Can. J. Math., Vol. XXXV, No. 1, 1983, pp. 145–166

WEIGHTED AND SUBSEQUENTIAL ERGODIC THEOREMS

J. R. BAXTER AND J. H. OLSEN

1. Introduction. Let $(X, \mathscr{F}, \mu)$ be a probability space, T a linear operator on $\mathscr{L}^p(X, \mathscr{F}, \mu)$, for some p, $1 \leqq p \leqq \infty$. Let a_n be a sequence of complex numbers, $n = 0, 1, \ldots,$ which we shall often refer to as *weights*. We shall say that the weighted pointwise ergodic theorem holds for T on $\mathscr{L}^p$, if, for every f in $\mathscr{L}^p$,

$$(1.1)\qquad \lim_{N\to\infty} N^{-1} \sum_{n=0}^{N-1} a_n T^n f(x) \quad \text{exists for } \mu - \text{a.e. } x.$$

Let $\mathbf{a}$ denote the sequence (a_n). If (1.1) holds we shall say that $\mathbf{a}$ is *Birkhoff* for T on $\mathscr{L}^p$, or, more briefly, that $(\mathbf{a}, T)$ is Birkhoff.

We are also interested in ergodic theorems for subsequences. Let $n(k)$ be a subsequence. We shall say the pointwise ergodic theorem holds for the subsequence $n(k)$ and the operator T if, for every f in $\mathscr{L}^p$,

$$(1.2)\qquad \lim_{K\to\infty} K^{-1} \sum_{k=0}^{K-1} T^{n(k)} f(x) \quad \text{exists for } \mu - \text{a.e. } x.$$

Obviously, if the subsequence $n(k)$ has nonzero density we can reduce (1.2) to (1.1) be letting $a_n = 1$ if $n(k) = n$ for some k, $a_n = 0$ otherwise. Subsequences that do not have a density could easily be dealt with by the methods of the present paper, after a suitable generalization of the definition of weights given above. For simplicity, however, we shall restrict ourselves to the kind of weights already described and thus to subsequences that have nonzero density.

Suppose we have defined a class of weights $\mathscr{W}$, that is, a collection of complex sequences $\mathbf{a}$ satisfying some condition. Let $\mathscr{T}(\mathscr{W})$ denote the class of all operators T such that $(\mathbf{a}, T)$ is Birkhoff for every $\mathbf{a}$ in $\mathscr{W}$. Given a class $\mathscr{W}$, one basic problem is to determine $\mathscr{T}(\mathscr{W})$, or at least to get some idea how large $\mathscr{T}(\mathscr{W})$ is. One might wish, for example, to show that $\mathscr{T}(\mathscr{W})$ contains all operators T induced by measure-preserving point transformations of X. A second basic problem is to determine what stationary stochastic processes produce sequences $\mathbf{a}$ in $\mathscr{W}$ (with probability one). Since any stationary stochastic process can be represented using a measure-preserving point transformation, we can

Received November 3, 1981. The research of the first author was supported in part by NSERC Grant No. A3974 and a National Science Foundation Grant, while that of the second was supported in part by NSERC Grant No. A3974.

https://doi.org/10.4153/CJM-1983-010-7 Published online by Cambridge University Press

expect to see two classes of measure-preserving point transformations associated with a class $\mathscr{W}$ of weights: first, those transformations inducing operators in $\mathscr{T}(\mathscr{W})$, and second, those transformations producing sequences in $\mathscr{W}$.

To motivate these problems, and to show the difficulties, we consider an example. Let τ be a measure-preserving point transformation on $(X, \mathscr{F}, \mu)$, T the operator induced by τ, so that $Tf(x) = f(\tau(x))$. Let $(Y, \mathscr{G}, \nu)$ be another probability space, σ a measure-preserving point transformation on $(Y, \mathscr{G}, \nu)$. Fix a bounded measurable function g on Y. Assume $\mathscr{L}^1(X, \mathscr{F}, \mu)$ is separable, and let H be a countable dense set. By the Birkhoff pointwise ergodic theorem, applied to $\sigma \times \tau$, we see there is a set A in $Y \times X$ with $\nu \times \mu(A^c) = 0$, such that, for every h in H,

$$N^{-1} \sum_{n=0}^{N-1} g(\sigma^n(y))h(\tau^n(x))$$

converges for every (y, x) in A. Let $B \subseteq Y$ be such that $\nu(B^c) = 0$ and

$$\mu(\{x : (y, x) \notin A\}) = 0 \text{ for every } y \text{ in } B.$$

Fix y in B. For every h in H,

$$N^{-1} \sum_{n=0}^{N-1} g(\sigma^n(y))h(\tau^n(x))$$

converges for μ-a.e. x. The usual maximal inequality argument can be applied, since g is bounded, to show that for every f in $\mathscr{L}^1(X, \mathscr{F}, \mu)$, $N^{-1} \sum_{n=0}^{N-1} g(\sigma^n(y)) f(\tau^n(x))$ converges for u – a.e. x. If we let

$$\mathscr{W} = \{\mathbf{a} : a_n = g(\sigma^n(y)), y \text{ in } B\},$$

we have shown that $T \in \mathscr{T}(\mathscr{W})$. However, given a particular sequence $\mathbf{a}$, it seems difficult to determine whether $\mathbf{a}$ is in $\mathscr{W}$ or not. Also, $\mathscr{W}$ depends on T and g. Thus $\mathscr{W}$ is not a very good class of weights to investigate. It appears that we must define our classes of weights by unambiguous analytical conditions, before attempting to solve the two problems described earlier.

In the present paper we shall concern ourselves with four main classes of weights: $\mathscr{W}_2$, the weights corresponding to the uniform subsequences of Brunel and Keane [**4**], $\tilde{\mathscr{W}}_0$, which is almost the same as the class of bounded Besicovitch weights introduced by Ryll-Nardzewski [**16**], and two new classes of weights, $\overline{\mathscr{W}}_3$ and $\overline{\mathscr{W}}_4$, defined in Section 4. We also consider some examples of the saturating subsequences introduced by Reich [**15**].

Much is already known concerning $\mathscr{W}_2$ and $\tilde{\mathscr{W}}_0$. Ryll-Nardzewski showed $\mathscr{W}_2 \subseteq \tilde{\mathscr{W}}_0$. $\mathscr{T}(\mathscr{W}_2)$ and $\mathscr{T}(\tilde{\mathscr{W}}_0)$ are both known to contain fairly general operators ([**6**], [**11**], [**12**], [**13**], [**16**], [**17**]). In the present paper we prove two more results in this direction. Theorem (6.1) implies that any reasonable positive operator for which the ordinary pointwise ergodic theorem holds must be in $\mathcal{T}(\mathcal{W}_2)$. Theorem (6.3) shows $(\mathcal{T}\tilde{\mathcal{W}}_0)$ contains any operator which is dominated by a positive $\mathcal{L}^p$-contraction, $1 < p < \infty$.

Any measure-preserving point transformation with discrete spectrum produces sequences in $\tilde{\mathcal{W}}_2$, with probability one (Theorem 5.2). Any mean-zero sequence produced by a weak mixing point transformation lies in $\overline{\mathcal{W}}_3$, with probability one, which is indeed why $\overline{\mathcal{W}}_3$ is introduced (Theorem 5.2). Unfortunately, $\mathcal{T}(\overline{\mathcal{W}}_3)$ is not known, and may be rather small. We can show, however, that if $\mathbf{a}$ is in $\overline{\mathcal{W}}_3$ then $N^{-1}\sum_{n=0}^{N-1} a_n T^n f$ converges in norm, for a fairly large class of operators (Theorem 6.23). $\overline{\mathcal{W}}_4$ is a subset of $\overline{\mathcal{W}}_3$. We show that any mean-zero sequence produced by the shift associated with an i.i.d. sequence of random variables lies in $\overline{\mathcal{W}}_4$, with probability one (Theorem 5.3). $\mathcal{T}(\overline{\mathcal{W}}_4)$ is quite large. We show that any operator T on $\mathcal{L}^\infty(X, \mathcal{F}, \mu)$ which is power bounded in $\mathcal{L}^\infty$-norm and power-bounded in one other $\mathcal{L}^p$-norm must be in $\mathcal{T}(\overline{\mathcal{W}}_4)$ (Theorem 6.4).

We note that discrete-spectrum transformations and Bernoulli shifts both produce weights such that (1.1) holds for a large class of T. The behavior of weights produced by transformations falling between these two categories remains an open question.

