STUDIA MATHEMATICA, T. LXXXV. (1987)

A note on martingale transforms and A_p-weights

by

RODRIGO BAÑUELOS (Pasadena, Cal.)

Abstract. We study martingale transforms and A_p-weights on the d-adic filtration. The functions in Muckenhoupt's A_p class are characterized by the boundedness of the d-adic Hilbert transform and more general nondegenerate martingale transforms.

A positive locally integrable function W in $\boldsymbol{R}^n$ is in the (Muckenhoupt) class $A_p(\boldsymbol{R}^n)$, $p > 1$, if

$$\sup_I \left(\frac{1}{|I|}\int_I W\,dx\right)\left(\frac{1}{|I|}\int_I W^{-1/(p-1)}\,dx\right)^{p-1} < \infty,$$

the sup being taking over all cubes I with sides parallel to the axes.

A well-known result of R. Hunt, B. Muckenhoupt, and R. Wheeden [5] says that the Hilbert transform H characterizes $A_p(\boldsymbol{R})$ in the following sense: $W \in A_p(\boldsymbol{R})$, $p > 1$, if and only if H is bounded on $L^p(W dx)$. In $\boldsymbol{R}^n$, $n > 1$, the Hilbert transform is replaced by the Riesz transforms and we have a similar characterization: $W \in A_p(\boldsymbol{R}^n)$, $p > 1$, if and only if the Riesz transforms are bounded on $L^p(W dx)$. The relationship between A_p-weights and more general singular integral operators was established by R. Coifman and C. Fefferman [3] who showed that if $W \in A_p(\boldsymbol{R}^n)$, $p > 1$, then any singular integral operator of Calderón–Zygmund type is bounded on $L^p(W dx)$. (See E. Stein [9] for the definitions and properties of Calderón–Zygmund operators.) On the other hand, given $p > 1$, it is not difficult to construct a $W \notin A_p(\boldsymbol{R}^n)$ and Calderón–Zygmund operators which are bounded on $L^p(W dx)$. Thus not every singular integral operator characterizes $A_p(\boldsymbol{R}^n)$. As far as we know, no one has given precise conditions on a Calderón–Zygmund kernel in order for the corresponding singular integral operator to characterize $A_p(\boldsymbol{R}^n)$. In this note we approach this problem in the martingale setting. We look for necessary and sufficient conditions on a collection of matrices in order for their martingale transforms to characterize the d-adic A_p. In § 1, we prove the analogue of the Hunt–Muckenhoupt–Wheeden result in terms of what we call the "d-adic Hilbert transform". In § 2, we define what it means for a collection of matrices $\{A_0, \ldots, A_m\}$ to be nondegenerate and prove that this is a sufficient condition for the boundedness of their martingale transforms to characterize the d-adic A_p. We also give examples which give strong indications (but no proof) that the nondegeneracy condition is necessary.

1. The d-adic Hilbert transform and A_p-weights. Let $(\Omega, \mathscr{F}, P)$ be a probability space and $\{\mathscr{F}_n\}$ an increasing sequence of σ-subfields in $\mathscr{F}$ such that $\mathscr{F}_n$ is generated by d^n disjoint atoms of probability d^{-n}, d a fixed integer greater than or equal to 3. Thus an atom Q of $\mathscr{F}_n$ is the union of d atoms of $\mathscr{F}_{n+1}$ which we denote by $Q_1, \ldots, Q_d$. There is no loss of generality in assuming that our filtration is the usual d-adic filtration of $(0, 1]$ with the Lebesgue measure dx as P. To obtain this filtration fix n and for each $1 \leqslant j \leqslant d^n$, let $A(j, n) = \left((j-1)/d^n, j/d^n\right]$. Let $\mathscr{F}_0$ be the trivial σ-field $\{\emptyset, (0, 1]\}$ and if $n \geqslant 1$, let $\mathscr{F}_n$ be the σ-field generated by the d^n intervals $A(j, n)$. Finally, let $\mathscr{F} = \sigma(\bigcup_{n \geqslant 0} \mathscr{F}_n)$.

