Transactions of the American mathematical society, 354 (2002), 2013-2033.

A_p weights for nondoubling measures in R^n and applications

Joan Orobitg § and Carlos Pérez ¶

Abstract: We study an analogue of the classical theory of $A_p(\mu)$ weights in $\mathbb{R}^n$ without assuming that the underlying measure μ is doubling. Then, we obtain weighted norm inequalities for the (centered) Hardy-Littlewood maximal function and corresponding weighted estimates for nonclassical Calderón-Zygmund operators (in the sense of [NTV1]). We also consider commutators of those Calderón-Zygmund operators with bounded mean oscillation functions (BMO), extending the main result from [CRW]. Finally, we study self–improving properties of Poincaré–B.M.O. type inequalities within this context, more precisely we show that if f is a locally integrable function satisfying $\frac{1}{\mu(Q)}\int_Q |f - f_Q|\, d\mu \leq a(Q)$ for all cubes Q, then it is possible to deduce higher L^p integrability result of f assuming certain simple geometric condition on the functional a.

1 Introduction

The classical theory of harmonic analysis for maximal functions and singular integrals on $(\mathbb{R}^n, \mu)$ has been developed under the assumption that the underlying measure μ satisfies the doubling property, i.e., there exists a constant $C > 0$ such that $\mu(B(x, 2r)) \leq C\ \mu(B(x, r))$ for every $x \in \mathbb{R}^n$ and $r > 0$. However, some recent results on Calderón-Zygmund operators ([NTV1], [NTV2], [T1], [T2]) and functions of bounded mean oscillation ([MMNO], [T3]) show that it should be possible to dispense with the doubling condition for most of the classical theory. The purpose of this paper is to present some results which strengthen this point of view.

The use of doubling measures in $\mathbb{R}^n$ has two main advantages: (a) one can work with the nested property of dyadic cubes and (b) the faces (or edges) of the cubes have measure zero. The easiness and utility of the dyadic scheme is well known. The profit of (b) is the continuity of the measure μ on cubes. That is, given cubes $R_0 \subset R_1$ one can find a monotone family of cubes $\{R_s\}$, $s \in [0, 1]$, such that $R_s \subset R_t$ if $s < t$ and the map $L(s) = \mu(R_s)$ is continuos on $[0, 1]$.

Following the previous paper [MMNO], we will renounce (a) but we shall maintain property (b). Thus, our point of departure is to consider nonnegative Radon measures μ in $\mathbb{R}^n$ without mass-points. Then a result of geometrical measure type (see [MMNO, Theorem 2]) assures that we may choose an orthonormal system in $\mathbb{R}^n$ so that any cube Q with sides parallel to the coordinates axes satisfies the above property (b)($\mu(\partial Q) = 0$). Through this work

§Research partially supported by CIRIT grant 2000 SGR00059 and by DGICYT grant BFM 2000-0361, Spain.

¶Research partially supported by DGESIC grant PB98-0106, Spain.

we shall assume (making a rotation if necessary) that the measure μ and the orthonormal system have this property. Moreover, we only shall consider cubes with sides parallel to the coordinate axes.

Our measure satisfies a Calderón-Zygmund decomposition (see [MMNO] or Section 2 later), which is one of the basic and most frequently used tools in the classical theory. This fact and the argument of its proof will allow us to recover many results without assuming that the measure μ is doubling. Related to the Calderón-Zygmund decomposition, the Hardy-Littlewood maximal operator also plays a central role. Given a locally integrable function f one defines the (centered) Hardy-Littlewood maximal function M as

$$Mf(x) = \sup_{r>0} \frac{1}{\mu(Q(x,r))} \int_{Q(x,r)} |f| \, d\mu,$$

where $Q(x,r)$ denotes the cube centered at x with sidelength equals to r. The noncentered maximal function N is defined as

$$Nf(x) = \sup_{Q \ni x} \frac{1}{\mu(Q)} \int_{Q} |f| \, d\mu,$$

where the supremum is taken over all cubes Q containing x. Clearly, $Mf(x) \leq Nf(x)$ and when the measure μ is doubling it also holds $Nf(x) \leq C\, Mf(x)$. However, if μ is nondoubling the maximal functions Mf and Nf may be very different. For instance, it is well known that the operator M acts on $L^p(\mu)$, $p > 1$, and from $L^1(\mu)$ to $L^{1,\infty}(\mu)$, whereas this is not the case in general for the operator N. On the other hand, weights for the noncentered case N has been studied and characterized by Jawerth [Ja].

The first part of this paper is devoted to develop an analogue of the classical theory of $A_p(\mu)$ weights with underlying measure μ as above. Then, we will obtain weighted norm inequalities for M and corresponding weighted estimates for Calderón-Zygmund operators. This result for singular integrals will follow as a consequence of a version of the classical estimate by Coifman proved in [C]. We will also consider commutators of Calderón-Zygmund operators with BMO extending the main result from [CRW].

In the second part of the paper we will study BMO–Poincaré type inequalities. We will obtain similar results to those obtained in [FPW] and [MP] where the underlying measure was assumed to be doubling. The main idea is as follows. Let $a : \mathcal{Q} \to [0, \infty)$ be a functional defined on the family of cubes with sides parallel to the coordinates axes. We want to show that if f is a locally integrable function satisfying

$$\frac{1}{\mu(Q)} \int_{Q} |f - f_Q| \, d\mu \leq a(Q)$$

for all cubes Q, then it is possible to deduce higher L^p integrability result of f assuming certain simple geometric condition on a (see (15) below). Of course, the case $a(Q) \equiv C$ corresponds to BMO but it is also related to Poincaré when considering

$$a(Q) = \frac{\ell(Q)}{\mu(Q)} \int_{Q} g \, d\mu.$$

The paper is organized as it follows. Section 2 is devoted to define the $A_p(\mu)$ class of weights and to study their properties. In Section 3 we shall prove the $L^p(w)$-boundedness of M when $w \in A_p(\mu)$ (Muckenhoupt's Theorem). Weights and commutators for nonclassical singular integral operators are discussed in Section 4. Self-improving properties are considered in Section 5.

Acknowledgment. We are grateful to J. Mateu and J. Verdera for their interest and discussions concerning the results of this paper. We thank E. Saksman for sending us his preprint [Sa]. The second author is very grateful for the hospitality of the Centre de Recerca Matemàtica at the Universitat Autònoma de Barcelona where this work was completed. Finally we are very grateful to the referee for the careful reading of the paper and for providing us with some enlightening examples (see Remark 2.4).

2 $A_p(\mu)$ theory of weights

The purpose of this section is to describe a A_p theory of weights adapted to our more general underlying measure μ. For this purpose we state a version of the classical Calderón-Zygmund type decomposition adapted to our situation.

A useful tool to prove the above lemma will be the following auxiliary maximal function. For a given cube Q and for each x in the interior of Q we define the basis

$$C_Q(x) = \{Q_x(r)\}$$

where $Q_x(r)$ is the unique cube with sidelength r contained in Q which minimizes the distance from x to the center of $Q_x(r)$. The radius of a cube Q is defined to be half of the sidelength. We define the corresponding maximal function

$$M_Q f(x) = \sup_{R: R \in C_Q(x)} \frac{1}{\mu(R)} \int_R |f(y)|\, d\mu(y),$$

Observe that the properties on μ imply that the function $h_x(r) := \frac{1}{\mu(Q_x(r))} \int_{Q_x(r)} |f|\, d\mu$ is continuous on $[0, \ell(Q)]$ for all x in the interior of Q.

We also denote by $\Omega_t = \{x \in Q : M_Q(g)(x) > t\}$ the level set of $M_Q(g)$.

Recall that a family of cubes $\{Q_j\}$ is quasidisjoint if there exists a universal constant C such that $\sum_j \chi_{Q_j} \leq C$, where χ_E denotes the characteristic function of the set E.

Lemma 2.1 (The Besicovitch-Calderón–Zygmund decomposition). *Let Q be a cube and let $g \in L^1(\mu)(Q)$ be a non-negative function. Also let t be a positive number such that $t > g_Q = \frac{1}{\mu(Q)} \int_Q g\, d\mu$ and such that Ω_t is not empty. Then there is a family of quasidisjoint cubes $\{Q_j\}$contained in Q satisfying*

$$\frac{1}{\mu(Q_j)} \int_{Q_j} g\, d\mu = t$$

for each j and such that

$$g(x) \leq t \qquad \textit{for } x \in Q \setminus \bigcup_j Q_j, \quad \mu\textit{–a.e.} \tag{1}$$

In fact we can write

$$\bigcup_j Q_j = \bigcup_{k=1}^{B(n)} \bigcup_{i \in \mathcal{F}_k} Q_i, \tag{2}$$

where each of the family $\{Q_i\}_{i\in\mathcal{F}_k}$, $k = 1, \cdots, B(n)$, is formed by pairwise disjoint cubes. $B(n) > 1$ is usually called the Besicovitch constant.

