JUHA KINNUNEN

Harmonic Analysis

Department of Mathematics, Aalto University
2024

Contents

1 Calderón-Zygmund decomposition

Dyadic cubes and the Calderón-Zygmund decomposition are very useful tools in harmonic analysis. The property of dyadic cubes, that either one is contained in the other or the interiors of the cubes are disjoint, is very useful in constructing coverings with pairwise disjoint cubes. The Calderón -Zygmund decomposition gives decompositions of sets and functions into good and bad parts, which can be considered separately using real variable and harmonic analysis techniques.

1.1 Dyadic subcubes of a cube

A closed cube is a bounded interval in $\mathbb{R}^n$, whose sides are parallel to the coordinate axes and equally long, that is,

$$Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$$

with $b_1 - a_1 = \ldots = b_n - a_n$. The side length of a cube Q is denoted by $l(Q)$. In case we want to specify the center, we write

$$Q(x,l) = \left\{ y \in \mathbb{R}^n : |y_i - x_i| \leqslant \tfrac{l}{2}, \, i = 1, \ldots, n \right\}$$

for a cube with center at $x \in \mathbb{R}^n$ and side length $l > 0$. If $Q = Q(x,l)$, we denote $\alpha Q = Q(x, \alpha l)$ for $\alpha > 0$. Thus αQ the cube with the same center as Q, but the side length multiplied by factor α. The integral average of $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ in a cube Q is denoted by

$$f_Q = \fint_Q f(x)\,dx = \frac{1}{|Q|} \int_Q f(x)\,dx.$$

Let $Q = [a_1, b_1] \times \ldots \times [a_n, b_n]$ be a closed cube in $\mathbb{R}^n$ with side length $l = b_1 - a_1 = \ldots = b_n - a_n$. We decompose Q into subcubes recursively. Denote $\mathscr{D}_0 = \{Q\}$. Bisect each interval $[a_i, b_i]$, $i = 1, 2, \dots$, in to two equal parts and obtain 2^n congruent subcubes of Q. Denote this collection of cubes by $\mathscr{D}_1$. The cubes in $\mathscr{D}_1$ is a partition of Q into dyadic subcubes with pairwise disjoint interiors, that is, only the boundaries of the cubes may overlap. This does not matter, since the union of the boundaries of the cubes in D_1 is a set of measure zero. Bisect every cube in $\mathscr{D}_1$ and obtain 2^n subcubes. Denote this collection of cubes by $\mathscr{D}_2$. By continuing this way, we obtain generations of dyadic cubes $\mathscr{D}_k$, $k = 0, 1, 2, \dots$. The dyadic subcubes in $\mathscr{D}_k$ are of the form

$$\left[a_1 + \frac{m_1 l}{2^k}, a_1 + \frac{(m_1+1)l}{2^k}\right] \times \dots \times \left[a_n + \frac{m_n l}{2^k}, a_n + \frac{(m_n+1)l}{2^k}\right],$$

where $k = 0, 1, 2, \dots$ and $m_j = 0, 1, \dots, 2^k - 1$, $j = 1, \dots, n$. The collection of all dyadic subcubes of Q is

$$\mathscr{D}(Q) = \mathscr{D} = \bigcup_{k=0}^{\infty} \mathscr{D}_k.$$

A cube $Q' \in \mathscr{D}$ is called a dyadic subcube of Q. Sometimes it is convenient to consider half open cubes of the type $[a_1, b_1) \times \dots \times [a_n, b_n)$ with $b_1 - a_1 = \dots = b_n - a_n$. The corresponding dyadic subcubes are pairwise disjoint and cover the original half open cube.

THE MORAL: For many phenomena in harmonic analysis it is enough to consider dyadic cubes instead of all cubes. Dyadic cubes have a rigid recursive structure.

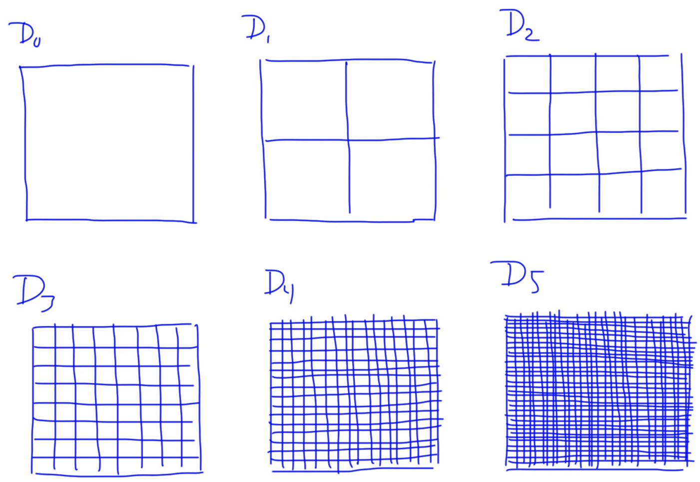

Figure 1.1: Collections of dyadic subcubes.

Remark 1.1. Dyadic subcubes of Q have the following properties:

(1) Every $Q' \in \mathscr{D}(Q)$ is a subcube of Q. Moreover, if $Q' \in \mathscr{D}(Q)$, then $\mathscr{D}(Q') \subset \mathscr{D}(Q)$.

(2) Cubes in $\mathscr{D}_k$ cover Q and the interiors of the cubes in $\mathscr{D}_k$ are pairwise disjoint for every $k = 0,1,2,\dots$.

(3) If $Q', Q'' \in \mathscr{D}$, either one is contained in the other or the interiors of the cubes are disjoint. This is called the nesting property, see Figure 1.2.

(4) If $Q' \in \mathscr{D}_k$ and $j < k$, there is exactly one ancestor cube in $\mathscr{D}_j$, which contains Q'. In particular, for every $k = 1,2,\dots$, there exists exactly one parent cube in $\mathscr{D}_{k-1}$, which contains Q'.

(5) Every cube $Q' \in \mathscr{D}_k$ is a union of exactly 2^n children cubes $Q'' \in \mathscr{D}_{k+1}$ with $|Q'| = 2^n |Q''|$.

(6) If $Q' \in \mathscr{D}_k$, then $l(Q') = 2^{-k} l(Q)$ and $|Q'| = 2^{-nk}|Q|$.

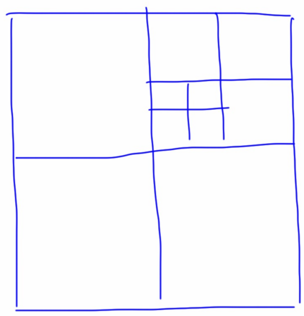

Figure 1.2: Nestedness property.

Assume that $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$. By the Lebesgue differentiation theorem

$$\lim_{r\to 0} \fint_{B(x,r)} |f(y) - f(x)|\,dy = 0 \quad \text{for almost every} \quad x \in \mathbb{R}^n. \tag{1.2}$$

A point $x \in \mathbb{R}^n$, at which (1.2) holds, is called a Lebesgue point of f. For every Lebesgue point x we have

$$\lim_{r\to 0} \fint_{B(x,r)} f(y)\,dy = f(x),$$

since

$$\left| \fint_{B(x,r)} f(y)\,dy - f(x) \right| \leqslant \fint_{B(x,r)} |f(y)-f(x)|\,dy \xrightarrow{r\to 0} 0.$$

Moreover, every Lebesgue point x of f is a Lebesgue point of $|f|$, since

$$\fint_{B(x,r)} ||f(y)|-|f(x)||\,dy \leqslant \fint_{B(x,r)} |f(y)-f(x)|\,dy \xrightarrow{r\to 0} 0.$$

We shall need the following version of the Lebesgue differentiation theorem. It can be proved by applying the weak type estimate for the dyadic maximal function, see Remarks 1.24 (1), but we show that it follows from the standard version of the Lebesgue differentiation theorem with balls.

Lemma 1.3. Assume that $x \in \mathbb{R}^n$ is a Lebesgue point of $f \in L^1_{\text{loc}}(\mathbb{R}^n)$. Then

$$\lim_{i\to\infty} \frac{1}{|Q_i|} \int_{Q_i} f(y)\,dy = f(x)$$

whenever $Q_1, Q_2, Q_3, \dots$ is any sequence of cubes containing x such that $\lim_{i\to\infty} |Q_i| = 0$.

THE MORAL: The Lebesgue differentiation theorem does not only hold for balls but also for cubes and dyadic cubes.

Proof. Let $Q_i = Q(x_i, l_i)$, where $x_i \in \mathbb{R}^n$ is the center and $l_i = l(Q_i)$ is the side lenght of the cube Q_i for every $i = 1,2,\dots$. We observe that $Q(x_i,l_i) \subset B(x,\sqrt{n}l_i)$ for every $i = 1,2,\dots$.

This implies

$$\begin{aligned}\left| \fint_{Q(x_i,l_i)} f(y)\,dy - f(x) \right| &\leqslant \fint_{Q(x_i,l_i)} |f(y)-f(x)|\,dy \\ &\leqslant \frac{|B(x,\sqrt{n}l_i)|}{|Q(x_i,l_i)|} \fint_{B(x,\sqrt{n}l_i)} |f(y)-f(x)|\,dy \\ &= |B(0,1)| n^{\frac{n}{2}} \fint_{B(x,\sqrt{n}l_i)} |f(y)-f(x)|\,dy \xrightarrow{i\to\infty} 0,\end{aligned}$$

since $l_i \to 0$ as $i \to \infty$. □

The following Calderón-Zygmund decomposition will be extremely useful in harmonic analysis.

Theorem 1.4 (Calderón-Zygmund decomposition of a cube (1952)). Assume that $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ and let Q be a cube in $\mathbb{R}^n$. Then for every

$$t \geqslant \fint_Q |f(y)|\,dy$$

there are countably many dyadic subcubes Q_i, $i = 1,2,\dots$, of Q such that

(1) the interiors of Q_i, $i = 1,2,\dots$, are pairwise disjoint,

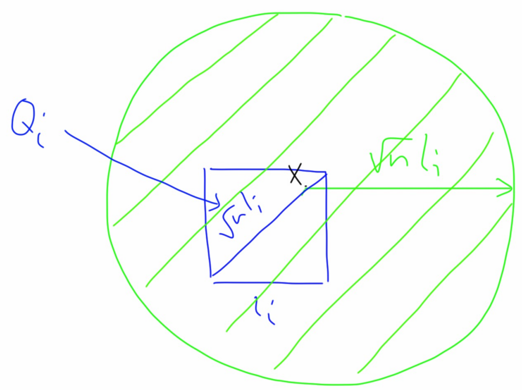

Figure 1.3: $Q(x_i, l_i) \subset B(x, \sqrt{n} l_i)$ for every $i = 1, 2, \ldots$.

(2) $t < \fint_{Q_i} |f(y)|\, dy \leqslant 2^n t$ for every $i = 1, 2, \ldots$ and

(3) $|f(x)| \leqslant t$ for almost every $x \in Q \setminus \bigcup_{i=1}^{\infty} Q_i$.

The collection of cubes Q_i, $i = 1, 2, \ldots$, is called the Calderón-Zygmund cubes in Q at level t.

T H E M O R A L : A cube can be divided into good and bad parts so that in the good part (complement of the Calderón-Zygmund cubes) the function is small and in the bad part (union of the Calderón-Zygmund cubes) the integral average of a function is in control. Note that the Calderón-Zygmund cubes cover the set $\{x \in Q : |f(x)| > t\}$, up to a set of measure zero, and thus the bad part contains the set where the function is unbounded.

Proof. The strategy of the proof is the following stopping time argument. For every $x \in Q$ such that $|f(x)| > t$ we choose the largest dyadic cube $Q' \in \mathscr{D}$ containing x such that

$$\fint_{Q'} |f(y)|\, dy > t.$$

Then we use the fact that for any collection of dyadic subcubes of Q there is a subcollection of dyadic cubes with disjoint interiors and with the same union as the original cubes. These are the desired Calderón-Zygmund cubes.

Then we give a rigorous argument. Consider (possibly empty) collection $\mathscr{Q}'$ of dyadic subcubes $Q' \in \mathscr{D}$ of Q, that satisfy

$$\fint_{Q'} |f(y)|\,dy > t. \tag{1.5}$$

□

The cubes in $\mathcal{Q}'$ are not necessarily pairwise disjoint, but we consider a collection of maximal dyadic cubes with respect to inclusion for which (1.5) holds true. For every $Q' \in \mathcal{Q}'$ we consider all cubes $Q'' \in \mathcal{Q}'$ with $Q' \subset Q''$. The maximal cube Q_i is the union of all dyadic subcubes of Q which satisfy (1.5) and contain Q'. Nestedness property of the dyadic subcubes, see Remark 1.1 (3), implies that $Q_i \in \mathscr{D}$. Let $\mathcal{Q} = \{Q_i\}_i$ be the collection of these maximal cubes. Maximality means that $\fint_R |f(y)|\,dy \leqslant t$ for every $R \supset Q_i$, $R \in \mathscr{D}$. Since $\fint_Q |f(y)|\,dy \leqslant t$, for every cube $Q' \in \mathcal{Q}'$ there exists a maximal cube $Q_i \in \mathcal{Q}'$. We show that this collection has the desired properties.

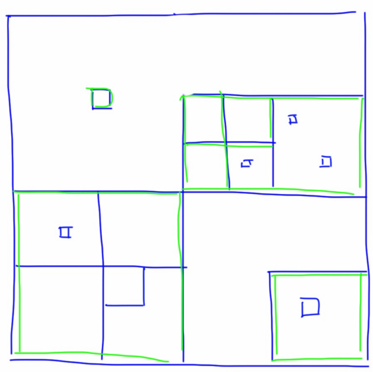

Figure 1.4: Collection of maximal subcubes.

$\boxed{(1)}$ This follows immediately from maximality of the cubes in $\mathcal{Q}'$ and nestedness property of the dyadic subcubes, see Remark 1.1 (3). Indeed, if the interiors of two different cubes in $\mathcal{Q}$ intersect then one is contained in the other, and hence one of them cannot be maximal, see Figure 1.4.

$\boxed{(2)}$ By (1.5) we have $Q \notin \mathcal{Q}'$. If $Q_i \in \mathcal{Q}' \cap \mathscr{D}_k$ for some k, then by properties (4) and (5) of the dyadic subcubes we conclude that Q_i is contained in some cube $Q' \in \mathscr{D}_{k-1}$ with $|Q'| = 2^n |Q_i|$, see Figure 1.5. Since Q_i maximal, cube Q' does not satisfy (1.5). Thus

$$t < \frac{1}{|Q_i|} \int_{Q_i} |f(y)|\,dy \leqslant \frac{|Q'|}{|Q_i|} \frac{1}{|Q'|} \int_{Q'} |f(y)|\,dy \leqslant 2^n t.$$

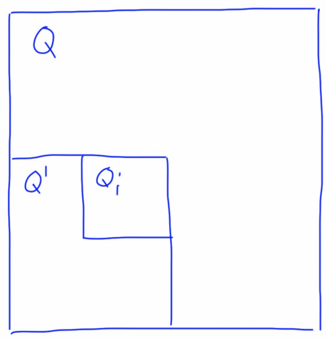

Figure 1.5: Q_i is contained in some cube $Q' \in \mathscr{D}_{k-1}$ with $|Q'| = 2^n |Q_i|$.

$\boxed{(3)}$ Assume that $x \in Q \setminus \bigcup_{i=1}^{\infty} Q_i$. By the beginning of the proof,

$$t \geqslant \fint_{Q'} |f(y)|\,dy$$

for every dyadic subcube $Q' \in \mathscr{D}$ containing point x. Thus there exist $Q'_k \in \mathscr{D}_k$ such that $x \in Q'_k$ for every $k = 1, 2, \dots$. Note that $Q'_1 \supset Q'_2 \supset Q'_3 \dots$ and $\bigcap_{i=1}^{\infty} Q'_i = \{x\}$, see the Figure 1.6. If x is a Lebesgue point of f, Lemma 1.3 implies

$$|f(x)| = \lim_{k \to \infty} \frac{1}{|Q'_k|} \int_{Q'_k} |f(y)|\,dy \leqslant t.$$

Remark 1.6. If $f \in L^\infty(Q)$ and $t \geqslant \operatorname{ess\,sup}_{y \in Q} |f(y)|$, then the collection Calderón-Zygmund cubes is empty, since

$$\fint_{Q'} |f(y)|\,dy \leqslant \operatorname*{ess\,sup}_{y \in Q} |f(y)| \leqslant t$$

for every dyadic subcube Q' of Q.

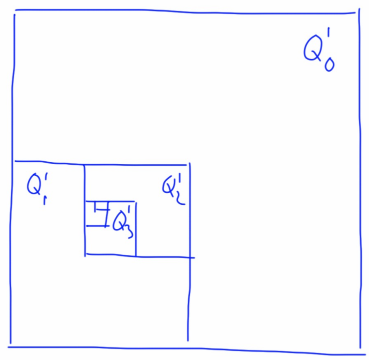

Figure 1.6: $Q_k' \in \mathscr{D}_k$ such that $x \in Q_k'$ for every $k = 1, 2, \ldots$.

Remark 1.7. The assumption $t \geqslant \fint_Q |f(y)|\,dy$ implies that the original cube Q cannot belong to the collection of the Calderón-Zygmund cubes at the level t. In other words, if the collection of the Calderón-Zygmund cubes $\{Q_i\}_i$ is nonempty, the cubes are proper subcubes of Q, that is $Q_i \subset Q$ and $Q_i \neq Q$ for every $i = 1, 2, \ldots$.

Remark 1.8. Let E be a measurable subset of a cube $Q \subset \mathbb{R}^n$ with $|E| \leqslant t|Q|$, $0 < t < \frac{1}{2^n}$. By applying the Calderón-Zygmund decomposition to $f = \chi_E$ at the level $t \geqslant \frac{|E|}{|Q|}$, we obtain countably many dyadic subcubes Q_i, $i = 1, 2, \ldots$, of Q such that

(1) the interiors of Q_i, $i = 1, 2, \ldots$, are pairwise disjoint,

(2) $t < \dfrac{|Q_i \cap E|}{|Q_i|} \leqslant 2^n t$ for every $i = 1, 2, \ldots$ and

(3) $\left|E \setminus \bigcup_{i=1}^\infty Q_i\right| = 0$.

THE MORAL: The set E can be covered, up to a set of measure zero, by pairwise almost disjoint dyadic subcubes Q_i, $i = 1, 2, \ldots$, of Q such that every cube Q_i intersects substantially both E and the complement of E, that is, $|Q_i \cap E| > t|Q_i|$ and $|Q_i \setminus E| = |Q_i| - |Q_i \cap E| \geqslant (1 - 2^n t)|Q_i|$ for every i.

Moreover, we have

$$t\left|\bigcup_{i=1}^\infty Q_i\right| = t\sum_{i=1}^\infty |Q_i| < \sum_{i=1}^\infty |Q_i \cap E| = \left|\bigcup_{i=1}^\infty (E \cap Q_i)\right| = |E|$$

and

$$\begin{aligned}|E| &= \left|E \cap \bigcup_{i=1}^{\infty} Q_i\right| + \left|E \setminus \bigcup_{i=1}^{\infty} Q_i\right| = \left|\bigcup_{i=1}^{\infty} (Q_i \cap E)\right| \\ &\leqslant \sum_{i=1}^{\infty} |Q_i \cap E| \leqslant 2^n t \sum_{i=1}^{\infty} |Q_i| = 2^n t \left|\bigcup_{i=1}^{\infty} Q_i\right|.\end{aligned}$$

THE MORAL: The Lebesgue measure of E is comparable to the Lebesgue measure of $\bigcup_{i=1}^{\infty} Q_i$.

1.2 Dyadic cubes of $\mathbb{R}^n$

Next we consider the dyadic cubes in $\mathbb{R}^n$ and a global version of the Calderón-Zygmund decomposition. A half open dyadic cube in $\mathbb{R}$ is an interval of the form

$$[m2^{-k}, (m+1)2^{-k}),$$

where $m, k \in \mathbb{Z}$. The advantage of considering half open cubes is that they are pairwise disjoint. A dyadic cube of $\mathbb{R}^n$ is a cartesian product of one-dimensional dyadic cubes

$$\prod_{j=1}^{n} [m_j 2^{-k}, (m_j+1)2^{-k}),$$

where $m_1, \dots, m_n$, $k \in \mathbb{Z}$. The collection of dyadic cubes $\mathcal{D}_k$, $k \in \mathbb{Z}$, consists of the dyadic cubes with the side length 2^{-k}. The collection of all dyadic cubes in $\mathbb{R}^n$ is

$$\mathcal{D}(\mathbb{R}^n) = \mathcal{D} = \bigcup_{k \in \mathbb{Z}} \mathcal{D}_k.$$

Observe that $\mathcal{D}_k$ consist of cubes whose vertices lie on the lattice $2^{-k}\mathbb{Z}^n$ and whose side length is 2^{-k}. The dyadic cubes in the kth generation can be defined as $\mathcal{D}_k = 2^{-k}([0,1)^n + \mathbb{Z}^n)$. The cubes in $\mathcal{D}_k$ cover the whole $\mathbb{R}^n$ and are pairwise disjoint, see Figure 1.7. Moreover, the dyadic cubes have the same properties (2)-(5) in Remark 1.1 as the dyadic subcubes of a given cube.

WARNING: It is not true that every cube is a subcube of a dyadic cube. For example, consider $[-1,1]^n$. However, there is a substitute for this property: For every cube Q there is a dyadic cube $Q' \in \mathcal{D}$ such that $Q' \subset Q \subset 3Q'$, see the discussion in Section 1.4.

Remarks 1.9:

(1) For any subcollection $\mathcal{Q} \subset \mathcal{D}$ of dyadic cubes whose union is a bounded set, or a set of finite Lebesgue measure, there exists a subcollection of pairwise disjoint maximal cubes with the same union (exercise). Note that $[0,2^k)^n$, $k = 0,1,2,\dots$, is a collection of dyadic cubes for which there does not exist

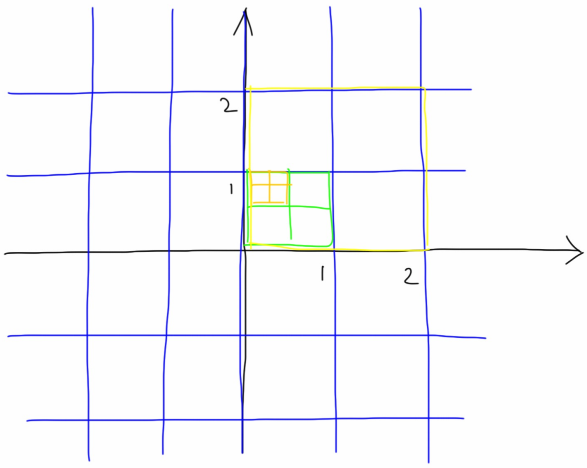

Figure 1.7: Dyadic cubes in $\mathbb{R}^n$.

a subcollection of pairwise disjoint maximal cubes with the same union. This shows that some kind of boundedness assumption is needed above. A cube $Q' \in \mathscr{Q}$ is called maximal, if there does not exist any strictly larger $Q \in \mathscr{Q}$ with $Q' \subset Q$, see Figure 1.8. A useful property is that the collection maximal cubes are always pairwise disjoint. This follows at once from nestedness property of the dyadic cubes. Indeed, if two different cubes in $\mathscr{Q}$ satisfy $Q \cap Q' \neq \emptyset$, then one is contained in the other, and hence one of them cannot be maximal.

(2) There are many ways to construct dyadic cubes in $\mathbb{R}^n$, see [15]. For example, we may start with any cube Q_0 and the corresponding dyadic subcubes $\mathscr{D}(Q_0)$. Then we may take any increasing sequence of cubes $Q_0 \subset Q_1 \subset Q_2 \subset \dots$ so that Q_k is a dyadic child of Q_{k+1} and $\bigcup_{k=1}^{\infty} Q_k = \mathbb{R}^n$. In other words, for any Q_k we choose one of 2^n dyadic parents Q_{k+1} in such a way that the the union if the selected cubes eventually covers the entire space. Note that $Q_k \in \mathscr{D}(Q_{k+1})$ and thus $\mathscr{D}(Q_k) \subset \mathscr{D}(Q_{k+1})$. Let $\mathscr{D}(\mathbb{R}^n) = \mathscr{D} = \bigcup_{k=0}^{\infty} \mathscr{D}(Q_k)$. The obtained collection $\mathscr{D}$ satisfies properties (2)-(5) in Remark 1.1 (exercise). In addition, for every compact set K, there exists a cube in $\mathscr{D}$ containing K. Note that the standard collection of dyadic cubes does not satisfy this property. For example, if $K = [-1,1]^n$, then there does not exist a standard dyadic cube that covers K.

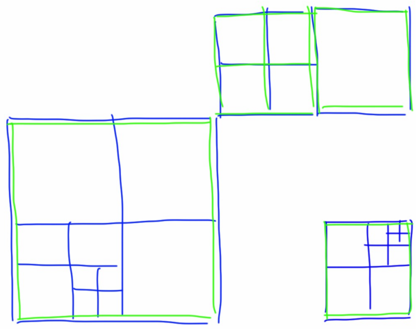

Figure 1.8: A collection of maximal cubes.

(3) Assume that $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$. Then

$$E_k f(x) = \sum_{Q \in \mathscr{D}_k} \left(\fint_Q f(y)\,dy \right) \chi_Q(x)$$

is the conditional expectation of f with respect to a σ-algebra generated by $\mathscr{D}_k$, $k \in \mathbb{Z}$. Note that

$$\begin{aligned}
\int_{\mathbb{R}^n} E_k f(x)\,dx &= \int_{\mathbb{R}^n} \sum_{Q \in \mathscr{D}_k} \left(\fint_Q f(y)\,dy \right) \chi_Q(x)\,dx \\
&= \sum_{Q \in \mathscr{D}_k} \left(\fint_Q f(y)\,dy \right) \int_{\mathbb{R}^n} \chi_Q(x)\,dx \\
&= \sum_{Q \in \mathscr{D}_k} \int_Q f(y)\,dy = \int_{\mathbb{R}^n} f(x)\,dx
\end{aligned}$$

for every $k \in \mathbb{Z}$ and E_k can be considered as a discrete analog of convolution approximation.

In the one dimensional case every nonempty open set is a union of countably many disjoint open intervals and the Lebesgue outer measure of an open set is the sum of volumes of these intervals. Next we consider this question in the higher dimensional case.

Lemma 1.10. Every nonempty open set in $\mathbb{R}^n$ is a union of countably many pairwise disjoint dyadic cubes.

Proof. Let Ω be a nonempty open set in $\mathbb{R}^n$. Consider dyadic cubes in $\mathscr{D}_1$ that are contained in Ω and denote $\mathscr{Q}_1 = \{Q \in \mathscr{D}_1 : Q \subset \Omega\}$. Then consider dyadic cubes in $\mathscr{Q}_2$ that are contained in Ω and do not intersect any of the cubes in $\mathscr{Q}_1$ and denote

$$\mathscr{Q}_2 = \{Q \in \mathscr{D}_2 : Q \subset \Omega,\ Q \cap J = \emptyset \text{ for every } J \in \mathscr{Q}_1\}.$$

Recursively define

$$\mathscr{Q}_k = \left\{Q \in \mathscr{D}_k : Q \subset \Omega,\ Q \cap J = \emptyset \text{ for every } J \in \bigcup_{i=1}^{k-1} \mathscr{Q}_i\right\}$$

for every $k = 2,3,\dots$. Then $\mathscr{Q} = \bigcup_{k=1}^{\infty} \mathscr{Q}_k$ is a countable collection of pairwise disjoint dyadic cubes.

Claim: $\Omega = \bigcup_{Q \in \mathscr{Q}} Q$.

Reason. It is clear from the construction that $\bigcup_{Q\in\mathscr{Q}} Q \subset \Omega$. For the reverse inclusion, let $x \in \Omega$. Let k be so large that the common diameter of the cubes in $\mathscr{D}_k$ is smaller than r, that is, $\sqrt{n}2^{-k} < r$. Since Ω is open, there exists a ball $B(x,r) \subset \Omega$ with $r > 0$. Since the dyadic $\mathscr{D}_k$ cubes cover $\mathbb{R}^n$, there exists a dyadic cube $Q \in \mathscr{D}_k$ with $x \in Q$ and $Q \subset B(x,r) \subset \Omega$. There are two possibilities $Q \in \mathscr{Q}_k$ or $Q \notin \mathscr{Q}_k$. If $Q \in \mathscr{Q}_k$, then $x \in Q \subset \bigcup_{Q\in\mathscr{Q}} Q$. If $Q \notin \mathscr{Q}_k$, there exists $J \in \bigcup_{i=1}^{k-1} \mathscr{Q}_i$ with $J \cap Q \neq \emptyset$. The nesting property of dyadic cubes implies $Q \subset J$ and $x \in Q \subset J \subset \bigcup_{Q\in\mathscr{Q}} Q$. ∎

Remark 1.11. The Whitney decomposition of a nonempty proper open subset Ω of $\mathbb{R}^n$ states that it can be represented as a union of countably many pairwise disjoint dyadic intervals whose side lengths are comparable to their distance to the boundary of the open set. More precisely, there are pairwise disjoint dyadic cubes Q_i, $i = 1,2,\dots$, such that

- $\Omega = \bigcup_{i=1}^{\infty} Q_i$,
- $\sqrt{n}\, l(Q_i) \leqslant \operatorname{dist}(Q_i, \mathbb{R}^n \setminus \Omega) \leqslant 4\sqrt{n}\, l(Q_i)$,
- if the boundaries of Q_i and Q_j touch, then $\frac{1}{4} \leqslant \frac{l(Q_i)}{l(Q_j)} \leqslant 4$,
- for every Q_i there exist at most 12^n cubes in the collection that touch it.

See [11, Proposition 7.3.4].

Theorem 1.12 (Global Calderón-Zygmund decomposition (1952)). Assume that $f \in L^1(\mathbb{R}^n)$. Then for every $t > 0$ there are countably or finitely many dyadic cubes Q_i, $i = 1,2,\dots$, in $\mathbb{R}^n$ such that

(1) Q_i, $i = 1,2,\dots$, are pairwise disjoint,

(2) $t < \fint_{Q_i} |f(y)|\,dy \leqslant 2^n t$ for every $i = 1,2,\dots$ and

(3) $|f(x)| \leqslant t$ for almost every $x \in \mathbb{R}^n \setminus \bigcup_{i=1}^{\infty} Q_i$.

The collection of cubes Q_i, $i = 1,2,\dots$, is called the Calderón-Zygmund cubes in $\mathbb{R}^n$ at level t.

THE MORAL: The difference to the Calderón-Zygmund in a cube is that we assume global integrability instead of local integrability. With this assumption, we obtain the Calderón-Zygmund decomposition at every level $t > 0$. Note that, if the function belongs to $L^\infty(\mathbb{R}^n)$, the Calderón-Zygmund decomposition is empty for $t \geqslant \|f\|_\infty$, see Remark 1.6.

Proof. As in the proof of Theorem 1.4, consider the (possibly empty) collection $\mathscr{D}'$ of dyadic cubes $Q' \in \mathscr{D}$ in $\mathbb{R}^n$ that satisfy $\fint_{Q'} |f(y)|\,dy > t$. Note that

$$l(Q')^n = |Q'| < \frac{1}{t}\int_{Q'} |f(y)|\,dy \leqslant \frac{1}{t}\int_{\mathbb{R}^n} |f(y)|\,dy$$

for every cube $Q' \in \mathscr{D}'$. Thus for every cube $Q' \in \mathscr{D}'$ there exists a maximal cube $Q_i \in \mathscr{D}'$. Otherwise, the proof is similar as the proof of Theorem 1.4. □

Remark 1.13. Instead of assuming $f \in L^1(\mathbb{R}^n)$ we may assume $f \in L^p(\mathbb{R}^n)$ for some $1 \leqslant p < \infty$ in Theorem 1.12. To see this, we apply Hölder's inequality to obtain

$$t < \fint_{Q'} |f(y)|\,dy \leqslant \left(\fint_{Q'} |f(y)|^p\,dy\right)^{\frac{1}{p}}$$

for every cube $Q' \in \mathscr{D}'$ in the proof of Theorem 1.12. Then the existence of a maximal cube can be concluded from

$$l(Q')^n = |Q'| < \frac{1}{t^p}\int_{Q'} |f(y)|^p\,dy \leqslant \frac{1}{t^p}\int_{\mathbb{R}^n} |f(y)|^p\,dy$$

for every cube $Q' \in \mathscr{D}'$.

Example 1.14. Consider the Calderón-Zygmund decomposition for $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \chi_{[0,1]^n}(x)$ at level $t > 0$. We may assume that $0 < t < 1 = \|f\|_\infty$, since $\fint_Q |f(x)|\,dx \leqslant 1$ for every interval $Q \subset \mathbb{R}^n$. In other words, if $t \geqslant 1$, there are no intervals Q in $\mathbb{R}^n$ for which $\fint_Q |f(x)|\,dx > t$ and the Calderón-Zygmund decomposition is empty. For $0 < t < 1$, choose $k \in \{0,1,2,\dots\}$ such that $2^{-n(k+1)} \leqslant t < 2^{-nk}$. We claim that the Calderón-Zygmund decomposition at level t consists only of one dyadic interval $[0,2^k)^n$. To see this, we observe that

$$\frac{1}{|[0,2^k)^n|}\int_{[0,2^k)^n} |f(x)|\,dx = 2^{-nk}\int_{[0,2^k)^n} \chi_{[0,1]^n}(x)\,dx = 2^{-nk} > t.$$

On the other hand, if Q' is a dyadic ancestor of Q, that is $Q \subset Q'$, $Q \neq Q'$, where $Q' \in \mathscr{D}$, then $Q' = [0,2^{k+l})^n$, $l \in \{1,2,\dots\}$, and thus

$$\frac{1}{|Q'|}\int_{Q'} |f(x)|\,dx = 2^{-n(k+l)} \leqslant 2^{-n(k+1)} \leqslant t$$

so that $[0,2^k)^n$ is the maximal dyadic cube with the property $\fint_Q |f(x)|\,dx > t$.

1.3 Calderón-Zygmund decomposition of a function

For a function $f \in L^1(\mathbb{R}^n)$, and any level $t > 0$, we have the decomposition

$$f = f\chi_{\{|f|\leqslant t\}} + f\chi_{\{|f|>t\}} \tag{1.15}$$

into good part $g = f\chi_{\{|f|\leqslant t\}}$, which is bounded, and the bad part $b = f\chi_{\{|f|>t\}}$. These parts can be analyzed separately using real variable techniques. For the good part we have the bounds

$$\|g\|_1 = \int_{\mathbb{R}^n} |g(x)|\,dx = \int_{\{|f|\leqslant t\}} |f(x)|\,dx \leqslant \int_{\mathbb{R}^n} |f(x)|\,dx = \|f\|_1 \quad \text{and} \quad \|g\|_\infty \leqslant t$$

and for the bad part

$$\|b\|_1 \leqslant \|f\|_1 \quad \text{and} \quad |\{x \in \mathbb{R}^n : b(x) \neq 0\}| \leqslant |\{x \in \mathbb{R}^n : |f(x)| > t\}| \leqslant \frac{1}{t}\|f\|_1.$$

The last bound follows from Chebyshev's inequality and tells that the measure of the support of the bad part is small. This truncation method is will be useful in later in connection with interpolation, see Lemma 2.1, but here we consider a more refined way to decompose an arbitrary integrable function into its good and large bad parts so that not only the absolute value but also the local oscillation is in control.

Theorem 1.16 (Calderón-Zygmund decomposition of a function (1952)). Assume that $f \in L^1(\mathbb{R}^n)$ and let $t > 0$. Then there are functions g and b, and countably or finitely many pairwise disjoint dyadic cubes Q_i, $i = 1, 2, \ldots$, such that

(1) $f = g + b$,

(2) $\|g\|_1 \leqslant \|f\|_1$,

(3) $\|g\|_\infty \leqslant 2^n t$,

(4) $b = \sum_{i=1}^{\infty} b_i$, where $b_i = 0$ in $\mathbb{R}^n \setminus Q_i$, $i = 1, 2, \ldots$,

(5) $\fint_{Q_i} b_i(x)\,dx = 0$, $i = 1, 2, \ldots$,

(6) $\fint_{Q_i} |b_i(x)|\,dx \leqslant 2^{n+1} t$ and

(7) $\left|\bigcup_{i=1}^{\infty} Q_i\right| \leqslant \frac{1}{t}\|f\|_1$.

THE MORAL: Any function $f \in L^1(\mathbb{R}^n)$ can be represented as a sum of a good and a bad function $f = g + b$, where g is bounded and $b = \sum_{i=1}^{\infty} b_i$, where b_i, $i = 1, 2, \ldots$, is a highly oscillating localized function with integral average zero. Note that the bad function b contains the unbounded part of function f.

Remarks 1.17:

(1) It follows from (1) and (2) that

$$\|b\|_1 \leqslant \|f-g\|_1 \leqslant \|f\|_1 + \|g\|_1 \leqslant 2\|f\|_1$$

and thus $b \in L^1(\mathbb{R}^n)$. This shows that the bad function b is integrable.

(2)

$$\begin{aligned}\|g\|_p &= \left(\int_{\mathbb{R}^n} |g(x)|^p\,dx\right)^{\frac{1}{p}} = \left(\int_{\mathbb{R}^n} |g(x)|^{p-1}|g(x)|\,dx\right)^{\frac{1}{p}} \\ &\leqslant \left(\int_{\mathbb{R}^n} \|g\|_\infty^{p-1}|g(x)|\,dx\right)^{\frac{1}{p}} \leqslant \|g\|_1^{\frac{1}{p}}\|g\|_\infty^{1-\frac{1}{p}} \\ &\leqslant \|g\|_1^{\frac{1}{p}}(2^n t)^{1-\frac{1}{p}} \leqslant \|f\|_1^{\frac{1}{p}}(2^n t)^{1-\frac{1}{p}}\end{aligned}$$

and thus $g \in L^p(\mathbb{R}^n)$ whenever $1 \leqslant p \leqslant \infty$. This shows that the good function g is essentially bounded and belongs to all L^p-spaces.

Proof. Let Q_i, $i = 1,2,\ldots$, be the Calderón-Zygmund cubes for f at level $t > 0$, see Theorem 1.12. Let

$$g(x) = \begin{cases} f(x), & x \in \mathbb{R}^n \setminus \bigcup_{i=1}^{\infty} Q_i, \\ \fint_{Q_i} f(y)\,dy, & x \in Q_i, \quad i = 1,2,\ldots, \end{cases}$$

that is,

$$\begin{aligned} g(x) &= f(x)\chi_{\mathbb{R}^n\setminus\bigcup_{i=1}^{\infty} Q_i}(x) + \sum_{i=1}^{\infty} f_{Q_i}\chi_{Q_i}(x) \\ &= f(x) - \sum_{i=1}^{\infty}(f(x) - f_{Q_i})\chi_{Q_i}(x) \end{aligned}$$

and

$$b(x) = f(x) - g(x) = \sum_{i=1}^{\infty}(f(x) - f_{Q_i})\chi_{Q_i}(x) = \sum_{i=1}^{\infty} b_i(x)$$

with $b_i(x) = (f(x) - f_{Q_i})\chi_{Q_i}(x)$, $i = 1,2,\ldots$.

THE MORAL: The function g is defined so that it is equal to f outside the Calderón-Zygmund cubes and in a Calderón-Zygmund cube the it is the average of the function in that cube.

$\boxed{(1)}$ $f = g + b$ follows from the construction above.

