TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 149, June 1970

ON ENTROPY AND GENERATORS OF MEASURE-PRESERVING TRANSFORMATIONS

BY
WOLFGANG KRIEGER

Abstract. Let T be an ergodic measure-preserving transformation of a Lebesgue measure space with entropy $h(T)$. We prove that T has a generator of size k where $e^{h(T)} \leqq k \leqq e^{h(T)}+1$.

1. **Introduction.** In this paper we are concerned with ergodic invertible measure-preserving transformations of a Lebesgue measure space $(E, \mathfrak{B}, p)$. By a partition $\{A_n : n \in \theta\}$ of E we shall mean a finite or countably infinite collection of disjoint sets $A_n \in \mathfrak{B}$ of positive measure such that

$$E = \bigcup_{n \in \theta} A_n.$$

We call a partition $\{A_n : n \in \theta\}$ a generator of an i.m.p.t. T of $(E, \mathfrak{B}, p)$ if $\mathfrak{B}$ is generated by

$$\bigcup_{i=-\infty}^{\infty} \{T^i A_n : n \in \theta\}.$$

For the theory of entropy and generators of i.m.p.t. we refer to [1], [4], [5] and [6]. It was proved by V. A. Rohlin that every aperiodic i.m.p.t. with finite entropy has a generator with finite entropy [6, 10.7]. We shall prove in §2 that every ergodic i.m.p.t. with finite entropy has a finite generator, thereby solving a problem that was posed by V. A. Rohlin [6, p. 30].

Throughout most of this paper we shall be given a finite or countably infinite state space Ω. For finite Ω we shall prove in §3 an approximation theorem for probability measures on $\Omega^{\boldsymbol{Z}}$ that are invariant under the shift S,

$$(Sx)_i = x_{i+1}, \qquad i \in \boldsymbol{Z}, \qquad x = (x_i)_{i=-\infty}^{\infty} \in \Omega^{\boldsymbol{Z}}.$$

This theorem will enable us to derive in §4 from the work of A. H. Zaslavskiĭ [7] a formula for the minimal number of elements that a generator of an ergodic i.m.p.t. can contain. Denote this number by $\Delta(T)$. If the entropy $h(T)$ of T is infinite then $\Delta(T)$ is also infinite, if $h(T) < \infty$, then $\Delta(T) \geqq e^{h(T)}$. Our result is

$$\Delta(T) \leqq e^{h(T)}+1.$$

Received by the editors March 18, 1969.

AMS Subject Classifications. Primary 2870; Secondary 5482.

Key Words and Phrases. Ergodic measure-preserving transformations, entropy, generators of ergodic measure-preserving transformations, shift-invariant measures, coding.

This answers for the ergodic case another question raised by Rohlin [6, p. 30]. In particular it follows that every ergodic i.m.p.t. with entropy zero has a generator with two elements. This was known before in the case of the quasi-discrete spectrum [3, p. 187].

2. **The existence of finite generators.**

(2.1) THEOREM. *Every ergodic i.m.p.t. with finite entropy has a finite generator.*

Proof. 1. Let $\{A_n : n \in \boldsymbol{N}\}$ be a partition of $(E, \mathfrak{B}, p)$ with finite entropy. Then there exists a mapping $n \to K_n \in \boldsymbol{N}$ $(n \in \boldsymbol{N})$ and a 1-1 mapping

$$\varphi: \boldsymbol{N} \to \bigcup_{k=1}^{\infty} \{1, 2, 3\}^k$$

where $\varphi(n) \in \{1, 2, 3\}^{K_n}$, $n \in \boldsymbol{N}$, such that

$$\sum_{n=1}^{\infty} K_n p(A_n) < \infty. \tag{1}$$

For a proof of this let $p(A_n) \geqq p(A_{n+1})$, $n \in \boldsymbol{N}$, and let $l(n)$, $n \in \boldsymbol{N}$, be nonnegative integers such that

$$-\log p(A_n) - 1 < l(n) \leqq -\log p(A_n), \qquad n \in \boldsymbol{N}. \tag{2}$$

Let further

$$n_1 = 1, \qquad n_m = \min \{n > n_{m-1} : l(n) > l(n_{m-1})\}, \qquad m > 1.$$

