The symbol ϱ denotes here a *defining function* for D, i.e. $\varrho \in C^{\infty}(\mathbf{R}^n)$, $D = \{x \in \mathbf{R}^n\colon \varrho(x) < 0\}$, $\operatorname{grad} \varrho \neq 0$ on ∂D.

Remark 2. All results of the present paper, excluding part (b) of Theorem 3, remain valid if the L^p, Sobolev and Bloch spaces of harmonic functions are replaced by the corresponding spaces of m-polyharmonic functions, i.e. functions u for which $\Delta^m u = 0$. The details will be given in a forthcoming paper *On duality and interpolation for spaces of polyharmonic functions*.

Added in proof (July 1987). Proposition 1 along with other results of this paper is valid in the case where D is any smooth bounded domain. Details will be given in our next paper.

References

[1] F. Beatrous and J. Burbea, *Sobolev spaces of holomorphic functions in the ball*, preprint, 1985.

[2] S. Bell, *A duality theorem for harmonic functions*, Michigan Math. J. 29 (1982), 123–128.

[3] J. Bergh and J. Löfström, *Interpolation Spaces. An Introduction*, Springer, 1976.

[4] S. G. Krein, Yu. I. Petunin and E. M. Semenov, *Interpolation of Linear Operators*, Nauka, Moscow 1978 (in Russian).

[5] E. Ligocka, *The Hölder duality for harmonic functions*, Studia Math. 84 (1986), 269–277.

[6] —, *The Bergman projection on harmonic functions*, ibid. 85 (1987), 229–246.

[7] —, *Estimates in Sobolev norms* $\|\cdot\|_p^s$ *for harmonic and holomorphic functions and interpolation between Sobolev and Hölder spaces of harmonic functions*, ibid. 86 (1987), 255–271.

INSTYTUT MATEMATYCZNY POLSKIEJ AKADEMII NAUK
INSTITUTE OF MATHEMATICS, POLISH ACADEMY OF SCIENCES
Śniadeckich 8, 00-950 Warszawa, Poland

Received December 3, 1985
Revised version April 30, 1986 (2118)

STUDIA MATHEMATICA, T. LXXXVII. (1987)

Weighted weak type inequalities for the ergodic maximal function and the pointwise ergodic theorem

by

F. J. MARTÍN-REYES and A. de la TORRE (Málaga)

Abstract. Let $(X, \mathcal{F}, \mu)$ be a σ-finite measure space and let $T\colon X \to X$ denote an invertible measure-preserving transformation. In this paper we characterize those pairs of positive functions u, v for which the maximal operator

$$Mf(x) = \sup_{k \geq 0} (k+1)^{-1} \sum_{i=0}^{k} |f(T^i x)|$$

is of weak type (1, 1) with respect to the measures $v\,d\mu$ and $u\,d\mu$. We also get a pointwise ergodic theorem for noninvertible T if $\mu(X) < \infty$. More precisely, we prove that $(k+1)^{-1} \sum_{i=0}^{k} f(T^i x)$ converges a.e. for every $f \in L^1(v\,d\mu)$ if and only if $\inf_{i \geq 0} v(T^i x) > 0$ a.e.

1. Introduction. Let $(X, \mathcal{F}, \mu)$ be a σ-finite measure space and let $T\colon X \to X$ denote an invertible measure-preserving transformation. For each pair of nonnegative integers, r and k, we consider the averages

$$T_{r,k} f(x) = (r+k+1)^{-1} \sum_{i=-r}^{k} f(T^i x)$$

where f is a measurable function. Associated to these averages we have the following maximal operators:

$$f^* = \sup_{r,k \geq 0} T_{r,k}|f|, \qquad Mf = \sup_{k \geq 0} T_{0,k}|f|.$$

The maximal ergodic theorem asserts that f^* and Mf satisfy weak type inequalities

$$\mu(\{x \in X\colon f^*(x) > \lambda\}) \leq 2\lambda^{-1} \int_X |f|\,d\mu,$$

$$\mu(\{x \in X\colon Mf(x) > \lambda\}) \leq \lambda^{-1} \int_X |f|\,d\mu$$

A.M.S. Subject classification (1980): Primary 28D05, Secondary 47A35.

Key words and phrases: maximal functions, ergodic maximal function, weighted inequalities, pointwise ergodic theorem.

for all $\lambda > 0$ and every f. In [4] the following weighted version of the inequality for f^* was proved: "Let u and v be nonnegative functions. There exists a constant $C > 0$ such that

$$\int_{\{x \in X:\ f^*(x) > \lambda\}} u\, d\mu \leqslant C\lambda^{-1} \int_X |f|\, v\, d\mu$$

if and only if $u^*(x) \leqslant Cv(x)$ a.e." (for a proof when T is ergodic, $(X, \mathscr{F}, \mu)$ a nonatomic probability space and $u = v$, see [1]). The purpose of this paper is to prove a similar result for Mf. We also get a characterization of those positive functions v for which the averages $T_{0,k} f$ converge a.e. for every $f \in L^1(v\, d\mu)$ if $\mu(X) < \infty$ and T is not necessarily invertible.

2. Weak type inequalities. Throughout this paper we shall consider two sets or two functions as equal if they agree up to a set of measure zero. As usual, C will denote a positive constant not necessarily the same at each occurrence. If c and d are integers with $c \leqslant d$ we will write $[c, d]$ for the set of integers j such that $c \leqslant j \leqslant d$, i.e., $[c, d]$ is an interval in the integers.

