PACIFIC JOURNAL OF MATHEMATICS
Vol. 43, No. 3, 1972

ENTROPY AND APPROXIMATION OF MEASURE PRESERVING TRANSFORMATIONS

T. SCHWARTZBAUER

The relationship between the entropy and rate of approximation of an automorphism was first discovered by A. B. Katok. He defined for each automorphism T an invariant $c(T)$ which depends only on the rate of approximation of T and then proved that $h(T) \leqq c(T) \leqq 2h(T)$ for any ergodic automorphism T, where $h(T)$ denotes the entropy of T. The proof which he gave that $c(T) \leqq 2h(T)$ can be generalized to the case where T is not ergodic, and it was asserted further that $c(T) = 2h(T)$ if T were ergodic, but the proof given was incomplete.

In this paper these results are generalized to the case of an arbitrary automorphism T.

We will extend the result $c(T) = 2h(T)$ to an arbitrary automorphism T by showing that $2h(T) \leqq c(T)$ for any automorphism.

In doing so we will apply the methods of approximation developed in [1], [5], and [6]. The general type of approximation introduced in [6], in particular, will allow us to reduce the case of a non-ergodic automorphism to the consideration of ergodic automorphisms.

2. Preliminary Definitions. We let $(X, \mathscr{F}, \mu)$ denote a normalized non-atomic Lebesgue space, that is, a measure space isomorphic to the unit interval, the Lebesgue measurable sets, and Lebesgue measure. An automorphism T will be an invertible transformation of X onto X such that $A \in \mathscr{F}$ if and only if $TA \in \mathscr{F}$, and for any $A \in \mathscr{F}$, we have

$$\mu(A) = \mu(TA) = \mu(T^{-1}A) .$$

If the equation $TA = A$ implies that $\mu(A) = 0$ or $\mu(A) = 1$, whenever $A \in \mathscr{F}$, then the automorphism T is said to be ergodic. A set $A \in \mathscr{F}$ is said to be invariant under T, or simply invariant if $TA = A$.

Henceforth all sets to which we refer will be assumed to be in $\mathscr{F}$.

A collection ξ of pairwise disjoint measurable sets whose union is X will be called *a partition*. If η and ξ are partitions and if every set in η is a union of sets in ξ we will write $\eta \leqq \xi$ and say that ξ is a finer partition than η, and that η is a coarser partition than ξ. If $\{\xi_n\}$ is a sequence of partitions, then the partition $\xi = \bigvee_{n=1}^{\infty} \xi_n$ is the coarsest partition which is finer than each ξ_n.

A partition ξ will be called a measurable partition if there exists a sequence $\{\xi_n\}$ of finite partitions such that $\xi_n \leqq \xi_{n+1}$ for each positive integer n, and $\xi = \bigvee_{n=1}^{\infty} \xi_n$.

If ξ is a partition then *the factor space of X with respect to ξ* is the measure space $(X/\xi, \mathscr{F}_\xi, \mu_\xi)$ where the points of X/ξ are the sets in ξ, where $\mathscr{F}_\xi$ consists of all measurable sets which are unions of sets in ξ, and where $\mu_\xi(A) = \mu(A)$ for any $A \in \mathscr{F}_\xi$.

If ξ is a finite or countable partition and if $A \in \mathscr{F}$ then among the sets in $\mathscr{F}_\xi$ there is at least one whose symmetric difference with A has minimal measure. We denote any one of these sets by $A(\xi)$. One may choose $A(\xi) = \bigcup B$ where $B \in \xi$ and $\mu(A \cap B) \geqq \mu(B)/2$.

If $\{\xi_n\}$ is a sequence of finite or countable partitions and if

$$\lim_{n \to \infty} (A \varDelta A(\xi_n)) = 0$$

for each $A \in \mathscr{F}$ then we write $\xi_n \to \varepsilon$, as $n \to \infty$. ($A \varDelta B$ denotes the symmetric difference of the sets A and B.)

