COMMUTING MEASURE-PRESERVING TRANSFORMATIONS

BY

YITZHAK KATZNELSON AND BENJAMIN WEISS

ABSTRACT

Let $\phi_1, \cdots, \phi_d$ be commuting measure-preserving transformations, $\phi^l \equiv \phi_1^{l_1} \phi_2^{l_2} \cdots \phi_d^{l_d}$, $\Phi = \{\phi^l\}$. The Kakutani-Rokhlin tower theorem is proved in a refined form for non-periodic groups Φ, and the Shannon-McMillan theorem is extended to ergodic groups. These results are used to extend recent isomorphism results to groups of transformations.

1. Introduction

Topological dynamics has long been studying the properties of a general group of homeomorphisms acting on a fixed space. The ergodic theory of measure-preserving transformations, on the other hand, has for the most part been developed only for the powers of a single transformation, i.e., actions of Z or R. Our goal here is to extend some of this theory to actions of Z^d with $d > 1$.

The problem which motivated this work was the question, raised in [12], as to whether entropy is a complete invariant for multi-parametered Bernoulli schemes. To give an affirmative answer to this question, some basic results of the one dimensional theory have to be extended to families of commuting transformations. Before describing these results, we introduce some notation. All transformations will be measure preserving and act on a fixed Lebesgue space $(X, \mathscr{B}, m)$. If $l = (l_1, \cdots, l_d) \in Z^d$, and ϕ_i, $1 \leqq i \leqq d$, are commuting transformations of X, ϕ^l will denote $\phi_1^{l_1}\phi_2^{l_2}\cdots\phi_d^{l_d}$, and Φ the group $\{\phi^l\}$.

In section 2 we prove a refined version of the basic Kakutani-Rokhlin tower theorem for groups Φ, and mention some applications. In section 3 the entropy of Φ is defined and in section 4 the Shannon-McMillan theorem is proved for

Received April 10, 1972

groups Φ. These results are then used in section 5 to prove the isomorphism of Bernoullian groups with the same entropy. The detail in which the arguments are given varies considerably; since section 2 should prove generally useful, it has been written *in extenso* while section 5 contains mainly hints for the initiated. Familiarity with [9] and [2] is assumed throughout.

2. Non-periodic transformations and the Kakutani-Rokhlin construction

The following theorem adumbrated by Kakutani, and first given explicitly by Rokhlin ([8], cf. [2]) is of basic importance in ergodic theory. A measure-preserving transformation ψ is said to *non-periodic* if for all $n \geqq 1$

$$m\{x: \phi^n x = x\} = 0. \tag{2.1}$$

THEOREM (K- R). *If ψ is non-periodic and $n \geqq 1$, $\varepsilon > 0$ are given, then there exists a set F such that $\{\psi^j F\}_{j=0}^n$ are disjoint and*

$$m\left\{\bigcup_{j=0}^{n} \psi^j F\mathrm{i}\right\} > 1 - \varepsilon. \tag{2.2}$$

REMARK. We shall have occasion to apply the K-R theorem to transformations that are only partially defined. To see that this is possible proceed as follows: If ψ is defined and one-to-one on $B \subset X$ and non-periodic in the sense that, for all n, $m\{x: \psi^i x \in B,\ 0 \leqq i \leqq n,\ \psi^n x = x\} = 0$, then ψ can be extended to $\tilde{\psi}$, a one-to-one non-periodic mapping on all of X. If $m(X \setminus B) < \varepsilon$, then the transformation induced by $\tilde{\psi}$ on B (see (2.6) below) differs from ψ on a set of measure at most ε. Now, this induced transformation is again non-periodic and the K-R theorem may be applied to it to yield an approximate tower for the original ψ.

Our aim in this section is to extend this theorem to groups Φ. We shall say that a group Φ of measure-preserving transformation is *non-periodic* if for every $l \in Z^d \setminus \{0\}$

$$m\{x: \phi^l x = x\} = 0 \tag{2.3}$$

The set of $l' \in Z^d$ that corresponds to $\{0, 1, \cdots, n\}$ in Z^1 is called the *l-rectangle* and written:

$$R_l = \{l': 0 \leqq l' \leqq l\}, \tag{2.4}$$

where, as usual, $l' \leqq l$ means $l'_i \leqq l_i$ for $1 \leqq i \leqq d$. It will be convenient to use the following notion:

DEFINITION A set A is said to be ε-*invariant* under ψ if

$$(2.5) \qquad \frac{m(A\Delta\psi A)}{m(A)} < \varepsilon$$

where $(A\Delta B = (A \cup B)\setminus(A \cap B))$.

