ON THE ENTROPY OF UNIQUELY ERGODIC TRANSFORMATIONS

BY

FRANK HAHN[1] AND YITZHAK KATZNELSON[2]

0. Introduction. Ergodic theory involves itself with the study of transformations of a measure space. Topological dynamics is involved with homeomorphisms of a topological space. The entropy of a measure preserving transformation is a gauge of its randomness. There are certain qualitative properties in topological dynamics which seem to play a similar role. For example, equicontinuity, distality, and minimality restrict the "randomness" of a transformation group. The question arises: If a measure preserving transformation is also a homeomorphism is there a strict relation between these notions? For instance it is known that equicontinuity and distality imply zero entropy. It was not known whether minimality implied zero entropy. W. Parry raised the following question: Let X be a compact metric space and T a continuous map of X into itself. Suppose there is a unique T invariant probability measure and suppose (X, T) is minimal (every orbit is dense). Can (X, T) have positive entropy?

The difficulty in answering such a question is the scarcity of a wide class of minimal, uniquely ergodic transformation groups (X, T). We give here a method of constructing minimal uniquely ergodic transformation groups (X, T). Furthermore our method is designed to show that there exist transformation groups (X, T), X a compact metric space, which are minimal and uniquely ergodic but have arbitrarily large finite entropy. Thus minimality and unique ergodicity do not bound the entropy of (X, T) and the answer to Parry's question is no.

We carry out our program by constructing closed shift invariant subspaces of a sequence space. In §1 we establish some notation and terminology. In §§2 and 3 we either modify or translate some known theorems into combinatorial terms which are to be used in the succeeding sections. In §4 we define the sequences which will be used to build the uniquely ergodic minimal pairs (X, T) of large but finite entropy. §§5, 6, 7, 8 are used to carry out the detailed examination of the properties of our sequence. §9 indicates the modification necessary in §4 to show that the same construction can be done on bisequences and thus T may also be a homeomorphism. The results are summarized in Corollaries 8.14 and 9.7.

The principal idea which makes our result possible is the weak law of large numbers. Lemmas 1.1 and 1.2 are merely forms of this law. The remainder is

---

application of these lemmas and technical computation. Since this phase is long, we have used an extensive numbering system with many parenthetic references where objects or notations are defined.

1. **Preliminaries.** Before we proceed we establish some of our terminology and two elementary lemmas which are the motivating ideas for the sequences con-constructed in §4.

Let $p \geq 1$ be an integer and let $a_0, a_1, \ldots, a_{p-1}$ be p distinguishable objects. The set $S = \{a_0, a_1, \ldots, a_{p-1}\}$ will be referred to as the symbol set on the p symbols $a_0, a_1, \ldots, a_{p-1}$. By a word of length $n \geq 1$ built on the symbol set S we mean an n-tuple $B = (b_0, b_1, \ldots, b_{n-1})$ where $b_i \in S$, $i = 0, 1, 2, \ldots, p-1$. We designate the length of B by $L(B)$. For the remainder of this section we will assume that all words are built on the symbol set S, and we often use the phrase "word" instead of "word built on the symbol set S."

If $B = (b_0, \ldots, b_{m-1})$ and $C = (c_0, \ldots, c_{n-1})$, $n \geq m$, are two words, then we say that B occurs in C at the ith position if $b_j = c_{i+j}$, $j = 0, 1, 2, \ldots, m-1$, and i is a fixed integer such that $0 \leq i \leq n - m$. The phrase "C contains B at the ith position" is to be equivalent to "B occurs in C at the ith position." If there is at least one i for which B occurs in C at the ith position, then we say that B occurs in C, or equivalently, C contains B.

LEMMA 1.1. *Let B be a fixed word on p symbols. The number of words C of length $k \cdot L(B)$ which contain B is asymptotic to $p^{k \cdot L(B)}$ as $k \to +\infty$.*

Proof. Let $L(B) = m$. Let $E(k)$ be the number of words of length km which contain B. Let $E'(k)$ be the number of words of length km which contain B in at least one position $i = jm$ where $0 \leq j < k$. Since the total number of words of length km is p^{km}, we have

$$p^{km} \geq E(k) \geq E'(k).$$

In order to prove our lemma it suffices to show that $E'(k) \approx p^{km}$ as $k \to +\infty$.

We compute $E'(k)$. Let C be a word of length km on the symbol set S. We let S' be the set whose elements are all words of length m on the symbol set S. There are p^m such words. C may be considered as a word of length k on the symbol set S'. $E'(k)$ is then the number of words of length k on the symbol set S' which contain the symbol B. The total number of words of length k on the symbol set S' is p^{mk}. The number of words of length k on the symbol set S' which fail to contain the symbol B is $(p^m - 1)^k$. Thus

$$E'(k) = p^{mk} - (p^m - 1)^k = p^{mk}(1 - (1 - 1/p^m)^k).$$

We see immediately that $E'(k) \approx p^{mk}$ as $k \to +\infty$ and our lemma is complete.

Let B be a word on the symbol set S. By $N(a_j, B)$ we mean the number of times that the symbol a_j occurs in B and this is defined to be the cardinality of the set $\{i : a_j$ occurs in B at the ith position$\}$. By $\alpha(a_j, B)$ we mean the frequency with which

a_j occurs in B and this is defined to be $(1/L(B)) \cdot N(a_j, B)$. We observe that $0 \leq N(a_j, B) \leq L(B)$ and $0 \leq \alpha(a_j, B) \leq 1$ for $0 \leq j < p$.

LEMMA 1.2. *Let $\varepsilon > 0$. The number of words C of length k having the property $|\alpha(a_j, C) - 1/p| < \varepsilon$ for $0 \leq j < p$ is asymptotic to p^k.*

Proof. This is a direct consequence of the weak law of large numbers, and we refer the reader to [5] or some other treatise on probability.

We now wish to define the space of sequences and the space of bisequences on the symbol set S. We let $\mathbf{Z}$ be the set of integers, $\mathbf{Z}^+$ the set of nonnegative integers, and when we wish to make a statement which applies to both $\mathbf{Z}$ and $\mathbf{Z}^+$ we will speak of $\mathbf{Z}'$. The space of all bisequences on the symbol set S will be denoted by $\Omega(S) = \prod_{\mathbf{Z}} S$. The space of all sequences in the symbol set S will be denoted by $\Omega^+(S) = \prod_{\mathbf{Z}^+} S$. Again, when we wish to make a statement which applies to either space we will speak of $\Omega'(S) = \prod_{\mathbf{Z}'} S$. The elements of $\Omega'(S)$ are functions $\omega : \mathbf{Z}' \to S$ and their value at $k \in \mathbf{Z}'$ will be denoted by ω_k.

We let $\Omega'(S)$ have the following topology: consider S as having the discrete topology and let $\Omega'(S)$ have the product topology. Since S is finite and thus compact, we see that $\Omega'(S)$ is compact. Since $\mathbf{Z}'$ is countable, it follows that $\Omega'(S)$ is metrizable. We will not be interested in a particular metric on $\Omega'(S)$, but we will make use of the neighborhood systems which we now describe. If $J \subset \mathbf{Z}'$ is any finite set $J = \{j_0 < j_1 < \cdots < j_{k-1}\}$ and if $\omega \in \Omega'(S)$ let $x(\omega, J) = (\omega_{j_0}, \omega_{j_1}, \ldots, \omega_{j_{k-1}})$. Thus $x(\omega, J)$ is a word of length $k =$ cardinality of J on the symbol set S. If B is a word of length k and $J = \{j_0 < j_1 < \cdots < j_{k-1}\} \subset \mathbf{Z}'$, then by a cylinder set determined by B and based on J we mean

$$(1.3) \qquad C(B, J) = \{\omega : \omega \in \Omega'(S) \text{ and } x(\omega, J) = B\}.$$

If we let $\mathfrak{B}'$ be the Borel field on $\Omega'(S)$ generated by the compact sets, then $\mathfrak{B}'$ is generated by all sets of the form (1.3).