To prove our main result (Theorem (2.4)) we shall need the following definition and the following lemma.

(2.1) Definition. Let $(X, \mathscr{F}, \mu)$ be a σ-finite measure space and let T be a measurable transformation from X into itself. Suppose k is a nonnegative integer. A measurable set $B \subset X$ is the *base* (with respect to T) *of an ergodic rectangle* of length $k+1$ if $T^{-i} B \cap T^{-j} B = \emptyset$, $i \neq j$, $0 \leqslant i, j \leqslant k$. In such a case, the set $R = \bigcup_{i=0}^{k} T^{-i} B$ will be called an *ergodic rectangle* with base B and length $k+1$.

Remark. If T is invertible and T^{-1} is a measurable transformation then it is clear that B is a base of a rectangle of length $k+1$ with respect to T if and only if B is a base of a rectangle of length $k+1$ with respect to T^{-1}. Thus, for T invertible with T^{-1} measurable, Definition (2.1) is equivalent to the definitions which appear in [1] and [4].

(2.2) Lemma. *Let $(X, \mathscr{F}, \mu)$ be a σ-finite measure space and let T be a measure-preserving transformation from X into itself. Let Y be a measurable subset of X and let k be a nonnegative integer. Then there exists a countable family $\{B_i\}_{i=0}^{\infty}$ of sets of finite measure such that*

(i) $Y = \bigcup_{i=0}^{\infty} B_i$.

(ii) $B_i \cap B_j = \emptyset$ *if* $i \neq j$.

(iii) *For every i, B_i is the base of an ergodic rectangle of length $1+s(i)$ with $0 \leqslant s(i) \leqslant k$ such that if $s(i) < k$ then $T^{-1-s(i)} A = A$ for every measurable set $A \subset B_i$. Consequently, for every measurable set $A \subset B_i$*

$$\sum_{j=0}^{k} \chi_{T^{-j}A} \leqslant C(i) \sum_{j=0}^{s(i)} \chi_{T^{-j}A} \tag{2.3}$$

$$= C(i)\, \chi_{\bigcup_{j=0}^{s(i)} T^{-j}A} \leqslant 2 \sum_{j=0}^{k} \chi_{T^{-j}A}$$

where $C(i)$ is the least integer not smaller than $(k+1)(1+s(i))^{-1}$.

This lemma is similar to Lemma (2.10) in [4] where T is an invertible measure-preserving transformation. To prove Lemma (2.2), just look at the proof of Lemma (2.10) in [4] and write it with T^h instead of T^{-h}.

Now, we shall state and prove our theorem.

(2.4) Theorem. *Let $(X, \mathscr{F}, \mu)$ be a σ-finite measure space and let $T\colon X \to X$ denote an invertible measure-preserving transformation. Suppose u and v are nonnegative measurable functions on X. The following statements are equivalent:*

(a) *There is a constant $C > 0$ such that*

$$\int_{\{x \in X:\ Mf(x) > \lambda\}} u\, d\mu \leqslant C\lambda^{-1} \int_{\{x \in X:\ Mf(x) > \lambda\}} |f|\, v\, d\mu$$

for all $\lambda > 0$ and for every f.

(b) *There is a constant $C > 0$ such that*

$$\sup_{k \geqslant 0} \int_{\{x \in X:\ |T_{0,k} f(x)| > \lambda\}} u\, d\mu \leqslant C\lambda^{-1} \int_{\{x \in X:\ Mf(x) > \lambda\}} |f|\, u\, d\mu$$

for all $\lambda > 0$ and for every f.

(c) *There is a constant $C > 0$ such that for every f*

$$\sup_{k \geqslant 0} \int_X |T_{0,k} f|\, u\, d\mu \leqslant C \int_X |f|\, v\, d\mu.$$

(d) *There is a constant $C > 0$ such that $T_{0,k} u(x) \leqslant Cv(T^k x)$ for all $k \geqslant 0$ and almost all $x \in X$.*

Proof. We begin by proving that (a), (b) and (d) are equivalent. It is clear that (a) implies (b). Assume that (b) holds. Let us fix a nonnegative integer k. Then $X = \bigcup_{i=0}^{\infty} B_i$ where the sets B_i are generated by Lemma (2.2) for X and k. Let B_i be fixed and let A be any measurable subset of B_i. Let R denote the set $\bigcup_{j=0}^{s(i)} T^{-j} A$. If χ_A is the characteristic function of A and $y \in R$ we get

$$T_{0,s(i)}\, \chi_A(y) \geqslant (s(i)+1)^{-1}.$$

Thus, by (b), we have

$$\int_R u\,d\mu \leqslant C\big(s(i)+1\big)\int_A v\,d\mu.$$

Since the sets $A, \ldots, T^{-s(i)}A$ are pairwise disjoint, the last inequality is the same as

$$\sum_{j=0}^{s(i)} \int_{T^{-j}A} u\,d\mu \leqslant C\big(s(i)+1\big)\int_A v\,d\mu.$$

Then the properties of the sets B_i give

$$\text{(2.5)}\qquad \sum_{j=0}^{k} \int_{T^{-j}A} u\,d\mu = \int_X u \sum_{j=0}^{k} \chi_{T^{-j}A}\,d\mu \leqslant C(i)\int_X u \sum_{j=0}^{s(i)} \chi_{T^{-j}A}\,d\mu$$
$$\leqslant C(i)\,C\big(s(i)+1\big)\int_A v\,d\mu \leqslant 2C(k+1)\int_A v\,d\mu.$$

Thus, since T preserves the measure μ, we have

$$\sum_{j=0}^{k} \int_{T^{-k}A} u(T^{k-j}x)\,d\mu \leqslant 2C(k+1)\int_{T^{-k}A} v(T^k x)\,d\mu.$$

Since $T^{-k}A$ is any subset of $T^{-k}B_i$ and $X = \bigcup_{i=0}^{\infty} T^{-k}B_i$ this inequality shows that (d) follows from (b).

