

Ergodic and mixing sequences of transformations

DANIEL BEREND AND VITALY BERGELSON

Department of Mathematics, University of California, Los Angeles, CA 90024, USA;
Institute of Mathematics, The Hebrew University of Jerusalem, Jerusalem, Israel

(Received 23 March 1983)

Abstract. The notions of ergodicity, strong mixing and weak mixing are defined and studied for arbitrary sequences of measure-preserving transformations of a probability space. Several results, notably ones connected with mean ergodic theorems, are generalized from the case of the sequence of all powers of a single transformation to this case. The conditions for ergodicity, strong mixing and weak mixing of sequences of affine transformations of compact groups are investigated.

1. Introduction

Ergodic theory is usually concerned with studying the properties of the sequence of powers of a measure-preserving transformation or, more generally, an arbitrary semigroup of measure-preserving transformations. The fact that the acting set of transformations is closed under composition turns out to play a crucial role in obtaining results of various kinds.

In this paper we introduce notions of ergodicity, strong mixing and weak mixing for arbitrary sequences $(T_n)_{n=1}^{\infty}$ of measure-preserving transformations of a probability space $(X, \mathcal{B}, \mu)$. These notions generalize those of the classical setup – T is ergodic (strongly mixing, weakly mixing) in the usual sense iff $(T^n)_{n=0}^{\infty}$ is such according to our definition.

In § 2 the definition of ergodicity is presented. Ergodicity is shown to be equivalent to the mean ergodic theorem (the limit function being constant). A criterion for verifying the ergodicity of given sequences is given.

Strongly mixing sequences are studied in § 3. A generalization of the Blum–Hanson theorem [2], by which a transformation T is strongly mixing iff

$$\frac{1}{N} \sum_{k=1}^N T^{n_k} f \xrightarrow[N \rightarrow \infty]{L^p} \int_X f d\mu$$

for any increasing sequence $(n_k)_{k=1}^{\infty}$ and $f \in L^p$, is established. A particular sequence, shown by Adler and Rivlin [1] to be strongly mixing in some sense, turns out to be strongly mixing in our (stronger) sense as well. A uniform sweeping out property, proved by Friedman [4] for a sequence of powers of a strongly mixing transformation, is valid for any strongly mixing sequence.

§ 4 deals with weakly mixing sequences. Equivalents of weak mixing in the regular case have sequential counterparts. Weakly mixing sequences have strongly mixing

subsequences. If T is weakly mixing then numerous subsequences of $(T^n)_{n=0}^\infty$, for example $(T^{n^2})_{n=0}^\infty$, form weakly mixing sequences.

In § 5 sequences of affine transformations of compact groups are studied. A sequence of translations is ergodic iff it corresponds to a uniformly distributed sequence in the group. Sequences of affine transformations are strongly (or weakly) mixing iff the sequences of epimorphisms lying below are such. Particular emphasis is given to sequences of epimorphisms, especially under further restrictions on the underlying group. As a special case we derive a result of Raikov [12], which states that a sequence of distinct endomorphisms of the circle group satisfies the mean ergodic theorem.

We wish to express our gratitude to H. Furstenberg for valuable discussions on this subject and to the referee for his remarks.

2. Ergodic sequences

In this section we extend the notion of ergodicity from the setup of a *dynamical system*, i.e. a quadruple $(X, \mathcal{B}, \mu, T)$ consisting of a probability space $(X, \mathcal{B}, \mu)$ and a measure-preserving transformation T , to a more general one.

Definition 2.1. A *sequential dynamical system* is a quadruple $(X, \mathcal{B}, \mu, \tilde{T})$, where $(X, \mathcal{B}, \mu)$ is a probability space and $\tilde{T} = (T_1, T_2, \dots)$ a sequence of measure-preserving transformations thereof.

Definition 2.2. A sequential dynamical system $(X, \mathcal{B}, \mu, \tilde{T})$ is *ergodic* if

$$\frac{1}{N^2} \sum_{m,n=1}^N \mu(T_m^{-1}A \cap T_n^{-1}B) \xrightarrow{N \rightarrow \infty} \mu(A)\mu(B), \quad A, B \in \mathcal{B}. \quad (2.1)$$

Alternatively, we say that $\tilde{T}$ is ergodic. Now we show that ergodicity is equivalent to the mean ergodic theorem. Denoting by $p(T)$ the sequence $(T^n)_{n=0}^\infty$ for a transformation T , we can easily infer from the following theorem that a dynamical system $(X, \mathcal{B}, \mu, T)$ is ergodic iff the corresponding sequential system $(X, \mathcal{B}, \mu, p(T))$ is.

THEOREM 2.1. Let $(X, \mathcal{B}, \mu, \tilde{T})$ be a sequential dynamical system. The following conditions are equivalent:

- (1) $\tilde{T}$ is ergodic.
- (2) $(1/N^2) \sum_{m,n=1}^N \mu(T_m^{-1}A \cap T_n^{-1}A) \xrightarrow{N \rightarrow \infty} \mu(A)^2, \quad A \in \mathcal{B}.$
- (3) For every $1 \leq p < \infty$ and $f \in L^p$

$$\frac{1}{N} \sum_{n=1}^N T_n f \xrightarrow{N \rightarrow \infty} \int_X f d\mu. \quad (2.2)$$

- (4) The former property holds for some $1 \leq p < \infty$.

Proof. The implications (1) $\Rightarrow$ (2) and (3) $\Rightarrow$ (4) are trivial. Now assume (2). For $f = 1_A$ we get (2.2) by

$$\left\| \frac{1}{N} \sum_{n=1}^N T_n 1_A - \int_X 1_A d\mu \right\|_p^{2p} \leq \left\| \frac{1}{N} \sum_{n=1}^N T_n 1_A - \mu(A) \right\|_1^2$$

$$\begin{aligned} &\leq \left\| \frac{1}{N} \sum_{n=1}^N T_n 1_A - \mu(A) \right\|_2^2 \\ &= \frac{1}{N^2} \sum_{m,n=1}^N \mu(T_m^{-1}A \cap T_n^{-1}A) - \mu(A)^2 \xrightarrow{N \rightarrow \infty} 0. \end{aligned}$$

The set of all $f \in L^p$ for which (2.2) holds forms a closed subspace of L^p . Since it contains the set of characteristic functions, this subspace is L^p itself. Thus (2) $\Rightarrow$ (3).

