ON THE ROOT PROBLEM IN ERGODIC THEORY

DONALD S. ORNSTEIN
STANFORD UNIVERSITY

1. Introduction

One of the unsolved problems in ergodic theory is the following. Let T be an invertible measure preserving transformation on the unit interval. When does T have a square root? When can T be imbedded in a flow? In his book on ergodic theory, Halmos asked, (1) if every weakly mixing transformation had a square root, (2) if every Bernoulli shift had a square root, and (3) if every Bernoulli shift could be imbedded in a flow. Chacon [1] showed that the answer to (1) was negative. We showed [5], [6] that the answer to (2) and (3) was yes. These results seem to indicate that "enough mixing" forces T to have a square root or to be imbeddable in a flow.

It is the purpose of this paper to give an example of a mixing transformation that has no square root. (T is mixing if and only if

$$\lim_{n\to\infty} m(T^n A \cap B) = m(A)m(B) \tag{1.1}$$

where $m(A)$ denotes the measure of the set A.) The transformation T that we will construct will not only lack a square root but will have the property that if S is a measure preserving transformation of the unit interval such that $ST = TS$, then $S = T^i$ for some integer i (possibly negative or 0).

It is still not known if every K automorphism has a square root. (K automorphisms have a stronger mixing property than "mixing." It was once conjectured that all K automorphisms were Bernoulli shifts but this is now known to be false [7].)

Before starting the construction of our example we shall prove the following theorem which we believe is of independent interest.

THEOREM 1.1. *If T is a measure preserving invertible transformation of* (0, 1) *such that every power of T is ergodic, and if T has the property that there is a constant K, $K > 1$, and* $\lim\sup_{n\to\infty} m(T^n A \cap B) < Km(A)m(B)$ *for all measurable sets A and B, then T is mixing.*

We could construct our example without the help of the above theorem but only at the cost of considerable additional complication.

The main motivation for proving Theorem 1.1, however, comes from the following conjecture of Kakutani. If there is a constant K, with $0 < K < 1$, such

This research was supported in part by a grant from the National Science Foundation under GP-8781.

that $\liminf_{n\to\infty} m(T^nA \cap B) \geqq Km(A)m(B)$, then T is mixing. (Clearly, all powers of T must be ergodic.) This conjecture was shown to be false [3].

Theorem 1.1 falls into a general group of theorems to the effect that if a transformation comes at all close to having a certain property, then it must have that property. Here are some examples. (1) The Birkhoff ergodic theorem implies that if for all A, B

$$\sup_n \frac{1}{n+1} \sum_0^n m(T^iA \cap B) > 0, \tag{1.2}$$

then

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_0^n m(T^iA \cap B) = m(A)m(B). \tag{1.3}$$

(2) England and Martin show that Von Neumann's mixing theorem implies the following. If for any two sets A and B there is a sequence of integers n_i of density 1 (depending on A and B) such that

$$\liminf_{i\to\infty} m(T^{n_i}A \cap B) > 0, \tag{1.4}$$

then for any two sets A and B there is a sequence m_j of density 1 (depending on A and B) such that

$$\lim_{j\to\infty} m(T^{m_j}A \cap B) = m(A)m(B). \tag{1.5}$$

2. A general theorem

THEOREM 2.1. *Let T be a 1–1, invertible measure preserving transformation of* (0, 1) *onto itself, and let m be Lebesgue measure. If* (*a*) *every power of T is ergodic, and* (*b*) *there is a K such that* $\limsup_{n\to\infty} m(T^nA \cap B) < Km(A)m(B)$ *for all measurable sets A and B, then T is mixing.*

LEMMA 2.1. *If T satisfies the hypothesis of Theorem* 2.1, *then T is weakly mixing.*

PROOF. Let $\tilde{T}$ be the unitary operator on L_2 given by $\tilde{T}f(x) = f[T(x)]$. If $\tilde{T}$ were not weakly mixing, then (by Von Neumann's theorem, see [4]) there would be a complex valued function g and a complex number α such that $\tilde{T}g = \alpha g$. Since $\tilde{T}$ is unitary, $|\alpha| = 1$. Since T is ergodic, $|g|$ is constant a.e. Since every power of T is ergodic, we have (∗) *any set on which g is constant has measure zero.*

Let F be the set of x such that $\Theta_1 \geqq \arg[g(x)] \geqq \Theta_2$. Then, given $\varepsilon > 0$, we can find arbitrarily large n such that T^nF is the set of x, $\Theta_1 + n\alpha \geqq \arg[g(x)] > \Theta_2 + n\alpha$, and $|n\alpha| < \varepsilon$. This and (∗) imply that given $\varepsilon' > 0$, we can find an arbitrary large n such that $|T^nF - F| < \varepsilon'$. Since $m(F)$ can be chosen as small as we want (by properly choosing Θ_1 and Θ_2), we have contradicted condition (b).