Suppose now that (d) holds. In order to prove (a) we shall need the following definition and the following lemmas.

(2.6) Definition. If a is a real-valued function on $\boldsymbol{Z}$ (the set of all integers) we define the maximal function $ma\colon \boldsymbol{Z}\to\boldsymbol{R}$ by

$$ma(j) = \sup_{k\geqslant 0}(k+1)^{-1}\sum_{i=0}^{k}|a(j+i)|.$$

(2.7) Lemma. *Let $a(0), a(1), \ldots, a(k)$ be nonnegative numbers. Then there exist $w(0), w(1), \ldots, w(k)$ such that*

(i) $0 \leqslant w(0) \leqslant w(1) \leqslant \ldots \leqslant w(k)$.

(ii) $w(i) \leqslant \max\{(r+1)^{-1}\sum_{j=0}^{r} a(i-j)\colon 0\leqslant r\leqslant i\}$ *for every i with* $0\leqslant i\leqslant k$.

(iii) $\sum_{j=0}^{k} a(j) \leqslant \sum_{j=\rho}^{k} w(j)$.

(2.8) Lemma. *Let $0\leqslant w(0)\leqslant w(1)\leqslant \ldots \leqslant w(k)$. Suppose that $a(0), a(1), \ldots, a(k)$ are nonnegative numbers such that*

$$0<\lambda<(k-i+1)^{-1}\sum_{j=i}^{k} a(j)$$

for every i with $0\leqslant i\leqslant k$. Then

$$\sum_{j=0}^{k} w(j) \leqslant \lambda^{-1}\sum_{j=0}^{k} a(j)\,w(j).$$

(2.9) Lemma. *If u and v are nonnegative functions on X and satisfy* (d) *then there is a constant $C>0$ such that*

$$\sum_{\{j\in\boldsymbol{Z}:\ ma(j)>\lambda\}} u(T^j x) \leqslant C\lambda^{-1}\sum_{\{j\in\boldsymbol{Z}:\ ma(j)>\lambda\}} |a(j)|\,v(T^j x)$$

for every real-valued function a on $\boldsymbol{Z}$, all $\lambda>0$ and almost all x in X.

Remark. Lemma (2.9) is the implication (d)$\Rightarrow$(a) in the case of the integers and it will follow from Lemma (2.8). Observe that (2.8) is nearly a particular case of (2.9).

Proof of Lemma (2.7). If $0\leqslant i\leqslant k$ we define

$$\tilde a(i) = \max\{(r+1)^{-1}\sum_{j=0}^{r} a(i-j)\colon 0\leqslant r\leqslant i\}$$

and $w(i) = \min\{\tilde a(j)\colon i\leqslant j\leqslant k\}$. Then it is clear that w satisfies (i) and (ii). Now we shall see that (iii) holds. Observe that, by the definition of w, it is possible to select intervals $J_1 = [0, i_1]$, $J_2 = [i_1+1, i_2], \ldots, J_n = [i_{n-1}+1, i_n]$, $i_n = k$, such that

$$w(j) = w(i_h) = \tilde a(i_h)\quad \text{for every } j\in J_h,\ h = 1, 2, \ldots, n.$$

Therefore, if $\# J_h$ stands for the number of elements of J_h, we have

$$\sum_{j=0}^{k} a(j) = \sum_{h=1}^{n}\sum_{j\in J_h} a(j) \leqslant \sum_{h=1}^{n} \# J_h\,\tilde a(i_h)$$
$$= \sum_{h=1}^{n}\sum_{j\in J_h} w(j) = \sum_{j=0}^{k} w(j).$$

This finishes the proof of Lemma (2.7).

Proof of Lemma (2.8). It is clear that

$$\lambda\sum_{i=0}^{k} w(i) = \lambda\sum_{i=0}^{k} w(i) + a(k)\,w(k) - a(k)\,w(k)$$
$$= \lambda\sum_{i=0}^{k-1} w(i) + a(k)\,w(k) + w(k)\big(\lambda - a(k)\big).$$

Since $\lambda - a(k) < 0$ and $w(k-1)\leqslant w(k)$ we have

$$\lambda\sum_{i=0}^{k} w(i) \leqslant \lambda\sum_{i=0}^{k-1} w(i) + w(k-1)\big(\lambda - a(k)\big) + a(k)\,w(k).$$

Now, the right-hand side of the last inequality is equal to

$$\lambda \sum_{i=0}^{k-2} w(i)+w(k-1)(2\lambda-a(k)-a(k-1))+a(k-1)w(k-1)+a(k)w(k).$$

Since $2\lambda < a(k)+a(k-1)$ and $w(k-2) \leqslant w(k-1)$ we get

$$\lambda \sum_{i=0}^{k} w(i) \leqslant \lambda \sum_{i=0}^{k-2} w(i)+w(k-2)(2\lambda-a(k)-a(k-1))+\sum_{i=k-1}^{k} a(i)w(i).$$

If we continue this process we have

$$\lambda \sum_{i=0}^{k} w(i) \leqslant \lambda w(0)+w(0)(k\lambda-\sum_{i=1}^{k} a(i))+\sum_{i=1}^{k} a(i)w(i)$$

$$= w(0)((k+1)\lambda-\sum_{i=0}^{k} a(i))+\sum_{i=0}^{k} a(i)w(i).$$

Now, the lemma follows from this inequality since $w(0) \geqslant 0$ and $(k+1)\lambda - \sum_{i=0}^{k} a(i) \leqslant 0$.