Assuming (4) we have in particular

$$\frac{1}{N} \sum_{n=1}^N T_n 1_A \xrightarrow[N \rightarrow \infty]{L^p} \mu(A),$$

for any $A \in \mathcal{B}$. In view of the Jensen inequality the same convergence takes place in L^1 , and hence in L^2 as well. Thus

$$\begin{aligned} \frac{1}{N^2} \sum_{m,n=1}^N \mu(T_m^{-1}A \cap T_n^{-1}B) &= \int_X \frac{1}{N} \sum_{m=1}^N T_m 1_A \cdot \frac{1}{N} \sum_{n=1}^N T_n 1_B d\mu \\ &\xrightarrow[N \rightarrow \infty]{} \int_X \mu(A) \mu(B) d\mu = \mu(A) \mu(B), \end{aligned}$$

which gives (1). This completes the proof. $\square$

Remark 2.1. In view of various equivalents of ergodicity in the usual case (see, for example, [3, Ch. 5]) another definition of ergodicity might have seemed rather more natural than ours. Namely, instead of the ‘double averaging’ of (2.1) consider the property

$$\frac{1}{N} \sum_{n=1}^N \mu(T_n^{-1}A \cap B) \xrightarrow[N \rightarrow \infty]{} \mu(A) \mu(B), \quad A, B \in \mathcal{B}. \quad (2.3)$$

While (2.1) can be easily shown to imply (2.3), the converse is false (see remark 3.1). We chose (2.1) as our definition of ergodicity so as to retain the equivalence to the mean ergodic theorem.

Denote by H_0 the closed subspace of L^2 consisting of those $f \in L^2$ with $\int_X f d\mu = 0$. The orthogonal complement of H_0 is the subspace H_C of all constant functions.

LEMMA 2.1. *Let F be a subset of H_0 spanning a dense subspace of H_0 . $\tilde{T}$ is ergodic iff*

$$\frac{1}{N^2} \sum_{m,n=1}^N \int_X f(T_m x) \overline{f(T_n x)} d\mu(x) \xrightarrow[N \rightarrow \infty]{} 0, \quad f \in F. \quad (2.4)$$

In fact, the validity of (2.4) for any $f \in H_0$ is equivalent to

$$\frac{1}{N} \sum_{n=1}^N T_n f \xrightarrow[N \rightarrow \infty]{L^2} 0.$$

Inasmuch as the set of all L^2 functions for which (2.4) holds is always a closed subspace of L^2 containing H_C , this proves the lemma.

PROPOSITION 2.1. *If $\tilde{T}$ is ergodic and $A \in \mathcal{B}$ with $\mu(A) > 0$ then*

$$\mu(\overline{\lim_{n \rightarrow \infty} T_n^{-1}A}) = 1.$$

The proof follows easily from theorem 2.1. We remark that in general the converse fails to be true. Moreover, one can construct sequences possessing the property in question admitting no ergodic subsequences.

3. Strongly mixing sequences

In this section the notion of strong mixing will be defined and studied for sequences of transformations.

Definition 3.1. Let A be any set and $B \subseteq A \times A$. B is of *bounded fibres* if there exists some c such that for every $a_1 \in A$ the set B contains at most c elements of the form (a_1, a_2) with $a_2 \in A$.

Definition 3.2. A sequential dynamical system $(X, \mathcal{B}, \mu, \tilde{T})$ is *strongly mixing* if for any $A, B \in \mathcal{B}$ and $\varepsilon > 0$ the set of solutions (m, n) of

$$|\mu(T_m^{-1}A \cap T_n^{-1}B) - \mu(A)\mu(B)| \geq \varepsilon \quad (3.1)$$

is of bounded fibres.

Evidently, a dynamical system $(X, \mathcal{B}, \mu, T)$ is strongly mixing iff the corresponding sequential dynamical system $(X, \mathcal{B}, \mu, p(T))$ is.

THEOREM 3.1. Let $(X, \mathcal{B}, \mu, \tilde{T})$ be a sequential dynamical system. The following conditions are equivalent:

- (1) $\tilde{T}$ is strongly mixing.
- (2) For any $A \in \mathcal{B}$ and $\varepsilon > 0$ the set of solutions (m, n) of

$$|\mu(T_m^{-1}A \cap T_n^{-1}A) - \mu(A)^2| \geq \varepsilon$$

is of bounded fibres.

- (3) For any $1 \leq p < \infty$, $f \in L^p$ and $\varepsilon > 0$ there exists a K such that if $N \geq K$ and $n_1 < n_2 < \dots < n_N$ are positive integers then

$$\left\| \frac{1}{N} \sum_{i=1}^N T_{n_i} f - \int_X f d\mu \right\|_p < \varepsilon.$$

- (4) Every subsequence of $\tilde{T}$ is ergodic.

Proof. The implication (1) $\Rightarrow$ (2) is trivial, and (2) $\Rightarrow$ (3) is proved as the analogue in theorem 2.1. (3) clearly implies that every subsequence of $\tilde{T}$ satisfies the mean ergodic theorem, so that every subsequence is ergodic. Thus (3) $\Rightarrow$ (4).

