_d^\mu f > \lambda\}) \leq \frac{1}{\lambda}\|f\|_{L^1(\mu)} := \frac{1}{\lambda}\int_{\mathbb{R}^d} |f| \,\mathrm{d}\mu.$$

Proof. An inspection of the proof of Proposition 1.5 shows that the same argument goes through verbatim, except for one point: in arguing that $\mathscr{Q}_\lambda$ does not have infinite increasing chains, we used a specific property of the Lebesgue measure that $|Q| \geq 2^d|Q'|$ for dyadic cubes $Q \supsetneq Q'$. This might fails for other measures, so we need to modify the argument. The problem is not serious however, and can be fixed by a simple localization.

Let $\mathscr{D}(n) := \{Q \in \mathscr{D} : \ell(Q) \leq 2^n\}$, and define $M_{d,n}^\mu f$ by replacing $\mathscr{D}$ in the definition of $M_d^\mu f$ by $\mathscr{D}(n)$. It is easy to check that $M_d^\mu f(x)$ is the monotone increasing limit of $M_{d,n}^\mu f(x)$ as $n \to \infty$. Now

$$\{M_{d,n}^\mu f > \lambda\} = \bigcup_{Q\in\mathscr{Q}_{\lambda,n}} Q, \quad \text{where} \quad \mathscr{Q}_{\lambda,n} := \{Q \in \mathscr{D}(n) : \frac{1}{\mu(Q)}\int_Q |f| \,\mathrm{d}\mu > \lambda\}$$

is nested and does not have infinite increasing chains (since the side-length in $\mathscr{D}(n)$ is bounded from above, but the side-length in an infinite increasing chain increases without limit). Thus a trivial modification of the existing argument shows that

$$\mu(\{M_{d,n}^\mu f > \lambda\}) \leq \frac{1}{\lambda}\int_{\{\{M_{d,n}^\mu f>\lambda\}} |f| \,\mathrm{d}\mu \leq \frac{1}{\lambda}\|f\|_{L^1(\mu)},$$

and the proposition follows by monotone convergence:

$$\mu(\{M_d^\mu f > \lambda\}) = \mu\Big(\bigcup_{n\in\mathbb{N}}\{M_{d,n}^\mu f > \lambda\}\Big) = \lim_{n\to\infty}\mu(\{M_{d,n}^\mu f > \lambda\}).$$ □

2. From dyadic to non-dyadic: shifted dyadic cubes

2.A. **Hardy–Littlewood maximal operators.** A general class of maximal operators can be defined as follows. Let $\mathfrak{Q} = \{\mathscr{Q}_x\}_{x\in\mathbb{R}^d}$ be a family of collections of sets, indexed by $x \in \mathbb{R}^d$. (I.e., for every $x \in \mathbb{R}^d$, $\mathscr{Q}_x$ is a collection of sets.) Then

$$M_{\mathfrak{Q}} f(x) := \sup_{Q\in\mathscr{Q}_x} \frac{1}{|Q|}\int_Q |f| \,\mathrm{d}y.$$

Clearly M_d is the special case with $\mathscr{Q}_x = \mathscr{D}_x := \{Q \in \mathscr{D} : Q \ni x\}$.

For two collections of sets $\mathscr{Q}$ and $\mathscr{R}$, we say that $\mathscr{R}$ *dominates* $\mathscr{Q}$ (with constant K) if

$$\forall\, Q \in \mathscr{Q} :\ \exists\, R \in \mathscr{R} : \begin{cases} R \supseteq Q, \text{ and} \\ |R| \leq K|Q|. \end{cases} \tag{2.1}$$

Exercise 2.2. Prove that if $\mathscr{R}_x$ dominates $\mathscr{Q}_x$, then $M_{\mathfrak{Q}} f(x) \leq K M_{\mathfrak{R}} f(x)$, where K is the domination constant.

In particular, if $\mathscr{R}_x$ dominates $\mathscr{Q}_x$, and $\mathscr{Q}_x$ dominates $\mathscr{R}_x$, then $M_{\mathfrak{Q}} f(x)$ and $M_{\mathfrak{R}} f(x)$ are comparable.

In many questions of Analysis, one of the following variants of the Hardy–Littlewood maximal operator plays a role:

(1) The *centred ball maximal operator* M_c corresponds to $\mathscr{Q}_x = \mathscr{B}_x^c := \{B(x,r) : r > 0\}$.
(2) The *non-centred ball maximal operator* M_b corresponds to $\mathscr{Q}_x = \mathscr{B}_x := \{B \text{ ball} : B \ni x\}$.
(3) The centred or non-centred cube maximal operators are defined in a similar way with cubes instead of balls.

Although they look different at first, these are actually all pointwise comparable. As an example, we check:

Lemma 2.3. *$M_c f(x) \leq M_b f(x) \leq 2^d M_c f(x)$.*

Proof. By Exercise 2.2, it suffices to check that $\mathscr{B}_x$ dominates $\mathscr{B}_x^c$ with constant 1, and $\mathscr{B}_x^c$ dominates $\mathscr{B}_x$ with constant 2^d. The first case is obvious, since $\mathscr{B}_x^c \subset \mathscr{B}_x$ (balls centred at x are special cases of balls containing x!). For the other case, let $B(z,r) \in \mathscr{B}_x$ be any ball containing x. We claim that $B(z,r) \subseteq B(x,2r) \in \mathscr{B}_x^c$; indeed, if $y \in B(z,r)$, then

$$|y-x| \leq |y-z| + |z-x| < r + r = 2r,$$

where we used $y \in B(z,r)$ and $z \in B(x,r)$. Thus $y \in B(x,2r)$, and since $\in B(z,r)$ was arbitrary, this shows that $B(z,r) \subseteq B(x,2r)$. On the other hand, we have

$$|B(x,2r)| = v_d(2r)^d = 2^d \cdot v_d r^d = 2^d |B(z,r)|,$$

where $v_d = |B(0,1)|$ is the measure of the d-dimensional unit ball. Thus we find that $\mathscr{B}_x^c$ dominates $\mathscr{B}_x$ with constant 2^d, as claimed. □

It is also easy to check that $\mathscr{B}_x$ dominates $\mathscr{D}_x$ with a dimensional constant, and thus $M_d f(x) \leq c_d M_b f(x) \leq 2^d c_d M_c f(x)$. However, the opposite estimate fails: It is easy to check that if f is nonzero function (more precisely: nonzero in a set of positive measure) then $M_c f(x) > 0$ at every point $x \in \mathbb{R}^d$. However, if f is supported in one of the "quadrants", say in the positive quadrant $[0,\infty)^d$, then so is $M_d f$. Thus, if f is nonzero and supported in $[0,\infty)^d$, and $x \in \mathbb{R}^d \setminus [0,\infty)^d$, then $M_d f(x) = 0 < M_c f(x)$. Hence $M_c f(x) \leq K M_d f(x)$ cannot be true with any constant K.

This seems to indicate that the dyadic maximal operator has no use in trying to estimate the Hardy–Littlewood maximal operator. However, the situation changes as soon as we consider *several dyadic systems*!

2.B. **Shifted dyadic systems.** There is a general constructing shifted dyadic systems. Let $\omega = (\omega_k)_{k\in\mathbb{Z}}$ be a sequence of binary vectors $\omega_k \in \{0,1\}^d$. For a dyadic cube $Q \in \mathscr{D}$, we define its shift by ω as a shift by a truncation of the formal (usually divergent) binary expansion $\sum_{k\in\mathbb{Z}} \omega_k 2^{-k}$:

$$Q \dot{+} \omega := Q + \omega(\ell(Q)) := Q + \sum_{j:2^{-j}<\ell(Q)} \omega_j 2^{-j}.$$

We define $\mathscr{D}_k^\omega := \{Q \dot{+} \omega : Q \in \mathscr{D}_k\}$ and $\mathscr{D}^\omega := \bigcup_{k\in\mathbb{Z}} \mathscr{D}_k^\omega$.

Lemma 2.4. *For every $\omega \in (\{0,1\}^d)^{\mathbb{Z}}$, the system $\mathscr{D}^\omega$ is another dyadic system in the sense that:*
(1) *Each $\mathscr{D}_k^\omega$ is a partition of $\mathbb{R}^d$.*
(2) *Each $\mathscr{D}_{k+1}^\omega$ is a refinement of the previous $\mathscr{D}_k^\omega$.*

Proof. (1) is obvious, since we just shift each element of the partition $\mathscr{D}_k$ by the same number $\omega(2^{-k})$.

(2) requires a little argument to check that the shifts by the different numbers $\omega(2^{-k})$ on $\mathscr{D}_k$, and $\omega(2^{-k-1})$ on $\mathscr{D}_{k+1}$, produce compatible partitions. The key is to observe that:

- $\omega(2^{-k}) = \omega(2^{-k-1}) + \omega_{k+1} 2^{-k-1}$, where $\omega_{k+1} \in \{0,1\}^d$, and
- we have

$$\begin{aligned}\mathscr{D}_{k+1} &= \{2^{-k-1}[0,1)^d + 2^{-k-1}m : m \in \mathbb{Z}^d\} \\ &= \{2^{-k-1}[0,1)^d + 2^{-k-1}(m + \omega_{k+1}) : m \in \mathbb{Z}^d\} = \mathscr{D}_{k+1} + 2^{-k-1}\omega_{k+1},\end{aligned}$$

  by simple re-parametrization, since $m + \omega_{k+1}$ also runs through $\mathbb{Z}^d$ when m does.

A combination of these points shows that

$$\mathscr{D}_{k+1} + \omega(2^{-k-1}) = \mathscr{D}_{k+1} + \omega_{k+1}2^{-k-1} + \omega(2^{-k-1}) = \mathscr{D}_{k+1} + \omega(2^{-k}).$$

This is obviously a refinement of $\mathscr{D}_k + \omega(2^{-k})$, since $\mathscr{D}_{k+1}$ is a refinement of $\mathscr{D}_k$. □

Exercise 2.5. Consider the dimension $d = 1$, and let $x \in \mathbb{R}$ be a fixed point. Consider the shifted system $\mathscr{D}^x := \{Q + x : Q \in \mathscr{D}\}$, where $Q + x$ is just a usual translation of the cube (interval) Q by x. Show that $\mathscr{D}^x$ is a special case of the shifted systems $\mathscr{D}^\omega$ for some $f\omega = \omega(x) \in \{0,1\}^{\mathbb{Z}}$, and compute this $\omega(x)$. Note that the formula should look slightly different for $x \geq 0$ and $x < 0$.

Exercise 2.6. From the result of the previous exercise, deduce a criterion for $\omega \in \{0,1\}^{\mathbb{Z}}$ to determine is it is of the form $\omega = \omega(x)$ for some $x \in \mathbb{R}^d$. Conclude that "most" $\omega \in \{0,1\}^{\mathbb{Z}}$ are not of this form.

2.C. **Special shifted systems,** $d = 1$. After the general definition, we concentrate for a while on dimension $d = 1$. It is obvious that $\omega_j \equiv 0$ gives the original dyadic system $\mathscr{D}^\omega = \mathscr{D}^0 = \mathscr{D}$. This also happens if $\omega_j \equiv 1$.

Our aim is now to construct a shifted system that, in some sense, is as far away from the original one as possible. Thus we choose ω_j where the values of 0 and 1 alternate, which gives two options: $\omega^1_j := 1_{2\mathbb{Z}}(j)$ and $\omega^2_j := 1_{2\mathbb{Z}+1}(j)$, i.e., ω^1_j is 1 for even j and 0 for odd j, and ω^2_j behaves in the opposite way. Let us abbreviate

$$\mathscr{D}^i := \mathscr{D}^{\omega^i}, \qquad i = 1, 2.$$

Lemma 2.7.

$$\begin{aligned}\omega^i(2^{-k}) &= \begin{cases} \frac{2}{3}\cdot 2^{-k}, & \textit{if } i,k \textit{ have same partity (both even or both odd)},\\ \frac{1}{3}\cdot 2^{-k}, & \textit{if } i,k \textit{ have different parity (one even and one odd)}.\end{cases}\\ &= \Big(1_{2\mathbb{Z}+1}(k) + (-1)^k\frac{i}{3}\Big)2^{-k}.\end{aligned}$$

Proof. By definition

$$\omega^i(2^{-k}) = \sum_{\substack{j>k\\ j-i \text{ odd}}} 2^{-j} = 2^{-k}\sum_{\substack{h>0\\ h-k-i \text{ odd}}} 2^{-h},$$

where the summation condition is h even, if $k - i$ is odd, and h odd, $k - i$ is even. It remains to observe that

$$\sum_{\substack{h>0\\ h \text{ even}}} 2^{-h} = \sum_{i=1}^{\infty} 2^{-2i} = \frac{1}{3}, \qquad \sum_{\substack{h>0\\ h \text{ odd}}} 2^{-h} = \sum_{i=1}^{\infty} 2^{-(2i-1)} = \frac{2}{3}. \qquad \square$$

This proves the first identity of the lemma, and the second one is easily checked by considering the four possibilities (k even or odd, $i = 1, 2$) case by case.

We have already pointed out before that $\mathscr{D}_k + 2^{-k}$ is simply $\mathscr{D}_k$. Thus it follows from the lemma that for $i = 1, 2$, we have

$$\mathscr{D}^i_k = \Big\{2^{-k}\Big([0,1) + m + (-1)^k\frac{i}{3}\Big) : m \in \mathbb{Z}\Big\}, \tag{2.8}$$

and this formula has the virtue of being obviously valid for $i = 0$ (with $\mathscr{D}^0 = \mathscr{D}$) as well. By Lemma 2.4, the new systems $\mathscr{D}^i$, $i = 1, 2$, share the key properties of the original $\mathscr{D}$. However, the true value of these shifted systems comes only when using them *simultaneously*:

Proposition 2.9. *Let $I \subset \mathbb{R}$ be any finite interval. Then for at least two (but not necessarily all three!) values of $i \in \{0, 1, 2\}$, there exists $J \in \mathscr{D}^i$ such that*

$$J \supseteq I, \qquad 3\ell(I) < \ell(J) \leq 6\ell(I).$$

Proof. Let k be the unique integer such that $3\ell(I) < 2^{-k} \leq 6\ell(I)$. (There is always a power of 2 in an interval $(t, 2t]$.) Let

$$\mathscr{E}_k^i := \{2^{-k}(m + (-1)^k i/3) : m \in \mathbb{Z}\}, \qquad i = 0, 1, 2,$$

be the set of end-points of the intervals $\mathscr{D}_k^i$. Then it is immediate that the sets $\mathscr{E}_k^i$, $i = 0, 1, 2$, are pairwise disjoint, and

$$\mathscr{E}_k := \mathscr{E}_k^0 \cup \mathscr{E}_k^1 \cup \mathscr{E}_k^2 = \{2^{-k}(m + i/3) : m \in \mathbb{Z}, i = 0, 1, 2\}$$

is a doubly infinite arithmetic sequence where the distance of consecutive points is $2^{-k}/3 > \ell(I)$. It follows that I contains at most one point of $\mathscr{E}_k$. By disjointness, there are at least two values $i \in \{0, 1, 2\}$ such that I does not contain any point of $\mathscr{E}_k^i$. Now, if $J \in \mathscr{D}_k^i$ (for either of these two values of i) is the unique interval that contains the centre of I, it must actually contain the whole of I. And we have $\ell(J) = 2^{-k} \in (3\ell(I), 6\ell(I)]$, as required. □

Let

$$\mathscr{I}_x := \{I \subset \mathbb{R} \text{ a finite interval} : I \ni x\},$$

and recall that

$$\mathscr{D}_x^i := \{J \in \mathscr{D}^i : J \ni x\}.$$

An immediate reformulation of Proposition 2.9 in the language of domination (recall (2.1)) is that $\mathscr{D}_x^i \cup \mathscr{D}_x^j$ dominates $\mathscr{I}_x$, whenever $i, j \in \{0, 1, 2\}$ and $i \neq j$. (Mostly we apply this with $(i, j) = (0, 1)$, i.e., we take the standard dyadic system, and one shifted copy of it.) Thus, an immediate consequence of Proposition 2.9 and Exercise 2.2 is that

$$\begin{aligned} M_b f(x) = \sup_{I \in \mathscr{I}_x} \frac{1}{|I|} \int_I |f| \, \mathrm{d}y &\leq 6 \sup_{J \in \mathscr{D}_x^i \cup \mathscr{D}_x^j} \frac{1}{|J|} \int_J |f| \, \mathrm{d}y \\ &\leq 6 \max(M_d^i f(x), M_d^j f(x)) \leq 6(M_d^i f(x) + M_d^j f(x)), \end{aligned}$$

where M_d^i is the dyadic maximal operator related to the dyadic system $\mathscr{D}^i$ in place of $\mathscr{D}$. (The notation M_b is used, since the intervals are one-dimensional "balls".) Thus, many results about the Hardy–Littlewood maximal operator can be deduced from results for the dyadic maximal operator, as soon as we use more than one dyadic system. We will later generalize this result to several dimensions as well.

Exercise 2.10. Consider the *triadic systems of intervals*

$$\mathscr{T}^i = \bigcup_{k \in \mathbb{Z}} \mathscr{T}_k^i, \qquad \mathscr{T}_k^i := \{3^{-k}([0, 1) + m + i/2) : m \in \mathbb{Z}\}, \qquad i = 0, 1.$$

Sketch a picture, and convince yourself (no need to make a detailed verification) of the facts that $\mathscr{T}_k^i$ is a partition of $\mathbb{R}$ and $\mathscr{T}_{k+1}^0$ refines $\mathscr{T}_k^0$. *Prove* that $\mathscr{T}_{k+1}^1$ refines $\mathscr{T}_k^1$.

Exercise 2.11. Prove an analogue of Proposition 2.9 for $\mathscr{T}^0$ and $\mathscr{T}^1$: For any finite interval $I \subseteq \mathbb{R}$, there exists $J \in \mathscr{T}^0 \cup \mathscr{T}^1$ such that $J \supseteq I$ and $\ell(I) \leq K\ell(J)$. Which estimate can you get for K?

3. Higher dimensions and first weighted bounds

3.A. **Special shifted systems,** $d \geq 1$**.** Recall that our general construction of shifted dyadic systems in any dimension was the formula

$$Q \dot{+} \omega = Q + \omega(\ell(Q)) = Q + \sum_{j : 2^{-j} < \ell(Q)} \omega_j 2^{-j},$$

where $\omega = (\omega_j)_{j \in \mathbb{Z}} \in (\{0, 1\}^d)^{\mathbb{Z}}$. Thus each ω_j is a d-vector $\omega_j = (\omega_j^i)_{i=1}^d$, and we can also from the sequences $\omega^i = (\omega_j^i)_{j \in \mathbb{Z}} \in \{0, 1\}^{\mathbb{Z}}$, which give rise to one-dimensional shifted dyadic systems.

If a cube $Q = I_1 \times \cdots \times I_d$ is written as a product of intervals, the shift takes the form

$$\begin{aligned} Q\dot{+}\omega &= I_1 \times \cdots \times I_d + \Big(\sum_{j:2^{-j}<\ell(Q)} \omega_j^i 2^{-j} \Big)_{i=1}^d \\ &= \Big(I_1 + \sum_{j:2^{-j}<\ell(I_1)} \omega_j^1 2^{-j} \Big) \times \cdots \times \Big(I_d + \sum_{j:2^{-j}<\ell(I_d)} \omega_j^d 2^{-j} \Big) \\ &= (I_1 \dot{+} \omega^1) \times \cdots \times (I_d \dot{+} \omega^d), \end{aligned} \tag{3.1}$$

a product of shifted intervals in each of the coordinate directions. (We used above the fact that $\ell(Q) = \ell(I_1) = \ldots = \ell(I_d)$: the side-length of a cube is equal to the length of any of its component intervals.) This establishes a one-to-one correspondence between shifted systems in d dimensions, and d-fold products of shifted systems in one dimensions.

If we take the products of the special shifted systems (2.8), we arrive at

$$\begin{aligned} \mathscr{D}^\alpha(\mathbb{R}^d) &= \Big\{ 2^{-k}([0,1)^d + m + \frac{1}{3}(-1)^k \alpha) : k \in \mathbb{Z}, m \in \mathbb{Z}^d \Big\}, \qquad \alpha = (\alpha_i)_{i=1}^d \in \{0,1,2\}^d, \\ &= \{ Q = I_1 \times \cdots \times I_d : I_i \in \mathscr{D}^{\alpha_i}(\mathbb{R}), \ell(I_1) = \ldots = \ell(I_d) \}, \end{aligned}$$

where we have indicated the underlying space of the dyadic systems in parentheses for clarity. Usually we work in one fixed dimension, so that there is no need to do so.

Proposition 3.2. *Let Q be any cube (with sides parallel to the coordinate axes) in $\mathbb{R}^d$. Then there exists $\alpha \in \{0,1\}^d$, and $R \in \mathscr{D}^\alpha(\mathbb{R}^d)$, such that*

$$R \supseteq Q, \qquad 3\ell(Q) < \ell(R) \leq 6\ell(Q).$$

Note that we need only the 2^d cases $\alpha \in \{0,1\}^d$ here. All the 3^d cases $\alpha \in \{0,1,2\}^d$ are used for a somewhat stronger conclusion formulated in the exercise below.

Proof. We write $Q = I_1 \times \ldots \times I_d$ as a product of intervals, and apply Proposition 2.9 in each coordinate direction. As stated, Proposition 2.9 says that for at least two choices of $\alpha_i \in \{0,1,2\}$, there is $J_i \in \mathscr{D}^{\alpha_i}(\mathbb{R})$ with $J_i \supset I_i$ and $3\ell(I_i) < \ell(J_i) \leq 6\ell(I_i)$. This means in particular that the same conclusion is true for at least one choice of $\alpha_i \in \{0,1\}$. Then it follows that the set $R := J_1 \times \cdots \times J_d$ contains $Q = I_1 \times \cdots \times I_d$.

To conclude observe that $\ell(I_i) = \ell(Q)$ for all i. Now $\ell(J_i) \in (3\ell(I_i), 6\ell(I_i)] = (3\ell(Q), 6\ell(Q)]$ and $\ell(J_i)$ is an integer power of two. There is exactly one such number in the interval $(3\ell(Q), 6\ell(Q)]$, and therefore all $\ell(J_i)$ are actually equal. Thus R is actually a cube, with $\ell(R) = \ell(J_1) = \ldots = \ell(J_d) \in (3\ell(Q), 6\ell(Q)]$, and more specifically it is a cube of the form appearing in (3.1). Thus $R \in \mathscr{D}^\alpha$, and we are done. $\square$

Exercise 3.3. Prove the following variant of Proposition 3.2: Let Q and P be any two cubes (with sides parallel to the coordinate axes) in $\mathbb{R}^d$. Then there exists $\alpha \in \{0,1,2\}^d$, and $R, S \in \mathscr{D}^\alpha(\mathbb{R}^d)$ (same α for both R and S!), such that

$$\begin{aligned} R \supseteq Q, &\qquad 3\ell(Q) < \ell(R) \leq 6\ell(Q), \\ S \supseteq P, &\qquad 3\ell(P) < \ell(S) \leq 6\ell(P). \end{aligned}$$

(Hint: Deduce the one-dimensional case from Proposition 2.9, then proceed to higher dimensions in a similar way as in the proof of Proposition 3.2.) Check that if $Q \subseteq P$, then $R \subseteq S$.