Proof of Lemma (2.9). Let

$$E = \{x \in X\colon\ T_{0,k}u(T^n x) \leqslant Cv(T^{k+n}x) \text{ for every } k \geqslant 0 \text{ and any integer } n\},$$

where C is the constant of statement (d) in Theorem (2.4). Then $\mu(X-E) = 0$. Let $x \in E$. For N a positive integer, we consider $b = a\chi_{[-N,N]}$ where $\chi_{[-N,N]}$ denotes the characteristic function of $[-N, N]$. Let $O_{\lambda,N} = \{j \in \boldsymbol{Z}\colon\ mb(j) > \lambda\}$ and let $\{I_r\}$ be the family of maximal intervals included in $O_{\lambda,N}$. Let $I = \{s, \ldots, s+k\}$ be any of the intervals I_r. Since $s+k+1 \notin O_{\lambda,N}$ it is clear that

$$\text{(2.10)} \qquad \lambda(k-i+1) < \sum_{j=i}^{k} |b(s+j)| \text{ for every } i \text{ with } 0 \leqslant i \leqslant k.$$

On the other hand, by Lemma (2.7), there exist numbers $w(s), \ldots, w(s+k)$ such that

(i) $0 \leqslant w(s) \leqslant w(s+1) \leqslant \ldots \leqslant w(s+k)$,

(ii) $w(s+i) \leqslant \max\{(r+1)^{-1} \sum_{j=0}^{r} u(T^{s+i-j}x)\colon\ 0 \leqslant r \leqslant i\}$ for every i with $0 \leqslant i \leqslant k$.

(iii) $\sum_{j\in I} u(T^j x) = \sum_{j=0}^{k} u(T^{s+j}x) \leqslant \sum_{j=0}^{k} w(s+j) = \sum_{j\in I} w(j)$.

Since u and v satisfy (d), property (ii) implies that $w(j) \leqslant Cv(T^j x)$ for every $j \in I$. Then Lemma (2.8) applied to $w(s), \ldots, w(s+k)$ and $|b(s)|, \ldots, |b(s+k)|$ (remember that (2.10) holds) gives

$$\sum_{j\in I} u(T^j x) \leqslant \sum_{j\in I} w(j) \leqslant \lambda^{-1} \sum_{j\in I} |b(j)|\, w(j)$$

$$\leqslant C\lambda^{-1} \sum_{j\in I} |b(j)|\, v(T^j x) \leqslant C\lambda^{-1} \sum_{j\in I} |a(j)|\, v(T^j x).$$

Since I is any of the intervals of the family $\{I_r\}$ it follows that

$$\sum_{j\in O_{\lambda,N}} u(T^j x) \leqslant C\lambda^{-1} \sum_{j\in O_{\lambda,N}} |a(j)|\, v(T^j x).$$

Letting N go to ∞ we have the inequality of Lemma (2.9) for every x in E. Since $\mu(X-E) = 0$ this finishes the proof of Lemma (2.9).

Once Lemma (2.9) has been proved, we will prove the implication (d) $\Rightarrow$ (a) using transference methods.

Let f be a measurable function from X into $\boldsymbol{R}$ and consider the truncated maximal operator

$$M_L f = \sup_{k\leqslant L} T_{0,k}|f|$$

where $L \geqslant 0$. Fix $\lambda > 0$ and denote by $O_{\lambda,L}$ the set $\{x \in X\colon\ M_L f(x) > \lambda\}$. Then since T preserves the measure μ we have

$$\text{(2.11)} \qquad \int_{O_{\lambda,L}} u\,d\mu = (k+1)^{-1} \sum_{j=0}^{k} \int_{T^{-j}O_{\lambda,L}} u(T^j x)\,d\mu.$$

If $f^x(j) = f(T^j x)$ the right-hand side of (2.11) is bounded by

$$\int_X (k+1)^{-1} \sum u(T^j x)\,d\mu$$

where the sum is extended over the integers j such that $m(f^x\chi_{[0,k+L]})(j) > \lambda$. Now Lemma (2.9) implies that the integrand is bounded by

$$C(k+1)^{-1}\lambda^{-1} \sum_{\{j\colon\ 0\leqslant j\leqslant k+L,\, m(f^x\chi_{[0,k+L]})(j)>\lambda\}} |f(T^j x)|\, v(T^j x)$$

$$\leqslant C(k+1)^{-1}\lambda^{-1} \sum_{\{j\colon\ 0\leqslant j\leqslant k+L,\, m(f^x)(j)>\lambda\}} |f(T^j x)|\, v(T^j x).$$

If we put these inequalities in (2.11) we have

$$\int_{O_{\lambda,L}} u\,d\mu \leqslant C\lambda^{-1}(k+L+1)(k+1)^{-1} \int_{\{x\colon\ Mf(x)>\lambda\}} |f|\, v\,d\mu.$$

Letting k and then L go to ∞ we get condition (a) in Theorem (2.4).