It remains to show that (4) $\Rightarrow$ (1). Suppose (1) is not satisfied. Select $A, B \in \mathcal{B}$ and $\varepsilon > 0$ such that the number of solutions n of (3.1) is unbounded as a function of m . Denote by $\#(F)$ the number of elements of a finite set F . Replacing B by B^c if necessary, we may assume that for any positive integer s there exist an m and a set S of positive integers with $\#(S) = s$ such that

$$\mu(T_m^{-1}A \cap T_n^{-1}B) - \mu(A)\mu(B) \geq \varepsilon, \quad n \in S. \quad (3.2)$$

This gives

$$\begin{aligned}
 \frac{1}{s^2} \sum_{i,j \in S} \mu(T_i^{-1}B \cap T_j^{-1}B) &= \mu(B)^2 + \frac{1}{s^2} \int_X \left(\sum_{i \in S} 1_{T_i^{-1}B} - s\mu(B) \right)^2 d\mu \\
 &\geq \mu(B)^2 + \frac{1}{s^2} \int_X \left(\sum_{i \in S} 1_{T_i^{-1}B} 1_{T_m^{-1}A} - s\mu(B) 1_{T_m^{-1}A} \right)^2 d\mu \\
 &\geq \mu(B)^2 + \frac{1}{s^2} \left(\sum_{i \in S} \mu(T_i^{-1}B \cap T_m^{-1}A) - s\mu(B)\mu(A) \right)^2 \\
 &\geq \mu(B)^2 + \varepsilon^2.
 \end{aligned} \tag{3.3}$$

Let $(m_k)_{k=1}^\infty$ be a sequence of integers and $(S_k)_{k=1}^\infty$ a sequence of finite sets of integers such that (3.2) holds with $m = m_k$ and $S = S_k$ for every k . The sets S_k may so be chosen that $\max S_k < \min S_{k+1}$. Consider the subsequence of $\tilde{T}$ corresponding to $\bigcup_{k=1}^\infty S_k$. In view of (3.3) if the sequence $(\#(S_k))_{k=1}^\infty$ is of sufficiently rapid growth then the resulting subsequence of $\tilde{T}$ is non-ergodic. It follows that (4) $\Rightarrow$ (1), which completes the proof. $\square$

COROLLARY I (Rényi [13]). *A dynamical system $(X, \mathcal{B}, \mu, T)$ is strongly mixing iff $\mu(A \cap T^{-n}A) \xrightarrow{n \rightarrow \infty} \mu(A)^2$ for every $A \in \mathcal{B}$.*

COROLLARY II (Blum–Hanson [2]). *$(X, \mathcal{B}, \mu, T)$ is strongly mixing iff*

$$\frac{1}{N} \sum_{k=1}^N T^{n_k} f \xrightarrow[N \rightarrow \infty]{L^p} \int_X f d\mu,$$

for every $1 \leq p < \infty$, $f \in L^p$ and increasing $(n_k)_{k=1}^\infty$.

LEMMA 3.1. *Let F be as in lemma 2.1. $\tilde{T}$ is strongly mixing iff for any $f \in F$ and $\varepsilon > 0$ the set*

$$\left\{ (m, n): \left| \int_X f(T_m x) \overline{f(T_n x)} d\mu(x) \right| \geq \varepsilon \right\}$$

is of bounded fibres.

The lemma follows from lemma 2.1 and theorem 3.1.

Example 3.1. In [1] the following sequence of transformations is examined:

$$T_n(x) = \cos(n \arccos x), \quad x \in [-1, 1], n = 1, 2, \dots$$

It is proved there that T_n is measure-preserving with respect to the measure given by

$$d\mu = \frac{1}{\pi \sqrt{1-x^2}} dx,$$

and also that

$$\mu(T_n^{-1}A \cap B) \xrightarrow{n \rightarrow \infty} \mu(A)\mu(B), \quad A, B \in \mathcal{B} \tag{3.4}$$

which property is there called strong mixing. In fact, employing lemma 3.1 and the basis of $L^2[-1, 1]$ considered in that paper we can show $\tilde{T}$ to be strongly mixing

in our sense as well. Another way of proving this fact is by noting that $\tilde{T}$ is a factor of the sequence $\tilde{\sigma}$ of endomorphisms of the circle group $\mathbb{T}$ given by $\sigma_n(x) = nx$ and using the results to be presented in § 5. In fact, the mapping $\varphi: \mathbb{T} \rightarrow [-1, 1]$ defined by $\varphi(x) = \cos(2\pi x)$, $x \in [0, 1]$, exhibits $\tilde{T}$ as a factor of $\tilde{\sigma}$.

Remark 3.1. The reason for taking definition 3.1 as our definition for strong mixing rather than the seemingly more natural property (3.4) is that this way the Blum–Hanson theorem remains valid for sequential systems. Property (3.4), on the other hand, was shown by Friedman [4, 4.6] not to imply the mean ergodic theorem even for sequences of strictly increasing powers of a single transformation.

PROPOSITION 3.1. *If $\tilde{T}$ is strongly mixing and $A \in \mathcal{B}$ with $\mu(A) > 0$ then for every $\varepsilon > 0$ there exists a K such that if $N \geq K$ and $n_1 < n_2 < \dots < n_N$ are positive integers then*

$$\mu\left(\bigcup_{i=1}^N T_{n_i}^{-1}A\right) > 1 - \varepsilon.$$

The proposition follows easily from theorem 3.1. It was proved by Friedman [4] for subsequences of $p(T)$, where T is strongly mixing.

4. Weakly mixing sequences

In this section we shall study weakly mixing systems. Several definitions are needed first. The *density* (*upper density*, *lower density*) of a set M of positive integers is the limit (limit superior, limit inferior, respectively) of the sequence

$$\left(\frac{1}{n} \#(M \cap \{1, 2, \dots, n\})\right)_{n=1}^{\infty}.$$

An analogous notion applies to subsets of $\mathbb{N}^2$, where now the proportion of the elements of M in squares $\{1, 2, \dots, n\}^2$ is examined. A double sequence $(a_{mn})_{m,n=1}^{\infty}$ converges in density to a if

$$a_{mn} \xrightarrow[m, n \rightarrow \infty]{(m, n) \in \mathbb{N}^2 - J} a$$

for some zero density set J in $\mathbb{N}^2$. In this case we write $a_{mn} \xrightarrow[m, n \rightarrow \infty]{D} a$. The following lemma is proved similarly to one-dimensional analogues.

LEMMA 4.1. *Let $(a_{mn})_{m,n=1}^{\infty}$ be a bounded sequence. The following conditions are equivalent:*

- (1) $a_{mn} \xrightarrow[m, n \rightarrow \infty]{D} a$.
- (2) The set $\{(m, n): |a_{mn} - a| \geq \varepsilon\}$ is of zero density for any $\varepsilon > 0$.
- (3) $(1/N^2) \sum_{m,n=1}^N |a_{mn} - a| \xrightarrow[N \rightarrow \infty]{} 0$.