Then

$$\sum_{m=1}^{\infty} (n_{m+1} - n_m) 3^{-l(n_m)} = \sum_{n=1}^{\infty} 3^{-l(n)},$$

and we see from (2) that

$$\sum_{m=1}^{\infty} (n_{m+1} - n_m) 3^{-l(n_m)} \leqq e.$$

Consequently, for some $m_0 \in \boldsymbol{N}$,

$$n_{m+1} - n_m < 3^{l(n_m)}, \qquad m \geqq m_0. \tag{3}$$

We set $K_n = l(n)$, $n \geqq n_{m_0}$. By (3) it is possible to assign to every $n \geqq n_{m_0}$ an element $\varphi(n)$ of $\{1, 2, 3\}^{K_n}$ such that $n \to \varphi(n)$, $n \geqq n_{m_0}$, is 1-1. The inequality (3) also shows that it is possible to define the $\varphi(n) \in \bigcup_{k=1}^{\infty} \{1, 2, 3\}^k$, $1 \leqq n < n_{m_0}$, in such a way that

$$\varphi: \boldsymbol{N} \to \bigcup_{k=1}^{\infty} \{1, 2, 3\}^k$$

is 1-1. In order to show that (1) holds it suffices to show that

$$\sum_{n=n_{m_0}}^{\infty} K_n p(A_n) < \infty,$$

and this follows from the finiteness of the entropy of $\{A_n : n \in \boldsymbol{N}\}$ and from (2):

$$\sum_{n=n_{m_0}}^{\infty} K_n p(A_n) \leqq \sum_{n=1}^{\infty} l(n) p(A_n) \leqq \sum_{n=1}^{\infty} -p(A_n) \log p(A_n) < \infty.$$

2. Let Ω be a finite set containing more than two elements. Let $\omega \in \Omega$, $C \in \boldsymbol{N}$, and let

$$X = \left\{ x = ((x_{i,j})_{j=1}^{D_i})_{i=-\infty}^{\infty} \in \left(\dot{\bigcup_{k=2}^{\infty}} \Omega^k \right)^{\boldsymbol{Z}} : \right.$$
$$\left. x_{i,j} \neq \omega,\ 1 \leqq j < D_i,\ x_{i,D_i} = \omega,\ -\infty < i < \infty \right\},$$

and

$$X_C = \bigcap_{k=-\infty}^{\infty} \bigcup_{l=1}^{\infty} \left\{ x = ((x_{i,j})_{j=1}^{D_i})_{i=-\infty}^{\infty} \in X : \sum_{m=k}^{k+l} (D_m - C) \leqq 0 \right\}.$$

We are going to construct a 1-1 Borel mapping $U \colon X_C \to (\Omega^C)^{\boldsymbol{Z}}$ that commutes with the shifts.

Let $x = ((x_{i,j})_{j=1}^{D_i})_{i=-\infty}^{\infty} \in X_C$ and let $\Gamma = \{ i \in \boldsymbol{Z} : D_i > C\}$. We define for $i \in \Gamma$, $C < j \leqq D_i$,

$$I(i,j) = \min \left\{ l > i : j - C + \sum_{i < m \leqq l} (D_m - C) \leqq 0 \right\}, \tag{4}$$

$$J(i,j) = j + \sum_{i < m \leqq I(i,j)} (D_m - C). \tag{5}$$

It follows that

$$D_{I(i,j)} < J(i,j) \leqq C, \qquad i \in \Gamma, \quad C < j < D_i.$$

The mapping

$$(i,j) \to (I(i,j), J(i,j)) \qquad (i \in \Gamma,\ C < j \leqq D_i)$$

is 1-1. Indeed, had we $i, i' \in \Gamma$, $C < j \leqq D_i$, $C < j' \leqq D_i'$,

$$(I(i,j), J(i,j)) = (I(i',j'), J(i',j')),$$

and say $i < i'$, then we could infer from (5) that

$$j + \sum_{i < m \leqq I(i,j)} D_m = J(i,j) + (I(i,j) - i)C,$$

$$\sum_{i' \leqq m \leqq I(i,j)} D_m \geqq J(i,j) + (I(i,j) - i')C,$$

and therefore that

$$j + \sum_{i < m < i'} D_m \leqq (i' - i)C,$$

in contradiction to $j > C$ or to (4). We define now

$$Ux = ((y_{i,j})_{j=1}^{C})_{i=-\infty}^{\infty} \in (\Omega^C)^{\boldsymbol{Z}}$$

by setting

$$y_{i,j} = x_{i,j}, \qquad \text{if } i \in \boldsymbol{Z},\quad 1 \leqq j \leqq \min(C, D_i),$$
$$y_{I(i,j),J(i,j)} = x_{i,j}, \qquad \text{if } i \in \Gamma,\quad C < j \leqq D_i,$$

and by setting $y_{i,j}=\alpha$, $\alpha \in \Omega$, $\alpha \neq \omega$, elsewhere. U is Borel and it commutes with the shifts. We prove now that it is 1-1 by showing that the D_i, $i \in \Gamma$, can be computed from Ux.

