A GEOMETRIC CONSTRUCTION OF MEASURE PRESERVING TRANSFORMATIONS

R. V. CHACON
UNIVERSITY OF MINNESOTA

1. Introduction

Our main purpose is to give an extension of a construction which we have obtained in [3]. A consequence of the discrete spectrum theorem is that transformations having discrete spectrum have a square root if and only if -1 is absent from the spectrum. One of the important problems in ergodic theory is to investigate to what extent the implications of the discrete spectrum theorem remain in the general case. Along these lines, Halmos [10] has asked whether every transformation with continuous spectrum has a square root. We remark that the question is raised (see [8]) for one-to-one transformations since otherwise the discrete spectrum theorem as well as the consequence it mentioned fail to hold. Note furthermore that the spectral theorem holds only for one-to-one transformations, and therefore we can expect the spectrum to determine the properties of the transformation in this case only.

In [3] an example is obtained which answers the question of Halmos to which we have referred. Here we extend the method to give an example of a one-to-one ergodic and measure preserving transformation which has continuous spectrum, is not strongly mixing, and which has no roots of any order. The fact that the transformation is not strongly mixing is of some interest, as in view of the fact that the shift transformations are strongly mixing.

The paper is divided into three sections. In the first we give a simple example due to Kakutani and von Neumann. In the second section we formulate and prove a general decomposition theorem for one-to-one measure preserving transformations. The sufficient part of this theorem is a summary and codification of ideas common to constructions given by several authors ([6], [9], [12], and [13]). The example given in the first section is the simplest as well as the earliest which exhibits the main features of the theorem. Our main intention in giving the decomposition theorem is to fix ideas for the construction of the counterexample which is the principal purpose of the paper. This construction is given in the third section and is independent of the first two.

This research was supported by a National Science Foundation Grant No. GP-3752. The main results were announced on March 15, 1965 in the *Notices Amer. Math. Soc.*, 12-4 (1965.)

In [4] we applied a small extension of the ideas suggested by the decomposition theorem to obtain what seems to us to be a substantial simplification of the example [13] of a transformation having no σ-finite invariant measure. In his thesis [5] N. Friedman has shown that these transformations are dense in the strong topology in the anti-periodic transformations, strengthening a result given in [14]. The method is sufficiently general so that one may show that this property is common to a large class of transformations, and indeed the interesting result was obtained in [1] that the transformations essentially given by P. R. Halmos in [8] having discrete spectrum and no roots are dense in the strong topology in the anti-periodic transformations. The class of transformations we construct in this paper also has the property that it is dense in the strong topology in the anti-periodic transformations, but since the proof is again like that in [5], we refrain from giving details.

2. An example having discrete spectrum

The example of Kakutani-von Neumann is of a measure preserving transformation having discrete spectrum with eigenvalues of the form $\exp(2\pi im/2^k)$, m, k integers. We give the example as follows. A *simple tower* T is a finite ordered partition $\{I_j, j = 1, \cdots, n\}$ of the unit interval which is composed of subintervals of equal length. The transformation τ_T induced by the simple tower T is defined by mapping I_k linearly onto I_{k+1}, $k = 1, \cdots, n-1$, so that the domain of τ_T is $\bigcup_{j=1}^{n-1} I_j$ and so that its range is $\bigcup_{j=2}^{n} I_j$. It is helpful to think of the tower T as an ordered stacking of the intervals I_j with I_1 on the bottom, I_n on the top, and with the other subintervals between them in order. We may then think of τ_T as mapping each point of the unit interval in the tower T to that point directly above, if any. Each simple tower T gives rise to a simple tower $S(T)$ as follows: write I_j as the sum of the two disjoint and consecutive intervals of equal length I_j^1 and I_j^2, and set

$$S(T) = \{I_j^1, j = 1, \cdots, n, I_j^2, j = 1, \cdots, n\}. \tag{2.1}$$

In terms of the description of T as a stack, $S(T)$ is the simple tower obtained by splitting T down the middle and putting the right-hand substack above the left-hand one. Note that $\tau_{S(T)}$ is an extension of τ_T.

Next we define a sequence of towers inductively by setting

$$T_2 = \{[0, \tfrac{1}{2}], [\tfrac{1}{2}, 1]\}, \qquad T_k = S(T_{k-1}), \qquad k \geq 3, \tag{2.2}$$

and set $\tau = \lim_{k\to\infty} \tau_{T_k}$. That τ is defined follows from the fact remarked above that $\tau_{S(T)}$ is an extension of τ_T, and that the measure of the top layer of the stack T_k tends to zero as k tends to infinity. To see that $\exp(2\pi im/2^k)$ is an eigenvalue, note that the top layer of T_k is mapped by τ onto the bottom layer. That τ has discrete spectrum follows from the fact that the subintervals of the stacks T_k generate the Borel field.

3. The decomposition theorem

Let (X, F, μ) be an interval, the Lebesgue sets and Lebesgue measure, respectively. Since we are interested in equivalence classes, modulo sets of measure zero, of transformations of X onto X, our equations are understood to hold almost everywhere. In particular, transformations need be defined only almost everywhere. The (invertible) transformation τ of X onto X is measurable if $A \in F$ implies $\tau(A) \in F$ and $\tau^{-1}(A) \in F$. A transformation τ of X onto X is non-singular, if it is measurable, and if $A \in F$ and $\mu(A) = 0$ implies that $\mu(\tau(A)) = \mu(\tau^{-1}(A)) = 0$. A transformation τ is measure preserving, if it is measurable, and if $A \in F$ implies $\mu(A) = \mu(\tau(A)) = \mu(\tau^{-1}(A))$.

We first generalize the definition of a simple tower: a *tower* T is an ordered class of subintervals $\{I_{j,k}, j = 1, \cdots, n; k = 1, \cdots, m(j)\}$ which forms a partition $P(T)$ of the unit interval and which have the property that for each j, $\mu(I_{j,k}) = \mu(I_{j,1})$, $k = 1, \cdots, m(j)$. The transformation τ_T induced by T is defined by mapping $I_{j,k}$ linearly onto

$$(3.1) \qquad I_{j,k+1}, \qquad 1 \leq k \leq m(j) - 1; \quad j = 1, \cdots, n.$$

The *base* of the tower is defined as the set $B(1, T) = \bigcup_{j=1}^{n} I_{j,1}$ and the *top* of the tower is defined as the set $B(2, T) = \bigcup_{j=1}^{n} I_{j,m(j)}$. Note that the domain of the transformation τ_T is $[0, 1] - B(2, T)$ and that its range is $[0, 1] - B(1, T)$. We say that a tower T_2 is *finer* than a tower T_1 if τ_{T_2} is an extension of τ_{T_1}, and if $P(T_2)$ is finer than $P(T_1)$.

It is helpful to think of the tower

$$(3.2) \qquad T = \{I_{j,k}, j = 1, \cdots, n; \quad k = 1, \cdots, m(j)\}$$

as composed of n ordered stacks of varying height $m(j)$. We may then regard τ_T as mapping each point of the unit interval in the tower T to that point directly above, if any. We note that if T_2 is finer than T_1, then T_2 can be obtained from T_1 by an application of the following operation. Cut T_1 by a finite number of vertical cuts, and then stack in groups those resulting substacks of the same width.

We define a *nested null sequence of towers* $\{T_n\}$ to be a sequence of towers such that T_k is finer than T_{k-1}, $k \geq 2$, and such that the measure of the sets $B(2, T_k)$ tends to zero, that is, the measure of the tops of the towers tends to zero. If the unit interval admits of a partition into two intervals $I(1)$ and $I(2)$, and if there exists an integer N such that $\{T_{n+N} \cap I(1)\}$ and $\{T_{n+N} \cap I(2)\}$ are nested null sequences of towers of $I(1)$ and of $I(2)$, respectively, then we say that the sequence $\{T_n\}$ is *divisible.*

If $\{T_n\}$ is a nested null sequence of towers, define $\tau_{\{T_n\}}$ as the limit, which clearly exists almost everywhere, given by $\tau_{\{T_n\}} = \lim_{n\to\infty} \tau_{T_n}$. We may now state the decomposition theorem as follows.

THEOREM. *If $\{T_n\}$ is a nested null sequence of towers, then $\tau_{\{T_n\}}$ is a one-to-one, anti-periodic, and measure preserving transformation of the unit interval. Con-*

versely, if τ is a one-to-one, anti-periodic, and measure preserving transformation of the unit interval, then there exists a nested null sequence of towers $\{T_n\}$ such that $\tau_{\{T_n\}}$ is isomorphic to τ. Furthermore, $\{T_n\}$ is in each case not divisible if and only if τ is ergodic.

The proof of the first part of the theorem is immediate. To prove the second part we need the following lemma.

LEMMA. *If τ is a one-to-one, anti-periodic and measure preserving transformation of the unit interval, then there exists a partition of the unit interval into three measurable sets A_1, A_2, and A_3 such that*

(i) $\tau(A_1) = A_2$,

(ii) $\tau([A_2 - \tau^{-1}(A_3)] + A_3) = A_1$.

PROOF OF THE LEMMA. Let $\mathfrak{A}$ be the class of measurable sets A such that $m(A \cap \tau(A)) = 0$. Order the class $\mathfrak{A}$ by set inclusion and apply Zorn's lemma to obtain a maximal set A_1 in the class. (We have made use of the fact that L_1 is a complete lattice.) Next, define $A_2 = \tau(A_1)$ and $A_3 = \tau^{-1}(A_1) - A_2$. That (i) and (ii) are satisfied follows from the maximality of A_1.

PROOF OF THE THEOREM. Apply the lemma to τ to obtain a tower, and then apply the lemma to the transformation induced on the base, and so on by induction. We may assume without loss of generality that the partitions generate the Borel field since we may always make them finer if necessary. We next apply the von Neumann isomorphism theorem to see that the partition elements may be supposed to be intervals, and the proof of the theorem is complete.

4. The main example

One-to-one and measure preserving transformations τ induce unitary operators T_τ by setting $T_\tau f(x) = f(\tau^{-1}(x))$ for $f \in L_2$. The transformation τ is said to have discrete spectrum if it is measure preserving and if its induced unitary operator has a complete orthonormal set of eigenfunctions. The transformation τ is said to have continuous spectrum if it is measure preserving and if its induced unitary operator has for its sole eigenfunction the function identically equal to one almost everywhere.

Our result may be stated as follows.