Definition 4.1. A sequential dynamical system $(X, \mathcal{B}, \mu, \tilde{T})$ is *weakly mixing* if

$$\mu(T_m^{-1}A \cap T_n^{-1}B) \xrightarrow[m, n \rightarrow \infty]{D} \mu(A)\mu(B) \quad A, B \in \mathcal{B}.$$

The product of two sequential systems $(X, \mathcal{B}, \mu, \tilde{T})$ and $(Y, \mathcal{C}, \nu, \tilde{S})$ is the system

$$(X \times Y, \mathcal{B} \times \mathcal{C}, \mu \times \nu, \tilde{T} \times \tilde{S})$$

where $(T_n \times S_n)(x, y) = (T_n x, S_n y)$.

THEOREM 4.1. Let $(X, \mathcal{B}, \mu, \tilde{T})$ be a sequential dynamical system. The following conditions are equivalent:

- (1) $\tilde{T}$ is weakly mixing.
- (2) $\mu(T_m^{-1}A \cap T_n^{-1}A) \xrightarrow{m, n \rightarrow \infty} \mu(A)^2$, $A \in \mathcal{B}$.
- (3) For any $A, B \in \mathcal{B}$ and $\delta, \varepsilon > 0$ there exists a K such that for any $N \geq K$ and m the inequality

$$|\mu(T_m^{-1}A \cap T_n^{-1}B) - \mu(A)\mu(B)| \geq \varepsilon$$

has at most δN solutions n with $1 \leq n \leq N$.

- (4) For any $1 \leq p < \infty$, $f \in L^p$ and $\delta, \varepsilon > 0$ there exists a K such that if $N \geq K$ and $n_1 < n_2 < \dots < n_N \leq N/\delta$ are positive integers then

$$\left\| \frac{1}{N} \sum_{i=1}^N T_{n_i} f - \int_X f d\mu \right\|_p < \varepsilon.$$

- (5) Every positive lower density subsequence of $\tilde{T}$ is ergodic.
- (6) $\tilde{T} \times \tilde{T}$ is ergodic.
- (7) $\tilde{T} \times \tilde{S}$ is ergodic for any ergodic $\tilde{S}$.

Proof. The implications $(1) \Rightarrow (2) \Rightarrow (4) \Rightarrow (5)$ are proved similarly to their analogues in theorem 2.1.

$(1) \Rightarrow (7)$ is routine, where we note that it suffices to check the ergodicity condition (2.1) for pairs of measurable rectangles in $\mathcal{B} \times \mathcal{C}$. Now assume (7). Taking $\tilde{S}$ to be trivial we see that $\tilde{T}$ is ergodic. Taking now $\tilde{S} = \tilde{T}$ we obtain (6). From (6) it follows, by just writing down the ergodicity condition for the pair of sets $A \times A$, $A \times A$, that

$$\frac{1}{N^2} \sum_{m, n=1}^N \mu(T_m^{-1}A \cap T_n^{-1}A)^2 \xrightarrow{N \rightarrow \infty} \mu(A)^4.$$

Hence:

$$\begin{aligned} & \frac{1}{N^2} \sum_{m, n=1}^N (\mu(T_m^{-1}A \cap T_n^{-1}A) - \mu(A)^2)^2 \\ &= \frac{1}{N^2} \sum_{m, n=1}^N \mu(T_m^{-1}A \cap T_n^{-1}A)^2 - \frac{2}{N^2} \mu(A)^2 \int_X \left(\sum_{n=1}^N 1_{T_n^{-1}A} \right)^2 d\mu + \mu(A)^4 \\ &\leq \frac{1}{N^2} \sum_{m, n=1}^N \mu(T_m^{-1}A \cap T_n^{-1}A)^2 - \frac{2}{N^2} \mu(A)^2 \left(\sum_{n=1}^N \int_X 1_{T_n^{-1}A} d\mu \right)^2 + \mu(A)^4 \xrightarrow{N \rightarrow \infty} 0. \end{aligned}$$

Being non-negative, the left hand side also converges to 0. By lemma 4.1 we get $\mu(T_m^{-1}A \cap T_n^{-1}A) \xrightarrow{m, n \rightarrow \infty} \mu(A)^2$, and so $(6) \Rightarrow (2)$.

$(3) \Rightarrow (1)$ is trivial, and so it only remains to show that $(5) \Rightarrow (3)$. Suppose (3) is not valid. Then there exist $A, B \in \mathcal{B}$, $\delta, \varepsilon > 0$ and sequences $(m_k)_{k=1}^\infty$ and $(N_k)_{k=1}^\infty$

with $N_{k+1}/N_k \xrightarrow[k \rightarrow \infty]{} \infty$, such that for each k the inequality

$$\mu(T_{m_k}^{-1}A \cap T_n^{-1}B) \geq \mu(A)\mu(B) + \varepsilon \quad (4.1)$$

has at least δN_k solutions $N_{k-1} < n \leq N_k$. Assume that for some fixed k (4.1) holds for $n = n_1, n_2, \dots, n_s$. From the proof of theorem 3.1 we see that

$$\frac{1}{s^2} \sum_{i,j=1}^s \mu(T_{n_i}^{-1}B \cap T_{n_j}^{-1}B) \geq \mu(B)^2 + \varepsilon^2.$$

Consider the subsequence of $\tilde{T}$ corresponding to all these solutions n taken for all positive integers k . This subsequence is obviously non-ergodic, but whereas it is necessarily of positive upper density it may be of zero lower density. Augment it therefore by adjoining, say, all the transformations $(T_{dn})_{n=1}^\infty$ for some positive integer d . The resulting subsequence is of positive lower density and, if d is chosen large enough, still non-ergodic. Thus (5) $\Rightarrow$ (3), which completes the proof. $\square$

A system $(X, \mathcal{B}, \mu, T)$ is weakly mixing iff $(X, \mathcal{B}, \mu, p(T))$ is. We conclude therefore

COROLLARY (Jones [9]). *$(X, \mathcal{B}, \mu, T)$ is weakly mixing iff*

$$\frac{1}{N} \sum_{k=1}^N T^{n_k} f \xrightarrow[N \rightarrow \infty]{L^p} \int_X f d\mu$$

for any $1 \leq p < \infty$, $f \in L^p$ and positive lower density sequence $(n_k)_{k=1}^\infty$.