3.B. **A maximal function estimate with different measures.** The action of maximal operators can be naturally extended to measures in place of functions. If ν is a positive Borel measure, we define

$$M_\mu^d \nu(x) := \sup_{\substack{Q \in \mathscr{D} \\ Q \ni x}} \frac{\nu(Q)}{\mu(Q)}. \tag{3.4}$$

It is immediate that

$$M_\mu^d \nu(x) = M_\mu^d \phi(x) \quad \text{if} \quad \nu(E) = \int_E \phi(x) \, \mathrm{d}\mu(x),$$

i.e., if ν is a *weighted measure* (with density $\phi \geq 0$) relative to the measure μ. The above relation of ν, μ and ϕ is often abbreviated as $\mathrm{d}\nu = \phi\,\mathrm{d}\mu$.

Proposition 3.5 (Fefferman–Stein [FS71])**.**

$$\nu(\{M_\mu^d f > \lambda\}) \leq \frac{1}{\lambda}\int_{\mathbb{R}^d} |f(x)| M_\mu^d \nu(x)\,\mathrm{d}\mu(x).$$

Proof. The proof is a simple adaptation of the case $\mu = \nu$. As before, consider

$$\mathscr{Q}_\lambda := \Big\{ Q \in \mathscr{D} : \frac{1}{\mu(Q)}\int_Q |f|\,\mathrm{d}\mu > \lambda \Big\}.$$

We write the proof in the case that $\mathscr{Q}_\lambda$ is without infinite increasing chains; in general, ensuring this would require a truncation argument, but we skip those details. Then, denoting by $\mathscr{Q}_\lambda$ the maximal cubes in $\mathscr{Q}_\lambda$, we have

$$\begin{aligned}
\nu(\{M_\mu^d f > \lambda\}) &= \nu\Big(\bigcup_{Q\in\mathscr{Q}_\lambda} Q\Big) = \nu\Big(\bigcup_{Q\in\mathscr{Q}_\lambda^*} Q\Big) = \sum_{Q\in\mathscr{Q}_\lambda^*} \nu(Q) \\
&= \sum_{Q\in\mathscr{Q}_\lambda^*} \frac{\nu(Q)}{\mu(Q)}\cdot\mu(Q) \leq \sum_{Q\in\mathscr{Q}_\lambda^*} \inf_{z\in Q} M_\mu^d\nu(z)\cdot\frac{1}{\lambda}\int_Q |f(x)|\,\mathrm{d}\mu(x) \\
&\leq \frac{1}{\lambda}\sum_{Q\in\mathscr{Q}_\lambda^*}\int_Q |f(z)| M_\mu^d\nu(x)\,\mathrm{d}\mu(x) \leq \frac{1}{\lambda}\int_{\mathbb{R}^d} |f(z)| M_\mu^d\nu(x)\,\mathrm{d}\mu(x).
\end{aligned}$$

Compared to the previous cases, the additional step was to multiply and divide by $\mu(Q)$, and to estimate the ratio $\nu(Q)/\mu(Q)$ in terms of the maximal function. □

3.C. **L^p bounds and interpolation.** Denoting

$$\mathrm{d}\omega := M_\mu^d\nu\,\mathrm{d}\mu,$$

the estimate just proven takes the form

$$\lambda\cdot\nu(\{M_\mu^d f > \lambda\}) \leq \|f\|_{L^1(\omega)}.$$

Defining

$$\|g\|_{L^{1,\infty}(\nu)} := \sup_{\lambda>0}\lambda\cdot\nu(\{|g| > \lambda\}),$$

this can be further written as

$$\|M_\mu^d f\|_{L^{1,\infty}(\nu)} \leq \|f\|_{L^1(\omega)},$$

which is known as a *weak-type* estimate. The space

$$L^{1,\infty}(\nu) := \{g : g \text{ is } \nu\text{-measurable and } \|g\|_{L^{1,\infty}}(\nu) < \infty\}$$

is called the weak L^1 space, and $\|\ \|_{L^{1,\infty}(\nu)}$ the weak L^1 norm, although it is actually not a norm. (It does not satisfy the triangle inequality.) Note that

$$\lambda\cdot\nu(\{|g| > \lambda\}) = \int_{\{|g|>\lambda\}} \lambda\,\mathrm{d}\nu \leq \int_{\{|g|>\lambda\}} |g|\,\mathrm{d}\nu \leq \|g\|_{L^1(\nu)},$$

so that $\|g\|_{L^{1,\infty}(\nu)} \leq \|g\|_{L^1(\nu)}$. In general the latter norm can be much larger. For example, on $\mathbb{R}$ with the Lebesgue measure, we have $1/x \in L^{1,\infty}(\mathbb{R})\setminus L^1(\mathbb{R})$. This explains the name "weak".

We would now like to consider L^p estimates for M_μ^d for $p > 1$. Let us first consider the extreme $p = \infty$:

Lemma 3.6.

$$\|M_\mu^d f\|_{L^\infty(\nu)} \leq \|f\|_{L^\infty(\omega)}, \qquad \mathrm{d}\omega = M_\mu^d\nu\,\mathrm{d}\mu.$$

Heuristic proof. If f is bounded by K, then

$$\frac{1}{\mu(Q)}\int_Q |f|\,\mathrm{d}\mu \leq \frac{1}{\mu(Q)}\int_Q K\,\mathrm{d}\mu = K\frac{\mu(Q)}{\mu(Q)} = K,$$

and hence $M_\mu^d f$ is also bounded by K. However, this is not a proper proof, since $\|f\|_{L^\infty(\omega)} = K$ does not mean that $|f(x)| \leq K$ at every x, only that this is true for ω-almost every x, and one needs some care with the zero sets of the different measures. □

Exercise 3.7. Give a proper proof of Lemma 3.6. (Hint: Recall that

$$\|f\|_{L^\infty(\omega)} := \inf\Big\{\lambda \in [0,\infty) : \omega(\{|f| > \lambda\}) = 0\Big\}.$$

Argue that it is enough to prove that $\nu(Q) = 0$ for all $Q \in \mathscr{Q}_\lambda$, where this set is defined as usual.) Also check that the above heuristic proof is almost correct in the case that $\nu = \mu$.

The L^p inequality for the maximal function reads as follows:

Theorem 3.8 (Dyadic maximal function inequalities with two measures)**.** *For two locally finite measures μ and ν on $\mathbb{R}^d$, we have the inequalities*

$$\begin{aligned}\|M_\mu^d f\|_{L^{1,\infty}(\nu)} &\leq \|f\|_{L^1(\omega)}, \qquad \mathrm{d}\omega = M_\mu^d \nu\,\mathrm{d}\mu,\\ \|M_\mu^d f\|_{L^p(\nu)} &\leq p'\|f\|_{L^p(\omega)}, \qquad \forall p \in (1,\infty], \quad p' = \frac{p}{p-1}.\end{aligned}$$

In particular, if $\mathrm{d}\nu = w\,\mathrm{d}\mu$ is a weighted measure with respect to μ, the second bound takes the form

$$\Big(\int_{\mathbb{R}^d}(M_\mu^d f)^p w\,\mathrm{d}\mu\Big)^{1/p} \leq p'\Big(\int_{\mathbb{R}^d}|f|^p M_\mu^d w\,\mathrm{d}\mu\Big)^{1/p}, \tag{3.9}$$

which will have an interesting application — completely unrelated to weights — in the following section.

Proof. The weak-type bound is Proposition 3.5, and the case $p = \infty$ is Lemma 3.6. The remaining cases are a consequence of these two via the following general interpolation theorem. □

Theorem 3.10 (Marcinkiewicz interpolation theorem, special case)**.** *Suppose that T is a sublinear operator (meaning that $|T(f+g)| \leq |Tf| + |Tg|$ pointwise) from $L^1(\omega) + L^\infty(\omega)$ to ν-measurable functions, such that*

$$\begin{aligned}\|Tf\|_{L^{1,\infty}(\nu)} &\leq A_1\|f\|_{L^1(\omega)},\\ \|Tf\|_{L^\infty(\nu)} &\leq A_\infty\|f\|_{L^\infty(\omega)}.\end{aligned}$$

Then

$$\|Tf\|_{L^p(\nu)} \leq p' A_1^{1/p} A_\infty^{1/p'}\|f\|_{L^p(\omega)}, \quad \forall p \in (1,\infty).$$

Proof. The proof depends on the important level set formula (applied to $\phi = Tf$)

$$\|\phi\|_{L^p(\nu)}^p = \int_0^\infty p\lambda^{p-1}\nu(\{|\phi| > \lambda\})\,\mathrm{d}\lambda. \tag{3.11}$$

This is most easily verified by observing that both sides are equal to the double integral (evaluated in one or the other order, which is legitimate by Fubini's theorem)

$$\int_0^\infty\int_{\mathbb{R}^d} p\lambda^{p-1}1_{\{(x,\lambda):|\phi(x)|>\lambda\}}(x,\lambda)\,\mathrm{d}\nu(x)\,\mathrm{d}\lambda.$$

Another tool is the splitting of f into parts where it is "small" or "large" relative to the level λ. A first attempt would

$$\tilde f_\lambda := 1_{\{|f|\leq A\lambda\}}f, \qquad \tilde f^\lambda := 1_{\{|f|>A\lambda\}}f,$$

where $\|\tilde f_\lambda\|_{L^\infty(\omega)} \leq A\lambda$. However, we can have this same bound even if we make $\tilde f_\lambda$ slightly larger, and thus $\tilde f^\lambda$ slightly smaller; namely, we define

$$f_\lambda := 1_{\{|f|\leq A\lambda\}}f + 1_{\{|f|>A\lambda\}}\frac{f}{|f|}A\lambda, \qquad f^\lambda := 1_{\{|f|>A\lambda\}}f\Big(1 - \frac{A\lambda}{|f|}\Big),$$

The parameter A will be chosen below.

We also split λ as $\lambda = \alpha\lambda + \beta\lambda$, where $\alpha + \beta = 1$, and this will be also chosen below. Now, if $|Tf| > \lambda$, then

$$\alpha\lambda + \beta\lambda = \lambda < |Tf| \leq |Tf_\lambda| + |Tf^\lambda|,$$

and we must have at least one of $|Tf_\lambda| > \alpha\lambda$ or $|Tf^\lambda| > \beta\lambda$.

We now choose the parameters so that the first case is impossible; namely, we have

$$\|Tf_\lambda\|_{L^\infty(\nu)} \leq A_\infty \|f_\lambda\|_{L^\infty(\omega)} \leq A_\infty A\lambda.$$

Thus we choose $A = \alpha/A_\infty$, and conclude that $|Tf_\lambda| > \alpha\lambda$ only in a set of measure zero. Hence

$$\nu(\{|Tf| > \lambda\}) \leq \nu(\{|Tf^\lambda| > \beta\lambda\}) \leq \frac{A_1}{\beta\lambda}\|f^\lambda\|_{L^1(\omega)} = \frac{A_1}{\beta\lambda}\int 1_{\{|f|>A\lambda\}}(|f| - A\lambda)\,\mathrm{d}\omega,$$

and thus

$$\begin{aligned}
\|Tf\|_{L^p(\nu)}^p &= \int_0^\infty p\lambda^{p-1}\nu(\{|Tf| > \lambda\})\,\mathrm{d}\lambda \\
&\leq \int_0^\infty p\lambda^{p-1}\nu(\{|Tf^\lambda| > \beta\lambda\})\,\mathrm{d}\lambda \\
&\leq \frac{pA_1}{\beta}\int_0^\infty \lambda^{p-2}\int 1_{\{|f|>A\lambda\}}(|f| - A\lambda)\,\mathrm{d}\omega\,\mathrm{d}\lambda \\
&\leq \frac{pA_1}{\beta}\int\int_0^{|f|/A} (\lambda^{p-2}|f| - A\lambda^{p-1})\,\mathrm{d}\lambda\,\mathrm{d}\omega \\
&\leq \frac{pA_1}{\beta}\int\Big(\frac{1}{p-1}\big(\frac{|f|}{A}\big)^{p-1}|f| - \frac{A}{p}\big(\frac{|f|}{A}\big)^p\Big)\,\mathrm{d}\omega \\
&= \frac{A_1}{\beta A^{p-1}(p-1)}\int |f|^p\,\mathrm{d}\omega = \frac{A_1 A_\infty^{p-1}}{(1-\alpha)\alpha^{p-1}(p-1)}\|f\|_{L^p(\omega)}^p,
\end{aligned}$$

where we took into account that $\beta = 1 - \alpha$ and $A = \alpha/A_\infty$. It is a high school exercise (find the zero of the derivative etc.) to optimize the right side in terms of α. The optimum is reached at $\alpha = (p-1)/p = 1/p'$, at which point the denominator above becomes $(1/p')^p$. Taking the pth roots we arrive at the asserted estimate for $\|Tf\|_{L^p(\nu)}$. $\square$

Exercise 3.12. Let us denote by

$$Mf(x) := \sup_{\substack{Q \text{ cube} \\ Q \ni x}} \frac{1}{|Q|}\int_Q |f(y)|\,\mathrm{d}y$$

the Hardy–Littlewood maximal function associated with the family of all cubes (with sides parallel to coordinate axes) in $\mathbb{R}^d$. Use Theorem 3.8 and shifted dyadic cubes to derive the analogous statements for the Hardy–Littlewood maximal operator: For any locally finite measure ν on $\mathbb{R}^d$,

$$\begin{aligned}
\|Mf\|_{L^{1,\infty}(\nu)} &\leq c_d\|f\|_{L^1(M\nu)}, \\
\|Mf\|_{L^p(\nu)} &\leq c_{d,p}\|f\|_{L^p(M\nu)}, \qquad p \in (1, \infty],
\end{aligned}$$

where

$$\|f\|_{L^p(w)} := \Big(\int |f|^p w\,\mathrm{d}x\Big)^{1/p}$$

(with $w(x) = M\nu(x)$, which is defined by an obvious modification of (3.4)) denotes the L^p norm weighted with respect to the Lebesgue measure. Pay attention to the weak-type estimate, bearing in mind that $\|\ \|_{L^{1,\infty}(\nu)}$ is *not* a proper norm.

4. Vector-valued maximal inequality

4.A. **Duality of L^p spaces.** The elementary inequality

$$ab \leq \frac{a^p}{p} + \frac{b^{p'}}{p'}, \qquad \text{for} \quad a, b \in [0, \infty), \quad p \in (1, \infty), \quad p' = \frac{p}{p-1},$$

can be easily verified as a high school exercise. If we apply this to $a = |f(x)|$ and $b = |g(x)|$ for each x, we find that

$$\int |f(x)g(x)| \, \mathrm{d}\mu(x) \leq \frac{1}{p} \int |f(x)|^p \, \mathrm{d}\mu(x) + \frac{1}{p'} \int |g(x)|^{p'} \, \mathrm{d}\mu(x).$$

If we replace f by λf and g by g/λ (for a constant $\lambda > 0$, the left side is unchanged, and the above estimate takes the form

$$\int |fg| \, \mathrm{d}\mu \leq \frac{\lambda^p}{p} \|f\|_p^p + \frac{\lambda^{-p'}}{p'} \|g\|_{p'}^{p'}.$$

If we optimize the right size with respect to λ, we arrive at the final form

$$\int |fg| \, \mathrm{d}\mu \leq \|f\|_p \|g\|_{p'}, \tag{4.1}$$

which is known as *Hölder's inequality.*

There is an important converse to this estimate, contained in the following:

Theorem 4.2. *Let μ be a σ-finite measure on $\mathbb{R}^d$. For any μ-measurable function f and $p \in (1, \infty)$, we have*

$$\|f\|_{L^p(\mu)} = \sup \Big\{ \Big| \int fg \, \mathrm{d}\mu \Big| : g \in L^{p'}(\mu) : \|g\|_{L^{p'}(\mu)} \leq 1 \Big\}. \tag{4.3}$$

The result is also true for $p = 1, \infty$, but requires a modification of the argument. The space $\mathbb{R}^d$ could be replaced by a more general measure space, with the same proof.

Proof. The bound "$\geq$" is already contained in (4.1), so it remains to consider "$\leq$".

Case $\|f\|_p = 0$. Then $f = 0$ almost everywhere, and the right side of (4.3) is zero as well.

Case $0 < \|f\|_p < \infty$. We have

$$\|f\|_p^p = \int |f|^p \, \mathrm{d}\mu = \int f \cdot \bar{f} |f|^{p-2} \, \mathrm{d}\mu =: \int f \cdot h_f \, \mathrm{d}\mu,$$

where

$$\|h_f\|_{p'}^{p'} = \int (|f|^{p-1})^{p'} \, \mathrm{d}\mu = \int |f|^p \, \mathrm{d}\mu = \|f\|_p^p;$$

thus

$$\|h_f\|_{p'} = \|f\|_p^{p/p'} = \|f\|_p^{p-1},$$

and hence

$$\|f\|_p = \frac{\|f\|_p^p}{\|f\|_p^{p-1}} = \frac{\int f \cdot h_f \, \mathrm{d}\mu}{\|h_f\|_{p'}} =: \int f \cdot g_f \, \mathrm{d}\mu,$$

where

$$g_f := \frac{h_f}{\|h_f\|_{p'}} = \frac{\bar{f}|f|^{p-2}}{\|f\|_p^{p-1}} \tag{4.4}$$

has $L^{p'}(\mu)$-norm equal to 1. So in this case, we not only have "$\leq$" in (4.3), but the supremum is actually reached with the specific function g_f.

Case $\|f\|_p = \infty$. The σ-finiteness assumption says that there are sets E_n of finite measure such that $E_n \uparrow \mathbb{R}^d$. Then also $F_n := \{|f| \leq n\} \cap E_n \uparrow \mathbb{R}^d$, and it is immediate that $\|1_{F_n} f\|_p < \infty$. Thus the previous case of the proof shows that there exists a function $g_n := g_{1_{F_n} f}$ with $\|g_n\|_{p'} = 1$ and

$$\int f \cdot 1_{F_n} g_n \, \mathrm{d}\mu = \int 1_{F_n} f \cdot g_n \, \mathrm{d}\mu = \|1_{F_n} f\|_p \to \|f\|_p = \infty$$

as $n \to \infty$, and $\|1_{F_n} g_n\|_p \leq \|g_n\|_p = 1$. (In fact, one can check that g_n is supported on F_n, so that $1_{F_n} g_n = g_n$, and its norm is exactly one.) This proves "$\leq$" in (4.3) even in this case. $\square$

4.B. **Vector-valued L^p spaces.** The space $L^p(\mu; \ell^q)$ consists of all sequences of functions $\vec{f} = (f_i)_{i=1}^\infty$ (or equivalently, sequence-valued functions) such that

$$\|\vec{f}\|_{L^p(\mu;\ell^q)} := \Big(\int \Big[\sum_{i=1}^\infty |f_i(x)|^q \Big]^{p/q} \mathrm{d}\mu(x) \Big)^{1/p} = \Big(\int |\vec{f}(x)|_{\ell^q}^p \, \mathrm{d}\mu(x) \Big)^{1/p} = \| |\vec{f}|_{\ell^q} \|_{L^p(\mu)}$$

is finite. Observe that if $p = q$, then

$$\begin{aligned} \|\vec{f}\|_{L^p(\mu;\ell^p)} &= \Big(\int \sum_{i=1}^\infty |f_i(x)|^p \, \mathrm{d}\mu(x) \Big)^{1/p} \\ &= \Big(\sum_{i=1}^\infty \int |f_i(x)|^p \, \mathrm{d}\mu(x) \Big)^{1/p} = \Big(\sum_{i=1}^\infty \|f_i\|_p^p \Big)^{1/p} \end{aligned} \tag{4.5}$$

can be written in terms of the L^p norms of the f_i, but this is the only such case.

The duality theorem extends to the vector-valued case as follows:

Theorem 4.6. *Let μ be a σ-finite measure on $\mathbb{R}^d$. For any μ-measurable sequence of functions $\vec{f} = (f_i)_{i=1}^\infty$ and $p, q \in (1, \infty)$, we have*

$$\|\vec{f}\|_{L^p(\mu;\ell^q)} = \sup \Big\{ \Big| \int \sum_{i=1}^\infty f_i g_i \, \mathrm{d}\mu \Big| : \vec{g} = (g_i)_{i=1}^\infty \in L^{p'}(\mu; \ell^{q'}) : \|\vec{g}\|_{L^{p'}(\mu;\ell^{q'})} \leq 1 \Big\}. \tag{4.7}$$

Proof. "$\geq$" consists of two applications of Hölder's inequality, one on sequences and the other on functions:

$$\int \sum_{i=1}^\infty |f_i(x) g_i(x)| \, \mathrm{d}\mu(x) \leq \int |\vec{f}(x)|_{\ell^q} |\vec{g}(x)|_{\ell^{q'}} \, \mathrm{d}\mu(x) \leq \| |\vec{f}|_{\ell^q} \|_{L^p(\mu)} \| |\vec{g}|_{\ell^{q'}} \|_{L^{p'}(\mu)}.$$

"$\leq$", case $0 < \|\vec{f}\|_{L^p(\mu;\ell^q)} < \infty$. We observe that the Duality Theorem 4.2 also applies to sequence space: either one can repeat the same proof, or observe that sequence spaces are actually special cases of function spaces, where the underlying measure $\mu = \sum_{i=1}^\infty \delta_i$ is supported on $\mathbb{Z}_+ = \{1, 2, \ldots\}$, and assigns the measure one to each $i \in \mathbb{Z}_+$. Thus, for each x, related to the sequence $\vec{f}(x) = (f_i(x))_{i=1}^\infty$, we can find another sequence $\vec{v}(x) = (v_i(x))_{i=1}^\infty$, given by an adaptation of the formula (4.4):

$$v_i(x) = \frac{\bar{f}_i(x) |f_i(x)|^{q-2}}{|\vec{f}(x)|_{\ell^q}^{q-1}},$$

so that $\sum_{i=1}^\infty f_i(x) v_i(x) = |\vec{f}(x)|_{\ell^q}$ and $|\vec{v}(x)|_{\ell^{q'}} = 1$. On the other hand, by the Duality Theorem 4.2 applied to the function $x \mapsto |\vec{f}(x)|_{\ell^q}$, we find another function u, again given by (4.4):

$$u(x) = \frac{|\vec{f}(x)|_{\ell^q}^{p-1}}{\| |\vec{f}|_{\ell^q} \|_{L^p(\mu)}^{p-1}},$$

such that $\int |\vec{f}|_{\ell^q} u \, \mathrm{d}\mu = \| |\vec{f}|_{\ell^q} \|_{L^p(\mu)}$. Putting things together, we find that

$$\|\vec{f}\|_{L^p(\mu;\ell^q)} = \int |\vec{f}(x)|\ell^q u(x) \, \mathrm{d}\mu(x) = \int \sum_{i=1}^\infty f_i(x) v_i(x) u(x) \, \mathrm{d}\mu(x) =: \int \sum_{i=1}^\infty f_i(x) g_i(x) \, \mathrm{d}\mu(x),$$

where

$$g_i(x) = v_i(x)u(x) = \frac{\bar{f}_i(x)|f_i(x)|^{q-2}}{|\vec{f}(x)|_{\ell^q}^{q-p}\|\vec{f}\|_{L^p(\mu;\ell^q)}^{p-1}}$$

satisfies

$$\||\vec{g}|_{\ell^{q'}}\|_{L^p(\mu)} = \|u|\vec{v}|_{\ell^{q'}}\|_{L^p(\mu)} = \|u\|_{L^p(\mu)} = 1. \qquad \square$$

Exercise 4.8. Complete the proof of Theorem 4.6 by considering the case that $\|\vec{f}\|_{L^p(\mu;\ell^q)} = \infty$. (Hint: In addition to the other approximations, it may be useful to truncate the sequences by considering approximations with $f_i \equiv 0$ for $i > n$.)