Proof: Since Ω_t is not empty, for any $x \in \Omega_t$ there is a cube $P_x \in C_Q(x)$ such that $(\mu(P_x))^{-1} \int_{P_x} g\, d\mu > t$. Therefore, since h_x is continuous, we have a cube $Q_x \in C_Q(x)$ satisfying

$$\frac{1}{\mu(Q_x)} \int_{Q_x} g\, d\mu = t$$

with $Q_x \subsetneq Q$. Now, observe that we can not apply directly the Besicovitch Covering Theorem since x may not be the center of Q_x. To overcome this obstacle we proceed as in [MMNO]. For any cube Q_x we define the rectangle R_x in $\mathbb{R}^n$ as the unique rectangle in $\mathbb{R}^n$ centered at x such that $R_x \cap Q = Q_x$. Clearly, the ratio of any two sidelengths of R_x is bounded by 2. So, by the Besicovitch Covering Theorem we have a countable collection of rectangles R_j such that they cover Ω_t, and every point of $\mathbb{R}^n$ belongs to at most $B(n)$ rectangles R_i. Replacing each R_j by its corresponding cube Q_j we get the cubes from the Lemma. Finally, (1) follows by the Lebesgue differentiation theorem since $x \in Q \setminus \Omega_t$ implies $g(x) \leq t$, μ-a.e.

□

A nonnegative, locally integrable function is called a weight. We will consider weights which satisfy the following conditions.

Definition 2.2 *Let $1 < p < \infty$ and $p' = p/(p-1)$. We say that a weight w satisfies the $A_p(\mu)$ condition if there exists a constant K such that for all cubes Q*

$$\left(\frac{1}{\mu(Q)} \int_Q w\, d\mu\right) \left(\frac{1}{\mu(Q)} \int_Q w^{1-p'}\, d\mu\right)^{p-1} \leq K. \tag{3}$$

We also say that a weight w satisfies the $A_1^s(\mu) = A_1(\mu)$ condition if there exists a constant K such that for all cubes Q,

$$\frac{1}{\mu(Q)} \int_Q w\, d\mu \leq K \, ess \inf_{x\in Q} w(x).$$

Finally, we define the $A_\infty(\mu)$ class as $A_\infty(\mu) = \bigcup_{p>1} A_p(\mu)$

Observe that trivially, $A_1(\mu) \subset A_p(\mu)$ for all $p > 1$ and $A_p(\mu) \subset A_q(\mu)$ if $p < q$.

The reason we use the notation $A_1^s(\mu) = A_1(\mu)$ (s from strong) is to distinguish this class from the class $A_1^w(\mu)$ (w from weak) of weights w such that for some constant C

$$Mw(x) \leq C\, w(x)$$

almost everywhere in x. Observe that $A_1^s(\mu) \subset A_1^w(\mu)$ and indeed, in the classical situation, i.e. if the underlying measure is doubling, both conditions are equivalent and hence $A_1^s(\mu) = A_1^w(\mu)$. However, and this is a big gap between the two theories, we will show in Section 3 that, in general, we have that $A_1^s(\mu) \subsetneq A_1^w(\mu)$. In fact the class $A_1^w(\mu)$ is too large since it is also shown there that it is not a subset of $A_\infty^w(\mu)$ in general.

We start by proving some of the classical results that hold in our more general situation. However, as we shall see in next section not all of them will be true.

We will use the standard notation $w(E) = \int_E w\, d\mu$, for any measurable set E.

Lemma 2.3 *For a weight w the following conditions are equivalent:*
a) $w \in A_\infty(\mu)$.
b) For every cube Q

$$\frac{1}{\mu(Q)} \int_Q w\, d\mu \approx \exp\left(\frac{1}{\mu(Q)} \int_Q \log w\, d\mu \right).$$

c) There are constants $0 < \alpha, \beta < 1$ such that for every cube Q

$$\mu(\{x \in Q : w(x) \leq \beta\, w_Q\}) \leq \alpha\, \mu(Q). \tag{4}$$

d) There are positive constants C and β such that for every cube Q and for every $\lambda > w_Q$

$$w(\{x \in Q : w(x) > \lambda\}) \leq C\lambda\, \mu(\{x \in Q : w(x) > \beta\, \lambda\}).$$

e) w satisfies a reverse Hölder inequality, namely there are positive constants c and δ such that for every cube Q

$$\left(\frac{1}{\mu(Q)} \int_Q w^{1+\delta}\, d\mu \right)^{\frac{1}{1+\delta}} \leq \frac{c}{\mu(Q)} \int_Q w\, d\mu.$$

f) There are positive constants c and ρ such that for any cube Q and any measurable set E contained in Q then

$$\frac{w(E)}{w(Q)} \leq c \left(\frac{\mu(E)}{\mu(Q)} \right)^\rho. \tag{5}$$

g) w satisfies the following condition: there are positive constants $\alpha, \beta < 1$ such that whenever E is a measurable set of a cube Q

$$\frac{\mu(E)}{\mu(Q)} < \alpha \quad \textit{implies} \quad \frac{w(E)}{w(Q)} < \beta. \tag{6}$$

Remark 2.4 *If one removes our standing assumption on the measure μ (that is, $\mu(\partial Q) = 0$ for any cube Q with sides parallel to the coordinates axes) then the Lemma 2.3 may be false, as the following examples show (similar one dimensional examples can be found in the recent paper [GMOPST]).*

The referee provided us with these examples. We thank his/her anonymous and lucid contribution.

First, consider on the line the Radon measure $\mu = \sum_{k\geq 1} 2^{-k^2}\delta_{u_k}$, where u_k is a decreasing sequence of positive numbers such that $u_k \downarrow 0$ and δ_x is the point mass at x. Also let w be the weight which takes on the value 2^{k^2-k} at u_k for each k. Then one can easily check that $w \in A_1(\mu)$ and hence for all $A_p(\mu)$, $p > 1$. Moreover, the maximal operator M acts from $L^1(w)$ to $L^{1,\infty}(w)$ (because $w \in A_1(\mu)$) and so M acts on $L^p(w)$, $p > 1$. However, it is clear that $w \notin L^{1+\varepsilon}_{loc}(\mu)$ for any $\varepsilon > 0$ and so w doesn't satisfy any reverse Hölder inequality.

The second example is a refinement of the previous one. We consider a Radon measure ν on the line that is not atom-free but it is fairly nice: it gives positive finite measure to all bounded intervals. Specifically, let ν be the sum of the Lebesgue measure on $\mathbb{R}$ and $\sum_{k\geq 1} 2^{-k^2}(\delta_{u_k} + \delta_{-u_k})$ where now $u_k = 2^{-k^3}$. Define $w(\pm u_k) = 2^{k^2-k}$ for each k and w equals 1 everywhere else. Then $w \in A_p(\nu)$ for all $p > 1$ ($w \notin A_1(\nu)$), as can be checked case-by-case (the point masses make only a bounded difference to the quantities on the left hand side of (3) except for small intervals Q close to the origin; in the exceptional case, the one or two point masses in the interval that are furthest from the origin dominate). Again, it is clear that $w \notin L^{1+\varepsilon}_{loc}(\mu)$ for any $\varepsilon > 0$.

As a third example, we define μ to be the measure on $\mathbb{R}^2$ which is the product of ν and Lebesgue measure on $\mathbb{R}$. Picking cubes oriented in coordinate directions, μ has the same behaviour as ν. The extra subtely is that μ is a non-atomic Radon measure to which the results in this paper apply. But it doesn't contradict our results because now we have chosen the one orientation for cubes that violates condition (b) on the Introduction.