We now define the martingale transforms. If $f \in L^1(\mathscr{F})$, we put $f_n = E(f|\mathscr{F}_n)$, the conditional expectation of f given $\mathscr{F}_n$. Observe that on any atom Q of $\mathscr{F}_n$, f_n is constant and f_{n+1} takes d values. Hence in studying one atom only the martingale difference $f_{n+1} - f_n$ may be regarded as a vector in $\boldsymbol{R}^d$ (we assume our functions to be real-valued), which is called the *local difference* of f on Q. It is easy to see that every local difference is actually a vector in the $(d-1)$-dimensional subspace $V = \{x \in \boldsymbol{R}^d: \sum_{i=1}^{d} x_i = 0\}$. Let A be a linear operator on V. Define $A * f$ to be the function whose local differences are obtained from those of f by the operator A. We also require our martingale transforms to start at 0, that is, $E(A * f) = 0$. As is customary, we extend A to be an operator from $\boldsymbol{R}^d$ to $\boldsymbol{R}^d$ and represent it by a matrix (a_{ij}) such that $\sum_i a_{ij} = \sum_j a_{ij} = 0$. In this way, for example, the identity from V to V is given by $I = (a_{ij})$ with $a_{ii} = (d-1)/d$ for all i and $a_{ij} = -1/d$ for $i \neq j$. The martingale transform that corresponds to the identity is $I * f = f - Ef$.

In what follows we will always assume that our martingales are uniformly integrable. We define $f_n^* = \sup_{k \leqslant n} |f_k|$ and $f^* = \sup_n |f_n|$. We denote the difference sequence by $d_k = \Delta f_k = f_k - f_{k-1}$ and set

$$S(f) = \left(\sum_{k=1}^{\infty} d_k^2\right)^{1/2}, \quad \sigma(f) = \left(\sum_{k=1}^{\infty} E(d_k^2|\mathscr{F}_{k-1})\right)^{1/2}.$$

$S(f)$ is called the *area function* and $\sigma(f)$ the *conditioned area function.*

If W is a positive function in $L^1(\mathscr{F})$ we say that $W \in A_p\{\mathscr{F}\}$, $p > 1$, if there exists a constant K such that

$$\sup_n \left\|\left[E\left((W_n/W)^{1/(p-1)} \middle| \mathscr{F}_n\right)\right]^{p-1}\right\|_{L^\infty(\mathscr{F})} < K.$$

When $p = 1$, we say $W \in A_1\{\mathscr{F}\}$ if $W^* \leqslant KW$ modulo null sets.

The A_p condition for general martingales was first considered by

M. Izumisawa and N. Kazamaki in [6]. See also [1] or [4] for more on A_p-weights and martingales.

We set some notation. Throughout this note, $\hat{E}$ will mean expectation with respect to the weighted measure $d\hat{P} = W dx$. K will always refer to some numerical constant which may not be the same at each occurrence but which is independent of the functions f involved. Finally, for any atom Q and integrable function f, $E(f; Q) = \int_Q f\,dx$, $E(f|Q) = |Q|^{-1} \int_Q f\,dx = f_Q$.

For $d \geqslant 3$ an odd integer we define the matrix H^d by

$$H^3 = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix}, \quad H^5 = \begin{bmatrix} 0 & -1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 & -1 \\ -1 & 1 & -1 & 1 & 0 \end{bmatrix}, \quad \ldots$$

We call H^d the *d-adic Hilbert transform.* As in the case of the classical Hilbert transform, H^d characterizes the d-adic BMO (see S. Janson [7] for this). We have the following analogue of the Hunt–Muckenhoupt–Wheeden result mentioned above. To simplify notation set $H = H^d$.

THEOREM 1. *The following are equivalent for* $p > 1$:

(a) $W \in A_p\{\mathscr{F}\}$.

(b) $\hat{E}|(H * f)^*|^p \leqslant K\hat{E}|f|^p$.

(c) $\hat{P}\{(H * f)^* > \lambda\} \leqslant \dfrac{K}{\lambda^p}\hat{E}|f|^p$.

(d) $\hat{E}|H * f|^p \leqslant K\hat{E}|f|^p$.

(e) $\hat{P}\{|H * f| > \lambda\} \leqslant \dfrac{K}{\lambda^p}\hat{E}|f|^p$.

When $p = 1$, (a), (c), *and* (e) *are equivalent.*

Before proving the theorem we need to recall a

LEMMA. *If* $W \in A_p\{\mathscr{F}\}$, $p > 1$, *then*

$$\hat{E}|\sigma(f)|^p \approx \hat{E}|S(f)|^p \approx \hat{E}|f^*|^p \approx \hat{E}|f|^p,$$

where $A \approx B$ *means that there exist constants* K_1 *and* K_2 *independent of* f *such that* $K_1 A \leqslant B \leqslant K_2 A$.