If $\omega^0 \in \Omega'(S)$ and J is described as before, we let

$$(1.4) \qquad U(\omega^0, J) = \{\omega : \omega \in \Omega'(S) \text{ and } x(\omega, J) = x(\omega^0, J)\}.$$

Each such set is a compact open neighborhood of ω^0. As J ranges over the finite subsets of $\mathbf{Z}'$ these sets form a fundamental system of neighborhoods of ω^0. There is an even smaller fundamental system of neighborhoods which we shall use. For each nonnegative integer n let

$$I(n) = \{j : j \in \mathbf{Z}, -n \leq j \leq n\}, \qquad I^+(n) = \{j : j \in \mathbf{Z}^+, 0 \leq j < n\}.$$

Again we use $I'(n)$ for statements about either $I(n)$ or $I^+(n)$. We let $M(I(n)) = 2n + 1$, $M^+(I^+(n)) = n$ and $M'(I'(n))$ be whichever of the two numbers is appropriate to the context. In general, if $I' \subset \mathbf{Z}'$ is an interval, we let $M'(I')$ be the number of elements it contains. If $\omega^0 \in \Omega'(S)$, then the system

$$(1.5) \qquad U(\omega^0, I'(n)), \qquad n = 1, 2, \ldots,$$

is a fundamental system of compact open neighborhoods of ω^0.

We now define the shift transformations $T: \Omega(S) \to \Omega(S)$ and $T^+: \Omega^+(S)$ $\to \Omega^+(S)$. Again statements made about $T': \Omega'(S) \to \Omega'(S)$ will be statements which hold for both T and T^+. If $\omega \in \Omega'(S)$, we define T' by

$$(1.6) \qquad\qquad (T'\omega)_k = \omega_{k+1}.$$

We observe that T' is continuous and surjective. T^+ is however not injective. T is injective and is thus a homeomorphism of $\Omega(S)$.

2. **Almost periodicity.**   In this section we will define what we mean by saying that $\omega \in \Omega'(S)$ is almost periodic. We will prove some lemmas which will be used in §§6, 7, and 8. Our definition is the same as that of [6] and [7]. In the case of $(\Omega(S), T)$ it reduces to the definition in [3] of an almost periodic point. The definition will be in terms of the combinatorial properties we wish to use. The lemmas are again phrased in these terms. Although they are special cases or modifications of theorems in [3], we present them for completeness sake.

We first must say what we mean by a word B occurring in $\omega \in \Omega'(S)$. Let $B = (b_0, \ldots, b_{n-1})$ be a word built on the symbol set S. Let $\omega \in \Omega'(S)$. By an interval $I' \subset Z'$ of length n we mean any set of the form $I' = \{k \leq j < k+n\}$. We say that B occurs in ω if there is an interval $I' \subset Z'$ for which $x(\omega, I') = B$.

DEFINITION 2.1.   Let $\omega \in \Omega'(S)$. We say that ω is almost periodic if for any word B which occurs in ω there is a positive integer $R(B)$ having the following property: If $I' \subset Z'$ is an interval of length no less than $R(B)$, then B occurs in $x(\omega, I')$.

If $\omega \in \Omega'(S)$, we let

$$(2.2) \qquad\qquad O(\omega) = \{T'^j\omega : j \in Z'\},$$

$$(2.3) \qquad\qquad O(\omega)^- = \text{closure of } O(\omega) \text{ in } \Omega'(S).$$

We observe that if $\omega \in \Omega(S)$, then $T(O(\omega)) = O(\omega)$ and $T(O(\omega)^-) = O(\omega)^-$. Similar equalities hold for T^{-1}. For T^+ we have $T^+(O(\omega)) \subset O(\omega)$ and $T^+(O(\omega)^-) \subset O(\omega)^-$.

LEMMA 2.4.   *Let $\omega \in \Omega'(S)$ be almost periodic. If $\omega^0 \in O(\omega)^-$, then $\omega \in O(\omega^0)^-$. Thus $O(\omega)^- = O(\omega^0)^-$.*

Proof. We first show that if B is a word occurring in ω, then B occurs in ω^0. Since ω is almost periodic, we choose an $R(B) = m$ satisfying Definition 2.1. Let I' be an interval of length m and consider the neighborhood $U(\omega^0, I')$ (see (1.5)). There exists a $j \in Z'$ for which $T^j\omega \in U(\omega^0, I')$. The word B occurs in $x(T^j\omega, I')$ because of the almost periodicity of ω. Thus B occurs in ω^0 since

$$x(T^j\omega, I') = x(\omega^0, I').$$

We now show that $\omega \in O(\omega^0)^-$. Let n be an arbitrary positive integer and consider the neighborhood $U(\omega, I'(n))$ of ω (see (1.5)). The word $B = x(\omega, I'(n))$ occurs in ω and thus it occurs in ω^0. Suppose $B = x(\omega, I'(n)) = x(\omega^0, J')$ where

$$J' = \{j : j \in Z', k \leq j < k + M'(I'(n))\}$$

(see paragraph preceding (1.5)). In either the case of Z or Z^+ there is a $t \in Z'$ for which $J' = I'(n) + t$. We see that $x(T^t\omega^0, I'(n)) = x(\omega^0, I'(n) + t) = x(\omega^0, J')$. Consequently $T^t\omega^0 \in U(\omega, I'(n))$. Since n was an arbitrary positive integer, we have $\omega \in O(\omega^0)^-$.

LEMMA 2.5. *Let $\omega \in \Omega'(S)$ be almost periodic and let $\omega^0 \in O(\omega)^-$. Then ω_0 is almost periodic and consequently any word which occurs in ω occurs in ω_0 and conversely.*

Proof. We see from the beginning of the proof of Lemma 2.4 it is only necessary to show that ω^0 is almost periodic. Suppose B' occurs in ω^0. Choose n large enough so B' occurs in $B = x(\omega^0, I'(n))$. It suffices to show that there is an $R(B)$ satisfying Definition 2.1 for ω^0. Since $\omega^0 \in O(\omega)^-$, there is a q for which $T^q\omega \in U(\omega^0, I'(n))$. Thus B occurs in $T^q\omega$ and a fortiori B occurs in ω. Since ω is almost periodic, choose $R(B)$ satisfying Definition 2.1 for ω. Consider any interval $J' \subset Z'$ of length no less than $R(B)$. There is an $n \in Z'$ for which $T^n\omega \in U(\omega^0, J')$. Since B occurs in $x(T^n\omega, J') = x(\omega^0, J')$ our lemma is complete.

LEMMA 2.6. *Let $\omega \in \Omega'(S)$ and let ω be almost periodic. Let B be a word and $I' \subset Z'$ an interval whose length is the length of B. With these hypotheses*

$$C(B, I') \cap O(\omega)^- \neq \varnothing$$

(see 1.3) if and only if B occurs in ω.