2. Preliminaries. Let $\mathbf{a}$ be a complex sequence. We say $\mathbf{a}$ has a *mean* if

$$\lim_{N\to\infty} N^{-1} \sum_{n=0}^{N-1} a_n \quad \text{exists,}$$

and we call this limit the mean of $\mathbf{a}$. Let

$$\|\mathbf{a}\|_\infty = \sup\{|a_n| : n = 0, 1, \dots\}.$$

For $1 \leq p < \infty$, define $\|\mathbf{a}\|_p$ by

$$\|\mathbf{a}\|_p^p = \limsup_{N\to\infty} N^{-1} \sum_{n=0}^{N-1} |a_n|^p.$$

$\|\cdot\|_p$ is a seminorm, $1 \leqq p \leqq \infty$, on $\{\mathbf{a} : \|\mathbf{a}\|_p < \infty\}$. Also, if $\|\mathbf{a}\|_p < \infty$, $\|\mathbf{b}\|_p < \infty$, then $\|\mathbf{a} + \mathbf{b}\|_p < \infty$.

(2.1) LEMMA. *Let* $\mathbf{a}(k)$, $\mathbf{a}$ *be complex sequences such that each* $\mathbf{a}(k)$ *has a mean. If* $\|\mathbf{a}(k) - \mathbf{a}\|_1 \to 0$ *as* $k \to \infty$, *then* $\mathbf{a}$ *has a mean.*

(2.2) LEMMA. *Let* $\mathbf{a}$, $\mathbf{b}$ *be complex sequences. Define* $\mathbf{ab} = \mathbf{c}$ *by* $c_n = a_n b_n$. *Let* $p^{-1} + q^{-1} = 1$, $1 \leqq p \leqq \infty$. *Then* $\|\mathbf{ab}\|_1 \leqq \|\mathbf{a}\|_p \|\mathbf{b}\|_q$.

(2.3) COROLLARY. *Let* $p^{-1} + q^{-1} = 1$, $1 \leqq p \leqq \infty$. *Suppose* $\|\mathbf{a}(k)\|_p$ *is bounded in* k, $\|\mathbf{b}(k)\|_q \to 0$. *Then* $\|\mathbf{a}(k)\mathbf{b}(k)\|_1 \to 0$.

https://doi.org/10.4153/CJM-1983-010-7 Published online by Cambridge University Press

Let T be a linear operator on $\mathscr{L}^p(X, \mathscr{F}, \mu)$, $1 \leq p < \infty$. Define the maximal operator M for T by

$$Mf(x) = \sup\left\{\left|N^{-1}\sum_{n=0}^{N-1} T^n f(x)\right| : N = 1, 2, \ldots\right\}.$$

We say a *maximal inequality* holds for T if

(2.4) $\alpha\mu(\{Mf > \alpha\}) \leq (\text{constant}) \|f\|_1$ for $f \in \mathscr{L}^p$, $\alpha > 0$.

We say a *dominated estimate* holds for T if

(2.5) $\|Mf\|_p \leq (\text{constant}) \|f\|_p$ for $f \in \mathscr{L}^p$.

(2.4) holds if T is Dunford–Schwartz, that is, if T is a contraction with respect to $\mathscr{L}^1$-norm and $\mathscr{L}^\infty$-norm (cf. [**8**], Chapter 8). (2.5) holds, for $p > 1$, if T is a positive $\mathscr{L}^p$-contraction ([**1**]), if T is positive, invertible, and power bounded on $\mathscr{L}^p$ ([**7**], [**9**]), or if T is a Lamperti contraction or $\mathscr{L}^p$ ([**9**]).

If (2.5) holds, then clearly

(2.6) $\alpha^p\mu(\{Mf > \alpha\}) \leqq (\text{constant}) \|f\|_p{}^p$ for $f \in \mathscr{L}^p$, $\alpha > 0$.

Actually we will need the slightly stronger maximal operator V defined by

$$Vf(x) = \sup\left\{N^{-1}\sum_{n=0}^{N-1} |T^n f(x)| : N = 1, 2, \ldots\right\}.$$

More generally, for $1 \leqq r < \infty$, let

(2.7) $$V_r f(x) = \sup\left\{N^{-1}\sum_{n=0}^{N-1} |T^n f(x)|^r : N = 1, 2, \ldots\right\}.$$

We will say that an operator T on $\mathscr{L}^p$ is *dominated* by an operator S if $|Tf| \leqq S|f|$ for all f in $\mathscr{L}^p$. S is of necessity then a positive operator. Since the various operators mentioned after equation (2.5) above are all dominated by positive operators of the same type, we have, for such operators,

(2.8) $\alpha^p\mu(\{Vf > \alpha\}) \leq (\text{constant}) \|f\|_p{}^p$ for $f \in \mathscr{L}^p$, $\alpha > 0$.

Also, we have

(2.9) LEMMA. *Let T be a linear operator on $\mathscr{L}^p(X, \mathscr{F}, \mu)$ for some p, $1 < p < \infty$, such that T is dominated by an $\mathscr{L}^p$-contraction. Then, for $1 \leqq r < p$, there exists C, depending only on p/r, such that:*

(2.10) $\int (V_r f)^{p/r} d\mu \leqq C \int |f|^p d\mu$, *for* $f \in \mathscr{L}^p$.

Proof. To say T is dominated by an $\mathscr{L}^p$-contraction means there exists an $\mathscr{L}^p$-contraction T' such that $|Tf| \leqq T'|f|$ for $f \in \mathscr{L}^p$. Then

https://doi.org/10.4153/CJM-1983-010-7 Published online by Cambridge University Press

T' is positive, and it is clearly enough to prove (2.10) when T itself is positive.

By [**1**] and [**2**], we can represent a positive $\mathscr{L}^p$-contraction T in terms of a conditional expectation and a positive $\mathscr{L}^p$-isometry on an appropriate $\mathscr{L}^p$-space. It follows easily from this representation that we need only prove (2.10) when T is a positive $\mathscr{L}^p$-isometry.

When T is a positive $\mathscr{L}^p$-isometry,we can at once construct a positive $\mathscr{L}^{p/r}$-isometry S, such that $|Tf|^r = S|f|^r$ for $f \in \mathscr{L}^p$ ([**3**]). Since $V_r|f| = M|f|^r$, where M is the maximal operator for S, (2.10) then follows from the dominated estimate for M ([**3**]).

(2.10) implies at once that

(2.11) $\alpha^{p/r}\mu(\{V_r f > \alpha\}) \leqq (\text{constant}) \|f\|_p{}^p$, for $f \in \mathscr{L}^p, \alpha > 0$.

We note that when T is induced by a measure preserving point transformation, (2.11) remains true when $r = p$, with a similar proof.

(2.12) Lemma. *Let T be a linear operator $\mathscr{L}^p(X, \mathscr{F}, \mu)$, $1 \leqq p < \infty$, and let r be such that (2.11) holds. Let $g \in \mathscr{L}^p(X, \mathscr{F}, \mu)$. Define $\mathbf{b}(x)$ by $b_n(x) = T^n g(x)$. Then $\|\mathbf{b}(x)\|_r < \infty$ for μ −* a.e. x. *Suppose*

$$g_k \in \mathscr{L}^p(X,\mathscr{F},\mu), \sum_{k=1}^{\infty} ||g_k - g||_p^p < \infty.$$

Define $\mathbf{b}(k)(x)$ by $b_n(k)(x) = T^n g_k(x)$. Then

$$\|\mathbf{b}(k)(x) - \mathbf{b}(x)\|_r \to 0 \text{ as } k \to \infty,$$

for u − a.e. x.

Proof. Let $A(\alpha) = \{x : \|\mathbf{b}(x)\|_r > \alpha\}$. Then

$$A(\alpha) \subseteq \{V_r g > \alpha^r\}.$$

Hence $\mu(A(\alpha)) \leqq \alpha^{-p}$ (constant) $\|g\|_p{}^p$. Thus $\|\mathbf{b}(x)\|_r < \infty$ μ − a.e. Fix $\delta > 0$. Let

$$B(k) = \{x : \|\mathbf{b}(k)(x) - \mathbf{b}(x)\|_r > \delta\}.$$

Then $B(k) \subseteq \{V_r(g_k - g) > \delta\}$, so

$$\mu(B(k)) \leqq \delta^{-p/r} (\text{constant}) \|g_k - g\|_p{}^p,$$

and hence

$$\sum_{k=1}^{\infty} \mu(B(k)) < \infty.$$

The lemma follows.

(2.13) Lemma. *Let a_n be a bounded complex sequence. Let T be an operator on $\mathscr{L}^p(X, \mathscr{F}, \mu)$, $1 \leqq p < \infty$, such that (2.8) holds. Suppose* *that* $N^{-1}\sum_{n=0}^{N-1} a_nT^nf$ *converges* μ − a.e. *for* f *in some dense subset of* $\mathscr{L}^p(X, \mathscr{F}, \mu)$. *Then* $(\mathbf{a}, T)$ *is Birkhoff, that is,* $N^{-1}\sum_{n=0}^{N-1} a_nT^nf$ *converges* μ − a.e. *for all* f *in* $\mathscr{L}^p(X, \mathscr{F}, \mu)$.

Proof. This follows at once from (2.12) with $r = 1$, (2.3), and (2.1).