For the proof of this lemma see [8]. We remark that the equivalence $\hat{E}|f^*|^p \approx \hat{E}|f|^p$ may not be true if the filtration $\{\mathscr{F}_n\}$ is not regular as shown by Uchiyama [10].

Proof of Theorem 1. Assume $p > 1$. Let A be any linear operator from V to V, not necessarily H. Let Δg_n denote the martingale differences of $A * f$. Then for any atom $Q \in \mathscr{F}_{n-1}$,

$$E(|\Delta g_n|^2|\mathcal{F}_{n-1})(Q)=E(|\Delta g_n|^2|Q)=\frac{1}{d}\sum_{i=1}^{d}E(|\Delta g_n|^2|Q_i)$$
$$=\frac{1}{d}\sum_{i=1}^{d}\Big|\sum_{j=1}^{d}a_{ij}\,\Delta f_n(Q_i)\Big|^2=\|A\|^2\frac{1}{d}\sum_{i=1}^{d}|\Delta f_n(Q_i)|^2$$
$$=\|A\|^2E(|\Delta f_n|^2|\mathcal{F}_{n-1})(Q),$$

where $\|A\|$ is the norm of the matrix A. From this it follows that $\sigma(A*f)\leqslant\|A\|\,\sigma(f)$. By the Lemma, if $W\in A_p\{\mathcal{F}\}$,

$$\hat{E}|(A*f)^*|^p\leqslant K\,\hat{E}|\sigma(A*f)|^p\leqslant\|A\|^pK\,\hat{E}|\sigma(f)|^p\leqslant\|A\|^pK\,\hat{E}|f|^p.$$

Thus (a) $\Rightarrow$ (b). Clearly (b) $\Rightarrow$ (c) $\Rightarrow$ (e) and (b) $\Rightarrow$ (d) $\Rightarrow$ (e). Therefore to complete the equivalence for $p>1$ we need to show (e) $\Rightarrow$ (a).

Let $Q\in\mathcal{F}_n$ be any atom and $Q_1,\ldots,Q_d$ the atoms in $\mathcal{F}_{n+1}$ whose union is Q. We claim there are constants K_1 and K_2 depending only on W such that

$$K_1E(W;Q_i)\leqslant E(W;Q_j)\leqslant K_2E(W;Q_i) \tag{1.1}$$

for any i and j. To see this assume without loss of generality that $i=1$ and let us show

$$E(W;Q_1)\leqslant K_1E(W;Q_j). \tag{1.2}$$

Set $f=\chi_{Q_j}-\chi_{Q_d}$. If $x\in Q_1$ an easy calculation shows that $(H*f)(x)=-2$ if $j=2$ and $(H*f)(x)=-1$ if $j\neq 2$. In any case $|H*f|\geqslant 1$ on Q_1 from which it follows that

$$\begin{aligned}E(W;Q_1)&\leqslant E(|H*f|^pW;Q_1)\leqslant\hat{E}|H*f|^p\leqslant K\,\hat{E}|f|^p\\&\leqslant K\,\hat{E}(W;Q_j)+K\,\hat{E}(W;Q_d).\end{aligned} \tag{1.3}$$

Next, set $f=\chi_{Q_j}$. Then $H*f=1$ on Q_d and

$$E(W;Q_d)\leqslant\hat{E}(|H*f|^p;Q_d)\leqslant\hat{E}|H*f|^p\leqslant K\,\hat{E}|f|^p\leqslant K\,E(W;Q_j).$$

The last inequality and (1.3) prove (1.2). The other side of (1.1) is proved similarly.

If we put $f=W^{\alpha}\chi_{Q_i}$, $\alpha=-1/(p-1)$, an easy calculation shows that $|H*(\lambda f/f_{Q_i})|\approx\lambda$ on Q_j for some $j\neq i$. From this we have

$$E(W;Q_j)\leqslant\hat{P}\{|H*(\lambda f/f_{Q_i})|>K\lambda\}\leqslant\frac{K}{\lambda^p}\hat{E}|\lambda f/f_{Q_i}|^p$$

from which it follows that

$$E(W;Q_j)[E(W^{-1/(p-1)}|Q_i)]^p\leqslant K\,E(W^{-1/(p-1)};Q_i).$$

This inequality together with (1.1) gives

$$E(W|Q_i)[E(W^{-1/(p-1)}|Q_i)]^p\leqslant K\,E(W^{-1/(p-1)}|Q_i) \tag{1.4}$$

from which the A_p condition follows and the theorem is proved for $p>1$.

