Weighted Inequalities in Analysis and Probability Theory

Adam Osękowski

CHAPTER 1

Preliminaries, some history and motivation

1.1. Notation and basic definitions. We will work primarily in $\mathbb{R}^d$ or on the subsets of this space. The norm in $\mathbb{R}^d$ will be denoted by $|\cdot|$ and the Lebesgue measure will be denoted by dx. For a measurable set $E \subset \mathbb{R}^d$, the symbol $|E|$ will stand for its Lebesgue measure and χ_E will be the associated indicator function. By a weight we will mean a nonnegative, locally integrable function w on $\mathbb{R}^d$. The induced measure on $\mathbb{R}^d$ will also be denoted by w, i.e., we define $w(E) = \int_E w \mathrm{d}x$ for any measurable set E. For $1 \leq p < \infty$, $L^p = L^p(\mathbb{R}^d, \mathrm{d}x)$ is the Banach function space with the norm

$$\|f\|_{L^p} = \left(\int_{\mathbb{R}^d} |f(x)|^p \mathrm{d}x \right)^{1/p},$$

with the usual extension to $p = \infty$: $\|f\|_{L^\infty(\mathbb{R}^d, dx)} = \operatorname{essup}|f|$. Analogous definitions hold in the case when $\mathbb{R}^d$ is replaced by a general measure space (X, μ). We will usually write $\|f\|_{L^p}$ and indicate the underlying measure space only if it is not clear from the context. One introduces the weighted versions of these spaces, by applying the integration with the underlying measure wdx. So, for $1 \leq p < \infty$, set

$$\|f\|_{L^p(w)} = \left(\int_{\mathbb{R}^d} |f(x)|^p w(x) \mathrm{d}x \right)^{1/p}, \qquad 1 \leq p < \infty,$$

with an analogous modification for $L^\infty(w)$.

Given $1 \leq p, q \leq \infty$ and a pair of weights (u, v), a bounded operator $T : L^p(\mathbb{R}^d, v) \to L^q(\mathbb{R}^d, u)$ is said to be of strong-type (p, q) (satisfy the strong type (p, q) inequality). This is equivalent to saying that there exists a finite constant C such that

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

for all $f \in L^p(v)$. If $u = v$ and $p = q$, T is said to be bounded on $L^p(u)$ and the infimum of the admissible constants C above is denoted by $\|T\|_{L^p(u)}$.

Given a locally integrable weight w, we define the weak space $L^{p,\infty}(w)$ to be the function space with the quasinorm given by

$$\|f\|_{L^{p,\infty}(w)} = \sup_{\lambda > 0} \lambda w\big(\{x \in \mathbb{R}^d : |f(x)| > \lambda\}\big)^{1/p}, \qquad 1 \leq p < \infty,$$

and we set $L^{\infty,\infty}(w) = L^\infty(w)$. An operator T is of weak type (p, q) if there is a finite constant C such that

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

for all $f \in L^p(v)$. This is equivalent to the existence of a constant $C < \infty$ such that for any $\lambda > 0$,

$$u\left(\{x \in \mathbb{R}^d : |Tf(x)| > \lambda\}\right) \leq C \left(\frac{1}{\lambda^p} \int_{\mathbb{R}^d} |f(x)|^p v(x) \mathrm{d}x \right)^{q/p}.$$

Recall the following classical statement, Marcinkiewicz interpolation theorem.

THEOREM 1.1. *Let (X,μ), (Y,ν) be two measure spaces. Suppose that T is a sublinear operator such that $M_i := \|T\|_{L^{p_i}(X,\mu)\to L^{p_i,\infty}(Y,\nu)} < \infty$, $i=0,1$, for some $1 \le p_0 < p_1 \le \infty$. Then $\|T\|_{L^p(X,\mu)\to L^p(Y,\nu)} < \infty$ for all $p \in (p_0,p_1)$.*

PROOF. Given $f \in L^p$ and $\lambda > 0$, we decompose f as $f = f_0 + f_1$, where

$$f_0 = f\chi_{\{|f|>c\lambda\}}, \qquad f_1 = f\chi_{\{|f|\le c\lambda\}}$$

and c will be specified below. Then $f_0 \in L^{p_0}$ and $f_1 \in L^{p_1}$, furthermore,

$$|Tf(x)| \le |Tf_0(x)| + |Tf_1(x)|,$$

so

$$\nu(\{|Tf > \lambda\}) \le \nu(\{Tf_0 > \lambda/2\}) + \nu(\{Tf_1 > \lambda/2\}). \tag{1.1}$$

We consider two cases separately.

The case $p_1 = \infty$. If we take $c = (2M_1)^{-1}$, then

$$\|Tf_1\|_{L^\infty} \le (2c)^{-1}\|f_1\|_{L^\infty} \le (2c)^{-1}\cdot c\lambda = \lambda/2$$

and hence $\nu(\{Tf_1 > \lambda/2\}) = 0$. Plugging this above and using the weak-type (p_0,p_0) inequality for T, we obtain

$$\begin{aligned}\nu(\{Tf > \lambda\}) \le \nu(\{Tf_0 > \lambda/2\}) &\le \left(\frac{2M_0}{\lambda}\right)^{p_0}\int_X |f_0|^{p_0}\mathrm{d}\mu \\ &= \left(\frac{2M_0}{\lambda}\right)^{p_0}\int_{\{|f|>c\lambda\}} |f|^{p_0}\mathrm{d}\mu.\end{aligned}$$

Multiplying both sides by $p\lambda^{p-1}$ and integrating over λ from 0 to ∞ yields

$$\|Tf\|_{L^p}^p \le A\int_X |f|^{p_0}\cdot|f|^{p-p_0}\mathrm{d}\mu = A\|f\|_{L^p}^p,$$

with

$$A = \frac{p}{p-p_0}(2M_0)^{p_0}(2M_1)^{p-p_0}.$$

The case $p_1 < \infty$. We have the estimates

$$\nu(\{Tf_i > \lambda/2\}) \le \left(\frac{2M_i}{\lambda}\|f_i\|_{L^{p_i}}\right)^{p_i}$$

for $i = 0, 1$. Plugging them into (1.1), multiplying both sides by $p\lambda^{p-1}$ and integrating as previously, we get

$$\|Tf\|_{L^p}^p \le \left(\frac{p2^{p_0}}{p-p_0}\frac{M_0^{p_0}}{c^{p-p_0}} + \frac{p2^{p_1}}{p_1-p}\frac{M_1^{p_1}}{c^{p-p_1}}\right)\|f\|_{L^p}^p.$$

Taking c such that $(2M_0c)^{p_0} = (2M_1c)^{p_1}$, the inequality simplifies to

$$\|Tf\|_{L^p} \le 2p^{1/p}\left(\frac{1}{p-p_0} + \frac{1}{p_1-p}\right)^{1/p} M_0^{1-\theta}M_1^{\theta}\|f\|_{L^p},$$

where $\theta \in (0,1)$ is uniquely determined by the equality

$$\frac{1}{p} = \frac{\theta}{p_1} + \frac{1-\theta}{p_0}.$$

□

1.2. Some history. The basic problems in the study of weighted norm inequalities are the strong- and weak-type estimates of the form

$$\int_{\mathbb{R}^d} |Tf(x)|^p u(x)\mathrm{d}x \leq C \int_{\mathbb{R}^d} |f(x)|^p v(x)\mathrm{d}x \tag{1.2}$$

or

$$u\left(\{x \in \mathbb{R}^d : |Tf(x)| > \lambda\}\right) \leq \frac{C}{\lambda^p} \int_{\mathbb{R}^d} |f(x)|^p v(x)\mathrm{d}x. \tag{1.3}$$

Here $1 \leq p < \infty$ and T is an operator, typically one of the classical operators of harmonic analysis: a maximal function, singular integral, square function, etc. Sometimes the spaces $\mathbb{R}^d$ are localized to some given cube Q (the most frequent choice is $[0,1)^d$), or are replaced by an arbitrary probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then there is an additional important choice for T: the martingale transform operator. The above problems are divided into two classes: the one-weight case, when we have just a single weight (that is, $u = v = w$) and the two-weight case, when we work with a pair (u, v) of weights. In these notes, we will focus mainly on the one-weight case, however, at some points we will present arguments from the two-weight setting.

The theory of one-weight inequalities started with the study of power weights $w(x) = |x|^\alpha$. The investigation of the general case (with an arbitrary w) goes back to the seminal works of Muckenhoupt from the seventies: during the study of (1.2) and (1.3) for the Hardy-Littlewood maximal function $\mathcal{M}$, Muckenhoupt introduced the class of the so-called A_p weights. For a given $1 < p < \infty$, a weight w satisfies the condtion A_p if

$$[w]_{A_p} := \left(\frac{1}{|Q|}\int_Q w\right)\left(\frac{1}{|Q|}\int_Q w^{1/(1-p)}\right)^{p-1} < \infty,$$

where the supremum is taken over the class of all cubes in $\mathbb{R}^d$ with sides parallel to the axes; there is a version for $p = 1$, which can be obtained directly by a simple limiting argument from the above case (see below). Muckenhoupt proved that the desired weighted strong- or weak-type (p,p) estimate for $\mathcal{M}$ holds if and only if $w \in A_p$. Soon after the appearance of the A_p weights it was realized that the condition characterizes the boundedness of much wider class of operators. At the beginning of eighties, Hunt, Muckenhoupt and Wheeden proved that the Hilbert transform

$$Hf(x) = \frac{1}{\pi}\text{p.v.}\int_{\mathbb{R}} \frac{f(y)}{x-y}\mathrm{d}y$$

satisfies the strong- or weak-type (p,p) estimate iff $w \in A_p$, $1 < p < \infty$. This statement was further generalized by Coifman and Fefferman to the class of any singular integral operators with sufficiently smooth kernel. Since then, (1.2) and (1.3) have been studied in numerous other contexts, there are lots of works which focus on the size of constants involved and aiming at the development of new methods.

Of course, the one-weight setting is a sub-case of the two-weight setting, and the latter case turns out to be much more challenging in general. Sawyer presented a successful approach to the characterization of pairs (u, v) satisfying (1.2) and (1.3) for the Hardy-Littlewood maximal operator. There are also many results for the so-called fractional operators, but the whole picture for general classes of operators (e.g., singular) is very far from being satisfactory. The best results known to date refer to the so-called testing conditions or provide only sufficient conditions for weights which guarantee the validity of strong and weak-type bounds.

1.3. Motivation: unweighted weak and strong type estimates for maximal operators. Let $d \geq 1$ be a fixed positive integer (dimension). We start with defining three types of maximal operators.

DEFINITION 1.2. The *uncentered* maximal operator $\mathcal{M}$ acts on locally integrable functions $f : \mathbb{R}^d \to \mathbb{R}$ by the formula

$$\mathcal{M}f(x) = \sup \frac{1}{|Q|} \int_Q |f(y)| \mathrm{d}y,$$

where the supremum is taken over all cubes $Q \subset \mathbb{R}^d$ containing x, with sides parallel to the axes.

DEFINITION 1.3. The *centered* maximal operator $\mathcal{M}_c$ acts on locally integrable functions $f : \mathbb{R}^d \to \mathbb{R}$ by the formula

$$\mathcal{M}_c f(x) = \sup \frac{1}{|Q|} \int_Q |f(y)| \mathrm{d}y,$$

where the supremum is taken over all cubes $Q \subset \mathbb{R}^d$ of center x, with sides parallel to the axes.

It is easy to see that both these objects are comparable, up to a multiplicative constant depending on the dimension d. Indeed, we obviously have $\mathcal{M}_c f \leq \mathcal{M}f$ for any f. Furthermore, for an arbitrary cube Q with $x \in Q$ and $|Q| = a^d$, we have $Q \subset [x-a, x+a]^d$ and consequently

$$\begin{aligned}\frac{1}{|Q|} \int_Q |f(y)| \mathrm{d}y &\leq \frac{1}{a^d} \int_{[x-a,x+a]^d} |f(y)| \mathrm{d}y \\ &\leq 2^d \cdot \frac{1}{|[x-a, x+a]^d|} \int_{[x-a,x+a]^d} |f(y)| \mathrm{d}y \leq \mathcal{M}_c f(x),\end{aligned}$$

so taking the supremum over all Q as above yields $\mathcal{M}f \leq 2^d \mathcal{M}_c f$.

Now we introduce the third type of the maximal function, the so-called *dyadic* maximal operator. To this end, we need to define the (standard) dyadic lattice $\mathfrak{D}$ in $\mathbb{R}^d$. It consists of all cubes of the form $[0, 2^k)^d + a \cdot 2^k$, where $k \in \mathbb{Z}$ and $a \in \mathbb{Z}^d$. It is easy to check that for any Q_1, $Q_2 \in \mathfrak{D}$, the cubes are either disjoint, or one is contained in the other.

DEFINITION 1.4. The *dyadic* maximal operator M acts on locally integrable functions $f : \mathbb{R}^d \to \mathbb{R}$ by the formula

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

where the supremum is taken over dyadic cubes Q (i.e., all $Q \in \mathfrak{D}$) containing x.

We can also define the version of M localized to a given and fixed cube $\mathfrak{Q}$. For such a cube, we define the natural dyadic lattice $\mathfrak{D}(\mathfrak{Q})$ (simply by bisecting the sides of $\mathfrak{Q}$). Then $M^{\mathfrak{Q}}$ acts of integrable functions $f : \mathfrak{Q} \to \mathbb{R}$ by the same formula as previously, however, in the supremum defining $Mf(x)$, only $Q \in \mathfrak{D}(\mathfrak{Q})$ containing x are considered. Actually, at many places in these notes we will be dealing with localized dyadic operators; the base cube $\mathfrak{Q}$ will always be given and fixed (actually, by standard dilation and translation argument, we can always assume that $\mathfrak{Q} = [0,1)^d$), and the superscript $\mathfrak{Q}$ will typically be omitted. There are several reasons for this distinction: the encoded dyadic lattice enables the use of additional and very efficient combinatorial arguments; second, the study of many classical operators in analysis (e.g. singular or fractional) is based on discretization (approximation) in terms of appropriate dyadic-type counterparts, whose behavior can be controlled by the dyadic maximal operator; third, the dyadic maximal function and other maximal operators defined above behave similarly, up to multiplicative constants. Finally, such objects are closely related to the probability, especially to the theory of martingales. Let us be more specific. Suppose that $\mathfrak{Q} = [0,1)^d$ is a given base cube, equipped with its Borel subsets $\mathcal{F} = \mathcal{B}(\mathfrak{Q})$ and Lebesgue's measure $|\cdot|$. The triple $(\mathfrak{Q}, \mathcal{F}, |\cdot|)$ is a probability space and it is equipped with the natural dyadic filtration $(\mathcal{F}_n)_{n\geq 0}$, where $\mathcal{F}_n$ is the σ-algebra generated by the dyadic subsets of $\mathfrak{Q}$ of measure 2^{-nd}. Any integrable function (random variable) f on $\mathfrak{Q}$ gives rise to the associated uniformly integrable martingale $f_n = \mathbb{E}(f|\mathcal{F}_n)$, $n = 0, 1, 2, \ldots$, which will also be denoted by f. For any such martingale, we define its maximal function $f^* = \sup_{n\geq 0} f_n$. The connection between this object and the dyadic maximal operator is given by the identity $Mf = |f|^*$.

Now we will prove two important inequalities for M, which serve as a nice motivation and will play an important role in our further analysis. The first result is the weak type (p,p) estimate for M.

THEOREM 1.5. *For any $1 \leq p < \infty$, any $\lambda > 0$ and any function $f \in L^p(\mathbb{R}^d)$ we have*

$$\lambda^p |\{x \in \mathbb{R}^d : Mf(x) > \lambda\}| \leq \int_{\{x\in\mathbb{R}^d : Mf(x)>\lambda\}} |f(y)|^p \, dy \leq \|f\|^p_{L^p(\mathbb{R}^d)}. \tag{1.4}$$

PROOF. Suppose that $Mf(x) > \lambda$. Then by the definition of M, there exists Q containing x such that

$$\lambda < \frac{1}{|Q|}\int_Q |f(y)|\mathrm{d}y \leq \left(\frac{1}{|Q|}\int_Q |f(y)|^p \mathrm{d}y\right)^{1/p},$$

and this inequality implies that actually all points from Q are contained in the set $\{Mf > \lambda\}$. Therefore, this set can be written as a union $\bigcup_j Q_j$ of disjoint dyadic cubes each of which satisfies the above inequality. Consequently,

$$\lambda^p |\{Mf > \lambda\}| = \lambda^p \sum_j |Q_j| \leq \sum_j \int_{Q_j} |f(y)|^p \leq \int_{\{Mf>\lambda\}} |f(y)|^p \mathrm{d}y,$$

which is the claim. $\square$

THEOREM 1.6. *For any $1 < p \leq \infty$ we have*

$$\|Mf\|_{L^p(\mathbb{R}^d)} \leq C_p \|f\|_{L^p(\mathbb{R}^d)}.$$

PROOF. If $p = \infty$, then the claim holds trivially with $C_p = 1$. For other p, the claim follows at once from the weak-type $(1,1)$ estimate proved above and Marcinkiewicz interpolation theorem. $\square$

CHAPTER 2

Weighted inequalities for maximal operator

Now we will characterize those nonnegative and locally integrable functions w on $\mathbb{R}^d$ such that the dyadic maximal operator on $\mathbb{R}^d$ is bounded as an operator from $L^p(w)$ to $L^{p,\infty}(w)$ or as an operator on $L^p(w)$. That is, we consider the following two questions:

(i) Let $1 \leq p < \infty$. Characterize those $w \in L^+_{loc}(\mathbb{R}^d)$ for which there is a finite constant $c = c_{p,w}$ such that

$$\lambda^p w(\{x : Mf(x) > \lambda\}) \leq c_{p,w} \|f\|^p_{L^p(w)}$$

for all λ and all locally integrable f on $\mathbb{R}^d$.

(ii) Let $1 < p < \infty$. Characterize those $w \in L^+_{loc}(\mathbb{R}^d)$ for which there is a finite constant $C = C_{p,w}$ such that

$$\|Mf\|_{L^p(w)} \leq C_{p,w} \|f\|_{L^p(w)}$$

for all locally integrable f on $\mathbb{R}^d$.

2.1. A_p **condition.** Our starting point is to assume that w satisfies the weak-type (p,p) inequality

$$\lambda^p w(\{x : Mf(x) > \lambda\}) \leq c_{p,w} \int_{\mathbb{R}^d} |f(y)|^p w(y) \mathrm{d}y \tag{2.1}$$

and deduce a certain necessary condition. Fix an arbitrary measurable set $S \subset \mathbb{R}^d$ with $|S| > 0$ and pick a dyadic cube Q such that $|Q \cap S| > 0$; we may actually assume that $S \subset Q$, replacing S with $Q \cap S$ if necessary. Setting $f = \chi_S$ and $\lambda < |S|/|Q|$, we see that on Q we have $Mf \geq \frac{1}{|Q|} \int_Q |f| \mathrm{d}y = |S|/|Q| > \lambda$ and hence, plugging this to the weak-type bound above, we get $\lambda^p w(Q) \leq c_{p,w} w(S)$ and therefore, letting $\lambda \to |S|/|Q|$,

$$w(Q) \left(\frac{|S|}{|Q|} \right)^p \leq c_{p,w} w(S). \tag{2.2}$$

This immediately implies that on each quadrant K of $\mathbb{R}^d$, the weight w is either identically zero or $w > 0$ almost everywhere. Indeed, if $w = 0$ on some set $S \subset K$ of positive measure (we may assume that it is bounded), then necessarily $w(Q) = 0$ for any $Q \subset K$ which contains S. Now we consider separately two cases.

1° The case $p = 1$. The inequality (2.2) becomes

$$\frac{w(Q)}{|Q|} \leq c_{1,w} \frac{w(S)}{|S|}.$$

If we put $a = \operatorname{essinf}_Q w$, then for any $\varepsilon > 0$ there is a measurable subset $S_\varepsilon \subset Q$ such that $|S_\varepsilon| > 0$ and $w(x) \le a + \varepsilon$ for any $x \in S_\varepsilon$. Therefore, for any $\varepsilon > 0$ we have

$$\frac{w(Q)}{|Q|} \le c_{1,w}(a+\varepsilon),$$

and hence letting $\varepsilon \to 0$ yields

$$\frac{1}{|Q|}\int_Q w \le c_{1,w} \operatorname*{essinf}_Q w \qquad \text{for any dyadic cube } Q.$$

This can be rewritten equivalently in the form

$$\frac{1}{|Q|}\int_Q w \le c_{1,w} w(x) \qquad \text{for almost all } x \in Q. \tag{2.3}$$

This is the so-called A_1 condition, and the weights which satisfy it will be referred to as A_1 weights. Furthermore, the smallest $c_{1,w}$ allowed above will be denoted by $[w]_{A_1}$ and called the A_1 characteristic of w. It is not difficult to see that $[w]_{A_1} \ge 1$ and the equality holds only for constant weights.

Let us also note here that the A_1 condition is equivalent to saying that

$$Mw(x) \le c_{1,w} w(x) \qquad \text{for almost all } x \in \mathbb{R}^d. \tag{2.4}$$

This is an easy exercise which is left to the reader. Let us rewrite the definition of the characteristic $[w]_{A_1}$:

$$[w]_{A_1} = \operatorname{essup} \frac{Mw(x)}{w(x)}.$$

2° The case $p > 1$. Plug in (2.1) the choice $f = \min\{w^{1-p'}, N\}\chi_Q$ and $\lambda < \int_Q f/|Q|$. On Q we have $Mf \ge \int_Q f/|Q| > \lambda$, so we obtain $\lambda^p w(Q) \le c_{p,w} \int_Q f^p w$. Letting $\lambda \to \int_Q f/|Q|$, the estimate becomes

$$\left(\frac{1}{|Q|}\int_Q w\right)\left(\frac{1}{|Q|}\int_Q f\right)^{p-1} \int_Q f \le c_{p,w} \int_Q f^p w.$$

But $f \le w^{1-p'}$, so $f^{p-1} \le w^{-1}$ and hence $f^p w \le f$. Consequently, the above bound yields

$$\left(\frac{1}{|Q|}\int_Q w\right)\left(\frac{1}{|Q|}\int_Q \min\{w^{1-p'}, N\}\right)^{p-1} \le c_{p,w}.$$

Letting $N \to \infty$ and applying Lebesgue's monotone convergence theorem, we get

$$\left(\frac{1}{|Q|}\int_Q w\right)\left(\frac{1}{|Q|}\int_Q w^{1-p'}\right)^{p-1} \le c_{p,w}. \tag{2.5}$$

This property is called the A_p condition, the weights satisfying it are referred to as A_p weights. The quantity

$$[w]_{A_p} = \sup_Q \left(\frac{1}{|Q|}\int_Q w\right)\left(\frac{1}{|Q|}\int_Q w^{1-p'}\right)^{p-1}$$

is called the A_p characteristic of w. Note that by Jensen's inequality we have $[w]_{A_p} \ge 1$, with equality for constant weights.

REMARK 2.1. The following geometrical-probabilistic interpretation of A_p condition is useful in many situations. Suppose that the underlying base cube is $[0,1)^d$ and let us treat it as a probability space (with the Borel sets as σ-algebra and Lebesgue measure as the probability), equipped with the dyadic filtration $(\mathcal{F}_n)_{n\geq 0}$. Assume first that $1<p<\infty$ and let $w\in A_p$ be a fixed weight. Let $w=(w_n)_{n\geq 0}$, $v=(v_n)_{n\geq 0}$ be martingales generated by w and $v=w^{1-p'}$ (i.e., set $w_n=\mathbb{E}(w|\mathcal{F}_n)$, $v_n=\mathbb{E}(w^{1-p'}|\mathcal{F}_n)$ for each $n\geq 0$). Then it is easy to see that

$$1\leq w_n v_n^{p-1}\leq [w]_{A_p}$$

almost surely for each n. Furthermore, we have the almost sure convergence $(w_n,v_n)\to(w,v)=(w,w^{1-p'})$. And conversely, any pair (w,v) of martingales, taking values in the set

$$\Omega_{p,c}=\{(x,y)\in(0,\infty)^2:1\leq xy^{p-1}\leq c\}$$

and terminating at the lower boundary $xy^{p-1}=1$, gives rise to the A_p weight w (equal to the first coordinate of the limit random variable) satisfying $[w]_{A_p}\leq c$.

In the case $p=1$ the analysis is somewhat different: given an A_1 weight w, we define the associated martingale $w=(w_n)_{n\geq 0}$ as above and let $v_n=\max_{0\leq k\leq n} w_k$ be the (partial) maximal function of w. Then the A_1 condition is equivalent to saying that

$$w_n\leq v_n\leq [w]_{A_1}w_n$$

almost surely for each n.

2.2. Estimates for M. Now we show that A_p conditions are also sufficient for the weak-type bounds.

THEOREM 2.2. *Let $1\leq p<\infty$. If a weight w satisfies the (dyadic) A_p condition, then* (2.1) *holds true with $c_{p,w}=[w]_{A_p}$.*

PROOF. Fix $f\in L^p(w)$ and $\lambda>0$. Arguing as in the unweighted case, the set $\{Mf>\lambda\}$ can be written as a union $\bigcup_j Q_j$ of pairwise disjoint dyadic cubes such that $\frac{1}{|Q_j|}\int_{Q_j}|f|\mathrm{d}y>\lambda$ for each j. If $p=1$, then we write

$$\int_{Q_j}|f|w\mathrm{d}x\geq\int_{Q_j}|f|\operatorname*{essinf}_{Q_j} w\mathrm{d}x\geq [w]_{A_1}^{-1}\int_{Q_j}|f|\mathrm{d}x\cdot\frac{w(Q_j)}{|Q_j|}\geq\lambda[w]_{A_1}^{-1}w(Q_j).$$

Summing over all j we get the desired weak-type bound. If $1<p<\infty$, then

$$\lambda^p w(Q_j)\leq [w]_{A_p}\cdot|Q_j|\cdot\frac{\left(\frac{1}{|Q_j|}\int_{Q_j}|f|\mathrm{d}x\right)^p}{\left(\frac{1}{|Q_j|}\int_{Q_j}w^{1-p'}\mathrm{d}x\right)^{p-1}}.$$

The function $B(x,y)=|x|^py^{1-p}$ is convex on $\mathbb{R}\times(0,\infty)$: its Hessian matrix

$$D^2B(x,y)=\begin{bmatrix} p(p-1)|x|^{p-2}y^{1-p} & p(1-p)|x|^{p-2}xy^{-p}\\ p(1-p)|x|^{p-2}xy^{-p} & p(p-1)|x|^py^{-1-p}\end{bmatrix}$$

is nonnegative-definite. Therefore, by Jensen's inequality, we have

$$B\left(\frac{1}{|Q_j|}\int_{Q_j}|f|\mathrm{d}x,\frac{1}{|Q_j|}\int_{Q_j}w^{1-p'}\mathrm{d}x\right)\leq\frac{1}{|Q_j|}\int_{Q_j}B(|f|,w^{1-p'})\mathrm{d}x,$$

which combined with the previous estimate yields

$$\begin{aligned}\lambda^p w(Q_j) &\le [w]_{A_p}\cdot\int_{Q_j} B(|f|, w^{1-p'})\mathrm{d}x = [w]_{A_p}\int_{Q_j}|f|^p w^{(1-p')(1-p)}\mathrm{d}x \\ &= [w]_{A_p}\int_{Q_j}|f|^p w\mathrm{d}x\end{aligned} \tag{2.6}$$

and it remains to sum over all j to get the assertion. □

We turn our attention to the following fundamental properties of A_p weights.

PROPOSITION 2.3.

(i) *We have $A_p \subset A_q$ if $1 \le p < q$; more precisely, we have $[w]_{A_q} \le [w]_{A_p}$.*
(ii) *If $1 < p < \infty$, then $w \in A_p$ if and only if $w^{1-p'} \in A_{p'}$. Furthermore, we have the equality $[w]_{A_p} = [w^{1-p'}]_{A_{p'}}^{p-1}$.*
(iii) *If w_0, $w_1 \in A_1$, then $w_0 w_1^{1-p} \in A_p$ and $[w_0 w_1^{1-p}]_{A_p} \le [w_0]_{A_1}[w_1]_{A_1}^{p-1}$.*

PROOF. (i) If $p = 1$, then for any dyadic cube Q,

$$\left(\frac{1}{|Q|}\int_Q w^{1-q'}\mathrm{d}x\right)^{q-1} \le \sup_{x\in Q} w(x)^{-1} = \left(\inf_{x\in Q} w(x)\right)^{-1} \le [w]_{A_1}\left(\frac{w(Q)}{|Q|}\right)^{-1},$$

which implies $[w]_{A_q} \le [w]_{A_1}$. If $p > 1$, the property follows at once from Hölder's inequality: we have

$$\frac{1}{|Q|}\int_Q w^{1-q'}\mathrm{d}x \le \left(\frac{1}{|Q|}\int_Q w^{1-p'}\mathrm{d}x\right)^{(1-p')/(1-q')},$$

or

$$\left(\frac{1}{|Q|}\int_Q w^{1-q'}\mathrm{d}x\right)^{q-1} \le \left(\frac{1}{|Q|}\int_Q w^{1-p'}\mathrm{d}x\right)^{p-1}.$$

(ii) This is straightforward: for any cube Q we have

$$\begin{aligned}&\left(\frac{1}{|Q|}\int_Q w\mathrm{d}x\right)\left(\frac{1}{|Q|}\int_Q w^{1-p'}\mathrm{d}x\right)^{p-1} \\ &\qquad = \left\{\left(\frac{1}{|Q|}\int_Q w^{1-p'}\mathrm{d}x\right)\left(\frac{1}{|Q|}\int_Q (w^{1-p'})^{1-(p')'}\right)^{p'-1}\right\}^{p-1},\end{aligned}$$

and taking the supremum over all Q yields the desired claim.

(iii) We must prove that

$$[w_0 w_1^{1-p}]_{A_p} = \sup_Q\left(\frac{1}{|Q|}\int_Q w_0 w_1^{1-p}\mathrm{d}x\right)\left(\frac{1}{|Q|}\int_Q w_0^{1-p'} w_1\right)^{p-1} < \infty.$$

Since w_0 and w_1 satisfy A_1 condition, we have

$$w_i(x)^{-1} \le \sup_{x\in Q} w_i(x)^{-1} = \left(\inf_{x\in Q} w_i(x)\right)^{-1} \le [w_i]_{A_1}\left(\frac{w_i(Q)}{|Q|}\right)^{-1},$$

which plugged into the previous supremum implies that the A_p characteristic of $w_0 w_1^{1-p}$ does not exceed

$$\sup_Q \frac{1}{|Q|}\int_Q w_0 \cdot [w_1]_{A_1}^{p-1}\left(\frac{w_1(Q)}{|Q|}\right)^{1-p} \mathrm{d}x \cdot \left(\frac{1}{|Q|}\int_Q [w_0]_{A_1}^{p'-1}\left(\frac{w_0(Q)}{|Q|}\right)^{1-p'} w_1\right)^{p-1}$$
$$= [w_0]_{A_1}[w_1]_{A_1}^{p-1}.$$

The proof is complete. □

We turn our attention to strong-type estimates.

