ESTIMATES FOR OPERATOR NORMS ON WEIGHTED SPACES AND REVERSE JENSEN INEQUALITIES

Stephen M. Buckley

Abstract. We examine the dependence on the A_p norm of w of the operator norms of singular integrals, maximal functions, and other operators in $L^p(w)$. We also examine connections between some fairly general reverse Jensen inequalities and the A_p and RH_p weight conditions.

1. Introduction

A question of considerable interest in harmonic analysis is, "What types of weights w have the property that T is bounded on $L^p(w)$?", where $1 < p < \infty$, and T is an operator which is bounded on the (unweighted) space L^p (typically T is the Hardy-Littlewood maximal operator, singular integral operators, or various related operators of interest in harmonic analysis). This type of question has been answered to a large extent by the work of Muckenhoupt, Hunt, Wheeden, Coifman, C. Fefferman, and others. In particular, it is known that Muckenhoupt's A_p condition is a necessary and sufficient condition for boundedness in the case of the Hardy-Littlewood maximal operator or singular integral operators (see [15], [13], and [4]). However, the dependence of the resulting operator norms on the "badness" of the A_p weight has never been adequately examined. We carry out this investigation in section 2, where we also give a new proof of the boundedness of the Hardy-Littlewood maximal operator on $L^p(w)$, for $w \in A_p$.

A_p and RH_p conditions are particular types of "reverse Jensen" inequalities which hold uniformly for all cubes. In section 3, we examine more general reverse Jensen inequalities (which hold uniformly for all cubes) with respect to some doubling measure μ on $\mathbf{R}^n$, and show how they are related to the usual $A_p(d\mu)$ and $RH_p(d\mu)$ conditions. Let us now introduce some notation and give some basic definitions.

For any set $S \subset \mathbf{R}^n$, $|S|$ is the Lebesgue measure of S. We will use the term "weight" to refer to any non-negative locally integrable function which is not everywhere zero. For any measure μ, we write $\fint_S g\, d\mu \equiv \frac{1}{\mu(S)} \int_S g\, d\mu$ (if μ is Lebesgue

Keywords and phrases. Singular integral, maximal function, weights.
1980 *Mathematics subject classifications* (1985 *Revision*). 42B20, 42B25.
The material in this work was drawn mainly from the author's Ph.D. thesis which was done under the supervision of Robert Fefferman at the University of Chicago.

Typeset by $\mathcal{AMS}$-TEX

measure, we write $g_S \equiv \fint_S g$. If w is a weight, we will write $w(S) = \int_S w$. By a "cube" in $\mathbf{R}^n$, we will mean an n-fold product of intervals of equal length (i.e. every face of the cube is perpendicular to a coordinate axis). If Q is a cube, rQ will denote the cube concentric with Q whose sidelength is r times that of Q (the "r-fold dilate" of Q). w will always denote a weight on $\mathbf{R}^n$ and p is a real number in the range $(1,\infty)$, unless otherwise stated. For any positive quantities X, Y, "$X \sim Y$" will mean "$1/C \le X/Y \le C$", where C is independent of the weight w (but may depend on n, p, and the operator T). For any exponent p, p' denotes the dual exponent $p/(p-1)$.

Definition. A singular integral operator is a principal value convolution operator $T : f \to K * f$ in $\mathbf{R}^n$, where the real-valued kernel K satisfies the following size and cancellation conditions:

$$
\begin{aligned}
||\hat{K}||_\infty &\le C \\
|K(x)| &\le \frac{C}{|x|^n} \\
|K(x) - K(x-y)| &\le \frac{C|y|}{|x|^{n+1}} \text{ for } |y| < \frac{|x|}{2}.
\end{aligned}
$$

T^* denotes the associated maximal singular integral operator which is defined by

$$T^* f(x) = \sup_{\epsilon>0} |(K.\chi_{\mathbf{R}^n \setminus B(0,\epsilon)}) * f(x)|.$$

Definition. If μ is a positive measure on $\mathbf{R}^n$, we say w is an $A_p(d\mu)$ weight (we write $w \in A_p(d\mu)$) if there is some $K > 0$ such that for all cubes $Q \in \mathbf{R}^n$,

$$\left(\fint_Q w \, d\mu\right) \left(\fint_Q w^{-1/(p-1)} \, d\mu\right)^{p-1} \le K. \tag{1.1}$$

We say w is an $A_1(d\mu)$ weight if, for all cubes $Q \in \mathbf{R}^n$,

$$\fint_Q w \, d\mu \le K \operatorname*{ess\,inf}_{x \in Q} w(x). \tag{1.2}$$

The smallest K for which (1.1) (or (1.2)) is true is referred to as the $A_p(d\mu)$-norm of w and will be denoted $K_{w,p;\mu}$ (resp. $K_{w,1;\mu}$) or simply $K_{w,\mu}$. We also write σ in place of $w^{-1/(p-1)}$ and refer to σ as the dual weight of w. It is easy to see that $w \in A_p(d\mu)$ if and only if $\sigma \in A_{p'}(d\mu)$ and that $K_{\sigma,p';\mu} = K_{w,p;\mu}^{p'-1}$. It is also clear that $w \in A_1(d\mu)$ if and only if $M_\mu w \le K w$ (M_μ indicates the Hardy-Littlewood maximal operator with respect to the measure μ).

Definition. We say w is an $A_\infty(d\mu)$ weight if, for all cubes Q, and all $E \subset Q$, we have

$$\frac{\nu(E)}{\nu(Q)} \leq C\left(\frac{\mu(E)}{\mu(Q)}\right)^{\epsilon} \tag{1.3}$$

for some $C, \epsilon > 0$, where $d\nu = w\,d\mu$.

Until section 3, we are interested only in $\mu =$ Lebesgue measure, and so we suppress references to μ (i.e. we write A_p, K_w, etc.). Weights of the form $w_r(x) = |x|^r$, the so-called power weights, provide the most basic examples of A_p weights; in fact $w_r \in A_p(\mathbf{R}^n)$ if and only if $-n < r < n(p-1)$. We have the following more precise estimates (the proof is straightforward and so we omit it).

Lemma 1.4. *If* $0 < \delta < 1$, *then* $u(x) \equiv |x|^{-n(1-\delta)} \in A_1$ *and* $K_{u,p} \sim \frac{1}{\delta}$, *for any* $p \geq 1$; *also,* $v(x) \equiv |x|^{n(p-1)(1-\delta)} \in A_p$ *and* $K_{v,p} \sim \frac{1}{\delta^{p-1}}$.

It is easy to prove that the dual space of $L^p(w)$ is $L^{p'}(\sigma)$. In addition, we have the following useful lemma, whose easy proof we also omit.

Lemma 1.5. *If a singular integral operator* T *is bounded on* $L^p(w)$ *and on* $L^{p'}(\sigma)$ *for some* $1 < p < \infty$, *then the two associated operator norms of* T *are equal.*

2. Bounds for operator norms

In this section, C *will denote a generic positive constant independent of everything, except possibly the dimension* n, *exponent* p, *and operator* T. *Also, for any weights given as examples,* δ *will denote a positive quantity which tends to 0.*

We now look at several important operators which are bounded on $L^p(w)$ spaces iff $w \in A_p$, and examine how the resulting operator norms depend on K_w, the A_p-norm of w. Our first main goal will be to do this for the Hardy-Littlewood maximal operator. We give a new proof of boundedness which gives a best possible dependence estimate. First of all, we need a few preparatory lemmas.

Lemma 2.1 [4]**.** *If* $w \in A_p$, *then* $w \in A_{p-\epsilon}$, *where* $\epsilon \sim K_{w,p}^{1-p'}$, *and* $K_{w,p-\epsilon} \leq CK_{w,p}$.

The next lemma, due to Besicovitch [1], is commonly referred to as the Besicovitch covering lemma. A proof of it can be found in [11, pp. 2-5]. Note that (ii) and (iii) just say that the sequence of cubes can be distributed into a bounded number of disjoint families.

Lemma 2.2. *Suppose that $A \subset \mathbf{R}^n$ is bounded and that for each $x \in A$, Q_x is a cube centered at x. Then we can choose, from among $\{Q_x : x \in A\}$, a (possibly finite) sequence $\{Q_i\}$ and an associated sequence of integers $\{m_i\}$ such that*

(i) $A \subset \bigcup_i Q_i$.
(ii) $1 \le m_i \le N_n$, *where N_n depends only on n.*
(iii) Q_i *and* Q_j *are disjoint if* $m_i = m_j$.

We say an operator is of weak-type p, with respect to the measure μ, if

$$\mu(\{Tf > \alpha\}) \le \left(\frac{C\|f\|_{L^p(d\mu)}}{\alpha} \right)^p .$$

The smallest such C is referred to as weak-type $L^p(d\mu)$-norm of T. We can now state a precise version of the Marcinkiewicz interpolation theorem with respect to a positive measure μ (the statement of this result given here, for μ being Lebesgue measure, is a special case of the result as proved by Zygmund [18]).