Proof. Suppose $C(B, I') \cap O(\omega)^- \neq \varnothing$. The set $C(B, I')$ is open in $\Omega'(S)$, thus if it meets $O(\omega)^-$, it must meet $O(\omega)$. Thus there is a k for which $T^k(\omega) \in C(B, I')$. It follows that B occurs in $T^k(\omega)$ and thus B occurs in ω. (Note we did not use the almost periodicity of ω.)

Suppose now that B occurs in ω. Let $I' = \{j: k \leq j < k + L(B)\}$. Since ω is almost periodic, it follows that there is an $s \in Z^+$ having the property that if

$$I'' = \{j: k + s \leq j < k + s + L(B)\},$$

then B occurs in ω at the interval I''. That is $x(\omega, I'') = B$. We then have $x(T^s(\omega), I') = x(\omega, s + I') = x(\omega, I'') = B$. But this merely says that $T^s(\omega) \in C(B, I')$. Thus $C(B, I') \cap O(\omega)^- \neq \varnothing$.

The next lemma is about bisequences.

LEMMA 2.7. *Let $\omega \in \Omega(S)$ be almost periodic. The set $\{T^j\omega: j \in Z^+\}$ is dense in $O(\omega)^-$.*

Proof. Let $\omega^0 \in O(\omega)^-$, let n be any positive integer. Consider the neighborhood $U(\omega^0, I(n))$. Let $B = x(\omega^0, I(n))$. We see that $C(B, I(n))$ meets $O(\omega)^-$ so the word B occurs in ω. Since ω is almost periodic, there is an interval $J \subset Z^+$ for which $B = x(\omega, J)$ and there is a $k \in Z^+$ such that $I(n) + k = J$. Thus

$$x(T^k\omega, I(n)) = x(\omega, I(n) + k) = x(\omega, J) = B.$$

We see that $T^k \omega \in U(\omega^0, I(n))$ for some $k \in Z^+$. Since n was an arbitrary positive integer, our lemma is proved.

3. **Unique ergodicity.** The labor of §5 will be to show that a given almost periodic sequence $\omega \in \Omega^+(S)$ is uniquely ergodic under the transformation T^+. In order to do this we take sufficient conditions for unique ergodicity, translate them into combinatorial terms and show that they hold for our given sequence ω. In this section we present these conditions as a series of lemmas. The lemmas are modifications of theorems in [2], [8]. In fact, the sufficient conditions given here are also necessary (see [2], [8]). We will not concern ourselves with this and only present those facts which we will use later.

Let X be a compact Hausdorff space and $\phi: X \to X$ a continuous mapping. Let $C(X)$ be the real Banach space of real valued continuous functions on X given their supremum norm. If $f \in C(X)$, we define ϕf by $\phi f = f \circ \phi$. Let P be the set of probability measures defined on the Borel subsets of X. We let ϕ induce a map of P into itself, which we again call ϕ. If $\mu \in P$ we define $\phi \mu$ by the equation $(f, \phi \mu) = (\phi f, \mu)$ for $f \in C(X)$. Here the symbol (f, μ) is defined as the integral of f with respect to the measure μ. We say that $\mu \in P$ is ϕ invariant if $\phi \mu = \mu$. It is a direct consequence of the Kakutani-Markov fixed point theorem that a ϕ invariant probability measure always exists. If this measure is also unique, then we say that the pair (X, ϕ) is uniquely ergodic. The terminology is standard [8] and is reasonable in the sense that the uniqueness of the invariant measure guarantees its ergodicity.

We shall use the following notation. If $f \in C(X)$ and n is a positive integer, then we let

$$(3.1) \qquad f_n(x) = \frac{1}{n} \sum_{j=0}^{n-1} \phi^j f(x) \in C(X).$$

LEMMA 3.2. *If for each $f \in C(X)$ and $x \in X$, $\lim_n f_n(x) = k(f) = $ constant, depending only on f, then (X, ϕ) is uniquely ergodic with unique invariant measure $\mu \in P$ and $(f, \mu) = k(f)$ for all $f \in C(X)$.*

Proof. We know there exist ϕ invariant measures P. Let ν be any such invariant measure. Using the Lebesgue dominated convergence theorem and the invariance of ν, we have, as $n \to +\infty$,

$$(f, \nu) = (f_n, \nu) \to (k(f), \nu) = k(f).$$

Thus for each invariant measure ν we have $(f, \nu) = k(f)$. This implies that there is a unique invariant measure μ and the lemma is complete.

LEMMA 3.3. *Suppose there is an $x_0 \in X$ such that $X = $ closure $\{\phi^j x_0 : j = 0, 1, 2, \ldots\}$. If for each $f \in C(X)$, $\lim_n (\phi^j f)_n(x_0)$ exists uniformly in $j = 0, 1, 2, \ldots$, then (X, ϕ) is uniquely ergodic.*

Proof. In view of the previous lemma we need only show that $\lim_n f_n(x) = k(f)$ for each $f \in C(X)$ and $x \in X$. We first show that our hypotheses imply that f_n

is a Cauchy sequence in $C(X)$. Given $\varepsilon > 0$, there is an N such that $n \geq N$ and $p \geq 1$ implies

$$|f_n(\phi^j x_0) - f_{n+p}(\phi^j x_0)| = |(\phi^j f)_n(x_0) - (\phi^j f)_{n+p}(x_0)| < \varepsilon/2 \qquad \text{for } j = 0, 1, 2, \ldots.$$

Since the set $\{\phi^j x_0 : j = 0, 1, 2, \ldots\}$ is dense in X, we see that if $n \geq N$ and $p \geq 1$ then $\|f_n - f_{n+p}\| < \varepsilon$. Thus f_n is a Cauchy sequence in $C(X)$ and consequently there is an $f^* \in C(X)$ such that $\lim_n f_n = f^*$.

To complete the lemma we must show that f^* is constant. Since for each $n = 1, 2, \ldots$ we have $|f_n(x) - f_n(\phi x)| < (2/n)\|f\|$, we see that

$$f^*(x_0) = f^*(\phi x_0) = \cdots = f^*(\phi^j x_0) = \cdots.$$

Thus f^* is constant on a dense subset of X. Since $f^* \in C(X)$, this implies that f^* is constant. This completes the proof of the lemma.

Fix an $\omega \in \Omega'(S)$. We are interested in a certain class of functions in $C(O(\omega)^-)$. If $\omega \in \Omega^+(S)$ and B is a word of length n, we let

(3.4) $$\chi_B = \text{characteristic function of } C(B, I^+(n)) \cap O(\omega)^-$$

(see (1.3)–(1.5)).

If $\omega \in \Omega(S)$ and B is a word of length $2n+1$, we let

(3.5) $$\chi_B = \text{characteristic function of } C(B, I(n)) \cap O(\omega)^-.$$

When we speak of a word B and the function χ_B, we will mean (3.4) or (3.5), whichever is appropriate to the context. In either case $\chi_B \in C(O(\omega)^-)$ since $C(B, I'(n))$ is both open and closed.