Denote

$$I_\omega(i) = I(i, D_i),\quad J_\omega(i) = J(i, D_i), \qquad i \in \Gamma,$$

and

$$N_\omega(i) = \sum_{j=1}^{C} \delta_{\omega, y_{i,j}}, \qquad i \in \boldsymbol{Z}.$$

We have for $i, i' \in \Gamma$

$$i < i' < I_\omega(i) \Rightarrow I_\omega(i') \leqq I_\omega(i), \tag{6}$$

$$i < I_\omega(i') < I_\omega(i) \Rightarrow i < i', \tag{7}$$

$$i < i',\quad I_\omega(i) = I_\omega(i') \Rightarrow J_\omega(i) > J_\omega(i'). \tag{8}$$

From these relations and since $i \to (I_\omega(i), J_\omega(i))$ $(i \in \Gamma)$ is 1-1 we have

$$\begin{aligned} &\sum_{i<m<I_\omega(i)} N_\omega(m) + \sum_{j=1}^{J_\omega(i)} \delta_{\omega, y_{I_\omega(i),j}} \\ &\quad = |\{i' \in \boldsymbol{Z}-\Gamma : i < i' \leqq I_\omega(i)\}| + |\{i' \in \Gamma : i < I_\omega(i') < I_\omega(i)\}| \\ &\qquad + |\{i' \in \Gamma : I_\omega(i') = I_\omega(i), J_\omega(i') < J_\omega(i)\}| + 1 \\ &\quad = |\{i' \in \boldsymbol{Z}-\Gamma : i < i' \leqq I_\omega(i)\}| + |\{i' \in \Gamma : i < i' < I_\omega(i)\}| + 1 \\ &\quad = I_\omega(i) - i + 1, \qquad i \in \Gamma. \end{aligned} \tag{9}$$

And we have from (7)

$$\begin{aligned} \sum_{i<m\leqq L} N_\omega(m) &= |\{i' \in \boldsymbol{Z}-\Gamma : i < i' \leqq L\}| + |\{i' \in \Gamma : i < I_\omega(i') \leqq L\}| \\ &= |\{i' \in \boldsymbol{Z}-\Gamma : i < i' \leqq L\}| + |\{i' \in \Gamma : i < i' < L\}| \\ &\leqq L - i, \qquad i \in \Gamma,\quad i < L < I_\omega(i). \end{aligned} \tag{10}$$

Now we see from (9) and (10) that

$$I_\omega(i) = \min\left\{L > i : \sum_{i<m\leqq L} N_\omega(m) > L - i\right\}, \qquad i \in \Gamma \tag{11}$$

and that

$$\begin{aligned} J_\omega(i) = \min\Bigg\{1 < l \leqq C : &\sum_{i<m<I_\omega(i)} N_\omega(m) \\ &+ \sum_{j=1}^{l} \delta_{\omega, y_{I_\omega(i),j}} = I_\omega(i) - i + 1\Bigg\}, \qquad i \in \Gamma. \end{aligned} \tag{12}$$

Next we observe that

$$(13) \qquad D_i = J_\omega(i)+(I_\omega(i)-i)C-\sum_{i<m\leqq I_\omega(i)} D_m, \qquad i\in\Gamma.$$

We know from (6) that

$$(14) \qquad i < i' < I_\omega(i) \Rightarrow I_\omega(i')-i' < I_\omega(i)-i, \qquad i, i'\in\Gamma.$$

It follows that if $i_0\in\Gamma$ is such that

$$I_\omega(i_0)-i_0 = \min\{I_\omega(i)-i : i\in\Gamma\}$$

then $i_0<i\leqq I_\omega(i_0) \Rightarrow i\in \boldsymbol{Z}-\Gamma$ and we see from (13) that D_{i_0} can be computed from the $y_{i,j}$, $1\leqq j\leqq C$, $i\in\boldsymbol{Z}$. Finally (14) implies also that (13) can be used as a recursion formula to compute all the D_i, $i\in\Gamma$, from Ux.

3. By Rohlin's result [6, 10.7] every ergodic i.m.p.t. is isomorphic to the shift on $\boldsymbol{N}^{\boldsymbol{Z}}$ together with an invariant probability measure μ such that the partition

$$\{(n_i)_{i=-\infty}^{\infty}\in \boldsymbol{N}^{\boldsymbol{Z}} : n_0 = m\}, \qquad m\in\boldsymbol{N},$$

has finite entropy. By part 1 of the proof there is a $C\in\boldsymbol{N}$ and a 1-1 mapping

$$n\to(x_{n,1},\ldots,x_{n,K_n})\in\bigcup_{k=1}^{\infty}\{1,2,3\}^k \qquad (n\in\boldsymbol{N})$$

such that

$$(15) \qquad \sum_{m=1}^{\infty} K_m\mu(\{(n_i)_{i=-\infty}^{\infty}\in\boldsymbol{N}^{\boldsymbol{Z}} : n_0 = m\}) < C-1.$$

We use this mapping to build a 1-1 mapping

$$V : (n_i)_{i=-\infty}^{\infty}\to((x_{n_i,1},\ldots,x_{n_i,K_{n_i}},\omega))_{i=-\infty}^{\infty}\in X \qquad ((n_i)_{i=-\infty}^{\infty}\in\boldsymbol{N}^{\boldsymbol{Z}})$$

that commutes with the shifts, where we can set $\Omega=\{1,2,3,\omega\}$. The individual ergodic theorem and (15) yield

$$\lim_{L\to\infty}\frac{1}{L}\sum_{i=1}^{L}(K_{n_i}+1-C) < 0, \qquad \text{for } \mu\text{-a.a. } (n_i)_{i=-\infty}^{\infty}\in\boldsymbol{N}^{\boldsymbol{Z}}.$$

Hence $\mu(V^{-1}X_C)=1$.