(2.14) *Remark.* Often we will show $(\mathbf{a}, T)$ is Birkhoff for some T on $\mathscr{L}^\infty(X, \mathscr{F}, \mu)$. If T extends to an operator $\tilde{T}$ on $\mathscr{L}^p(X, \mathscr{F}, \mu)$, $1 \leqq p < \infty$, such that (2.8) holds, and if $\mathbf{a}$ is bounded, then (2.13) shows $(\mathbf{a}, \tilde{T})$ is Birkhoff. Since this type of generalization is standard, we will not always explicitly note it.

Let $\mathscr{W}$ be a collection of complex sequences. Let

(2.15) $\quad \overline{\mathscr{W}} = \|\cdot\|_1$ − closure of $\mathscr{W}$.

2.16) Lemma. *Let* T *be a power-bounded operator on* $\mathscr{L}^\infty(X, \mathscr{F}, \mu)$. *If* $T \in \mathscr{T}(\mathscr{W})$ *then* $T \in \mathscr{T}(\overline{\mathscr{W}})$.

Proof. This follows at once from (2.3), (2.1).

We shall abbreviate measure preserving point transformation as mppt.

(2.17) Lemma. *Let* $\mathscr{W}$ *be a collection of complex sequences. Let* σ *be a* mppt *on a probability space* $(Y, \mathscr{G}, \nu)$. *Let* $C \subseteq \mathscr{L}^1(Y, \mathscr{G}, \nu)$. *Suppose for every* $g \in C$, *for* v − a.e. y, *the sequence* $\mathbf{a}$ *defined by* $a_n = g(\sigma^n(y))$ *lies in* $\mathscr{W}$. *Then for every* $g \in \overline{C}$, *for* ν − a.e. y, *the sequence* $\mathbf{a}$ *defined by* $a_n = g(\sigma^n(y))$ *lies in* $\overline{\mathscr{W}}$.

Proof. This follows from (2.12).

Let $(Z, \mathscr{H}, \lambda)$ be a probability space. Let

$$X = \prod_{n=-\infty}^{\infty} Z_n, \quad \mathscr{F} = \prod_{n=-\infty}^{\infty} \mathscr{H}_n, \quad \mu = \prod_{n=-\infty}^{\infty} \lambda_n,$$

where $Z_n = Z$, $\mathscr{H}_n = \mathscr{H}$, $\lambda_n = \lambda$ for every n. Define the shift $\tau : X \to X$ by $(\tau x)_n = (x)_{n+1}$. τ is clearly an invertible mppt. We will refer to τ as an iid *shift*, and $(X, \mathscr{F}, \mu, \tau)$ as an iid shift space.

(2.18) Lemma. *Let* $(X, \mathscr{F}, \mu, \tau)$ *be an* iid *shift space. Define* T *on* $\mathscr{L}^1(X, \mathscr{F}, \mu)$ *by* $Tf = f \circ \tau$. *Let* $\mathbf{a}$ *be a bounded complex sequence that has a mean. Then* $(\mathbf{a}, T)$ *is Birkhoff.*

Proof. Let f be a bounded function of finitely many coordinates. The strong law of large numbers shows easily that

$$N^{-1}\sum_{n=0}^{N-1} a_nT^nf \quad \text{converges } \mu - \text{a.e.}$$

The result then follows from Lemma (2.13).

https://doi.org/10.4153/CJM-1983-010-7 Published online by Cambridge University Press

Lemma (2.18) shows there is not much sport to be had in proving weighted ergodic theorems for operators induced by iid shifts. However, we will later study the weights produced by iid shifts.

The next result shows we need not consider the class of Dunford–Schwartz operators separately from the class of operators induced by point transformations, as far as weighted ergodic theorems are concerned.

(2.19) THEOREM. *Let $\mathscr{W}$ be a collection of bounded complex sequences. Suppose $\mathscr{T}(\mathscr{W})$ contains all operators induced on $\mathscr{L}^1$-spaces by invertible* mppt'*s. Then $\mathscr{T}(\mathscr{W})$ contains all Dunford–Schwartz operators.*

Proof. (i) $\mathscr{T}(\mathscr{W})$ contains all operators induced on $\mathscr{L}^1$-spaces by arbitrary mppt's, not necessarily invertible. This follows in the usual way by building one-sided and two-sided shifts.

(ii) $\mathscr{T}(\mathscr{W})$ contains all operators T on $\mathscr{L}^1$-spaces, T of the following sort: τ a mppt on a probability space $(X, \mathscr{F}, \mu)$, v real, measurable on X, $0 \leqq v \leqq 1$, T defined by $Tf = vf \circ \tau$. To see this, let

$$Y = \prod_{n=0}^{\infty} I_n, v = \prod_{n=0}^{\infty} \lambda_n,$$

where $I_n = [0, 1]$, and λ_n = Lebesgue measure on I_n, for each n. Let ξ_n be the nth coordinate function on Y. Let

$$\varphi : X \times Y \rightarrow \{0, 1, \ldots, \infty\}$$

be defined by $\varphi(x, y)$ = first n such that $\xi_n(y) > v(\tau^n(x))$, with $\varphi(x, y) = \infty$ if no such n exists. Clearly

$$\lambda(\{y : \varphi(x, y) > n\}) = \prod_{k=0}^{n} v(\tau^k(x)),$$

so

$$T^n f(x) = \int \chi_{\{\varphi > n-1\}}(y) f(\tau^n(x)) d\lambda(y).$$

Let $\mathbf{a} \in \mathscr{W}$. By Lemma (2.13), to show $(\mathbf{a}, T)$ is Birkhoff it is enough to show $N^{-1} \sum_{n=0}^{N-1} a_n T^n f$ converges for f bounded measurable.

$$N^{-1} \sum_{n=0}^{N-1} a_n T^n f = \int \left[N^{-1} \sum_{n=0}^{N-1} \chi_{\{\varphi > n-1\}}(y) a_n f(\tau^n(x)) \right] d\lambda(y),$$

and the integrand converges λ — a.e. for μ — a.e. x, by (i), so (ii) holds.

(iii) $\mathscr{T}(\mathscr{W})$ contains all operators T on $\mathscr{L}^1$-spaces, T of the following sort: τ a mppt on a probability space $(X, \mathscr{F}, \mu)$, v complex, measurable on X, $|v| \leqq 1$, T defined by $Tf = vf \circ \tau$. To see this, we use a trick of Ryll-Nardzewski [**16**]. Let

$$Y = X \times [0, 2\pi),$$
$$B = \mathscr{F} \times (\text{Borel sets of } [0, 2\pi)),$$
$$\lambda = \mu \times (\text{normalized Lebesgue measure on } [0, 2\pi)).$$

https://doi.org/10.4153/CJM-1983-010-7 Published online by Cambridge University Press

Write

$$v(x) = r(x)e^{i\varphi(x)}.$$

Define $\sigma: Y \to Y$ by

$$\sigma(x, \theta) = (\tau(x), \varphi(x) + \theta),$$

where $\varphi(x) + \theta$ is to be interpreted modulo 2π. Define V on $\mathscr{L}^1(Y, \mathscr{B}, \lambda)$ by

$$Vg(x, \theta) = r(x)g(\sigma(x, \theta)).$$

Given $f \in \mathscr{L}^1(X, \mathscr{F}, \mu)$, let

$$g(x, \theta) = f(x)e^{i\theta}.$$

Let $\mathbf{a} \in \mathscr{W}$. By (ii),

$$\lim_{N\to\infty} N^{-1} \sum_{n=0}^{N-1} a_n V^n g(x, \theta) \quad \text{exists for } \lambda - \text{a.e. } (x, \theta).$$

Since

$$V^n g(x, \theta) = e^{i\theta} T^n f(x),$$

(iii) follows.

(iv) Now let T be a Dunford–Schwartz operator on $\mathscr{L}^1(X, \mathscr{F}, \mu)$. We use an old trick of Doob. First, by a routine argument, we may assume that X is a compact metric space and $\mathscr{F}$ is the Borel σ-algebra. Imitating the proof of the existence of regular conditional probabilities, we may assume that there exists a proper Markov transition function p on $X \times \mathscr{F}$, and a complex $\mathscr{F} \times \mathscr{F}$-measurable function G on $X \times X$, such that $|G| \leqq 1$,

$$\int p(x, B)\mu(dx) = \mu(B) \text{ for every } B \in \mathscr{F}, \text{ and}$$
$$Tf(x) = \int p(x, dy)G(x, y)f(y) \text{ for } \mu - \text{a.e. } x,$$

for every $f \in \mathscr{L}^1(X, \mathscr{F}, \mu)$. Let

$$\Omega = \prod_{n=0}^{\infty} X_n,$$

where $X_n = X$ for each n, and let $(\Omega, \mathscr{H}, \mathscr{H}_n, \xi_n, Pr^x, \theta)$ be the usual Markov process with transition function p, where θ is the shift on Ω. We note that Pr^μ is invariant under θ. Let $v = G(\xi_0, \xi_1)$. Define W on $\mathscr{L}^1(\Omega, \mathscr{H}, Pr^\mu)$ by $Wg = vg \circ \theta$. Let $\mathbf{a} \in \mathscr{W}$. By Lemma (2.13) to show $(\mathbf{a}, T)$ is Birkhoff it is enough to show

$$N^{-1} \sum_{n=0}^{N-1} a_n T^n f \quad \text{converges } \mu - \text{a.e.}$$

for f bounded measurable. For such an f, let $g = f \circ \xi_0$. By (iii),

$$N^{-1} \sum_{n=0}^{N-1} a_n W^n g \quad \text{converges } Pr^{\mu} - \text{a.e.}$$

Since $T^n f(x) = E^x[W^n g]$ the theorem follows.