There is another way of describing this near invariance. Recall that ψ_A, the transformation induced by ψ on A, is defined on $x \in A$ by

$$(2.6) \qquad \psi_A(x) = \psi^{\tau_A(x)}(x)$$

where $\tau_A(x) = \min\{k: k \geqq 1, \psi^k x \in A\}$. The recurrence theorem guarantees that τ_A is finite a.e. Defining $m_A(B) = m(A \cap B)\setminus m(A)$, then up to a constant (2.5) is equivalent to

$$(2.7) \qquad m_A\{x: \psi_A(x) \neq \psi(x)\} < \varepsilon.$$

Note too, for later use, that if ψ is non-periodic so is ψ_A. We can now state a strong version of the K-R theorem for groups Φ:

THEOREM 1. *If Φ is non-periodic and $l \geqq 0$, $\varepsilon > 0$ are given, then there is a set $E \subset X$ that satisfies*

$$(2.8) \qquad \{\phi^{l'}E\}_{l' \in R_l} \text{ is a disjoint family;}$$

$$(2.9) \qquad m\left\{\bigcup_{l' \in R_l} \phi^{l'}E\right\} > 1 - \varepsilon;$$

$$(2.10) \qquad E \text{ is } \varepsilon\text{-invariant under } \phi_i^{l_i+1},\ 1 \leqq i \leqq d.$$

The theorem will follow from the special case $l_1 = 1$, $l_j = 0$, $2 \leqq j \leqq d$, which we formulate as

LEMMA 1. *Let Φ be non-periodic and $\varepsilon > 0$ given, then there is a set $A \subset X$ such that*

$$(2.8)' \qquad \phi_1(A) \cap A = \varnothing$$

$$(2.9)' \qquad m(A) > \tfrac{1}{2} - \varepsilon$$

$$(2.10)' \qquad A \text{ is } \varepsilon\text{-invariant under } \phi_j,\ 2 \leqq j \leqq d.$$

An induction argument will proceed by assuming Theorem 1 to hold for $d-1$ (note that for $d = 1$, (2.10) is obvious so we have just the K-R theorem) in order to prove Lemma 1 for d. Then, Lemma 1 for d together with the Theorem for $d-1$ will be used to prove the Theorem for d. In order to prove Lemma 1, we first prove a weaker version, and, to make the special role of ϕ_1 more explicit, we denote it by a different letter $\psi \equiv \phi_1$.

LEMMA 2. *Let Φ be non-periodic ($\phi_1 \equiv \psi$), and $\varepsilon > 0$ given. Then there exists a set $A \subset X$ such that*

$$\psi(A) \cap A = \emptyset \tag{2.8}''$$

$$m(\psi^{-1}(A) \cap A \cap \psi(A)) > 1 - \varepsilon \tag{2.9}''$$

$$A \text{ is } \varepsilon\text{-invariant under } \phi_j, 2 \leqq j \leqq d. \tag{2.10}''$$

PROOF OF LEMMA 2. Choose a $N \gg \varepsilon$ and $\varepsilon_1 > 0$ (precisely how large N must be and how small ε_1 will become clear in the course of the proof). Let $\tilde{l}$ denote an element of Z^{d-1}, $\tilde{l} = (l_2, l_3, \cdots, l_d)$, $\phi^{\tilde{l}} = \phi_2^{l_2}, \cdots, \phi_d^{l_d}$. Denote the vector $\tilde{l}$ with $l_j = N$, $2 \leqq j \leqq d$ by $\tilde{N}$ and let $\tilde{R}_N = \{\tilde{l}: 0 \leqq \tilde{l} \leqq \tilde{N}\}$. By Theorem 1 (for $d-1$) there exists a set A_0 with

$$\{\phi^{\tilde{l}} A_0\}_{\tilde{l} \in \tilde{R}_N} \text{ disjoint,} \tag{2.11}$$

$$m\left(\bigcup_{\tilde{l} \in \tilde{R}_N} \phi^{\tilde{l}} A_0\right) > 1 - \varepsilon_1, \tag{2.12}$$

$$A_0 \text{ is } \varepsilon_1\text{-invariant under } \phi_j^{N+1}, 2 \leqq j \leqq d. \tag{2.13}$$

Define ψ^* on A_0 as follows

$$\psi^*(x) = \tilde{\phi}^{-\tilde{l}} \psi(x) \tag{2.14}$$

where $\tilde{l} \in \tilde{R}_N$ is chosen so that $\psi^*(x) \in A_0$. This defines ψ^* on A_0 up to a set of measure less than ε_1; one checks that ψ^* is one-to-one and non-periodic in the sense of the Remark. Thus by the Remark we can find an $A_1 \subset A_0$ with

$$m(A_1) > \tfrac{1}{2} m(A_0) - 2\varepsilon_1, \quad \psi^*(A_1) \cap A_1 = \emptyset. \tag{2.15}$$