THEOREM 2.4. *For $1 < p < \infty$, M is bounded as an operator on $L^p(w)$ if and only if $w \in A_p$.*

Since strong-type estimates imply the corresponding weak-type inequalities, the necessity of the A_p condition follows at once from the previous argumentation. Therefore, the nontrivial part is the sufficiency of this condition, and we will exploit interpolation arguments of Marcinkiewicz. So, suppose that w is a nonzero A_p weight. Then $w > 0$ almost everywhere, so $L^\infty(w) = L^\infty$ and $\|Mf\|_{L^\infty(w)} = \|Mf\|_{L^\infty}$. On the other hand, we have $\|M\|_{L^p(w)\to L^{p,\infty}(w)} < \infty$, as follows from the above discussion. Consequently, Marcinkiewicz interpolation theorem implies that M is bounded as an operator on $L^r(w)$ for all $r > p$. This is not good enough: the desired sufficiency would follow from the above argument if we could show that M satisfies the weak-type (q, q) bound for some $q < p$. To this end, we will show that any $w \in A_p$, $p > 1$ is actually an A_q weight for some $q < p$. We first prove the following statement.

THEOREM 2.5 (Reverse Hölder inequality). *Let $w \in A_p$, $1 \leq p < \infty$. Then there exist a constant $\varepsilon > 0$ depending only on p and the A_p characteristic of w and a constant C depending only on the dimension, such that for any dyadic cube Q,*

$$\left(\frac{1}{|Q|}\int_Q w^{1+\varepsilon}\right)^{1/(1+\varepsilon)} \leq \frac{C}{|Q|}\int_Q w.$$

PROOF: INFORMAL SKETCH. We will consider only the case $p > 1$. Set $c = [w]_{A_p} \in [1, \infty)$. If $c = 1$, then w is constant almost everywhere and there is nothing to prove (the assertion holds with any $\varepsilon > 0$ and $C = 1$). So, suppose that $c > 1$ and consider the domain

$$\Omega_{p,c} = \{(w, v) \in (0, \infty)^2 \, : \, 1 \leq wv^{p-1} \leq c\}.$$

Introduce the function $B : \Omega_{p,c} \to [0, \infty)$ given by the formula

$$B(w, v) = w^r \varphi(t), \qquad t = wv^{p-1} \in [1, c],$$

where $r = 1 + (6c)^{-1}$ and $\varphi(t) = 2 - t^{-1} - (2c)^{-1}\ln t$. Note that on the lower boundary $\{(w, v) \in (0, \infty)^2 : wv^{p-1} = 1\}$ of the set $\Omega_{p,c}$, we have $B(w, v) = w^r$. Furthermore, we have $B(w, v) \leq 2w^r$ on the whole domain. Finally, the function B is concave, since its Hessian matrix

$$\begin{bmatrix} w^{r-2}[r(r-1)\varphi + 2rt\varphi' + t^2\varphi''] & (p-1)w^r v^{p-2}[(r+1)\varphi' + t\varphi''] \\ (p-1)w^r v^{p-2}[(r+1)\varphi' + t\varphi''] & (p-1)w^{r+1}v^{p-3}[(p-2)\varphi' + (p-1)t\varphi''] \end{bmatrix}$$

(the functions φ, φ', φ'' are evaluated at the point t) is nonpositive-definite. Consequently, combining the above observations with Jensen's inequality, we obtain

$$2\left(\frac{1}{|Q|}\int_Q w\right)^r \geq B\left(\frac{1}{|Q|}\int_Q w, \frac{1}{|Q|}\int_Q w^{1-p'}\right)$$
$$\geq \frac{1}{|Q|}\int_Q B(w, w^{1-p'}) = \frac{1}{|Q|}\int_Q w^r.$$

This is precisely the claim. □

The above reasoning has the following gap: the domain $\Omega_{p,c}$ is not convex, so we cannot apply Jensen's inequality directly. However, with some additional technical effort and at the cost of an additional factor depending only on the dimension, this obstacle can be removed. See [**2**].

We return to the characterization of strong type estimates. The following statement, combined with the interpolation argument described above, immediately yields Theorem 2.4.

Corollary 2.6.

(i) *We have* $A_p = \bigcup_{q<p} A_q$*;*
(ii) *If* $w \in A_p$*,* $1 \leq p < \infty$*, then there exists* $\varepsilon > 0$ *such that* $w^{1+\varepsilon} \in A_p$*.*
(iii) *If* $w \in A_p$*,* $1 \leq p < \infty$*, then there exist* C*,* $\delta > 0$ *such that given a dyadic cube* Q *and a measurable set* $S \subset Q$*,*

$$\frac{w(S)}{w(Q)} \leq C\left(\frac{|S|}{|Q|}\right)^\delta. \tag{2.7}$$

Remark 2.7. If a weight w satisfies (2.7), we say that it belongs to the class A_∞.

Proof. (i) If $w \in A_p$, then $w^{1-p'} \in A_{p'}$. Therefore, it satisfies reverse Hölder's inequality for some $\varepsilon > 0$: if Q is an arbitrary dyadic cube, then

$$\left(\frac{1}{|Q|}\int_Q w^{(1-p')(1+\varepsilon)}\right)^{1/(1+\varepsilon)} \leq \frac{C}{|Q|}\int_Q w^{1-p'}.$$

Let q be given by the equation $q' - 1 = (p' - 1)(1 + \varepsilon)$. Then $q' > p'$, so $q < p$ and the above bound implies

$$\left(\frac{1}{|Q|}\int_Q w\right)\left(\frac{1}{|Q|}\int_Q w^{1-q'}\right)^{q-1} = \left(\frac{1}{|Q|}\int_Q w\right)\left(\frac{1}{|Q|}\int_Q w^{(1-p')(1+\varepsilon)}\right)^{(p-1)/(1+\varepsilon)}$$
$$\leq C\left(\frac{1}{|Q|}\int_Q w\right)\left(\frac{1}{|Q|}\int_Q w^{1-p'}\right)^{p-1} \leq C[w]_{A_p}.$$

This implies that $w \in A_q$.

(ii) If $p > 1$, then the claim follows at once if we choose $\varepsilon > 0$ small enough so that both w and $w^{1-p'}$ satisfy reverse Hölder's inequality with exponent $1 + \varepsilon$. If $p = 1$, then there are C, $\varepsilon > 0$ such that for any dyadic cube Q and almost every $x \in Q$,

$$\frac{1}{|Q|}\int_Q w^{1+\varepsilon} \leq C\left(\frac{1}{|Q|}\int_Q w\right)^{1+\varepsilon} \leq Cw^{1+\varepsilon}.$$

(iii) Fix $S \subset Q$ and suppose that w satisfies the reverse Hölder inequality with exponent $1+\varepsilon$. Then

$$w(S) = \int_Q \chi_S w \le \left(\int_Q w^{1+\varepsilon}\right)^{1/(1+\varepsilon)} |S|^{\varepsilon/(1+\varepsilon)} \le Cw(Q)\left(\frac{|S|}{|Q|}\right)^{\varepsilon/(1+\varepsilon)},$$

which is (2.7) with $\delta = \varepsilon/(1+\varepsilon)$. □

REMARK 2.8. The above proof yields that for any $p>1$, there are a constant κ_p depending only on p and a constant C_d depending only on the dimension such that the following holds. If w is an A_p weight on an interval, then w is an $A_{p-\varepsilon}$ weight, where $\varepsilon = \kappa_p[w]_{A_p}^{-1/(p-1)}$. Moreover, $[w]_{A_{p-\varepsilon}} \le C_d[w]_{A_p}$.

2.3. On sharp dependence on the characteristic. There is an interesting quantitative question about the constants in the weak- and strong-type estimates, concerning the optimal dependence of these constants on the characteristic of the weight involved. For instance, fix $1 \le p < \infty$, an A_p weight w and consider the inequality

$$\|Mf\|_{L^{p,\infty}(w)} \le C_{p,w}\|f\|_{L^p(w)}, \qquad 1 \le p < \infty.$$

The problem is to identify the least exponent α_p such that $C_{p,w} \le C_p[w]_{A_p}^{\alpha_p}$, where C_p depends only on p. A glance at (2.6) gives $\alpha_p \le 1/p$ and a glance at (2.5) implies that this value is optimal. What about an analogous question for the strong-type estimate? To see what the above proof yields, we start with the following observation.

So, suppose that w is an A_p weight. By the above remark, it is also an $A_{p-\varepsilon}$ weight for an appropriate ε. Marcinkiewicz interpolation theorem implies that M is bounded on $L^p(w)$ with the constant behaving with respect to $[w]_{A_p}$ as

$$\frac{p}{p-(p-\varepsilon)}\left([w]_{A_p}^{1/(p-\varepsilon)}\right)^p \lesssim [w]_{A_p}^{1/(p-1)}[w]_{A_p}^{p/(p-\varepsilon)} \lesssim [w]_{A_p}^{1/(p-1)}.$$

The exponent $\alpha_p = 1/(p-1)$ can also be shown to be the best possible. We take the opportunity to present a different proof of this fact, which is of independent interest and connections.

THEOREM 2.9. *For any $1<p<\infty$ and any A_p weight w, we have*

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

PROOF. It is convenient to split the reasoning into a few parts.

Step 1. Let $\mathfrak{Q}$ be a given base dyadic cube. We will first construct inductively a family $\mathfrak{F}$ of certain special dyadic cubes. We start with $\mathfrak{F} = \emptyset$ and let $\mathfrak{S}$ to be the collection of all dyadic cubes contained in $\mathfrak{Q}$. The algorithm is as follows: at each step we add to $\mathfrak{F}$ all the maximal elements from $\mathfrak{S}$ and then remove from $\mathfrak{S}$ all the cubes Q' for which there is $Q \in \mathfrak{F}$ such that $\frac{1}{|Q'|}\int_{Q'} f \le \frac{2}{|Q|}\int_Q f$.

For $Q \in \mathfrak{S}$, let $\operatorname{ch} Q$ denote the class of *children* of Q: all maximal $Q' \in \mathfrak{S}$ strictly contained in Q. Observe that

$$\begin{aligned}\sum_{Q'\in \operatorname{ch} Q} |Q'| &= \sum_{Q'\in \operatorname{ch} Q}\left(\frac{1}{|Q'|}\int_{Q'} f\right)^{-1}\int_{Q'} f \\ &< \frac12 \sum_{Q'\in \operatorname{ch} Q}\left(\frac{1}{|Q|}\int_Q f\right)^{-1}\int_{Q'} f \le \frac12\left(\frac{1}{|Q|}\int_Q f\right)^{-1}\int_Q f = \frac12|Q|.\end{aligned}$$

Consequently, the sets $E(Q) = Q \setminus \bigcup \operatorname{ch} Q$ satisfy $|E(Q)| \geq \frac{1}{2}|Q|$ and they are pairwise disjoint. In addition, observe that the above construction yields that the localized maximal operator $M = M^{\mathfrak{Q}}$ satisfies

$$Mf \leq 2 \sum_{Q \in \mathfrak{F}} \left(\frac{1}{|Q|} \int_Q f \right) \chi_{E(Q)}. \tag{2.8}$$

Step 2. We proceed to the L^p bound for M. Set $v = w^{1/(1-p)}$. By (2.8),

$$\begin{aligned}
\|Mf\|_{L^p(w)}^p &\leq 2^p \int_{\mathfrak{Q}} \left(\sum_{Q \in \mathfrak{S}} \left(\frac{1}{|Q|} \int_Q f \right) \chi_{E(Q)} \right)^p w \mathrm{d}x \\
&= 2^p \sum_{Q \in \mathfrak{S}} \left(\frac{1}{|Q|} \int_Q f \right)^p \int_{\mathfrak{Q}} \chi_{E(Q)} w \mathrm{d}x \\
&= 2^p \sum_{Q \in \mathfrak{S}} \left(\frac{1}{v(Q)} \int_Q (fv^{-1})v \right)^p \cdot \left(\frac{1}{|Q|} \int_Q v \right)^p w(E(Q)) \\
&\leq 2^p \left\{ \sum_{Q \in \mathfrak{S}} \left(\frac{1}{v(Q)} \int_Q (fv^{-1})v \right)^p v(E(Q)) \right\} \\
&\qquad \cdot \sup_Q \left(v(E(Q))^{-1} \left(\frac{1}{|Q|} \int_Q v \right)^p w(E(Q)) \right).
\end{aligned}$$

Step 3. Let us analyze the two factors of the latter expression. The sets $E(Q)$ are disjoint, so the term in the parentheses satisfies

$$\begin{aligned}
\sum_{Q \in \mathfrak{S}} \left(\frac{1}{v(Q)} \int_Q (fv^{-1})v \right)^p v(E(Q)) &\leq \|M_v(fv^{-1})\|_{L^p(v)}^p \\
&\leq C_p^p \|fv^{-1}\|_{L^p(v)}^p = C_p^p \|f\|_{L^p(w)}^p.
\end{aligned}$$

To control the second factor, we use the fact that

$$|Q| \leq 2|E(Q)| = \int_{E(Q)} w^{1/p} v^{1/p'} \leq w(E(Q))^{1/p} v(E(Q))^{1/p'}$$

which implies $|Q|^{p'} v(E(Q))^{-1} \leq w(E(Q))^{p'/p}$ and

$$\begin{aligned}
&\sup_Q \left(v(E(Q))^{-1} \left(\frac{1}{|Q|} \int_Q v \right)^p w(E(Q)) \right) \\
&= \sup_Q \left(|Q|^{-p'} |Q|^{p'} v(E(Q))^{-1} \left(\frac{1}{|Q|} \int_Q v \right)^p w(E(Q)) \right) \\
&\leq \sup_Q \left(|Q|^{-p'} \left(\frac{1}{|Q|} \int_Q v \right)^p w(E(Q))^{1+p'/p} \right) \\
&\leq \sup_Q \left(|Q|^{-p'} \left(\frac{1}{|Q|} \int_Q v \right)^p w(Q)^{1+p'/p} \right) = \sup_Q \left(\frac{1}{|Q|} \int_Q v \right)^p \left(\frac{1}{|Q|} \int_Q w \right)^{p'} = [w]_{A_p}^{p'}.
\end{aligned}$$

Putting all the above facts together and letting $|\mathfrak{Q}| \to \infty$, we get the claim. $\square$

2.4. A_∞ weights. Now we will study separately the class of weights satisfying the condition A_∞. Although we have already encountered this condition earlier (see (2.7)), we will start with a different definition. Namely, letting $p \to \infty$ in (2.5), it is natural to consider the following.

DEFINITION 2.10. A weight w on $\mathbb{R}^d$ belongs to the (dyadic) class A_∞, if

$$[w]_{A_\infty} = \sup_Q \left(\frac{1}{|Q|} \int_Q w \right) \exp \left(\frac{1}{|Q|} \int_Q \log w^{-1} \right) < \infty,$$

where the supremum is taken over the class of all dyadic cubes.

The following theorem is the main result of this subsection.

THEOREM 2.11. *Suppose that w is a weight. Then w is in A_∞ if and only if one of the following conditions holds.*

(a) There exist γ, $\delta \in (0,1)$ such that for all dyadic cubes Q we have

$$\left| \left\{ x \in Q : w(x) \leq \gamma \frac{1}{|Q|} \int_Q w \right\} \right| \leq \delta |Q|.$$

(b) There exist α, $\beta \in (0,1)$ such that for all dyadic cubes Q and all measurable subsets A of Q we have

$$|A| \leq \alpha |Q| \qquad \Rightarrow \qquad w(A) \leq \beta w(Q).$$

(c) The reverse Hölder condition holds: there exist positive constants C and ε such that for all dyadic cubes Q,

$$\left(\frac{1}{|Q|} \int_Q w^{1+\varepsilon} \right)^{1/(1+\varepsilon)} \leq \frac{C}{|Q|} \int_Q w.$$

(d) There exist positive constants C_2, δ such that for all dyadic cubes Q and all measurable subsets A of Q we have

$$\frac{w(A)}{w(Q)} \leq C_2 \left(\frac{|A|}{|Q|} \right)^\delta .$$

(e) There exist α', $\beta' \in (0,1)$ such that for all dyadic cubes Q and all measurable subsets A of Q we have

$$w(A) \leq \alpha' w(Q) \qquad \Rightarrow \qquad |A| \leq \beta' |Q|.$$

(f) The weight w belongs to A_p for some $p < \infty$.

PROOF. $w \in A_\infty \Rightarrow$ *(a).* Fix a cube Q. It is easy to see that multiplication of an A_∞ weight by a constant does not alter the A_∞ characteristic, and hence we may assume that $\int_Q \log w = 0$ and hence $\frac{1}{|Q|}\int_Q w \le [w]_{A_\infty}$. Consequently,

$$\begin{aligned}\left|\left\{x \in Q : w(x) \le \gamma \cdot \frac{1}{|Q|}\int_Q w\right\}\right| &\le |\{x \in Q : w(x) \le \gamma [w]_{A_\infty}\}| \\ &\le \left|\left\{x \in Q : \log\left(1+\frac{1}{w(x)}\right) \ge \log\left(1+\frac{1}{\gamma[w]_{A_\infty}}\right)\right\}\right| \\ &\le \frac{1}{\log\left(1+(\gamma[w]_{A_\infty})^{-1}\right)}\int_Q \log\left(1+\frac{1}{w}\right) \\ &= \frac{1}{\log\left(1+(\gamma[w]_{A_\infty})^{-1}\right)}\int_Q \log(1+w) \\ &\le \frac{1}{\log\left(1+(\gamma[w]_{A_\infty})^{-1}\right)}\int_Q w \\ &\le \frac{[w]_{A_\infty}|Q|}{\log\left(1+(\gamma[w]_{A_\infty})^{-1}\right)} \\ &= \frac{1}{2}|Q|,\end{aligned}$$

if $\gamma = [w]_{A_\infty}^{-1}(e^{2[w]_{A_\infty}}-1)^{-1}$.

(a) $\Rightarrow$ *(b)* Fix a cube Q and a subset $A \subset Q$ with $w(A) \ge \beta w(Q)$, for some β to be chosen later. If $S = Q \setminus A$, then $w(S) < (1-\beta)w(Q)$. Consider the decomposition $S = S_1 \cup S_2$, where

$$S_1 = \left\{x \in S : w > \gamma\frac{1}{|Q|}\int_Q w\right\}, \qquad S_2 = \left\{x \in S : w \le \gamma\frac{1}{|Q|}\int_Q w\right\}.$$

By (a), we know that $|S_2| \le \delta|Q|$. For S_1, we apply Chebyshev's inequality to get

$$|S_1| \le \frac{1}{\gamma\frac{1}{|Q|}\int_Q w}\int_S w = \frac{|Q|w(S)}{\gamma w(Q)} \le \frac{1-\beta}{\gamma}|Q|.$$

Consequently,

$$|S| \le |S_1| + |S_2| \le \frac{1-\beta}{\gamma}|Q| + \delta|Q| = \left(\delta + \frac{1-\beta}{\gamma}\right)|Q|.$$

It remains to choose α and β such that $\delta + \frac{1-\beta}{\gamma} = 1-\alpha$: for instance, $\alpha = (1-\delta)/2$ and $\beta = 1-(1-\delta)\gamma/2$, to obtain $|S| \le (1-\alpha)|Q|$, that is, $|A| > \alpha|Q|$.

(b) $\Rightarrow$ (c) Fix a cube Q, denote $\alpha_0 = \frac{1}{|Q|}\int_Q w$ and introduce the increasing sequence $\alpha_k = (2^d\alpha^{-1})^k\alpha_0$. Now we will consider the so-called Calderón-Zygmund decomposition of w at height α_k, where k is a fixed nonnegative integer. This decomposition is based on a selection of certain dyadic subcubes of Q. Namely, at the first step we do not choose Q. Next we consider 2^d children Q_1, Q_2, ..., Q_{2^d} of Q and select those for which we have

$$\frac{1}{|Q_j|}\int_{Q_j} w > \alpha_k.$$

Next, each unselected cube is divided into 2^d subcubes, from which those satisfying the above estimate are selected; and so on. Let $\{Q_{k,j}\}_j$ be the collection of all cubes selected. Then following properties hold:

(1) We have $\alpha_k < \frac{1}{|Q_{k,j}|}\int_{Q_{k,j}} w \le 2^d\alpha_k$.

(2) If $U_k = \bigcup_j Q_{k,j}$, then $w \leq \alpha_k$ outside U_k.
(3) Each $Q_{k+1,j}$ is contained in some $Q_{k,\ell}$.

The first condition is obvious, since the dyadic parent R of $Q_{k,j}$ was not selected in the procedure:

$$2^d \alpha_k \geq \frac{2^d}{|R|} \int_R w = \frac{1}{|Q_{k,j}|} \int_R w \geq \frac{1}{|Q_{k,j}|} \int_{Q_{k,j}} w.$$

The second inequality follows from Lebesgue's differentiation theorem: for any $x \notin U_k$ and any dyadic cube R containing x we have $\frac{1}{|R|} \int_R w \leq \alpha_k$, and hence also $w(x) \leq \alpha_k$. The final property is trivial: since the average of w over $Q_{k+1,j}$ is bigger than α_{k+1}, it is also bigger than α_k.

Let us now compute the portion of $Q_{k,\ell}$ which is covered by $\{Q_{k+1,j}\}$. We have

$$\begin{aligned} 2^d \alpha_k &\geq \frac{1}{|Q_{k,\ell}|} \int_{Q_{k,\ell}} w \\ &\geq \frac{1}{|Q_{k,\ell}|} \sum_{j : Q_{k+1,j} \subseteq Q_{k,\ell}} |Q_{k+1,j}| \frac{1}{|Q_{k+1,j}|} \int_{Q_{k+1,j}} w \\ &\geq \frac{|Q_{k,\ell} \cap U_{k+1}|}{|Q_{k,\ell}|} \cdot \alpha_{k+1} \\ &= \frac{|Q_{k,\ell} \cap U_{k+1}|}{|Q_{k,\ell}|} \cdot 2^d \alpha^{-1} \alpha_k. \end{aligned}$$

Hence $|Q_{k,\ell} \cap U_{k+1}| \leq \alpha |Q_{k,\ell}|$ which, by (b), implies

$$w(Q_{k,\ell} \cap U_{k+1}) \leq \beta w(Q_{k,\ell}).$$

Summing over all ℓ, we get $w(U_{k+1}) \leq \beta w(U_k)$, so in particular $w(U_k) \leq \beta^k w(U_0)$. We also have $|U_{k+1}| \leq \alpha |U_k|$ and hence $|U_k| \to 0$ as $k \to \infty$. Therefore we may write, up to a set of Lebesgue's measure zero, that

$$Q = (Q \setminus U_0) \cup (U_0 \setminus U_1) \cup (U_1 \setminus U_2) \cup \ldots.$$

Consequently, since $w \leq \alpha_{k+1}$ outside U_{k+1},

$$\begin{aligned} \int_Q w^{1+\varepsilon} &= \int_{Q \setminus U_0} w^\varepsilon w + \sum_{k=0}^{\infty} \int_{U_k \setminus U_{k+1}} w^\varepsilon w \\ &\leq \alpha_0^\varepsilon w(Q \setminus U_0) + \sum_{k=0}^{\infty} \alpha_{k+1}^\varepsilon w(U_k) \\ &\leq \alpha_0^\varepsilon w(Q \setminus U_0) + \sum_{k=0}^{\infty} \left((2^d \alpha^{-1})^{k+1} \alpha_0\right)^\varepsilon \beta^k w(U_0) \\ &\leq \alpha_0^\varepsilon \left(1 + (2^d \alpha^{-1})^\varepsilon \sum_{k=0}^{\infty} (2^d \alpha^{-1})^{k\varepsilon} \beta^k \right) w(Q) \\ &= \left(\frac{1}{|Q|} \int_Q w\right)^\varepsilon \left(1 + \frac{(2^d \alpha^{-1})^\varepsilon}{1 - (2^d \alpha^{-1})^\varepsilon \beta}\right) \int_Q w, \end{aligned}$$

provided ε is chosen sufficiently small. This yields the reverse Hölder property.

(c) $\Rightarrow$ (d) We have already seen the proof above (see (2.7)).

(d) $\Rightarrow$ (e) Pick $\alpha'' \in (0,1)$ such that $\beta'' = C_2(\alpha'')^\delta < 1$. Then, by (d),

$$|A| < \alpha''|Q| \qquad \Rightarrow \qquad w(A) \le \beta'' w(Q).$$

Replacing A with $Q \setminus A$, the above implication can be rewritten as

$$|A| \ge (1-\alpha'')|Q| \qquad \Rightarrow \qquad w(A) \ge (1-\beta'')w(Q).$$

Therefore,

$$w(A) < (1-\beta'')w(Q) \qquad \Rightarrow \qquad |A| < (1-\alpha'')|Q|,$$

which is the desired property, with $\alpha' = 1-\beta''$ and $\beta' = 1-\alpha''$.

(e) $\Rightarrow$ (f). Left as an exercise.

(f) $\Rightarrow w \in A_\infty$. This follows at once from the inequality $[w]_{A_\infty} \le [w]_{A_p}$. □

Problems

1. Prove that if w is an A_p weight and u, $u^{-1} \in L^\infty$, then uw is an A_p weight.

2. Suppose that w is an A_p weight for some $p \ge 1$ and let $\delta \in (0,1)$. Prove that $w^\delta \in A_q$, where $q = \delta p + 1 - \delta$ and $[w^\delta]_{A_q} \le [w]_{A_p}^\delta$.

3. Show that if the A_p characteristic constants of a weight w are uniformly bounded on $(1,\infty)$, then w is an A_1 weight.

4. Let $w_0 \in A_{p_0}$ and $w_1 \in A_{p_1}$ for some $1 \le p_0 < p_1 < \infty$. For a given $\theta \in (0,1)$, define

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1} \qquad \text{and} \qquad w^{\frac{1}{p}} = w_0^{\frac{1-\theta}{p_0}} w_1^{\frac{\theta}{p_1}}.$$

Prove that w is an A_p weight and

$$[w]_{A_p} \le [w_0]_{A_{p_0}}^{(1-\theta)\frac{p}{p_0}} [w_1]_{A_{p_1}}^{\theta\frac{p}{p_1}}.$$

5. Let w be an A_p weight for some $1 \le p < \infty$ and $k \ge 1$ be a constant. Show that $\min\{w,k\}$ is in A_p and satisfies

$$[\min\{w,k\}]_{A_p} \le c_p([w]_{A_p} + 1)$$

where $c_p = 1$ for $p \le 2$ and $c_p = 2^{p-2}$ for $p > 2$.

6. Let $\mathcal{M}$ be an uncentered maximal operator and let M^α be the dyadic maximal operator on $\mathbb{R}^d$ shifted by $\alpha \in \{0, \frac{1}{3}\}^d$, i.e., corresponding to the grid

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

Prove the pointwise bound

$$\mathcal{M}f(x) \le 6^d \sum_{\alpha \in \{0,\frac{1}{3}\}^d} M^\alpha f(x).$$

7. Let $\mathcal{M}$ be the uncentered maximal operator on $\mathbb{R}^d$ and let w be a positive function on $\mathbb{R}^d$. Prove that $\mathcal{M}$ is bounded as an operator on $L^p(w)$ if and only if

$$[w]_{A_p} = \sup \left(\frac{1}{|Q|}\int_Q w\right)\left(\frac{1}{|Q|}\int_Q w^{-1/(p-1)}\right)^{p-1} < \infty,$$

where the supremum is taken over all cubes Q with sides parallel to the axes.

8. Prove the following alternative characterization of the dyadic A_p constant:

$$[w]_{A_p} = \sup \frac{\left(\frac{1}{|Q|}\int_Q f(x)\mathrm{d}x\right)^p}{\frac{1}{w(Q)}\int_Q f(x)^p w(x)\mathrm{d}x},$$

where the supremum is taken over all dyadic cubes Q and all strictly positive functions $f \in L^p(Q;w)$.

9. Let d be a fixed dimension. Characterize those $\alpha \in \mathbb{R}$, for which the function $w(x) = |x|^\alpha$, $x \in [0,1)^d$, is a dyadic A_p weight.

10. Prove that for any nonnegative function f and $\alpha \in (0,1)$, the function $w = (Mf)^\alpha$ is an A_1 weight and $[w]_{A_1} \lesssim (1-\alpha)^{-1}$.

11. For a weight w on $\mathbb{R}^d$, we define the A_∞ characteristic by $[w]^*_{A_\infty} = \sup_{Q\in\mathfrak{D}} \frac{1}{w(Q)}\int_Q M(w\chi_Q)$. Prove that for any $1 \le p < \infty$ we have $[w]^*_{A_\infty} \lesssim [w]_{A_p}$.

12. Let w be a weight on $\mathbb{R}^d$ and let $Q \in \mathfrak{D}$ be a maximal cube satisfying $\frac{1}{|Q|}\int_Q w > 1$. Prove that $Mw = M(w\chi_Q)$ on Q, $\int_Q Mw \le 2^d[w]^*_{A_\infty}|Q|$ and $w(Q) \le 2^d|Q|$.

13. Suppose that w is a weight in $\mathbb{R}^d$ and set $\varepsilon = (2^{d+1}[w]^*_{A_\infty} - 1)^{-1}$. Prove that for any $Q \in \mathfrak{D}$ we have

$$\frac{1}{|Q|}\int_Q (Mw\chi_Q)^{1+\varepsilon} \le 2\left(\frac{1}{|Q|}\int_Q w\right)^\varepsilon \cdot \frac{1}{|Q|}\int_Q M(w\chi_Q) \le 2[w]^*_{A_\infty}\left(\frac{1}{|Q|}\int_Q w\right)^{1+\varepsilon}$$

and

$$\frac{1}{|Q|}\int_Q M(w\chi_Q)^\varepsilon w \le 2\left(\frac{1}{|Q|}\int_Q w\right)^{1+\varepsilon}.$$

14. Let $1 < p < \infty$ and $w \in A_p(\mathbb{R}^d)$. Prove that $[w]_{A_q} \lesssim [w]_{A_p}$, where

$$q = p - \frac{p-1}{2^{d+1}[w^{1-p'}]^*_{A_\infty}}.$$

15. Prove that if $[w]^*_{A_\infty} < \infty$, then $[w]_{A_\infty} < \infty$.

16. Prove that w belongs to A_p if and only if w and $w^{-1/(p-1)}$ belong to A_∞.

17. Show that the product of two A_1 weights may not be an A_∞ weight.

18. Let $f \in L^p(w)$ for some $1 \le p \le \infty$ and some $w \in A_p$. Prove that $f \in L^1_{loc}(\mathbb{R}^d)$.

19. Prove that a weight w belongs to A_∞ if and only if there exist positive constants γ, C such that for all cubes Q,

$$w(\{x \in Q : w(x) > \lambda\}) \le C\lambda|\{x \in Q : w(x) > \gamma\lambda\}|$$

for all $\lambda > \frac{1}{|Q|}\int_Q w$.