3. Weight classes $\mathscr{W}_0, \mathscr{W}_1, \mathscr{W}_2$. Let

(3.1) $\mathscr{W}_0$ = complex-linear span $\{\mathbf{a}$: for some λ, $|\lambda| = 1$, $a_n = \lambda^n$ for all $n\}$.

The bounded sequences in $\overline{\mathscr{W}}_0$ are referred to by Ryll-Nardzewski as *bounded Besicovitch* sequences [**16**]. We shall also study some slightly smaller classes of weights. For any collection $\mathscr{W}$ of bounded complex sequences, let

(3.2) $\tilde{\mathscr{W}}$ = sup-dominated $\|\cdot\|_2$-closure $\mathscr{W}$.

This means $\tilde{\mathscr{W}}$ is the set of $\mathbf{a}$ for which there exists $\mathbf{a}(k) \in \mathscr{W}$ with $\|\mathbf{a}(k)\|_\infty \leqq \|\mathbf{a}\|_\infty$ for all k and $\|\mathbf{a}(k) - \mathbf{a}\|_1 \to 0$ as $k \to \infty$.

Since $\|\mathbf{a}\|_\infty \leqq \|\mathbf{b}\|_\infty$ is a transitive relation on sequences, we see easily that

(3.3) $(\tilde{\mathscr{W}})^\sim = \tilde{\mathscr{W}}$.

Clearly, if $\mathscr{W}$ is a real-linear space of complex sequences then

(3.4) $\|\cdot\|_\infty$ – closure $\mathscr{W} \subseteq \tilde{\mathscr{W}}$, $\tilde{\mathscr{W}} \subseteq \|\cdot\|_p$ – closure $\mathscr{W}$, $1 \leqq p < \infty$.

Now let $\mathscr{W}$ be a real-linear space of complex sequences containing the constant sequences 1. For any bounded complex sequence $\mathbf{a}$, let

(3.5) $\mathscr{W}$-seminorm $(\mathbf{a}) = \inf\{\|\mathbf{b}\|_1 : \mathbf{b} \in \mathscr{W}, |a_n| \leqq b_n \text{ for all } n\}$.

Clearly

(3.6) $\|\mathbf{a}\|_1 \leqq \mathscr{W}$-seminorm $(\mathbf{a}) \leqq \|\mathbf{a}\|_\infty$.

Let $\mathscr{V}$ be a collection of bounded complex sequences. Let

(3.7) $\mathscr{W}(\mathscr{V})$ = sup-dominated $\mathscr{W}$-seminorm closure $\mathscr{V}$.

That is, $\mathscr{W}(\mathscr{V})$ consists of those sequences $\mathbf{a}$ for which there exists $\mathbf{a}(k) \in \mathscr{V}$ with $\|\mathbf{a}(k)\|_\infty \leqq \|\mathbf{a}\|_\infty$ for all k and $\mathscr{W}$-seminorm $(\mathbf{a}(k) - \mathbf{a}) \to 0$ as $k \to \infty$.

Just as for (3.3), we see that

(3.8) $\mathscr{W}(\mathscr{W}(\mathscr{V})) = \mathscr{W}(\mathscr{V})$.

By (3.6),

(3.9) $\mathscr{W}(\mathscr{V}) \subseteq \tilde{\mathscr{V}} \subseteq \overline{\mathscr{V}}$,

https://doi.org/10.4153/CJM-1983-010-7 Published online by Cambridge University Press

and if $\mathscr{V}$ is real-linear then

(3.10) $\|\cdot\|_\infty$-closure $\mathscr{V} \subseteq \mathscr{W}(\mathscr{V})$.

Let $\mathscr{W}$ and $\mathscr{U}$ be real-linear spaces of complex sequences both containing the constant sequence 1. If $\mathscr{W} \subseteq \|\cdot\|_\infty$-closure $\mathscr{U}$, then

(3.11) $\mathscr{W}$-seminorm $\geqq$ $\mathscr{U}$-seminorm.

If (3.11) holds, then for any collection $\mathscr{V}$ of bounded complex sequences.

(3.12) $\mathscr{W}(\mathscr{V}) \subseteq \mathscr{U}(\mathscr{V})$.

We define

(3.13) $\mathscr{W}' = \mathscr{W}(\mathscr{W})$.

$\mathscr{W}_0'$ is a very well-behaved set of sequences. We will show it contains some interesting examples.

Let Y be a compact metric space, $\mathscr{B}$ the Borel sets of Y. Let σ be a homeomorphism of Y such that $\{\sigma^n : n \text{ an integer}\}$ is an equicontinuous family. We may choose an equivalent metric such that σ is an isometry (cf. [**10**]). (Actually it is enough to have $\{\sigma^n : n \text{ a positive integer}\}$ equicontinuous. Then the rest follows.) Let there exist some point in Y with a dense orbit. It is then easy to see that every point in Y has a dense orbit, that there is a unique σ-invariant probability ν on $(Y, \mathscr{B})$, and that for $g \in C(Y)$,

(3.14) $N^{-1} \sum_{n=0}^{N-1} g \circ \sigma^n$ converges uniformly to $\bar{g}$,

where $\bar{g}$ is the constant function $= \int g d\nu$. A system (Y, σ) of the sort just described is called *strictly $\mathscr{L}$-stable*. Let (Y, σ) be any strictly $\mathscr{L}$-stable system, g any function in $C(Y)$, y any point in Y. The sequence of numbers $a_n = g(\sigma^n(y))$ will be called a *continuous* sequence. Let

(3.15) $\mathscr{W}_1$ = complex-linear span $\{\mathbf{a} : \mathbf{a} \text{ is continuous}\}$.

Let $C_1(Y)$ denote the collection of functions g on Y such that g is continuous except on at most a closed ν-null set. For any strictly $\mathscr{L}$-stable system (Y, σ), any $g \in C_1(Y)$, and any $y \in Y$, the sequence $a_n = g(\sigma^n(y))$ is called *uniform* ([**6**]). Let

(3.16) $\mathscr{W}_2$ = complex-linear span $\{\mathbf{a} : \mathbf{a} \text{ uniform}\}$.

Clearly

(3.17) $\mathscr{W}_0 \subseteq \mathscr{W}_1 \subseteq \mathscr{W}_2$.

Since the span of the continuous eigenfunctions of the unitary operator induced by σ on $\mathscr{L}^2(Y, \mathscr{B}, \nu)$ is dense in $C(Y)$ in sup-norm (cf[**10**], [**16**]), we have

(3.18) $\mathscr{W}_1 \subseteq \|\cdot\|_\infty$-closure $\mathscr{W}_0$.

Let $h \in C_1(Y)$. Let B be a closed ν-null set such that h is continuous on B^c. We can find sequences $g_k, f_k \in C(Y)$ such that

$$\|g_k\|_{\text{sup}} \leqq \|h\chi_{B^c}\|_{\text{sup}},$$
$$|g_k - h| \leqq f_k \text{ on } Y \text{ for all } k,$$
$$\int f_k dv \to 0 \text{ as } k \to \infty.$$

Let $y \in Y$. Let $a_n = h(\sigma^n(y))$, $a_n(k) = g_k(\sigma^n(y))$.

Since y has a dense orbit, it is not hard to see that $\{\sigma^n(y) : n = 0, 1, \dots\}$ is dense. Hence

$$\|\mathbf{a}\|_\infty \geqq \|h\chi_{B^c}\|_{\text{sup}},$$

and so $\|\mathbf{a}\|_\infty \geqq \|\mathbf{a}(k)\|_\infty$ for all k.

$$\|\mathbf{b}(k)\|_1 = \int f_k dv,$$

by (3.14), where $b_n(k) = f_k(\sigma^n(y))$. Hence

$$\mathscr{W}_1\text{-seminorm } (\mathbf{a}(k) - \mathbf{a}) \leqq \int f_k d\nu \to 0 \text{ as } k \to \infty.$$

Thus

(3.19) $\mathscr{W}_2 \subseteq \mathscr{W}_1(\mathscr{W}_1) \equiv \mathscr{W}_1'$.