PROOF OF THEOREM 2.1. (1) We can pick a sequence of integers n_i such that if C and D are intervals with rational end points, then $\lim_{i\to\infty} m[(T^{n_i}C) \cap D]$ exists. This follows from a standard diagonal procedure since there are only a countable number of such C, D.

(2) If T is not mixing, then n_i can be chosen so that, in addition to satisfying (1), there is one pair of intervals with rational end points C_1, D_1 and

$$\lim_{i\to\infty} m[(T^{n_i}C_1) \cap D_1] \neq m(C_1)m(D_1). \tag{2.1}$$

(3) There is a measure u on $(0, 1) \times (0, 1)$ such that u is absolutely continuous with respect to Lebesgue measure on $(0, 1) \times (0, 1)$, and if C and D are intervals with rational end points, then

$$u(C \times D) = \lim_{i\to\infty} m[(T^{n_i}C) \cap D]. \tag{2.2}$$

This requires some proof. Order the pairs C, D and let F_n be the field of sets in $(0, 1) \times (0, 1)$ generated by the first n, $C \times D$. Define f_n to be the function on $(0, 1) \times (0, 1)$ that is constant on each atom of F_n and such that if $C \times D$ is in F_n, then

$$\iint_{C\times D} f_n \, dmdm = \lim_{i\to\infty} m[(T^{n_i}C) \cap D]. \tag{2.3}$$

Also, $0 \leqq f_n \leqq K$ by (b). The f_n form a Martingale, and hence $f_n \to f$ a.e. and f will be the derivative of u. (It is not really necessary to use the Martingale convergence theorem here. We could have defined u on the algebra generated by the $C \times D$ with rational end points, used (b) to show that u is countably additive when restricted to this algebra, and hence, by a theorem in [8], shown that u can be extended to a measure on $(0, 1) \times (0, 1)$.)

(4) If A and B are any measurable sets in $(0, 1)$, then

$$u(A \times B) = \lim_{i\to\infty} m[(T^{n_i}A) \cap B]. \tag{2.4}$$

This holds if A and B are each the union of a finite number of intervals with rational end points. Let A_n and B_n be a sequence of such sets approaching A and B, respectively. Then

$$\begin{aligned} u(A \times B) &= \lim_{n\to\infty} u(A_n \times B_n), \\ u(A_n \times B_n) &= \lim_{i\to\infty} m[(T^{n_i}A_n) \cap B_n]. \end{aligned} \tag{2.5}$$

However, condition (b) implies that

$$\begin{aligned} &\limsup_{i\to\infty} m[(T^{n_i}A_n) \cap B_n] - m[(T^{n_i}A) \cap B] \\ &\qquad < K(m|A_n - A|m(B) + m|B_n - B|m(A) + m|B_n - B|m|A_n - A|) \end{aligned} \tag{2.6}$$

where $|A_n - A|$ denotes the symmetric difference of A_n and A.

(5) The transformation T induces another transformation $\bar{T}$ on $(0, 1) \times (0, 1)$ by the relation $\bar{T}(x, y) = (Tx, Ty)$. The measure u is invariant under $\bar{T}$. To see this, we need only check it on sets of the form $A \times B$. This gives

$$(2.7)\qquad \begin{aligned} u(A \times B) &= \lim_{i\to\infty} m[(T^{n_i}A) \cap B], \\ u(TA \times TB) &= \lim_{i\to\infty} m[(T^{n_i}TA) \cap TB] = \lim_{i\to\infty} m[T([T^{n_i}A] \cap B)]. \end{aligned}$$

(6) Since T is weakly mixing, $\bar{T}$ is ergodic (with respect to Lebesgue measure on $(0, 1) \times (0, 1)$). Since u is invariant under $\bar{T}$ and since u is absolutely continuous with respect to Lebesgue measure, u must be a multiple of Lebesgue measure. This multiple is 1 since $u[(0, 1) \times (0, 1)] = 1$. This gives a contradiction since $u(C_1 \times D_1) \neq m(C_1)m(D_1)$, using (2).