For each $2 \leqq j \leqq d$ set $A_{1,0}^{(j)} = \phi_j^{-N-1}(A_1) \cap A_1$, $A_{1,1}^{(j)} = A_1 \backslash A_{1,0}^{(j)}$, $B_{1,0}^{(j)} = \psi^*(A_{1,0}^{(j)})$, $B_{1,1}^{(j)} = \psi^*(A_{1,1}^{(j)})$, and $B_1 = \psi^* A_1$. If $x \in A_{1,0}^{(j)}$ then $\psi^* \phi_j^{N+1}(x) = \phi_j^{N+1} \psi^*(x) \in B_1$, and hence

$$\phi_j^{N+1}(B_{1,0}^{(j)}) \subset B_1. \tag{2.16}$$

Similarly one sees

$$\phi_j^{N+1}(A_{1,1}^{(j)}) \subset B_1, \ \phi_j^{N+1}(B_{1,1}^{(j)}) \subset A_1. \tag{2.17}$$

Hence ψ_j may be defined on A_1 to A_1 by

$$\psi_j(x) = \begin{cases} \phi_j^{N+1}(x), & x \in A_{1,0}^{(j)} \\ \phi_j^{N+1} \psi^*(x), & x \in A_{1,1}^{(j)} \end{cases} \tag{2.18}$$

One checks again that ψ_j is non-periodic on A_1. Hence by the K-R theorem, there are sets $A_2^{(j)}$ such that

$$m\left(A_2^{(j)}\right) > \left(\tfrac{1}{2} - \frac{\varepsilon_1}{2d}\right) m(A_1), \quad A_2^{(j)} \cap \psi_j(A_2^{(j)}) = \varnothing. \tag{2.19}$$

Omitting a set of measure less than ε_1, we have $(d-1)$ partitions of A_1 into two sets. Hence if G is the largest atom of $\bigvee_2^d \{A_2^{(j)}, \psi_j A_2^{(j)}\}$ then

$$m(G) > 2^{1-d}(1-\varepsilon_1)m(A_1). \tag{2.20}$$

If $C = \bigcup_{\bar{\imath} \in \tilde{R}_N} \phi^{\bar{\imath}} \tilde{C}$ then from (2.20)

$$m(C) > 2^{-d}(1-\varepsilon_1), \tag{2.21}$$

while from (2.19) and (2.18) it follows that $\psi(C) \cap C = \varnothing$ and C is $1/N$-invariant under ϕ_j 2, $\leqq j \leqq d$.

The natural thing to do once we know how to get a definite fraction as in (2.21) is to iterate. However, a preliminary iteration is necessary to obtain that the complement of $\psi^{-1}(C) \cup C \cup \psi(C)$ is not only nearly ϕ_j-invariant, $2 \leqq j \leqq d$, but also nearly ψ-invariant so that the construction in the first part may be repeated. Therefore put

$$\begin{aligned} D_0 &= \psi^{-1}(C) \cup C \cup \psi(C), \quad C_1' = (X \setminus D_0) \cap \psi^{-1}(D_0), \\ C_1 &= C \cup C_1' \text{ and in general} \\ D_i &= \psi^{-1}(C_i) \cup C_i \cup \psi(C_i) \\ C_{i+1}' &= \psi^{-1}(D_i) \cap (X \setminus D_i), \quad C_{i+1} = C_i \cup C_{i+1}'. \end{aligned} \tag{2.22}$$

After k-iterations, $E = C_k$ is k^2/N-invariant under ϕ_j, $2 \leqq j \leqq d$, $1/k$-invariant under ψ and $E \cap \psi(E) = \varnothing$. This implies that $F = X \setminus (\psi^{-1}(E) \cup E \cup \psi(E))$ is $1/k$-invariant under ψ as well as k^2/N-invariant under ϕ_j, $2 \leqq j \leqq d$. If k is chosen large enough and then N still larger so that k^2/N is small, we can iterate the entire construction of the first part (naturally with a smaller N) to get E_1 and F_1. Now since the entire construction need only be iterated L-times, where $(1-2^{-d})^L < \varepsilon$ in order for (2.9)″ to hold, the initial N can be chosen large enough and ε_1 small enough in order to retain the necessary almost invariance for L-steps. We put finally $A = E \cup E_1 \cup \cdots \cup E_L$ and thus complete the proof. □