DEFINITION 2.1. Let $\{f(n)\}$ be a monotonic sequence of positive numbers such that $\lim_{n\to\infty} f(n) = 0$. We say that the automorphism *T admits an approximation with speed $f(n)$* if for each positive integer n there exists a partition $\xi_n = \{C_i(n), i = 1, 2, \cdots, q(n)\}$ and an invertible mapping T_n from X/ξ_n onto X/ξ_n such that if we let $\tilde{C}_i(n) = T_n C_i(n)$ then

1. $\xi_n \to \varepsilon$ as $n \to \infty$, and
2. $\sum_{i=1}^{q(n)} \mu(TC_i(n) \cap \tilde{C}_i'(n)) < f(q(n))$

where $\tilde{C}_1'(n)$ denotes the complement of the set $\tilde{C}_i(n)$ with respect to the whole space.

If T_n is a cyclic mapping then, by reordering the elements of ξ_n if necessary, we may assume that $\tilde{C}_i(n) = C_{i+1}(n)$ for $i = 1, 2, \cdots, q(n)$ where $C_{q(n)+1}(n) = C_1(n)$. In this case the second condition above takes the form $\sum_{i=1}^{q(n)} \mu(TC_i(n) \cap (C_{i+1}'(n)) < f(q(n))$, and we say that *$T$ admits a cyclic approximation with speed $f(n)$*.

If T_n is measure preserving for each positive integer n then we say that *T admits a measure preserving approximation with speed $f(n)$*.

This particular type of approximation was introduced in [6], and a discussion of the properties of an automorphism which admits this kind of approximation can be found in that reference.

3. Approximation and invariance.

THEOREM. *If* $(X, \mathscr{F}, \mu)$ *is a Lebesgue space and if* β *is a measurable partition of* X *then*

(1) *the factor space* $(X/\beta, \mathscr{F}_\beta, \mu_\beta)$ *is a Lebesgue space,*

(2) *for almost all* $B \in X/\beta$ *there exists a sigma field* $\mathscr{F}_B$ *of subsets of* B *and a measure* μ_B *such that* $(B, \mathscr{F}_B, \mu_B)$ *is a Lebesgue space, and*

(3) *for any set* $A \in \mathscr{F}$ *we have* $A \cap B \in \mathscr{F}_B$ *for almost all* $B \in X/\beta$. *The function* $\mu_B(A \cap B)$ *is a measurable function on* X/β *and*

$$\mu(A) = \int_{X/\beta} \mu_B(A \cap B)\, d\mu_\beta(B) .$$

This theorem, as well as any properties of measurable partitions or entropy of which we may make use, can be found in [4].

THEOREM 3.1. [6] *If the automorphism* T *admits an approximation with speed* $f(n)$ *and if* β *is a measurable partition such that every set* $B \in \beta$ *is invariant under* T, *then* T *restricted to* B, *denoted by* T_B, *admits an approximation with speed* $f_B(n)$ *for almost all* $B \in X/\beta$ *where*

$$\int_{X/\beta} f_B(n) d\mu_B(n) < f(n) .$$

4. Entropy and approximation of ergodic automorphisms.

We now discuss the relationship between the entropy of an ergodic automorphism and its speed of approximation. In doing so, we will rely heavily on the proofs of the analogous results obtained by Katok in [2] in the case of a measure preserving approximation.

DEFINITION 4.1. If T is an automorphism, then $c(T)$ will denote the greatest lower bound of the set of numbers $2c$ such that T admits a measure preserving approximation with speed $c/(\ln n)$. If T admits no such approximation for any number c, then $c(T) = \infty$.

(The type of approximation defined in [2] uses $\mu(TC_i(n) \Delta \widetilde{C}_i(n))$ where we have used $\mu(TC_i(n) \cap \widetilde{C}'_i(n))$. The effect of this is to multiply all speeds by a factor of two as is discussed in reference [5]. This accounts for the factor of two in the above definition.)

DEFINITION 4.2. If T is an automorphism, then $b(T)$ will denote the greatest lower bound of the set of numbers $2b$ such that T admits an approximation with speed $b/(\ln n)$. If T admits no such approximation for any number b, then we set $b(T) = \infty$.