By part 2 of the proof there is a 1-1 Borel mapping

$$U: X_C\to\Omega^{\boldsymbol{Z}}$$

that commutes with the shifts. If we set for a Borel set $F\subset\Omega^{\boldsymbol{Z}}$,

$$\nu(F) = \mu(V^{-1}U^{-1}F),$$

then we find that $(\boldsymbol{N}^{\boldsymbol{Z}},\mu,S)$ is isomorphic to $(\Omega^{\boldsymbol{Z}},\nu,S)$. (If a finite Borel measure on a polish space is transported via a 1-1 Borel mapping to another polish space then the Borel mapping becomes an isomorphism between the measure space given by μ and the measure space that is given by the transported measure. This can be seen from the fact that every analytic subset of a polish space is measurable with respect to every finite Borel measure (see e.g. [2, §6, n°9]). Q.E.D.

3. **An approximation theorem for shift-invariant measures.** Let Ω be a state space containing a finite number n of elements, $n \geqq 2$. We define for a probability measure μ on Ω^I, $I \in \boldsymbol{N}$, such that $\mu(a) > 0$ for all $a \in \Omega^I$

$$\tilde{h}(\mu) = -\sum_{a=(a_i)_{i=1}^I \in \Omega^I} \mu(a) \log \frac{\mu(a)}{\sum_{\alpha \in \Omega} \mu((a_1, \ldots, a_{I-1}, \alpha))}.$$

We denote by $\mathfrak{M}_I$, $I \in \boldsymbol{N}$, the set of all probability measures μ on Ω^I such that $\mu(a) > 0$, $a \in \Omega^l$, and

$$\mu(\{b \in \Omega^I : a = (b_i)_{i=1}^l\}) = \mu(\{b \in \Omega^I : a = (b_{i+m})_{i=1}^l\}),$$
$$1 \leqq m \leqq I-l, \quad a \in \Omega^l, \quad 1 \leqq l < I.$$

Further we set

$$Z_a = \{x \in \Omega^{\boldsymbol{Z}} : a = (x_i)_{i=1}^I\}, \qquad a \in \Omega^I, \quad I \in \boldsymbol{N}.$$

For $\mu \in \mathfrak{M}_I$, $I \in \boldsymbol{N}$, we define a shift-invariant probability measure $\hat{\mu}$ on $\Omega^{\boldsymbol{Z}}$ by

$$\hat{\mu}(Z_a) = \mu(a), \qquad a \in \Omega^I,$$

$$\hat{\mu}(Z_{(\alpha_j)_{j=1}^J}) = \frac{\hat{\mu}(Z_{(\alpha_j)_{j=2}^J})\hat{\mu}(Z_{(\alpha_j)_{j=1}^{J-1}})}{\hat{\mu}(Z_{(\alpha_j)_{j=2}^{J-1}})}, \qquad (\alpha_j)_{j=1}^J \in \Omega^J, \quad J > I.$$

We note that (see [6, 5.10]) $h(\hat{\mu}) = \tilde{h}(\mu)$, $\mu \in \mathfrak{M}_I$, $I \in \boldsymbol{N}$, and that the $\hat{\mu}$ are ergodic. Indeed, the systems $(\hat{\mu}, S^1)$, $\mu \in \mathfrak{M}_I$, arise from indecomposable Markov chains. For probability measures μ, ν on Ω^I, $I \in \boldsymbol{N}$, we use the metric

$$|\mu, \nu| = \max_{a \in \Omega^I} |\mu(a) - \nu(a)|.$$

Let $I, N \in \boldsymbol{N}$, $I < N$. We define for $x \in \Omega^{N+I-1}$ a probability measure $\lambda_x^{(I)}$ on Ω^I by

$$\lambda_x^{(I)}(a) = N^{-1} \sum_{j=1}^{N} \delta_{a,(x_{j+i-1})_{i=1}^I}, \qquad a \in \Omega^I.$$

We set also

$$A(I, \mu, \delta, N) = \{x \in \Omega^{N+I-1} : |\lambda_x^{(I)}, \mu| < \delta\}, \qquad \mu \in \mathfrak{M}_I, \quad \delta > 0.$$

(3.1) Lemma. *Let $\mu \in \mathfrak{M}_I$, $I \in \boldsymbol{N}$ and let ε, $\delta > 0$. Then there is an $L > I$ such that*

$$|A(I, \mu, \delta, N)| > \exp[(\tilde{h}(\mu) - \varepsilon)N], \qquad N \geqq L.$$