3. Construction of a special mixing transformation

Definition of Class 1. This is simply an explicit construction for a transformation that can be approximated by periodic transformations. The first step in our construction will be to describe a class of transformations that we will call Class 1. Each transformation T in Class 1 will be defined on $(0, 1)$ and will be obtained as a limit of a sequence of transformations T_n on subsets X_n of $(0, 1)$. The transformations T_n will have the following form: X_n will be the union of $h(n)$ disjoint intervals ${}_nJ_i$, with $i = 1, \cdots, h(n)$; all of the ${}_nJ_i$ will have the same length; and T_n will map ${}_nJ_i$, with $1 \leqq i < h(n)$, linearly onto ${}_nJ_{i+1}$. T_n will not be defined on ${}_nJ_{h(n)}$.

The X_{n+1} and T_{n+1} will be obtained from X_n and T_n as follows. Divide ${}_nJ_1$ into $p(n)$ disjoint intervals ${}_nJ_1^i$, $i = 1, \cdots, p(n)$, of the same length. Let ${}_nJ_\ell^i = T_n^{\ell-1}({}_nJ_1^i)$, with $\ell = 1, \cdots, h(n)$. For each i, with $1 \leqq i < p(n)$, we will pick an integer $a(i, n)$ with the property that $a(i, n) < h(n - 1)$. For each $a(i, n)$ we will pick $a(i, n) + h(n - 1)$ disjoint intervals of the same length as ${}_nJ_1^i$ in $(0, 1) - \bigcup_{i=1}^n X_i$. Call these intervals ${}_nJ_{h(n)+1}^i, \cdots, {}_nJ_{h(n)+a(i,n)+h(n-1)}^i$. These intervals will be added to X_n to get X_{n+1}. Let T_{n+1} map ${}_nJ_j^i$ linearly onto ${}_nJ_{j+1}^i$ if $j < h(n) + a(i, n) + h(n - 1)$, and let T_{n+1} map ${}_nJ_{h(n)+a(i,n)+h(n-1)}^i$ linearly onto ${}_nJ_1^{i+1}$ for $i < p(n)$. T_{n+1} will not be defined on ${}_nJ_{h(n)+a(i,n)}^{p(n)}$.

Let ${}_nJ_i^j = {}_{n+1}J_\ell$ where

$$(3.1)\qquad \ell = \sum_{k=1}^{j-1} a(k, n) + (j - 1)[h(n) + h(n - 1)] + i.$$

Then

$$(3.2)\qquad h(n + 1) = p(n)[h(n) + h(n + 1)] + \sum_{i=1}^{p(n)} a(i, n),$$

and $T_{n+1}({}_{n+1}J_i) = {}_{n+1}J_{i+1}$ for $i < h(n + 1)$. The transformation T_{n+1} is not defined on ${}_{n+1}J_{h(n+1)}$.

THEOREM 3.1. *Class* 1 *contains a mixing transformation.*

The rest of this section will be devoted to a proof of this theorem. Lemma 2.1 will be the crucial lemma. Here we construct a sequence with certain properties. It is hard to do this explicitly but we will show by a probability argument that most sequences have these properties.

We will start with a lemma which is a form of the law of large numbers. It is well known (and easy), and we include it only for the sake of completeness.

LEMMA 3.1. *Fix α, with $0 < \alpha < 1$, and give each sequence of zeros and ones of length n the measure $\alpha^m(1-\alpha)^{n-m}$ where m is the number of zeros in the sequence. This puts a probability measure on the space of such sequences. Let $P_n(\varepsilon)$ be the measure of the set of sequences of length n having more than $(\alpha+\varepsilon)n$ zeros. Then for fixed $\varepsilon > 0$, $\limsup_{n\to\infty} [P_n(\varepsilon)]^{1/n} < 1$.*

PROOF. Let r_m be the probability that there are exactly m zeros.

$$r_m = \frac{n(n-1)\cdots(n-m+1)}{m!}[\alpha^m(1-\alpha)^{n-m}]. \tag{3.3}$$

If $m > r\alpha n$, $r > 1$, then

$$\frac{r_{m+1}}{r_m} \leqq \frac{1-r\alpha}{r\alpha}\left(\frac{\alpha}{1-\alpha}\right) < \frac{1}{r}. \tag{3.4}$$

This follows from (3.3).

If $\beta_1 > \beta_2 > r\alpha$, then

$$r_{[\beta_1 n]} \leqq \left(\frac{1}{r}\right)^{(\beta_1-\beta_2)n} r_{[\beta_2 n]} \leqq \left[\left(\frac{1}{r}\right)^{(\beta_1-\beta_2)}\right]^n. \tag{3.5}$$

This follows from (3.4). Also

$$\sum_{[\beta_1 n]}^{\infty} r_m \leqq \frac{1}{1-1/r} r_{[\beta_1 n]}. \tag{3.6}$$

This follows from (3.4).