PROOF OF LEMMA 1. Pick ε_1 small and let A_0 be given by Lemma 2 for this ε_1. Put

$$B = A_0 \cap \psi^{-3} A_0 \tag{2.23}$$

If the measure of B is less than $\varepsilon/3$, i.e., is sufficiently small, then we can use A_0 directly to obtain the desired A. Bearing in mind then that B is not small, we build a skyscraper on B of $2k+1$ levels, i.e., set $B_1 = B$,

$$B_{i+1} = \psi(B_i) \cap (A_0 \setminus B_1)$$

Each even level is mapped entirely onto the next one, as is B_1, while from the odd levels $1 < i < 2k+1$, part is mapped by ψ into B_1, and the rest is mapped onto B_{i+1}. All levels B_i, $1 \leqq i \leqq 2k+1$ are easily seen to be $0(\varepsilon_1)$-invariant under ϕ_j, $2 \leqq j \leqq d$, (if necessary a smaller ε_1 may be taken to ensure this last fact). Now apply Lemma 1 again to B with $\phi_2, \cdots, \phi_d$ and the mapping $\bar{\psi}$ induced by ψ on B. The fact that ψ and ϕ_j's commute on B is seen, as in the argument preceding (2.16). We obtain a set $E \subset B$, which is $0(\varepsilon_1)$-invariant under ϕ_j, $2 \leqq j \leqq d$, $\bar{\psi}(E) \cap E = \varnothing$, $m(E) > \frac{1}{4} m(B)$. Set

$$G = \bigcup_{n=0}^{k} (\psi^{2n}(E) \cap B_{2n+1}) \cup \bigcup_{n=1}^{k} (\psi^{2n-1}(B \setminus E) \cap B_{2n}) \tag{2.24}$$

Now G has the same properties as A_0 with $\varepsilon_2 > \varepsilon_1$ which can be estimated as a function of ε_1 and k and which can be made arbitrarily small by choosing ε_1 small enough. The improvement that G has over A_0 is that

$$m(X \setminus G \cup \psi(G)) < \tfrac{3}{4} m(X \setminus (A_0 \cup \psi(A_0))) \tag{2.25}$$

The reader is advised to sketch a diagram to see that (2.24) leads to (2.25). We must repeat this L times where $(\frac{3}{4})^L < \varepsilon$, and thus the initial ε_1 can be taken small enough to allow L iterations and obtain the desired A. □

Proof of Theorem 1. We will use Lemma 1 to prove the theorem when $l_i = 2^k$, $1 \leqq i \leqq d$ for arbitrary k. For arbitrary l, the theorem follows by choosing k sufficiently large and then collecting, together with the base E, all $\phi^{l'}E$ with $0 \leqq l'_i \leqq 2^k - 1$ and $l'_i = n_i l_i$, $n \in Z^d$, $1 \leqq i \leqq d$. Suppose then that k and $\varepsilon > 0$ are given. Choose ε_1 sufficiently small so that

(i) if B is $k\varepsilon_1$ invariant under ϕ_j, $2 \leqq j \leqq d$, then the theorem for $(d-1)$, and $2^k - 1$ can be applied to B.

(ii) $\varepsilon_1 \ll \varepsilon \cdot k^{-2} \cdot 2^{-k}$.

Apply Lemma 1, k times with $\varepsilon_1, \phi_2, \cdots, \phi_d$ and ϕ_1 replaced by $\phi_1^{2^m - 1}$, $1 \leqq m \leqq k$, and obtain sets A_m which are ε_1-invariant under ϕ_j, $2 \leqq j \leqq d$, and satisfy

$n(A_m) > \frac{1}{2} - \varepsilon_1$, $\phi_1^{2^{m-1}}(A_m) \cap A_m = \varnothing$. Set $B = \bigcap_{m=1}^k A_m$. One checks that $\phi_1^n(B)$ are disjoint for $n = 0, \cdots, 2^k - 1$,

$$m\left(\bigcup_{n=0}^{2^k-1} \phi_1^n(B)\right) > 1 - k \cdot 2^k \varepsilon_1, \tag{2.26}$$

and B is $k\varepsilon_1$-invariant under ϕ_j, $2 \leqq j \leqq d$.

