A dominated ergodic estimate for $L_p$ spaces with weights

by

E. ATENCIA and A. DE LA TORRE (Malaga)

Abstract. In this note we characterize those positive functions $w$ such that the ergodic maximal function associated to an invertible, measure preserving ergodic transformation on a probability space is a bounded operator in $L_p(wd\mu)$.

1. Introduction. Let $(X, \mathcal{B}, \mu)$ be a non-atomic probability space and let $T: X \to X$ be an ergodic, invertible measure preserving transformation.

For each pair of non-negative integers $n$, $m$ we define the operator $T_{n,m}$, acting on measurable functions, as

$$T_{n,m}f(x) = (n+m+1)^{-1} \sum_{k=m}^{n} |f(T^k x)|.$$

It is well known that in order to study the a.e. convergence of the averages $T_{n,m}$, it is enough to prove a Dominated Ergodic Estimate (D.E.E.) with respect to the measure $\mu$, i.e. if

$$f^*(x) = \sup_{n,m \geq 0} T_{n,m}f(x),$$

then there exists a constant, namely $p/(p-1)$, such that

$$\|f^*\|_p \leq \frac{p}{p-1} \|f\|_p$$

for all $f \in L_p(d\mu)$

which certainly holds for $p > 1$ [8].

Our aim is to study the a.e. convergence of $T_{n,m}f$ but with respect to another measure $wd\mu$ where $w$ is a positive integrable function. We are thus led to try and characterize those positive functions $w$ such that the D.E.E. holds but with respect to the measure $wd\mu$. Let us fix $p > 1$. We will say that $T$ satisfies the D.E.E. with respect to the weight $w$ if

$$\frac{1}{k} \int f^p w d\mu \leq C_p \frac{1}{k} \int |f|^p w d\mu$$

for all $f$ in $L_p(wd\mu)$. 

(1.1)
Our main result is given by the following:

**Theorem.** In the above situation (1.1) holds if and only if \( w \) satisfies the condition:

\[
(A'_k) \quad \text{There exists a constant } M \text{ such that for a.e. } x
\]

\[
k^{-1} \sum_{j=0}^{k-1} w(T^j x)(k^{-1} \sum_{j=0}^{k-1} w(T^j x))^{1/(p-1)} \leq M
\]

for all positive integers \( k \).

Condition \((A'_k)\) is nothing but the condition in Theorem 10 of [2] with a constant independent of \( a \) and the natural analogues of the Markov property for the Hardy-Littlewood Maximal Operator [4].

Observe that if \( w \) satisfies \((A'_k)\), we have a D.E.E. for \( T \) as an operator from \( L_p(wd\mu) \) into \( L_p(wd\mu) \). \( T \) is obviously a positive operator but it is not, in general, a contraction in \( L_p(wd\mu) \) and its powers do not form, in general, a uniformly bounded group of positive operators, i.e., we obtain a D.E.E. for an operator \( T \) which, even though it separates supports, is not power-bounded as in [3].

2. Main results. In this section we will prove our result using the ideas in [2] adapted to our situation.

Our main tool will be the idea of “ergodic rectangle”.

**Definition.** Let \( B \) be a subset of \( X \) with positive measure and \( k \) a positive integer such that

\[
(T^kB \cap T^jB) = 0, \quad i \neq j, \quad 0 \leq i, j \leq k-1.
\]

Then the set \( B = \bigcup_{i=0}^{k-1} T^iB \) will be called an (ergodic) rectangle with base \( B \) and length \( k \).

Obviously \( \mu(B) = k\mu(B) \).

In the proof of the theorem we will need the following two results:

(2.1) **Proposition.** Let \( k \) be a positive integer and let \( A \subset X \) be a subset of positive measure. Then there exists \( B \subset A \) such that \( B \) is base of a rectangle of length \( k \).

(2.2) **Lemma.** \( X \) can be written as a countable union of bases of rectangles of length \( k \).

**Proof of the proposition.** First we will consider the case \( k = 2 \). We may assume \( \mu(A) \approx 1 \). Since \( T \) is ergodic, \( A \) is not invariant. If \( \mu(A \cap TA) = 0 \), then we choose \( B = A \) and we are done.

If \( \mu(A \cap TA) > 0 \), then \( \mu(A - (A \cap TA)) > 0 \) since otherwise \( A \) would be invariant. So now we pick \( B = A - (A \cap TA) \); obviously \( \mu(B) > 0 \), and \( B \cap TB = \emptyset \).