Lemma 3.1 follows from (3.5) and (3.6).

LEMMA 3.2. *Given $\varepsilon > 0$ and even positive integers N and K and $\alpha > 1$, we can find an $m > N$ and a sequence $\{a_i\}$, with $i = 1, \cdots, m$, of integers such that*

(i) $|\Sigma_j^{j+k} a_i| \leqq K$ *for all* $1 \leqq j \leqq j+k \leqq m$.

(ii) *Let $H(\ell, k)$ be the number of j such that $\Sigma_j^{j+k} a_i = \ell$ where $1 \leqq j \leqq j+k \leqq m$. If $k < (1-\varepsilon)m$, then $H(\ell, k) < \alpha(K)^{-1}(m-k)$.*

PROOF. (1) Let Y be the space of all sequences of integers $\{s_i(w)\}$, with $i = 1, \cdots, m$; $w \in Y$, where $|s_i(w)| \leqq \frac{1}{2}K$. We can put a probability measure on Y giving all sequences equal probability.

(2) For fixed i, k, and ℓ,

$$\text{Prob}\,\{s_{i+k}(w) - s_i(w) = \ell\} \leqq \frac{1}{K}. \tag{3.7}$$

(3) Now consider a fixed $k \leqq (1-\varepsilon)m$. We can divide the pairs of integers $(i, i+k)$ into two disjoint sets $E(k)$ and $E^1(k)$ in such a way that no integer occurs in more than one pair in $E(k)$, and the same is true of $E^1(k)$. Furthermore, $|E(k)| > \frac{1}{4}\varepsilon m$ and $|E^1(k)| > \frac{1}{4}\varepsilon m$ where $|E(k)|$ denotes the number of elements in $E(k)$. (This is easy to see, and we omit the details.)

(4) There exists a γ, with $0 \leqq \gamma < 1$, such that for all m large enough, all ℓ, and all $k < (1 - \varepsilon)m$, we have

$$\text{Prob}\,\{(\text{number of } i \text{ such that } (i, i + k) \in E(k) \text{ and } s_{i+k}(w) - s_i(w) = \ell) > \alpha(K)^{-1}|E(k)|\} < \gamma^m. \tag{3.8}$$

This follows because our choice of $E(k)$ which makes all the $s_{i+k}(w) - s_i(w)$, where $(i, i + k) \in E(k)$, independent and makes $|E(k)| > \varepsilon m$, allows us to apply (2) and Lemma 3.1.

(5) Let $a_i(w) = s_i(w) - s_{i-1}(w)$, $\big(a_1(w) = s_1(w)\big)$. Then $s_{i+k}(w) - s_i(w) = \Sigma_{i+1}^{i+k} a_j(w)$. If m is large enough, then (4) tells us that the probability that $\{a_i(w)\}$ does not satisfy (ii) is $< 2mK\gamma^m$. (To see (5), we note that (ii) is automatically true for $|\ell| > K$; therefore, if (ii) fails, (4) fails for one of the $(1 - \varepsilon)mK$ pairs k, ℓ and $E(k)$ or $E^1(k)$.) Lemma 3.2 follows immediately from (5) since $\lim_{n\to\infty} 2mK\gamma^m = 0$.

Construction of T. The construction will be determined if we decide at the nth stage what $p(n)$ and the $a(i, n)$ should be. Choose $N > 10^n$, $K = h(n - 1)$, $\varepsilon = 10^{-n-3}$, and $\alpha = 5/4$. Apply Lemma 3.2 to get $m > N$ and a sequence a_i, with $i = 1, \cdots, m$. Let $p(n) = m$ and let $a_i = a(i, n)$, $i < m$. The integer $a[p(n), n]$ will be a positive integer < 5 and will be determined later.

LEMMA 3.3. *Pick an integer M and intervals ${}_nJ_\alpha$ and ${}_{n-1}J_\beta$. We will make the following assumptions.* (i) $h(n - 1) < \alpha < h(n) - h(n - 1)$. (ii) *Define r and k as follows: let $M = k[h(n) + h(n - 1)] + r$ where k and r are nonnegative integers and $r < [h(n) + h(n - 1)]$. Then $k < (1 - 10^{-n})p(n)$.* (iii) *Either* (a) *holds, or* (b) *and* (c) *hold, where*

(a) $\alpha + r > h(n) + h(n - 1)$,

(b) $\alpha + r < h(n) - h(n - 1)$, *and*

(c) $k \geqq 1$.