Apply the theorem for $d - 1$ to B, to get a subset $C \subset B$ satisfying

$$\phi^{\bar{l}} C \text{ are disjoint if } \bar{l} \in R_{2^k - 1}, \tag{2.27}$$

$$m\left(\bigcup_{\bar{l} \in R2^k - 1} \phi^{\bar{l}} C\right) > (1 - \varepsilon/2) m(B) \tag{2.28}$$

$$C \text{ is } \tfrac{1}{2}\,\varepsilon\text{-invariant under } \phi^{\,2^k}_{j},\ 2 \leqq j \leqq d. \tag{2.29}$$

From (2.26) and what precedes it and (2.27) to (2.29), one sees that C satisfies the requirements (2.8) to (2.10) for $l_i = 2^k - 1$, $1 \leqq i \leqq d$ and ε. □

For applications especially in the course of the proof of the isomorphism theorem, the following refinement of Theorem 1 is very useful. It can be used to replace mixing by non-periodicity in some of Ornstein's arguments in [5]–[7]. We shall use the notation $\alpha = \{A_1, \cdots, A_r\}$ for partitions, dist $\alpha = (m(A_1), \cdots, m(A_r))$ and if E is any subset

$$\operatorname{dist} \alpha \,\big|\, E = \left(\frac{m(A_r \cap E)}{m(E)}, \cdots \frac{m(A_r \cap E)}{m(E)}\right).$$

THEOREM 2. *If Φ is non-periodic and l, $\varepsilon > 0$ and a finite partition α are given, then there is a set E such that* (2.8) *to* (2.10) *hold and in addition*

$$\operatorname{dist} \alpha \,\big|\, \phi^{l'} E = \operatorname{dist} \alpha, \qquad l' \in R_l. \tag{2.30}$$

PROOF. By replacing α by $\bigvee_{l' \in R_l} \phi^{-l'} \alpha$, it suffices to show that (2.30) holds for $l' = 0$. For the reader's convenience, we will write out the proof for $d = 1$ since there is no change in the argument for $d > 1$. So suppose that ε, N and α are given. Choose n so that n/N is an integer and

$$N/n < \varepsilon/8, \tag{2.31}$$

Apply Theorem 1 with n and $\varepsilon/8$ to get a set E with $\{\phi^i E\}_{i=0}^n$ disjoint and $m(\bigcup_{i=0}^n \phi^i E) > 1 - \varepsilon/8$. It follows that if $p = \operatorname{dist} \alpha \,\big|\, \bigcup_{i=0}^n \phi_i E$

$$\| p - \operatorname{dist} \alpha \|_1 < \varepsilon/4. \tag{2.32}$$

Now for each $x \in \phi^j E$ $(j = 0, \cdots N - 1)$, let $v(x)$ represent the empirical distribution according to α of $\{x, \phi^N x, \phi^{2N} \cdots x_j \cdots \phi^{kN} x\}$ where $k = n/N - 1$. This means that the ith component of $v(x)$ is the number of elements of $\{x, \phi^N x, \cdots, \phi^{kN} x\}$ that belong to A_i, divided by n/N. It follows then that

$$\sum_{j=0}^{N-1} \int_{\phi^j E} v(x) m(dx) = p. \tag{2.33}$$

We shall now find subsets $B_j \subset \phi^j E$ so that $\phi^{-j} B_j$ partition E and

$$\left\| \sum_{j=0}^{N-1} \int_{B_j} v(x) m(dx) - p/N \right\|_1 < \varepsilon/4N. \tag{2.34}$$

To this end, simply partition E into sets C so that the variation of $v(x)$ on $C, \phi C, \cdots, \phi^{N-1} C$ is less than $\varepsilon/8N$. Since ϕ is non-periodic, the measure m is non-atomic so that C may be divided into N sets of equal measure $C_0, C_1, \cdots, C_{N-1}$. Assign to B_j the set $\phi^j C_j$ and then (2.34) is readily seen to hold. Set now $F_0 = \bigcup_{j=0}^{N-1} B_j$, $F_m = \phi^{N \cdot m} F_0$, $0 < m < n/N$, and $F = \bigcup_{0 \leq m < n/N} F_m$. From (2.34) and the definition of $v(x)$ we see that

$$\| \operatorname{dist} \alpha \,|\, F - p \|_1 < \varepsilon/4, \tag{2.35}$$

$\{\phi^j F\}_{j=0}^{N-1}$ are disjoint and $m(\bigcup_{j=0}^{N-1}) \phi^i F) > 1 - \varepsilon/2$. Now by (2.32) and (2.35), we can remove from F a set H of measure less than $\varepsilon/2 \cdot m(F)$ so that if $G = F \setminus H$

$$\operatorname{dist} \alpha \,|\, G = \operatorname{dist} \alpha. \tag{2.36}$$

Clearly $\{\phi^j G\}_{j=0}^{N-1}$ are disjoint and $m(\bigcup_0^{N-1} \phi^j G) > 1 - \varepsilon$. This completes the proof. □

We close this section by remarking that it is now a routine task to verify that the Halmos-Rokhlin theorem, that in general Φ is or is not mixing, holds for $d > 1$. Needless to say the approximation of non-periodic transformations by strongly periodic transformations is also a straightforward consequence of Theorem 1. We leave to the reader the exercise of carrying these things out.