4.C. **The Fefferman–Stein inequality.**

Theorem 4.9 (Dyadic Fefferman–Stein inequality [FS71])**.**

$$\|(M_\mu^d f_i)_{i=1}^\infty\|_{L^p(\mu;\ell^q)} \le c_{pq}\|(f_i)_{i=1}^\infty\|_{L^p(\mu;\ell^q)}$$

Proof. We divide the proof into three cases according to the relative size of p and q:

Case $p = q$. This is the easy case: From (4.5), we find that

$$\begin{aligned}\|(M_\mu^d f_i)_{i=1}^\infty\|_{L^p(\mu;\ell^p)} &\overset{(4.5)}{=} \Big(\sum_{i=1}^\infty \|M_\mu^d f_i\|_{L^p(\mu)}^p\Big)^{1/p} \\ &\overset{(\star)}{\le} \Big(\sum_{i=1}^\infty (p'\|f_i\|_{L^p(\mu)})^p\Big)^{1/p} \overset{(4.5)}{=} p'\|(f_i)_{i=1}^\infty\|_{L^p(\mu;\ell^p)},\end{aligned}$$

where $(\star)$ was an application of the usual dyadic maximal inequality in each component f_i separately.

Case $p > q$. Now

$$\|(h_i)_{i=1}^\infty\|_{L^p(\mu;\ell^q)} = \Big(\int \Big[\sum_{i=1}^\infty |h_i|^q\Big]^{p/q} \mathrm{d}\mu\Big)^{q/p\times 1/q} = \Big\|\sum_{i=1}^\infty |h_i|^q\Big\|_{L^{p/q}(\mu)}^{1/q}, \tag{4.10}$$

where $L^{p/q}(\mu)$ is another Lebesgue space with exponent $p/q \in (1,\infty)$, where we can apply the Duality Theorem 4.2. Thus

$$\|(M_\mu^d f_i)_{i=1}^\infty\|_{L^p(\mu;\ell^q)}^q = \sup\Big\{\int \Big[\sum_{i=1}^\infty (M_\mu^d f_i)^q\Big] g\,\mathrm{d}\mu : \|g\|_{L^{(p/q)'}(\mu)} \le 1\Big\},$$

where

$$\begin{aligned}\int \Big[\sum_{i=1}^\infty (M_\mu^d f_i)^q\Big] g\,\mathrm{d}\mu &= \sum_{i=1}^\infty \int (M_\mu^d f_i)^q g\,\mathrm{d}\mu \le \sum_{i=1}^\infty (q')^q \int |f_i|^q M_\mu g\,\mathrm{d}\mu \\ &= (q')^q \int \Big[\sum_{i=1}^\infty |f_i|^q\Big] M_\mu g\,\mathrm{d}\mu \le (q')^q \Big\|\sum_{i=1}^\infty |f_i|^q\Big\|_{L^{p/q}(\mu)} \|M_\mu g\|_{L^{(p/q)'}(\mu)} \\ &\le (q')^q \|(f_i)_{i=1}^\infty\|_{L^p(\mu)}^q \frac{p}{q}\|g\|_{L^{(p/q)'}(\mu)}.\end{aligned}$$

The last factor is one in the supremum under consideration, so altogether we conclude that

$$\|(M_\mu^d f_i)_{i=1}^\infty\|_{L^p(\mu;\ell^q)} \le q'\Big(\frac{p}{q}\Big)^{1/q}\|(f_i)_{i=1}^\infty\|_{L^p(\mu;\ell^q)}, \qquad 1 < q < p < \infty.$$

Case $p < q$. Our strategy is to reduce this to the previous case by duality: observe that now $p' > q'$. However, there is a slight technical complication which requires to choose an additional auxiliary exponent $s \in (1,p)$. We then observe that

$$\|(h_i)_{i=1}^\infty\|_{L^p(\mu;\ell^q)} = \Big(\int \Big[\sum_{i=1}^\infty |h_i|^{s\times q/s}\Big]^{(p/s)/(q/s)} \,\mathrm{d}\mu\Big)^{s/p\times 1/s} = \|(|h_i|^s)_{i=1}^\infty\|_{L^{p/s}(\mu;\ell^{q/s})}^{1/s}. \tag{4.11}$$

By the vector-valued Duality Theorem 4.6,

$$\|(M_\mu^d f_i)_{i=1}^\infty\|_{L^p(\mu;\ell^q)}^s = \sup\Big\{\int \sum_{i=1}^\infty (M_\mu^d f_i)^s g_i \,\mathrm{d}\mu : \|\vec{g}\|_{L^{(p/s)'}(\mu;\ell^{(q/s)'}} \leq 1\Big\},$$

where

$$\begin{aligned}\int \sum_{i=1}^\infty (M_\mu^d f_i)^s g_i \,\mathrm{d}\mu &= \sum_{i=1}^\infty \int (M_\mu^d f_i)^s g_i \,\mathrm{d}\mu \leq \sum_{i=1}^\infty (s')^s \int |f_i|^s M_\mu^d g_i \,\mathrm{d}\mu \\ &\leq (s')^s \|(|f_i|^s)_{i=1}^\infty\|_{L^{(p/s)}(\mu;\ell^{q/s})} \|(M_\mu^d g_i)_{i=1}^\infty\|_{L^{(p/s)'}(\mu;\ell^{(q/s)'})} \\ &\leq (s')^s \|(f_i)_{i=1}^\infty\|_{L^p(\mu;\ell^q)}^s \frac{q}{s}\Big(\frac{(p/s)'}{(q/s)'}\Big)^{1-s/q} \|(g_i)_{i=1}^\infty\|_{L^{(p/s)'}(\mu;\ell^{(q/s)'})},\end{aligned}$$

where we used (4.11) to the f factor, and the previous case of the Theorem to the g factor, with $(p/s)' > (q/s)'$ in place of $p > q$. Altogether, it follows that

$$\|(M_\mu^d f_i)_{i=1}^\infty\|_{L^p(\mu;\ell^q)} \leq s'\Big(\frac{q}{s}\Big)^{1/s}\Big(\frac{p(q-s)}{q(p-s)}\Big)^{1/s-1/q} \|(f_i)_{i=1}^\infty\|_{L^p(\mu;\ell^q)}, \tag{4.12}$$

which is valid for any $s \in (1,p)$. □

Remark 4.13. We followed the (main lines of) the original proof of Fefferman and Stein for the case $p > q$, but not for $p < q$. Originally, they handled this with the help of a vector-valued L^1–$L^{1,\infty}$ estimate. This weak-type strategy is perhaps a bit more complicated, but it provides additional information and a better bound for the constants, even if one tries to optimize with respect to the parameter $s \in (1,p)$ in (4.12).

Exercise 4.14. Use Theorem 4.9 to derive the usual form of the Fefferman–Stein inequality for the Hardy–Littlewood maximal operator M and the Lebesgue measure:

$$\|(Mf_i)_{i=1}^\infty\|_{L^p(\mathbb{R}^d;\ell^q)} \leq c_{pqd}\|(f_i)_{i=1}^\infty\|_{L^p(\mathbb{R}^d;\ell^q)}.$$

Pay attention to which versions of the triangle inequality you need to get this.

5. The median

Our next major goal is to prove a certain formula, discovered by A. Lerner [Ler10], which provides very useful and precise information about a measurable function in terms of its "local oscillations". Before we can state the formula, we need to introduce some auxiliary concepts, the first of which is the notion of a median of a measurable function.

5.A. **Lebesgue's differentiation theorem.**

The following result, on the one hand, is one of the main applications of the maximal function. On the other hand, it will be soon needed in our study of the median.

Here we always consider the Lebesgue measure on $\mathbb{R}^d$.

Theorem 5.1 (Lebesgue)**.** *Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$. Then for a.e. $x \in \mathbb{R}^d$, we have*

$$\lim_{\substack{Q\ni x\\ \ell(Q)\to 0}} \frac{1}{|Q|}\int_Q f\,\mathrm{d}y = f(x),$$

where the limit is along all cubes that contain x and whose sidelength shrinks to zero.

Proof. If f is continuous at x, the claim is immediate from the definition of continuity.

Case $f \in L^1(\mathbb{R}^d)$. The claim is equivalent to

$$Lf(x) := \limsup_{\substack{Q \ni x \\ \ell(Q)\to 0}} \Big| \frac{1}{|Q|} \int_Q f \, \mathrm{d}y - f(x) \Big| = 0$$

almost everywhere. We make the following observations concerning the operator L:

- L is sublinear: $L(f+g) \leq Lf + Lg$.
- If g is continuous, then $Lg = 0$.
- $Lf \leq Mf + |f|$, where M is the Hardy–Littlewood maximal operator (with respect to cubes).

For a given $f \in L^1(\mathbb{R}^d)$, we want to prove that $Lf = 0$ a.e. We first estimate the size of the set $\{Lf > \epsilon\}$ for $\epsilon > 0$. Given $\delta > 0$, we can find a continuous f_δ (even compactly supported if we like) such that $\|f - f_\delta\|_1 < \delta$. (Recall from Real Analysis that such functions are *dense* in $L^1(\mathbb{R}^d)$.) Now

$$Lf(x) \leq L(f - f_\delta)(x) + Lf_\delta(x) \leq M(f - f_\delta)(x) + |f(x) - f_\delta(x)| + 0,$$

and hence

$$\begin{aligned} |\{Lf > \epsilon\}| &\leq |\{M(f - f_\delta) > \epsilon/2\}| + |\{|f - f_\delta| > \epsilon/2\}| \\ &\leq \frac{2}{\epsilon}\|M(f - f_\delta)\|_{L^{1,\infty}} + \frac{2}{\epsilon}\|f - f_\delta\|_{L^{1,\infty}} \\ &\leq \frac{2}{\epsilon} C\|f - f_\delta\|_1 + \frac{2}{\epsilon}\|f - f_\delta\|_1 \leq \frac{2}{\epsilon}(C+1)\delta. \end{aligned}$$

This is true for any $\delta > 0$, and the left side is independent of δ. Taking the limit $\delta \to 0$, we find that $|\{Lf > \epsilon\}| = 0$. Thus also

$$|\{Lf > 0\}| = \Big| \bigcup_{n=1}^\infty \{Lf > \frac{1}{n}\} \Big| = 0.$$

Case $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$. If $x \in B(0,n)$ (ball of radius n centred at the origin), then $Lf(x) = L(1_{B(0,2n)} f)(x) = 0$ at almost every such x, since $1_{B(0,2n)} f \in L^1(\mathbb{R}^d)$. Since $\mathbb{R}^d = \bigcup_{n=1}^\infty B(0,n)$, the claim follows. $\square$

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

$$|Q \cap \{f > m_f(Q)\}| \leq \tfrac{1}{2}|Q|, \tag{5.2}$$

$$|Q \cap \{f < m_f(Q)\}| \leq \tfrac{1}{2}|Q|. \tag{5.3}$$

Exercise 5.4. Show that every measurable function $f : Q \to \mathbb{R}$ has a median, and that the set of all medians is a closed interval. (Hint: Show that the set of numbers $m_f(Q)$ that satisfy (5.2) is of the form $[a, \infty)$ for some $a \in \mathbb{R}$; in particular, it is nonempty. Similarly, show that the set of numbers $m_f(Q)$ satisfying (5.3) is of the form $(-\infty, b]$ for some $b \in \mathbb{R}$. Finally show that $a \leq b$.)

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

The following simple observation is handy for estimating the median:

Lemma 5.5. *The following claims hold for all medians $m_f(Q)$ and real numbers α:*

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

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

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

a contradiction.

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

There is a median analogue of Lebesgue's differentiation Theorem 5.1:

Proposition 5.6 (Fujii [Fuj91]). *Let $f : \mathbb{R}^d \to \mathbb{R}$ be measurable. Then for almost every $x \in \mathbb{R}^d$, we have*

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

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

Proof. We introduce the auxiliary functions

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

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

$$\frac{|Q \cap E_{kj}|}{|Q|} = \frac{1}{|Q|} \int_Q 1_{E_{kj}} \,\mathrm{d}x \underset{\substack{Q \ni x \\ \ell(Q) \to 0}}{\longrightarrow} 1_{E_{kj}}(x) \tag{5.7}$$

for almost every $x \in \mathbb{R}^d$. Let us explicitly denote the exceptional null set by N_{kj}, so the above convergence holds for every $x \in N_{kj}^c$. Let $N := \bigcup_{k,j} N_{k,j}$. This is another null set, and the convergence (5.7) holds for every $x \in N^c$ and every $k, j \in \mathbb{Z}$.

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

$$\forall \varepsilon > 0 \ \exists \delta > 0 : \forall Q \ni x, \ell(Q) < \delta \Rightarrow |m_f(Q) - f(x)| < \varepsilon, \tag{5.8}$$

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

Let $x \in N^c$ and $\varepsilon > 0$ be given. We choose k so that $2^{-k} < \varepsilon$. There is a unique j (determined by x and k) such that $x \in E_{kj}$. By (5.7), we have the existence of a $\delta > 0$ such that

$$\frac{|Q \cap E_{kj}|}{|Q|} > \frac{2}{3} \qquad \forall Q \ni x \text{ such that } \ell(Q) < \delta. \tag{5.9}$$

We now check that this same δ is also good for (5.8). So fix any cube Q as in (5.8). Recalling that $E_{kj} \subseteq \{f \geq j2^{-k}\} \cap \{f \leq (j+1)2^{-k}\}$, Lemma 5.5 and (5.9) imply that

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

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

6. Local oscillations and Lerner's formula

We proceed to introduce more notation needed for Lerner's formula.

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

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

We make the following observations:

- f^* is non-increasing.
  Indeed, if $s > t$, the condition $|\{|f| > \alpha\}| \leq s$ is easier to satisfy than $|\{|f| > \alpha\}| \leq t$. So the set of admissible α's is bigger for s, and the infimum of a bigger set is smaller.
- The set inside the infimum is of the form $[\alpha_0, \infty)$ (or $\varnothing$). Hence the infimum is reached as a minimum; in particular, $f^*(t)$ itself is an admissible value of α, so that

$$|\{|f| > f^*(t)\}| \leq t. \tag{6.1}$$

  Indeed, if the set is nonempty and α belongs to this set, then every $\alpha' > \alpha$ satisfies

$$|\{|f| > \alpha'\}| \leq |\{|f| > \alpha\}| \leq t,$$

  so also α' belongs to the set. So it remains to show that the infimum α_0 also belongs to the set. This follows from $\{|f| > \alpha_0\} = \bigcup_{j=1}^{\infty}\{|f| > \alpha_0 + j^{-1}\}$ and the monotonicity of the measure,

$$|\{|f| > \alpha_0\}| = \lim_{j\to\infty} |\{|f| > \alpha_0 + j^{-1}\}| \leq t.$$

- We have $(f1_Q)^*(t) = \inf\{\alpha \geq 0 : |Q \cap \{|f| > \alpha\}| \leq t\}$.
  It suffices to check that $Q \cap \{|f| > \alpha\} = \{|1_Q f| > \alpha\}$. But this is easy to see.

Exercise 6.2. Prove the following alternative formula for the decreasing rearrangement:

$$f^*(t) = \inf_{|E| \leq t} \|1_{E^c} f\|_\infty,$$

where the infimum is taken over all measurable sets with $|E| \leq t$. Show that the infimum is reached as a minimum with the set $E = \{|f| > f^*(t)\}$.

Exercise 6.3. Let f, g, h be measurable functions and c a constant. Prove the pointwise bounds and identities

$$|f| \leq |h| \quad \Rightarrow \quad f^*(t) \leq h^*(t),$$

$$(f+g)^*(t) \leq f^*(\lambda t) + g^*((1-\lambda)t), \qquad t \in [0, \infty),\ \lambda \in [0,1].$$

$$(f+c)^*(t) \leq f^*(t) + |c|.$$

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

Lemma 6.4. *The following estimate holds for all $\lambda \in (0, \frac{1}{2})$ and all medians $m_f(Q)$:*

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

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

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

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

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

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

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

Remark 6.5. The limiting case $\lambda = \frac{1}{2}$ of the previous estimate is more tricky. It is only true that *there exists* a median $m_f(Q)$ such that $|m_f(Q)| \leq (f1_Q)^*(\frac{1}{2}|Q|)$, but this need not be the case for all medians. Therefore we prefer to work with the more flexible estimate with $\lambda < \frac{1}{2}$, where we do not need to specify the choice of the median which we work with.

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

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

$$\omega_\lambda(f;Q) := \inf_{c\in\mathbb{R}} \big((f-c)1_Q\big)^*(\lambda|Q|).$$

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

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

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

Lemma 6.6 (Quasiminimizer lemma). *For $\lambda \in (0, \frac{1}{2})$, we have*

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

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

$$m_{f-c}(Q) = m_f(Q) - c \tag{6.7}$$

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

Proof of Lemma 6.6. By triangle inequality, (6.7) and Lemma 6.4, we have

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

(Here, given a median $m_f(Q)$ of f, we have that $m_f(Q)-c$ is a median of $f-c$, and it is important that the bound of Lemma 6.4 holds for all these medians; this is ensured by $\lambda < \frac{1}{2}$.)

From the results of Exercise 6.3 it then follows that

$$(1_Q(f-m_f(Q)))^*(t) \leq (1_Q(f-c))^*(t) + \big((f-c)1_Q\big)^*(\lambda|Q|),$$

and substituting $t = \lambda|Q|$ gives the claim. □

6.C. **Lerner's formula.** Now we are fully prepared for our next main result:

Theorem 6.8. *For any measurable function f on a cube $Q^0 \subset \mathbb{R}^d$, we have*

$$|f(x) - m_f(Q^0)| \leq 2 \sum_{L\in\mathscr{L}} \omega_\lambda(f;L)1_L(x), \qquad \lambda = 2^{-d-2},$$

where $\mathscr{L} \subset \mathscr{D}(Q^0)$ is sparse*: there are pairwise disjoint* major subsets $E(L) \subset L$ *with* $|E(L)| \geq \gamma|L|$. *In fact, we can take $\gamma = \frac{1}{2}$.*

Here $\mathscr{D}(Q^0)$ means the dyadic subcubes of Q^0 *relative to* Q^0: i.e., the subcubes obtained by repeatedly bisecting each side of Q^0 into halves.

Although we still need to prove Theorem 6.8, let us already look at a motivating application:

Exercise 6.9. (Of course, you are allowed to use Lerner's formula!) Let $f : Q^0 \to \mathbb{R}$ (where $Q^0 \subset \mathbb{R}^d$) be a function for which the local oscillations are uniformly bounded: $\omega_\lambda(f;Q) \leq K$ for all $Q \subseteq Q^0$ (with $\lambda = 2^{-2-d}$). Show that in this case f is *exponentially integrable* on Q^0: for some constants $\epsilon > 0$ and $C < \infty$ (try to get some estimates for their size; these should depend on the dimension d and the oscillatory bound K),

$$\frac{1}{|Q^0|}\int_{Q^0} \exp\Big(\epsilon|f(x) - m_f(Q)|\Big)\,\mathrm{d}x \leq C. \tag{6.10}$$

(Hint: Note that $\sum_{L\in\mathscr{L}} 1_L(x)$ is counts the number of cubes $L \in \mathscr{L}$ that meet the point x. For each $k = 1, 2, \ldots$, estimate the size of the set, where this number is equal to k.) Observe that the local oscillation bound follows in particular if we assume that $f \in L^1(Q^0)$ and

$$\frac{1}{|Q|}\int_Q |f(x) - \langle f\rangle_Q|\,\mathrm{d}x \leq K', \qquad \langle f\rangle_Q := \frac{1}{|Q|}\int_Q f\,\mathrm{d}y \qquad \forall Q \subseteq Q^0. \tag{6.11}$$

((6.11) is known as the BMO condition. The implication (6.11) $\Rightarrow$ (6.10) is known as the John–Nirenberg inequality.)

Exercise 6.12. Let $\mathscr{L}$ be a sparse collection of dyadic cubes with parameter γ, and let $a_L \geq 0$ be nonnegative constants. Consider the functions $\phi := \sum_{L\in\mathscr{L}} a_L 1_L$ and $\Phi := \sup_{L\in\mathscr{L}} a_L 1_L$. For another function $g \geq 0$, prove that

$$\int \phi \cdot g\,\mathrm{d}x \leq c_\gamma \int \Phi \cdot M_d g\,\mathrm{d}x,$$

where c_γ depends only on γ (how?). Deduce (by the duality of L^p spaces) that $\|\phi\|_p \leq c_{\gamma p}\|\Phi\|_p$ for $p \in (1,\infty)$. Denoting

$$M^{\#}_{\lambda,Q^0} f(x) := \sup_{Q\in\mathscr{D}(Q^0)} 1_Q(x)\omega_\lambda(f;Q),$$

conclude from Lerner's formula that

$$\|1_{Q^0}(f - m_f(Q^0))\|_p \leq c_p\|M^{\#}_{\lambda,Q^0} f\|_p, \qquad \lambda = 2^{-2-d}, \quad p \in (1,\infty).$$

The first step of the proof of Lerner's formula is as follows:

Lemma 6.13. *For any measurable function f on a cube $Q^0 \subset \mathbb{R}^d$, let $Q^1_j \in \mathscr{D}(Q^0)$ be the maximal dyadic subcubes of Q^0 such that*

$$\max_{Q'\in\mathrm{ch}(Q^1_j)} |m_f(Q') - m_f(Q^0)| > (1_{Q^0}(f - m_f(Q^0)))^*(\lambda|Q^0|), \tag{6.14}$$

where $\mathrm{ch}(Q) := \{Q' \in \mathscr{D}(Q) : \ell(Q') = \frac{1}{2}\ell(Q)\}$ is the collection of dyadic children of Q. Then

$$1_{Q^0}(x)|f(x) - m_f(Q^0)| \leq 1_{Q^0}(x)\tilde{\omega}_\lambda(f;Q^0) + \sum_j 1_{Q^1_j}(x)|f(x) - m_f(Q^1_j)|. \tag{6.15}$$

Proof. For any family of pairwise disjoint subcubes Q^1_j of Q^0, we can write the decomposition

$$\begin{aligned} 1_{Q^0}(f - m_f(Q^0)) &= 1_{Q^0\setminus\bigcup Q^1_j}(f - m_f(Q^0)) + \sum_j 1_{Q^1_j}(m_f(Q^1_j) - m_f(Q^0)) \\ &\quad + \sum_j 1_{Q^1_j}(f - m_f(Q^1_j)) =: I + II + III. \end{aligned} \tag{6.16}$$

We apply this with the cubes Q^1_j defined as in the statement of the lemma. From the maximality it follows that Q^1_j in place of $Q' \in \mathrm{ch}(Q^1_j)$ satisfies the opposite estimate, namely

$$|m_f(Q) - m_f(Q^0)| \leq (1_{Q^0}(f - m_f(Q^0)))^*(\lambda|Q^0|) \leq \tilde{\omega}_\lambda(f;Q^0) \tag{6.17}$$

for $Q = Q^1_j$, and hence the term II in (6.16) is dominated by

$$|II| \leq 1_{\bigcup Q^1_j}\tilde{\omega}_\lambda(f;Q^0).$$

On the other hand, if $x \in Q^0 \setminus \bigcup Q_j^1$, then the estimate (6.17) holds for all dyadic $Q \ni x$. By the result of Fujii (Proposition 5.6) we know that $m_f(Q) \to f(x)$ as $Q \to x$ for almost every x, and hence also the term I in (6.16) satisfies

$$|I| \leq 1_{Q^0 \setminus \bigcup Q_j^1} \tilde{\omega}_\lambda(f; Q^0).$$

Substituting the bounds for I and II into (6.16), observing that $1_{Q^0 \setminus \bigcup Q_j^1} + 1_{\bigcup Q_j^1} = 1_{Q^0}$, we arrive at the claim of the lemma.