Next, suppose $p=1$. We first show (a) $\Rightarrow$ (c). Again, assume A is any linear operator on V. $W_n\leqslant K\,W$ implies

$$[E((W_n/W)^{1/(p-1)}|\mathcal{F}_n)]^{p-1}\leqslant K \tag{1.5}$$

and therefore if $W\in A_1\{\mathcal{F}\}$, $W\in A_p\{\mathcal{F}\}$ for any $p>1$. Set $|f|_n=E(|f||\mathcal{F}_n)$ and observe that $W\in A_1\{\mathcal{F}\}$ implies

$$E(|f||\mathcal{F}_n)\leqslant K\,E\Big(\frac{W}{W_n}|f|\,\Big|\,\mathcal{F}_n\Big)=K\,\hat{E}(|f||\mathcal{F}_n) \tag{1.6}$$

where the last inequality follows from the fact that if X is an integrable random variable,

$$\hat{E}(X|\mathcal{F}_n)=E\Big(\frac{W}{W_n}X\,\Big|\,\mathcal{F}_n\Big)\quad\text{a.e.} \tag{1.7}$$

under both measures P and $\hat{P}$. So from (1.6) we have $|f|^*\leqslant K|f|^{**}$ where $|f|^{**}=\sup\hat{E}(|f||\mathcal{F}_n)$. Applying Doob's inequality with respect to $\hat{P}$, we have

$$\hat{P}\{|f|^*>\lambda\}\leqslant\frac{K}{\lambda}\hat{E}|f|. \tag{1.8}$$

Let $\tau=\inf\{n\colon |f|_n>\lambda\}$. Then

$$\begin{aligned}\lambda\hat{P}\{|f|^*\leqslant\lambda,\,(A*f)^*>\lambda\}&=\lambda\hat{P}\{\tau=\infty,\,(A*f)^*>\lambda\}\\&\leqslant\lambda\hat{P}\{(A*f)^*_\tau>\lambda\}.\end{aligned} \tag{1.9}$$

It follows from Chebyshev's inequality that

$$\lambda^2\hat{P}\{((A*f)^*_\tau)^2>\lambda^2\}\leqslant K\,\hat{E}(|(A*f)^*_\tau|^2). \tag{1.10}$$

Since $W\in A_2\{\mathcal{F}\}$ the first part of the theorem can be applied and we find

$$K\,\hat{E}|(A*f)^*_\tau|^2\leqslant\|A\|^2K\,\hat{E}(|f|_\tau^2). \tag{1.11}$$

Since $|f|_n\leqslant d|f|_{n-1}$ a.e. by breaking our stopping time τ according to its values we also have $|f|_\tau\leqslant d|f|_{\tau-1}$ a.e. So,

$$\|A\|^2K\,\hat{E}(|f|_\tau^2)\leqslant K\|A\|^2\hat{E}(|f|_{\tau-1}^2)\leqslant\lambda K\|A\|^2\hat{E}(|f|_{\tau-1}) \tag{1.12}$$

where the last inequality follows from the definition of τ. From (1.5), $|f|_n\leqslant K\,\hat{E}(|f||\mathcal{F}_n)$ and applying $\hat{E}$ to both sides gives $\hat{E}(|f|_n)\leqslant K\,\hat{E}|f|$. Thus $\hat{E}(|f|_{\tau-1})\leqslant K\,\hat{E}|f|$.

Putting this together with (1.10), (1.11), and (1.12) gives

$$\lambda\hat{P}\{((A*f)^*_\tau)^2>\lambda^2\}\leqslant K\,\hat{E}|f|.$$

This inequality, (1.8), and (1.9) show

$$\hat{P}\{(A*f)^* > \lambda\} \leqslant \frac{K}{\lambda}\hat{E}|f|$$

and (a) $\Rightarrow$ (c).

The proof of the last implication follows the proof of Lemma 1 in D. Burkholder [2] very closely.