We have already shown that (a), (b) and (d) are equivalent. Thus to finish the proof of Theorem (2.4) it suffices to prove (c) $\Rightarrow$ (b) and (d) $\Rightarrow$ (c). The first implication is well known. On the other hand,

$$\text{(2.12)} \qquad \int_X |T_{0,k}f|\, u\,d\mu \leqslant \int_X T_{0,k}|f|\, u\,d\mu$$

$$= \int_X |f|\, T_{k,0}u\,d\mu = \int_X |f|(x)\, T_{0,k}u(T^{-k}x)\,d\mu.$$

If we assume (d) then $T_{0,k} u(T^{-k} x) \leqslant Cv(x)$ a.e. and therefore (c) follows from this inequality and (2.12).

Remarks.

(2.13) From the proof of Theorem (2.4) it can be seen that the theorem remains valid if $T^{-1}\mathscr{F} = \mathscr{F}$ (up to sets of measure zero) even if T is not invertible. We shall now give an example which shows that (a) does not imply (d) when T is not one-to-one and $T^{-1}\mathscr{F} \neq \mathscr{F}$. Let

$$X = \{(n, m) \in \boldsymbol{Z} \times \boldsymbol{Z}:\ n > 0,\ m < 0\} \cup \{(n, m) \in \boldsymbol{Z} \times \boldsymbol{Z}:\ n = 1,\ m \geqslant 0\},$$

Let $\mathscr{F}$ be the power set of X, μ the measure determined by

$$\mu\{(n, m)\} = 2^{-n} \quad \text{if } m < 0,$$
$$\mu\{(1, m)\} = 1 \quad \text{if } m \geqslant 0,$$

and T: $X \to X$ the transformation given by

$$T(n, m) = (n, m+1) \quad \text{if } m \neq -1,$$
$$T(n, -1) = (1, 0).$$

It is clear that T is not one-to-one, $T^{-1}\mathscr{F} \neq \mathscr{F}$ and T preserves the measure μ. Let us consider the function $u = v$ defined by

$$u(n, m) = 0 \quad \text{if } m < -1,$$
$$u(n, -1) = n,$$
$$u(1, m) = 1 \quad \text{if } m \geqslant 0.$$

Condition (d) does not hold in this case since $T_{0,1} u(n, -1) = \frac{1}{2}(u(n, -1) + u(1, 0)) = \frac{1}{2}(n+1)$ and there is no constant C such that $\frac{1}{2}(n+1) \leqslant Cu(1, 0) = C$ for every n. We will now see that (a) in Theorem (2.4) holds. Let f be a function from X into $\boldsymbol{R}$. It is clear that $\int_{O_\lambda} u\,d\mu = A + B$ where

$$O_\lambda = \{X:\ Mf(x) > \lambda\},\quad A = \sum_{\{n:\ (n,-1) \in O_\lambda\}} n2^{-n},\quad B = \#\{m \geqslant 0:\ (1, m) \in O_\lambda\}.$$

If $B = 0$ we have $(1, 0) \notin O_\lambda$ and thus $|f(n, -1)| > \lambda$ for every n such that $(n, -1) \in O_\lambda$. Then

$$\int_{O_\lambda} u\,d\mu = A < \lambda^{-1} \sum_{\{n:\ (n,-1) \in O_\lambda\}} |f(n, -1)|\, n2^{-n} = \lambda^{-1} \int_{O_\lambda} |f|\, u\,d\mu.$$

If $B \neq 0$ we have

$$A \leqslant \sum_{n=1}^{\infty} n2^{-n} \leqslant CB.$$

Then

$$\int_{O_\lambda} u\,d\mu \leqslant (C+1)B \leqslant (C+1)\lambda^{-1} \sum_{\{m \geqslant 0:\ (1,m) \in O_\lambda\}} |f(1, m)| \leqslant (C+1)\lambda^{-1} \int_{O_\lambda} |f|\, u\,d\mu$$

where the second inequality follows from the maximal ergodic theorem (classical case).

(2.14) In any case (for T invertible or not), assertions (b) and (c) in Theorem (2.4) are equivalent, and each one is equivalent to

(d′) $$\sup_{k \geqslant 0} (k+1)^{-1} \sum_{j=0}^{k} \int_{T^{-j}A} u\,d\mu \leqslant C \int_A v\,d\mu$$

for any measurable subset A.

It is clear that (c) implies (b). On the other hand, statement (c) with $f = \chi_A$ is statement (d′). Thus if we assume (d′) then (c) holds for characteristic functions, hence for simple functions and finally for every measurable function. In order to see that (b) implies (d′) we follow the proof of (b) $\Rightarrow$ (d) to get (2.5). Then (d′) follows from this and Lemma (2.2).

Assume now $u = v$ and set $\nu = u\,d\mu$. Then condition (d′) can be written in the following way:

$$\text{“}\sup_{k \geqslant 0} (k+1)^{-1} \sum_{j=0}^{k} \nu(T^{-j} A) \leqslant C\nu(A) \quad \text{for any measurable subset } A.\text{”}$$

This condition can be found in a paper of N. Dunford and D. S. Miller [2] in which they characterize the finite measures ν such that $\{T_{0,k} f\}$ converges in $L^1(d\nu)$ for every $f \in L^1(d\nu)$ (T is not necessarily one-to-one and does not preserve the measure ν). Besides it is shown that mean convergence is equivalent to uniform boundedness of the averages (our condition (c) with $u = v$) and it implies a.e. convergence.