20. Prove that $[w]^*_{A_\infty} \le e[w]_{A_\infty}$.

21. Prove that for any nonnegative function f and any $r > 1$, $\alpha < 0$, the weight $(Mf^r)^\alpha$ belongs to the class A_∞.

CHAPTER 3

Extrapolation and Factorization

3.1. Extrapolation. We will now study a very beautiful and powerful property of weighted inequalities. Roughly speaking, if $1 \le p_0 < \infty$ is a fixed parameter and T is an operator acting on measurable functions on $\mathbb{R}^d$ which is bounded on $L^{p_0}(w)$ for all A_{p_0} weights, then T is automatically bounded on $L^p(w)$ for all p and all A_p weights w (in particular, it is bounded on L^p). This surprising fact was first proved for $p_0 = 2$ ('there are no L^p spaces, just weighted L^2').

We start with a procedure, known as Rubio de Francia algorithm.

THEOREM 3.1. *Let $p > 1$ and suppose that w is an A_p weight and f is a nonnegative function from $L^p(w)$. For any integer k, let M^k denote the k-th iteration of M, $M^0 = I$. Define*

$$Rf(x) = \sum_{k=0}^{\infty} \frac{M^k f(x)}{(2\|M\|_{L^p(w)\to L^p(w)})^k}.$$

Then $f(x) \le Rf(x)$ almost everywhere, $\|Rf\|_{L^p(w)} \le 2\|f\|_{L^p(w)}$ and Rf is an A_1 weight satisfying $[Rf]_{A_1} \le 2\|M\|_{L^p(w)}$.

PROOF. The inequality $f \le Rf$ is evident. The second claim is also easy: simply use the bound $\|M^k\|_{L^p(w)\to L^p(w)} \le \|M\|^k_{L^p(w)\to L^p(w)}$ and sum the geometric series. Finally, note that by sublinearity of the maximal operator,

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

This yields the assertion. □

THEOREM 3.2. *Assume that for some pair (f, g) of nonnegative functions, for some $p_0 \in [1, \infty)$ and all $w \in A_{p_0}$ we have*

$$\|g\|_{L^{p_0}(w)} \le CN([w]_{A_{p_0}})\|f\|_{L^{p_0}(w)},$$

where N is an increasing function and C does not depend on w. Then for all $1 < p < \infty$ and all $w \in A_p$ we have

$$\|g\|_{L^p(w)} \le CK(w)\|f\|_{L^p(w)},$$

where

$$K(w) = \begin{cases} N\big([w]_{A_p}(2\|M\|_{L^p(w)\to L^p(w)})^{p_0-p}\big) & \text{if } p < p_0, \\ N\left([w]_{A_p}^{\frac{p_0-1}{p-1}}(2\|M\|_{L^{p'}(w^{1/(1-p)})\to L^{p'}(w^{1/(1-p)})})^{\frac{p-p_0}{p-1}}\right) & \text{if } p > p_0. \end{cases}$$

In particular, $K(w) \le C_1 N\left(C_2[w]_{A_p}^{\max\{\frac{p_0-1}{p-1},1\}}\right)$ *for* $w \in A_p$.

Proof for $p < p_0$. For $f \in L^p(w)$, let Rf be the A_1 weight built by the above Rubio de Francia algorithm. We have

$$\begin{aligned}
\int_{\mathbb{R}^d} g^p w &= \int_{\mathbb{R}^d} g^p w (Rf)^{\frac{p}{p_0}(p-p_0)} (Rf)^{\frac{p}{p_0}(p_0-p)} \\
&\le \left(\int_{\mathbb{R}^d} g^{p_0} w (Rf)^{p-p_0}\right)^{\frac{p}{p_0}} \left(\int_{\mathbb{R}^d} (Rf)^p w\right)^{1-\frac{p}{p_0}} \\
&\le CN\left([w(Rf)^{p-p_0}]_{A_{p_0}}\right)^p \left(\int_{\mathbb{R}^d} f^{p_0} w (Rf)^{p-p_0}\right)^{\frac{p}{p_0}} \left(\int_{\mathbb{R}^d} f^p w\right)^{1-\frac{p}{p_0}} \\
&\le CN\left([w]_{A_p}[Rf]_{A_1}^{p_0-p}\right)^p \int_{\mathbb{R}^d} f^p w \\
&\le CN\left([w]_{A_p}(2\|M\|_{L^p(w)\to L^p(w)})^{p_0-p}\right)^p \int_{\mathbb{R}^d} f^p w.
\end{aligned}$$

We can now use the estimate $\|M\|_{L^p(w)\to L^p(w)} \lesssim_p [w]_{A_p}^{1/(p-1)}$ to obtain the bound $C_1 N\left(C_2[w]_{A_p}^{\frac{p_0-1}{p-1}}\right)$. □

Proof for $p > p_0$. By duality, we can write

$$\left(\int_{\mathbb{R}^d} g^p w\right)^{\frac{p_0}{p}} = \sup\left\{\int_{\mathbb{R}^d} g^{p_0} h w \,:\, h \in L^{\frac{p}{p-p_0}}(w) \text{ of norm } 1\right\}.$$

Fix a function h as above; we may and do assume that it is nonnegative. Define H by the identity $H^{p'} w^{1-p'} = h^{\frac{p}{p-p_0}} w$. It is easy to see that $\|H\|_{L^{p'}(w^{1-p'})} = 1$. Let RH be the A_1 weight built by Rubio de Francia algorithm applied to p' and the weight $w^{1-p'}$. Since $H \le RH$, we get

$$\begin{aligned}
&\int_{\mathbb{R}^d} g^{p_0} h w \\
&\le \int_{\mathbb{R}^d} g^{p_0} w^{\frac{p_0-1}{p-1}} (RH)^{\frac{p-p_0}{p-1}} \\
&\le CN\left(\left[w^{\frac{p_0-1}{p-1}} (RH)^{\frac{p-p_0}{p-1}}\right]_{A_{p_0}}\right)^{p_0} \int_{\mathbb{R}^d} f^{p_0} w^{\frac{p_0-1}{p-1}} (RH)^{\frac{p-p_0}{p-1}} \\
&\le CN\left([w]_{A_p}^{\frac{p_0-1}{p-1}} \left(2\|M\|_{L^{p'}(w^{1-p'})}\right)^{\frac{p-p_0}{p-1}}\right)^{p_0} \left(\int_{\mathbb{R}^d} f^p w\right)^{\frac{p_0}{p}} \left(\int_{\mathbb{R}^d} (RH)^{p'} w^{1-p'}\right)^{1-\frac{p_0}{p}}.
\end{aligned}$$

This is precisely the desired assertion. □

The above statement does not refer to any operator T. This abstract formulation allows for a number of important examples.

Corollary 3.3. *Suppose that T is an operator acting on measurable functions on $\mathbb{R}^d$ and let $p_0 \in [1,\infty)$ be a fixed parameter. Let C be a finite constant and let N be an increasing function on $[1,\infty)$.*

(i) If we have the estimate $\|T\|_{L^{p_0}(w)\to L^{p_0}(w)} \le CN([w]_{A_{p_0}})$ for all weights $w \in A_{p_0}$, then $\|T\|_{L^p(w)\to L^p(w)} \le CK(w)$, where K is as in the previous statement.

(ii) If we have the estimate $\|T\|_{L^{p_0}(w)\to L^{p_0,\infty}(w)} \le CN([w]_{A_{p_0}})$ *for all weights* $w \in A_{p_0}$*, then* $\|T\|_{L^p(w)\to L^{p,\infty}(w)} \le CK(w)$*, where* K *is as in the previous statement.*

(iii) Suppose that $\|T\|_{L^{p_0}(w)\to L^{p_0}(w)} \le CN([w]_{A_{p_0}})$ *for all weights* $w \in A_{p_0}$*. Consider the vector-valued extension of* T*, given by* $T\mathtt{f} = (Tf_i)_{i\ge 0}$*, where* $\mathtt{f} = (f_0, f_1, f_2, \ldots)$*. Then for any* $1 \le p,\, q < \infty$ *we have*

$$\int_{\mathbb{R}^d} \|Tf(x)\|_{\ell^q}^p w(x)\,dx \le L \int_{\mathbb{R}^d} \|f(x)\|_{\ell^q}^p w(x)\,dx.$$

Proof. (i) Simply apply the previous theorem with f and $g = |Tf|$.

(ii) Fix $\lambda > 0$ and note that

$$\|\lambda\chi_{\{|Tf|>\lambda\}}\|_{L^{p_0}(w)} = \lambda\left(w_0(|Tf|>\lambda)\right)^{1/p_0} \le \|Tf\|_{L^{p_0}(w)} \le CN([w]_{A_{p_0}})\|f\|_{L^{p_0}(w)}$$

for any A_p weight w. Therefore, an application of the previous theorem gives

$$\|\lambda\chi_{\{|Tf|>\lambda\}}\|_{L^p(w)} \le CK(w)\|f\|_{L^p(w)},$$

and taking the supremum over λ on the left yields the claim.

(iii) By extrapolation, T is bounded on $L^q(w)$ whenever $w \in A_q$ and hence

$$\int_{\mathbb{R}^d} |Tf_i(x)|^q w(x)\mathrm{d}x \le C^q K(w)^q \int_{\mathbb{R}^d} |f_i(x)|^q w(x)\mathrm{d}x.$$

Summing over all i, we get

$$\int_{\mathbb{R}^d} \|T\mathtt{f}(x)\|_{\ell^q}^q w(x)\mathrm{d}x \le C^q K(w)^q \int_{\mathbb{R}^d} \|\mathtt{f}(x)\|_{\ell^q}^q w(x)\mathrm{d}x$$

for all A_q weights w. Applying the extrapolation again, with $f = \|\mathtt{f}\|_{\ell^q}$ and $g = \|T\mathtt{f}\|_{\ell^q}$, we get the claim.

□

3.2. Factorization. It is easy to show, using Hölder's inequality, that if w_1, w_2 are A_1 weights, then their product $w_1 w_2^{1-p}$ is an A_p weight. We will now turn our attention to the reverse statement, often referred to as the factorization of A_p weights.

Theorem 3.4. *Suppose that* w *is an* A_p *weight for some* $1 < p < \infty$*. Then there exist* A_1 *weights* w_1 *and* w_2 *such that* $w = w_1 w_2^{1-p}$*.*

Proof. Fix $p \ge 2$ and define the operator T by the formula

$$T(f) = (w^{-1/p} M(f^{p-1} w^{1/p}))^{1/(p-1)} + w^{1/p} M(f w^{-1/p}).$$

Then T is bounded on (unweighted) L^p, since each of the two parts of T have this property. Indeed, $v = w^{-1/(p-1)}$ is an $A_{p'}$ weight, so

$$\begin{aligned}\left\|(w^{-1/p} M(f^{p-1} w^{1/p}))^{1/(p-1)}\right\|_{L^p} &= \|M(f^{p-1} w^{1/p})\|_{L^{p'}(v)} \\ &\le \|M\|_{L^{p'}(v)\to L^{p'}(v)} \|f^{p-1} w^{1/p}\|_{L^{p'}(v)} \\ &\lesssim_p [v]_{A_{p'}}^{1/(p'-1)} \|f\|_{L^p} \\ &= [w]_{A_p} \|f\|_{L^p}.\end{aligned}$$

The second part of T is handled similarly:

$$\left\|w^{1/p} M(f w^{-1/p})\right\|_{L^p} \le \|M\|_{L^p(w)\to L^p(w)} \|f\|_{L^p} \lesssim_p [w]_{A_p}^{1/(p-1)} \|f\|_{L^p}$$

and hence we finally get $\|T\|_{L^p\to L^p} \lesssim_p [w]_{A_p}$. Next, observe that the operator T is nonnegative and sublinear: $T(f+g) \le Tf + Tg$ and $T(\alpha f) = \alpha Tf$ for any functions f, $g \ge 0$ and any $\alpha \in \mathbb{R}_+$. This follows at once from the fact that for each cube Q, the operator

$$f \mapsto \left(\frac{1}{|Q|}\int_Q f^{p-1} w^{1/p} \mathrm{d}x\right)^{1/(p-1)}$$

is sublinear (here we use the assumption $p \ge 2$).

Now, fix $f_0 \in L^p$ with $\|f_0\|_{L^p} = 1$ and define φ as the sum of L^p convergent series

$$\varphi = \sum_{j=1}^{\infty} \frac{T^j(f_0)}{(2\|T\|_{L^p\to L^p})^j}.$$

Put $w_1 = w^{1/p}\varphi^{p-1}$ and $w_2 = w^{-1/p}\varphi$, so that $w = w_1 w_2^{1-p}$. Now we will prove that w_1 and w_2 are A_1 weights. Applying T to φ, we get

$$T\varphi \le 2\|T\|_{L^p\to L^p} \sum_{j=1}^{\infty} \frac{T^{j+1} f_0}{(2\|T\|_{L^p\to L^p})^{j+1}} \le 2\|T\|_{L^p\to L^p}\varphi,$$

that is, equivalently,

$$(w^{-1/p} M(\varphi^{p-1} w^{1/p}))^{1/(p-1)} + w^{1/p} M(\varphi w^{-1/p}) \le 2\|T\|_{L^p\to L^p}\varphi.$$

But $\varphi = (w^{-1/p} w_1)^{1/(p-1)} = w^{1/p} w_2$, so the above estimate yields $M(w_1) \le (2\|T\|_{L^p\to L^p})^{p-1} w_1$ and $Mw_2 \le 2\|T\|_{L^p\to L^p} w_2$. This gives the claim for $p \ge 2$; note that we obtain $[w_1]_{A_1} \le (2\|T\|_{L^p\to L^p})^{p-1} \lesssim_p [w]_{A_p}^{p-1}$ and $[w_2]_{A_2} \le 2\|T\|_{L^p\to L^p} \lesssim_p [w]_{A_p}$.

We pass to the case $p < 2$. Given a weight $w \in A_p$, we consider its dual $v = w^{1-p'} = w^{1/(1-p)}$ which is in $A_{p'}$. Using the claim for $p \ge 2$ (which has already been established above), we write $w^{1-p'} = v_1 v_2^{1-p'}$, where v_1, v_2 are A_1 weights. Consequently, we see that $w = v_1^{1-p} v_2$, which is the desired decomposition. Observe that

$$[v_1]_{A_1} \lesssim_p [v]_{A_{p'}}^{p'-1} = [w]_{A_p}^{1/(p-1)^2} \quad \text{and} \quad [v_2]_{A_1} \lesssim_p [v]_{A_{p'}} = [w]_{A_p}^{1/(p-1)}. \quad \square$$

Problems

1. Let (X, μ) be a measure space and let T be a sublinear positive operator on $L^p(X,\mu)$ for some $1 < p < \infty$. Prove that for all $f_0 \in L^p(X,\mu)$ there exists a function $f \in L^p(X,\mu)$ such that

(a) $f_0 \le f$ almost everywhere on X.

(b) $\|f\|_{L^p(X)} \le 2\|f_0\|_{L^p(X)}$.

(c) $T(f) \le 2\|T\|_{L^p\to L^p} f$ almost everywhere on X.

2. Suppose that $1 < p < \infty$ and let w be a weight. Prove that there is a nonnegative sublinear operator R for which

(i) $h \le R(h)$,

(ii) $\|R(h)\|_{L^p(w)} \le 2\|h\|_{L^p(w)}$,

(iii) $[R(h)w^{1/p}]_{A_1} \le cp'$ for some universal constant c.

3. Let $1 < p < \infty$ and $w \in A_p$. Show that there is a nonnegative sublinear operator R bounded on $L^{p'}(w)$ such that for any nonnegative $h \in L^{p'}(w)$ we have

(i) $h \le R(h)$;

(ii) $\|R(h)\|_{L^{p'}(w)} \leq 2\|h\|_{L^{p'}(w)}$;
(iii) $[R(h)w]_{A_1} \leq C_p[w]_{A_p}$.

4. Let T be a sublinear operator defined on $\bigcup_{q>2} L^q$ and suppose that for all f and u we have

$$\int_{\mathbb{R}^d} |T(f)|^2 u \mathrm{d}x \leq \int_{\mathbb{R}^d} |f|^2 M u \mathrm{d}x.$$

Prove that T maps L^p to itself for all $2 < p < \infty$.

5. Let T be a sublinear operator defined on $\bigcup_{q>2} L^q(w)$ and suppose that for all weights w satisfying $w^{-1} \in A_1$ we have

$$\int_{\mathbb{R}^d} |Tf|^2 w \leq C[w^{-1}]_{A_1}^{\alpha} \int_{\mathbb{R}^d} |f|^2 w$$

for some constants C, α independent of w. Prove that T maps L^p to itself for all $1 < p < 2$.

6. Let w be an A_1 weight and suppose that a sublinear operator T satisfies

$$\|Tf\|_{L^2(w)} \leq C[w]_{A_1}^{\alpha} \|f\|_{L^2(w)}$$

for some constant C independent of w. Prove that for $p > 2$, T is bounded on L^p and satisfies $\|T\|_{L^p \to L^p} \leq O(p)$.

7. Let T be an operator acting on measurable functions on $\mathbb{R}^d$. Given $r > 0$, suppose that for some $p_0 \geq r$, T is bounded on $L^{p_0}(w)$ whenever $w \in A_{p_0/r}$:

$$\int_{\mathbb{R}^d} |Tf|^{p_0} w \leq CN([w]_{A_{p_0/r}}) \int_{\mathbb{R}^d} |f|^{p_0} w.$$

Prove that for all $p > r$, the operator T is bounded on $L^p(w)$ if $w \in A_{p/r}$.

8. Let S, T be a pair of operators acting on measurable functions on $\mathbb{R}^d$. Suppose that for some $0 < p_0 < \infty$ and all $w \in A_\infty$ we have the estimate $\|Tf\|_{L^{p_0}(w)} \leq C\|Sf\|_{L^{p_0}(w)}$, where C does not depend on w. Prove that for all $0 < p < \infty$ we have $\|Tf\|_{L^p(w)} \leq C'\|Sf\|_{L^p(w)}$ whenever $w \in A_\infty$, where C' does not depend on w.

CHAPTER 4

Weighted inequalities for singular integrals

Now we will use the facts developed in the previous chapters to study some weighted estimates for Calderón-Zygmund singular integral operators. The most fundamental example of such an operator is the so-called Hilbert transform

$$Hf(x) = \frac{1}{\pi}\int_{\mathbb{R}} \frac{f(y)}{x-y}\mathrm{d}y,$$

for any integrable function f on the line. Since $1/x$ is not integrable, the above definition must be treated with care (and corrected a little bit). To avoid the singularity, one studies the truncated Hilbert transforms

$$H^{(\varepsilon)}f(x) = \frac{1}{\pi}\int_{y:|x-y|>\varepsilon} \frac{f(y)}{x-y}\mathrm{d}y$$

and shows that for $f \in L^1(\mathbb{R})$, the limit $\lim_{\varepsilon\to 0} H^{(\varepsilon)}f$ exists almost surely. This limiting procedure is called the Cauchy principal integral: so, we have

$$Hf(x) = \text{p.v.}\int_{\mathbb{R}} \frac{f(y)}{x-y}\mathrm{d}x = \lim_{\varepsilon\to 0}\int_{y:|x-y|>\varepsilon} \frac{f(y)}{|x-y|}\mathrm{d}y.$$

An analogous approach allows the study of much wider convolution operators on $\mathbb{R}^d$, given by

$$Tf(x) = T_K f(x) = \text{p.v.}\int_{\mathbb{R}^d} K(x-y)f(y)\mathrm{d}y.$$

We will assume that the kernel K satisfies the growth conditions

(i) $\|\hat{K}\|_\infty \leq C$,
(ii) $|K(x)| \leq C|x|^{-d}$,
(iii) $|K(x) - K(x-y)| \leq C|y||x|^{-1-d}$ for $|y| < |x|/2$.

These requirements enable an efficient study of various boundedness properties of T_K. In particular, the condition (i) implies that T_K is bounded on L^2. Indeed, by Plancherel's theorem,

$$\|Tf\|_{L^2(\mathbb{R}^d)} = \|\widehat{Tf}\|_{L^2(\mathbb{R}^d)} \leq \|\widehat{K}\widehat{f}\|_{L^2(\mathbb{R}^d)} \leq \|\widehat{K}\|_{L^\infty(\mathbb{R}^d)}\|\widehat{f}\|_{L^2(\mathbb{R}^d)} \leq C\|f\|_{L^2(\mathbb{R}^d)}.$$

This property can be significantly improved. Consider the maximal truncation

$$T^*f(x) = \sup_{\varepsilon>0}|T^{(\varepsilon)}f(x)| = \sup_{\varepsilon>0}\left|\int_{y:|y-x|>\varepsilon} K(x-y)f(y)\mathrm{d}y\right|.$$

The so-called Cotlar's inequality asserts that there is a finite constant depending only on K and the dimension such that

$$T^*f(x) \leq M(Tf)(x) + CMf(x)$$

almost everywhere. By the L^2 boundedness of T and M, we obtain that T^* is also bounded on L^2.

Now we will establish the following weak-type estimate.

THEOREM 4.1. *There is a constant C depending only on K and d such that for any $f \in L^2(\mathbb{R}^d)$,*

$$\lambda|\{x \in \mathbb{R}^d : |Tf(x)| > \lambda\}| \leq C\|f\|_{L^1(\mathbb{R}^d)}.$$

PROOF. Consider the Calderón-Zygmund decomposition of f on the level λ; as the result, we obtain a family $\{Q_j\})_j$ of pairwise disjoint cubes such that

$$\lambda < \frac{1}{|Q_j|}\int_{Q_j} |f|\mathrm{d}x \leq 2^d\lambda$$

and such that $|f| \leq \lambda$ on the set $U = \bigcup_j Q_j$. In particular, the above bound implies

$$|U| = \sum_j |Q_j| \leq \sum_j \frac{1}{\lambda}\int_{Q_j} |f|\mathrm{d}x \leq \frac{\|f\|_{L^1(\mathbb{R}^d)}}{\lambda}. \tag{4.1}$$

Consider the decomposition of f given by

$$g(x) = \begin{cases} \frac{1}{|Q_j|}\int_{Q_j} f & \text{if } x \in Q_j, \\ f(x) & \text{if } x \notin U, \end{cases} \qquad b(x) = f(x) - g(x).$$

By the triangle inequality, we have $|\{|Tf| > \lambda\}| \leq |\{|Tg| > \lambda/2\}| + |\{|Tb| > \lambda/2\}|$. The function g is bounded by $2^d\lambda$, so using the fact that $\|T\|_{L^2\to L^2} < \infty$, we obtain

$$\begin{aligned} |\{|Tg| > \lambda/2\}| \leq \frac{4\|g\|^2_{L^2(\mathbb{R}^d)}}{\lambda^2} &= 4\lambda^{-2}\left(\int_U g^2 + \int_{\mathbb{R}^d\setminus U} g^2\right) \\ &\leq 4\lambda^{-2}\left(2^{2d}\lambda^2|U| + \int_{\mathbb{R}^d\setminus U} f^2\right) \\ &\leq \frac{4(2^{2d}+1)\|f\|_{L^1(\mathbb{R}^d)}}{\lambda}. \end{aligned}$$

To handle the term involving Tb, we write

$$|\{|Tb| > \lambda/2\}| \leq \left|\bigcup_j Q_j^*\right| + \left|\left\{x \in \mathbb{R}^d \setminus \bigcup_j Q_j^* : |Tb(x)| > \lambda/2\right\}\right|,$$

where Q_j^* stands for the double of Q_j, i.e., the cube of the same center as Q_j, but twice larger side. Arguing as in (4.1), we see that the first term on the right is bounded by $C\|f\|_{L^1(\mathbb{R}^d)}/\lambda$ and it is enough to estimate the second. We have $b = \sum_j b_j$, where $b_j(x) = \left[f(x) - \frac{1}{|Q_j|}\int_{Q_j} f\right]\chi_{Q_j}(x)$ (the series converges in L^1). Each b_j has integral zero, so

$$Tb_j(x) = \int_{\mathbb{R}^d} K(x-y)b_j(y)\mathrm{d}y = \int_{\mathbb{R}^d}(K(x-y) - K(x - x_{Q_j}))b_j(y)\mathrm{d}y,$$

where x_{Q_j} is the center of Q_j. Note that the principal values are not necessary: b_j is supported on Q_j and x belongs to the compliment of Q_j^*. We see that $|(x-y) - (x - x_Q)| = |x_Q - y| < \operatorname{diam} Q_j/2 < |x - x_Q|/2$, so the assumption (iii) on the kernel K gives

$$|Tb_j(x)| \leq C\int_{Q_j} \frac{\operatorname{diam} Q_j \cdot |b_j(y)|}{|x - x_{Q_j}|^{d+1}}\mathrm{d}y.$$

Therefore, integrating and summing over j, we get

$$\begin{aligned}\int_{\mathbb{R}^d\setminus\bigcup_j Q_j^*}|Tb(x)|\mathrm{d}x &\le C\sum_j\int_{\mathbb{R}^d\setminus\bigcup_j Q_j^*}\int_{Q_j}\frac{\operatorname{diam}Q_j\cdot|b_j(y)|}{|x-x_{Q_j}|^{d+1}}\mathrm{d}y\mathrm{d}x\\ &= C\sum_j\int_{Q_j}\int_{\mathbb{R}^d\setminus\bigcup_j Q_j^*}\frac{\operatorname{diam}Q_j\cdot|b_j(y)|}{|x-x_{Q_j}|^{d+1}}\mathrm{d}x\mathrm{d}y\\ &\le C\sum_j\int_{Q_j}|b_j(y)|\mathrm{d}y\le C\|f\|_{L^1(\mathbb{R}^d)}.\end{aligned}$$

It remains to apply Chebyshev's inequality to get

$$\left|\left\{x\in\mathbb{R}^d\setminus\bigcup_j Q_j^*:|Tb(x)|>\lambda/2\right\}\right|\le\frac{C\|f\|_{L^1(\mathbb{R}^d)}}{\lambda}.$$

Putting all the above facts together, we get the desired assertion. □

Marcinkiewicz interpolation theorem allows now to prove that T is bounded on L^p, $1<p<2$. The boundedness for $p>2$ is shown by duality. Namely, for any f, $g\in L^2(\mathbb{R}^d)$ we have

$$\int_{\mathbb{R}^d}T_Kf(x)g(x)\mathrm{d}x=\int_{\mathbb{R}^d}f(x)T_{K'}g(x)\mathrm{d}x,$$

where the kernel K', given by $K'(x)=K(-x)$, also satisfies the assumptions (i), (ii) and (iii). Consequently, if $p>2$, $f\in L^p(\mathbb{R}^d)\cap L^2(\mathbb{R}^d)$ and $g\in L^{p'}(\mathbb{R}^d)\cap L^2(\mathbb{R}^d)$, then

$$\int_{\mathbb{R}^d}T_Kf(x)g(x)\mathrm{d}x\le\|f\|_{L^p(\mathbb{R}^d)}\|T_{K'}g\|_{L^{p'}(\mathbb{R}^d)}\le C\|f\|_{L^p(\mathbb{R}^d)}\|g\|_{L^{p'}(\mathbb{R}^d)}.$$

This yields $\|T_Kf\|_{L^p(\mathbb{R}^d)}\le C\|f\|_{L^p(\mathbb{R}^d)}$ and standard approximation allows to extend this estimate for an arbitrary $f\in L^p(\mathbb{R}^d)$. The aforementioned Cotlar's inequality yields the validity of these L^p bounds for the maximal truncation. We would like to conclude the discussion by saying that a little more effort leads to the related weak-type bound for T^*:

$$\lambda|\{x\in\mathbb{R}^d:|T^*f(x)|>\lambda\}|\le C\|f\|_{L^1(\mathbb{R}^d)}.$$

We will study the above estimates in the weighted context. We start with a preparatory lemma, the so-called Whitney decomposition.

LEMMA 4.2. *Suppose that U is an open subset of $\mathbb{R}^d$ of finite Lebesgue's measure. Then there exists a collection $\{Q_j\}_j$ of pairwise disjoint dyadic cubes such that $U=\bigcup_j Q_j$ and*

$$\frac{1}{2}\operatorname{dist}(Q_j,\mathbb{R}^d\setminus U)\le\operatorname{diam}Q_j.$$

PROOF. The desired collection is the family of maximal dyadic cubes contained in U. The first two properties are then obvious, and the estimate follows from the fact that for each Q from the collection, its parent intersects $\mathbb{R}^d\setminus U$. □

REMARK 4.3. Sometimes Whitney's decomposition refers to a splitting of the given open set U into cubes Q_j for which

$$c_1\operatorname{dist}(Q_j,\mathbb{R}^d\setminus U)\le\operatorname{diam}Q_j\le c_2\operatorname{dist}(Q_j,\mathbb{R}^d\setminus U),$$

for some given positive constants $c_1 < c_2$. Such a decomposition is obtained by considering the class of all dyadic cubes Q_j contained in U, which satisfy $\operatorname{diam} Q_j \le c_2 \operatorname{dist}(Q_j, \mathbb{R}^d \setminus U)$, and then choosing the maximal subcollection.

We return to the context of singular integral operators. Fix an arbitrary function $f \in L^1(\mathbb{R}^d)$, a positive number $\lambda > 0$ and consider the Whitney decomposition $\{Q_j\}_j$ of the open set $U_\lambda = \{T^* f > \lambda\}$. Then there are points $x_j \in \mathbb{R}^d \setminus U_\lambda$ such that the distance from x_j to Q_j is less than $2d_j$, where d_j is the diameter of Q_j. Denote by $\overline{Q_j}$ the cube centered at x_j, with diameter $20d_j$: then $Q_j^* \le \overline{Q_j}$.

We will establish the following good-λ inequality. Let $\mathcal{M}$ be the uncentered maximal operator.