LEMMA 3.6. *Let $\omega \in \Omega'(S)$ be almost periodic. If $\lim_n (T''^j \chi_B)_n(\omega)$ converges uniformly in j for each appropriate word B (either of length q or $2q+1$) and $C(B, I'(q)) \cap O(\omega)^- \neq \varnothing$, then $(O(\omega)^-, T')$ is uniquely ergodic.*

Proof. Because of Lemma 3.3 and Lemma 2.7 we need only show for each $f \in C(O(\omega)^-)$ that $\lim_n (T''^j f)_n(\omega)$ exists uniformly in j. Choose an $f \in C(O(\omega)^-)$ and let $\varepsilon > 0$. Because of uniform continuity of f there is a q such that if B is a word of length q (or $2q+1$), then the oscillation of f over $C(B, I'(q)) \cap O(\omega)^-$ is less than $\varepsilon/6$. It is sufficient to show that $(T''^j f)_n(\omega)$ is a Cauchy sequence in n uniformly in j.

For each B such that $C(B, I'(q)) \cap O(\omega)^- \neq \varnothing$ we let

$$m_B = \inf\{f(\omega') : \omega' \in C(B, I'(q)) \cap O(\omega)^-\},$$

$$M_B = \sup\{f(\omega') : \omega' \in C(B, I'(q)) \cap O(\omega)^-\},$$

$$m = \sum_B m_B \chi_B \in C(O(\omega)^-),$$

$$M = \sum_B M_B \chi_B \in C(O(\omega)^-)$$

where the sum is taken over all words B of length q (or $2q+1$) for which

$$C(B, I'(q)) \cap O(\omega)^- \neq \varnothing.$$

We observe that $m \leq f \leq M$ and

$$\|M - m\| \leq \varepsilon/6 \quad \text{and} \quad \|f - m\| < \varepsilon/3.$$

From our hypothesis it follows that $(T''^j m)_n(\omega)$ is a Cauchy sequence in n uniformly in j. Thus there is an N such that $n \geq N$ and $s \geq 0$ imply

$$|(T''^j m)_n(\omega) - (T''^j m)_{n+s}(\omega)| < \varepsilon/3 \quad \text{for } j = 0, 1, 2, \ldots.$$

We also have $T''^j m \leq T''^j f \leq T''^j M$,

$$\|T'^j M - T''^j m\| \leq \varepsilon/6 \quad \text{and} \quad \|T''^j f - T''^j m\| < \varepsilon/3 \quad \text{for } j = 0, 1, 2, \ldots.$$

Thus if $n \geq N$ and $s \geq 0$ we have

$$
\begin{aligned}
|(T''^j f)_n(\omega) - (T''^j f)_{n+s}(\omega)| &\leq |(T''^j f)_n(\omega) - (T''^j m)_n(\omega)| \\
&\quad + |(T''^j m)_n(\omega) - (T''^j m)_{n+s}(\omega)| \\
&\quad + |(T''^j m)_{n+s}(\omega) - (T''^j f)_{n+s}(\omega)| \\
&< \varepsilon.
\end{aligned}
$$

Thus $(T''^j f)_n(\omega)$ is a Cauchy sequence in n uniformly in j and our lemma is complete.

We wish to translate Lemma 3.6 into combinatorial terms. Let B and C be two words built on the symbol set S. We define $N(B, C)$ to be the number of occurrences of B in C. This is the cardinality of the set $\{i: B \text{ occurs in } C \text{ at the } i\text{th position}\}$ (see §1). We observe that if $L(B) > L(C)$, then $N(B, C) = 0$ and if $L(B) \leq L(C)$, then $0 \leq N(B, C) \leq L(C) - L(B) + 1 \leq L(C)$. We define the frequency of occurrence of B in C to be

$$(3.7) \qquad\qquad \alpha(B, C) = (1/L(C)) N(B, C)$$

and observe that $0 \leq \alpha(B, C) \leq 1$.

If $\omega \in \Omega'(S)$ and if $j \in Z$ and $k \in Z^+$, $k \neq 0$, we let

$$(3.8) \qquad\qquad D_\omega(j, k) = D(j, k) = x(\omega, [j, j+k))$$

where $[j, j+k) = \{i: i \in Z, j \leq i < j+k\}$. Thus $D_\omega(j, k)$ is the word of length k which occurs in ω at the position j.

THEOREM 3.9. *If $\omega \in \Omega'(S)$ is almost periodic and if $\lim_k \alpha(B, D(j, k))$ exists uniformly in j for each word B of length q (or $2q+1$), $q = 1, 2, 3, \ldots$, which appears in ω, then $(O(\omega)^-, T')$ is uniquely ergodic. Moreover, this limit depends only on B and is the measure of the set of all sequences in $O(\omega)^-$ which begin with B (have B centered at zero).*

Proof. It follows from Lemma 2.6 that B occurs in ω if and only if

$$C(B, I'(q)) \cap O(\omega)^- \neq \varnothing.$$

Thus by Lemma 3.6 we need only show that $\lim_n (T''^j \chi_B)_n(\omega)$ converges uniformly in j for each word B of length q (or $2q+1$), $q = 1, 2, 3, \ldots$, for which

$$C(B, I'(q)) \cap O(\omega)^- \neq 0.$$

We observe that

$$(T''^j \chi_B)_n(\omega) = \frac{1}{n} \sum_{i=0}^{n-1} T''^{j+i} \chi_B(\omega)$$

$$= \frac{1}{n} \sum_{i=j}^{j+n-1} T''^i \chi_B(\omega) = \frac{1}{n} N(B, D(j, n+L(B)-1))$$

$$= \frac{n+L(B)-1}{n} \cdot \frac{N(B, D(j, n+L(B)-1))}{L(D(j, n+L(B)-1))}$$

$$= \frac{n+L(B)-1}{n} \cdot \alpha(B, D(j, n+L(B)-1)).$$

Our hypothesis asserts that $\lim_k \alpha(B, D(j, k))$ exists uniformly in j. It follows from the above equation that $\lim_n (T''^j \chi_B)(\omega)$ exists uniformly in j. Thus $(O(\omega)^-, T')$ is uniquely ergodic. The $\lim_n (T''^j \chi_B)_n(\omega)$ is equal to $\lim_k \alpha(B, D(j, k))$ and depends only on B. This limit is the integral of χ_B with respect to the unique invariant measure. Thus our theorem is complete.

The next lemma will not be used until §8. For this lemma we let X_i, $i = 1, 2$, be compact metric spaces. Let $T_i : X_i \to X_i$ be continuous and let $\pi : X_1 \to X_2$ such that $\pi T_1 = T_2 \pi$. Let (X_2, T_2) be uniquely ergodic with unique invariant probability measure μ_2.

LEMMA 3.10. *If μ_1 is a T_1 invariant probability measure, then $\pi \mu_1 = \mu_2$. If $E \subset X_2$ is a Borel set, then $\mu_2(E) = \mu_1(\pi^{-1}E)$.*

Proof. It is clear that $\pi \mu_1$ is a probability measure on X_2. One sees easily that $T_2 \pi \mu_1 = \pi T_1 \mu_1 = \pi \mu_1$ and then since μ_2 is the unique invariant probability on X_2, we have $\pi \mu_1 = \mu_2$. It suffices to show that $\mu_2(E) = \mu_1(\pi^{-1}E)$ for each compact subset E of X_2. To do this we choose a sequence $f_n \in C(X_2)$ for which $f_n \downarrow \chi_E$. We then have

$$(f_n, \mu_2) \downarrow (\chi_E, \mu_2) = \mu_2(E),$$

$$(f_n, \mu_2) = (f_n, \pi \mu_1) = (\pi f_n, \mu_1) \downarrow (\pi \chi_E, \mu_1)$$

$$= \mu_1(\pi^{-1}E).$$

Thus $\mu_2(E) = \mu_1(\pi^{-1}E)$ and our lemma is complete.