The general case follows by applying the same method.

**Proof of the lemma.** Let \( \mathcal{B}_2 = \{ B \subset X : B \text{ is base of a rectangle of length } k \} \). Because of the proposition, \( \mathcal{B}_2 \) is not empty. Let \( \eta_1 = \sup_{B \in \mathcal{B}_2} \mu(B) \).

Clearly \( 0 < \eta_1 \leq 1 \). We pick \( B^1 \in \mathcal{B}_2 \), \( \mu(B^1) > \eta_1/2 \). Let

\[
\mathcal{B}_1 = \{ B \in \mathcal{B}_2 : B \cap B^1 = \emptyset \}, \quad \eta_1 = \sup_{B \in \mathcal{B}_1} \mu(B),
\]

and choose \( B^2 \in \mathcal{B}_1 \), \( \mu(B^2) > \eta_1/2 \). We proceed by induction and define

\[
\mathcal{B}_n = \{ B \in \mathcal{B}_1 : B \cap \bigcup_{i=1}^{n-1} B^i = \emptyset \}, \quad \eta_n = \sup_{B \in \mathcal{B}_n} \mu(B)
\]

and choose \( B^n \in \mathcal{B}_n \), \( \mu(B^n) > \eta_n/2 \).

If for some \( n \) \( \mathcal{B}_n \) is empty, then \( X = \bigcup_{i=1}^{n-1} B^i \) a.e. Indeed, if \( X = \bigcup_{i=1}^{n-1} B^i \) is a.e. and \( \mu(A) > 0 \), then by the proposition there is \( B \subset A \), \( B \in \mathcal{B}_1 \), and obviously \( B \cap \bigcup_{i=1}^{n-1} B^i = \emptyset \) against \( \mathcal{B}_n \) being empty.

If no \( \mathcal{B}_n \) is empty, we obtain an infinite pairwise disjoint sequence \( B^1, B^2, \ldots, B^n, \ldots \) and we claim that

\[
X = \bigcup_{i=1}^\infty B^i.
\]

Let us prove it. First of all note that \( \lim \mu(B^n) = 0 \) since the sets are disjoint and \( \mu(X) \) is finite. If \( X = \bigcup_{i=1}^\infty B^n = A \) and \( \mu(A) > 0 \), we choose \( B \in \mathcal{B}_1 \), \( B \subset A \), \( \mu(B) = \mu \geq 0 \). Then there is \( n_0 \) such that \( \mu(B_{n_0}) < \mu/3 \) and observe that \( B \in \mathcal{B}_{n_0} \) which means \( \mu(B_{n_0}) > \mu/3 \) by the method of choosing \( B_{n_0} \) it should be \( \mu(B_{n_0}) > \mu/2 \) against \( \mu(B_{n_0}) < \mu/3 \).

Condition (1.1) implies \( w \) satisfies \((A'_k)\); let \( k \) be a non-negative integer and let us fix a rectangle with base \( B \) and length \( k \). For each integer \( n \) we consider the subset of \( B \)

\[
B_n = \{ x \in B : \mu(B_n x) > \mu/3 \}.
\]

Clearly \( B = \bigcup_{n=1}^\infty B_n \).

Let us fix \( n \) and let \( A \) be an arbitrary measurable subset of \( B_n \) with \( \mu(A) > 0 \). Let \( \mathcal{B} \) be the rectangle with base \( A \) and length \( k \). From the definition of our maximal operator it is obvious that

