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 $e^{h(T)} \leq k \leq e^{h(T)} + 1$.

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

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

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 $Q$. For finite $Q$ we shall prove in §3 an approximation theorem for probability measures on $Q$ that are invariant under the shift $S$,

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

This theorem will enable us to derive in §4 from the work of A. H. Zaslavskii [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 $\Delta(T) \geq e^{h(T)}$. Our result is

$$\Delta(T) \leq e^{h(T)} + 1.$$
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 \( \{A_n : n \in N\} \) be a partition of \((E, \mathcal{B}, p)\) with finite entropy. Then there exists a mapping \( n \rightarrow K_n \in N (n \in N) \) and a 1-1 mapping

\[ \varphi : N \rightarrow \bigcup_{k=1}^{\infty} \{1, 2, 3\}^k \]

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

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

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

(2) \[ -\log p(A_n) - 1 < l(n) \leq -\log p(A_n), \quad n \in N. \]

Let further

\[ n_1 = 1, \quad n_m = \min \{n > n_{m-1} : l(n) > l(n_{m-1})\}, \quad m > 1. \]

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

\[ \sum_{m=1}^{\infty} (n_{m+1} - n_m) 3^{-l(n_m)} \leq e. \]

Consequently, for some \( m_0 \in N, \)

(3) \[ n_{m+1} - n_m < 3^{l(n_m)}, \quad m \geq m_0. \]

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

\[ \varphi : N \rightarrow \bigcup_{k=1}^{\infty} \{1, 2, 3\}^k \]

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 \( \{A_n : n \in \mathbb{N} \} \) and from (2):

\[
\sum_{n = n_0}^{\infty} K_n p(A_n) \leq \sum_{n = 1}^{\infty} l(n) p(A_n) \leq \sum_{n = 1}^{\infty} -p(A_n) \log p(A_n) < \infty.
\]

2. Let \( \Omega \) be a finite set containing more than two elements. Let \( \omega \in \Omega, C \in \mathbb{N} \), and let

\[
X = \left\{ x = ((x_{i,j})_{j=1}^\infty)_{i=-\infty}^{\infty} \in \left( \bigcup_{k=2}^{\infty} \Omega^k \right)^2 : x_{i,j} \neq \omega, 1 \leq j < D_i, x_{i,D_i} = \omega, -\infty < i < \infty \right\},
\]

and

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

We are going to construct a 1-1 Borel mapping \( U : X_C \to (\Omega^C)^2 \) that commutes with the shifts.

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

\[
I(i, j) = \min \left\{ l > j - C : \sum_{i < m \leq l} (D_m - C) \leq 0 \right\},
\]

and

\[
J(i, j) = j + \sum_{i < m \leq l(i, j)} (D_m - C).
\]

It follows that

\[
D_{I(i, j)} < J(i, j) \leq C, \quad i \in \Gamma, \quad C < j \leq D_i.
\]

The mapping

\[
(i, j) \to (I(i, j), J(i, j)) \quad (i \in \Gamma, C < j \leq D_i)
\]

is 1-1. Indeed, had we \( i, i' \in \Gamma, C < j \leq D_i, C < j' \leq D_{i'} \),

\[
(I(i, j), J(i, j)) = (I(i', j'), J(i', j')),
\]

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

\[
j + \sum_{i < m \leq l(i, j)} D_m \geq J(i, j) + (I(i, j) - i) C,
\]

and therefore that

\[
j + \sum_{i < m < i'} D_m \leq (i' - i) C,
\]

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

\[
Ux = ((y_{i,j})_{j=1}^\infty)_{i=-\infty}^{\infty} \in (\Omega^C)^2
\]
by setting
\[ y_{i,f} = x_{i,f}, \quad \text{if } i \in \mathbb{Z}, \quad 1 \leq j \leq \min(C, D), \]
\[ y_{i,J(i) \cap U, f} = x_{i,f}, \quad \text{if } i \in \Gamma, \quad C < j \leq D, \]
and by setting \( y_{1,f} = \alpha, \alpha \in \Omega, \alpha \neq \omega, \) 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
\[ I_o(i) = I(i, D_i), \quad J_o(i) = J(i, D_i), \quad i \in \Gamma, \]
and
\[ N_o(i) = \sum_{j=1}^{C} \delta_{o,y_{i,f}}, \quad i \in \mathbb{Z}. \]