$\boxed{(2)}$

$$\begin{aligned}
\int_{\mathbb{R}^n} |g(x)|\,dx &= \int_{\mathbb{R}^n\setminus\bigcup_{i=1}^\infty Q_i} |g(x)|\,dx + \int_{\bigcup_{i=1}^\infty Q_i} |g(x)|\,dx \\
&= \int_{\mathbb{R}^n\setminus\bigcup_{i=1}^\infty Q_i} |f(x)|\,dx + \sum_{i=1}^\infty \int_{Q_i} |g(x)|\,dx \\
&= \int_{\mathbb{R}^n\setminus\bigcup_{i=1}^\infty Q_i} |f(x)|\,dx + \sum_{i=1}^\infty \int_{Q_i} |f_{Q_i}|\,dx \\
&\leqslant \int_{\mathbb{R}^n\setminus\bigcup_{i=1}^\infty Q_i} |f(x)|\,dx + \sum_{i=1}^\infty \int_{Q_i} |f|_{Q_i}\,dx \\
&= \int_{\mathbb{R}^n\setminus\bigcup_{i=1}^\infty Q_i} |f(x)|\,dx + \sum_{i=1}^\infty \fint_{Q_i} |f(x)|\,dx|Q_i| \\
&= \int_{\mathbb{R}^n\setminus\bigcup_{i=1}^\infty Q_i} |f(x)|\,dx + \int_{\bigcup_{i=1}^\infty Q_i} |f(x)|\,dx = \int_{\mathbb{R}^n} |f(x)|\,dx.
\end{aligned}$$

$\boxed{(3)}$ By Theorem 1.12, we have $|f(x)| \leqslant t$ for almost every $x \in \mathbb{R}^n \setminus \bigcup_{i=1}^\infty Q_i$ and

$$\left|\fint_{Q_i} f(y)\,dy\right| \leqslant \fint_{Q_i} |f(y)|\,dy \leqslant 2^n t, \quad i = 1,2,\dots.$$

This implies that $|g(x)| \leqslant 2^n t$ for almost every $x \in \mathbb{R}^n$.

$\boxed{(4)}$ See the proof of (1).

$\boxed{(5)}$

$$\begin{aligned}
\fint_{Q_i} b_i(x)\,dx &= \fint_{Q_i} (f(x) - f_{Q_i})\chi_{Q_i}(x)\,dx \\
&= \fint_{Q_i} f(x)\,dx - f_{Q_i} = 0, \quad i = 1,2,\dots.
\end{aligned}$$

$\boxed{(6)}$ By Theorem 1.12 , we have

$$\begin{aligned}
\int_{Q_i} |b_i(x)|\,dx &\leqslant \int_{Q_i} (|f(x)| + |f|_{Q_i})\,dx \\
&\leqslant 2\int_{Q_i} |f(x)|\,dx \leqslant 2^{n+1} t |Q_i|, \quad i = 1,2,\dots.
\end{aligned}$$

$\boxed{(7)}$

$$\left|\bigcup_{i=1}^\infty Q_i\right| = \sum_{i=1}^\infty |Q_i| \leqslant \frac{1}{t}\sum_{i=1}^\infty \int_{Q_i} |f(y)|\,dy = \frac{1}{t}\|f\|_1.$$

□

1.4 Dyadic maximal function on $\mathbb{R}^n$

There is an interpretation of the Calderón-Zygmund decomposition in terms of maximal functions. The dyadic maximal function of $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ is

$$M_d f(x) = \sup \fint_Q |f(y)|\,dy, \tag{1.18}$$

where the supremum is taken over all dyadic cubes Q containing x. By Lemma 1.3, for almost every $x \in \mathbb{R}^n$, we have

$$|f(x)| = \lim_{k\to\infty} \fint_{Q_k} |f(y)|\,dy \leqslant M_d f(x),$$

where $x \in Q_k \in \mathscr{D}_k$. Thus the dyadic maximal function is bigger than the absolute value of the function almost everywhere. This explains the name maximal function.

Example 1.19. Let $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = \chi_{[0,1]^n}(x)$ as in Example 1.14. For

$$x \in \{y = (y_1,\dots,y_n) \in \mathbb{R}^n : y_i \geqslant 0 \text{ for every } i = 1,\dots,n\},$$

consider the smallest integer $k \in \{0,1,2,\dots\}$ with $x \in [0,2^k)^n$, that is, choose the smallest dyadic cube $[0,2^k)^n$ with with side length at least one and the corner at the origin containing the point x. Then (exercise)

$$M_d f(x) = \fint_{[0,2^k)^n} |f(y)|\,dy = 2^{-nk}$$

and $M_d f(x) = 0$ for every

$$x \in \mathbb{R}^n \setminus \{y = (y_1,\dots,y_n) \in \mathbb{R}^n : y_i \geqslant 0 \text{ for every } i = 1,\dots,n\}.$$

WARNING: The dyadic maximal function is not comparable to the standard Hardy-Littlewood maximal function

$$Mf(x) = \sup \fint_Q |f(y)|\,dy, \tag{1.20}$$

where the supremum is taken over all cubes Q in $\mathbb{R}^n$ containing x.

Example 1.21. It is clear that $M_d f(x) \leqslant Mf(x)$ for every $x \in \mathbb{R}^n$, but the inequality in the reverse direction does not hold. Consider $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = \chi_{[0,1]^n}(x)$ as in Example 1.19. Then $M_d f(x) = 0$ and $Mf(x) > 0$ for every

$$x \in \mathbb{R}^n \setminus \{y = (y_1,\dots,y_n) \in \mathbb{R}^n : y_i \geqslant 0 \text{ for every } i = 1,\dots,n\}.$$

Note that it is possible to compare integral averages over cubes $Q_1 \subset Q \subset Q_2$ by observing that

$$\frac{|Q_1|}{|Q|} \fint_{Q_1} |f(y)|\,dy \leqslant \fint_Q |f(y)|\,dy \leqslant \frac{|Q_2|}{|Q|} \fint_{Q_2} |f(y)|\,dy.$$

The challenge is to bound the volume ratios $\frac{|Q_1|}{|Q|}$ and $\frac{|Q_2|}{|Q|}$. For a cube $Q = Q(x,l)$, $l > 0$, we may take a dyadic cube $Q_1 \in \mathscr{D}_k$ containing the center x of Q with the side length 2^{-k} where $\frac{l}{4} < 2^{-k} \leqslant \frac{l}{2}$. Then $Q_1 \subset Q$ and

$$0 < \frac{1}{4^n} \leqslant \frac{|Q_1|}{|Q|} \leqslant \frac{1}{2^n} \leqslant 1.$$

In general, there does not exist a dyadic cube Q_2 with $Q \subset Q_2$. Even in the case such a cube exists, the side length of Q_2 may be arbitrarily many times lager than the side length of Q and there does not exist a uniform upper bound for $\frac{|Q_2|}{|Q|}$ (exercise).

However, it is possible to cover

$$Q = Q(x,l) = \left[x_1 - \frac{l}{2}, x_1 + \frac{l}{2}\right] \times \cdots \times \left[x_n - \frac{l}{2}, x_n + \frac{l}{2}\right]$$

by a finitely many dyadic cubes of comparable size. We may take a dyadic cube

$$Q_2 = \prod_{j=1}^{n} [m_j 2^{-k}, (m_j + 1)2^{-k}), \quad m_1, \dots, m_n, k \in \mathbb{Z},$$

containing the corner $(x_1 - \frac{l}{2}, \dots, x_n - \frac{l}{2})$ of Q with the side length 2^{-k} where $\frac{l}{2} < 2^{-k} \leqslant l$. Then $Q \subset \widetilde{Q}_2$, where

$$\widetilde{Q}_2 = \prod_{j=1}^{n} [m_j 2^{-k}, (m_j + 3)2^{-k})$$

has the same corner as $Q_2 \in \mathscr{D}_k$ but the side length is $3 \cdot 2^{-k}$. The cube $\widetilde{Q}_2$ is not a dyadic cube, but it is a union of 3^n pairwise disjoint dyadic cubes in $\mathscr{D}_k$. This approach can be developed further, see [15]. We shall return to this in the proof of Lemma 1.26.

Lemma 1.22. Assume that $f \in L^1(\mathbb{R}^n)$ and let

$$E_t = \{x \in \mathbb{R}^n : M_d f(x) > t\}, \quad t > 0.$$

Then E_t is the union of pairwise disjoint dyadic Calderón-Zygmund cubes Q_i, $i = 1, 2, \dots$, given by Theorem 1.12. In particular, the cubes Q_i, $i = 1, 2, \dots$, satisfy properties (1)-(3) in Theorem 1.12.

Proof. We show that $E_t = \bigcup_{i=1}^{\infty} Q_i$. If $x \in E_t$, then $M_d f(x) > t$ and there exists a dyadic cube $Q \in \mathscr{D}$ such that $x \in Q$ and

$$\fint_Q |f(y)|\,dy > t.$$

The Calderón-Zygmund cubes Q_i, $i = 1, 2, \dots$, given by Theorem 1.12 is the collection of maximal dyadic cubes with this property. This implies that $x \in \bigcup_{i=1}^{\infty} Q_i$ and thus $E_t \subset \bigcup_{i=1}^{\infty} Q_i$.

On the other hand, if $x \in \bigcup_{i=1}^{\infty} Q_i$, then $x \in Q_i$ for some $i = 1, 2, \dots$ and by the Calderón-Zygmund decomposition

$$M_d f(x) \geqslant \fint_{Q_i} |f(y)|\,dy > t.$$

This shows that $x \in E_t$ and thus $\bigcup_{i=1}^{\infty} Q_i \subset E_t$. This completes the proof. □

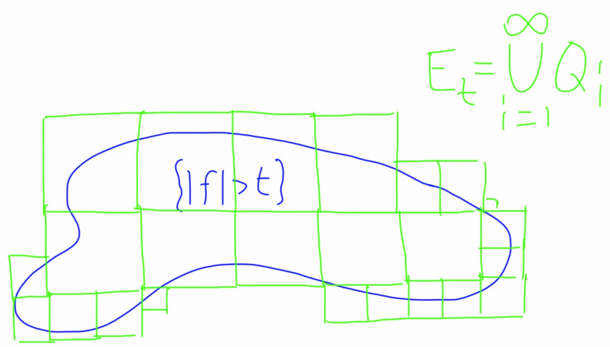

Figure 1.9: The distribution set of the dyadic maximal function.

THE MORAL: The union of the Calderón-Zygmund cubes is the distribution set of the dyadic maximal function. This means that the Calderón-Zygmund decomposition is more closely related to $f = f\chi_{\{M_d f \leqslant t\}} + f\chi_{\{M_d f > t\}}$ instead of $f = f\chi_{\{|f|\leqslant t\}} + f\chi_{\{|f|>t\}}$ in (1.15). Note carefully, that this is not the Calderón-Zygmund decomposition of a function constructed in the proof of Theorem 1.16, but Lemma 1.22 shows that $\{M_d f > t\}$ is the union of the Calderón-Zygmund cubes. This suggest another point of view to the Calderón-Zygmund decomposition, in which we analyse the distribution set of the dyadic maximal function, for example, using the Whitney covering theorem, see Remark 1.11.

Example 1.23. Consider $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = \chi_{[0,1]^n}(x)$ as in Example 1.14 and Example 1.19. If $t \geqslant 1 = \|f\|$, the Calderón-Zygmund decomposition is empty and $\{x \in \mathbb{R}^n : M_d f(x) > t\} = \emptyset$. For $0 < t < 1$ the Calderón-Zygmund decomposition consist of one dyadic cube $[0,2^k)$, $k \in \{0,1,2,\dots\}$ and $\{x \in \mathbb{R}^n : M_d f(x) > t\} = [0,2^k)$.

Remarks 1.24:

(1) By applying Theorem 1.12 (2) and summing up over all Calderón-Zygmund cubes, we have

$$|E_t| = \sum_{i=1}^{\infty} |Q_i| \leqslant \frac{1}{t} \sum_{i=1}^{\infty} \int_{Q_i} |f(y)|\,dy = \frac{1}{t} \int_{E_t} |f(y)|\,dy.$$

This is a weak type estimate for the dyadic maximal function. Observe that in contrast with the standard Hardy-Littlewood maximal function, there is no dimensional constant in the estimate. This estimate also holds true for other measures than Lebesgue measure.

(2) We also have an inequality to the reverse direction, since by Theorem 1.12 (2) we have

$$\int_{E_t} |f(y)|\,dy = \sum_{i=1}^{\infty} \int_{Q_i} |f(y)|\,dy \leqslant 2^n t \sum_{i=1}^{\infty} |Q_i| = 2^n t |E_t|. \tag{1.25}$$

This is a reverse weak type inequality for the dyadic maximal function.

(3) If $t > s$, then $E_t \subset E_s$ and, by maximality of the Calderón-Zygmund cubes, each cube in the decomposition at level t is contained in a cube in the decomposition at level s. In this sense, the Calderón-Zygmund decompositions are nested.

(4) Instead of assuming that $f \in L^1(\mathbb{R}^n)$ in Lemma 1.22 we may assume that $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ is such that the set $E_t = \{x \in \mathbb{R}^n : M_d f(x) > t\}$, $t > 0$, has finite measure. If $x \in E_t$, then $M_d f(x) > t$ and there exists a dyadic cube $Q \in \mathscr{D}$ such that $x \in Q$ and

$$\fint_Q |f(y)|\,dy > t.$$

Observe that

$$M_d f(z) \geqslant \fint_Q |f(y)|\,dy > t$$

for every $z \in Q$ and thus $Q \subset E_t$. In particular, this implies that the union of such dyadic cubes is a set of finite measure. By Remark 1.9 (1) there exists a subcollection of pairwise disjoint maximal cubes Q_i, $i = 1, 2, \dots$, with the same union. By maximality, the parent cube Q_i' of every Q_i, $i = 1, 2, \dots$, intersects $\mathbb{R}^n \setminus E_t = \{x \in \mathbb{R}^n : M_d f(x) \leqslant t\}$. This implies that, for every $i = 1, 2, \dots$, there exists a point $z_i \in Q_i'$ with $M_d f(z_i) \leqslant t$. It follows that

$$t < \fint_Q |f(y)|\,dy \leqslant \frac{|Q'|}{|Q|} \fint_{Q'} |f(y)|\,dy \leqslant 2^n M_d f(z_i) \leqslant 2^n t.$$

This shows that the Calderón-Zygmund decomposition in Theorem 1.12 can be obtained under the assumption that $|\{x \in \mathbb{R}^n : M_d f(x) > t\}| < \infty$.

Next we show how we can use the Calderón-Zygmund decomposition to obtain estimates for the standard maximal function defined by (1.20).

Lemma 1.26. Assume that $f \in L^1(\mathbb{R}^n)$ and let Q_i, $i = 1, 2, \dots$, be the Calderón-Zygmund cubes of f at level $t > 0$ given by Theorem 1.12. Then

(1) $\bigcup_{i=1}^{\infty} Q_i \subset \{x \in \mathbb{R}^n : Mf(x) > t\}$ and

(2) $\{x \in \mathbb{R}^n : Mf(x) > 4^n t\} \subset \bigcup_{i=1}^{\infty} 3Q_i$.

THE MORAL: The first claim is essentially a restatement of the fact that $M_d f(x) \leqslant Mf(x)$ for every $x \in \mathbb{R}^n$ and thus

$$\{x \in \mathbb{R}^n : M_d f(x) > t\} \subset \{x \in \mathbb{R}^n : Mf(x) > t\}.$$

By Example 1.21 the does not exist a constant c such that $Mf(x) \leqslant cM_d f(x)$ for every $x \in \mathbb{R}^n$, but the second claim implies the following distributional inequality in the reverse direction

$$\begin{aligned} |\{x \in \mathbb{R}^n : Mf(x) > 4^n t\}| &\leqslant \sum_{i=1}^{\infty} |3Q_i| = 3^n \sum_{i=1}^{\infty} |Q_i| \\ &= 3^n |\{x \in \mathbb{R}^n : M_d f(x) > t\}| \end{aligned} \tag{1.27}$$

for every $t > 0$. In particular, this gives the weak type estimate for the standard Hardy-Littlewood maximal function as well, since

$$|\{x \in \mathbb{R}^n : Mf(x) > t\}| \leqslant 3^n \left|\{x \in \mathbb{R}^n : M_d f(x) > \tfrac{t}{4^n}\}\right| \leqslant \frac{3^n 4^n}{t} \int_{\mathbb{R}^n} |f(y)|\,dy$$

for every $t > 0$. Thus the weak type estimate for the standard Hardy-Littlewood maximal function follows from the corresponding estimate for the dyadic maximal function. Moreover, by Cavalieri's principle we obtain the correspding estimate for the L^p norms, see Example 2.6 (2). This shows that information on dyadic cubes can be used to obtain information over all cubes, see also Example 2.6 (2).

THE MORAL: Even though two functions are not comparable pointwise, their distribution functions and L^p norms may be comparable.

Proof. $\boxed{(1)}$ If $x \in \bigcup_{i=1}^{\infty} Q_i$, then $x \in Q_i$ for some $i = 1, 2, \ldots$. By Theorem 1.12

$$Mf(x) \geqslant \fint_{Q_i} |f(y)|\,dy > t$$

and thus $\bigcup_{i=1}^{\infty} Q_i \subset \{x \in \mathbb{R}^n : Mf(x) > t\}$.

$\boxed{(2)}$ Assume that $x \in \mathbb{R}^n \setminus \bigcup_{i=1}^{\infty} 3Q_i$ and let Q any closed cube in $\mathbb{R}^n$ containing x. Choose $k \in \mathbb{Z}$ such that $2^{-k-1} < l(Q) \leqslant 2^{-k}$. Then there exists at most 2^n such dyadic cubes $R_1, \ldots, R_m \in \mathscr{D}_k$, which intersect the interior of Q. We note that $Q \subset 3R_j$ for every $j = 1, \ldots, m$. Each cube R_j, $j = 1, \ldots, m$, cannot be a subset of any of the cubes Q_i, $i = 1, 2, \ldots$, since otherwise $x \in Q \subset 3R_j \subset 3Q_i$ for some $i = 1, 2, \ldots$, which is not possible, since $x \in \mathbb{R}^n \setminus \bigcup_{i=1}^{\infty} 3Q_i$. Since R_j is not contained in the union of the Calderón-Zygmund cubes, by the proof of Theorem 1.12 and Theorem 1.4, we have

$$\frac{1}{|R_j|} \int_{R_j} |f(y)|\,dy \leqslant t, \quad i = 1, \ldots, m.$$

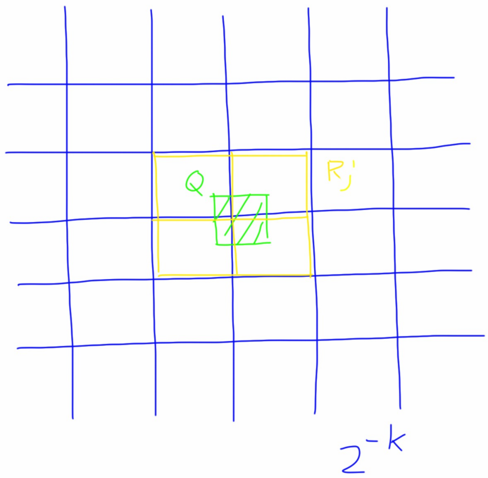

Figure 1.10: At most 2^n dyadic cubes $R_1,\dots,R_m \in \mathscr{D}_k$ intersect the interior of Q.

On the other hand, $|R_j| = 2^{-kn} = 2^n 2^{-(k+1)n} \leqslant 2^n l(Q)^n \leqslant 2^n |Q|$ and $1 \leqslant m \leqslant 2^n$, thus

$$\begin{aligned}\frac{1}{|Q|}\int_Q |f(y)|\,dy &= \frac{1}{|Q|}\sum_{j=1}^m \int_{Q\cap R_j} |f(y)|\,dy \\ &\leqslant \sum_{j=1}^m \frac{|R_j|}{|Q|}\frac{1}{|R_j|}\int_{R_j}|f(y)|\,dy \\ &\leqslant m2^n t \leqslant 4^n t.\end{aligned}$$

Since this holds true for every cube Q containing x, we have $Mf(x) \leqslant 4^n t$ for every $x \in \mathbb{R}^n \setminus \bigcup_{i=1}^\infty 3Q_i$. In other words,

$$\mathbb{R}^n \setminus \bigcup_{i=1}^\infty 3Q_i \subset \{x \in \mathbb{R}^n : Mf(x) \leqslant 4^n t\},$$

from which the claim follows. □

Remark 1.28. By Remark 1.24 (4) it is enough to assume that $|\{x \in \mathbb{R}^n : M_d f(x) > t\}| < \infty$ in Lemma 1.26.

1.5 Dyadic maximal function on a cube

Next we discuss briefly the dyadic maximal function with respect to the dyadic subcubes of a cube. Let $Q_0 \subset \mathbb{R}^n$ be a cube and assume that $f \in L^1(Q_0)$. The dyadic maximal function $M_{d;Q_0}f$ at $x \in Q_0$ is

$$M_{d,Q_0}f(x) = \sup_{Q \ni x} \fint_Q |f(y)|\,dy, \tag{1.29}$$

where the supremum is taken over all dyadic cubes $Q \in \mathscr{D}(Q_0)$ with $x \in Q$.

Let $f, g \in L^1(Q_0)$ and $x \in Q_0$. It follows immediately from the definition that $M_{d,Q_0}f(x) \geqslant 0$,

$$M_{d,Q_0}(f+g)(x) \leqslant M_{d,Q_0}f(x) + M_{d,Q_0}g(x),$$

and

$$M_{d,Q_0}(af)(x) = |a| M_{d,Q_0}f(x)$$

for every $a \in \mathbb{R}$.

Let $E_t = \{x \in Q_0 : M_{d,Q_0}f(x) > t\}$, $t > 0$. For every $t \geqslant |f|_{Q_0}$ the set E_t is the union of pairwise disjoint dyadic Calderón-Zygmund cubes Q_i, $i = 1, 2, \dots$, given by Theorem 1.4. In particular, cubes Q_i, $i = 1, 2, \dots$, satisfy properties (1)-(3) in Theorem 1.4. For $0 < t < |f|_{Q_0}$, we have $E_t = Q_0$.

Theorem 1.4 gives simple proofs for norm estimates for the dyadic maximal function. The next result is a weak type estimate.

Lemma 1.30. Let $Q_0 \subset \mathbb{R}^n$ be a cube. Assume that $f \in L^1(Q_0)$ and let $E_t = \{x \in Q_0 : M_{d,Q_0}f(x) > t\}$. Then

$$|E_t| \leqslant \frac{1}{t} \int_{E_t} |f(x)|\,dx$$

for every $t > 0$.

Proof. Let $t > 0$. If $t < |f|_{Q_0}$, then $E_t = Q_0$ and thus

$$|E_t| = |Q_0| \leqslant \frac{1}{t} \int_{Q_0} |f(x)|\,dx = \frac{1}{t} \int_{E_t} |f(x)|\,dx.$$

Then assume that $t \geqslant |f|_{Q_0}$. Let Q_i, $i = 1, 2, \dots$ be the collection of dyadic subcubes of Q_0 given by Theorem 1.4. By using the properties of the Calderón-Zygmund cubes, we obtain

$$|E_t| = \left|\bigcup_{i=1}^{\infty} Q_i\right| = \sum_{i=1}^{\infty} |Q_i| \leqslant \frac{1}{t} \sum_{i=1}^{\infty} \int_{Q_i} |f(x)|\,dx = \frac{1}{t} \int_{E_t} |f(x)|\,dx. \qquad \square$$

There is also a reverse weak type estimate for the dyadic maximal function.

Lemma 1.31. Let $Q_0 \subset \mathbb{R}^n$ be a cube. Assume that $f \in L^1(Q_0)$ is a nonnegative function and let $E_t = \{x \in Q_0 : M_{d,Q_0}f(x) > t\}$. Then

$$\int_{E_t} |f(x)|\,dx \leqslant 2^n t |E_t|$$

for every $t \geqslant |f|_{Q_0}$.

Proof. Let Q_i, $i = 1,2,\dots$ be the collection of dyadic subcubes of Q_0 given by Theorem 1.4. By using the properties of the Calderón-Zygmund cubes, we obtain

$$\int_{E_t} |f(x)|\,dx = \sum_{i=1}^{\infty} \int_{Q_i} f(x)\,dx \leqslant 2^n t \sum_{i=1}^{\infty} |Q_i| = 2^n t |E_t|.$$

□

2

Marcinkiewicz interpolation theorem

Interpolation of operators is an important tool in harmonic analysis. Consider an operator, which maps Lebesgue measurable functions to functions. A typical example is the maximal operator. The rough idea of interpolation is that if we know that the operator is a bounded in two different function spaces, then it is bounded in the intermediate function spaces.

We are mainly interested in $L^p(\mathbb{R}^n)$ spaces with $1 \leqslant p \leqslant \infty$ and we begin with a useful decomposition of an $L^p(\mathbb{R}^n)$ function into two parts. To this end, we define $L^{p_1}(\mathbb{R}^n) + L^{p_2}(\mathbb{R}^n)$, $1 \leqslant p_1 < p_2 \leqslant \infty$, to be the space of all functions of the form $f = f_1 + f_2$, where $f_1 \in L^{p_1}(\mathbb{R}^n)$ and $f_2 \in L^{p_2}(\mathbb{R}^n)$.

Lemma 2.1. Let $1 \leqslant p_1 < p_2 \leqslant \infty$ and $p_1 \leqslant p \leqslant p_2$. Then $L^p(\mathbb{R}^n) \subset L^{p_1}(\mathbb{R}^n) + L^{p_2}(\mathbb{R}^n)$.

THE MORAL: Every $L^p(\mathbb{R}^n)$ function can be written as a sum of an $L^{p_1}(\mathbb{R}^n)$ function and an $L^{p_2}(\mathbb{R}^n)$ function whenever $p_1 \leqslant p \leqslant p_2$.

Proof. If $p = p_1$ or $p = p_2$, there is nothing to prove, since $f = f + 0$. Thus we assume that $p_1 < p < p_2$. Assume that $f \in L^p(\mathbb{R}^n)$ and let $t > 0$. Define

$$f_1(x) = f(x)\chi_{\{|f|>t\}}(x) = \begin{cases} f(x), & \text{if} \quad |f(x)| > t, \\ 0, & \text{if} \quad |f(x)| \leqslant t, \end{cases}$$

and

$$f_2(x) = f(x)\chi_{\{|f|\leqslant t\}}(x) = \begin{cases} f(x), & \text{if} \quad |f(x)| \leqslant t, \\ 0, & \text{if} \quad |f(x)| > t. \end{cases}$$

Clearly

$$f(x) = f(x)\chi_{\{|f|>t\}}(x) + f(x)\chi_{\{|f|\leqslant t\}}(x) = f_1(x) + f_2(x),$$

see (1.15).

First we show that $f_1 \in L^{p_1}(\mathbb{R}^n)$. Since $p_1 < p$, we obtain

$$\int_{\mathbb{R}^n} |f_1(x)|^{p_1}\,dx = \int_{\{|f|>t\}} |f(x)|^{p_1}\,dx = \int_{\{|f|>t\}} |f(x)|^{p_1-p}|f(x)|^p\,dx \leqslant t^{p_1-p}\int_{\mathbb{R}^n} |f(x)|^p\,dx \leqslant t^{p_1-p}\|f\|_p^p < \infty.$$

Then we show that $f_2 \in L^{p_2}(\mathbb{R}^n)$. Since $p_2 > p$, we have

$$\int_{\mathbb{R}^n} |f_2(x)|^{p_2}\,dx = \int_{\{|f|\leqslant t\}} |f(x)|^{p_2}\,dx = \int_{\{|f|\leqslant t\}} |f(x)|^{p_2-p}|f(x)|^p\,dx \leqslant t^{p_2-p}\int_{\mathbb{R}^n} |f(x)|^p\,dx \leqslant t^{p_2-p}\|f\|_p^p < \infty.$$

Thus $f = f_1 + f_2$ with $f_1 \in L^{p_1}(\mathbb{R}^n)$ and $f_2 \in L^{p_2}(\mathbb{R}^n)$, as required. □

Definition 2.2. Let T be an operator from $L^p(\mathbb{R}^n)$ to Lebesgue measurable functions on $\mathbb{R}^n$.

(1) T is sublinear, if for every $f, g \in L^p(\mathbb{R}^n)$,

$$|T(f+g)(x)| \leqslant |Tf(x)| + |Tg(x)|$$

and

$$|T(af)(x)| = |a||Tf(x)|, \quad a \in \mathbb{R},$$

for almost every $x \in \mathbb{R}^n$.

(2) T is of strong type (p,p), $1 \leqslant p \leqslant \infty$, if there exists a constant c, independent of the function f, such that

$$\|Tf\|_p \leqslant c\|f\|_p$$

for every $f \in L^p(\mathbb{R}^n)$.

(3) T is of weak type (p,p), $1 \leqslant p < \infty$, if there exists a constant c, independent of the function f, such that

$$|\{x \in \mathbb{R}^n : |Tf(x)| > t\}| \leqslant \left(\frac{c}{t}\|f\|_p\right)^p$$

for every $t > 0$ and $f \in L^p(\mathbb{R}^n)$.

THE MORAL: Operator is of strong type (p,p) if and only if it is a bounded operator from $L^p(\mathbb{R}^n)$ to $L^p(\mathbb{R}^n)$. The corresponding weak type condition is a substitute for this for several operators in harmonic analysis which fail to be bounded in certain $L^p(\mathbb{R}^n)$ spaces. For example, the Hardy-Littlewood maximal operator is a sublinear operator which is not of strong type $(1,1)$ but it is of weak type $(1,1)$.

Remarks 2.3:

(1) Every linear operator T is sublinear, since

$$|T(f+g)(x)| = |Tf(x)+Tg(x)| \leqslant |Tf(x)| + |Tg(x)|$$

and

$$|T(af)(x)| = |aTf(x)| = |a||Tf(x)|.$$

(2) The notion of strong type (p,p) is stronger than weak type (p,p). If $\|Tf\|_p \leqslant c\|f\|_p$ for every $f \in L^p(\mathbb{R}^n)$, by Chebyshev's inequality

$$|\{x \in \mathbb{R}^n : |Tf(x)| > t\}| \leqslant \frac{1}{t^p}\int_{\mathbb{R}^n} |Tf(x)|^p\,dx = \frac{1}{t^p}\|Tf\|_p^p \leqslant \left(\frac{c}{t}\|f\|_p\right)^p.$$

Theorem 2.4 (Marcinkiewicz interpolation theorem (1939)). Let $1 \leqslant p_1 < p_2 \leqslant \infty$ and assume that T is a sublinear operator from $L^{p_1}(\mathbb{R}^n) + L^{p_2}(\mathbb{R}^n)$ to Lebesgue measurable functions on $\mathbb{R}^n$, which is simultaneously of weak type (p_1,p_1) and (p_2,p_2). Then T is of strong type (p,p) whenever $p_1 < p < p_2$.

THE MORAL: Weak type estimates at the endpoint spaces imply strong type estimates spaces between.

Proof. $\boxed{p_2 < \infty}$ Assume that if there exist constant c_1 and c_2, independent of the function f, such that

$$|\{x \in \mathbb{R}^n : |Tf(x)| > t\}| \leqslant \left(\frac{c_1}{t}||f||_{p_1}\right)^{p_1}, \quad t > 0,$$

for every $f \in L^{p_1}(\mathbb{R}^n)$ and

$$|\{x \in \mathbb{R}^n : |Tf(x)| > t\}| \leqslant \left(\frac{c_2}{t}||f||_{p_2}\right)^{p_2}, \quad t > 0.$$

for every $f \in L^{p_2}(\mathbb{R}^n)$. Consider the decomposition

$$f = f_1 + f_2 = f\chi_{\{|f|>t\}} + f\chi_{\{|f|\leqslant t\}},$$

where $f_1 \in L^{p_1}(\mathbb{R}^n)$ and $f_2 \in L^{p_2}(\mathbb{R}^n)$, given by Lemma 2.1. Sublinearity $|Tf(x)| \leqslant |Tf_1(x)| + |Tf_2(x)|$ implies that for almost every x for which $|Tf(x)| > t$, either $|Tf_1(x)| > \frac{t}{2}$ or $|Tf_2(x)| > \frac{t}{2}$. Thus

$$\begin{aligned}
|\{x \in \mathbb{R}^n : |Tf(x)| > t\}| &\leqslant \left|\left\{x \in \mathbb{R}^n : |Tf_1(x)| > \tfrac{t}{2}\right\} \cup \left\{x \in \mathbb{R}^n : |Tf_2(x)| > \tfrac{t}{2}\right\}\right| \\
&\leqslant \left|\left\{x \in \mathbb{R}^n : |Tf_1(x)| > \tfrac{t}{2}\right\}\right| + \left|\left\{x \in \mathbb{R}^n : |Tf_2(x)| > \tfrac{t}{2}\right\}\right| \\
&\leqslant \left(\frac{c_1}{\frac{t}{2}}||f_1||_{p_1}\right)^{p_1} + \left(\frac{c_2}{\frac{t}{2}}||f_2||_{p_2}\right)^{p_2} \\
&\leqslant \left(\frac{2c_1}{t}\right)^{p_1}\int_{\{x\in\mathbb{R}^n:|f(x)|>t\}} |f(x)|^{p_1}\,dx \\
&\quad + \left(\frac{2c_2}{t}\right)^{p_2}\int_{\{x\in\mathbb{R}^n:|f(x)|\leqslant t\}} |f(x)|^{p_2}\,dx.
\end{aligned}$$

By Cavalieri's principle

$$\begin{aligned}\int_{\mathbb{R}^n} |Tf(x)|^p\,dx &= p\int_0^\infty t^{p-1}|\{x\in\mathbb{R}^n : |Tf(x)|>t\}|\,dt\\ &\leqslant (2c_1)^{p_1}p\int_0^\infty t^{p-p_1-1}\int_{\{x\in\mathbb{R}^n:|f(x)|>t\}}|f(x)|^{p_1}\,dx\,dt\\ &\quad+(2c_2)^{p_2}p\int_0^\infty t^{p-p_2-1}\int_{\{x\in\mathbb{R}^n:|f(x)|\leqslant t\}}|f(x)|^{p_2}\,dx\,dt,\end{aligned}$$

where the integrals on the right-hand side are computed by Fubini's theorem as

$$\begin{aligned}\int_0^\infty t^{p-p_1-1}\int_{\{x\in\mathbb{R}^n:|f(x)|>t\}}|f(x)|^{p_1}\,dx\,dt &= \int_{\mathbb{R}^n}|f(x)|^{p_1}\int_0^{|f(x)|}t^{p-p_1-1}\,dt\,dx\\ &=\frac{1}{p-p_1}\int_{\mathbb{R}^n}|f(x)|^{p-p_1}|f(x)|^{p_1}\,dx\\ &=\frac{1}{p-p_1}\int_{\mathbb{R}^n}|f(x)|^p\,dx\end{aligned}$$

and

$$\begin{aligned}\int_0^\infty t^{p-p_2-1}\int_{\{x\in\mathbb{R}^n:|f(x)|\leqslant t\}}|f(x)|^{p_2}\,dx\,dt &= \int_{\mathbb{R}^n}|f(x)|^{p_2}\int_{|f(x)|}^{\infty}t^{p-p_2-1}\,dt\,dx\\ &=\frac{1}{p_2-p}\int_{\mathbb{R}^n}|f(x)|^{p_2}|f(x)|^{p-p_2}\,dx\\ &=\frac{1}{p_2-p}\int_{\mathbb{R}^n}|f(x)|^p\,dx.\end{aligned}$$

Thus we arrive at

$$\begin{aligned}||Tf||_p^p &= \int_{\mathbb{R}^n}|Tf(x)|^p\,dx\\ &\leqslant (2c_1)^{p_1}\frac{p}{p-p_1}\int_{\mathbb{R}^n}|f(x)|^p\,dx+(2c_2)^{p_2}\frac{p}{p_2-p}\int_{\mathbb{R}^n}|f(x)|^p\,dx\\ &= p\left(\frac{(2c_1)^{p_1}}{p-p_1}+\frac{(2c_2)^{p_2}}{p_2-p}\right)||f||_p^p.\end{aligned}$$

$\boxed{p_2=\infty}$ Assume that $||Tf||_\infty \leqslant c_2||f||_\infty$ for every $f\in L^\infty(\mathbb{R}^n)$ and write

$$f = f_1+f_2 = f\chi_{\{|f|>\frac{t}{2c_2}\}}+f\chi_{\{|f|\leqslant\frac{t}{2c_2}\}}.$$

Then $f_1\in L^{p_1}(\mathbb{R}^n)$ as in Lemma 2.1 and $f_2\in L^\infty(\mathbb{R}^n)$, since $\|f_2\|_\infty\leqslant\frac{t}{2c_2}$. We apply strong (∞,∞) estimate for f_2 and obtain

$$|Tf_2(x)|\leqslant||Tf_2||_\infty\leqslant c_2||f_2||_\infty\leqslant c_2\frac{t}{2c_2}=\frac{t}{2}$$

for almost every $x\in\mathbb{R}^n$ and, consequently,

$$\left|\{x\in\mathbb{R}^n : |Tf_2(x)|>\tfrac{t}{2}\}\right| = 0.$$

Thus

$$\begin{aligned}
|\{x\in\mathbb{R}^n : |Tf(x)|>t\}| &\leqslant \left|\left\{x\in\mathbb{R}^n : |Tf_1(x)|>\tfrac{t}{2}\right\}\right| + \left|\left\{x\in\mathbb{R}^n : |Tf_2(x)|>\tfrac{t}{2}\right\}\right| \\
&\leqslant \left|\left\{x\in\mathbb{R}^n : |Tf_1(x)|>\tfrac{t}{2}\right\}\right| \\
&\leqslant \left(\frac{c_1}{\frac{t}{2}}\|f_1\|_{p_1}\right)^{p_1} \\
&\leqslant \left(\frac{2c_1}{t}\right)^{p_1}\int_{\{x\in\mathbb{R}^n : |f(x)|>\frac{t}{2c_2}\}}|f(x)|^{p_1}\,dx
\end{aligned}$$

By Cavalieri's principle

$$\begin{aligned}
\|Tf\|_p^p &= \int_{\mathbb{R}^n}|Tf(x)|^p\,dx \\
&= p\int_0^\infty t^{p-1}|\{x\in\mathbb{R}^n : |Tf(x)|>t\}|\,dt \\
&\leqslant (2c_1)^{p_1}p\int_0^\infty t^{p-p_1-1}\int_{\{x\in\mathbb{R}^n : |f(x)|>\frac{t}{2c_2}\}}|f(x)|^{p_1}\,dx\,dt \\
&= (2c_1)^{p_1}p\int_{\mathbb{R}^n}|f(x)|^{p_1}\int_0^{2c_2|f(x)|}t^{p-p_1-1}\,dt\,dx \\
&= \frac{(2c_1)^{p_1}p}{p-p_1}\int_{\mathbb{R}^n}|2c_2f(x)|^{p-p_1}|f(x)|^{p_1}\,dx \\
&= (2c_1)^p c_2^{p-p_1}\frac{p}{p-p_1}\int_{\mathbb{R}^n}|f(x)|^p\,dx \\
&= p\frac{(2c_1)^p c_2^{p-p_1}}{p-p_1}\|f\|_p^p. \qquad \square
\end{aligned}$$

Remarks 2.5:

(1) In particular, if T is a linear operator which is strong type (p_1,p_1) and (p_2,p_2), then it is strong type (p,p) whenever $p_1\leqslant p\leqslant p_2$. For linear operators there are better Riesz-Thorin type interpolation results, see [2] and [22].