By (3.17), (3.18), and 3.11),

(3.20) $\mathscr{W}_0$-seminorm $= \mathscr{W}_1$-seminorm.

Hence

$$\mathscr{W}_1(\mathscr{W}_1) = \mathscr{W}_0(\mathscr{W}_1) \subseteq \mathscr{W}_0(\|\cdot\|_\infty\text{-closure } \mathscr{W}_0)$$
$$\subseteq \mathscr{W}_0(\mathscr{W}_0(\mathscr{W}_0)) = \mathscr{W}_0(\mathscr{W}_0) = \mathscr{W}_0'.$$

Also

$$\mathscr{W}_0(\mathscr{W}_0) \subseteq \mathscr{W}_0(\mathscr{W}_1) = \mathscr{W}_1(\mathscr{W}_1) = \mathscr{W}_1'.$$

Thus

(3.21) $\mathscr{W}_0' = \mathscr{W}_1'$,

and so

(3.22) $\mathscr{W}_2 \subseteq \mathscr{W}_0'$.

We will use (3.22) in Corollary (6.2). Actually all that is needed there is that $\mathscr{W}_2 \subseteq \mathscr{W}_0$-seminorm closure $\mathscr{W}_0$.

(3.23) LEMMA. *Let $\mathscr{W}$ be any complex algebra of bounded sequences (componentwise operations), closed under conjugation. Then $\tilde{\mathscr{W}}$ = the set of bounded sequences in $\bar{\mathscr{W}}$.*

Proof. This is similar to that of Lemma 5.1, using the Weierstrass approximation theorem. By (3.23), $\tilde{\mathscr{W}}_0$ is exactly the set of bounded Besicovitch sequences.

4. Weight classes $\mathscr{W}_3, \mathscr{W}_4$. We now define classes of sequences which are very different from $\mathscr{W}_0, \mathscr{W}_1, \mathscr{W}_2$. Let $\mathbf{a}$ be a complex sequence. Suppose there exists, for each positive integer K, a positive integer $N(K)$, and a set $I(K)$ of nonnegative integers, such that:

(4.1) $\quad N(K) \to \infty$ as $K \to \infty$,

(4.2) $\quad N(K + 1)/N(K) \to 1$ as $K \to \infty$,

(4.3) $\quad N(K)^{-1} \sup I(K) \to 0$ as $K \to \infty$,

(4.4) $\quad |I(K)| \to \infty$ as $K \to \infty$,

where $|I|$ denotes the number of elements in I, and such that for every $\epsilon > 0$, for K sufficiently large (depending on ϵ), if $i, j \in I(K)$, $i \neq j$, then

$$\left| \sum_{n=0}^{N(K)-1} \overline{a_{n+i}} a_{n+j} \right| \leqq \epsilon N(K). \tag{4.5}$$

Let

$$\mathscr{W}_3 = \text{complex-linear span } \{\mathbf{a} : \|\mathbf{a}\|_2 < \infty, \text{ (4.1)–(4.5) hold}\}. \tag{4.6}$$

We will say that sequences in $\mathscr{W}_3$ are *shift-orthogonal*.

If we apply Theorem (6.23) below to the operator T on the functions of a one-point space defined by $Tf = \lambda f$, where $|\lambda| = 1$, we see that for $\mathbf{a} \in \overline{\mathscr{W}}_3$,

$$N^{-1} \sum_{n=0}^{N-1} a_n \lambda^n \to 0 \quad \text{as } N \to \infty.$$

On the other hand, if $\mathbf{b} \in \|\cdot\|_2$-closure of $\mathscr{W}_0$, the usual arguments show we can find a countable family λ_j, $|\lambda_j| = 1$, and coefficients c_j, such that

$$c_j = \lim_{N\to\infty} N^{-1} \sum_{n=0}^{N-1} b_n \bar{\lambda}_j^{\,n}, \quad \text{and}$$

$$\|\mathbf{b} - \mathbf{b}(k)\|_2 \to 0 \text{ as } k \to \infty,$$

where

$$b_n(k) = \sum_{j=1}^{k} c_j \lambda_j^{\,n}.$$

Thus

$$\overline{\mathscr{W}}_3 \cap (\|\cdot\|_2\text{-closure } \mathscr{W}_0) = \{\mathbf{a} : \|\mathbf{a}\|_2 = 0\}.$$

We also wish to consider sequences $\mathbf{a}$ with (4.4) replaced by the stronger assumption:

$$\sum_{K=1}^{\infty} |I(K)|^{-1} < \infty. \tag{4.7}$$

Let

$$\mathscr{W}_4 = \text{complex linear span } \{\mathbf{a} : \|\mathbf{a}\|_2 < \infty, \text{ (4.1)–(4.3), (4.5), (4.7) hold}\}. \tag{4.8}$$

https://doi.org/10.4153/CJM-1983-010-7 Published online by Cambridge University Press

5. Outputs of mppt.

(5.1) LEMMA. *Let τ be a* mppt *on a probability space* $(X, \mathscr{F}, \mu)$, *T the induced operator on* $\mathscr{L}^2$. *Let V be the closed linear span of the eigenfunctions of T. Let A be the span of the bounded eigenfunctions of T. Then* $V = \bar{A}$. *A is an algebra closed under complex conjugation. If* $f \in \mathscr{L}^\infty(X, \mathscr{F}, \mu) \cap V$, *then there exists* $f_k \in A$ *with* $\|f_k\|_\infty \leqq \|f\|_\infty$ *for all n,* $\|f_k - f\|_2 \to 0$ *as* $k \to \infty$. *There exists a σ-algebra* $\mathscr{D} \subseteq \mathscr{F}$ *such that* $V = \mathscr{L}^2(X, \mathscr{D}, \mu)$. *In particular, orthogonal projection onto V defines an* $\mathscr{L}^\infty$*-contraction.*

Proof. This is a straightforward application of the Weierstrass approximation theorem.

(5.2) THEOREM. *Let* $(X, \mathscr{F}, \mu)$, τ, T, V *be as in Lemma* (5.1). *Assume τ is ergodic. Let* $W = V^\perp$. *Let f be bounded, measurable. Then* $f = g + h$, *where g, h are bounded, measurable,* $g \in V$, $h \in W$. *For* $\mu -$ a.e. x, *the sequence* $g(\tau^n(x))$ *is in* $\tilde{\mathscr{W}}_0$, *and the sequence* $h(\tau^n(x))$ *is in* $\mathscr{W}_3$.

Proof. g and h are bounded by Lemma (5.1). For $\mu -$ a.e. x, $g(\tau^n(x) \in \tilde{\mathscr{W}}_0$ by Lemma (5.1) and Lemma (2.12). (We note that since τ is ergodic, if $a_n(x) = g(\tau^n(x))$, then $\|\mathbf{a}(x)\|_\infty = \|g\|_\infty$ for $\mu -$ a.e. x.)

Since the spectral measure of h is continuous,

$$\lim_{M\to\infty} M^{-1} \sum_{m=1}^{M} \left| \int \bar{h}(h \circ \tau^m) d\mu \right|^2 = 0.$$

For any positive integer L, choose a positive integer $Q = Q(L)$, such that

$$Q^{-1} \sum_{m=1}^{L^2 Q} \left| \int \bar{h}(h \circ \tau^m) d\mu \right|^2 < L^{-1}.$$

Consider the sets $J_t = \{k, 2k, \ldots, Lk\}$, $k = LQ - t$, where t runs through the values $0, 1, \ldots, Q - 1$. It is easy to see that the sets J_t are pairwise disjoint, and each $J_t \subseteq \{1, 2, \ldots, L^2 Q\}$. If each J_t contains some m with

$$|\textstyle\int \bar{h}(h \circ \tau^m) d\mu|^2 \geqq L^{-1},$$

then, since there are Q sets J_t, a contradiction results. Hence for some set J_t,

$$|\textstyle\int \bar{h}(h \circ \tau^m) d\mu|^2 < L^{-1}, \text{ for every } m \in J_t.$$

Let $J(L)$ be one such set J_t. We have

$$|J(L)| = L, \text{ and}$$
$$|\textstyle\int \bar{h}(h \circ \tau^m) d\mu|^2 < L^{-1} \text{ for each } m \in J(L),$$

hence

$$|\textstyle\int \bar{h} \circ \tau^i)(h \circ \tau^j) d\mu|^2 < L^{-1} \text{ for each } i, j \in J(L) \text{ with } i \neq j.$$

https://doi.org/10.4153/CJM-1983-010-7 Published online by Cambridge University Press

Let $L_x(N)$ be the largest $L \leqq N$ such that

$$\sup J(L) \leqq N^{1/2} \text{ and}$$

$$\left| N^{-1} \sum_{n=0}^{N-1} \bar{h}(\tau^{i+n}(x))h(\tau^{j+n}(x)) \right| < L^{-1}$$

for each $i, j \in J(L)$, $i \neq j$. If no such L exists, let $L_x(N) = 1$. For μ — a.e. x,

$$\int (\bar{h} \circ \tau^i)(h \circ \tau^j) d\mu = \lim_{N\to\infty} N^{-1} \sum_{n=0}^{N-1} \bar{h}(\tau^{i+n}(x))h(\tau^{j+n}(x)),$$

for all i, j. Hence, for μ — a.e. x, any L, for N sufficiently large L satisfies the conditions just stated. Thus $L \leqq L_x(N)$. Hence $L_x(N) \to \infty$, μ — a.e. Let $I(N) = J(L_x(N))$. Then for μ — a.e. x, if $a_n = h(\tau^n(x))$, then (4.1)–(4.5) hold (with $N(K) = K$), so $\mathbf{a} \in \mathscr{W}_3$, and the theorem is proved.