Proof of Lemma 2.3: We will write the complete proof of this lemma even though several of the implications are trivial. In the literature (the Lebesgue measure case) there are different ways to prove the main implication *a) ⇒ e)* , but only the one that we present can be adapted to our setting. Here we combine the methods from [CF] and [GCRdF].

a) ⇒ b)

By Jensen's inequality it is enough to show that

$$\frac{1}{\mu(Q)}\int_Q w\, d\mu \leq C \exp\left(\frac{1}{\mu(Q)}\int_Q \log w\, d\mu\right)$$

Since the A_p classes are increasing on p, if $w \in A_\infty(\mu)$ there exists some $p_0 > 1$ such that $w \in A_p$ for $p \geq p_0$. Then, there exists a constant K such that for $p \geq p_0$

$$\left(\frac{1}{\mu(Q)}\int_Q w\, d\mu\right)\left(\frac{1}{\mu(Q)}\int_Q w^{1-p'}\, d\mu\right)^{p-1} \leq K.$$

Letting p tend to ∞ we obtain b).

b) ⇒ c)

Dividing w by an appropiate constant (to be precise $\exp\left(\frac{1}{\mu(Q)}\int_Q \log w\, d\mu\right)$) we may assume that $\int_Q \log w\, d\mu = 0$ and, consequently $w_Q \le C$.

$$\mu(\{x \in Q : w(x) \le \beta\, w_Q\}) \le \mu(\{x \in Q : w(x) \le \beta\, C\})$$
$$= \mu\left(\{x \in Q : \log(1 + \frac{1}{w(x)}) \ge \log(1 + \frac{1}{\beta\, C})\}\right)$$
$$\le \frac{1}{\log(1+\frac{1}{\beta\, C})} \int_Q \log(1+\frac{1}{w})\, d\mu = \frac{1}{\log(1+\frac{1}{\beta\, C})} \int_Q \log(1+w)\, d\mu$$

since $\int_Q \log w\, d\mu = 0$. Now, since $\log(1+t) \le t$, $t \ge 0$, we get:

$$\mu(\{x \in Q : w(x) \le \beta\, w_Q\}) \le \frac{1}{\log(1+\frac{1}{\beta\, C})} \int_Q w\, d\mu \le \frac{C\,\mu(Q)}{\log(1+\frac{1}{\beta\, C})} \le \frac{1}{2}\,\mu(Q),$$

if we choose β small enough.

c) ⇒ d)

Since we assume that $\lambda > w_Q$ we may consider the Besicovitch-Calderón–Zygmund decomposition $\{Q_j\}$ of w and we find a family of quasidisjoint cubes satisfying

$$\lambda < \frac{1}{\mu(Q_j)} \int_{Q_j} w\, d\mu \le 2\,\lambda$$

for each j. By the properties of the cubes combined with (4) we have

$$w(\{x \in Q : w(x) > \lambda\}) \le \sum_{k=1}^{B(n)} \sum_{i \in \mathcal{F}_k} w(Q_i)$$

$$\le 2\lambda \sum_{k=1}^{B(n)} \sum_{i \in \mathcal{F}_k} \mu(Q_i) \le \frac{2\,\lambda}{1-\alpha} \sum_{k=1}^{B(n)} \mu(\{x \in Q_i : w(x) > \beta\, w_{Q_i}\})$$
$$\le \frac{2\,\lambda\, B(n)}{1-\alpha}\,\mu(\{x \in Q : w(x) > \beta\, w_Q\})$$

since $w_{Q_i} > \lambda > w_Q$.

d) ⇒ e)

We will be using the formula

$$\int_X f(x)^p\, d\nu = p \int_0^\infty \lambda^p \nu(\{x \in X : f(x) > \lambda\})\, \frac{d\lambda}{\lambda}$$

which holds for every nonnegative measurable function f and in any arbitrary measure space (X,ν) with nonnegative measure ν. Then for arbitrary positive δ we have

$$\frac{1}{\mu(Q)}\int_Q w^{1+\delta}\,d\mu=\frac{\delta}{\mu(Q)}\int_0^\infty \lambda^\delta w(\{x\in Q: w(x)>\lambda\})\,\frac{d\lambda}{\lambda}$$

$$=\frac{\delta}{\mu(Q)}\int_0^{w_Q} \lambda^\delta w(\{x\in Q: w(x)>\lambda\})\,\frac{d\lambda}{\lambda}+\frac{\delta}{\mu(Q)}\int_{w_Q}^\infty \lambda^\delta w(\{x\in Q: w(x)>\lambda\})\,\frac{d\lambda}{\lambda}$$

$$\leq (w_Q)^{1+\delta}+\frac{\delta}{\mu(Q)}\int_{w_Q}^\infty \lambda^\delta w(\{x\in Q: w(x)>\lambda\})\,\frac{d\lambda}{\lambda}$$

$$\leq (w_Q)^{1+\delta}+\frac{C\,\delta}{\mu(Q)}\int_{w_Q}^\infty \lambda^{\delta+1} \mu(\{x\in Q: w(x)>\beta\,\lambda\})\,\frac{d\lambda}{\lambda}$$

$$\leq (w_Q)^{1+\delta}+\frac{C\,\delta}{\beta^{1+\delta}}\frac{1}{\mu(Q)}\int_{w_Q\,\beta}^\infty \lambda^{\delta+1} \mu(\{x\in Q: w(x)>\lambda\})\,\frac{d\lambda}{\lambda}$$

$$\leq (w_Q)^{1+\delta}+\frac{C\,\delta}{\beta^{1+\delta}}\frac{1}{\mu(Q)}\int_Q w^{1+\delta}\,d\mu.$$

If we choose δ small enough such that $\frac{C\,\delta}{\beta^{1+\delta}}<1$ the last term can be absorbed by the first term of the string of inequalities.

e) $\Rightarrow$ *f)*

This is just Hölder's inequality with $r=1+\delta$. Indeed if $E\subset Q$

$$\frac{w(E)}{\mu(Q)}=\frac{1}{\mu(Q)}\int_Q \chi_E w\,d\mu\leq\left(\frac{1}{\mu(Q)}\int_Q w^r\,d\mu\right)^{1/r}\left(\frac{\mu(E)}{\mu(Q)}\right)^{1/r'}\leq$$

$$\frac{C}{\mu(Q)}\int_Q w\,d\mu\left(\frac{\mu(E)}{\mu(Q)}\right)^{1/r'}$$

and this implies the A_∞ condition (5) with $\rho=1/r'$.

f) $\Rightarrow$ *g)*

This is immediate.

g) $\Rightarrow$ *c)*

First observe that condition (6) is equivalent to saying that there are positive constants $\alpha',\beta'<1$ such that whenever E is a measurable set of a cube Q

$$\frac{w(E)}{w(Q)}<\alpha' \quad \text{implies} \quad \frac{\mu(E)}{\mu(Q)}<\beta'. \tag{7}$$

Then let $E=\{x\in Q: w(x)>bw_Q\}$ where $b\in(0,1)$ is going to be chosen now and let $E'=Q\setminus E=\{x\in Q: w(x)\leq bw_Q\}$. Then $w(E')\leq b\,w_Q\mu(E')\leq b\,w(Q)$. Then if we take $b=\beta'$ we have that $\mu(E')\leq\alpha'\,\mu(Q)$. This yields (4).

Therefore we have shown that c) $\Leftrightarrow$ d) $\Leftrightarrow$ e) $\Leftrightarrow$ f) $\Leftrightarrow$ g) *c)* $\Rightarrow$ *a)*

We use again that condition e) is symmetric, namely that condition (7) holds. We also use that the measure $w\,d\mu$ does not see hyperplanes parallel to the axes since the same is true for $d\mu$. Now, if we write $d\mu = w^{-1}wd\mu$ and since c) $\Leftrightarrow$ e) we have that there are positive constants c and δ such that

$$\left(\frac{1}{w(Q)}\int_Q (w^{-1})^{1+\delta}\, wd\mu\right)^{\frac{1}{1+\delta}} \leq \frac{c}{w(Q)}\int_Q w^{-1}\, wd\mu.$$

Hence

$$\frac{w(Q)}{\mu(Q)}\left(\frac{1}{\mu(Q)}\int_Q w^{-\delta}\, d\mu\right)^{1/\delta} \leq C.$$

Then, if we let $\delta = \frac{1}{p-1}$, that is, $p = \frac{1}{\delta} + 1 > 1$ we have that $w \in A_p$.

The proof of the Lemma is now complete.

$\square$

Now $w \in A_p(\mu)$, $p > 1$, obviously implies that $w^{1-p'} \in A_{p'}(\mu)$. Consequently, it is easy to deduce the following corollary.