Proof. The mean ergodic theorem and the Shannon-McMillan theorem [1, p. 129] show that there is an $L \in \boldsymbol{N}$ such that for all $N \geqq L$

$$\hat{\mu}(\{x \in \Omega^{\boldsymbol{Z}} : |\lambda^{(I)}_{(x_i)_{i=1}^{N+I-1}}, \mu| < \delta\}$$
$$\cap \{x \in \Omega^{\boldsymbol{Z}} : |N^{-1} \log \hat{\mu}(Z_{(x_i)_{i=1}^{N+I-1}}) + \tilde{h}(\mu)| < \varepsilon\}) > e^{-\varepsilon}.$$

We infer from this that

$$|A(I, \mu, \delta, N)| > \exp[(\tilde{h}(\mu)-\varepsilon)N-\varepsilon], \qquad N \geqq L. \qquad \text{Q.E.D.}$$

We say that an $a \in \Omega^l$, $l \in \boldsymbol{N}$, is a coding sequence if

$$(a_i)_{i=1}^m \neq (a_{l-m+i})_{i=1}^m, \qquad 1 \leqq m < l.$$

We say that an $a \in \Omega^l$ is an α-coding sequence of length l if $a_i=\alpha$, $1 \leqq i < l$ and $a_l=\beta\neq\alpha$. We set for $I, N \in \boldsymbol{N}$, $I<N$, $\mu \in \mathfrak{M}_I$ and $\delta > 0$,

$$B_a(I, \mu, \delta, N) = \{x \in A(I, \mu, \delta, N) : (x_{m+i})_{i=1}^l \neq a, \ 0 \leqq m \leqq N-I+1-l\}.$$

(3.2) Lemma. *Let $\mu \in \mathfrak{M}_I$, $I \in \boldsymbol{N}$, and let δ, $\varepsilon>0$. Then there exists a $K \in \boldsymbol{N}$ with the following property: For all α-coding sequences a of length $L \geqq K$*

$$|B_a(I, \mu, \delta, N)| > \exp[(\tilde{h}(\mu)-\varepsilon)N], \qquad N \geqq L.$$

Proof. By (3.1) we can find an $M \in \boldsymbol{N}$ such that

$$(1) \qquad |A(I, \mu, \delta/2, N)| > \exp[(\tilde{h}(\mu)-\varepsilon)N], \qquad N \geqq M.$$

We claim that any $K \in \boldsymbol{N}$ such that

$$(2) \qquad K > 4(M+I)\varepsilon^{-1}\delta^{-1}n$$

has the property that is stated in the lemma. Indeed, if a is an α-coding sequence of length L then with $\beta\neq\alpha$

$$(3) \qquad \begin{aligned} B_a(I, \mu, \delta, N) \supset {} & A(I, \mu, \delta, N) \\ & \cap \{x \in \Omega^{N-I+1} : x_{k(L-2)} = \beta, \ 1 \leqq k \leqq (N+I-1)(L-2)^{-1}\}. \end{aligned}$$

If $L \geqq K$, then $L-2-I > M$. Hence, by (1), (2) and (3),

$$\begin{aligned} |B_a(I, \mu, \delta, N)| &> \exp[(\tilde{h}(\mu)-\varepsilon)(1-\varepsilon\tilde{h}(\mu)^{-1})N] \\ &> \exp[(\tilde{h}(\mu)-2\varepsilon)N]. \end{aligned} \qquad \text{Q.E.D.}$$

We set for $I, N \in \boldsymbol{N}$

$$\mathfrak{N}(I, N) = \left\{k = (k_a)_{a\in\Omega^I} \in \boldsymbol{Z}^{n^I} : k_a > 0, \ a \in \Omega^I, \ \sum_{a\in\Omega^I} k_a = N\right\},$$

and for $k \in \mathfrak{N}(I, N)$

$$\tilde{h}(k) = \tilde{h}((N^{-1}k_a)_{a\in\Omega^I}),$$

$$C(I, N, k) = \{x \in \Omega^{N+I-1} : k_a = N\lambda_x^{(I)}(a), \ a \in \Omega^I\}.$$

(3.3) Lemma. *For all $k \in \mathfrak{N}(I, N)$*

$$|C(I, N, k)| < \exp(\tilde{h}(k)N) \prod_{a\in\Omega^I} \left(\frac{N}{k_a}\right)^{1/2}.$$

Proof. It is

$$|C(I, N, k)| \leqq n^{I-1} \prod_{a \in \Omega^{I-1}} \frac{\left(\sum_{\alpha \in \Omega} k_{(a_1,\dots,a_{I-1},\alpha)}\right)!}{\prod_{\alpha \in \Omega} k_{(a_1,\dots,a_{I-1},\alpha)}!}$$

$$= n^{I-1} \prod_{a \in \Omega^{I-1}} \left(\sum_{\alpha \in \Omega} k_{(a_1,\dots,a_{I-1},\alpha)}\right)! \left(\prod_{b \in \Omega^I} k_b!\right)^{-1}.$$

The lemma follows from this by an application of Stirling's formula. Q.E.D.