The definitions and properties of entropy and factor automorphisms of which we will make use can be found in [4].

THEOREM 4.1. *If T is an ergodic automorphism then*

$$h(T) \leqq b(T) \leqq c(T) .$$

It is obvious from the definitions of the invariants $b(T)$ and $c(T)$ that $b(T) \leqq c(T)$ for any automorphism T, so that we have only to prove the first inequality. The proof we will give will be a modification of the proof given in [2] to show that $h(T) \leqq c(T)$ for any ergodic automorphism T. We prefix two lemmas to the proof of the theorem in which we will require the use of Sinai's weak isomorphism theorem:

THEOREM (Sinai [7]). *If T is an ergodic automorphism and if S is a Bernoulli automorphism such that $h(S) \leqq h(T)$, then there exists a factor automorphism of the automorphism T which is isomorphic to the automorphism S.*

Since the result of Theorem 4.1 is trivially true if $h(T) = 0$, we assume in the following that $0 < h(T) < \infty$. The case $h(T) = \infty$ will be considered subsequently.

LEMMA 4.1. *If an ergodic automorphism T admits an approximation with speed $b/(\ln n)$, then*

$$b \geqq \max_{0<\alpha<(p(T))/(p(T)+1)} 2\alpha(h(T) + \alpha \ln \alpha + (1-\alpha)\ln(1-\alpha) - \alpha \ln p(T)) ,$$

where $p(T)$ denotes the greatest integer which is strictly less than $e^{h(T)}$.

REMARKS. By an application of the weak isomorphism theorem we see that if S is a Bernoulli automorphism with a state space of $p(T)+1$ elements and probability distribution $\pi_1, \pi_2, \cdots, \pi_{p(T)+1}$ such that $h(S) = h(T)$ then there exists a partition η of X,

$$\eta = \{\beta_1, \beta_2, \cdots, \beta_{p(T)+1}\} ,$$

such that

$$\mu(\beta_{i_0} \cap T^{-1}\beta_{i_1} \cap \cdots \cap T^{-k}\beta_{i_k}) = \pi_{i_0}\pi_{i_1} \cdots \pi_{i_k}$$

for any integer k and any choice of indices $i_0, i_1, \cdots, i_k$. Since $\xi_n \to \varepsilon$ we may assume that if $\delta > 0$ then for n sufficiently large there exists a partition $\eta^{(n)} = \{\gamma_1^{(n)}, \gamma_2^{(n)}, \cdots, \gamma_{p(T)+1}^{(n)}\}$ such that $\eta^{(n)} \leqq \xi_n$ and such that if $D_n = \bigcup_{i=1}^{p(T)+1}(\gamma_i(n)\Delta\beta_i)$ then $\mu(D_n) < \delta$.

We consider also the partition $\eta_k = \bigvee_{l=0}^{k} T^{-l}\eta$ and we let B_θ^k denote the union of those elements of η_k whose measures exceed $e^{-k(h(T)-\theta)}$. Since T is ergodic we may apply the Shannon-McMillan-Breiman theorem to obtain $\lim_{k\to\infty} \mu(B_\theta^k) = 0$ for any $\theta > 0$. We fix $\theta > 0$ and given $\delta > 0$ we choose k and n sufficiently large in what follows so that $\mu(B_\theta^k) < \delta$ and $\mu(D_n) < \delta$.

Before proving Lemma 4.1, we state and prove another auxiliary lemma.

LEMMA 4.2. *If an ergodic automorphism T admits an approximation with speed $b/(\ln n)$, then for n sufficienty large there exists an element $C^*(n) \in \xi_n$ such that $\mu(C^*(n)) > 1/(q(n)\ln q(n))$ and*

$$\frac{k}{\sqrt{\delta}}\mu\Big(C^*(n) \cap B_\theta^k) + \sum_{l=0}^{k-1} \mu(T_n^l C^*(n) \cap D_n)\Big) + \sum_{l=0}^{k-1} \mu(T^l C^*(n) \cap (T_n^l C^*(n))') \leqq (2k\sqrt{\delta}(1+\delta) + \frac{b(1+\delta)k(k+1)}{2\ln q(n)}\mu(C^*(n)) \,.$$