If we consider $h \in \mathscr{L}^2(X, \mathscr{F}, \mu)$, rather than h bounded as in the theorem, the same proof just given shows $a_n = h(\tau^n(x))$ defines a sequence $\mathbf{a} \in \mathscr{W}_3$ for μ — a.e. x, once we note that $\|\mathbf{a}\|_2 < \infty$ by Lemma (2.12) with $r = p = 2$.

(5.3) THEOREM. *Let $\{\xi_n\}$ be a sequence of complex valued random variables. Let $r > 0$, $s < 2$. For each N, let $J(N)$ be a set of nonnegative integers. Suppose that*

(5.4) $\quad N^{-1} \sup J(N) \to 0$ *as* $N \to \infty$,

and that for N sufficiently large

(5.5) $\quad |J(N)| > N^r$

(5.6) $\quad E\left| \sum_{n=0}^{N-1} \bar{\xi}_{n+i}\xi_{n+j} \right|^2 \leqq$ (*constant*) N^s, *for all* $i, j \in J(N)$, $i \neq j$.

Then, with probability one, the sequence $a_n = \xi_n(\omega)$ lies in $\mathscr{W}_4$ if $\|\mathbf{a}\|_2 < \infty$.

Proof. We may assume $s + 6r < 2$, by decreasing r. Discarding part of $J(N)$, we may assume $|J(N)| \leqq N^{2r}$ for N sufficiently large. Let $t = r \wedge (2 - s - 6r)$. Let $N(K)$ = the least integer greater than $K^{2/t}$. For every K, and every $i, j \in J(N(K))$, $i \neq j$, let

$$A(K, i, j) = \left\{ \omega : \left| \sum_{n=0}^{N-1} \bar{\xi}_{n+i}(\omega)\xi_{n+j}(\omega) \right| \geqq N(K)^{1-r} \right\}.$$

Then

$$Pr(A(K, i, j)) \leqq (\text{constant})\ N(K)^{s+2r-2}.$$

https://doi.org/10.4153/CJM-1983-010-7 Published online by Cambridge University Press

Let $A(K) = \cup A(K, i, j)$, where the union is over all $i, j \in J(N(K))$ with $i \neq j$. Then

$$Pr(A(K)) \leqq (\text{constant})\ N(K)^{s+6r-2} \text{ for } K \text{ large.}$$

Hence

$$\sum_{K-1}^{\infty} Pr(A(K)) < \infty .$$

Let $I(K) = J(N(K))$. Then $|I(K)| \geqq K^2$ so (4.7) holds. Since with probability one each ω is in at most finitely many sets $A(K)$, it is easy to see that the sequence $a_n = \xi_n(\omega)$ satisfies (4.5). Thus the theorem is proved.

(5.7) COROLLARY. *Let* $(X, \mathscr{F}, \mu, \tau)$ *be an* iid *shift space. For any* $f \in \mathscr{L}^1(X, \mathscr{F}, \mu)$ *with* $\int f d\mu = 0$, *for* μ – a.e. x *the sequence* $a_n = f(\tau^n(x))$ *lies in* $\overline{\mathscr{W}}_4$.

Proof. If f is a bounded measurable function of finitely many coordinates, the statement follows readily from Theorem (5.3). Such functions are dense in $\mathscr{L}^1$, so the corollary follows from Lemma (2.12).

6. Convergence theorems. We will use the notation of Section 3. Let T be an operator on $\mathscr{L}^p(X, \mathscr{F}, \mu)$ which satisfies the ordinary pointwise ergodic theorem. The trick of Ryll-Nardzewski [**16**], already used in the proof of Theorem (2.19), suggests that we should then have $T \in \mathscr{T}(\mathscr{W}_0)$. The next theorem shows that, under mild restrictions on T, $T \in \mathscr{T}(\mathscr{W}_0)$ implies

$$T \in \mathscr{T}(\mathscr{W}_0\text{-seminorm closure } \mathscr{W}_0).$$

(6.1) THEOREM. *Let* $\mathscr{W}$ *and* $\mathscr{V}$ *be collections of complex sequences, where* $\mathscr{W}$ *is a real-linear space of complex sequence containing the constant sequence* 1. *Let* S *and* T *be linear operators on* $\mathscr{L}^p(X, \mathscr{F}, \mu)$ *for some* p, $1 \leqq p < \infty$. *Suppose*

$$|T^n f| \leqq (\text{constant}) S^n |f| \text{ for } f \in \mathscr{L}^p(X, \mathscr{F}, \mu) \text{ and all } n,$$

and suppose S *is power-bounded from* $\mathscr{L}^p(X, \mathscr{F}, \mu)$ *to* $\mathscr{L}^1(X, \mathscr{F}, \mu)$. *If* $T \in \mathscr{T}(\mathscr{V})$, $S \in \mathscr{T}(\mathscr{W})$, *then*

$$T \in \mathscr{T}(\mathscr{W}\text{-seminorm closure } \mathscr{V}).$$

Proof. Let $\mathbf{a} \in \mathscr{W}$-seminorm closure $\mathscr{V}$, $f \in \mathscr{L}^p(X, \mathscr{F}, \mu)$. Let

$$A(N)(x) = N^{-1} \sum_{n=0}^{N-1} a_n T^n f(x).$$

We must show $A(N)$ converges μ – a.e. Let $\mathbf{a}(k) \in \mathscr{V}$ with $\mathscr{W}$-seminorm

$(\mathbf{a}(k) - \mathbf{a}) \to 0$ as $k \to \infty$. Let

$$A(N, k)(x) = N^{-1} \sum_{n=0}^{N-1} a_n(k) T^n f(x).$$

Let

$$U(N) = \sup \{|A(M) - A(L)| : M, L \geqq N\},$$
$$U(N, k) = \sup \{|A(M, k) - A(L, k)| : M, L \geqq N\}.$$

Let

$$R(N, k)(x) = N^{-1} \sum_{n=0}^{N-1} |a_n(k) - a_n| \, |T^n f(x)|.$$

Let

$$Q(N, k) = \sup \{R(M, k) : M \geqq N\}.$$

We find at once

$$U(N) \leqq U(N, k) + 2Q(N, k).$$

$U(N)$, $Q(N, k)$ are decreasing in N, with limits U, $Q(k)$ respectively. Since $U(N, k) \to 0$ as $N \to \infty$, $U \leqq 2Q(k)$. We must show $U = 0$ $\mu -$ a.e. Thus it is enough to show

$$\int U d\mu = 0, \text{ or } \lim_{k \to \infty} \int Q(k) d\mu = 0.$$

Let $\mathbf{b} \in \mathscr{W}$ with $|a_n(k) - a_n| \leqq b_n$ for all n. Then we find

$$Q(k) \leqq (\text{constant}) \lim_{N \to \infty} N^{-1} \sum_{n=0}^{N-1} b_n S^n |f|, \ \mu - \text{a.e.}$$

Hence

$$\int Q(k) d\mu \leqq (\text{constant}) \overline{\lim_{N \to \infty}} N^{-1} \sum_{n=0}^{N-1} b_n ||S^n f||_1 \leqq (\text{constant}) \, ||b||_1.$$

Hence

$$\int Q(k) d\mu \leqq (\text{constant}) \mathscr{W}\text{-norm } (\mathbf{a}(k) - \mathbf{a}),$$

and the theorem follows.

(6.2) COROLLARY. *Let S, T be as in Theorem* (6.1). *If $S, T \in \mathscr{T}(\mathscr{W}_0)$ then $S, T \in \mathscr{T}(\mathscr{W}_2)$.*

Proof. This follows at once from (3.22).