THEOREM 4.2. *A weakly mixing sequence on a separable probability space admits a strongly mixing subsequence.*

Proof. Suppose $(X, \mathcal{B}, \mu)$ is separable and $\tilde{T}$ weakly mixing. Let $(A_n)_{n=1}^\infty$ be dense in $\mathcal{B}$. Define a sequence $(n_k)_{k=1}^\infty$ inductively. Choose n_1 arbitrarily. Suppose $n_1, n_2, \dots, n_k$ have been selected. Condition (3) of theorem 4.1 secures the existence of some $n_{k+1} > n_k$ with

$$|\mu(T_{n_i}^{-1}A_j \cap T_{n_{k+1}}^{-1}A_j) - \mu(A_j)^2| < 1/k, \quad 1 \leq i, j \leq k.$$

The sequence $(T_{n_k})_{k=1}^\infty$ is obviously strongly mixing. This proves the theorem. $\square$

Remark 4.1. It may happen that a weakly mixing sequence admits only zero density strongly mixing subsequences. Such is the case, for example, with the sequence $\tilde{\sigma}$ of epimorphisms of $\mathbb{T}$ given by

$$\sigma_n(x) = [n^{1/2}]x, \quad x \in \mathbb{T}, n = 1, 2, \dots$$

One can construct examples of sequences $(n_k)_{k=1}^\infty$ of zero density with the property that $(T^{n_k})_{k=1}^\infty$ is a weakly mixing sequence for any weakly mixing transformation T . For example, if p is a non-constant polynomial with rational coefficients such that $p(n) \in \mathbb{N}$ for every $n \in \mathbb{N}$ then $(T^{p(n)})_{n=1}^\infty$ is weakly mixing if T is. The proof of this fact goes along the lines elaborated in [5, § 3.4].

5. Sequences of affine transformations

Throughout this section G denotes a compact Hausdorff group, $\mathcal{B}$ the σ -algebra generated by the open sets and μ the Haar measure. A transformation of G of the form $\pi(x) = \sigma(x)\alpha$, where σ is a continuous endomorphism of G and $\alpha \in G$ is

measure-preserving iff σ is epic. In this case π is called an *affine transformation*. Let $\tilde{\pi} = \tilde{\sigma}\tilde{\alpha}$ be a sequence of affine transformations, that is

$$\pi_n(x) = \sigma_n(x)\alpha_n, \quad x \in G, n = 1, 2, \dots \quad (5.1)$$

In this section we investigate the conditions for ergodicity and mixing of $\tilde{\pi}$, especially under further restrictions on $\tilde{\pi}$ and G .

$\mathcal{U}$ will denote the set of all continuous irreducible unitary representations of G . I stands for the trivial representation. Given a representation $U \in \mathcal{U}$ we denote by d_U the dimension of the representation space of U and by u_{ij} , $1 \leq i, j \leq d_U$, the coordinate functions corresponding to U . If U is the only relevant representation we put $d = d_U$.

LEMMA 5.1. $\tilde{\pi}$ is ergodic iff

$$\frac{1}{N^2} \sum_{m,n=1}^N \sum_{s,t=1}^d u_{sj}(\alpha_m) \overline{u_{it}(\alpha_n)} \int_G u_{is}(\sigma_m(x)) \overline{u_{it}(\sigma_n(x))} d\mu(x) \xrightarrow{N \rightarrow \infty} 0, \quad (5.2)$$

for $I \neq U \in \mathcal{U}$, $1 \leq i, j \leq d$.

The lemma follows in a straightforward manner from lemma 2.1 observing that, in view of the Peter–Weyl theorem [7], the set of coordinate functions corresponding to all the non-trivial representations in $\mathcal{U}$ enjoys the properties of the set F of that lemma.

THEOREM 5.1. A sequence of translations of G

$$\pi_n(x) = x\alpha_n, \quad x \in G, n = 1, 2, \dots$$

is ergodic iff $\tilde{\alpha}$ is uniformly distributed in G .

Proof. By lemma 5.1 the ergodicity criterion in our case takes the form:

$$\frac{1}{N^2} \sum_{m,n=1}^N \sum_{s=1}^d \frac{1}{d} u_{sj}(\alpha_m) \overline{u_{sj}(\alpha_n)} \xrightarrow{N \rightarrow \infty} 0, \quad I \neq U \in \mathcal{U}, 1 \leq j \leq d.$$

Equivalently:

$$\frac{1}{d} \sum_{s=1}^d \left| \frac{1}{N} \sum_{n=1}^N u_{sj}(\alpha_n) \right|^2 \xrightarrow{N \rightarrow \infty} 0, \quad I \neq U \in \mathcal{U}, 1 \leq j < d.$$

Thus $\tilde{\pi}$ is ergodic iff

$$\frac{1}{N} \sum_{n=1}^N u_{sj}(\alpha_n) \xrightarrow{N \rightarrow \infty} 0, \quad I \neq U \in \mathcal{U}, 1 \leq j, s \leq d$$

which is precisely the condition for $\tilde{\alpha}$ to be uniformly distributed [10, Ch. 4]. $\square$

Remark 5.1. The ‘if’ part of the theorem can be proved without invoking the ergodicity criterion. In fact, if $\tilde{\alpha}$ is uniformly distributed then

$$\frac{1}{N} \sum_{n=1}^N f(x\alpha_n) \xrightarrow{N \rightarrow \infty} \int_G f d\mu, \quad f \in C(G), x \in G$$

where $C(G)$ is the space of complex-valued continuous functions on G . The same convergence holds in L^2 . It follows that the mean ergodic theorem is satisfied for

every $f \in L^2(G)$, so that the sequence is ergodic. The theorem shows that in our case these types of convergence, namely that of pointwise convergence for continuous functions and that of mean convergence for L^2 functions, are equivalent.