3. Entropy

This section is devoted to extending the definitions and basic properties of the entropy of a single transformation to a group Φ. Such extensions have been mentioned before (see [3] and [10]) and no novelties arise in the fundamentals. However since difficulties crop up in extending the Shannon-McMillan-Breiman theorem, some details are given here.

In addition to the coordinatewise partial ordering on Z^d, we shall need the lexicographical total ordering. Using the notation $\hat{l} = (l_2, l_3, \cdots, l_d) \in Z^{d-1}$, this ordering may be inductively defined by

$$l \prec l' \text{ if } \begin{cases} l_1 < l_1' l, \\ l_1 = \hat{l}_1' \text{ and } \prec \hat{l}' \end{cases} \tag{3.1}$$

for $d = 1$, $l \prec l'$ is the same as $l < l'$, The indices corresponding to the *past* will be denoted by P_d, where

$$P_d = \{l \in Z^d \colon l \prec 0\}. \tag{3.2}$$

If α is a partition, $Q \subset Z^d$ any subset and Φ is given,

$$\alpha(Q) = \bigvee_{l \in Q} \Phi^l \alpha \tag{3.3}$$

Strictly speaking we should be writing $\alpha_\Phi(Q)$ to indicate the dependence on Φ. In case Q is infinite, the symbol for the join on the right hand side of (3.3) will mean the smallest σ-algebra generated by $\{\phi^l \alpha\}_{l \in Q}$.

DEFINITION. The *entropy* of Φ with respect to α is defined by:

$$h(\Phi, \alpha) = H(\alpha \,|\, \alpha(P_d)). \tag{3.4}$$

The right hand side is the usual conditional entropy, its properties as set forth in [9], for example, will be used without comment. The number of elements of Q will be denoted by $|Q|$. To state the basic computational result for $h(\Phi, \alpha)$ we need one more piece of notation for a set of partitions whose limit is P_d. If $(a, b) = (a_1, \cdots, a_d, b_2, \cdots, b_d) \in Z_+^{2d+1}$ then set

$$P(a, b) = \{l \in P_d \colon -a_1 \leqq l_1, -\hat{a} \leqq \hat{l} \leqq b\}. \tag{3.5}$$

If $a_i = b_i = n$ for all n_i write simply $P(n)$.

THEOREM 3. (i) $h(\Phi, \alpha) = \lim_{n \to \infty} H(\alpha \,|\, \alpha(P(n)))$

(ii) $h(\Phi, \alpha) = \lim_{(a,b) \to \infty} H(\alpha \,|\, P(a, b))$

(iii) $h(\Phi, \alpha) = \lim_{l \to \infty} 1 / |R_l| \; H(\alpha(R_l))$.

PROOF. (i) is just the usual fact that the conditional entropy of an increasing sequence of σ-algebras is continuous in the limit. The second statement (ii) then follows from (i) and the monotonicity of $H(\alpha \,|\, \cdot\,)$ as a function of the conditioning σ-algebra. To apply the standard computation and prove (iii) define

$$Q_{l'}^{l} = \{l'' \in R_l \colon l'' \prec l'\} \tag{3.6}$$

and observe that

$$Q_{l'}^{l} - l' = P(l', \hat{l} - \hat{l}'). \tag{3.7}$$

Since $\prec$ is a total ordering we have

$$H(\alpha(R_l)) = \sum_{l' \in R} H(\phi^{l'} \alpha \mid \alpha(Q_{l'}^{l})) \tag{3.8}$$

where $\alpha(\emptyset)$ is the trivial partition. Now since Φ is measure-preserving we have

$$H(\phi^{l'}\alpha \mid \alpha(Q_{l'}^{l})) = H(\alpha \mid \phi^{-l'}\alpha(Q_{l'}^{l})) = H(\alpha \mid P(l', \hat{l} - \hat{l}') \tag{3.9}$$

by (3.7).

For any $(n, n, \cdots n) \in Z^{2d-1}$, l can be chosen sufficiently large so that for most $l' \in R_l$, $(l', \hat{l} - \hat{l}') \geqq (n, n, \cdots n)$. Thus dividing both sides of (3.8) by $|R_l|$ and using (3.9) we obtain (iii). □

Having established Theorem 3, there is no difficulty in defining the *entropy* of Φ by means of

$$h(\Phi) = \sup_{\alpha, H(\alpha) < \infty} h(\Phi, \alpha), \tag{3.10}$$

and obtaining the basic results that if α is a generator (i.e. $\alpha(Z^d) = \mathscr{B}$) then $h(\Phi) = h(\Phi, \alpha)$ and that if $\alpha_n \uparrow \mathscr{B}$ then $h(\Phi, \alpha_n) \uparrow h(\Phi)$. In addition, having at our disposal the Kakutani-Rokhlin theorem we can prove the existence of generators with finite entropy in case $h(\Phi) < \infty$. Concerning the existence of finite generators see below.