It is clear by what we have already shown that (a) implies (d′) for T invertible or not. However, we do not know whether (d′) implies (a) (observe that conditions (d) and (d′) are not equivalent as the example in Remark (2.13) shows). Of course, assertions (d) and (d′) are equivalent in the invertible case or if $T^{-1}\mathscr{F} = \mathscr{F}$. Thus it seems reasonable to conjecture that (d′) is the condition which characterizes weak type (1, 1) of the ergodic maximal operator associated to a general transformation T.

(2.15) It is clear that, for a general transformation (invertible or not), condition (d) implies (d′). On the other hand, the proof of the implication (d) $\Rightarrow$ (a) of Theorem (2.4) works for a general transformation. Thus we have the following theorem.

(2.16) Theorem. *Let $(X, \mathscr{F}, \mu)$ be a σ-finite measure space and let T: $X \to X$ denote a measure-preserving transformation. Suppose u and v are* *nonnegative measurable functions on X. If $T_{0,k}u(x) \leqslant Cv(T^k x)$ for all $k > 0$ and almost all $x \in X$ then there exists a constant $C > 0$ such that*

$$\sup_{k \geqslant 0} \int_X |T_{0,k} f| u \, d\mu \leqslant C \int_X |f| v \, d\mu \quad \text{and}$$

$$\int_{\{x:\ Mf(x) > \lambda\}} u \, d\mu \leqslant C\lambda^{-1} \int_{\{x:\ Mf(x) > \lambda\}} |f| v \, d\mu$$

for every $\lambda > 0$ and any f.

(2.17) Theorem (2.4) remains valid if we set X instead of $\{x:\ Mf(x) > \lambda\}$ on the right-hand side of the inequalities of (a) and (b).

(2.18) Once Theorem (2.4) has been proved the result of [4] mentioned in the introduction follows easily as a corollary of (2.4).

3. The a.e. convergence. We will now characterize, for a general transformation (invertible or not), those positive functions v for which the averages $T_{0,k} f$ converge a.e. for every $f \in L^1(v d\mu)$ assuming $\mu(X) < \infty$.

(3.1) THEOREM. *Let $(X, \mathscr{F}, \mu)$ be a finite measure space and let T: $X \to X$ denote a measure-preserving transformation. Suppose v is a positive measurable function on X. The following statements are equivalent*:

(a) *The averages $T_{0,k} f$ converge a.e. for every $f \in L^1(v d\mu)$.*

(b) *$Mf(x) < \infty$ a.e. for every $f \in L^1(v d\mu)$.*

(c) *There exist a positive measurable function u and a constant $C > 0$ such that for every f and $\lambda > 0$*

$$\int_{\{x:\ Mf(x) > \lambda\}} u \, d\mu \leqslant C\lambda^{-1} \int_X |f| v \, d\mu.$$

(d) *There exist a positive measurable function u on X and a constant $C > 0$ such that for all $k \geqslant 0$ and almost all $x \in X$, $T_{0,k} u(x) \leqslant Cv(T^k x)$.*

(e) $0 < \inf_{i \geqslant 0} v(T^i x)$ *a.e.*

Proof. It is clear that (a) implies (b) and (c) implies (b). On the other hand, (d) implies (c) by Theorem (2.16) and (d) follows from (e) by taking $u(x) = \inf_{i \geqslant 0} v(T^i x)$. Thus it remains to prove (b) $\Rightarrow$ (e) and (e) $\Rightarrow$ (a).

(b) $\Rightarrow$ (e). Let $E = \{x:\ \inf_{i \geqslant 0} v(T^i x) = 0\}$. Suppose $\mu(E) > 0$. For $m > n$ we denote by $V_{m,n}$ the set $\{x:\ 2^{-m} \leqslant v(x) < 2^{-n}\}$. It is clear that

$$E = \bigcup_{n=2}^{\infty} E \cap \Big[\bigcup_{i=0}^{\infty} T^{-i} V_{n,1}\Big].$$

Given $(1-2^{-1})\mu(E)$ there exists $m(1) \geqslant 2$ such that

$$\mu\Big(E \cap \Big[\bigcup_{i=0}^{\infty} T^{-i} V_{m(1),1}\Big]\Big) > (1-2^{-1})\mu(E).$$

Now, observe that

$$E = \bigcup_{n=m(1)+1}^{\infty} E \cap \Big[\bigcup_{i=1}^{\infty} T^{-i} V_{n,m(1)}\Big].$$

Given $(1-2^{-2})\mu(E)$ we choose $m(2) \geqslant m(1)+1$ such that

$$\mu\Big(E \cap \Big[\bigcup_{i=0}^{\infty} T^{-i} V_{m(2),m(1)}\Big]\Big) > (1-2^{-2})\mu(E).$$

In this way we choose measurable sets $E_1 = V_{m(1),1}$, $E_2 = V_{m(2),m(1)}, \ldots, E_{n+1} = V_{m(n+1),m(n)}, \ldots$, such that $v(x) < 2^{-n}$ on E_n and

$$\mu(E \cap A_n) > (1-2^{-n})\mu(E) \quad \text{where} \quad A_n = \bigcup_{i=0}^{\infty} T^{-i} E_n. \tag{3.2}$$

Since $T^{-1} A_n \subset A_n$, we may assume clearly that A_n is invariant under T, i.e., $T^{-1} A_n = A_n$. Let $\mathscr{I}$ be the σ-field of all measurable subsets of A_n that are invariant under T, and define γ on $E_n \cap \mathscr{I} = \{E_n \cap A:\ A \in \mathscr{I}\}$ by

$$\gamma(E_n \cap A) = \mu(A) \quad \text{for } A \in \mathscr{I}.$$

We shall now see that γ is well defined. Let A and B be sets of the σ-field $\mathscr{I}$ such that $E_n \cap A = E_n \cap B$. It is clear that $E_n \subset A \cup (A_n - B) \subset A_n$ and $A \cup (A_n - B)$ is an invariant set. Since A_n is the smallest invariant set containing E_n we get $A \cup (A_n - B) = A_n$. Then $\mu(A_n) \leqslant \mu(A) + \mu(A_n) - \mu(B)$. In the same way $\mu(A_n) \leqslant \mu(B) + \mu(A_n) - \mu(A)$. Both inequalities give $\mu(A_n) = \mu(A) + \mu(A_n) - \mu(B)$. Thus $\mu(A) = \mu(B)$ and γ is well defined.