Altogether, we find that

$$|1_{Q^0}(f - m_f(Q^0))| \leq 1_{Q^0} \cdot 2\omega_\lambda(f; Q^0) + \sum_j |1_{Q_j^1}(f - m_f(Q_j^1))|,$$

where the terms in the sum are of the same form as the left side, with Q^0 replaced by Q_j^1, and we are in a position to iterate. □

Observe that in (6.15) the terms in the sum on the right have exactly the same form as the expression on the left. This gives us a change to iterate the decomposition.

Lemma 6.18. *The maximal dyadic subcubes $Q_j^1 \subseteq Q^0$ with* (6.14) *have the estimate*

$$\sum_j |Q_j^1| \leq \frac{1}{2}|Q^0|.$$

Proof. We abbreviate $f_0 := f - m_f(Q^0)$. Then the stopping condition gives for some $Q' \in \mathrm{ch}(Q_j^1)$ and any $\nu \in (0, \frac{1}{2})$ the estimate

$$\alpha := (1_{Q^0} f_0)^*(\lambda|Q^0|) < |m_{f_0}(Q')| \leq (1_{Q'} f_0)^*(\nu|Q'|) \leq (1_{Q_j^1} f_0)^*(\nu 2^{-d}|Q_j^1|).$$

Thus $|Q_j^1 \cap \{|f_0| > \alpha\}| \geq \nu 2^{-d}|Q_j^1|$, and hence

$$\nu 2^{-d} \sum_j |Q_j^1| \leq \sum_j |Q_j^1 \cap \{|f_0| > \alpha\}| \leq |Q^0 \cap \{|f_0| > \alpha\}| \leq \lambda|Q^0| = 2^{-d-2}|Q^0|.$$

Letting $\nu \to \frac{1}{2}$, we get $\sum_j |Q_j^1| \leq \frac{1}{2}|Q^0|$, as claimed. □

Proof of Theorem 6.8 (Lerner's formula). A combination of the two lemmas shows that, given Q^0, there exist disjoint dyadic subcubes Q_j^1 such that

$$1_{Q^0}|f - m_f(Q^0)| \leq 1_{Q^0} \tilde{\omega}_\lambda(f, Q^0) + \sum_j 1_{Q_j^1}|f - m_f(Q_j^1)|, \qquad \sum_j |Q_j^1| \leq \frac{1}{2}|Q^0|.$$

We apply the same result to each Q_j^1 in place of Q^0, yielding further subcubes Q_i^2 such that

$$1_{Q^0}|f - m_f(Q^0)| \leq 1_{Q^0} \tilde{\omega}_\lambda(f, Q^0) + \sum_j 1_{Q_j^1} \tilde{\omega}_\lambda(f; Q_j^1) + \sum_i 1_{Q_i^2}|f - m_f(Q_i^2)|,$$

where

$$\sum_{i: Q_i^2 \subseteq Q_j^1} |Q_i^2| \leq \frac{1}{2}|Q_j^1|, \qquad \sum_i |Q_i^2| = \sum_j \sum_{i: Q_i^2 \subseteq Q_j^1} |Q_i^2| \leq \frac{1}{2} \sum_j |Q_j^1| \leq \frac{1}{4}|Q^0|.$$

By induction, we obtain consecutive generations of stopping cubes Q_j^n such that for every N,

$$1_{Q^0}|f - m_f(Q^0)| \leq \sum_{n=0}^{N-1} \sum_j 1_{Q_j^n} \tilde{\omega}_\lambda(f, Q_j^n) + \sum_i 1_{Q_i^N}|f - m_f(Q_i^N)|, \tag{6.19}$$

where

$$\sum_{i: Q_i^N \subseteq Q_j^{N-1}} |Q_i^N| \leq \frac{1}{2}|Q_j^{N-1}|, \qquad \sum_i |Q_i^N| \leq 2^{-N}|Q^0|. \tag{6.20}$$

The "error term" (i.e., the last term) in (6.19) is supported on $\Omega^N := \bigcup_i Q_i^N$ whose measure is bounded by $|\Omega^N| \leq 2^{-N}|Q^0|$. Note also that $\Omega^N \subseteq \Omega^{N-1}$ by the construction. If we pass to the limit $N \to \infty$ in (6.19), we see that the limit of the error term is supported on the intersection $\Omega^\infty = \bigcap_{N=1}^\infty \Omega^N$. Since $|\Omega^\infty| \le |\Omega^N| \le 2^{-N}|Q^0|$ for any N, this has measure zero. Thus, in the limit (and by the Quasiminimizer Lemma), we deduce that

$$1_{Q^0}|f - m_f(Q^0)| \le \sum_{n=0}^\infty \sum_j 1_{Q_j^n} \tilde{\omega}_\lambda(f, Q_j^n) \le 2 \sum_{n=0}^\infty \sum_j 1_{Q_j^n} \omega_\lambda(f, Q_j^n) \qquad \text{a.e.}$$

It only remains to check that the collection $\mathscr{L} := \{Q_j^n\}_{n,j}$ is sparse. But this is immediate from (6.20), since it shows that

$$E(Q_j^n) := Q_j^n \setminus \bigcup_{i: Q_i^{n+1} \subseteq Q_j^n} Q_i^{n+1}$$

satisfies $|E(Q_j^n)| \ge \frac{1}{2}|Q_j^n|$, and clearly these sets are pairwise disjoint. $\square$

Remark 6.21. Our proof of Lerner's formula actually provided a slightly stronger form of sparseness; namely, the disjoint parts $E(L)$ are not just arbitrary subsets of L, but they are given by

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

It is not automatic from the definition of sparseness that the sets $E(L)$ can be chosen to have this form. For instance, the collection $\mathscr{L} = \{[0,1), [0,\frac{1}{2}), [\frac{1}{2},1)\}$ is $\frac{1}{2}$-sparse (the disjoint parts may be taken to be $\{[\frac{1}{4},\frac{3}{4}), [0,\frac{1}{4}), [\frac{3}{4},1)\}$), but $\dot{E}([0,1)) = [0,1) \setminus ([0,\frac{1}{2}) \cup [\frac{1}{2},1)) = \varnothing$.

We say that $\mathscr{L} \subset \mathscr{D}$ is strongly γ-sparse, if $|\dot{E}(L)| \ge \gamma|L|$ for all $L \in \mathscr{L}$. We say that $\mathscr{L}$ satisfies the *Carleson condition* with Carleson constant K if

$$\sum_{\substack{L' \subseteq L \\ L' \in \mathscr{L}}} |L'| \le K|L| \qquad \forall L \in \mathscr{L}.$$

Note that $K \ge 1$, since the sum contains at least the term $|L|$.

Exercise 6.22. Prove that if $\mathscr{L}$ is γ-sparse, then it satisfies the Carleson condition with $K \le 1/\gamma$. Prove that this bound is optimal in the following sense: for every $\gamma \in (0,1]$, give an example of a collection $\mathscr{L}$ that is (even strongly) γ-sparse and that Carleson constant exactly $K = 1/\gamma$. (Hint for the example: think about the Cantor set.)

Exercise 6.23. Let $\mathscr{L}$ be a collection that satisfies the Carleson condition with constant K, and let all $L \in \mathscr{L}$ be contained in a single maximal $L_0 \in \mathscr{L}$. Prove that, for a finite N depending only on K, there is a disjoint decomposition $\mathscr{L} = \bigcup_{n=0}^N \mathscr{L}_n \cup \mathscr{L}'$, where each $\mathscr{L}_n$ is pairwise disjoint, and

$$\Big| \bigcup_{L' \in \mathscr{L}'} L' \Big| \le \frac{1}{2}|L|.$$

Hint: Let $\mathscr{L}_0 := \{L\}$, and set first $\mathscr{L}' := \mathscr{L} \setminus \mathscr{L}_0$ (this might be redefined later). Then (why?)

$$\theta := \frac{1}{|L|}\Big| \bigcup_{L' \in \mathscr{L}'} L' \Big| \le \frac{1}{|L|} \sum_{L' \in \mathscr{L}'} |L'| \le K - 1. \tag{6.24}$$

If $\theta \le \frac{1}{2}$, we are done. If $\theta > \frac{1}{2}$, let $\mathscr{L}_1$ consist of all maximal elements in $\mathscr{L}'$, and remove $\mathscr{L}_1$ from $\mathscr{L}'$. Check that with this new $\mathscr{L}'$, the bound (6.24) holds with $K - \frac{3}{2}$ in place of $K - 1$. Consider again the options that $\theta \le \frac{1}{2}$ and $\theta > \frac{1}{2}$, and continue similarly. Write down the induction step and show that the process must terminate (with $\theta \le \frac{1}{2}$) after boundedly many steps, and give an estimate for this number. Show that this gives the required decomposition.

Exercise 6.25. Let $\mathscr{L} \subset \mathscr{D}$ be without infinite increasing chains (so that each $L \in \mathscr{L}$ is contained in some maximal $L^* \in \mathscr{L}^*$). If $\mathscr{L}$ satisfies the Carleson condition, prove that $\mathscr{L}$ can be divided into a bounded number of subcollections that are strongly $\frac{1}{2}$-sparse. (Hint: Apply the result of the previous exercise recursively, starting from the maximal cubes $L^* \in \mathscr{L}^*$.)

7. Calderón–Zygmund operators

Definition 7.1. A *Calderón–Zygmund operator* (CZO) is a bounded linear operator $T : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ with the integral representation

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

where $K : \mathbb{R}^d \times \mathbb{R}^d \setminus \{(x,x) : x \in \mathbb{R}^d\}$ is called the *kernel* and satisfies the *standard estimates*

$$|K(x,y)| \leq \frac{C}{|x-y|^d}, \tag{7.2}$$

$$|K(x,y) - K(x',y)| + |K(y,x) - K(y,x')| \leq C\frac{|x-x'|^\alpha}{|x-y|^{d+\alpha}}, \quad \text{if} \quad |x-x'| < \frac{1}{2}|x-y|, \tag{7.3}$$

where $\alpha \in (0,1]$ is the regularity exponent.

Lemma 7.4. *The conditions* (7.2) *and* (7.3) *are equivalent to* (7.2) *and* (7.5)*, where the last condition is defined by*

$$|K(x,y) - K(x',y)| + |K(y,x) - K(y,x')| \leq \frac{C|x-x'|^\alpha}{\min(|x-y|, |x'-y|)^{d+\alpha}}. \tag{7.5}$$

Note that (7.5) is assumed for all x, x', y, in contrast to (7.3), which is only assumed for certain triplets x, x', y.

Proof. We consider two cases:

Case $|x-x'| < \frac{1}{2}|x-y|$. Then $|x'-y| \leq |x-y| + |x'-x| < \frac{3}{2}|x-y|$ and $|x'-y| \geq |x-y| - |x'-x| > \frac{1}{2}|x-y|$, so that $|x-y| \eqsim |x'-y|$ and in particular $\min(|x-y|, |x'-y|) \eqsim |x-y|$. Thus the estimates (7.3) and (7.5) are exactly the same estimate in this case.

Case $|x-x'| \geq \frac{1}{2}|x-y|$. Then $|x'-y| \leq |x'-x| + |x-y| \leq 3|x-x'|$, so that $|x-x'|$ dominates both $|x-y|$ and $|x'-y|$ (up to numerical constants) in this case. Since (7.3) says nothing now, we need to show that (7.5) follows from (7.2). Indeed, by (7.2),

$$\begin{aligned} |K(x,y) - K(x',y)| + |K(y,x) - K(y,x')| &\leq |K(x,y)| + |K(x',y)| + |K(y,x)| + |K(y,x')| \\ &\leq \frac{C}{|x-y|^d} + \frac{C}{|x'-y|^d}. \end{aligned}$$

Moreover, since $|x-x'|/|x-y| \geq \frac{1}{2}$, we have

$$\frac{C}{|x-y|^d} \leq \frac{C}{|x-y|^d}\Big(\frac{|x-x'|}{|x-y|}\Big)^\alpha = C\frac{|x-x'|^\alpha}{|x-y|^{d+\alpha}}$$

and similarly $C/|x'-y|^d \leq C|x'-x|^\alpha/|x'-y|^{d+\alpha}$. Finally, observe that

$$\frac{1}{|x-y|^{d+\alpha}} + \frac{1}{|x'-y|^{d+\alpha}} \leq 2\max\Big(\frac{1}{|x-y|^{d+\alpha}}, \frac{1}{|x'-y|^{d+\alpha}}\Big) = \frac{2}{\min(|x-y|, |x'-y|)^{d+\alpha}},$$

which completes the proof. □

Remark 7.6. A typical case, where we use the regularity of the kernel, is as follows. Let Q be a cube, $x, x' \in Q$ and $y \in (2Q)^c$. (We denote by $2Q$ the cube with the same centre and double the sidelength of Q.) In this case $|x-x'| \leq \sqrt{d}\ell(Q)$ (which is the length of the diagonal of Q), whereas $|x-y|, |x'-y| \leq \frac{1}{2}\ell(Q)$, where equality can be (almost) achieved, so that the condition of (7.3), that $|x-x'| < \frac{1}{2}|x-y|$, is not necessarily guaranteed. However, the condition (7.5) can be conveniently applied. Notice that

$$|x-y| \leq |x-x'| + |x'-y| \leq \sqrt{d}\ell(Q) + |x'-y| \leq (2\sqrt{d}+1)|x'-y|$$

and by symmetry also $|x'-y| \leq (2\sqrt{d}+1)|x-y|$ so in fact $|x-y| \eqsim |x'-y|$, up to numerical constants. Thus we find that

$$|K(x,y) - K(x',y)| + |K(y,x) - K(y,x')| \leq C\frac{\ell(Q)^\alpha}{|x'-y|^{d+\alpha}}, \quad \text{if} \quad x, x' \in Q, \quad y \in (2Q)^c. \tag{7.7}$$

Lemma 7.8. *Let b be a function with $\operatorname{spt} b \subseteq Q$ and $\int_Q b \,\mathrm{d}x = 0$. If T is a Calderón–Zygmund operator, then*

$$\int_{(2Q)^c} |Tb| \,\mathrm{d}y \leq C\|b\|_1.$$

Proof. We can write

$$\begin{aligned}\int_{(2Q)^c} |Tb(y)| \,\mathrm{d}y &= \int_{(2Q)^c} \Big| \int_Q K(y,x)b(x) \,\mathrm{d}x \Big| \,\mathrm{d}y \\ &= \int_{(2Q)^c} \Big| \int_Q [K(y,x) - K(y,c_Q)]b(x) \,\mathrm{d}x \Big| \,\mathrm{d}y,\end{aligned}$$

where c_Q is the centre of Q, and we used the observation that

$$\int_Q K(y,c_Q)b(x) \,\mathrm{d}x = K(y,c_Q) \int_Q b(x) \,\mathrm{d}x = 0.$$

Using (7.7) with $x' = c_Q$, we find that

$$\int_Q |K(y,x) - K(y,c_Q)||b(x)| \,\mathrm{d}x \leq C \frac{\ell(Q)^\alpha}{|y - c_Q|^{d+\alpha}} \int_Q |b(x)| \,\mathrm{d}x.$$

It remains to integrate

$$\int_{(2Q)^c} \frac{\mathrm{d}y}{|y - c_Q|^{d+\alpha}} \leq C \int_{\ell(Q)}^\infty \frac{r^{d-1} \,\mathrm{d}r}{r^{d+\alpha}} = C \int_{\ell(Q)}^\infty r^{-1-\alpha} \,\mathrm{d}r = C \Big|_{\ell(Q)}^\infty \frac{r^{-\alpha}}{-\alpha} = \frac{C}{\alpha} \ell(Q)^{-\alpha},$$

where we changed to polar coordinates centred at c_Q, and observed that $(2Q)^c \subseteq B(c_Q, \ell(Q))^c$.

A combination of these estimates proves the claim. □

Exercise 7.9. Let T be an operator whose kernel $K(x,y)$ satisfies *Hörmander's condition*

$$\int_{(2Q)^c} |K(y,x) - K(y,x')| \,\mathrm{d}y \leq C \quad \forall x,x' \in Q, \ \forall \text{ cubes } Q.$$

Prove that

(1) The conclusion of Lemma 7.8 is valid for all operators with Hörmander's condition.
(2) All Calderón–Zygmund operators satisfy Hörmander's condition.

Theorem 7.10 (Calderón–Zygmund)**.** *If T is a Calderón–Zygmund operator, then $T : L^1(\mathbb{R}^d) \to L^{1,\infty}(\mathbb{R}^d)$ boundedly.*

The proof is based on:

Proposition 7.11 (Calderón–Zygmund decomposition)**.** *Let $f \in L^1(\mathbb{R}^d)$ and $\lambda > 0$. Then there is a decomposition $f = g + b$ ("good + bad"), where*

- *The good part satisfies $\|g\|_\infty \leq 2^d\lambda$ and $\|g\|_1 \leq \|f\|_1$.*
- *The bad part is $b = \sum_{Q\in\mathscr{Q}} b_Q$, where $\mathscr{Q}$ is a disjoint collection of dyadic cubes such that*

$$\sum_{Q\in\mathscr{Q}} |Q| \leq \frac{1}{\lambda}\|f\|_1, \quad \operatorname{spt} b_Q \subseteq Q, \quad \int b_Q \,\mathrm{d}x = 0, \quad \int |b_Q| \,\mathrm{d}x \leq 2 \int_Q |f| \,\mathrm{d}x.$$

Proof. Let $\mathscr{Q}$ be the maximal dyadic cubes Q with the property that $|Q|^{-1} \int_Q |f| \,\mathrm{d}x > \lambda$. Then they are pairwise disjoint and their union is

$$\Omega := \bigcup_{Q\in\mathscr{Q}} Q = \{M_d f > \lambda\}.$$

From the weak-type inequality for the maximal function it follows that

$$|\Omega| = \sum_{Q\in\mathscr{Q}} |Q| = |\{M_d f > \lambda\}| \leq \frac{1}{\lambda}\|f\|_1.$$

The good part. We define

$$g := 1_{\Omega^c} f + \sum_{Q\in\mathscr{Q}} 1_Q \langle f\rangle_Q, \qquad \langle f\rangle_Q := \frac{1}{|Q|}\int_Q f\,\mathrm{d}x.$$

For almost every $x \in \Omega^c$, we have

$$|g(x)| = |f(x)| = \lim_{\substack{Q\in\mathscr{D},\ Q\ni x\\ \ell(Q)\to 0}} \Big|\frac{1}{|Q|}\int_Q f\,\mathrm{d}y\Big| \leq \sup_{\substack{Q\in\mathscr{D}\\ Q\ni x}} \Big|\frac{1}{|Q|}\int_Q f\,\mathrm{d}y\Big| \leq M_d f(x) \leq \lambda,$$

since $\Omega^c = \{M_d f \leq \lambda\}$.

If $x \in \Omega$, then $x \in Q$ for exactly one $Q \in \mathscr{Q}$, and then $g(x) = \langle f\rangle_Q$. Since Q is maximal with the property $\langle |f|\rangle_Q > \lambda$, its parent dyadic cube $Q^{(1)}$ satisfies the opposite inequality $\langle |f|\rangle_{Q^{(1)}} \leq \lambda$, and hence

$$|\langle f\rangle_Q| \leq \langle |f|\rangle_Q = \frac{1}{|Q|}\int_Q |f|\,\mathrm{d}y \leq \frac{2^d}{|Q^{(1)}|}\int_{Q^{(1)}} |f|\,\mathrm{d}y \leq 2^d\lambda.$$

Thus $|g(x)| = |\langle f\rangle_Q| \leq 2^d\lambda$ also for $x \in \Omega$.

For the L^1 bound, note that

$$\int_{\mathbb{R}^d} |g|\,\mathrm{d}x = \int_{\Omega^c} |f|\,\mathrm{d}x + \sum_{Q\in\mathscr{Q}} |Q||\langle f\rangle_Q| \leq \int_{\Omega^c} |f|\,\mathrm{d}x + \sum_{Q\in\mathscr{Q}} \int_Q |f|\,\mathrm{d}x = \int_{\mathbb{R}^d} |f|\,\mathrm{d}x.$$

The bad part. To have $f = g + b$, this has to be defined by

$$b := \sum_{Q\in\mathscr{Q}} 1_Q (f - \langle f\rangle_Q) =: \sum_{Q\in\mathscr{Q}} b_Q.$$

The required properties of b_Q, namely $\operatorname{spt} b_Q \subseteq Q$, $\int b_Q\,\mathrm{d}x = 0$ and $\|b_Q\|_1 \leq \|1_Q f\|_1$, are immediate to check. □

Proof of Theorem 7.10. Given $f \in L^1(\mathbb{R}^d)$ and $\lambda > 0$, we need to estimate $|\{|Tf| > \lambda\}|$. We write the Calderón–Zygmund decomposition $f = g + b$ and observe that

$$\begin{aligned} |\{|Tf| > \lambda\}| &\leq |\{|Tg| > \tfrac{1}{2}\lambda\}| + |\{|Tb| > \tfrac{1}{2}\lambda\}| \\ &\leq |\{|Tg| > \tfrac{1}{2}\lambda\}| + |2\Omega| + |\{|Tb| > \tfrac{1}{2}\lambda\} \setminus 2\Omega| =: I + II + III, \end{aligned} \tag{7.12}$$

where $2\Omega := \bigcup_{Q\in\mathscr{Q}} 2Q$ for the cubes appearing in the Calderón–Zygmund decomposition.