4. Construction of sequences. In this section we construct by induction a family of symbol sets S_j, $j = 0, 1, 2, 3, \ldots$, and a family of sequences $\omega^j \in \Omega^+(S_j)$. These will be the sequences examined in §§5, 6, 7, and 8. The purpose of our inductive procedure is to see that in each $\omega^j \in \Omega^+(S_j)$ extremely many words of a given length

are used, that words occur with great regularity, and that ω^j is almost periodic. The idea which allows us to do this is the weak law of large numbers (Lemma 1.2) and the fact that almost all sequences contain a given symbol (Lemma 1.1).

We use the following notation: if $B=(b_0, b_1, \ldots, b_{m-1})$ and $C=(c_0, c_1, \ldots, c_{n-1})$ are two words built on a symbol set S, then BC is the word of length $m+n$ built on S given by $BC=(b_0, b_1, \ldots, b_{m-1}, c_0, c_1, \ldots, c_{n-1})$. If $B_1, B_2, \ldots, B_m$ are words built on S, then it is clear what we mean by the word $B_1 B_2 \cdots B_m$ built on S. We now begin our construction.

Let $0 < \delta < 1$ and choose $\delta_0, \delta_1, \ldots$ such that $0 < \delta_i < 1$ and $\prod_{i=0}^{\infty} (1-\delta_i) \downarrow 1-\delta$. The number δ is subject only to the restriction $0 < \delta < 1$ and remains fixed throughout our arguments.

For $j=0$ let $S_0 = \{B(0, 0), B(0, 1), \ldots, B(0, n_0-1)\}$ where the $B(0, t)$ are distinct symbols and $n_0 \geq 2$. We define the word

$$A_0 = B(0, 0)B(0, 1) \cdots B(0, n_0-1).$$

Let $\eta_0 = 1$ and let $k_0 = 1$.

Suppose now that for all j, $0 \leq j < m$, we have defined $\eta_j > 0$, $k_j \in \mathbf{Z}^+$, $k_j \neq 0$ and $S_j = \{B(j, 0), B(j, 1), \ldots, B(j, n_j-1)\}$ and the word

$$(4.1) \qquad A_j = B(j, 0)B(j, 1) \cdots B(j, n_j-1)$$

satisfying the following conditions.

(4.2) $n_j \geq 2$ and for $j>0$ each $B(j, t)$, $0 \leq t < n_j$, is a word built on the symbol set $S_{j-1} = \{B(j-1, 0), \ldots, B(j-1, n_{j-1})\}$.

(4.3) A_{j-1} occurs in the oth place of $B(j, 0)$. That is, A_{j-1} is the initial word of $B(j, 0)$.

(4.4) $L_{j-1}(B(j, t)) = n_{j-1}k_j = r_j$, $j>0$, for $0 \leq t < n_j$, where $L_{j-1}(B(j, t))$ is the length of $B(j, t)$ in terms of the symbols of S_{j-1}.

(4.5) For $j>0$, S_j is the set of all distinct words on the symbol set S_{j-1} whose length in terms of these symbols is $r_j = n_{j-1} \cdot k_j$ and which have the following properties:

(a) A_{j-1} occurs in each $B(j, t) \in S_j$;

(b) $|\alpha(B(j-1, t), B(j, s)) - 1/(n_{j-1})| < \eta_{j-1}/n_{j-1}$ for $0 \leq t < n_{j-1}$, $0 \leq s < n_j$. Here $\alpha(B(j-1, t), B(j, s))$ is the frequency with which the symbol $B(j-1, t)$ occurs in the word $B(j, s)$ (see Lemma 1.2 and the paragraph preceding it).

$$(4.6) \qquad 0 < \eta_j < (1/n_j^2) \cdot (1/2^j) \qquad \text{for } j > 0, \eta_0 = 1.$$

$$(4.7) \qquad \begin{aligned} &n_j \geq n_{j-1}^{r_j}(1-\delta_j) \quad \text{for } j > 0, \text{ where } r_j = n_{j-1} \cdot k_j. \\ &k_j \geq n_{j-1}^2 \qquad\qquad \text{for } j > 0, \\ &k_0 = 1, \\ &n_j > n_{j-1} \qquad\qquad \text{for } j > 0. \end{aligned}$$

We now proceed on the induction step of our construction. Consider all words B of length $r_m = n_{m-1} \cdot k_m$ built on the symbol set S_{m-1}. (The number k_m is an integer which is yet to be fixed.) We choose from these all those words B which have the additional properties:

(4.8) $$A_{m-1} \text{ occurs in } B,$$

(4.9) $$|\alpha(B(m-1, t), B) - 1/(n_{m-1})| < \eta_{m-1}/n_{m-1} \quad \text{for } 0 \leq t < n_{m-1}.$$

We let S_m be the set of all these words and label them in the following fashion

(4.10) $$S_m = \{B(m, 0), B(m, 1), \ldots, B(m, n_m - 1)\}.$$

We stipulate that A_{m-1} is the initial word of $B(m, 0)$.

We now fix k_m. Lemmas 1.1 and 1.2 insure us that we may choose k_m such that

(4.11)
$$k_m \geqq n_{m-1}^2,$$
$$n_m > n_{m-1},$$
$$n_m \geqq n_{m-1}^{r_m}(1 - \delta_m) \quad \text{where } r_m = n_{m-1} \cdot k_m.$$

We now choose η_m such that

(4.12) $$0 < \eta_m < (1/n_m^2) \cdot (1/2^m).$$

We let A_m be the word

(4.13) $$A_m = B(m, 0)B(m, 1) \cdots B(m, n_m - 1).$$

We see that k_m, η_m, S_m, and A_m satisfy 4.1 through 4.7 and our induction is complete.

Before we define the sequences $\omega^j \in \Omega^+(S_j)$, we make a few observations which will be helpful later.

(4.14) If $j > 0$ and $1 \leqq r \leqq j$, then we may look at each $B(j, t)$, $0 \leqq t < n_j$, as a word built on the symbol set S_{j-r}. This is obvious from our construction. Each $B(j, t)$ is a word built on the symbol set S_{j-1}, but each symbol in this set is a word built on the symbol set S_{j-2}. We proceed inductively. Thus we see that if $j \geqq 0$ and $0 \leqq k \leqq j$, then A_j is a word built on the symbol set S_k.

(4.15) Considering A_{j+k}, $j \in \mathbf{Z}^+$, $k \in \mathbf{Z}^+$, $k \neq 0$, as a word built on the symbol set S_j, we see that A_j occurs in the 0th position of A_{j+k}. That is, A_j is the initial word of A_{j+k}.