\[
(2.3) \quad (w^{-1/(p-1)} A)^\ast(T^j x) \geq k^{-1} \sum_{m=0}^{k-1} w^{-1/(p-1)}(T^j x) \geq 2^n, \quad x \in A, \quad 0 \leq j < k.
\]
The last inequality on the right holds since \( a \in A \subset B_a \). Raising to the power \( p \), multiplying by \( w(T^ax) \), and integrating over \( A \) we obtain
\[
\int_A \left( \int w^{-1/p-1} \right) \mu \left( T^ax \right) w(T^ax) \, d\mu \geq 2^{np} \int_A w(T^ax) \, d\mu.
\]
Adding up in \( j \) from 0 to \( k-1 \) and keeping in mind that \( \mu(T^jA \cap T^kA) = 0 \), \( 0 \leq i, j \leq k-1 \), we have
\[
2^{np} \int \frac{w(y)}{\lambda} \, d\mu \leq \int \left( \int w^{-1/(p-1)}(\lambda) \right) \frac{w(y) \, d\mu}{\lambda} \, d\mu.
\]
But using (1.1) the last term is majorized by
\[
\sum_{\lambda} \left( \int w^{-1/(p-1)}(\lambda) \right) \frac{w(y) \, d\mu}{\lambda} = \sum_{\lambda} \int w^{-1/(p-1)} \, d\mu,
\]
I.e.
\[
2^{np} \int \frac{w(y)}{\lambda} \, d\mu \leq \sum_{\lambda} \int w^{-1/(p-1)} \, d\mu
\]
and raising to the power \( p \) and using (2.5) we get
\[
\mu(A)^{-1} \int \left( \int w^{-1/(p-1)}(\lambda) \right) \frac{w(y) \, d\mu}{\lambda} \, d\mu \leq 2^{np+1}
\]
and raising to the power \( p \) and using (2.5) we get
\[
\mu(A)^{-1} \int \left( \int w^{-1/(p-1)}(\lambda) \right) \frac{w(y) \, d\mu}{\lambda} \, d\mu \leq 2^{np+1}
\]
which can be written as
\[
(\lambda) \mu(\lambda)^{-1} \int \frac{w(y) \, d\mu}{\lambda} \, d\mu \leq 2^{np+1}.
\]
We call it \( A_\lambda \) because it looks like condition \( A_\lambda \) in [4] but with the special rectangles \( B_\lambda \) instead of the cubes of the classical case.

Write \( A_\lambda \) as
\[
\mu(A)^{-1} \int \frac{1}{\lambda} \sum_{\lambda} w(T^ax) \, d\mu \mu(A)^{-1} \int \frac{1}{\lambda} \sum_{\lambda} w^{-1/(p-1)}(T^ax) \, d\mu \, d\mu \leq C_p 2^{np}.
\]
Since this holds for every \( A \), arbitrary measurable subset of positive measure of \( B_a \), we easily obtain \( A_\lambda \) for almost all \( x \) in \( B_a \). A straightforward application of the proposition and lemma gives \( A_\lambda \).

Before proving the converse we will state some results that will be needed in the proof. These results are a discrete version of the Calderón–Zygmund decomposition [9] and of some results in [1]. The proof follows the same pattern as in [1] and we will include it only to make the article self-contained.

**Calderón–Zygmund Decomposition.** Let us fix the integers 0, 1, 2, ..., \( k-1 \) and let \( \lambda \) be a real number such that
\[
\lambda > k^{-1} \sum_{\lambda} w(T^ax)
\]
where \( x \) is a fixed point of \( X \). Then for the set of integers 0, 1, 2, ..., \( k-1 \) we can choose a (possibly empty) family of disjoint subsets \( I_1, ..., I_l \) each of them made up of consecutive integers and such that the following holds:

(a) For each \( I_i \), \( i = 1, ..., l \)
\[
\lambda < \frac{1}{|I_i|} \sum_{\lambda} w(T^ax) \leq 3\lambda
\]
where \( |I_i| \) denotes the number of integers in \( I_i \).

(b) If \( j \not\in \bigcup_{i=1}^{l} I_i \), \( 0 < j < k-1 \), then \( w(T^ax) < \lambda \).

Proof. Let us call a set of consecutive integers an interval. Split 0, 1, ..., \( k-1 \) into two disjoint intervals \( I_1, I_2 \) where \( I_1 = 0, 1, ..., \lfloor (k-1)/2 \rfloor \). Now consider
\[
\frac{1}{|I_i|} \sum_{\lambda} w(T^ax), \quad i = 1, 2
\]
If this average is bigger than \( \lambda \), we select this interval and we have
\[
\frac{1}{|I_i|} \sum_{\lambda} w(T^ax) \leq \frac{k}{|I_i|} \frac{1}{k} \sum_{\lambda} w(T^ax) \leq 3\lambda.
\]
If this average is not bigger than \( \lambda \), we repeat the process. This process will finish in a finite number of steps. The chosen intervals satisfy (a) and if an integer \( r \) is left out, then obviously

\[
w(T^r x) \leq \lambda.
\]

In what follows we will often use for the averages the notation established in the introduction. In particular remember that

\[
k^{-1} \sum_{i=0}^{k-1} w(T^i x) = T_{k-1} \cdot w(x).
\]

Lemma. Let \( w \) satisfy \( A' \); then there exist positive constants \( a, \beta \) depending only on the constant \( M \) of condition \( A'_p \) such that if

\[
E = \{ i : 0 \leq i < k-1 : w(T^i x) > \beta k^{-1} \sum_{i=0}^{k-1} w(T^i x) \}
\]

then \( \# E > a k \) (\( \# E \) is the number of integers in \( E \)).