Corollary 2.5 *Let $p > 1$ and let $w \in A_p(\mu)$, then:*
(i) There is $\varepsilon > 0$ such that $w \in A_{p-\varepsilon}(\mu)$ and therefore

$$A_p(\mu) = \bigcup_{q<p} A_q(\mu)$$

(ii) There is $\eta > 0$ such that $w^{1+\eta} \in A_p(\mu)$.

It is well known that there is an intimate relationship between the A_p weights and the John-Nirenberg space $BMO(\mu)$ of locally integrable functions with bounded mean oscillation, namely if

$$\sup_Q \frac{1}{\mu(Q)}\int_Q |f - f_Q|\, d\mu < \infty,$$

where the supremum is taken over all cubes Q with sides parallel to the cordinates axes. As a consequence of the John-Nirenberg property for $BMO(\mu)$ ([MMNO]) and the above lemma, we have the following relationship between weights and BMO.

Corollary 2.6
(i) If $w \in A_\infty(\mu)$ then $\log(w) \in BMO(\mu)$.
(ii) Fix $p > 1$ and let $b \in BMO(\mu)$. Then there exists $\varepsilon > 0$ depending upon the $BMO(\mu)$ constant of b such that $e^{xb} \in A_p(\mu)$ for $|x| < \varepsilon$.

We will skip the proof because the classical one (e.g. [GCRdF, chapter IV]) also works in our setting.

3 Muckenhoupt's Theorem

The purpose of this section is to state show the following result similar to the classical theorem of Muckenhoupt.

THEOREM 3.1 *Let $1 < p < \infty$ and suppose that $w \in A_p(\mu)$. Then, there exists a constant C such that for all functions f*

$$\int_{\mathbb{R}^n} Mf(x)^p\, w(x) d\mu(x) \leq C \int_{\mathbb{R}^n} |f(x)|^p\, w(x) d\mu(x).$$

Further, suppose that $w \in A_1^s(\mu)$, Then, there exists a constant C such that for all functions f

$$w(\{x \in \mathbb{R}^n : Mf(x) > \lambda\}) \leq \frac{C}{\lambda} \int_{\mathbb{R}^n} |f(x)|\, w(x) d\mu(x).$$

Recall that this theorem is well known within the classical situation of $\mathbb{R}^n$, when the underlying measure is the Lebesgue measure, or more generally when the underlying measure is doubling. Again we will omit its proof because it follows from the standard arguments. The A_p condition (3) and Lemma 2.1 give the weak type (p,p) boundedness of M. Then Corollary 2.5 and the Marcinkiewicz interpolation theorem complete the proof. On the other hand, Jawerth [Ja] and Christ and Feffermann [ChF] gave a shorter proof of Muckenhoupt result without using the reverse Hölder inequality; this approach doesn't work in our context.

In contrast with Theorem 3.1 we have the following negative results showing that the class A_1^w is not the right class for the centered maximal function.

Remark 3.2 *Let $1 < p < \infty$. The following inequality is* **false** *in general:*

$$\int_{\mathbb{R}^n} Mf(x)^p\, w(x) d\mu(x) \leq C \int_{\mathbb{R}^n} |f(x)|^p\, Mw(x) d\mu(x). \tag{8}$$

As a consequence, the following inequality is also **false** *in general:*

$$w(\{x \in \mathbb{R}^n : Mf(x) > \lambda\}) \leq \frac{C}{\lambda} \int_{\mathbb{R}^n} |f(x)|\, Mw(x) d\mu(x).$$

Example (suggested to us by F. Soria and A. Vargas):

Take μ on $\mathbb{R}^n$ defined as $d\mu = \exp(-\sum_{i=1}^n x_i)dx$, where dx denotes the Lebesgue measure on $\mathbb{R}^n$, and let $w(x) = \exp(\sum_{i=1}^n x_i)$. Thus, $w\, d\mu = dx$. Again, $Q(x,r)$ denotes the cube centered at x with sidelength equals to r. Therefore by trivial computations we get that

$$\frac{w(Q(x,r))}{\mu(Q(x,r))} = r^n \left(\prod_{i=1}^n e^{-x_i}(e^{r/2} - e^{-r/2})\right)^{-1} = w(x) \left(\frac{r}{e^{r/2} - e^{-r/2}}\right)^n$$

and

$$Mw(x) = \sup_{r>0} \frac{w(Q(x,r))}{\mu(Q(x,r))} = a_n w(x).$$

Observe that this means that $w \in A_1^w(\mu)$. In this example the inequality (8) is equals to

$$\int_{\mathbb{R}^n} (Mf)^p \, dx \leq C \int_{\mathbb{R}^n} |f(x)|^p \, dx. \tag{9}$$

Let f be the characteristic function of Q_0, the cube of sidelength 1 and centered at $(1/2, \dots, 1/2)$. Clearly, $\int |f(x)|^p \, dx = 1$. For each positive integer j we define $Q_j = (j, \dots, j) + Q_0$ and let P_j be the cube centered at $(j+1, \dots, j+1)$ with sidelength $2(j+2)$. Simply geometry and easy computations give that if $x \in Q_j$

$$Mf(x) \geq \frac{\mu(Q_0)}{\mu(P_j)} \geq \frac{(1-e^{-1})^n}{e^n}.$$

Consequently

$$\int (Mf)^p \, dx = \infty$$

and (9) doesn't hold.

Note that this example shows that

$$w \in A_1^w(\mu) \setminus A_\infty(\mu).$$

Remark 3.3 *As a consequence of the above example, the following Wiener type inequality is* **false** *in general: There exists a constant C such that for any cube Q (including $\mathbb{R}^n$ as degenerate case with $\mu(R^n) = \infty$) and for $\lambda > |f|_Q$*

$$\frac{C}{\lambda} \int_{\{x \in Q: |f(x)| > \lambda\}} |f(x)| \, d\mu \leq \mu(\{x \in Q : Mf(x) > \lambda\}). \tag{10}$$

Proof: Assume that (10) is true and take $f = w \in A_1^w(\mu)$, $Mw(x) \leq A\, w(x)$. Then for $\lambda > w_Q$

$$\frac{1}{\lambda} \int_{\{x \in Q: \, w(x) > \lambda\}} w(x) \, d\mu \leq C\mu(\{x \in Q : w(x) > A^{-1}\lambda\}).$$

This is precisely condition d) in Lemma 2.3. So that w would belong to $A_\infty(\mu)$, but as we remarked in the above example the class $A_1^w(\mu)$ is not always contained in $A_\infty(\mu)$.

$\square$

We finish this Section by noting some gaps of our approach. The first one is that we don't know if the $L^p(w)$-boundedness of M implies that $w \in A_p(\mu)$. We only can obtain the weaker condition

$$\left(\frac{1}{\mu(3Q)} \int_Q w \, d\mu \right) \left(\frac{1}{\mu(3Q)} \int_Q w^{1-p'} \, d\mu \right)^{p-1} \leq K.$$

Let $f \in L^1_{loc}(\mu)$ and take $s \in (0,1)$. It is well known that if the measure μ is doubling then $(Mf)^s$ belongs to the class A_1 ([CR].) In our setting we don't know if this result is true.

If we have two $A_1^s(\mu)$ weights w_1 and w_2 and if $1 < p < \infty$, then it is easy to check that $w_1 w_2^{1-p}$ belongs to the class $A_p(\mu)$. However, we don't know if the converse is true. Of course, if the factorization theorem were true then we would get the distance in $BMO(\mu)$ to $L^\infty(\mu)$, that is, for $f \in BMO(\mu)$ we would have

$$\inf \{\|f-g\|_* : \quad g \in L^\infty(\mu)\} \simeq \left(\sup\{\lambda > 0 : \quad e^{\lambda f} \in A_2(\mu)\}\right)^{-1}.$$

When the measure μ is the Lebesgue measure this result is known as Garnett-Jones formula (see [GCRdF, p. 445]).

4 Weights and singular integral operators

This section is devoted to deduce some weighted inequalities for non-classical Calderón-Zygmund integral operators. The theory of Calderón-Zygmund operators on nonhomogeneous spaces has been developed by Nazarov, Treil and Volberg (see [NTV1] and [NTV2]). We mention that Tolsa has also constructed a satisfactory theory for the particular case of the Cauchy integral operator ([T1], [T2]).