(4.16) As we observed in (4.14), if $j > 0$ and $0 \leqq r \leqq j$, then $B(j, t)$ is a word built on the symbol set S_{j-r}. We let $L_{j-r}(B(j, t))$ be its length in terms of these symbols. An easy induction gives us (see (4.4))

$$L_{j-r}(B(j, t)) = r_j \cdot r_{j-1} \cdots r_{j-r+1}$$
$$= k_j \cdots k_{j-r+1} \cdot n_{j-1} \cdots n_{j-r} \quad \text{for } 0 \leqq t < n_j.$$

We immediately have

$$L_{j-r}(A_j) = k_j \cdots k_{j-r+1} \cdot n_j \cdot n_{j-1} \cdots n_{j-r}.$$

Thus we see that if $n > 0$ and $0 \leqq j < n$, then

$$L_j(A_n) \to +\infty \qquad \text{as } n \to +\infty.$$

(4.17) We now define $\omega^j \in \Omega^+(S_j)$ for $j \in Z^+$. Fix a $j \in Z^+$. For each $m \geqq j$ we may consider A_m as a word built on the symbol set S_j. To define the ith coordinate of ω^j which is designated by $(\omega^j)_i$, we choose any $m \geqq j$ for which $L_j(A_m) \geqq i$. We let $(\omega^j)_i$ be the ith symbol in the word A_m. Our remarks in 4.15 assure us that $(\omega^j)_i$ is unambiguously defined. We see that $\omega^j \in \Omega^+(S_j)$. For the sake of brevity we will use the imprecise but suggestive notation

$$\omega^j = \bigcup_{i=j}^{\infty} A_i.$$

We remark here that the distinction between the ω^j, $j = 0, 1, 2, \ldots$, is mostly a formal one. We see that ω^0 is a sequence built on the symbol set S_0. The sequence ω^j is in a sense the same sequence as ω^0, except that we think of it as being built of the blocks of ω^0 of length r_j which are located at $0, r_j, 2r_j, \ldots$. The distinction between ω^0 and ω^j is technical but lends itself to the computations which we must perform in §§7 and 8.

The various parts of our inductive definition and the remarks made here will be used primarily as follows: In §5 we use (4.4), (4.5), (4.16) to show that each ω^j is almost periodic. In §6 we use the δ, $0 < \delta < 1$, and (4.7) to bound the combinatorial entropy. We do not use (4.6) until §8 when we bound the entropy of ω^j.

5. Almost periodicity of ω^j.

THEOREM 5.1. *For each $j \in Z^+$ the sequence $\omega^j \in \Omega^+(S_j)$ is almost periodic.*

Proof. Suppose B is a word built on the symbol set S_j and B occurs in ω^j. We must show that there is an integer $R(B)$ with the following property: If $I \subset Z^+$ is an interval of length $R(B)$, then B occurs in $x(\omega^j, I)$ (see Definition 2.1). Since by (4.16) $L_j(A_n) \to +\infty$ as $n \to +\infty$, there is an $n > j$ for which B occurs in A_n.

Clearly it suffices to show that there is an integer $R(A_n)$ with the property: If $I \subset Z^+$ is an interval of length $R(A_n)$, then A_n occurs in $x(\omega^j, I)$. This is almost immediate from our construction of ω^j. In our construction of ω^j we see that each A_m, $m \geqq n+1$, is a word in the symbols of S_{n+1}. Each symbol $B(n+1, t) \in S_{n+1}$ must by (4.5) contain A_n. By (4.16) $L_j(B(n+1, t)) = r_{n+1} \cdots r_{j+1}$. Thus for every $s \in Z^+$ we see that A_n occurs in

$$x(\omega^j, [sr_{n+1} \cdots r_{j+1}, (s+1)r_{n+1} \cdots r_{j+1})).$$

These are blocks in ω^j which begin at $sr_{n+1} \cdots r_{j+1}$ and end at $(s+1)r_{n+1} \cdots r_{j+1} - 1$. Clearly, if we set $R(A_n) = 3r_{n+1} \cdots r_{j+1}$, our theorem is completed.

6. The combinatorial entropy of ω^j. The notion of combinatorial entropy of a sequence appears in [9] (referred to as absolute entropy) and is a special case of the topological entropy of [1]. In §8 we will see how it is related to the measure theoretic entropy. In this section we wish to show that the combinatorial entropy of ω^j, denoted by $K(\omega^j)$, has the property that $\log n_j(1-\delta) \leq K(\omega^j) < \log n_j$. We recall that δ was fixed at the beginning of our construction of ω^j and subject only to the condition $0 < \delta < 1$. We remark, of course, that the sequence ω^j depends on the δ chosen.

We now define the combinatorial entropy of an arbitrary sequence or bisequence. Let $S = \{a_0, a_1, \ldots, a_{p-1}\}$ be a symbol set on p symbols. Let $\omega \in \Omega'(S)$ and let $\theta(n, \omega)$ be the number of distinct words of length n which occur in ω (see paragraph preceding Definition 2.1). We see immediately

(6.1) $$0 \leq \theta(n, \omega) \leq p^n.$$

It is not difficult to see that

(6.2) $$\theta(m+n, \omega) \leq \theta(m, \omega) \cdot \theta(n, \omega).$$

From this it follows that $\log \theta(n, \omega)$ is subadditive in n. A known theorem asserts that the following limit exists

(6.3) $$\lim_n \frac{\log \theta(n, \omega)}{n} = \inf_n \frac{\log \theta(n, \omega)}{n}.$$

DEFINITION 6.4. If $\omega \in \Omega'(S)$, we define the combinatorial entropy of ω to be

$$K(\omega) = \lim_n \frac{\log \theta(n, \omega)}{n}.$$

From (6.1) we see that

(6.4) $$0 \leq K(\omega) \leq \log p.$$

LEMMA 6.5. *If $\omega \in \Omega'(S)$ and if ω is almost periodic, then $K(\omega) < \log p$.*

Proof. First of all we may assume that each symbol of S occurs in ω. If this were not the case, then $\omega \in \Omega'(S')$, where S' has p' elements and $p' < p$. The truth of the lemma would assert $K(\omega) < \log p' < \log p$.

Thus the symbol a_0 occurs in ω. Since ω is almost periodic, there is a positive integer m having the property that if I' is an interval of length m in Z', then a_0 occurs in $x(\omega, I')$. That is, a_0 occurs in any word B of length m which occurs in ω. In particular the word $B = (a_1, \ldots, a_1)$ of length m cannot occur in ω. Let k be a positive integer. By the same argument used in Lemma 1.2 we see that the number of words in ω of length km is bounded above by $(p^m - 1)^k$. That is, $\theta(km) \leq (p^m - 1)^k$. Thus

(6.6) $$(\log \theta(km))/km \leq (\log (p^m - 1))/m.$$

Taking limits on k, we observe that

$$K(\omega) \leqq (\log (p^m - 1))/m < \log p$$

and our lemma holds.

THEOREM 6.7. *For the ω^j defined in §4 we have*

$$\log n_j(1 - \delta) \leqq K(\omega^j) < \log n_j$$

for each $j \in Z^+$, where δ is the constant described in §4 and is subject only to the condition $0 < \delta < 1$.