Proof. Observe that for any positive \( \beta \) if \( E' = (0, 1, \ldots, k-1) - E \), then

\[
\beta^{-1} k^{-1} \# E' \leq T_{k-1} \cdot w(x) \left( k^{-1} \sum_{i=0}^{k-1} w(T^i x) \right)^{p-1}
\]

this is because in \( E' \) is \( w(T^i x) \geq \beta k^{-1} \sum_{i=0}^{k-1} w(T^i x) \). But the last term in (2.7) is, obviously, dominated by

\[
T_{k-1} \cdot w(x) \left( k^{-1} \sum_{i=0}^{k-1} w(T^i x) \right)^{p-1} < M \quad \text{since} \quad A'_p.
\]

Choose \( \beta < M^{-1} \), \( a = 1 - (M\beta)^{p-1} \) and the lemma is proved.

Notes. If instead of \( 0 \leq i < k-1 \) we start with any other interval \( I \), then we have

\[
\# I \left( i \in I : w(T^i x) > \beta |I|^{-1} \sum_{i=0}^{k-1} w(T^i x) \right) > a |I|.
\]

The Calderon–Zygmund decomposition and the preceding lemma allow us to prove, in our context, the "reverse Hölder inequality".

Lemma. Let \( w \) satisfy \( A'_p \), \( 1 < p < \infty \); then there exist positive constants \( C, \delta \) such that

\[
k^{-1} \sum_{i=0}^{k-1} w(T^i x) |T^{i+1} x|^{1+p} \leq C k^{-1} \sum_{i=0}^{k-1} w(T^i x)
\]

for every \( k \) and \( x \).

\[\text{Proof. Let } \lambda \text{ be a positive number such that } \lambda > T_{k-1} \cdot w(x). \]

We want to estimate \( \sum w(T^i x) \) extended to those \( i \)'s, \( 0 \leq i < k \), where \( w(T^i x) > \lambda \). Using the Calderon–Zygmund decomposition for this \( \lambda \), we have a family of disjoint intervals \( I_j \) satisfying (a) and (b) of the said decomposition, so

\[
A(\lambda) = \{ i : 0 \leq i < k : w(T^i x) > \lambda \} = \bigcup_j I_j.
\]

Now

\[
\sum_{i=0}^{k-1} w(T^i x) \leq \sum_{i=0}^{k-1} \sum_{i=0}^{k-1} w(T^i x) \leq M \sum_{i=0}^{k-1} |I_i| \leq M \sum_{i=0}^{k-1} \sum_{i=0}^{k-1} w(T^i x) > \beta k^{-1} \sum_{i=0}^{k-1} w(T^i x)
\]

this is because in \( E' \) is \( w(T^i x) \geq \beta k^{-1} \sum_{i=0}^{k-1} w(T^i x) \). But the last term in (2.7) is, obviously, dominated by

\[
T_{k-1} \cdot w(x) \left( k^{-1} \sum_{i=0}^{k-1} w(T^i x) \right)^{p-1} < M \quad \text{since} \quad A'_p.
\]

Choose \( \beta < M^{-1} \), \( a = 1 - (M\beta)^{p-1} \) and the lemma is proved.

Notes. If instead of \( 0 \leq i < k-1 \) we start with any other interval \( I \), then we have

\[
\# I \left( i \in I : w(T^i x) > \beta |I|^{-1} \sum_{i=0}^{k-1} w(T^i x) \right) > a |I|.
\]

The Calderon–Zygmund decomposition and the preceding lemma allow us to prove, in our context, the "reverse Hölder inequality".