Term I. This uses the L^2-bound of T. We have

$$\begin{aligned} |\{|Tg| > \tfrac{1}{2}\lambda\}| &= \int_{\{|Tg|>\frac{1}{2}\lambda\}} \mathrm{d}x \leq \int \Big(\frac{|Tg|}{\frac{1}{2}\lambda}\Big)^2 \mathrm{d}x = \frac{4}{\lambda^2}\int |Tg|^2\,\mathrm{d}x \leq \frac{C}{\lambda^2}\int |g|^2\,\mathrm{d}x \\ &\leq \frac{C}{\lambda^2}\int \lambda|g|\,\mathrm{d}x = \frac{C}{\lambda}\|g\|_1 \leq \frac{C}{\lambda}\|f\|_1. \end{aligned}$$

Term II. This uses only the properties of the Calderón–Zygmund cubes:

$$|2\Omega| \leq \sum_{Q\in\mathscr{Q}} |2Q| = \sum_{Q\in\mathscr{Q}} 2^d|Q| \leq \frac{2^d}{\lambda}\|f\|_1.$$

Term III. This is the most difficult part, which uses the kernel bounds for the Calderón–Zygmund operator. We begin with

$$\begin{aligned} |\{|Tb| > \tfrac{1}{2}\lambda\} \setminus 2\Omega| &= \int_{\{|Tb|>\frac{1}{2}\lambda\}\setminus 2\Omega} \mathrm{d}x \leq \int_{(2\Omega)^c} \frac{|Tb|}{\frac{1}{2}\lambda}\,\mathrm{d}x = \frac{2}{\lambda}\int_{(2\Omega)^c} |Tb|\,\mathrm{d}x \\ &\leq \frac{2}{\lambda}\sum_{Q\in\mathscr{Q}} \int_{(2\Omega)^c} |Tb_Q|\,\mathrm{d}x \leq \frac{2}{\lambda}\sum_{Q\in\mathscr{Q}} \int_{(2Q)^c} |Tb_Q|\,\mathrm{d}x, \end{aligned}$$

where in the last step we observed that $2Q \subseteq 2\Omega$ implies $(2\Omega)^c \subseteq (2Q)^c$. Lemma 7.8 shows that $\int_{(2Q)^c} |Tb_Q| \,\mathrm{d}x \le C\|b_Q\|_1$, and then by the properties of the bad part we conclude that

$$|\{|Tb| > \tfrac{1}{2}\lambda\} \setminus 2\Omega| \le \frac{C}{\lambda} \sum_{Q\in\mathscr{Q}} \|b_Q\|_1 \le \frac{C}{\lambda}\|f\|_1.$$

We have shown that all three terms in (7.12) have an upper bound $C\lambda^{-1}\|f\|_1$, and this is what we had to prove. □

The next result about the oscillations of a Calderón–Zygmund operator connects the topics of this and the previous section.

Proposition 7.13 (Jawerth–Torchinsky [JT85])**.**

$$\omega_\lambda(Tf, Q) \le C \sum_{k=1}^{\infty} 2^{-k\alpha} \frac{1}{|2^k Q|} \int_{2^k Q} |f| \,\mathrm{d}x.$$

For the proof, we need a simple relation of the decreasing rearrangement and the $L^{1,\infty}$ norm:

Lemma 7.14.

$$f^*(t) \le \frac{1}{t}\|f\|_{L^{1,\infty}}.$$

Proof. If $|\{|f| > \alpha\}| \le t$, then $f^*(t)$ (which is the infimum over all such α) is at most α. So it suffices to check that with $\alpha = t^{-1}\|f\|_{L^{1,\infty}}$, we have

$$|\{|f| > \alpha\}| \le \frac{1}{\alpha}\|f\|_{L^{1,\infty}} = \frac{t}{\|f\|_{L^{1,\infty}}}\|f\|_{L^{1,\infty}} = t. \qquad \square$$

Exercise 7.15. For any $p \in (0,\infty)$, the weak L^p "norm" is $\|f\|_{L^{p,\infty}} := \sup_{t>0} t \cdot |\{|f| > t\}|^{1/p}$. Prove that $\|f\|_{L^{p,\infty}} \le \|f\|_{L^p}$, and estimate $f^*(t)$ in terms of $\|f\|_{L^{p,\infty}}$ for any $p \in (0,\infty)$.

Proof of Proposition 7.13. For $x \in Q$, we have

$$Tf(x) = T(1_{2Q}f)(x) + [T(1_{(2Q)^c}f)(x) - T(1_{(2Q)^c}f)(c_Q)] + \text{const},$$

where we subtracted and added the same constant const $= T(1_{(2Q)^c}f)(c_Q)$. Hence

$$|1_Q(x)(Tf(x) - \text{const})| \le |T(1_{2Q}f)(x)| + \|1_Q[T(1_{(2Q)^c}f)(\cdot) - T(1_{(2Q)^c}f)(c_Q)]\|_\infty,$$

and therefore (using Exercise 6.3)

$$\begin{aligned}\omega_\lambda(Tf,Q) &= \inf_c (1_Q(Tf - c))^*(\lambda|Q|) \\ &\le (T(1_{2Q}f))^*(\lambda|Q|) + \|1_Q[T(1_{(2Q)^c}f)(\cdot) - T(1_{(2Q)^c}f)(c_Q)]\|_\infty = I + II.\end{aligned}$$

By Lemma 7.14 ($f^*(t) \le t^{-1}\|f\|_{L^{1,\infty}}$) and Theorem 7.10 ($T : L^1 \to L^{1,\infty}$), the first term is estimated by

$$I = (T(1_{2Q}f))^*(\lambda|Q|) \le \frac{1}{\lambda|Q|}\|T(1_{2Q}f)^*\|_{L^{1,\infty}} \le \frac{C}{|Q|}\|1_{2Q}f\|_{L^1} \le \frac{C}{|2Q|}\int_{2Q}|f| \,\mathrm{d}x.$$

The argument for the second term II is similar to the proof of Lemma 7.8: For $x \in Q$,

$$\begin{aligned}|T(1_{(2Q)^c}f)(x) - T(1_{(2Q)^c}f)(c_Q)| &= \Big|\int_{(2Q)^c} K(x,y)f(y)\,\mathrm{d}y - \int_{(2Q)^c} K(c_Q,y)f(y)\,\mathrm{d}y\Big| \\ &\le \int_{(2Q)^c} |K(x,y) - K(c_Q,y)||f(y)|\,\mathrm{d}y \\ &\le \int_{(2Q)^c} \frac{C\ell(Q)^\alpha}{|y - c_Q|^{d+\alpha}}|f(y)|\,\mathrm{d}y.\end{aligned}$$

We then split the integration domain $(2Q)^c$ into cubic annuli $(2Q)^c = \bigcup_{k=2}^{\infty} 2^k Q \setminus 2^{k-1}Q$, and observe that $|y - c_Q| \eqsim 2^k \ell(Q)$ for $y \in 2^k Q \setminus 2^{k-1}Q$. Hence

$$\begin{aligned} II &\leq \int_{(2Q)^c} \frac{C\ell(Q)^\alpha}{|y - c_Q|^{d+\alpha}} |f(y)| \,\mathrm{d}y \leq \sum_{k=2}^{\infty} \int_{2^k Q \setminus 2^{k-1}Q} \frac{C\ell(Q)^\alpha}{(2^k \ell(Q))^{d+\alpha}} |f(y)| \,\mathrm{d}y \\ &\leq \sum_{k=2}^{\infty} C 2^{-k\alpha} \frac{1}{|2^k Q|} \int_{2^k Q} |f(y)| \,\mathrm{d}y. \end{aligned}$$

A combination of the above bounds for I and II gives the asserted bound for $\omega_\lambda(Tf, Q) \leq I + II$. □

8. The dyadic domination theorem

We formulate the following theorem in the general set-up of a *Banach function space.* We postpone its definition for the moment, and content ourselves by pointing out that all $L^p(w)$ spaces with $p \in [1, \infty)$ and a weight $w \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ are particular examples.

Theorem 8.1 (Lerner [Ler13]). *Let T be a Calderón–Zygmund operator, $f : \mathbb{R}^d \to \mathbb{R}$ be bounded and compactly supported, and X be a Banach function space of $\mathbb{R}^d$. Then*

$$\|Tf\|_X \leq c_T \sup_{\mathscr{D}, \mathscr{S}} \|A_{\mathscr{S}} |f|\|_X,$$

where the supremum is over all 3^d dyadic systems

$$\mathscr{D} = \mathscr{D}^\alpha = \Big\{ 2^{-k}([0,1)^d + m + (-1)^k \tfrac{1}{3}\alpha) : k \in \mathbb{Z}, m \in \mathbb{Z}^d \Big\}, \quad \alpha \in \{0,1,2\}^d,$$

and all $\frac{1}{2}$-sparse subcollections $\mathscr{S} \subset \mathscr{D}$, and $A_{\mathscr{S}}$ is the averaging operator

$$A_{\mathscr{S}} f = \sum_{S \in \mathscr{S}} \frac{1_S}{|S|} \int_S f \,\mathrm{d}x.$$

The main applications of this theorem are in weighted norm inequalities. Here is, however, a simple consequence for unweighted spaces. It is already a nontrivial result, although it could be proven in many other ways as well:

Exercise 8.2. Prove that $\|A_{\mathscr{S}} f\|_p \leq C\|f\|_p$, where C depends only on $p \in (1, \infty)$ and the sparseness parameter γ, but not on the particular sparse collection $\mathscr{S}$; conclude from Theorem 8.1 that all Calderón–Zygmund operators map $T : L^p \to L^p$ for all $p \in (1, \infty)$. (Hint: Exercise 6.12.)

The proof of Theorem 8.1 begins with the following lemma. We now denote the Hölder exponent of the kernel of T by $\beta \in (0,1]$, in order not to confuse it with the parameter $\alpha \in \{0,1,2\}^d$ that indexes the different systems of cubes.

Lemma 8.3. *If $Q^0 \supset \operatorname{spt} f$ is a cube, then*

$$1_{Q^0} |Tf| \leq C \sum_{L \in \mathscr{L}} \sum_{k=0}^{\infty} 2^{-k\beta} \frac{1_L}{|2^k L|} \int_{2^k L} |f| \,\mathrm{d}x, \tag{8.4}$$

for a $\frac{1}{2}$-sparse collection $\mathscr{L}$ of dyadic subcubes of Q^0.

Proof. We apply Lerner's formula to Tf and Q^0. This gives

$$1_{Q^0} |Tf| \leq 1_{Q^0} |m_{Tf}(Q^0)| + \sum_{L \in \mathscr{L}} 2\omega_\lambda(Tf; L) 1_L,$$

where $\mathscr{L}$ is a $\frac{1}{2}$-sparse subcollection of dyadic subcubes of Q^0, with $Q^0 \in \mathscr{L}$. For the oscillatory terms, Proposition 7.13 shows that

$$\omega_\lambda(Tf; L) \leq C \sum_{k=1}^{\infty} 2^{-k\beta} \frac{1}{|2^k L|} \int_{2^k L} |f| \,\mathrm{d}x.$$

For the median term, we can estimate, for $\nu \in (0, \frac{1}{2})$,

$$|m_{Tf}(Q_0)| \leq (1_{Q_0} Tf)^*(\nu|Q^0|) \leq \frac{1}{\nu|Q^0|}\|Tf\|_{L^{1,\infty}} \leq \frac{C}{|Q^0|}\|f\|_{L^1} = \frac{C}{|Q^0|}\int_{Q^0}|f|\,\mathrm{d}x,$$

recalling that $\operatorname{spt} f \subseteq Q^0$. This is of the same form as terms bounding $\omega_\lambda(Tf; L)$, with $k = 0$ and $L = Q^0$. A combination of these bounds proves the assertion. □

We next recall Exercise 3.3, which says that for any two cubes Q, P, there exist $\alpha \in \{0,1,2\}^d$, and $R, S \in \mathscr{D}^\alpha$ (same α for both R and S!), such that

$$\begin{aligned} R \supseteq Q, \qquad & 3\ell(Q) < \ell(R) \leq 6\ell(Q), \\ S \supseteq P, \qquad & 3\ell(P) < \ell(S) \leq 6\ell(P). \end{aligned}$$

We apply this with $Q = L$ and $P = 2^k L$. Since the side-lengths of R and S are powers of 2, it follows that $\ell(S) = 2^k\ell(R)$. Since $R \cap S \supseteq Q \cap P = L \neq \varnothing$, it follows that the bigger cube must contain the smaller, i.e., $S \supseteq R$. So in fact $S = R^{(k)}$ is the kth dyadic ancestor of R. Summarizing this argument, we have:

Lemma 8.5. *For any cube L and $k \in \mathbb{N}$, there is $\alpha \in \{0,1,2\}^d$ and $R \in \mathscr{D}^\alpha$ such that $R \supseteq L$, $R^{(k)} \supseteq 2^k L$, and $\ell(R) \in (3\ell(L), 6\ell(L)]$.*

Let us denote the corresponding α and R by $\alpha(L,k)$ and $R(L,k) \in \mathscr{D}^{\alpha(L,k)}$. For this $R = R(L,k)$, we have

$$1_L \cdot \frac{1}{|2^k L|}\int_{2^k L}|f|\,\mathrm{d}x \leq 1_R \cdot \frac{6^d}{|R^{(k)}|}\int_{R^{(k)}}|f|\,\mathrm{d}x.$$

We use this estimate and reorganize the sum in (8.4) according to the value of $\alpha(L,k) \in \{0,1,2\}^d$, and then according to the value of $R(L,k) \in \mathscr{R}_k^\alpha := \{R(L,k) : L \in \mathscr{L}, \alpha(L,k) = \alpha\}$:

$$\begin{aligned} 1_{Q^0}|Tf| &\leq C\sum_{L\in\mathscr{L}}\sum_{k=0}^\infty 2^{-k\beta}\frac{1_L}{|2^k L|}\int_{2^k L}|f|\,\mathrm{d}x \\ &\leq C\sum_{\alpha\in\{0,1,2\}^d}\sum_{k=0}^\infty 2^{-k\beta}\sum_{\substack{L\in\mathscr{L}\\ \alpha(L,k)=\alpha}}\frac{1_{R(L,k)}}{|R(L,k)^{(k)}|}\int_{R(L,k)^{(k)}}|f|\,\mathrm{d}x \\ &= C\sum_{\alpha\in\{0,1,2\}^d}\sum_{k=0}^\infty 2^{-k\beta}\sum_{R\in\mathscr{R}_k^\alpha}\sum_{\substack{L\in\mathscr{L}\\ \alpha(L,k)=\alpha\\ R(L,k)=R}}\frac{1_R}{|R^{(k)}|}\int_{R^{(k)}}|f|\,\mathrm{d}x \end{aligned} \tag{8.6}$$

To proceed, we need a bound for how many different L can give the same $R = R(L,k)$:

Lemma 8.7. *For fixed α and R, we have*

$$\sum_{\substack{L\in\mathscr{L}\\ \alpha(L,k)=\alpha\\ R(L,k)=R}} 1 \leq 6^d.$$

Proof. First recall that $R(L,k) = R$ means that $\ell(L) \in [\frac{1}{6}\ell(R), \frac{1}{3}\ell(R))$, and there is a unique power of 2 in this interval. Thus all L with $R(L,k) = R$ are dyadic cubes of the same sidelength, so they are pairwise disjoint. They are also all contained in R, and each of them has measure $|L| = \ell(L)^d \geq (\frac{1}{6}\ell(R))^d = 6^{-d}|R|$. Thus

$$|R| \geq \sum_{\substack{L\in\mathscr{L}\\ \alpha(L,k)=\alpha\\ R(L,k)=R}}|L| \geq \sum_{\substack{L\in\mathscr{L}\\ \alpha(L,k)=\alpha\\ R(L,k)=R}} 6^{-d}|R|.$$

Division by $6^{-d}|R|$ gives the claim. □

Exercise 8.8. The above proof used the fact that $\ell(L) \in [\frac{1}{6}\ell(R), \frac{1}{3}\ell(R))$. Suppose that we only know that $\ell(L) \geq \frac{1}{6}\ell(R)$, so we cannot conclude that all L have the same size. Check that the conclusion of the lemma is still true, although maybe with a different number in place of 6^d.

Thus we can continue the computation in (8.6) with

$$1_{Q^0}|Tf| \leq C \sum_{\alpha\in\{0,1,2\}^d} \sum_{k=0}^{\infty} 2^{-k\beta} \sum_{R\in\mathscr{R}_k^\alpha} \frac{1_R}{|R^{(k)}|} \int_{R^{(k)}} |f| \,\mathrm{d}x. \tag{8.9}$$

Lemma 8.10. *Each collection $\mathscr{R}_k^\alpha \subset \mathscr{D}^\alpha$ is γ-sparse with $\gamma = 2^{-1}6^{-d}$.*

Proof. By definition, for every $R \in \mathscr{R}_k^\alpha$, there exists at least one $L \in \mathscr{L}$ such that $R = R(L,k)$. Let us choose one such L and denote it by $L(R)$. Note that $L(R) \neq L(R')$ if $R \neq R'$. We can then define $E(R) := E(L(R))$. These sets are pairwise disjoint (by the property of the sets $E(L)$, $L \in \mathscr{L}$), and moreover

$$|E(R)| = |E(L(R))| \geq \frac{1}{2}|L(R)| \geq \frac{1}{2}\frac{1}{6^d}|R|. \qquad \square$$

We are in a good position towards Lerner's dyadic domination theorem, but there is still work to do. At this point, the nature of the considerations changes slightly, and we begin a new section:

9. Duality and dyadic operators

The bound (8.9) with Lemma 8.10 is as far as we can proceed with a *pointwise* estimate. But recall that we are eventually interested in a norm inequality, and thus we next estimate the integral pairing

$$\Big|\int Tf\cdot g \,\mathrm{d}x\Big| \leq \int |Tf|\cdot|g|\,\mathrm{d}x,$$

where g is another bounded compactly supported function. If we take Q^0 big enough to also contain the support of g, then

$$\begin{aligned}\int |Tf|\cdot|g|\,\mathrm{d}x &= \int 1_{Q^0}|Tf|\cdot|g|\,\mathrm{d}x \\ &\leq C \sum_{\alpha\in\{0,1,2\}^d} \sum_{k=0}^{\infty} 2^{-k\beta} \sum_{R\in\mathscr{R}_k^\alpha} \frac{1}{|R^{(k)}|}\int_{R^{(k)}} |f|\,\mathrm{d}y \cdot \int_R |g|\,\mathrm{d}x \\ &= C \sum_{\alpha\in\{0,1,2\}^d} \sum_{k=0}^{\infty} 2^{-k\beta} \int |f| \Big(\sum_{R\in\mathscr{R}_k^\alpha} \frac{1_{R^{(k)}}}{|R^{(k)}|}\int_R |g|\,\mathrm{d}x\Big)\,\mathrm{d}y \\ &=: C \sum_{\alpha\in\{0,1,2\}^d} \sum_{k=0}^{\infty} 2^{-k\beta} \int |f|\cdot A_k^\alpha|g|\,\mathrm{d}y.\end{aligned} \tag{9.1}$$

We next analyse the operator A_k^α, or more generally any

$$\begin{aligned} A_{\mathscr{S},k}g &:= \sum_{S\in\mathscr{S}} \frac{1_{S^{(k)}}}{|S^{(k)}|}\int_S g\,\mathrm{d}x \\ &= \sum_{Q\in\mathscr{D}} \sum_{\substack{S\in\mathscr{S}\\ S^{(k)}=Q}} \frac{1_Q}{|Q|}\int_S g\,\mathrm{d}x = \sum_{Q\in\mathscr{D}} \frac{1_Q}{|Q|}\int_Q \eta_Q g\,\mathrm{d}x, \quad \eta_Q := \sum_{\substack{S\in\mathscr{S}\\ S^{(k)}=Q}} 1_S, \end{aligned} \tag{9.2}$$

where $\mathscr{S} \subset \mathscr{D}$ is γ-sparse, $\mathscr{D}$ is a dyadic system, and $k \in \mathbb{N}$. (Of course, A_k^α corresponds to the case $\mathscr{R}_k^\alpha \subset \mathscr{D}^\alpha$, which is $2^{-1}6^{-d}$-sparse.) Note that the $A_{\mathscr{S}}$ in the dyadic domination theorem that we are trying to prove corresponds to the case $k = 0$ in the above notation $A_{\mathscr{S},k}$. Thus we still need to do something to reduce everything to this case.

Lemma 9.3. $\|A_{\mathscr{S},k}f\|_p \leq \gamma^{-1}pp'\|f\|_p$ *for $p \in (1,\infty)$.*

Note in particular that the estimate is independent of k.

Proof. We have

$$\begin{aligned}\int A_{\mathscr{S},k}f\cdot h\,\mathrm{d}x&=\sum_{S\in\mathscr{S}}\frac{1}{|S^{(k)}|}\int_{S^{(k)}}h\,\mathrm{d}x\cdot\frac{1}{|S|}\int_S f\,\mathrm{d}x\cdot|S|\\&\leq\sum_{S\in\mathscr{S}}\inf_{y\in S^{(k)}}M_dh(y)\cdot\inf_{z\in S}M_df(z)\cdot\frac{1}{\gamma}\int_{E(S)}1\,\mathrm{d}x\\&\leq\frac{1}{\gamma}\sum_{S\in\mathscr{S}}\int_{E(S)}M_dh(x)M_df(x)\,\mathrm{d}x\\&\leq\frac{1}{\gamma}\int_{\mathbb{R}^d}M_dh(x)M_df(x)\,\mathrm{d}x\\&\leq\frac{1}{\gamma}\|M_dh\|_{p'}\|M_df\|_p\leq\frac{1}{\gamma}\cdot p\|h\|_{p'}\cdot p'\|f\|_p.\end{aligned}$$

Taking the supremum over $\|h\|_{p'}\leq 1$ and using the other direction of L^p duality, we find that $\|A_{\mathscr{S},k}f\|_p\leq pp'/\gamma$. □

Proposition 9.4. $\|A_{\mathscr{S},k}f\|_{L^{1,\infty}}\leq c(1+k)\|f\|_{L^1}$*, where* $c=c_\gamma$.

Here we introduce a dependence on k, but a reasonably moderate one. The exponential decay $2^{-k\beta}$ in (9.1) will be enough to take care of this linear growth in k.