Proof. Consider the space $X \times Y$ where $Y = \{y_1, \cdots, y_k\}$ is the discrete space of k elements, each element having measure one. Let $\hat{\mu}$ indicate the product measure on $X \times Y$ We define the function $G(x, y)$ on $X \times Y$ so that

$$G(x, y_l) = \frac{\frac{1}{\sqrt{\delta}}\mu(C_i(n) \cap B_\theta^k) + \mu(T_n^l C_i(n) \cap D_n) + \mu(T^l C_i(n) \cap (T_n^l C_i(n))')}{\mu(C_i(n))}$$

if and only if $x \in C_i(n)$. It is easily seen that

$$\int_{X\times Y} G(x, y)d\hat{\mu} = \frac{k\mu(B_\theta^k)}{\sqrt{\delta}} + \sum_{l=0}^{k-1}\Big(\sum_{i=1}^{q(n)} \mu(T_n^l C_i(n) \cap D_n)\Big) + \sum_{l=0}^{k-1}\Big(\sum_{i=1}^{q(n)} \mu(T^l C_i(n) \cap (T_n^l C_i(n))')\Big) \,.$$

Since

$$\sum_{i=1}^{q(n)} \mu(T_n^l C_i(n) \cap D_n) = \mu(D_n) < \delta$$

and

$$\sum_{i=1}^{q(n)} \mu(T^l C_i(n) \cap (T_n^l C_i(n))') \leqq l \sum_{i=1}^{q(n)} \mu(T C_i(n) \cap (T_n C_i(n))') \leqq \frac{lb}{\ln q(n)}$$

we have

$$\int_{X\times Y} G(x, y)d\hat{\mu} < k\sqrt{\delta} + k\delta + \frac{bk(k+1)}{2\ln q(n)} < 2k\sqrt{\delta} + \frac{bk(k+1)}{2\ln q(n)} .$$

If we let F_n denote the union of those elements of ξ_n whose measures are not greater than $1/(q(n)\ln q(n))$ then $\mu(F_n) \leqq 1/(\ln q(n))$. We define $E_n = X - F_n$. If we let $\varepsilon_n = \hat{\mu}(F_n \times Y)$ then $\lim_{n\to\infty} \varepsilon_n = \lim_{n\to\infty} \mu(F_n) = 0$. Now there must exist an element $C^*(n) \subset E_n$ such that

$$\int_{C^*(n)\times Y} G(x, y)\, d\hat{\mu} \leqq \left(2k\sqrt{\delta} + \frac{bk(k+1)}{2\ln q(n)}\right)\frac{\mu(C^*(n))}{1-\varepsilon_n}$$

for otherwise we would have

$$\begin{aligned}\int_{X\times Y} G(x, y)\, d\hat{\mu} &\geqq \int_{E_n\times Y} G(x, y)\, d\hat{\mu} > \left(2k\sqrt{\delta} + \frac{bk(k+1)}{2\ln q(n)}\right)\frac{\hat{\mu}(E_n)}{1-\varepsilon_n}\\ &\geqq 2k\sqrt{\delta} + \frac{b\,k(k+1)}{2\ln q(n)}\end{aligned}$$

which we know to be false.

The lemma follows by noting that $C^*(n) \subset E_n$ implies that

$$\mu(C^*(n)) > \frac{1}{q(n)\ln q(n)} ,$$

that n sufficiently large will insure that $1/(1-\varepsilon_n) < 1 + \delta$, and that

$$\begin{aligned}\int_{C^*(n)\times Y} G(x, y)d\hat{\mu} &= \frac{k}{\sqrt{\delta}}\mu(C^*(n) \cap B_\delta^k) + \sum_{l=0}^{k-1}\mu(T_n^l C^*(n) \cap D_n)\\ &\quad + \sum_{l=0}^{k-1}\mu(T^l C^*(n) \cap (T_n^l C^*(n)') .\end{aligned}$$

Proof of Lemma 4.1. We let $\gamma_{j_l}^{(n)}$ denote the element of the partition $\eta^{(n)}$ which contains $T_n^l C^*(n)$ for $l = 0, 1, \cdots, k-1$ and we let β_{j_l} denote the corresponding element of the partition η.