(2) If T is a linear operator instead of sublinear, then it is enough to assume weak type (p_1,p_1) and (p_2,p_2) estimates for simple functions.

(3) Note that the constant in strong type (p,p) estimate blows up as $p\to p_1$ and $p\to p_2$, when $p_2<\infty$.

(4) By considering $f=f_1+f_2=f\chi_{\{|f|>\gamma t\}}+f\chi_{\{|f|\leqslant\gamma t\}}$, where $\gamma>0$ is chosen appropriately, we obtain strong type bound $\|Tf\|_p\leqslant c\|f\|_p$, $p_1<p<p_2$, with

$$c = 2\left(\frac{p}{p-p_1}+\frac{p}{p_2-p}\right)^{\frac{1}{p}} c_1^{\frac{\frac{1}{p}-\frac{1}{p_2}}{\frac{1}{p_1}-\frac{1}{p_2}}} c_2^{\frac{\frac{1}{p_1}-\frac{1}{p}}{\frac{1}{p_1}-\frac{1}{p_2}}}, \quad \text{if} \quad p_2<\infty,$$

and

$$c = 2\left(\frac{p}{p-p_1}\right)^{\frac{1}{p}} c_1^{\frac{p_1}{p}} c_2^{1-\frac{p_1}{p}}, \quad \text{if} \quad p_2=\infty.$$

(5) The proof is based only on Cavalieri's principle, from which we conclude that the Marcinkewicz interpolation theorem holds for other measures than the Lebesgue measure as well. In this case we assume that the weak type estimates hold with respect to the given measure. Moreover, the Marcinkewicz interpolation theorem holds in more general spaces than the Euclidean spaces.

(6) There are more general versions of the Marcinkiewicz interpolation theorem in Lorenz spaces, see [2] and [22].

(7) There is a general theory of interpolation of operators in Banach spaces and more general topological spaces. The Marcinkiewicz interpolation theorem has lead to real method of interpolation and Thorin's method has lead to complex method of interpolation, see [2] and references therein.

Examples 2.6:

(1) The dyadic maximal function defined in (1.18) is of strong type (∞,∞), since for every $x\in\mathbb{R}^n$, we have

$$M_d f(x) = \sup \fint_Q |f(y)|\,dy \leqslant \|f\|_\infty,$$

where the supremum is taken over all dyadic cubes Q containing x. Thus $\|M_d f\|_\infty \leqslant \|f\|_\infty$ for every $f\in L^\infty(\mathbb{R}^n)$. In fact, we have $\|M_d f\|_\infty = \|f\|_\infty$ for every $f\in L^\infty(\mathbb{R}^n)$ (exercise). On the other hand, by Remark 1.24 (1) we have

$$|\{x\in\mathbb{R}^n : M_d f(x) > t\}| \leqslant \frac{1}{t}\int_{\mathbb{R}^n} |f(y)|\,dy, \quad t>0,$$

for every $f\in L^1(\mathbb{R}^n)$. Since the dyadic maximal function is sublinear, by the Marcinkiewicz interpolation theorem we conclude that it is of strong type (p,p) for every $1<p<\infty$ and

$$\|M_d f\|_p \leqslant \frac{p2^p}{p-1}\|f\|_p, \quad 1<p<\infty.$$

When $p=\infty$, the bound holds with constant one.

(2) Note that by Lemma 1.26 and Remark 1.24 (4), we obtain similar bounds for the standard Hardy-Littlewood maximal function defined by (1.20) as well. If

$$|\{x\in\mathbb{R}^n : M_d f(x) > t\}| < \infty$$

for every $t>0$, then by (1.27), we have

$$|\{x\in\mathbb{R}^n : Mf(x) > 4^n t\}| \leqslant 3^n |\{x\in\mathbb{R}^n : M_d f(x) > t\}|$$

for every $t > 0$ and

$$\begin{aligned}\|Mf\|_p^p &= p\int_0^\infty t^{p-1}|\{x\in\mathbb{R}^n : Mf(x) > t\}|\,dt\\ &= 4^n p\int_0^\infty t^{p-1}|\{x\in\mathbb{R}^n : Mf(x) > 4^n t\}|\,dt\\ &\leqslant 4^n 3^n p\int_0^\infty t^{p-1}|\{x\in\mathbb{R}^n : M_d f(x) > t\}|\,dt\\ &= 12^n\|M_d f\|_p^p.\end{aligned}$$

In particular, by (1) it follows that

$$\|Mf\|_p \leqslant c\|f\|_p \quad \text{whenever} \quad 1 < p \leqslant \infty.$$

This gives a dyadic proof for the Hardy-Littlewood-Wiener maximal function theorem.

Next we demonstrate how to apply the Marcinkiewicz interpolation theorem together with the Calderón-Zygmund decomposition to prove a strong type estimate for certain sublinear operators, see [3, Lemma 2.1] and [18, Lemma 2.10]. Let Q_i, $i = 1,2,\dots$, be the Calderón-Zygmund cubes for $f \in L^p(\mathbb{R}^n)$ at level $t > 0$, see Theorem 1.12 and Remark 1.13. We consider the Calderón-Zygmund decomposition of $f = g + b$ in to good and bad parts as in Theorem 1.16.

Theorem 2.7. Let $1 < p < \infty$ and assume that T is a sublinear operator of weak type (p,p) such that $Tb(x) = 0$ for almost every $x \in \mathbb{R}^n \setminus \bigcup_{i=1}^\infty Q_i$. Then T is of strong type (q,q) for $1 < q < p$.

WARNING: The condition $Tb(x) = 0$ for almost every $x \in \mathbb{R}^n \setminus \bigcup_{i=1}^\infty Q_i$ is relatively restrictive. However, a similar technique can be applied for other operators as well.

Proof. We show that T is of weak type (1,1). Let $f \in L^1(\mathbb{R}^n)$. As in the proof of the Marcinkiewicz interpolation theorem (Theorem 2.4), we have

$$|\{x\in\mathbb{R}^n : |Tf(x)| > t\}| \leqslant \left|\left\{x\in\mathbb{R}^n : |Tg(x)| > \tfrac{t}{2}\right\}\right| + \left|\left\{x\in\mathbb{R}^n : |Tb(x)| > \tfrac{t}{2}\right\}\right|.$$

Since T is of weak type (p,p), by Theorem 1.16 (3) and (2) we have

$$\begin{aligned}\left|\left\{x\in\mathbb{R}^n : |Tg(x)| > \tfrac{t}{2}\right\}\right| &\leqslant \left(\frac{c}{\frac{t}{2}}\|g\|_p\right)^p \leqslant \frac{2^p c^p}{t^p}\|g\|_\infty^{p-1}\|g\|_1\\ &\leqslant \frac{2^p c^p}{t^p}(2^n t)^{p-1}\|f\|_1\\ &\leqslant \frac{2^{p(n+1)}c^p}{t}\|f\|_1.\end{aligned}$$

On the other hand, since $Tb(x) = 0$ for almost every $x \in \mathbb{R}^n \setminus \bigcup_{i=1}^{\infty} Q_i$, by Theorem 1.16 (7) we obtain

$$\left|\left\{x \in \mathbb{R}^n : |Tb(x)| > \tfrac{t}{2}\right\}\right| = \left|\bigcup_{i=1}^{\infty} Q_i\right| \leqslant \frac{1}{t}\|f\|_1.$$

This implies that

$$|\{x \in \mathbb{R}^n : |Tf(x)| > t\}| \leqslant \frac{2^{p(n+1)}c^p + 1}{t}\|f\|_1, \quad t > 0,$$

and thus T is of weak type (1,1). Since T is of weak type (1,1) and of weak type (p,p), the Marcinkiewicz interpolation theorem (Theorem 2.4) implies that T is of strong type (q,q), $1 < q < p$. □

The Marcinkiewicz interpolation theorem shows that if a sublinear operator satisfies weak type conditions at the end points of exponents, then it is satisfies the strong type condition for all exponents in between. The next result describes the difference between the weak type and strong type conditions. This result will be used in the proof of Theorem 5.1 later.

Theorem 2.8 (Kolmogorov). Let $1 \leqslant p < \infty$. Assume that T is a sublinear operator from $L^p(\mathbb{R}^n)$ to Lebesgue measurable functions on $\mathbb{R}^n$.

(1) If T is of weak type (p,p), then for all $0 < q < p$ and $A \subset \mathbb{R}^n$ with $0 < |A| < \infty$, there exists $c < \infty$ such that

$$\int_A |Tf(x)|^q \, dx \leqslant c|A|^{1-\frac{q}{p}}\|f\|_p^q. \tag{2.9}$$

(2) If there exists $0 < q < p$ and constant c such that (2.9) holds for every $A \subset \mathbb{R}^n$ with $0 < |A| < \infty$, then T is of weak type (p,p).

THE MORAL: This shows that the weak type condition of a sublinear operator at a given exponent is essentially equivalent to the strong type condition for all strictly smaller exponents.

Proof. $\boxed{(1)}$ Since T is of weak type (p,p), we have

$$|\{x \in A : |Tf(x)| > t\}| \leqslant |\{x \in \mathbb{R}^n : |Tf(x)| > t\}| \leqslant \frac{c}{t^p}\|f\|_p^p, \quad t > 0.$$

Thus for every $0 < q < p$, we obtain

$$\begin{aligned}
\int_A |Tf(x)|^q \, dx &= q\int_0^{\infty} t^{q-1}|\{x \in A : |Tf(x)| > t\}| \, dt \\
&\leqslant \int_0^{\infty} t^{q-1}\min\left\{|A|, \frac{c}{t^p}\|f\|_p^p\right\} dt \\
&= q\int_0^{c\|f\|_p|A|^{-1/p}} t^{q-1}|A| \, dt + cq\int_{c\|f\|_p|A|^{-1/p}}^{\infty} t^{q-1-p}\|f\|_p^p \, dt \\
&\leqslant c|A|^{1-\frac{q}{p}}\|f\|_p^q.
\end{aligned}$$

$\boxed{(2)}$ Let $t > 0$ and $A = \{x \in \mathbb{R}^n : |Tf(x)| > t\}$. Then A is a measurable set and $|A| < \infty$. If $|A| = \infty$, there are sets $A_k \subset A$ with $|A_k| = k$ for $k = 1, 2, \ldots$, such that

$$t^q k = t^q |A_k| \leqslant \int_{A_k} |Tf(x)|^q \, dx \leqslant c|A_k|^{1-\frac{q}{p}} \|f\|_p^q = ck^{1-\frac{q}{p}} \|f\|_p^q.$$

This is impossible. Thus by (2.9), we obtain

$$t^q |A| \leqslant \int_A |Tf(x)|^q \, dx \leqslant c|A|^{1-\frac{q}{p}} \|f\|_p^q.$$

It follows that

$$|\{x \in \mathbb{R}^n : |Tf(x)| > t\}| = |A| \leqslant \frac{c}{t^p} \|f\|_p^p.$$

Thus T is of weak type (p, p). $\square$

3 Bounded mean oscillation

The space of functions of bounded mean oscillation (BMO) turns out to be a natural substitute for $L^\infty(\mathbb{R}^n)$ in harmonic analysis. It consists of functions, whose mean oscillation over cubes is uniformly bounded. Every bounded function belongs to BMO, but there exist unbounded functions with bounded mean oscillation. Such functions typically blow up logarithmically as shown by the John-Nirenberg theorem. The relevance of BMO is attested by the fact that classical singular integral operators fail to map $L^\infty(\mathbb{R}^n)$ to $L^\infty(\mathbb{R}^n)$, but instead they map $L^\infty(\mathbb{R}^n)$ to BMO. Moreover, BMO is the dual space of the Hardy space H^1. BMO also plays a central role in the regularity theory for nonlinear partial differential equations.

3.1 Basic properties of BMO

The mean oscillation of a function $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ in a cube $Q \subset \mathbb{R}^n$ is

$$\fint_Q |f(x) - f_Q|\,dx = \frac{1}{|Q|}\int_Q |f(x) - f_Q|\,dx.$$

The mean oscillation tells that how much function differs in average from its integral average in Q. Observe that, if $f \in L^\infty_{\mathrm{loc}}(\mathbb{R}^n)$, then we have (exercise)

$$\fint_Q |f(x) - f_Q|\,dx \leqslant \operatorname*{ess\,sup}_Q f - \operatorname*{ess\,inf}_Q f.$$

T H E   M O R A L : The mean oscillation is bounded by the pointwise oscillation.

Definition 3.1. Assume that $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and let

$$\|f\|_* = \sup_Q \fint_Q |f(x) - f_Q|\,dx,$$

where the supremum is taken over all cubes Q in $\mathbb{R}^n$. If $\|f\|_* < \infty$, we say that f has bounded mean oscillation. The number $\|f\|_*$ is called the BMO norm of f. The class of functions of bounded mean oscillation is denoted as

$$\mathrm{BMO} = \left\{f \in L^1_{\mathrm{loc}}(\mathbb{R}^n) : \|f\|_* < \infty\right\}.$$

THE MORAL: A function $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ belongs to BMO, if there exists constant $M < \infty$, independent of the cube $Q \subset \mathbb{R}^n$, such that inequality

$$\fint_Q |f(x) - f_Q|\,dx \leqslant M$$

holds for every cube $Q \subset \mathbb{R}^n$. The BMO norm is the smallest constant M for which this is true.

If $f, g \in \mathrm{BMO}$, then

$$\begin{aligned}
\|f+g\|_* &= \sup_Q \fint_Q |f(x)+g(x)-f_Q-g_Q|\,dx \\
&\leqslant \sup_Q \fint_Q |f(x)-f_Q|\,dx + \sup_Q \fint_Q |g(x)-g_Q|\,dx \\
&= \|f\|_* + \|g\|_* < \infty.
\end{aligned}$$

Thus $f+g \in \mathrm{BMO}$ and $\|f+g\|_* \leqslant \|f\|_* + \|g\|_*$. Moreover, if $a \in \mathbb{R}$, then

$$\begin{aligned}
\|af\|_* &= \sup_Q \fint_Q |af(x) - af_Q|\,dx \\
&= |a| \sup_Q \fint_Q |f(x)-f_Q|\,dx = |a|\|f\|_* < \infty.
\end{aligned}$$

Thus $af \in \mathrm{BMO}$ and $\|af\|_* = |a|\|f\|_*$. This shows that BMO is a vector space. However, the BMO norm is not a norm, it is only a seminorm. The reason is that $\|f\|_* = 0$ does not imply that $f = 0$. In fact, we have

$$\|f\|_* = 0 \Longleftrightarrow f = c \quad \text{for almost every} \quad x \in \mathbb{R}^n,$$

where $c \in \mathbb{R}$ is a constant. To see this, for every constant function $f = c$, we have

$$\|f\|_* = \sup_Q \fint_Q |f(x)-f_Q|\,dx = \sup_Q \fint_Q |c-c|\,dx = 0.$$

Conversely, if $\|f\|_* = 0$, then for every cube $Q \subset \mathbb{R}^n$, we have

$$\fint_Q |f(x) - f_Q|\,dx = 0.$$

This implies $|f(x) - f_Q| = 0$ and thus $f(x) = f_Q$ for almost every $x \in Q$. This implies that $f(x) = c$ for almost every $x \in \mathbb{R}^n$. To see this, observe that $f(x) = f_{Q(0,1)}$ for almost every $x \in Q(0,1)$. Moreover, $f(x) = f_{Q(0,k)}$ for almost every $x \in Q(0,k)$,

$k = 1,2,\ldots$. Thus $f_{Q(0,k)} = f_{Q(0,1)}$, $k = 1,2,\ldots$, and $f(x) = f_{Q(0,1)}$ for almost every $x \in Q(0,k)$, $k = 1,2,\ldots$. Since $\mathbb{R}^n = \bigcup_{k=1}^{\infty} Q(0,k)$, this implies $f(x) = f_{Q(0,1)} = c$ for almost every $x \in \mathbb{R}^n$.

THE MORAL: The previous discussion shows that functions f and $f + c$, $c \in \mathbb{R}$, have the same BMO norm.

We may overcome this by identifying all BMO functions, whose difference is constant by considering the equivalence relation

$$f \sim g \Longleftrightarrow f - g = c, \quad c \in \mathbb{R}.$$

The space of corresponding equivalence classes $\widetilde{f}$ is a normed space with norm $\|\widetilde{f}\|_* = \|f\|_*$. Instead of considering the equivalence classes, we identify functions whose difference is constant almost everywhere. In this sense a function in BMO is defined only up to an additive constant.

Next we study basic properties of BMO functions. The following lemma will be a useful tool in showing that certain functions belong to BMO.

Lemma 3.2. Assume that for every cube $Q \subset \mathbb{R}^n$, there exists a constant c_Q, which may depend on Q, such that

$$\fint_Q |f(x) - c_Q| \, dx \leqslant M,$$

where $M < \infty$ is a constant that does not depend on Q. Then $f \in \mathrm{BMO}$ and $\|f\|_* \leqslant 2M$. Moreover,

$$\frac{1}{2}\|f\|_* \leqslant \sup_Q \inf_{c_Q \in \mathbb{R}} \fint_Q |f(x) - c_Q| \, dx \leqslant \|f\|_*.$$

THE MORAL: The integral average in the mean oscillation can be replaced with any other number depending on the cube in the definition of BMO.

Proof. Let Q be a cube in $\mathbb{R}^n$. Then

$$\begin{aligned}
\fint_Q |f(x) - f_Q| \, dx &\leqslant \fint_Q |f - c_Q| \, dx + \fint_Q |f_Q - c_Q| \, dx \\
&\leqslant \fint_Q |f - c_Q| \, dx + |f_Q - c_Q| \\
&= \fint_Q |f - c_Q| \, dx + \left| \fint_Q (f(x) - c_Q) \, dx \right| \\
&\leqslant \fint_Q |f - c_Q| \, dx + \fint_Q |f(x) - c_Q| \, dx \\
&\leqslant 2 \fint_Q |f - c_Q| \, dx \leqslant 2M.
\end{aligned}$$

This implies that

$$\|f\|_* = \sup_Q \fint_Q |f(x) - f_Q| \, dx \leqslant 2M.$$

First by taking infimum over $c_Q \in \mathbb{R}$ for a fixed cube Q and then taking supremum over cubes Q in the estimate above, we obtain

$$\|f\|_* \leqslant 2 \sup_Q \inf_{c_Q \in \mathbb{R}} \fint_Q |f(x) - c_Q| \, dx.$$

On the other hand, it is clear that

$$\inf_{c_Q \in \mathbb{R}} \fint_Q |f(x) - c_Q| \, dx \leqslant \fint_Q |f(x) - f_Q| \, dx.$$

This shows that

$$\sup_Q \inf_{c_Q \in \mathbb{R}} \fint_Q |f(x) - c_Q| \, dx \leqslant \|f\|_*.$$

□

Remark 3.3. For a cube Q a constant c_Q, for which $\inf_{c_Q \in \mathbb{R}} \fint_Q |f(x) - c_Q| \, dx$ is attained, satisfies

$$|\{x \in Q : f(x) > c_Q\}| \leqslant \frac{1}{2}|Q| \quad \text{and} \quad |\{x \in Q : f(x) < c_Q\}| \leqslant \frac{1}{2}|Q|.$$

Even if it is not unique, we call such a constant c_Q the median of f in Q. On the other hand, the constant c_Q for which $\inf_{c_Q \in \mathbb{R}} \fint_Q |f(x) - c_Q|^2 \, dx$ is attained is f_Q (exercise).

THE MORAL: The median minimizes the L^1 mean oscillation and the integral mean value minimizes the L^2 mean oscillation.

Remark 3.4. We use Lemma 3.2 to show that $f \in \mathrm{BMO}$ implies $|f| \in \mathrm{BMO}$. Indeed, by the triangle inequality

$$\fint_Q \big||f(x)| - |f_Q|\big| \, dx \leqslant \fint_Q |f(x) - f_Q| \, dx \leqslant \|f\|_*$$

for every cube Q in $\mathbb{R}^n$. Lemma 3.2 with $c_Q = |f_Q|$ implies $|f| \in \mathrm{BMO}$ and $\||f|\|_* \leqslant 2\|f\|_*$.

Examples 3.5:

(1) We note that $L^\infty(\mathbb{R}^n) \subset \mathrm{BMO}$ and $\|f\|_* \leqslant 2\|f\|_\infty$. This follows, since

$$\begin{aligned}\fint_Q |f(x) - f_Q| \, dx &\leqslant \fint_Q (|f(x)| + |f_Q|) \, dx \leqslant \fint_Q |f(x)| \, dx + |f_Q| \\ &\leqslant \fint_Q |f(x)| \, dx + |f|_Q \leqslant 2 \fint_Q |f(x)| \, dx \leqslant 2\|f\|_\infty\end{aligned}$$

for every cube Q.

THE MORAL: Every bounded function is in BMO.

(2) Let $f:\mathbb{R}^n \to [-\infty,\infty]$, $f(x)=\log|x|$. We show that $f\in \mathrm{BMO}$. By Lemma 3.2 it is enough to show that for every cube $Q(x,l)$, with $x\in\mathbb{R}^n$ and $l>0$, there exists a constant $c_{Q(x,l)}\in\mathbb{R}$ so that

$$\sup_{Q(x,l)} ⨍_{Q(x,l)} |\log|y|-c_{Q(x,l)}|\,dy<\infty.$$

We consider two cases.

Case 1: First assume that $|x|<\sqrt{n}l$. In this case, we choose $c_{Q(x,l)}=\log l$ and by change of variables $y=lz, dy=l^n\,dz$, we obtain

$$\begin{aligned}\frac{1}{l^n}\int_{Q(x,l)} |\log|y|-c_{Q(x,l)}|\,dy &= \int_{Q(\frac{x}{l},1)} |\log(l|z|)-\log l|\,dz\\ &= \int_{Q(\frac{x}{l},1)} |\log|z||\,dz\end{aligned}$$

Claim: $\displaystyle\int_{Q(\frac{x}{l},1)} |\log|z||\,dz \leqslant \int_{B(0,2\sqrt{n})} |\log|z||\,dz<\infty$ whenever $|x|<\sqrt{n}l$.

Reason.

$$|z|\leqslant \frac{|x|}{l}+\frac{\sqrt{n}}{2}<\sqrt{n}+\frac{\sqrt{n}}{2}\leqslant 2\sqrt{n}$$

for every $z\in Q(\frac{x}{l},1)$. Thus $Q(\frac{x}{l},1)\subset B(0,2\sqrt{n})$. ■

Case 2: Assume then that $|x|\geqslant\sqrt{n}l$. In this case, we choose $c_{Q(x,l)}=\log|x|$, do the same change of variables as above, and obtain

$$\begin{aligned}\frac{1}{l^n}\int_{Q(x,l)} |\log|y|-c_{Q(x,l)}|\,dy &= \int_{Q(\frac{x}{l},1)} |\log l|z|-\log|x||\,dz\\ &= \int_{Q(\frac{x}{l},1)} \left|\log\frac{l|z|}{|x|}\right|\,dz\end{aligned}$$

Claim: $\displaystyle\int_{Q(\frac{x}{l},1)} \left|\log\frac{l|z|}{|x|}\right|\,dz\leqslant\log 2<\infty$ whenever $|x|\geqslant\sqrt{n}l$.

Reason. We note that

$$|z|\geqslant\frac{|x|}{l}-\frac{\sqrt{n}}{2}\Longrightarrow\frac{l|z|}{|x|}\geqslant 1-\frac{l}{|x|}\frac{\sqrt{n}}{2}\geqslant 1-\frac{1}{2}=\frac{1}{2}$$

for every $z\in Q(\frac{x}{l},1)$ and

$$|z|\leqslant\frac{|x|}{l}+\frac{\sqrt{n}}{2}\Longrightarrow\frac{l|z|}{|x|}\leqslant 1+\frac{l}{|x|}\frac{\sqrt{n}}{2}\leqslant 1+\frac{1}{2}=\frac{3}{2}$$

for every $z\in Q(\frac{x}{l},1)$. Thus

$$\log\frac{1}{2}\leqslant\log\frac{l|z|}{|x|}\leqslant\log\frac{3}{2}$$

for every $z\in Q(\frac{x}{l},1)$. This implies

$$\left|\log\frac{l|z|}{|x|}\right|\leqslant-\log\frac{1}{2}=\log 2$$

for every $z\in Q(\frac{x}{l},1)$. ■

We conclude that in both cases the mean oscillation is uniformly bounded and thus $\log|x| \in \mathrm{BMO}$.

THE MORAL: $\log|x| \in \mathrm{BMO}$, but $\log|x| \notin L^\infty(\mathbb{R}^n)$. This is an example of an unbounded BMO function. Thus inclusion $L^\infty(\mathbb{R}^n)$ is a proper subset of BMO.

(3) Let $f : \mathbb{R} \to \mathbb{R}$,

$$f(x) = \begin{cases} \log|x|, & x \leqslant 0, \\ -\log|x|, & x > 0. \end{cases}$$

Since f is an odd function, we have

$$\frac{1}{2a} \int_{-a}^{a} f(x)\,dx = 0$$

for every interval $[-a,a] \subset \mathbb{R}$. For $0 < a < 1$ we have

$$\begin{aligned} \|f\|_* &\geqslant \frac{1}{2a} \int_{-a}^{a} |f(x)|\,dx = \frac{1}{a} \int_0^a -\log x\,dx \\ &= \frac{1}{a} \lim_{\varepsilon \to 0+} \Big/_{\varepsilon}^{a} (x - x\log x) = \frac{1}{a}(a - a\log a) \\ &= 1 - \log a \xrightarrow{a \to 0+} \infty. \end{aligned}$$

Thus $f \notin \mathrm{BMO}$, even though by (2) and Remark 3.4, we have $|f(x)| = |\log|x|| \in \mathrm{BMO}$.

THE MORAL: $|f| \in \mathrm{BMO}$ does not imply that $f \in \mathrm{BMO}$. Since $f = |f| \operatorname{sgn} f$ with $|f| \in \mathrm{BMO}$ and $\operatorname{sgn} f \in L^\infty(\mathbb{R}^n) \subset \mathrm{BMO}$, this also shows that product of two functions in BMO does not necessarily belong to BMO.

(4) Consider dyadic BMO, where the supremum of the mean oscillation is taken only over dyadic cubes. It is clear that BMO is a subset of dyadic BMO, but the converse inclusion is not true. For example, the function in (3) belongs to dyadic BMO, but it does not belong to BMO (exercise).

THE MORAL: BMO is a proper subset of dyadic BMO.

(5) Let $g : \mathbb{R} \to \mathbb{R}$,

$$g(x) = \begin{cases} \log x, & x > 0, \\ 0, & x \leqslant 0. \end{cases}$$

Let f be as in (3). Since $\log|x| \in \mathrm{BMO}$ by (2) and $f \notin \mathrm{BMO}$ by (3), the difference

$$g(x) = \frac{1}{2}(\log|x| - f(x)) = \chi_{(0,\infty)} \log|x| \notin \mathrm{BMO}.$$

THE MORAL: $f \in \mathrm{BMO}$ does not imply that $f\chi_A \in \mathrm{BMO}$, $A \subset \mathbb{R}^n$.

Theorem 3.6. Assume that $f,g \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$.

(1) If $f,g \in \mathrm{BMO}$, then $\max(f,g) \in \mathrm{BMO}$ and $\min\{f,g\} \in \mathrm{BMO}$. Moreover, we have

$$\|\max\{f,g\}\|_* \leqslant \frac{3}{2}\|f\|_* + \frac{3}{2}\|g\|_* \quad \text{and} \quad \|\min\{f,g\}\|_* \leqslant \frac{3}{2}\|f\|_* + \frac{3}{2}\|g\|_*.$$

(2) If $f \in \mathrm{BMO}$ and $h \in \mathbb{R}^n$, then function $\tau_h f \in \mathrm{BMO}$, where $\tau_h f(x) = f(x+h)$. Moreover, we have $\|\tau_h f\|_* = \|f\|_*$.

(3) If $f \in \mathrm{BMO}$ and $a \in \mathbb{R}$, $a \neq 0$, then function $\delta_a f \in \mathrm{BMO}$, where $\delta_a f(x) = f(ax)$. Moreover, we have $\|\delta_a f\|_* = \|f\|_*$.

Proof. $\boxed{(1)}$ Since $\max\{f,g\} = \frac{1}{2}(f+g+|f-g|)$, by Remark 3.4, we obtain

$$\begin{aligned}\|\max\{f,g\}\|_* &\leqslant \frac{1}{2}\||f-g|\|_* + \frac{1}{2}\|f\|_* + \frac{1}{2}\|g\|_* \\ &\leqslant \|f-g\|_* + \frac{1}{2}\|f\|_* + \frac{1}{2}\|g\|_* \\ &\leqslant \frac{3}{2}\|f\|_* + \frac{3}{2}\|g\|_*.\end{aligned}$$

On the other hand, $\min\{f,g\} = \frac{1}{2}(f+g-|f-g|)$ and a similar argument as above shows that $\min\{f,g\} \in \mathrm{BMO}$.

$\boxed{(2)}$ Change of variables $y = x+h$, $dx = dy$, gives

$$(\tau_h f)_Q = \frac{1}{|Q|}\int_Q \tau_h f(x)\,dx = \frac{1}{|Q|}\int_Q f(x+h)\,dx = \frac{1}{|Q+h|}\int_{Q+h} f(y)\,dy = f_{Q+h}$$

for every cube Q in $\mathbb{R}^n$. Here $Q+h = \{z+h : z \in Q\}$. Thus

$$\begin{aligned}\|\tau_h f\|_* &= \sup_Q \frac{1}{|Q|}\int_Q |\tau_h f(x) - (\tau_h f)_Q|\,dx \\ &= \sup_Q \frac{1}{|Q|}\int_Q |f(x+h) - f_{Q+h}|\,dx \\ &= \sup_Q \frac{1}{|Q+h|}\int_{Q+h} |f(y) - f_{Q+h}|\,dy \\ &= \sup_Q \frac{1}{|Q|}\int_Q |f(y) - f_Q|\,dy = \|f\|_*.\end{aligned}$$

$\boxed{(3)}$ Change of variables $y = ax$, $dx = |a|^{-n}\,dy$ gives

$$\begin{aligned}(\delta_a f)_Q &= \frac{1}{|Q|}\int_Q \delta_a f(x)\,dx = \frac{1}{|Q|}\int_Q f(ax)\,dx \\ &= \frac{1}{|Q|}\int_{aQ} |a|^{-n} f(y)\,dy = \frac{1}{|aQ|}\int_{aQ} f(y)\,dy = f_{aQ},\end{aligned}$$

where $aQ = \{az : z \in Q\}$. Thus

$$
\begin{aligned}
\|\delta_a f\|_* &= \sup_Q \frac{1}{|Q|} \int_Q |\delta_a f(x) - (\delta_a f)_Q| \, dx \\
&= \sup_Q \frac{1}{|Q|} \int_Q |f(ax) - f_{aQ}| \, dx \\
&= \sup_Q \frac{1}{|aQ|} \int_{aQ} |f(y) - f_{aQ}| \, dy \\
&= \sup_Q \frac{1}{|Q|} \int_Q |f(y) - f_Q| \, dy = \|f\|_*. \qquad \square
\end{aligned}
$$

Remark 3.7. Every function $f \in \mathrm{BMO}$ can be approximated pointwise by an increasing sequence of bounded BMO functions, since truncations

$$
f_k(x) = \min\{\max\{f(x), -k\}, k\} = \begin{cases} k, & f(x) > k, \\ f(x), & -k \leqslant f(x) \leqslant k, \\ -k, & f(x) < -k, \end{cases}
$$

$k = 1, 2, \ldots$, belong to BMO with $\|f_k\| \leqslant \frac{9}{4}\|f\|_*$ for every $k = 1, 2, \ldots$, $f_k \to f$ pointwise and $f_k \to f$ in $L^1_{\mathrm{loc}}(\mathbb{R}^n)$ as $k \to \infty$ (exercise).

3.2 Completeness of BMO.

We show that BMO is a Banach space using the argument from [17]. The proof is based on the Riesz-Fischer theorem, which asserts that the L^p space is complete.

Theorem 3.8. BMO is a Banach space with the norm $\|\cdot\|_*$. Here we consider BMO as a space of equivalence classes of functions up to an additive constant.

Proof. Let $(f_i)_i$ be a Cauchy sequence in BMO. Fix a cube Q and consider the representative $f_i - (f_i)_Q$ of equivalence class $\widetilde{f}_i$. We claim that $(f_i - (f_i)_Q)_i$ is a Cauchy sequence in $L^1(Q)$. Since (f_i) is a Cauchy sequence in BMO, for every $\varepsilon > 0$ there exists i_ε such that $\|f_i - f_j\|_* < \varepsilon$, if $i, j \geqslant i_\varepsilon$. We observe that

$$
\begin{aligned}
\big\|\big(f_i - (f_i)_Q\big) - \big(f_j) - (f_j)_Q\big)\big\|_{L^1(Q)} &= \int_Q \big|\big(f_i(x) - (f_i)_Q\big) - \big(f_j(x) - (f_j)_Q\big)\big| \, dx \\
&= |Q| \fint_Q \big|\big(f_i(x) - (f_i)_Q\big) - \big(f_j(x) - (f_j)_Q\big)\big| \, dx \\
&= |Q| \fint_Q \big|\big(f_i(x) - f_j(x)\big) - (f_i - f_j)_Q\big| \, dx \\
&\leqslant |Q| \|f_i - f_j\|_* < \varepsilon |Q| \quad \text{when} \quad i, j \geqslant i_\varepsilon.
\end{aligned}
$$

This shows that $(f_i - (f_i)_Q)_i$ is a Cauchy sequence in $L^1(Q)$. Since $L^1(Q)$ a complete space, we conclude that sequence $(f_i - (f_i)_Q)_i$ converges in $L^1(Q)$ to function

$g^Q \in L^1(Q)$ and thus

$$\big\|\big(f_i - (f_i)_Q\big) - g^Q\big\|_{L^1(Q)} \xrightarrow{i\to\infty} 0. \tag{3.9}$$

Let $Q' \supset Q$ be a cube containing Q. As above, the sequence $(f_i - (f_i)_{Q'})_i$ converges in $L^1(Q')$ to a function $g^{Q'}$ and thus

$$\big\|\big(f_i - (f_i)_{Q'}\big) - g^{Q'}\big\|_{L^1(Q')} \xrightarrow{i\to\infty} 0. \tag{3.10}$$

Then we consider the sequence $((f_i)_Q - (f_i)_{Q'})_i$ of real numbers, whose terms can be also interpreted as constant functions on cube Q. It follows from (3.9) and (3.10) that

$$\begin{aligned}
&\big\|\big((f_i)_Q - (f_i)_{Q'}\big) - (g^{Q'} - g^Q)\big)\big\|_{L^1(Q)} \\
&\quad = \big\|\big(\big(f_i - (f_i)_{Q'}\big) - g^{Q'}\big) - \big(\big(f_i - (f_i)_Q\big) - g^Q\big)\big\|_{L^1(Q)} \\
&\quad \leqslant \big\|\big(f_i - (f_i)_{Q'}\big) - g^{Q'}\big\|_{L^1(Q)} + \big\|\big(f_i - (f_i)_Q\big) - g^Q\big\|_{L^1(Q)} \xrightarrow{i\to\infty} 0.
\end{aligned}$$

Since the sequence $((f_i)_Q - (f_i)_{Q'})_i$ of constant functions converges to $g^{Q'} - g^Q$ in $L^1(Q)$ as $i \to \infty$, there exists a subsequence that converges almost everywhere in Q. This implies that $g^{Q'} - g^Q \in L^1(Q)$ is a constant function in Q. Let

$$g^{Q'} - g^Q = C(Q,Q') \quad \text{in } Q, \tag{3.11}$$

where $C(Q,Q')$ is a constant. On the other hand, by (3.10) we have

$$\begin{aligned}
\left| \fint_Q \big(f_i - (f_i)_{Q'}\big)\,dx - \fint_Q g^{Q'}\,dx \right| &\leqslant \fint_Q \big|\big(f_i - (f_i)_{Q'}\big) - g^{Q'}(x)\big|\,dx \\
&= \frac{1}{|Q|} \big\|\big(f_i - (f_i)_{Q'}\big) - g^{Q'}\big\|_{L^1(Q')} \xrightarrow{i\to\infty} 0.
\end{aligned}$$

It follows that

$$(f_i)_Q - (f_i)_{Q'} = \fint_Q \big(f_i - (f_i)_{Q'}\big)\,dx \xrightarrow{i\to\infty} \fint_Q g^{Q'}\,dx$$

and thus

$$C(Q,Q') = \fint_Q g^{Q'}\,dx. \tag{3.12}$$

We define the function f as follows. Let $Q_k = Q(0,k)$, $k = 1,2,\dots$ and

$$f = g^{Q_k} - C(Q_1,Q_k) \quad \text{in } Q_k. \tag{3.13}$$

In principle, this definition makes sense, since every $x \in \mathbb{R}^n$ belongs to Q_k for k large enough, but we have to show that f is well defined, that is, if $1 < k < k'$, then

$$g^{Q_{k'}} - C(Q_1,Q_{k'}) = g^{Q_k} - C(Q_1,Q_k) \quad \text{in } Q_k.$$

By (3.11) this is equivalent to showing that

$$C(Q_1,Q_{k'}) - C(Q_1,Q_k) = C(Q_k,Q_{k'})$$

whenever $1<k<k'$. However, this follows from (3.12) and (3.11), since

$$\begin{aligned} C(Q_1,Q_{k'})-C(Q_1,Q_k) &= \fint_{Q_1} g^{Q_{k'}}\,dx - \fint_{Q_1} g^{Q_k}\,dx \\ &= \fint_{Q_1} \big(g^{Q_{k'}}-g^{Q_k}\big)\,dx \\ &= \fint_{Q_1} C(Q_k,Q_{k'})\,dx = C(Q_k,Q_{k'}) \end{aligned}$$

This shows that f is well defined.