Lemma 2.3. *Suppose $1 \le p_0 < p_1 < \infty$ and that T is a sublinear operator of weak-type p_0 and p_1, with respect to the measure μ, with norms R_0 and R_1 respectively, then T is actually bounded on $L^p(d\mu)$ for all $p_0 < p < p_1$. In fact, for any $0 < t < 1$,*

$$\|Tf\|_{L^{p_t}(d\mu)} \le C_t R_0^{1-t} R_1^t \|f\|_{L^{p_t}(d\mu)}$$

where $\dfrac{1}{p_t} = \dfrac{1-t}{p_0} + \dfrac{t}{p_1}$ *and* $C_t^{p_t} = \dfrac{2^{p_t}}{p_t} \left(\dfrac{p_1}{p_1 - p_t} + \dfrac{p_0}{p_t - p_0} \right)$.

We shall only need to apply this lemma where p_0 and p_1 are in some fixed interval $[1, S]$, $t = 1/2$ and $R_1, R_2 \le R$. In this case, writing $p = p_{1/2}$, we simply get the inequality

$$\|Tf\|_{L^p(d\mu)} \le \frac{CR}{(p_1 - p_0)^{1/p}} \|f\|_{L^p(d\mu)}$$

where C depends only on S.

Lemma 2.4. *If $f \in L^p(w)$ and $f_{Q_k} \ge \alpha > 0$ for each of the disjoint cubes $\{Q_k\}$, then*

$$\sum_k w(Q_k) \le K_w \left(\frac{\|f\|_{L^p(w)}}{\alpha} \right)^p .$$

Proof. We can assume that $f(x) \ge 0$ and that $\|f\|_{L^p(w)} = 1$, without loss of generality. Now,

$$\begin{aligned} \sum_k w(Q_k) &\le \int \sum_k \frac{w(Q_k)}{\alpha |Q_k|} \chi_{Q_k} f \\ &\le \left\| \sum_k \frac{w(Q_k)}{\alpha |Q_k|} \chi_{Q_k} \right\|_{L^{p'}(\sigma)} \|f\|_{L^p(w)} \end{aligned}$$

$$\leq \left(\sum_k \frac{w(Q_k)^{p'}}{\alpha^{p'} |Q_k|^{p'}} \sigma(Q_k) \right)^{1/p'}$$

$$\leq \frac{K_\sigma^{1/p'}}{\alpha} \left(\sum_k w(Q_k) \right)^{1/p'}, \qquad \text{since } \sigma \in A_{p'}$$

and so,

$$\sum_k w(Q_k) \leq \frac{K_\sigma^{p/p'}}{\alpha^p} = \frac{K_w}{\alpha^p}. \quad \square$$

Using lemma 2.4, our first main theorem is now easy to state and prove.

Theorem 2.5. *If $w \in A_p$, then $\|Mf\|^p_{L^p(w)} \leq CK_w^{p'} \|f\|^p_{L^p(w)}$. The power $K_w^{p'}$ is best possible.*

Proof. First, we show that for $1 \leq p < \infty$,

$$w(\{Mf > \alpha\}) \leq CK_w(\|f\|_{L^p(w)}/\alpha)^p. \tag{2.6}$$

Without loss of generality, we assume that $f(x) \geq 0$ and that $\|f\|_{L^p(w)} = 1$. Suppose that $Mf(x) > \alpha > 0$ so that $f_{Q_k} \geq \alpha$ for some cube Q_x centered at x. Let $A_r = \{x : |x| < r, Mf(x) > \alpha\}$. The Besicovich covering lemma tells us that A_r can be covered by the union of N_n collections of disjoint cubes, on each of which the mean value of f is at least α. Choose the collection $\{Q_k\}$, whose union has maximal w-measure. Thus, $w(A_r) \leq N_n w\left(\bigcup_k Q_k\right) \leq CK_w/\alpha^p$, by lemma 2.4. Letting $r \to \infty$, we get (2.6).

Suppose now that $p > 1$. By lemma 2.1, w is also an $A_{p-\epsilon}$ weight with comparable norm, where $\epsilon \sim K_{w,p}^{1-p'}$ and, trivially, w is an $A_{p+\epsilon}$ weight, with norm no larger than $K_{w,p}$. Applying the Marcinkiewicz interpolation theorem to the corresponding weak-type results at $p - \epsilon$ and $p + \epsilon$, we get the strong-type result we require with the indicated bound for the operator norm.

To see that the power $K_w^{p'}$ is best possible, we give an example for $\mathbf{R}$ (a similar example works in $\mathbf{R}^n$ for any n). Let $w(x) = |x|^{(p-1)(1-\delta)}$, so that $K_w \sim 1/\delta^{p-1}$ by lemma 1.4. Now, $f(x) = |x|^{-1+\delta}\chi_{[0,1]} \in L^p(w)$. It is easy to see that $Mf \geq \dfrac{f}{\delta}$ and so, $|Mf\|^p_{L^p(w)}/\|f\|^p_{L^p(w)} \geq C\delta^{-p} \sim K_w^{p'}$. $\square$

Remark 2.7. The proof of Coifman and Fefferman [4], will also give the best possible exponent $K_w^{p'}$, when the proof is examined closely, but some other proofs of the boundedness of M, e.g. [14], will not do so. The dependence in the weak-(p,p) inequality (2.6) was found and shown to be best possible by Muckenhoupt [15].

Remark 2.8. It is easy to prove, using (2.6), that if $w \in A_q$ for some $q < p$, then $\|Mf\|^p_{L^p(w)} \le CK_{w,q}\|f\|^p_{L^p(w)}$, where $C = C_{p,q}$ now depends on q as well as p (and $C_{p,q}$ gets very large when q is very close to p).

Theorem 2.5 neatly sews up the dependence for the Hardy-Littlewood maximal operator. The dependence for singular integral operators is not at all as easy to handle and, in fact, we shall not be able to find the best power of K_w. The best we can do is as follows, which actually takes care of the maximal singular integral operator T^*.

Theorem 2.9. *If $w \in A_p$, then $\|T^*f\|^p_{L^p(w)} \le CK_w^{p'+p}\|f\|^p_{L^p(w)}$. The best power of K_w in this inequality must lie in the interval* $[\max(p, p'), p' + p]$.

Proof. The proof of the boundedness of T^* on $L^p(w)$ for $w \in A_p$ given in [4] will give the required exponent, as long as we sharpen one of the inequalities used, namely the good-λ inequality

$$|\{x \in Q : T^*f > 2\alpha, Mf < \gamma\alpha\}| \le C\gamma|Q|,$$

which holds for any cube Q in the Whitney decomposition of $\{T^*f > \alpha\}$. We replace it by the sharp good-λ inequality

$$|\{x \in Q : T^*f > 2\alpha, Mf < \gamma\alpha\}| \le Ce^{-c/\gamma}|Q|, \tag{2.10}$$

for such cubes, which is proven in lemma 2.13 below.

To see that the dependence is best possible, we give examples on $\mathbf{R}$ (similar examples can be found in $\mathbf{R}^n$ for $n > 1$). Choose $w(x) = |x|^{(p-1)(1-\delta)}$ and $f(x) = |x|^{-1+\delta}\chi_{[0,1]}$, so that $\int f^p w = 1/\delta$. For $x > 2$, $Hf(x) \sim 1/\delta x$ and so

$$\int_2^\infty |Hf|^p w \sim 1/\delta^{p+1} \sim \delta^{-p}\|f\|^p_{L^p(w)}.$$

Since $\delta^{-p} \sim K_w^{p'}$ the best power must be at least p'. Since the operator norm for $T : L^p(w) \to L^p(w)$ can be at least $CK_w^{p'/p}$, the operator norm for $T : L^{p'}(\sigma) \to L^{p'}(\sigma)$ is also at least $CK_w^{p'/p} = CK_\sigma$. Thus, the best power in our estimate must be at least $\max(p, p')$ (an explicit example is provided by $f(x) = x^\delta\chi_{[0,1]}(x)$ and $T = H$; it is easy to show that $\|Hf\|_{L^p(w,[-1,0])} \ge CK_w\|f\|_{L^p(w)}$. $\square$

We must now prove the sharp good-λ inequality (2.10). The proof is a modification, using standard good-λ techniques, of Hunt's main result in [12] which deals with the conjugate function on the unit circle. First we state an elementary lemma which is needed.