Let H be the part of ${}_nJ_\alpha$ on which T^M_{n+1} is defined. Then

$$m[(T^M_{n+1}H) \cap {}_{n-1}J_\beta] \leqq \tfrac{3}{2}m(H)m({}_{n-1}J_\beta). \tag{3.9}$$

PROOF. (1) If ${}_nJ^i_\alpha \subset H$, then $T^M_{n+1}({}_nJ^i_\alpha) = {}_nJ^{s(i,\alpha)}_{t(i,\alpha)}$. If condition (a) holds, then $s(i, \alpha) = i + k + 1$ and $t(i, \alpha) = \alpha + r + \Sigma_{\ell=i}^{i+k} a(\ell, n)$. If conditions (b) and (c) hold, then $s(i, \alpha) = i + k$ and $t(i, \alpha) = \alpha + r + \Sigma_{\ell=i}^{i+k-1} a(\ell, n)$. These formulas follow from the definition of transformations in Class 1.

(2) ${}_nJ^i_\alpha \subset H$ if $i < p(n) - k - 1$ and ${}_nJ^i_\alpha \not\subset H$ if $i > p(n) - k$. (One i is not accounted for; to account for it, one would have to see which condition (a) or (b) holds. It will not, however, be necessary to bother with this.)

(3) There are at most two numbers γ_1 and γ_2 such that

$$\alpha + r - h(n - 1) \leqq \gamma_i \leqq \alpha + r + h(n - 1), \qquad i = 1, 2 \tag{3.10}$$

and ${}_nJ_{\gamma_i} \in {}_{n-1}J_\beta$, with $i = 1, 2$. This follows easily from the description of the construction of transformations in Class 1.

(4) Let ξ be the number of integers i such that $1 \leqq i \leqq p(n) - k$ and $t(i, \alpha) = \gamma_1$ or $t(i, \alpha) = \gamma_2$. Using the formula for $t(i, \alpha)$ in (1) and Lemma 3.2 (i), we see that $\alpha + r - h(n-1) \leqq t(i, \alpha) \leqq \alpha + r + h(n-1)$. Applying (3), hypothesis (ii) and Lemma 3.2 (ii), we get

$$\frac{\xi}{p(n) - k} \leqq \tfrac{10}{4}[h(n-1)]^{-1}. \tag{3.11}$$

Rewriting (4) we get

$$\frac{m[T^M_{n+1}(H) \cap {}_{n-1}J_\beta]}{m(H)} < \tfrac{10}{4}[h(n-1)]^{-1} < 3m({}_{n-1}J_\beta). \tag{3.12}$$

Lemma 3.4. *Let M be an integer such that $h(n) + h(n-1) \leqq M < h(n+1) + h(n)$. Then the following sets D_i all have measure $< 10^{-n+2}$.*

D_1 is the union of the ${}_nJ_\alpha$ which fail to satisfy either (i) *and* (a) *or* (i) *and* (*b*) *of Lemma* 3.3.

D_2 is the union of the ${}_{n+1}J_\alpha$ which fail to satisfy either (i) *and* (a) *or* (i) *and* (b) *of Lemma* 3.3. *By* (i) *we mean $h(n) < \alpha < h(n+1) - h(n)$. Let $M = k[h(n+1) + h(n)] + r'$; then $k' = 0$, $r' = M$,* (a) *says $\alpha + M > h(n+1) + h(n)$, and* (b) *says $\alpha + M < h(n+1) - h(n)$.*

D_3 is the part of X_{n+2} on which T^M_{n+2} is not defined.

Proof. We have constructed T in such a way that $h(n+1) > 10^n h(n)$. It follows immediately from this that the measures of D_1 and D_2 are $< 10^{-n+2}$.

The transformation T^M_{n+2} is defined on ${}_{n+2}J_i$ if $i \leqq h(n+2) - M$. Since $M < 2h(n+1)$, this shows that $m(D_3) < 10^{-n+2}$.

Lemma 3.5. *Let A be a measurable set in X. Then we can find a sequence of sets A_n such that each A_n is a union of some of the ${}_nJ_i$, and $|A_n - A|$ tends to* 0 *as n tends to ∞.*

The proof is standard and will be omitted.