We have for \( i, i' \in \Gamma \)
\[ i < i' < I_o(i) \Rightarrow I_o(i') \leq I_o(i), \tag{6} \]
\[ i < I_o(i') < I_o(i) \Rightarrow i < i', \tag{7} \]
\[ i < i', \quad I_o(i) = I_o(i') \Rightarrow J_o(i) > J_o(i'). \tag{8} \]

From these relations and since \( i \to (I_o(i), J_o(i)) (i \in \Gamma) \) is 1-1 we have
\[ \sum_{i < m < I_o(i)} N_o(m) + \sum_{j=1}^{I_o(i)} \delta_{o,y_{i,f}} = |\{i' \in \mathbb{Z} - \Gamma : i < i' \leq I_o(i)\}| + |\{i' \in \Gamma : i < I_o(i') < I_o(i)\}| \]
\[ + \{i' \in \Gamma : I_o(i') = I_o(i), J_o(i') < J_o(i)\} + 1 \]
\[ = |\{i' \in \mathbb{Z} - \Gamma : i < i' \leq I_o(i)\}| + |\{i' \in \Gamma : i < i' < I_o(i)\}| + 1 \]
\[ = I_o(i) - i + 1, \quad i \in \Gamma. \tag{9} \]

And we have from (7)
\[ \sum_{i < m < L} N_o(m) = |\{i' \in \mathbb{Z} - \Gamma : i < i' \leq L\}| + |\{i' \in \Gamma : i < I_o(i') \leq L\}| \]
\[ = |\{i' \in \mathbb{Z} - \Gamma : i < i' \leq L\}| + |\{i' \in \Gamma : i < i' < L\}| \]
\[ \leq L - i, \quad i \in \Gamma, \quad i < L < I_o(i). \tag{10} \]

Now we see from (9) and (10) that
\[ I_o(i) = \min \left\{ L > i : \sum_{i < m < L} N_o(m) > L - i \right\}, \quad i \in \Gamma \tag{11} \]
and that
\[ J_o(i) = \min \left\{ 1 < l \leq C : \sum_{i < m < I_o(i)} N_o(m) \right. \]
\[ + \sum_{j=1}^{I_o(i)} \delta_{o,y_{i,f}} = I_o(i) - i + 1 \}, \quad i \in \Gamma. \tag{12} \]
Next we observe that

\[ D_i = I_0(i) + (I_0(i) - i)C - \sum_{t < n \in I_0(i)} D_t, \quad i \in \Gamma. \]

We know from (6) that

\[ i < i' < I_0(i) \Rightarrow I_0(i') - i' < I_0(i) - i, \quad i, i' \in \Gamma. \]

It follows that if \( i_0 \in \Gamma \) is such that

\[ I_0(i_0) - i_0 = \min \{ I_0(i) - i : i \in \Gamma \} \]

then \( i_0 < i \leq I_0(i_0) = i \in \mathbb{Z} - \Gamma \) and we see from (13) that \( D_{i_0} \) can be computed from the \( y_{i,j} \), \( 1 \leq j \leq C \), \( i \in \mathbb{Z} \). 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 \( N^\mathbb{Z} \) together with an invariant probability measure \( \mu \) such that the partition

\[ \{(n_i)_{i=-\infty}^{\infty} \in N^\mathbb{Z} : n_0 = m \}, \quad m \in \mathbb{N}, \]