Lemma 2.11. *Suppose $f \in L^\infty(Q)$ and that T is an operator for which*

$$|\{x : Tg(x) > \alpha\}| \le \left(\frac{Cp\,\|g\|_p}{\alpha}\right)^p,$$

for all $g \in L^p$ and sufficiently large p and α, C being a constant independent of p. Then,

$$|\{x : Tf(x) > \alpha\}| \le Ce^{-\frac{\alpha}{e\|f\|_\infty}}|Q|.$$

Proof. It suffices to prove this result for large α, since the result is trivial otherwise. Let $K = \|f\|_\infty$. Since $f \in L^\infty(Q)$, $f \in L^p(Q)$ for all $1 < p < \infty$ and $\|f\|_p \le |Q|^{1/p}K$, and so

$$|\{x \;:\; Tf(x) > \alpha\}| \le \left(\frac{Cp\|f\|_p}{\alpha}\right)^p \le C|Q|\left(\frac{pK}{\alpha}\right)^p.$$

Letting $p = \dfrac{\alpha}{eK}$, we get the required result. $\square$

Remark 2.12. If T is the maximal operator, a singular integral operator, or a maximal singular integral operator, then it satisfies the condition of the above lemma (see [16, p. 48]).

Lemma 2.13. *Let $\Omega = \bigcup Q_j$ be the Whitney covering of $\{T^*f > \alpha\}$. Then*

$$|\{x \in Q_j \;:\; T^*f(x) > 2\alpha, Mf(x) \le \gamma\alpha\}| \le Ce^{-c/\gamma}|Q_j|.$$

Proof. We can clearly assume that $Mf(x_0) < \gamma\alpha$ for some $x_0 \in Q_j$, and that γ is small. We write $f = f_1 + f_2$, where $f_1 = f_{100Q_j}$, and $f_2 = f_{\mathbf{R}^n\setminus 100Q_j}$. By standard estimation (as in the proof of theorem III in [4]), we get that, for $x \in Q_j$, $T^*f_2(x) \le \alpha + C\gamma\alpha < \dfrac{3\alpha}{2}$, if γ is small enough.

To handle f_1, we first let $\Omega' = \bigcup P_k$ be the Whitney decomposition of $\{Mf_1 > 2^n\gamma\alpha\}$, where n is the dimension. Note that $\|f_1\| \le (101)^n\gamma\alpha|Q_j|$, and so $\Omega' \subset 200Q_j$. Let

$$g(x) = \begin{cases} f_1(x), & x \notin \Omega' \\ (f_1)_{P_k}, & x \in P_k \end{cases}$$

and $b = f_1 - g$. Then g is supported in $200Q_j$, $\|g\|_\infty \le C\gamma\alpha$ and so, by lemma 2.11,

$$|\{x \in Q_j : T^*g > \frac{\alpha}{4}\}| \le Ce^{-c/\gamma}|Q_j|.$$

As for b, let us define $\Omega'' = \bigcup 2P_k$. Since $\int_{P_k} b = 0$, we have, for $x \notin \Omega''$,

$$T^*b(x) \le \sum_k \int_{P_k} |b(t)||K(x-t) - K(x-t_k)|,$$

where t_k is the centre of P_k. It follows that

$$\begin{aligned} T^*b(x) &\le \sum_k \int_{P_k} |b(t)| \left(\frac{\delta_k}{\delta_k^{n+1} + (t_k - x)^{n+1}} \right) dt \\ &\le C\gamma\alpha \sum_k \frac{\delta_k}{\delta_k^{n+1} + (t_k - x)^{n+1}} \equiv C\gamma\alpha\Delta(x) \end{aligned}$$

where δ_k is the diameter of P_k.

Carleson's [2] exponential estimate of Δ tells us that

$$|\{x \in Q_j : \Delta(x) > \frac{c}{\gamma}\}| \le Ce^{-c/\gamma}|Q_j|,$$

and so, since $Mf(x) > \gamma\alpha$ if $x \in \Omega''$,

$$|\{x \in Q_j : T^*b(x) > \alpha/4, Mf(x) \le \gamma\alpha\}| \le Ce^{-c/\gamma}|Q_j|.$$

This, together with our estimates for f_2 and g, is easily seen to imply the desired result. □

Let us now examine K_w-dependence of operator norms for a particular class of weights, namely power weights. In the case of the Hardy-Littlewood maximal operator on power weighted spaces, we can clearly do no better than theorem 2.5 (or remark 2.8 for negative power weights, which are in A_1), since all our examples so far have involved power weights. However, in contrast to the case of general A_p weights, we can also give a best possible dependence result for singular integral operators.

Theorem 2.14. *If T is a singular integral operator on $\mathbf{R}^n$ and $0 < \delta < 1$, then*

(i) $w(\{T^*f > \alpha\}) \le \dfrac{CK_w}{\alpha} \int |f|w,$ *if* $w(x) = |x|^{-n(1-\delta)}$.

(ii) $\int |T^*f|^p w \le CK_w^p \int |f|^p w,$ *if* $w(x) = |x|^{-n(1-\delta)}$.

(iii) $\int |T^*f|^p w \le CK_w^{p'} \int |f|^p w,$ *if* $w(x) = |x|^{n(p-1)(1-\delta)}$.

The exponents in (i)-(iii) are best possible.

Proof. We first prove (i). By normalization, we can assume that $\|f\|_{L^1(w)} = 1$. We write $A_j = \{x \in \mathbf{R}^n : 2^j \le |x| < 2^{j+1}\}$, $f_j = f\chi_{A_j}$, $f_{j,1} = f\chi_{\{|x|<2^{j+2}\}}$, and $f_{j,2} = f - f_{j,1}$. Clearly,

$$\begin{aligned} w(\{T^*f > \alpha\}) &= \sum_{j=-\infty}^{\infty} w(\{T^*f > \alpha\} \cap A_j) \\ &\le \sum_{i=1}^{2} \sum_{j=-\infty}^{\infty} w(\{T^*f_{j,i} > \alpha/2\} \cap A_j) = \sum_{i=1}^{2} S_i, \qquad \text{say.} \end{aligned}$$

Now,

$$
\begin{aligned}
S_1 &\le \sum_{j=-\infty}^{\infty} 2^{-jn(1-\delta)} |\{T^* f_{j,1} > \alpha/2\} \cap A_j| \\
&\le C \left(\sum_{j=-\infty}^{\infty} \frac{2^{-jn(1-\delta)}}{\alpha} \int |f_{j,1}| \right), \qquad \text{by the unweighted theory} \\
&= \frac{C}{\alpha} \sum_{j=-\infty}^{\infty} 2^{-jn(1-\delta)} \sum_{k \le j+1} \int |f_k| \\
&= \frac{C}{\alpha} \sum_{k=-\infty}^{\infty} \int |f_k| \left(\sum_{j \ge k-1} 2^{-jn(1-\delta)} \right) \\
&\le \frac{C}{\alpha} \sum_{k=-\infty}^{\infty} 2^{-nk(1-\delta)} \int |f_n| \le \frac{C}{\alpha}.
\end{aligned}
$$

As for S_2, we note first that,

$$
\begin{aligned}
1 = \int |f| w = \sum_k \int_{A_k} |f| w &\ge C \sum_{k=-\infty}^{\infty} 2^{-kn(1-\delta)} \int_{A_k} |f| \\
&\ge C \sum_{k=-\infty}^{\infty} 2^{kn\delta} ⨍_{A_k} |f|
\end{aligned}
$$

and so, if $x \in A_j$ then

$$
\begin{aligned}
T^* f_{j,2}(x) \le \sum_{k > j+1} T^* f_k(x) &\le C \left(\sum_{k > j+1} \int_{A_k} \frac{|f(y)|}{|x-y|^n} \right) \\
&\le C \left(\sum_{k > j+1} ⨍_{A_k} |f| \right) \\
&\le C 2^{-jn\delta}.
\end{aligned}
$$

But if $2^{-jn\delta} > c\alpha$, then $j < \dfrac{-\log_2 c\alpha}{n\delta} \equiv j_0$. It follows that

$$
S_2 \le \int_{|x| < 2^{j_0}} |x|^{-n(1-\delta)}\, dx \sim \frac{2^{n\delta j_0}}{\delta} \sim \frac{1}{\alpha\delta} \sim \frac{K_w}{\alpha}.
$$

We next prove (iii). Here $w(x) = |x|^{(p-1)(1-\delta)}$. We define A_j as before, but now we define $f_{j,1} = f\chi_{\{|x| < 2^{j-1}\}}$, and $f_{j,2} = f - f_{j,1}$. Now, as in the S_1 case of (i),

$$
\begin{aligned}
\sum_{j=-\infty}^{\infty} \int_{A_j} |T^* f_{j,2}|^p w &\le C \sum_{j=-\infty}^{\infty} 2^{jn(p-1)(1-\delta)} \int_{A_j} |T^* f_{j,2}|^p \\
&\le C \sum_{j=-\infty}^{\infty} 2^{jn(p-1)(1-\delta)} \int |f_{j,2}|^p
\end{aligned}
$$

$$
\begin{aligned}
&= C \sum_{k=-\infty}^{\infty} \int |f_k| \left(\sum_{j \le k+1} 2^{jn(p-1)(1-\delta)} \right) \\
&\le C \sum_{k=-\infty}^{\infty} 2^{kn(p-1)(1-\delta)} \int |f_n| \\
&\le C \int |f|^p w.
\end{aligned}
$$

As for the other terms, it is easy to see that if $x \in A_j$ then $T^* f_{j,1} \le CMf_{j,1} \le CMf(x)$. Now using theorem 2.5, we get

$$\sum_{j=-\infty}^{\infty} \int_{A_j} |T^* f_{j,1}|^p w \le C \int |Mf|^p w \le CK_w^{p'} \int |f|^p w.$$

Part (iii) now follows readily from the estimates of the last two paragraphs.