Lemma 3.6. *Let M be an integer such that $h(n) + h(n-1) \leqq M < h(n+1) + h(n)$. Let A and B be measurable sets in X, and define A_n and B_n as in Lemma* 3.5. *Then $m[(T^M A_n) \cap B_{n-1}] \leqq 4m(A_n)m(B_{n-1})$.*

Proof. Let H be the union of the ${}_nJ_i$ in A_n which are not in D_1. Note that H can be regarded as a union of ${}_{n+1}J_i$. Let H_1 be the union of the ${}_{n+1}J_i$ in H on which T^M_{n+1} is defined. If $m(H_1) > 10^{-n+1}$, then condition (ii) of Lemma 3.3 is also satisfied. Also (c) holds because $h(n) + h(n-1) \leqq M$. We can then apply Lemma 3.3 to get

$$m[(T^M_{n+1}H_1) \cap B'_{n-1}] \leqq 3m(H_1)m(B'_{n-1}). \tag{3.13}$$

Let H'_2 be the part of $H - H_1$ that is not in D_2. Since T^M_{n+1} is not defined on H'_2, we have that for each ${}_{n+1}J_\alpha \subset H'_2$, $\alpha + M > h(n+1)$. If $M = k'[h(n+1) + h(n)] + r'$, then $k' = 0$ and $r' = M$. Since either (b) or (a) holds, (a) must hold and we do not have to worry about (c). Condition (ii) holds trivially. We can therefore apply Lemma 3.3 again to get the following. Let H_2 be the part of H_2' on which T_{n+2}^M is defined; then

$$m[(T_{n+2}^M H_2) \cap B_{n-1}''] \leqq 3m(H_2)m(B_{n-1}''). \tag{3.14}$$

(3.13) and (3.14) imply Lemma 3.6.

PROOF OF THEOREM 3.1. Lemma 3.5 says that, given A, B and ε, there is an N such that if $n > N$, then $|A_n - A| < \varepsilon$ and $|B_n - B| < \varepsilon$. If $M > h(N)$, then there is an $n' \geqq N$ such that $h(n') + h(n' - 1) \leqq M < h(n' + 1) + h(n')$. Lemma 3.6 says that $m[(T^M A_{n'}) \cap B_{n'-1}] \leqq 4m(A_{n'})m(B_{n'-1})$. We then get

(1) $\limsup_{M\to\infty} m[T^M A) \cap B] \leqq 4m(A)m(B)$. If we show that every power of T is ergodic, we can then apply Theorem 1.1 to show that T is mixing.

(2) T is ergodic. If this were not true, there would be an invariant set A, $m(A) \leqq 1/2$, and an n and i such that $m({}_nJ_i \cap A) > (9/10)m({}_nJ_i)$. This implies that $m({}_nJ_j \cap A) > (9/10)m({}_nJ_j)$ for $1 \leqq j \leqq h(n)$, contradicting $m(A) \leqq 1/2$.

(3) Every power of T is ergodic. If this were false, we could use (1) to show that there is a minimal set, A, invariant under T^i. We would then get an α such that T^jA are disjoint for $a \leqq j < \alpha$ and $A = T^\alpha A$. Note that α must be $\leqq 4$ by (1) and $\cup\, T^jA = X$ by (2). By the argument in (2), for n large enough, more than 9/10 of each ${}_nJ_i$ is in one of the T^jA, with $0 \leqq j < \alpha$, and if $i = k \pmod{\alpha}$, then 9/10 of both ${}_nJ_i$ and ${}_nJ_k$ are in the same T^jA. Since this is true for n and $n + 1$, we get that more than half of the $h(n - 1) + a(i, n)$, with $1 \leqq i \leqq p(n)$, must be congruent to 0 mod α. Furthermore, the above statement must hold for all n large enough. By changing $a[p(n - 1), n - 1]$, we change $h(n - 1)$ and we can thus insure that for each integer $\alpha \leqq 4$, there are infinitely many n such that more than half of the $h(n - 1) + a(i, n)$ are not congruent to 0 mod α.

4. A proof that nothing commutes with the transformation constructed in the previous section

THEOREM 4.1. *Let T be a mixing transformation in Class* 1. *Let S be a measure preserving transformation of* (0, 1) *onto itself such that $ST(x) = TS(x)$ for a.e. x in* (0, 1). *Then there is an integer i (possibly negative or* 0) *such that $S(x) = T^i(x)$ a.e.* (All we use is that T is approximable by periodic transformations.)