To complete the proof, it remains to show (e) $\Rightarrow$ (a), since (c) $\Rightarrow$ (e) trivially. This is done exactly as in the classical case (see [5]) and therefore we shall be very brief. We have to show $W^* \leqslant K W$. Let Q_{i_0} be any atom in $\mathscr{F}_n$ and Q the atom in $\mathscr{F}_{n-1}$ such that $Q_{i_0} \subset Q$ and $Q_1, \ldots, Q_{i_0-1}, Q_{i_0+1}, \ldots, Q_d$ are the atoms in $\mathscr{F}_n$ such that $\bigcup_{i=1}^{d} Q_i = Q$. If $\operatorname{ess\,inf}\{W(x)\colon x \in Q_{i_0}\} = \infty$, then we are done. If not, given $\varepsilon > 0$ we can find a subset $\tilde{Q} \subset Q_{i_0}$ of positive measure such that $W(\tilde{x}) \leqslant \varepsilon + \operatorname{ess\,inf}\{W(x)\colon x \in Q_{i_0}\}$, $\tilde{x} \in \tilde{Q}$. Put $f = \chi_{\tilde{Q}}$. Then $H*f = |\tilde{Q}|/|Q_{i_0}|$ on Q_j, where Q_j is one of the brothers of Q_{i_0} in $\mathscr{F}_n$. So,

$$E(W;\, Q_j) \leqslant K\frac{|Q_{i_0}|}{|\tilde{Q}|}E(W;\, \tilde{Q}) \leqslant K|Q_i|\,[\varepsilon + \operatorname{ess\,inf}\{W(x)\colon x \in Q_{i_0}\}].$$

Let $\varepsilon \to 0$ to get $E(W|Q_j) \leqslant K \operatorname{ess\,inf}\{W(x)\colon x \in Q_j\}$; applying (1.1) gives the result. This completes the proof of Theorem 1.

Remark. Suppose $W = 1$. A minor modification of the argument above gives the weak type (1, 1) inequality $P\{(A*f)^* > \lambda\} \leqslant (4d/\lambda)\|A\|\,E|f|$ and if we remove the maximal operator $*$, the 4 can be changed to a 2. (The constant given by S. Janson [7] is $5d\|A\|$.) It is then natural to ask whether these operators have a weak type (1, 1) inequality with constant independent of d. We have not been able to answer this question.

2. Nondegenerate transforms and A_p-weights. Let (a_{ij}) be the $d \times d$ matrix which represents the linear operator $A\colon V \to V$. We will say that A is i-*degenerate* if $a_{ij} = -a_{ii}/(d-1)$ or $a_{ji} = a_{ii}/(d-1)$ for every $j \neq i$. If no such i exists, we will say that A is *nondegenerate*. A collection of operators $\{A_0, \ldots, A_m\}$ is *degenerate* if there exists an i such that A_j is i-degenerate for every $j = 0, \ldots, m$. If no such i exists $\{A_0, \ldots, A_m\}$ is *nondegenerate*.

THEOREM 2. *Let $\{A_0, \ldots, A_m\}$ be nondegenerate and suppose $\hat{E}|A_r*f|^p \leqslant K\hat{E}|f|^p$ for all $r = 0, \ldots, m$ and for $p > 1$. Then $W \in A_p\{\mathscr{F}\}$. In the other direction, there exist $\{A_0, \ldots, A_m\}$ degenerate and $W \notin A_p\{\mathscr{F}\}$ such that $\hat{E}|A_r*f|^p \leqslant K\hat{E}|f|^p$ for all r.*

Proof. The case $d = 4$ already shows the general argument and therefore we only do this case. We first recall a few facts. Set $A_r = (a_{ij}^r)$.

Let q be the conjugate exponent of p and denote by A_r^T the transpose of A_r. We have

(2.1) $$\hat{E}|A_r*f|^p \leqslant K\hat{E}|f|^p \;\Leftrightarrow\; E(|A_r^T*f|^q W^{-1/(p-1)}) \leqslant K\,E(|f|^q W^{-1/(p-1)}).$$

This fact is an easy consequence of the fact that $(L^p(W dx))^*$ is isomorphic to $L^q(W^{-1/(p-1)}\,dx)$ and we leave it to the interested reader.

Now let Q_2 be any atom in $\mathscr{F}_n$ and Q_1, Q_3, and Q_4 the atoms in $\mathscr{F}_n$ such that $\bigcup_{i=1}^{4} Q_i = Q$, Q the atom in $\mathscr{F}_{n-1}$ containing Q_2. For the rest of the paper $\alpha = -1/(p-1)$. Our goal is to show

(*) $$E(W|Q_2)(E(W^\alpha|Q_2))^{p-1} \leqslant K$$

with K independent of Q_2 and n. If g is a positive function and $f = \chi_{Q_i} g$, an easy computation shows that $A_r*f = (A_r*\chi_{Q_i})g_{Q_i}$ on Q_j, $j \neq i$, and