On the other hand, γ and μ restricted to $E_n \cap \mathscr{I}$ are equivalent measures. Thus, by the Radon–Nikodym theorem, if we put $f = d\gamma/d(\mu|_{E_n \cap \mathscr{I}})$ then $0 \leqslant f \in L^1(E_n, E_n \cap \mathscr{I}, \mu)$ and $\|f\|_1 = \mu(A_n) \leqslant \mu(X)$. Write

$$\tilde{f}(x) = \lim_{k \to \infty} T_{0,k} f(x).$$

Then, by the ergodic theorem, for all $A \in \mathscr{I}$,

$$\int_A \tilde{f} \, d\mu = \int_A f \, d\mu = \int_{E_n \cap A} f \, d\mu = \gamma(E_n \cap A) = \mu(A)$$

and this implies $\tilde{f} = 1$ a.e. on A_n. Let $f_n(x) = nf(x)$ if $x \in E_n$ and $f_n(x) = 0$ if $x \notin E_n$. Then $f_n \in L^1(E_n)$ and, since $v < 2^{-n}$ on E_n,

$$\int_{E_n} f_n v \, d\mu \leqslant n 2^{-n} \int_{E_n} f \, d\mu \leqslant n 2^{-n} \mu(X).$$

Thus the function $g = \sum_{n=1}^{\infty} f_n$ is in $L^1(v d\mu)$, and for almost all $x \in A_n$ we have $Mg(x) \geqslant Mf_n(x) \geqslant n$. This inequality together with (3.2) proves that $Mg(x) = \infty$ a.e. on E, a contradiction to (b).

(e) $\Rightarrow$ (a). Let

$$A_n = \{x:\ \inf_{i \geqslant 0} v(T^i x) < 2^{-n}\} = \bigcup_{i=0}^{\infty} T^{-i}(\{x:\ v(x) < 2^{-n}\}).$$

Observe that A_n, and therefore $X - A_n$, is invariant under T and since

$v(x) \geqslant 2^{-n}$ on $X-A_n$ we have $L^1(X-A_n, v\,d\mu) \subset L^1(X-A_n, d\mu)$. Then the ergodic theorem shows that $T_{0,k}f$ converge a.e. on $X-A_n$ for all $f \in L^1(X-A_n, v\,d\mu)$. Since $\lim \mu(A_n) = 0$ it is clear that (a) follows from this.

The implications (e) $\Rightarrow$ (d) and (d) $\Rightarrow$ (c) do not use the fact that the measure is finite. On the other hand, the convergence of the averages follows from (c), the density of $L^1(v\,d\mu) \cap L^1(d\mu)$ in $L^1(v\,d\mu)$ and the Banach principle even if the measure is not finite. Therefore the following theorem holds.

(3.3) THEOREM. *Let $(X, \mathscr{F}, \mu)$ be a σ-finite measure space and let $T\colon X \to X$ denote a measure-preserving transformation. Suppose v is a positive measurable function on X. If $\inf_{i \geqslant 0} v(T^i x) > 0$ a.e. then $T_{0,k}f$ converges a.e. for every $f \in L^1(v\,d\mu)$.*

The converse of Theorem (3.3) does not hold (in the case $\mu(X) = \infty$) as the following example shows.

(3.4) EXAMPLE. Let X be the set of all integers with the counting measure μ. Let T be the shift transformation $Tx = x+1$ and v the function defined by

$$v(x) = 1 \quad \text{if } x \neq 2^n \text{ for every positive integer } n,$$
$$v(2^n) = 1/n.$$

It is clear that $\inf_{i \geqslant 0} v(T^i x) = 0$ for every x. Now, let f be a positive function in $L^1(v\,d\mu)$. We will prove that $T_{0,k}f(x)$ converges for every x. To prove this we consider $f_1(x) = f(x)$ if $x = 2^n$ for some n and $f_1(x) = 0$ otherwise. Let $f_2 = f - f_1$. The function f_2 is in $L^1(d\mu)$ and thus $T_{0,k}f_2(x)$ converges for every x. To finish the proof, we will see that

$$\lim_{k \to \infty} T_{0,k} f_1(x) = 0 \quad \text{for every } x.$$

It will suffice to get

$$\lim_{n \to \infty} T_{0,2^n} f_1(x) = 0.$$

Let n be a positive integer such that $x \leqslant 2^n$. Then

$$0 \leqslant T_{0,2^n} f_1(x) \leqslant (2^n+1)^{-1} \|f_1\|_{1,v\,d\mu} \sum_{k=1}^{n+1} k$$

and since

$$\lim_{n \to \infty} (2^n+1)^{-1} \sum_{k=1}^{n+1} k = 0$$

we have $\lim_{n \to \infty} T_{0,2^n} f_1(x) = 0$. Therefore $T_{0,k}f(x)$ converges for every x and $f \in L^1(v\,d\mu)$ even though $\inf_{i \geqslant 0} v(T^i x) = 0$.