Lemma. Let \( w \) satisfy \( A'_p \), \( 1 < p < \infty \); then there exist positive constants \( C, \delta \) such that

\[
k^{-1} \sum_{i=0}^{k-1} w(T^i x) |T^{i+1} x|^{1+p} \leq C k^{-1} \sum_{i=0}^{k-1} w(T^i x)
\]

for every \( k \) and \( x \).

\[\text{Proof. Let } \lambda \text{ be a positive number such that } \lambda > T_{k-1} \cdot w(x). \]

We want to estimate \( \sum w(T^i x) \) extended to those \( i \)'s, \( 0 \leq i < k \), where \( w(T^i x) > \lambda \). Using the Calderon–Zygmund decomposition for this \( \lambda \), we have a family of disjoint intervals \( I_j \) satisfying (a) and (b) of the said decomposition, so

\[
A(\lambda) = \{ i : 0 \leq i < k : w(T^i x) > \lambda \} = \bigcup_j I_j.
\]

Now

\[
\sum_{i=0}^{k-1} w(T^i x) \leq \sum_{i=0}^{k-1} \sum_{i=0}^{k-1} w(T^i x) \leq M \sum_{i=0}^{k-1} |I_i| \leq M \sum_{i=0}^{k-1} \sum_{i=0}^{k-1} w(T^i x) > \beta k^{-1} \sum_{i=0}^{k-1} w(T^i x)
\]

this is because in \( E' \) is \( w(T^i x) \geq \beta k^{-1} \sum_{i=0}^{k-1} w(T^i x) \). But the last term in (2.7) is, obviously, dominated by

\[
T_{k-1} \cdot w(x) \left( k^{-1} \sum_{i=0}^{k-1} w(T^i x) \right)^{p-1} < M \quad \text{since} \quad A'_p.
\]

Choose \( \beta < M^{-1} \), \( a = 1 - (M\beta)^{p-1} \) and the lemma is proved.

Notes. If instead of \( 0 \leq i < k-1 \) we start with any other interval \( I \), then we have

\[
\# I \left( i \in I : w(T^i x) > \beta |I|^{-1} \sum_{i=0}^{k-1} w(T^i x) \right) > a |I|.
\]

The Calderon–Zygmund decomposition and the preceding lemma allow us to prove, in our context, the "reverse Hölder inequality".

Lemma. Let \( w \) satisfy \( A'_p \), \( 1 < p < \infty \); then there exist positive constants \( C, \delta \) such that

\[
k^{-1} \sum_{i=0}^{k-1} w(T^i x) |T^{i+1} x|^{1+p} \leq C k^{-1} \sum_{i=0}^{k-1} w(T^i x)
\]

for every \( k \) and \( x \).

\[\text{Proof. Let } \lambda \text{ be a positive number such that } \lambda > T_{k-1} \cdot w(x). \]

We want to estimate \( \sum w(T^i x) \) extended to those \( i \)'s, \( 0 \leq i < k \), where \( w(T^i x) > \lambda \). Using the Calderon–Zygmund decomposition for this \( \lambda \), we have a family of disjoint intervals \( I_j \) satisfying (a) and (b) of the said decomposition, so

\[
A(\lambda) = \{ i : 0 \leq i < k : w(T^i x) > \lambda \} = \bigcup_j I_j.
\]

Now

\[
\sum_{i=0}^{k-1} w(T^i x) \leq \sum_{i=0}^{k-1} \sum_{i=0}^{k-1} w(T^i x) \leq M \sum_{i=0}^{k-1} |I_i| \leq M \sum_{i=0}^{k-1} \sum_{i=0}^{k-1} w(T^i x) > \beta k^{-1} \sum_{i=0}^{k-1} w(T^i x)
\]

this is because in \( E' \) is \( w(T^i x) \geq \beta k^{-1} \sum_{i=0}^{k-1} w(T^i x) \). But the last term in (2.7) is, obviously, dominated by

\[
T_{k-1} \cdot w(x) \left( k^{-1} \sum_{i=0}^{k-1} w(T^i x) \right)^{p-1} < M \quad \text{since} \quad A'_p.
\]

Choose \( \beta < M^{-1} \), \( a = 1 - (M\beta)^{p-1} \) and the lemma is proved.