(2.2) $$A_r*\chi_{Q_i} = a_{ji} + (\tilde{\alpha}a_{11}^r + \beta a_{22}^r + \gamma a_{33}^r + \delta a_{44}^r) \quad \text{on } Q_j$$

where $\tilde{\alpha} + \beta + \gamma + \delta = 1/4 + 1/4^2 + \ldots + 1/4^{n-1}$. Note that the term in parenthesis in (2.2) does not change if we replace A_r by A_r^T. For convenience we denote it by η_r. Also if $\sum_{r=0}^{m}|a_{ij}^r - a_{ik}^r| > \varepsilon > 0$, then $\sum_{r=0}^{m}|a_{ij}^r + \eta_r| + \sum_{r=0}^{m}|a_{ik}^r + \eta_r| > \varepsilon$. So one of these two sums must be greater than $\varepsilon/2$. Hence, if $\{A_0, \ldots, A_m\}$ is nondegenerate there exists $\varepsilon > 0$ depending only on the collection $\{A_0, \ldots, A_m\}$ such that given any i there exists a $j \neq i$ for which

(2.3) $$\sum_{r=0}^{m}|a_{ij}^r + \eta_r| > \varepsilon > 0.$$

Also, Hölder's inequality gives

(2.4) $$1 \leqslant E(W|Q_i)(E(W^{-1/(p-1)}|Q_i))^{p-1} \quad \text{for all } i.$$

We are now ready to prove (*).

Put $f = \chi_{Q_2}$. We have $A_r*f = a_{12}^r + \eta_r$ on Q_i. By our observation above there exists $i \neq 2$, which we may assume without loss of generality to be 1, such that

(2.5) $$\sum_{r=0}^{m}|a_{12}^r + \eta_r| > \varepsilon.$$

So, there exists $r_0 \in \{0, 1, \ldots, m\}$ such that

(2.6) $$\varepsilon/(m+1) \leqslant |a_{12}^{r_0} + \eta_{r_0}|.$$

Setting $f = \chi_{Q_2} W^\alpha$, we have

(2.7) $$|A_{r_0}*f| = |(A_{r_0}*\chi_{Q_2})\cdot(W^\alpha)_{Q_2}| \geqslant \frac{\varepsilon}{m+1}(W^\alpha)_{Q_2} = \frac{\varepsilon}{m+1}E(W^\alpha|Q_2) \quad \text{on } Q_1.$$

Therefore

$$\left(\frac{\varepsilon}{m+1}\right)^p E(W[E(W^\alpha|Q_2)]^p; Q_1) \leqslant E(W|A_{r_0} * f|^p) = \hat{E}|A_{r_0} * f|^p$$

$$\leqslant K\hat{E}|f|^p = K\hat{E}(W^{\alpha p}; Q_2) = KE(W^\alpha; Q_2)$$

or

$$E(W; Q_1)[E(W^\alpha|Q_2)]^p \leqslant KE(W^\alpha; Q_2)$$

from which we get

$$E(W|Q_1)[E(W^\alpha|Q_2)]^{p-1} \leqslant K. \tag{2.8}$$

Next we set $f = \chi_{Q_2} W$. Then $|A_r^T * f| = |a_{ji}^r + \eta_r| W_{Q_2}$ on Q_j, $j \neq 2$. So as above, there exists an $i_0 \neq 2$ such that on Q_{i_0} we have

$$\sum_{r=0}^{m} |A_r^T * f| = \sum_{r=0}^{m} |a_{i_0 i}^r + \eta_r| W_{Q_2} > \varepsilon W_{Q_2}. \tag{2.9}$$

Thus there exists $r_{i_0} \in \{0, 1, \ldots, m\}$ such that

$$\frac{\varepsilon}{m+1} W_{Q_2} \leqslant |A_{r_{i_0}}^T * f| \quad \text{on } Q_{i_0}. \tag{2.10}$$

From this we have

$$\begin{aligned} E(W_{Q_2}^q \cdot W^\alpha; Q_{i_0}) &\leqslant KE(|A_{r_{i_0}}^T * f|^q W^\alpha; Q_{i_0}) \leqslant KE(|A_{r_{i_0}}^T * f|^q W^\alpha) \\ &\leqslant KE(|f|^q W^\alpha) = KE(W^q \cdot W^\alpha; Q_2), \end{aligned} \tag{2.11}$$

where the second to the last inequality follows from (2.1). From (2.11) it follows that

$$E(W|Q_2)[E(W^\alpha|Q_{i_0})]^{p-1} \leqslant K. \tag{2.12}$$

Next we consider the different possibilities for i_0.