THEOREM. *There exists an explicit construction of an ergodic, invertible, and measure preserving transformation τ of the unit interval such that* (i) *τ has continuous spectrum,* (ii) *τ is not strongly mixing, and* (iii) *τ has no nonsingular roots of any order.*

It is not hard to see that a nonsingular p-th root of an invertible, ergodic, and measure preserving transformation is also measure preserving; we do not use this fact in the sequel.

In order to make the proof as transparent as possible we first prove the following theorem given in [3].

THEOREM. *There exists an explicit construction of an ergodic, invertible, and measure preserving transformation τ of the unit interval such that* (i) *τ has con-*

tinuous spectrum, (ii) τ *is not strongly mixing, and* (iii) τ *has no nonsingular square roots.*

The proof we give here is an adaptation of the proof given in [3]. It is changed in such a way that it can be readily extended to give the general result. We indicate the extensions needed in the general case after the square root case is carried out in detail.

Definition of the transformation in the square root case. Given a sequence $\{j(k), k = 1, 2, \cdots\}$ of integers with $j(k) \geq 1, k \geq 1$, we shall define an invertible measure preserving transformation $\tau = \tau(\{j(k)\})$ having the desired properties. For convenience the transformation τ is defined on the union of the unit interval with another interval N disjoint from it; the length of N depends on the sequence $\{j(k)\}$.

The transformation is defined inductively. For $r = k$, we suppose that the transformation is defined on part of the space, and we extend the domain of definition for $r = k + 1$.

For $r = k$, we suppose the transformation given as follows: let

$$m(k) = 3m(k-1)j(k) + 6, \qquad m(0) = 2, \tag{4.1}$$

and suppose that there exist $m(k)$ intervals of equal length $A_1(k), \cdots, A_{m(k)}(k)$, each a subinterval of the unit interval or of N, and that τ maps $A_{i-1}(k)$ linearly onto $A_i(k), i = 2, \cdots, m(k)$.

For $r = k + 1$, we need to extend the definition of τ in such a way that it admits a representation as in the preceding paragraph, with k replaced by $k + 1$. This is accomplished in two steps.

The first step is the following. Divide $A_i(k)$, $(1 \leq i \leq m(k))$, into $j(k+1)$ consecutive subintervals of equal length, denoting them

$$B_{i,1}(k+1), \cdots, B_{i,j(k+1)}(k+1), \qquad (1 \leq i \leq m(k)). \tag{4.2}$$

Since $A_{i-1}(k)$ is mapped linearly onto $A_i(k)$, $(2 \leq i \leq m(k))$, τ also maps $B_{i-1,j}(k+1)$ linearly onto

$$B_{i,j}(k+1), \qquad (2 \leq i \leq m(k),\ 1 \leq j \leq j(k+1)). \tag{4.3}$$

In this first step we extend τ by mapping $B_{m(k),j}(k+1)$ linearly onto

$$B_{1,j+1}(k+1), \qquad (1 \leq j \leq j(k+1) - 1). \tag{4.4}$$

There are $m(k)$ $A_i(k)$'s and each $A_i(k)$ is divided into $j(k+1)$ $B_{i,j}(k+1)$'s so that there are $m(k)j(k+1)$ $B_{i,j}(k+1)$'s altogether. We write these sets with a single subscript $B_i(k+1)$, $(i = 1, \cdots, m(k)j(k+1))$, in such a way that τ as thus far extended maps $B_{i-1}(k+1)$ linearly onto

$$B_i(k+1), \qquad (2 \leq i \leq m(k)j(k+1)). \tag{4.5}$$

The second step in the extension of τ is obtained by dividing $B_i(k+1)$ into three consecutive subintervals of equal length $B_i^1(k+1)$, $B_i^2(k+1)$, $B_i^3(k+1)$. The transformation τ, as already defined, maps $B_{i-1}^u(k+1)$ linearly onto

$$B_i^u(k+1), \qquad (2 \leq i \leq m(k)j(k+1),\ 1 \leq u \leq 3). \tag{4.6}$$

We take from the subinterval of N, where τ is as yet undefined, six consecutive subintervals of equal length, equal to the length of the $B_i^u(k+1)$, and denote these six subintervals of N by $E_j(k+1)$, $(1 \leq j \leq 6)$. (We suppose that N is just big enough so that at each stage there is an interval left over.) Then extend τ so that the sequence of $3m(k)j(k+1)+6$ sets (with $(k+1)$ deleted for simplicity), with $s = m(k)j(k+1)$,

$$E_1, B_1^1, \cdots, B_s^1, E_2, E_3, B_1^2, \cdots, B_s^2, E_4, E_5, E_6, B_1^3, \cdots, B_s^3 \tag{4.7}$$

has the property that τ maps each set, except for the last, linearly onto the next set. We denote this sequence by $A_i(k+1)$, $(1 \leq i \leq m(k+1))$, where $m(k+1) = 3m(k)j(k+1)+6$.

To complete the induction it remains to define τ for $r = 0$. We set $A_1(0) = (0, \frac{1}{2})$ and $A_2(0) = (\frac{1}{2}, 1)$ and define τ initially as the linear map of $A_1(0)$ onto $A_2(0)$. An intuitive description of the transformation is the following.

At the k-th stage there is a stack of $m(k)$ intervals of equal length, and τ maps each point to the one directly above, so that the points of the top interval are not yet mapped anywhere.

The first step of the extension of τ is obtained as follows. The original stack is split into $j(k+1)$ equal and consecutive stacks, and these are stacked in order. This has the effect of mapping a $j(k+1) - 1/j(k+1)$ part of the top interval of the original stack into the bottom interval of the original stack. (The symmetric difference of the image under τ of the top interval and the bottom interval has measure which tends to zero as $j(k+1)$ tends to infinity, therefore.)

In the second step of the extension, we take the resulting stack, which is clearly composed of $m(k)j(k+1)$ intervals, and divide it into three consecutive and equal substacks. We then take six consecutive and equal subintervals of N where τ is as yet undefined, of length equal to the length of each of the intervals of the three stacks, and put one under the first stack, two under the second stack, and three under the third stack. The thus modified three stacks are then stacked in order.

Note that if we regard the union of the first interval of each of the modified stacks as a set, then this set is mapped upward until it gets to the top, which is one-third covered by a further subinterval. At the next step, one-third returns to the bottom, and the part covered by the further subinterval returns to the bottom in two steps.

Properties of the transformation. The transformation is defined inductively as we have seen. In what follows we do not suppose that the measure space is normalized. We deal with statements depending on $\{j(k)\}$ and on r, and usually concerning the transformation $\tau = \tau(\{j(k)\})$ of the following sort.

The statement will be shown to be true for $r = 0$ and the truth of the statement for $r = 0, \cdots, k-1$, with a certain choice of $j(1), \cdots, j(k-1)$, will imply that there exists a $K = K(j(1), \cdots, j(k-1))$ such that if $j(k) \geq K$, then the statement is true for $r = k$ with that choice of $j(1), \cdots, j(k)$. This clearly means that there exists a sequence $\{j(k)\}$ such that the statement is true for all r. It is furthermore clear that if there is a finite number of such statements, then there is a single sequence $\{j(k)\}$ such that all the statements will be true for all r. We shall refer to this state of affairs by saying that a certain statement holds for all r if $\{j(k)\}$ tends to infinity sufficiently rapidly. It is clear from the discussion that if a finite number of statements hold for separate $\{j(k)\}$ tending to infinity sufficiently rapidly, that then there exists a single sequence $\{j(k)\}$ tending to infinity sufficiently rapidly so that all the statements are true for all r.

For the sake of simplicity, we say that the set A equals the set B with an error of δ, if the measure of the symmetric difference of A and B is no more than δ, and we write

$$A = B + E(\delta). \tag{4.8}$$

LEMMA 1. *Let $\{E_{j,k}, 1 \leq j \leq r(k), 1 \leq k < +\infty\}$ generate the Borel field and suppose that for each k, $1 \leq k < +\infty$, the sets $\{E_{j,k}, 1 \leq j \leq r(k)\}$ are pairwise disjoint. Suppose also that each $E_{j,k}$ is the union of a finite number of sets from the class $\{E_{j,k+1}, 1 \leq j \leq r(k+1)\}$. Then for each $\epsilon > 0$ and A in the Borel field, there exists a set of indices $G(\epsilon, A)$ and an integer $k(\epsilon, A)$ such that A equals $\bigcup_{j \in G(\epsilon,A)} E_{j,k(\epsilon,A)}$ with an error of ϵ, and*

$$\mu(A \cap E_{j,k(\epsilon,A)}) \geq (1 - \epsilon)\mu(E_{j,k(\epsilon,A)}) \tag{4.9}$$

for $j \in G(\epsilon, A)$.

PROOF. This is a straightforward result in measure theory.

LEMMA 2. *The sets $\{A_j(k), 1 \leq j \leq m(k), 1 \leq k < +\infty\}$ generate the Borel field. Furthermore, each $A_j(k)$ is the union of $3j(k+1)$ sets of the class $\{A_j(k+1), 1 \leq j \leq m(k+1)\}$.*

PROOF. The lemma follows at once from the construction.

LEMMA 3. *The sets $\{B_j(k), 1 \leq j \leq m(k-1)j(k), 1 \leq k < +\infty\}$ generate the Borel field. Furthermore, each $B_j(k)$ is the union of $3j(k+1)$ sets of the class $\{B_j(k+1), 1 \leq j \leq m(k)j(k+1)\}$.*

PROOF. Lemma 3 also follows at once from the construction.

LEMMA 4. *Given $\{\delta(k)\} \downarrow 0$ and if $\{j(k)\}$ tends to infinity sufficiently rapidly, then for all k,*

(i) $\mu(\bigcup_{j=1}^{m(k)} A_j(k)) \geq (1 - \delta(k))\mu(X + N)$,

(ii) $\mu(\tau A_{m(k)}(k) \Delta A_1(k)) \leq (1 - \delta(k))\mu(A_1(k))$,

(iii) *$m(k)$ depends only on $\delta(0), \cdots, \delta(k-1)$.*

PROOF. By $A \Delta B$ we mean the symmetric difference of the sets A and B. The lemma follows at once from the construction, since $m(k) = 3m(k-1)j(k) + 6$.