Proof. It follows from Theorem 5.1 that ω^j is almost periodic. Thus the previous lemma asserts

$$K(\omega^j) < \log n_j.$$

Fix a $j \in Z^+$. Then $\omega^j \in \Omega^+(S_j)$ and by (4.16) $L_j(B(j+s, t)) = r_{j+s} \cdot r_{j+s-1} \cdots r_{j+1}$ for $s \geqq 1$ and $0 \leqq t < n_{j+s}$. Recall this is the length of the word $B(j+s, t)$ in terms of the symbols S_j. Thus the number of words of length $r_{j+s} \cdots r_{j+1}$ is no less than the number of symbols in S_{j+s}. We thus have

(6.8) $\theta(r_{j+s} \cdots r_{j+1}) \geqq n_{j+s}.$

(4.7) asserts that, for each $m \geqq 1$,

(6.9) $n_m \geqq n_{m-1}^{r_m}(1 - \delta_m).$

A simple induction and the fact that $0 < 1 - \delta_m \leqq 1$ gives us

(6.10)
$$n_{j+s} \geqq (n_j(1 - \delta_{j+1})(1 - \delta_{j+2}) \cdots (1 - \delta_{j+s}))^{r_{j+s} \cdots r_{j+1}}$$
$$\geqq (n_j \cdot (1 - \delta))^{r_{j+s} \cdots r_{j+1}}.$$

Now as $s \to +\infty$, $r_{j+s} \cdots r_{j+1} \to +\infty$. Using 6.8 and 6.10 and the definition of $K(\omega^j)$, we obtain

$$K(\omega^j) \geqq \log n_j(1 - \delta),$$

which completes our theorem.

7. **Unique ergodicity of ω^j.** The purpose of this section is to show that for each j the system $(O(\omega^j)^-, T_j^+)$ is uniquely ergodic. T_j^+ is the shift transformation of §1 (see (1.6)) defined on the space $\Omega^+(S_j)$. Recall that in §3 (see (3.7)) we defined $\alpha(B, C)$ to be the frequency with which the word B occurs in the word C. For any $\omega^j, j \in Z^+$ we let

(7.1) $D_{\omega^j}(k, n) = D^j(k, n) = x(\omega^j, [k, k+n)), \quad k \in Z^+, n \in Z^+, n \neq 0 \quad$ (see (3.8)).

Lemma 3.10 tells us that if for each word C which occurs in ω^j it is true that $\lim_n \alpha(C, D^j(k, n))$ exists uniformly in k, then $(O(\omega^j)^-, T_j^+)$ is uniquely ergodic. We will show that this is actually the case.

In order to have a suggestive terminology, we make the following two definitions.

DEFINITION 7.2. *If C occurs in ω^j, then we say that C has a density d in ω^j if $\lim_n \alpha(C, D^j(k, n)) = d$ for $k \in \mathbf{Z}^+$.*

DEFINITION 7.3. *If C occurs in ω^j, then we say that C has a uniform density d in ω^j if $\lim_n \alpha(C, D^j(k, n)) = d$ uniformly in $k \in \mathbf{Z}^+$.*

Our problem is thus reduced to showing that if C is a word on the symbol set S_j, then C has a uniform density in ω^j. The remaining part of this section is devoted to this. It is technical but not difficult. We will first show that if $C = B(j, k)$ is a word of length one built on the symbol set S_j, then $B(j, k)$ has a uniform density $d(j, k)$ in ω^j. This is done by first showing that the frequency of occurrence of $B(j, k)$ in $B(j+t, s)$ converges as $t \to +\infty$ and that this limit is independent of s. This is then used to show the existence of the uniform density $d(j, k)$. We now embark on this phase of the problem. To prevent formulas from becoming too cumbersome, we will streamline our notation but will attempt to retain its suggestiveness.

In (3.7) we defined the number of occurrences of B in C and this number was designated by $N(B, C)$. We let

(7.4)
$$N(B(j, k), B(j+t, s)) = N(j, k; j+t, s)$$
$$\text{for } j, k \in \mathbf{Z}^+, 0 \leq k < n_j \text{ and } 0 \leq s < n_{j+t}.$$

In (4.16) we let $L_{j-r}(B(j, k))$ be the length of $B(j, k)$ in terms of the symbols of $S_{j-r}, j \in \mathbf{Z}^+, 0 \leq r \leq j$. We let

(7.5) $$L_j(B(j+t, s)) = L(j, j+t) \quad \text{for } j, t \in \mathbf{Z}^+ \text{ and } 0 \leq s < n_{j+t}.$$

Recall that by our construction and (4.16) this length is independent of s. It also follows immediately that

(7.6) $$L(j, j+t) = L(j+t-1, j+t) \cdot L(j, j+t-1).$$

This is merely saying that the length of $B(j+t, s)$ in terms of the symbols of S_j is the length of $B(j+t, s)$ in terms of the symbols of S_{j+t-1} times the length of these symbols in terms of the symbols S_j.

We now designate the frequency with which $B(j, k)$ occurs in $B(j+t, s)$ by

(7.7)
$$\alpha(B(j, k), B(j+t, s)) = \alpha(j, k; j+t, s) = \frac{1}{L(j, j+t)} \cdot N(j, k; j+t, s),$$
$$0 \leq \alpha(j, k; j+t, s) \leq 1.$$

Henceforth, we will always assume that the indices used in our formulas are constrained in such a manner that the objects are defined. In the case of statements about limits the objects need only be ultimately defined. For the remainder of this section we think of our words and sequences as being built on the symbol set S_j.

LEMMA 7.8. *If $j, t \in Z^+$ and $t \neq 0$, then the following statements hold:*

(a) $$N(j, k; j+t, s) = \sum_{m=0}^{n_{j+t-1}-1} N(j+t-1, m; j+t, s)N(j, k; j+t-1, m),$$

(b) $$\alpha(j, k; j+t, s) = \sum_{m=0}^{n_{j+t-1}-1} \alpha(j+t-1, m; j+t, s)\alpha(j, k; j+t-1, m).$$

Proof. (a) is immediate. It says that the number of occurrences of $B(j, k)$ in $B(j+t, s)$ is equal to the sum of the number of occurrences of $B(j, k)$ in a symbol of S_{j+t-1} multiplied by the occurrences of that symbol in $B(j+t, s)$. (b) is obtained from (a) by applying (7.7) and (7.6).

For the next lemma we estimate the difference between the frequency of occurrence of $B(j, k)$ in $B(j+t, s)$ and the frequency of occurrence of $B(j, k)$ in $B(j+t, s')$.

LEMMA 7.9. *The following inequality holds*

$$|\alpha(j, k; j+t, s) - \alpha(j, k; j+t, s')| < 2\eta_{j+t-1}.$$

Proof. We apply Lemma 7.8(b) and (7.7) and (4.5b) to obtain

$$|\alpha(j, k; j+t, s) - \alpha(j, k; j+t, s')|$$

$$\leqq \sum_{m=0}^{n_{j+t-1}-1} |\alpha(j+t-1, m; j+t, s) - \alpha(j+t-1, m; j+t, s')| |\alpha(j, k; j+t-1, m)|$$

$$< 2\eta_{j+t-1}.$$

For the next lemma we estimate how fast the frequency of occurrence of $B(j, k)$ in $B(j+t, s)$ increases as t increases.

LEMMA 7.10. *The following inequality holds*

$$|\alpha(j, k; j+t+1, s) - \alpha(j, k; j+t, s)| < \eta_{j+t} + 2\eta_{j+t-1}.$$

Proof. From Lemma 7.8(b) we have

$$\alpha(j, k; j+t+1, s) = \sum_{m=0}^{n_{j+t}-1} \alpha(j+t, m; j+t+1, s) \cdot \alpha(j, k; j+t, m).$$

We also have

$$\alpha(j, k; j+t, s) = \frac{1}{n_{j+t}} \sum_{m=0}^{n_{j+t}-1} \alpha(j, k; j+t, s).$$

We thus obtain (using (7.7) and (4.5) and Lemma 7.9)

$$|\alpha(j, k; j+t+1, s) - \alpha(j, k; j+t, s)|$$

$$\leqq \sum_{m=0}^{n_{j+t}-1} \left| \alpha(j+t, m; j+t+1, s) - \frac{1}{n_{j+t}} \right| |\alpha(j, k; j+t, m)|$$

$$+ \frac{1}{n_{j+t}} \sum_{m=0}^{n_{j+t}-1} |\alpha(j, k; j+t, m) - \alpha(j, k; j+t, s)|$$

$$< \eta_{j+t} + 2\eta_{j+t-1}.$$

We are now prepared to examine the limiting properties of the frequency of occurrence of $B(j, k)$ in $B(j+t, s)$ as $t \to +\infty$.