Case 1: $i_0 = 1$. If this happens, (2.12) together with (2.8) implies

$$E(W|Q_1)[E(W^\alpha|Q_1)]^{p-1} E(W|Q_2)[E(W^\alpha|Q_2)]^{p-1} \leqslant K \tag{2.13}$$

which combined with (2.4) gives (*).

Case 2: $i_0 = 3$. This means $|A_{r_3}^T * f| = |a_{23}^{r_3} + \eta_{r_3}| > \varepsilon/(m+1)$ on Q_3 for some $r_3 \in \{0, 1, \ldots, m\}$. But if we set $f = \chi_{Q_3}$, then $|A_{r_3} * f| = |a_{23}^{r_3} + \eta_{r_3}|$ on Q_2. Thus

$$\begin{aligned} E(W; Q_2) &\leqslant KE[|A_{r_3} * f|^p W; Q_2] \leqslant K\hat{E}|A_{r_3} * f|^p \\ &= K\hat{E}|f|^p = K\hat{E}(W; Q_3). \end{aligned} \tag{2.14}$$

Setting $f = \chi_{Q_1}$, we get $A_r * f = a_{i1}^r + \eta_r$ on Q_i. If we repeat the argument above, we find an $i_1 \neq 1$ and an $r_{i_1} \in \{0, 1, \ldots, m\}$ such that $|a_{i_1 1}^{r_{i_1}} + \eta_{r_1}| > \varepsilon/(m+1)$. If $i_1 = 2$, we get as above

$$E(W; Q_2) \leqslant KE(W; Q_1) \tag{2.15}$$

and (2.15) together with (2.8) implies (*). If $i_1 = 3$, we get

$$E(W; Q_3) \leqslant KE(W; Q_1). \tag{2.16}$$

From (2.16) and (2.14) we have

$$E(W; Q_2) \leqslant KE(W; Q_1) \tag{2.17}$$

and again (*) follows. Thus we may assume $i_1 = 4$. That is, $|a_{41}^{r_4} + \eta_{r_4}| > \varepsilon/(m+1)$ for some $r_4 \in \{0, 1, \ldots, m\}$. Arguing as above we find that

$$E(W; Q_4) \leqslant KE(W; Q_1). \tag{2.18}$$

Set $f = \chi_{Q_4} W$. Then

$$|A_{r_4}^T * f| = |A_{r_4}^T * \chi_{Q_4}| W_{Q_4} > \frac{\varepsilon}{m+1} W_{Q_4} \quad \text{on } Q_1. \tag{2.19}$$

The same argument as in (2.11) shows that (2.19) implies

$$E(W|Q_4)[E(W^\alpha|Q_1)]^{p-1} \leqslant K. \tag{2.20}$$

Putting (2.8), (2.12), and (2.20) together we get (recall we are in case $i_0 = 3$)

$$A_1(W) \cdot A_2(W) \cdot E(W|Q_4)[E(W^\alpha|Q_3)]^{p-1} \leqslant K^3 = K \tag{2.21}$$

where we have used the notation

$$A_j(W) = E(W|Q_j)[E(W^\alpha|Q_j)]^{p-1}.$$

Set $f = \chi_{Q_3}$. Then $A_r^T * f = a_{3i}^r + \eta_r$ on Q_i. As above, we find that there exists an $i_2 \neq 3$ such that $|a_{3i_2}^{r_{i_3}} + \eta_{r_3}| > \varepsilon/(m+1)$ for some $r_{i_3} \in \{0, 1, \ldots, m\}$. If $i_2 = 1$, we find that $E(W; Q_3) \leqslant KE(W; Q_1)$ which is just (2.16) and therefore we get (*). If $i_2 = 4$, a similar computation leads to

$$E(W^\alpha|Q_4) \leqslant KE(W^\alpha|Q_3) \tag{2.22}$$

and so (2.21) implies that $A_1(W) \cdot A_2(W) \cdot A_4(W) \leqslant K$ and (*) follows from (2.4). Finally, suppose $i_2 = 2$. This leads to

$$E(W^\alpha|Q_2) \leqslant KE(W^\alpha|Q_3) \tag{2.23}$$

which together with (2.12) (we are in case $i_0 = 3$) gives (*). The case $i_0 = 4$ follows similarly and (*) is proved and we have half of the theorem.