We turn to study the conditions for ergodicity of sequences of epimorphisms. First we need:

LEMMA 5.2. *Let $(a_n)_{n=1}^\infty$ be a sequence in some set A . The following conditions are equivalent:*

$$(1) \max_{a \in A} \# \{n \mid 1 \leq n \leq N, a_n = a\} = o(N). \quad (5.3)$$

$$(2) \# \{(m, n) \mid 1 \leq m, n \leq N, a_m = a_n\} = o(N^2). \quad (5.4)$$

The proof is routine.

Definition 5.1. A sequence $(a_n)_{n=1}^\infty$ in a set A is of *reasonable multiplicity* if it satisfies either one of the equivalent conditions (5.3) and (5.4).

Let Γ denote the dual object of G , i.e. the set of all equivalence classes of representations in $\mathcal{U}$ (see, for example, [7]). An epimorphism σ of G gives rise to a dual one-to-one transformation of $\mathcal{U}$, which induces in turn a transformation of Γ . The dual of σ is usually also denoted by σ , so to avoid confusion we write $\sigma(x)$ for the action of σ on $x \in G$ and $U\sigma$ or $\gamma\sigma$ for its action on representations.

THEOREM 5.2. *A sequence $\vec{\sigma}$ of epimorphisms of G is ergodic iff the sequence $\gamma\vec{\sigma}$ is of reasonable multiplicity for any non-trivial $\gamma \in \Gamma$.*

Proof. The sufficiency is a consequence of lemma (5.1) using the variant (5.4) of reasonable multiplicity. For the converse direction assume that for some non-trivial $\gamma \in \Gamma$ the sequence $\gamma\vec{\sigma}$ is not of reasonable multiplicity. Let $U \in \gamma$. Denote by χ_U the character corresponding to U . A direct computation shows that $f = \chi_U$ does not satisfy (2.4), and so $\vec{\sigma}$ is non-ergodic. This proves the theorem. $\square$

In what follows, $\vec{\pi}$ is always as in (5.1).

PROPOSITION 5.1. *If $\vec{\sigma}$ is ergodic then $\vec{\pi}$ is ergodic as well.*

The proposition follows easily from lemma 5.1 and theorem 5.2.

Definition 5.2. A sequence $(a_n)_{n=1}^\infty$ in a set A is of *bounded multiplicity* if there exists a K such that no $a \in A$ occurs more than K times in the sequence.

THEOREM 5.3. *The following conditions are equivalent:*

- (1) $\vec{\sigma}$ is strongly mixing.
- (2) $\vec{\pi}$ is strongly mixing.
- (3) The sequence $\gamma\vec{\sigma}$ is of bounded multiplicity for any non-trivial $\gamma \in \Gamma$.

The proof is omitted; the required ideas appear in the proof of:

THEOREM 5.4. *The following conditions are equivalent:*

- (1) $\vec{\sigma}$ is ergodic.
- (2) $\vec{\sigma}$ is weakly mixing.
- (3) $\vec{\pi}$ is weakly mixing.

Proof. If $\tilde{\sigma}$ is ergodic then employing theorem 5.2 we observe that any positive lower density subsequence is also ergodic. By proposition 5.1 the same applies to $\tilde{\pi}$. From this we conclude that (1) $\Rightarrow$ (2) $\Rightarrow$ (3).

It remains to show that (3) $\Rightarrow$ (1). Suppose $\tilde{\pi}$ is weakly mixing while $\tilde{\sigma}$ is not ergodic. Consequently, there exist $\delta > 0$, a non-trivial U and a sequence $(U_k)_{k=1}^\infty$ in $\mathcal{U}$, a sequence $(N_k)_{k=1}^\infty$ of positive integers and integers n_{ik} , $k \in \mathbb{N}$, $1 \leq i \leq s_k$, such that:

$$1 < N_{k+1}/N_k \xrightarrow[k \rightarrow \infty]{} \infty;$$

$$N_k < n_{1k} < n_{2k} < \cdots < n_{s_k k} \leq N_{k+1};$$

$$s_k \geq \delta N_{k+1};$$

$$U\sigma_{n_{ik}} \sim U_k, \quad i = 1, 2, \dots, s_k;$$

for all k .

Take some $\varepsilon > 0$. Given a matrix $A = (a_{ij})_{i,j=1}^d$ put

$$\|A\| = \max \{|a_{ij}| : 1 \leq i, j \leq d\}.$$

Write the group of $d \times d$ unitary matrices as a finite union $\bigcup_{i=1}^r W_i$ such that if $w_{i1}, w_{i2} \in W_i$ then $\|w_{i1}w_{i2}^{-1} - I\| < \varepsilon$. Set $\alpha_{ik} = \alpha_{n_{ik}}$. For some $1 \leq l \leq r$ we have infinitely many numbers k with at least s_k/r of the matrices $U(\alpha_{ik})$, $1 \leq i \leq s_k$, belonging to W_l . Consider the sequence of positive integers consisting of all those n_{ik} with $U(\alpha_{ik}) \in W_l$. Its upper density is at least δ/r . Denote by $\tilde{\alpha}'$, $\tilde{\sigma}'$, and $\tilde{\pi}'$ the subsequences of $\tilde{\alpha}$, $\tilde{\sigma}$ and $\tilde{\pi}$ corresponding to this sequence. Select some $\alpha_0 \in W_l$. Consider the sequence $\tilde{\pi}''$ defined by

$$\pi_n''(x) = \pi_n'(x)\alpha_0^{-1}, \quad x \in G, n = 1, 2, \dots;$$

$$t_k = \# \{j : 1 \leq j \leq s_k, U(\alpha_{jk}) \in W_l\}, \quad k = 1, 2, \dots;$$

$$\beta_n = \alpha_n' \alpha_0^{-1}, \quad n = 1, 2, \dots.$$

Fix momentarily a k with $t_k \geq s_k/r$ and put $T = \sum_{j=1}^k t_j$, $t = \sum_{j=1}^{k-1} t_j + 1$. Then