Fix $d > 0$ (not necessarily integer). Throughout this section, μ will denote a non-negative "d-dimensional" Borel measure, i.e., a measure satisfying

$$\mu(B(x,r)) \leq r^d \qquad \text{for all} \quad x \in \mathbb{R}^n, r > 0.$$

Given a kernel K on $\mathbb{R}^n \times \mathbb{R}^n$ —i.e. a locally integrable, complex-valued function defined off the diagonal— we say that it satisfies the standard "d-dimensional" estimates if there exist $\delta \in (0,1]$ and $A > 0$ such that

1. $|K(x,y)| \leq A|x-y|^{-d}$
2. $|K(x,y) - K(z,y)| \leq A\dfrac{|x-z|^\delta}{|x-y|^{d+\delta}}$
3. $|K(y,x) - K(y,z)| \leq A\dfrac{|x-z|^\delta}{|z-y|^{d+\delta}}$
   whenever $x, y, z \in \mathbb{R}^n$ and $|x-z| \leq \frac{1}{2}|x-y|$.

A bounded linear operator T on $L^2(\mu)$ is called a Calderón-Zygmund integral operator with Calderón-Zygmund kernel K if for every $f \in L^2(\mu)$,

$$Tf(x) = \int_{\mathbb{R}^n} K(x,y) f(y)\, d\mu(y)$$

for μ-almost every $x \in \mathbb{R}^n \setminus \operatorname{supp} f$.

A way to define the L^2-boundedness is as follows. Consider the family of the truncated operators T_ε,

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

We say that T is bounded in $L^2(\mu)$ if all T_ε are uniformly (in ε) bounded in $L^2(\mu)$. In [NTV1] it is proved that this holds if and only if the T_ε (and its adjoint) are uniformly bounded on characteristic functions of squares. Moreover, the L^2-boundedness is equivalent to the L^p-boundedness, for any $p \in (1,\infty)$. We refer the reader to the cited works of Nazarov, Treil and Volberg to get an exhaustive information on Calderón-Zygmund operators on non-homogeneous spaces.

Our starting point is that we have a Calderón-Zygmund integral operator T with a kernel K as before. Therefore, the operator T is bounded on $L^p(\mu)$ and we want to conclude that T is bounded on $L^p(w)$ if $w \in A_p(\mu)$. Precisely,

$$\int_{\mathbb{R}^n} |T_\varepsilon f(x)|^p w(x) d\mu(x) \le C \int_{\mathbb{R}^n} |f(x)|^p w(x) d\mu(x) \tag{11}$$

where C is a constant independent of ε and f. However, not all weights satisfying (11) belong to the class $A_p(\mu)$. There is an example of this fact in [Sa]. In the example of E. Saksman the operator T is the Hilbert transform and the measure μ is the Lebesgue measure restricted to a particular open set of $\mathbb{R}$.

We have all ingredients to prove (11), in fact we have the well–known method (e.g. [St, p. 205–209]), the classical "good λ inequality" (as used in [V]) and weights for the centered maximal function (our contribution).

Without loss of generality we assume that $\mu(\partial Q) = 0$ for any cube Q, (we need this property to apply our weight's properties). For technical reasons we redefine the truncated operators T_ε,

$$T_\varepsilon f(x) = \int_{y \notin Q(x,\varepsilon)} K(x,y) f(y) d\mu(y)\,,$$

where $Q(x,\varepsilon)$ is the cube centered at x and sidelength ε. Observe that the difference between the truncated operator using a ball of radius ε and the trucated operator using a cube of sidelength ε is pointwise bounded by the centered maximal function. Therefore, they have the same behavior with respect to the L^p-boundedness.

For each $\varepsilon > 0$ and $f \in L^p(\mu)$ we define the maximal operator

$$T^*_\varepsilon f(x) := \sup_{\delta > \varepsilon} |T_\delta f(x)|$$

(It is known that $T^*_\varepsilon f \in L^p(\mu)$ and the operator T^*_ε is weak type (1,1) [NTV2], [T2]).

Then we will prove that for any $w \in A_\infty$ and for appropiate constants a, β and γ we have that

$$w\left(\{x : T^*_\varepsilon f(x) > (1+\beta)t, \quad Mf(x) \le \gamma t\}\right) \le a w\left(\{x : T^*_\varepsilon f(x) > t\}\right) \tag{12}$$

holds for all $t > 0$. Therefore, if $a^p < (1+\beta)^{-1}$ this relative distributional inequality easily gives:

THEOREM 4.1 *Let $0 < p < \infty$ and suppose that $w \in A_\infty(\mu)$ then the inequality*

$$\int_{\mathbb{R}^n} |T_\varepsilon^* f(x)|^p w(x) d\mu(x) \leq C \int_{\mathbb{R}^n} Mf(x)^p w(x) d\mu(x)\,,$$

holds for every f for which the left hand side is finite.

Clearly, (11) is a corollary of this theorem.

Since the statement of inequality (12) is somewhat simpler when $w \equiv 1$ (and since its proof easily implies the general case), we first consider that special situation. The set $\Omega = \{x \in \mathbb{R}^n : T_\varepsilon^* f(x) > t\}$ is open (by definition of T_ε^* and because $\mu(\partial Q) = 0$ for any cube Q). Therefore we can decompose it as a disjoint union $\Omega = \cup Q_j$ of Whitney cubes: they are mutually disjoint and $2\,\mathrm{diam}\,(Q_j) \leq \mathrm{dist}\,(Q_j, \Omega^c) \leq 8\,\mathrm{diam}\,(Q_j)$.

Moreover, the family $4Q_j$ is almost disjoint with constant 4^n and obviously $4Q_j \subset \Omega$. We are going to show that, given $\beta > 0$ and $0 < \alpha < 1$ there exists $\gamma = \gamma(\beta, \alpha, n)$ such that for all j

$$\mu\left(\{x \in Q_j : T_\varepsilon^* f(x) > (1+\beta)t \quad \text{and} \quad Mf(x) \leq \gamma t\}\right) \leq \alpha \mu(4Q_j). \tag{13}$$

Then summing over j,

$$\mu\left(\{x : T_\varepsilon^* f(x) > (1+\beta)t \quad \text{and} \quad Mf(x) \leq \gamma t\}\right) \leq \alpha 4^n \mu(\Omega).$$

Choosing α so that $\alpha 4^n < 1$ we get (12) for the special case $w \equiv 1$. For general w, recall that if w belongs to $A_\infty(\mu)$ there are positive constants c and ρ such that for all cubes Q and all subsets $E \subset Q$,

$$\frac{w(E)}{w(Q)} \leq c \left(\frac{\mu(E)}{\mu(Q)}\right)^\rho .$$

Looking back at (13) we obtain

$$w\left(\{x \in Q_j : T_\varepsilon^* f(x) > (1+\beta)t \quad \text{and} \quad Mf(x) \leq \gamma t\}\right) \leq c\alpha^\rho w(4Q_j)\,.$$

Summing again over j,

$$w\left(\{x : T_\varepsilon^* f(x) > (1+\beta)t \quad \text{and} \quad Mf(x) \leq \gamma t\}\right) \leq c\alpha^\rho 4^n w(\Omega)\,.$$

Choosing α so that $c\alpha^\rho 4^n < (1+\beta)^{-1}$ we would finally get (12).