We claim that sequence $(f_i)_i$ converges to f defined by (3.13) in BMO. It follows from (3.13) that $f\in L^1_{\mathrm{loc}}(\mathbb{R}^n)$. Then we show that f is the required limit function. Let $\varepsilon>0$. Since $(f_i)_i$ is a Cauchy sequence in BMO, there exists i_ε such that $\|f_i-f_j\|_*<\varepsilon$, if $i,j\geqslant i_\varepsilon$, or, equivalently, for every cube Q in $\mathbb{R}^n$, we have

$$\fint_Q \big|\big(f_i-(f_i)_Q\big)-\big(f_j-(f_j)_Q\big)\big|\,dx<\varepsilon, \quad \text{if} \quad i,j\geqslant i_\varepsilon.$$

By letting $j\to\infty$ and using (3.9), for every cube Q in $\mathbb{R}^n$, we have

$$\fint_Q \big|\big(f_i-(f_i)_Q\big)-g^Q\big|\,dx\leqslant\varepsilon, \quad \text{if} \quad i\geqslant i_\varepsilon.$$

Every cube Q in $\mathbb{R}^n$ is contained in Q_k for k large enough, so that by (3.13), (3.12) and (3.11), we obtain

$$\begin{aligned} \fint_Q \big|(f_i-f)-(f_i-f)_Q\big|\,dx &= \fint_Q \big|f_i-g^{Q_k}+\underbrace{C(Q_1,Q_k)}_{=(g^{Q_k})_{Q_1}}-(f_i)_Q+f_Q\big|\,dx \\ &= \fint_Q \big|f_i-g^{Q_k}+(g^{Q_k})_{Q_1}-(f_i)_Q+f_Q\big|\,dx \\ &= \fint_Q \big|f_i-(f_i)_Q-g^Q+\underbrace{\big((g^{Q_k})_{Q_1}-\big(g^{Q_k}-g^Q\big)+f_Q\big)}_{=0}\big|\,dx \\ &= \fint_Q \big|f_i-(f_i)_Q-g^Q\big|\,dx\leqslant\varepsilon, \quad \text{if} \quad i\geqslant i_\varepsilon. \end{aligned}$$

Here we used the fact that

$$\begin{aligned} (g^{Q_k})_{Q_1}-\big(g^{Q_k}-g^Q(x)\big)+f_Q &= C(Q_1,Q_k)-C(Q,Q_k)+\fint_Q f\,dx \\ &= C(Q_1,Q_k)-C(Q,Q_k)+\fint_Q \big(g^{Q_k}(x)-C(Q_1,Q_k)\big)\,dx \\ &= C(Q_1,Q_k)-C(Q,Q_k)+(g^{Q_k})_Q-C(Q_1,Q_k) \\ &= C(Q_1,Q_k)-C(Q,Q_k)+C(Q,Q_k)-C(Q_1,Q_k)=0. \end{aligned}$$

This shows that $f\in\mathrm{BMO}$ and

$$\|f_i-f\|_*=\sup_Q \fint_Q \big|(f_i-f)-(f_i-f)_Q\big|\,dx \xrightarrow{i\to\infty} 0.$$

$\square$

3.3 The John-Nirenberg inequality

We begin with an example.

Example 3.14. Consider $f:\mathbb{R}\to\mathbb{R}$, $f(x)=\log|x|$, which is an example of an unbounded function in BMO, see Example 3.5 (2). Then

$$f_{[-a,a]}=\frac{1}{2a}\int_{-a}^{a}\log|x|\,dx=\frac{1}{a}\int_{0}^{a}\log x\,dx=\frac{1}{a}\Big/_{0}^{a}(x\log x-x)=\log a-1,\quad a>0,$$

and thus for $x\in[-a,a]$ and $t>1$, we have

$$|f(x)-f_{[-a,a]}|>t\iff|\log|x|-(\log a-1)|>t\iff\left|\log\frac{e|x|}{a}\right|>t\iff|x|<\frac{a}{e}e^{-t}.$$

This implies

$$|\{x\in[-a,a]:|f(x)-f_{[-a,a]}|>t\}|=2ae^{-t-1},\quad t>1,$$

THE MORAL: The distribution function decays exponentially.

The John-Nirenberg inequality gives a similar exponential estimate for the distribution function of oscillation of an arbitrary BMO function. The proof that we present here is based on a recursive use of the Calderón-Zygmund decomposition.

Theorem 3.15 (The John-Nirenberg lemma (1961)). There exists constants c_1 and c_2, depending only on dimension n, such that if $f\in\mathrm{BMO}$, then

$$|\{x\in Q:|f(x)-f_Q|>t\}|\leqslant c_1\exp\left(-\frac{c_2t}{\|f\|_*}\right)|Q|,\quad t>0,$$

for every cube $Q\subset\mathbb{R}^n$.

THE MORAL: The John-Nirenberg inequality tells that logarithmic blowup, as for $f(x)=\log|x|$, is the worst possible behaviour for a general BMO function. In this sense the John-Nirenberg inequality is the best possible result we can hope for.

Proof. Let $Q\subset\mathbb{R}^n$ be a cube and let $s>\|f\|_*$ to be a parameter, which is to be chosen later. Then

$$\fint_Q|f(y)-f_Q|\,dy\leqslant\|f\|_*<s.$$

Apply the Calderón-Zygmund decomposition for $f-f_Q$ at level s in Q, see Theorem 1.4. We obtain pairwise disjoint dyadic subcubes Q_{i_1}, $i_1=1,2,\dots$, of Q such that

$$s<\fint_{Q_{i_1}}|f(y)-f_Q|\,dy\leqslant 2^ns,\quad i_1=1,2,\dots,$$

and

$$|f(x)-f_Q| \leqslant s \quad \text{for almost every} \quad x \in Q \setminus \bigcup_{i_1=1}^{\infty} Q_{i_1}. \tag{3.16}$$

This implies

$$|f_{Q_{i_1}} - f_Q| = \left| \fint_{Q_{i_1}} (f(y)-f_Q)\,dy \right| \leqslant \fint_{Q_{i_1}} |f(y)-f_Q|\,dy \leqslant 2^n s, \quad i_1 = 1,2,\dots. \tag{3.17}$$

Since cubes Q_{i_1} are pairwise disjoint subcubes of Q and $f \in \mathrm{BMO}$, we obtain

$$\sum_{i_1=1}^{\infty} |Q_{i_1}| \leqslant \frac{1}{s} \sum_{i_1=1}^{\infty} \int_{Q_{i_1}} |f(y)-f_Q|\,dy \leqslant \frac{1}{s} \int_Q |f(y)-f_Q|\,dy \leqslant \frac{\|f\|_*}{s} |Q|. \tag{3.18}$$

We proceed recursively. Since $f \in \mathrm{BMO}$, we have

$$\fint_{Q_{i_1}} |f(y)-f_{Q_{i_1}}|\,dy \leqslant \|f\|_* < s \quad \text{for every} \quad i_1 = 1,2,\dots.$$

We apply the Calderón-Zygmund decomposition for $f - f_{Q_{i_1}}$ at level s in every cube Q_{i_1}, $i_1 = 1,2,\dots$. We obtain pairwise disjoint dyadic subcubes Q_{i_1,i_2}, $i_2 = 1,2,\dots$, of Q_{i_1} such that

$$s < \fint_{Q_{i_1,i_2}} |f(y)-f_{Q_{i_1}}|\,dy \leqslant 2^n s, \quad i_2 = 1,2,\dots$$

and

$$|f(x)-f_{Q_{i_1}}| \leqslant s \quad \text{for almost every} \quad x \in Q_{i_1} \setminus \bigcup_{i_2=1}^{\infty} Q_{i_1,i_2}. \tag{3.19}$$

This gives

$$\begin{aligned} \sum_{i_2=1}^{\infty} |Q_{i_1,i_2}| &\leqslant \frac{1}{s} \sum_{i_2=1}^{\infty} \int_{Q_{i_1,i_2}} |f(y)-f_{Q_{i_1}}|\,dy \\ &\leqslant \frac{1}{s} \int_{Q_{i_1}} |f(y)-f_{Q_{i_1}}|\,dy \leqslant \frac{\|f\|_*}{s} |Q_{i_1}|, \quad i_1 = 1,2,\dots. \end{aligned} \tag{3.20}$$

By (3.17) and (3.19), for every $i_1 = 1,2,\dots$, we have

$$\begin{aligned} |f(x)-f_Q| &\leqslant |f(x)-f_{Q_{i_1}}| + |f_{Q_{i_1}} - f_Q| \\ &\leqslant s + 2^n s \leqslant 2 \cdot 2^n s \quad \text{for almost every} \quad x \in Q_{i_1} \setminus \bigcup_{i_2=1}^{\infty} Q_{i_1,i_2}. \end{aligned}$$

By (3.16), we obtain

$$|f(x)-f_Q| \leqslant 2 \cdot 2^n s \quad \text{for almost every} \quad x \in Q \setminus \bigcup_{i_1,i_2=1}^{\infty} Q_{i_1,i_2}$$

and by (3.18) and (3.20), we have

$$\sum_{i_1,i_2=1}^{\infty} |Q_{i_1,i_2}| = \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} |Q_{i_1,i_2}| \leqslant \sum_{i_1=1}^{\infty} \frac{\|f\|_*}{s} |Q_{i_1}| \leqslant \left(\frac{\|f\|_*}{s} \right)^2 |Q|.$$

At kth step we observe that

$$\fint_{Q_{i_1,\dots,i_{k-1}}} |f(y)-f_{Q_{i_1,\dots,i_{k-1}}}|\,dy \leqslant \|f\|_* < s \quad \text{for every} \quad i_1,\dots,i_{k-1}=1,2,\dots.$$

We apply the Calderón-Zygmund decomposition for $f-f_{Q_{i_1,\dots,i_{k-1}}}$ at level s in every cube $Q_{i_1,\dots,i_{k-1}}$, $i_1,\dots,i_{k-1}=1,2,\dots$. We obtain pairwise disjoint dyadic subcubes $Q_{i_1,\dots,i_k}$, $i_k=1,2,\dots$, of $Q_{i_1,\dots,i_{k-1}}$ such that

$$s < \fint_{Q_{i_1,\dots,i_k}} |f(y)-f_{Q_{i_1,\dots,i_{k-1}}}|\,dy \leqslant 2^n s, \quad i_k=1,2,\dots$$

and

$$|f(x)-f_{Q_{i_1,\dots,i_{k-1}}}| \leqslant s \quad \text{for almost every} \quad x \in Q_{i_1,\dots,i_{k-1}} \setminus \bigcup_{i_k=1}^{\infty} Q_{i_1,\dots,i_k}.$$

As above, we have

$$\sum_{i_1,\dots,i_k=1}^{\infty} |Q_{i_1,\dots,i_k}| \leqslant \left(\frac{\|f\|_*}{s}\right)^k |Q|,$$

and

$$|f(x)-f_Q| \leqslant k2^n s \quad \text{for almost every} \quad x \in Q \setminus \bigcup_{i_1,\dots,i_k=1}^{\infty} Q_{i_1,\dots,i_k}.$$

In other words, almost every point of the set $\{x\in Q : |f(x)-f_Q| > k2^n s\}$ belongs to some of the cubes $Q_{i_1,\dots,i_k}$, $i_1,\dots,i_k=1,2,\dots$.

Let us then complete the proof of the exponential estimate for the distribution function. To this end, first assume that $t \geqslant 2^n s$. Then we choose an integer k with $k2^n s \leqslant t < (k+1)2^n s$. By the beginning of the proof, we have

$$\{x\in Q : |f(x)-f_Q| > t\} \subset \{x \in Q : |f(x)-f_Q| > k2^n s\} \subset \bigcup_{i_1,\dots,i_k=1}^{\infty} Q_{i_1,\dots,i_k},$$

where the last inclusion holds up to a set of measure zero. Thus

$$\begin{aligned} |\{x\in Q : |f(x)-f_Q| > t\}| &\leqslant \left| \bigcup_{i_1,\dots,i_k=1}^{\infty} Q_{i_1,\dots,i_k} \right| = \sum_{i_1,\dots,i_k=1}^{\infty} |Q_{i_1,\dots,i_k}| \\ &\leqslant \left(\frac{\|f\|_*}{s}\right)^k |Q| \leqslant \exp\left(-k\log\frac{s}{\|f\|_*}\right)|Q|. \end{aligned}$$

Since

$$t < (k+1)2^n s = k2^n s + 2^n s \leqslant 2k2^n s = k2^{n+1}s,$$

we have

$$k > \frac{t}{2^{n+1}s},$$

from which it follows that

$$\exp\left(-k\log\frac{s}{\|f\|_*}\right) \leqslant \exp\left(-\frac{t}{2^{n+1}s}\log\frac{s}{\|f\|_*}\right).$$

Recall that $s > \|f\|_*$ is a free parameter and it can be chosen as we want. By choosing $s = 2\|f\|_*$, we obtain

$$|\{x \in Q : |f(x) - f_Q| > t\}| \leqslant \exp\left(-\frac{t}{2^{n+1}2\|f\|_*}\log\frac{2\|f\|_*}{\|f\|_*}\right)|Q| = \exp\left(-\frac{c_2 t}{\|f\|_*}\right)|Q|,$$

where $c_2 = \frac{\log 2}{2^{n+2}}$. This proves the claim in the case $t \geqslant 2^n s$.

Then assume that $0 < t < 2^n s = 2^{n+1}\|f\|_*$. In this case

$$\frac{t}{2^{n+1}\|f\|_*} < 1,$$

and thus

$$|\{x \in Q : |f(x) - f_Q| > t\}| \leqslant |Q| e \cdot e^{-1} \leqslant e \cdot \exp\left(-\frac{t}{2^{n+1}\|f\|_*}\right)|Q|.$$

Thus, in both cases, we have

$$|\{x \in Q : |f(x) - f_Q| > t\}) \leqslant c_1 \exp\left(-\frac{c_2 t}{\|f\|_*}\right)|Q|,$$

with $c_1 = e$ and $c_2 = \frac{\log 2}{2^{n+2}}$. □

Remarks 3.21:

(1) The John-Nirenberg lemma can be stated in the following form: There exists a constant $c_1 = c_1(n) > 0$ such that

$$|\{x \in Q : |f(x) - f_Q| > t\}| \leqslant c_1 \exp\left(-\frac{t}{c_1\|f\|_*}\right)|Q|, \quad t > 0,$$

for every cube $Q \subset \mathbb{R}^n$.

(2) The John-Nirenberg lemma gives a characterization of BMO. Assume that there exist constants c_1 and c_2, independent of cube Q, such that

$$|\{x \in Q : |f(x) - f_Q| > t\}| \leqslant c_1 \exp(-c_2 t)|Q|, \quad t > 0,$$

for every cube Q in $\mathbb{R}^n$. Then

$$\begin{aligned}\frac{1}{|Q|}\int_Q |f(x) - f_Q|\,dx &= \frac{1}{|Q|}\int_0^\infty |\{x \in Q : |f(x) - f_Q| > t\}|\,dt \\ &\leqslant c_1 \int_0^\infty \exp(-c_2 t)\,dt = -\frac{c_1}{c_2}\Big/_0^\infty \exp(-c_2 t) = \frac{c_1}{c_2}\end{aligned}$$

for every cube Q. Thus $f \in \mathrm{BMO}$ with $\|f\|_* \leqslant \frac{c_1}{c_2}$.

(3) If a function f satisfies the BMO condition with the average f_Q replaced with some other constants c_Q, see Lemma 3.2, the same is true for the John-Nirenberg inequality. Assume that for every cube $Q \subset \mathbb{R}^n$, there is constant c_Q, which may depend on Q, such that

$$\fint_Q |f(x) - c_Q|\,dx \leqslant M,$$

where $M < \infty$ is a constant that does not depend on Q. Then Lemma 3.2 implies $f \in \mathrm{BMO}$ and $\|f\|_* \leqslant 2M$ and the John-Nirenberg lemma applies. Moreover,

$$\begin{aligned}|f(x)-c_Q| &\leqslant |f(x)-f_Q| + |f_Q - c_Q| \\ &\leqslant |f(x)-f_Q| + \fint_Q |f-c_Q|\,dx \leqslant |f(x)-f_Q| + M,\end{aligned}$$

and thus

$$\begin{aligned}|\{x \in Q : |f(x)-c_Q| > t\}| &\leqslant |\{x \in Q : |f(x)-f_Q| + M > t\}| \\ &\leqslant c_1 \exp\left(-\frac{c_2(t-M)}{\|f\|_*}\right)|Q| = c_1' \exp\left(-\frac{c_2 t}{\|f\|_*}\right)|Q|\end{aligned}$$

for every $t > M$ and every cube $Q \subset \mathbb{R}^n$. Here c_1' depends on n and $\|f\|_*$. For $0 < t \leqslant M$ we have

$$|\{x \in Q : |f(x)-c_Q| > t\}| \leqslant |Q| \leqslant c_1' \exp(-c_2 M)|Q| \leqslant c_1 \exp(-c_2 t)|Q|,$$

where c_1' depends on n and $\|f\|_*$.

Remark 3.22. The proof of Theorem 3.15 can be modified to give a slightly better estimate than in Theorem 3.15, see [13]. By applying (3.18) in the form

$$\sum_{i_1=1}^{\infty} |Q_{i_1}| \leqslant \frac{1}{s}\int_Q |f(y)-f_Q|\,dy$$

we obtain the bound

$$\begin{aligned}\sum_{i_1,\dots,i_k=1}^{\infty} |Q_{i_1,\dots,i_k}| &\leqslant \left(\frac{\|f\|_*}{s}\right)^{k-1} \frac{1}{s}\int_Q |f(y)-f_Q|\,dy \\ &= \left(\frac{\|f\|_*}{s}\right)^{k} \frac{1}{\|f\|_*}\int_Q |f(y)-f_Q|\,dy.\end{aligned}$$

By choosing $s = 2\|f\|_*$ as in the proof of Theorem 3.15, we obtain

$$|\{x \in Q : |f(x)-f_Q| > t\}| \leqslant \exp\left(-\frac{c_2 t}{\|f\|_*}\right) \frac{1}{\|f\|_*}\int_Q |f(y)-f_Q|\,dy, \tag{3.23}$$

for every $t \geqslant 2^n s = 2^{n+1}\|f\|_*$ with $c_2 = \frac{\log 2}{2^{n+2}}$.

Let $p > 1$ and $s_0 = 2^{n+1}\|f\|_*$. Denote $E_t = \{x \in Q : |f(x)-f_Q| > t\}$, $t > 0$. Then

$$\begin{aligned}\int_Q |f(x)-f_Q|^p\,dx &= p\int_0^{\infty} t^{p-1}|E_t|\,dt \\ &= p\int_0^{s_0} t^{p-1}|E_t|\,dt + p\int_{s_0}^{\infty} t^{p-1}|E_t|\,dt\end{aligned}$$

For $0 < t \leqslant s_0$ we have $t^{p-1} \leqslant s_0^{p-1}$, which implies that

$$\begin{aligned}p\int_0^{s_0} t^{p-1}|E_t|\,dt &\leqslant p s_0^{p-1}\int_0^{s} |E_t|\,dt \leqslant p s_0^{p-1}\int_0^{\infty} |E_t|\,dt \\ &= p s_0^{p-1}\int_Q |f(x)-f_Q|\,dx.\end{aligned}$$

By (3.23), we obtain

$$\begin{aligned}
p\int_{s_0}^{\infty} t^{p-1}|E_t|\,dt &\leqslant \frac{p}{\|f\|_*}\int_Q |f(x)-f_Q|\,dx \int_0^{\infty} t^{p-1}\exp\left(-\frac{c_2 t}{\|f\|_*}\right)dt\\
&= \frac{p}{\|f\|_*}\int_Q |f(x)-f_Q|\,dx \left(\frac{\|f\|_*}{c_2}\right)^{p-1}\int_0^{\infty}\left(\frac{c_2 t}{\|f\|_*}\right)^{p-1}\exp\left(-\frac{c_2 t}{\|f\|_*}\right)dt\\
&= \frac{p}{\|f\|_*}\int_Q |f(x)-f_Q|\,dx \left(\frac{\|f\|_*}{c_2}\right)^{p}\int_0^{\infty} s^{p-1}e^{-s}\,ds\\
&= \frac{p\Gamma(p)}{c_2^p}\|f\|_*^{p-1}\int_Q |f(x)-f_Q|\,dx
\end{aligned}$$

for every cube Q in $\mathbb{R}^n$. Here Γ is the gamma function. This implies

$$\begin{aligned}
\int_Q |f(x)-f_Q|^p\,dx &\leqslant p\left(s_0^{p-1}+\frac{\Gamma(p)}{c_2^p}\|f\|_*^{p-1}\right)\int_Q |f(x)-f_Q|\,dx\\
&\leqslant p\left(2^{(n+1)(p-1)}+\frac{\Gamma(p)}{c_2^p}\right)\|f\|_*^{p-1}\int_Q |f(x)-f_Q|\,dx.
\end{aligned}$$

This proves that if $f\in \mathrm{BMO}$, there exists a constant $c=c(n,p)$ such that

$$\int_Q |f(x)-f_Q|^p\,dx \leqslant c\|f\|_*^{p-1}\int_Q |f(x)-f_Q|\,dx.$$

for every cube $Q\subset\mathbb{R}^n$. This implies that $f\in L^p_{\mathrm{loc}}(\mathbb{R}^n)$ for every $p>1$ with the estimate above, compare with Theorem 3.28 below. This shows that, in a certain sense, BMO functions satisfy a reverse Hölder inequality.

3.4 Alternative proofs for the John-Nirenberg inequality

We discuss two other proofs for the John-Nirenberg lemma. First we present the original proof of the John-Nirenberg lemma in [13].

Proof (The original proof of the John-Nirenberg lemma). Assume that $f\in\mathrm{BMO}$ and let Q be a cube in $\mathbb{R}^n$. By considering $\frac{f}{\|f\|_*}$ with $\|f\|_*>0$ we may assume that $\|f\|_*=1$. We show that there exists constants c_1 and c_2, depending only on dimension n such that

$$|\{x\in Q : |f(x)-f_Q|>t\}| \leqslant c_1\exp(-c_2 t)|Q|, \quad t>0.$$

By replacing f with $f-f_Q$, we could also assume $f_Q=0$. However, we want to keep track on the oscillation in the argument. Denote $E_Q(t)=|\{x\in Q : |f(x)-f_Q|>t\}|$, $t>0$. We consider a pointwise infimum of functions F satisfying

$$|E_Q(t)| \leqslant F(t)\int_Q |f-f_Q|\,dx$$

for every $t > 0$ and every cube $Q \subset \mathbb{R}^n$. With a slight abuse of notation we denote this minimal function by F. By Chebyshev's inequality

$$|E_Q(t)| \leqslant \frac{1}{t} \int_Q |f - f_Q|\,dx$$

for every $t > 0$ and $Q \subset \mathbb{R}^n$ and thus $F(t) \leqslant \frac{1}{t}$ for every $t > 0$.

Let $s \geqslant \|f\|_* = 1$. Then

$$\fint_Q |f(y) - f_Q|\,dy \leqslant \|f\|_* \leqslant s.$$

Apply the Calderón-Zygmund decomposition for function $f - f_Q$ at level s in Q, see Theorem 1.4. We obtain pairwise disjoint dyadic subcubes Q_i, $i = 1, 2, \dots$, of Q such that

$$s < \fint_{Q_i} |f(y) - f_Q|\,dy \leqslant 2^n s, \quad i = 1, 2, \dots,$$

and

$$|f(x) - f_Q| \leqslant s \quad \text{for almost every} \quad x \in Q \setminus \bigcup_{i=1}^{\infty} Q_i.$$

For $t > 2^n s$ we have

$$\begin{aligned}
|E_Q(t)| &= \sum_{i=1}^{\infty} |\{x \in Q_i : |f(x) - f_Q| > t\}| \\
&\leqslant \sum_{i=1}^{\infty} |\{x \in Q_i : |f(x) - f_{Q_i}| + |f_{Q_i} - f_Q| > t\}| \\
&\leqslant \sum_{i=1}^{\infty} \left| \left\{ x \in Q_i : |f(x) - f_{Q_i}| + \fint_{Q_i} |f(y) - f_Q|\,dy > t \right\} \right| \\
&\leqslant \sum_{i=1}^{\infty} |\{x \in Q_i : |f(x) - f_{Q_i}| > t - 2^n s\}| \\
&= \sum_{i=1}^{\infty} |E_{Q_i}(t - 2^n s)|.
\end{aligned}$$

By the definition of F and the assumption $\|f\|_* = 1$, we have

$$\begin{aligned}
|E_Q(t)| &\leqslant \sum_{i=1}^{\infty} |E_{Q_i}(t - 2^n s)| \\
&\leqslant \sum_{i=1}^{\infty} F(t - 2^n s) \int_{Q_i} |f - f_{Q_i}|\,dx \\
&\leqslant F(t - 2^n s) \|f\|_* \sum_{i=1}^{\infty} |Q_i| \\
&\leqslant \frac{F(t - 2^n s)}{s} \sum_{i=1}^{\infty} \int_{Q_i} |f - f_Q|\,dx \\
&\leqslant \frac{F(t - 2^n s)}{s} \int_Q |f - f_Q|\,dx
\end{aligned}$$

By minimality of F we have

$$F(t) \leqslant \frac{F(t-2^n s)}{s} \tag{3.24}$$

for $t \geqslant 2^n s$.

A recursive application of (3.24) and the fact that F is a non-increasing function implies that F is exponentially decreasing. Let $s = e$ and $t \geqslant 2^n s = 2^n e$. Let k be an integer with $k2^n e \leqslant t < (k+1)2^n e$. By applying (3.24) $k-1$ times, we have

$$F(t) \leqslant F(k2^n e) \leqslant e^{-(k-1)} F(2^n e) \leqslant e^{-k+1} \frac{1}{2^n e} = 2^{-n} e^{-k} \leqslant 2^{-n} \exp\left(1 - \frac{t}{2^n e}\right)$$

for every $t \geqslant 2^n e$. This implies

$$|E_Q(t)| \leqslant F(t) \int_Q |f - f_Q| \, dx \leqslant 2^{-n} \exp\left(1 - \frac{t}{2^n e}\right) \int_Q |f - f_Q| \, dx$$

for every $t \geqslant 2^n e$. For $0 < t < 2^n e$ we have

$$|E_Q(t)| \leqslant |Q| \leqslant \exp\left(1 - \frac{t}{2^n e}\right) |Q|.$$

This proves the required inequality with $c_1 = e$ and $c_2 = (2^n e)^{-1}$. □

Next we discuss an alternative proof of the John-Nirenberg lemma by A.P. Calderón. He never published the proof himself but it is contained in a paper by U. Neri [17]. Calderón's method is very flexible and it applies also in the parabolic case.

Proof (Calderón's proof of the John-Nirenberg lemma). Assume that $f \in \mathrm{BMO}$ and let Q be a cube in $\mathbb{R}^n$. By considering $\frac{f}{\|f\|_*}$ with $\|f\|_* > 0$ we may assume that $\|f\|_* = 1$. By replacing f with $f - f_Q$, we may also assume $f_Q = 0$.

Let $t \geqslant \|f\|_* = 1$. Then

$$\fint_Q |f(y) - f_Q| \, dy = \fint_Q |f(y)| \, dy \leqslant \|f\|_* \leqslant t.$$

Apply the Calderón-Zygmund decomposition for f at level t in Q, see Theorem 1.4. We obtain pairwise disjoint dyadic subcubes Q_i, $i = 1, 2, \dots$, of Q such that

$$t < \fint_{Q_i} |f(y)| \, dy \leqslant 2^n t, \quad i = 1, 2, \dots,$$

and

$$|f(x)| \leqslant t \quad \text{for almost every} \quad x \in Q \setminus \bigcup_{i=1}^{\infty} Q_i.$$

It follows that

$$|\{x \in Q : |f(x)| > t\}| \leqslant \sum_{i=1}^{\infty} |Q_i|.$$

We consider $|E(t)| = \sum_{i=1}^{\infty} |Q_i|$ as a function of $t \geqslant 1$. Note that this is the measure of the union of the Calderón-Zygmund cubes at level t.

Let $t > s \geqslant 1$ and denote $\{Q_i\} = \{Q_i(t)\}$ and $\{Q_j\} = \{Q_j(s)\}$ the corresponding Calderón-Zygmund cubes at levels t and s respectively. Index i refers to a cube at level t and index j refers to a cube at level s. Observe that each cube Q_i is contained in a unique cube Q_j. In particular, the function $|E(t)|$ is non-increasing.

By the properties of the Calderón-Zygmund cubes

$$t\sum_{i=1}^{\infty}|Q_i| \leqslant \sum_{i=1}^{\infty}\int_{Q_i}|f(y)|\,dy.$$

For every $j = 1,2,\dots$ denote $I_j = \{i : Q_i \subset Q_j\}$ and rewrite the inequality above as

$$t\sum_{i=1}^{\infty}|Q_i| \leqslant \sum_{j=1}^{\infty}\sum_{i\in I_j}\int_{Q_i}|f(y)|\,dy.$$

Each Q_j was obtained by subdividing a parent dyadic cube Q'_j in the previous generation with $|f_{Q'_j}| \leqslant |f|_{Q'_j} \leqslant s$ and $|Q'_j| = 2^n|Q_j|$. Thus

$$\begin{aligned}
\sum_{i\in I_j}\int_{Q_i}|f(y)|\,dy &\leqslant \sum_{i\in I_j}\int_{Q_i}\left(|f(y)|+s-|f_{Q'_j}|\right)dy\\
&\leqslant \sum_{i\in I_j}\int_{Q_i}\left||f(y)|+s-|f_{Q'_j}|\right|dy\\
&\leqslant \sum_{i\in I_j}\int_{Q_i}\left||f(y)|-|f_{Q'_j}|\right|dy+\sum_{i\in I_j}\int_{Q_i}|s|\,dy\\
&\leqslant \sum_{i\in I_j}\int_{Q_i}|f(y)-f_{Q'_j}|\,dy+s\sum_{i\in I_j}|Q_i|\\
&\leqslant \int_{Q'_j}|f(y)-f_{Q'_j}|\,dy+s\sum_{i\in I_j}|Q_i|\\
&\leqslant \|f\|_*|Q'_j|+s\sum_{i\in I_j}|Q_i|\\
&= 2^n|Q_j|+s\sum_{i\in I_j}|Q_i|.
\end{aligned}$$

Here we also used the assumption $\|f\|_* = 1$. By summing over j we have

$$\begin{aligned}
t\sum_{i=1}^{\infty}|Q_i| &\leqslant \sum_{j=1}^{\infty}\sum_{i\in I_j}\int_{Q_i}|f(y)|\,dy\\
&\leqslant 2^n\sum_{j=1}^{\infty}|Q_j|+s\sum_{j=1}^{\infty}\sum_{i\in I_j}|Q_i|\\
&\leqslant 2^n\sum_{j=1}^{\infty}|Q_j|+s\sum_{j=1}^{\infty}|Q_j|.
\end{aligned}$$

This implies

$$|E(t)| \leqslant \frac{2^n}{t-s}|E(s)| \tag{3.25}$$

for $t > s \geqslant 1$.

Let $a = 2^{n+1}$ and apply (3.25) to have

$$|E(t+a)| \leqslant \frac{|E(t)|}{2}$$

for every $t \geqslant 1$. For $t \geqslant a$ we choose an integer k with $ka \leqslant t < (k+1)a$. Together with the fact that $|E(t)|$ is non-decreasing, this implies

$$|E(t)| \leqslant |E(ka)| \leqslant 2^{-(k-1)}|E(a)| \leqslant \exp\left(\left(2-\frac{t}{a}\right)\log 2\right)|Q| = 4\exp\left(-\frac{t\log 2}{a}\right)|Q|.$$

For $0 < t < a$, we have

$$|E(t)| \leqslant |Q| \leqslant 4e^{-\log 2}|Q| \leqslant 4\exp\left(-\frac{t\log 2}{a}\right)|Q|.$$

This proves the John-Nirenberg inequality with $c_1 = 4$ and $c_2 = \frac{\log 2}{2^{n+1}}$. □

Remark 3.26. Let $Q_0 \subset \mathbb{R}^n$ be a cube and assume that $f \in L^1(Q_0)$. Consider the dyadic maximal function $M_{d,Q_0}f$ defined by (1.29). Let $E_t = \{x \in Q_0 : M_{d,Q_0}f(x) > t\}$, $t > 0$. For every $t \geqslant |f|_{Q_0}$ the set E_t is the union of pairwise disjoint dyadic Calderón-Zygmund cubes Q_i, $i = 1,2,\dots$, given by Theorem 1.4. This implies $|E_t| = \sum_{i=1}^{\infty}|Q_i|$. Thus Calderón's proof of the John-Nirenberg lemma gives

$$|\{x \in Q_0 : M_{d,Q_0}(f - f_{Q_0})(x) > t\}| \leqslant c_1 \exp\left(-\frac{c_2 t}{\|f\|_*}\right)|Q|, \quad t > 0,$$

This is a stronger assertion than in the Jon-Nirenberg lemma, since by the Lebesgue differentiation theorem

$$|f(x) - f_{Q_0}| \leqslant M_{d,Q_0}(f - f_{Q_0})(x)$$

for almost every $x \in Q_0$.

3.5 Consequences of the John-Nirenberg inequality

Next we consider two consequences of the John-Nirenberg inequality. Assume that $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and let $1 \leqslant p < \infty$. Let

$$\|f\|_{*,p} = \sup_Q \left(\fint_Q |f(x) - f_Q|^p \, dx\right)^{\frac{1}{p}},$$

where the supremum is taken over all cubes Q in $\mathbb{R}^n$, and the corresponding function space

$$\mathrm{BMO}_p = \left\{f \in L^1_{\mathrm{loc}}(\mathbb{R}^n) : \|f\|_{*,p} < \infty\right\}.$$

T H E M O R A L : BMO_p is an L^p version of BMO. The standard BMO corresponds to the case $p = 1$.

Note that if $1 \leqslant p \leqslant q$, by Jensen's (or Hölder's) inequality

$$\fint_Q |f(x)-f_Q|\,dx \leqslant \left(\fint_Q |f(x)-f_Q|^p\,dx\right)^{\frac{1}{p}} \leqslant \left(\fint_Q |f(x)-f_Q|^q\,dx\right)^{\frac{1}{q}}$$

for every cube Q in $\mathbb{R}^n$. This implies

$$\|f\|_* \leqslant \|f\|_{*,p} \leqslant \|f\|_{*,q} \tag{3.27}$$

and thus $\mathrm{BMO}_q \subset \mathrm{BMO}_p \subset \mathrm{BMO}$ whenever $1 \leqslant p \leqslant q$.

The next result shows that these are same spaces.

Theorem 3.28. Assume that $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$. For every p with $1 < p < \infty$ there exists a constant $c = c(n,p)$ such that

$$\|f\|_* \leqslant \|f\|_{*,p} \leqslant c\|f\|_*.$$

T H E M O R A L : Norms $\|f\|_{*,p}$ and $\|f\|_*$ are equivalent and thus $\mathrm{BMO}_p = \mathrm{BMO}$ for every $1 \leqslant p < \infty$.

Proof. The first inequality in the claim follows from Jensen's inequality as in (3.27).

On the other hand, by the John-Nirenberg lemma, there exist constants c_1 and c_2, depending only on dimension n, such that

$$\begin{aligned}
\fint_Q |f(x)-f_Q|^p\,dx &= \frac{p}{|Q|}\int_0^\infty t^{p-1}|\{x \in Q : |f(x)-f_Q| > t\}|\,dt \\
&\leqslant pc_1 \int_0^\infty t^{p-1}\exp\left(-\frac{c_2 t}{\|f\|_*}\right)dt \\
&= pc_1\left(\frac{\|f\|_*}{c_2}\right)^{p-1}\int_0^\infty \left(\frac{c_2 t}{\|f\|_*}\right)^{p-1}\exp\left(-\frac{c_2 t}{\|f\|_*}\right)dt \\
&= pc_1\left(\frac{\|f\|_*}{c_2}\right)^{p}\int_0^\infty s^{p-1}e^{-s}\,ds \\
&= pc_1\left(\frac{\|f\|_*}{c_2}\right)^{p}\Gamma(p)
\end{aligned}$$

for every cube Q in $\mathbb{R}^n$. Here Γ is the gamma function. This implies

$$\|f\|_{*,p} \leqslant \frac{(pc_1\Gamma(p))^{\frac{1}{p}}}{c_2}\|f\|_*.$$

□

T H E M O R A L : The proof shows that, in a certain sense, BMO functions satisfy a reverse Hölder inequality.

Remark 3.29. In particular, it follows that $\mathrm{BMO} \subset L^p_{\mathrm{loc}}(\mathbb{R}^n)$ for every $1 < p < \infty$. To see this, we note that

$$\begin{aligned}\left(\fint_Q |f(x)|^p\,dx\right)^{\frac{1}{p}} &= \left(\fint_Q |f(x)-f_Q+f_Q|^p\,dx\right)^{\frac{1}{p}}\\ &\leqslant \left(\fint_Q |f(x)-f_Q|^p\,dx\right)^{\frac{1}{p}} + |f_Q|\\ &\leqslant \|f\|_{*,p} + |f|_Q \leqslant c\|f\|_* + |f|_Q < \infty\end{aligned}$$

for every cube Q in $\mathbb{R}^n$. Recall that by Hölder's inequality, $L^\infty_{\mathrm{loc}}(\mathbb{R}^n) \subset L^p_{\mathrm{loc}}(\mathbb{R}^n) \subset L^1_{\mathrm{loc}}(\mathbb{R}^n)$. In other words, every BMO function is locally integrable to arbitrarily large power p, but is not necessarily locally bounded as $f(x) = \log|x|$ shows.

We show that BMO functions are locally exponentially integrable. This could be done by using the Taylor series expansion of the exponential function, but we apply Cavalieri's principle instead.

Theorem 3.30. Assume that $f \in \mathrm{BMO}$. There exist a constant $c = c(n) > 0$ such that

$$\fint_Q \exp\left(\frac{|f(x)-f_Q|}{c\|f\|_*}\right) dx \leqslant c$$

for every cube Q in $\mathbb{R}^n$.

THE MORAL: BMO is a substitute for $L^\infty(\mathbb{R}^n)$ in the sense that every function in BMO is locally exponentially integrable. In particular, this implies that every function in BMO is locally integrable to any power.

Proof. Let $c_1 = c_1(n) > 0$ be the constant in the John-Nirenberg inequality such that

$$|\{x \in Q : |f(x)-f_Q| > t\}| \leqslant c_1 \exp\left(-\frac{t}{c_1\|f\|_*}\right)|Q|, \quad t > 0,$$

for every cube $Q \subset \mathbb{R}^n$. By denoting

$$g(x) = \frac{|f(x)-f_Q|}{2c_1\|f\|_*}$$

for every $x \in \mathbb{R}^n$, we have

$$|\{x \in Q : g(x) > t\}| \leqslant c_1 \exp(-2t)|Q|, \quad t > 0,$$

for every cube $Q \subset \mathbb{R}^n$.

Cavalieri's principle and the John-Nirenberg lemma imply

$$\begin{aligned}
\int_Q \exp(g(x))\,dx &= \int_0^\infty |\{x \in Q : \exp(g(x)) > t\}|\,dt \\
&= \int_0^1 |\{x \in Q : \exp(g(x)) > t\}|\,dt + \int_1^\infty |\{x \in Q : \exp(g(x)) > t\}|\,dt \\
&\leqslant \int_0^1 |Q|\,dt + \int_0^\infty e^t |\{x \in Q : g(x) > t\}|\,dt \\
&\leqslant c_1|Q| \int_0^\infty \exp(-t)\,dt + |Q| = (c_1+1)|Q|.
\end{aligned}$$

The claim holds with $c = \max\{c_1+1, 2c_1\}$. □

Remark 3.31. We have already several times used the fact that

$$\|f\|^p_{L^p(A)} = \int_A |f(x)|^p\,dx = p\int_0^\infty t^{p-1}|\{x \in A : |f(x)| > t\}|\,dt, \quad 1 \leqslant p < \infty.$$

This can be easily generalized by replacing t^p with any other increasing and differentiable function $\varphi : [0,\infty) \to [0,\infty)$, when we obtain

$$\int_A \varphi(|f|)\,dx = \int_0^\infty \varphi'(t)|\{x \in A : |f(x)| > t\}|\,dt + \varphi(0)|A|. \tag{3.32}$$

In particular, we can choose $\varphi(t) = e^t$ (exercise). Instead of a change of variables we could have used this version of Cavalieri's principle in the previous proof.

Remark 3.33. It can be shown that the following claims are equivalent for $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ (exercise):

(1) $f \in \text{BMO}$,

(2) there are positive constants c_1 and c_2, independent of cube Q, such that

$$\fint_Q e^{c_1|f(x)-f_Q|}\,dx \leqslant c_2$$

for every cube Q in $\mathbb{R}^n$,

(3) there are positive constants c_1 and c_2, independent of cube Q, such that

$$\fint_Q e^{c_1 f(x)}\,dx \fint_Q e^{-c_1 f(x)}\,dx \leqslant c_2$$

for every cube Q in $\mathbb{R}^n$.