We next prove (ii). To see this, let $w(x) = |x|^{n(p-1)(1-\delta)}$. By (iii), we have

$$\|T^* f\|_{L^p(w)} \le CK_w^{p'/p} \|f\|_{L^p(w)},$$

and so by lemma 1.5,

$$\|T^* f\|_{L^{p'}(\sigma)} \le CK_w^{p'/p} \|f\|_{L^{p'}(\sigma)}.$$

But $K_{w,p}^{p'/p} = K_{\sigma,p'}$ and $\sigma(x) = |x|^{-n(1-\delta)}$, giving us our required result on $L^{p'}(\sigma)$.

We are left with giving examples to show that the exponents in (i)-(iii) are best possible. In (i), we let $f = \chi_{[1,2]}$, $T = H$, the Hilbert transform, and $\alpha = 1/3$. Then $|Hf(x)| > 1/2$ for $x \in [0,1]$ and $\int_0^1 |x|^{-1+\delta}\, dx = \frac{1}{\delta} \sim K_w$. In (ii) and (iii), the examples given in the proof of theorem 2.9 suffice. □

We now turn our attention to the Marcinkiewicz integral operator J_Ω, associated with an open set Ω of finite measure, which is defined for all $f : \Omega \to [0, \infty)$ by the equation

$$J_\Omega(f)(x) = \int_\Omega f(y) \frac{\delta(y)}{\delta(y)^{n+1} + |x-y|^{n+1}}\, dy,$$

where $\delta(y) = \operatorname{dist}(y, \Omega^c)$. This is the version of the Marcinkiewicz integral operator used by Carleson in [2]. J_Ω is an important tool for controlling singular integral operators (see [16]). The following result summarizes the dependence of the resulting operator norm on the A_p norm of w.

Theorem 2.15. *If $1 \le p < \infty$, then J_Ω is bounded on $L^p(w)$ uniformly for all open Ω of finite measure if and only if $w \in A_p$. Furthermore, $\|J_\Omega f\|_{L^p(w,\mathbf{R}^n)} \le CK_w \|f\|_{L^p(w,\Omega)}$. The dependence on K_w is best possible.*

To prove theorem 2.15, we first need the following lemma.

Lemma 2.16. *For any functions $f, g \geq 0$,*

$$\int_{\mathbf{R}^n} (J_\Omega f) g \leq C \int_\Omega f M g.$$

Proof. By Fubini's theorem,

$$\int_{\mathbf{R}^n} (J_\Omega f)(x) g(x)\, dx = \int_\Omega f(y)\delta(y) \left(\int_{\mathbf{R}^n} \frac{g(x)}{\delta(y)^{n+1} + |x-y|^{n+1}}\, dx \right) dy.$$

Letting $\mathcal{A}_k = \{u \in \mathbf{R}^n : 2^k \delta(y) \leq |u| \leq 2^{k+1}\delta(y)\}$ and making the change of variable $u = x - y$, we get

$$\begin{aligned} \int_{\mathbf{R}^n} \frac{g(x)}{\delta(y)^{n+1} + |x-y|^{n+1}}\, dx &\leq \int_{|u| \leq \delta(y)} \frac{g(y+u)\, du}{\delta(y)^{n+1}} + \sum_{k=0}^{\infty} \int_{\mathcal{A}_k} \frac{g(y+u)\, du}{|u|^{n+1}} \\ &\leq \frac{CMg(y)}{\delta(y)} \left(1 + \sum_{k=0}^{\infty} 2^{-k} \right) \\ &\leq \frac{CMg(y)}{\delta(y)}. \end{aligned}$$

The required result now follows easily.

Proof of theorem 2.15. Suppose J_Ω is bounded on $L^p(w)$ uniformly for all open Ω of finite measure. Fix a cube Q, and let $\Omega = 2Q$. Then, for any non-negative function f supported on Q, $J_{2Q} f(x) \sim f_Q$ for all $x \in Q$. If $p > 1$, let $f = \chi_Q w^{-1/(p-1)}$. Because J_{2Q} is bounded on $L^p(w)$, it follows that

$$\left(\int_Q w \right) \left(⨍_Q w^{-1/(p-1)} \right)^p \leq C \int_Q w^{-1/(p-1)}$$

which clearly implies $w \in A_p$.

For the case $p = 1$, let $f = \chi_S$ for an arbitrary measurable subset S of Q. The boundedness of J_{2Q} on $L^p(w)$ now implies that

$$\frac{|S|}{|Q|} w(Q) \leq C w(S).$$

If we take $S = \{x \in Q : w(x) < a\}$ for any $a > a_0 \equiv \operatorname{ess\,inf}_Q w$, and then let $a \to a_0$, we get $w_Q \leq C a_0$, i.e. $w \in A_1$.

In proving the converse, we may assume, without loss of generality, that f is supported on Ω and that $\|f\|_{L^p(w)} = 1$. The case $p = 1$ follows by letting $g = w$ in lemma 2.16, so we assume $1 < p < \infty$. If g is a function for which $\|g\|_{L^{p'}(\sigma)} = 1$, then lemma 2.16 tells us that,

$$\begin{aligned}\int_{\mathbf{R}^n} (Jf)g &\le C \int_\Omega fMg \\ &\le C\left(\int_\Omega f^p w\right)^{1/p}\left(\int_\Omega (Mg)^{p'}\sigma\right)^{1/p'} \\ &\le CK_{\sigma,p'}^{p/p'} = CK_{w,p}.\end{aligned}$$

The required boundedness follows by duality. To see that this dependence is best possible, we let $w(x) = |x|^{-1+\delta}$, $f(x) = x^\delta \chi_{[0,1]}(x)$, and $\Omega = (0,1)$. Now for $0 < x < 1$, it is clear that

$$J_\Omega f(-x) > \frac{1}{2}\int_x^1 y^{\delta-1}\,dy = \frac{(1-x^\delta)}{2\delta}$$

and so,

$$\|Jf\|^p_{L^p(w,[-1,0])} > \frac{1}{2^p\delta^p}\int_{-1}^0 (|x|^{-1+\delta} - p|x|^{-1+(p+1)\delta})\,dx = \frac{1}{(p+1)2^p\delta^{p+1}}$$

whereas $\|f\|^p_{L^p(w)} = 1/(p+1)\delta$. Since $K_w \sim 1/\delta$, it follows that $\|Jf\|_{L^p}/\|f\|_{L^p(w)} \ge CK_w$, as required. □

Remark 2.17. Our three operator dependence results

$$\begin{aligned}\|Mf\|^p_{L^p(w)} &\le CK_w^{p'}\|f\|^p_{L^p(w)} \\ \|Tf\|^p_{L^p(w)} &\le CK_w^{p+p'}\|f\|^p_{L^p(w)} \\ \|Jf\|^p_{L^p(w)} &\le CK_w^p\|f\|^p_{L^p(w)}\end{aligned}$$

tie together well intuitively because, if f is a function of bounded support B then, roughly speaking, Jf can be as "nasty" as Tf near B, but tends to be smaller than it far from B, whereas Mf can control Tf far away from B, but not near B.

By way of contrast with the K_w-dependence of the above operators, let us finish by looking at simple averaging operators of the form $T_Q(f) = f * \dfrac{\chi_Q}{|Q|}$, where Q is some cube. T_Q is of course dominated by the maximal operator, which proves that for any $w \in A_p$, T_Q is bounded on $L^p(w)$ (at least for $1 < p < \infty$) with norm-dependence on w of the form $K_w^{p'}$. Intuitively, however, T_Q is so "close" to the identity operator that we expect to be able to get a better exponent than p'. The following lemma shows that this is indeed the case (simple examples show it is best possible).

Lemma 2.18. *If $1 \le p < \infty$, T_Q is bounded on $L^p(w)$ uniformly for all cubes Q if and only if $w \in A_p$. Furthermore, for any cube Q centered at 0, $\|T_Q f\|^p_{L^p(w)} \le CK_w \|f\|^p_{L^p(w)}$.*

Proof. We may assume, without loss of generality, that $f \ge 0$. This allows us to also assume Q is centered at 0, since otherwise we can bound $T_Q f$ by a constant (dependent on Q) times $T_{Q'} f$ where Q' is the smallest cube centered at 0 containing Q.