LEMMA 4.4. *There is a constant C depending only on K and the dimension d, such that the following holds. For any $\gamma > 0$ and for any cube Q_i from the above Whitney decomposition, we have*

$$|\{x \in Q_i : T^* f(x) > 2\lambda \text{ and } \mathcal{M} f(x) \le \gamma\lambda\}| \le C\gamma |Q_i|. \tag{4.2}$$

PROOF. We may and do assume that $\mathcal{M} f(\xi_i) \le \gamma\lambda$ for at least one point $\xi_i \in Q_i$, since otherwise there is nothing to prove. We may also assume that γ is sufficiently small: indeed, the claim is obvious for $C \ge \gamma^{-1}$. We write $f = f_1 + f_2$, where $f_1 = f\chi_{\overline{Q_i}}$ and $f_2 = f\chi_{\mathbb{R}^d \setminus \overline{Q_i}}$. Because $\xi \in \overline{Q_i}$, we get

$$\frac{1}{|\overline{Q_i}|} \int_{\mathbb{R}^d} |f_1(y)| \mathrm{d}y = \frac{1}{|\overline{Q_i}|} \int_{\overline{Q_i}} |f(y)| \mathrm{d}y \le \mathcal{M} f(\xi_i) \le \gamma\lambda,$$

so that the weak-type inequality for T^* yields

$$|\{T^* f_1 > \lambda/2\}| \le \frac{2C}{\lambda} \int_{\mathbb{R}^d} |f_1(y)| \mathrm{d}y \le C\gamma |Q_i|. \tag{4.3}$$

Now we will prove that

$$T^* f_2(x) \le \lambda + C\gamma\lambda \quad \text{for } x \in Q_i. \tag{4.4}$$

To this end, we fix a cube Q_x centered at x and let Q_{x_i} be the cube of the same size, centered at x_i (this point lies outside U_λ but its distance from Q_i is not bigger than $2d_i$: see above). Then we have

$$\begin{aligned}
&\left| \int_{\mathbb{R}^d \setminus Q_x} K(x-y) f_2(y) \mathrm{d}y \right| \\
&\le \left| \int_{\mathbb{R}^d \setminus Q_{x_i}} K(x-y) f_2(y) \mathrm{d}y \right| + \int_{Q_{x_i} \Delta Q_x} |K(x-y)||f_2(y)| \mathrm{d}y \\
&\le \left| \int_{\mathbb{R}^d \setminus Q_{x_i}} K(x_i-y) f_2(y) \mathrm{d}y \right| \\
&\quad + \int_{\mathbb{R}^d \setminus Q_{x_i}} |K(x_i-y) - K(x-y)||f_2(y)| \mathrm{d}y + \int_{Q_{x_i} \Delta Q_x} |K(x-y)||f_2(y)| \mathrm{d}y \\
&= I_1 + I_2 + I_3.
\end{aligned}$$

Observe that $I_1 \le T^* f(x_i) \le \lambda$, since $x_i \notin U_\lambda$. Furthermore, we have $|(x_i - y) - (x-y)| = |x_i - x| < 3d_i < |x_i - y|/2$ for any $y \notin \overline{Q_i}$; consequently, by the condition

(iii) on the kernel, we see that

$$I_2 \leq \int_{\mathbb{R}^d \setminus \overline{Q_{x_i}}} \frac{2C|x_i - x|}{|x_i - y|^{d+1}} |f_2(y)| \mathrm{d}y.$$

Now, let $Q_{2^j x_i}$ be a cube of center x and the volume equal to $2^{jd}|\overline{Q_i}|$. We may write the telescoping sum

$$\begin{aligned}\int_{\mathbb{R}^d \setminus \overline{Q_{x_i}}} \frac{|x_i - x||f_2(y)|}{|x_i - y|^{d+1}} \mathrm{d}y &= \sum_{j=0}^{\infty} \int_{Q_{2^{j+1} x_i} \setminus Q_{2^j x_i}} \frac{|x_i - x||f_2(y)|}{|x_i - y|^{d+1}} \mathrm{d}y \\ &\lesssim \sum_{j=0}^{\infty} \frac{1}{|Q_{2^{j+1} x_i}|} \int_{Q_{2^{j+1} x_i} \setminus Q_{2^j x_i}} \frac{|x_i - x||f_2(y)|}{|x_i - y|} \mathrm{d}y \\ &\lesssim \sum_{j=0}^{\infty} \frac{1}{2^{j+1}} \cdot \frac{1}{|Q_{2^{j+1} x_i}|} \int_{Q_{2^{j+1} x_i}} |f_2(y)| \mathrm{d}y \leq \mathcal{M}f(\xi_i),\end{aligned}$$

since $\xi_i \in Q_i \subset Q_{2^{j+1} x_i}$. The term I_3 can be handled similarly and (4.4) follows. Combining this estimate with (4.3) gives

$$|\{x \in Q_i : T^* f(x) > \lambda/2 + \lambda + C\gamma\lambda, \;\; f^* \leq \gamma\lambda\}| \leq C\gamma |Q_i|,$$

which is the claim (for $C\gamma = 1/2$). □

We are ready for the weighted estimates.

THEOREM 4.5. *Let $1 < p < \infty$ and let w be a (full, non-dyadic) A_∞ weight. Then for any kernel K satisfying (i), (ii) and (iii) we have*

$$\|T^* f\|_{L^p(w)} \leq C_{p,w} \|\mathcal{M}f\|_{L^p(w)}.$$

In particular, if w belongs to (the full) A_p class, then $\|T^*\|_{L^p(w) \to L^p(w)} < \infty$.

PROOF. Fix $\lambda > 0$ and let $\{Q_j\}_j$ be the Whitney decomposition of the set $\{T^* f > \lambda\}$. Since w is in a full A_∞, it is also a dyadic weight, for an arbitrary dyadic lattice. Therefore, by (4.2) and Theorem 2.11 (d), we have

$$\begin{aligned}&\frac{w(\{x \in Q_i : T^* f(x) > 2\lambda, \mathcal{M}f(x) \leq \gamma\lambda\})}{w(Q_i)} \\ &\leq C' \left(\frac{|\{x \in Q_i : T^* f(x) > 2\lambda, \mathcal{M}f(x) \leq \gamma\lambda\}|}{|Q_i|} \right)^{\delta} \leq C\gamma^\delta.\end{aligned}$$

Therefore, summing over i, we get

$$w(T^* f > 2\lambda, \mathcal{M}f(x) \leq \gamma\lambda) \leq C\gamma^\delta w(T^* f > \lambda)$$

and hence

$$w(T^* f > 2\lambda) \leq C\gamma^\delta w(T^* f > \lambda) + w(\mathcal{M}f > \gamma\lambda).$$

Multiply both sides by $p\lambda^{p-1}$ and integrate over λ from 0 to infinity: as the result, we arrive at

$$2^{-p} \int_{\mathbb{R}^d} |T^* f|^p w \leq C\gamma^\delta \int_{\mathbb{R}^d} |T^* f|^p w + \gamma^{-p} \int_{\mathbb{R}^d} (\mathcal{M}f)^p w,$$

which is the claim, after picking γ sufficiently small. □

Now we will prove that the A_p condition is necessary for the validity of the weak L^p estimates for the class of Riesz transforms, given by

$$R_j f(x) = \frac{\Gamma\left(\frac{d+1}{2}\right)}{\pi^{\frac{d+1}{2}}} \,\mathrm{p.\,v.} \int_{\mathbb{R}^d} \frac{x_j - y_j}{|x-y|^{d+1}} f(y)\mathrm{d}y.$$

THEOREM 4.6. *Suppose that w is a weight on $\mathbb{R}^d$ and let $d \geq 2$, $1 \leq p < \infty$. Assume further that each of the Riesz transforms R_j is of weak type (p,p) with respect to w. Then w is an A_p weight.*

PROOF. Let $Q = [a_1 + m] \times [a_2 + m] \times [a_d + m]$ be an arbitrary cube in $\mathbb{R}^d$ and let f be a nonnegative function supported on Q, with nonzero integral. Consider another cube $Q' = [a_1 + m, a_1 + 2m] \times [a_2 + m, a_2 + 2m] \times [a_d + m, a_d + 2m]$. Observe that for each $x \in Q'$ and $y \in Q$ we have $x_j \geq y_j$ and therefore

$$\left|\sum_{j=1}^d R_j f(x)\right| = \frac{\Gamma\left(\frac{d+1}{2}\right)}{\pi^{\frac{d+1}{2}}} \sum_{j=1}^d \int_Q \frac{x_j - y_j}{|x-y|^{d+1}} f(y)\mathrm{d}y \geq \frac{\Gamma\left(\frac{d+1}{2}\right)}{\pi^{\frac{d+1}{2}}} \int_Q \frac{f(y)}{|x-y|^d}\mathrm{d}y.$$

Note that if $x \in Q'$ and $y \in Q$, then $|x-y| \leq 2\sqrt{d}m$ and hence $|x-y|^{-d} \geq (2\sqrt{d})^{-d}|Q|$. Putting $C_n = \Gamma\left(\frac{d+1}{2}\right)\pi^{-\frac{d+1}{2}}(2\sqrt{d})^{-d}$, we see that the above arguments yield

$$Q' \subset \left\{ x \in \mathbb{R}^d \,:\, \left|\sum_{j=1}^d R_j f(x)\right| > \alpha \right\}$$

for any $0 < \alpha < C_n \cdot \frac{1}{|Q|}\int_Q f$. We have assumed that R_j's satisfy the weak-type bound with respect to w (denote the corresponding uniform constant by C), so

$$w(Q') \leq \frac{C^p}{\alpha^p} \int_Q f(x)^p w(x)\mathrm{d}x,$$

for each α as above. Consequently,

$$\left(\frac{1}{|Q|}\int_Q f\right)^p \leq \frac{C_n^{-p}C^p}{w(Q')} \int_Q f(x)^p w(x)\mathrm{d}x.$$

The same estimate holds if we switch Q and Q': for any nonnegative function g supported on Q',

$$\left(\frac{1}{|Q'|}\int_{Q'} g\right)^p \leq \frac{C_n^{-p}C^p}{w(Q)} \int_{Q'} g(x)^p w(x)\mathrm{d}x.$$

In particular, if we take $g = \chi_{Q'}$, we obtain $w(Q) \leq C_n^p C^p w(Q')$. Plugging this into the previous bound, we get

$$\left(\frac{1}{|Q|}\int_Q f\right)^p \leq \frac{(C_n^{-p}C^p)^2}{w(Q)} \int_Q f(x)^p w(x)\mathrm{d}x.$$

This implies $w \in A_p$: see Exercise 8 in Chapter 2. □

Problems

1. Let T be a singular integral operator with the kernel K satisfying conditions (i), (ii) and (iii). Prove that for any $1 < p < \infty$ and any $\varepsilon > 0$ we have the estimate

$$\int_{\mathbb{R}^d} |T^* f|^p u \mathrm{d}x \leq C_{K,\varepsilon,p,d} \int_{\mathbb{R}^d} |f|^p \mathcal{M}(u^{1+\varepsilon})^{1/(1+\varepsilon)} \mathrm{d}x.$$

2. Let $f = g + b$ be the Calderón-Zygmund decomposition of f at a given height λ. Prove that for any A_1 weight w we have

$$\|g\|_{L^1(w)} \leq [w]_{A_1} \|f\|_{L^1(w)}, \qquad \text{and} \qquad \|b\|_{L^1(w)} \leq 2[w]_{A_1} \|f\|_{L^1(w)}.$$

3. For a fixed $1 < p < \infty$, let w be an A_p weight and let $v = w^{1-p'}$ be the dual to w. Prove that if a linear operator T is bounded on $L^p(w)$, then its adjoint T^* is bounded on $L^{p'}(v)$, and the two norms coincide.

4. Let T be a singular integral operator on $\mathbb{R}^d$. For $t \geq 1$, define

$$\varphi_T(t) = \sup_{w:[w]_{A_1} \leq t} \|T\|_{L^1(w) \to L^{1,\infty}(w)}.$$

Let w be an A_1 weight and define $\sigma = w^{-1}$.

a) For any nonnegative function g on $\mathbb{R}^d$ with $\|g\|_{L^2(\sigma)} = 1$, prove that there is Rg such that $g \leq Rg$, $\|Rg\|_{L^2(\sigma)} \leq 2$ and $[Rg]_{A_1} \leq 2\|M\|_{L^2(\sigma) \to L^2(\sigma)}$.

b) Prove that for any function f on $\mathbb{R}^d$ we have

$$\|Tf\|_{L^{2,\infty}(w)} \leq 2\varphi_T(2\|M\|_{L^2(\sigma) \to L^2(\sigma)}) \|f\|_{L^2(w)}.$$

5. Let w be an A_p weight and let T be a singular integral operator. Prove that for any $u \in L^{p'}_+(w)$, there is a nonnegative function U such that

$$\|U\|_{L^{p'}(w)} \leq \|u\|_{L^{p'}(w)}$$

and such that for each f,

$$\int_{\mathbb{R}^d} |Tf| uw \leq C \int_{\mathbb{R}^d} (Mf) Uw,$$

where C depends only on T and the dimension.

6. Let w be an A_p weight and let T be a positive operator, whose adjoint is bounded by maximal operator. Prove that for any $u \in L^{p'}_+(w)$, there is a nonnegative function U such that

$$\|U\|_{L^{p'}(w)} \leq \|u\|_{L^{p'}(w)}$$

and such that for each f,

$$\int_{\mathbb{R}^d} |Tf| uw \leq C \int_{\mathbb{R}^d} |f| Uw,$$

where C depends only on T and the dimension.

7. Let T be an operator acting on measurable functions on $\mathbb{R}^d$, such that $\|T\|_{L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d)} = O(p)$ as $p \to \infty$. Prove that if for any A_1 weight we have $\|T\|_{L^2(w) \to L^2(w)} \leq C[w]^{\alpha}_{A_1}$, then $\alpha \geq 1$.

8. Let H be the Hilbert transform on the real line. For any $1 < p < \infty$, prove that the estimate

$$\|Hf\|_{L^p(w)} \leq C_p \|Mf\|_{L^p(Mw)}$$

does not hold with any finite constant C_p depending only on p.

9. A collection $\{Q_j\}$ of dyadic cubes in $\mathbb{R}^d$ is called *sparse*, if there is a family $\{E(Q_j)\}$ of pairwise disjoint sets such that $E(Q_j) \subseteq Q_j$ and $|E(Q_j)| \geq |Q_j|/2$ for each j. Consider the associated operator T acting on locally integrable functions on $\mathbb{R}^d$ by the formula

$$Tf(x) = \sum_j \frac{1}{|Q_j|} \int_{Q_j} f \mathrm{d}x \cdot \chi_{Q_j}.$$

Prove that there is a constant C such that

$$\|Tf\|_{L^1(w)} \leq C\|f\|_{L^1(Mw)},$$

whenever f, w are locally integrable.

10. Suppose that U is an open subset of $\mathbb{R}^d$ of finite Lebesgue's measure. Then there exists a collection $\{Q_j\}_j$ of pairwise disjoint dyadic cubes such that $U = \bigcup_j Q_j$ and

$$\frac{1}{2}\operatorname{dist}(Q_j, \mathbb{R}^d \setminus U) \leq \operatorname{diam} Q_j \leq \operatorname{dist}(Q_j, \mathbb{R}^d \setminus U),$$

11. Let $1 < p < \infty$ and let w be a (full, non-dyadic) A_∞ weight.

(i) Show that for any kernel K satisfying (i), (ii) and (iii) we have

$$\|T^* f\|_{L^{p,\infty}(w)} \leq C_{p,w}\|\mathcal{M}f\|_{L^{p,\infty}(w)}.$$

(ii) Prove that if w belongs to A_p class, then $\|M\|_{L^{p,\infty}(w)\to L^{p,\infty}(w)} < \infty$.

CHAPTER 5

Weighted inequalities for martingales

Much of the material discussed so far concerned analytic weights and their properties. We start with the introduction of the corresponding objects in the probabilistic setting; as we shall see, there are certain obstacles which need to be overcome. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, equipped with $(\mathcal{F}_t)_{t\geq 0}$, a continuous-time, right-continuous filtration such that $\mathcal{F}_0$ contains only Ω, $\emptyset$ and all $\mathbb{P}$-null sets of $\mathcal{F}$. Furthermore, we assume that all adapted martingales have continuous trajectories. For instance, this is the case if $(\mathcal{F}_t)_{t\geq 0}$ is the natural filtration of Brownian motion. Note that the assumption on $\mathcal{F}_0$ implies that $\mathcal{F}_0$-measurable random variables are constant almost surely. For any adapted martingale $X = (X_t)_{t\geq 0}$, its maximal function is given by $X^* = \sup_{t\geq 0} |X_t|$. We will also use the notation $X_t^* = \sup_{0\leq s\leq t} |X_s|$ for the truncated maximal function. Any nonnegative and integrable random variable is called a weight. Any weight W gives rise to the associated measure on $(\Omega, \mathcal{F})$ (also denoted by W), given by $W(A) = \mathbb{E}\chi_A W$. It also gives rise to the martingale $(W_t)_{t\geq 0} = (\mathbb{E}(W|\mathcal{F}_t))_{t\geq 0}$, which, with no risk of confusion, will also be denoted by W.

1. Probabilistic A_p weights

As in the analytic setting, we start with the notion of Muckenhoupt's condition. Motivated by Remark 2.1, we propose the following.

Definition 5.1. (A_1) A weight $W = (W_t)_{t\geq 0}$ is said to satisfy Muckenhoupt's A_1 condition (belong to the A_1 class) if there is a finite deterministic constant C such that

$$W_t^* \leq CW_t \qquad \text{almost surely for all } t \geq 0.$$

The smallest C with the above property is denoted by $[W]_{A_1}$ and called the A_1 characteristic of W.

(A_p) Let $1 < p < \infty$. A weight $W = (W_t)_{t\geq 0}$ is said to satisfy Muckenhoupt's A_p condition (belong to the A_p class) if the weight $V = W^{-1/(p-1)}$ is integrable and there is a finite constant C such that

$$W_t V_t^{p-1} \leq C \qquad \text{almost surely for all } t \geq 0.$$

The smallest C with the above property is denoted by $[W]_{A_p}$ and called the A_p characteristic of W.

(A_∞) A weight $W = (W_t)_{t\geq 0}$ is said to satisfy Muckenhoupt's A_∞ condition (belong to the A_∞ class) if the weight $V = \log W^{-1}$ is integrable and there is a finite constant C such that

$$W_t \exp(V_t) \leq C \qquad \text{almost surely for all } t \geq 0.$$

The smallest C with the above property is denoted by $[W]_{A_\infty}$ and called the A_∞ characteristic of W.

Some comments and observations are in order.

REMARK 5.2. (i) The condition A_1 can be equivalently formulated as follows: there is a finite constant C such that for all t,

$$W_t \leq C \operatorname*{essinf}_{s\geq t} W_s \qquad \text{almost surely.} \tag{5.1}$$

Furthermore, the smallest C above is equal to $[W]_{A_1}$. To see this, suppose that W belongs to the A_1 class. Fix $t \geq 0$, $\varepsilon > 0$, $\omega \in \Omega$ and suppose that $u \geq t$ satisfies $W_u(\omega) \leq \operatorname{essinf}_{s\geq t} W_s(\omega) + \varepsilon$. Then

$$W_t(\omega) \leq W_u^*(\omega) \leq [W]_{A_1} W_u(\omega) \leq [W]_{A_1}(\operatorname*{essinf}_{s\geq t} W_s(\omega) + \varepsilon).$$

Since ε was arbitrary, (5.1) follows with $C = [W]_{A_1}$. To show the reverse implication, suppose that (5.1) holds. For fixed $t \geq 0$, $\varepsilon > 0$ and $\omega \in \Omega$, let $u \in [0, t]$ be a time parameter for which $W_u(\omega) \geq W_t^*(\omega) - \varepsilon$. Then

$$W_t^*(\omega) \leq W_u(\omega) + \varepsilon \leq C \operatorname*{essinf}_{s\geq u} W_s(\omega) + \varepsilon \leq C W_t(\omega) + \varepsilon$$

and it remains to let $\varepsilon \to 0$.

(ii) The A_p characteristic can be defined by

$$[W]_{A_p} := \operatorname*{esssup}_{\Omega} \sup_{t\geq 0} \mathbb{E}(W|\mathcal{F}_t) \left(\mathbb{E}(W^{-1/(p-1)}|\mathcal{F}_t)\right)^{p-1},$$

which is more common in the literature (sometimes the time t above is replaced by an arbitrary stopping time τ).

(iii) By Jensen's inequality, the reverse estimates to those defining the A_p conditions are true, with constant 1. Furthermore, if the A_p characteristic is equal to 1, then the weight is constant almost surely.

(iv) Again by Jensen's inequality, as in the analytic setting, we have $[W]_{A_p} \leq [W]_{A_q}$ whenever $1 \leq q \leq p \leq \infty$.

(v) As in Remark 2.1, it is convenient to interpret A_p weights as appropriate two-dimensional processes, as we will describe now. Any A_1 weight W can be identified with the pair (W, W^*) which takes values in the angle $\{(x, y) \in (0, \infty) : x \leq y \leq [W]_{A_1} x\}$. Furthermore, this pair has the following dynamics: when away from the diagonal $x = y$, the process moves horizontally "in a martingale manner"; if $W = W^*$, then the second coordinate might increase infinitesimally. For $1 < p < \infty$, we identify an A_p weight W with the martingale pair (W, V), which takes values in the hyperbolic strip $\{(x, y) \in (0, \infty)^2 : 1 \leq xy^{p-1} \leq [W]_{A_p}\}$. Finally, any A_∞ weight can be regarded as a martingale pair (W, V) taking values in the logarithmic strip $\{(x, y) \in (0, \infty) \times \mathbb{R} : 1 \leq xe^y \leq [W]_{A_\infty}\}$.

Now we will postpone the discussion on the further properties of probabilistic A_p weights. This will be taken up later, when we have introduced an important type of approach.

2. The Bellman function method

The Bellman function method is a powerful technique which can be used in the study of various semimartingale inequalities. We will concentrate on estimates for continuous-time processes, but it should be emphasized that the method originates from the discrete-time, where it is also fully understood.

The general concept is very simple. Suppose that D is some given open subset of $\mathbb{R}^d$ and let $\overline{D}$ denote its closure. Assume further that $\mathcal{V} : \overline{D} \to \mathbb{R}$ is a given function, K is some fixed constant and suppose that we are interested in showing the estimate

$$\mathbb{E}\mathcal{V}(S_t) \leq K \qquad \text{for all } t \geq 0, \tag{5.2}$$

where $S = (S_t)_{t\geq 0}$ is assumed to belong to the class $\mathcal{S}$ of continuous-time Markov processes taking values in $\overline{D}$, with some prescribed conditions on the infinitesimal generator (which describe the "evolution rules"). Before we proceed, let us list several important examples.

EXAMPLE 5.1. *1°: maximal inequalities.* Suppose that we are interested in Doob's maximal inequality

$$\mathbb{E}(X_t^*)^p - \left(\frac{p}{p-1}\right)^p \mathbb{E}|X_t|^p \leq 0, \qquad t \geq 0.$$

This corresponds to the choices $D = \{(x, y) \in \mathbb{R}^2 : y > |x|\}$, $\mathcal{V}(x, y) = y^p - \left(\frac{p}{p-1}\right)^p |x|^p$, $K = 0$ and $\mathcal{S}$ denotes the class of all processes of the form (X, X^*), where X is a martingale. Of course, the consideration of other functions V lead to other maximal inequalities.

2°: optimal stopping for Brownian motion. Assume that we are interested in proving Burkholder-Davis-Gundy inequalities

$$\mathbb{E}|B_\tau|^p \leq C_p \mathbb{E}\tau^{p/2}$$

for any stopping time τ of B. This corresponds to the choice $D = (0, \infty) \times \mathbb{R}$, $\mathcal{V}(x, y) = |y|^p - C_p x^{p/2}$, $K = 0$ and $\mathcal{S}$ is the class of the stopped processes $((\tau \wedge t, B_{\tau\wedge t}))_{t\geq 0}$, where τ ranges over all stopping times as above. As previously, different functions $\mathcal{V}$ lead to other optimal stopping problems. This example can be further generalized by replacing (τ, B^τ) by the pair $(\langle X\rangle, X)$, where X is an arbitrary continuous-time martingale and $\langle X\rangle$ is the associated square bracket. This choice yields the corresponding square-function estimates.

3°: estimates for stochastic integrals (martingale transforms). Suppose that X is a continuous-time martingale and let Y be the stochastic integral, with respect to X, of a certain predictable process $H = (H_t)_{t\geq 0}$ with values in $[-1, 1]$:

$$Y_t = H_0 X_0 + \int_{0+}^t H_s \mathrm{d}X_s, \qquad t \geq 0.$$

The problem of comparing the sizes of X and Y (measured in terms of various norms) has gained a lot of interest in the literature. For instance, fix $1 < p < \infty$ and consider the L^p estimate

$$\mathbb{E}|Y_t|^p - (p^* - 1)^p \mathbb{E}|X_t|^p \leq 0, \qquad t \geq 0,$$

where $p^* = \max\{p, p/(p-1)\}$. This estimate is of the form (5.2): we take $D = \mathbb{R}^2$, $\mathcal{V}(x, y) = |y|^p - (p^* - 1)^p |x|^p$ and $\mathcal{S}$ is the class of all X, Y as above.

4°: inequalities for A_p weights, $p = 1$. Assume that we want to prove the reverse Hölder inequality

$$\left(\mathbb{E}W_t^{1+\varepsilon}\right)^{1/(1+\varepsilon)} \leq C\mathbb{E}W_t, \qquad t \geq 0$$

where W is an A_1 weight and the constants C, ε depend on the A_1 characteristic of W. To see that the above approach works, we specify some information on the characteristic: we assume that $[W]_{A_1} \leq c$ and aim at finding C and ε depending on c. Furthermore, we may assume that the $\mathbb{E}W_t = W_0 = w_0$, the starting point of W is given, as long as the final bound is independent of this parameter. Then we set $D = \{(w, v) \in (0, \infty)^2 : w < v < cw\}$, $\mathcal{V}(w, v) = w^{1+\varepsilon}$ and $K = C^{1+\varepsilon} w_0^{1+\varepsilon}$. Obviously, $\mathcal{S}$ is the class of all pairs of the form (W, W^*), where W is an A_1 weight with $[W]_{A_1} \leq c$.

5°: inequalities for A_p weights, $1 < p < \infty$. The same approach works for other Muckenhoupt's classes. For instance, the choice $D = \{(w, v) \in (0, \infty)^2 : 1 \leq wv^{p-1} \leq c\}$ and $\mathcal{S}$ consisting of all processes of the form (W, V), where W is an A_p weight of characteristic not exceeding c and $V = W^{-1/(p-1)}$ lead to estimates for A_p weights. A similar choice gives the corresponding results in the case $p = \infty$ as well.

6°: weighted inequalities for maximal functions. This example is more complicated and shows that we might merge the contexts described above. Suppose that we want to prove the probabilistic version of Muckenhoupt's estimate:

$$\mathbb{E}(X_t^*)^p W_t - C_{p,w} \mathbb{E}|X_t|^p W_t \leq 0, \qquad t \geq 0,$$

where X is a martingale and W is an A_p weight (the constant depends only on the parameters indicated). To analyze this problem by means of (5.2), we first restrict ourselves to weights of bounded characteristic: $[W]_{A_p} \leq c$, and then consider the four-dimensional domain

$$D = \{(x, y, w, v) \in \mathbb{R} \times (0, \infty)^3 : |x| < y,\ 1 \leq wv^{p-1} \leq c\}.$$

Then Muckenhoupt's estimate corresponds to $\mathcal{V}(x, y, w, v) = y^p w - C_{p,w}|x|^p w$ and $K = 0$. (Although the variable v does not appear in the definition of V, we cannot get rid of it, as it is needed to control the behavior of the weight W.) The class $\mathcal{S}$ consists of all quadruples (X, X^*, W, V), where X is an arbitrary martingale, W is an A_p weight satisfying $[W]_{A_p} \leq c$ and $V = W^{-1/(p-1)}$ is the dual weight to W. Obvious modifications lead to the corresponding problems for $p = 1$ or $p = \infty$.

We return to the general context. Bellman function method rests on the following simple observation: to handle (5.2), it suffices to construct a function U on $\overline{D}$ such that

1° $U(S_0) \leq K$ almost surely;

2° $U \geq \mathcal{V}$ on $\overline{D}$;

3° for any $S \in \mathcal{S}$, the process $(U(S_t))_{t \geq 0}$ is a supermartingale.

Indeed, once we have constructed such an object, the inequality (5.2) follows:

$$\mathbb{E}\mathcal{V}(S_t) \leq \mathbb{E}U(S_t) \leq EU(S_0) \leq K.$$

How to check that a given function U satisfies 1°, 2° and 3°? Typically, the first two conditions are easy and reduce to some more or less elementary pointwise bounds. The main difficulty lies in the third property, which in most cases can be handled with the use of Itô's formula. To this end, one needs to make sure that U has sufficient regularity (for instance, it is of class C^2 and defined on some neighborhood of $\overline{D}$) and that the generator L of S satisfies $LU \leq 0$. This, in turn, in most cases means that U satisfies a certain partial differential inequality. Summarizing, in order to study (5.2), we need to construct a certain special function

U, which is neither too big nor too small (which is captured by 1° and 2°), and satisfies a certain concavity-type property (described by $LU \le 0$). As we shall see, in the search of such a function U, it is often fruitful to look at the condition $LU = 0$ and try to understand its geometric meaning.

Let us provide a comprehensive example. Suppose that we are interested in Doob's inequality

$$\mathbb{E}(X_t^*)^p - C_p \mathbb{E}|X_t|^p \le 0, \qquad t \ge 0,$$

for some constant C_p to be found; here X is an arbitrary L^p bounded continuous-time martingale. As we have already noted above, this corresponds to the choice $D = \{(x,y) \in \mathbb{R}^2 : y > |x|\}$, $\mathcal{V}(x,y) = y^p - C_p|x|^p$ and $K = 0$. We want to find a special function $U : \overline{D} \to \mathbb{R}$ such that

1° $U(x,|x|) \le 0$ for all $x \in \mathbb{R}$;
2° $U(x,y) \ge y^p - C_p|x|^p$ for all $x \in \mathbb{R}, y \ge 0$;
3° $(U(X_t, X_t^*))_{t\ge 0}$ is a martingale for any martingale X.

To understand what 3° means, we *assume* that U is of class C^2 and apply Itô's formula to obtain

$$U(X_t, X_t^*) = U(X_0, X_0^*) + \int_{0+}^t U_x(X_s, Y_s)\mathrm{d}X_s + I + \frac{II}{2},$$

where

$$I = \int_{0+}^t U_y(X_s, X_s^*)\mathrm{d}X_s^*, \qquad II = \int_{0+}^t U_{xx}(X_s, Y_s)\mathrm{d}\langle X\rangle_s.$$

The supermartingale property will follow if we show that I and II are nonincreasing processes. Note the measure dX_s^* increases on the set $\{s \ge 0 : X_s^* = |X_s|\}$, so I will have the monotonicity property if $U_y(x,|x|) \le 0$. Furthermore, II will have this property if $U_{xx}(x,y) \le 0$ for all (x,y). Summarizing, we will be done if we find a C^2 function for which

1° $U(x,|x|) \le 0$ for all $x \in \mathbb{R}$;
2° $U(x,y) \ge y^p - C_p|x|^p$ for all $x \in \mathbb{R}, y \ge 0$;
3° $U_y(x,|x|) \le 0$ and $U_{xx}(x,y) \le 0$.