$$\begin{aligned} & \frac{1}{T^2} \sum_{m,n=1}^T \int_G \chi_U(\pi_m''(x)) \overline{\chi_U(\pi_n''(x))} d\mu(x) \\ & \geq \frac{1}{T^2} \sum_{m,n=t}^T \int_G \chi_U(\pi_m''(x)) \overline{\chi_U(\pi_n''(x))} d\mu(x) - \frac{1}{T^2} 2tTd^2 \\ & = \frac{1}{T^2} \sum_{m,n=t}^T \sum_{p,q,s,v=1}^d \int_G u_{pq}(\sigma_m'(x)) u_{qp}(\beta_m) \overline{u_{sv}(\sigma_n'(x)) u_{vs}(\beta_n)} d\mu(x) - 2td^2/T \\ & \geq \frac{1}{T^2} \sum_{m,n=t}^T \sum_{p,q,s,v=1}^d \int_G u_{pq}(\sigma_m'(x)) \delta_{qp} \overline{u_{sv}(\sigma_n'(x)) \delta_{vs}} d\mu(x) - \frac{1}{T^2} 3\varepsilon T^2 d^4 - 2td^2/T \\ & = \frac{1}{T^2} \sum_{m,n=t}^T \int_G \chi_U(\sigma_m'(x)) \overline{\chi_U(\sigma_n'(x))} d\mu(x) - 3\varepsilon d^4 - 2td^2/T \\ & \geq (T-t)^2/T^2 - 3\varepsilon d^4 - 2d^2(N_{k-1}+1)/(\delta N_k/r) \xrightarrow[k \rightarrow \infty]{} 1 - 3\varepsilon d^4 > 0, \end{aligned}$$

for sufficiently small $\varepsilon > 0$. Adjoining to $\tilde{\pi}''$ an arithmetic subsequence of $\tilde{\pi}$ of sufficiently small density, also translated by α_0^{-1} , we still have a non-ergodic sequence.

Thus $\tilde{\pi}$ contains a non-ergodic subsequence of positive lower density, which shows that $\tilde{\pi}$ is not weakly mixing. Thus (3) $\Rightarrow$ (1), and thereby the theorem is proved. $\square$

For the remainder of this section G is assumed to be abelian. Γ is the dual group of G . For a sequence $\tilde{\sigma}$ of epimorphisms the condition for ergodicity is that every non-zero character γ should have an orbit of reasonable multiplicity in Γ under the action of the dual sequence of monomorphisms of Γ . Our results yield the following theorem, due to Raikov [12]:

THEOREM. *Let $(n_k)_{k=1}^\infty$ be a sequence of distinct positive integers and $f \in L^2[0, 1)$. Then:*

$$\frac{1}{N} \sum_{k=1}^N f(\{n_k x\}) \xrightarrow[N \rightarrow \infty]{L^2} \int_0^1 f(t) dt.$$

More generally, we get

THEOREM 5.5. *Let G be a connected compact abelian group, $(n_k)_{k=1}^\infty$ a sequence of non-zero integers. The sequence $\tilde{\sigma}$ of epimorphisms of G given by $\sigma_k(x) = n_k x$ is:*

- (1) *ergodic – iff it is of reasonable multiplicity;*
- (2) *strongly mixing – iff it is of bounded multiplicity.*

Remark 5.2. Unlike the situation in the case of translations, where ergodicity can be deduced from known results concerning uniform distribution (see remark 5.1), this is impossible in general for epimorphisms. In fact, already for the circle group $\mathbb{T}$ the ergodicity condition, as given in the last theorem, is slightly weaker than the conditions known to ensure that a sequence $(n_k)_{k=1}^\infty$ has the property that $(n_k x)_{k=1}^\infty$ is uniformly distributed in $\mathbb{T}$ for almost every $x \in \mathbb{T}$ (see [10, § 1.4]). The gap between the two types of convergence (mean convergence for square-integrable functions and a.e. convergence for continuous functions) is even wider in the case of tori (see theorem 5.6 and [11, Th. 2]).

Now let G be a finite-dimensional connected compact abelian group. Such a group is characterized by its dual Γ being a torsion-free discrete abelian group of finite rank (see [6, Chs. 23–24]). We may accordingly assume that Γ is a subgroup of $\mathbb{Q}^r$ containing $\mathbb{Z}^r$ for some r , endowed with discrete topology. Thus G is a quotient group of $\hat{\mathbb{Q}}^r$.

Let σ be an endomorphism of G . Its dual $\hat{\sigma}$, being an endomorphism of Γ , is represented by an $r \times r$ rational matrix. Evidently, $\hat{\sigma}$ can be uniquely extended to $\mathbb{Q}^r$. By duality it follows that σ can be uniquely lifted to an endomorphism of $\hat{\mathbb{Q}}^r$, to be denoted by $\tilde{\sigma}$. Note that if σ is epic then $\tilde{\sigma}$ is epic as well. If σ and τ commute then $\hat{\sigma}$ and $\hat{\tau}$ commute, hence their extensions to $\mathbb{Q}^r$ commute, which implies that $\tilde{\sigma}$ and $\tilde{\tau}$ also commute.

LEMMA 5.3. *Let $\tilde{\sigma}$ be a sequence of epimorphisms of G and $\vec{\tilde{\sigma}}$ the sequence of their corresponding liftings to $\hat{\mathbb{Q}}^r$. Then $\tilde{\sigma}$ is ergodic iff $\vec{\tilde{\sigma}}$ is.*

The proof is straightforward.

Given a sequence $\tilde{\sigma}$ of epimorphisms of G , we put

$$\Delta_0(\tilde{\sigma}) = \{(m, n) \mid \det(\sigma_m - \sigma_n) = 0\}.$$

$\text{spec}(\sigma)$ is the finite sequence $\lambda_{1\sigma}, \lambda_{2\sigma}, \dots, \lambda_{r\sigma}$ of eigenvalues of an epimorphism σ (arbitrarily ordered). $\text{spec}(\vec{\sigma})$ is the sequence formed of all the finite sequences $\text{spec}(\sigma_n)$, i.e. $\text{spec}(\vec{\sigma}) = (\text{spec}(\sigma_1), \text{spec}(\sigma_2), \dots)$.

THEOREM 5.6. *Let $\vec{\sigma}$ be a sequence of commuting epimorphisms of G . Then:*

- (1) *The following conditions are equivalent:*
 - (a) $\vec{\sigma}$ is ergodic;
 - (b) $\Delta_0(\vec{\sigma})$ is a zero density subset of $\mathbb{N}^2$;
 - (c) $\text{spec}(\vec{\sigma})$ is of reasonable multiplicity.
- (2) *The following conditions are equivalent:*
 - (a) $\vec{\sigma}$ is strongly mixing;
 - (b) $\Delta_0(\vec{\sigma})$ is of bounded fibres;
 - (c) $\text{spec}(\vec{\sigma})$ is of bounded multiplicity.