Divide $\mathbf{R}^n$ into the unique mesh M of cubes of equal sidelength and disjoint interiors for which $Q \in M$. Suppose first that f is supported in some $Q_0 \in M$ and so $T_Q f$ is supported in $3Q_0$. If $p > 1$ then by Hölder's inequality,

$$\begin{aligned}\int_{3Q_0} \left(\frac{\chi_Q}{|Q|} * f \right)^p w &\le \frac{1}{|Q|^p} \int_{3Q_0} \left(\int_{(Q+x)\cap Q_0} f^p(y) w(y)\, dy \right) \sigma(Q_0)^{p-1} w(x)\, dx \\ &= \frac{\sigma(Q_0)^{p-1}}{|Q|^p} \left(\int_{3Q_0} w(x)\, dx \right) \int f^p w \le CK_w \int f^p w,\end{aligned}$$

as required. In the case $p = 1$, we simply estimate

$$\begin{aligned}\frac{1}{|Q|} \int_{3Q_0} \left(\int_{Q+x} f(y)\, dy \right) w(x)\, dx &\le \frac{C}{|Q|} \left(\int fw \right) \left(\int_{3Q_0} w(x)\, dx \right) / \operatorname*{ess\,inf}_{y \in Q_0} w(y) \\ &\le CK_w \left(\int fw \right).\end{aligned}$$

For a general function f, we simply decompose $f = \sum_{C \in M} f\chi_C$, and we get the required result because of the limited amount of overlap among the supports of the functions $\{\frac{\chi_Q}{|Q|} * (f\chi_C)\}_{C \in M}$. $\square$

3. Reverse Jensen Inequalities

In this section, we examine some rather general reverse Jensen inequalities and show their connection to the conditions $A_p(d\mu)$ and $RH_p(d\mu)$. The $RH_p(d\mu)$ condition (defined below) was first examined by Gehring [10] (in the case $\mu =$ Lebesgue measure), and it was Coifman and C. Fefferman [4] who first showed the close relation between RH_p and A_p conditions (they showed that a weight is in some A_p space if and only if it is in some RH_q space, but there is no possible relationship between p and q).

Since then, the RH_p condition has become important in its own right in the theory of elliptic operators on Lipschitz spaces. Dahlberg [5] showed that the Dirichlet problem for such operators is solvable with L^p boundary values if and only if harmonic measure is in $RH_p(d\sigma)$, where $d\sigma$ is surface measure. For further results in this direction, see [6], [7] and [8].

We say a positive measure μ is a doubling measure, $\mu \in \mathbf{D}$, if $\mu(2Q) \leq C\mu(Q)$ for all cubes Q. We say w is a doubling weight if $w\,dx$ is a doubling measure. If Q is a cube we denote by $l(Q)$ the sidelength of Q. We define $\log^+ x \equiv \log(2+x)$.

Definition. If $\mu \in \mathbf{D}$ and $1 < p < \infty$, we say that w is a $RH_p(d\mu)$ weight if

$$\left(\fint_Q w^p \, d\mu\right)^{1/p} \leq K \fint_Q w \, d\mu \tag{3.1}$$

for all cubes Q. The smallest such K is referred to as the $RH_p(d\mu)$-norm of w.

Condition (3.1) is often called a "reverse Hölder inequality", because it is Hölder's inequality with the direction of the inequality reversed (Hölder's inequality is of course true with $K = 1$). More generally, if $\|.\|_{1,Q}$ and $\|.\|_{2,Q}$ are norms for functions defined on an arbitrary cube Q, and Jensen's inequality implies that $\|f\|_{1,Q} \leq C_1 \|f\|_{2,Q}$, then we refer to the condition $\|w\|_{2,Q} \leq C_2 \|w\|_{1,Q}$ as a reverse Jensen inequality (we will only be interested in such inequalities when they hold uniformly for all cubes Q).

If μ_1 and μ_2 are positive doubling measures, we say that μ_1 is comparable to μ_2 if there exist $\alpha, \beta \in (0,1)$ such that $\dfrac{\mu_1(E)}{\mu_1(Q)} < \beta$ whenever $\dfrac{\mu_2(E)}{\mu_2(Q)} < \alpha$ for every $E \subseteq Q$, and every cube Q. Let us now state a result taken directly from [4], which is very useful for our purposes.

Lemma 3.2. *If μ_1 and μ_2 are positive doubling measures, the following are equivalent*

(i) *There exists $C, \delta > 0$ such that for every $E \subseteq Q \subset \mathbf{R}^n$,*

$$\frac{\mu_2(E)}{\mu_2(Q)} \leq C \left(\frac{\mu_1(E)}{\mu_1(Q)}\right)^{\delta}.$$

(ii) *μ_2 is comparable to μ_1.*
(iii) *μ_1 is comparable to μ_2.*
(iv) *$d\mu_2 = w(x)\, d\mu_1(x)$ and for every cube Q,*

$$\left(\fint_Q w^{1+\delta} \, d\mu_1\right)^{\frac{1}{1+\delta}} \leq C \fint_Q w \, d\mu_1.$$

Lemma 3.2 allows us to prove the following lemma, which generalizes to $A_p(d\mu)$ and $RH_p(d\mu)$, results which are well-known for $\mu =$ Lebesgue measure. For the rest of the section, μ is an arbitrary but fixed doubling measure on $\mathbf{R}^n$, and $d\nu = w\, d\mu$.

Lemma 3.3. *If $w \in A_\infty(d\mu)$ then $\nu \in \mathbf{D}$. Furthermore,*

$$A_\infty(d\mu) = \bigcup_{1<p<\infty} RH_p(d\mu) = \bigcup_{1<q<\infty} A_q(d\mu).$$

Proof. If we prove that $\nu \in \mathbf{D}$, the rest of the lemma follows fairly easily from lemma 3.2 (alternatively, it is implied by theorem 15 in chapter 1 of Strömberg and Torchinsky [17]), so we shall confine ourselves to proving that $\nu \in \mathbf{D}$.

First note that the condition $\mu \in \mathbf{D}$ is equivalent to assuming there is some $C_1 > 0$ for which

$$\mu(Q') \leq C_1 \mu(Q) \tag{3.4}$$

for all cubes Q, Q' which are adjacent and of equal size.

We will now show, roughly speaking, that a very thin slice from a side of a cube has very small μ-measure compared with the full cube. For simplicity, we will prove this for the cube $Q_0 = \{x : |x_i| \leq 1\}$ and the slice $S_\epsilon \equiv \{x : |x_i| \leq 1, x_1 > 1 - \epsilon\}$.

We divide Q_0 into 2^n cubes of sidelength 1, half of which are in the slice S_1. Applying the estimate (3.4) to each subcube in S_1 and its adjacent subcube in $Q_0 \backslash S_1$, gives us the inequality $\mu(S_1) \leq \dfrac{C_1}{C_1+1}\mu(Q_0)$. This process can be continued to give $\mu(S_{2^{-(k+1)}}) \leq \dfrac{C_1}{C_1+1}\mu(S_{2^{-k}})$ (to see this, simply divide $S_{2^{-k}}$ into 2^{kn+n-k} cubes of sidelength $2^{-(k+1)}$, half of which are in $S_{2^{-(k+1)}}$, and half in $S_{2^{-k}} \backslash S_{2^{-(k+1)}}$). Thus $\mu(S_{2^{-k}}) \leq \left(\dfrac{C_1}{C_1+1}\right)^{k+1} \mu(Q_0)$, and so

$$\mu(S_\epsilon)/\mu(Q) \to 0 \qquad (\epsilon \to 0). \tag{3.5}$$

Clearly, the above argument will work equally well if we let Q_0 be an arbitrary cube and S_ϵ be a slice of thickness $\dfrac{l(Q)}{2\epsilon}$ and, in fact, the convergence in (3.5) is uniform for all such cubes and slices.

We are now ready to show that $\nu \in \mathbf{D}$. Given a cube Q_0, let us write $Q_\epsilon \equiv (1+\epsilon)Q_0$ for any $\epsilon > 0$. Since one can get Q_0 from Q_ϵ by removing $2n$ slices of thickness $\epsilon l(Q)$, it follows that

$$\frac{\mu(Q_\epsilon \backslash Q_0)}{\mu(Q_\epsilon)} \to 0 \qquad (\epsilon \to 0)$$

uniformly for all cubes Q_0. Using the fact that $\nu \in A_\infty(d\mu)$ we conclude that $\dfrac{\nu(Q_\epsilon \backslash Q_0)}{\nu(Q_\epsilon)} < \dfrac{1}{2}$ for some sufficiently small $\epsilon > 0$. Thus $\nu(Q_\epsilon) < 2\nu(Q_0)$ which we can iterate to get the doubling condition $\nu(Q_1) < 2^k \nu(Q_0)$ for any $k > \log_{1+\epsilon}(2)$. $\square$

Given exponents $0 < q < p < \infty$, it is natural to consider the more general reverse Hölder's inequality,

$$\left(\fint_Q w^p \, d\mu \right)^{1/p} \leq K \left(\fint_Q w^q \, d\mu \right)^{1/q}. \tag{3.6}$$

Let us denote by $RH_{p,q}(d\mu)$ the class of weights satisfying (3.6) for all cubes Q. In fact we have not introduced anything new: if $p > 1$ then $RH_{p,q}(d\mu) = RH_p(d\mu)$ for any $0 < q < p$. This follows as a special case of the "self-improving" nature of these weights: if $w \in RH_{p,q}(d\mu)$, then $w \in RH_{p,r}(d\mu)$ for any $0 < r < q$. To see this, we use both reverse and normal Hölder inequalities to get

$$\begin{aligned} \left(\fint_Q w^p \, d\mu \right)^{q/p} &\leq K \left(\fint_Q w^q \, d\mu \right) \\ &\leq K \left(\fint_Q w^p \, d\mu \right)^{\frac{q-r}{p-r}} \left(\fint_Q w^r \, d\mu \right)^{\frac{p-q}{p-r}} \end{aligned}$$

which clearly implies that $w \in RH_{p,r}(d\mu)$.