This is still quite a difficult problem, but there are two additional observations which will simplify it considerably. First, in many situations it is fruitful that in the concavity condition 3° we have equality: $U_y(x,|x|) = 0$ and $U_{xx}(x,y) = 0$. Second, note that , Doob's inequality is homogeneous of order p, and V also has this property. Furthermore, V is symmetric with respect to x. It is plausible to conjecture that U should also enjoy these structural conditions:

$$U(x,y) = y^p \varphi\left(\frac{|x|}{|y|}\right) \qquad \text{for all } y \ge |x|$$

for some function $\varphi : [0,1] \to \mathbb{R}$ to be found. Let us translate the conditions 1°, 2° and (strengthened) 3° into the language of φ. We obtain four estimates: $\varphi(1) \le 0$, $\varphi(s) \ge 1 - C_p s^p$ for all $s \in [0,1]$, $p\varphi(1) - \varphi'(1) = 0$; $\varphi''(s) = 0$. So, the function φ is linear: $\varphi(s) = As + B$, and the inequality $p\varphi(1) - \varphi'(1) = 0$ means that $B = -\frac{p-1}{p}A$, so

$$\varphi(s) = A\left(s - \frac{p-1}{p}\right).$$

Let us take a look at the majorization: $A\left(s-\frac{p-1}{p}\right)\geq 1-C_ps^p$ for $s\in[0,1]$, or

$$C_p \geq \frac{1-A\left(s-\frac{p-1}{p}\right)}{s^p} \qquad \text{for } s\in(0,1].$$

Note that A must be nonnegative, since otherwise this majorization does not hold for $s=0$. The right-hand side, considered as a function of s, attains its maximum for $s_0 = 1+\frac{p}{A(p-1)}$, and this maximal value is

$$\frac{1}{(p-1)s_0^{p-1}(1-s_0)}. \tag{5.3}$$

Let us now *minimize* this expression over s_0; this will lead us to the choice of A for which the constant C_p is as small as the approach can give. A simple analysis of the derivative returns the value $s_0=(p-1)/p$, for which (5.3) is equal to $\left(\frac{p}{p-1}\right)^p$. So, we have obtained the function

$$U(x,y)=y^p\varphi\left(\frac{|x|}{y}\right)^p = py^{p-1}\left(y-\frac{p}{p-1}|x|\right).$$

Although the function U we discovered is not of class C^2, it is not difficult to check, using refined versions of Itô's formula, that the conditions 1°, 2° and 3° are satisfied and hence the estimate holds.

3. Reverse Hölder inequality/self-improving property for A_p weights

As we already know, any A_p weight w (for $1<p\leq\infty$) in the analytic setting is actually an $A_{p-\varepsilon}$ weight, for some $\varepsilon>0$ depending on p and the A_p characteristic of w. We will prove the probabilistic counterpart of this result, using Bellman function method. Let us start with the description of the appropriate adjustment of the approach. We pick an A_p weight W and we are interested in finding $r\in(1,p)$ and $C<\infty$ such that

$$W_T\mathbb{E}(W^{1/(1-r)}|\mathcal{F}_T)^{r-1}\leq C \tag{5.4}$$

almost surely for any T: this will mean that $[W]_{A_r}\leq C$. Let us restrict ourselves to weights W satisfying $[W]_{A_p}\leq c$, for some fixed $c\geq 1$. If $c=1$, then W is constant and there is nothing to prove; so, from now on, we assume that c is strictly bigger than 1. As we know, we can identify the weight with the two-dimensional martingale (W,V), where $V=W^{1/(1-p)}$, taking values in the domain $\overline{D}$, where $D=\{(x_1,x_2): 1<x_1x_2^{p-1}<c\}$. How to address (5.4) with the use of Bellman functions? Let us rewrite it in the form

$$\mathbb{E}(W^{1/(1-r)}|\mathcal{F}_T)\leq CW_T^{1/(1-r)}. \tag{5.5}$$

This inequality is a little reminiscent of (5.2), if we replace the expectation there with the conditional expectation with respect to $\mathcal{F}_T$ here; *this corresponds to the fact that the initial time will not be equal to zero, but to T).* Let us try to repeat the reasoning we have successfully applied above. We have already defined the domain D, the class $\mathcal{S}$ is also obvious - it consists of all A_p weights (W,V) with $[W]_{A_p}\leq c$. We also let $V(x_1,x_2)=x_1^{1/(1-r)}$ and $K=W_T^{1/(1-r)}$. Suppose we have successfully constructed a function $U:\overline{D}\to\mathbb{R}$ satisfying

1° $U(S_T)\leq K$ for all $S\in\mathcal{S}$;

2° $U(x_1,x_2) \geq x_1^{1/(1-r)}$ on $\overline{D}$;
3° $(U(S_T))_{t\geq T}$ is a supermartingale.

Then (5.5) will follow. Indeed, for any $t \geq T$ we have

$$\mathbb{E}(W_t^{1/(1-r)}|\mathcal{F}_T) \leq \mathbb{E}(U(W_t,V_t)|\mathcal{F}_T) \leq U(W_T,V_T) \leq W_T^{1/(1-r)}.$$

Let us comment on the properties 1°, 2° and 3°. The first two of them are actually pointwise bounds, the main difficulty lies in the understanding of the third one. As we have mentioned above, the class $\mathcal{S}$ consists of two-dimensional martingales moving in $\overline{D}$, with no other restrictions. Thus, by Itô's formula, we need to ensure that U is locally concave in $\overline{D}$:

$$D^2U = \begin{bmatrix} U_{x_1x_1} & U_{x_1x_2} \\ U_{x_1x_2} & U_{x_2x_2} \end{bmatrix} \leq 0. \tag{5.6}$$

5.1. On solutions to Monge–Ampère equation in planar domains. When treating the concavity conditions, it is often fruitful to assume that the corresponding differential inequalities actually become equalities. This general observation, applied to (5.6), leads us to the question about the structure of the solutions to the equation

$$\det D^2U(x_1,x_2) = U_{x_1x_1}(x_1,x_2)U_{x_2x_2}(x_1,x_2) - (U_{x_1x_2}(x_1,x_2))^2 = 0, \tag{5.7}$$

the so-called Monge-Ampère equation. We will establish the following statement.

THEOREM 5.1. *Suppose that D is an open connected subset of $\mathbb{R}^2$ and assume that $U : D \to \mathbb{R}$ is a C^2 function satisfying* (5.7) *such that its Hessian matrix is non-degenerate for each $(x,y) \in D$. Then the domain D can be foliated, i.e., it can be written as a union of pairwise disjoint line segments along which the function U is linear. Furthermore, the partial derivatives of U are constant on the leaves of the foliation.*

PROOF. Fix a point $\mathtt{x} \in D$. For each point (x_1,x_2) belonging to some neighborhood $\mathcal{V}$ of $\mathtt{x}$ there is a nonzero vector $v = v(x_1,x_2)$ such that $D^2U(x_1,x_2)v = 0$; we can assume that v forms a continuous vector field on $\mathcal{V}$. Let $\varphi = (\varphi_1,\varphi_2)$ be the integral curve of this vector field, passing through (x_1,x_2): that is, φ is a function defined on some interval I containing 0 such that $\varphi(0) = (x_1,x_2)$ and $D^2U(\varphi(t))\varphi'(t) = 0$ for each $t \in I$. Observe that U_{x_1} and U_{x_2} are constant along φ: we have

$$\frac{d}{dt}U_{x_1}(\varphi(t)) = U_{x_1x_1}(\varphi(t))\varphi_1'(t) + U_{x_1x_2}(\varphi(t))\varphi_2'(t) = (D^2U(\varphi(t))\varphi'(t))_1 = 0 \tag{5.8}$$

and

$$\frac{d}{dt}U_{x_2}(\varphi(t)) = U_{x_1x_2}(\varphi(t))\varphi_1'(t) + U_{x_2x_2}(\varphi(t))\varphi_2'(t) = (D^2U(\varphi(t))\varphi'(t))_2 = 0.$$

By the non-degeneration condition, there is a function F such that $U_{x_2}(x_1,x_2) = F(U_{x_1}(x_1,x_2))$. A direct differentiation yields

$$U_{x_1x_2}(x_1,x_2) = F'(U_{x_1}(x_1,x_2))U_{x_1x_1}(x_1,x_2),$$

which gives that on any integral curve of v the ratio $U_{x_1x_2}/U_{x_1x_1}$ is constant. Hence, by (5.8), the integral curves must be line segments. A similar calculation shows that the function $(x_1,x_2) \mapsto U(x_1,x_2) - x_1U_{x_1}(x_1,x_2) - x_2U_{x_2}(x_1,x_2)$ is also constant along φ. Indeed, if $\xi(t) = U(\varphi(t)) - \varphi_1(t)U_{x_1}(\varphi(t)) - \varphi_2(t)U_{x_2}(\varphi(t))$, then we easily check that

$$\xi'(t) = -\varphi_1(t) \cdot \frac{d}{dt}U_{x_1}(\varphi(t)) - \varphi_2(t) \cdot \frac{d}{dt}U_{x_2}(\varphi(t)) = 0.$$

Thus we have proved that

$$U(x_1, x_2) = U_{x_1}(x_1, x_2)x_1 + U_{x_2}(x_1, x_2)x_2 + c(x_1, x_2),$$

where c is some function which is constant on any integral curve of the vector field v. This is precisely the desired linearity property of U. ☐

5.2. Certain geometric parameters and the main result. To formulate the statement, we need some auxiliary geometric objects. Let $c \geq 1$ and $1 < p < \infty$ be fixed. Then the line, tangent to the curve $x_1x_2^{p-1} = c$ at the point $(1, c^{1/(p-1)})$, intersects the curve $x_1x_2^{p-1} = 1$ at one point (if $c = 1$) or two points (if $c > 1$). Take the intersection point with larger x_1-coordinate, and denote this coordinate by $1 + d(p, c)$. Formally, $d = d(p, c)$ is the unique number in $[0, p-1)$ satisfying the equation

$$c(1 + d)(p - 1 - d)^{p-1} = (p - 1)^{p-1}. \tag{5.9}$$

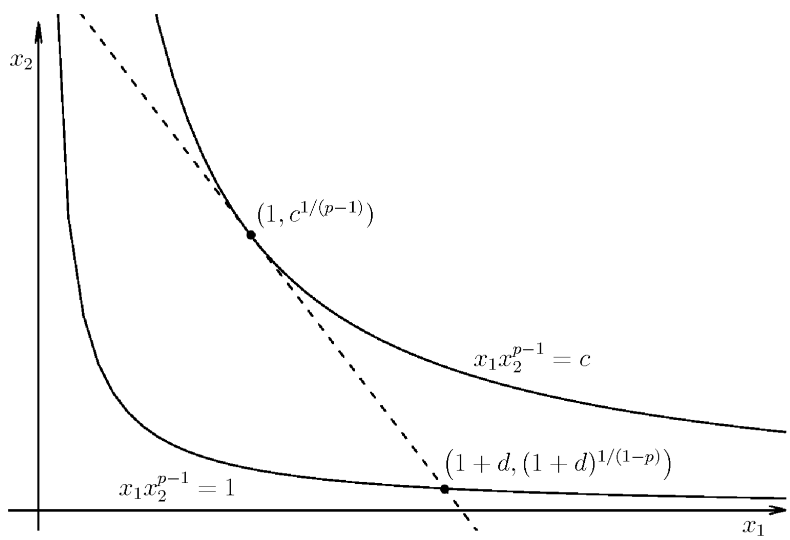

FIGURE 1. The geometric interpretation of the number $d = d(p, c)$.

We are ready for our main result.

THEOREM 5.2. *Let $1 < p < \infty$. Suppose that W is an A_p weight, put $c = [W]_{A_p}$ and let $d = d(p, c)$ be the positive constant given by (5.9). Then for $r \in (d + 1, p)$ we have*

$$[W]_{A_r} \leq \left(\frac{(1 + d)^{1/(1-r)}(1 - r)}{d + 1 - r} \right)^{r-1}.$$

The bound on the right can actually be shown to be the best possible. A little more careful analysis of (5.9) leads to the following corollary.

COROLLARY 5.3. *Suppose that W is an A_p weight. If we set $q = p - (p-1)(p[W]_{A_p})^{-1/(p-1)}/2$, then*

$$[W]_{A_q} \leq 2^{q-1}p^2[W]_{A_p}.$$

Let us also relate the inequality (5.4) with the reverse Hölder inequality. Suppose that W is an A_p weight and let $V = W^{1/(1-p)}$ stand for its dual: then $V \in A_{p'}$ and $[V]_{A_{p'}} = [W]_{A_p}^{1/(p-1)}$. By the self-improving inequality (5.5) applied to V, we get the existence of $r < p'$ and $C < \infty$ such that for all $T \geq 0$,

$$\mathbb{E}(V^{1/(1-r)}|\mathcal{F}_T) \leq CV_T^{1/(1-r)}.$$

Plugging $V = W^{1/(1-p)}$ and denoting $\delta = (p-1)^{-1}(r-1)^{-1} > 1$, we see that the above bound is equivalent to $\mathbb{E}(W^\delta|\mathcal{F}_T)^{1/\delta} \leq CW_T$. It remains to note that the inequality $r' < p$ implies $\delta > 1$.

5.3. A special function and its properties. We work in the domain

$$\overline{D} = \{(x_1, x_2) \in \mathbb{R}_+^2 : 1 \leq x_1x_2^{p-1} \leq c\},$$

foliated by the family of curves $\gamma_b = \{(x_1, x_2) \in \mathbb{R}_+^2 : x_1x_2^{p-1} = b\}$, $1 \leq b \leq c$. Then γ_1 and γ_c are the lower and the upper parts of the boundary of $\overline{D}$. We will also need the following geometrical object. For any $x > 0$, consider the line ℓ tangent to the upper boundary, passing through the point $(x, (c/x)^{1/(p-1)})$. This line has the equation

$$x_2 = -\frac{c^{1/(p-1)}x^{-p/(p-1)}}{p-1}x_1 + \frac{p}{p-1}\left(\frac{c}{x}\right)^{1/(p-1)} \tag{5.10}$$

and intersects the lower boundary at two points. Take the point with the larger x_1-coordinate: since the point lies on γ_1, its coordinates can be expressed in the form

$$(x(1+d), (x(1+d))^{1/(1-p)})$$

for some $d > 0$. It is straightforward to check that d does not depend on x and hence (by taking $x = 1$) it must be equal to $d(p, c)$ defined in (5.9). Let I_x be the line segment tangent to γ_c with the endpoints $(x, (c/x)^{1/(p-1)})$ and $(x(1+d), (x(1+d))^{1/(1-p)})$.

We are ready to introduce the special function. For a given $d + 1 < r < p$, consider $U : \overline{D} \to \mathbb{R}$ uniquely determined by the following three requirements:

(i) For any $(x_1, x_2) \in \gamma_c$ we have

$$U(x_1, x_2) = \frac{(1+d)^{1/(1-r)}(1-r)}{d+1-r}x_1^{1/(1-r)}.$$

(ii) For any $(x_1, x_2) \in \gamma_1$ we have

$$U(x_1, x_2) = x_1^{1/(1-r)}.$$

(iii) The function U is linear along any line segment I_x, $x > 0$.

It is not difficult to check that U is continuous and of class C^∞ in the interior of $\overline{D}$. Let us turn our attention to 1°, 2° and 3°.

LEMMA 5.4. *For any p, c and $r \in (d(p,c)+1, p)$, we have*

$$x_1^{1/(1-r)} \le U(x_1, x_2) \le \frac{(1+d)^{1/(1-r)}(1-r)}{d+1-r} x_1^{1/(1-r)}. \tag{5.11}$$

PROOF. Fix x_1 and let $y > 0$, $z > 0$ be chosen so that $(x_1, y) \in \gamma_c$ and $(x_1, z) \in \gamma_1$; then the desired inequality is equivalent to $U(x_1, z) \le U(x_1, x_2) \le U(x_1, y)$. Therefore, we will be done if we show that the function $t \mapsto U(x_1, t)$ is nondecreasing on the interval $\{t : (x_1, t) \in \overline{D}\} = [z, y]$. Pick $x_1/(1+d) < x < x_1$. Using the equation (5.10), we see that the point

$$P = \left(x_1, -\frac{c^{1/(p-1)} x^{-p/(p-1)}}{p-1} x_1 + \frac{p}{p-1} \left(\frac{c}{x}\right)^{1/(p-1)} \right)$$

lies on I_x and hence, by the definition of U,

$$\begin{aligned} U(P) &= \frac{x_1 - x}{dx} U\left(x(1+d), (x(1+d))^{1/(1-p)}\right) + \frac{x(1+d) - x_1}{dx} U\left(x, \left(\frac{c}{x}\right)^{1/(p-1)}\right) \\ &= \frac{x_1 - x}{dx} (x(1+d))^{1/(1-r)} + \frac{x(1+d) - x_1}{dx} \cdot \frac{(1+d)^{1/(1-r)}(1-r)}{d+1-r} x^{1/(1-r)}. \end{aligned}$$

Differentiating this equality with respect to x and calculating a little bit, we get

$$U_{x_2}(P) \frac{p}{(p-1)^2} \left(\frac{c}{x}\right)^{1/(p-1)} \cdot \frac{x_1 - x}{x^2} = \frac{r(x(1+d))^{1/(1-r)}}{(1-r)(d+1-r)} \cdot \frac{x_1 - x}{x^2}.$$

This gives the claim, since taking all x ranging from $x_1/(1+d)$ to x_1, we get all points from the interior of $\mathcal{D}_{p,c}$ with the first coordinate equal to x_1. □

The above lemma gives 1° and 2°. We only need to prove the third condition.

LEMMA 5.5. *For any p, c and $r \in (d(p,c)+1, p)$, the function U is locally concave, i.e., concave along any line segment entirely contained in $\overline{D}$.*

PROOF (SKETCH): By Sylvester's criterion, it is enough to check that for each point $x \in D$, the Hessian matrix $D^2U(x)$ is nonpositive-definite. We already know that the determinant of the matrix is zero: this follows from the fact that there is a line segment passing through x along which U is linear. Therefore, to establish the estimate $D^2U \le 0$, it is enough to check that $U_{x_2x_2} \le 0$. This can be done with the use of similar (but a little longer) calculations to those above. □

4. Martingale inequalities

5.1. Weighted inequalities for martingale maximal functions. We are ready to complicate things a little bit and study weak- and strong-type estimates for the maximal function X^* of a given continuous-path martingale X. We know already how to adjust the above general setup to this particular context. Fix $1 < p < \infty$, let $D = \{(x, y, w, v) \in \mathbb{R} \times (0, \infty) : |x| < y, 1 < wv^{p-1} < c\}$ and let $\mathcal{V} : \overline{D} \to \mathbb{R}$ be a given function. In order to prove the estimate

$$\mathbb{E}\mathcal{V}(X_t, X_t^*, W_t, V_t) \le 0, \qquad t \ge 0,$$

where $X = (X_t)_{t\ge0}$ is a continuous-path martingale and $W = (W_t)_{t\ge0}$ is an A_p weight satisfying $[W]_{A_p} \le c$ (and, as usual, V is the martingale generated by the dual weight $W^{1/(1-p)}$), it suffices to construct a C^2 function $U : \overline{D} \to \mathbb{R}$ such that

1° $U(x, |x|, w, v) \le 0$ for all $(x, |x|, w, v) \in D$;
2° $U \ge V$ on D;

3° $U(X, X^*, W, V)$ is a supermartingale for any (X, X^*, W, V) as above.

By Itô's formula, the requirement 3° will be satisfied if we ensure that

$$\int_0^t U_y(X_s, X_s^*, W_s, V_s)\mathrm{d}X_s^*, \qquad \int_0^t D^2_{x,w,v}U(X_s, X_s^*, W_s, V_s)\mathrm{d}[X, W, V]_s$$

are nonincreasing processes which, in turn, is guaranteed by the set of estimates

$$U_y \leq 0, \qquad D^2_{x,w,v}U \leq 0 \qquad \text{on } D.$$

Equipped with the above observations, we turn to the weak-type estimate

$$W(X^* > 1) \leq C_{p,W}\mathbb{E}|X|^p W, \qquad p \geq 1,$$

or rather its "local" version $W_t(X_t^* > 1) \leq C_{p,W}\mathbb{E}|X_t|^p W_t$. Suppose that $[W]_{A_p} \leq c$, then the bound corresponds to the choice $\mathcal{V}(x, y, w, v) = w1_{\{y>1\}} - C_{p,W}|x|^p w$. We will take $C_{p,W} = c$; the special function U is given by the formula

$$U(x, y, w, v) = \begin{cases} 0 & \text{if } y < 1, \\ w - c|x|^p v^{1-p} & \text{if } y \geq 1. \end{cases}$$

Observe that 1° holds: we have $U(x, |x|, w, v) = 0$ if $|x| < 1$ and, for remaining x,

$$U(x, |x|, w, v) = w - c|x|^p v^{1-p} \leq w - cv^{1-p} \leq 0.$$

The majorization is also straightforward: it is obvious for $y < 1$, while for $y \geq 0$ it is equivalent to $w - c|x|^p v^{1-p} \geq w - c|x|^p w$, or $wv^{p-1} \geq 1$. Note that an even stronger bound $U(x, y, w, v) \geq 1_{\{y \geq 1\}} - c|x|^p v^{1-p}$ is satisfied. Finally, we have $U_y = 0$ on D, $D^2_{x,w,v} = 0$ for $y < 1$ and

$$D^2_{x,y,w}U = \begin{bmatrix} -cp(p-1)|x|^{p-2}v^{1-p} & 0 & cp(p-1)|x|^{p-1}\operatorname{sgn} x \cdot v^{-p} \\ 0 & 0 & 0 \\ cp(p-1)|x|^{p-1}\operatorname{sgn} x \cdot v^{-p} & 0 & -cp(p-1)|x|^p v^{-1-p} \end{bmatrix} \leq 0$$

for the remaining y. Thus, U meets all the requirements and the estimate holds.

Actually, there is an obstacle, since U is not of class C^1 (even continuous): there is a singularity at $y = 1$, but it can be handled with the use of auxiliary stopping arguments. Let us briefly discuss this. Consider the stopping time $\tau = \inf\{t : |X_t| > 1\}$ and consider Itô's (or Itô-Tanaka, for $1 < p < 2$) formula on the time interval $[\tau, \tau \vee t]$, for some given $t > 0$. On this time interval we have $X^* \geq 1$, so (X, X^*, W, V) takes values in the set where U is sufficiently regular (and hence the application of the formula is permitted). By the above considerations, we get

$$\mathbb{E}U(X_{\tau\vee t}, X^*_{\tau\vee t}, W_{\tau\vee t}, V_{\tau\vee t})1_{\{\tau<\infty\}} \leq \mathbb{E}U(X_\tau, X^*_\tau, W_\tau, V_\tau)1_{\{\tau<\infty\}}.$$

Note that $|X_\tau| = X^*_\tau$, so the right-hand side is nonpositive. Applying the stronger version of 2° to the left-hand side, we obtain

$$W_{\tau\vee t}(\tau < \infty, X^*_{\tau\vee t} > 1) - c\mathbb{E}|X_{\tau\vee t}|^p V^{1-p}_{\tau\vee t} \leq 0.$$

Now, the function $(x, v) \mapsto |x|^p v^{1-p}$ is convex (we checked this above). Consequently, we have $\mathbb{E}|X_{\tau\vee t}|^p V^{1-p}_{\tau\vee t} \leq \mathbb{E}|X|^p V^{1-p} = \mathbb{E}|X|^p W$ and it remains to let $t \to \infty$ to get the claim, by virtue of Fatou's lemma.

Summarizing, we have shown the following result.

Theorem 5.1. *Suppose that X is a continuous-path uniformly integrable martingale and W is an A_p weight. Then*

$$\|X^*\|_{L^{p,\infty}(W)} \leq [W]^{1/p}_{A_p}\|X\|_{L^p(W)}.$$

Now we will prove the corresponding strong-type bound. First we will show it for a slightly more restrictive class of weights.

THEOREM 5.2. *For any $1 < q < p < \infty$, any continuous-path uniformly integrable martingale X and any A_q weight W we have the estimate*

$$||X^*||_{L^p(W)} \leq \left(\frac{p}{p-q}\right)^{1/q} [W]_{A_q}^{1/q} ||X||_{L^p(W)}.$$

PROOF. Consider the domain $D = \{(x,y,w,v) : |x| < y, 1 < wv^{q-1} < c\}$. The strong-type inequality corresponds to the choice $\mathcal{V}(x,y,w,v) = y^p w - \left(\frac{p}{p-q}\right)^{p/q} c^{p/q} x^p w$; actually, for convenience, we will take a slightly different function, with an additional multiplicative constant:

$$\mathcal{V}(x,y,w,v) = \frac{p}{q}\left(y^p w - \left(\frac{p}{p-q}\right)^{p/q} c^{p/q} x^p w\right).$$

Let $U : \overline{D} \to \mathbb{R}$ be given by the formula

$$U(x,y,w,v) = y^p w - \frac{p}{p-q} c|x|^q y^{p-q} v^{1-q}$$

To check 1°, note that

$$U(x,|x|,w,v) = |x|^p v^{1-q}\left[wv^{q-1} - \frac{p}{p-q}c\right] \leq |x|^p v^{1-q}\left[c - \frac{p}{p-q}c\right] \leq 0.$$

The majorization 2° follows at once from the mean-value property of the convex function $s \mapsto s^{p/q}$: indeed, after the substitution $x' = \left(\frac{p}{p-q}c\right)^{1/q} x$, the inequality is equivalent to $y^{p/q} - (x')^{p/q} \leq (p/q)y^{p/q-1}(y - x')$. We turn our attention to the concavity condition. First note that

$$U_y(x,|x|,w,v) = p|x|^{p-1}v^{1-q}\left(wv^{q-1} - c\right) \leq 0.$$

The requirement on the Hessian is evident (it is equivalent to saying that the function $(x,v) \mapsto |x|^q v^{1-q}$ is convex). This gives the desired bound. □

To complete the proof of the strong-type estimate in the case when $W \in A_p$, we need to use Corollary 5.3 and combine it with the above statement. As the result, we get

$$||X^*||_{L^p(W)} \leq \frac{p^{(3p-2)/(p-1)}}{p-1}[W]_{A_p}^{1/(p-1)}||X||_{L^p(W)}.$$

5.2. Square function inequalities. Now we will focus on weak- and strong-type estimates for martingale square bracket $[X,X]$ (or rather its quadratic root). Recall that $[X,X]_t$ is the limit in probability of the sums

$$\sum_{k=0}^{k_n} |X_{t_{k+1}^{(n)}} - X_{t_k^{(n)}}|^2,$$

where $(t_k^{(n)})_{k=0}^{k_n}$ is the sequence of partitions of the interval $[0,t]$ with mesh converging to zero. In particular, specifying X to be a stopped Brownian motion $B^\tau = (B_{\tau\wedge t})_{t\geq 0}$, we will prove weighted estimates between τ and B_τ.

As previously, we discuss the adjustment of the general approach to this particular context. If we want to study estimates for A_p weights W satisfying $[W]_{A_p} \leq c$, we need to consider the region $D = \{(x, y, w, v) \in \mathbb{R} \times (0, \infty)^3 : 1 < wv^{p-1} < c\}$. The inequality

$$\mathbb{E}\mathcal{V}(X_t, [X, X]_t, W_t, V_t) \leq 0, \qquad t \geq 0,\ [W]_{A_p} \leq c,$$

will follow if we prove the existence of a function U satisfying the following conditions.

1° $U(x, x^2, w, v) \leq 0$;
2° $U \geq \mathcal{V}$ on D;
3° $U(X, [X, X], W, V)$ is a supermartingale for all X, W, V as above.

By Itô's formula, the last requirement will be satisfied if the matrix

$$\begin{bmatrix} U_{xx} + 2U_y & U_{xw} & U_{wv} \\ U_{wx} & U_{ww} & U_{wv} \\ U_{vx} & U_{vw} & U_{vv} \end{bmatrix}$$

is negative semidefinite.

Our starting point is the following auxiliary estimate.

LEMMA 5.3. *For any $1 < q < 2$ and any A_q weight W we have*

$$\|[X, X]^{1/2}\|_{L^2(W)} \leq \sqrt{\frac{q}{2-q}}[W]_{A_q}^{1/2}\|X\|_{L^2(W)}. \tag{5.12}$$

PROOF. Let $c = [W]_{A_q}$ and consider the domain $D = \{(x, y, w, v) \in \mathbb{R} \times (0, \infty)^3 : 1 < wv^{q-1} < c\}$. The estimate corresponds to the choice $\mathcal{V}(x, y, w, v) = yw - \frac{q}{2-q}cx^2w$. Consider the function $U(x, y, w, v) = yw - \frac{q}{2-q}cx^2v^{1-q}$. Then

$$U(x, x^2, w, v) = x^2v^{1-q}\left(wv^{q-1} - \frac{q}{2-q}c\right) \leq 0.$$

The majorization 2° is equivalent to $wv^{q-1} \geq 1$. The concavity requirement becomes

$$\begin{bmatrix} -\frac{2qc}{2-q}v^{1-q} + w & 0 & -\frac{2q(1-q)c}{2-q}xv^{-q} \\ 0 & 0 & 0 \\ -\frac{2q(1-q)c}{2-q}xv^{-q} & 0 & -\frac{q^2(q-1)c}{2-q}x^2v^{-q-1} \end{bmatrix}.$$

Since $wv^{q-1} \leq c$, the entry in the upper-left corner is not bigger than $\frac{(2-3q)c}{2-q}v^{1-q}$ and, to check that the matrix is negative semidefinite, it is enough to show that the determinant of the reduced matrix

$$\begin{bmatrix} \frac{(2-3q)c}{2-q}v^{1-q} & -\frac{2q(1-q)c}{2-q}xv^{-q} \\ -\frac{2q(1-q)c}{2-q}xv^{-q} & -\frac{q^2(q-1)c}{2-q}x^2v^{-q-1} \end{bmatrix}$$

is nonnegative. After some straightforward manipulations, this is equivalent to $q \leq 2$. Therefore, the estimate holds. □

REMARK 5.4. Actually, a slightly stronger estimate can be extracted from the above reasoning. Namely, if we skip the majorization (equivalently, set $V := U$), then Bellman function method shows that for each t,

$$\mathbb{E}[X, X]_tW_t \leq \frac{q}{2-q}[W]_{A_q}\mathbb{E}X_t^2V_t^{1-q}. \tag{5.13}$$

The same holds if we replace t with an arbitrary stopping time τ.

The above estimate yields weighted weak-type bounds for $[X,X]^{1/2}$, as demonstrated below.