We have only to prove (13). Fix j and set $Q = Q_j$ and $r = l(Q)$. Assume that there exists $b \in Q$ so that $Mf(b) \leq \gamma t$ (if not the set appearing in (13) would be empty). Let z be a point in Ω^c (that is, $T_\varepsilon^* f(z) \leq t$) such that $\mathrm{dist}\,(z, Q) = \mathrm{dist}\,(Q, \Omega^c)$. Now turn our attention to some simple geometric facts about the cube Q and observe that

$$Q \subset P \equiv Q(b, \frac{5}{2}r) \subset 4Q \subset B \equiv Q(z, 18r)\,.$$

Set $f_1 = f\chi_B$ and $f_2 = f - f_1$. Then, for $x \in Q$ and $\delta > \varepsilon$,

$$\begin{aligned}
|T_\delta f_1(x)| &\leq |T_\delta(f\chi_P)(x)| + \frac{C}{r^d}\int_B |f(y)|\, d\mu(y) \\
&\leq T_\varepsilon^*(f\chi_P)(x) + CMf(b) \\
&\leq T_\varepsilon^*(f\chi_P)(x) + C\gamma t
\end{aligned}$$

and so

$$|T_\delta f(x)| \leq |T_\delta f_2(x)| + T_\varepsilon^*(f\chi_P)(x) + C\gamma t.$$

To compare $T_\delta f_2(x)$ with $T_\delta f_2(z)$ we use the standard arguments (see [St, p. 208]). We get

$$|T_\delta f_2(x) - T_\delta f_2(z)| \leq CMf(b)$$

and

$$|T_\delta f_2(z)| \leq T_\varepsilon^* f(z) \leq t.$$

Therefore

$$T_\varepsilon^* f(x) \leq T_\varepsilon^*(f\chi_P)(x) + (1 + C\gamma)t\,, \quad x \in Q\,.$$

Now choose γ so that $2C\gamma \leq \beta$ and consequently

$$\{x \in Q : T_\varepsilon^* f(x) > (1+\beta)t \quad \text{and} \quad Mf(x) \leq \gamma t\} \subset \{x \in Q : T_\varepsilon^*(f\chi_P)(x) > \frac{\beta}{2}t\}$$

Finally, using that T_ε^* is weak type (1,1) we have

$$\begin{aligned}
\mu\left(\{x \in Q : T_\varepsilon^*(f\chi_P)(x) > \frac{\beta}{2}t\}\right) &\leq \frac{C}{\beta t}\int_P |f(y)| d\mu(y) = \\
&= \frac{C}{\beta t}\frac{\mu(P)}{\mu(P)}\int_P |f(y)| d\mu(y) \leq \frac{C}{\beta t}\mu(P)Mf(b) \leq \\
&\leq \frac{C}{\beta}\gamma\mu(4Q) \leq \alpha\mu(4Q),
\end{aligned}$$

always provided that γ is chosen small enough so that $C\beta^{-1}\gamma \leq \alpha$.

As an application of our result we shall prove that the commutator $[b, T_\varepsilon]$ defined as

$$[b, T_\varepsilon]f = b \cdot T_\varepsilon f - T_\varepsilon(b \cdot f)$$

is a bounded operator in $L^p(\mu)$ when $b \in BMO(\mu)$. This is an extension of the classical result of Coifman, Rochberg and Weiss [CRW].

THEOREM 4.2 *Let $b \in BMO(\mu)$, $w \in A_p(\mu)$, $p > 1$, and let T be a Calderón-Zygmund operator. Then*

$$||[b, T_\varepsilon]f||_{L^p(w)} \leq C||b||_*||f||_{L^p(w)}.$$

Our proof is very close to the less known proof of the theorem in [CRW], see for instance [GCRdF, p. 473]. Tolsa [T3] has proved that the same result holds when $w \equiv 1$ if one replaces $b \in BMO(\mu)$ by $b \in RBMO(\mu)$, another space of functions of bounded mean oscillation more adapted to work with singular integrals and sharp maximal functions. His proof doesn't use weights, but it is based on the use of a sharp maximal operator.

Proof: To simplify notation we will write T instead of T_ε. By Corollary 2.5 ii) there is $\eta > 0$ such that $w^{1+\eta} \in A_p(\mu)$. Then using Corollary 2.6 (ii) we choose $\delta > 0$ such that $\exp(s\,p\,b(1+\eta)/\eta) \in A_p(\mu)$ if $0 \le s(1+\eta)/\eta < \delta$ with uniform constant. For $z \in \mathbb{C}$ we define the operator

$$S_z f = e^{zb} T(e^{-zb} f).$$

We claim that

$$\|S_z f\|_{L^p(w)} \le C \|f\|_{L^p(w)}$$

uniformly on $|z| \le s < \delta\eta/(1+\eta)$.

The function $z \mapsto S_z f$ is analytic, and by the Cauchy theorem, if $s < \delta\eta/(1+\eta)$,

$$\tfrac{d}{dz} S_z f\big|_{z=0} = \frac{1}{2\pi i} \int_{|z|=s} \frac{S_z f}{z^2}\, dz.$$

Observing that

$$\tfrac{d}{dz} S_z f\big|_{z=0} = [b, T] f$$

and applying the Minkowski inequality to the previous equality we conclude

$$\| \, [b,T]f \, \|_{L^p(w)} \le \frac{1}{2\pi} \int_{|z|=s} \frac{\|S_z f\,\|_{L^p(w)}}{s^2}\, |dz| \le \frac{C}{s} \, \| \, f \, \|_{L^p(w)} \, .$$

Thus, we are left with proving the claim, which is equivalent to

$$\int |Tf(x)|^p \exp(\Re(z) p b(x))\, w(x)\, d\mu(x) \le C \int |f(x)|^p \exp(\Re(z) p b(x))\, w(x)\, d\mu(x) \qquad (14)$$

We write $w_0 := \exp(\Re(z) p\, b(1+\eta)/\eta)$ and $w_1 := w^{1+\eta}$. Since w_0 and w_1 belong to $A_p(\mu)$ we have

$$\int |Tf(x)|^p w_0(x)\, d\mu(x) \le C \int |f(x)|^p w_0(x)\, d\mu(x)$$

and

$$\int |Tf(x)|^p w_1(x)\, d\mu(x) \le C \int |f(x)|^p w_1(x)\, d\mu(x).$$

Now, by the Stein-Weiss interpolation theorem (e.g. [BeL, p.115]) we have that

$$\int |Tf(x)|^p w_0^{1-\theta} w_1^{\theta}\, d\mu(x)$$

$$\le C \int |f(x)|^p w_0^{1-\theta} w_1^{\theta}\, d\mu(x)$$

and taking $\theta = (1+\eta)^{-1}$ we get (14).

□

5 BMO–Poincaré inequalities

In this section we shall apply the results from Section 2 to extend the main theorem from [FPW] and [MP] to our setting. It has been shown in these papers that there is a unified theory of some well known classical results concerning L^p properties of functions with some kind of smoothness, more precisely with control on the oscillation. In particular it is shown for instance both the classical Sobolev theorem and the L^p property of BMO functions are part of a more general phenomenon of self–improving properties.

As in [FPW] and [MP] we impose the following discrete condition on the functional a relative to a locally integrable weight function w.

Recall that a functional a is a function $a : \mathcal{Q} \to [0, \infty)$ where $\mathcal{Q}$ denotes the family of cubes with sides parallel to the coordinates axes. Recall that we use the notation $w(E) = \int_E w\, d\mu$.

Definition 5.1 *Let $0 < r < \infty$ and w be a weight function. We say that the functional a satisfies the weighted D_r condition if there exists a finite constant C such that for each cube Q and any family Δ of pairwise* **disjoint** *subcubes of Q,*

$$\sum_{P \in \Delta} a(P)^r w(P) \le C^r\, a(Q)^r w(Q). \tag{15}$$

We denote by $\|a\|$ the best constant C.

We introduce the notation

$$\|g\|_{L^{r,\infty}(Q,w)} = \sup_{t>0} t \left(\frac{w(\{x \in Q : |g(x)| > t\})}{w(Q)} \right)^{1/r}$$

for the normalized weak or Marcinkiewicz L^r norm.

Before we present our result we need to make some observations in order to adapt to our setting some well known properties of the so–called optimal polynomials defined as follows. Fix a cube Q and a nonnegative integer m. The space $\mathcal{P}_m$ of real-valued polynomials of degree at most m is a Hilbert space with the inner product

$$\frac{1}{\mu(Q)} \int_Q f\, g\, d\mu.$$

Consider the orthonormal basis $\{\varphi_\nu\}$, $|\nu| \le m$, obtained by applying the Gram–Schmidt orthonormalization process to the power functions $\{x^\nu\}$, $|\nu| \le m$. Observe that

$$\|\varphi_\nu\|_{L^\infty(Q)} \le C \left(\frac{1}{\mu(Q)} \int_Q |\varphi_\nu|^2\, d\mu \right)^{1/2} = C \tag{16}$$

since the space $\mathcal{P}_m$ is finite dimensional, and so all norms on it are equivalent. The constant C depends only on m.