It is reasonable to extend the definition of $RH_p(d\mu)$ so that it is defined for all $p > 0$ by the equation $RH_p(d\mu) \equiv RH_{p,q}(d\mu)$ for any $q < p$. The next lemma, which links $RH_p(d\mu)$ with $A_\infty(d\mu)$, is now easy to prove (this lemma is also to be found in Strömberg and Torchinsky[17]).

Lemma 3.7. *If $0 < p < \infty$, then $w \in RH_p(d\mu)$ if and only if $w^p \in A_\infty(d\mu)$.*

Proof. If $w^p \in A_\infty(d\mu)$ then, by lemma 3.3, $w^p \in RH_q(d\mu)$ for some $q > 1$. Thus $w \in RH_{pq}(d\mu) \subset RH_p(d\mu)$.

For the converse, we may assume $p = 1$. If $w \in RH_1(d\mu)$ then, for any $0 < q < 1$, $w^q \in RH_{1/q}(d\mu)$ and so $w^q \in A_p(d\mu)$ for some $1 < p < \infty$ by lemma 3.3. It follows that

$$\left(\fint_Q w \, d\mu \right)^q \left(\fint_Q w^{\frac{-q}{p-1}} \, d\mu \right)^{p-1} \leq C \left(\fint_Q w^q \, d\mu \right) \left(\fint_Q w^{\frac{-q}{p-1}} \, d\mu \right)^{p-1} \leq C.$$

The inequality between the first and last terms is essentially the defining inequality for $w \in A_{\frac{p-1+q}{q}}(d\mu)$, and so $w \in A_\infty(d\mu)$. $\square$

Remark 3.8. We showed at the beginning of this proof that if $w \in RH_p(d\mu)$, then $w \in RH_{p+\epsilon}(d\mu)$ for some $\epsilon > 0$. This analog for RH_p of lemma 2.1 was first proved by Gehring [10] in the case $\mu =$ Lebesgue measure.

The following lemma gives a couple of useful alternative characterizations of $A_\infty(d\mu)$ (the first of which is a reverse Jensen inequality). Part (i), for $\mu =$ Lebesgue measure, is due to García-Cuerva and Rubio de Francia [9], and part (ii) is due to Coifman and C. Fefferman [4].

Lemma 3.9. *Each of the following is equivalent to $w \in A_\infty(d\mu)$.*

(i) *For all cubes Q,*

$$\fint_Q w\, d\mu \le C \exp\left(\fint_Q \log w\, d\mu\right).$$

(ii) *There are constants α and β such that for all cubes Q,*

$$\mu(\{x \in Q : w(x) \ge \beta\nu(Q)/\mu(Q)\}) > \alpha\mu(Q). \tag{3.10}$$

Proof. We prove only (i), as the easy proof of (ii) for Lebesgue measure in [4] can be readily modified to handle the more general case. To prove (i), suppose that $w \in A_\infty(d\mu)$. Then

$$\fint_Q w\, d\mu \le C\left(\fint_Q w^{-\epsilon}\, d\mu\right)^{-1/\epsilon} \le C\exp\left(\fint_Q \log w\, d\mu\right)$$

where the first inequality is because $w \in A_q(d\mu)$ for some $1 < q$, and the second inequality is by Jensen's lemma (since $\log x^{-1/\epsilon}$ is convex).

Conversely, if (i) is satisfied, then we can apply Jensen's inequality with respect to the convex function $e^{x/2}$ to get

$$\fint_Q w\, d\mu \le C\exp\left(\fint_Q \log w\, d\mu\right) \le C\left(\fint_Q w^{1/2}\, d\mu\right)^2$$

which implies $w \in A_\infty(d\mu)$. $\square$

We shall now examine more general reverse Jensen inequalities, but first we need to introduce some notation. Let F be the class of continuous increasing functions mapping $[0,\infty)$ onto itself. If $\phi \in F$, we define

$$\|f\|_{\phi(L_Q)(d\mu)} = \inf\left\{C > 0 : \fint_Q \phi\left(\frac{|f(x)|}{C}\right) d\mu(x) \le 1\right\}$$

if it exists. If ϕ is convex, this is the usual Orlicz norm with respect to Q-normalized Lebesgue measure. In other cases, this "norm" can still be defined but it does not satisfy the triangle inequality.

If $\phi_1, \phi_2 \in F$, $\phi_2 \circ \phi_1^{-1}$ is convex, and $\phi_2(2x)/\phi_2(x) > 1+\epsilon_2$, then it follows from Jensen's inequality that, for all cubes Q, $\|f\|_{\phi_1(L_Q)(d\mu)} \le C\|f\|_{\phi_2(L_Q)(d\mu)}$, C being a constant that depends only on $\phi_2 \circ \phi_1^{-1}(1)$ and ϵ_2 (the ϵ_2 condition is unnecessary if $\phi_2 \circ \phi_1^{-1}(1) \ge 1$). We are interested in the connection between conditions involving $A_p(d\mu)$ or $RH_p(d\mu)$ and inequalities of reverse Jensen type which hold uniformly for all cubes, i.e. inequalities of the form

$$\|w\|_{\phi_2(L_Q)(d\mu)} \le C_0\|w\|_{\phi_1(L_Q)(d\mu)} \qquad \text{for all cubes } Q, \tag{3.11}$$

where w is some weight, C_0 is some constant and $\phi_2 \circ \phi_1^{-1}$ is convex (or satisfies some related condition). For example, if $\phi_2(x) = x^2$, and $\phi_1(x) = x$, then (3.11) is the defining condition for $w \in RH_2(d\mu)$.

We are mainly interested in functions which "grow like powers of x" (as opposed to exponentially, or logarithmically, or other such growth), so we will make assumptions such as $\phi_i(2x) \le C_i\phi_i(x)$ or $\phi_i(2x) \ge (1+\epsilon_i)\phi_i(x)$ whenever they are useful for our purposes.

If there is some $c > 0$ for which

$$\phi_1(x) > \phi_2(cx) \qquad \text{for all } x > 0, \tag{3.12}$$

then (3.11) is trivially true, so we confine our interest to the case where

$$\phi_1(x)/\phi_2(cx) \to 0 \quad (x \to \infty) \qquad \text{for all } c > 0. \tag{3.13}$$

This is not a very restrictive assumption because, if $\phi_2 \circ \phi_1^{-1}$ is convex and if $(1+\epsilon_2)\phi_2(x) \le \phi_2(2x) \le C_2\phi_2(x)$ (in which case (3.13) can be written simply as $\phi_1(x)/\phi_2(x) \to 0$ $(x \to \infty)$), it is easily seen that (3.13) is true whenever (3.12) is false. Interestingly, (3.13) makes superfluous the assumption that $\phi_2 \circ \phi_1^{-1}$ is convex. In fact, our first result is the following.

Proposition 3.14. *If ϕ_1 is convex, and (3.13) and (3.11) are both satisfied, then $w \in A_\infty(d\mu)$.*

Proof. Suppose that $w \notin A_\infty(d\mu)$. Let us fix $0 < \epsilon < 1/4$ and let m be so large that $\dfrac{\phi_1(x)}{\phi_2(x/C_0)} < \epsilon$ whenever $x > \phi_1^{-1}\left(\dfrac{m}{4}\right)$. Then, by lemma 3.9, there is a cube Q for which $\mu(S) > (1 - \dfrac{1}{m})\mu(Q)$, where

$$S = \{x \in Q \,:\, w(x) \le \frac{\phi_1^{-1}(1/2)}{\phi_1^{-1}(1)} \fint_Q w\,d\mu\}.$$

Letting $\widehat{w} = \dfrac{w}{\|w\|_{\phi_1(L_Q)(d\mu)}}$, we see that $\fint_Q \widehat{w}\,d\mu \le \phi_1^{-1}(1)$ by Jensen's lemma. Thus, if $x \in S$, $\phi_1(\widehat{w}(x)) \le 1/2$ and so $\int_{Q\backslash S} \phi_1 \circ \widehat{w} \ge 1/2$. Since $\mu(Q\backslash S) \le \mu(Q)/m$, it follows that

$$\fint_L \phi_1 \circ \widehat{w}\,d\mu \ge \frac{1}{4}$$

where $L = \{x \in Q : \phi_1(\widehat{w}(x)) > \frac{m}{4}\}$. From our definition of m, we get

$$\fint_Q \phi_2\left(\frac{\widehat{w}}{C_0}\right) d\mu > \frac{1}{4\epsilon} > 1.$$

This contradicts (3.11), and so $w \in A_\infty(d\mu)$, as required. $\square$

As an example of this proposition, the case $\phi_1(x) = x$, $\phi_2(x) = x \log^+ x$, $\mu =$ Lebesgue measure, is to be found in [7]. Proposition 3.14 says that weaker conditions, such as that given by $\phi_1(x) = x$, $\phi_2(x) = x \log^+ \log^+ x$ are also sufficient to guarantee $x \in A_\infty(d\mu)$.