THEOREM 5.5. *For any* $1<p<2$ *we have*

$$W([X,X]>1)\leq \frac{2}{2-p}[W]_{A_p}\mathbb{E}|X|^pW. \tag{5.14}$$

PROOF. Consider an auxiliary stopping time $\tau=\inf\{t:|X_t|\geq 1\}$ and write

$$W([X,X]\geq 1)\leq W([X,X]\geq 1,\,\tau<\infty)+W([X,X]\geq 1,\,\tau=\infty).$$

By the weighted weak-type inequality for the maximal function, we get

$$W([X,X]\geq 1,\,\tau<\infty)\leq W(\tau<\infty)=W(X^*>1)\leq [W]_{A_p}\mathbb{E}|X|^pW.$$

On the other hand, by (5.13) and Chebyshev's inequality,

$$\begin{aligned} W([X,X]\geq 1,\,\tau=\infty)&=W([X^\tau,X^\tau]\geq 1,\,\tau=\infty)\\ &\leq W([X^\tau,X^\tau]\geq 1)\\ &\leq \mathbb{E}[X^\tau,X^\tau]W\\ &\leq \frac{p}{2-p}[W]_{A_p}\mathbb{E}|X^\tau|^2V_\tau^{1-p}\leq \frac{p}{2-p}[W]_{A_p}\mathbb{E}|X^\tau|^pV_\tau^{1-p}. \end{aligned}$$

Now, (X_τ,V_τ) is the conditional expectation of (X,V) with respect to $\mathcal{F}_\tau$ and the function $(x,v)\mapsto |x|^pv^{1-p}$ is convex. Consequently,

$$W([X,X]\geq 1,\,\tau=\infty)\leq \frac{p}{2-p}[W]_{A_p}\mathbb{E}|X|^pV^{1-p}=\frac{p}{2-p}[W]_{A_p}\mathbb{E}|X|^pW.$$

Combining the above facts we get the assertion. □

A similar reasoning leads to the following version for $p=2$.

THEOREM 5.6. *For any* $W\in A_2$,

$$W([X,X]>1)\leq C[W]_{A_2}^2\mathbb{E}|X|^2W.$$

PROOF. Let $W\in A_2$. We start as previously, defining τ and splitting $W([X,X]>1)$ into two parts. The part

$$W([X,X]\geq 1,\,\tau<\infty)\leq W(\tau<\infty)=W(X^*>1)\leq [W]_{A_2}\mathbb{E}|X|^2W$$

is handled as before. Now we make use of Corollary 5.3: we have $W\in A_p$, with $p=2-(4[W]_{A_2})^{-1}$. Consequently, the above arguments give

$$W([X,X]\geq 1,\,\tau=\infty)\leq \frac{p}{2-p}[W]_{A_p}\mathbb{E}|X^\tau|^2V_\tau^{1-p}\lesssim [W]_{A_p}^2\mathbb{E}|X|^2W.$$

The proof is complete. □

REMARK 5.7. The linear dependence on $[W]_{A_p}$ in (5.14) is optimal. For $p=2$, the above argument can be improved to get the dependence of order $[W]_{A_2}\log^{1/2}([W]_{A_2}+1)$. The question whether this result is optimal is open.

For $p>2$, we can study the weak-type estimates with the use of extrapolation. As we have shown above, given $p_0\in(1,2)$, we have

$$\|[X,X]^{1/2}\|_{L^{p_0,\infty}}\leq C[W]_{A_{p_0}}^{1/p_0}\|X\|_{L^{p_0}}.$$

Consequently, extrapolating from p_0, we get that for each $p>p_0$ we have

$$\|[X,X]^{1/2}\|_{L^{p,\infty}}\leq C_p[W]_{A_p}^{1/p_0}\|X\|_{L^p}.$$

The dependence $[W]_{A_p}^{1/p_0}$ can be improved to $[W]_{A_p}^{1/2}$ for $p > 2$, but we will not present the detailed discussion here.

We turn to the strong-type inequalities. We will first prove the appropriate weighted L^p bounds between $[X, X]$ and X^*. Our starting point is the following statement.

LEMMA 5.8. *Suppose that X is a continuous-path martingale starting from* 0. *Then for any $\beta > 1$ and δ, $\lambda > 0$ we have*

$$\mathbb{P}([X,X]^{1/2} > \beta\lambda,\ X^* \le \delta\lambda) \le \frac{\delta^2}{\beta^2 - 1}\mathbb{P}([X,X]^{1/2} > \lambda) \tag{5.15}$$

and

$$\mathbb{P}(X^* > \beta\lambda,\ [X,X]^{1/2} \le \delta\lambda) \le \frac{\delta^2}{(\beta - 1)^2}\mathbb{P}(X^* > \lambda). \tag{5.16}$$

PROOF OF (5.15). We may assume that X is bounded, by an appropriate localization. Introduce the stopping times

$$\begin{aligned} \mu &= \inf\{t : [X,X]_t^{1/2} > \lambda\}, \\ \nu &= \inf\{t : [X,X]_t^{1/2} > \beta\lambda\}, \\ \sigma &= \inf\{t : |X_t| > \delta\lambda\}. \end{aligned}$$

Then we obviously have

$$\begin{aligned} \mathbb{P}([X,X]^{1/2} > \beta\lambda,\ X^* \le \delta\lambda) &= \mathbb{P}(\mu \le \nu < \infty,\ \sigma = \infty) \\ &= \mathbb{P}([X,X]_{\nu\wedge\sigma} - [X,X]_{\mu\wedge\sigma} \ge \beta^2\lambda^2 - \lambda^2) \\ &\le \frac{1}{(\beta^2 - 1)\lambda^2}\mathbb{E}([X,X]_{\nu\wedge\sigma} - [X,X]_{\mu\wedge\sigma}). \end{aligned}$$

On the set $\{\mu = \infty\}$, we have $\nu = \infty$ and hence $[X,X]_{\nu\wedge\sigma} - [X,X]_{\mu\wedge\sigma} = 0$. In addition, since X is bounded, the process $X^2 - [X,X]$ is a martingale. Consequently, we may write

$$\begin{aligned} \mathbb{E}([X,X]_{\nu\wedge\sigma} - [X,X]_{\mu\wedge\sigma}) &= \mathbb{E}([X,X]_{\nu\wedge\sigma} - [X,X]_{\mu\wedge\sigma})1_{\{\mu<\infty\}} \\ &= \mathbb{E}\Big[\mathbb{E}\big([X,X]_{\nu\wedge\sigma} - [X,X]_{\mu\wedge\sigma}\big|\mathcal{F}_\mu\big)\Big]1_{\{\mu<\infty\}} \\ &= \mathbb{E}\Big[\mathbb{E}\big(X_{\nu\wedge\sigma}^2 - X_{\mu\wedge\sigma}^2\big|\mathcal{F}_\mu\big)\Big]1_{\{\mu<\infty\}} \\ &= \mathbb{E}\big(X_{\nu\wedge\sigma}^2 - X_{\mu\wedge\sigma}^2\big)1_{\{\mu<\infty\}} \\ &\le \mathbb{E}X_{\nu\wedge\sigma}^2 1_{\{\mu<\infty\}} \le \delta^2\mathbb{P}(\mu < \infty), \end{aligned}$$

where the last estimate follows from the definition of σ. Putting all the facts together, we get the claim. □

PROOF OF (5.16). The reasoning is similar. Assume that X is bounded and introduce the stopping times

$$\begin{aligned} \mu &= \inf\{t : |X_t| > \lambda\}, \\ \nu &= \inf\{t : |X_t| > \beta\lambda\}, \\ \sigma &= \inf\{t : [X,X]_t^{1/2} > \delta\lambda\}. \end{aligned}$$

Observe that

$$\begin{aligned}\mathbb{P}(X^* > \beta\lambda,\, [X,X]^{1/2} \leq \delta\lambda) &= \mathbb{P}(\mu \leq \nu < \infty,\, \sigma = \infty) \\ &= \mathbb{P}(X_{\nu\wedge\sigma} - X_{\mu\wedge\sigma} \geq (\beta-1)\lambda) \\ &\leq \frac{1}{(\beta-1)^2\lambda^2}\mathbb{E}(X_{\nu\wedge\sigma} - X_{\mu\wedge\sigma})^2.\end{aligned}$$

Now, on the set $\{\mu = \infty\}$ we have $\nu = \infty$ and $X_{\nu\wedge\sigma} - X_{\mu\wedge\sigma} = 0$. Since $X^2 - [X,X]$ is a mean-zero martingale, an argument similar to that above yields

$$\begin{aligned}\mathbb{E}(X_{\nu\wedge\sigma} - X_{\mu\wedge\sigma})^2 &= \mathbb{E}(X_{\nu\wedge\sigma} - X_{\mu\wedge\sigma})^2 1_{\{\mu<\infty\}} \\ &= \mathbb{E}\Big[\mathbb{E}\big((X_{\nu\wedge\sigma} - X_{\mu\wedge\sigma})^2 | \mathcal{F}_\mu\big)\Big] 1_{\{\mu<\infty\}} \\ &= \mathbb{E}\Big[\mathbb{E}\big(X_{\nu\wedge\sigma}^2 - X_{\mu\wedge\sigma}^2 | \mathcal{F}_\mu\big)\Big] 1_{\{\mu<\infty\}} \\ &\leq \mathbb{E}\Big[\mathbb{E}\big(X_{\nu\wedge\sigma}^2 | \mathcal{F}_\mu\big)\Big] 1_{\{\mu<\infty\}} \\ &= \mathbb{E}[X,X]_{\nu\wedge\sigma} 1_{\{\mu<\infty\}} \leq \delta^2\mathbb{P}(\mu < \infty).\end{aligned}$$

Combining this with the previous estimates gives the assertion. □

Now, if W is an arbitrary A_∞ weight, then W satisfies the reverse Hölder inequality: there are constants C and $\varepsilon > 0$ depending on W such that

$$\big(\mathbb{E}W^{1+\varepsilon}\big)^{1/(1+\varepsilon)} \leq C\mathbb{E}W.$$

See (5.5) and Exercise 4 below. Let μ be a stopping time. It is easy to check that $W1_{\mu<\infty}$ is an A_∞ weight on a modified probability space $(\{\mu < \infty\}, \mathcal{F}, \mathbb{P}/\mathbb{P}(\mu < \infty))$ equipped with the filtration $\mathcal{G}_t = \mathcal{F}_{\mu\wedge t}$, $t \geq 0$. Consequently, for any event E contained in $\{\mu < \infty\}$, we see that

$$\begin{aligned}\frac{W(E)}{\mathbb{P}(\mu<\infty)} = \frac{\mathbb{E}W1_{\{\mu<\infty\}}1_E}{\mathbb{P}(\mu<\infty)} &\leq \left(\frac{\mathbb{E}(W1_{\{\mu<\infty\}})^{1+\varepsilon}}{\mathbb{P}(\mu<\infty)}\right)^{1/(1+\varepsilon)} \left(\frac{\mathbb{P}(E)}{\mathbb{P}(\mu<\infty)}\right)^{\varepsilon/(1+\varepsilon)} \\ &\leq C\frac{\mathbb{E}W1_{\{\mu<\infty\}}}{\mathbb{P}(\mu<\infty)} \cdot \left(\frac{\mathbb{P}(E)}{\mathbb{P}(\mu<\infty)}\right)^{\varepsilon/(1+\varepsilon)},\end{aligned}$$

that is,

$$W(E) \leq CW(\mu<\infty) \cdot \left(\frac{\mathbb{P}(E)}{\mathbb{P}(\mu<\infty)}\right)^{\varepsilon/(1+\varepsilon)}. \tag{5.17}$$

This observation, together with the preceding lemma, lead to the following result.

THEOREM 5.9. *Suppose that W is an A_∞ weight and X is an arbitrary martingale. Then for any $0 < p < \infty$ we have*

$$c_{p,W}^{-1}\|[X,X]^{1/2}\|_{L^p(W)} \leq \|X^*\|_{L^p(W)} \leq c_{p,W}\|[X,X]^{1/2}\|_{L^p(W)},$$

where the constant $c_{p,W}$ is finite and depends only on the parameters indicated.

PROOF. We will only prove the left estimate, the right one can be established in a similar manner. We may assume that X starts from 0, replacing it with $X - X_0$

if necessary: then

$$\begin{aligned}\|[X,X]^{1/2}\|_{L^p(W)} &\leq \|X_0\|_{L^p(W)} + \|[X-X_0,X-X_0]\|_{L^p(W)}\\ &\leq \|X^*\|_{L^p(W)} + c_{p,W}\|(X-X_0)^*\|_{L^p(W)}\\ &\leq (1+2c_{p,W})\|X^*\|_{L^p(W)}.\end{aligned}$$

Let $\beta > 1$, let δ, $\lambda > 0$ and let μ be the stopping time as in the proof of (5.15) above. The event

$$E = \{[X,X]^{1/2} > \beta\lambda,\ X^* \leq \delta\lambda\}$$

is contained in $\{\mu < \infty\}$ and hence (5.17) yields

$$W(E) \leq CW(\mu<\infty)\cdot\left(\frac{\mathbb{P}(E)}{\mathbb{P}(\mu<\infty)}\right)^{\varepsilon/(1+\varepsilon)}.$$

Combining this with (5.15), we get

$$W([X,X]^{1/2} > \beta\lambda,\ X^* \leq \delta\lambda) \leq CW([X,X]^{1/2} > \lambda)\cdot\left(\frac{\delta^2}{\beta^2-1}\right)^{\varepsilon/(1+\varepsilon)}.$$

Therefore, we may write

$$\begin{aligned}W([X,X]^{1/2} > \beta\lambda) &\leq W(X^* > \delta\lambda) + W([X,X]^{1/2} > \beta\lambda,\ X^* \leq \delta\lambda)\\ &\leq W(X^* > \delta\lambda) + CW([X,X]^{1/2} > \lambda)\cdot\left(\frac{\delta^2}{\beta^2-1}\right)^{\varepsilon/(1+\varepsilon)}.\end{aligned}$$

Multiply both sides by $p\lambda^{p-1}$ and integrate over λ from 0 to ∞. As the result, we get

$$\beta^{-p}\|[X,X]^{1/2}\|^p_{L^p(W)} \leq \delta^{-p}\|X^*\|^p_{L^p(W)} + C\left(\frac{\delta^2}{\beta^2-1}\right)^{\varepsilon/(1+\varepsilon)}\|[X,X]^{1/2}\|^p_{L^p(W)}.$$

Now we move the last term on the right to the left-hand side. If δ is sufficiently small (so that $\beta^{-p} > C(\delta^2/(\beta^2-1))^{\varepsilon/(1+\varepsilon)}$), then the estimate can be rewritten in the form

$$\|[X,X]^{1/2}\|^p_{L^p(W)} \leq \frac{\delta^{-p}}{\beta^{-p} - C\left(\frac{\delta^2}{\beta^2-1}\right)^{\varepsilon/(1+\varepsilon)}}\|X^*\|^p_{L^p(W)}.$$

The proof is complete. □

The above statement, combined with the weighted estimates for the martingale maximal function, immediately gives the following consequence, the desired class of estimates for square functions.

COROLLARY 5.10. *Let X be a continuous-path martingale.*
(i) If W is an A_∞ weight, then for any $0 < p < \infty$,

$$\|X\|_{L^p(W)} \leq c_{p,W}\|[X,X]^{1/2}\|_{L^p(W)}.$$

(ii) If $1 < p < \infty$ and W is an A_p weight, then

$$\|[X,X]^{1/2}\|_{L^p(W)} \leq c_{p,W}\|X\|_{L^p(W)}.$$

REMARK 5.11. In analogy to the previous considerations, one may ask about the optimal dependence on the A_p characteristic of the underlying weight. It can be proved that

$$\|[X,X]^{1/2}\|_{L^p(W)} \leq c_p[W]_{A_p}^{\max\{1/2,1/(p-1)\}}\|X\|_{L^p(W)}.$$

For some partial results in this direction, see Exercise 7 below.

5.3. Inequalities for stochastic integrals/martingale transforms. The next context we study concerns the case when X is an arbitrary continuous-path martingale and Y is the stochastic integral, with respect to X, of some predictable process H with values in $[-1,1]$:

$$Y_t = Y_0H_0 + \int_0^t H_s \mathrm{d}X_s, \qquad t \geq 0.$$

We will deal with the weak-type estimates first. We start with an auxiliary statement.

LEMMA 5.12. *Let W be an A_p weight for some $1 \leq p < 2$ and let $\alpha \in (0,1)$. Then for X and Y as above, we have*

$$\|Y\|_{L^2(W^\alpha)} \leq \frac{K_p[W]_{A_p}^{1/2}}{1-\alpha}\|X\|_{L^2(W^\alpha)}.$$

PROOF. This will be proved with the use of Bellman function method. Let us restrict ourselves to weights satisfying $[W]_{A_p} \leq c$. Up to some standard limiting arguments, the claim is equivalent to

$$\mathbb{E}\mathcal{V}(X_t,Y_t,W_t,V_t) \leq 0, \qquad t \geq 0,$$

where $\mathcal{V}$ is a function on the domain

$$D = \{(x,y,w,v) : 1 \leq wv^{p-1} \leq c\},$$

given by the formula $\mathcal{V}(x,y,w,v) = |y|^2w^\alpha - \frac{Kc}{1-\alpha}|x|^2w^\alpha$. We need to find a function U which satisfies

1° $U(x,y,w,v) \leq 0$ provided $|y| \leq |x|$,
2° $U \geq V$ on D,
3° $U(X,Y,W,V)$ is a supermartingale.

The last requirement, by Itô's formula, will follow if we ensure that for any $(x,y,w,v) \in D$ and any h, k, r, s with $|k| \leq |h|$, the function $t \mapsto U(x+th, y+tk, w+tr, v+ts)$ is concave: here t is such that $(x+th, y+tk, w+tr, v+ts) \in D$. What is the differential version of this property? If we write $k = \alpha h$, $\gamma \in [-1,1]$, this gives rise to the following requirement for Hessian matrix:

$$\begin{bmatrix} U_{xx}+2\gamma U_{xy}+\gamma^2U_{yy} & U_{xw}+\gamma U_{yw} & U_{xv}+\gamma U_{yv} \\ U_{wx}+\gamma U_{wy} & U_{ww} & Uwv \\ U_{vx}+\gamma U_{vy} & U_{vw} & U_{vv} \end{bmatrix} \leq 0.$$

Note that if $U_{yy} \geq 0$, then it is enough to check the above condition for $\gamma = \pm 1$. Indeed, having checked this, we can majorize the above matrix from above by a convex combination of the matrices corresponding to $\gamma = \pm 1$.

The function U is given by $U(x,y,w,v) = y^2w^\alpha - c^2x^2v^{(1-p)\alpha}$. The first two requirements 1°, 2° are straightforward and their verification is left to the reader.

To check 3°, we study the terms $u_1(x,y,w,v) = y^2w^\alpha$ and $u_2(x,y,w,v) = x^2v^{(1-p)\alpha}$ separately. The matrix corresponding to u_1 equals

$$\begin{bmatrix} 2\gamma^2 w^\alpha & 2\alpha\gamma y w^{\alpha-1} & 0 \\ 2\alpha\gamma y w^{\alpha-1} & \alpha(\alpha-1)y^2w^{\alpha-2} & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Note that if the entry in the upper-left corner was equal to $4\alpha(\alpha-1)^{-1}\gamma^2w^\alpha$, then this matrix would be negative semidefinite. However, the term $2\gamma^2w^\alpha$ is bigger (it is even positive), and the required "compensation" will come from the function u_2. The Hessian corresponding to this function is given by

$$\begin{bmatrix} 2v^{(1-p)\alpha} & 0 & 2(1-p)\alpha x v^{(1-p)\alpha-1} \\ 0 & 0 & 0 \\ 2(1-p)\alpha x v^{(1-p)\alpha-1} & 0 & (p-1)\alpha\big((p-1)\alpha+1\big)x^2v^{(1-p)\alpha-2} \end{bmatrix}.$$

Now, observe that the modified matrix

$$\begin{bmatrix} \frac{4(p-1)\alpha}{(p-1)\alpha+1}v^{(1-p)\alpha} & 0 & 2(1-p)\alpha x v^{(1-p)\alpha-1} \\ 0 & 0 & 0 \\ 2(1-p)\alpha x v^{(1-p)\alpha-1} & 0 & (p-1)\alpha\big((p-1)\alpha+1\big)x^2v^{(1-p)\alpha-2} \end{bmatrix}$$

is positive semidefinite (the trace is positive, the first and the third row are linearly dependent). Comparing this to the previous matrix, we see the difference

$$2v^{(1-p)\alpha} - \frac{4(p-1)\alpha}{(p-1)\alpha+1}v^{(1-p)\alpha} = \frac{2(1-(p-1)\alpha)}{1+(p-1)\alpha}v^{(p-1)\alpha} > 0$$

on the first coordinate. Therefore, the previous matrix is also positive semidefinite, actually, we have the surplus in the upper-left entry which can be used to handle the Hessian of u_1. Indeed, it is enough to observe that

$$\begin{aligned} \frac{Kc}{1-\alpha}\frac{2(1-(p-1)\alpha)}{1+(p-1)\alpha}v^{(p-1)\alpha} &= \frac{K}{1-\alpha}\frac{2(1-(p-1)\alpha)}{1+(p-1)\alpha}\cdot cv^{(p-1)\alpha} \\ &\geq 4\alpha(1-\alpha)^{-1}\gamma^2w^\alpha \end{aligned}$$

for some K depending only on p: for instance, $K = 4/(2-p)$ suffices. For this choice of K, the Hessians of u_1 and u_2 match appropriately and produce a function U which satisfies the condition 3°. □

REMARK 5.13. Actually, U satisfies the stronger domination condition: we have

$$U(x,y,w,v) \geq y^2w^\alpha - \frac{Kc}{1-\alpha}x^2v^{(1-p)\alpha},$$

and therefore Bellman function method gives the slightly stronger bound

$$\mathbb{E}Y_t^2W_t^\alpha \leq \frac{Kc}{1-\alpha}\mathbb{E}X_t^2V_t^{(1-p)\alpha}, \qquad t \geq 0.$$

This is also true if we replace t with an arbitrary stopping time.

Now we will prove the weak-type bound between X and Y.

THEOREM 5.14. *For any $1 < p < 2$, X, Y as above and $W \in A_p$ we have*

$$W(Y^* > 1) \leq K_p[W]_{A_p}^2\mathbb{E}|X|^pW. \tag{5.18}$$

Proof. Fix $T \geq 0$. Arguing as in the proof of the analogous bound for the square function, we may assume that $X \equiv 0$. Introduce the stopping time $\tau = \inf\{t : |X_t| \geq 1\}$ and write

$$W(|Y_T| \geq 1) \leq W(|Y_T| \geq 1, \, \tau < \infty) + W(|Y_T| \geq 1, \, \tau = \infty).$$

By the weighted weak-type inequality for the maximal function, we get

$$W(|Y_T| \geq 1, \, \tau < \infty) \leq W(\tau < \infty) = W(X^* > 1) \leq [W]_{A_p} \mathbb{E}|X|^p W.$$

On the other hand, by (5.13) and Chebyshev's inequality,

$$W(|Y_T| \geq 1, \, \tau = \infty) = W(|Y_{\tau \wedge T}| \geq 1, \, \tau = \infty) \leq W(|Y_{\tau \wedge T}| \geq 1) \leq \mathbb{E}|Y_{\tau \wedge T}|^2 W.$$

Now, by Exercise 9 below, we have $W = U^\alpha$ for some $U \in A_p$ and $\alpha \in (0,1)$ satisfying $1 - \alpha = c_p/[W]_{A_p}$. By Jensen's inequality and (5.13),

$$\mathbb{E}|Y_{\tau \wedge T}|^2 W \leq \mathbb{E}|Y_{\tau \wedge T}|^2 U^\alpha_{\tau \wedge T} \leq K_p [W]^2_{A_p} \mathbb{E} X^2_{\tau \wedge T} V^{(1-p)\alpha}_{\tau \wedge T},$$

where V is the weight conjugate to U. Applying Jensen's inequality again (this time to the convex function $(x, v) \mapsto x^p v^{(1-p)\alpha}$), we obtain

$$\begin{aligned} \mathbb{E}|Y_{\tau \wedge T}|^2 W \leq K_p [W]^2_{A_p} \mathbb{E} X^2_\tau V_\tau^{(1-p)\alpha} &\leq K_p [W]^2_{A_p} \mathbb{E}|X_\tau|^p V_\tau^{(1-p)\alpha} \\ &\leq K_p [W]^2_{A_p} \mathbb{E}|X|^p V^{(1-p)\alpha} \\ &= K_p [W]^2_{A_p} \mathbb{E}|X|^p W. \end{aligned}$$

Putting the above facts together, we obtain $W(|Y_T| \geq 1) \leq K_p [W]^2_{A_p} \mathbb{E}|X|^p W$. To pass from Y_T to Y^* above, consider the stopping time $\sigma = \inf\{t \geq 0 : |Y_t| > 1\}$. On the set $\{\sigma < \infty\}$ we have $|Y_\sigma| = 1$ and hence

$$W(Y^* > 1) = W(\sigma < \infty) \leq W(|Y_\sigma| = 1) \leq \lim_{T \to \infty} W(|Y_{\sigma \wedge T}| \geq 1).$$

It remains to note that the stopped process Y^σ is a stochastic integral, with respect to X, of a predictable process H' given by $H'_t = H_t$ if $t \leq \sigma$ and $H'_t = 0$ elsewhere. Therefore $W(|Y_{\sigma \wedge T}| \geq 1) \leq K_p [W]^2_{A_p} \mathbb{E}|X|^p W$ and letting $T \to \infty$ yields the claim. □

Remark 5.15. A stronger statement can be proved: for $1 < p < \infty$ and X, Y, W as above, we have

$$W(Y^* > 1) \leq K_p [W]^p_{A_p} \mathbb{E}|X|^p W,$$

and the exponent p is the best possible.

We turn to the strong-type estimates. We will need the following good-λ inequality for stochastic integrals.

Lemma 5.16. *Let X, Y be as above. Then for any $\beta > 1$ and δ, $\lambda > 0$,*

$$\mathbb{P}(Y^* > \beta\lambda, \, X^* \leq \delta\lambda) \leq \frac{\delta^2}{\beta^2 - 1} \mathbb{P}(Y^* > \lambda). \tag{5.19}$$

Proof. We may assume that X is bounded, by an appropriate localization. Consider the stopping times

$$\begin{aligned} \mu &= \inf\{t : |Y_t| > \lambda\}, \\ \nu &= \inf\{t : |Y_t| > \beta\lambda\}, \\ \sigma &= \inf\{t : |X_t| > \delta\lambda\}. \end{aligned}$$

Then we obviously have

$$\begin{aligned}\mathbb{P}(Y^* > \beta\lambda,\, X^* \leq \delta\lambda) &= \mathbb{P}(\mu \leq \nu < \infty,\, \sigma = \infty)\\ &= \mathbb{P}(|Y_{\nu\wedge\sigma}| - |Y_{\mu\wedge\sigma}| \geq \beta^2\lambda^2 - \lambda^2)\\ &\leq \mathbb{P}(|Y_{\nu\wedge\sigma} - Y_{\mu\wedge\sigma}| \geq (\beta^2-1)\lambda^2)\\ &\leq \frac{1}{(\beta^2-1)\lambda^2}\mathbb{E}(Y_{\nu\wedge\sigma} - Y_{\mu\wedge\sigma})^2.\end{aligned}$$

On the set $\{\mu = \infty\}$, we have $\nu = \infty$ and hence $Y_{\nu\wedge\sigma} - Y_{\mu\wedge\sigma} = 0$. In addition, since X is bounded and H is bounded by 1, the process $Y^2 - X^2$ is a supermartingale. Consequently,

$$\begin{aligned}\mathbb{E}(Y_{\nu\wedge\sigma} - Y_{\mu\wedge\sigma})^2 &= \mathbb{E}(Y_{\nu\wedge\sigma} - Y_{\mu\wedge\sigma})^2 1_{\{\mu<\infty\}}\\ &= \mathbb{E}\Big[\mathbb{E}\big((Y_{\nu\wedge\sigma} - Y_{\mu\wedge\sigma})^2 \big| \mathcal{F}_\mu\big)\Big] 1_{\{\mu<\infty\}}\\ &\leq \mathbb{E}\Big[\mathbb{E}\big((X_{\nu\wedge\sigma} - X_{\mu\wedge\sigma})^2 \big| \mathcal{F}_\mu\big)\Big] 1_{\{\mu<\infty\}}\\ &= \mathbb{E}(X_{\nu\wedge\sigma} - X_{\mu\wedge\sigma})^2 1_{\{\mu<\infty\}}\\ &= \mathbb{E}\big(X^2_{\nu\wedge\sigma} - X^2_{\mu\wedge\sigma}\big) 1_{\{\mu<\infty\}}\\ &\leq \mathbb{E}X^2_{\nu\wedge\sigma} 1_{\{\mu<\infty\}} \leq \delta^2\mathbb{P}(\mu<\infty),\end{aligned}$$

where the last passage is due to the definition of σ. Putting all the facts together, we get the claim. □

THEOREM 5.17. *For $0 < p < \infty$, any X, Y as above and any A_∞ weight W, we have*

$$\|Y^*\|_{L^p(W)} \leq K_{p,W}\|X^*\|_{L^p(W)}.$$

PROOF. Arguing as for the square-function estimates, we show that the good-λ inequality (5.19) extends to the case of A_∞ weights and hence

$$\begin{aligned}W(Y^* > \beta\lambda) &\leq W(X^* > \delta\lambda) + W(Y^* > \beta\lambda,\, X^* \leq \delta\lambda)\\ &\leq W(X^* > \delta\lambda) + CW(Y^* > \lambda)\cdot\left(\frac{\delta^2}{\beta^2-1}\right)^{\varepsilon/(1+\varepsilon)},\end{aligned}$$

where C, ε are the constants coming from the reverse-Hölder inequality for W. It remains to multiply both sides by $p\lambda^{p-1}$, integrate over λ from 0 to ∞, rearrange terms to obtain the claim (for appropriately chosen β and δ). We omit the straightforward details. □

COROLLARY 5.18. *If $1 < p < \infty$, $W \in A_p$ and X, Y are as above, then*

$$\|Y\|_{L^p(W)} \leq K_{p,W}\|X\|_{L^p(W)}.$$

REMARK 5.19. It can be proved (with a significant effort), that the optimal dependence on the A_p characteristic is given by

$$\|Y\|_{L^p(W)} \leq K_p[W]_{A_p}^{\max\{1,1/(p-1)\}}\|X\|_{L^p(W)}.$$

5. Problems

1. Suppose that W is an A_p weight. Find the best constant in the inequality

$$\mathbb{E}(W_t^*)^p \leq C\mathbb{E}W_t^p, \qquad t \geq 0.$$

2. Using the Bellman function method find, for any $c \geq 1$ and $1 < p < \infty$, the best constant $C = C(p, c)$ such that the following holds. If W is an A_1 weight with $[W]_{A_1} \leq c$ and $\mathbb{E}W = 1$, then $\mathbb{E}W_t^p \leq C$ for all $t \geq 0$.