(Naturally we can take $S = T$ and conclude $S \in \mathscr{T}(\mathscr{W}_2)$.) This result shows that for a wide class of operators, proving the ergodic theorem for uniform subsequences is no harder than proving the ordinary ergodic theorem. Bounded Besicovitch sequences do not seem as well behaved. However, we can prove the following result (we recall that by T *dominated* by S we mean $|Tf| \leqq S|f|$ for all f):

(6.3) THEOREM. *Let T be a linear operator on $\mathscr{L}^p(X, \mathscr{F}, \mu)$, for some p, $1 < p < \infty$, which is dominated by a positive $\mathscr{L}^p$-contraction. Let $p^{-1} + q^{-1} = 1$. Let $\mathbf{a} \in \|\cdot\|_s$-closure $\mathscr{W}_0$, for some $s > q$. Then $(\mathbf{a}, T)$ is Birkhoff.*

Proof. Let $r^{-1} + s^{-1} = 1$. Then $1 < r < p$. Let $f \in \mathscr{L}^p$. Let $\mathbf{b}(x)$ be defined by $b_n(x) = T^n f(x)$. By Lemma (2.12),

$$\|\mathbf{b}(x)\|_r < \infty \text{ for } \mu - \text{a.e. } x.$$

Let $\mathbf{a}(k) \in \mathscr{W}_0$, $\|\mathbf{a}(k) - \mathbf{a}\|_s \to 0$ as $k \to \infty$. $(\mathbf{a}(k), T)$ is Birkhoff since T is dominated by a positive $\mathscr{L}^p$-contraction. Thus for μ − a.e. x, $\mathbf{a}(k)\ \mathbf{b}(x)$ has a mean for all k, so $\mathbf{ab}(x)$ has a mean by (2.3) and (2.1), and the theorem is proved.

In particular the conclusion of Theorem (6.3) holds when $\mathbf{a} \in \tilde{\mathscr{W}}_0$.

(6.4) THEOREM. *Let T be a linear operator on $\mathscr{L}^\infty(X, \mathscr{F}, \mu)$ which is power bounded with respect to $\mathscr{L}^\infty$-norm and one other $\mathscr{L}^p$-norm, $1 \leq p < \infty$. Then $T \in \mathscr{T}(\mathscr{W}_4)$.*

Proof. By (2.16) it is enough to show $T \in \mathscr{T}(\mathscr{W}_4)$. Let $\mathbf{a}$ be a complex sequence with $\|\mathbf{a}\|_2 < \infty$, such that (4.1), (4.2), (4.3), (4.5), (4.7) hold. We will show first that

$$(6.5)\qquad \lim_{K\to\infty} N(K)^{-1} \sum_{n=0}^{N(K)-1} a_n T^n f = 0 \quad \mu - \text{a.e.}$$

Let C be such that

$$(6.6)\qquad \|T^n g\|_\infty \leqq C\|g\|_\infty,\ \|T^n g\|_p \leqq C\|g\|_p \text{ for } g \in \mathscr{L}^\infty.$$

We may assume

$$(6.7)\qquad \sum_{n=0}^{N-1} |a_n|^2 \leqq N \quad \text{for all } N.$$

Fix $f \in \mathscr{L}^\infty$. We may assume

$$(6.8)\qquad |T^n f| \leqq 1 \text{ everywhere on } X, \text{ for all } n.$$

Since $\bar{a}_n$ satisfies the same assumptions as a_n, to prove (6.5) it is enough to prove

$$(6.9)\qquad \lim_{K\to\infty} N(K)^{-1} \sum_{n=0}^{N(K)-1} \bar{a}_n T^n f = 0 \quad \mu - \text{a.e.}$$

For each K, define sequences $G(K)$, $G(K, j)$ by

$$(6.10)\qquad G_n(K) = a_n,\ n = 0, \ldots, N(K) - 1,\ G_n(K) = 0 \text{ for } n \geqq N(K),$$

$$(6.11)\qquad G_n(K, j) = a_{n+j},\ n = 0, \ldots, N(K) - 1,\ G_n(K, j) = 0 \quad \text{for } n \geqq N(K).$$

https://doi.org/10.4153/CJM-1983-010-7 Published online by Cambridge University Press

For any sequences G, F, write

$$\langle G, F\rangle = \sum_{n=0}^{\infty} \bar{G}_n F_n.$$

Let $F_n(x) = T^n f(x)$. Then (6.9) becomes

(6.12) $\lim_{K\to\infty} N(K)^{-1}\langle G(K), F\rangle = 0\ \mu - \text{a.e.}$

For any $\epsilon > 0$, by (4.5), for sufficiently large K, if $i, j \in I(K)$, $i \neq j$,

(6.13) $|\langle G(K, i), G(K, j)\rangle| \leqq \epsilon N(K)$.

Fix $\delta > 0$. Let $\epsilon = (\delta/3C)^2$. Let

(6.14) $A(K) = \{x : |\langle G(K), F(x)\rangle| > 2\delta N(K)\}$,

(6.15) $B(K, j) = \{x : |\langle G(K, j), F(x)\rangle| > 2\sqrt{\epsilon}N(K)\}$.

Fix x. Let J be a subset of $I(K)$, such that $x \in B(K, j)$ for each $j \in J$. Let v be the number of elements in J. Then $|\langle G(K, j), F(x)\rangle| > 2\sqrt{\epsilon}N(K)\ \forall\, j \in J$, so we can choose $\lambda(j)$, $|\lambda(j)| = 1$, such that if we set

$$H = \sum \lambda(j) G(K, j),$$

where the summation is over all $j \in J$, then

$$\langle H, F(x)\rangle > 2v\sqrt{\epsilon}N(K).$$

By (6.8),

$$\langle H, H\rangle \geqq 4v^2 \epsilon N(K).$$

By (6.7) and (4.3), for large K,

$$|\langle G(K, j), G(K, j)\rangle| \leqq 2N(K) \text{ for all } j \in I(K).$$

Hence, for large K, using (6.13) we have

$$\langle H, H\rangle \leqq 2vN(K) + v^2\epsilon N(K),$$

hence

(6.16) $v < \epsilon^{-1}$.

By (6.7), (6.8), and (4.3), for large K, for every $j \in I(K)$

$$\left|\sum_{n=0}^{N-1} a_{n+j} T^{n+j} f - \sum_{n=0}^{N-1} a_n T^n f\right| \leqq C\sqrt{\epsilon}N(K) \quad \text{everywhere on } X.$$

That is,

(6.17) $|T^j \langle G(K, j), F\rangle - \langle G(K), F\rangle| \leqq C\sqrt{\epsilon}N(K)$ everywhere on X.

Let $g_j = \langle G(K, j), F\rangle$. Clearly, for $j \in I(K)$

$$|T^j g_j| \leq |T^j \chi_{B(K,j)} g_j| + 2C\sqrt{\epsilon}N(K) \text{ a.e. on } X,$$

https://doi.org/10.4153/CJM-1983-010-7 Published online by Cambridge University Press

so

(6.18) $|T^j \chi_{B(K,j)} g_j| \geqq |\langle G(K), F\rangle| - 3C\sqrt{\epsilon}N(K)$ a.e. on X.

Hence

(6.19) $|T^j \chi_{B(K,j)} g_j| \geqq \delta N(K)$ a.e. on $A(K)$.

Hence

$$C\|\chi_{B(K,j)} g_j\|_p \geqq \delta N(K)(\mu(A(K)))^{1/p},$$

or, for large K, by (6.7) and (4.3), for each $j \in I(K)$,

$$C^p(2N(K))^p \mu(B(K,j)) \geq \delta^p(N(K))^p \mu(A(K)).$$

Summing over $j \in I(K)$, and using (6.16),

(6.20) $\delta^p |I(K)| \mu(A(K)) \leqq 2^p C^p \epsilon^{-2}$.

By (4.7),

$$\sum_{K=1}^{\infty} \mu(A(K)) < \infty,$$

so (6.9) holds, and hence (6.5) holds.

Let

$$S(N) = N^{-1} \sum_{n=0}^{N-1} a_n T^n f.$$

By (6.7) and (6.8), for $M \leqq N$ we have

(6.21) $|S(M) - S(N)| \leqq ((N - M)/N) + ((N - M)/N)^{1/2}$.

By (4.2), we see that

(6.22) $\lim_{N\to\infty} \sum_{n=0}^{N-1} a_n T^n f = 0 \quad \mu -$ a.e.

so the theorem is proved.

As usual, Remark (2.14) applies to Theorem (6.4).

(6.23) THEOREM. *Let T be as in Theorem* (6.4), $\mathbf{a} \in \overline{\mathscr{W}}_3$. *Then*

$$N^{-1} \sum_{n=0}^{N-1} a_n T^n f \to 0 \text{ in } \mathscr{L}^p\text{-norm as } N \to \infty,$$

for every $f \in \mathscr{L}^\infty$.

Proof. As usual, we need only consider $\mathbf{a} \in \mathscr{W}_3$. We find (6.20) still holds just as before, and so by (4.4), $\mu(A(K)) \to 0$ as $K \to \infty$. This implies

$$N(K)^{-1} \sum_{n=0}^{N(K)-1} a_n T^n f \to 0 \quad \text{in } \mathscr{L}^p\text{-norm as } K \to \infty.$$

The theorem then follows from (6.21) and (4.2).