Ideally, we would like to generalize lemma 3.7 by eliminating the hypothesis that ϕ_1 is convex from the above theorem and proving that, assuming (3.13), $\phi_2 \circ w \in A_\infty(d\mu)$ if and only if (3.11) holds. Unfortunately, this is not true. For example, let $\mu =$ Lebesgue measure, $w(x) = \min(1, x^{-1/2})$, $\phi_1(x) = x^{1/2}$ and

$$\phi_2(x) = \begin{cases} x, & x \geq 10^{-10} \\ 10^{10}x^2, & x < 10^{-10} \end{cases}.$$

Then (3.11) is true, but $\phi_2(w(x))$ behaves like $\dfrac{1}{x}$ for large x, and so it is not an A_∞ function.

Upon reflection, this counterexample reveals why we cannot prove such a result. If $\phi(x)$ is very small for $x < x_0$, the exact values of $\phi(x)$ for $x < x_0$ have very little effect on the $\phi(L_Q)$-norm of a function, whereas the $A_\infty(d\mu)$ condition is very much dependent on the relative size of the weight at different points, but independent of the average value of the weight in the interval. Also, whatever result we can get should reflect the invariance of reverse Jensen inequalities (involving a weight w) under the transformations $w \mapsto bw$ $(b > 0)$ and the invariance of the condition $\phi \circ w \in A_\infty(d\mu)$ under the transformations $\phi \mapsto r\phi$ $(r > 0)$. The next theorem is fairly close to the result we want; it has the advantage of being true, but the disadvantage of involving a whole family of reverse Hölder inequalities, and thus being a more difficult condition to verify.

Theorem 3.15. *If $\phi_1, \phi_2 \in F$, $\phi_1(x)/x \to 0$ $(x \to \infty)$, and $(1 + \epsilon_2)\phi_2(x) \leq \phi_2(2x) \leq C_2\phi_2(x)$, then the following are equivalent:*

(i) $\|w\|_{r\phi_2(L_Q)(d\mu)} \leq C_0\|w\|_{(\phi_1 \circ (r\phi_2))(L_Q)(d\mu)}$, *for all $r > 0$, and all cubes Q.*
(ii) $\phi_2(bw) \in A_\infty$ *for every $b > 0$.*

Proof. Suppose (i) is true, but, for some fixed $b > 0$, $w_2 \equiv \phi_2(bw) \notin A_\infty(d\mu)$. We may assume $C_0 \geq 1$ without loss of generality. Let us fix $\epsilon > 0$ and choose m so large that $\dfrac{\phi_1(x)}{x} < \epsilon$ whenever $x > \phi_1^{-1}\left(\dfrac{m}{4}\right)$.

Since $w_2 \notin A_\infty$, there is a cube Q for which $\mu(S) > (1 - \dfrac{1}{m})\mu(Q)$, where

$$S = \{x \in Q \ : \ w_2(x) \leq \frac{\phi_1^{-1}(1/2)}{K} \int_Q w_2 \, d\mu\}$$

and $K = C_2^{\log_2(C_0)+1}$. We choose r so that $\int_Q \phi_1(rw_2)\, d\mu = \mu(Q)$. It follows from our hypotheses that $\int_Q rw_2 \, d\mu \leq K\mu(Q)$ and so, for all $x \in S$, $\phi_1(rw_2(x)) \leq 1/2$. Arguing as in proposition 3.14, we get

$$\int_Q rw_2 \, d\mu > 1/4\epsilon,$$

which is a contradiction for $\epsilon \le 1/4K$.

Conversely, if (ii) is true, we show that (i) is true for fixed but arbitrary $r > 0$. Since (ii) is true for all $b > 0$, we can assume $\|w\|_{(\phi_1 \circ r\phi_2)(L_Q)(d\mu)} = 1$, without loss of generality . But now, by lemma 3.9, we get that

$$\mu(\{x \in Q \,:\, r\phi_2(w(x)) > \beta(r\phi_2 \circ w)_Q\}) > \alpha\mu(Q)$$

for some α and β. It follows that $\fint_Q r\phi_2(w)\,d\mu \le \frac{1}{\beta}\phi_1^{-1}(2/\alpha)$ because, if this were not so, then

$$\mu(\{x \in Q \,:\, \phi_1(r\phi_2(w(x))) > 2/\alpha\}) > \alpha\mu(Q),$$

which contradicts the assumption $\|w\|_{\phi_1 \circ r\phi_2(L_Q)(d\mu)} = 1$. It now follows from the "$1 + \epsilon_2$" rate of growth assumption on ϕ_2 that $\|w\|_{r\phi_2(L_Q)(d\mu)}$ is bounded, as required. $\square$

In the case $\phi_2(x) = x$, the parameters r and b in theorem 3.15 become superfluous, and so we get the following corollary.

Corollary 3.16. *If $\phi_1 \in F$, $\phi_1(x)/x \to 0$ $(x \to \infty)$ then $w \in A_\infty \iff \|w\|_{L^1_Q(d\mu)} \le C_0\|w\|_{\phi_1(L_Q)(d\mu)}$.*

Let us now look at a class of inequalities that generalize the definitions of $A_p(d\mu)$ and $RH_p(d\mu)$. We will replace the function $x \mapsto x^p$ by a whole class of similar functions, and associate a "norm" with each of these functions. We then define a partial ordering on these functions which has the property that if one function precedes another, its associated norm dominates the other; furthermore, for a particular weight w, there is a reverse inequality between these norms of w holding uniformly for all cubes if and only if a particular power of w is in $A_p(d\mu)$ for a particular p $(1 < p \le \infty)$.

We first define the class of functions $G_1 \subset F$. If $\phi \in F$ then $\phi \in G_1$ if there are constants $a, \epsilon > 0$ for which:

(a) $ax \le \phi(x)$,
(b) $\phi(x)/x^r \to 0$ $(x \to \infty)$ for all $r > 1$,
(c) $\phi(2x) \ge (1 + \epsilon)\phi(x)$.

For example, the functions $x \mapsto x(\log^+ x)^r$, and $x \mapsto x(\log^+ \log^+ x)^r$ are in G_1 for any $r \ge 0$. We then define

$$\begin{aligned} G_p &= \{\, x \mapsto \phi(x^p) \,:\, \phi \in G_1 \,\}, \qquad \text{for all } p \ne 0 \\ G_0 &= \{\, \log \,\} \\ G &= \bigcup_{p \in \mathbf{R}} G_p \end{aligned}$$

If $\phi \in G_p$, then ϕ has domain $[0, \infty)$ if $p > 0$, and ϕ has domain $(0, \infty)$ if $p \le 0$.

Suppose $\phi \in G_p$. If $p \neq 0$, and so $\phi(x) = \phi_1(x^p)$ for some $\phi_1 \in G_1$, we define $\|w\|_{\phi(L_Q)(d\mu)}$ to mean $\|w^p\|^{1/p}_{\phi_1(L_Q)(d\mu)}$, where the latter norm is previously defined because $\phi_1 \in F$. If $\phi \in F$, then this definition of an ϕ-norm is consistent with the previous definition. We define $\|w\|_{\log L_Q(d\mu)}$ in the obvious way, namely

$$\|w\|_{\log L_Q(d\mu)} \equiv \exp\left(\fint_Q \log w\, d\mu\right).$$

We also define a partial ordering $\prec$ on G. Let $\phi_i \in G_{p_i}$ for $i = 1, 2$. Then

(i) For $p_1, p_2 > 0$, $\quad \phi_1 \prec \phi_2$ whenever $\dfrac{\phi_1(x)}{\phi_2(x)} \to 0 \quad (x \to \infty)$.

(ii) For $p_1, p_2 < 0$, $\quad \phi_1 \prec \phi_2$ whenever $\dfrac{\phi_2(x)}{\phi_1(x)} \to 0 \quad (x \to 0)$.

(iii) For $p_1 < 0 < p_2$, $\quad \phi_1 \prec \log \prec \phi_2$.