Notes. If instead of \( 0 \leq i < k-1 \) we start with any other interval \( I \), then we have

\[
\# I \left( i \in I : w(T^i x) > \beta |I|^{-1} \sum_{i=0}^{k-1} w(T^i x) \right) > a |I|.
\]}
is not less than
\[ \delta^{-1} \sum_{k=1}^{k-1} w(T^k x) \left[ |w(T^k x)|^p - |T_{k-1} w(x)|^p \right]; \]
so we obtain
\[ \delta^{-1} - C(1 + \delta^{-1}) \sum_{k=1}^{k-1} w(T^k x) \delta^{-1} \leq \left( \delta^{-1} \sum_{k=1}^{k-1} w(T^k x) \right)^{1+\delta}\]
and the lemma follows by choosing \( \delta \) small enough to make
\[ \delta^{-1} - C(1 + \delta^{-1}) > 0. \]

(2.10) Lemma. Let \( w \) satisfy \( A^p \), then there exists \( \epsilon > 0 \) so that \( w \) satisfy \( A^{p-\epsilon} \).

Proof. Check first that if \( w \) satisfies \( A^p \), then \( v = w^{-1/p} \) satisfies \( A^p \) with \( p^{-1} + q^{-1} = 1 \). Applying now the preceding lemma to \( v \) we have for some \( \delta > 0 \)
\[ |T_{k-1}^k v(x)|^{1/1+\delta} \leq C T_{k-1}^k v(x) \]
replacing \( v \) by \( w^{-1/p} \) and taking \( \epsilon = (p-1)/\delta(1 + \delta^{-1}) \) we have
\[ \sum_{k=1}^{k-1} w(T^k x) \left( \delta^{-1} \sum_{k=1}^{k-1} w(T^k x) \right)^{1/1+\delta} \leq C^{1-\delta} \sum_{k=1}^{k-1} w(T^k x) \left( \delta^{-1} \sum_{k=1}^{k-1} w(T^k x) \right)^{1/1+\delta} \leq C^{1-\delta} M. \]

The following maximal function appears in a natural way associated to the weight \( w \)
\[ M_w f(x) = \sup_{n \in \mathbb{N}, \|w\| < \infty} \frac{\sum_{k=1}^{n} |f(T^k x)| w(T^k x)}{\sum_{k=1}^{n} w(T^k x)}. \]

As we will see this maximal function controls \( f^* \). Indeed, if \( p^{-1} + q^{-1} = 1 \), we have
\[ (n+m+1)^{-1} \sum_{k=n}^{m} |f(T^k x)| \]
\[ = (n+m+1)^{-1} \sum_{k=n}^{m} |f(T^k x)| w^{1/p}(T^k x) w^{-1/p}(T^k x) \]
\[ \leq (n+m+1)^{-1} \sum_{k=n}^{m} |f(T^k x)| w(T^k x) \left( (n+m+1)^{-1} \sum_{k=n}^{m} w^{-1/p}(T^k x) \right)^{1/p} \]
\[ \leq M \left( \frac{(n+m+1)^{-1} \sum_{k=n}^{m} |f(T^k x)| w(T^k x)}{\sum_{k=n}^{m} w(T^k x)} \right)^{1/p} \leq M(M_w f(x))^{1/p}. \]

A dominated ergodic estimate
\[ \left( f^*(x) \right)^p w(x) d\mu \leq \int X f^*(x)^p w(x) d\mu \]
where \( p/q > 1 \); so if we prove that the maximal operator \( M_w f \) is bounded in \( L^p(w d\mu) \) for all \( r > 1 \) we will have
\[ \int X f^*(x)^p w(x) d\mu \leq C \int X f^*(x)^p w(x) d\mu \]
and we will be done. Since \( M_w f \) is obviously bounded in \( L^p \), it will be enough to prove weak type \((1, 1)\) and use the Marcinkiewicz interpolation theorem.

(2.13) Theorem. The maximal operator with weight defined by
\[ M_w f(x) = \sup_{n \in \mathbb{N}} \frac{\sum_{k=1}^{n} |f(T^k x)| w(T^k x)}{\sum_{k=1}^{n} w(T^k x)} \]
is of weak type \((1, 1)\) with respect to the measure \( w d\mu \).