Denote

$$X_a = \{x \in \Omega^{\mathbf{Z}} : S^i x \in Z_a,\ S^{-j} x \in Z_a, \text{ for infinitely many } i, j \in \mathbf{N}\}, \quad a \in \Omega^I,\ I \in \mathbf{N}.$$

(3.4) THEOREM. *Let μ be an ergodic shift-invariant probability measure on $\Omega^{\mathbf{Z}}$ such that*

$$\mu(Z_a) > 0, \qquad a \in \Omega^I, \quad I \in \mathbf{N},$$

and let $\nu \in \mathfrak{M}_I$, $I \in \mathbf{N}$, $\tilde{h}(\nu) \geqq h(\mu)$. Let $\varepsilon > 0$. Then there exist coding sequences b and c and a homeomorphism U: $X_b \to X_c$ that commutes with the shift, such that

$$|\mu(U^{-1} Z_a) - \nu(a)| < \varepsilon, \qquad a \in \Omega^I.$$

Proof. We remark first that we can restrict attention to the case $h(\mu) < \tilde{h}(\nu)$. Indeed, if $h(\mu) = \log n$, then

$$\mu(Z_a) = \nu(a), \qquad a \in \Omega^I,$$

and if $h(\mu) = \tilde{h}(\nu) < \log n$, then there is a $\nu' \in \mathfrak{M}_I$ such that

$$h(\mu) < \tilde{h}(\nu'),$$

and $|\nu'(a) - \nu(a)| < \varepsilon/2$, $a \in \Omega^I$.

Let therefore

$$4\xi = \tilde{h}(\nu) - h(\mu) > 0.$$

We choose an $I' \geqq I$ such that $\tilde{h}(\nu) - \tilde{h}(\mu') < \xi$, where $\mu'(a) = \mu(Z_a)$, $a \in \Omega^{I'}$. Let

$$6\varepsilon' = n^{I-I'}\varepsilon. \tag{4}$$

Let also

$$\nu'(a) = \hat{\nu}(Z_a), \qquad a \in \Omega^{I'}.$$

We set $2\delta = \min_{a \in \Omega^{I'}} \mu(Z_a)$ and

$$F_N = \left\{x \in \Omega^{N+I'-1} : h(\lambda_x^{(I')}) - h(\mu) < 2\xi,\ \min_{a \in \Omega^{I'}} \lambda_x^{(I')}(a) > \delta\right\}, \qquad N > I'.$$

As a consequence of the individual ergodic theorem there is an $M \in \mathbf{N}$ such that

$$\mu\left(\bigcap_{M', -M'' \geqq M} \{x \in \Omega^{\mathbf{Z}} : (x_i)_{i=M''}^{M'} \in F_{M'-M''-I'+2}\}\right) > 1 - \varepsilon'. \tag{5}$$

By (3.2) we can also find an $L \in \mathbf{N}$, $L \geqq n$, such that for all γ-coding sequences c of length $L' \geqq L$

$$(6) \qquad |B_c(I', \nu', \varepsilon', N)| > \exp[(h(\nu)-\xi)N], \qquad N \geqq L'.$$

Let further $J \in \mathbf{N}$, $J > \xi^{-1}$, be such that

$$(7) \qquad J^{-n^{I'}}\delta^{(n^{I'}/2)} \exp(J\xi) > 1.$$

We choose now a $K \geqq L$ and $\alpha_1, \ldots, \alpha_K \in \Omega$ such that

$$(8) \qquad (I'+M+J+K)n^{-K} < \varepsilon',$$

and

$$(9) \qquad \mu(Z_{(\alpha_i)_{i=1}^K}) < n^{-K}.$$

Let $b \in \Omega^{2K+1}$ be the coding sequence that is given by

$$\begin{aligned} b_k &= \alpha_k, && \text{if } 1 \leqq k \leqq K, \\ &= \alpha_1, && \text{if } k = K+1, \\ &= \gamma \neq \alpha_1, && \text{if } K+1 < k \leqq 2K+1, \end{aligned}$$

and set

$$Y = \Omega^{\mathbf{Z}} - \bigcup_{i=1}^{2K+1} S^i Z_b.$$

We have from (8) and (9)

$$(10) \qquad \mu(Y) > 1-2\varepsilon'.$$

We define for $x \in Y$

$$i^+(x) = \min\{i \geqq 0 : S^i x \in Z_b\}, \qquad i^-(x) = \min\{i \geqq 0 : S^{-i-2K-2} \in Z_b\}.$$