In particular, it follows from the above and properties (a) and (b) of G_1 that if $p_1 < p_2$ then $\phi_1 \prec \phi_2$. Also, $x \mapsto x^p$ is a minimal element in G_p for all $p > 0$ and a maximal element in G_p for all $p < 0$. The next lemma shows that this partial ordering is indeed very natural for our purposes.

Lemma 3.17. *If $\phi_1, \phi_2 \in G$ and $\phi_1 \prec \phi_2$ then*

$$\|w\|_{\phi_1(L_Q)(d\mu)} \leq C\|w\|_{\phi_2(L_Q)(d\mu)}.$$

Proof. Suppose $\phi_i \in G_{p_i}$ for $i = 1, 2$. If $p_1 > 0$, then $\phi_1, \phi_2 \in F$, and so the ϕ_i-norm is as defined for $\phi \in F$. Now, since $\phi_1 \prec \phi_2$, it follows that $\phi_1(x) < C\phi_2(x)$ for all $x \geq \phi_1^{-1}(1/2)$. Letting $\widehat{w} = \dfrac{w}{\|w\|_{\phi_1(L_Q)(d\mu)}}$ and $L = \{x \in Q : \phi_1(\widehat{w}) > 1/2\}$, it follows that

$$\fint_Q \phi_2(\widehat{w})\, d\mu > \frac{1}{C}\fint_L \phi_1(\widehat{w})\, d\mu \geq \frac{1}{2C}.$$

Since property (c) of G_1 clearly extends to G_p for all $p > 0$, the desired conclusion follows easily.

If $p_2 < 0$, we can reduce to the first case by letting $\widetilde{\phi}_i(x) = \phi_i(1/x)$, because $\widetilde{\phi}_i \in G_{-p_i}$, $\widetilde{\phi}_2 \prec \widetilde{\phi}_1$, and

$$\|w\|_{\phi_i(L_Q)(d\mu)} = \|1/w\|^{-1}_{\widetilde{\phi}_i(L_Q)(d\mu)}.$$

If $p_1 = 0$ (so $\phi_1 = \log$), then we can choose p such that $0 < p < p_2$, and it follows from Jensen's inequality, and the previously handled "$0 < p_1 < p_2$" case, that

$$\|w\|_{\log(L_Q)(d\mu)} \leq \|w\|_{L^p_Q(d\mu)} \leq C\|w\|_{\phi_2(L_Q)(d\mu)}.$$

We can reduce the case $p_2 = 0$ to the case $p_1 = 0$ by taking reciprocal functions $\widetilde{\phi}_i(x)$, as before. Finally, the case $p_1 < 0 < p_2$ follows by combining the last two cases. $\square$

We are now ready to state and prove the main theorem which classifies all "reverse Jensen" inequalities involving functions in G.

Theorem 3.18. *Suppose $\phi_i \in G_{p_i}$ for $i = 1, 2$, and $\phi_1 \prec \phi_2$. Then, the inequality*

$$\|w\|_{\phi_2(L_Q)(d\mu)} \le C\|w\|_{\phi_1(L_Q)(d\mu)} \qquad \text{for all cubes } Q \tag{3.19}$$

is equivalent to

(i) *$w^{p_2} \in A_\infty(d\mu)$, if $p_1 \ge 0$ (equivalently, $w \in RH_{p_2}(d\mu)$).*
(ii) *$w^{p_2} \in A_r(d\mu)$, if $p_2 > 0 > p_1$ (where $r = \dfrac{p_1 - p_2}{p_1}$).*
(iii) *$w^{p_1} \in A_\infty(d\mu)$, if $p_2 \le 0$ (equivalently, $w^{-1} \in RH_{-p_1}(d\mu)$).*

Proof. Let us first prove (i). It is sufficient to prove it in the case $p_2 = 1$, because of the way we defined $\|w\|_{\phi(L_Q)(d\mu)}$ for $\phi \in G_p$ for $p \neq 0$. Suppose w satisfies (3.19). If $p_1 < 1$, then by property (a) of G_1 we see that

$$\|w\|_{L^1_Q} \le C\|w\|_{\phi_2(L_Q)(d\mu)} \le C\|w\|_{\phi_1(L_Q)(d\mu)}$$

and so property (b) of G_1 and corollary 3.16 together imply that $w \in A_\infty(d\mu)$, as required. If $p_1 = 1$, then we can argue as in proposition 3.14 that $w \in A_\infty$ (in proposition 3.14, we assumed ϕ_1 is convex, but we only used convexity to prove that $\displaystyle\fint_Q \widehat{w} \le C$, where $\widehat{w} = \dfrac{w}{\|w\|_{\phi_1(L_Q)(d\mu)}}$, a fact that follows easily from property (a) of G_1).

Conversely, if $w \in A_\infty(d\mu)$, then $w \in RH_p(d\mu)$ for some $p > 1$, and so by lemmas 3.9 and 3.17,

$$\|w\|_{\phi_2(L_Q)(d\mu)} \le C\|w\|_{L^p_Q} \le C\|w\|_{\log(L_Q)(d\mu)} \le C\|w\|_{\phi_1(L_Q)(d\mu)},$$

as required.

Next, we prove (ii). If (3.19) is true then, by property (a) of G_1,

$$\|w\|_{L^{p_2}_Q(d\mu)} \le C\|w\|_{\phi_2(L_Q)(d\mu)} \le C\|w\|_{\phi_1(L_Q)(d\mu)} \le C\|w\|_{L^{p_1}_Q(d\mu)}$$

and the inequality between the first and last norms implies that $w^{p_2} \in A_r(d\mu)$, where $r = \dfrac{p_1 - p_2}{p_1}$. Conversely, if $w^{p_2} \in A_r(d\mu)$, then $w^{p_2+\epsilon} \in A_r(d\mu)$, and so

$$\|w\|_{\phi_2(L_Q)(d\mu)} \le C\|w\|_{L^{p_2+\epsilon}_Q(d\mu)} \le C\|w\|_{L_Q^{p_1(1+\frac{\epsilon}{p_2})}(d\mu)} \le C\|w\|_{\phi_1(L_Q)(d\mu)}.$$

Finally, (iii) follows from (i) by taking reciprocal functions $\widetilde{\phi}_i$, as in the proof of lemma 3.17. $\square$

References

[**1**] A. Besicovitch, *A general form of the covering principle and relative differentiation of additive functions*, Fund. Math. **41** (1945), 103–110.

[**2**] L. Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Math. **116** (1966), 135–157.

[**3**] M. Christ and R. Fefferman, *A note on weighted norm inequalities for the Hardy-Littlewood maximal operator*, Proc. Amer. Math. Soc. **87** (1983), 447–448.

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

[**5**] B. Dahlberg, *On the Poisson integral for Lipschitz and C^1 domains*, Studia Math. **66** (1979), 7–24.

[**6**] B. Dahlberg, *On the absolute continuity of elliptic measures*, Amer. J. Math. **108** (1986), 1119–1138.

[**7**] R. Fefferman, *A criterion for the absolute continuity of the harmonic measure associated with an elliptic operator*, Jour. Amer. Math. Soc. **2** (1989), 127–135.

[**8**] R. Fefferman, C. Kenig, and J. Pipher, *The theory of weights and the Dirichlet problem for elliptic equations*, to appear.

[**9**] J. García-Cuerva and J. Rubio de Francia, *Weighted norm inequalities and related topics*, mathematics studies, vol. 116, North-Holland, 1985.

[**10**] F. Gehring, *The L^p-integrability of the partial derivatives of a quasi-conformal mapping*, Acta Math. **130** (1973), 265–277.

[**11**] M. de Guzman, *Differentiation of integrals in $\mathbf{R}^n$*, Lecture Notes in Mathematics 481, Springer-Verlag, 1975.

[**12**] R. Hunt, *An estimate of the conjugate function*, Stud. Math. **44** (1972), 371–377.

[**13**] R. Hunt, B. Muckenhoupt, and R. Wheeden, *Weighted norm inequalities for the conjugate function and Hilbert transform*, Trans. Amer. Math. Soc. **176** (1973), 227–251.

[**14**] B. Jawerth, *Weighted Inequalities for maximal operators: linearization, localization, and factorization*, Amer. J. Math. **108** (1986), 361–414.

[**15**] B. Muckenhoupt, *Weighted norm inequalities for the Hardy maximal function*, Trans. Amer. Math. Soc. **165** (1972), 207–226.

[**16**] E. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, 1970.

[**17**] J. Strömberg and A. Torchinsky, *Weighted Hardy Spaces*, Lecture Notes in Mathematics 1381, Springer-Verlag, 1989.

[**18**] A. Zygmund, *On a theorem of Marcinkiewicz concerning interpolation of operators*, Jour. de Math. **35** (1956), 223–248.

Department of Mathematics, St. Patrick's College, Maynooth, Co. Kildare, Ireland; e-mail: sbuckley@maths.may.ie