LEMMA 5. *Let $\{j(k)\}$ be given. Then $\tau(\{j(k)\})$ has continuous spectrum.*

PROOF. Suppose to the contrary. Then there exists a constant λ of absolute value one and a function f of absolute value one, almost everywhere such that (τ is clearly ergodic),

$$T_\tau f = \lambda f, \qquad \lambda \neq 1. \tag{4.10}$$

It follows from lemmas 1 and 3 that, given $\epsilon_1, \epsilon_2 > 0$, there exists a set $B = B^{\cdot}_{(\epsilon_1,\epsilon_2)}(k(\epsilon_1, \epsilon_2))$ such that it has a subset A having the property

$$\mu(A) \geq (1 - \epsilon_1)\mu(B), \tag{4.11}$$

and, for $x \in A$,

$$f(x) = e^{i(\theta + \delta(x))}, \qquad |\delta(x)| \leq \epsilon_2. \tag{4.12}$$

In extending τ from $k = k(\epsilon_1, \epsilon_2)$ to $k + 1$, we split, according to the construction, the stack composed of $B_j(k)$, $j = 1, \cdots, m(k)j(k + 1)$, into three substacks and put one extra interval at the beginning of the first substack, two extra intervals at the beginning of the second substack, three extra intervals at the beginning of the third substack, and then stack these modified substacks in order. This means that the interval $B_j(k)$ is split into three intervals of equal length, $B_j^1(k)$, $B_j^2(k)$, and $B_j^3(k)$ such that

$$\begin{aligned} \tau^{m(k)+2}B_j^1(k) &= B_j^2(k) \subset B_j(k), \\ \tau^{m(k)+1}B_j^2(k) &= B_j^3(k) \subset B_j(k). \end{aligned} \tag{4.13}$$

It follows from (4.10) that

$$\begin{aligned} T_\tau^{-(m(k)+2)}f(x) &= \lambda^{-(m(k)+2)}f(x), \\ T_\tau^{-(m(k)+1)}f(x) &= \lambda^{-(m(k)+1)}f(x). \end{aligned} \tag{4.14}$$

Putting (4.11), (4.12), (4.13), and (4.14) together we see that if ϵ_1 is small enough, then there exist two sets of positive measure $B^1 = B_j^1(k) \cap A$, $B^2 = B_j^2(k) \cap A$ such that, for $x \in B^1$,

$$e^{i(\theta+\delta(x))} = \lambda^{m(k)+2}e^{i(\theta+\delta(\tau^{m(k)+2}x))}, \tag{4.15}$$

and for $x \in B^2$,

$$e^{i(\theta+\delta(x))} = \lambda^{m(k)+1}e^{i(\theta+\delta(\tau^{m(k)+1}x))}. \tag{4.16}$$

Equations (4.15), (4.16), and (4.12), writing $\lambda = e^{i\psi}$, imply, modulo 2π,

$$|\psi| \leq 4\epsilon_2. \tag{4.17}$$

Since ϵ_2 can be chosen arbitrarily small, this implies that $\lambda = 1$.

LEMMA 6. *The transformation $\tau(\{j(k)\})$ is not strongly mixing.*

PROOF. Let $B = B_1(k + 1)$. From the construction we have

$$\mu(B \cap \tau^{m(r)j(r)+2}B) \geq \tfrac{1}{3}\mu(B), \tag{4.18}$$

for $r \geq k$. Since $\mu(B_1(k + 1))$ tends to zero as k tends to infinity, the normalized measure of $B_1(k + 1)$ also has this property. The proof is complete if we take k sufficiently large, since strong mixing implies that

$$\lim_{j\to\infty} \frac{\mu(B \cap \tau^j B)}{\mu(X + N)} = \left[\frac{\mu(B)}{\mu(X + N)}\right]^2. \tag{4.19}$$

In the sequel it will be convenient to define $1A = A$, $0A = \varnothing$, where A is a set, and $\varnothing$ is the empty set.

LEMMA 7. *Let $A = (\frac{1}{2}, 1)$, and let $\{j(k)\}$ be given. Then for each k there exists a sequence $\{x_i(k), 1 \leq i \leq m(k)\}$ of zeros and ones such that*

$$A = \sum_{i=1}^{m(k)} x_i(k) A_i(k). \tag{4.20}$$

Furthermore, we have

$$\tau^j A = \sum_{i=1}^{m(k)} x_{i-j} A_i(k) + E(\eta(j(k+1))), \tag{4.21}$$

for $j = 1, \cdots, m(k)$, *where* $\eta(j(k+1))$ *tends to zero as* $j(k+1)$ *tends to infinity, where the subscripts are taken modulo* $m(k)$.

PROOF. The first part of the lemma follows at once from the construction. The second part follows since τ maps $A_{m(k)}(k)$ onto $A_1(k)$, with error tending to zero as $j(k+1)$ tends to infinity.

It is convenient in what follows to define (a, β), where $\beta = (a_1, \cdots, a_n)$, to be $(a, a_1, \cdots, a_n)$, with analogous meaning for similar expressions.

LEMMA 8. *Let* $\alpha(k) = (x_1(k), \cdots, x_{m(k)}(k))$ *where the sequences* $(x_1(k), \cdots, x_{m(k)}(k))$ *are those of lemma 7. Then the* $\alpha(k)$ *are given inductively by*

(4.22)

$$\alpha(0) = (0, 1)$$

$$\alpha(1) = \Big(0, \overbrace{\alpha(0), \cdots, \alpha(0)}^{j(1)}, 0, 0, \overbrace{\alpha(0), \cdots, \alpha(0)}^{j(1)}, 0, 0, 0, \overbrace{\alpha(0), \cdots, \alpha(0)}^{j(1)}\Big)$$

$\cdot$
$\cdot$
$\cdot$

$$\alpha(k) = \Big(0, \overbrace{\alpha(k-1), \cdots, \alpha(k-1)}^{j(k)}, 0, 0, \overbrace{\alpha(k-1), \cdots, \alpha(k-1)}^{j(k)}, 0, 0, 0, \overbrace{\alpha(k-1), \cdots, \alpha(k-1)}^{j(k)}\Big).$$

PROOF. This again follows from the construction of the transformation.

We note that subscripts on the elements of $\alpha(k)$ are understood to be modulo $m(k)$, so that letting $i - j = r$ modulo $m(k)$, then $x_{i-j}(k) = x_r(k)$. Furthermore, note that $m(k)$, $k = 0, 1, 2, \cdots$, are even numbers, so that with the convention we use, $x_N(k)$ has an even subscript if and only if N is even, for each k.

Recall the definition of $\alpha(k)$ in terms of $\alpha(k-1)$:

$$\alpha(k) = \Big(\underbrace{0, \overbrace{\alpha(k-1), \cdots, \alpha(k-1)}^{j(k)}, 0, 0, \overbrace{\alpha(k-1), \cdots, \alpha(k-1)}^{j(k)}, 0, 0, 0, \overbrace{\alpha(k-1), \cdots, \alpha(k-1)}^{j(k)}}_{3j(k)m(k-1)+6 \text{ elements}}\Big), \tag{4.23}$$

and introduce the vector $\beta(k)$:

$$\beta(k) = \Big(\underbrace{\overbrace{\alpha(k-1), \cdots, \alpha(k-1)}^{j(k)}, \overbrace{0, \cdots\cdots, 0}^{2j(k)m(k-1)+6}}_{3j(k)m(k-1)+6 \text{ elements}}\Big), \tag{4.24}$$

so that we may write

$$\alpha = S\beta + S^{j(k)m(k-1)+3}\beta + S^{2j(k)m(k-1)+6}\beta, \tag{4.25}$$

where S is the one-step shift operator in the space of sequences of length $3j(k)\, m(k-1) + 6$ (deleting (k) following α and β in (4.25)).

LEMMA 9. *Let M be an integer, and let $N = \langle M\rangle$, $r(1) = \langle N + m(k-1)j(k) + 1\rangle$, and $r(2) = \langle N + 2m(k-1)j(k) + 3\rangle$, where $\langle a\rangle$ is that integer equal to a modulo $m(k)$, which satisfies $-m(k)/2 \le \langle a\rangle < m(k)/2$. Then there exist three positive numbers K_1, K_2, K_3 (depending on $j(k)$ such that*

$$\lim_{j(k)\to\infty} \frac{3(K_1 + K_2 + K_3)m(k-1)}{m(k)} = 1, \tag{4.26}$$

and for $t = 0, 1$,

$$\begin{aligned}\sum_{i=1}^{m(k)} x_i(k)x_{i-M-t}(k) \le{} & 3K_1 \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-N-t}(k-1) \\ & + K_2\Bigg[\sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(1)-t}(k-1) + \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(1)-1-t}(k-1) \\ & \qquad + \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(1)-2-t}(k-1)\Bigg] \\ & + K_3\Bigg[\sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(2)-t}(k-1) + \sum_{i=1}^{m(k-1)} x_ik-1x_{i-r(2)-1-t}(k-1) \\ & \qquad + \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(2)-2-t}(k-1)\Bigg].\end{aligned} \tag{4.27}$$

PROOF. Fix the value of k, and using the notation of (4.24) and (4.25), we see that

$$\sum_{i=1}^{m(k)} x_i(k)x_{i-M}(k) = \alpha S^N\alpha, \tag{4.28}$$

and

$$\begin{aligned}\alpha S^N\alpha = {} & (S\beta + S^{j(k)m(k-1)+3}\beta + S^{2j(k)m(k-1)+6}\beta) \\ & \cdot (S^{N+1}\beta + S^{N+j(k)m(k-1)+3}\beta + S^{N+2j(k)m(k-1)+6}\beta).\end{aligned} \tag{4.29}$$

Using $\gamma S\delta = S^{-1}\gamma\delta$ and $S^r = S^{3j(k)m(k-1)+6+r}$, it follows from (4.29) that

$$\begin{aligned}\alpha S^N\alpha &= \beta S^N\beta + \beta S^{N+2m(k-1)j(k)+4}\beta + \beta S^{N+m(k-1)j(k)+1}\beta \\ &\quad + \beta S^{N+m(k-1)j(k)+2}\beta + \beta S^N\beta + \beta S^{N+2m(k-1)j(k)+3}\beta \\ &\quad + \beta S^{N+2m(k-1)j(k)+5}\beta + \beta S^{N+m(k-1)j(k)+3}\beta + \beta S^N\beta \\ &= \beta S^N\beta + \beta S^{r(2)+1}\beta + \beta S^{r(1)}\beta \\ &\quad + \beta S^{r(1)+1}\beta + \beta S^N\beta + \beta S^{r(2)}\beta \\ &\quad + \beta S^{r(2)+2}\beta + \beta S^{r(1)+2}\beta + \beta S^N\beta.\end{aligned} \tag{4.30}$$