NOTATION. Pick K disjoint sets $A, B \cdots$ in (0, 1) each of measure $(K)^{-1}$. Let $A' = S^{-1}A, \cdots$. Let A_n and A_n' be sequences of sets where A_n, (A_n') is a union of ${}_nJ_i$ and $A_n \to A$ $(A_n' \to A')$ as in Lemma 3.5. By $T_n^\ell(E)$ we will mean T_n^ℓ applied to the part of E on which T_n^ℓ is defined.

REMARK. It is no loss of generality to assume that $TS(x) = ST(x)$ for all x.

LEMMA 4.1. *Assume K is already given as above. We can then choose N such that if $n > N$ and $|\ell| > N$, then* (i) $|A_n' - S^{-1}A_n| < 10^{-10}K^{-1}$; (ii) $m(T_n^\ell A_n') \cap A_n) < 2K^{-2}$. *Furthermore,* (i) *and* (ii) *still hold if we substitute any of the other K sets for A. (By $|A - B|$ we mean $m(A \cup B - A \cap B)$.)*

The proof is obvious.

DEFINITION. *If $x \in X_n \cap S^{-1}X_n$, define ${}_nf(x)$ to be an integer such that $T^{({}_nf(x))}(x)$ and $S(x)$ are in the same ${}_nJ_i$.*

LEMMA 4.2. *Let x be a point in ${}_nJ_i \cap S^{-1}X_n$. Let j be an integer (not necessarily positive) such that $1 \leqq i + j \leqq k(n)$ and $1 \leqq j + i + {}_nf(x) \leqq h(n)$. Then ${}_nf(x) = {}_nf[T_n^j(x)]$.*

PROOF. This follows immediately from the definition of ${}_nf$ and the fact that $ST = TS$.

LEMMA 4.3. *Let K and N be defined as in Lemma 4.1. Assume that the set where $|{}_nf(x)| < N$ has measure $< \frac{1}{2}m(X_n)$. Then, for all n large enough, we can find an ℓ, $\frac{1}{4}h(n) \leqq \ell \leqq \frac{3}{4}h(n)$ such that if we let E be the subset of ${}_nJ_\ell$ on which $|{}_nf(x)| > N$, then $m(E) > 2^{-3}m({}_nJ_\ell)$.*

Proof. (1) As $n \to \infty$, the measure of the set where ${}_nf$ is not defined tends to 0. We can, therefore, find an ℓ, $\frac{1}{4}h(n) \leqq \ell \leqq \frac{3}{4}h(n)$ such that if E' is the part of ${}_nJ_\ell$ on which ${}_nf$ is defined, then $m(E') > (99/100)m({}_nJ_\ell)$.

(2) Let E be the part of E' where $|{}_nf| > N$. Now $|{}_nf| < N$ on $E' - E$, and if $m(E) < 2^{-3}m({}_nJ_\ell)$, then $m(E' - E) > \frac{3}{4}m({}_nJ_\ell)$.

(3) We can now apply Lemma 4.2 and (2) to show that $|{}_nf| < N$ on more than 3/4 of each ${}_nJ_i$ where $N < i < h(n) - N$. If n is large enough, with respect to N, this contradicts the assumption of Lemma 4.3.

The following lemma contains the whole idea of the proof of Theorem 4.1.

LEMMA 4.4. *Let K and N be defined as in Lemma 4.1. Then for n large enough, the set where $|{}_nf(x)| < N$ has measure $> \frac{1}{2}m(X_n)$.*

PROOF. We will start by assuming that Lemma 4.4 is false.

(1) We can therefore apply Lemma 4.3 to get $\frac{1}{4}h(n) \leqq \ell \leqq \frac{3}{4}h(n)$ such that the subset E, of ${}_nJ_\ell$ where $|{}_nf(x)| > N$, has measure $> 2^{-3}m({}_nJ_\ell)$. Let E' be the part of E where ${}_nf(x) \leqq -N$. We can assume $m(E') \geqq \frac{1}{2}m(E) > 2^{-4}m({}_nJ_\ell)$. (The other case, where the part of E in which ${}_nf(x) \geqq N$ has measure $> \frac{1}{2}m(E)$, will follow by exactly the same argument.)

(2) One of the sets, A, B, C, and so forth, (we will call it A for the sake of notation) has the following property. Let ${}_1A'_n$ be the union of those ${}_nJ_i$ such that ${}_nJ_i \subset A'_n$ and $\ell \leqq i \leqq h(n)$. Then $m({}_1A'_n) > 2^{-3}m(A') = 2^{-3}K^{-1}$.