We let P_Q be the operator defined by

$$P_Q f(x) = \sum_{|\nu| \le m} \frac{1}{\mu(Q)} \int_Q f \varphi_\nu \, d\mu \; \varphi_\nu(x),$$

which is a projection from $L^1(Q)$ onto $\mathcal{P}_m$. By (16) we have the following key property:

$$\|P_Q f\|_{L^\infty(Q)} \le \frac{\gamma}{\mu(Q)} \int_Q |f(y)| \, d\mu(y), \tag{17}$$

where $\gamma = C^2$. Observe that when $m = 0$, $P_Q f = f_Q = \frac{1}{\mu(Q)} \int_Q f \, d\mu$. These polynomials $P_Q f$ are optimal in the sense that

$$\inf_{\pi \in \mathcal{P}_m} \frac{1}{\mu(Q)} \int_Q |f - \pi| \, d\mu \approx \frac{1}{\mu(Q)} \int_Q |f - P_Q f| \, d\mu.$$

In fact we may replace the L^1 norm by any L^p norm, $1 < p < \infty$. Indeed, the inequality in the direction "$\le$" is trivial. To prove the opposite inequality, observe that since P_Q is a projection we have $P_Q \pi = \pi$ for any polynomial of degree at most m, and therefore

$$\frac{1}{\mu(Q)} \int_Q |f - P_Q f| \, d\mu \le \frac{1}{\mu(Q)} \int_Q (|f - \pi| + |P_Q(f - \pi)|) \, d\mu$$
$$\le \frac{1}{\mu(Q)} \int_Q |f - \pi| \, d\mu + \|P_Q(f - \pi)\|_{L^\infty(Q)} \le \frac{1 + \gamma}{\mu(Q)} \int_Q |f - \pi| \, d\mu$$

by (17).

THEOREM 5.2 *Let μ be a measure as above and let w be an $A_\infty(\mu)$ weight. Let a be a functional satisfying the D_r condition (15). Suppose that f is a locally integrable function such that for all cubes Q in $\mathbb{R}^n$*

$$\frac{1}{\mu(Q)} \int_Q |f - P_Q f| \, d\mu \le a(Q) \tag{18}$$

Then there exists a constant C such that for all the cubes Q in $\mathbb{R}^n$

$$\|f - P_Q f\|_{L^{r,\infty}(Q,w)} \le C \, \|a\| \, a(Q).$$

Corollary 5.3 *Under the same hypothesis of the Theorem, if $0 < p < r$, then there exists a constant $C = C(p)$ independent of f and Q such that*

$$\left(\frac{1}{w(Q)} \int_Q |f - P_Q f|^p \, w d\mu \right)^{1/p} \le C \, \|a\| \, a(Q).$$

This is a consequence of the well-known inequality

$$\left(\frac{1}{\nu(E)}\int_E |h|^p \, wd\mu\right)^{1/p} \leq C_{p,r}\,\|h\|_{L^{r,\infty}(E)},$$

which holds in any measure space of finite measure, and whenever p lies between zero and r.

Proof of Theorem 5.2. To prove the theorem we adapt the method considered in [FPW, Appendix] which again is based on the good–λ method.

The first step. For a fixed cube Q we let $E(Q,t) = \{x \in Q : |f(x) - P_Q f(x)| > t\}$ and $\Omega(Q,t) = \{x \in Q : M_Q(f - P_Q f)(x) > t\}$. Observe that by the Lebesgue differentation theorem $E(Q,t) \subset \Omega(Q,t)$, μ almost everywhere. We want to prove that

$$\sup_{t>0} t^r \frac{w(\Omega(Q,t))}{w(Q)} \leq C\, a(Q)^r \tag{19}$$

with a constant C independent of $t > 0$. First observe that we may assume that $t > a(Q)$ since otherwise (19) is easy. With this assumption the Besicovitch-Calderón–Zygmund decomposition of $|f - P_Q f|$ gives us a family $\{Q_i^t\}$ of cubes strictly contained in Q, such that

$$\frac{1}{\mu(Q_i^t)}\int_{Q_i^t} |f - P_Q f|\, d\mu = t \tag{20}$$

and such that $\Omega(Q,t) \subset \cup_i Q_i^t$ μ–almost everywhere. Since w is absolutely continuous with respect to μ we have

$$w(E(Q,t)) \leq \sum_i w(Q_i^t).$$

To any of these cubes Q_i^t we perform again the corresponding Besicovitch-Calderón–Zygmund decomposition of $|f - P_Q f|$ at level qt with $q > 1$ and we obtain another family of subcubes $\{Q_j^{qt}\}$, strictly contained in Q_i^t, such that for each j

$$\frac{1}{\mu(Q_j^{qt})}\int_{Q_j^{qt}} |f - P_Q f|\, d\mu = \, q\, t \tag{21}$$

and hence for each i, $\{x \in Q_i^t : |f(x) - P_Q f(x)| > qt\} \subset \bigcup_j Q_j^{q\,t}$ (μ–almost everywhere). On the other hand, if $x \in Q \setminus \cup_i Q_i^t$ then $x \notin \cup_j Q_j^{q\,t}$ and hence $|f(x) - P_Q f(x)| \leq qt$ μ–almost everywhere. Therefore

$$E(Q,qt) = \bigcup_i \{x \in Q_i^t : |f(x) - P_Q f(x)| > qt\} \subset \bigcup_j Q_j^{qt} \tag{22}$$

μ–almost everywhere and consequently

$$w(E(Q,qt)) \leq \sum_j w(Q_j^{q\,t}).$$

The second step (good-lambda inequality).

Recall that $\|a\|, \gamma, B(n) \geq 1$ denote the constants in (15), (17) and the Besicovitch constant respectively. Also, ρ denotes the constant in the A_∞ condition (5).

We claim the following:

There exists a constant c such that for each $q > \gamma$ and $0 < \epsilon \leq \|a\|$ and $t > 0$ the following estimate holds

$$\sum_j w(Q_j^{q\,t}) \leq [B(n)]^2 \left(\frac{\epsilon^\rho\, c}{(q-\gamma)^\rho} \sum_i w(Q_i^t) + \frac{\|a\|^r}{\epsilon^r t^r} a(Q)^r w(Q) \right). \qquad (23)$$

Let us observe first that for $t \leq a(Q)$ the inequality (23) trivially holds. Indeed, since the family $\{Q_j^{qt}\}$ has bounded overlap with constant $B(n)^2$ we have

$$\sum_j w(Q_j^{q\,t}) \leq B(n)^2\, w(Q) \leq [B(n)]^2\, \frac{\|a\|^r}{\epsilon^r t^r} a(Q)^r w(Q)$$

which is smaller than the right side of the inequality.

¿From now on we fix $\epsilon > 0$ and $t > a(Q)$. Using (2) it is easy to see that we may assume that the cubes $\{Q_i^t\}$ are pairwise disjoint as well as the cubes $\{Q_j^{tq}\}$ inside each cube Q_i^t. Therefore we only have to prove (23) without the constant $(B(n))^2$ on its right hand side.

We split the family $\{Q_i^t\}$ in two:
(i) $i \in I$ if

$$\frac{1}{|Q_i^t|} \int_{Q_i^t} |f - P_{Q_i^t} f| \, < \epsilon\, t$$

or (ii) $i \in II$ if

$$\frac{1}{|Q_i^t|} \int_{Q_i^t} |f - P_{Q_i^t} f| \, \geq \epsilon\, t.$$

Then

$$\sum_j w(Q_j^{q\,t}) = \sum_{i\in I} \sum_{Q_j^{q\,t} \subset Q_i^t} w(Q_j^{qt}) + \sum_{i\in II} \sum_{Q_j^{q\,t} \subset Q_i^t} w(Q_j^{qt}) = I \,+ II$$

To estimate the second sum II we will use (18) and (15):

$$II \leq \sum_{i\in II} w(Q_i^t) \leq \sum_{i\in II} \left(\frac{1}{\epsilon t |Q_i^t|} \int_{Q_i^t} |f - P_{Q_i^t} f| \right)^r w(Q_i^t)$$

$$\leq \frac{1}{\epsilon^r t^r} \sum_i a(Q_i^t)^r w(Q_i^t) \leq \frac{\|a\|^r}{\epsilon^r t^r} a(Q)^r w(Q).$$

To estimate I we use that $w \in A_\infty(\mu)$ and therefore it satisfies (5). Thus

$$I=\sum_{i\in I} w(\bigcup_{Q_j^{qt}\subset Q_i^t} Q_j^{qt}) \leq c\sum_{i\in I}\left(\frac{\mu(\bigcup_{Q_j^{qt}\subset Q_i^t} Q_j^{qt})}{\mu(Q_i^t)}\right)^{\rho} w(Q_i^t)$$