Proof. We need to prove that

$$|\{|A_{\mathscr{S},k}f|>\lambda\}|\leq\frac{c(1+k)}{\lambda}\|f\|_1.$$

We use again the Calderón–Zygmund decomposition to write $f=g+b$, where $\|g\|_\infty\leq 2^d\lambda$ and $\|g\|_1\leq\|f\|_1$, where as the bad part satisfies

$$b=\sum_{Q\in\mathscr{Q}}b_Q,\quad b_Q=1_Q(f-\langle f\rangle_Q),\quad |\Omega|:=\Big|\bigcup_{Q\in\mathscr{Q}}Q\Big|=\sum_{Q\in\mathscr{Q}}|Q|\leq\frac{\|f\|_1}{\lambda},$$

where the cubes $Q\in\mathscr{Q}$ are pairwise disjoint cubes in the same dyadic system $\mathscr{D}$ that contains the sparse collection $\mathscr{S}$.

We estimate

$$|\{|A_{\mathscr{S},k}f|>\lambda\}|\leq|\{|A_{\mathscr{S},k}g|>\tfrac12\lambda\}|+|\Omega|+|\{|A_{\mathscr{S},k}b|>\tfrac12\lambda\}\cap\Omega^c|=:I+II+III.$$

For I, we have

$$I\leq\int\Big(\frac{|A_{\mathscr{S},k}g|}{\frac12\lambda}\Big)^2\mathrm{d}x=\frac{4}{\lambda^2}\|A_{\mathscr{S},k}g\|_2^2\leq\frac{C}{\lambda^2}\int|g|^2\,\mathrm{d}x\leq\frac{C}{\lambda^2}\int\lambda|g|\,\mathrm{d}x\leq\frac{C}{\lambda}\|f\|_1$$

by the L^2-boundedness of $A_{\mathscr{S},k}$ and the properties of the good part. Of course $II\leq\|f\|_1/\lambda$ by the properties of the decomposition. Hence we only need to estimate III.

For this, we have

$$III\leq\int_{\Omega^c}\frac{|A_{\mathscr{S},k}b|}{\frac12\lambda}\,\mathrm{d}x\leq\frac{2}{\lambda}\sum_{Q\in\mathscr{Q}}\int_{\Omega^c}|A_{\mathscr{S},k}b_Q|\,\mathrm{d}x\leq\frac{2}{\lambda}\sum_{Q\in\mathscr{Q}}\int 1_{Q^c}|A_{\mathscr{S},k}b_Q|\,\mathrm{d}x.$$

Now let us look at

$$1_{Q^c}A_{\mathscr{S},k}b_Q=1_{Q^c}\sum_{R\in\mathscr{D}}\frac{1_R}{|R|}\int_R\eta_Rb_Q\,\mathrm{d}x.$$

For the cube R to contribute to this sum, there are two obvious requirements:

- $R\cap Q\neq\varnothing$, since otherwise the domain of integration R does not meet the support of the function b_Q.
- $R\cap Q^c\neq\varnothing$, since otherwise the product of indicators $1_{Q^c}1_R$ is zero.

The only way that a dyadic R meets both Q and Q^c is that $R \supsetneq Q$.

But there is yet another, slightly less trivial condition: Since b_Q is supported on Q and its integral is zero, η_R *must not be constant on* Q (since if $\eta_R = c$ throughout Q, then $\int_R \eta_R b_Q \,\mathrm{d}x = \int_Q \eta_R b_Q \,\mathrm{d}x = c\int_Q b_Q \,\mathrm{d}x = c \cdot 0 = 0$). But observe from the defining formula (9.2) that η_R, as a sum of indicators, is constant on dyadic cubes that have sidelength $2^{-k}\ell(R)$ (or smaller). So for nonzero contribution, we *must have* $\ell(R) < 2^k\ell(Q)$, and previously we observed that we must also have $R \supsetneq Q$. This only leaves a finite set of possibilities: $R = Q^{(j)}$ is a dyadic ancestor of Q that preceeds Q by j generations, where $j = 1, \ldots, k-1$. So in fact

$$\begin{aligned}\int 1_{Q^c}|A_{\mathscr{S},k}b_Q|\,\mathrm{d}x &= \int \Big|1_{Q^c}\sum_{R\in\mathscr{D}}\frac{1_R}{|R|}\int_R \eta_R b_Q \,\mathrm{d}y\Big|\,\mathrm{d}x \\ &= \int \Big|1_{Q^c}\sum_{\substack{R=Q^{(j)}\\ j=1,\ldots,k-1}}\frac{1_R}{|R|}\int_R \eta_R b_Q \,\mathrm{d}y\Big|\,\mathrm{d}x \\ &\leq \sum_{\substack{R=Q^{(j)}\\ j=1,\ldots,k-1}}\int_R |b_Q|\,\mathrm{d}y \leq (k-1)_+\|b_Q\|_1 \leq 2(k-1)_+\|1_Q f\|_1.\end{aligned}$$

(Here $(k-1)_+ := \max(k-1, 0)$; if $k = 0$, the sum is empty, but we still don't get a negative upper bound!) From the disjointness of the cubes $Q \in \mathscr{Q}$ it then follows that

$$III \leq \frac{2}{\lambda}\sum_{Q\in\mathscr{Q}}\int 1_{Q^c}|A_{\mathscr{S},k}b_Q|\,\mathrm{d}x. \leq \frac{C}{\lambda}(k-1)_+\sum_{Q\in\mathscr{Q}}\|1_Q f\|_1 \leq \frac{C}{\lambda}(k-1)_+\|f\|_1,$$

and the combination of the bounds for I, II and III completes the proof. □

Exercise 9.5. Suppose that $\mathscr{S}$ is a sparse collection with the additional property that its *scales are k-separated*, i.e.: for some fixed $j \in \{0, 1, \ldots, k-1\}$, all $S \in \mathscr{S}$ have length of the form $\ell(S) = 2^{nk+j}$, where $n \in \mathbb{Z}$ is a variable that can be different for different S, but j (and k) is the same fixed quantity for all S. Prove that in this case the bound of Proposition 9.4 can be improved to $\|A_{\mathscr{S},k}f\|_{L^{1,\infty}} \leq c\|f\|_{L^1}$, i.e., a bound independendent of k. (Hint: you don't need to redo the whole proof; concentrate on the part, where the k-dependence came from.)

Corollary 9.6.

$$\omega_\lambda(A_{\mathscr{S},k}f; L) \leq \frac{c(1+k)}{|L|}\int_L |f|\,\mathrm{d}x.$$

Proof. We have

$$1_L A_{\mathscr{S},k}f = 1_L\sum_{Q\in\mathscr{D}}\frac{1_Q}{|Q|}\int_Q \eta_Q f\,\mathrm{d}x = \sum_{Q\subseteq L}\frac{1_Q}{|Q|}\int_Q \eta_Q f\,\mathrm{d}x + 1_L\sum_{Q\supsetneq L}\frac{1}{|Q|}\int_Q \eta_Q f\,\mathrm{d}x,$$

where we used the fact that $1_L 1_Q \neq 0$ only if $Q \subseteq L$ or $Q \supsetneq L$. A key point is that the last term is 1_L times a constant c (not a function), so that, with this constant,

$$1_L|A_{\mathscr{S},k}f - c| \leq \Big|\sum_{Q\subseteq L}\frac{1_Q}{|Q|}\int_Q \eta_Q f\,\mathrm{d}x\Big| \leq \sum_{Q\subseteq L}\frac{1_Q}{|Q|}\int_Q \eta_Q 1_L|f|\,\mathrm{d}x \leq A_{\mathscr{S},k}(1_L|f|).$$

Thus

$$\begin{aligned}\omega_\lambda(A_{\mathscr{S},k}f; L) &:= \inf_c (1_L(A_{\mathscr{S},k}f - c))^*(\lambda|L|) \\ &\leq (A_{\mathscr{S},k}(1_Q|f|))^*(\lambda|L|) \leq \frac{1}{\lambda|L|}\|A_{\mathscr{S},k}(1_L|f|)\|_{L^{1,\infty}} \leq \frac{c(1+k)}{|L|}\|1_L f\|_{L^1},\end{aligned}$$

which is what we claimed. □

We now return to the formula (9.1). We want to apply Lerner's formula to the function $A_k^\alpha|g| = A_{\mathscr{R}_k^\alpha,k}|g|$.

Lemma 9.7. *Every cube $R^{(k)}$ with $R \in \mathscr{R}_k^\alpha$ is contained in a maximal R^* of the same form.*

Proof. Recall that $\mathscr{R}_k^\alpha = \{R(L,k) : L \in \mathscr{L}, \alpha(L,k) = \alpha\}$. Since $\mathscr{L}$ is bounded from above (by the initial cube Q^0) and $\ell(R(L,k)) \leq 6\ell(L)$, the cubes in $\mathscr{R}_k^\alpha$ also have sidelengths bounded from above. And then so do the cubes $R^{(k)}$, $R \in \mathscr{R}_k^\alpha$, for any fixed k. Thus there cannot be infinite increasing chains among these cubes, and hence each of them will be contained in a maximal one. □

In every such R^* as in the Lemma, we apply Lerner's formula to $A_k^\alpha|g|$ to obtain a $\frac{1}{2}$-sparse collection $\mathscr{S}(R^*)$ of dyadic subcubes of R^* such that

$$1_{R^*}A_k^\alpha|g| \leq 1_{R^*}|m_{A_k^\alpha|g|}(R^*)| + \sum_{S\in\mathscr{S}(R^*)} 2\omega_\lambda(A_k^\alpha|g|;S)\cdot 1_S.$$

For the median term, we have

$$\begin{aligned}|m_{A_k^\alpha|g|}(R^*)| &\leq (1_{R^*}A_k^\alpha|g|)^*(\tfrac{1}{4}|R^*|) = (A_k^\alpha(1_{R^*}|g|))^*(\tfrac{1}{4}|R^*|)\\ &\leq \frac{1}{\frac{1}{4}|R^*|}\|A_k^\alpha(1_{R^*}|g|)\|_{L^{1,\infty}} \leq \frac{c(1+k)}{|R^*|}\|1_{R^*}g\|_{L^1},\end{aligned}$$

where the equality is easy by recalling the definition of A_k^α and the maximality of R^* among the relevant cubes $R^{(k)}$ appearing there.

The oscillatory terms we already estimate above by

$$\sum_{S\in\mathscr{S}(R^*)} 2\omega_\lambda(A_k^\alpha|g|;S)\cdot 1_S. \leq c(1+k)\sum_{S\in\mathscr{S}(R^*)}\frac{1_S}{|S|}\int_S|f|\,\mathrm{d}x,$$

and the estimate for the median term was exactly of the same form and can be absorbed in this sum. So altogether

$$1_{R^*}A_k^\alpha|g| \leq c(1+k)\sum_{S\in\mathscr{S}(R^*)}\frac{1_S}{|S|}\int_S|f|\,\mathrm{d}x,$$

where $\mathscr{S}(R^*)$ is a $\frac{1}{2}$-sparse collection of subcubes of R^*. If we sum over all maximal R^*, on the left we get simply $A_k^\alpha|g|$, and on the right a similar sum with $\mathscr{S}_k^\alpha = \bigcup_{R^*}\mathscr{S}(R^*)$ in place of $\mathscr{S}(R^*)$. Now $\mathscr{S}_k^\alpha$ is also $\frac{1}{2}$-sparse, since the maximal cubes R^* of the different $\mathscr{S}(R^*)$ are pairwise disjoint:

Exercise 9.8. Let $\mathscr{S}_i$, $i \in \mathscr{I}$, be γ-sparse collections such that $S_i \cap S_j = \varnothing$ whenever $S_k \in \mathscr{S}_k$ for both $k \in \{i,j\} \subseteq \mathscr{I}$ and $i \neq j$. Prove then $\bigcup_{i\in\mathscr{I}}\mathscr{S}_i$ is also γ-sparse. (Hint: This is easy.)

Substituting back to (9.1), we get

$$\begin{aligned}\int |Tf|\cdot|g|\,\mathrm{d}x &\leq C\sum_{\alpha\in\{0,1,2\}^d}\sum_{k=0}^\infty 2^{-k\beta}\int|f|\cdot A_k^\alpha|g|\,\mathrm{d}y\\ &\leq C\sum_{\alpha\in\{0,1,2\}^d}\sum_{k=0}^\infty 2^{-k\beta}(1+k)\sum_{S\in\mathscr{S}_k^\alpha}\frac{1}{|S|}\int_S|f|\,\mathrm{d}x\cdot\int_S|g|\,\mathrm{d}y\\ &\leq C\sum_{\alpha\in\{0,1,2\}^d}\sum_{k=0}^\infty 2^{-k\beta}(1+k)\sup_{\mathscr{D},\mathscr{S}}\sum_{S\in\mathscr{S}}\frac{1}{|S|}\int_S|f|\,\mathrm{d}x\int_S|g|\,\mathrm{d}y\\ &\leq C\sup_{\mathscr{D},\mathscr{S}}\sum_{S\in\mathscr{S}}\frac{1}{|S|}\int_S|f|\,\mathrm{d}x\int_S|g|\,\mathrm{d}y,\end{aligned}\tag{9.9}$$

where the supremum is over all relevant dyadic systems and their $\frac{1}{2}$-sparse subcollections, and in the last step we observed that the sum over k converges and the sum over α is finite in any case.

The estimate (9.9) is almost as useful as Theorem 8.1 itself.

Lemma 9.10. *Let $w : \mathbb{R}^d \to (0,\infty)$, and denote $L^p(w) := L^p(w\,\mathrm{d}x)$ the L^p-space with respect to the weighted measure $w\,\mathrm{d}x$. Then*

$$\|f\|_{L^p(w)} = \sup\Big\{\Big|\int f\cdot h\,\mathrm{d}x\Big| : \|h\|_{L^{p'}(w^{1-p'})} \leq 1\Big\}, \qquad p\in(1,\infty). \tag{9.11}$$

Proof. Recall that

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

for any σ-finite measure μ and $p\in(1,\infty)$. We apply this to a weighted measure $\mathrm{d}\mu = w\,\mathrm{d}x$ to find that

$$\|f\|_{L^p(w)} = \sup\Big\{\Big|\int f\cdot g\cdot w\,\mathrm{d}x\Big| : \|g\|_{L^{p'}(w)} \leq 1\Big\},$$

which looks a bit different than the claim. Now denote $h := g\cdot w$, so that $g = h/w$ and $\|g\|_{L^{p'}(w)} = \|h/w\|_{L^{p'}(w)} = \|h\|_{L^{p'}(w^{1-p'})}$. Observe that $g\mapsto h$ is an isometric bijection between the spaces $L^{p'}(w)$ and $L^{p'}(w^{1-p'})$ by this computation. Thus the previous display is actually the same as the claim of the lemma after the change of variable $h = g\cdot w$. □

Exercise 9.12. Show that formula (4.3) is still valid, if we only take the supremum over all bounded and compactly supported h which satisfy $\|h\|_{L^{p'}(w^{1-p'})}\leq 1$. Show that formula is also true for $p=1$ id we replace the condition on h by just requiring that h is compactly supported and $|h|\leq 1$ everywhere.

From (9.9), we can now easily give:

Proof of Theorem 8.1 for $X = L^p(w)$, $p\in[1,\infty)$. By the previous lemma and exercise, we have

$$\|Tf\|_{L^p(w)} = \sup_h \Big|\int Tf\cdot h\,\mathrm{d}x\Big| \leq \sup_h \int |Tf|\cdot|h|\,\mathrm{d}x,$$

where the supremum is over an appropriate set of bounded and compactly supported functions h. By (9.9) and the other direction of the weighted duality (9.11), we have

$$\begin{aligned}\sup_h \int |Tf|\cdot|h|\,\mathrm{d}x &\leq C\sup_h\sup_{\mathscr{D},\mathscr{S}}\sum_{S\in\mathscr{S}}\frac{1}{|S|}\int_S|f|\,\mathrm{d}x\cdot\int_S|h|\,\mathrm{d}y \\ &= C\sup_{\mathscr{D},\mathscr{S}}\sup_h\int\Big(\sum_{S\in\mathscr{S}}\frac{1_S}{|S|}\int_S|f|\,\mathrm{d}x\Big)|h|\,\mathrm{d}y \\ &= C\sup_{\mathscr{D},\mathscr{S}}\|\sum_{S\in\mathscr{S}}\frac{1_S}{|S|}\int_S|f|\,\mathrm{d}x\|_{L^p(w)} = C\sup_{\mathscr{D},\mathscr{S}}\|A_{\mathscr{S}}f\|_{L^p(w)},\end{aligned}$$

and this completes the proof. □

A general Banach function space is essentially just an abstract framework where we can repeat a similar reasoning.

10. A_p weights and the maximal operator

We next study several inequalities in the weighted spaces $L^p(w)$, where $w\in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ is positive and finite almost everywhere. The A_p class of weights consists of precisely those weights for which many such inequalities are valid. Instead of defining this class right away, we shall see how the A_p condition naturally arises from questions of boundedness for the Hardy–Littlewood maximal operator

$$Mf(x) = \sup_{\substack{Q \text{ cube}\\ Q\ni x}}\frac{1}{|Q|}\int_Q|f|\,\mathrm{d}y.$$

The question we want to address is: What are the weights w for which one of the following inequalities holds?

$$\|Mf\|_{L^{p,\infty}(w)} \leq C\|f\|_{L^p(w)} \tag{10.1}$$

or

$$\|Mf\|_{L^p(w)} \leq C\|f\|_{L^p(w)}. \tag{10.2}$$

Notice that we have the same weight w on both sides of (10.1) and (10.2). This is in contrast to the Fefferman–Stein inequalities

$$\|Mf\|_{L^{1,\infty}(w)} \leq C\|f\|_{L^1(Mw)},\qquad \|Mf\|_{L^{p,\infty}(w)} \leq \|Mf\|_{L^p(w)} \leq C\|f\|_{L^p(Mw)},\quad p\in(1,\infty),$$

where we have w on the left and Mw on the right. However, these provide a simple *sufficient* condition for (10.1) and (10.2): if the weight satisfies

$$Mw \leq cw, \tag{10.3}$$

the (10.1) holds for all $p \in [1, \infty)$, and (10.2) holds for all $p \in (1, \infty)$. The condition (10.3) is called the A_1 condition, and we define

$$[w]_{A_1} := \Big\| \frac{Mw}{w} \Big\|_\infty \tag{10.4}$$

and $w \in A_1$ if and only if $[w]_{A_1} < \infty$.

What about necessary conditions? We have the following:

Lemma 10.5. *If* (10.1) *holds for some* $p \in [1, \infty)$*, then*

$$[w]_{A_p} \leq \|M\|^p_{L^p(w) \to L^{p,\infty}(w)},$$

where $[w]_{A_1}$ *is defined in* (10.4)*, and*

$$[w]_{A_p} := \sup_{Q \text{ cube}} \Big(\frac{1}{|Q|} \int_Q w \Big) \Big(\frac{1}{|Q|} \int_Q w^{-1/(p-1)} \Big)^{p-1}, \quad p \in (1, \infty).$$

Proof. Let us abbreviate $N := \|M\|_{L^p(w) \to L^{p,\infty}(w)}$. Recall that

$$\|f\|_{L^{p,\infty}(w)} = \sup_{\lambda > 0} \lambda w(\{|f| > \lambda\})^{1/p} = \sup_{\lambda > 0} \lambda w(\{|f| \geq \lambda\})^{1/p},$$

where the first equality is the definition and the second is easy to check. (It is not necessarily true for a fixed λ, only for the supremum!)

Consider a function f that is nonnegative and supported in a cube Q, and write $\lambda := |Q|^{-1} \int_Q f \, \mathrm{d}y$. Then $Mf(x) \geq \lambda$ for all $x \in Q$, and thus

$$\|Mf\|_{L^{p,\infty}(w)} \geq \lambda w(Q)^{1/p} = \frac{1}{|Q|} \int_Q f \, \mathrm{d}y \cdot w(Q)^{1/p}.$$

Thus (10.1) gives that

$$\frac{1}{|Q|} \int_Q f \, \mathrm{d}y \cdot w(Q)^{1/p} \leq N \|f\|_{L^p(w)} = N \Big(\int_Q f^p w \, \mathrm{d}y \Big)^{1/p}. \tag{10.6}$$

Case $p > 1$. Basically, we want to make the two integrals equal, i.e., we want that $f = f^p w$, which means that we would choose $f = w^{-1/(p-1)}$. If we substitute this into (10.6), we get

$$\frac{1}{|Q|} \int_Q w^{-1/(p-1)} \, \mathrm{d}y \cdot w(Q)^{1/p} \leq N \Big(\int_Q w^{-1/(p-1)} \, \mathrm{d}y \Big)^{1/p}.$$

If $(\int_Q w^{-1/(p-1)} \, \mathrm{d}y)^{1/p} < \infty$, we can divide both sides by this quantity to get

$$\frac{1}{|Q|} \Big(\int_Q w^{-1/(p-1)} \, \mathrm{d}y \Big)^{1/p'} \cdot w(Q)^{1/p} \leq N. \tag{10.7}$$

Raising to power p and taking the supremum over cubes Q, this gives $[w]_{A_p} \leq N^p$, as claimed.

However, we do not know a priori whether the division that we made is correct, and need to proceed more carefully. So we choose instead $f = (w + \varepsilon)^{-1/(p-1)}$, which is bounded from above and thus certainly locally integrable. Then (10.6) gives

$$\begin{aligned} \frac{1}{|Q|} \int_Q (w + \varepsilon)^{-1/(p-1)} \, \mathrm{d}y \cdot w(Q)^{1/p} &\leq N \Big(\int_Q (w + \epsilon)^{-p/(p-1)} w \, \mathrm{d}y \Big)^{1/p} \\ &\leq N \Big(\int_Q (w + \epsilon)^{-p/(p-1)} (w + \varepsilon) \, \mathrm{d}y \Big)^{1/p} = N \Big(\int_Q (w + \epsilon)^{-1/(p-1)} \, \mathrm{d}y \Big)^{1/p}. \end{aligned}$$

Dividing by $(\int_Q (w + \epsilon)^{-1/(p-1)} \, \mathrm{d}y)^{1/p}$ (which is certainly legal), we arrive at

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

Passing to the limit $\varepsilon \to 0$, we obtain (10.7) by dominated convergence, and we complete the proof as indicated above.