There are extensions of the notions of K-automorphisms to the situation of a group Φ, and an analogue of the Pinsker-Rokhlin-Sinai theorem identifying them with those Φ such that $h(\Phi, \alpha) > 0$ for any non-trivial α. We leave these side roads however, and go on to developing the main theme.

4. The Shannon-McMillan theorem

The most important property of $h(\Phi, \alpha)$ is that in case Φ is ergodic it measures the typical size of an atom in $R_l(\alpha)$. In the case of Z^1, this is made precise by the Shannon-McMillan-Breiman theorem. We are unable at the present to give a analogue of the Breiman contribution, namely the a.e. convergence, because of the difficulties that arise in the pointwise behavior of martingales over general directed families of σ-algebras. However the mean convergence is not very difficult to obtain and this is all that is really needed in most applications.

Recall that Φ is said to be *ergodic* if only the constants satisfy

$$f(x) = f(\phi^l x), \text{ for all } l \in Z^d. \tag{4.1}$$

A word of caution—in case $d = 1$, if m is assumed non-atomic then ergodicity implies non-periodicity. This is *not* the case for groups Φ. For simple counter-examples let $d = 2$ and let $\phi_1 = \phi_2$ be some ergodic transformation. Clearly the *group* is ergodic but not non-periodic, since, for example, $\phi_1\phi_2^{-1} = \text{identity}$. If ϕ^l is not the identity for $l \neq 0$ then ergodicity does imply non-periodicity. Another way of saying this is if Φ is isomorphic to Z^d then ergodicity implies non-periodicity.

For ergodic groups Φ, the mean ergodic theorem takes the form:

THEOREM 4. *If $f \in L_p$ $(1 \leqq p \leqq \infty)$ and Φ is ergodic then*

$$\lim_{l \to \infty} \Big\| 1/|R_l| \sum_{l' \in R_l} f(\phi^{l'} x) - \int f\,dm \Big\|_p = 0. \tag{4.2}$$

This result is quite general and doesn't depend upon the fact that the ϕ_i commute (see [1, chapter 8]. We shall need the notion of information in the proof of the Shannon-McMillan theorem. Recall that the *information* of one partition α with respect to another, say β is defined by

$$I(\alpha|\beta) = -\sum_i \mathscr{I}_{A_i} \log m(A_i|\beta) \tag{4.3}$$

where $\mathscr{I}_A$ is the indicator function of A and $m(A|\beta) = E\,\{\mathscr{I}_A|\beta\}$ with $E\{f|\beta\}$ the usual conditional expectation operator. Note that $\int I(\alpha|\beta)\,dm = H(\alpha|\beta)$.

THEOREM 5 (S-M). *If Φ is ergodic and $H(\alpha) < +\infty$ then*

$$\lim_{l \to \infty} \big\| 1/|R_l|\, I(\alpha(R_l)) - h(\Phi, \alpha) \big\|_1 = 0. \tag{4.4}$$

PROOF. We shall use the notation of the proof of Theorem 3 (see (3.6) to (3.8). We have

$$I(\alpha(R_l)) = \sum_{l' \in R_l} I(\phi^{l'}\alpha | \alpha(Q^l_{l'})). \tag{4.5}$$

For the terms in the right-hand side we have

$$I(\phi^{l'}\alpha | \alpha(Q^l_{l'})) \cdot (x) = I(\alpha | \phi^{-l'} \alpha(Q^l_{l'})) \cdot (\phi^{-l'} x) \equiv f_{l,l'}(\phi^{-l'} x). \tag{4.6}$$

Now we need to know that $f_{l,l'} \to f = I(\alpha | \alpha(P_d))$ in L^1-norm. To see this, it suffices to observe that for any cofinal totally ordered subset of indices, this norm convergence holds, since then we are in the situation of an increasing sequence of algebras and the usual theory applies. Hence

$$\lim_{l,l' \to \infty} \|f_{l,l'} - f\|_1 = 0. \tag{4.7}$$

Thus we have the estimates:

$$
\begin{aligned}
&\left\| 1/|R_l|\ I(\alpha(R_l)) - h(\Phi,\alpha) \right\|_1 \\
&= \Big\| 1/|R_l| \sum_{l' \in R_l} f_{l,l'}(\phi^{-l'}x) - h(\Phi,\alpha) \Big\|_1 \\
&\leqq \Big\| 1/|R_l| \sum_{l' \in R_l} f(\phi^{-l'}x) - h(\Phi,\alpha) \Big\|_1 \\
&+ 1/|R_l| \sum_{l' \in R_l} \| f_{l',l} - f \|_1
\end{aligned}
\tag{4.8}
$$