3.6 The sharp maximal function

We define the sharp maximal function $f^\# : \mathbb{R}^n \to [0,\infty]$ of $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ by

$$f^\#(x) = \sup_Q \fint_Q |f(y) - f_Q|\,dy, \tag{3.34}$$

where the supremum is taken over all cubes Q in $\mathbb{R}^n$ containing x.

THE MORAL: The difference compared to the Hardy-Littlewood maximal function is that instead of the integral averages we maximize the mean oscillation over cubes.

The connection of the sharp maximal function to BMO is that

$$f \in \mathrm{BMO} \iff \|f\|_* = \|f^\#\|_\infty < \infty.$$

Since

$$\fint_Q |f(y) - f_Q| \, dy \leqslant \frac{1}{|Q|} \int_Q |f(y)| \, dy + |f_Q| \leqslant 2 \fint_Q |f(y)| \, dy$$

for every cube $Q \subset \mathbb{R}^n$, by the definition of the Hardy-Littlewood maximal function

$$f^\#(x) \leqslant 2Mf(x) \quad \text{for every} \quad x \in \mathbb{R}^n$$

and, by the maximal function theorem, for every p with $1 < p \leqslant \infty$, there exists constant c, depending only on n and p such that

$$\|f^\#\|_p \leqslant 2\|Mf\|_p \leqslant c\|f\|_p.$$

THE MORAL: The sharp maximal operator is a bounded operator on $L^p(\mathbb{R}^n)$ whenever $1 < p \leqslant \infty$.

There is no pointwise inequality to the reverse direction $Mf(x) \leqslant cf^\#(x)$, as can be seen by considering constant functions. However, it turns out that the L^p norms of the sharp maximal function and the Hardy-Littlewood maximal functions are comparable under certain assumptions. In order to prove this, we consider certain useful inequalities for the distribution functions. This kind of inequalities, with t replaced by λ, are sometimes called as good lambda inequalities.

Lemma 3.35. Assume that $f, g : \mathbb{R}^n \to [0, \infty]$ are measurable functions with the property that there exist nonnegative constants a, b and c for which

$$|\{x \in \mathbb{R}^n : f(x) > t, \, g(x) \leqslant ct\}| \leqslant a|\{x \in \mathbb{R}^n : f(x) > bt\}|$$

for every $t > 0$. Moreover, we assume that $f \in L^p(\mathbb{R}^n)$ with $1 \leqslant p < \infty$. Then there exists constant $A = A(a, b, c, p)$, such that $\|f\|_p \leqslant A\|g\|_p$, whenever $a < b^p$.

THE MORAL: L^p-bounds can be proved using distribution function inequalities instead of pointwise inequalities.

Proof. Since

$$\{x \in \mathbb{R}^n : f(x) > t\} = \{x \in \mathbb{R}^n : f(x) > t, \, g(x) \leqslant ct\} \cup \{x \in \mathbb{R}^n : f(x) > t, \, g(x) > ct\},$$

the assumption implies

$$|\{x \in \mathbb{R}^n : f(x) > t\}| \leqslant a|\{x \in \mathbb{R}^n : f(x) > bt\}| + |\{x \in \mathbb{R}^n : g(x) > ct\}|$$

for every $t > 0$. Thus

$$\begin{aligned}\int_{\mathbb{R}^n} f(x)^p\,dx &= p\int_0^\infty t^{p-1}|\{x\in\mathbb{R}^n : f(x) > t\}|\,dt\\ &\leqslant pa\int_0^\infty t^{p-1}|\{x\in\mathbb{R}^n : f(x) > bt\}|\,dt + p\int_0^\infty t^{p-1}|\{x\in\mathbb{R}^n : g(x) > ct\}|\,dt\\ &= \frac{a}{b^p}\int_{\mathbb{R}^n} f(x)^p\,dx + \frac{1}{c^p}\int_{\mathbb{R}^n} g(x)^p\,dx. \qquad \square\end{aligned}$$

Since $\|f\|_p < \infty$, we may absorb term $\frac{a}{b^p}\int_{\mathbb{R}^n} f(x)^p\,dx < \infty$ into the left-hand side. This implies

$$\Big(1-\frac{a}{b^p}\Big)\int_{\mathbb{R}^n} f(x)^p\,dx \leqslant \frac{1}{c^p}\int_{\mathbb{R}^n} g(x)^p\,dx$$

and, since $a < b^p$, we arrive at

$$\int_{\mathbb{R}^n} f(x)^p\,dx \leqslant \frac{b^p}{c^p(b^p-a)}\int_{\mathbb{R}^n} g(x)^p\,dx.$$

Remark 3.36. If $b \leqslant 1$ in Lemma 3.35, we may replace assumption $\|f\|_p < \infty$ with $\|f\|_{p_0} < \infty$ for some p_0 with $0 < p_0 < p$. To prove this, let

$$I_k = p\int_0^k t^{p-1}|\{x\in\mathbb{R}^n : f(x) > t\}|\,dt, \quad k = 1,2,\dots.$$

Chebyshev's inequality gives

$$|\{x\in\mathbb{R}^n : f(x) > t\}| \leqslant \frac{1}{t^{p_0}}\int_{\mathbb{R}^n} f(x)^{p_0}\,dx < \infty$$

for every $t > 0$, and since $p_1 < p$, we have

$$I_k \leqslant p\|f\|_{p_0}^{p_0}\int_0^k t^{p-p_0-1}\,dt < \infty, \quad k = 1,2,\dots.$$

As in the proof of Lemma 3.35, we obtain

$$I_k \leqslant \frac{a}{b^p} I_{Nb} + \frac{1}{c^p}\|g\|_p^p, \quad k = 1,2,\dots.$$

If $b \leqslant 1$, then $I_{kb} \leqslant I_k$ and

$$I_k \leqslant \frac{a}{b^p} I_k + \frac{1}{c^p}\|g\|_p^p.$$

Since $a < b^p$, we have

$$\Big(1-\frac{a}{b^p}\Big) I_k \leqslant \frac{1}{c^p}\|g\|_p^p$$

and, equivalently, $I_k \leqslant A\|g\|_p^p$. The claim follows by letting $k \to \infty$.

Recall that the dyadic maximal function defined in (1.18) is

$$M_d f(x) = \sup \fint_Q |f(y)|\,dy,$$

where the supremum is taken over all dyadic cubes Q containing x. Analogously, we define the dyadic sharp maximal function as

$$f_d^\#(x) = \sup \fint_Q |f(y) - f_Q|\,dy,$$

where the supremum is taken over all dyadic cubes Q containing x. We observe that $f_d^\#(x) \leqslant 2_d Mf(x)$ for every $x \in \mathbb{R}^n$ and, as before, that there is no pointwise inequality in the reverse direction. However, we have the following good lambda inequality for the distribution functions.

Lemma 3.37. Assume that $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, $c > 0$ and $0 < b < 1$. Then

$$|\{x \in \mathbb{R}^n : M_d f(x) > t, f_d^\#(x) \leqslant ct\}| \leqslant a|\{x \in \mathbb{R}^n : M_d f(x) > bt\}|$$

for every $t > 0$ with $a = \frac{2^n c}{1-b}$.

Proof. We may assume that $|\{x \in \mathbb{R}^n : M_d f(x) > bt\}| < \infty$, since otherwise there is nothing to prove. Let Q_i, $i = 1, 2, \dots$, be the collection of Calderón-Zygmund cubes at level $bt > 0$ as in Lemma 1.22. See also Remark 1.24 (4). In particular, we have

$$\{x \in \mathbb{R}^n : M_d f(x) > bt\} = \bigcup_{i=1}^\infty Q_i.$$

We show that

$$|\{x \in Q_i : M_d f(x) > t, f_d^\#(x) \leqslant ct\}| \leqslant a|Q_i|, \quad i = 1, 2, \dots.$$

Let $\widetilde{Q}_i$ be the unique parent dyadic cube of Q_i, $i = 1, 2, \dots$. By maximality of the Calderón-Zygmund cubes, as in Theorem 1.12 and Lemma 1.22, we have

$$\fint_{\widetilde{Q}_i} |f(y)|\,dy \leqslant bt. \tag{3.38}$$

We claim that

$$M_d(f\chi_{Q_i})(x) > t \quad \text{and} \quad M_d((f - f_{\widetilde{Q}_i})\chi_{Q_i})(x) > (1-b)t$$

for every $x \in Q_i$ with $M_d f(x) > t$. To prove the first inequality, we note that

$$\fint_{Q'} |f(y)|\,dy \leqslant bt \leqslant t$$

for all dyadic cubes Q' containing Q_i. Here we used maximality of the Calderón-Zygmund cubes again. Thus all dyadic cubes Q containing x with

$$\fint_Q |f(y)|\,dy > t$$

are subcubes of Q_i and consequently

$$M_d(f\chi_{Q_i})(x) = M_d f(x) > t. \tag{3.39}$$

To prove the second inequality, by sublinearity of the dyadic maximal operator, we have

$$M_d(f\chi_{Q_i})(x) \leqslant M_d((f - f_{\widetilde{Q}_i})\chi_{Q_i})(x) + M_d(f_{\widetilde{Q}_i}\chi_{Q_i})(x)$$

for every $x \in Q_i$. We also note that

$$M_d(f_{\widetilde{Q}_i}\chi_{Q_i})(x) = |f_{\widetilde{Q}_i}| \quad \text{for every} \quad x \in Q_i.$$

Thus (3.39) and (3.38) imply

$$\begin{aligned} M_d((f - f_{\widetilde{Q}_i})\chi_{Q_i})(x) &\geqslant M_d(f\chi_{Q_i})(x) - M_d(f_{\widetilde{Q}_i}\chi_{Q_i})(x) \\ &> t - |f_{\widetilde{Q}_i}| \geqslant t - |f|_{\widetilde{Q}_i} \geqslant t - bt \end{aligned}$$

for every $x \in Q_i$ with $M_d f(x) > t$.

By the weak type estimate for the dyadic maximal function, see Remark 1.24 (1), we obtain

$$\begin{aligned} |\{x \in Q_i : M_d f(x) > t, f_d^\#(x) \leqslant ct\}| &\leqslant |\{x \in Q_i : M_d f(x) > t\}| \\ &\leqslant |\{x \in Q_i : M_d((f - f_{\widetilde{Q}_i})\chi_{Q_i})(x) > (1-b)t\}| \\ &\leqslant |\{x \in \mathbb{R}^n : M_d((f - f_{\widetilde{Q}_i})\chi_{Q_i})(x) > (1-b)t\}| \\ &\leqslant \frac{1}{(1-b)t} \int_{\mathbb{R}^n} |f(y) - f_{\widetilde{Q}_i}|\chi_{Q_i}(y)\,dy \\ &\leqslant \frac{1}{(1-b)t} \int_{\widetilde{Q}_i} |f(y) - f_{\widetilde{Q}_i}|\,dy \\ &\leqslant \frac{|\widetilde{Q}_i|}{(1-b)t} \inf_{x \in Q_i} f_d^\#(x) \\ &\leqslant \frac{|\widetilde{Q}_i|}{(1-b)t} ct = \frac{2^n c}{1-b}|Q_i| = a|Q_i|, \end{aligned}$$

if there exists $x \in Q_i$ such that $f_d^\#(x) \leqslant ct$. Otherwise the set is empty and there is nothing to prove. Here we also used the fact that, by the definition of the dyadic sharp maximal function, we have

$$f_d^\#(x) \geqslant \frac{1}{|\widetilde{Q}_i|} \int_{\widetilde{Q}_i} |f(y) - f_{\widetilde{Q}_i}|\,dy \quad \text{for every} \quad x \in Q_i$$

and thus

$$\inf_{x \in Q_i} f_d^\#(x) \geqslant \frac{1}{|\widetilde{Q}_i|} \int_{\widetilde{Q}_i} |f(y) - f_{\widetilde{Q}_i}|\,dy.$$

By summing over the Calderón-Zygmund cubes we arrive at

$$\begin{aligned}
|\{x \in \mathbb{R}^n : M_d f(x) > t, f_d^\#(x) \leqslant ct\}| &\leqslant |\{x \in \mathbb{R}^n : M_d f(x) > bt, f_d^\#(x) \leqslant ct\}| \\
&= \left| \left\{ x \in \bigcup_{i=1}^\infty Q_i : M_d f(x) > bt, f_d^\#(x) \leqslant ct \right\} \right| \\
&= \sum_{i=1}^\infty |\{x \in Q_i : M_d f(x) > bt, f_d^\#(x) \leqslant ct\}| \\
&\leqslant a \sum_{i=1}^\infty |Q_i| = a \left| \bigcup_{i=1}^\infty Q_i \right| \\
&= a |\{x \in \mathbb{R}^n : M_d f(x) > bt\}|.
\end{aligned}$$

This completes the proof. □

Now we are ready to prove that, under certain assumptions, the sharp maximal function and the Hardy-Littlewood maximal function are comparable on L^p level even though they are not comparable pointwise.

Theorem 3.40 (Fefferman-Stein (1972)). Let $1 < p < \infty$. Assume that $f \in L^{p_0}(\mathbb{R}^n)$ for some $1 \leqslant p_0 \leqslant p$ and $f^\# \in L^p(\mathbb{R}^n)$, then $f \in L^p(\mathbb{R}^n)$ and

$$\|Mf\|_p \leqslant c\|f^\#\|_p,$$

where $c = c(n,p)$.

THE MORAL: Under the assumptions in the Fefferman-Stein theorem,

$$\|f\|_p \leqslant \|Mf\|_p \leqslant c\|f^\#\|_p \leqslant c\|Mf\|_p \leqslant c\|f\|_p,$$

that is, the L^p norms of f, Mf and $f^\#$ are comparable. Here we used the facts that $f(x) \leqslant Mf(x)$ for almost every $x \in \mathbb{R}^n$, $f^\#(x) \leqslant 2Mf(x)$ for every $x \in \mathbb{R}^n$ and $\|Mf\|_p \leqslant c\|f\|_p$ with $1 < p < \infty$. The last fact is the Hardy-Littlewood-Wiener maximal function theorem, see Example 2.6.

Remark 3.41. The assumption $f \in L^{p_0}(\mathbb{R}^n)$ for some $1 \leqslant p_0 \leqslant p$ cannot be omitted. For example, if f is a nonzero constant function, then $f^\# = 0$ and the claim does not hold.

Proof. Let $c = 2^{-n-p-2}$, $b = \frac{1}{2}$ and $a = \frac{2^n c}{1-b} = 2^{n+1}c$. By Lemma 3.37,

$$|\{x \in \mathbb{R}^n : M_d f(x) > t, f_d^\#(x) \leqslant ct\}| \leqslant a|\{x \in \mathbb{R}^n : M_d f(x) > bt\}|$$

for every $t > 0$. Since $a = 2^{n+1}c = 2^{-p-1} < 2^{-p} = b^p$, by Lemma 3.35 and Remark 3.36 we obtain

$$\|M_d f\|_p \leqslant c\|f_d^\#\|_p,$$

with $c = c(n,p)$. This proves claim for the corresponding dyadic maximal functions. It is clear that $f_d^\#(x) \leqslant f^\#(x)$ for every $x \in \mathbb{R}^n$ and thus

$$\|M_d f\|_p \leqslant \|f_d^\#\|_p \leqslant c\|f^\#\|_p.$$

On the other hand, by Example 2.6 (2) we have

$$\|Mf\|_p \leqslant 12^{\frac{n}{p}} \|M_d f\|_p \leqslant 12^{\frac{n}{p}} c\|f^\#\|_p.$$

This completes the proof. □

Remark 3.42. Alternatively, we could conclude in the proof that

$$\|Mf\|_p \leqslant c\|f\|_p \leqslant c\|M_d f\|_p \leqslant c\|f_d^\#\|_p \leqslant c\|f^\#\|_p$$

for some constant $c = c(n,p)$.

3.7 BMO and interpolation

Many interesting linear operators in harmonic analysis, as the Hilbert and Riesz transforms, are bounded in $L^p(\mathbb{R}^n)$ with $1 < p < \infty$, but they fail to be bounded in the limiting cases $p = 1$ and $p = \infty$. The substitute in the case $p = 1$ is that they map $L^1(\mathbb{R}^n)$ to weak $L^1(\mathbb{R}^n)$. It turns out that the substitute in the case $p = \infty$ is that they map $L^\infty(\mathbb{R}^n)$ to BMO. In this sense, the space BMO plays a similar role for $L^\infty(\mathbb{R}^n)$ as weak $L^1(\mathbb{R}^n)$ plays for $L^1(\mathbb{R}^n)$.

BMO is a natural substitute for $L^\infty(\mathbb{R}^n)$ also in the following sense. If $f \in L^{p_0}(\mathbb{R}^n)$ for some $1 \leqslant p_0 \leqslant p$, by Theorem 3.40 we have

$$f \in L^p(\mathbb{R}^n) \Longleftrightarrow Mf \in L^p(\mathbb{R}^n) \Longleftrightarrow f^\# \in L^p(\mathbb{R}^n), \quad 1 < p < \infty.$$

The situation is different in the case $p = \infty$. Then

$$f \in L^\infty(\mathbb{R}^n) \Longleftrightarrow Mf \in L^\infty(\mathbb{R}^n), \quad \text{but} \quad f^\# \in L^\infty(\mathbb{R}^n) \Longleftrightarrow f \in \mathrm{BMO}.$$

The following result is a BMO version of the Marcinkiewicz interpolation theorem for linear operators, see Theorem 2.4.

Theorem 3.43 (The Stampacchia interpolation theorem (1965)). Let $1 < p_1 < \infty$ and assume that T is a linear operator from $L^{p_1}(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)$ to Lebesgue measurable functions on $\mathbb{R}^n$, which is of strong type (p_1, p_1) and bounded from $L^\infty(\mathbb{R}^n)$ to BMO. Then there exists a constant c such that $\|Tf\|_p \leqslant c\|f\|_p$ for every $f \in L^{p_1}(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$, that is, T satisfies the strong type (p,p) estimate for functions $f \in L^{p_1}(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$.

THE MORAL: BMO can be used as a substitute for $L^\infty(\mathbb{R}^n)$ also in interpolation.

Proof. Denote $Sf = (Tf)^\#$ and let $f, g \in L^{p_1}(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)$. Since T is linear, we have

$$S(f+g) = (T(f+g))^\# = (Tf + Tg)^\# \leqslant (Tf)^\# + (Tg)^\#,$$

which implies that S is a sublinear operator. By the assumption

$$\|Sf\|_\infty = \|(Tf)^\#\|_\infty = \|Tf\|_* \leqslant c\|f\|_\infty$$

for every $f \in L^\infty(\mathbb{R}^n)$. Thus S is of strong type (∞,∞).

On the other hand, by the Hardy-Littlewood-Wiener maximal function theorem, see Example 2.6 (2), we have

$$\|Sf\|_{p_1} = \|(Tf)^\#\|_{p_1} \leqslant 2\|M(Tf)\|_{p_1} \leqslant c\|Tf\|_{p_1} \leqslant c\|f\|_{p_1}$$

for every $f \in L^{p_1}(\mathbb{R}^n)$. Thus S is of strong type (p_1,p_1) with $p_1 > 1$. The Marcinkiewicz interpolation theorem, see Theorem 2.4, implies

$$\|Sf\|_p \leqslant c\|f\|_p$$

for every $f \in L^p(\mathbb{R}^n)$ whenever $1 < p_1 \leqslant p < \infty$.

Let $f \in L^{p_1}(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$. Since T is of strong type (p_1,p_1), we conclude that $Tf \in L^{p_1}(\mathbb{R}^n)$. On the other hand,

$$\|(Tf)^\#\|_p = \|Sf\|_p \leqslant c\|f\|_p < \infty$$

for every $f \in L^p(\mathbb{R}^n)$, which implies $(Tf)^\# \in L^p(\mathbb{R}^n)$. By Theorem 3.40, we obtain

$$\|Tf\|_p \leqslant \|M(Tf)\|_p \leqslant c\|(Tf)^\#\|_p \leqslant c\|f\|_p \tag{3.44}$$

for every $f \in L^{p_1}(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$. □

Remark 3.45. The claim of the Stampacchia interpolation theorem can be extended to all $f \in L^p(\mathbb{R}^n)$ by the following argument. Let $f \in L^p(\mathbb{R}^n)$. Since $L^{p_1}(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$, there exists a sequence (f_i) of functions $f_i \in L^{p_1}(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$, $i = 1,2,\dots$, such that $f_i \to f$ in $L^p(\mathbb{R}^n)$ as $i \to \infty$. For example, we may consider $f_i = f\chi_{\{|f|>\frac{1}{i}\}}$, $i = 1,2,\dots$. We could also use the fact that compactly supported continuous functions are dense in $L^p(\mathbb{R}^n)$ to obtain the approximation. By (3.44), we see that

$$\|Tf_i - Tf_j\|_p = \|T(f_i - f_j)\|_p \leqslant c\|f_i - f_j\|_p$$

which implies that (Tf_i) is a Cauchy sequence in $L^p(\mathbb{R}^n)$. Since $L^p(\mathbb{R}^n)$ is a complete space, there exists $h \in L^p(\mathbb{R}^n)$ such that $Tf_i \to h$ in $L^p(\mathbb{R}^n)$ as $i \to \infty$.

We claim that the function h is independent of the approximating sequence. To see this let (f_i) and (g_i) be two sequences of functions $f_i, g_i \in L^{p_1}(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$, $i = 1,2,\dots$, such that $f_i \to f$ in $L^p(\mathbb{R}^n)$ and $g_i \to f$ in $L^p(\mathbb{R}^n)$ as $i \to \infty$. As above, we conclude that (Tf_i) and (Tg_i) are Cauchy sequences in $L^p(\mathbb{R}^n)$ and there exist $h_1 \in L^p(\mathbb{R}^n)$ and $h_2 \in L^p(\mathbb{R}^n)$ such that $Tf_i \to h_1$ and $Tg_i \to h_2$ in $L^p(\mathbb{R}^n)$ as $i \to \infty$.

By (3.44), we see that

$$\begin{aligned}
\|h_1 - h_2\|_p &= \|Tf_i - Tg_i - (Tf_i - h_1) + (Tg_i - h_2)\|_p \\
&\leqslant \|Tf_i - Tg_i\|_p + \|Tf_i - h_1\|_p + \|Tg_i - h_2\|_p \\
&= \|T(f_i - g_i)\|_p + \|Tf_i - h_1\|_p + \|Tg_i - h_2\|_p \\
&\leqslant c\|f_i - g_i\|_p + \|Tf_i - h_1\|_p + \|Tg_i - h_2\|_p \xrightarrow{i\to\infty} 0.
\end{aligned}$$

This implies that $\|h_1 - h_2\|_p = 0$ and thus $h_1 = h_2$ almost everywhere.

We may extend the operator T from $L^{p_1}(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$ to $L^p(\mathbb{R}^n)$ by setting $Tf = h$. We claim that this operator is bounded on $L^p(\mathbb{R}^n)$. By Minkowski's inequality

$$|\|Tf\|_p - \|Tf_i\|_p| \leqslant \|Tf - Tf_i\|_p = \|T(f - f_i)\|_p \leqslant c\|f - f_i\|_p \xrightarrow{i\to\infty} 0$$

Since $f_i \to f$ in $L^p(\mathbb{R}^n)$ as $i \to \infty$, by using (3.44) once more, we conclude that

$$\|Tf\|_p = \lim_{i\to\infty} \|Tf_i\|_p \leqslant c \lim_{i\to\infty} \|f_i\|_p = c\|f\|_p.$$

This shows that the extended operator T is of strong type (p,p) whenever $1 < p_1 \leqslant p < \infty$. Observe that T is the unique bounded linear operator on $L^p(\mathbb{R}^n)$ that extends the corresponding operator on $L^{p_1}(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$.

THE MORAL: A bounded linear operator on a dense subspace of L^p can be uniquely extended to a bounded linear operator on L^p.

4
Muckenhoupt weights

In this chapter we study the theory of Muckenhoupt's A_p weights and weighted norm inequalities. Some of the results can be used to give characterizations of BMO functions. The main goal is to show that the Hardy-Littlewood maximal operator is of weighted strong type (p,p) with $1 < p < \infty$ if and only if the weight satisfies Muckenhoupt's A_p condition. Weighted norm inequalities arise in Fourier analysis, but these techniques play an important role also in harmonic analysis and partial differential equations.

4.1 The A_p condition

Any nonnegative locally integrable function w on $\mathbb{R}^n$ is called a weight, that is, $w \in L^1_{\text{loc}}(\mathbb{R}^n)$ and $w \geqslant 0$ almost everywhere in $\mathbb{R}^n$. There is a natural measure associated with a weight, since every weight w gives rise to the measure

$$\mu(E) = \int_E w(x)\,dx,$$

where E is a Lebesgue measurable subset of $\mathbb{R}^n$. Since $w \in L^1_{\text{loc}}(\mathbb{R}^n)$, we have $\mu(K) < \infty$ for every compact set $K \subset \mathbb{R}^n$. Note that measure μ associated with weight w is absolutely continuous with respect to the Lebesgue measure, that is, it satisfies the property $|E| = 0$ implies $\mu(E) = 0$. Conversely, we observe that every Radon measure, which is absolutely continuous with respect to the Lebesgue measure, is given by a weight. Let μ be a Radon measure on $\mathbb{R}^n$, which is absolutely continuous with respect to the Lebesgue measure. Recall that a Radon measure is a Borel regular outer measure with the property that the measure of every compact set is finite. By the Radon-Nikodym theorem, there exists a function

$w \in L^1_{\text{loc}}(\mathbb{R}^n)$, $w \geqslant 0$, such that

$$\mu(E) = \int_E w(x)\,dx$$

for every Lebesgue measurable set E in $\mathbb{R}^n$.

THE MORAL: Every Radon measure which is absolutely continuous with respect to the Lebesgue measure is given by a weight.

By measure and integration theory, the set function μ is a measure on Lebesgue measurable sets and

$$\int_E f(x)\,d\mu(x) = \int_E f(x)w(x)\,dx.$$

For $w = 1$ we have the standard Lebesgue measure and $L^p(\mathbb{R}^n)$. The weighted space Lebesgue space $L^p(\mathbb{R}^n; w)$, with $1 \leqslant p < \infty$, is the space of Lebesgue measurable functions $f : \mathbb{R}^n \to [-\infty, \infty]$ for which

$$\|f\|_{L^p(\mathbb{R}^n;w)} = \left(\int_{R^n} |f(x)|^p w(x)\,dx\right)^{\frac{1}{p}} < \infty.$$

THE MORAL: $L^p(\mathbb{R}^n; w)$ is a space where the standard Lebesgue measure is replaced with a measure given by a weight. The weight function describes the nonhomogeneous mass distribution on $\mathbb{R}^n$.

We study weighted norm inequalities for the Hardy-Littlewood maximal operator

$$Mf(x) = \sup \fint_Q |f(y)|\,dy,$$

where the supremum is taken over all cubes Q in $\mathbb{R}^n$ containing x. We have already seen that the Hardy-Littlewood maximal operator is of weak type $(1,1)$ and strong type (p,p) with $1 < p \leqslant \infty$, see Example 2.6. Next we consider the similar estimates in weighted spaces. We use a relatively standard notation

$$w(E) = \mu(E) = \int_E w(x)\,dx,$$

where E is a Lebesgue measurable subset of $\mathbb{R}^n$.

We discuss the following questions:

(1) For which weights w we have the weighted strong type (p,p) estimate

$$\int_{\mathbb{R}^n} (Mf(x))^p w(x)\,dx \leqslant c \int_{\mathbb{R}^n} |f(x)|^p w(x)\,dx, \tag{4.1}$$

with $1 < p < \infty$, for every $f \in L^1_{\text{loc}}(\mathbb{R}^n)$?

(2) For which weights w we have the corresponding weak type (p,p) estimate

$$\begin{aligned} w(\{x \in \mathbb{R}^n : Mf(x) > t\}) &= \int_{\{x\in\mathbb{R}^n : Mf(x)>t\}} w(x)\,dx \\ &\leqslant \frac{c}{t^p} \int_{\mathbb{R}^n} |f(x)|^p w(x)\,dx, \quad t > 0, \end{aligned} \tag{4.2}$$

with $1 \leqslant p < \infty$, for every $f \in L^1_{\text{loc}}(\mathbb{R}^n)$?

THE MORAL: These estimates hold with $w = 1$, but do they hold for some other weights?

We note that the strong type estimate implies the weak type estimate, that is, (4.1) implies (4.2). This follows from Chebyshev's inequality, since

$$\begin{aligned} w(\{x \in \mathbb{R}^n : Mf(x) > t\}) &\leqslant \frac{1}{t^p} \int_{\mathbb{R}^n} (Mf(x))^p w(x)\,dx \\ &\leqslant \frac{c}{t^p} \int_{\mathbb{R}^n} |f(x)|^p w(x)\,dx, \quad t > 0. \end{aligned}$$

THE MORAL: The strong type estimate implies the weak type estimate also in the weighted case.

We begin with a weighted weak type estimate by Fefferman and Stein. The argument is similar to the proof of the standard weak type estimate for the maximal function.

Theorem 4.3. Let w be a weight and $f \in L^1_{\text{loc}}(\mathbb{R}^n)$.

(1) Let $1 \leqslant p < \infty$. There exists a constant $c = c(n)$ such that

$$w(\{x \in \mathbb{R}^n : Mf(x) > t\}) \leqslant \frac{c}{t^p} \int_{\mathbb{R}^n} |f(x)|^p Mw(x)\,dx, \quad t > 0.$$

(2) Let $1 < p < \infty$. There exists a constant $c = c(n,p)$ such that

$$\int_{\mathbb{R}^n} (Mf(x))^p w(x)\,dx \leqslant c \int_{\mathbb{R}^n} |f(x)|^p Mw(x)\,dx.$$

THE MORAL: The weights appearing on the both sides are different. The weight on the left-hand side is w and on the right-hand side Mw. For weights satisfying the pointwise inequality $Mw(x) \leqslant cw(x)$ for almost every $x \in \mathbb{R}^n$, we obtain (4.1) and (4.2). This condition, called the Muckenhoupt A_1, condition will be discussed later.

Proof. Let $E_t = \{x \in \mathbb{R}^n : Mf(x) > t\}$. For every $x \in E_t$, there exists a cube Q_x containing x such that

$$\fint_{Q_x} |f(y)|\,dy > t.$$

Let K be a compact subset of E_t. Since K is compact, it can be covered by a finite number of cubes as above. By the Vitali covering theorem, we obtain pairwise disjoint cubes $Q(x_i,l_i)$, $i=1,2,\dots,N$, such that $K\subset\bigcup_{i=1}^N Q(x_i,5l_i)$. This implies

$$\begin{aligned}
w(K)=\int_K w(z)\,dz &\leqslant \sum_{i=1}^N \int_{Q(x_i,5l_i)} w(z)\,dz\\
&\leqslant \sum_{i=1}^N \int_{Q(x_i,5l_i)} w(z)\,dz \frac{1}{t} \fint_{Q(x_i,l_i)} |f(y)|\,dy\\
&\leqslant \frac{5^n}{t}\sum_{i=1}^N \fint_{Q(x_i,5l_i)} w(z)\,dz \int_{Q(x_i,l_i)} |f(y)|\,dy\\
&= \frac{5^n}{t}\sum_{i=1}^N \int_{Q(x_i,l_i)} \left(|f(y)| \fint_{Q(x_i,5l_i)} w(z)\,dy\right)dy\\
&\leqslant \frac{5^n}{t}\sum_{i=1}^N \int_{Q(x_i,l_i)} |f(y)|Mw(y)\,dy,
\end{aligned}$$

since

$$Mw(x)\geqslant \fint_{Q(x_i,5l_i)} w(z)\,dz$$

for every $x\in Q(x_i,l_i)$. Since $Q(x_i,l_i)\subset E_t$, $i=1,2,\dots,N$, are pairwise disjoint, we conclude that

$$w(K)\leqslant \frac{5^n}{t}\int_{E_t} |f(y)|Mw(y)\,dy$$

for every compact subset K of E_t. The claim follows from this, since

$$\begin{aligned}
w(\{x\in\mathbb{R}^n : Mf(x)>t\}) &= \sup\{w(K) : K\subset E_t, K \text{ compact}\}\\
&\leqslant \frac{5^n}{t}\int_{E_t} |f(y)|Mw(y)\,dy.
\end{aligned}$$

A similar proof gives the weak type (p,p) estimate, since by Hölder's inequality

$$t<\fint_{Q_x} |f(y)|\,dy \leqslant \left(\fint_{Q_x} |f(y)|^p\,dy\right)^{\frac{1}{p}}.$$

The corresponding strong type estimate can be proved as in the proof of the standard Hardy-Littlewood-Wiener theorem (exercise). □

We begin with deriving necessary conditions for (4.1) and (4.2). Since the strong type estimate implies the weak type, a necessary condition for the weak type estimate is also a necessary condition for the strong type estimate. Thus we assume that (4.2) holds.

Assume that f is a Lebesgue measurable function on $\mathbb{R}^n$, let Q be a cube in $\mathbb{R}^n$ with $|f|_Q>0$ and let $0<t<|f|_Q$. If $|f|_Q=0$ there is nothing to prove. Since

$$M(f\chi_Q)(x)\geqslant \fint_Q |f(y)|\,dy \quad \text{for every} \quad x\in Q,$$

we have $Q \subset \{x \in \mathbb{R}^n : M(f\chi_Q)(x) > t\}$, and thus (4.2) implies

$$\begin{aligned} w(Q) &\leqslant w(\{x \in \mathbb{R}^n : M(f\chi_Q)(x) > t\}) \\ &\leqslant \frac{c}{t^p} \int_{\mathbb{R}^n} |f(x)\chi_Q(x)|^p w(x)\,dx \\ &= \frac{c}{t^p} \int_Q |f(x)|^p w(x)\,dx. \end{aligned}$$

Since this holds for $0 < t < |f|_Q$, by letting $t \to |f|_Q$, we obtain

$$\left(\frac{1}{|Q|}\int_Q |f(x)|\,dx\right)^p w(Q) \leqslant c \int_Q |f(x)|^p w(x)\,dx, \tag{4.4}$$

which is equivalent to

$$\begin{aligned} \fint_Q |f(x)|\,dx &= \frac{1}{|Q|}\int_Q |f(x)|\,dx \\ &\leqslant c\left(\frac{1}{w(Q)}\int_Q |f(x)|^p w(x)\,dx\right)^{\frac{1}{p}} \\ &= c\left(\fint_Q |f(x)|^p\,d\mu\right)^{\frac{1}{p}}, \end{aligned}$$

where μ the measure associated with w. Here we assume that $w(Q) > 0$, but this is not a serious issue, as we shall see. Observe that if $w = 1$, the inequality above follows from Hölder's (or Jensen's) inequality.

THE MORAL: This is Hölder's (or Jensen's) inequality with the Lebesgue measure on the left-hand side and the weighted measure on the right-hand side. In particular, this implies that every function in $L^p(\mathbb{R}^n; w)$ is locally integrable.

For a measurable set $E \subset Q$, we may choose $f = \chi_E$ in (4.4) and obtain

$$w(Q)\left(\frac{|E|}{|Q|}\right)^p \leqslant c\,w(E), \tag{4.5}$$

which is equivalent to

$$\left(\frac{|E|}{|Q|}\right)^p \leqslant c\,\frac{w(E)}{w(Q)}.$$

Here we assume that $w(Q) > 0$. Observe that the constants in (4.4) and (4.5) are independent of the cube Q.

THE MORAL: This is a quantitative version of absolute continuity of the Lebesgue measure with respect to the weighted measure: For every $\varepsilon > 0$ there exists $\delta > 0$, independent of a cube $Q \subset \mathbb{R}^n$, such that $\frac{|E|}{|Q|} < \varepsilon$ for every measurable set $E \subset Q$ with $\frac{w(E)}{w(Q)} < \delta$. Compare with Remark 4.21 below.

We conclude certain elementary properties of weights directly from (4.5).

Lemma 4.6. The following properties hold for a nonnegative Lebesgue measurable function w that satisfies (4.5).

(1) Either $w = 0$ or $w > 0$ almost everywhere in $\mathbb{R}^n$.

(2) Either $w \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ or $w = \infty$ almost everywhere in $\mathbb{R}^n$.

THE MORAL: We may assume that $w > 0$ almost everywhere in $\mathbb{R}^n$ and that $w \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, since otherwise we get a trivial theory. In this case, the Lebesgue measure and the weighted measure have the same sets of measure zero, the same classes of measurable sets and measurable functions.

Proof. $\boxed{(1)}$ We observe that either $w > 0$ almost everywhere in $\mathbb{R}^n$ or $w = 0$ on a set of positive measure. Let $E = \{x \in \mathbb{R}^n : w(x) = 0\}$ and assume that $|E| > 0$. We begin with showing that we can assume that E is bounded. Since

$$0 < |E| = \left| \bigcup_{k=1}^{\infty} (Q(0,k) \cap E) \right| = \lim_{k\to\infty} |E \cap Q(0,k)|,$$

by choosing k_0 large enough, we have $|E \cap Q(0,k_0)| > 0$. We obtain from (4.5) that

$$\begin{aligned}\int_{Q(0,k)} w(x)\,dx = w(Q(0,k)) &\leqslant c\left(\frac{|Q(0,k)|}{|E\cap Q(0,k_0)|}\right)^p w(E\cap Q(0,k_0))\\ &\leqslant c\left(\frac{|Q(0,k)|}{|E\cap Q(0,k_0)|}\right)^p \int_E w(x)\,dx = 0\end{aligned}$$

for every $k \geqslant k_0$. It follows that

$$\int_{\mathbb{R}^n} w(x)\,dx = \int_{\bigcup_{k=k_0}^{\infty} Q(0,k)} w(x)\,dx = \lim_{k\to\infty} \int_{Q(0,k)} w(x)\,dx = 0.$$

Since $w \geqslant 0$, we conclude that $w = 0$ almost everywhere in $\mathbb{R}^n$.

$\boxed{(2)}$ We observe that either $w = \infty$ almost everywhere in $\mathbb{R}^n$ or $w < \infty$ on a set of positive measure. Assume that $|\{x \in \mathbb{R}^n : w(x) < \infty\}| > 0$. Then

$$\begin{aligned}0 < |\{x \in \mathbb{R}^n : w(x) < \infty\}| &= \left| \bigcup_{k=1}^{\infty} \{x \in Q(0,k) : w(x) < \infty\} \right|\\ &= \lim_{k\to\infty} |\{x \in Q(0,k) : w(x) < \infty\}|.\end{aligned}$$

Thus there exists k_0 such that $|\{x \in Q(0,k_0) : w(x) < \infty\}| > 0$. Then

$$\begin{aligned}0 < |\{x \in Q(0,k_0) : w(x) < \infty\}| &= \left| \bigcup_{i=1}^{\infty} \{x \in Q(0,k_0) : w(x) < i\} \right|\\ &= \lim_{i\to\infty} |\{x \in Q(0,k_0) : w(x) < i\}|.\end{aligned}$$

Thus there exists i_0 such that $|\{x \in Q(0,k_0) : w(x) < i_0\}| > 0$. Let

$$E = \{x \in Q(0,k_0) : w(x) < i_0\}$$

and recall that $|E|>0$. We obtain from (4.5) that

$$\begin{aligned}
\int_{Q(0,k)} w(x)\,dx = w(Q(0,k)) &\leqslant c\left(\frac{|Q(0,k)|}{|E|}\right)^p w(E)\\
&\leqslant c\left(\frac{|Q(0,k)|}{|E|}\right)^p w(\{x\in Q(0,k_0): w(x)<i_0\})\\
&= c\left(\frac{|Q(0,k)|}{|E|}\right)^p \int_{\{x\in Q(0,k_0): w(x)<i_0\}} w(x)\,dx\\
&\leqslant c\left(\frac{|Q(0,k)|}{|E|}\right)^p i_0|\{x\in Q(0,k_0): w(x)<i_0\}|\\
&\leqslant c\left(\frac{|Q(0,k)|}{|E|}\right)^p i_0|Q(0,k_0)|<\infty
\end{aligned}$$

for $k\geqslant k_0$. Thus $w\in L^1(Q(0,k))$ for every $k\geqslant k_0$ and consequently $f\in L^1_{\mathrm{loc}}(\mathbb{R}^n)$. □

We continue deriving necessary conditions for (4.2) and consider the cases $p=1$ and $1<p<\infty$ separately.