(3.5) Remark. Suppose that $Mf(x) < \infty$ a.e. for every $f \in L^1(v\,d\mu)$. By Nikishin's theorem [3] there exists a positive function u such that

$$\int_{\{x:\ Mf(x) > \lambda\}} u\,d\mu \leqslant \lambda^{-1} \int_X |f|\,v\,d\mu$$

for every $\lambda > 0$ and any f. Therefore Theorem (3.1) says that, in the case of finite measure, the Nikishin function is $u(x) = \inf_{i \geqslant 0} v(T^i x)$ (see (e) $\Rightarrow$ (d) in Theorem (3.1)).

4. On the ergodic Hilbert transform. Suppose that T is an invertible measure-preserving transformation. Then the proof of (e) $\Rightarrow$ (a) in Theorem (3.1) can be applied to the ergodic Hilbert transform $H(T)f$,

$$H(T)f(x) = \lim_{k \to \infty} \sum_{i=1}^{k} \frac{f(T^i x) - f(T^{-i} x)}{i}.$$

Thus the following theorem holds.

(4.1) THEOREM. *Let $(X, \mathscr{F}, \mu)$ be a finite measure space and let $T\colon X \to X$ denote an invertible measure-preserving transformation. Let v be a positive measurable function. If $\inf_{i \geqslant 0} v(T^i x) > 0$ a.e. then for a.e. x and any $f \in L^1(v\,d\mu)$ the limit*

$$\lim_{k \to \infty} \sum_{i=1}^{k} \frac{f(T^i x) - f(T^{-i} x)}{i} = H(T)f(x)$$

exists.

Acknowledgement. We want to thank the referee for some suggestions that led to an improvement of our original results. More precisely, we had proved Theorem (3.1) under the additional assumption that T is invertible. The referee pointed out that this condition could be eliminated.

References

[1] E. Atencia and F. J. Martín-Reyes, *Weak type inequalities for the maximal ergodic function and the maximal ergodic Hilbert transform in weighted spaces*, Studia Math. 78 (1984), 231–244.

[2] N. Dunford and D. S. Miller, *On the ergodic theorem*, Trans. Amer. Math. Soc. 60 (1946), 538–549.

[3] J. E. Gilbert, *Nikishin–Stein theory and factorization with applications*, in: Proc. Sympos. Pure Math. 35 (2), Amer. Math. Soc., Providence, R. I., 1979, 233–267.
[4] F. J. Martín-Reyes, *Inequalities for the ergodic maximal function and convergence of the averages in weighted L^p-spaces*, Trans. Amer. Math. Soc. 296 (1986), 61–82.

DEPARTAMENTO DE ANÁLISIS MATEMÁTICO
FACULTAD DE CIENCIAS
UNIVERSIDAD DE MÁLAGA
Málaga, Spain

Received December 9, 1985
Revised version May 28, 1986 (2121)

STUDIA MATHEMATICA, T. LXXXVII. (1987)

On a ratio ergodic theorem

by

RYOTARO SATO (Okayama)

Abstract. Let $(X, \mathcal{F}, \mu)$ be a σ-finite measure space and Q the positive isometry on $L_1(\mu)$ induced by a measurable transformation T and a positive function h; thus $Qf(x) = h(x) f(Tx)$. Clearly Q acts on functions on X. In this paper the ratio ergodic theorem is discussed for Q but the underlying space $L_1(\mu)$ is replaced by $L_1(wd\mu)$, where w is a positive function. The results obtained generalize some recent ones due to Martín-Reyes and de la Torre; they considered the case where T is measure-preserving and $h = 1$.

Introduction and results. Let T be a measurable transformation from X into itself such that $\mu(A) = 0$ if and only if $\mu(T^{-1}A) = 0$, and h a positive measurable function on X such that $\int_{T^{-1}A} h\,d\mu = \mu(A)$ for all $A \in \mathcal{F}$. If we define

$$Qf(x) = h(x) f(Tx)$$

for measurable functions f on X, then Q is positive and satisfies $\|Qf\|_1 = \|f\|_1$ for all $f \in L_1(\mu)$. (In view of the Banach–Lamperti work [3], Q represents a wide class of positive L_1-isometries.) It follows from the Chacon–Ornstein ergodic theorem [2] that for all f and e in $L_1(\mu)$ with $e \geq 0$, the ratios

$$R_0^n(f, e)(x) = \sum_{k=0}^{n} Q^k f(x) / \sum_{k=0}^{n} Q^k e(x)$$

converge almost everywhere on $\{x: \sum_0^\infty Q^k e(x) > 0\}$. But if we replace the measure μ by a measure $wd\mu$, where w is a positive measurable function, and if we assume that f and e are in $L_1(wd\mu)$, then the a.e. convergence of $R_0^n(f, e)(x)$ is unknown. (In case T is measure-preserving and $h = 1$ on X, this has recently been solved by Martín-Reyes and de la Torre [4], [5].) This is the starting point for the study in this paper. For this purpose we may and do fix an e with $0 < e \in L_1(\mu)$, as is easily seen from the Chacon identification theorem [1]. We characterize the positive functions w with the property that if $f \in L_1(wd\mu)$ then the ratios $R_0^n(f, e)(x)$ converge almost everywhere.