Proof. We may assume \( f \) is non-negative. Let \( \lambda \) be a positive number bigger than
\[ \epsilon = \frac{\int \epsilon f(s) d\mu}{\int \epsilon w d\mu} \]
and let
\[ O_0 = \{ s \in X ; M_w f(s) > \lambda \}. \]
The set \( O_0 \) is, clearly, measurable. For any \( x \in X \) we consider the orbit of \( x \) in \( O_0 \), that we denote by \( J_x \), i.e.
\[ J_x = \{ T^k x \in O_0, k \in \mathbb{Z} \}. \]
We associate, in a natural way, to the orbit of \( x \) in \( O_0, J_x \), the subset of the integers given by
\[ \{ k : T^k x \in O_0 \} \]
that we can express as a countable union of disjoint intervals \( \bigcup I^k \).
Let prove that, for almost all \( x \), no \( \mathcal{H} \) has infinite number of integers. The individual ergodic theorem tells us that

\[
\lim_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-1} \mathcal{V}(T^j x) \mathcal{W}(T^j x) = \frac{1}{\lambda} \int \mathcal{V}(x) \mathcal{W}(x) \, d\mu \quad \text{a.e.}
\]

If for some \( i \), \( \mathcal{H} = \{ i, i+1, i+2, \ldots \} \), then, by the above-mentioned theorem, we have

\[
\lim_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-1} \mathcal{V}(T^j x) \mathcal{W}(T^j x) = \mathcal{V}(x) \mathcal{W}(x) \quad (\text{a.e.})
\]

Thus, being \( \lambda > \mathcal{V} \), there exists a positive integer \( K \) such that

\[
\sum_{j=0}^{K-1} \mathcal{V}(T^j x) \mathcal{W}(T^j x) < \lambda \sum_{j=0}^{K-1} \mathcal{W}(T^j x) \quad (K > K).
\]

Clearly, by the definitions of \( \mathcal{R} \) and \( \mathcal{O} \), there exists \( r \) verifying

\[
\sum_{j=0}^{r-1} \mathcal{V}(T^j x) \mathcal{W}(T^j x) > \sum_{j=0}^{r-1} \mathcal{W}(T^j x) \quad (T^r x \in \mathcal{O}_1)
\]

where, by (2.15), \( r < K \). Considering now \( T^{i+1} x \) that belongs to \( \mathcal{O}_1 \) there exists \( r_1 \gg r+1 \) such that

\[
\sum_{j=0}^{r_1+1} \mathcal{V}(T^j x) \mathcal{W}(T^j x) > \lambda \sum_{j=0}^{r_1+1} \mathcal{W}(T^j x)
\]

and applying the same process we obtain a sum of the type

\[
\sum_{j=0}^{s} \mathcal{V}(T^j x) \mathcal{W}(T^j x) > \lambda \sum_{j=0}^{s} \mathcal{W}(T^j x)
\]

where \( s > K \), in contradiction with (2.15). Therefore, there are not intervals of infinite length.

Choose now an interval \( I_i \), namely

\[
I_i = \{ i, i+1, \ldots, i+m \}
\]

which means that \( T^ix, \ldots, T^{i+m}x \in \mathcal{O}_i \) and \( T^{i+m+1}x \notin \mathcal{O}_i \). The aim is to prove

\[
\sum_{j=0}^{m} \mathcal{V}(T^j x) \mathcal{W}(T^j x) \geq \lambda \sum_{j=0}^{m} \mathcal{W}(T^j x)
\]

If this were not the case, then for every positive integer \( r \) we would have

\[
\sum_{j=0}^{m} \mathcal{V}(T^j x) \mathcal{W}(T^j x) = \sum_{j=0}^{m} \mathcal{V}(T^j x) \mathcal{W}(T^j x) + \sum_{j=m+1}^{m+r} \mathcal{V}(T^j x) \mathcal{W}(T^j x)
\]

\[
< \lambda \sum_{j=0}^{m} \mathcal{W}(T^j x) + \lambda \sum_{j=0}^{m} \mathcal{W}(T^j x)
\]

where we have used that \( T^{i+m+1}x \notin \mathcal{O}_i \), so for every positive integer \( r \) we would have

\[
\sum_{j=0}^{m} \mathcal{V}(T^j x) \mathcal{W}(T^j x) < \lambda \sum_{j=0}^{m} \mathcal{W}(T^j x).
\]

On the other hand, for some \( h \), clearly not bigger than \( m \) it should be

\[
\sum_{j=0}^{h} \mathcal{V}(T^j x) \mathcal{W}(T^j x) > \lambda \sum_{j=0}^{h} \mathcal{W}(T^j x),
\]

if we call

\[
\mathcal{h} = \max\{ \mathcal{h} : \mathcal{V}(T^j x) \mathcal{W}(T^j x) > \lambda \sum_{j=0}^{h} \mathcal{W}(T^j x) \}
\]

then we claim that \( \mathcal{h} < m \).