Now to estimate the inner unweighted part we first observe that by (17) and (20)

$$|P_{Q_i^t}f-P_Qf|=|P_{Q_i^t}(f-P_Qf)|\leq \frac{\gamma}{\mu(Q_i^t)}\int_{Q_i^t}|f-P_Qf|\,d\mu=\gamma t$$

and hence by (21)

$$qt=\frac{1}{\mu(Q_j^{qt})}\int_{Q_j^{qt}}|f-P_Qf|\,d\mu\leq \frac{1}{\mu(Q_j^{qt})}\int_{Q_j^{qt}}|f-P_{Q_i^t}f|\,d\mu+\gamma t$$

and then

$$\mu(Q_j^{qt})\leq \frac{1}{(q-\gamma)t}\int_{Q_j^{qt}}|f-P_{Q_i^t}f|\,d\mu.$$

Therefore

$$\mu(\bigcup_{Q_j^{qt}\subset Q_i^t} Q_j^{qt})=\sum_{Q_j^{qt}\subset Q_i^t}\mu(Q_j^{qt})\leq \sum_{Q_j^{qt}\subset Q_i^t}\frac{1}{(q-\gamma)t}\int_{Q_j^{qt}}|f-P_{Q_i^t}f|\,d\mu$$

$$\leq \frac{1}{(q-\gamma)t}\int_{Q_i^t}|f-P_{Q_i^t}f|\,d\mu\leq \frac{\epsilon}{q-\gamma}\mu(Q_i^t)$$

since $i\in I$. Combining this estimate we get

$$I\leq \frac{\epsilon^{\rho}\,c}{(q-\gamma)^{\rho}}\sum_{i\in\mathcal{F}_k} w(Q_i^t)$$

and finally

$$\sum_j w(Q_j^{q\,t})\leq \left(\frac{\epsilon^{\rho}\,c}{(q-\gamma)^{\rho}}\sum_i w(Q_i^t)+\frac{\|a\|}{\epsilon^r t^r}a(Q)^r w(Q)\right),$$

as desired.

The third step. We now are ready to prove (19). Let $t_0>0$ be a constant to be chosen in a moment and let q such that $q-\gamma>1$. Also, we denote $E(Q,t_0)=E_0$ and $E(Q,q^m t_0)=E_m$ for $m=1,2,\cdots$. We have the inclusions

$$E_0\supset E_1\supset\cdots E_m\supset\cdots.$$

To each set E_m we assign and fix a family of cubes $\{Q_i^m\}$ following the method described above in the first step of the proof of the theorem. We start by taking the family $\{Q_i^0\}=$

$\{Q_i^{t_0}\}$, corresponding to the first level set E_0, and from this we obtain $\{Q_i^1\}$ and we keep repeating the procedure. Then we have

$$\bigcup_i Q_i^0 \supset \bigcup_i Q_i^1 \supset \bigcup_i Q_i^2 \supset \cdots \supset \bigcup_i Q_i^m \supset \cdots$$

and by (22) we have, μ–almost everywhere,

$$E_m \subset \bigcup_i Q_i^m.$$

Therefore

$$w(E_m) \le \sum_i w(Q_i^m) = I(m).$$

Now, since $q > \gamma$ we have by the good-λ inequality (23) that

$$I(m) \le C\,\epsilon^\rho I(m-1) + C\,\frac{\|a\|^r}{\epsilon^r (q^{m-1}t_0)^r} a(Q)^r w(Q),$$

and hence for each $m = 1, 2, \cdots$

$$\frac{(q^m t_0)^r I(m)}{w(Q)} \le C\,q^r \epsilon^\rho \frac{(q^{m-1}t_0)^r I(m-1)}{w(Q)} + C\,q^r \frac{\|a\|^r}{\epsilon^r} a(Q)^r. \tag{24}$$

Also for $m = 0$ there is a corresponding inequality:

$$\frac{t_0^r I(0)}{w(Q)} \le \frac{C\,q^r \|a\|^r}{\epsilon^r} a(Q)^r,$$

using that $I(0) \le C\,w(Q)$, $\epsilon \le \|a\|$, and by choosing $t_0 = q\,a(Q)$.

Now for each $N = 1, 2, \cdots$, we define

$$\varphi(N) = \sup_{m=0,1,\cdots,N} \frac{(q^m t_0)^r I(m)}{w(Q)} < \infty.$$

Then, combining the last inequality with (24) we get

$$\varphi(N) \le C\,q^r \epsilon^\rho \varphi(N) + C\,q^r \frac{\|a\|^r}{\epsilon^r} a(Q)^r$$

and we can choose $\epsilon > 0$ sufficiently small such that

$$\varphi(N) \le C\,\|a\|^r\,a(Q)^r.$$

This means that

$$t^r \frac{w(E_t)}{w(Q)} \le C\,\|a\|^r\,a(Q)^r,$$

for t of the form $t = q^m t_0$, $m = 0, 1, \cdots$. This yields immediately the result for arbitrary t concluding the proof of the theorem.

□

References

[BeL] J. Bergh and J. Löfström, "*Interpolation Spaces*," Springer-Verlag, 1976.

[ChF] M. Christ and R. Fefferman, *A note on weighted norm inequalities for the Hardy-Littlewood maximal operator*, Proc. Amer. Math. Soc. **12** (1983), 447–448.

[C] R. Coifman, *Distribution function inequalities for singular integrals*, Proc. Acad. Sci. U.S.A. **69** (1972), 2838–2839.

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

[CRW] R. Coifman, R. Rochberg and G. Weiss, *Factorization theorems for Hardy spaces in several variables*, Ann. of Math. **103** (1976), 611–635.

[CR] R. Coifman and R. Rochberg, *Another characterization of BMO*, Proc. Amer. Math. Soc. **79** (1980), 249–254.

[FPW] B. Franchi, C. Pérez and R. L. Wheeden, *Self– Improving Properties of John–Nirenberg and Poincaré Inequalities on Spaces of Homogeneous Type*, J. Funct. Anal. **153** (1998), 108–146.

[GCRdF] J. García-Cuerva and J.L. Rubio de Francia, "*Weighted Norm Inequalities and Related Topics*," North Holland Math. Studies 116, North Holland, Amsterdam, 1985.

[GMOPST] P. Gurka, P., F. J. Martín-Reyes, P. Ortega, L. Pick, M. D. Sarrión and A. de la Torre, *Good and bad measures*, J. London Math. Soc. (2) **61** (2000), no. 1, 123–138.

[Ja] B. Jawerth, *Weighted norm inequalities: linearization, localization, and factorization*, Amer. J. Math. **108** (1986), 361-414.

[MP] P. MacManus and C. Pérez, *Generalized Poincaré inequalities: Sharp self–improving properties*, Internat. Math. Res. Notices, **1998** (2), 101– 116.

[MMNO] J. Mateu, P. Mattila, A. Nicolau and J. Orobitg, *BMO for nondoubling measures*, Duke Math. J. **102** (2000), 533–565.

[NTV1] F. Nazarov, S. Treil and A. Volberg, Cauchy integral and Calderón-Zygmund operators on nonhomogeneous spaces, Internat. Math. Res. Notices **1997** (15), 703–726.

[NTV2] F. Nazarov, S. Treil and A. Volberg, Weak type estimates and Cotlar inequalities for Calderón-Zygmund operators on nonhomogeneous spaces, Internat. Math. Res. Notices **1998** (9), 463–487.

[Sa] E. Saksman, *The local mapping properties of the Hilbert transform*, preprint 1999.

[St] E. M. Stein, "*Harmonic Analysis*," Princeton Univ. Press, Princeton, NJ, 1995.

[T1] X. Tolsa, *L^2-boundedness of the Cauchy integral operator for continuous measures*, Duke Math. J. **98** (1999), 269–304.

[T2] X. Tolsa, *Cotlar's inequality and existence of principal values for the Cauchy integral wihout the doubling condition*, J. Reine Angew. Math. **502** (1998), 199–235.

[T3] X. Tolsa, *BMO, H^1 and Calderón-Zygmund operators for non doubling measures*, to appear in Math. Ann.

[V] J. Verdera, *On the $T(1)$-Theorem for the Cauchy integral*, Ark. Mat. **38** (2000), 183–199.

Joan Orobitg
Departament de Matemàtiques
Universitat Autònoma de Barcelona
08193 Bellaterra, Spain
orobitg@mat.uab.es

Carlos Prez
Departamento de Matemáticas
Universidad Autónoma de Madird
28049 Madrid, Spain
carlos.perez@uam.es