Case $p = 1$. Now it is not possible to choose f with $f = fw$, so we take a different approach. Let $f = 1_R$ for a subcube $R \subset Q$. Then (10.6) (with $p = 1$) gives

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

We rearrange this estimate and pass to the limit where R shrinks to a point $x \in Q$. Lebesgue's differentiation theorem then guarantees that

$$\frac{w(Q)}{|Q|} \leq N \lim_{\substack{R \ni x \\ \ell(R) \to 0}} \frac{w(R)}{|R|} = N \cdot w(x)$$

at almost every $x \in Q$. If we consider a countable family $\mathscr{Q}$ of cubes (say all cubes with rational centres and sidelengths), then the union of the exceptional null sets is also a null set, and we find that for almost every x, the above bound holds for all $Q \in \mathscr{Q}$ with $Q \in x$. Thus, for almost every $x \in \mathbb{R}^d$, we have

$$N \cdot w(x) \geq \sup_{\substack{Q \in \mathscr{Q} \\ Q \ni x}} \frac{w(Q)}{|Q|} = \sup_{\substack{Q \text{ cube} \\ Q \ni x}} \frac{w(Q)}{|Q|} = M w(x),$$

where the equality of the supremums follows easily by observing that the ratio $w(Q)/|Q|$ for arbitrary cubes can be approximated arbitrarily well by the corresponding ratio for the cubes $Q \in \mathscr{Q}$. But the above bound precisely shows that $[w]_{A_1} \leq N$, as claimed. □

The necessary condition thus obtained turns out to be sufficient as well:

Theorem 10.8 (Muckenhoupt [Muc72])**.** *Let w be a weight and $p \in (1, \infty)$. Then the following conditions are equivalent:*

(1) *$M : L^p(w) \to L^{p,\infty}(w)$ boundedly.*
(2) *$M : L^p(w) \to L^p(w)$ boundedly.*
(3) *$w \in A_p$*

For $p = 1$, the conditions (1) *and* (3) *are still equivalent.*

We already proved case $p = 1$ completely. For $p \in (1, \infty)$, Lemma 10.5 shows that (1) $\Rightarrow$ (3), whereas (2) $\Rightarrow$ (1) is clear. It remains to prove that (3) $\Rightarrow$ (2), and we observe that it is enough to do this for the dyadic version M_d in place of M. For this, we use an argument of Lerner [Ler08]:

Lemma 10.9 (Lerner)**.**

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

Proof.

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

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

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

Lerner's proof of Theorem 10.8, (3) $\Rightarrow$ (2).

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

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

$$\|M_d f\|_{L^p(w)} \le e \cdot p' \cdot [w]_{A_p}^{1/(p-1)} \|f\|_{L^p(w)}. \qquad \square$$

The proofs given above already indicated a more quantitative formulation of Theorem 10.8, formulate below:

Theorem 10.10 (Buckley [Buc93]). *For a weight w, the following estimates hold:*

$$[w]_{A_p}^{1/p} \le \|M\|_{L^p(w)\to L^{p,\infty}(w)} \le c_p [w]_{A_p}^{1/p}, \qquad p \in [1,\infty) \tag{10.11}$$

$$\|M\|_{L^p(w)\to L^p(w)} \le c_p [w]_{A_p}^{1/(p-1)}, \qquad p \in (1,\infty), \tag{10.12}$$

and this is sharp in the following sense: if $\|M\|_{L^p(w)\to L^p(w)} \le \phi([w]_{A_p})$ for some increasing positive ϕ, then $\phi(t) \ge ct^{1/(p-1)}$.

We already proved the lower bound in (10.11), the upper bound in (10.11) for $p = 1$, and the bound (10.12). The remaining claims are proven in the following exercises:

Exercise 10.13. Prove the upper bound in (10.11) for $p \in (1,\infty)$. Hint: Adapt the proof of the $p = 1$ case, namely the Fefferman–Stein inequality. After using the condition $\lambda < |Q|^{-1}\int_Q f$, use Hölder appropriately to estimate $\int_Q f$ in terms of $\int_Q f^p w$.

Exercise 10.14. Consider the *power weights* $x \in \mathbb{R}^d \mapsto w(x) = |x|^\alpha$, where $\alpha \in \mathbb{R}$. For $p \in (1,\infty)$, prove that $w \in A_p$ if and only if $-d < \alpha < d(p-1)$, and in this range

$$c_{p,d}[w]_{A_p} \le \frac{1}{(d+\alpha)(d(p-1)-\alpha)^{p-1}} \le C_{p,d}[w]_{A_p}.$$

Exercise 10.15. For every $\varepsilon \in (0,1)$, consider the function $f(x) = 1_B(x)|x|^{\varepsilon-d}$, where $B = B(0,1)$ is the unit ball of $\mathbb{R}^d$. Prove that $Mf(x) \ge c_d\varepsilon^{-1}f(x)$. Conclude that $\|M\|_{L^p(w)\to L^p(w)} \ge c_d\epsilon^{-1}$ whenever w is a weight such that $f \in L^p(w)$. Find a power weight with this property, and such that $[w]_{A_p} \le C_{p,d}\epsilon^{-(p-1)}$. Use this this example to derive the sharpness claim of Theorem 10.10.

11. A_2 weights and Calderón–Zygmund operators

The surprising feature of the A_p class is its *universality.* We already saw that this same condition characterizes both the strong and the weak type inequalities for the maximal operator. The next result shows that the same class is also the correct one for Calderón–Zygmund operators:

Theorem 11.1 (Hunt, Muckenhoupt, Wheeden [HMW73]; Coifman, Fefferman [CF74]). *For a weight $w \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ and $p \in (1,\infty)$, the following are equivalent:*

(1) *$w \in A_p$.*
(2) *$T : L^p(w) \to L^p(w)$ boundedly for all Calderón–Zygmund operators T.*
(3) *If $d = 1$, then $H : L^p(w) \to L^p(w)$ boundedly for the Hilbert transform*

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

It is immediate to check the kernel $K(x,y) = 1/(x-y)$ of the Hilbert transform satisfies the Calderón–Zygmund standard estimates. We will take for granted in these lectures that $H : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is bounded (This can be proven in several different ways; it has been done for instance in the courses "Martingales and harmonic analysis" and "Singular integrals and the Tb theorem".), and therefore H is a Calderón–Zygmund operator. Thus (2) $\Rightarrow$ (3) is obvious.

In the rest of this section, we prove the implications (3) $\Rightarrow$ (1) and (1) $\Rightarrow$ (2) in the case $p = 2$.

11.A. **Dual weight formulation.** When studying an inequality of the form

$$\|Tf\|_{L^p(w)} \leq N\|f\|_{L^p(w)}, \tag{11.2}$$

it is often convenient to make a substitution $f = g\sigma$, where σ is almost everywhere positive and finite. Then the previous bound takes the form

$$\|T(g\sigma)\|_{L^p(w)} \leq N\|g\sigma\|_{L^p(w)} = N\|g\|_{L^p(\sigma^p w)}.$$

We then want to choose σ so that $\sigma^p w = \sigma$, i.e., $\sigma = w^{-1/(p-1)}$, which indeed satisfies the condition of almost everywhere positivity and finiteness, since w has these properties. With this choice, observe that the relation $f = g\sigma$ establishes a one-to-one isometric correspondence between all $f \in L^p(w)$ and all $g \in L^p(\sigma)$. Thus, the bound (11.2) for all $f \in L^p(w)$ is equivalent to the bound

$$\|T(g\sigma)\|_{L^p(w)} \leq N\|g\|_{L^p(\sigma)} \tag{11.3}$$

for all $g \in L^p(\sigma)$, where $\sigma = w^{-1/(p-1)} = w^{1-p'}$ is the *dual weight*. Observe that this same dual weight also appears both in the A_p condition, and in the duality of weighted L^p spaces! By duality, yet another equivalent form is given by

$$\Big|\int T(g\sigma)\cdot hw\Big| \leq N\|g\|_{L^p(\sigma)}\|h\|_{L^{p'}(w)}. \tag{11.4}$$

11.B. **The necessity of the A_2 condition for the Hilbert transform.** We study the dual weight formulation for (11.4) for $T = H$. Let g and h be supported on disjoint adjacent intervals I, J of equal length $|I| = |J|$. Then every $x \in J$ is outside the support of g, and we can apply the kernel representation of the Hilbert transform (without the principal values):

$$\Big|\int_J \Big(\int_I \frac{g(y)\sigma(y)\,\mathrm{d}y}{x-y}\Big)h(x)w(x)\,\mathrm{d}x\Big| \leq N\|g\|_{L^p(\sigma)}\|h\|_{L^{p'}(w)}.$$

Now let us assume that g and h are nonnegative, and observe that $x - y$ has a constant sign throughout the double integral: it is positive if J is on the right of I, and negative if J is on the left of I. Thus $x - y = \varepsilon|x-y|$, where $\varepsilon = \pm 1$ is a constant. But this constant is killed by the absolute values outside the integral above, and we conclude that

$$\int_J\int_I \frac{g(y)\sigma(y)\,\mathrm{d}y}{|x-y|}h(x)w(x)\,\mathrm{d}x \leq N\|g\|_{L^p(\sigma)}\|h\|_{L^{p'}(w)}.$$

Observe next that $|x-y| \leq 2|I|$ (using that $x \in J$, $y \in I$, and these intervals are adjacent and have equal length), so that $1/|x-y| \geq 1/(2|I|)$. Then we choose $h = 1_J$ and $g = 1_{I\cap\{\sigma\leq k\}}$ (since we do not yet know that the dual weight is locally integrable). This gives

$$\frac{1}{2|I|}\sigma(I\cap\{\sigma\leq k\})w(J) \leq N\sigma(I\cap\{\sigma\leq k\})^{1/p}w(J)^{1/p'}.$$

Suppose that both factors on the right are non-zero. (Their finiteness is automatic, since $w \in L^1_{\mathrm{loc}}$ by assumption, and $\sigma(I\cap\{\sigma\leq k\}) \leq k|I|$.) Dividing both sides, we arrive at

$$\sigma(I\cap\{\sigma\leq k\})^{1/p'}w(J)^{1/p} \leq 2N|I|.$$

We can then pass to the limit $k \to \infty$ and apply monotone convergence to conclude that

$$\sigma(I)^{1/p'}w(J)^{1/p} \leq 2N|I|.$$

Since $w(J) > 0$, this shows in particular that $\sigma(I) < \infty$ and thus $\sigma \in L^1_{\mathrm{loc}}$ as well. Note that if $J = I$, then the above bound (raised to power p and reorganized) would be precisely the A_p condition for w. However, we now have it for adjacent intervals, not equal intervals.

We now turn to the case $p = 2$, and show how to derive the A_2 condition from

$$\sigma(I)w(J) \leq 4N^2|I|^2 =: M|I||J|, \qquad I, J \text{ adjacent intervals of equal length.} \tag{11.5}$$

We need the following geometric observation:

Lemma 11.6. *If $K = [a, b) \subset \mathbb{R}$ is an interval, then we have the disjoint union*

$$\{(x, y) \in K \times K : x \neq y\} = \bigcup_{\substack{I \subseteq K \\ dyadic \\ subinterval}} \Big((I_{\text{left}} \times I_{\text{right}}) \cup (I_{\text{right}} \times I_{\text{left}}) \Big)$$

Proof. The inclusion $\supseteq$ is easy to see. For the other direction, if $x, y \in K$ and $x \neq y$, then there is a minimal dyadic subinterval $I \subseteq K$ such that $x, y \in I$. (Indeed, there are some intervals like this, at least the interval K itself, and there cannot be arbitrarily small such intervals since $x \neq y$; thus there is a minimal one among all I that contain both x and y.) Since I is minimal, it cannot be that both $x, y \in I_u$ for $u \in \{\text{left}, \text{right}\}$. Thus the only possibility is that $x \in I_u$ and $y \in I_v$ for different halves of the interval I. This proves $\subseteq$.

The disjointness follows from the same consideration: if $(x, y) \in I_u \times I_v$ for $u \neq v$, $u \in \{\text{left}, \text{right}\}$, then I contains both x and y, and it is the minimal interval with this property, thus unique. $\square$

Now we can prove:

Lemma 11.7. *If* (11.5) *holds for some weights $\sigma, w \in L^1_{\text{loc}}(\mathbb{R})$, then they also satisfy*

$$\sigma(K)w(K) \leq M|K|^2$$

for all intervals $K \subset \mathbb{R}$.

Proof. We interpret both the assumption and the claim as estimates for *product measures* in the plane, namely, we assume that

$$(\sigma \times w)(I \times J) \leq M|I \times J|,$$

for equal adjacent intervals, where on the right we have the Lebesgue area measure of the square $I \times J$, and want to prove that

$$(\sigma \times w)(K \times K) \leq M|K \times K|.$$

We can apply Lemma 11.6, once we realize that the diagonal $\{(x, x) : x \in K\}$ has area zero, and therefore also $\sigma \times w$-measure zero, since this is a locally finite weighted measure $\sigma(x)w(y)\,\mathrm{d}x\,\mathrm{d}y$ with respect to the area measure. Thus

$$\begin{aligned}
(\sigma \times w)(K \times K) &= (\sigma \times w)(\{(x, y) \in K \times K : x \neq y\}) \\
&= \sum_{\substack{I \subseteq K \\ \text{dyadic} \\ \text{subinterval}}} \Big((\sigma \times w)(I_{\text{left}} \times I_{\text{right}}) + (\sigma \times w)(I_{\text{right}} \times I_{\text{left}}) \Big) \\
&\leq \sum_{\substack{I \subseteq K \\ \text{dyadic} \\ \text{subinterval}}} \Big(M|I_{\text{left}} \times I_{\text{right}}| + M|I_{\text{right}} \times I_{\text{left}}| \Big) \\
&= M|\{(x, y) \in K \times K : x \neq y\}| = M|K \times K|.
\end{aligned}$$ $\square$

Combining the above estimates, we have shown:

$$[w]_{A_2} \leq 4 \cdot \|H\|^2_{L^2(w) \to L^2(w)},$$

which in particular confirms (3) $\Rightarrow$ (1) of Theorem 11.1 for $p = 2$.

11.C. **The A_2 theorem.** We will prove (1) $\Rightarrow$ (2) (still for $p=2$) in the following sharp quantitative form:

Theorem 11.8 (The A_2 theorem [Hyt12]). *For any Calderón–Zygmund operator T, and any $w \in A_2$, we have*
$$\|Tf\|_{L^2(w)} \leq c_T [w]_{A_2} \|f\|_{L^2(w)}.$$

The proof that we present makes the theorem look rather simple. However, when the theorem was first proven, Lerner's dyadic domination theorem was not yet available, and the proof had to proceed by a different route.

By the dyadic domination theorem, it is enough to prove that
$$\|A_{\mathscr{S}} f\|_{L^2(w)} \leq c \|f\|_{L^2(w)},$$
uniformly for all $\frac{1}{2}$-sparse collection $\mathscr{S}$. By the dual weight formulation, this is equivalent to showing that
$$\int A_{\mathscr{S}}(f\sigma) \cdot gw \leq c\|f\|_{L^2(w)} \|g\|_{L^2(\sigma)}. \tag{11.9}$$

Proof of (11.9) *by Cruz-Uribe, Martell & Pérez* [CUMP10].
$$\begin{aligned}
\int A_{\mathscr{S}}(f\sigma) \cdot gw &= \sum_{S \in \mathscr{S}} \frac{1}{|S|} \int_S f\sigma \cdot \int_S gw \\
&= \sum_{S \in \mathscr{S}} \frac{\sigma(S) w(S)}{|S|^2} \cdot \frac{1}{\sigma(S)} \int_S f\sigma \cdot \frac{1}{w(S)} \int_S gw \cdot |S| \\
&\leq \sum_{S \in \mathscr{S}} [w]_{A_2} \cdot \inf_{y \in S} M_\sigma^d f(y) \cdot \inf_{z \in S} M_w^d g(z) \cdot 2|E(S)| \\
&\leq 2[w]_{A_2} \sum_{S \in \mathscr{S}} \int_{E(S)} M_\sigma^d f(x) \cdot M_w^d g(x) \,\mathrm{d}x \\
&\leq 2[w]_{A_2} \int_{\mathbb{R}^d} M_\sigma^d f(x) \cdot M_w^d g(x) \cdot \sigma(x)^{1/2} w(x)^{1/2} \,\mathrm{d}x \\
&\leq 2[w]_{A_2} \Big(\int_{\mathbb{R}^d} (M_\sigma^d f)^2 \sigma \Big)^{1/2} \Big(\int_{\mathbb{R}^d} (M_w^d g)^2 w \Big)^{1/2} \\
&\leq 2[w]_{A_2} \cdot 2\|f\|_{L^2(\sigma)} \cdot 2\|g\|_{L^2(w)} = 8[w]_{A_2} \|f\|_{L^2(\sigma)} \|g\|_{L^2(w)}.
\end{aligned}$$
□

Note that this proof of (11.9) existed before Theorem 11.8; since the dyadic domination theorem was not yet available, it was not known that (11.9) can be used to get Theorem 11.8 at that point. The particular case of the Hilbert transform is now a corollary to Theorem 11.8, but historically, it was already known earlier:

Theorem 11.10 (Petermichl [Pet07]). *For the Hilbert transform H, and any $w \in A_2$, we have*
$$\|Hf\|_{L^2(w)} \leq c[w]_{A_2} \|f\|_{L^2(w)},$$
and this estimate is sharp in the following sense: if $\|H\|_{L^2(w) \to L^2(w)} \leq \phi([w]_{A_2})$ for some increasing positive ϕ, then $\phi(t) \geq ct$.

Exercise 11.11. Let $p \in (1, \infty)$. Show that if $\|H\|_{L^p(w) \to L^p(w)} \leq \phi([w]_{A_p})$, then $\phi(t) \geq ct^{1/(p-1)}$. (Hint: This is similar to the case of the maximal function. Let $f(x) = |x|^{-\alpha} 1_{(-1,0)}(x)$, and estimate the size of $Hf(x)$ for $x \in (0,1)$. Make a conclusion about the size of $\|H\|_{L^p(w) \to L^p(w)}$ for $w(x) = |x|^\beta$ for those β for which $f \in L^p(w)$.)

Exercise 11.12. Let $p \in (1, \infty)$. Show that if $\|H\|_{L^p(w) \to L^p(w)} \leq \phi([w]_{A_p})$, then also $\phi(t) \geq ct$. (Hint: Use duality. The Hilbert transform satisfies $\int Hf \cdot g \,\mathrm{d}x = -\int f \cdot Hg \,\mathrm{d}x$; check this by a formal computation. Use this and the weighted duality to conclude that $\|H\|_{L^p(w) \to L^p(w)} = \|H\|_{L^{p'}(\sigma) \to L^{p'}(\sigma)}$ for the dual weight $\sigma = w^{1-p'}$. Then use the result of the previous exercise.

A combination of the two exercises shows that $\|H\|_{L^p(w)\to L^p(w)} \leq c_p[w]_{A_p}^{\max(1,1/(p-1))}$ is the best possible bound for the Hilbert transform for any $p \in (1,\infty)$. In the next section, we will prove this optimal bound by an *extrapolation method* from Theorem 11.10.

12. The extrapolation theorem of Rubio de Francia

Theorem 12.1. *Let T be an operator, not necessarily linear, and $r \in (1,\infty)$. Suppose that T satisfies the estimate*

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

for all $w \in A_r$ and all $f \in L^r(w)$, where ϕ_r is a nonnegative increasing function. Then it satisfies

$$\|Tf\|_{L^p(w)} \leq \phi_p([w]_{A_p})\|f\|_{L^r(w)}$$

for all $p \in (1,\infty)$, all $w \in A_p$ and all $f \in L^p(w)$, where each ϕ_p is a nonnegative increasing function.

In particular, if $\phi_r(t) = ct^\tau$, then $\phi_p(t) \leq c_p t^{\tau \max(\frac{r-1}{p-1},1)}$.

The proof naturally splits into two cases, $p \in (1,r)$ and $p \in (r,\infty)$. We first give the beginning of the proof, which motivates a certain auxiliary construction that is needed to complete it.

12.A. **Proof of Theorem 12.1, case $p \in (1,r)$, beginning.** In order to use the assumed L^r inequality, we apply Hölder, after dividing and multiplying by an auxiliary function ψ, yet to be chosen:

$$\begin{aligned}\|Tf\|_{L^p(w)} &= \Big(\int\Big(\frac{|Tf|}{\psi}\Big)^p\cdot\psi^p w\Big)^{1/p}\\ &\leq \Big(\int\Big(\frac{|Tf|}{\psi}\Big)^r\cdot\psi^p w\Big)^{1/r}\Big(\int\psi^p w\Big)^{1/p-1/r}.\end{aligned}$$

Now, we would like to apply the L^r boundedness of T to the first factor, which would require that $\psi^{-r/p}w \in A_r$, and we would like to estimate the second factor by $\|f\|_{L^p(w)}$. We try to achieve this by some $\psi = Rf$. Thus, we would get

$$\begin{aligned}\|Tf\|_{L^p(w)} &\leq \Big(\int|Tf|^r(Rf)^{p-r}w\Big)^{1/r}\Big(\int(Rf)^p w\Big)^{1/p-1/r}\\ &\leq \phi_r\big([(Rf)^{p-r}w]_{A_r}\big)\Big(\int|f|^r(Rf)^{p-r}w\Big)^{1/r}\|S\|^{1-p/r}_{L^p(w)\to L^p(w)}\Big(\int|f|^p w\Big)^{1/p-1/r} \qquad (12.2)\\ &\leq \phi_r\big([(Rf)^{p-r}w]_{A_r}\big)\|R\|^{1-p/r}_{L^p(w)\to L^p(w)}\Big(\int|f|^p w\Big)^{1/p},\end{aligned}$$

provided that $Rf \geq |f|$, so that $(Rf)^{p-r} \leq |f|^{p-r}$ in the last step.

Altogether, we would like to have an operator R such that: $Rf \geq |f|$ pointwise, R is bounded on $L^p(w)$, and $(Rf)^{p-r}w \in A_r$ for all $w \in A_p$. Such an operator is constructed next:

12.B. **The Rubio de Francia algorithm.** The following general construction is the key to our problem. It has many other applications as well.

Proposition 12.3 (Rubio de Francia algorithm)**.** *For $\epsilon > 0$, consider the operator*

$$Rg = R_\epsilon g := \sum_{k=0}^\infty \epsilon^k M^k g,$$

where $M^0 g := |g|$, $M^1 g := Mg$ is the maximal function of g, and $M^k g := M(M^{k-1}g)$ is the k-fold iteration of M acting on g. Then this satisfies

(i) $|g| \leq Rg$,

(ii) $\|Rg\|_{L^p(w)} \leq \Big(\sum_{k=0}^\infty \epsilon^k\|M\|^k_{L^p(w)\to L^p(w)}\Big)\|g\|_{L^p(w)}$,

(iii) $[Rg]_{A_1} \leq \epsilon^{-1}$.

In particular, if $\epsilon = \epsilon(p,w) := \frac{1}{2}\|M\|^{-1}_{L^p(w)\to L^p(w)}$, then

(ii') $\|Rg\|_{L^p(w)} \leq 2\|g\|_{L^p(w)}$
(iii') $[Rg]_{A_1} \leq 2\|M\|_{L^p(w)\to L^p(w)} \leq c_p[w]_{A_p}^{1/(p-1)}$.

Here (i) says that Rg is bigger than g, but not too much bigger by (ii) or especially (ii'). The hole point of the construction is (iii) (or (iii')), which shows that Rg is an A_1 weight, with control on the A_1 constant.