(3.3) together with (6) and (7) implies that for a γ-coding sequence c of length $2K+1$

$$\begin{aligned} |B_c(I', \nu', \varepsilon', N)|\, |F_N|^{-1} &> \exp[(\tilde{h}(\nu)-\xi)N]N^{-n^{I'}}\delta^{(n^{I'}/2)} \exp[-(h(\mu)+2\xi)N] \\ &= N^{-n^{I'}}\delta^{(n^{I'}/2)} \exp(\xi N) > 1, \qquad N \geqq J+2K+1. \end{aligned}$$

We see now that there are mappings φ_N, $N \in \mathbf{N}$, of Ω^N onto itself such that

$$\varphi_{N+I'-1}(\{a \in F_N : (a_i)_{i=l}^{l+2K} \neq b,\ 1 \leqq l \leqq N+I'-2K\}) \subset B_c(I', \nu', \varepsilon', N), \qquad N \geqq J+2K+1,$$

and such that

$$\varphi_N(\{a \in \Omega^N : (a_i)_{i=l}^{l+2K} \neq b,\ 1 \leqq l \leqq N-2K\}) = \{a \in \Omega^N : (a_i)_{i=l}^{l+2K} \neq c,\ 1 \leqq l \leqq N-2K\}, \qquad N \geqq 2K+1.$$

We define now a homeomorphism $U\colon X_b \to X_c$, that commutes with S by setting for $x \in Z_b \cap X_b$, $Ux=y$, where

$$y_i = c_i, \qquad 1 \leqq i \leqq 2K+1$$
$$(y_i)_{i=-i^-(x)}^{0} = \varphi_{i^-(x)+1}(x_i)_{i=-i^-(x)}^{0}.$$

To conclude the proof of the theorem we use (5), (8), (9) and (10) to get

$$\begin{aligned}\mu(\{x \in Y : i^+(x)+i^-(x) \geqq J+2K,\ i^+(x) \geqq I'-1,\ (x_i)_{i=-i^-(x)}^{i^+(x)} \in F_{i^+(x)+i^-(x)-I'+2}\}) \\ > \mu(\{x \in Y : i^+(x),\ i^-(x) \geqq I'+M+J+K\})-\varepsilon' \\ > 1-2(I'+M+J+K)n^{-K}-3\varepsilon' \\ > 1-5\varepsilon'.\end{aligned}$$

We infer from this by applying the individual ergodic theorem that

$$(\nu'(a)-\varepsilon')(1-5\varepsilon') < \mu(U^{-1}Z_a) < \nu'(a)+6\varepsilon', \qquad a \in \Omega^{I'}.$$

Finally by (4)

$$|\mu(U^{-1}Z_a)-\nu(a)| < \varepsilon, \qquad a \in \Omega^I. \qquad \text{Q.E.D.}$$

We want to point out the following consequence of (3.4). Let

$$X = \bigcap_{I=1}^{\infty} \bigcap_{a\in\Omega^I} X_a,$$

and let $\mathfrak{M}_h$ be the set of shift-invariant ergodic probability measures μ on $\Omega^{\boldsymbol{Z}}$ such that $\mu(Z_a)>0$, $a \in \Omega^I$, $I \in \boldsymbol{N}$, and $\mu(X)=1$, $h(\mu)=h$, $0 \leqq h \leqq \ln n$.

The $\mathfrak{M}_h$ with the weak topology are polish spaces. The group $\mathfrak{G}$ of homeomorphisms of X that commute with the shift acts on $\mathfrak{M}_h$ by $\mu \to U\mu$, $\mu \in \mathfrak{M}_h$, $U \in \mathfrak{G}$, where

$$U\mu(Z_a) = \mu(U^{-1}Z_a), \qquad a \in \Omega^I, \quad I \in \boldsymbol{N}.$$

The homeomorphism that we have constructed in the proof of (3.4) maps X onto X. It follows therefore from (3.4) that the transformation groups $(\mathfrak{G}, \mathfrak{M}_h)$ are minimal, $0 \leqq h \leqq \ln n$.

4. **An estimate for** Δ.

(4.1) LEMMA. *For every ergodic shift-invariant probability measure μ on $\Omega^{\boldsymbol{Z}}$ there exists a shift-invariant probability measure ν on $\Omega^{\boldsymbol{Z}}$ such that for all $a \in \Omega^I$, $I \in \boldsymbol{N}$, $\nu(Z_a)>0$, and such that the systems $(\Omega^{\boldsymbol{Z}}, \mu, S)$ and $(\Omega^{\boldsymbol{Z}}, \nu, S)$ are isomorphic.*

Proof. If there is a $d \in \bigcup_{I=1}^{\infty} \Omega^I$, such that $\mu(Z_d)=0$, then we can assign in a 1-1 manner to every $a \in \bigcup_{I=1}^{\infty} \Omega^I$ a coding sequence $b(a)$ that contains a as a subsequence such that $\mu(Z_{b(a)})=0$. Let $L(a)$ be the length of $b(a)$. We can find Borel sets $A_a \subset \Omega^{\boldsymbol{Z}}$ such that for all $a, a' \in \bigcup_{I=1}^{\infty} \Omega^I$

$$\mu(S^l A_a \cap S^{l'} A_a) = 0, \qquad 0 \leqq l, \quad l' \leqq 2L(a),$$

and

$$\mu\left(\left(\bigcup_{l=0}^{2L(a)} S^l A_a\right) \cap \left(\bigcup_{l=0}^{2L(a')} S^{l'} A_{a'}\right)\right) = 0.$$

Choose $c(a) \in \Omega^{L(a)}$ such that $\mu(Z_{c(a)} \cap A_a) > 0$.