We now denote the elements of β as follows:

$$\beta(k) = (y_1(k), \cdots, y_{m(k)}(k)). \tag{4.31}$$

It can then be obtained easily from (4.30) and (4.31) that

$$\alpha S^N \alpha = 3 \sum_{i=1}^{m(k)} y_i y_{i-N} + \sum_{i=1}^{m(k)} y_i y_{i-r(1)} + \sum_{i=1}^{m(k)} y_i y_{i-(r(1)+1)} + \sum_{i=1}^{m(k)} y_i y_{i-(r(1)+2)} + \sum_{i=1}^{m(k)} y_i y_{i-r(2)} + \sum_{i=1}^{m(k)} y_i y_{i-(r(2)+1)} + \sum_{i=1}^{m(k)} y_i y_{i-(r(2)+2)}. \tag{4.32}$$

For the fixed value of k we are working with, we define

$$\bar{\alpha}(k) = \Big(0, \overbrace{\bar{\alpha}(k-1), \cdots, \bar{\alpha}(k-1)}^{j(k)}, 0, 0, \overbrace{\bar{\alpha}(k-1), \cdots, \bar{\alpha}(k-1)}^{j(k)}, 0, 0, 0, \overbrace{\bar{\alpha}(k-1), \cdots, \bar{\alpha}(k-1)}^{j(k)}\Big), \tag{4.33}$$

where, however,

$$\bar{\alpha}(k-1) = \Big(\overbrace{1, 1, \cdots, 1}^{m(k-1)}\Big). \tag{4.34}$$

All the previous formulas hold with bars on the various terms if the bar vectors are defined in the obvious way. Our purpose in introducing them is to make certain estimates, as follows. Let G_1 be the set of indices where both $\bar{y}_i$ and $\bar{y}_{i-N}$ are equal to one. Let $G_2, G_3, \cdots, G_7$ be similarly defined using the pairs $\bar{y}_i$ and $\bar{y}_{i-r(1)}$, $\bar{y}_i$ and $\bar{y}_{i-(r(1)+1)}$, $\bar{y}_i$ and $\bar{y}_{i-(r(2)+2)}$, $\bar{y}_i$ and $\bar{y}_{i-r(2)}$, $\bar{y}_i$ and $\bar{y}_{i(r(2)+1)}$, and $\bar{y}_i$ and $\bar{y}_{i-(r(2)+2)}$. It follows from (4.32) with bars that

$$\lim_{j(k)\to\infty} \frac{3n(G_1) + n(G_2) + n(G_3) + n(G_4) + n(G_5) + n(G_6) + n(G_7)}{m(k)} = 1 \tag{4.35}$$

when $n(G_i)$ is the number of elements in G_1. Furthermore, in (4.32) without bars the sums can be taken respectively over $G_1, \cdots, G_7$ since $y_s = 1$ implies $\bar{y}_s = 1$. We note also that for each i the set G_i is a set of consecutive integers, and that G_2, G_3, and G_4 as well as G_5, G_6, and G_7 differ by at most one element. This implies that

$$\sum_{i=1}^{m(k)} y_i(k) y_{i-N-t}(k) \leq K_1^* \sum_{i=1}^{m(k-1)} x_i(k-1) x_{i-N-t}(k-1), \qquad t = 0, 1, \tag{4.36}$$

where K_1^* is not larger than 2 plus the whole number of times that $m(k-1)$ goes into the number of elements in G_1. (We had denoted the number of elements in G_1 by $n(G_1)$.)

Similarly, we have that

$$\begin{aligned} &\sum_{i=1}^{m(k)} y_i(k) y_{i-r(1)-t}(k) \leq K_2^* \sum_{i=1}^{m(k-1)} x_i(k-1) x_{i-r(1)-t}(k-1), \\ &\sum_{i=1}^{m(k)} y_i(k) y_{i-r(1)-1-t}(k) \leq K_3^* \sum_{i=1}^{m(k-1)} x_i(k-1) x_{i-r(1)-1-t}(k-1), \end{aligned} \tag{4.37}$$

and that

$$(4.38)\qquad \sum_{i=1}^{m(k)} y_i(k)y_{i-r(1)-2-t}(k) \leq K_4^* \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(1)-2-t}(k-1),$$

$t = 0, 1$, where K_i^* is not larger than 2 more than the number of times $m(k-1)$ goes into $n(G_i)$, $i = 2, 3, 4$.

Similarly,

$$(4.39)\qquad \begin{aligned} \sum_{i=1}^{m(k)} y_i(k)y_{i-r(2)-t}(k) &\leq K_5^* \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(2)-t}(k-1),\\ \sum_{i=1}^{m(k)} y_i(k)y_{i-r(2)-1-t}(k) &\leq K_6^* \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(2)-1-t}(k-1), \end{aligned}$$

and

$$(4.40)\qquad \sum_{i=1}^{m(k)} y_i(k)y_{i-r(2)-2-t}(k) \leq K_7^* \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(2)-2-t}(k-1),$$

where $t = 0, 1$, and where K_5^*, K_6^*, and K_7^* are defined as in the previous cases and are not larger, respectively, than 2 plus the number of times that $m(k-1)$ goes into $n(G_5)$, $n(G_6)$, and $n(G_7)$. We see from the remarks following (4.35) that we may take $K_1 = K_1^*$, $K_2 = \max(K_2^*, K_3^*, K_4^*)$, and $K_3 = \max(K_5^*, K_6^*, K_7^*)$.

Before proceeding to the next lemma it is convenient to introduce the following notation:

$$(4.41)\qquad M(k) = \frac{1}{3m(k)} \sup_r \left[\sum_{i=1}^{m(k)} x_i(k)x_{i-r}(k) + \sum_{i=1}^{m(k)} x_i(k)x_{i-r-1}(k) + \sum_{i=1}^{m(k)} x_i(k)x_{i-r-2}(k)\right]$$

noting that $M(0) = \frac{1}{3}$.

LEMMA 10. *Let $j(1), \cdots, j(k-1)$ be given so that*

$$(4.42)\qquad M(r) < \tfrac{1}{3} + \epsilon, \qquad r = 1, \cdots, k-1.$$

Then there exists N so that $j(k) \geq N$ implies that

$$(4.43)\qquad M(k) < \tfrac{1}{3} + \epsilon.$$

PROOF. The lemma is an immediate consequence of lemma 9.

LEMMA 11. *Given $\epsilon > 0$, then if $\{j(k)\}$ tends to infinity sufficiently rapidly, we have for all n*

$$(4.44)\qquad \frac{1}{m(n)} \sum_{i=1}^{m(n)} x_i(n)x_{i-j}(n) < \tfrac{1}{3} + \epsilon,$$

for all j an odd integer.

PROOF. The result holds for $n = 0$, as we easily see. We suppose that $j(1), \cdots, j(k-1)$ has been chosen so that the result holds for $n = 1, \cdots, k-1$, and also so that

$$(4.45)\qquad M(n) < \tfrac{1}{3} + \epsilon,$$

holds also for $n = 1, \cdots, k-1$. It then follows from lemma 9 that

$$\sum_{i=1}^{m(k)} x_i(k)x_{i-j}(k) \le 3K_1 \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-(j)}(k-1) \tag{4.46}$$
$$+ K_2\left[\sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(1)}(k-1) + \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(1)-1}(k-1)\right.$$
$$\left. + \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(1)-2}(k-1)\right]$$
$$+ K_3\left[\sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(2)}(k-1) + \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(2)-1}(k-1)\right.$$
$$\left. + \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(2)-2}(k-1)\right].$$

Applying the induction hypothesis and (4.45) to the right side of (4.46), we obtain for $j(k) \ge N(j(1), \cdots, j(k-1)$ that it is less than or equal to

$$3K_1m(k-1)\,(\tfrac{1}{3}+\epsilon) + 3K_2m(k-1)\,(\tfrac{1}{3}+\epsilon) + 3K_3m(k-1)\,(\tfrac{1}{3}+\epsilon). \tag{4.47}$$

If we now divide through by $m(k)$ and apply the first part of lemma 9, we see that if $j(k)$ is chosen sufficiently large, the conclusion of the lemma holds for $n = k$. This concludes the proof of the lemma.

LEMMA 12. *Given $\epsilon > 0$, then if $\{j(k)\}$ tends to infinity sufficiently rapidly we have*

$$\mu(A \cap \tau^j(A)) < \tfrac{1}{3} + \epsilon, \tag{4.48}$$

for all j an odd integer.

PROOF. This follows directly from lemma 11 and from lemmas 7 and 8.

As we have remarked earlier, $\tau A_{m(k)}(k)$ equals $A_1(k)$, with error tending to zero as $j(k)$ tends to infinity. (Recall that this means that the measure of the symmetric difference of the sets tends to zero.)

Before proceeding to the next lemma, we introduce more notation: let $G_j(k)$, $1 \le k \le m(k)$, be that subset of $A_j(k)$ which is mapped by $\tau^{m(k)-j+1}$ into $A_1(k)$. Then we have $G_j(k) = \tau^{j-m(k)}G_{m(k)}(k)$. We also let

$$G^*(k) = \bigcup_{j=1}^{m(k)} G_j(k), \qquad A^*(k) = \bigcup_{j=1}^{m(k)} A_j(k). \tag{4.49}$$

It follows from our preceding remarks that $G^*(k)$, which is obviously contained in $A^*(k)$, equals $A^*(k)$ with error tending to zero as $j(k+1)$ tends to infinity. Next, let $A = (\tfrac{1}{2}, 1)$ as in lemma 7. We have then, as before,

$$A = \sum_{i=1}^{m(k)} x_i(k)A_i(k), \tag{4.50}$$

and, for any transformation σ,

$$\sigma(A) = \sum_{i=1}^{m(k)} x_i(k)\sigma(A_i(k)). \tag{4.51}$$

Recall that we write $A = B + E(\delta)$ for two sets whose symmetric difference has measure less than δ.

LEMMA 13. *Let σ be an arbitrary measurable transformation which commutes with τ and let $A = (\frac{1}{2}, 1)$. We have then*

$$\sigma(A \cap G^*(k)) = \sum_{i=1}^{m(k)} x_i(k)\tau^{i-1}(\sigma(G_1(k))) \cap G^*(k) + E(\delta), \tag{4.52}$$

where δ tends to zero uniformly in σ as $j(k+1)$ tends to infinity (and where the $x_i(k)$ is given as in lemma 7).