The case $p=1$. If $|E|>0$, then (4.5) gives

$$\frac{1}{|Q|}\int_Q w(x)\,dx \leqslant c\frac{w(E)}{|E|}.$$

Denote $a=\operatorname{ess\,inf}_{x\in Q} w(x)$ and let $\varepsilon>0$. Recall that

$$\operatorname*{ess\,inf}_{x\in Q} w(x)=\sup\{m\in\mathbb{R}: w(x)\geqslant m \text{ for almost every } x\in Q\}.$$

There exists $E_\varepsilon\subset Q$ such that $|E_\varepsilon|>0$ and $w(x)<a+\varepsilon$ for every $x\in E_\varepsilon$. Thus

$$\frac{1}{|Q|}\int_Q w(x)\,dx\leqslant c\frac{1}{|E_\varepsilon|}\int_{E_\varepsilon}(a+\varepsilon)\,dx=c(a+\varepsilon)=c(\operatorname*{ess\,inf}_{x\in Q} w(x)+\varepsilon),$$

from which it follows that

$$\fint_Q w(x)\,dx\leqslant c\operatorname*{ess\,inf}_{x\in Q} w(x).$$

This leads to the definition of the Muckenhoupt class A_1.

Definition 4.7. A weight $w\in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, with $w(x)>0$ for almost every $x\in\mathbb{R}^n$, for which there exists a constant c, independent of cube Q, such that

$$\fint_Q w(x)\,dx\leqslant c\operatorname*{ess\,inf}_{x\in Q} w(x). \tag{4.8}$$

for every cube Q in $\mathbb{R}^n$ is called an A_1 weight. The smallest constant c for which (4.8) holds is called the A_1 constant of w and it is denoted by $[w]_{A_1}$.

Remark 4.9. Equivalently, (4.8) can be written in the form

$$\fint_Q w(x)\,dx\leqslant cw(x)\quad\text{for almost every}\quad x\in Q, \tag{4.10}$$

for every cube Q in $\mathbb{R}^n$.

Theorem 4.11. A weight $w \in A_1$ if and only if there exist a constant c such that

$$Mw(x) \leqslant cw(x) \quad \text{for almost every} \quad x \in \mathbb{R}^n. \tag{4.12}$$

Furthermore, we can choose $c = [w]_{A_1}$, if $w \in A_1$.

THE MORAL: By the Lebesgue differentiation theorem $w(x) \leqslant Mw(x) \leqslant cw(x)$ for almost every $x \in \mathbb{R}^n$. Thus an A_1 weight is pointwise comparable to its maximal function. This is a maximal function characterization of A_1.

Proof. $\boxed{\Longleftarrow}$ It is clear that (4.12) implies (4.10), since

$$\fint_Q w(y)\,dy \leqslant Mw(x) \leqslant cw(x) \quad \text{for almost every} \quad x \in \mathbb{R}^n.$$

$\boxed{\Longrightarrow}$ Assume $w \in A_1$ with the constant $[w]_{A_1}$. We claim that

$$|\{x \in \mathbb{R}^n : Mw(x) > [w]_{A_1} w(x)\}| = 0.$$

Let $x \in \mathbb{R}^n$ with $Mw(x) > [w]_{A_1} w(x)$. There exists a cube Q containing x such that

$$\fint_Q w(y)\,dy > [w]_{A_1} w(x). \tag{4.13}$$

For every $\varepsilon > 0$ there is a cube $\widetilde{Q}$, whose corners have rational coordinates with $Q \subset \widetilde{Q}$ and $|\widetilde{Q} \setminus Q| < \varepsilon$. Note that $|\widetilde{Q}| = |Q| + |\widetilde{Q} \setminus Q| < |Q| + \varepsilon$. By choosing $\varepsilon > 0$ small enough, we have

$$\fint_{\widetilde{Q}} w(y)\,dy = \frac{1}{|\widetilde{Q}|} \int_{\widetilde{Q}} w(y)\,dy \geqslant \frac{1}{|Q| + \varepsilon} \int_Q w(y)\,dy > [w]_{A_1} w(x).$$

Thus we may assume that the corner points of the cubes Q satisfying (4.13) are rational. The A_1 condition in (4.8) and (4.13) imply

$$[w]_{A_1} w(x) < \fint_Q w(y)\,dy \leqslant [w]_{A_1} \operatorname*{ess\,inf}_{y \in Q} w(y)$$

and thus

$$w(x) < \operatorname*{ess\,inf}_{y \in Q} w(y).$$

This implies that $x \in E$, $E \subset Q$ with $|E| = 0$. Since there are at most countable many cubes Q_i, $i = 1, 2, \dots$, with rational corners, satisfying (4.13), we have

$$\{x \in \mathbb{R}^n : Mw(x) > [w]_{A_1} w(x)\} \subset \bigcup_{i=1}^{\infty} (Q_i \cap E_i),$$

where $E_i \subset Q_i$ with $|E_i| = 0$, $i = 1, 2, \dots$. This implies

$$|\{x \in \mathbb{R}^n : Mw(x) > [w]_{A_1} w(x)\}| \leqslant \sum_{i=1}^{\infty} |Q_i \cap E_i| \leqslant \sum_{i=1}^{\infty} |E_i| = 0.$$

$\square$

We have shown above, that $w \in A_1$ is a necessary condition for the weighted weak type estimate. By Theorem 4.3 (1) with $p = 1$ and Theorem 4.11, we have

$$\begin{aligned} w(\{x \in \mathbb{R}^n : Mf(x) > t\}) &\leqslant \frac{c}{t} \int_{\mathbb{R}^n} |f(x)| Mw(x)\, dx \\ &\leqslant \frac{c}{t} \int_{\mathbb{R}^n} |f(x)| w(x)\, dx, \quad t > 0. \end{aligned}$$

T H E M O R A L : The Muckenhoupt A_1 condition is a necessary and sufficient condition for the weighted weak type estimate with $p = 1$.

The case $1 < p < \infty$. Next we derive necessary conditions for (4.2) in the case $1 < p < \infty$. We would like to choose $f = w^{1-p'}$ in (4.4), where p' is the conjugate exponent given by $\frac{1}{p} + \frac{1}{p'} = 1$, but we do not know whether $w^{1-p'} \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and thus some of the quantities that we encounter may be infinite. To overcome this problem, consider the sequence (w_k) defined by

$$w_k(x) = \left(w(x) + \frac{1}{k}\right)^{1-p'}, \quad k = 1, 2, \dots.$$

Since $w(x) < \infty$ for almost every $x \in \mathbb{R}^n$, we have $w_k(x) > 0$ for almost every $x \in \mathbb{R}^n$ and thus

$$0 < \int_Q w_k(x)\, dx = \int_Q \left(w(x) + \frac{1}{k}\right)^{1-p'} dx \leqslant \int_Q \left(\frac{1}{k}\right)^{1-p'} dx < \infty$$

for every cube Q in $\mathbb{R}^n$. From this we conclude that $w_k \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, (w_k) is an increasing sequence and

$$\lim_{k \to \infty} w_k(x) = w(x)^{1-p'} \quad \text{for every} \quad x \in \mathbb{R}^n.$$

Since $(1 - p')p = -p'$, by (4.4) with $f = w_k$, we obtain

$$\begin{aligned} w(Q) \left(\frac{1}{|Q|} \int_Q w_k(x)\, dx \right)^p &\leqslant c \int_Q w_k(x)^p w(x)\, dx \\ &\leqslant c \int_Q \left(w(x) + \frac{1}{k}\right)^{(1-p')p} \left(w(x) + \frac{1}{k}\right) dx \\ &= c \int_Q \left(w(x) + \frac{1}{k}\right)^{1-p'} dx \\ &= c \int_Q w_k(x)\, dx \end{aligned}$$

for every $k = 1, 2, \dots$. Since $0 < \int_Q w_k(x)\, dx < \infty$, we obtain

$$w(Q)|Q|^{-p} \left(\int_Q w_k(x)\, dx \right)^{p-1} \leqslant c$$

By passing $k \to \infty$, the Lebesgue monotone convergence theorem implies

$$w(Q)|Q|^{-p} \left(\int_Q w(x)^{1-p'}\, dx \right)^p \leqslant c$$

and consequently

$$\frac{1}{|Q|}\int_Q w(x)\,dx\left(\frac{1}{|Q|}\int_Q w(x)^{1-p'}\,dx\right)^{p-1} \leqslant c$$

for every cube Q in $\mathbb{R}^n$. Note carefully that exponent $1-p' = \frac{1}{1-p}$ with $1<p<\infty$ is a negative number. In particular, we have $w^{1-p'} \in L^1_{\text{loc}}(\mathbb{R}^n)$. Now we are ready to define the Muckenhoupt condition A_p for $1<p<\infty$.

Definition 4.14. Let $1<p<\infty$. A weight $w \in L^1_{\text{loc}}(\mathbb{R}^n)$, with $w(x)>0$ for almost every $x \in \mathbb{R}^n$, for which there exists a constant c, independent of cube Q, such that

$$\fint_Q w(x)\,dx\left(\fint_Q w(x)^{\frac{1}{1-p}}\,dx\right)^{p-1} \leqslant c \tag{4.15}$$

for every cube Q in $\mathbb{R}^n$, is called an A_p weight. In this case, we denote $w \in A_p$. The smallest constant c for which this holds, is called the A_p constant of w and it is denoted by $[w]_{A_p}$.

T H E M O R A L : The Muckenhoupt A_p condition is a necessary condition for the weighted weak type (p,p) estimate and thus and thus also for the weighted strong type (p,p).

Remarks 4.16:

(1) The A_2 condition reads

$$\fint_Q w(x)\,dx \fint_Q \frac{1}{w(x)}\,dx \leqslant c$$

for every cube Q in $\mathbb{R}^n$.

(2) If $w \in A_p$ with $1<p<\infty$, we have $w \in L^1_{\text{loc}}(\mathbb{R}^n)$ and $w^{1-p'} \in L^1_{\text{loc}}(\mathbb{R}^n)$ by (4.15). This implies $0<w(x)<\infty$ for almost every $x \in \mathbb{R}^n$ and thus

$$0 < \fint_Q w(x)\,dx < \infty \quad \text{and} \quad 0 < \fint_Q w(x)^{1-p'}\,dx < \infty.$$

T H E M O R A L : There is no danger, for example, to divide by an integral average over a cube since 0 and ∞ do not occur.

(3) By Hölder's inequality

$$\begin{aligned}
|Q| &= \int_Q w(x)^{\frac{1}{p}}\left(\frac{1}{w(x)}\right)^{\frac{1}{p}}\,dx \\
&\leqslant \left(\int_Q w(x)^{\frac{1}{p}\cdot p}\,dx\right)^{\frac{1}{p}}\left(\int_Q\left(\frac{1}{w(x)}\right)^{\frac{1}{p}\cdot p'}\,dx\right)^{\frac{1}{p'}} \\
&= \left(\int_Q w(x)\,dx\right)^{\frac{1}{p}}\left(\int_Q\left(\frac{1}{w(x)}\right)^{\frac{1}{p-1}}\,dx\right)^{\frac{1}{p'}} \\
&= \left(\int_Q w(x)\,dx\right)^{\frac{1}{p}}\left(\int_Q w(x)^{\frac{1}{1-p}}\,dx\right)^{\frac{1}{p'}},
\end{aligned}$$

for every cube Q in $\mathbb{R}^n$, from which it follows that

$$\fint_Q w(x)\,dx \left(\fint_Q w(x)^{1-p'}\,dx\right)^{p-1} \geqslant 1.$$

THE MORAL: The A_p constant $[w]_{A_p}$ is bigger or equal than one.

(4) By Hölder's (Jensen's) inequality

$$\left(\fint_Q w(x)^p\,dx\right)^{\frac{1}{p}} \leqslant \left(\fint_Q w(x)^q\,dx\right)^{\frac{1}{q}}, \quad -\infty < p \leqslant q < \infty, \quad pq \neq 0.$$

Moreover,

$$\lim_{p\to-\infty}\left(\fint_Q w(x)^p\,dx\right)^{\frac{1}{p}} = \operatorname*{ess\,inf}_{x\in Q} w(x)$$

and

$$\lim_{p\to\infty}\left(\fint_Q w(x)^p\,dx\right)^{\frac{1}{p}} = \operatorname*{ess\,sup}_{x\in Q} w(x).$$

It can be also shown that

$$\lim_{p\to 0}\left(\fint_Q w(x)^p\,dx\right)^{\frac{1}{p}} = \exp\left(\fint_Q \log w(x)\,dx\right).$$

Condition (4.15) can be written in the form

$$\fint_Q w(x)\,dx \leqslant c\left(\fint_Q w(x)^{\frac{1}{1-p}}\,dx\right)^{1-p}$$

for every cube Q in $\mathbb{R}^n$. This is a reverse Hölder inequality with a negative power on the right-hand side.

The A_1-condition can be seen as a limit of the A_p conditions as $p \to 1^+$, since

$$\lim_{p\to 1^+}\left(\fint_Q w(x)^{\frac{1}{1-p}}\,dx\right)^{1-p} = \operatorname*{ess\,inf}_{x\in Q} w(x).$$

(5) If $w \in A_p$ and $a \geqslant 0$, then function $aw \in A_p$. Moreover, we the A_p constants of aw and w are same (exercise).

(6) If $w \in A_p$ and $h \in \mathbb{R}^n$, then function $\tau_h w \in A_p$, where $\tau_h w(x) = w(x+h)$. Moreover, we the A_p constants of $\tau_h w$ and w are same (exercise).

(7) If $w \in A_p$ and $a \in \mathbb{R}$, $a \neq 0$, then function $\delta_a w \in A_p$, where $\delta_a w(x) = w(ax)$. Moreover, we the A_p constants of $\delta_a w$ and w are same (exercise).

Example 4.17. Let $w : \mathbb{R}^n \to [0,\infty]$, $w(x) = |x|^\alpha$ with $\alpha > -n$. Note that in this range of α, we have $w \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$. Then $w \in A_p$ with $1 < p < \infty$ whenever $-n < \alpha < n(p-1)$. Moreover, $w \in A_1$ whenever $-n < \alpha \leqslant 0$. See also Remark 5.4. On the other hand, the measure associated with w is doubling for $-n < \alpha < \infty$, see Definition 4.23. Thus for $\alpha > n(p-1)$ this gives an example of a doubling weight which does not belong to A_p. (Exercise)

4.2 Properties of A_p weights

The following basic properties of A_p weights follow rather directly from the definitions and Hölder's inequality.

Theorem 4.18. (1) If $1 \leqslant p < q < \infty$, then $A_p \subset A_q$.

(2) If $p > 1$, then $w \in A_p$ if and only if $w^{1-p'} \in A_{p'}$. Here p' is the conjugate exponent given by $\frac{1}{p} + \frac{1}{p'} = 1$.

(3) If $w_1, w_2 \in A_1$, then $w_1 w_2^{1-p} \in A_p$.

THE MORAL: The first claim shows that A_p classes are nested and that A_1 is the strongest condition. The second claim is an interpretation of duality. The third claim gives a method to construct A_p weights from A_1 weights. Later we shall see that all A_p weights can be written in this fashion.

Proof. $\boxed{(1)}$ If $1 = p < q < \infty$ and $w \in A_1$, then

$$\begin{aligned}\left(\fint_Q w(x)^{1-q'}\,dx\right)^{q-1} &= \left(\fint_Q \left(\frac{1}{w(x)}\right)^{\frac{1}{q-1}} dx\right)^{q-1} \leqslant \operatorname*{ess\,sup}_{x\in Q} \frac{1}{w(x)} \\ &= \frac{1}{\operatorname{ess\,inf}_{x\in Q} w(x)} \leqslant \frac{[w]_{A_1}}{\fint_Q w(x)\,dx}\end{aligned}$$

for every cube Q in $\mathbb{R}^n$. This shows that $w \in A_q$.

Assume then $1 < p < q < \infty$ and $w \in A_p$. By Hölder's inequality

$$\begin{aligned}\left(\fint_Q w(x)^{\frac{1}{1-q}}\,dx\right)^{q-1} &= \left(\frac{1}{|Q|}\int_Q \left(\frac{1}{w(x)}\right)^{\frac{1}{q-1}} dx\right)^{q-1} \\ &\leqslant \left(\frac{1}{|Q|}\right)^{q-1} \left(\int_Q \left(\left(\frac{1}{w(x)}\right)^{\frac{1}{q-1}}\right)^{\frac{q-1}{p-1}} dx\right)^{(q-1)\frac{p-1}{q-1}} \left(\int_Q 1^{\frac{q-1}{q-p}}\,dx\right)^{(q-1)\frac{q-p}{q-1}} \\ &= \left(\frac{1}{|Q|}\right)^{q-1} \left(\int_Q \left(\frac{1}{w(x)}\right)^{\frac{1}{p-1}} dx\right)^{p-1} \left(\int_Q 1\,dx\right)^{q-p} \\ &= \left(\int_Q \left(\frac{1}{w(x)}\right)^{\frac{1}{p-1}} dx\right)^{p-1} |Q|^{1-p} \\ &= \left(\fint_Q w(x)^{\frac{1}{1-p}}\,dx\right)^{p-1}.\end{aligned}$$

This implies

$$\fint_Q w(x)\,dx \left(\fint_Q w(x)^{\frac{1}{1-q}}\,dx\right)^{q-1} \leqslant \fint_Q w(x)\,dx \left(\fint_Q w(x)^{\frac{1}{1-p}}\,dx\right)^{p-1} \leqslant [w]_{A_p}$$

for every cube Q in $\mathbb{R}^n$. In the last inequality we used the fact that $w \in A_p$. This shows that $w \in A_q$.

$\boxed{(2)}$ $\boxed{\Longrightarrow}$ Assume that $w \in A_p$ with $1 < p < \infty$. Then

$$\begin{aligned}
&\fint_Q w(x)^{1-p'}\,dx \left(\fint_Q \left(w(x)^{1-p'}\right)^{\frac{1}{1-p'}}\,dx\right)^{p'-1} \\
&\qquad = \fint_Q w(x)^{1-p'}\,dx \left(\fint_Q w(x)\,dx\right)^{\frac{1}{p-1}} \\
&\qquad = \fint_Q w(x)^{\frac{1}{1-p}}\,dx \left(\fint_Q w(x)\,dx\right)^{\frac{1}{p-1}} \leqslant [w]_{A_p}^{\frac{1}{p-1}}
\end{aligned}$$

for every cube Q in $\mathbb{R}^n$. This shows that $w^{1-p'} \in A_{p'}$.

$\boxed{\Longleftarrow}$ Assume that $w^{1-p'} \in A_{p'}$ with $1 < p < \infty$. As above

$$\begin{aligned}
&\fint_Q w(x)^{\frac{1}{1-p}}\,dx \left(\fint_Q w(x)\,dx\right)^{\frac{1}{p-1}} \\
&\qquad = \fint_Q w(x)^{1-p'}\,dx \left(\fint_Q w(x)\,dx\right)^{\frac{1}{p-1}} \\
&\qquad = \fint_Q w(x)^{1-p'}\,dx \left(\fint_Q \left(w(x)^{1-p'}\right)^{\frac{1}{1-p'}}\,dx\right)^{p'-1} \leqslant [w^{1-p'}]_{A_{p'}}
\end{aligned}$$

for every cube Q in $\mathbb{R}^n$. This shows that $w \in A_p$.

$\boxed{(3)}$ Assume that $w_1 \in A_1$ and $w_2 \in A_1$. The A_p condition for $w_1 w_2^{1-p}$ is

$$\fint_Q w_1(x) w_2(x)^{1-p}\,dx \left(\fint_Q w_1(x)^{\frac{1}{1-p}} w_2(x)\,dx\right)^{p-1} \leqslant c.$$

Since $w_1 \in A_1$ and $w_2 \in A_1$,

$$\frac{1}{w_i(x)} \leqslant [w_i]_{A_1} \frac{|Q|}{w_i(Q)} \quad \text{for almost every} \quad x \in Q, \quad i = 1,2.$$

This implies

$$\begin{aligned}
\fint_Q w_1(x) w_2(x)^{1-p}\,dx &= \fint_Q w_1(x) \left(\frac{1}{w_2(x)}\right)^{p-1} dx \\
&\leqslant [w_2]_{A_1}^{p-1} \left(\frac{|Q|}{w_2(Q)}\right)^{p-1} \fint_Q w_1(x)\,dx \\
&= [w_2]_{A_1}^{p-1} \left(\frac{|Q|}{w_2(Q)}\right)^{p-1} \frac{w_1(Q)}{|Q|}
\end{aligned}$$

and

$$\begin{aligned}
\left(\fint_Q w_1(x)^{\frac{1}{1-p}} w_2(x)\,dx\right)^{p-1} &= \left(\fint_Q \left(\frac{1}{w_1(x)}\right)^{\frac{1}{p-1}} w_2(x)\,dx\right)^{p-1} \\
&\leqslant [w_1]_{A_1} \frac{|Q|}{w_1(Q)} \left(\fint_Q w_2(x)\,dx\right)^{p-1} \\
&= [w_1]_{A_1} \frac{|Q|}{w_1(Q)} \left(\frac{w_2(Q)}{|Q|}\right)^{p-1}.
\end{aligned}$$

Thus

$$\fint_Q w_1(x)w_2(x)^{1-p}\,dx\left(\fint_Q w_1(x)^{\frac{1}{1-p}}w_2(x)\,dx\right)^{p-1}\leqslant [w_1]_{A_1}[w_2]_{A_1}^{p-1}$$

for every cube Q in $\mathbb{R}^n$. □

The following result shows that the class of Muckenhoupt weights is closed under maximum and minimum.

Theorem 4.19. Let $1<p<\infty$ and assume that $v,w\in A_p$. Then $\max\{v,w\}\in A_p$ and $\min\{v,w\}\in A_p$.

Proof. $\boxed{(1)}$ Note that $\max\{v(x),w(x)\}>0$ for almost every $x\in\mathbb{R}^n$ and $\max\{v,w\}\in L^1_{\mathrm{loc}}(\mathbb{R}^n)$. By (4.15) we have

$$\begin{aligned}\fint_Q \max\{v(x),w(x)\}\,dx &\leqslant \fint_Q v(x)\,dx+\fint_Q w(x)\,dx\\ &\leqslant [v]_{A_p}\left(\fint_Q v(x)^{\frac{1}{1-p}}\,dx\right)^{1-p}+[w]_{A_p}\left(\fint_Q w(x)^{\frac{1}{1-p}}\,dx\right)^{1-p}\\ &\leqslant ([v]_{A_p}+[w]_{A_p})\left(\fint_Q \max\{v(x),w(x)\}^{\frac{1}{1-p}}\,dx\right)^{1-p}\end{aligned}$$

for every cube $Q\subset\mathbb{R}^n$. Thus $\max\{v,w\}\in A_p$.

$\boxed{(2)}$ Let $\sigma(x)=\min\{v(x),w(x)\}$ for $x\in\mathbb{R}^n$. Clearly $\sigma(x)>0$ for almost every $x\in\mathbb{R}^n$ and $\sigma\in L^1_{\mathrm{loc}}(\mathbb{R}^n)$. By Theorem 4.18 (2), we have $v^{\frac{1}{1-p}}\in A_{\frac{p}{p-1}}$ and $w^{\frac{1}{1-p}}\in A_{\frac{p}{p-1}}$. By the beginning of the proof

$$\sigma^{\frac{1}{1-p}}=(\min\{v,w\})^{\frac{1}{1-p}}=\max\{v^{\frac{1}{1-p}},w^{\frac{1}{1-p}}\}\in A_{\frac{p}{p-1}},$$

and $\min\{v,w\}=\sigma\in A_p$ follows from Theorem 4.18 (2). □

The next lemma gives a weighted Hölder type inequality for integral averages. The estimate below also implies that $f\in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ whenever $f\in L^1_{\mathrm{loc}}(\mathbb{R}^n;w)$ if $w\in A_p$.

Lemma 4.20. Let $1\leqslant p<\infty$ and assume that $w\in A_p$. Let $f\in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ be a nonnegative function. Then

$$\fint_Q |f(x)|\,dx\leqslant [w]_{A_p}^{\frac{1}{p}}\left(\frac{1}{w(Q)}\int_Q |f(x)|^p w(x)\,dx\right)^{\frac{1}{p}}$$

for every cube $Q\subset\mathbb{R}^n$.

THE MORAL: This shows that every $w\in A_p$, $1\leqslant p<\infty$, satisfies (4.4). Recall that in (4.4) we assumed that the weight satisfies a weak type estimate in (4.2). Thus we obtain (4.4) from (4.2) and $w\in A_p$, separately.

Proof. $\boxed{p=1}$ Assume that $w \in A_1$. By (4.8) we have

$$\frac{w(Q)}{|Q|} \leqslant [w]_{A_1} \operatorname*{ess\,inf}_{x\in Q} w(x)$$

for every cube Q in $\mathbb{R}^n$. This implies

$$\begin{aligned}\left(\frac{1}{|Q|}\int_Q |f(x)|\,dx\right) w(Q) &= \int_Q |f(x)|\frac{w(Q)}{|Q|}\,dx\\ &\leqslant [w]_{A_1}\int_Q |f(x)| \operatorname*{ess\,inf}_{x\in Q} w(x)\,dx\\ &\leqslant [w]_{A_1}\int_Q |f(x)|w(x)\,dx\end{aligned}$$

for every cube Q in $\mathbb{R}^n$.

$\boxed{1<p<\infty}$ Assume that $w \in A_p$ with $1<p<\infty$. Hölder's inequality and (4.15) imply

$$\begin{aligned}\fint_Q |f(x)|\,dx &= \frac{1}{|Q|}\int_Q |f(x)|w(x)^{\frac{1}{p}} w(x)^{-\frac{1}{p}}\,dx\\ &\leqslant \frac{1}{|Q|}\left(\int_Q w(x)^{\frac{1}{1-p}}\,dx\right)^{\frac{p-1}{p}}\left(\int_Q |f(x)|^p w(x)\,dx\right)^{\frac{1}{p}}\\ &= \left(\frac{w(Q)}{|Q|}\right)^{\frac{1}{p}}\left(\fint_Q w(x)^{\frac{1}{1-p}}\,dx\right)^{\frac{p-1}{p}}\left(\frac{1}{w(Q)}\int_Q |f(x)|^p w(x)\,dx\right)^{\frac{1}{p}}\\ &\leqslant [w]_{A_p}^{\frac{1}{p}}\left(\frac{1}{w(Q)}\int_Q |f(x)|^p w(x)\,dx\right)^{\frac{1}{p}}.\end{aligned}$$

□

Remark 4.21. Assume that $w \in A_p$ with $1 \leqslant p < \infty$. For a cube $Q \subset \mathbb{R}^n$ and a measurable set $E \subset Q$ we may choose $f = \chi_E$ in Lemma 4.20 and obtain

$$\left(\frac{|E|}{|Q|}\right)^p \leqslant [w]_{A_p}\frac{w(E)}{w(Q)}, \tag{4.22}$$

compare to (4.5). In particular, we may apply Lemma 4.6 and conclude that the assumptions $w \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and $w>0$ almost everywhere in $\mathbb{R}^n$ in the definition of Muckenhoupt A_p weights (Definition 4.14) are natural.

The following doubling condition is useful in harmonic analysis.

Definition 4.23. A Borel measure μ on $\mathbb{R}^n$ is doubling, if there exists a constant c such that

$$\mu(B(x,2r)) \leqslant c\mu(B(x,r))$$

for every $x \in \mathbb{R}^n$ and $r>0$. A weight w in $\mathbb{R}^n$ is doubling, if the measure associated with w is doubling, that is, there exists a constant c such that $w(B(x,2r)) \leqslant cw(B(x,r))$ for every $x \in \mathbb{R}^n$ and $r>0$.

THE MORAL: The doubling condition gives a scale and location invariant bound for the measure of a ball with the radius doubled.

Remarks 4.24:

(1) The Lebesgue measure with $w = 1$ is doubling since

$$w(B(x,2r)) = \int_{B(x,2r)} 1\,dy = |B(x,2r)| = 2^n|B(x,r)| = 2^n w(B(x,r))$$

for every $x \in \mathbb{R}^n$ and $r > 0$.

(2) The doubling condition in Definition 4.23 can be equivalently stated in the form that there exists a constant c such that $w(Q(x,2l)) \leqslant cw(Q(x,l))$ for every $x \in \mathbb{R}^n$ and $l > 0$ (exercise). Recall that $Q(x,l)$ is a cube with center at $x \in \mathbb{R}^n$ and side length $l > 0$.

As an application of Lemma 4.20 we conclude that Muckenhoupt weights are doubling.

Theorem 4.25. Let $1 < p < \infty$ and assume that $w \in A_p$. Then w is doubling, that is, there exists a constant $c = c(n,p,[w]_{A_p})$ such that $w(B(x,2r)) \leqslant cw(B(x,r))$ for every $x \in \mathbb{R}^n$ and $r > 0$.

Proof. Let $x \in \mathbb{R}^n$ and $r > 0$. Since $Q(x,\frac{4r}{\sqrt{n}}) \subset B(x,2r) \subset Q(x,4r)$, Lemma 4.20 with $f = \chi_{B(x,r)}$ implies

$$\begin{aligned}
\frac{1}{2^n} = \frac{|B(x,r)|}{|B(x,2r)|} &\leqslant \frac{|B(x,r)|}{|Q(x,\frac{4r}{\sqrt{n}})|} = n^{\frac{n}{2}}\frac{|B(x,r)|}{|Q(x,4r)|} \\
&= n^{\frac{n}{2}} \fint_{Q(x,4r)} f(y)\,dy \\
&\leqslant n^{\frac{n}{2}}[w]_{A_p}^{\frac{1}{p}} \left(\frac{1}{w(Q(x,4r))}\int_{Q(x,4r)} f(y)^p w(y)\,dy\right)^{\frac{1}{p}} \\
&\leqslant n^{\frac{n}{2}}[w]_{A_p}^{\frac{1}{p}} \left(\frac{w(B(x,r))}{w(B(x,2r))}\right)^{\frac{1}{p}}.
\end{aligned}$$

The claim follows by raising both sides to power p and reorganizing the terms. □

Let w be a doubling weight in $\mathbb{R}^n$ with a constant c_1. For a measurable function $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n;w)$ the weighted maximal function $M^w f(x)$ at $x \in \mathbb{R}^n$ is

$$M^w f(x) = \sup_{r>0} \frac{1}{w(B(x,r))}\int_{B(x,r)} |f(y)|w(y)\,dy. \tag{4.26}$$

where the supremum is taken over all cubes $Q \subset \mathbb{R}^n$ with $x \in Q$. Note that

$$M^w f(x) = \sup_{r>0} \frac{1}{\mu(B(x,r))}\int_{B(x,r)} |f(y)|\,d\mu(y)$$

where μ is the measure associated with w. For $w = 1$ we obtain the standard Hardy-Littlewood maximal function Mf.

Let $1 < p \leqslant \infty$. There exists a constant $c = c(n, p, c_1)$ such that

$$\|M^w f\|_{L^p(\mathbb{R}^n;w)} \leqslant c\|f\|_{L^p(\mathbb{R}^n;w)} \tag{4.27}$$

for every $f \in L^p(\mathbb{R}^n; w)$. The proof of this statement is a straightforward adaptation of the proof of the maximal function theorem by applying a weak type $(1,1)$ estimate for $M^w f$ (exercise).

THE MORAL: The weighted weak type $(1,1)$ estimate and the weighted strong type (p,p) estimate for $M^w f$ can be proved is the same way as in the nonweighted case, if the weighted measure is doubling. Note carefully the difference to (4.1) and (4.2), where we consider weighted norm inequalities for the nonweighted Hardy-Littlewood maximal operator.

Remark 4.28. Let w be a doubling weight in $\mathbb{R}^n$ and let $f \in L^1_{\text{loc}}(\mathbb{R}^n; w)$. There exists a measurable set E such that $w(E) = |E| = 0$ and

$$\lim_{r\to 0} \frac{1}{w(B(x,r))} \int_{B(x,r)} |f(y) - f(x)| w(y)\, dy = 0$$

for every $x \in \mathbb{R}^n \setminus E$. As a consequence, if $x \in \mathbb{R}^n \setminus E$ and $(Q_j)_{j\in\mathbb{N}}$ is a sequence of cubes such that $x \in Q_j$ for every $j \in \mathbb{N}$ and $l(Q_j) \to 0$ as $j \to \infty$, then

$$f(x) = \lim_{j\to\infty} \frac{1}{w(Q_j)} \int_{Q_j} f(y) w(y)\, dy.$$

The proof of this statement is a straightforward adaptation of the proof of the Lebesgue differentiation theorem (exercise).

4.3 A weak type characterization of A_p

We have already shown that the Muckenhoupt A_1 condition is a necessary and sufficient condition for the weighted weak type estimate in (4.2) with $p = 1$. Next we discuss the corresponding characterization for A_p weights with $1 \leqslant p < \infty$.

Theorem 4.29 (Charaterization of the weak type (p,p) estimate). Assume that $w \in L^1_{\text{loc}}(\mathbb{R}^n)$ such that $w(x) > 0$ almost every $x \in \mathbb{R}^n$ and let $1 \leqslant p < \infty$. Then $w \in A_p$ if and only if there exists constant $c < \infty$ such that

$$w(\{x \in \mathbb{R}^n : Mf(x) > t\}) \leqslant \frac{c}{t^p} \int_{\mathbb{R}^n} |f(x)|^p w(x)\, dx$$

for every $t > 0$ and $f \in L^1_{\text{loc}}(\mathbb{R}^n)$.

THE MORAL: The A_p condition with $1 \leqslant p < \infty$ is equivalent with weighted weak type (p,p) estimate in (4.2). In other words, the Muckenhoupt condition characterizes weights for which a weighted weak type estimate holds for the Hardy-Littlewood maximal function.

Proof. $\boxed{\Longleftarrow}$ We have already shown that the A_p condition is necessary for the weighted weak type (p,p) estimate.

$\boxed{\Longrightarrow}$ Assume that $w \in A_p$. By Theorem 4.25, the weight w is doubling. We would like to apply Lemma 1.26 in which we assume that the function is integrable. Thus we consider the truncated functions

$$f_k = f\chi_{B(0,k)} \in L^1(\mathbb{R}^n), \quad k = 1,2,\dots.$$

Lemma 1.26 gives

$$\{x \in \mathbb{R}^n : Mf_k(x) > 4^n t\} \subset \bigcup_{i=1}^{\infty} 3Q_i$$

where Q_i, $i = 1,2,\dots$, are the Calderón-Zygmund cubes for f_k at level $t > 0$. Since the Calderón-Zygmund cubes are pairwise disjoint, we obtain

$$w(\{x \in \mathbb{R}^n : Mf_k(x) > 4^n t\}) \leqslant \sum_{i=1}^{\infty} w(3Q_i).$$

Moreover, by Lemma 1.22 and Theorem 1.12, we have

$$\fint_{Q_i} |f_k(x)|\,dx > t, \quad i = 1,2,\dots.$$

Using the doubling property of w and Lemma 4.20, we conclude

$$\begin{aligned}
w(\{x \in \mathbb{R}^n : Mf_k(x) > 4^n t\}) &\leqslant \sum_{i=1}^{\infty} w(3Q_i) \leqslant c\sum_{i=1}^{\infty} w(Q_i) \\
&\leqslant c\sum_{i=1}^{\infty} \left(\fint_{Q_i} |f_k(x)|\,dx\right)^{-p} \int_{Q_i} |f_k(x)|^p w(x)\,dx \\
&\leqslant \frac{c}{t^p} \sum_{i=1}^{\infty} \int_{Q_i} |f_k(x)|^p w(x)\,dx \\
&\leqslant \frac{c}{t^p} \int_{\mathbb{R}^n} |f_k(x)|^p w(x)\,dx.
\end{aligned}$$

Then we would like to pass $k \to \infty$. Since $Mf_k(x) \leqslant Mf_{k+1}(x)$ for every $x \in \mathbb{R}^n$ and $k = 1,2,\dots$, we have

$$w(\{x \in \mathbb{R}^n : Mf_k(x) > t\}) \leqslant w(\{x \in \mathbb{R}^n : Mf_{k+1}(x) > t\}), \quad k = 1,2,\dots.$$

We claim that

$$\{x \in \mathbb{R}^n : Mf(x) > t\} = \bigcup_{k=1}^{\infty} \{x \in \mathbb{R}^n : Mf_k(x) > t\}.$$

First we note that $Mf_k(x) \leqslant Mf(x)$ for every $x \in \mathbb{R}^n$ and $k = 1,2,\dots$ and thus $\{x \in \mathbb{R}^n : Mf_k(x) > t\} \subset \{x \in \mathbb{R}^n : Mf(x) > t\}$. For the reverse inclusion, let $x \in \mathbb{R}^n$ such that $Mf(x) > t$. Then there exist a cube Q containing x such that $\fint_Q |f(y)|\,dy > t$. Choose k large enough so that $Q \subset B(0,k)$. Then

$$Mf_k(x) \geqslant \fint_Q |f_k(y)|\,dy = \fint_Q |f(y)|\,dy > t$$

and thus $\{x \in \mathbb{R}^n : Mf(x) > t\} \subset \{x \in \mathbb{R}^n : Mf_k(x) > t\}$. Finally, by the monotone convergence theorem

$$\begin{aligned} w(\{x \in \mathbb{R}^n : Mf(x) > t\}) &= w\left(\bigcup_{k=1}^{\infty} \{x \in \mathbb{R}^n : Mf_k(x) > t\}\right) \\ &= \lim_{k\to\infty} w(\{x \in \mathbb{R}^n : Mf_k(x) > t\}) \\ &\leqslant \lim_{k\to\infty} \frac{c}{t^p} \int_{\mathbb{R}^n} |f_k(x)|^p w(x)\,dx \\ &= \frac{c}{t^p} \int_{\mathbb{R}^n} |f(x)|^p w(x)\,dx. \end{aligned}$$ □

Remarks 4.30:

(1) The weak type estimate for $f \in L^p(\mathbb{R}^n;\mu)$ follows from Theorem 4.29 by considering sequence (w_k), with $w_k(x) = w(x) + \frac{1}{k}$, $k = 1,2,\dots$.