3. For any $c \geq 1$ and $1 < p < \infty$, find the best constant $C(\lambda, c)$ such that the following holds. If W is an A_2 weight satisfying $[W]_{A_2} \leq c$ and $\mathbb{E}W = 1$, then $\mathbb{E}W_t^p \leq C$ for all $t \geq 0$.

4. Using Bellman function method, prove the analogue of Theorem 5.2 for $p = \infty$.

5. Using Bellman function method, prove the estimate

$$\|X^*\|_{L^p(W)} \leq \frac{p[W]_{A_1}}{p-1}\|X\|_{L^p(W)}, \qquad 1 < p < \infty.$$

6. Using Bellman function method, prove the estimate

$$\|[X, X]^{1/2}\|_{L^{1,\infty}(W)} \leq [W]_{A_1}\|X\|_{L^1(W)}.$$

7. Prove that for $W \in A_2$ and any martingale X we have

$$\|[X, X]\|_{L^2(W)} \leq c[W]_{A_2}^{1/2}\|X\|_{L^2(W)}.$$

Prove that the exponent $1/2$ is the best possible.

8. Prove that for $W \in A_1$ and any martingale X we have

$$\|[X, X]^{1/2}\|_{L^1(W)} \leq C[W]_{A_1}^{1/2}\|X\|_{L^1(W)}.$$

Prove that the exponent $1/2$ is the best possible.

9. Prove that for any $1 < p < \infty$ there is a constant $c_p > 0$ depending only on p such that $W^{1+c_p/[W]_{A_p}} \in A_p$.

10. Let $W \in A_1$. Show that if Y is the stochastic integral, with respect to X, of some predictable process with values in $[-1, 1]$, then

$$\|Y\|_{L^p(W)} \leq c_p[W]_{A_1}^{\alpha_p}\|X\|_{L^p(W)}$$

for some exponent α_p. (It can be shown that $\alpha_p = 1$ is the best possible). Deduce the weak-type estimate

$$\|Y\|_{L^{1,\infty}(W)} \leq C[W]_{A_1}\log(1 + [W]_{A_1})\|X\|_{L^1(W)}.$$

11. Prove that

$$\|[X, X]^{1/2}\|_{L^1(W)} \leq [W]_{A_1}^{1/2}\|X^*\|_{L^1(W)}.$$

12. Prove that if X, Y are as above, then

$$\|Y\|_{L^1(W)} \leq [W]_{A_1}^{1/2}\|X^*\|_{L^1(W)}.$$

CHAPTER 6

Inequalities for sparse and fractional operators

In this chapter we will study weighted estimates for the dyadic counterparts of Calderón-Zygmund operators and Riesz potentials. Recall that a Calderón-Zygmund operator is an $L^2(\mathbb{R}^d)$ bounded linear operator, associated with a standard kernel K, i.e., given by the formula

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

for $f \in L^2_c(\mathbb{R}^d)$ and $x \notin \operatorname{supp} f$. The collection of fractional integral operators (or Riesz potentials) is defined for $0 < \alpha < d$ by

$$I_\alpha f(x) = \int_{\mathbb{R}^d} \frac{f(y)}{|x-y|^{d-\alpha}}\mathrm{d}y.$$

Let us turn our attention to the appropriate discrete versions of these objects. Let $\mathcal{D}$ be the dyadic grid in $\mathbb{R}^d$. Define the fractional maximal operator

$$M_\alpha f(x) = \sup_{Q\in\mathcal{D}} \frac{1}{|Q|^{1-\frac{\alpha}{d}}} \int_Q |f|\mathrm{d}y,$$

where the supremum is taken over all dyadic cubes Q containing x. Note that if $\alpha = 0$, then this is the usual dyadic maximal operator we worked with above.

To introduce the appropriate versions of T and I_α, we need the following notion.

Definition 6.1. A collection $\{Q_j\}_{j\in\mathcal{S}}$ of dyadic cubes in $\mathbb{R}^d$ is called *sparse*, if there is a family $\{E(Q_j)\}_{j\in\mathcal{S}}$ of pairwise disjoint sets such that $E(Q_j) \subseteq Q_j$ and $|E(Q_j)| \geq |Q_j|/2$ for each j.

Given a dyadic grid $\mathcal{D}$ and $0 < \alpha < d$, the fractional integral operator is

$$I_\alpha^{\mathcal{D}} f := \sum_{Q\in\mathcal{D}} |Q|^{\alpha/d-1} \int_Q f\mathrm{d}x \cdot \chi_Q.$$

To describe the connection between I_α and $I_\alpha^{\mathcal{D}}$, consider the dyadic grids

$$\mathcal{D}^t = \{2^k([0,1)^d + m + (-1)^k t) : k \in \mathbb{Z},\, m \in \mathbb{Z}^d\}, \quad t \in \{0, 1/3\}^d.$$

We have the following fact.

Lemma 6.2. *Given any cube Q in $\mathbb{R}^n$, there exists $t \in \{0,1/3\}^d$ and a cube $Q_t \in \mathcal{D}^t$ such that $Q \subseteq Q_t$ and $\ell(Q_t) \leq 6\ell(Q)$.*

Proof. We will prove the claim for the case $d = 1$ only. Let us understand the structure of $\mathcal{D}^0$ and $\mathcal{D}^{1/3}$. For any $k \in \mathbb{Z}$, the k-th generation of $\mathcal{D}^0$ is the interval $[0, 2^k)$ and its translations by multiplicities of 2^k. On the other hand, the k-th generation of $\mathcal{D}^{1/3}$ is the interval $\left[\frac{(-2)^k}{3}, 2^k + \frac{(-2)^k}{3}\right)$ and its translations by multiplicities of 2^k. Note that the interval $[0, 2^k)$ and $\left[\frac{(-2)^k}{3}, 2^k + \frac{(-2)^k}{3}\right)$ have nontrivial intersection, whose length is equal to $\frac{2}{3} \cdot 2^k$, i.e., $\frac{2}{3}$ of the length of each interval. By translation, we see the following property of $\mathcal{D}^0$ and $\mathcal{D}^{1/3}$: for any $I \in \mathcal{D}^0$ there is $J \in \mathcal{D}^{1/3}$ with $\ell(I) = \ell(J)$ such that $\inf J < \inf I$ and $\ell(I \cap J) \geq \ell(I)/3$, $\ell(J \setminus I) \geq \ell(I)/3$.

Now, pick an arbitrary cube (interval) Q and let k be the smallest integer such that $\ell(Q) \leq 2^k/3$; note that then $2^{k-1}/3 < \ell(Q)$, i.e., $2^k \leq 6\ell(Q)$. Take $I \in \mathcal{D}^0$, $\ell(I) = 2^k$, such that $I \cap Q \neq \emptyset$ and $\sup I \geq \sup Q$. If $Q \subset I$, then we are done: $\ell(I) = 2^k \leq 6\ell(Q)$, as we have shown above. If Q is not contained in I, then let $J \in \mathcal{D}^{1/3}$ be the interval guaranteed by the above reasoning (i.e., satisfying $\ell(I) = \ell(J)$, $\inf J < \inf I$, etc.). It suffices to note that

$$\inf J \leq \inf I - \frac{1}{3}\ell(I) \leq \sup Q - \ell(Q) = \inf Q$$

and

$$\sup J \geq \inf I + \frac{1}{3}\ell(I) \geq \sup Q - \ell(Q) + \frac{1}{3}|I| \geq \sup Q.$$

Thus J has all the required properties. The claim is proved. □

Equipped with the above statement, we can prove the following.

THEOREM 6.3. *Given $0 < \alpha < d$ and $f \geq 0$, we have*

$$I_\alpha f(x) \lesssim \max_{t \in \{0,1/3\}^d} I_\alpha^{\mathcal{D}} f(x).$$

PROOF. Let $Q(x, r)$ be the cube of side $2r$ centered at x. We have

$$I_\alpha f(x) \leq 2^{d-\alpha} \sum_{k \in \mathbb{Z}} (2^{-k})^{d-\alpha} \int_{Q(x,2^k)} f(y) \mathrm{d}y.$$

Indeed:

$$\begin{aligned} I_\alpha f(x) &= \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|^{d-\alpha}} \mathrm{d}y \\ &= \sum_{k \in \mathbb{Z}} \int_{Q(x,2^{k+1}) \setminus Q(x,2^k)} \frac{f(y)}{|x-y|^{d-\alpha}} \mathrm{d}y \\ &\leq \sum_{k \in \mathbb{Z}} (2^{-k})^{d-\alpha} \int_{Q(x,2^{k+1}) \setminus Q(x,2^k)} f(y) \mathrm{d}y \\ &\leq \sum_{k \in \mathbb{Z}} (2^{-k})^{d-\alpha} \int_{Q(x,2^{k+1})} f(y) \mathrm{d}y \\ &= 2^{n-\alpha} \sum_{k \in \mathbb{Z}} (2^{-k})^{d-\alpha} \int_{Q(x,2^k)} f(y) \mathrm{d}y. \end{aligned}$$

By the previous lemma, for each $k \in \mathbb{Z}$ there exists $t \in \{0, 1/3\}^d$ and $Q_t \in \mathcal{D}^t$ such that $Q(x, 2^k) \subseteq Q_t$ and

$$2^{k+1} = \ell(Q(x, 2^k)) \leq \ell(Q_t) \leq 6\ell(Q(x, 2^k)) = 12 \cdot 2^k.$$

But $\ell(Q_t) = 2^j$ for some integer j, so $2^{k+1} \leq \ell(Q_t) \leq 2^{k+3}$. Consequently,

$$\begin{aligned} I_\alpha f(x) &\leq 2^{d-\alpha} \sum_{k\in\mathbb{Z}} (2^{-k})^{d-\alpha} \int_{Q(x,2^k)} f(y)\mathrm{d}y \\ &\lesssim \sum_{k\in\mathbb{Z}} \sum_{t\in\{0,1/3\}^d} \sum_{\substack{Q\in\mathcal{D}^t,\\ 2^{k+1}\leq\ell(Q)\leq 2^{k+3}}} \frac{1}{|Q|^{1-\alpha/d}} \int_Q f(y)\mathrm{d}y \cdot \chi_Q(x) \\ &\lesssim \sum_{t\in\{0,1/3\}^d} \sum_{Q\in\mathcal{D}^t} \frac{1}{|Q|^{1-\alpha/d}} \int_Q f(y)\mathrm{d}y \cdot \chi_Q(x) \\ &\lesssim \max_{t\in\{0,1/3\}^d} I_\alpha^{\mathcal{D}^t} f(x). \end{aligned}$$

This completes the proof. □

Thus, the problem of handling I_α is reduced to that of controlling $I_\alpha^{\mathcal{D}}$. The operator $I_\alpha^{\mathcal{D}}$ can be further decomposed with the use of sparse families. Let $f \geq 0$ be a bounded, compactly supported function. Then there exists a sparse family $\mathcal{S} = \mathcal{S}(f)$ (depending on f!), such that

$$I_\alpha^{\mathcal{D}} f \simeq \sum_{Q\in\mathcal{S}} |Q|^{\alpha/d-1} \int_Q f\mathrm{d}x \cdot \chi_Q =: I_\alpha^{\mathcal{S}} f.$$

The implicit constants depend on d and α, but not on the function. Here is the formal statement.

THEOREM 6.4. *Given a bounded, nonnegative function f with compact support and a dyadic grid $\mathcal{D}$, there exists a sparse family $\mathcal{S}$ such that for all α, $0 < \alpha < d$, we have $I_\alpha^{\mathcal{D}} f(x) \lesssim I_\alpha^{\mathcal{S}} f(x)$.*

PROOF. Let $a = 2^{d+1}$. For each $k \in \mathbb{Z}$ define

$$\mathcal{Q}^k = \left\{ P \in \mathcal{D} : a^k < \frac{1}{|P|} \int_P f\mathrm{d}y \leq a^{k+1} \right\}.$$

Then for every $P \in \mathcal{D}$ satisfying $\frac{1}{|P|}\int_P f\mathrm{d}y \neq 0$ there exists a unique k such that $P \in \mathcal{Q}^k$. Consequently,

$$\begin{aligned} I_\alpha^{\mathcal{D}} f(x) &= \sum_{P\in\mathcal{D}} \frac{1}{|P|^{1-\alpha/d}} \int_P f\mathrm{d}y \cdot \chi_P(x) \\ &= \sum_k \sum_{P\in\mathcal{Q}^k} \frac{1}{|P|^{1-\alpha/d}} \int_P f\mathrm{d}y \cdot \chi_P(x) \leq \sum_k a^{k+1} \sum_{P\in\mathcal{Q}^k} |P|^{\alpha/d} \cdot \chi_P(x). \end{aligned}$$

Let $\mathcal{S}_k$ be the collection of disjoint, maximal cubes $Q \in \mathcal{D}$ such that

$$\frac{1}{|Q|} \int_Q f\mathrm{d}x > a^k.$$

Such a collection exists since $\mathcal{D}$ is a dyadic grid and f is bounded with compact support. Put $\mathcal{S} = \bigcup_k \mathcal{S}_k$. Then for every $P \in \mathcal{Q}^k$ there exists $Q \in \mathcal{S}_k$ containing P. Consequently,

$$I_\alpha^{\mathcal{D}} f(x) \leq a \sum_k a^k \sum_{Q\in\mathcal{S}_k} \sum_{P\in\mathcal{D},\, P\subset Q} |P|^{\alpha/d} \cdot \chi_P(x).$$

We evaluate the inner sum as follows:

$$\sum_{P\in\mathcal{D},\, P\subset Q} |P|^{\alpha/d}\cdot\chi_P(x) = \sum_{r=0}^{\infty}\ \sum_{P\in\mathcal{D}:P\subseteq Q,\, \ell(P)=2^{-r}\ell(Q)} |P|^{\alpha/d}\cdot\chi_P(x) = \frac{1}{1-2^{-\alpha}}|Q|^{\alpha/d}\chi_Q(x).$$

Furthermore, since $a^k < \frac{1}{|Q|}\int_Q f\mathrm{d}y$ whenever $Q\in\mathcal{S}_k$, we thus obtain

$$I_\alpha^{\mathcal{D}} f(x) \lesssim I_\alpha^{\mathcal{S}} f(x).$$

It remains to show that $\mathcal{S}$ is sparse. If $Q\in\mathcal{S}$, then $Q\in\mathcal{S}_k$ for some $k\in\mathbb{Z}$. Hence, by the maximality of the cubes in $\mathcal{S}$,

$$\begin{aligned}\Bigg|\bigcup_{Q'\in\mathcal{S},\, Q'\subsetneq Q} Q'\Bigg| &= \sum_{Q'\in\mathcal{S}_{k+1},\, Q'\subset Q} |Q'| < \frac{1}{a^{k+1}} \sum_{Q'\in\mathcal{S}_{k+1},\, Q'\subset Q} \int_{Q'} f\mathrm{d}x \\ &\le \frac{1}{a^{k+1}}\int_Q f\mathrm{d}x \le \frac{2^d}{a}|Q| = \frac{1}{2}|Q|.\end{aligned}$$

Therefore, the family $E(Q) = Q\setminus \bigcup_{Q'\in\mathcal{S},\, Q'\subsetneq Q} Q'$, $Q\in\mathcal{S}$, enjoys all the relevant properties. □

Thus, estimates for $I_\alpha^{\mathcal{S}}$ immediately yield the corresponding bounds for Riesz potentials. The discrete versions of Calderón-Zygmund operators are the so-called dyadic shifts (or dyadic shift operators), which are defined by

$$T^{\mathcal{S}} f := \sum_{Q\in\mathcal{S}} \frac{1}{|Q|}\int_Q f\mathrm{d}x\cdot\chi_Q,$$

where $\mathcal{S}$ is a given sparse family. The following statement relates $T^{\mathcal{S}}$ to T.

THEOREM 6.5. *For any Banach function spaces X, Y on $\mathbb{R}^d$ and any Calderón-Zygmund operator T, there exists a constant c_T depending only on Y such that*

$$\|T\|_{\mathcal{B}(X,Y)} \le c_T \sup_{\mathcal{S}\subset\mathfrak{D}} \|T^{\mathcal{S}}\|_{\mathcal{B}(X,Y)}.$$

We start the analysis of the boundedness of the above operators with M_α.

THEOREM 6.6. *If $0\le\alpha<d$, $1<p\le\frac{d}{\alpha}$ and $\frac{1}{q}=\frac{1}{p}-\frac{\alpha}{d}$, then*

$$\|M_\alpha f\|_{L^q} \le \left(1+\frac{p'}{q}\right)^{1-\frac{\alpha}{d}} \|f\|_{L^p}.$$

PROOF. *Step 1.* First we will establish the weak-type inequality

$$|\{x : M_\alpha f(x) > \lambda\}| \le \left(\frac{1}{\lambda}\int_{\{M_\alpha f>\lambda\}} |f(x)|\mathrm{d}x\right)^{\frac{d}{d-\alpha}}. \tag{6.1}$$

The argument is standard: for any x such that $M_\alpha f(x)>\lambda$, let Q_x be the maximal dyadic cube containing x such that

$$\frac{1}{|Q_x|^{1-\frac{\alpha}{d}}}\int_{Q_x} |f|\mathrm{d}y > \lambda.$$

Any point $y \in Q_x$ satisfies $M_\alpha f(y) > \lambda$, so the collection $\{Q_x\}$ gives rise to the splitting of the set $\{M_\alpha f > \lambda\}$ into dyadic cubes: $\bigcup_{Q \in S} Q$. For each cube $Q \in S$, we have $\int_Q |f| \mathrm{d}y > \lambda |Q|^{1-\frac{\alpha}{d}}$, or

$$|Q| < \left(\frac{1}{\lambda} \int_Q |f| \mathrm{d}y \right)^{\frac{d}{d-\alpha}}.$$

Therefore, summing over $Q \in S$, we get

$$|\{M_\alpha f > \lambda\}| \leq \sum_{Q \in S} \left(\frac{1}{\lambda} \int_Q |f| \mathrm{d}y \right)^{\frac{d}{d-\alpha}} \leq \left(\frac{1}{\lambda} \sum_{Q \in S} \int_Q |f| \mathrm{d}y \right)^{\frac{d}{d-\alpha}},$$

which is (6.1).

Step 2. Note that $q > \frac{d}{d-\alpha}$, so we may write

$$\begin{aligned}
\int_{\mathbb{R}^d} (M_\alpha f)^q \mathrm{d}x &= q \int_0^\infty \lambda^{q-1} |\{M_\alpha f > \lambda\}| \mathrm{d}\lambda \\
&\leq q \int_0^\infty \lambda^{q-1} \left(\frac{1}{\lambda} \int_{\{M_\alpha f > \lambda\}} |f| \mathrm{d}x \right)^{\frac{d}{d-\alpha}} \mathrm{d}\lambda \\
&\leq q \left(\int_{\mathbb{R}^d} |f(x)| \left(\int_0^{M_\alpha f(x)} \lambda^{q - \frac{d}{d-\alpha} - 1} \mathrm{d}\lambda \right)^{(d-\alpha)/d} \mathrm{d}x \right)^{\frac{d}{d-\alpha}} \\
&= \frac{q}{q - \frac{d}{d-\alpha}} \left(\int_{\mathbb{R}^d} |f(x) (M_\alpha f(x))^{q/p'} \mathrm{d}x \right)^{\frac{d}{d-\alpha}} \\
&\leq \frac{q}{q - \frac{d}{d-\alpha}} \|f\|_{L^p}^{\frac{d}{d-\alpha}} \|M_\alpha f\|_{L^q}^{\frac{q}{p'} \frac{d}{d-\alpha}}.
\end{aligned}$$

Since $q/(q - \frac{d}{d-\alpha}) = 1 + \frac{p'}{q}$, the claim is established. □

Now we will prove the L^p boundedness for dyadic shifts.

THEOREM 6.7. *Suppose that $\mathcal{D}$ is a dyadic grid, $\mathcal{S} \subset \mathcal{D}$ is a sparse family, $1 < p < \infty$ and $w \in A_p$. Then we have*

$$\|T^{\mathcal{S}}\|_{L^p(w) \to L^p(w)} \leq c_p [w]_{A_p}^{\max\{1, \frac{p'}{p}\}},$$

where $c_p = pp' 2^{\max\{p-1, (p-1)^{-1}\}}$.

PROOF. Note that $T^{\mathcal{S}}$ is a positive operator, so it is enough to study the action of T on $f \geq 0$. Suppose that $p \geq 2$ and let $\sigma = w^{1-p'}$. We will show that

$$\|T^{\mathcal{S}}\|_{L^p(w) \to L^p(w)} = \|T^{\mathcal{S}}(\cdot \sigma)\|_{L^p(\sigma) \to L^p(w)}. \tag{6.2}$$

Indeed, the left-hand side is equal to $\sup \int_{\mathbb{R}^d} T^{\mathcal{S}} f g w \mathrm{d}x$, where the supremum is taken over all $f \in L^p(w)$ and $g \in L^{p'}(w)$ of norms one. Putting $f = h\sigma$, we see that $\|h\|_{L^p(\sigma)} = 1$ and thus, reverting, we get the right-hand side of (6.2).

Therefore, by duality, it is enough to estimate appropriately the expression

$$\int_{\mathbb{R}^n} T^{\mathcal{S}}(f\sigma) g w \mathrm{d}x = \sum_{Q \in \mathcal{S}} \frac{1}{|Q|} \int_Q f \sigma \mathrm{d}x \int_Q g w \mathrm{d}x.$$

We write

$$\begin{aligned}
&\sum_{Q\in\mathcal{S}} \frac{1}{|Q|}\int_Q f\sigma \mathrm{d}x \int_Q gw\mathrm{d}x\\
&= \sum_{Q\in\mathcal{S}} \frac{w(Q)\sigma(Q)^{p-1}}{|Q|^p}\cdot \frac{|Q|^{p-1}}{w(Q)\sigma(Q)^{p-1}}\int_Q f\sigma \mathrm{d}x \int_Q gw\mathrm{d}x\\
&\le [w]_{A_p}\sum_{Q\in\mathcal{S}} \frac{|Q|^{p-1}}{w(Q)\sigma(Q)^{p-1}}\int_Q f\sigma \mathrm{d}x \int_Q gw\mathrm{d}x\\
&= [w]_{A_p}\sum_{Q\in\mathcal{S}} \frac{1}{\sigma(Q)}\int_Q f\sigma \mathrm{d}x\cdot \frac{1}{w(Q)}\int_Q gw\mathrm{d}x \cdot |Q|^{p-1}\sigma(Q)^{2-p}\\
&\le 2^{p-1}[w]_{A_p}\sum_{Q\in\mathcal{S}} \frac{1}{\sigma(Q)}\int_Q f\sigma \mathrm{d}x\cdot \frac{1}{w(Q)}\int_Q gw\mathrm{d}x \cdot |E(Q)|^{p-1}\sigma(Q)^{2-p},
\end{aligned}$$

where in the last passage we have used the estimate $|Q|\le 2|E(Q)|$. Since $p\ge 2$ and $E(Q)\subset Q$, we have $\sigma(Q)^{2-p}\le \sigma(E(Q))^{2-p}$ and hence

$$\begin{aligned}
&\int_{\mathbb{R}^n} T^{\mathcal{S}}(f\sigma)gw\mathrm{d}x\\
&\le 2^{p-1}[w]_{A_p}\sum_{Q\in\mathcal{S}} \frac{1}{\sigma(Q)}\int_Q f\sigma \mathrm{d}x\cdot \frac{1}{w(Q)}\int_Q gw\mathrm{d}x \cdot |E(Q)|^{p-1}\sigma(E(Q))^{2-p}.
\end{aligned}$$

By Hölder's inequality, we get

$$|E(Q)|\le w(E(Q))^{\frac{1}{p}}\sigma(E(Q))^{\frac{1}{p'}},$$

so

$$|E(Q)|^{p-1}\sigma(E(Q))^{2-p}\le \sigma(E(Q))^{\frac{1}{p}}w(E(Q))^{\frac{1}{p'}}.$$

Plugging these observations above and applying Hölder's inequality, we get

$$\begin{aligned}
&\sum_{Q\in\mathcal{S}} \frac{1}{\sigma(Q)}\int_Q f\sigma \mathrm{d}x\cdot \frac{1}{w(Q)}\int_Q gw\mathrm{d}x \cdot |E(Q)|^{p-1}\sigma(E(Q))^{2-p}\\
&\le \sum_{Q\in\mathcal{S}} \frac{1}{\sigma(Q)}\int_Q f\sigma \mathrm{d}x\cdot \frac{1}{w(Q)}\int_Q gw\mathrm{d}x \cdot \sigma(E(Q))^{\frac{1}{p}}w(E(Q))^{\frac{1}{p'}}\\
&\le \left(\sum_{Q\in\mathcal{S}}\left(\frac{1}{\sigma(Q)}\int_Q f\sigma\mathrm{d}x\right)^p\sigma(E(Q))\right)^{\frac{1}{p}}\left(\sum_{Q\in\mathcal{S}}\left(\frac{1}{w(Q)}\int_Q gw\mathrm{d}x\right)^{p'}w(E(Q))\right)^{\frac{1}{p'}}\\
&\le \|M_\sigma f\|_{L^p(\sigma)}\|M_w g\|_{L^{p'}(w)}\\
&\le pp'\|f\|_{L^p(\sigma)}\|g\|_{L^{p'}(w)}.
\end{aligned}$$

This is exactly what we need. The case $1<p<2$ follows by duality, since $T^{\mathcal{S}}$ is self-adjoint. □

Now we turn our attention to the L^p-boundedness of fractional integrals. For brevity, we will write I_α instead of $I_\alpha^{\mathcal{S}}$.

THEOREM 6.8. *Suppose that $\mathcal{D}$ is a dyadic grid, $\mathcal{S}\subset\mathcal{D}$ is a sparse family, $0<\alpha<d$, $1<p<d/\alpha$ and $1/q=1/p-\alpha/d$. If $\min\{\frac{p'}{q},\frac{q}{p'}\}\le 1-\frac{\alpha}{d}$ and w*

satisfies

$$[w]_{A_{p,q}} := \sup_{Q\in\mathcal{D}} \left(\frac{1}{|Q|}\int_Q w^q\right)\left(\frac{1}{|Q|}\int_Q w^{-p'}\right)^{q/p'} < \infty,$$

then

$$\|I_\alpha\|_{L^p(w^p)\to L^q(w^q)} \le c_{p,\alpha}[w]_{A_{p,q}}^{(1-\frac{\alpha}{d})\max\{1,\frac{p'}{q}\}},$$

where

$$c_{p,\alpha} = p'\left(1+\frac{q}{p'}\right)^{1-\frac{\alpha}{d}} 2^{(1-\frac{\alpha}{d})\max\{\frac{q}{p'},\frac{p'}{q}\}}.$$

PROOF. Suppose that $\frac{p'}{q} \le 1-\frac{\alpha}{d}$ and denote $u = w^q$, $\sigma = w^{-p'}$, so that

$$[w]_{A_{p,q}} = \sup_{Q\in\mathcal{D}} \frac{u(Q)\sigma(Q)^{\frac{q}{p'}}}{|Q|^{1+\frac{q}{p'}}} = [u]_{A_{1+\frac{q}{p'}}}.$$

Put $r = 1+\frac{q}{p'}$ and observe that $r' = 1+\frac{p'}{q}$. We start with noticing that

$$\|I_\alpha\|_{L^p(w^p)\to L^q(w^q)} = \|I_\alpha(\cdot\sigma)\|_{L^p(\sigma)\to L^q(u)}.$$

Next, we proceed as follows: for $g\in L^{q'}(u)$,

$$\begin{aligned}
&\int_{\mathbb{R}^d} I_\alpha(f\sigma)gu\,\mathrm{d}x\\
&= \sum_{Q\in\mathcal{S}} |Q|^{\alpha/d-1}\int_Q f\sigma\,\mathrm{d}x\cdot\int_Q gu\,\mathrm{d}x\\
&= \sum_{Q\in\mathcal{S}} \frac{u(Q)^{1-\frac{\alpha}{d}}\sigma(Q)^{\frac{q}{p'}(1-\frac{\alpha}{d})}}{|Q|^{(1+\frac{q}{p'})(1-\frac{\alpha}{d})}}\cdot\frac{|Q|^{\frac{q}{p'}(1-\frac{\alpha}{d})}}{u(Q)^{1-\frac{\alpha}{d}}\sigma(Q)^{\frac{q}{p'}(1-\frac{\alpha}{d})}}\int_Q f\sigma\,\mathrm{d}x\cdot\int_Q gu\,\mathrm{d}x\\
&\le [w]_{A_{p,q}}^{1-\frac{\alpha}{d}} \sum_{Q\in\mathcal{S}} \frac{1}{\sigma(Q)}\int_Q f\sigma\,\mathrm{d}x\cdot\frac{1}{u(Q)^{1-\frac{\alpha}{d}}}\int_Q gu\,\mathrm{d}x\cdot|Q|^{\frac{q}{p'}(1-\frac{\alpha}{d})}\sigma(Q)^{1-\frac{q}{p'}(1-\frac{\alpha}{d})}.
\end{aligned}$$

Now, since $\frac{p'}{q}\le 1-\frac{\alpha}{d}$ and $E(Q)\subset Q$, we have

$$\sigma(Q)^{1-\frac{q}{p'}(1-\frac{\alpha}{d})} \le \sigma(E(Q))^{1-\frac{q}{p'}(1-\frac{\alpha}{d})},$$

which combined with the estimate $|Q|\le 2|E(Q)|$ gives

$$\begin{aligned}
\int_{\mathbb{R}^d} I_\alpha(f\sigma)gu\,\mathrm{d}x \le 2^{\frac{q}{p'}(1-\frac{\alpha}{d})}[w]_{A_{p,q}}^{1-\frac{\alpha}{d}} \sum_{Q\in\mathcal{S}} \frac{1}{\sigma(Q)}\int_Q f\sigma\,\mathrm{d}x\cdot\frac{1}{u(Q)^{1-\frac{\alpha}{d}}}\int_Q gu\,\mathrm{d}x\\
\cdot|E(Q)|^{\frac{q}{p'}(1-\frac{\alpha}{d})}\sigma(E(Q))^{1-\frac{q}{p'}(1-\frac{\alpha}{d})}.
\end{aligned}$$

By Hölder's inequality with r and r' we have

$$|E(Q)| \le u(E(Q))^{\frac{1}{r}}\sigma(E(Q))^{\frac{1}{r'}}.$$

Since

$$\frac{q}{p'}\left(1-\frac{\alpha}{d}\right) = \frac{q}{p'}\left(\frac{1}{q}+\frac{1}{p'}\right) = \frac{1}{p'}\left(1+\frac{q}{p'}\right) = \frac{r}{p'},$$

we obtain

$$\begin{aligned}
&|E(Q)|^{\frac{q}{p'}(1-\frac{\alpha}{d})}\sigma(E(Q))^{1-\frac{q}{p'}(1-\frac{\alpha}{d})} = |E(Q)|^{\frac{r}{p'}}\sigma(E(Q))^{1-\frac{r}{p'}}\\
&\le u(E(Q))^{\frac{1}{p'}}\sigma(E(Q))^{\frac{r}{r'p'}}\sigma(E(Q))^{1-\frac{r'}{p}} = u(E(Q))^{\frac{1}{p'}}\sigma(E(Q))^{\frac{1}{p}}.
\end{aligned}$$

Combining all the above facts, we get

$$\sum_{Q\in\mathcal{S}} \frac{1}{\sigma(Q)}\int_Q f\sigma \mathrm{d}x \cdot \frac{1}{u(Q)^{1-\frac{\alpha}{d}}}\int_Q gu\mathrm{d}x \cdot |E(Q)|^{\frac{q}{p'}(1-\frac{\alpha}{d})}\sigma(E(Q))^{1-\frac{q}{p'}(1-\frac{\alpha}{d})}$$
$$\leq \sum_{Q\in\mathcal{S}} \frac{1}{\sigma(Q)}\int_Q f\sigma \mathrm{d}x \cdot \frac{1}{u(Q)^{1-\frac{\alpha}{d}}}\int_Q gu\mathrm{d}x \cdot \sigma(E(Q))^{\frac{1}{p}}u(E(Q))^{\frac{1}{p'}}$$
$$\leq \left(\sum_{Q\in\mathcal{S}}\left(\frac{1}{\sigma(Q)}\int_Q f\sigma \mathrm{d}x\right)^p \sigma(E(Q))\right)^{\frac{1}{p}} \left(\sum_{Q\in\mathcal{S}}\left(\frac{1}{u(Q)^{1-\frac{\alpha}{d}}}\int_Q gu\mathrm{d}x\right)^{p'} u(E(Q))\right)^{\frac{1}{p'}}$$
$$\leq \|M_\sigma f\|_{L^p(\sigma)}\|M_{\alpha,u}g\|_{L^{p'}(u)}$$
$$\leq p'\left(1+\frac{q}{p'}\right)^{1-\frac{\alpha}{d}}\|f\|_{L^p(\sigma)}\|g\|_{L^{q'}(u)}.$$

Here in the last passage we have used the boundedness of maximal and the fractional maximal operators. The case $\frac{p'}{q} \geq (1-\frac{\alpha}{d})^{-1}$ follows from duality (see Exercise 1 below). The proof is complete. □

Problems

1. Show the identity $\|I_\alpha\|_{L^p(w^p)\to L^q(w^q)} = \|I_\alpha\|_{L^{q'}(w^{-q'})\to L^{p'}(w^{-p'})}$.

2. Prove that if $\|I_\alpha\|_{L^p(w^p)\to L^{q,\infty}(w^q)} < \infty$, then

$$[w]_{A_{p,q}} := \sup_{Q\in\mathcal{D}}\left(\frac{1}{|Q|}\int_Q w^q\right)\left(\frac{1}{|Q|}\int_Q w^{-p'}\right)^{q/p'} < \infty.$$

3. The family $\mathcal{S}\subset\mathcal{D}$ is sparse. Is the family $\{Q^*\}_{Q\in\mathcal{S}}$, where Q^* denotes the dyadic parent of Q, sparse? If not, is this family contained in a finite number of spase families?