We now note a simple lemma, which is sometimes useful in verifying that a sequence is in $\mathscr{W}_4$.

(6.24) LEMMA. *Let $N(K)$, $M(K)$ be sequences of positive real numbers, such that*

$$N(K) \to \infty, \; N(K+1)/(N(K) \to 1, \; M(K) \to \infty, \qquad M(K+1)/M(K) \to 1,$$

as $K \to \infty$. Then we can choose a subsequence $K(L)$ such that

$$N(K(L+1)) - N(K(L)) > 1, \; N(K(L)) \geqq M(L) \textit{ for all } L, \textit{ and}$$
$$N(K(L+1))/N(K(L)) \to 1 \textit{ as } L \to \infty.$$

7. Saturating sequences. Saturating sequences were defined by Reich in [**15**]. In that paper the convergence of $K^{-1} \sum_{k=0}^{K-1} T^{n(k)} f$ is considered, when $n(k)$ is a general sequence of integers, not necessarily increasing or positive. Such general sequences do not seem to fit naturally into the approach of the present paper, and we will confine outselves here to the case that $n(k)$ is an increasing subsequence of nonnegative integers. We will also assume that $n(k)$ has a nonzero density. Even in this case, we will not consider general saturating sequences, but only those satisfying a uniform order condition, as defined in [**15**], on every closed interval in $(0, 2\pi)$. We will use the terminology of weights rather than subsequences, as explained in Section 1.

Let $\mathbf{a}$ be a complex sequence. Let $N(K)$ be a sequence of positive integers such that

(7.1) $\quad N(K) \to \infty, \; N(K+1)/N(K) \to 1$ as $K \to \infty$.

For each $\epsilon > 0$, let

(7.2) $$B(K, \epsilon) = \sup \left\{ \left| N(K)^{-1} \sum_{n=0}^{N(K)-1} a_n \lambda^n \right|^2 : |\lambda| = 1, \; |\lambda - 1| > \epsilon \right\}.$$

Suppose

(7.3) $$\sum_{K=1}^{\infty} B(K, \epsilon) < \infty.$$

Let

(7.4) $\quad \mathscr{W}_5 =$ complex linear span $\{\mathbf{a} : \|\mathbf{a}\|_2 < \infty, \; (4.1)\text{–}(4.3) \text{ hold}\}$.

https://doi.org/10.4153/CJM-1983-010-7 Published online by Cambridge University Press

Applying the method of [**15**] one can show (at least) that $\mathscr{T}(\mathscr{W}_5)$ contains any T on $\mathscr{L}^\infty(X, \mathscr{F}, \mu)$ which is power bounded in $\mathscr{L}^\infty$-norm and unitary in $\mathscr{L}^2$-norm. The methods of [**15**] and [**4**] (cf. also [**5**]) give that if $\{\xi_n\}$ is an independent mean zero uniformly bounded sequence of complex valued random variables, then $a_n = \xi_n(\omega)$ defines a sequence in $\mathscr{W}_5$ with probability one.

The following example is considered in [**15**]. Let $\{\zeta_k\}$ be an iid sequence of positive integer valued random variables, with $E|\zeta_1| < \infty$. Suppose the distribution of ζ_1 is not supported on any proper subgroup of the integers. Let

$$n(k) = \zeta_1 + \ldots + \zeta_k.$$

It is shown that with probability one, the subsequence $\{n(k)\}$ is such that, for every operator T induced on $\mathscr{L}^1$ by a mppt,

$$(7.5) \qquad \lim_{K\to\infty} K^{-1} \sum_{k=0}^{K-1} T^{n(k)}f \quad \text{exists a.e., for } f \in \mathscr{L}^1.$$

For each sequence $n(k)$, let $a_n = 1$ if $n = n(k)$ for some k, $a_n = 0$ otherwise. As noted in Section 1, (7.5) is equivalent (since $\{n(k)\}$ has positive density) to

$$(7.6) \qquad \lim_{N\to\infty} N^{-1} \sum_{n=0}^{N-1} a_n T^n f \quad \text{exists a.e., for } f \in \mathscr{L}^1.$$

The proof of (7.5) in [**15**] (Theorem 6.4 of that paper) shows that $\mathbf{a} \in \mathscr{W}_5$. We note that also the sequence $\mathbf{b}$ obtained by subtracting the mean of $\mathbf{a}$ lies in $\overline{\mathscr{W}}_4$. To see this, we observe that the ζ_k may be regarded as return times for a Markov process. The result of Ornstein and Shields [**14**] then permits us to apply Corollary (5.7).

One step needed in [**4**] and [**5**] is an estimate for an expression of the form

$$E\left|\sum_{n=1}^{N} \exp\,(i\alpha S_n)\right|^4,$$

where $S_n = \zeta_1 + \ldots + \zeta_n$, for an iid sequence $\{\zeta_n\}$. Since the results of [**4**] and [**5**] are very interesting, it may be worthwhile to note a lemma which simplifies estimation of higher moments:

(7.7) LEMMA. *Let $\{\xi_n\}$ be a sequence of complex valued random variables such that for each m and n, $|\xi_{n+1} + \ldots + \xi_{n+m}|$ is independent of $|\xi_1 + \ldots + \xi_n|$, and $|\xi_{n+1} + \ldots + \xi_{n+m}|$ has the same moments as $|\xi_1 + \ldots + \xi_m|$. Suppose*

$$E|\xi_1 + \ldots + \xi_n|^2 \leqq An \text{ for all } n, \text{ and}$$

$$E|\xi_1|^l \leqq B^l, \text{ for some integer } l > 2.$$

https://doi.org/10.4153/CJM-1983-010-7 Published online by Cambridge University Press

Then

$$E|\xi_1 + \ldots + \xi_n|^l \leqq C_l(A^{l/2} + B^l)n^{l/2} \text{ for all } n,$$

where C_l *depends only on* l.

Proof. This follows by induction on l. To pass from l to $l + 1$, one considers $n = 2^j$, and again uses induction, on j.

References

1. M. A. Akcoglu, *A pointwise ergodic theorem in* $\mathscr{L}_p$*-spaces*, Can. J. Math. *27* (1975), 1975–1982.
2. M. A. Akcoglu and L. Sucheston, *Dilations of positive contractions on* $\mathscr{L}_p$*-spaces*, Can. J. Math. Bull. *20* (1977), 285–292.
3. A. Bellow, *Ergodic properties of isometries in* $\mathscr{L}_p$*-spaces*, $1 < p < \infty$, Bull. A.M.S. *70* (1964), 366–371.
4. J. R. Blum and R. Cogburn, *On ergodic sequences of measures*, Proc. A.M.S. *51* (1975), 359–365.
5. J. R. Blum and J. E. Reich, *p-sets for random walks*, Z. Wahrscheinlichkeitstheorie and Verw. Gebiete *48* (1979), 193–200.
6. A. Brunel and M. Keane, *Ergodic theorems for operator sequences*, Z. Wahrscheinlichkeitstheorie and Verw. Gebiete *12* (1969), 231–240.
7. A. de la Torre, *A simple proof of the maximal ergodic theorem*, Can. J. Math. *28* (1976), 1073–1075.
8. N. Dunford and J. Schwartz, *Linear operators I* (John Wiley, New York, 1958).
9. C. Kan, *Ergodic properties of Lamperti operators*, Can. J. Math. *30* (1978), 1206–1214.
10. P. Halmos and J. Von Neumann, *Operator methods in classical mechanics II*, Annals of Math. *43* (1942), 333–350.
11. J. H. Olsen, *Akcoglu's ergodic theorem for uniform sequences*, Can. J. Math. *32* (1980), 880–884.
12. ——— *The individual weighted ergodic theorem for bounded Besicovitch sequences*, Can. Bull. Math. *25* (1982), 468–471.
13. ——— *The individual ergodic theorem for Lamperti contractions*, C. R. Math. Rep. Acad. Sci. Canada *3* (1981), 113–118.
14. D. S. Ornstein and P. C. Shields, *Mixing Markov shifts of kernel type are Bernoulli*, Adv. Math. *10* (1973), 143–146.
15. J. I. Reich, *On the individual ergodic theorem for subsequences*, Annals of Prob. *5* (1977), 1039–1046.
16. C. Ryll-Nardzewski, *Topics in ergodic theory*, in Proceedings of the Winter School in Probability, Karpacz, Poland, 131–156, Lecture Notes in Mathematics *472* (Springer-Verlag, Berlin 1975).
17. R. Sato, *Operator averages for subsequences*, Math. J. of Okayama University *22* (1980).

University of Minnesota,
Minneapolis, Minnesota;
North Dakota State University,
Fargo, North Dakota

https://doi.org/10.4153/CJM-1983-010-7 Published online by Cambridge University Press