Proof: In view of lemma 5.3 we may assume that $G = \hat{\mathbb{Q}}^r$. The implications (c) $\Rightarrow$ (b) $\Rightarrow$ (a) are easy for both parts, while (2a) $\Rightarrow$ (2c) follows from (1a) $\Rightarrow$ (1c) and theorem 3.1. It remains therefore to show that (1a) $\Rightarrow$ (1c).

Since $\vec{\sigma}$ is commutative there exists an $r \times r$ matrix τ over the field $\mathbb{A}$ of algebraic numbers such that for every σ in the sequence

$$\tau^{-1}\sigma\tau = \begin{pmatrix} (s_{1\sigma}) & & & 0 \\ & (s_{2\sigma}) & & \\ & & \ddots & \\ 0 & & & (s_{h\sigma}) \end{pmatrix}, \quad (s_{j\sigma}) = \begin{pmatrix} \lambda_{j\sigma} & & & * \\ & \lambda_{j\sigma} & & \\ & & \ddots & \\ 0 & & & \lambda_{j\sigma} \end{pmatrix}$$

(see, for example, [8, p. 134]). Suppose (1.c) is not fulfilled. Then there exists some $1 \leq j \leq h$ and $\varepsilon > 0$ such that for infinitely many positive integers N we have $\lambda_{j\sigma_n} = \lambda_N$ for at least εN numbers $1 \leq n \leq N$ for appropriately chosen $\lambda_N \in \mathbb{A}$. The triangulation of $\vec{\sigma}$ implies the existence of a row vector $v \in \mathbb{A}^r$ with $v\sigma_n = \lambda_{j\sigma_n}v$ for all n . Put $K = \mathbb{Q}(\{\lambda_{j\sigma_n} | n \in \mathbb{N}\})$. The extension $K/\mathbb{Q}$ is finite since $K \subseteq \mathbb{Q}(\{v_i | 1 \leq i \leq r\})$, where v_i , $1 \leq i \leq r$, are the components of v . We can select v so as to have $K = \mathbb{Q}(\{v_i | 1 \leq i \leq r\})$.

Let $v^{(1)} = v, v^{(2)}, \dots, v^{(t)}$ be all the conjugates of v over $\mathbb{Q}$. All these vectors clearly form common eigenvectors of $\vec{\sigma}$. Set $e = \sum_{i=1}^t v^{(i)}$. We have $e \in \mathbb{Q}^r$, and for infinitely many numbers N the equation $e\sigma_m = e\sigma_n$ has at least $\varepsilon^2 N^2$ solutions (m, n) with $1 \leq m, n \leq N$. To show that $\vec{\sigma}$ is non-ergodic it suffices therefore to prove that $e \neq 0$. Pick $\lambda \in K$ with $K = \mathbb{Q}(\lambda)$. λ is a polynomial in some eigenvalues $\lambda_{j\sigma_n}$. It follows that $v\sigma = \lambda v$, where σ is the corresponding polynomial in the matrices σ_n . The vectors $v^{(1)}, \dots, v^{(t)}$ are eigenvectors of σ , and the corresponding eigenvalues are all the conjugates of λ . Thus $v^{(1)}, \dots, v^{(t)}$ are linearly independent, whence $e \neq 0$. This completes the proof. $\square$

Example 5.1. The commutativity assumption in the last theorem is essential. In fact, consider the sequence of epimorphisms of $\mathbb{T}^2$ given by

$$\sigma_n = \begin{pmatrix} 1+n & n^2 \\ -1 & 1-n \end{pmatrix}, \quad n \in \mathbb{N},$$

It is easy to verify that $\gamma\sigma_m = \gamma\sigma_n$ for $0 \neq \gamma \in \mathbb{Z}^2$ implies $m = n$, so that $\bar{\sigma}$ is strongly mixing. Yet the only eigenvalue of any σ_n is 1. Note also that the sequence is strongly mixing even though no σ_n is even ergodic.

REFERENCES

- [1] R. L. Adler & T. J. Rivlin. Ergodic and mixing properties of Chebyshev polynomials. *Proc. Amer. Math. Soc.* **15** (1964), 794–796.
- [2] J. R. Blum & D. L. Hanson. On the mean ergodic theorem for subsequences. *Bull. Amer. Math. Soc.* **66** (1960), 308–311.
- [3] N. A. Friedman. *Introduction to Ergodic Theory*. Van Nostrand: Princeton, New Jersey, 1970.
- [4] N. A. Friedman. Mixing on sequences. *Canad. J. Math.* **35** (1983), 339–352.
- [5] H. Furstenberg. *Recurrence in Ergodic Theory and Combinatorial Number Theory*. Princeton University Press: Princeton, New Jersey, 1981.
- [6] E. Hewitt & K. A. Ross. *Abstract Harmonic Analysis*, Vol. I. Springer-Verlag: Berlin, 1963.
- [7] E. Hewitt & K. A. Ross. *Abstract Harmonic Analysis*, Vol. II. Springer-Verlag: Berlin, 1970.
- [8] N. Jacobson. *Lectures in Abstract Algebra*, vol. II. Van Nostrand: Princeton, New Jersey, 1953.
- [9] L. K. Jones. A mean ergodic theorem for weakly mixing operators. *Advances in Math.* **7** (1971), 211–216.
- [10] L. Kuipers & H. Niederreiter. *Uniform Distribution of Sequences*. Wiley: New York, 1974.
- [11] W. Phillip. An n -dimensional analogue of a theorem of H. Weyl. *Compositio Math.* **16** (1964), 161–163.
- [12] D. A. Raikov. On some arithmetical properties of summable functions (Russian). *Mat. Sb.* **1** (43), (1936) 377–384.
- [13] A. Rényi. On mixing sequences of sets, *Acta. Math. Acad. Sci. Hungar.* **9** (1958), 215–228.