4. Consider the probability space $([0,1]^d, \mathcal{B}([0,1]^d), |\cdot|)$ equipped with its dyadic filtration. Find the sparse dominating operator/a family of sparse dominating operators for:

a) a martingale variable,
b) a stopped martingale variable,
c) the martigale maximal function,
d) the martingale transform,
e) the martingale square function.

CHAPTER 7

Elements of two-weight theory

Now we turn our attention to the estimates which involve different weights on the left- and the right-hand side. This context has proved to be much more difficult and challenging than the one-weight setting. For simplicity, we will mainly focus on boundedness of the dyadic maximal operator M. As usual, we will study the weak- and strong-type estimates; so, the questions we want to answer are the following.

I. Given $1 \le p < \infty$, characterize those pairs (w, v) of weights in $\mathbb{R}^d$ such that $\|M\|_{L^p(v)\to L^{p,\infty}(w)} < \infty$.

II. Given $1 \le p < \infty$, characterize those pairs (w, v) of weights in $\mathbb{R}^d$ such that $\|M\|_{L^p(v)\to L^p(w)} < \infty$.

As we will show now, the answer to the first question is very simple and natural, from the viewpoint of the one-weight theory. We will show the following statement.

THEOREM 7.1. *Fix $1 < p < \infty$. Then $\|M\|_{L^p(v)\to L^{p,\infty}(w)} < \infty$ if and only if*

$$[w, v]_{A_p} = \sup \left(\frac{1}{|Q|}\int_Q w\right)\left(\frac{1}{|Q|}\int_Q v^{1/(1-p)}\right)^{p-1} < \infty,$$

where the supremum is taken over all dyadic cubes in $\mathbb{R}^d$. For $p = 1$, the validity of the weak-type bound is characterized by the condition

$$[w, v]_{A_1} = \sup \left(\frac{1}{|Q|}\int_Q w\right)\left(\operatorname*{essinf}_Q v\right)^{-1} < \infty.$$

PROOF. We will study the case $p > 1$ only, the proof for $p = 1$ requires only small modifications. We start with the necessity. Suppose that for a pair w, v of weights, we have $C = \|M\|_{L^p(v)\to L^{p,\infty}(w)} < \infty$. Then for each $\lambda > 0$ and any $f \in L^p(v)$ we have the estimate

$$\lambda^p w(Mf \ge \lambda) \le C \int_{\mathbb{R}^d} |f|^p v.$$

Pick an arbitrary dyadic cube Q and put $f = v^{1/(1-p)}\chi_{Q\cap\{v>\varepsilon\}}$, $\lambda = \frac{1}{|Q|}\int_Q f$. Then $Mf \ge \lambda$ on Q by the very definition of the maximal operator, so

$$\left(\frac{1}{|Q|}\int_Q f\right)^p \cdot w(Q) \le \lambda^p w(Mf \ge \lambda) \le C\int_{\mathbb{R}^d} |f|^p v = C\int_Q f.$$

This is equivalent to saying that

$$\left(\frac{1}{|Q|}\int_Q w\right)\left(\frac{1}{|Q|}\int_Q v^{1/(1-p)}\chi_{\{v>\varepsilon\}}\right)^{p-1} \le C.$$

Letting $\varepsilon \to 0$ and using Lebesgue's monotone convergence theorem, we see that $[w, v]_{A_p} \le C$, since the cube Q was arbitrary.

Now, we will prove the sufficiency part. Suppose that $[w,v]_{A_p}<\infty$ and fix an arbitrary function $f\in L^p(v)$. By the standard stopping argument, the set $\{Mf>\lambda\}$ is a union of pairwise disjoint dyadic cubes: $\{Mf>\lambda\}=\bigcup_j Q_j$ such that $\frac{1}{|Q_j|}\int_{Q_j}|f|>\lambda$. Therefore, for each j we may write

$$\lambda^p w(Q_j)\leq [w,v]_{A_p}|Q_j|\cdot\frac{\left(\frac{1}{|Q_j|}\int_{Q_j}|f|\right)^p}{\left(\frac{1}{|Q_j|}\int_{Q_j}v^{1/(1-p)}\right)^{p-1}}.$$

However, the function $B(x,y)=|x|^py^{1-p}$ is convex on $\mathbb{R}\times(0,\infty)$, so

$$B\left(\frac{1}{|Q_j|}\int_{Q_j}|f|,\frac{1}{|Q_j|}\int_{Q_j}v^{1/(1-p)}\right)\leq\frac{1}{|Q_j|}\int_{Q_j}B(|f|,v^{1/(1-p)}),$$

which implies

$$\lambda^p w(Q_j)\leq [w,v]_{A_p}\int_{Q_j}|f|^pv.$$

Summing over all j, we get the desired weak-type bound; actually, we get the slightly stronger estimate

$$\lambda^p w(Mf>\lambda)\leq [w,v]_{A_p}\int_{\{Mf>\lambda\}}|f|^pv. \tag{7.1}$$

□

REMARK 7.2. The above proof shows that $\|M\|_{L^p(v)\to L^{p,\infty}(w)}\leq [w,v]_{A_p}$. This bound is sharp: for a large class of pairs (w,v), the equality is attained.

Theorem 7.1 answers the question **I** above, so we can turn to the second problem. Of course, if $\|M\|_{L^p(v)\to L^p(w)}<\infty$, then also $\|M\|_{L^p(v)\to L^{p,\infty}(w)}<\infty$ and hence we must have $[w,v]_{A_p}<\infty$. A little surprising and unexpected fact (in the light of the one-weight theory) is that the condition $[w,v]_{A_p}<\infty$ is not sufficient. More precisely, we have the following fact.

THEOREM 7.3. *For any $1<p<\infty$ there is a pair (w,v) of weights in $\mathbb{R}$ such that $[w,v]_{A_p}<\infty$ but $\|M\|_{L^p(v)\to L^p(w)}=\infty$.*

PROOF. We will handle the case $p=2$ only. For example, one can consider the weights $w(x)=-x\log x\chi_{(0,1/4)}(x)$, $v(x)=x\log^2x\,\chi_{(0,1/4)}(x)$. The details are left to the reader (Exercise 1 below). □

Nevertheless, the condition $[w,v]_{A_p}<\infty$ is very close to the characterization. First we will show the following result.

THEOREM 7.4. *If $[w,v]_{A_q}<\infty$ for some $1<q<p$, then*

$$\|M\|_{L^p(w)\to L^p\left(v^{\frac{p}{q}}w^{\frac{q-p}{q}}\right)}\leq\left(\frac{p}{p-q}[w,v]_{A_q}\right)^{p/q}.$$

PROOF. By (7.1), we have

$$\lambda^q w(Mf>\lambda)\leq [w,v]_{A_q}\int_{\{Mf>\lambda\}}|f|^qv.$$

Multiply both sides by $p\lambda^{p-q-1}$ and integrate over λ from 0 to infinity, to obtain

$$\|Mf\|^p_{L^p(w)}\leq\frac{p}{p-q}[w,v]_{A_q}\int_{\mathbb{R}^d}Mf^{p-q}|f|^qv\mathrm{d}x.$$

Applying Hölder's inequality, we get

$$\|Mf\|_{L^p(w)}^p \leq \frac{p}{p-q}[w,v]_{A_q}\|Mf\|_{L^p(w)}^{\frac{p-q}{p}}\|f\|_{L^p\left(v^{\frac{p}{q}}w^{\frac{q-p}{q}}\right)}^{\frac{q}{p}},$$

which is equivalent to the claim. □

The above statement can be improved to the following.

THEOREM 7.5. *If* $[w,v]_{A_q}<\infty$ *for some* $1<q<p$, *then* $\|M\|_{L^p(w)\to L^p(v)}<\infty$.

PROOF. This follows from Marcinkiewicz interpolation theorem. Note that the condition $[w,v]_{A_q}<\infty$ implies, by Lebesgue's differentiation theorem, the estimate $w\leq [w,v]_{A_q}v$ almost everywhere. Consequently, we have $\|M\|_{L^\infty(w)\to L^\infty(v)}<\infty$, which combined with the weak-type (q,q) estimate yields the claim. □

Of course, the above theorem gives only a partial answer to the question **II**, as still we do not have the appropriate characterization of the pairs (w,v) (just the sufficient condition). Nowadays, there are two types of approach to the problem. The first is to develop the so-called testing conditions, while the second is to introduce the so-called bump conditions. Let us start with the first approach.

THEOREM 7.6. *Let* $1<p<\infty$ *and let* (w,v) *be a pair of dyadic weights on* $\mathbb{R}^d$. *The inequality*

$$\left(\int_{\mathbb{R}^d}(Mf)^p w\,dx\right)^{1/p} \leq C_{p,w,v}\left(\int_{\mathbb{R}^d}|f|^p v\,dx\right)^{1/p} \tag{7.2}$$

holds true for all $f\in L^p(v)$ *if and only if for any dyadic cube* Q *we have the estimate*

$$\left(\int_Q (M(v^{1/(1-p)}\chi_Q)^p w\,dx\right)^{1/p} \leq c_{p,w,v}\left(\int_Q v^{1/(1-p)}\,dx\right)^{1/p}. \tag{7.3}$$

PROOF. The necessity of the condition (7.3) is evident: fir any dyadic Q, simply apply (*test*) (7.2) with $f=v^{1/(1-p)}\chi_Q$ (and replace $\mathbb{R}^d$ on the left by Q) to obtain (7.3) with $c_{p,w,v}=C_{p,w,v}$. The nontrivial part is the sufficiency of the testing condition. So, suppose that (w,v) satisfies (7.3) and let f be an arbitrary element of $L^p(v)$. With no loss of generality, by a simple approximation argument, we may assume that f is measurable with respect to some dyadic generation $\mathcal{D}^N$. We will use the linearization of the maximal function: we have

$$Mf=\sum_{Q\in\mathcal{D}}\left(\frac{1}{|Q|}\int_Q|f|\right)\chi_{E(Q)},$$

where the sets $E(Q)$ are defined as follows. By simplicity of f, for any x there is a cube $Q=Q(x)$ for which $Mf(x)=\frac{1}{|Q|}\int_Q|f|$; if there are several Q with this property, we choose Q with the largest measure. Then we set $E(Q)=\{x:Q(x)=Q\}$. Note that the simplicity of f (measurability with respect to $mathcalD^N$) implies that $E(Q)=\emptyset$ if $|Q|<2^{-Nd}$. Substituting $f=gv^{1/(1-p)}$, we see that the desired estimate (7.2) takes the form

$$\sum_{Q\in\mathcal{D}}\left(\frac{1}{|Q|}\int_Q|g|v^{1/(1-p)}\mathrm{d}x\right)^p w(E(Q)) \leq C_{p,w,v}^p\int_{\mathbb{R}^d}|g|^p v^{1/(1-p)}\mathrm{d}x,$$

or

$$\sum_{Q\in\mathcal{D}}\left(\frac{1}{v^{1/(1-p)}(Q)}\int_Q |g|v^{1/(1-p)}\mathrm{d}x\right)^p\cdot\left(\frac{v^{1/(1-p)}(Q)}{|Q|}\right)^p w(E(Q))$$
$$\leq C_{p,w,v}^p\int_{\mathbb{R}^d}|g|^p v^{1/(1-p)}\mathrm{d}x.$$

As we will prove in a lemma below, the testing condition implies that there are pairwise disjoint cubes sets $e(Q)$ satisfying $e(Q)\subseteq Q$ and

$$\left(\frac{v^{1/(1-p)}(Q)}{|Q|}\right)^p w(E(Q)) = c_{p,w,v}^p v^{1/(1-p)}(e(Q)).$$

Plugging this above, we see that the left-hand side equals

$$c_{p,w,v}^p\sum_{Q\in\mathcal{D}}\left(\frac{1}{v^{1/(1-p)}(Q)}\int_Q |g|v^{1/(1-p)}\mathrm{d}x\right)^p\cdot v^{1/(1-p)}(e(Q))$$
$$\leq c_{p,w,v}^p\|M_{v^{1/(1-p)}}g\|_{L^p(v^{1/(1-p)})}^p,$$

which, by the unweighted L^p estimate for maximal functions, is bounded from above by

$$c_{p,w,v}^p\left(\frac{p}{p-1}\right)^p\|g\|_{L^p(v^{1/(1-p)})}^p.$$

In other words, we get the estimate (7.2) with $C_{p,w,v}=\frac{p}{p-1}c_{p,w,v}$. □

So, the only thing which needs to be established is the following.

LEMMA 7.7. *Suppose that a pair* (w,v) *of weights satisfies* (7.3) *and let* $(E(Q))_{Q\in\mathcal{D}}$ *be a family of sets as in the previous theorem. Then there exists a family* $\{e(Q)\}_{Q\in\mathcal{D}}$ *of pairwise disjoint sets such that* $e(Q)\subseteq Q$ *for each* Q *and*

$$\left(\frac{v^{1/(1-p)}(Q)}{|Q|}\right)^p w(E(Q)) = c_{p,w,v}^p v^{1/(1-p)}(e(Q)).$$

PROOF. Recall that $E(Q)=\emptyset$ for $|Q|<2^{-Nd}$; therefore, we may set $e(Q)=\emptyset$ for such Q. For the remaining cubes, we do the following. By (7.3), for any Q we have

$$\sum_{R\subseteq Q, R\in\mathcal{D}}\left(\frac{v^{1/(1-p)}(R)}{|R|}\right)^p w(E(R))\leq\int_Q\left[M(v^{1/(1-p)}\chi_Q)\right]^p w\mathrm{d}x$$
$$\leq c_{p,w,v}^p\int_Q v^{1/(1-p)}\mathrm{d}x. \tag{7.4}$$

Therefore, we proceed by the following inductive argument. Pick an arbitrary cube Q with $|Q|=2^{-Nd}$. By the above estimate, we have

$$\left(\frac{v^{1/(1-p)}(Q)}{|Q|}\right)^p w(E(Q))\leq c_{p,w,v}^p v^{1/(1-p)}(Q),$$

so we may choose a subset $e(Q)$ of Q such that the left-hand side is equal to $c_{p,w,v}^p v^{1/(1-p)}(e(Q))$. Next, suppose we have successfully defined the sets $e(Q)$ for some family $\mathcal{F}$ of cubes. Let $Q \notin \mathcal{F}$ be a cube which is a direct parent of some element of $\mathcal{F}$. By (7.4), we see that

$$\left(\frac{v^{1/(1-p)}(Q)}{|Q|}\right)^p w(E(Q)) + \sum_{R \subseteq Q, R \in \mathcal{F}} \left(\frac{v^{1/(1-p)}(R)}{|R|}\right)^p w(E(R)) \leq c_{p,w,v}^p v^{1/(1-p)}(Q),$$

which implies

$$\left(\frac{v^{1/(1-p)}(Q)}{|Q|}\right)^p w(E(Q)) + c_{p,w,v}^p v^{1/(1-p)} \left(\bigcup_{R \subseteq Q, R \in \mathcal{F}} e(R)\right) \leq c_{p,w,v}^p v^{1/(1-p)}(Q),$$

or

$$\left(\frac{v^{1/(1-p)}(Q)}{|Q|}\right)^p w(E(Q)) \leq c_{p,w,v}^p v^{1/(1-p)} \left(Q \setminus \bigcup_{R \subseteq Q, R \in \mathcal{F}} e(R)\right).$$

In other words, there is a room for a set $e(Q)$, disjoint with all the sets $e(R)$, $R \in \mathcal{F}$, such that

$$\left(\frac{v^{1/(1-p)}(Q)}{|Q|}\right)^p w(E(Q)) = c_{p,w,v}^p v^{1/(1-p)}(e(Q)).$$

This allows to define $e(Q)$ for the full dyadic class $\mathcal{D}$. $\square$

REMARK 7.8. The above theorem has an interesting consequence for the one-weight theory. As we have already proved, the boundedness of M as the operator on the weighted space $L^p(w)$ holds if and only if $w \in A_p$. Therefore, we get that $w \in A_p$ if and only if for any cube Q we have

$$\left(\int_Q (M(w^{1/(1-p)}\chi_Q)^p w \mathrm{d}x\right)^{1/p} \leq c_{p,w} \left(\int_Q w^{1/(1-p)} \mathrm{d}x\right)^{1/p},$$

or passing to the dual weights $v = w^{1/(1-p)}$,

$$\left(\int_Q (M(v\chi_Q)^p w \mathrm{d}x\right)^{1/p} \leq c_{p,w} \left(\int_Q v \mathrm{d}x\right)^{1/p}.$$

The testing conditions are typically not easy to verify in practice; in many cases the problem of proving the testing conditions is actually of the same difficulty as that of proving the general weighted L^p bound. This gives rise to the question about other conditions, which are maybe not equivalent, but sufficient and more applicable. We have already seen above that the condition $[w,v]_{A_q} < \infty$ for some $q < p$ does the job (and it is indeed much simpler). The so-called bumping conditions are of similar flavor. To introduce them from a convenient perspective, let us rewrite

$$[w,v]_{A_p}^{1/p} = \sup_Q \|w^{1/p}\|_{L^p(Q)} \|v^{-1/p}\|_{L^{p'}(Q)} < \infty.$$

We will prove the following fact.

THEOREM 7.9. *Given a pair (w,v) of weights, there is a weight $u \in A_p$ and positive constants c_1, c_2 such that $c_1 w \leq u \leq c_2 v$ if and only if there is a constant $r > 1$ such that*

$$\sup_Q \|w^{1/p}\|_{L^{rp}(Q)} \|v^{-1/p}\|_{L^{rp'}(Q)} < \infty.$$

Of course, this theorem allows to deduce many two-weight estimates from their one-weight counterparts. To prove the above statements, we need some preliminary observations.

LEMMA 7.10. *Suppose that*

$$\|M\|_{L^p(v)\to L^p(w)} = A \quad \text{and} \quad \|M\|_{L^{p'}(w^{1/(1-p)})\to L^{p'}(v^{1/(1-p)})} = B.$$

Then there are functions $u_j \geq 0$ *such that* $w^{1/p}Mu_j \leq C_j u_j v^{1/p}$, $j = 1, 2$, *and* $w^{1/p}v^{1/p'} = u_1 u_2^{1-p}$.

PROOF. We may assume that $p \geq 2$, since in the case $1 < p < 2$ we simply interchange the roles of A and B. The proof will rest on Rubio de Francia algorithm. Set $s = p/p' = p-1$ and put $M_s f = [M(f^s)]^{1/s}$. The operator $S(f) = w^{1/p}M(fv^{-1/p}) + v^{-1/ps}M_s(fw^{1/ps})$ is bounded on L^p: indeed, we have

$$\|w^{1/p}M(fv^{-1/p})\|_{L^p} = \|M(fv^{-1/p})\|_{L^p(w)} \leq A\|fv^{-1/p}\|_{L^p(v)} = A\|f\|_{L^p}$$

and similarly

$$\begin{aligned}\|v^{-1/ps}M_s(fw^{1/ps})\|_{L^p} &= \|M_s(fw^{1/ps})\|_{L^p(v^{-1/s})} \\ &= \|M(f^s w^{1/p})\|^{1/(p-1)}_{L^{p'}(v^{1/(1-p)})} \\ &\leq B^{1/(p-1)}\|f^s w^{1/p}\|^{1/(p-1)}_{L^{p'}(w^{1/(1-p)})} = B^{1/(p-1)}\|f\|_{L^p}.\end{aligned}$$

Let $C = \|S\|_{L^p\to L^p}$, pick $\tilde{C} > C$ and for $u_0 \in L^p_+$, define $U = \sum_{j=0}^{\infty} S^j(u_0)/\tilde{C}^j$. Then $U \in L^p$ and $S(U) \leq \tilde{C}U$, so

$$M(Uv^{-1/p})w^{1/p} \leq S(U) \leq \tilde{C}U \quad \text{and} \quad v^{-1/ps}M_s(Uw^{1/ps}) \leq \tilde{C}U.$$

Therefore, if $u_2 = Uv^{-1/p}$, then $w^{1/p}Mu_2 \leq \tilde{C}u_2 v^{1/p}$ and if we let $u_1 = U^s w^{1/p}$, then $w^{1/p}Mu_1 \leq \tilde{C}^s u_1 v^{1/p}$ and $w^{1/p}v^{1/p'} = u_1 u_2^{1-p}$. □

LEMMA 7.11. *Suppose that* $[w, v]_{A_p} < \infty$ *and* $0 < \delta < 1$. *Then there exists a weight* $u = u_\delta \in A_p$ *such that* $c_1 w^\delta \leq u \leq c_2 v^\delta$.

PROOF. Choose $0 < \varepsilon, \eta < 1$ such that $\delta = \varepsilon\eta$. By Exercise 2 below, we know

$$\|M\|_{L^p(v^\varepsilon)\to L^p(w^\varepsilon)} < \infty, \qquad \|M\|_{L^p(w^{\varepsilon/(1-p)})\to L^p(v^{\varepsilon/(1-p)})} < \infty.$$

Therefore, by the previous lemma, we obtain $w^{\varepsilon/p}v^{\varepsilon/p'} = u_1 u_2^{1-p}$, where $Mu_j \leq c_j u_j (v/w)^{\varepsilon/p}$. Observe that

$$v^\varepsilon = u_1 \left(\frac{v}{w}\right)^{\varepsilon/p} u_2^{1-p} \geq cMu_1(Mu_2)^{1-p} \geq cu_1 u_2^{1-p}\left(\frac{v}{w}\right)^{\varepsilon(1-p)/p} = cw^\varepsilon,$$

so $c_1 w^\delta \leq (Mu_1)^\eta (Mu_2)^{\eta(1-p)} \leq c_2 v^\delta$. Since $0 < \eta < 1$, we have $(Mu_j)^\eta \in A_1$, and hence $u = (Mu_1)^\eta(Mu_2)^{\eta(1-p)} \in A_p$, by factorization. □

The above two lemmas easily give the assertion of Theorem 7.9.

PROOF OF THEOREM 7.9. Suppose that an A_p weight u can be inserted. Both u and $u^{1/(1-p)}$ enjoy the reverse Hölder inequality, so there is $r > 1$ such that $u^r \in A_p$ and therefore

$$\left(\frac{1}{|Q|}\int_Q w^r\right)\left(\frac{1}{|Q|}\int_Q v^{r/(1-p)}\right)^{p-1} \leq c_1^{-r}c_2^r\left(\frac{1}{|Q|}\int_Q u^r\right)\left(\frac{1}{|Q|}\int_Q u^{r/(1-p)}\right)^{p-1}$$

and $[w^r, v^r]_{A_p} < \infty$. Conversely, if $[w^r, v^r]_{A_p} < \infty$, then the desired u is guaranteed by the previous lemma, applied with $\delta = r^{-1}$. □

We should point out here that the assertion can be significantly strengthened to the setting of Orlicz spaces.

DEFINITION 7.12. A function $B : [0, \infty) \to [0, \infty)$ is a Young function if it is continuous, convex, increasing and satisfying $B(0) = 0$, $B(t) \to \infty$ as $t \to \infty$.

DEFINITION 7.13. Let (X, μ) be a measure space and let B be a Young function. The Orlicz space $L^B(X, \mu)$ consists of all μ-measurable functions f such that

$$\int_X B\left(\frac{|f(y)|}{\lambda}\right) \mathrm{d}\mu(y) < \infty$$

for some $\lambda > 0$. The space $L^B(X, \mu)$ can be normed with the use of Luxemburg norm

$$\|f\|_{B,\mu} = \inf\left\{\lambda > 0 \, : \, \int_X B\left(\frac{|f(y)|}{\lambda}\right) \mathrm{d}\mu \leq 1\right\}.$$

In the case when $X = \mathbb{R}^d$ and μ is the Lebesgue measure, we will simply write $\|f\|_B$ instead of $\|f\|_{B,\mu}$.

It can be shown that under some mild assumptions on the growth of B, we have $X^* = L^{B^*}(X, \mu)$, where B^* is the conjugate function given by

$$B^*(t) = \sup_{s>0} \big[ts - B(s)\big].$$

DEFINITION 7.14. For a Young function B, we define the associated maximal operator acting on locally integrable functions on $\mathbb{R}^d$ by

$$M_B f(x) = \sup_{Q \ni x} \|f\|_{B,Q}.$$

Here

$$\|f\|_{B,Q} = \inf\left\{\lambda > 0 : \frac{1}{|Q|} \int_Q B\left(\frac{|f(y)|}{\lambda}\right) \mathrm{d}y \leq 1\right\}.$$

The following result holds true.

THEOREM 7.15. *Let $1 < p < \infty$ and let B be a Young function such that M_{B^*} is bounded on L^p. Suppose in addition that (w, v) is a pair of weights such that*

$$\sup_Q \left(\frac{1}{|Q|} \int_Q w\right) \|v^{-1/p}\|_{B,Q}^p < \infty. \tag{7.5}$$

Then $\|M\|_{L^p(v) \to L^p(w)} < \infty$.

For example, if take $B(t) = t^q$, then $B^* \sim t^{q'}$, $\|f\|_{B^*} \sim \left(\frac{1}{|Q|}\int_Q |f|^{q'}\right)^{1/q'}$ and M_{B^*} is given by

$$M_{B^*} f(x) \sim \sup_{Q \ni x} \left(\frac{1}{|Q|} \int_Q |f|^{q'}\right)^{1/q'}.$$

This maximal function is bounded on L^p if and only if $p/q' > 1$, that is, we must have $q > p'$. Then the condition (7.5), combined with the identity $\|f\|_{B,Q} =$

$\left(\frac{1}{|Q|}\int_Q |f|^q\right)^{1/q}$, becomes

$$\sup_Q \left(\frac{1}{|Q|}\int_Q w\right)\left(\frac{1}{|Q|}\int_Q v^{-q/p}\right)^{p/q} < \infty,$$

which is equivalent to $[w,v]_{A_{p/q+1}} < \infty$. Since $q > p'$, we have $p/q + 1 < p$ and we recover the assertion of Theorem 7.5.

Other examples in which the above statement can be applied, include the functions

$$B(t) = t^{p'} \log^{p'-1+\delta}(1+t)$$

and

$$B(t) = t^{p'} \log^{p'-1}(1+t)[\log\log(e+t)]^{p'-1+\delta}.$$

Problems

1. Prove Theorem 7.3 for $p = 2$.

2. Show that if $[w,v]_{A_p} < \infty$ and $\delta \in (0,1)$, then $\|M\|_{L^p(v^\delta)\to L^p(w^\delta)} < \infty$.

3. Let w be an arbitrary nonnegative function. Prove that $[w, Mw]_{A_p} \le 1$ for any $p < \infty$.

4. Let $1 < p < \infty$ and suppose that T is an operator on $L^p(w)$, $w \in A_p$, such that $\|T\|_{L^p(w)\to L^p(w)} \approx_p [w]_{A_p}^{\alpha_p}$, $\alpha_p > 1/p$. Prove that there exists no constant $C_{p,c}$ such that

$$\|T\|_{L^p(w)\to L^p(v)} \le C_{p,[w,v]_{A_p}}.$$

5. For a pair w, v of weights, define

$$[w,v]^* = \sup_Q \left(\frac{1}{|Q|}\int_{\mathbb{R}^d} [M\chi_Q]^p w\right)\left(\frac{1}{|Q|}\int_Q v^{1/(1-p)}\right)^{p-1}.$$

Prove that if $\|M\|_{L^p(v)\to L^p(w)} < \infty$, then $[w,v]^* < \infty$. Prove that the reverse implication is not true.

Bibliography

[1] B. Muckenhoupt, *Weighted norm inequalities for the Hardy maximal function*, Trans. Amer. Math. Soc. 165 (1972), 207–226.

[2] V. Vasyunin, *The exact constant in the inverse Hölder inequality for Muckenhoupt weights* (Russian), Algebra i Analiz 15 (2003), 73–117; translation in St. Petersburg Math. J. 15 (2004), 49–79.