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

\[ n \rightarrow (x_{n,1}, \ldots, x_{n,K_n}) \in \bigcup_{k=1}^\infty \{1, 2, 3\}^k \quad (n \in \mathbb{N}) \]

such that

\[ \sum_{m=1}^\infty K_m \mu(\{(n_i)_{i=-\infty}^{\infty} \in N^\mathbb{Z} : n_0 = m \}) < C - 1. \]

We use this mapping to build a 1-1 mapping

\[ V : (n_i)_{i=-\infty}^{\infty} \rightarrow ((x_{n_i,1}, \ldots, x_{n_i,K_n}, \omega))_{i=-\infty}^{\infty} \in X \quad ((n_i)_{i=-\infty}^{\infty} \in N^\mathbb{Z}) \]

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

\[ \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{i=1}^L (K_i + 1 - C) < 0, \quad \text{for } \mu\text{-a.a. } (n_i)_{i=-\infty}^{\infty} \in N^\mathbb{Z}. \]

Hence \( \mu(V^{-1}X_\mathcal{O}) = 1 \).

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

\[ U : X_\mathcal{O} \rightarrow \Omega^\mathbb{Z} \]

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

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

then we find that \((N^\mathbb{Z}, \mu, S)\) is isomorphic to \((\Omega^\mathbb{Z}, \nu, S)\). (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 $\mu$ 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 $\Omega$ be a state space containing a finite number $n$ of elements, $n \geq 2$. We define for a probability measure $\mu$ on $\Omega^I$, $I \in N$, such that $\mu(a) > 0$ for all $a \in \Omega^I$

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

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

$$\mu(\{b \in \Omega^I : a = (b_i)_{i=1}^I\}) = \mu(\{b \in \Omega^I : a = (b_{i+m})_{i=1}^I\}),$$

$$1 \leq m \leq I-I, \quad a \in \Omega^I, \quad 1 \leq I < I.$$

Further we set

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

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

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

$$\hat{\mu}(Z_{(a_i)_{i=1}^J}) = \frac{\hat{\mu}(Z_{(a_i)_{i=2}^J}) \hat{\mu}(Z_{(a_i)_{i=1}^J})}{\hat{\mu}(Z_{(a_i)_{i=2}^J})}, \quad (a_i)_{i=1}^J \in \Omega^I, \quad J > I.$$ 

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

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

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

$$\lambda_x^{(I)}(a) = N^{-1} \sum_{j=1}^N \delta_{a, (x_j)_{j=1}^I}, \quad a \in \Omega^I.$$ 

We set also

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

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

$$|A(I, \mu, \delta, N)| > \exp [(h(\mu) - \epsilon)N], \quad N \geq L.$$ 

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

$$\hat{\mu}(\{x \in \Omega^\infty : |\lambda_x^{(I+N+I-1)}, \mu| < \delta\} \cap \{x \in \Omega^\infty : |N^{-1} \log \hat{\mu}(Z_{(x_i)_{i=1}^{N+I-1}}, \hat{h}(\mu)| < \epsilon\}) > e^{-\epsilon}. $$
We infer from this that

\[ |A(I, \mu, \delta, N)| > \exp \left[ (\bar{h}(\mu) - \varepsilon)N - \varepsilon \right], \quad N \geq L. \quad \text{Q.E.D.} \]

We say that an \( a \in \Omega^l \), \( l \in \mathbb{N} \), is a coding sequence if

\[ (a_i)_{i=1}^l \neq (a_{i-m})_{i=m+1}^l, \quad 1 \leq m < l. \]

We say that an \( a \in \Omega^l \) is an \( \alpha \)-coding sequence of length \( l \) if \( a_i = \alpha \), \( 1 \leq i < l \) and \( a_l = \beta \neq \alpha \). We set for \( I, N \in \mathbb{N}, I < N, \mu \in \mathcal{M}_I \); and \( \delta > 0 \),