(3) Let ${}_2A'_n$ be the part of ${}_1A'_n$ on which ${}_nf < -N$. Let $E_{i,j}$ be the part of ${}_2A'_n \cap {}_nJ_j$ where ${}_nf = -i$.

(4) $E' = \bigcup_i E_{i,\ell}$ and $E_{i,j} = T_n^{j-\ell}(E_{i,\ell})$. (The first statement is just the definition of E' and the second comes immediately from Lemma 4.2. This holds for those j such that ${}_nJ_j \cap {}_1A'_n \neq 0$.)

(5) $m({}_2A'_n) > 2^{-7}K^{-1}$. This comes from (1), (2) and (4). We are now ready to get (a) the part of ${}_2A'_n$ not in $S^{-1}A_n$ has measure $< 10^{-10}K^{-1} < 10^{-3}m({}_2A'_n)$. (The first inequality comes from Lemma 4.1 (a) and the second from (5).) The rest of this proof will de devoted to proving (b) the part of ${}_2A'_n$ not in $S^{-1}A_n$ has measure $> \frac{1}{2}m({}_2A'_n)$.

(6) If $\alpha > N$, then Lemma 4.1 (ii) implies

$$m({}_1A'_n \cap T_n^\alpha A_n) < aK^{-2} < 2^3[2K^{-1}m({}_1A'_n)]. \tag{4.1}$$

(The last inequality comes from (2).) We will now use (4.1) to get: if $\alpha > N$,

$$m\left(\bigcup_{i=\alpha} E_{i,j} \cap S^{-1}A_n\right) < 2^4K^{-1}m\left(\bigcup_{i=\alpha} E_{i,j}\right). \tag{4.2}$$

To get (4.2) note that (4) implies $m(E_{\alpha,j})/m({}_nJ_j)$ does not depend on j if ${}_nJ_j \subset {}_1A'_n$. We also have $E_{\alpha,j} \subset S^{-1}A_n$ or $E_{\alpha,j} \cap S^{-1}A_n = \varnothing$, and that the former occurs if and only if ${}_nJ_j \subset T_n^{\alpha}A_n$. These two statements combined with (4.1) give (4.2).

If K were chosen so that $2^4K^{-1} < \frac{1}{4}$, (4.2) would imply (ii) by summing over α.

Proof of Theorem 4.1. Lemma 4.4 implies that for all n large enough, the part of X where $|{}_nf(x)| < N$ has measure $> \frac{1}{4}$. This implies that there is a set $F \subset X$, $m(F) > \frac{1}{4}$ such that if $x \in F$, then $|{}_nf(x)| \leqq N$ for infinitely many n. This implies that there is an $|M(x)| \leqq N$ such that ${}_nf(x) = M(x)$ for infinitely many n, and all $x \in F' \subset F$ where $m(F') \neq 0$. We therefore have that $S(x)$ and $T^{M(x)}(x)$ are in the same ${}_nJ_i$ for infinitely many n, for $x \in F'$. But the ${}_nJ_i$ are intervals whose lengths are tending to zero which implies that $S(x) = T^{M(x)}(x)$ for $x \in F'$.

For fixed M, the set of x such that $S(x) = T^M(x)$ is invariant under T. Since T is ergodic, this proves Theorem 4.1.

Remark. It is not necessary to assume that S is measure preserving and onto. It is enough to assume that $S^{-1}A$ is measurable if A is measurable, that there is no set A such that $m(A) = 0$ and $m(S^{-1}A) = 1$, and that the range of S is a measurable set of nonzero measure. It is then easy to see that the assumption that S commutes with an ergodic measure preserving transformation forces S to be measure preserving and onto.

REFERENCES

[1] R. V. Chacon, "A geometric construction of measure preserving transformations," *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1965, Vol. 2, Part 2, pp. 335–360.

[2] N. Friedman, *Notes on Ergodic Theory*, to appear.

[3] N. Friedman and D. Ornstein, "On partially mixing transformations," to appear.

[4] P. R. Halmos, *Lectures on Ergodic Theory*, New York, Chelsea, 1960.

[5] D. S. Ornstein, "Bernoulli shifts with the same entropy are isomorphic," *Advances in Math.*, Vol. 4 (1970), pp. 337–352.

[6] ———, "Imbedding Bernoulli shifts in flows," *Lecture Notes in Mathematics, Contributions to Ergodic Theory and Probability*, Vol. 160, 1970, pp. 178–218.

[7] ———, "A Kolmogorov automorphism that is not a Bernoulli shift," to appear.

[8] H. L. Royden, *Real Analysis*, New York, Macmillan, 1963.