A Borel mapping $U: \Omega^{\boldsymbol{Z}} \to \Omega^{\boldsymbol{Z}}$ that commutes with the shift can be defined by

$$(Ux)_i = x_i, \qquad \text{if } S^i x \notin \bigcup_{I=1}^{\infty} \bigcup_{a \in \Omega^I} (Z_{c(a)} \cap A_a),$$

and

$$Ux \in Z_{c(a)}, \qquad \text{if } x \in Z_{c(a)} \cap A_a, \quad a \in \Omega^I, \quad I \in \boldsymbol{N}.$$

Setting for a Borel set $F \in \Omega^{\boldsymbol{Z}}$, $\nu(F) = \mu(U^{-1}F)$ proves the lemma. Q.E.D.

(4.2) Lemma. *Let T be an ergodic i.m.p.t. of $(E, \mathfrak{B}, p)$ with a generator*

$$\{A_0, \ldots, A_m\}, \qquad m > 1,$$

such that

$$p(A_0) > p(A_1) + 2p(A_2).$$

Then $\Delta(T) \leqq m$.

Proof. This lemma follows from a slightly generalized version of a theorem of A. H. Zaslavskiĭ [7, p. 295]. Q.E.D.

(4.3) Theorem. *Let T be an ergodic i.m.p.t. Then $\Delta(T) \leqq e^{h(T)} + 1$.*

Proof. By (2.1) there exist a state space $\Omega = \{0, \ldots, m\}$, $m \in \boldsymbol{N}$, and a shift-invariant probability measure μ on $\Omega^{\boldsymbol{Z}}$ such that T is isomorphic to the system $(\Omega^{\boldsymbol{Z}}, \mu, S)$. By (4.1) we can assume here that $\mu(Z_a) > 0$, $a \in \Omega^I$, $I \in \boldsymbol{N}$. If now $m > e^{h(T)}$, then we can find a q, $0 < q < (2m)^{-1}$ such that

$$h((\lambda_k)_{k=0}^m) > h(T),$$

where

$$\lambda_0 = n^{-1} + q, \ \lambda_1 = n^{-1} - 2q, \ \lambda_2 = q, \ \lambda_k = n^{-1}, \qquad 2 < k \leqq m.$$

From (3.4) we see now that there is a shift-invariant probability measure ν on $\Omega^{\boldsymbol{Z}}$ such that $(\Omega^{\boldsymbol{Z}}, \mu, S)$ is isomorphic to $(\Omega^{\boldsymbol{Z}}, \nu, S)$ and such that

$$|\nu(Z_{(k)}) - \lambda_k| < q/4, \qquad 1 \leqq k \leqq m.$$

It is then

$$\nu(Z_{(0)}) > \nu(Z_{(1)}) + 2\nu(Z_{(2)})$$

and the theorem follows by means of (4.2). Q.E.D.

(4.4) Corollary. *Let T be the cartesian product of the n-shift with entropy* $\ln n$, $n \geqq 2$, *and an ergodic i.m.p.t. with entropy zero. Then $\Delta(T) = n + 1$.*

References

1. P. Billingsley, *Ergodic theory and information*, Wiley, New York, 1965. MR **33** #254.
2. N. Bourbaki, *Topologie générale*, 2ème éd., Livre III Chapitre 9: *Utilisation des nombres réels en topologie générale*, Actualités Sci. Indust., no. 1045, Hermann, Paris, 1958. MR **30** #3439.
3. F. Hahn and W. Parry, *Some characteristic properties of dynamical systems with quasi-discrete spectra*, Math. Systems Theory **2** (1968), 179–190. MR **37** #6435.
4. K. Jacobs, *Lecture notes on ergodic theory*, Aarhus Universitet, 1962/63.
5. W. Parry, *Entropy and generators in ergodic theory*, Benjamin, New York, 1969.
6. V. A. Rohlin, *Lectures on the entropy theory of transformations with invariant measure*, Uspehi Mat. Nauk **22** (1967), no. 5 (137), 3–56 = Russian Math. Surveys **22** (1967), no. 5, 1–52. MR **36** #349.
7. A. H. Zaslavskiĭ, *On the isomorphism problem for stationary processes*, Teor. Verojatnost i Primenen. **9** (1964), 318–326 = Theor. Probability Appl. **9** (1964), 291–298. MR **29** #1674.

Ohio State University,
Columbus, Ohio 43210