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

(3.2) Lemma. Let \( \mu \in \mathcal{M}_I, I \in \mathbb{N}, \) and let \( \delta, \varepsilon > 0 \). Then there exists a \( K \in \mathbb{N} \) with the following property: For all \( \alpha \)-coding sequences \( a \) of length \( L \geq K \)

\[ |B_a(I, \mu, \delta, N)| > \exp \left[ (\bar{h}(\mu) - \varepsilon)N \right], \quad N \geq L. \]

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

(1) \[ |A(I, \mu, \delta/2, N)| > \exp \left[ (\bar{h}(\mu) - \varepsilon)N \right], \quad N \geq M. \]

We claim that any \( K \in \mathbb{N} \) such that

(2) \[ K > 4(M + 1)e^{-1} \delta^{-1} \]

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

\[ B_a(I, \mu, \delta, N) \supset A(I, \mu, \delta, N) \]

\[ \cap \{ x \in \Omega^{N-I+1} : x_{\sigma L-2} = \beta, 1 \leq k \leq (N+I)(L-2)^{-1} \}. \]

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

\[ |B_a(I, \mu, \delta, N)| > \exp \left[ (\bar{h}(\mu) - \varepsilon)(1 - e\bar{h}(\mu)^{-1})N \right] \]

\[ > \exp \left[ (\bar{h}(\mu) - 2\varepsilon)N \right]. \quad \text{Q.E.D.} \]

We set for \( I, N \in \mathbb{N} \)

\[ \mathcal{R}(I, N) = \left\{ k = (k_a)_{a \in \Omega^l} \in \mathbb{Z}^{\Omega^l} : k_a > 0, \ a \in \Omega^l, \sum_{a \in \Omega^l} k_a = N \right\}, \]

and for \( k \in \mathcal{R}(I, N) \)

\[ \bar{h}(k) = \bar{h}(N^{-1}k_a)_{a \in \Omega^l}, \]

\[ C(I, N, k) = \{ x \in \Omega^{N+I-1} : k_a = N\lambda^I_N(a), \ a \in \Omega^l \}. \]

(3.3) Lemma. For all \( k \in \mathcal{R}(I, N) \)

\[ |C(I, N, k)| < \exp \left( \bar{h}(k)N \right) \prod_{a \in \Omega^l} \left( \frac{N}{k_a} \right)^{1/2}. \]
Proof. It is

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

$$= n^I - 1 \prod_{a \in \Omega^I} \left( \frac{\sum_{a \in \Omega^I} k(a_1, \ldots, a_{I-1}, a)}{\sum_{a \in \Omega} k(a_1, \ldots, a_{I-1}, a)} \right)! \left( \prod_{a \in \Omega^I} k_a \right)^{-1}.$$

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

Denote

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

(3.4) THEOREM. Let \( \mu \) be an ergodic shift-invariant probability measure on \( \Omega^Z \) such that

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

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

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

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

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

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

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

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

Let therefore

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

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

$$6\epsilon' = n^{I-I'} \epsilon.$$

Let also

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

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

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

As a consequence of the individual ergodic theorem there is an \( M \in \mathbb{N} \) such that

$$\mu_{M', \cap M'} \left( x \in \Omega^Z : (x_i)_{m^* - 1}^{m^*} \in F_{M'^* - M'^* + 2} \right) > 1 - \epsilon'.$$
By (3.2) we can also find an \( L \in N, L \geq n \), such that for all \( \gamma \)-coding sequences \( c \) of length \( L' \geq L \)

\[
|B_c(I', \nu', \varepsilon', N)| > \exp\left(\frac{(h(\nu) - \xi)N}{n}\right), \quad N \geq L'.
\]

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

\[
J^{-n'/\xi(n'/2)} \exp(J\xi) > 1.
\]

We choose now a \( K \geq L \) and \( \alpha_1, \ldots, \alpha_K \in \Omega \) such that

\[
(I' + M + J + K)n^{-K} < \varepsilon',
\]

and

\[
\mu(Z(\alpha)_n) < n^{-K}.
\]