PROOF. First we establish that

$$G^*(k) = X + N + E(\eta) \tag{4.53}$$

where η tends to zero as $j(k+1)$ tends to infinity. To see this note that we may write

$$cG^*(k) = cA^*(k) + \sum_{j=0}^{m(k)} \tau^{-j}[A_{m(k)}(k) - G_{m(k)}(k)], \tag{4.54}$$

and consequently,

$$\begin{aligned}\mu(cG^*(k)) &\leq \mu(cA^*(k)) + m(k)\mu(A_{m(k)}(k) - G_{m(k)}(k)) \\ &\leq \mu(cA^*(k)) + \frac{m(k)\mu(A_{m(k)}(k))}{j(k+1)}.\end{aligned} \tag{4.55}$$

This implies (4.53) as we may readily see from the construction.

Next we note that (4.53) implies

$$\sigma(G_1(k)) \cap G^* = \sigma(G_1(k)) + E(\delta/m(k)), \tag{4.56}$$

where δ tends to zero uniformly in σ as $j(k+1)$ tends to infinity. It follows from the definition that

$$A \cap G^*(k) = \sum_{i=1}^{m(k)} x_i(k)G_i(k) = \sum_{i=1}^{m(k)} x_i\tau^{i-1}G_1(k). \tag{4.57}$$

Equation (4.56) and the fact that τ is measure preserving imply that

$$\tau^{i-1}\sigma(G_1(k)) = \tau^{i-1}(\sigma(G_1(k)) \cap G^*) + E(\delta/m(k)). \tag{4.58}$$

Equations (4.57) and (4.58) together with the fact that σ and τ commute yield

$$\sigma(A \cap G^*(k)) = \sum_{i=1}^{m(k)} x_i\tau^{i-1}(\sigma(G_1(k)) \cap G^*(k)) + E(\delta), \tag{4.59}$$

where δ tends to zero uniformly in σ as $j(k+1)$ tends to infinity.

LEMMA 14. *If σ is a measurable square root of τ, then for each k there exist $\{H_i(k), 1 \leq i \leq m(k)\}$, pairwise disjoint subsets of $G_1(k)$, such that*

$$\begin{aligned}G^*(k) \cap \sigma(G_1(k)) &= \sum_{i=1}^{m(k)} \tau^{i-1}H_i(k), \\ H_i(k) &= \tau^{1-i}(G_i(k) \cap \sigma(G_1(k))).\end{aligned} \tag{4.60}$$

PROOF. If we define the sets by the second formula of (4.60) it is clear that they satisfy the first. It only remains to show that the sets are pairwise disjoint.

In order to see this, suppose that there exists $i(1) < i(2)$ and a point $x \in H_{i(1)} \cap H_{i(2)}$. It follows from the definitions that this implies that there exist two points x_1 and x_2 in $G_1(k)$ such that $\sigma(x_1) = \tau^{i(2)-i(1)}\sigma(x_2)$. Applying σ to both sides of the equation, we find that x belongs to $G_{i(2)-i(1)}(k)$, which yields a contradiction.

In the following lemma, as in lemma 13, the $x_i(k)$ are given as in lemma 7.

LEMMA 15. *If σ is a measurable square root of τ, then there exist $\{J_i(k), 1 \le i \le m(k)\}$ pairwise disjoint subsets of $A_1(k)$ such that*

$$\sigma(A \cap G^*(k)) = \sum_{j=1}^{m(k)} \sum_{i=1}^{m(k)} x_i(k)\tau^{(i+j-2) \bmod m(k)} J_j(k) + E(\delta), \tag{4.61}$$

where δ tends to zero uniformly in σ as $j(k+1)$ tends to infinity.

PROOF. We note that as the $j(k)$ are changed, the transformation τ is changed, and consequently, so are its square roots. The uniformity condition given by the lemma is with respect to this changing class of square roots.

From lemmas 13 and 14, with $J_j(k) = H_j(k)$, we have

$$\sigma(A \cap G^*(k)) = \sum_{j=1}^{m(k)} \sum_{i=1}^{m(k)} x_i(k)\tau^{i-1}\tau^{j-1}J_j(k) + E(\delta_1), \tag{4.62}$$

where δ_1 tends to zero uniformly in σ as $j(k+1)$ tends to infinity. Note also, as we shall show, that

$$\tau^{m(k)}J_i(k) \cap J_j(k) = \varnothing \tag{4.63}$$

if $i \ne j$, $1 \le i, j \le m(k)$. If not, there exist $i(1)$ and $i(2)$ such that

$$\tau^{m(k)-(i(1)-1)+(i(2)-1)}\tau^{i(1)-1}J_{i(1)} \cap \tau^{i(2)-1}J_{i(2)} \ne \varnothing. \tag{4.64}$$

This implies that there exist points x_1 and x_2 in $G_1(k)$ such that

$$\tau^{m(k)-(i(1)-1)+(i(2)-1)}x_1 = x_2. \tag{4.65}$$

Equation (4.65) and the fact that the points are in $G_1(k)$ imply that the exponent in (4.65) is equal to zero modulo $m(k)$. This is impossible, and we have (4.63).

Equation (4.63) implies that

$$\tau^j J_i = \tau^{j \bmod m(k)} J_i + E(\delta_i), \qquad 0 \le j \le 2m(k), \tag{4.66}$$

where the sum of the δ_i is bounded by the measure of $cG_1(k) \cap A_1(k)$. This in turn implies that

$$\begin{aligned} &\sum_{j=1}^{m(k)} \sum_{i=1}^{m(k)} x_i(k)\tau^{i-1}\tau^{j-1}J_j(k) \\ &\qquad = \sum_{j=1}^{m(k)} \sum_{i=1}^{m(k)} x_i(k)\tau^{(i+j-2) \bmod m(k)}J_j(k) + E[m(k)\mu(cG_1(k) \cap A_1(k))], \end{aligned} \tag{4.67}$$

which, together with (4.62), yields the lemma, since the measure of the set $cG_1(k) \cap A_1(k)$ tends to zero as $j(k+1)$ tends to infinity.

The next lemma is purely combinatorial, as are lemmas 8 through 11. It will enable us to make estimates on the size of the intersection of the set $A = (\frac{1}{2}, 1)$

with odd translates of itself (that is, translates under τ^{2n-1}) under the assumption that a nonsingular square root exists. The estimate will be that the size of the intersection of A with certain of its odd translates is nearly as big as A. This, of course, is sufficient to prove the theorem. The sequences we refer to in the lemma which follows need not be given as in lemma 7, although it will only be applied to such sequences. Finally, remark that the set A was selected to equal $(\frac{1}{2}, 1)$ for convenience. It could have been chosen equally well to be $(0, \frac{1}{2})$ for example. For a set G, we again let $n(G)$ be the number of elements in it.

LEMMA 16. *Let $\alpha_i = (x_{i1}, \cdots, x_{in})$, $i = 1, \cdots, k$ be k sequences of zeros and ones of length n. Let $\{b_i, i = 1, \cdots, k\}$ be nonnegative numbers adding to unity, and suppose that there is a subset G of $\{1, \cdots, n\}$ such that*

$$\sum_{i=1}^{k} b_i x_{ij} \geq 1 - \eta \qquad \text{for all } j \text{ in } G. \tag{4.68}$$

Then there exists an integer w, $1 \leq w \leq k$, such that

$$\sum_{j \text{ in } G} \sum_{i=1}^{k} b_i x_{ij} x_{wj} \geq n(G)(1 - 2\eta). \tag{4.69}$$

PROOF. We may assume, without loss of generality, that $b_i = 1/k$, if we proceed as follows. First note that we can certainly suppose them to be rational numbers. Writing these numbers over a common denominator, we have $b_i = m_i/m$. Then consider the problem for the $\bar{\alpha}$'s, obtained by taking each α_i m_i times; this reduces the lemma to the case where $\bar{k}$ equals the sum of the m_i's, and where all the $\bar{b}$'s are equal to $1/m$.

Let $B(i)$ be the subset of G such that for j in $B(i)$

$$x_{ij} = 0. \tag{4.70}$$

Let

$$\delta = \inf_{1 \leq i \leq k} \frac{n(B(i))}{n(G)}, \tag{4.71}$$

where $n(B(i))$ and $n(G)$ stand for the number of elements in $B(i)$ and G respectively. By changing the order of summation from (4.71), we have

$$\sum_{j \text{ in } G} \left[\sum_{i=1}^{k} \psi_{B(i)}(j) \right] \geq kn(G)\delta. \tag{4.72}$$

It then follows easily from (4.72) that for at least one value of j, say $j(0)$,

$$\sum_{i=1}^{k} \psi_{B(i)}(j(0)) \geq k\delta. \tag{4.73}$$

Equation (4.73) implies that $[k\delta]$ of the sets B_i intersects (here $[k\delta]$ means the least integer greater than or equal to $k\delta$), say $B_{i(1)}, B_{i(2)}, \cdots, B_{i([k\delta])}$, and $j(0) \in C = \bigcap_{j=1}^{[k\delta]} B_{i(j)}$. We then have

$$\sum_{j=1}^{k} x_{jj(0)} = \sum_{j \in (i(1), \cdots, i([k\delta]))} x_{jj(0)} + \sum_{j \notin (i(1), \cdots, i([k\delta]))} x_{jj(0)}. \tag{4.74}$$

It follows from the definition of $B_{i(r)}$ that $x_{jj(0)} = 0$ for $j \in (i(1), \cdots, i([k\delta]))$, and consequently, (4.74) implies that

$$\sum_{j=1}^{k} x_{jj(0)} \leq k - [k\delta] \leq k(1 - \delta). \tag{4.75}$$

Since we have assumed that $b_j = 1/k$, (4.75) implies

$$\sum_{j=1}^{k} b_j x_{jj(0)} \leq (1 - \delta), \tag{4.76}$$

which together with (4.68) and (4.71), yields

$$\inf_{1 \leq i \leq n} \frac{n(B(i))}{n(G)} = \delta \leq \eta. \tag{4.77}$$

Equation (4.77) in turn yields that there exists an integer w such that

$$\frac{n(B(w))}{n(G)} \leq \eta. \tag{4.78}$$

Now,

$$\begin{aligned} \sum_{j \in G} \sum_{i=1}^{k} x_{ij} x_{wj} &= \sum_{i=1}^{k} \sum_{j \in G} x_{ij} x_{wj} \\ &= \sum_{i=1}^{k} \sum_{j \in G - B(w)} x_{ij} x_{wj} + \sum_{i=1}^{k} \sum_{j \in B(w)} x_{ij} x_{wj} \\ &= \sum_{i=1}^{k} \sum_{j \in G - B(w)} x_{ij} x_{wj} = \sum_{i=1}^{k} \sum_{j \in G - B(w)} x_{ij}, \end{aligned} \tag{4.79}$$

so that

$$\begin{aligned} \sum_{j \in G} \sum_{i=1}^{k} x_{ij} x_{wj} &= \sum_{i=1} \sum_{j \in G - B(w)} x_{ij} \\ &\quad + \sum_{i=1}^{k} n(B(w)) - \sum_{i=1}^{k} n(B(w)) \\ &\geq \sum_{i=1}^{k} \sum_{j \in G} x_{ij} - kn(B(w)). \end{aligned} \tag{4.80}$$