The first term tends to zero by the mean ergodic theorem (Theorem 4) whereas the second term tends to zero by (4.7). Note that $\int I(\alpha \mid \alpha(P_d))\, dm = H(\alpha \mid \alpha(P_d)) = h(\Phi,\alpha)$. □

The following corollary is an immediate consequence:

COROLLARY. *If Φ is ergodic and $H(\alpha) < \infty$, then given $\varepsilon > 0$, there is an l_0, such that for $l \geqq l_0$ ε-almost every atom A of $\alpha(R_l)$ satisfies*

$$\exp[-|R_l|(h(\Phi,\alpha)+\varepsilon)] \leqq m(A) \leqq \exp[-|R_l|(h(\Phi,\alpha)-\varepsilon)], \tag{4.9}$$

and if N is the number of atoms of $\alpha(R_l)$ that satisfy (4.17) *then*

$$\exp[|R_l|(h(\Phi,\alpha)-\varepsilon)] \leqq N \leqq \exp[|R_l|(h(\Phi,\alpha)+\varepsilon)]. \tag{4.10}$$

5. Isomorphism results

A partition α is said to be *independent* for Φ if $\{\phi^l\alpha\}_{l \in Z^d}$ form an independent family of partitions. If in addition, α is a generator for Φ, then we say that Φ is Bernoullian. Two groups Φ and Ψ are *isomorphic* if there is a one-to-one measure-preserving mapping between the spaces on which they operate which takes the action of Φ into the action of Ψ.

THEOREM. *If Φ and Ψ are Bernoullian, then Φ and Ψ are isomorphic if and only if $h(\Phi) = h(\Psi)$.*

For the proof of the theorem one simply rewrites the proof of D. Ornstein [5], making use of the Kakutani-Rokhlin theorem of section 2 and the S-M theorem of section 4 in the appropriate places. A number of minor modifications, which may be left to the reader, need to be made. By a suitable definition of *finitely determined* the results of [7] may be extended to yield, for example, the fact that any factor of a Bernoullian group is Bernoullian.

The original proof of Krieger [4] that transformations with finite entropy have finite generators does not seem to extend to groups Φ, but the techniques of

D. Ornstein ([5] to [7]) allow an alternate proof which can be carried out even in the case of a general group. For details of this kind of approach in case $d = 1$, see [11, chapter 9].

Postscript. After submitting this paper, we learned of the researches of J. P. Conze, "Entropie d'un Groupe Abelien de Transformations" and J. P. Thouvenot, "Convergence en Moyenne de l'Information pour l'Action de Z^2". Taken together, these papers overlap a great many of the results that we obtained here. It should be noted that in contrast to Conze's briefer proof, our proof of Theorem 1 is constructive.

References

1. N. Dunford and J. T. Schwartz, *Linear Operators I*, (Chapter 8), N. Y., 1958.

2. P. Halmos, *Lectures on Ergodic Theory*, The Mathematical Society of Japan, Tokyo, 1956.

3. A. A. Kirillov, *Dynamical systems, factors and group representations*, Uspehi Mat. Nauk **22** (1967), 67–80.

4. W. Krieger, *On entropy and generators of measure-preserving transformations*, TAMS **149** (1970), 453–464.

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

6. D. S. Ornstein, *Two Bernoulli shifts with infinite entropy are isomorphic*, Advances in Math. **6** (1970), 339–348.

7. D. S. Ornstein, *Factors of Bernoulli shifts are Bernoulli shifts*, Advances in Math. **5** (1970), 349–364.

8. V. A. Rokhlin, *A general measure-preserving transformation is not mixing*, Dokl. Akad. Nauk SSSR **60** (1948), 349–351.

9. V. A. Rokhlin, *Lectures on the entropy theory of transformations with invariant measure*, Uspehi Math. Nauk **22** (1967), 3–56.

10. D. Ruelle, *Statistical Mechanics*, N.Y., 1969.

11. M. Smorodinsky, *Ergodic Theory, Entropy*, Lecture Notes in Mathematics, No. 214, 1971.

12. B. Weiss, *The isomorphism problem in ergodic theory*, BAMS (to appear).

THE HEBREW UNIVERSITY OF JERUSALEM