(2) In the beginning of this section, we showed

$$(4.2) \Longrightarrow (4.4) \Longrightarrow w \in A_p.$$

On the other hand, the previous theorem implies $w \in A_p \Longrightarrow$ (4.2). Thus

$$w \in A_p \Longleftrightarrow (4.2) \Longleftrightarrow (4.4).$$

4.4 A strong type characterization of A_p

By the Marcinkiewicz interpolation theorem and Theorem 4.29 we may conclude that the Hardy-Littlewood maximal operator is bounded on $L^p(\mathbb{R}^n;w)$, with $1 < p < \infty$, whenever $w \in \bigcup_{q<p} A_q$, see Section 4.7. Remark 4.45 below asserts that $\bigcup_{q<p} A_q = A_p$, but this is based on a deep self-improving property of Muckenhoupt weights. In this section we discuss a direct proof by Lerner [14], which shows that the Hardy-Littlewood maximal operator is bounded in $L^p(\mathbb{R}^n;w)$ whenever $w \in A_p$ with $1 < p < \infty$, see (4.1).

Theorem 4.31 (Charaterization of the strong type (p,p) estimate). Assume that $w \in L^1_{\text{loc}}(\mathbb{R}^n)$ such that $w(x) > 0$ almost every $x \in \mathbb{R}^n$ and let $1 < p < \infty$. Then $w \in A_p$ if and only if there exists a constant $c < \infty$ such that

$$\int_{\mathbb{R}^n} (Mf(x))^p w(x)\,dx \leqslant c \int_{\mathbb{R}^n} |f(x)|^p w(x)\,dx$$

for every $f \in L^1_{\text{loc}}(\mathbb{R}^n)$.

THE MORAL: The A_p condition with $1 < p < \infty$ is equivalent with weighted strong type (p,p) estimate (4.1). In other words, the Muckenhoupt condition characterizes weights for which a weighted maximal function theorem holds true.

Proof. $\boxed{\Longleftarrow}$ Since weak type (p,p) estimate (4.2) follows from strong type (p,p) estimate (4.1), necessity have been proved in the beginning of the chapter, see also Theorem 4.29.

$\boxed{\Longrightarrow}$ Let $\sigma = w^{\frac{1}{1-p}}$. Theorem 4.18 (2) implies that $\sigma \in A_{p'}$, where $p' = \frac{p}{p-1}$. Assume that $f \in L^p(\mathbb{R}^n;w)$. Let $x \in \mathbb{R}^n$ and consider a ball $B = B(x_0,r)$ with $x \in B$. Then

$$\fint_B |f(y)|\,dy = c\left(\frac{|B|}{w(B(x_0,3r))}\left(\frac{1}{\sigma(B(x_0,3r))}\int_B |f(y)|\,dy\right)^{p-1}\right)^{\frac{1}{p-1}},$$

where

$$\begin{aligned}
c &= \left(\frac{w(B(x_0,3r))\sigma(B(x_0,3r))^{p-1}}{|B|^p}\right)^{\frac{1}{p-1}}\\
&\leqslant c(n,p)\left(\frac{w(Q(x_0,6r))\sigma(Q(x_0,6r))^{p-1}}{|Q(x_0,6r)|^p}\right)^{\frac{1}{p-1}}\\
&= c(n,p)\left(\fint_{Q(x_0,6r)} w(y)\,dy\left(\fint_{Q(x_0,6r)} w(y)^{\frac{1}{1-p}}\,dy\right)^{p-1}\right)^{\frac{1}{p-1}}\\
&\leqslant c(n,p)[w]_{A_p}^{\frac{1}{p-1}}.
\end{aligned}$$

For $z \in B$, we have $B \subset B(z,2r) \subset B(x_0,3r)$ and thus

$$\begin{aligned}
\frac{1}{\sigma(B(x_0,3r))}\int_B |f(y)|\,dy &\leqslant \frac{1}{\sigma(B(z,2r))}\int_{B(z,2r)} |f(y)\sigma(y)^{-1}|\sigma(y)\,dy\\
&\leqslant M^\sigma(f\sigma^{-1})(z),
\end{aligned}$$

where M^σ is the weighted maximal function as in (4.26). This implies

$$\begin{aligned}
|B|\left(\frac{1}{\sigma(B(x_0,3r))}\int_B |f(y)|\,dy\right)^{p-1} &= \int_B\left(\frac{1}{\sigma(B(x_0,3r))}\int_B |f(y)|\,dy\right)^{p-1} dz\\
&\leqslant \int_B \big(M^\sigma(f\sigma^{-1})(z)\big)^{p-1}\,dz\\
&= \int_B \big(M^\sigma(f\sigma^{-1})(z)\big)^{p-1} w(z)^{-1}w(z)\,dz
\end{aligned}$$

and consequently

$$\begin{aligned}
\fint_B |f(y)|\,dy &\leqslant c\left(\frac{1}{w(B(x_0,3r))}\int_B \big(M^\sigma(f\sigma^{-1})(z)\big)^{p-1} w(z)^{-1}w(z)\,dz\right)^{\frac{1}{p-1}}\\
&\leqslant c\left(\frac{1}{w(B(x,2r))}\int_{B(x,2r)} \big(M^\sigma(f\sigma^{-1})(z)\big)^{p-1} w(z)^{-1}w(z)\,dz\right)^{\frac{1}{p-1}}\\
&\leqslant c\big(M^w\big(M^\sigma(f\sigma^{-1})^{p-1}w^{-1}\big)(x)\big)^{\frac{1}{p-1}}
\end{aligned}$$

with $c = c(n,p,[w]_{A_p})$. By taking supremum over all balls B with $x \in B$, we obtain

$$M^* f(x) \leqslant c\big(M^w\big(M^\sigma(f\sigma^{-1})^{p-1}w^{-1}\big)(x)\big)^{\frac{1}{p-1}}$$

for every $x \in \mathbb{R}^n$ with $c = c(n,p,[w]_{A_p})$. Here M^* is the noncentered maximal function of $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ defined by

$$M^* f(x) = \sup_{B \ni x} \fint_B |f(y)|\,dy,$$

where the supremum is taken over all balls B with $x \in B$.

Observe that $w \in A_p$ and $\sigma \in A_{p'}$ are doubling weights by Theorem 4.25, with constants depending on n, p and $[w]_{A_p}$. The strong type estimate (4.27) for the weighted maximal function M^w with the exponent $\frac{p}{p-1} > 1$ and M^σ with the exponent $p > 1$ imply

$$\begin{aligned}
\int_{\mathbb{R}^n} (M^* f(x))^p w(x)\,dx &\leqslant c \int_{\mathbb{R}^n} \big(M^w\big(M^\sigma(f\sigma^{-1})^{p-1} w^{-1}\big)(x)\big)^{\frac{p}{p-1}} w(x)\,dx \\
&\leqslant c \int_{\mathbb{R}^n} \big(\big(M^\sigma(f\sigma^{-1})^{p-1}(x) w^{-1}(x)\big)\big)^{\frac{p}{p-1}} w(x)\,dx \\
&= c \int_{\mathbb{R}^n} \big(M^\sigma(f\sigma^{-1})(x)\big)^p \sigma(x)\,dx \\
&\leqslant c \int_{\mathbb{R}^n} \big(f(x)\sigma(x)^{-1}\big)^p \sigma(x)\,dx \\
&= c \int_{\mathbb{R}^n} |f(x)|^p \sigma(x)^{1-p}\,dx \\
&= c \int_{\mathbb{R}^n} |f(x)|^p w(x)\,dx,
\end{aligned}$$

with $c = c(n,p,[w]_{A_p})$, where we used the facts that $\sigma = w^{\frac{1}{1-p}}$ and $w = \sigma^{1-p}$. $\square$

4.5 A_∞ and reverse Hölder inequalities

This section discusses the limiting Muckenhoupt condition with $p = \infty$. There are several equivalent versions of the definition. The Muckenhoupt condition below is a reverse inequality to (4.22), which holds for A_p weights with $1 < p < \infty$.

Definition 4.32. weight $w \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, with $w(x) > 0$ for almost every $x \in \mathbb{R}^n$, belongs to the Muckenhoupt class A_∞, if there exist constants $c, \delta > 0$ such that

$$\frac{w(E)}{w(Q)} \leqslant c \left(\frac{|E|}{|Q|}\right)^\delta \tag{4.33}$$

whenever $Q \subset \mathbb{R}^n$ is a cube and $E \subset Q$ is a measurable set. The smallest constant c for which this holds, is called the A_∞ constant of w and it is denoted by $[w]_{A_\infty}$.

T H E M O R A L : The A_∞ condition gives a quantitative version of absolute continuity of the weighted measure with respect to the Lebesgue measure. If $w \in A_\infty$, for every $\varepsilon > 0$ there exists $\delta > 0$, independent of a cube $Q \subset \mathbb{R}^n$, such that $\frac{w(E)}{w(Q)} < \varepsilon$ for every measurable set $A \subset Q$ with $\frac{|E|}{|Q|} < \delta$.

By Theorem 4.18 (1), we have $A_p \subset A_q$ for $1 \leqslant p < q < \infty$. We show that $A_\infty = \bigcup_{1<p<\infty} A_p$. Neither inclusion is trivial. The main tool is a reverse Hölder inequality, see Theorem 4.40 (3). The first step is to show that an A_∞ weight is doubling.

Theorem 4.34. Assume that $w \in A_\infty$. Then w is doubling, that is, there exists a constant $c = c(n,w)$ such that $w(B(x,2r)) \leqslant cw(B(x,r))$ for every $x \in \mathbb{R}^n$ and $r > 0$.

Proof. Let $[w]_{A_\infty}$ and δ be the constants in (4.33). Let $0 < \beta < 1$ with $0 < [w]_{A_\infty}\beta^\delta < 1$ and $\alpha = 1 - [w]_{A_\infty}\beta^\delta > 0$. Let $Q \subset \mathbb{R}^n$ be a cube and $E \subset Q$ a measurable set with $|E| \geqslant (1-\beta)|Q|$. Then $|Q \setminus E| = |Q| - |E| \leqslant \beta|Q|$ and (4.33) implies

$$w(Q \setminus E) \leqslant [w]_{A_\infty} \left(\frac{|Q \setminus E|}{|Q|} \right)^\delta w(Q) \leqslant [w]_{A_\infty}\beta^\delta w(Q).$$

Consequently,

$$w(E) = w(Q) - w(Q \setminus E) \geqslant (1 - [w]_{A_\infty}\beta^\delta) w(Q) = \alpha w(Q). \tag{4.35}$$

Let $x \in \mathbb{R}^n$ and $r > 0$. Let $l_1 = 4r$ and

$$l_{j+1} = (1-\beta)^{\frac{1}{n}} l_j \quad \text{for every} \quad j = 1, \dots, k-1,$$

where $k = k(n,\beta)$ is the smallest integer for which $Q(x,l_k) \subset B(x,r)$. Then

$$Q(x,l_{j+1}) \subset Q(x,l_j) \quad \text{and} \quad |Q(x,l_{j+1})| = (1-\beta)|Q(x,l_j)|$$

for every $j = 1, \dots, k-1$. Hence we may apply (4.35) iteratively for these cubes, and obtain

$$\begin{aligned} w(B(x,2r)) &\leqslant w(Q(x,4r)) = w(Q(x,l_1)) \leqslant \alpha^{-1} w(Q(x,l_2)) \\ &\leqslant \dots \leqslant \alpha^{1-k} w(Q(x,l_k)) \leqslant \alpha^{1-k} w(B(x,r)). \end{aligned}$$

This proves the claim with the constant $c(n,w) = \alpha^{1-k}$. $\square$

Let $Q \subset \mathbb{R}^n$ be a cube, σ be a weight and $f \in L^1(Q;\sigma)$ be a nonnegative function. We denote the σ-weighted average of f by

$$f_{Q;\sigma} = \frac{1}{\sigma(Q)} \int_Q f(x)\sigma(x)\,dx.$$

Next provide a sufficient condition for a weighted reverse Hölder inequality. Lemma 4.36 and its proof are interesting already in the unweighted case $\sigma = 1$, see Remark 4.39 below.

Lemma 4.36. Let $\sigma \in A_\infty$ and let $w \geqslant 0$ be a measurable function such that $w\sigma \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$. Assume that there exist $\alpha > 0$ and $0 < \beta < 1$ such that

$$\sigma(\{x \in Q : w(x) > \beta w_{Q;\sigma}\}) \geqslant \alpha\sigma(Q)$$

for every cube $Q \subset \mathbb{R}^n$. Then there exist constants $q = q(n,\sigma,\alpha,\beta) > 1$ and $c = c(n,\sigma,\alpha,\beta)$ such that

$$\left(\frac{1}{\sigma(Q_0)}\int_{Q_0} w(x)^q \sigma(x)\,dx\right)^{\frac{1}{q}} \leqslant \frac{c}{\sigma(Q_0)}\int_{Q_0} w(x)\sigma(x)\,dx$$

for every cube $Q_0 \subset \mathbb{R}^n$.

Proof. Let $Q_0 \subset \mathbb{R}^n$ be a cube. For $t > w_{Q_0;\sigma}$, let $\mathscr{D}_t \subset \mathscr{D}(Q_0)$ be the collection of maximal dyadic subcubes $Q \subset Q_0$ such that $w_{Q;\sigma} > t$, compare the proof of Theorem 1.4. Observe that $Q \subset Q_0$ with $Q \neq Q_0$ for every $Q \in \mathscr{D}_t$, since $t > w_{Q_0;\sigma}$, see Remark 1.7.

Let $Q \in \mathscr{D}_t$ and let $Q' \in \mathscr{D}(Q_0)$ be the dyadic parent cube of Q, which satisfies the conditions $Q \subset Q'$ and $2^n|Q| = |Q'|$. Theorem 4.34 implies

$$\frac{\sigma(Q')}{\sigma(Q)} \leqslant c(n,\sigma).$$

By maximality of $Q \in \mathscr{D}_t$, we have $w_{Q';\sigma} \leqslant t$ and

$$\begin{aligned} t < w_{Q;\sigma} &= \frac{1}{\sigma(Q)}\int_Q w(x)\sigma(x)\,dx \\ &\leqslant \frac{\sigma(Q')}{\sigma(Q)}\frac{1}{\sigma(Q')}\int_{Q'} w(x)\sigma(x)\,dx \leqslant ct. \end{aligned} \tag{4.37}$$

with $c = c(n,\sigma)$. On the other hand, if $x \in Q_0 \setminus \bigcup_{Q\in\mathscr{D}_t} Q$, then $w_{Q';\sigma} \leqslant t$ for every $Q' \in \mathscr{D}(Q_0)$ with $x \in Q'$. Otherwise $Q' \subset Q$ for some $Q \in \mathscr{D}_t$, and this would contradict the choice of x. Since there are arbitrarily small such cubes Q' containing x, the Lebesgue differentiation theorem for the doubling weight σ, see Remark 4.28, shows that $w(x) \leqslant t$ for almost every $x \in Q_0 \setminus \bigcup_{Q\in\mathscr{D}_t} Q$ with respect to the measure associated with σ.

Let μ be the measure associated with the weight $w\sigma$, that is,

$$\mu(A) = \int_A w(x)\sigma(x)\,dx$$

for every Lebesgue measurable subset A of $\mathbb{R}^n$. We claim that μ is absolutely continuous with respect to the measure induced by σ. If $A \subset \mathbb{R}^n$ such that $\sigma(A) = \int_A \sigma(x)\,dx = 0$, then $\sigma(x) = 0$ almost every $x \in A$ and

$$\mu(A) = \int_A w(x)\sigma(x)\,dx = 0.$$

By (4.37) and the assumption we obtain

$$
\begin{aligned}
\mu(\{x \in Q_0 : w(x) > t\}) &\leqslant \sum_{Q \in \mathscr{D}_t} \mu(Q) = \sum_{Q \in \mathscr{D}_t} \int_Q w(x)\sigma(x)\,dx \\
&\leqslant ct \sum_{Q \in \mathscr{D}_t} \sigma(Q) \\
&\leqslant \frac{ct}{\alpha} \sum_{Q \in \mathscr{D}_t} \sigma(\{x \in Q : w(x) > \beta w_{Q;\sigma}\}) \qquad (4.38)\\
&\leqslant \frac{ct}{\alpha} \sum_{Q \in \mathscr{D}_t} \sigma(\{x \in Q : w(x) > \beta t\}) \\
&\leqslant \frac{ct}{\alpha} \sigma(\{x \in Q_0 : w(x) > \beta t\})
\end{aligned}
$$

with $c = c(n,\sigma)$. The final inequality holds since the cubes in $\mathscr{D}_t \subset \mathscr{D}(Q_0)$ are pairwise disjoint.

Next we multiply (4.38) by t^{q-2}, where $1 < q < 2$ is to be specified later, and then integrate the resulting estimate from $w_{Q_0;\sigma}$ to $t_0 > w_{Q_0;\sigma}$. This gives

$$
\begin{aligned}
&\int_{w_{Q_0;\sigma}}^{t_0} t^{q-2} \mu(\{x \in Q_0 : w(x) > t\})\,dt \\
&\leqslant \frac{c}{\alpha\beta^q} \int_0^{t_0} (\beta t)^q \sigma(\{x \in Q_0 : w(x) > \beta t\}) \frac{dt}{t} \\
&= \frac{c}{\alpha\beta^q} \int_0^{t_0} (\beta t)^q \sigma(\{x \in Q_0 : \min\{w(x), t_0\} > \beta t\}) \frac{dt}{t} \\
&\leqslant \frac{c}{\alpha\beta^q} \int_0^{\infty} (\beta t)^q \sigma(\{x \in Q_0 : \min\{w(x), t_0\} > \beta t\}) \frac{dt}{t} \\
&= \frac{c}{q\alpha\beta^q} \int_{Q_0} \min\{w(x), t_0\}^q \sigma(x)\,dx \\
&\leqslant \frac{c}{q\alpha\beta^q} \int_{Q_0} \min\{w(x), t_0\}^{q-1} w(x)\sigma(x)\,dx < \infty
\end{aligned}
$$

with $c = c(n,\sigma)$. By Fubini's theorem,

$$
\begin{aligned}
&\int_{w_{Q_0;\sigma}}^{t_0} t^{q-2}\mu(\{x \in Q_0 : w(x) > t\})\,dt \\
&\quad = \int_{w_{Q_0;\sigma}}^{t_0} t^{q-2}\mu(\{x \in Q_0 : \min\{w(x), t_0\} > t\})\,dt \\
&\quad = \int_{\{x\in Q_0 : w(x) > w_{Q_0;\sigma}\}} \left(\int_{w_{Q_0;\sigma}}^{\min\{w(x),t_0\}} t^{q-2}\,dt\right) w(x)\sigma(x)\,dx \\
&\quad = \frac{1}{q-1}\int_{\{x\in Q_0 : w(x) > w_{Q_0;\sigma}\}} \left(\min\{w(x), t_0\}^{q-1} - (w_{Q_0;\sigma})^{q-1}\right) w(x)\sigma(x)\,dx \\
&\quad = \frac{1}{q-1}\int_{Q_0} w(x)\min\{w(x), t_0\}^{q-1}\sigma(x)\,dx \\
&\qquad - \frac{1}{q-1}\int_{\{x\in Q_0 : w(x) \leqslant w_{Q_0;\sigma}\}} \min\{w(x), t_0\}^{q-1} w(x)\sigma(x)\,dx \\
&\qquad - \frac{1}{q-1}\int_{\{x\in Q_0 : w(x) > w_{Q_0;\sigma}\}} (w_{Q_0;\sigma})^{q-1} w(x)\sigma(x)\,dx \\
&\quad \geqslant \frac{1}{q-1}\int_{Q_0} \min\{w(x), t_0\}^{q-1} w(x)\sigma(x)\,dx - \frac{(w_{Q_0;\sigma})^{q-1}}{q-1}\int_{Q_0} w(x)\sigma(x)\,dx.
\end{aligned}
$$

Combining the estimates above, we arrive at

$$
\begin{aligned}
\left(\frac{1}{q-1} - \frac{c(n,\sigma)}{q\alpha\beta^q}\right)\int_{Q_0} \min\{w(x), t_0\}^{q-1} w(x)\sigma(x)\,dx &\leqslant \frac{(w_{Q_0;\sigma})^{q-1}}{q-1}\int_{Q_0} w(x)\sigma(x)\,dx \\
&= \frac{1}{q-1}\left(\frac{1}{\sigma(Q_0)}\int_{Q_0} w(x)\sigma(x)\,dx\right)^q \sigma(Q_0)
\end{aligned}
$$

for every $t_0 > w_{Q_0;\sigma}$. By choosing $1 < q < 2$ to be so small that

$$
\frac{1}{q-1} - \frac{c(n,\sigma)}{q\alpha\beta^q} \geqslant 1,
$$

we have

$$
\frac{1}{\sigma(Q_0)}\int_{Q_0} \min\{w(x), t_0\}^{q-1} w(x)\sigma(x)\,dx \leqslant c\left(\frac{1}{\sigma(Q_0)}\int_{Q_0} w(x)\sigma(x)\,dx\right)^q,
$$

with $c = c(n,\sigma,\alpha,\beta)$. Letting $t_0 \to \infty$ and applying Fatou's lemma we have

$$
\begin{aligned}
\frac{1}{\sigma(Q_0)}\int_{Q_0} w(x)^q\sigma(x)\,dx &\leqslant \liminf_{t_0\to\infty} \frac{1}{\sigma(Q_0)}\int_{Q_0} \min\{w(x), t_0\}^{q-1} w(x)\sigma(x)\,dx \\
&\leqslant c\left(\frac{1}{\sigma(Q_0)}\int_{Q_0} w(x)\sigma(x)\,dx\right)^q
\end{aligned}
$$

with $c = c(n,\sigma,\alpha,\beta)$. □

Remark 4.39. Before discussing a general result related to reverse Hölder inequalities, we discuss the special case of A_1. Assume that $w \in A_1$. By Theorem 4.11, we have

$$
M_{d,Q}w(x) \leqslant Mw(x) \leqslant [w]_{A_1} w(x)
$$

for almost every $x \in Q$. Here $M_{d,Q}$ is the dyadic maximal function in a cube Q as in (1.29). Chebyshev's inequality and Lemma 1.31 imply

$$\begin{aligned}
|\{x \in Q : w(x) > t\}| &\leqslant \frac{1}{t} \int_{\{x \in Q : w(x) > t\}} w(x)\,dx \\
&\leqslant \int_{\{x \in Q : M_{d,Q} w(x) > t\}} w(x)|\,dx \\
&\leqslant 2^n t |\{x \in Q : M_{d,Q} w(x) > t\}| \\
&\leqslant 2^n t |\{x \in Q : w(x) > [w]_{A_1}^{-1} t\}|
\end{aligned}$$

for every $t \geqslant w_{Q_0}$. This is similar to (4.38) with μ and σ equal to the Lebesgue measure. Proceeding as in the proof of Lemma 4.36, we conclude that there exist constants $q = q(n, [w]_{A_1}) > 1$ and $c = c(n, [w]_{A_1})$ such that

$$\left(\fint_Q w(x)^q \,dx \right)^{\frac{1}{q}} \leqslant c \fint_Q w(x)\,dx$$

for every cube $Q \subset \mathbb{R}^n$. This shows that every A_1 weight satisfies a reverse Hölder inequality. A similar argument can be applied to show that every A_p weight, with $1 \leqslant p < \infty$, satisfies a reverse Hölder inequality, see [8].

Next we discuss a general result related to reverse Hölder inequalities.

Theorem 4.40. Assume that $w \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, with $w(x) > 0$ for almost every $x \in \mathbb{R}^n$, Then the following three conditions are equivalent:

(1) $w \in A_\infty$,

(2) $w \in A_p$ for some $1 < p < \infty$,

(3) there are constants c and $q > 1$ such that

$$\left(\fint_Q w(x)^q \,dx \right)^{\frac{1}{q}} \leqslant c \fint_Q w(x)\,dx$$

for every cube $Q \subset \mathbb{R}^n$.

Moreover, the constants and exponents in each of the conditions only depend on n, w, and each other.

THE MORAL: A weight belongs to a Muckenhoupt class if and only if it satisfies a reverse Hölder inquality. Moreover, $A_\infty = \bigcup_{1<p<\infty} A_p$.

Proof. $\boxed{(1) \Longrightarrow (2)}$ Let $\sigma \in A_\infty$ and $w = \sigma^{-1}$. Let $0 < \beta = \beta(\sigma) < 1$ so small that $[w]_{A_\infty} \beta^\delta < 1$, where $[w]_{A_\infty} > 0$ and $\delta = \delta(\sigma) > 0$ are the constants appearing in (4.33) for σ. Let $\alpha = 1 - [w]_{A_\infty} \beta^\delta > 0$. Let $Q \subset \mathbb{R}^n$ be a cube and $E = \{x \in Q :$

$w(x) > \beta w_{Q;\sigma}\}$. Since $w(x) \leqslant \beta w_{Q;\sigma}$ for every $x \in Q \setminus E$, we obtain

$$\begin{aligned}\frac{|Q\setminus E|}{|Q|} &= \frac{1}{|Q|}\int_{Q\setminus E} w(x)w(x)^{-1}\,dx = \frac{1}{|Q|}\int_{Q\setminus E} w(x)\sigma(x)\,dx \\ &\leqslant \frac{\beta w_{Q;\sigma}}{|Q|}\int_{Q\setminus E}\sigma(x)\,dx = \frac{\beta}{|Q|}\frac{1}{\sigma(Q)}\int_Q w(x)\sigma(x)\,dx\int_{Q\setminus E}\sigma(x)\,dx \\ &= \frac{\beta}{|Q|}\frac{|Q|}{\sigma(Q)}\sigma(Q\setminus E) \leqslant \beta.\end{aligned}$$

By (4.33), we have

$$1 = \frac{\sigma(E)+\sigma(Q\setminus E)}{\sigma(Q)} \leqslant \frac{\sigma(E)}{\sigma(Q)} + [w]_{A_\infty}\left(\frac{|Q\setminus E|}{|Q|}\right)^{\delta} \leqslant \frac{\sigma(E)}{\sigma(Q)} + [w]_{A_\infty}\beta^{\delta}, \tag{4.41}$$

and thus

$$\alpha\sigma(Q) = \left(1-[w]_{A_\infty}\beta^{\delta}\right)\sigma(Q) \leqslant \sigma(E).$$

From Lemma 4.36 we obtain $q = q(n,\sigma,\alpha,\beta) = q(n,\sigma) > 1$ and $c = c(n,\sigma,\alpha,\beta) = c(n,\sigma)$ such that

$$\begin{aligned}\frac{1}{\sigma(Q)^{\frac{1}{q}}}\left(\int_Q \sigma(x)^{1-q}\,dx\right)^{\frac{1}{q}} &= \left(\frac{1}{\sigma(Q)}\int_Q w(x)^q\sigma(x)\,dx\right)^{\frac{1}{q}} \\ &\leqslant \frac{c}{\sigma(Q)}\int_Q w(x)\sigma(x)\,dx \\ &= c\frac{|Q|}{\sigma(Q)} = c\frac{|Q|^{\frac{1}{q}}|Q|^{\frac{q-1}{q}}}{\sigma(Q)^{\frac{1}{q}}\sigma(Q)^{\frac{q-1}{q}}}.\end{aligned}$$

Reorganization of the terms gives

$$|Q|^{-\frac{q-1}{q}}\sigma(Q)^{\frac{q-1}{q}}|Q|^{-\frac{1}{q}}\left(\int_Q \sigma(x)^{1-q}\,dx\right)^{\frac{1}{q}} \leqslant c(n,\sigma).$$

With $p = \frac{q}{q-1}$ we have

$$\fint_Q \sigma(x)\,dx\left(\fint_Q \sigma(x)^{\frac{1}{1-p}}\,dx\right)^{p-1} \leqslant c.$$

Hence $\sigma \in A_p$ with $p = p(n,\sigma)$ and the A_p constant $c = c(n,\sigma)$.

$\boxed{(2) \Longrightarrow (3)}$ Let $1 < p < \infty$ and assume that $w \in A_p$. Let $0 < \beta = \beta(p,w) < 1$ be so small that $(\beta[w]_{A_p})^{\frac{1}{p-1}} < 1$. Let $\alpha = 1-(\beta[w]_{A_p})^{\frac{1}{p-1}} > 0$. Let Q be a cube and $E = \{x \in Q : w(x) > \beta w_Q\}$. Since $w \in A_p$, we have

$$\begin{aligned}\frac{1}{\beta}\left(\frac{|Q\setminus E|}{|Q|}\right)^{p-1} &= w_Q\left(\frac{1}{|Q|}\int_{Q\setminus E}(\beta w_Q)^{\frac{1}{1-p}}\,dx\right)^{p-1} \\ &\leqslant w_Q\left(\fint_Q w(x)^{\frac{1}{1-p}}\,dx\right)^{p-1} \leqslant [w]_{A_p}.\end{aligned}$$

As in (4.41), we obtain

$$|E| \geqslant \left(1 - (\beta[w]_{A_p})^{\frac{1}{p-1}}\right)|Q| = \alpha|Q|.$$

From Lemma 4.36, with $\sigma(x) = 1$ for every $x \in \mathbb{R}^n$, we obtain $q = q(n,\alpha,\beta) = q(n,p,w) > 1$ and $c = c(n,\alpha,\beta) = c(n,p,w) > 0$ such that

$$\left(\fint_Q w(x)^q \, dx\right)^{\frac{1}{q}} \leqslant c \fint_Q w(x)\, dx$$

for every cube $Q \subset \mathbb{R}^n$, and thus assertion (3) holds.

$\boxed{(3) \Longrightarrow (1)}$ Let w be a weight and assume that there are constants c and $q > 1$ such that (3) holds. Let $Q \subset \mathbb{R}^n$ be a cube and $E \subset Q$ be a measurable set. By Hölder's inequality and the reverse Hölder inequality, we obtain

$$\begin{aligned} w(E) &= \int_Q w(x)\,dx \leqslant |E|^{\frac{q-1}{q}} \left(\int_Q w(x)^q \, dx\right)^{\frac{1}{q}} \\ &\leqslant c|E|^{\frac{q-1}{q}} |Q|^{\frac{1}{q}} \fint_Q w(x)\,dx = c\left(\frac{|E|}{|Q|}\right)^{\frac{q-1}{q}} w(Q). \end{aligned}$$

Thus w satisfies the A_∞ condition (4.33) with $\delta = \frac{q-1}{q} > 0$, and (1) holds. □

Remark 4.42. By Remark 4.21 we have a reverse inequality to (4.33) for $w \in A_p$. Theorem 4.40 implies that the Lebesgue measure and the weighted measure with $w \in A_\infty$ are mutually absolutely continuous with quantitative estimates.

4.6 Self-improving properties of A_p

From Theorem 4.40 we obtain a self-improveving property for Muckenhoupt weights.

Theorem 4.43. Let $1 < p < \infty$ and assume that $w \in A_p$. There exists a constant $\varepsilon = \varepsilon(n,p,[w]_{A_p}) > 0$ such that $w \in A_{p-\varepsilon}$, with $[w]_{A_{p-\varepsilon}}$ depending only on n, p, and $[w]_{A_p}$.

T H E M O R A L : Assume that $w \in A_p$. Then $w \in A_{p+\varepsilon}$ for every $\varepsilon > 0$, but there also exists a small $\varepsilon > 0$ for which $w \in A_{p-\varepsilon}$.

Proof. By Theorem 4.18 (2), we have $w^{\frac{1}{1-p}} \in A_{p'}$ with $p' = \frac{p}{p-1}$, and $[w]_{A_{p'}}$ depends only on p and $[w]_{A_p}$. Theorem 4.40 implies that there exist $c = c(n,p,[w]_{A_p})$ and $q = q(n,p,[w]_{A_p}) > 1$ such that

$$\left(\fint_Q w(x)^{\frac{q}{1-p}} \, dx\right)^{\frac{1}{q}} \leqslant c \fint_Q w(x)^{\frac{1}{1-p}}\, dx \tag{4.44}$$

for every cube $Q \subset \mathbb{R}^n$. Let $0 < \varepsilon = \varepsilon(p,q) = \varepsilon(n,p,[w]_{A_p}) < p-1$ be determined by $\frac{q}{p-1} = \frac{1}{p-\varepsilon-1}$. By (4.44) and $w \in A_p$, we have

$$\begin{aligned}\left(\fint_Q w(x)^{\frac{1}{1-(p-\varepsilon)}}\,dx\right)^{p-\varepsilon-1} &= \left(\fint_Q w(x)^{\frac{q}{1-p}}\,dx\right)^{\frac{p-1}{q}}\\ &\leqslant c\left(\fint_Q w(x)^{\frac{1}{1-p}}\,dx\right)^{p-1}\\ &\leqslant c\left(\fint_Q w(x)\,dx\right)^{-1}\end{aligned}$$

for every cube $Q \subset \mathbb{R}^n$ with $c = c(n,p,[w]_{A_p})$. This shows that $w \in A_{p-\varepsilon}$ with a constant $c(n,p,[w]_{A_p})$. □

Remark 4.45. For $1 < p < \infty$ we have $A_p = \bigcup_{q<p} A_q$.

The reverse Hölder inequality implies the following self-improving property of A_p weights.

Theorem 4.46. If $w \in A_p$ with $1 \leqslant p < \infty$, then there exists $\varepsilon = \varepsilon(p,q) = \varepsilon(n,p,[w]_{A_p}) > 0$ such that $w^{1+\varepsilon} \in A_p$.

Proof. If $w \in A_p$ with $1 \leqslant p < \infty$, then by Theorem 4.40 there exists $\varepsilon = \varepsilon(n,p,[w]_{A_p}) > 0$ and $c = c(n,p,[w]_{A_p})$ such that

$$\fint_Q w(x)^{1+\varepsilon}\,dx \leqslant c\left(\fint_Q w(x)\,dx\right)^{1+\varepsilon} \tag{4.47}$$

for every cube $Q \subset \mathbb{R}^n$.

$\boxed{p=1}$ Assume that $w \in A_1$. Then (4.47) and the A_1 condition (4.8) imply

$$\fint_Q w(x)^{1+\varepsilon}\,dx \leqslant c\left(\fint_Q w(x)\,dx\right)^{1+\varepsilon} \leqslant c(\operatorname*{ess\,inf}_{x\in Q} w(x))^{1+\varepsilon} = c\operatorname*{ess\,inf}_{x\in Q} w(x)^{1+\varepsilon}$$

This shows that $w^{1+\varepsilon} \in A_1$.

$\boxed{1<p<\infty}$ Assume that $w \in A_p$ with $1 < p < \infty$. Theorem 4.18 (2) implies $w^{1-p'} \in A_{p'}$. By Theorem 4.40 we may choose $\varepsilon > 0$, so that both w and $w^{1-p'}$ satisfy a revere Hölder inequality with the same exponent $1+\varepsilon$, that is, (4.47) holds and

$$\fint_Q w(x)^{(1-p')(1+\varepsilon)}\,dx \leqslant c\left(\fint_Q w(x)^{1-p'}\,dx\right)^{1+\varepsilon}$$

for every cube $Q \subset \mathbb{R}^n$. Together with (4.15), this implies

$$\begin{aligned}&\fint_Q w(x)^{1+\varepsilon}\,dx\left(\fint_Q w(x)^{(1+\varepsilon)(1-p')}\,dx\right)^{p-1}\\ &\qquad\leqslant c\left(\fint_Q w(x)\,dx\right)^{1+\varepsilon}\left(\fint_Q w(x)^{1-p'}\,dx\right)^{(1+\varepsilon)(p-1)} \leqslant c < \infty.\end{aligned}$$

This shows that $w^{1+\varepsilon} \in A_p$. □

4.7 Strong type characterization revisited

In this section we give two proofs for the strong type estimate in Theorem 4.31 based on the self-improving property of Muckenhoupt weights in Theorem 4.43. Assume $w \in A_p$.

$\boxed{(1)}$ The first argument is based on the Marcinkiewicz interpolation theorem, see Theorem 2.4, which is applied with the weighted measure

$$w(E) = \mu(E) = \int_E w(x)\,dx.$$

Observe that the Marcinkewicz interpolation theorem holds for other measures than the Lebesgue measure as well, see Remark 2.5 (5). The Hardy-Littlewood maximal function is of strong type (∞,∞) with respect to the Lebesgue measure, since

$$\|Mf\|_{L^\infty(\mathbb{R}^n)} \leqslant \|f\|_{L^\infty(\mathbb{R}^n)},$$

see the discussion in Example 2.6. We claim that the Hardy-Littlewood maximal function is also of weighted type (∞,∞). Since $w(x) > 0$ for almost every $x \in \mathbb{R}^n$, we observe that $w(A) = 0$ if and only if $|A| = 0$. This implies

$$\begin{aligned}\|f\|_{L^\infty(\mathbb{R}^n;w)} &= \inf\{t : w(\{x \in \mathbb{R}^n : |f(x)| > t\}) = 0\}\\ &= \inf\{t : |\{x \in \mathbb{R}^n : |f(x)| > t\}| = 0\} = \|f\|_{L^\infty(\mathbb{R}^n)}.\end{aligned}$$

Thus

$$\|Mf\|_{L^\infty(\mathbb{R}^n;w)} = \|Mf\|_{L^\infty(\mathbb{R}^n)} \leqslant \|f\|_{L^\infty(\mathbb{R}^n)} = \|f\|_{L^\infty(\mathbb{R}^n;w)}.$$

By Theorem 4.43, we have $w \in A_q$ for some $q < p$, so that by Theorem 4.29 we see that the Hardy-Littlewood maximal operator is of weak type (q,q), that is,

$$w(\{x \in \mathbb{R}^n : Mf(x) > t\}) \leqslant c \int_{\mathbb{R}^n} |f(x)|^q w(x)\,dx \quad t > 0.$$

Since the Hardy-Littlewood maximal operator is of weak type (q,q) and strong type (∞,∞), the Marcinkiewicz interpolation theorem, see Theorem 2.4, shows that it is of strong type (p,p) whenever $q < p < \infty$.

$\boxed{(2)}$ Then we consider the second approach. Without loss of generality, we may assume that the right-hand side of the strong type estimate in Theorem 4.31 is finite. By Theorem 4.43, there exists $1 < q = q(n,p,[w]_{A_p}) < p$ such that $w \in A_q$. Let $x \in \mathbb{R}^n$, $r > 0$ and $Q = Q(x,2r)$. Lemma 4.20 implies

$$\begin{aligned}\fint_{B(x,r)} |f(y)|\,dy &\leqslant c(n) \fint_Q |f(y)|\,dy\\ &\leqslant c(n,q,w)\left(\frac{1}{w(Q)}\int_Q |f(y)|^q w(y)\,dy\right)^{\frac{1}{q}}\\ &\leqslant c(n,q,w)\big(M^w(|f|^q)(x)\big)^{\frac{1}{q}},\end{aligned}$$

where M^w is the weighted maximal function, see (4.26). By taking supremum over all $r > 0$, we obtain

$$M|f|(x) \leqslant c(n,q,w)\big(M^w(|f|^q)(x)\big)^{\frac{1}{q}},$$

where M is the standard centered maximal function over balls. By (4.27) we arrive at

$$\begin{aligned}\int_{\mathbb{R}^n}(Mf(x))^p w(x)\,dx &\leqslant c(n,p,q,w)\int_{\mathbb{R}^n}\big(M^w(|f|^q)(x)\big)^{\frac{p}{q}} w(x)\,dx\\ &\leqslant c(n,p,w)\int_{\mathbb{R}^n}|f(x)|^p w(x)\,dx.\end{aligned}$$

Finally, the A_q constant of w depends on w only through $[w]_{A_p}$, see Theorem 4.43. The constant in Lemma 4.20 only depends on $[w]_{A_q}$ and q. The constant in (4.27) depends on n, $\frac{p}{q}$ and the doubling constant of w, which in turn only depends on $[w]_{A_q}$, n and q, see the proof of Theorem 4.25. Hence we conclude that the claim holds with $c = c(n,p,w) = c(n,p,[w]_{A_p})$.