Keeping in mind that $b_i = 1/k$, this implies

$$\begin{aligned} \sum_{j \in G} \sum_{i=1}^{k} b_i x_{ij} x_{wj} &\geq \sum_{j \in G} \sum_{i=1}^{k} b_i x_{ij} - n(B(w)) \\ &\geq n(G)(1 - \eta) - n(B(w)) \\ &= n(G) \left\{ 1 - \eta - \frac{n(B(w))}{n(G)} \right\} \\ &= n(G)\{1 - 2\eta\}. \end{aligned} \tag{4.81}$$

LEMMA 17. *If $\{j(k)\}$ tends to infinity sufficiently rapidly, and if $\tau(j(k))$ has a nonsingular square root σ, then there exists an integer $N(\eta)$ such that*

$$\mu(\sigma(A) - \tau^N(A)) \leq \eta \tag{4.82}$$

for each $\eta > 0$.

PROOF. If $\{j(k)\}$ tends to infinity sufficiently rapidly, then there exists a sequence $\{\delta(k)\} \downarrow 0$ such that

$$\lim_{k\to\infty} \delta(k)/\mu(A_1(k)) = 0, \tag{4.83}$$

$$\sigma(A \cap G^*(k)) = \sum_{j=1}^{m(k)} \sum_{i=1}^{m(k)} x_{i-j+1}(k) D_{i,j}(k) + E(\delta(k)), \tag{4.84}$$

where $D_{i,j}(k) = \tau^{i-1}J_j(k)$, and consequently, where

$$D_{i,j}(k) \subset A_i(k), \qquad 1 \leq i, \quad j \leq m(k), \tag{4.85}$$

and

$$\tau D_{i,j}(k) = D_{i+1,j}(k), \qquad 1 \leq j \leq m(k) - 1, \tag{4.86}$$

as we may see from lemma 15. It follows from lemma 7 that we may also satisfy, at the same time, the conditions

$$G^*(k) = X + N + E(\delta(k)) \tag{4.87}$$

and

$$\int \left| \left[\sum_{i=1}^{m(k)} \psi A_i x_{i-j} \right] - \psi_{\tau^j A} \right| \leq \delta(k), \qquad j = 0, 1, \cdots, m(k). \tag{4.88}$$

We have from lemmas 1 and 2 that for $\lambda > 0$ and $M > 0$ there exists an $n \geq M$ and an index set G such that

$$\sigma(A) = \bigcup_{i\in G} A_i(n) + E(\lambda), \tag{4.89}$$

$$\sigma(A) \cap A_i(n) = A_i(n) + E(\lambda\mu(A_i(n))), \qquad i \in G. \tag{4.90}$$

From (4.84), (4.87), and (4.90) we obtain

$$\sum_{j=1}^{m(k)} x_{i-j+1}(n) D_{i,j}(n) = A_i(n) + E(2\delta(n) + \lambda\mu(A_i(n))). \tag{4.91}$$

We then define $b_j = \mu(D_{i,j})/\mu(\cup_{j=1}^{m(k)} D_{i,j})$ and note that (4.91) yields

$$\sum_{i=1}^{m(k)} x_{i-j+1} b_j \geq 1 - \left[\frac{2\delta(n)}{\mu(A_i(n))} + \lambda\right], \qquad i \in G. \tag{4.92}$$

Since the sequence $\{\delta(k)\}$ has property (4.83), we have that if M is sufficiently large, then

$$\sum_{j=1}^{m(k)} x_{i-j+1} b_j \geq 1 - 2\lambda, \qquad i \in G. \tag{4.93}$$

We are now in a position to apply lemma 16 with $\alpha_i = \{x_{j-i+1}, 1 \leq j \leq m(k)\}$ in order to find that there exists a w such that

$$\sum_{i\in G} \sum_{j=1}^{m(k)} b_j x_{i-j+1} x_{i-w+1} \geq n(G)(1 - 4\lambda). \tag{4.94}$$

In order to evaluate the left side of equation (4.94), we keep in mind the definition of the constants x_i and b_j, and the fact that the sets $A_i(k)$, $i = 1, 2, \cdots, m(k)$ are pairwise disjoint, and that

$$D_{i,j}(k) \subset A_i(k). \tag{4.95}$$

We see that it is given by

$$\frac{1}{\mu\left(\sum_{j=1}^{m(k)} D_{i,j}\right)} \int \left\{\sum_{i=1}^{m(k)} \psi_{A_i} x_{i-w+1}\right\}\left\{\sum_{i \in G} \sum_{j=1}^{m(k)} x_{i-j+1}\psi_{D_{i,j}}\right\}. \tag{4.96}$$

Further, it follows from (4.91) that

$$\sum_{j=1}^{m(k)} D_{i,j}(n) = A_1(n) + E(2\delta(n) + \lambda\mu(A_1(n))), \tag{4.97}$$

and thus, that for M large enough,

$$\mu\left(\sum_{j=1}^{m(k)} D_{i,j}\right) \geq \mu(A_1)(1 - \lambda). \tag{4.98}$$

Formulas (4.89) and (4.98) imply

$$\mu\left(\sum_{j=1}^{m(k)} D_{1,j}\right) n(G)(1 - 4\lambda) \geq (\mu(\sigma(A)) - \lambda)(1 - \lambda)(1 - 4\lambda)). \tag{4.99}$$

Furthermore we have that (4.84), (4.87), and (4.89) imply that

$$\sigma(A) = \sum_{i \in G} \sum_{j=1}^{m(k)} x_{i-j+1} D_{i,j}(n) + E(\lambda + 2\delta(n)), \tag{4.100}$$

and consequently, that with M sufficiently large,

$$\int \left| \sum_{i \in G} \sum_{j=1}^{m(k)} x_{i-j+1}\psi D_{i,j} - \psi_{\sigma(A)} \right| \leq \lambda + 2\delta(n) \leq 2\lambda. \tag{4.101}$$

Putting (4.94), (4.96), (4.99), (4.88), and (4.101) together, we obtain that

$$\int \psi_{\tau^{w-1}(A)}\psi_{\sigma(A)} \geq \mu(\sigma(A)) + \epsilon(\lambda), \tag{4.102}$$

where $\epsilon(\lambda)$ tends to zero as λ tends to zero, and we see that the lemma follows by taking λ sufficiently small, with $N = w - 1$.

Proof of the theorem in the square root case. To see that the theorem holds, we choose a sequence $\{j(k)\}$ tending to infinity sufficiently rapidly for lemmas 5, 6, 12, and 17 to hold. We then have a transformation having continuous spectrum which is not strongly mixing and which satisfies the condition that

(i) $\mu(A \cap \tau^j A) < \frac{1}{3} + \epsilon$,

for all odd j, for $A = (\frac{1}{2}, 1)$. Furthermore, if the transformation has a nonsingular square root, then for each $\eta > 0$ there is an $N = N(\eta)$ such that

(ii) $\mu(\sigma(A) - \tau^N(A)) \leq \eta$.

Part (ii) means that $\sigma(A) \subset \tau^N(A) + F$, where the measure of F is less than or equal to η, and thus

$$\begin{aligned} \sigma\sigma(A) &\subset \tau^N(\sigma(A)) + \sigma(F) \\ &\subset \tau^N(\tau^N(A) + F) + \sigma(F) \\ &= \tau^{2N}(A) + \tau^N(F) + \sigma(F). \end{aligned} \tag{4.103}$$

Since σ is nonsingular and τ is measure preserving, this implies that

$$A = \tau^{2N-1}(A) + E(\eta) + E(\epsilon(\eta)), \tag{4.104}$$

where $\epsilon(\eta)$ tends to zero as η tends to zero, contradicting (i), if ϵ is sufficiently small and η is sufficiently small.

Definition of the transformation in the general case. Let $p(n)$ be the $(n + 1)$-th prime, so that $p(1) = 2$, and let $\pi(n) = p(1) \cdots p(n)$. Given a sequence $\{j(k), k = 1, 2, \cdots\}$ of integers with $j(k) \geq 1, k \geq 1$, we shall define an invertible measure preserving transformation $\tau = \tau(\{j(k)\})$ having the desired properties. For convenience, the transformation τ is defined on the union of the unit interval with another interval N disjoint from it; the length of N depends on the sequence $\{j(k)\}$.

The transformation is defined inductively. For $r = k$, we suppose that the transformation is defined on part of the space, and we extend the domain of definition for $r = k + 1$.

For $r = k$, we suppose the transformation given as follows. Let $m(k) = 3m(k - 1)j(k) + 3\pi(k)$, $m(0) = 2$, and suppose that there exist $m(k)$ intervals of equal length $A_1(k), \cdots, A_{m(k)}(k)$, each a subinterval of the unit interval or of N, and that τ maps $A_{i-1}(k)$ linearly onto $A_i(k)$, $i = 2, \cdots, m(k)$.

For $r = k + 1$, we need to extend the definition of τ in such a way that it admits a representation, as in the preceding paragraph, with k replaced by $k + 1$. This is accomplished in two steps.