Clearly

$$\mu(T^l C^*(n) \cap T_n^l C^*(n)) = \mu(C^*(n)) - \mu(T^l C^*(n) \cap (T_n^l C^*(n))')$$

so that

$$\begin{aligned}\mu(T^l C^*(n) \cap \beta_{j_l}) &\geqq \mu(T^l C^*(n) \cap T_n^l C^*(n)) - \mu(T_n^l C^*(n) \cap \beta'_{j_l})\\ &\geqq \mu(C^*(n)) - \mu(T^l C^*(n) \cap (T_n^l C^*(n))')\\ &\quad - \mu(T_n^l C^*(n) \cap D_n) .\end{aligned}$$

Thus using the results of Lemma 4.2 we get

$$\sum_{l=0}^{k-1}\mu(T^l C^*(n) \cap \beta_{j_l}) \geqq k\mu(C^*(n)\left(1 + 2\sqrt{\delta}(1+\delta) + \frac{b(1+\delta)(k+1)}{2\ln q(n)}\right) .$$

Let M_m indicate the set of points $x \in C^*(n)$ such that $T^l x \notin \beta_l$ for at least m values of l, where $l = 0, 1, , \cdots, k-1$. It is easily seen that

$$m\mu(M_m) + \sum_{l=0}^{k-1} \mu(T^l C^*(n) \cap \beta_{j_l}) \leqq k\mu(C^*(n)) \ .$$

Combining this with the preceding inequality give us that

$$\mu(M_m) \leqq \frac{k}{m}\, \mu(C^*(n)) \left(2\sqrt{\delta}\,(1+\delta) + \frac{b(1+\delta)(k+1)}{2 \ln q(n)}\right) .$$

We can write $C^*(n)$ as the disjoint union of three sets:

$$C^*(n) \cap (B_\theta^k) \ , \quad M_m \cap (B_\theta^k)' \ ,$$

and N_m, where N_m is the complement with respect to $C^*(n)$ of the union of the two preceding sets. If follows from Lemma 4.2 that

$$\mu(C^*(n) \cap B_\theta^k) \leqq (2\delta(1+\delta) + \frac{b\sqrt{\delta}\,(1+\delta)(k+1)}{2 \ln q(n)} \mu(C^*(n))$$

so that we have only to estimate $\mu(N_m)$.

If we let $\hat{N}_m$ denote the set of points $x \in C^*(n)$ such that $T^l x \notin \beta_{j_l}$ for no more than m values of l, then $N_m \subset \hat{N}_m \cap (B_\theta^k)'$. Now this last set consists of elements of the partition η_k whose measures are not greater than $e^{-k(h(T)-\theta)}$. Determining the number of elements in $\hat{N}_m \cap (B_\theta^k)'$ is accomplished as follows. Let E be such an element, than E can be identified by telling in which element of

$$\eta = \{\beta_1, \beta_2, \cdots, \beta_{p(T)+1}\}$$

the set $T^l E$ is found for every value of $l = 0, 1, \cdots, k-1$. Thus each set E determines a sequence of length k whose l-th entry is the index of the element of the partition η in which $T^l E$ is found. Conversely every such sequence determines a possible element of η_k. In order for such an element to be in $\hat{N}_m \cap (B_\theta^k)'$ it is necessary that the given index j_l occurs as the l-th entry in the sequence for at least $k - m$ entries. The number of sequences in which the given indices occur in exactly $k - m$ places is $\binom{k}{k-m} (p(T))^m$. Assume that $m < (p(T))/(p(T)+1)k$ then if $0 < s \leqq m$ we have

$$\binom{k}{k-m+s}(p(T))^{m-s} < \binom{k}{k-m}(p(T))^m \ .$$

Thus the number of sequences in which we are interested does not exceed $m\binom{k}{k-m}(p(T))^m = m\binom{k}{m}(p(T))^m$. Therefore

$$\mu(N_m) \leqq e^{-k(h(T)-\theta)} m \binom{k}{m} (p(T))^m .$$

If we let α be determined by the relation $m = \alpha k$ and if $\nu \geqq 0$, then if we apply Stirling's formula we obtain for sufficiently large k and m that $\mu(N_{\alpha k}) < e^{-k(h(T)+\alpha \ln \alpha + (1-\alpha) \ln (1-\alpha) - \alpha \ln p(T) - \nu - \theta)}$.