Suppose that \( \mathcal{h} < m \). Then \( T^{i+1}x \notin \mathcal{O}_i \), fact that implies the existence of a \( i \gg \mathcal{h} + 1 \) such that

\[
\sum_{j=0}^{\mathcal{h}+1} \mathcal{V}(T^j x) \mathcal{W}(T^j x) > \lambda \sum_{j=0}^{\mathcal{h}+1} \mathcal{W}(T^j x),
\]

and

\[
\sum_{j=0}^{\mathcal{h}} \mathcal{V}(T^j x) \mathcal{W}(T^j x) > \lambda \sum_{j=0}^{\mathcal{h}} \mathcal{W}(T^j x)
\]

adding up (2.17) and (2.18) we obtain a contradiction with the assumption that \( \mathcal{h} \) was the considered maximum. Therefore \( \mathcal{h} = m \), and the inequality (2.16) follows.

Call now

\[
B_i = \{ x \in \mathcal{O}_i : x, \ldots, T^{i+1}x \in \mathcal{O}_i, T^{i+2}x \notin \mathcal{O}_i \}
\]

\( B_i \) is clearly measurable.

Let \( B_i \) be defined by

\[
B_i = B_i \cup TB_i \cup \ldots \cup T^{i-1}B_i.
\]

It is obvious that \( T^m B_i \cap T^m B_i = \varnothing \), \( 0 \leq m, \), \( m \leq i-1, \), \( m \neq m \) and that

\[
O_i = \bigcup B_i.
\]

Consider now \( x \in B_i \):

\[
\sum_{j=0}^{i-1} \mathcal{W}(T^j x) \leq \lambda^{-1} \sum_{j=0}^{i-1} \mathcal{V}(T^j x) \mathcal{W}(T^j x).
\]

Therefore

\[
\int \mathcal{W}(x) \, d\mu \leq \int \mathcal{V}(x) \mathcal{W}(x) \, d\mu.
\]
and summing over \( i \) we have:

\[
\int_{E_i} w \, d\mu \leq \lambda^{-1} \int_{E_i} f \, d\mu
\]

and the theorem is proved.

To prove that \( R_w \) is also of weak type \((1,1)\) we just observe that \( R_w f \) is dominated by \( R_w f - R_{w,1} f \) where \( R_w f(x) \) is what we called \( R_w f(x) \) while \( R_{w,1} f(x) \) is the corresponding operator defined as \( R_{w,1} f(x) \) but using \( T^{-1} \) instead of \( T \).

3. Final remarks. If \( w \) is constant, then clearly satisfies \( A_p' \) for any \( p \). But apart from this trivial case the natural question is if there exists a non-constant \( w \) satisfying \( A_p' \). In the case of the Hardy–Littlewood maximal function in \( B^p \) Stein provides [5] an example: \( |w|^p, -1 < p < 1 \). Unfortunately this does not make any sense in our context. There is another way of producing good weights. It is to find a function \( w \) such that \( w^\alpha \) is essentially dominated by \( Cw \).

In [7] it is proved that if \( f \) is any function in \( L^1(X) \) and we construct

\[
w = (f^\alpha)^{1/\alpha},
\]

then

(2.19) \[
w^\alpha(x) \leq C w(x)
\]

for any \( x \) where \( C \) is an universal constant.

Since \( T:a \leq w^\alpha(T^a) \), \( 0 \leq i \leq k-1 \), then by (2.19) we have

\[
T:a \leq C w(T^a), \quad 0 \leq i \leq k-1
\]

and this certainly implies \( A_p' \) for any \( p > 1 \) with \( M = C \). Observe also that \( A_p' \) implies easily that \( w(T^a) \leq C w(x) \) where \( C \) depends only on the constant in \( A_p' \). This means what we said in the introduction, that the operator

\[
f \rightarrow T f \quad (T f(x) = f(Tx))
\]

maps \( L^p(w \, d\mu) \) into \( L^p(w \, d\mu) \) but, since \( C \) is in general bigger than 1, this is neither a contraction nor a uniformly bounded group of operators.

References