Remark 4.48. The previous proof shows that standard and weighted L^∞ norms are same if the measures are mutually absolutely continuous. Thus weights for which we have a weighted strong type (∞,∞) are precisely those weights. This shows that the question of characterizing weights for which we have a weighted strong type (∞,∞) is much easier than the corresponding question for $1 \leqslant p < \infty$. In this sense the A_p condition can be seen as a quantitative version of absolute continuity, see Remark 4.21 and (4.33).

5

A_p and BMO

5.1 Characterizations of A_p

In this section we discuss two characterizations of Muckenhoupt weights. We begin with a characterization of the A_1 weights by the Hardy-Littlewood maximal function

$$Mf(x) = \sup_{Q \ni x} \fint_Q |f(y)|\,dy,$$

where the supremum is taken over all cubes Q in $\mathbb{R}^n$ containing x. This characterization is based on the argument in the proof of Theorem 2.8 (Kolmogorov). The following result has been originally shown by Coifman and Rochberg in [6]. Later we shall apply this result to give a characterization of BMO by the Hardy-Littlewood maximal function.

Theorem 5.1 (Coifman-Rochberg (1980)). (1) Assume that $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ is such that $Mf(x) < \infty$ for almost every $x \in \mathbb{R}^n$ and let $0 < \delta < 1$. Then $w(x) = (Mf(x))^\delta$ is an A_1 weight with $[w]_{A_1}$ depending only on n and δ.

(2) Conversely, if $w \in A_1$, then there exist $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, $0 < \delta < 1$ and b, with $b, \frac{1}{b} \in L^\infty(\mathbb{R}^n)$, such that $w(x) = b(x)(Mf(x))^\delta$ for almost every $x \in \mathbb{R}^n$.

THE MORAL: This result asserts that every A_1 weight is essentially of the form $(Mf(x))^\delta$ for some $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and $0 < \delta < 1$. Moreover, this gives a useful method to construct A_1 weights.

Proof. $\boxed{(1)}$ We show that there exists constant $c = c(n,\delta)$, such that for every cube Q and for every $x \in Q$ we have

$$\fint_Q Mf(y)^\delta\,dy \leqslant cMf(x)^\delta.$$

This implies

$$\fint_Q Mf(y)^\delta \, dy \leqslant c \inf_{x\in Q} Mf(x)^\delta \leqslant c \operatorname*{ess\,inf}_{x\in Q} Mf(x)^\delta,$$

which means that $(Mf)^\delta \in A_1$.

Let Q be a cube in $\mathbb{R}^n$ and $x \in Q$. Decompose f as $f(x) = f_1(x) + f_2(x)$ with

$$f_1(x) = f(x)\chi_{2Q}(x) \quad \text{and} \quad f_2(x) = f(x)\chi_{\mathbb{R}^n \setminus 2Q}(x).$$

We may consider $(Mf_1)^\delta$ and $(Mf_2)^\delta$ separately, since by sublinearity of the maximal operator $Mf(y) \leqslant Mf_1(y) + Mf_2(y)$ and by the elementary inequality $(a+b)^\delta \leqslant a^\delta + b^\delta$, $a, b > 0$, $0 < \delta < 1$, we obtain

$$Mf(y)^\delta \leqslant Mf_1(y)^\delta + Mf_2(y)^\delta$$

for every $y \in \mathbb{R}^n$. This implies that

$$\fint_Q Mf(y)^\delta \, dy \leqslant \fint_Q Mf_1(y)^\delta \, dy + \fint_Q Mf_2(y)^\delta \, dy.$$

Thus is is enough to show

$$\fint_Q Mf_i(y)^\delta \, dy \leqslant cMf(x)^\delta, \quad c = c(n,\delta), \quad i = 1,2. \tag{5.2}$$

We may assume that $\|f_1\|_1 > 0$ for $i = 1,2$, since otherwise $f_1 = 0$ almost everywhere in $\mathbb{R}^n$ and $Mf_1 = 0$ almost everywhere in $\mathbb{R}^n$. By the Cavalieri principle and the weak type (1,1) estimate for the maximal operator, there exists a constant $c = c(n,p)$ such that

$$\begin{aligned}
\fint_Q Mf_1(y)^\delta \, dy &= \frac{\delta}{|Q|} \int_0^\infty t^{\delta-1} |\{x \in Q : Mf_1(x) > t\}| \, dt \\
&\leqslant \frac{\delta}{|Q|} \int_0^\infty t^{\delta-1} \min\left\{|Q|, \frac{c}{t}\|f_1\|_1\right\} dt \\
&= \frac{\delta}{|Q|} \left(\int_0^{t_0} \dots dt + \int_{t_0}^\infty \dots dt \right) \\
&\leqslant \frac{\delta}{|Q|} \left(\int_0^{t_0} |Q| t^{\delta-1} \, dt + \int_{t_0}^\infty c\|f_1\|_1 t^{\delta-2} \, dt \right) \\
&= \Big/_{0}^{t_0} t^\delta + c \frac{\|f_1\|_1}{|Q|} \frac{\delta}{\delta - 1} \Big/_{t_0}^{\infty} t^{\delta-1} \\
&= t_0^\delta \left(1 + c \frac{\|f_1\|_1}{|Q|} \frac{\delta}{1-\delta} t_0^{-1} \right).
\end{aligned}$$

The integral above has been divided into two parts by using an arbitrary parameter

$t_0 > 0$. By choosing $t_0 = \frac{\|f_1\|_1}{|Q|}$, we arrive at

$$\begin{aligned}
\fint_Q Mf_1(y)^\delta \, dy &\leqslant \frac{\|f_1\|_1^\delta}{|Q|^\delta}\left(1 + c\frac{\|f_1\|_1}{|Q|}\frac{\delta}{1-\delta}\frac{|Q|}{\|f_1\|_1}\right) \\
&= \left(1+\frac{c\delta}{1-\delta}\right)\left(\frac{1}{|Q|}\int_{2Q}|f(y)|\,dy\right)^\delta \\
&= 2^{n\delta}\left(1+\frac{c\delta}{1-\delta}\right)\left(\frac{1}{|2Q|}\int_{2Q}|f(y)|\,dy\right)^\delta \\
&\leqslant cMf(x)^\delta
\end{aligned}$$

with $c = c(n,p,\delta)$. This proves (5.2) for $(Mf_1)^\delta$. Note that we could have used Theorem 2.8 (Kolmogorov) directly in the argument as well.

To obtain a similar estimate for $(Mf_2)^\delta$ we derive a pointwise estimate in Q. Let $y \in Q$ and let R be any cube containing y. If $R \subset 2Q$, then

$$\frac{1}{|R|}\int_R |f_2(z)|\,dz = 0.$$

Thus we may assume that R intersects the complement of $2Q$ and this gives a lower bound $l(R) > \frac{1}{2}l(Q)$ for the side length of R. This implies $Q \subset 5R$. Thus

$$\begin{aligned}
\frac{1}{|R|}\int_R |f_2(z)|\,dz &\leqslant \frac{5^n}{|5R|}\int_{5R}|f_2(z)|\,dz \\
&\leqslant \frac{5^n}{|5R|}\int_{5R}|f(z)|\,dz \\
&\leqslant 5^n Mf(x).
\end{aligned}$$

By taking supremum over all cubes R containing y, we obtain $Mf_2(y) \leqslant 5^n Mf(x)$ for every $y \in Q$. This implies that

$$\fint_Q Mf_2(y)^\delta \, dy \leqslant 5^{n\delta} Mf(x)^\delta$$

which proves (5.2) for $(Mf_2)^\delta$.

$\boxed{(2)}$ If $w \in A_1$, by Theorem 4.18 (1) we have $w \in A_p$ for every $p > 1$. By Theorem 4.40, there exist $c = c(n,[w]_{A_1}) > 0$ and $\varepsilon = \varepsilon(n,[w]_{A_1}) > 0$ such that

$$\left(\fint_Q w(y)^{1+\varepsilon}\,dy\right)^{\frac{1}{1+\varepsilon}} \leqslant c \fint_Q w(y)\,dy$$

for every cube Q in $\mathbb{R}^n$. Together with (4.12), this implies

$$\begin{aligned}
M(w^{1+\varepsilon})(x)^{\frac{1}{1+\varepsilon}} &= \left(\sup_{Q\ni x}\fint_Q w(y)^{1+\varepsilon}\,dy\right)^{\frac{1}{1+\varepsilon}} = \sup_{Q\ni x}\left(\fint_Q w(y)^{1+\varepsilon}\,dy\right)^{\frac{1}{1+\varepsilon}} \\
&\leqslant c\sup_{Q\ni x}\fint_Q w(y)\,dy = cMw(x) \leqslant cw(x)
\end{aligned}$$

for almost every $x \in \mathbb{R}^n$. By Hölder's inequality

$$\begin{aligned} w(x) \leqslant Mw(x) &= \sup_{Q \ni x} \fint_Q w(y)\,dy \\ &\leqslant \left(\sup_{Q \ni x} \fint_Q w(y)^{1+\varepsilon}\,dy \right)^{\frac{1}{1+\varepsilon}} = M(w^{1+\varepsilon})(x)^{\frac{1}{1+\varepsilon}} \end{aligned}$$

for almost every $x \in \mathbb{R}^n$. Thus we have

$$w(x) \leqslant Mf(x)^\delta \leqslant cw(x) \quad \text{for almost every} \quad x \in \mathbb{R}^n,$$

where $f(x) = w(x)^{1+\varepsilon}$ and $\delta = \frac{1}{1+\varepsilon}$. Then

$$w(x) = \frac{w(x)}{Mf(x)^\delta} Mf(x)^\delta = b(x) Mf(x)^\delta$$

for almost every $x \in \mathbb{R}^n$ with

$$b(x) = \frac{w(x)}{Mf(x)^\delta}.$$

Note that $0 < \frac{1}{c} \leqslant \|b\|_\infty \leqslant 1$ by the estimate above. □

Remark 5.3. The claim (1) in Theorem 5.1 does not hold for $\delta = 1$. In fact, if f is a compactly supported measurable function with $0 < \|f\|_\infty < \infty$, then the function Mf cannot be an A_1 weight nor an A_2 weight. To see this, assume that $Mf \in A_1$. Theorem 4.18 (1) implies $A_1 \subset A_2$ and thus we have $Mf \in A_2$. Theorem 4.18 (2) with $p = 2$ implies that $w = (Mf)^{-1} \in A_2$. In particular, we have

$$0 < w(x) = \frac{1}{Mf(x)} \leqslant \frac{1}{f(x)}$$

for almost every $x \in \mathbb{R}^n$. By the boundedness of the maximal function in $L^2(\mathbb{R}^n; w)$, see Theorem 4.31, we obtain

$$\begin{aligned} \int_{\mathbb{R}^n} Mf(x)\,dx &= \int_{\mathbb{R}^n} (Mf(x))^2 w(x)\,dx \\ &\leqslant c \int_{\mathbb{R}^n} |f(x)|^2 w(x)\,dx \\ &\leqslant c \int_{\mathbb{R}^n} |f(x)|\,dx < \infty. \end{aligned}$$

However, by the properties of the maximal function, the left-hand side is finite only if $f(x) = 0$ for almost every $x \in \mathbb{R}^n$.

Remark 5.4. The original result of Coifman and Rochberg is more general. Assume that μ is a positive Borel measure on $\mathbb{R}^n$ and define the maximal function

$$M\mu(x) = \sup_{Q \ni x} \frac{1}{|Q|} \int_Q 1\,d\mu(x) = \sup_{Q \ni x} \frac{\mu(Q)}{|Q|}$$

for every $x \in \mathbb{R}^n$. If $0 < M\mu(x) < \infty$ for almost every $x \in \mathbb{R}^n$ and $0 < \delta < 1$, we can show as above that $w = (M\mu)^\delta$ is an A_1 weight. Claim (1) in the previous theorem follows by considering the measure $\mu(E) = \int_E |f(x)|\,dx$ for Lebesgue measurable $E \subset \mathbb{R}^n$. In particular, if $\mu = \delta$, where δ is the Dirac measure at the origin, then

$$M\delta(x) = c|x|^{-n}.$$

Thus $|x|^\alpha \in A_1$, if $-n < \alpha \neq 0$ and, by Theorem 4.18 (3), we conclude $|x|^\alpha \in A_p$, if $-n < \alpha < n(p-1)$. See also Example 4.17.

By Theorem (4.18) (3), we know that $w_1, w_2 \in A_1$ implies $w_1 w_2^{1-p} \in A_p$. The next result shows that the converse holds true as well. Thus we obtain a characterization for A_p weights with $1 < p < \infty$. This was first proved by Jones [12]. The proof presented here is from [5].

Theorem 5.5 (Jones factorization (1980)). Let w be a weight and $1 < p < \infty$. Then $w \in A_p$ if and only if $w = w_1 w_2^{1-p}$, where $w_1, w_2 \in A_1$.

THE MORAL: Every weight $w \in A_p$ with $1 < p < \infty$ can be written as a product of two A_1 weights in the form $w_1 w_2^{1-p}$, where $w_1, w_2 \in A_1$. This is a factorization result for Muckenhoupt weights.

Proof. First we consider the case $p = p' = 2$. Assume that $w \in A_2$. We claim that $w = w_1 w_2^{-1}$, where $w_1, w_2 \in A_1$. Consider operator

$$Sf = w^{-\frac{1}{2}} M(w^{\frac{1}{2}} f).$$

Since $w \in A_2$, we also have $w^{-1} \in A_2$ and Theorem 4.31 implies that $M : L^2(\mathbb{R}^n; w^{-1}) \to L^2(\mathbb{R}^n; w^{-1})$ is a bounded operator, that is,

$$\int_{\mathbb{R}^n} (Mf(x))^2 w(x)^{-1}\,dx \leqslant c \int_{\mathbb{R}^n} |f(x)|^2 w(x)^{-1}\,dx$$

for every $f \in L^2(\mathbb{R}^n; w^{-1})$. This implies that $S : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is bounded. Indeed, if $f \in L^2(\mathbb{R}^n)$, then $w^{\frac{1}{2}} f \in L^2(\mathbb{R}^n; w^{-1})$ and

$$\begin{aligned}\int_{\mathbb{R}^n} (Sf(x))^2\,dx &= \int_{\mathbb{R}^n} M(w^{\frac{1}{2}} f)(x)^2 w(x)^{-1}\,dx \\ &\leqslant c \int_{\mathbb{R}^n} |w(x)^{\frac{1}{2}} f(x)|^2 w(x)^{-1}\,dx \\ &= c \int_{\mathbb{R}^n} |f(x)|^2\,dx.\end{aligned}$$

By switching the roles of w and w^{-1} we see that also $f \mapsto w^{\frac{1}{2}} M(w^{-\frac{1}{2}} f)$ is a bounded operator $L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$.

Let

$$Tf = w^{-\frac{1}{2}} M(w^{\frac{1}{2}} f) + w^{\frac{1}{2}} M(w^{-\frac{1}{2}} f).$$

Then $T : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is a bounded operator. Moreover, since the Hardy-Littlewood maximal function is sublinear, we see that T is subadditive. Thus

$$\|Tf\|_2 \leqslant c\|f\|_2 \quad \text{and} \quad T(f+g) \leqslant Tf + Tg$$

for every $f, g \in L^2(\mathbb{R}^n)$.

Let $f \in L^2\mathbb{R}^n)$ and

$$\eta = \sum_{k=1}^{\infty} (2c)^{-k} T^k f, \tag{5.6}$$

where $T^k = T^{k-1} \circ T$ is the iterated operator and c is the constant in the L^2 bound above. Since

$$\|T^k f\|_2 \leqslant c^k \|f\|_2, \quad k = 1, 2, \dots,$$

the series (5.6) converges absolutely and by the completeness of $L^2(\mathbb{R}^n)$ we conclude that $\eta \in L^2(\mathbb{R}^n)$. By the properties of T, we obtain

$$T\eta \leqslant \sum_{k=1}^{\infty} (2c)^{-k} T^{k+1} f = \sum_{k=2}^{\infty} (2c)^{1-k} T^k f = 2c \sum_{k=2}^{\infty} (2c)^{-k} T^k f \leqslant 2c\eta.$$

This implies that $w_1 = w^{\frac{1}{2}}\eta$ is an A_1 weight, since

$$\begin{aligned} Mw_1 = M(w^{\frac{1}{2}}\eta) &\leqslant M(w^{\frac{1}{2}}\eta) + wM(w^{-\frac{1}{2}}\eta) \\ &= w^{\frac{1}{2}} T\eta \leqslant 2c\eta w^{\frac{1}{2}} = 2cw_1. \end{aligned}$$

In the same way we see that $w_2 = w^{-\frac{1}{2}}\eta \in A_1$ (exercise). This proves the claim is the case $p = 2$, since

$$w = w^{\frac{1}{2}}\eta(w^{-\frac{1}{2}}\eta)^{-1} = w_1 w_2^{-1} = w_1 w_2^{1-p}.$$

Assume then that $w \in A_p$ with $p \geqslant 2$. Set

$$Tf = \left(w^{-\frac{1}{p}} M(w^{\frac{1}{p}} f^{\frac{p}{p'}})\right)^{\frac{p'}{p}} + w^{\frac{1}{p}} M(w^{-\frac{1}{p}} f).$$

Since $w \in A_p$, Theorem 4.18 (2) implies that $w^{1-p'} = w^{-\frac{p'}{p}} \in A_{p'}$. Theorem 4.31 implies that $M : L^{p'}(\mathbb{R}^n; w^{-p'/p}) \to L^{p'}(\mathbb{R}^n; w^{-p'/p})$ is a bounded operator and $M : L^p(\mathbb{R}^n; w) \to L^p(\mathbb{R}^n; w)$ is a bounded operator. As above, we see that $T : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$ is a bounded operator. Since $\frac{p}{p'} \geqslant 1$, Minkowski's inequality implies that T is sublinear. Consider η defined by (5.6), then $\eta \in L^p(\mathbb{R}^n)$. The claim follows (exercise) by choosing

$$w_1 = w^{\frac{1}{p}} \eta^{\frac{p}{p'}} \quad \text{and} \quad w_2 = w^{-\frac{1}{p}} \eta.$$

Assume then that $1 < p \leqslant 2$. Since $w^{1-p'} = w^{-\frac{p'}{p}} \in A_{p'}$, with $p' \geqslant 2$, as above we obtain

$$w^{-\frac{p'}{p}} = w_1 w_2^{1-p'},$$

where $w_1, w_2 \in A_1$. The claim follows by raising this equality to the power $-\frac{p}{p'}$. □

5.2 Characterizations of BMO

In this section we discuss the connections between functions of bounded mean oscillation and Muckenhoupt's weights. In particular, we discuss two characterizations of BMO.

Theorem 5.7. Assume that $p > 1$.

(1) If $w \in A_p$, then $\log w \in \mathrm{BMO}$. Moreover, if $w \in A_2$, then $\|\log w\|_* \leqslant \log(2[w]_{A_2})$.

(2) Conversely, if $f \in \mathrm{BMO}$, then $\exp(\delta f) \in A_p$ for $\delta = \delta(n, \|f\|_*) > 0$ small enough.

THE MORAL: This gives the following characterization for BMO: If $p > 1$, then

$$\mathrm{BMO} = \{a \log w : a \geqslant 0, w \in A_p\}.$$

Proof. $\boxed{(1)}$ Assume first that $w \in A_p$ with $1 \leqslant p \leqslant 2$. Let $Q \subset \mathbb{R}^n$ be a cube. By Theorem 4.18 (1) we have $w \in A_2$. Let $f = \log w$. Then $\exp f = w \in A_2$ and by the A_2 condition we have

$$\fint_Q \exp(f(x))\,dx \fint_Q \exp(-f(x))\,dx = \fint_Q w(x)\,dx \fint_Q w(x)^{-1}\,dx \leqslant [w]_{A_2}.$$

This implies

$$\begin{aligned}
&\fint_Q \exp(f(x) - f_Q)\,dx \fint_Q \exp(f_Q - f(x))\,dx \\
&\qquad = \exp(-f_Q)\exp(f_Q) \fint_Q \exp(f(x))\,dx \fint_Q \exp(-f(x))\,dx \leqslant [w]_{A_2}.
\end{aligned}$$

By Jensen's inequality

$$1 = \exp(0) = \exp(f_Q - f_Q) = \exp\left(\fint_Q (f_Q - f(x))\,dx\right) \leqslant \fint_Q \exp(f_Q - f(x))\,dx$$

and thus

$$\begin{aligned}
\fint_Q \exp(f(x) - f_Q)\,dx &\leqslant [w]_{A_2}\left(\fint_Q \exp(f_Q - f(x))\,dx\right)^{-1} \\
&\leqslant [w]_{A_2}\left(\exp\left(\fint_Q (f_Q - f(x))\,dx\right)\right)^{-1} = [w]_{A_2}.
\end{aligned}$$

Similarly

$$1 = \exp(0) = \exp\left(\fint_Q (f(x) - f_Q)\,dx\right) \leqslant \fint_Q \exp(f(x) - f_Q)\,dx$$

and

$$\begin{aligned}
\fint_Q \exp(f_Q - f(x))\,dx &\leqslant [w]_{A_2}\left(\fint_Q \exp(f(x) - f_Q)\,dx\right)^{-1} \\
&\leqslant [w]_{A_2}\left(\exp\left(\fint_Q (f(x) - f_Q)\,dx\right)\right)^{-1} = [w]_{A_2}.
\end{aligned}$$

Thus

$$\fint_Q \exp|f(x)-f_Q|\,dx = \fint_Q \max\{\exp(f(x)-f_Q), \exp(f_Q-f(x))\}\,dx$$
$$\leqslant \fint_Q \exp(f(x)-f_Q)\,dx + \fint_Q \exp(f_Q-f(x))\,dx \leqslant 2[w]_{A_2}$$

By Jensen's inequality

$$\exp\left(\fint_Q |f(x)-f_Q|\,dx\right) \leqslant \fint_Q \exp|f(x)-f_Q|\,dx \leqslant 2[w]_{A_2}$$

and thus

$$\fint_Q |f(x)-f_Q|\,dx \leqslant \log(2[w]_{A_2}).$$

This shows that $f \in \mathrm{BMO}$ whenever $1 \leqslant p \leqslant 2$.

If $p > 2$, then $p' < 2$ Theorem 4.18 (2) and (1) we have $w^{1-p'} \in A_{p'} \subset A_2$. The claim follows from this as above.

$\boxed{(2)}$ Assume that $f \in \mathrm{BMO}$ and first consider the case $p \geqslant 2$. We may assume that $\|f\|_* > 0$. By Theorem 3.30 there exists a constant $c = c(n) > 0$ such that

$$\fint_Q \exp\left(\frac{|f(x)-f_Q|}{c\|f\|_*}\right) dx \leqslant c$$

for every cube $Q \subset \mathbb{R}^n$. Let $Q \subset \mathbb{R}^n$ be a cube and let $g = \delta f$ with $\delta = \frac{1}{c\|f\|_*}$. Then

$$\fint_Q \exp(|g(x)-g_Q|)\,dx = \fint_Q \exp(\delta|f(x)-f_Q|)\,dx \leqslant c.$$

Thus

$$\fint_Q \exp(g(x)-g_Q)\,dx \fint_Q \exp(g_Q-g(x))\,dx \leqslant c^2,$$

and observing that $\exp(g_Q)\exp(-g_Q) = 1$ gives

$$\fint_Q \exp(g(x))\,dx \fint_Q \exp(-g(x))\,dx \leqslant c^2.$$

This is the A_2 condition for $\exp(g)$ on the cube Q. The claim follows by taking supremum over all cubes $Q \subset \mathbb{R}^n$ and observing that $A_2 \subset A_p$ for $p \geqslant 2$. If $1 \leqslant p < 2$, then $p' > 2$ and by the above argument $\exp(\delta f) \in A_2$ for small enough $\delta > 0$. Theorem 4.18 (2) implies that $\exp(-\delta f) \in A_2 \subset A_{p'}$. Theorem 4.18 (2) implies $e^{\delta(p-1)f} \in A_p$. □

Remarks 5.8:

(1) The proof shows that $\mathrm{BMO} = \{a \log w : a \geqslant 0, w \in A_2\}$.

(2) The parameter δ is necessary in Theorem 5.7. Indeed, we have $-n\log|x| \in \mathrm{BMO}(\mathbb{R}^n)$ but $|x|^{-n} = \exp(-n\log|x|) \notin A_2$.

We define a class of functions of bounded lower oscillation, which is closely related to BMO. Since $\operatorname{ess\,inf}_{x\in Q} f(x) \leqslant f(y)$ for almost every $y \in Q$ we have

$$\fint_Q |f(y) - \operatorname*{ess\,inf}_{x\in Q} f(x)|\,dx = f_Q - \operatorname*{ess\,inf}_{x\in Q} f(x).$$

Definition 5.9. We say that f of bounded lower oscillation, denoted $f \in \mathrm{BLO}$, if there exists constant c such that

$$f_Q - \operatorname*{ess\,inf}_{x\in Q} f(x) \leqslant c$$

for every cube Q in $\mathbb{R}^n$.

T H E M O R A L : In the definition of BLO we compare the mean oscillation to the essential infimum instead of the mean value as in the definition of BMO.

Remarks 5.10:

(1) Lemma 3.2 implies $\mathrm{BLO} \subset \mathrm{BMO}$.

(2) BLO is not a vector space, but it is closed under adding and multiplication by a nonnegative factor. In fact, we have $\mathrm{BLO} \cap -\mathrm{BLO} = L^\infty(\mathbb{R}^n)$. Indeed, if $f \in \mathrm{BLO}$ and $-f \in \mathrm{BLO}$, then simultaneously

$$f_Q - \operatorname*{ess\,inf}_{x\in Q} f(x) \leqslant c \quad \text{and} \quad -f_Q + \operatorname*{ess\,sup}_{x\in Q} f(x) \leqslant c$$

for every cube Q in $\mathbb{R}^n$. By adding up, we obtain

$$\operatorname*{ess\,sup}_{x\in Q} f(x) - \operatorname*{ess\,inf}_{x\in Q} f(x) \leqslant c$$

for every cube Q in $\mathbb{R}^n$, which implies $f \in L^\infty(\mathbb{R}^n)$.

The class BLO is connected to A_1 in the same way as BMO is connected to A_p with $p > 1$ in Theorem 5.7.

Theorem 5.11. (1) If $w \in A_1$, then $\log w \in \mathrm{BLO}$.

(2) Conversely, if $f \in \mathrm{BLO}$, then $\exp(\delta f) \in A_1$ for $\delta > 0$ small enough.

T H E M O R A L : This gives the following characterization for BLO:

$$\mathrm{BLO} = \{a \log w : a \geqslant 0, w \in A_1\}.$$

Proof. $\boxed{(1)}$ Assume that $w \in A_1$ and let Q be a cube in $\mathbb{R}^n$. Then

$$\frac{w(Q)}{|Q|} \leqslant [w]_{A_1} w(x) \quad \text{for almost every} \quad x \in Q.$$

This is equivalent with

$$\operatorname*{ess\,sup}_{x\in Q}(w(x)^{-1}) \leqslant [w]_{A_1} \frac{|Q|}{w(Q)}. \tag{5.12}$$

We denote $\log w = f$ and thus $w = \exp(f)$. Since

$$\operatorname*{ess\,sup}_{x\in Q}(\exp(-f(x))) = \exp(-\operatorname*{ess\,inf}_{x\in Q} f(x)),$$

(5.12) can be written as

$$\exp(-\operatorname*{ess\,inf}_{x\in Q} f(x)) \leqslant [w]_{A_1}\frac{|Q|}{w(Q)}. \tag{5.13}$$

Jensen's inequality together with (5.13) gives

$$\begin{aligned}\exp(f_Q - \operatorname*{ess\,inf}_{x\in Q} f(x)) &= \exp(f_Q)\exp(-\operatorname*{ess\,inf}_{x\in Q} f(x))\\ &\leqslant [w]_{A_1}\fint_Q \exp(f(x))\,dx\frac{|Q|}{w(Q)} = [w]_{A_1}.\end{aligned}$$

Thus

$$f_Q - \operatorname*{ess\,inf}_{x\in Q} f(x) \leqslant [w]_{A_1},$$

and $f \in \mathrm{BLO}$.

$\boxed{(2)}$ Assume that $f \in \mathrm{BLO}$. If we denote $a = \operatorname{ess\,inf}_{x\in Q} f(x)$, then by the definition of BLO we have

$$f_Q \leqslant c + a. \tag{5.14}$$

If $0 < \delta < \frac{c_2}{\|f\|_*}$, where $c_2 = \frac{1}{2^n e}$, then by the proof of Theorem 3.30, we have

$$\fint_Q \exp(\delta(f(x) - f_Q))\,dx \leqslant \fint_Q \exp(\delta|f(x) - f_Q|)\,dx \leqslant c',$$

where c' is independent of Q. This implies

$$\fint_Q \exp(\delta f(x))\,dx \leqslant c'\exp(\delta f_Q).$$

By (5.14) we arrive at

$$\begin{aligned}\fint_Q \exp(\delta f(x))\,dx &\leqslant c'\exp(\delta f_Q) \leqslant c'\exp(\delta(c+a))\\ &= c'\exp(\delta c)\exp(\delta a) = c'\exp(\delta c)\operatorname*{ess\,inf}_{x\in Q}(\exp(\delta f(x))),\end{aligned}$$

which shows that $\exp(\delta f) \in A_1$. □

Remark 5.15. By Remark 5.4 we have $|x|^\alpha \in A_1$, if $-n < \alpha \neq 0$. See also Example 4.17. Thus by Theorem 5.11 (1), we have $\log|x| \in \mathrm{BMO}$. This is a way to avoid the direct computation in Example 3.5 (2).

Next we clarify the connection between BLO and BMO.

Lemma 5.16. Every function in BMO can be represented as a difference of two functions in BLO.

Proof. Assume that $f \in \mathrm{BMO}$. By Theorem 5.7 we have $f = a \log w$, where $w \in A_2$ and $a \geqslant 0$. By Theorem 5.5 every A_2 weight w can be represented as $w = w_1 w_2^{-1}$, where $w_1, w_2 \in A_1$. Thus

$$f = a \log w = a \log(w_1 w_2^{-1}) = a \log w_1 - a \log w_2.$$

By Theorem 5.11 we have $\log w_1, \log w_2 \in \mathrm{BLO}$ and $a \geqslant 0$. This completes the proof. □

Now we are ready for a maximal function characterization of BMO.

Theorem 5.17. Every function f in BMO can be represented as

$$f = \alpha \log Mg - \beta \log Mh + b,$$

where $g, h \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, $b \in L^\infty(\mathbb{R}^n)$ and $\alpha, \beta > 0$.

THE MORAL: Every function in BMO can be represented as a difference logarithms of maximal functions plus a bounded function. Moreover, this gives a useful tool to construct BMO functions.

Proof. By Lemma 5.16

$$f(x) = a \log w_1(x) - a \log w_2(x),$$

where $w_1, w_2 \in A_1$. By Theorem 5.1

$$w_1(x) = b_1(x)(Mg(x))^{\delta_1} \quad \text{and} \quad w_2(x) = b_2(x)(Mh(x))^{\delta_2}, \tag{5.18}$$

where $g, h \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, $\delta_1, \delta_2 \in (0,1)$ and $b_1, b_2, \frac{1}{b_1}, \frac{1}{b_2} \in L^\infty(\mathbb{R}^n)$. Thus

$$\begin{aligned} f(x) &= a \log w_1(x) - a \log w_2(x) \\ &= a \log(b_1(x)(Mg(x))^{\delta_1}) - a \log(b_2(x)(Mh(x))^{\delta_2}) \\ &= \alpha \log Mg(x) - \beta \log Mh(x) + b(x), \end{aligned}$$

where $\alpha = \delta_1 a > 0$, $\beta = \delta_2 a > 0$ and $b = a(\ln b_1 - \ln b_2) \in L^\infty(\mathbb{R}^n)$. □

5.3 Maximal functions and BMO

The main result of this section is a boundedness result for a maximal operator on $\mathrm{BMO}(\mathbb{R}^n)$. Using Theorem 5.7, which was based on the John-Nirenberg inequality, and the Coifman-Rochberg lemma, see Theorem 5.1, we obtain the following boundedness result for the maximal operator on $\mathrm{BMO}(\mathbb{R}^n)$. The proof is based on [4]. This result was first proved by Bennett, DeVore and Sharpley in [1], see also [2, Theorem 7.18].

Theorem 5.19 (Bennett, DeVore and Sharpley (1981)). Let $f \in \mathrm{BMO}(\mathbb{R}^n)$ and assume that $Mf < \infty$ for almost every $x \in \mathbb{R}^n$. Then $Mf \in \mathrm{BMO}(\mathbb{R}^n)$ and there exists a constant $c = c(n)$ such that

$$\|Mf\|_* \leqslant c\|f\|_*.$$

Proof. We may assume that $\|f\|_* > 0$, and by Remark 3.4, we may also assume that $f \geqslant 0$. By Theorem 3.30 there exists a constant $c_1 = c_1(n) > 0$ such that

$$\fint_Q \exp\left(\frac{|f(x)-f_Q|}{c_1\|f\|_*}\right) dx \leqslant c_1 \tag{5.20}$$

for every cube $Q \subset \mathbb{R}^n$. Let $Q \subset \mathbb{R}^n$ be a cube and let $g = \delta f$ with $\delta = \frac{1}{c_1\|f\|_*}$. By (5.20), we have

$$\begin{aligned}\exp(-g_Q)\fint_Q \exp(g(x))\,dx &= \fint_Q \exp(g(x)-g_Q)\,dx \leqslant \fint_Q \exp(|g(x)-g_Q|)\,dx\\ &= \fint_Q \exp(\delta|f(x)-f_Q|)\,dx \leqslant c_1,\end{aligned}$$

which implies

$$\fint_Q \exp(g(x))\,dx \leqslant c_1 \exp(g_Q).$$

By Jensen's inequality,

$$\exp(g_Q) = \exp\left(\fint_Q g(x)\,dx\right) \leqslant \fint_Q \exp(g(x))\,dx \leqslant c_1 \exp(g_Q).$$

This holds for all cubes $Q \subset \mathbb{R}^n$, and thus

$$\exp(Mg(x)) \leqslant M(\exp g)(x) \leqslant c_1 \exp(Mg(x)) \tag{5.21}$$

for every $x \in \mathbb{R}^n$. Since $\exp g \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and $Mg(x) = \delta Mf(x) < \infty$ for almost every $x \in \mathbb{R}^n$, Theorem 5.1 shows that

$$w = \big(M(\exp g)\big)^{\frac{1}{2}}$$

is an A_1 weight with $[w]_{A_1} \leqslant c_2 = c_2(n)$. Let

$$v = (\exp(Mg))^{\frac{1}{2}} = \exp\big(\tfrac{1}{2}Mg\big).$$

The inequalities in (5.21) imply $v(x) \leqslant w(x) \leqslant c_1^{\frac{1}{2}} v(x)$ for every $x \in \mathbb{R}^n$ and thus

$$\begin{aligned}\fint_Q v(x)\,dx &\leqslant \fint_Q w(x)\,dx \leqslant [w]_{A_1} \operatorname*{ess\,inf}_{x\in Q} w(x)\\ &\leqslant c_1^{\frac{1}{2}}[w]_{A_1} \operatorname*{ess\,inf}_{x\in Q} v(x) \leqslant c(c_1,c_2) \operatorname*{ess\,inf}_{y\in Q} v(y)\end{aligned}$$

for every cube $Q \subset \mathbb{R}^n$. Thus v is an A_1 weight and $[v]_{A_1} \leqslant c(c_1,c_2) = C(n)$. Using $v = \exp(\frac{1}{2}Mg)$, Theorem 5.7 and the fact that $v \in A_2$ with $[v]_{A_2} \leqslant [v]_{A_1}$, we obtain

$$\frac{\|Mg\|_*}{2} = \|\log v\|_* \leqslant \log(2[v]_{A_2}) \leqslant \log(2[v]_{A_1}) \leqslant c(n).$$

Since $g = \delta f$ with $\delta = \frac{1}{c_1 \|f\|_*}$, we have

$$\|Mf\|_* = \frac{\|Mg\|_*}{\delta} \leqslant \frac{2c(n)}{\delta} = c(n)\|f\|_*. \qquad \square$$

THE END

Bibliography

[1] C. Bennett, R.A. DeVore and R. Sharpley, Weak L^∞ and BMO, Ann. of Math. (2) 113 (1981), 601–611.

[2] C. Bennett and R. Sharpley, Interpolation of Operators, Academic Press, 1988.

[3] O. Blasco, Dyadic BMO, paraproducts and Haar multipliers. Interpolation theory and applications, 11–18, Contemp. Math., 445, Amer. Math. Soc., 2007.

[4] F. Chiarenza and M. Frasca, Morrey spaces and Hardy-Littlewood maximal function. Rend. Mat. Appl. 7 (1988), 273–279.

[5] R. R. Coifman, P. W. Jones and J. L. Rubio de Francia, Constructive decomposition of BMO functions and factorization of A_p weights, Proc. Amer. Math. Soc. 87 (1983), 665–666.

[6] R. R. Coifman and R. Rochberg, Another Characterization of BMO, Proc. Amer. Math. Soc. 79 (1980), 249–254.

[7] J. Duoandikoetxea, Fourier Analysis, American Mathematical Society, 2001.

[8] J. Duoandikoetxea, Forty Years of Muckenhoupt Weights, Function Spaces and Inequalities, Lecture Notes Paseky nad Jizerou 2013, pages 23–75. Matfyzpress, Praga, 2013.

[9] J. García-Cuerva and J. L. Rubio de Francia, Weighted Norm Inequalities and Related Topics, North-Holland, 1985.

[10] L. Grafakos, Classical Fourier Analysis, Springer, 2008.

[11] L. Grafakos, Modern Fourier Analysis, Springer, 2008.

[12] P. W. Jones, Factorization of A_p weights, Ann. of Math. 111 (1980), 511–530.

[13] F. John and L. Nirenberg, On functions of bounded mean oscillation, Comm. Pure Appl. Math. 14 (1961), 391–413.

[14] A. Lerner, An elementary approach to several results on the Hardy-Littlewood maximal operator, Proc. Amer. Math. Soc. 136 (2008), 2829–2833.

[15] A. Lerner and F. Nazarov, Intuitive dyadic calculus: the basics, Expo. Math. 37 (2019), 225–265.

[16] S. Lu, Y. Ding and D. Yan, Singular Integrals and Related Topics, World Scientific, 2007.

[17] U. Neri, Some properties of functions of bounded mean oscillation, Studia Math. 61 (1977), 63–75.

[18] M. C. Pereyra, Lecture notes on dyadic harmonic analysis. Second Summer School in Analysis and Mathematical Physics (Cuernavaca, 2000), 1–60, Contemp. Math., 289, Amer. Math. Soc., 2001.

[19] B. Simon, Harmonic Analysis, A Comprehensive Course in Analysis, Part 3, American Mathematical Society, 2015.

[20] E. M. Stein, Singular Integrals and Differentiability Properties of Functions, Princeton University Press, 1970.

[21] E. M. Stein, Harmonic Analysis, Real Variable Methods, Orthogonality, and Oscillatory Integrals, Princeton University Press, 1993.

[22] E. M. Stein and G. Weiss, Fourier Analysis on Euclidean Spaces, Princeton University Press, 1970.

[23] A. Torchinsky, Real Variable Methods in Harmonic Analysis, Academic Press, 1986.