If α is chosen so that

$$h(T) + \alpha \ln \alpha + (1 - \alpha) \ln (1 - \alpha) - \alpha \ln p(T) - \nu - \theta > 0$$

and k is then chosen so that

$$(1) \qquad k \geqq \frac{(1 + \delta) \ln q(n)}{h(T) + \alpha \ln \alpha + (1 - \alpha) \ln (1 - \alpha) - \alpha \ln p(T) - \nu - \theta}$$

it will follow that

$$\mu(N_{\alpha k}) \leqq \frac{1}{q(n)^{1+\delta}} .$$

Since for n sufficiently large

$$\frac{1}{q(n)^{1+\delta}} < \frac{\delta}{q(n) \ln q(n)} < \delta \mu(C^*(n)) ,$$

we have

$$\begin{aligned} \mu(C^*(n)) &= \mu(C^*(n) \cap B_\delta^k) + \mu(M_m \cap (B_\delta^k)') + \mu(N_m) \\ &\leqq \Big(2\delta(1 + \delta) + \frac{b\sqrt{\delta}(1 + \delta)(k + 1)}{2 \ln q(n)} + \frac{2\sqrt{\delta}(1 + \delta)}{\alpha} \\ &\qquad + \frac{b(1 + \delta)(k + 1)}{2\alpha \ln q(n)} + \delta\Big) \mu(C^*(n)) . \end{aligned}$$

Now if the term inside the parenthesis on the right of this inequality is less than one, we will have the contradiction

$$\mu(C^*(n)) < \mu(C^*(n)) .$$

It is easily seen that this term will be less than one if

$$k < \frac{2\alpha \ln q(n)}{b(\alpha\sqrt{\delta} + 1)(1 + \delta)} \left(1 - 3\delta - 2\delta^2 - \frac{2\sqrt{\delta}(1 + \delta)}{\alpha}\right) - 1 .$$

Therefore this inequality and inequality (1) above cannot be simultaneously satisfied for any values of δ, ν, θ, and α. This implies that

$$\frac{(1+\delta)\ln q(n)}{h(T)+\alpha\ln\alpha+(1-\alpha)\ln(1-\alpha)-\alpha\ln p(T)-\nu-\theta}$$
$$> \frac{2\alpha\ln q(n)}{b(\alpha\sqrt{\delta}+1)(1+\delta)}\left(1-3\delta-2\delta^2-\frac{2\sqrt{\delta}(1+\delta)}{\alpha}\right)-2\,.$$

Since δ, ν, and θ may be taken arbitrarily small and n may be chosen arbitrarily large this implies that

$$b \geqq 2\alpha(h(T)+\alpha\ln\alpha+(1-\alpha)\ln(1-\alpha)-\alpha\ln p(T))$$

for all α such that $0<\alpha<(p(T))/(p(T)+1)$.

REMARK. If $h(T)=\infty$ then we may choose S to be any Bernoulli automorphism with finite entropy and, by adapting the above proof to this case, we would arrive at

$$b \geqq \max_{0<\alpha<p(T)/(p(T)+1)} 2\alpha(h(S)+\alpha\ln\alpha+(1-\alpha)\ln(1-\alpha)-\alpha\ln p(T)$$

where $p(T)$ is the greatest integer which is strictly less than $e^{h(s)}$.

Proof of Theorem 4.1. The following argument is exactly the same as that found in [2]; however, we state it for the sake of completeness.

There are two cases to consider.