The first step is the following. Divide $A_i(k)$, $(1 \leq i \leq m(k))$ into $j(k + 1)$ consecutive subintervals of equal length denoting them

$$B_{i,1}(k + 1), \cdots, B_{i,j(k+1)}(k + 1), \qquad (1 \leq i \leq m(k)). \tag{4.105}$$

Since $A_{i-1}(k)$ is mapped linearly onto $A_i(k)$, $(2 \leq i \leq m(k))$, τ also maps $B_{i-1,j}(k + 1)$ linearly onto

$$B_{i,j}(k + 1), \qquad (2 \leq i \leq m(k), 1 \leq j \leq j(k + 1)). \tag{4.106}$$

In this first step we extend τ by mapping $B_{m(k),j}(k + 1)$ linearly onto

$$B_{i,j+1}(k + 1), \qquad (1 \leq j \leq j(k + 1) - 1). \tag{4.106a}$$

There are $m(k)$ $A_i(k)$'s, and each $A_i(k)$ is divided into $j(k + 1)$ $B_{i,j}(k + 1)$'s so that there are $m(k)j(k + 1)B_{i,j}(k + 1)$'s altogether. We write these sets with a single subscript $B_i(k + 1)$, $(i = 1, \cdots, m(k)j(k + 1))$, in such a way that τ, as thus far extended, maps $B_{i-1}(k + 1)$ linearly onto

$$B_i(k + 1), \qquad (2 \leq i \leq m(k)j(k + 1)). \tag{4.106b}$$

The second step in the extension of τ is obtained by dividing $B_i(k + 1)$ into three consecutive subintervals of equal length $B_i^1(k + 1)$, $B_i^2(k + 1)$, $B_i^3(k + 1)$. The transformation τ, as already defined, maps $B_{i-1}^u(k + 1)$ linearly onto

$$B_i^u(k + 1), \qquad (2 \leq i \leq m(k)j(k + 1), 1 \leq u \leq 3). \tag{4.107}$$

We take from that subinterval of N, where τ is as yet undefined, $3\pi(k + 1)$ consecutive subintervals of equal length equal to the length of the $B_i^u(k + 1)$ and denote these $3\pi(k+1)$ subintervals of N by $E_j(k+1)$, $(1 \le j \le 3\pi(k+1))$. (We suppose that N is just big enough for there to be an interval left over at each stage.) Then extend τ so that the sequence of $3m(k)j(k+1) + 3\pi(k+1)$ sets (with $(k+1)$ deleted for simplicity), with $s = m(k)j(k+1)$,

$$E_1, \cdots, E_{\pi(k+1)-1}, B_1^1, \cdots, B_s^1, E_{\pi(k+1)}, \cdots, E_{2\pi(k+1)-1}, B_1^2, \cdots, B_s^2, E_{2\pi(k+1)}, \cdots, E_{3\pi(k+1)}, B_1^3, \cdots, B_s^3, \tag{4.108}$$

having the property that τ maps each set, except for the last, linearly onto the next set. We denote this sequence by $A_i(k+1)$, $(1 \le i \le m(k+1))$, where $m(k+1) = 3m(k)j(k+1) + 3\pi(k+1)$.

To complete the induction it remains to define τ for $r = 0$. We set $A_1(0) = (0, \frac{1}{2})$ and $A_2(0) = (\frac{1}{2}, 1)$ and define τ initially as the linear map of $A_1(0)$ onto $A_2(0)$. An intuitive description of the transformation is similar to the square root case and is the following.

At the k-th stage there is a stack of $m(k)$ intervals of equal length, and τ maps each point to the one directly above so that the points of the top interval are not yet mapped anywhere.

The first step of the extension of τ is obtained as follows. The original stack is split into $j(k+1)$ equal and consecutive substacks, and these are stacked in order. This has the effect of mapping a $j(k+1) - 1/j(k+1)$ part of the top interval of the original stack into the bottom interval of the original stack.

In the second step of the extension, we take the resulting stack, which is composed of $m(k)j(k+1)$ intervals, and divide it into three consecutive and equal substacks. We then take $3\pi(k+1)$ consecutive and equal subintervals of N, where τ is as yet undefined, of length equal to the length of each of the intervals of the three stacks, and put $\pi(k+1) - 1$ under the first stack, $\pi(k+1)$ under the second stack, and $\pi(k+1) + 1$ under the third stack. The thus modified three stacks are then stacked in order.

Properties of the transformation in the general case. The transformation is again defined inductively. In what follows we again do not suppose that the measure space is normalized. We need to make one unimportant remark before proceeding. The sequence $\{j(k)\}$ must tend to infinity sufficiently rapidly so that N has finite measure. This was true automatically before, since $\pi(k) \equiv 2$ then. We divide the proof into several lemmas. When the proofs are exactly as in the square root case we omit them.

LEMMA 1. *Let $\{E_{j,k}, 1 \le j \le r(k), 1 \le k < +\infty\}$ generate the Borel field, and suppose that for each k, $1 \le k < +\infty$, the sets $\{E_{j,k}, 1 \le j \le r(k)\}$ are pairwise disjoint. Suppose also that each $E_{j,k}$ is the union of a finite number of sets from the class $\{E_{j,k+1}, 1 \le j \le r(k+1)\}$. Then for each $\epsilon > 0$ and A in the Borel field, there exists a set of indices $G(\epsilon, A)$ and an integer $k(\epsilon, A)$ such that A equals $\bigcup_{j \in G(\epsilon,A)} E_{j,k(\epsilon,A)}$ with an error of ϵ, and*

$$\mu(A \cap E_{j,k(\epsilon,A)}) \ge (1-\epsilon)\mu(E_{j,k(\epsilon,A)}) \tag{4.109}$$

for $j \in G(\epsilon, A)$.

LEMMA 2. *The sets $\{A_j(k), 1 \le j \le m(k), 1 \le k < +\infty\}$ generate the Borel* *field; furthermore, each* $A_j(k)$ *is the union of* $3j(k+1)$ *sets of the class* $\{A_j(k+1), 1 \leq j \leq m(k+1)\}$.

LEMMA 3. *The sets* $\{B_j(k), 1 \leq j \leq m(k-1)j(k), 1 \leq k < +\infty\}$ *generate the Borel field; furthermore, each* $B_j(k)$ *is the union of* $3j(k+1)$ *sets of the class* $\{B_j(k+1), 1 \leq j \leq m(k)j(k+1)\}$.

LEMMA 4. *Given* $\{\delta(k)\} \downarrow 0$, *and if* $\{j(k)\}$ *tends to infinity sufficiently rapidly, then for all* k,

(i) $\mu(\bigcup_{j=1}^{m(k)} A_j(k)) \geq (1 - \delta(k))\mu(X + N)$,

(ii) $\mu(\tau A_{m(k)}(k) \triangle A_1(k)) \leq (1 - \delta(k))\mu(A_1(k))$,

(iii) $m(k)$ *depends only on* $\delta(0), \cdots, \delta(k-1)$.

LEMMA 5. *Let* $\{j(k)\}$ *be given. Then* $\tau(\{j(k)\})$ *has continuous spectrum.*

LEMMA 6. *The transformation* $\tau(\{j(k)\})$ *is not strongly mixing.*

PROOF. Let $B = B_1(k+1)$. We have from the construction that

$$\mu(B \cap \tau^{m(r)j(r)+\pi(r)}B) \geq \tfrac{1}{3}\mu(B), \tag{4.110}$$

for $r \geq k$. Since $\mu(B_1(k+1))$ tends to zero as k tends to infinity, the normalized measure of $B_1(k+1)$ also has this property. The proof is complete if we take k sufficiently large since strong mixing implies that

$$\lim_{j\to\infty} \frac{\mu(B \cap \tau^j B)}{\mu(X+N)} = \left[\frac{\mu(B)}{\mu(X+N)}\right]^2. \tag{4.111}$$

We again define $1A = A$, $0A = \varnothing$, where A is a set, and $\varnothing$ is the empty set. We fix the prime p and prove that the transformation has no p-th root.

LEMMA 7. *Let* $A = A_{m(p)}(p)$, *and let* $\{j(k)\}$ *be given. Then for each* $k \geq p$ *there exists a sequence* $\{x_i(k), 1 \leq i \leq m(k)\}$ *of zeros and ones such that*

$$A = \sum_{i=1}^{m(k)} x_i(k)A_i(k). \tag{4.112}$$

Furthermore, we have

$$\tau^j A = \sum_{i=1}^{m(k)} x_{i-j}A_i(k) + E(\eta(j(k+1))), \tag{4.113}$$

for $j = 1, \cdots, m(k)$, *where* $\eta(j(k+1))$ *tends to zero as* $j(k+1)$ *tends to infinity, where the subscripts are taken modulo* $m(k)$.

LEMMA 8. *Let* $\alpha(k) = (x_1(k), \cdots, x_{m(k)}(k))$ *where the sequences* $(x_1(k), \cdots, x_{m(k)}(k))$ *are those of lemma 7. Then the* $\alpha(k)$ *are given inductively by*

$$\alpha(p) = \Big(\overbrace{0, \cdots, 0, 1}^{m(p)}\Big) \tag{4.114}$$

$$\vdots$$

$$\alpha(k) = \Big(\overbrace{0, \cdots, 0}^{\pi(k)-1}, \overbrace{\alpha(k-1), \cdots, \alpha(k-1)}^{j(k)}, \overbrace{0, \cdots, 0}^{\pi(k)}, \overbrace{\alpha(k-1), \cdots, \alpha(k-1)}^{j(k)}, \overbrace{0, \cdots, 0}^{\pi(k)+1}, \overbrace{\alpha(k-1), \cdots, \alpha(k-1)}^{j(k)}\Big).$$

We note that subscripts on the elements of $\alpha(k)$ are understood to be modulo $m(k)$, so that letting $i - j = r$ modulo $m(k)$, then $x_{i-j}(k) = x_r(k)$. Furthermore, note that $m(k)$, $k \geq p$, are numbers divisible by p so that with the convention we use, $x_N(k)$ has a subscript divisible by p if and only if N is divisible by p for each $k \geq p$.