Proof. (i) is clear, since Rg is a sum of nonnegative terms, and the zeroth term is $|g|$. (ii) follows from the triangle inequality in $L^p(w)$ and iteration of the estimate

$$\|M^k g\|_{L^p(w)} \leq \|M\|_{L^p(w)\to L^p(w)} \|M^{k-1}g\|_{L^p(w)}.$$

To prove (iii), recall that $[w]_{A_1} = \|Mw/w\|_\infty$. And indeed, by the sublinearity of M, we have

$$M(Rg) \leq \sum_{k=0}^\infty \epsilon^k M^{k+1}g = \epsilon^{-1}\sum_{k=0}^\infty \epsilon^{k+1}M^{k+1}g = \epsilon^{-1}\sum_{k=1}^\infty \epsilon^k M^k g \leq \epsilon^{-1}Rg,$$

so that $M(Rg)/Rg \leq \epsilon^{-1}$. □

Remark 12.4. In applications of the Rubio de Francia algorithm, the following reformulation of the A_1 condition is often handy. Recall that the A_1 condition says that $Mw(x) \leq [w]_{A_1}w(x)$, which by the definition of the maximal function is equivalent to

$$\langle w\rangle_Q \leq [w]_{A_1}w(x), \qquad \langle w\rangle_Q := \frac{1}{|Q|}\int_Q w$$

whenever Q is a cube that contains x. Dividing both sides, this transform into

$$w(x)^{-\alpha} \leq [w]_{A_1}^\alpha \langle w\rangle_Q^{-\alpha}, \qquad \alpha \geq 0, \qquad x \in Q. \tag{12.5}$$

(Actually, one needs slight care with the fact that the above conditions hold for all Q but only a.e. $x \in Q$. In the dyadic case there is no problem, since there are only countably many dyadic cubes altogether. In general, one might like to observe that the A_1 condition over all cubes is equivalent to the same condition, say, for all cubes with rational centres and sidelengths, which is again a countable family.)

Lemma 12.6. *Let $1 < p < r < \infty$. Let $w \in A_p$, $\epsilon = \epsilon(p,w)$ and $f \in L^p(w)$. Then $W := (Rf)^{p-r}w = (R_\epsilon f)^{p-r}w$ satisfies $[W]_{A_r} \leq c_p^{r-p}[w]_{A_p}^{(r-1)/(p-1)}$.*

Substituting this into (12.2), we obtain

$$\|Tf\|_{L^p(w)} \leq \phi_r\big(c_p^{r-p}[w]_{A_p}^{(r-1)/(p-1)}\big)\cdot 2^{1-p/r}\cdot\|f\|_{L^p(w)},$$

which completes the proof of Theorem 12.1 in the case $p \in (1,r)$.

Proof of the Lemma. By the definition of A_r, we need to estimate $\langle W\rangle_Q\langle W^{-1/(r-1)}\rangle_Q^{r-1}$. By (12.5) with $\alpha = r-p > 0$, the first factor is

$$\langle W\rangle_Q = \langle (Rf)^{p-r}w\rangle_Q \leq [Rf]_{A_1}^{r-p}\langle Rf\rangle_Q^{p-r}\langle w\rangle_Q,$$

whereas by Hölder's inequality, the second factor is

$$\begin{aligned}\langle W^{-1/(r-1)}\rangle_Q^{r-1} &= \langle (Rf)^{(r-p)/(r-1)}w^{-1/(r-1)}\rangle_Q^{r-1} \\ &\leq \langle Rf\rangle_Q^{r-p}\langle w^{-1/(p-1)}\rangle_Q^{p-1}.\end{aligned}$$

Forming the product, the factors involving $\langle Rf\rangle_Q$ cancel out, and we are left with

$$\begin{aligned}\langle W\rangle_Q\langle W^{-1/(r-1)}\rangle_Q^{r-1} &\leq [Rf]_{A_1}^{r-p}\langle w\rangle_Q\langle w^{-1/(p-1)}\rangle_Q^{p-1} \\ &\leq (c_p[w]_{A_p}^{1/p-1})^{r-p}[w]_{A_p} = c_p^{r-p}[w]_{A_p}^{(r-1)/(p-1)}.\end{aligned}$$

□

12.C. **Proof of Theorem 12.1, case $p \in (r, \infty)$.** In the attempt to argue by Hölder's inequality as before, we face the problem that "Hölder increases the exponent", while we now would like to use information about the smaller exponent r to get bounds for $p > r$. We circumvent this problem by duality:

$$\|Tf\|_{L^p(w)} = \sup\Big\{ \int |Tf| \cdot g : \|g\|_{L^{p'}(\sigma)} \leq 1 \Big\}, \qquad \sigma = w^{-1/(p-1)}.$$

Here it is useful to observe that

$$[w]_{A_p}^{1/p} = [\sigma]_{A_{p'}}^{1/p'}. \tag{12.7}$$

With a good faith in the Rubio de Francia algorithm $R = R_\epsilon$ with $\epsilon = \epsilon(p', \sigma)$ so that

$$\epsilon^{-1} = 2\|M\|_{L^{p'}(\sigma) \to L^{p'}(\sigma)} \leq c_{p'} [\sigma]_{A_{p'}}^{1/(p'-1)} = c_{p'} [\sigma]_{A_{p'}}^{1/p' \cdot p'/(p'-1)} = c_{p'} [w]_{A_p}^{1/p \cdot p} = c_{p'} [w]_{A_p}, \tag{12.8}$$

we estimate

$$\begin{aligned} \int |Tf| \cdot g \leq \int |Tf| \cdot Rg = \int |Tf| \cdot \frac{Rg}{w} \cdot w = \int |Tf| \cdot (\frac{Rg}{w})^{1-u} \cdot (\frac{Rg}{w})^u \cdot w \\ \leq \Big(\int |Tf|^r (\frac{Rg}{w})^{(1-u)r} w \Big)^{1/r} \Big(\int (\frac{Rg}{w})^{ur'} w \Big)^{1/r'}. \end{aligned}$$

We demand that $ur' = p'$, so that $u = p'/r'$ and

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

Thus

$$\begin{aligned} \int |Tf| \cdot g &\leq \Big(\int |Tf|^r (Rg)^{\frac{p-r}{p-1}} w^{\frac{r-1}{p-1}} \Big)^{1/r} \Big(\int (Rg)^{p'} w^{1-p'} \Big)^{1/r'} \\ &\leq \phi_r([W]_{A_r}) \Big(\int |f|^r (Rg)^{\frac{p-r}{p-1}} w^{r/p} w^{\frac{r-p}{p(p-1)}} \Big)^{1/r} (2\|g\|_{L^{p'}(\sigma)})^{p'/r'} \\ &\leq \phi_r([W]_{A_r}) \Big(\int |Tf|^p w \Big)^{1/p} \Big(\int (Rg)^{\frac{p-r}{p-1} \cdot (p/r)'} w^{\frac{r-p}{p(p-1)} \cdot (p/r)'} \Big)^{1/r - 1/p} 2^{p'/r'} \\ &= \phi_r([W]_{A_r}) \Big(\int |Tf|^p w \Big)^{1/p} \Big(\int (Rg)^{p'} w^{-\frac{1}{p-1}} \Big)^{1/r-1/p} 2^{p'/r'} \\ &\leq \phi_r([W]_{A_r}) \Big(\int |Tf|^p w \Big)^{1/p} 2^{p'(1/r-1/p)} 2^{p'/r'} = 2\phi_r([W]_{A_r}) \|Tf\|_{L^p(w)}, \end{aligned}$$

where $W := (Rg)^{\frac{p-r}{p-1}} w^{\frac{r-1}{p-1}}$. The proof of Theorem 12.1 is then completed by the following Lemma:

Lemma 12.9. *Let $1 < r < p < \infty$, let $w \in A_p$, $g \in L^{p'}(w^{1-p'})$ and $R = R_\epsilon$ with $\epsilon = \epsilon(p', w^{1-p'})$. Then $W := (Rg)^{\frac{p-r}{p-1}} w^{\frac{r-1}{p-1}} \in A_r$, and $[W]_{A_r} \leq c_{pr} [w]_{A_p}$.*

Proof. We estimate $\langle W \rangle_Q$ by Hölder's inequality:

$$\langle W \rangle_Q \leq \langle Rg \rangle_Q^{(p-r)/(p-1)} \langle w \rangle_Q^{(r-1)/(p-1)},$$

and $\langle W^{-1/(r-1)} \rangle_Q^{r-1}$ with the help of $Rg \in A_1$ via (12.5):

$$\langle W^{-1/(r-1)} \rangle_Q^{r-1} = \langle (Rg)^{-\frac{p-r}{(p-1)(r-1)}} w^{-\frac{1}{p-1}} \rangle_Q^{r-1} \leq [Rg]_{A_1}^{\frac{p-r}{p-1}} \langle Rg \rangle_Q^{-(p-r)/(p-1)} \langle w^{-1/(p-1)} \rangle_Q^{r-1}.$$

Forming the product, observing that terms with $\langle Rg \rangle_Q$ cancel out, and using $[Rg]_{A_1} \leq \epsilon^{-1}$ given by (12.8), we are left with

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

which is what we wanted. $\square$

13. Some two-weight theory

For the sake of simplicity, we here concentrate on $p = 2$, although some of the results discussed have natural extensions to all $p \in (1, \infty)$.

Recall that the weighted norm inequality

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

can be equivalently reformulated as

$$\|T(f\sigma)\|_{L^2(w)} \leq C\|f\|_{L^2(\sigma)}, \tag{13.1}$$

where $\sigma = 1/w$. However, once we have arrived at this formulation, we may also study (13.1) on its own right for two arbitrary weights (w, σ), no more assuming the condition that $\sigma = 1/w$. This general *two-weight problem* has some important particular instances:

13.A. **The characterization problem:** Describe all pairs of weights (w, σ) so that (13.1) is valid

- for the Hilbert transform $T = H$, or another particular singular integral operator, or
- for all Calderón–Zygmund operators T.

This is a difficult problem that has been only recently solved in the case of the Hilbert transform by Lacey, Sawyer, Shen and Uriarte-Tuero [Lac13, LSSUT12]:

Theorem 13.2. *The inequality* (13.1) *holds for $T = H$, if and only if*

$$\|H(1_I\sigma)\|_{L^2(w)} \leq C\|1_I\|_{L^2(\sigma)} = C\sigma(I)^{1/2}, \quad \textit{and}$$
$$\|H(1_I w)\|_{L^2(\sigma)} \leq C\|1_I\|_{L^2(w)} = Cw(I)^{1/2}$$

for all finite intervals $I \subset \mathbb{R}$.

That is, one only needs to "test" the estimate (13.1), and its dual version, for all indicators 1_I instead of all functions f. The dual weight formulation (13.1) is also meaningful for general (say, Radon) *measures* w, σ instead of weights with respect to the Lebesgue measure. The result of Lacey et al. also covers many but not all such situations. For general measures, Theorem 13.2 has been established in [Hyt13]. Already for weights, Theorem 13.2 is a difficult result, and we will not say more about it here.

13.B. **Bump theory.** Another line of research is trying to find simple sufficient conditions for (13.1) conditions, in the style of the classical A_2 condition. Recall that

$$[w]_{A_2} = \sup_Q \frac{1}{|Q|}\int_Q w \cdot \frac{1}{|Q|}\int_Q \sigma, \qquad \sigma = \frac{1}{w}.$$

Again, this naturally generalizes to the two-weight case, just by dropping the condition that $\sigma = 1/w$:

$$[w, \sigma]_{A_2} := \sup_Q \frac{1}{|Q|}\int_Q w \cdot \frac{1}{|Q|}\int_Q \sigma.$$

The proof that we gave for the necessity of $[w]_{A_2} < \infty$ for the $L^2(w)$ boundedness of the Hilbert transform easily extends to show that $[\sigma, w]_{A_2} < \infty$ is a necessary condition for (13.1) with $T = H$, in the case of arbitrary two weights. However, it is also known that this condition is not sufficient; in the case of general measures, it is actually not necessary either, so it is completely unrelated to (13.1) in the full generality. Nevertheless, the success of the A_2 theory in the one-weight case has encouraged attempts to find some similar theory even in the two-weight case. An early result in this direction is the following:

Theorem 13.3 (Neugebauer)**.** *Suppose that two weights w, σ satisfy the "bumped up" A_2 condition*

$$\mathscr{A}_2 := \sup_Q \Big(\frac{1}{|Q|}\int_Q w^r\Big)^{1/r}\Big(\frac{1}{|Q|}\int_Q \sigma^r\Big)^{1/r} \tag{13.4}$$

for some $r > 1$. Then (13.1) *holds for all Calderón–Zygmund operators T.*

There is a lot of theory towards more general sufficient bump conditions in the style of (13.4). We will prove Theorem 13.3 as a special case of an abstract formulation due to Lerner. This requires some preparations.

First, it is convenient to reformulate (13.4) slightly: note that

$$\Big(\frac{1}{|Q|}\int_Q w^r\Big)^{1/r} = \Big(\frac{1}{|Q|}\int_Q (w^{1/2})^{2r}\Big)^{2/2r} = \|w^{1/2}\|^2_{L^{2r}(Q)},$$

where we use the normalized norm

$$\|f\|_{L^q(Q)} := \Big(\frac{1}{|Q|}\int_Q |f|^q\Big)^{1/q}.$$

Now suppose that X_Q is a Banach space of functions on Q, with a "dual" space X'_Q in the sense that

$$\frac{1}{|Q|}\int_Q f\cdot g \leq \|f\|_{X_Q}\|g\|_{X'_Q}.$$

In the above formalism, $(L^q(Q))' = L^{q'}(Q)$ for the usual dual exponent.

Related to the dual spaces, we consider the maximal function

$$M_{X'}f(x) := \sup_{Q\ni x}\|f\|_{X'_Q}.$$

If $X'_Q = L^s(Q)$, then

$$M_{L^s}f(x) = \sup_{Q\ni x}\Big(\frac{1}{|Q|}\int_Q |f|^s\Big)^{1/s} = M(|f|^s)(x)^{1/s},$$

where M is the usual maximal function. We observe that

$$\|M_{L^s}f\|_{L^p} = \|M(|f|^s)\|^{1/s}_{L^{p/s}} \leq \|M\|^{1/s}_{L^{p/s}\to L^{p/s}}\||f|^s\|^{1/s}_{L^{p/s}} \leq \|M\|^{1/s}_{L^{p/s}\to L^{p/s}}\|f\|_{L^p}.$$

Since M is bounded on L^p if and only if $p > 1$, we see that M_{L^s} is bounded on L^p if and only if $p > s$. Now we are ready to formulate:

Theorem 13.5 (Bump theorem of Lerner; conjectured by Cruz-Uribe)**.** *For each cube Q, let X_Q, Y_Q be Banach function spaces with duals X'_Q, Y'_Q. Let w, σ be weights. Suppose that*

$$\mathscr{A}_2 := \sup_Q \|w^{1/2}\|^2_{X_Q}\|\sigma^{1/2}\|^2_{Y_Q} < \infty$$

and that $M_{X'}$ and $M_{Y'}$ are both bounded on (the unweighted) L^2. Then (13.1) *holds for all Calderón–Zygmund operators, and more precisely*

$$\|T(f\sigma)\|_{L^2(w)} \leq c_T\mathscr{A}_2^{1/2}\|M_{X'}\|_{L^2\to L^2}\|M_{Y'}\|_{L^2\to L^2}\|f\|_{L^2(\sigma)}. \tag{13.6}$$

Let us first observe that this contains Neugebauer's theorem. Indeed, in that case $X_Q = Y_Q = L^{2r}(Q)$, so that $M_{X'} = M_{Y'} = M_{L^{(2r)'}}$. Since $2r > 2$, we have $(2r)' < 2' = 2$, and hence $M_{L^{(2r)'}}$ is bounded on L^2.

Proof of Theorem 13.5. By Lerner's dyadic domination theorem, it is enough to prove (13.6) with the averaging operators $A_{\mathscr{S}}$ in place of T. By duality, we need to prove that

$$\int A_{\mathscr{S}}(f\sigma)\cdot gw \leq c\mathscr{A}_2^{1/2}\|M_{X'}\|_{L^2\to L^2}\|M_{Y'}\|_{L^2\to L^2}\|f\|_{L^2(\sigma)}\|g\|_{L^2(w)}. \tag{13.7}$$

Let us do this.

Recalling that

$$A_{\mathscr{S}}f = \sum_{S\in\mathscr{S}}\frac{1_S}{|S|}\int_S f,$$

where the cubes S have pairwise disjoint subsets $E(S)$ with $|E(S)| \geq \frac{1}{2}|S|$, we compute

$$\begin{aligned}
\int A_{\mathscr{S}}(f\sigma)\cdot gw &= \sum_{S\in\mathscr{S}} \frac{1}{|S|}\int_S f\sigma \cdot \int_S gw = \sum_{S\in\mathscr{S}} \frac{1}{|S|}\int_S f\sigma^{1/2}\sigma^{1/2}\cdot \frac{1}{|S|}\int_S gw^{1/2}w^{1/2}\cdot |S| \\
&\leq \sum_{S\in\mathscr{S}} \|f\sigma^{1/2}\|_{Y_S'}\|\sigma^{1/2}\|_{Y_S}\|gw^{1/2}\|_{X_S'}\|w^{1/2}\|_{X_S}\cdot 2|E(S)| \\
&\leq 2\Big(\sup_{S\in\mathscr{S}}\|\sigma^{1/2}\|_{Y_S}\|w^{1/2}\|_{X_S}\Big)\sum_{S\in\mathscr{S}}\|f\sigma^{1/2}\|_{Y_S'}\|gw^{1/2}\|_{X_S'}\cdot|E(S)| \\
&\leq 2\mathscr{A}_2^{1/2}\sum_{S\in\mathscr{S}}\inf_{y\in S}M_{Y'}(f\sigma^{1/2})(y)\inf_{z\in S}M_{X'}(gw^{1/2})(z)\int_{E(S)}\mathrm{d}x \\
&\leq 2\mathscr{A}_2^{1/2}\sum_{S\in\mathscr{S}}\int_{E(S)}M_{Y'}(f\sigma^{1/2})(x)M_{X'}(gw^{1/2})(x)\,\mathrm{d}x \\
&\leq 2\mathscr{A}_2^{1/2}\int_{\mathbb{R}^d}M_{Y'}(f\sigma^{1/2})(x)M_{X'}(gw^{1/2})(x)\,\mathrm{d}x \\
&\leq 2\mathscr{A}_2^{1/2}\|M_{Y'}(f\sigma^{1/2})\|_{L^2}\|M_{X'}(gw^{1/2})\|_{L^2} \\
&\leq 2\mathscr{A}_2^{1/2}\|M_{Y'}\|_{L^2\to L^2}\|f\sigma^{1/2}\|_{L^2}\|M_{X'}\|_{L^2\to L^2}\|gw^{1/2}\|_{L^2} \\
&= 2\mathscr{A}_2^{1/2}\|M_{X'}\|_{L^2\to L^2}\|M_{Y'}\|_{L^2\to L^2}\|f\|_{L^2(\sigma)}\|g\|_{L^2(w)}.
\end{aligned}$$

This proves (13.7), and therefore the theorem. □

13.C. **Some open problems.** After Lerner's proof of the bump theorem above, the question arose, whether its assumptions may be weakened as follows. Assume only the *one-sided* bump condition

$$\sup_Q \|w^{1/2}\|_{X_Q}^2\langle\sigma\rangle_Q + \sup_Q\langle w\rangle_Q\|\sigma^{1/2}\|_{Y_Q}^2 < \infty,$$

as well as $M_{X'}$ and $M_{Y'}$ bounded on L^2 as before. Is this enough to guarantee (13.1)? There are some partial positive results for particular forms of the bump spaces X_Q, Y_Q, but the problem is open in the stated generality.

There is also the following quantitative variant: Under the same assumptions as in Lerner's bump theorem, is it true that

$$\|T(f\sigma)\|_{L^2(w)} \leq c_T\mathscr{A}_2^{1/2}(\|M_{X'}\|_{L^2\to L^2} + \|M_{Y'}\|_{L^2\to L^2})\|f\|_{L^2(\sigma)}? \tag{13.8}$$

The difference is that we have replaced the product of the maximal norms by their sum.

If true, even just for the classical bump (13.4), the estimate (13.8) would imply the A_2 theorem when specialized to $\sigma = 1/w$. Namely, in the one-weight case there is a powerful *reverse Hölder inequality* which provides a certain self-improvement of the A_2 condition: the classical A_2 condition already implies the stronger version (13.4), even with comparable constant $\mathscr{A}_2 \leq 2[w]_{A_2}$, provided that $r = 1+\epsilon$ with $\epsilon \leq c_d/[w]_{A_2}$ for some small dimensional constant c_d. In this case, one can check that $\|M_{L^{(2r)'}}\|_{L^2\to L^2}$ is of the order $\epsilon^{-1/2} \eqsim [w]_{A_2}^{1/2}$. Thus, the Bump Theorem 13.5, when specialized to the one-weight case, would only give

$$\|Tf\|_{L^2(w)} \leq c_T[w]_{A_2}^{3/2}\|f\|_{L^2(w)},$$

whereas (13.8) would recover the sharp form of the A_2 theorem. At the present, the two-weight bump theory is not strong enough to recover the sharp results in the one-weight theory.

References

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

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

[CUMP10] D. Cruz-Uribe, J. M. Martell, and C. Pérez. Sharp weighted estimates for approximating dyadic operators. *Electron. Res. Announc. Math. Sci.*, 17:12–19, 2010.

[FS71] C. Fefferman and E. M. Stein. Some maximal inequalities. *Amer. J. Math.*, 93:107–115, 1971.

[Fuj91] N. Fujii. A condition for a two-weight norm inequality for singular integral operators. *Studia Math.*, 98(3):175–190, 1991.

[HMW73] R. Hunt, B. Muckenhoupt, and R. Wheeden. Weighted norm inequalities for the conjugate function and Hilbert transform. *Trans. Amer. Math. Soc.*, 176:227–251, 1973.

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

[Hyt13] T. Hytönen. The two-weight inequality for the hilbert transform with general measures. 2013. Preprint, arXiv:1312.0843.

[JT85] B. Jawerth and A. Torchinsky. Local sharp maximal functions. *J. Approx. Theory*, 43(3):231–270, 1985.

[Lac13] M. T. Lacey. Two weight inequality for the Hilbert transform: A real variable characterization, II. 2013. Preprint, arXiv:1301.4663.

[Ler08] A. K. Lerner. An elementary approach to several results on the Hardy-Littlewood maximal operator. *Proc. Amer. Math. Soc.*, 136(8):2829–2833, 2008.

[Ler10] A. K. Lerner. A pointwise estimate for the local sharp maximal function with applications to singular integrals. *Bull. Lond. Math. Soc.*, 42(5):843–856, 2010.

[Ler13] A. K. Lerner. On an estimate of Calderón-Zygmund operators by dyadic positive operators. *J. Anal. Math.*, 121:141–161, 2013.

[LSSUT12] M. T. Lacey, E. T. Sawyer, C.-Y. Shen, and I. Uriarte-Tuero. Two weight inequality for the Hilbert transform: A real variable characterization, I. 2012. Preprint, arXiv:1201.4319.

[Muc72] B. Muckenhoupt. Weighted norm inequalities for the Hardy maximal function. *Trans. Amer. Math. Soc.*, 165:207–226, 1972.

[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.