(1) If $h(T)=\infty$ then for any integer p there is a factor automorphism of T which is isomorphic to the Bernoulli automorphism S with a state space of p elements and with propability distribution $\pi_1=1/p, \pi_2=1/p, \cdots, \pi_p=1/p$. Since $h(S)=\ln p$, by the remark following the proof of Lemma 4.1, we have

$$b \geqq \max_{0<\alpha<(p-1)/p} 2\alpha(\ln p+\alpha\ln\alpha+(1-\alpha)\ln(1-\alpha)-\alpha\ln(p-1))\,.$$

If we set $\alpha=1/2$ this implies that $2b \geqq \ln p - 2\ln 2$ for all positive integers p. Therefore $b(T)=\infty$.

(2) Assume that T is an ergodic automorphism such that

$$0<h(T)<\infty$$

and T admits an approximation with speed $(\theta h(T))/(\ln n)$. We consider the automorphism T^m where m is a positive integer. Since $\mu(T^mC_i(n)\cap(T_n^mC_i(n))') \leqq \sum_{j=0}^{m-1}\mu(TC_{i+j}(n)\cap(T_nC_{i+j}(n))')$ for any element $C_i(n)$ in the partition ξ_n it is easily seen that T^m admits an approximation with speed $m(\theta h(T))/(\ln n)$. It is, of course, well known that $h(T^m)=m\,h(T)$. We wish to apply the result of Lemma 4.1 to T^m. This is not immediately possible, however, since T^m may not be ergodic. Nevertheless, it is clear that S^m is isomorphic to a factor automorphism of T^m since S is isomorphic to a factor automorphism of T. We can thus replace the partition η by $\bigvee_{j=0}^{m-1} T^j\eta$ and carry out the proof of Lemma 4.1 to obtain

$$\begin{aligned}&\theta mh(T)\\ &\geqq \max_{0<\alpha<p(T^m)/(p(T^m)+1)} 2\alpha(mh(T)+\alpha\ln\alpha+(1-\alpha)\ln(1-\alpha)-\alpha\ln p(T^m)) .\end{aligned}$$

Since $\lim_{m\to\infty} p(T^m)/(p(T^m)+1)=1$ we can assume for m sufficiently large that $p(T^m)/(p(T^m)+1)>1/2$ so that we may set $\alpha=1/2$ and obtain

$$\theta mh(T)\geqq mh(T)-\ln 2-\frac{1}{2}\ln p(T^m)>\frac{m}{2}h(T)-\ln 2 .$$

That is,

$$\theta h(T)>\frac{h(T)}{2}-\frac{\ln 2}{m}$$

for m sufficiently large. This is only possible if $\theta\geqq 1/2$, and this implies that $b(T)\geqq h(T)$.

Added note [4-25-72]. The above proof can be strengthened to obtain the result $2h(T)\leqq b(T)\leqq c(T)$ by making the following change originally devised by Katok.

We may assume that the integer k is even, and we replace the partition $\eta_k=\bigvee_{l=0}^{k}T^{-l}\eta$ by $\eta_k=\bigvee_{l=-k/2}^{k/2}T^{-l}\eta$. It follows then in the proof of Lemma 4.2 that

$$\begin{aligned}\int_{X\times Y}G(x,y)d\hat{\mu}&<k\sqrt{\delta}+k\delta+\frac{b}{\ln q(n)}\sum_{l=-k/2}^{k/2}|l|\\ &<2k\sqrt{\delta}+\frac{b\,k(k+2)}{4\ln q(n)} .\end{aligned}$$

Thus we arrive at

$$b\geqq \max_{0<\alpha<p(T)/(p(T)+1)} 4\alpha(h(T)+\alpha\ln\alpha+(1-\alpha)\ln(1-\alpha)-\alpha\ln p(T))$$

which leads to $b(T)\geqq 2h(T)$.

5. Entropy and approximation of measure preserving transformations. We now turn our attention to the case of an arbitrary automorphism T. It is stated in [2] that $c(T)\leqq 2h(T)$ for any automorphism T although the proof is given only in the ergodic case. By applying Theorem 3.1 we will extend the result of Section 4 to an arbitrary automorphism T and obtain $2\,h(T)\leqq b(T)$.