LEMMA 9. *Let M be an integer, $k \geq p + 1$, and let $N = \langle M\rangle$, $r(1) = \langle N + m(k-1)j(k) + 1\rangle$, and $r(2) = \langle N + 2m(k-1)j(k) + 3\rangle$, where $\langle a\rangle$ is that integer equal to a modulo $m(k)$ which satisfies $-m(k)/2 \leq \langle a\rangle < m(k)/2$. Then there exist three positive numbers K_1, K_2, K_3 (depending on $j(k)$) such that*

$$\lim_{j(k)\to\infty} \frac{3(K_1 + K_2 + K_3)m(k-1)}{m(k)} = 1, \tag{4.115}$$

and such that, for $t = 0, 1$,

$$\begin{aligned}\sum_{i=1}^{m(k)} x_i(k)x_{i-M-t}(k) \leq{}& 3K_1 \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-N-t}(k-1)\\ &+ K_2\Bigg[\sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(1)-t}(k-1) + \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(1)-1-t}(k-1)\\ &\qquad + \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(1)-2-t}(k-1)\Bigg]\\ &+ K_3\Bigg[\sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(2)-t}(k-1) + \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(2)-1-t}(k-1)\\ &\qquad + \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(2)-2-t}(k-1)\Bigg].\end{aligned} \tag{4.116}$$

We introduce the following notation:

$$M(k) = \frac{1}{3m(k)} \sup_r \left[\sum_{i=1}^{m(k)} x_i(k)x_{i-r}(k) + \sum_{i=1}^{m(k)} x_i(k)x_{i-r-1}(k) + \sum x_i(k)x_{i-r-2}(k)\right]. \tag{4.117}$$

for $k \geq p > 2$, noting that $M(p) = \frac{1}{3}m(p)$.

LEMMA 10. *Let $j(p), \cdots, j(k-1)$ be given so that*

$$M(r) < \frac{1}{3m(p)} + \epsilon, \qquad r = p, \cdots, k-1. \tag{4.118}$$

Then there exists N so that $j(k) \geq N$ implies

$$M(p) < \frac{1}{3m(p)} + \epsilon. \tag{4.119}$$

LEMMA 11. *Given $\epsilon > 0$, then if $\{j(k)\}$ tends to infinity sufficiently rapidly, for all $n \geq p$ we have*

$$\frac{1}{m(n)} \sum_{i=1}^{m(n)} x_i(n)x_{i-j}(n) < \frac{1}{3m(p)} + \epsilon, \tag{4.120}$$

for all j not divisible by p.

PROOF. The result holds for $n = p$, as we easily see. We suppose that $j(p), \cdots, j(k-1)$ has been chosen so that the result holds for $n = p, \cdots, k-1$, and

$$M(n) < \frac{1}{3m(p)} + \epsilon, \tag{4.121}$$

holds also for $n = p, \cdots, k-1$. It then follows from lemma 9 that

$$\begin{aligned}\sum_{i=1}^{m(k)} x_i(k)x_{i-j}(k) \leq 3K_1 \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-\langle j\rangle}(k-1) \\ + K_2\Bigg[\sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(1)}(k-1) + \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(1)-1}(k-1) \\ + \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(1)-2}(k-1)\Bigg] \\ + K_3\Bigg[\sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(2)}(k-1) + \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(2)-1}(k-1) \\ + \sum_{i=1}^{m(k-1)} x_i(k-1)x_{i-r(2)-2}(k-1)\Bigg].\end{aligned} \tag{4.122}$$

Applying the induction hypothesis and (4.121) to the right side of (4.122) we obtain for $j(k) \geq N(j(p), \cdots, j(k-1))$ that it is less than or equal to

$$\begin{aligned}3K_1 m(k-1)\left(\tfrac{1}{3}m(p) + \epsilon\right) + 3K_2 m(k-1)\left(\tfrac{1}{3}m(p) + \epsilon\right) \\ + 3K_3 m(k-1)\left(\tfrac{1}{3}m(p) + \epsilon\right).\end{aligned} \tag{4.123}$$

If we now divide through by $m(k)$ and apply the first part of lemma 9, we see that if $j(k)$ is chosen sufficiently large, the conclusion of the lemma holds for $n = k$. This concludes the proof of the lemma.

LEMMA 12. *Given $\epsilon > 0$, then if $\{j(k)\}$ tends to infinity sufficiently rapidly, we have*

$$\mu(A \cap \tau^j(A)) < \frac{1}{3m(p)} + \epsilon, \tag{4.124}$$

for all j not divisible by p.

Recall that $G_j(k)$, $1 \leq k \leq m(k)$ is that subset of $A_j(k)$ which is mapped by $\tau^{m(k)-j+1}$ into $A_1(k)$. We then have $G_j(k) = \tau^{j-m(k)}G_{m(k)}(k)$. We also let

$$G^*(k) = \bigcup_{j=1}^{m(k)} G_j(k), \qquad A^*(k) = \bigcup_{j=1}^{m(k)} A_j(k). \tag{4.125}$$

It follows from our preceding remarks that $G^*(k)$, which is obviously contained in $A^*(k)$, equals $A^*(k)$ with error tending to zero as $j(k+1)$ tends to infinity. Next, let $A = A_{m(p)}(p)$ as in lemma 7. Then, as before, we have

$$A = \sum_{i=1}^{m(k)} x_i(k)A_i(k), \tag{4.126}$$

and, for any transformation σ,

$$\sigma(A) = \sum_{i=1}^{m(k)} x_i(k)\sigma(A_i(k)). \tag{4.127}$$

LEMMA 13. *Let σ be an arbitrary measurable transformation which commutes with τ, and let $A = A_{m(p)}(p)$. We then have*

$$\sigma(A \cap G^*(k)) = \sum_{i=1}^{m(k)} x_i(k)\tau^{i-1}(\sigma(G_1(k)) \cap G^*(k)) + E(\delta), \tag{4.128}$$

where δ tends to zero uniformly in σ as $j(k+1)$ tends to infinity (and where the $x_i(k)$ are given as in lemma 7).

LEMMA 14. *If σ is a measurable p-th root of τ, then for each k there exist $\{H_i(k),\ 1 \le i \le m(k)\}$ pairwise disjoint subsets of $G_1(k)$, such that*

$$\begin{aligned} G^*(k) \cap \sigma(G_1(k)) &= \sum_{i=1}^{m(k)} \tau^{i-1}H_i(k), \\ H_i(k) &= \tau^{1-i}(G_i(k) \cap \sigma(G_1(k))). \end{aligned} \tag{4.129}$$

LEMMA 15. *If σ is a measurable p-th root of τ, then there exist $\{J_i(k),\ 1 \le i \le m(k)\}$ pairwise disjoint subsets of $A_1(k)$ such that*

$$\sigma(A \cap G^*(k)) = \sum_{j=1}^{m(k)} \sum_{i=1}^{m(k)} x_i(k)\tau^{(i+j-2) \bmod m(k)}J_j(k) + E(\delta), \tag{4.130}$$

where δ tends to zero uniformly in σ as $j(k+1)$ tends to infinity.

LEMMA 16. *Let $\alpha_i = (X_{i1}, \cdots, x_{in})$, $i = 1, \cdots, k$ be k sequences of zeros and ones of length n. Let $\{b_i, i = 1, \cdots, k\}$ be nonnegative numbers whose sum equals one, and suppose that there is a subset G of $\{1, \cdots, n\}$ such that*

$$\sum_{i=1}^{k} b_i x_{ij} \ge 1 - \eta, \qquad \text{for all } j \text{ in } G. \tag{4.131}$$

Then there exists an integer w, $1 \le w \le k$, such that

$$\sum_{j \text{ in } G} \sum_{i=1}^{k} b_i x_{ij} x_{wj} \ge n(G)(1 - 2\eta). \tag{4.132}$$

LEMMA 17. *If $\{j(k)\}$ tends to infinity sufficiently rapidly and if $\tau(j(k))$ has a nonsingular p-th root σ, then there exists an integer $N(\eta)$ such that*

$$\mu(\sigma(A) - \tau^N(A)) \le \eta \tag{4.133}$$

for each $\eta > 0$.

Proof of the theorem in the general case. To see that the theorem holds, choose a sequence $\{j(k)\}$ tending to infinity sufficiently rapidly for lemmas 5, 6, 12, and 17 to hold. We then have a transformation having continuous spectrum which is not strongly mixing and which satisfies the condition that

(i) $\mu(A \cap \tau^j A) < \dfrac{1}{3m(p)} + \epsilon$

for all odd j not divisible by p, for $A = A_{m(p)}(p)$. Furthermore, if the transformation has a nonsingular p-th root, then for each $\eta > 0$ there is an $N = N(\eta)$ such that

(ii) $\mu(\sigma(A) - \tau^N(A)) \le \eta$.

Part (ii) means that $\sigma(A) \subset \tau^N(A) + F$, where the measure of F is less than or equal to η, and thus

$$\sigma^p(A) \subset \tau^{pN}(A) + \sigma^{p-1}(F) + \tau^N(\sigma^{p-2}(F)) + \cdots + \tau^{(p-1)N}(F). \tag{4.134}$$

Since σ is nonsingular and τ is measure preserving, this implies that

$$A = \tau^{pN-1}(A) + E(\eta) + E(\delta(\eta)), \tag{4.135}$$

where, if ϵ is sufficiently small and η is sufficiently small, $\epsilon(\eta)$ tends to zero as η tends to zero, contradicting (i).

REFERENCES

[1] J. R. Blum and N. Friedman, "On roots of transformations," to appear.

[2] R. V. Chacon, "Weakly mixing transformations which are not strongly mixing," to appear.

[3] ———, "Transformations having continuous spectrum," to appear.

[4] ———, "A class of linear transformations," *Proc. Amer. Math. Soc.*, Vol. 15 (1964), pp. 560–564.

[5] R. V. Chacon and N. Friedman, "Approximation and invariant measures," *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, Vol. 3 (1965), pp. 286–295.

[6] Y. N. Dowker, "On measurable transformations in finite measure spaces," *Ann. of Math.*, Vol. 62 (1955), pp. 504–516.

[7] Y. N. Dowker and P. Erdős, "Some examples in ergodic theory," *Proc. London Math. Soc.*, Vol. 39 (1959), pp. 227–241.

[8] P. R. Halmos, "Square roots of measure preserving transformations," *J. Math.*, Vol. 64 (1942), pp. 153–166.

[9] ———, "In general a measure preserving transformation is mixing," *Ann. of Math.*, 45 (1944), pp. 786–792.

[10] ———, *Lectures on Ergodic Theory*, Mathematical Society of Japan, Publication 3, 1956.

[11] S. Kakutani, "Induced measure preserving transformations," *Proc. Imp. Acad. Tokyo*, Vol. 19 (1943), pp. 635–641.

[12] S. Kakutani and J. von Neumann, unpublished manuscript.

[13] D. S. Ornstein, "On invariant measures," *Bull. Amer. Math. Soc.*, Vol. 66 (1960), pp. 197–200.

[14] A. Ionescu Tulcea, "On the category of certain classes of transformations in ergodic theory," *Trans. Amer. Math. Soc.*, Vol. 114-1 (1965), pp. 261–279.