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

\[
b_k = \alpha_k, \quad \text{if } 1 \leq k \leq K,
\]

\[
= \alpha_1, \quad \text{if } k = K+1,
\]

\[
= \gamma \neq \alpha_1, \quad \text{if } K+1 < k \leq 2K+1,
\]

and set

\[
Y = \Omega^Z - \bigcup_{i=1}^{2K+1} S'_i Z_b.
\]

We have from (8) and (9)

\[
\mu(Y) > 1 - 2\varepsilon'.
\]

We define for \( x \in Y \)

\[
i^+(x) = \min\{i \geq 0 : S^i x \in Z_b\}, \quad i^-(x) = \min\{i \geq 0 : S^{-i-2K-2} \in Z_b\}.
\]

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

\[
|B_c(I', \nu', \varepsilon', N)| \exp\left(\frac{(\tilde{\nu}(\nu) - \tilde{\xi})N}{n'/\xi(n'/2)}\right) \exp\left[-(h(\mu)+2\xi)N\right]
\]

\[
= N^{-n'/\xi(n'/2)} \exp(\xi N) > 1, \quad N \geq J+2K+1.
\]

We see now that there are mappings \( \varphi_N, N \in N, \) of \( \Omega^N \) onto itself such that

\[
\varphi_{N+I'-1}\{a \in F_N : (a_l)_{l=1}^{N+2K} \neq b, 1 \leq l \leq N+I'-2K\} \subseteq B_c(I', \nu', \varepsilon', N),
\]

\[
N \geq J+2K+1, \quad N \geq J+2K+1,
\]

and such that

\[
\varphi_N\{a \in \Omega^N : (a_l)_{l=1}^{N+2K} \neq b, 1 \leq l \leq N-2K\}
\]

\[
= \{a \in \Omega^N : (a_l)_{l=1}^{N+2K} \neq c, 1 \leq l \leq N-2K\}, \quad N \geq 2K+1.
\]
We define now a homeomorphism $U : X_b \to X_c$, that commutes with $S$ by setting for $x \in Z_b \cap X_b$, $Ux = y$, where

$$y_i = c_i, \quad 1 \leq i \leq 2K + 1$$

$$(y_i)^{p_{-i}} = q_{i-1}(x)_1 (x_i)^{p_{-i}} = q_{i-1}(x).$$

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

$$\mu(\{x \in Y : i^+(x) + i^-(x) \geq J + 2K, i^+(x) \geq I' - 1, (x_i)^{p_{-i}} = p_{i-1}(x), i^-(x) = i^+(x) + 1 + K\}) - \epsilon'$$

$$> 1 - 2(I' + M + J + K)n^{-K} - 3\epsilon'$$

$$> 1 - 5\epsilon'.$$

We infer from this by applying the individual ergodic theorem that

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

Finally by (4)

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

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

$$X = \bigcap_{i=1}^{\infty} X_i$$

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

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

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

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 $(\Theta, \mathcal{M}_h)$ are minimal, $0 \leq h \leq \ln n$.

4. An estimate for $\Delta$.

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

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

$$\mu(S^l A_a \cap S'^{l'} A_a) = 0, \quad 0 \leq l, \quad l' \leq 2L(a).$$
Choose $c(a) \in \Omega^{Z(a)}$ such that $\mu(Z_{c(a)} \cap A_a) > 0$.

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

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

and

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

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

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

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

such that

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

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

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

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

Proof. By (2.1) there exist a state space $\Omega = \{0, \ldots, m\}$, $m \in \mathbb{N}$, and a shift-invariant probability measure $\mu$ on $\Omega^Z$ such that $T$ is isomorphic to the system $(\Omega^Z, \mu, S)$. By (4.1) we can assume here that $\mu(Z_a) > 0, a \in \Omega^I, I \in \mathbb{N}$. 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

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

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

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

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 $\ln n, n \geq 2$, and an ergodic i.m.p.t. with entropy zero. Then $\Delta(T) = n + 1$.  

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
References


Ohio State University,
Columbus, Ohio 43210