The following result is found in [4].

THEOREM. *If β is a measurable partition such that every set $B \in \beta$ is invariant under T, then*

$$h(T) = \int_{X/\beta} h(T_B) d\mu_\beta(B)$$

where $h(T_B)$ devotes the entropy of the automorphism T_B on the measure space $(B, \mathscr{F}_B, \mu_B)$.

It is known (see reference [4]) that for any automorphism T there exists a maximal invariant measurable partition, that is, a measurable partition β such that every set $B \in \beta$ is invariant under T and T_B is ergodic for almost all $B \in X/\beta$.

THEOREM 5.1. *If T is an arbitrary automorphism then $2h(T) \leqq b(T)$.*

Proof. If T admits an approximation with speed $b/\ln n$ and if β is the maximal invariant measurable partition for T, then Theorem 3.1 states that T_B admits an approximation with speed $f_B(n)$ such that

$$\int_{X/\beta} f_B(n) d\mu_\beta(B) \leqq \frac{b}{\ln n} .$$

It follows then that

$$\int_{X/\beta} \liminf_{n\to\infty} (\ln n) f_B(n) d\mu_\beta(B) \leqq \liminf_{n\to\infty} \int_{X/\beta} (\ln n) f_B(n) \mu_\beta(B) \leqq b .$$

If $b(T_B)$ indicates the greatest lower bound of the numbers $2a$ such that T_B admits an approximation with speed $a/(\ln n)$ then certainly $b(T_B) \leqq 2 \lim_{n\to\infty} \inf (\ln n) f_B(n)$, so that $\int_{X/\beta} b(T_B) \leqq 2b$.

Since T_B is ergodic for almost all B, we may apply Theorem 4.1 to obtain $2h(T) = \int_{X/\beta} 2h(T_B) d\mu_\beta(B) \leqq \int_{X/\beta} b(T_B) d\mu_\beta(B) \leqq 2b$, from which it immediately follows that $2h(T) \leqq b(T)$.

Combining this with the results from [2], we obtain the following corollary.

COROLLARY 5.1. *For any automorphism T, $2h(T) = b(T) = c(T)$.*

COROLLARY 5.2. *If an automorphism T admits an approximation with speed $b/(\ln n)$, then T admits a measure preserving approximation with speed $(b + \varepsilon)/(\ln n)$ for each $\varepsilon > 0$.*

COROLLARY 5.3. *An automorphism T has infinite entropy if and only if it admits no approximation with speed $b/(\ln n)$ for any b.*

In view of Ornstein's Isomorphy Theorem for Bernoulli shifts [8] we have the following corollary:

CORORARY 5.4. *If S and T are Bernoulli shifts then S and T are isomorphic if and only if $c(S) = c(T)$.*

REFERENCES

1. R. V. Chacon and T. Schwartzbauer, *Approximation and Invariance; Contributions to Ergodic Theory and Probability*, Proceedings of the First Midwestern Conference on Ergodic Theory, Springer, New York (1970).

2. A. B. Katok, *Entropy and Approximation of dynamical systems by periodic transformations*, Funktsional'nyi Analiz, (1) **1**, (1967), 75-85.

3. V. A. Rokhlin, *On the fundamental ideas of measure theory*, Math. Sbor. **67**, (1949), 107-150. Amer. Math. Soc. Trans. No. **71** (1952).

4. V. A. Rokhlin, *Lectures on the entropy theory of measure preserving transformations*, Russian Math. Surveys, (5) **22**, (1967), 1-52.

5. T. Schwartzbauer, *A general method for approximating measure preserving transformations*, Proc. Amer. Math. Soc., **24**, (1970), 643-648.

6. ———, *Approximation of measure preserving transformations*, to appear.

7. Ya. G. Sinai, *Weak isomorphism of transformations with invariant measure*, Math. Sbor., **63** (105), No. 1, (1963), 23-42.

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

Received April 19, 1972.

THE OHIO STATE UNIVERSITY