Next let us fix $n_0 > 1$. Let Q be the leftmost atom in $\mathcal{F}_{n_0}$. In other words, $Q = A(1, n_0)$. Let $A_r = (a_{ij}^r)$ be linear operators on V such that for every r, $a^r_{1j} = 0$ for all $j = 1, \ldots, d$. It is easy to see that $A_r * f = A_r * (f\chi_Q)$ on Q for every $f \in L^p(dx)$, $p > 1$. Setting $W = \chi_Q$ we have

$$\hat{E}|A_r * f|^p = E(|A_r * f|^p; Q) = E(|A_r * (f\chi_Q)|^p; Q)$$
$$\leqslant E|A_r * (f\chi_Q)|^p \leqslant K_r E|f\chi_Q|^p = K_r \hat{E}|f|^p$$

where the last inequality is due to the usual boundedness of martingale transforms. If we let K be the largest K_r we have $\hat{E}|A_r * f|^p \leqslant K\hat{E}|f|^p$ for all r. Clearly $W \notin A_p\{\mathscr{F}\}$ for any p. This completes the proof of the theorem.

We note that the example given above does not quite show that nondegeneracy is a necessary condition for the boundedness of martingale transforms to characterize A_p. To do this we need to show that for any collection of degenerate operators we can find a $W \notin A_p\{\mathscr{F}\}$ with the martingale transforms bounded on $L^p(Wdx)$. The example, however, comes very close to doing this. For suppose we have one operator A which is degenerate. Then we can write $A = \lambda I + D$ where λ is a real number, I is the matrix representing the identity and D is a matrix with a zero row or a zero column. From (2.1) we may assume that D has a zero row. A minor modification of the above example (depending on which row of D is zero) gives a $W \notin A_p\{\mathscr{F}\}$ for which $\hat{E}|D * f|^p \leqslant K\hat{E}|f|^p$. Since $A * f = \lambda(f - Ef) + D * f$, we see that if $Ef = 0$ we have $\hat{E}|A * f|^p \leqslant K\hat{E}|f|^p$ or, if we apply this to $f - Ef$, we get $\hat{E}|A * f|^p \leqslant K\hat{E}|f - Ef|^p$. It is clear, though, that to have $\hat{E}|f - Ef|^p \leqslant K\hat{E}|f|^p$, W cannot vanish on a set of positive measure and our example above does. So we cannot conclude that $\hat{E}|A * f|^p \leqslant K\hat{E}|f|^p$.

Martingale transforms in this setting were introduced by S. Janson [7] who showed that the Fefferman–Stein decomposition of BMO functions in terms of Riesz transforms has an analogue for the d-adic BMO in terms of these operators. Furthermore, Janson gave a simple necessary and sufficient condition for this decomposition to hold: the matrices should have no common real eigenvector. An important observation we wish to make is that our nondegeneracy condition is weaker than the Janson condition for the characterization of BMO. For example, it is easy to write down a symmetric matrix which is nondegenerate. The fact that the condition for the characterization of A_p-weights should be weaker than that for the characterization of BMO, should not be surprising. It is easy to show that in $\boldsymbol{R}^n$ the boundedness of any Riesz transform on $L^p(Wdx)$ implies that W is an A_p-function. On the other hand, it is shown in [7] that to characterize BMO we must have all the Riesz transforms.

References

[1] A. Bonami et D. Lepingle, *Fonction maximale et variation quadratique des martingales en présence d'un poids*, in: Sém. de Probabilités XIII, Lecture Notes in. Math. 721, Springer, 1979, 294–306.

[2] D. L. Burkholder, *A sharp inequality for martingale transforms*, Ann. Probab. 7 (1979), 858–863.

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

[4] C. Doléans-Dade et P. A. Meyer, *Inégalités de normes avec poids*, in: Sém. de Probabilités XIII, Lecture Notes in Math. 721, Springer, 1979, 313–331.

[5] 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.

[6] M. Izumisawa and N. Kazamaki, *Weighted norm inequalities for martingales*, Tôhoku Math. J. 29 (1977), 115–124.

[7] S. Janson, *Characterizations of H^1 by singular integral transforms on martingales and $\boldsymbol{R}^n$*, Math. Scand. 41 (1977), 140–152.

[8] Long Jui-Lin, *Martingales régulières et Φ-inégalités avec poids entre f^*, $S(f)$ et $\sigma(f)$*, C. R. Acad. Sci. Paris Sér. A 291 (1980), 31–34.

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

[10] A. Uchiyama, *Weighted functions on probability spaces*, Tôhoku Math. J. 30 (1978), 463–470.

DEPARTMENT OF MATHEMATICS 253-37
CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California 91125, U.S.A.

Received May 21, 1985
Revised version November 11, 1985 (2060)