LEMMA 7.11. *For each $j \in Z^+$ and $0 \leq k < n_j$ there exists a number $d(j, k)$ for which*

$$\lim_{t \to \infty} \alpha(j, k; t, s) = d(j, k)$$

uniformly in j, k, and s, and $0 \leq d(j, k) \leq 1$.

Proof. Let $\varepsilon > 0$ be arbitrary. It follows from (4.6) that $\sum_{i=0}^{\infty} \eta_i < +\infty$. From Lemma 7.10 we obtain ($t > j$, $r \geq 0$)

$$|\alpha(j, k; t, s) - \alpha(j, k; t+r, s)| \leq \sum_{i=0}^{r-1} |\alpha(j, k; t+i, s) - \alpha(j, k; t+i+1, s)|$$

(7.12)
$$< \sum_{i=0}^{r-1} \eta_{t+i} + 2 \sum_{i=0}^{r-1} \eta_{t+i-1}$$

$$\leq 3 \sum_{i=t-1}^{\infty} \eta_i.$$

Since $\sum_{i=0}^{\infty} \eta_i < +\infty$, we see that for t sufficiently large the difference (7.12) is $< \varepsilon$. Thus $\alpha(j, k; t, s)$ is a Cauchy sequence in t, uniformly in j, k and s. As $t \to +\infty$, the limit must exist. Lemma 7.9 assures us the limit is independent of s. Call this limit $d(j, k)$. If in (7.12) we let $r \to +\infty$, we obtain the estimate

(7.13)
$$|\alpha(j, k; t, s) - d(j, k)| \leq 3 \sum_{i=t-1}^{\infty} \eta_i.$$

We conclude our lemma by taking limits on t and observing that the inequality $0 \leq d(j, k) \leq 1$ is forced upon us by (7.7).

We can now prove the fact that each $B(j, k)$ has a uniform density in ω^j.

THEOREM 7.14. *Each $B(j, k) \in S_j$ has uniform density $d(j, k)$ in ω^j. That is,*

$$\lim_{t} \alpha(B(j, k), D^j(s, s+t)) = d(j, k)$$

uniformly in s (see (7.1)).

Proof. Let $\varepsilon > 0$ be arbitrary. Choose an $n > k$ for which

(7.15)
$$|\alpha(j, k; n, m) - d(j, k)| < \varepsilon/5$$

for $0 \leq m < n_m$.

Let $L = L(j, n) = L_j(B(n, m))$ (see (7.5)) and choose $t_0 \in Z^+$, $t_0 \neq 0$, such that $L/t_0 < \varepsilon/5$. Suppose $t \in Z^+$ and $t > t_0$. We wish to show that

(7.16)
$$|\alpha(B(j, k), D^j(s, s+t)) - d(j, k)| < \varepsilon$$

independently of s. This will complete our theorem.

There exist r and $u \in Z^+$, $r \neq 0$, such that

$$(r-1)L < s \leq rL < (r+1)L < \cdots < (r+u)L < s+t \leq (r+u+1)L.$$

We observe that

$$(7.17) \quad N(B(j,k), D^j(s, s+t)) = L \sum_{i=0}^{u-1} \frac{1}{L} N(B(j,k), D^j(r+i)L, (r+i+1)L) + H(s, t)$$

where $0 \leq H(s, t) \leq 2L$ is the contribution due to the occurrences of $B(j, k)$ in the intervals $[s, rL)$ and $[(r+u)L, s+t)$.

We divide both sides of (7.17) by t, use the definition of $\alpha(j, k; n, m)$ and the inequality (7.15) to obtain

$$\alpha(B(j,k), D^j(s, s+t)) = d(j, k) + \left(\frac{Lu-t}{t}\right) d(j, k) + \frac{Lu}{t} G(j, k) + \frac{1}{t} H(s, t)$$

where $0 \leq G(j, k) < \varepsilon/5$ and is obtained from (7.15). Since $0 \leq Lu/t \leq 1$, $0 \leq d(i, j) \leq 1$, $0 \leq |(Lu-t)/t| \leq 2L/t < 2\varepsilon/5$, we obtain (7.16) and our theorem is complete.

We now embark on the final stage of proving that $(O(\omega^j)^-, T_j^+)$ is uniquely ergodic. We need to take an arbitrary word C which occurs in ω^j and show that it has a uniform density $d^j(C)$. To each $B(t, s)$, $t \geq j$, we assign the frequency $\alpha(C, B(t, s))$. We then average this over ω^t to obtain a number $d_t^j(C)$. We then show that $d_t^j(C) \to d^j(C)$ as $t \to +\infty$. Our final step is to show that this is the uniform density of C in ω^j.

Let us now fix a word C which occurs in ω^j. Let $t \in Z^+$ and $t \geq j$. Define

$$(7.18) \qquad f_i(t) = \alpha(C, B(t, s)) \qquad \text{when } (\omega^t)_i = B(t, s).$$

In words this says that $f_i(t)$ is the frequency of C in the ith symbol of ω^t.

LEMMA 7.19.

$$\lim_n \frac{1}{n} \sum_{i=m}^{m+n-1} f_i(t) = \sum_{s=0}^{n_t-1} \alpha(C, B(t, s)) d(t, s) =_{\text{def}} d_t^j(C)$$

uniformly in m.

Proof. Our proof follows quickly from Lemma 7.14. For we have

$$\lim_n \frac{1}{n} \sum_{i=m}^{m+n-1} f_i(t) = \lim_n \frac{1}{n} \sum_{s=0}^{n_t-1} \alpha(C, B(t, s)) \cdot N(B(t, s), D^t(m, m+n))$$

$$= \lim_n \sum_{s=0}^{n_t-1} \alpha(C, B(t, s)) \alpha(B(t, s), D^t(m, m+n))$$

$$= \sum_{s=0}^{n_t-1} \alpha(C, B(t, s)) d(t, s)$$

uniformly in m.

LEMMA 7.20. *For each word C occurring in ω^j the following limit exists*

$$\lim_t d_t^j(C) = d^j(C) \qquad \text{(see Lemma 7.19).}$$

Proof. We need only show that $d_t^j(C)$ is a Cauchy sequence in t, $t \geq j$. In order to do this it suffices to show

$$\sum_{t=j}^{\infty} |d_t^j(C) - d_{t+1}^j(C)| < +\infty.$$

We do this by estimating the differences of terms which converge to $d_t^j(C)$ and $d_{t+1}^j(C)$ (see Lemma 7.19). We first observe

$$\alpha(C, B(t+1, s)) = \frac{1}{L(j, t+1)} N(C, B(t+1, s))$$

(7.21)
$$= \frac{1}{L(j, t+1)} \sum_{s'=0}^{n_t-1} N(C, B(t, s')) \cdot N(B(t, s'), B(t+1, s))$$

$$+ \frac{1}{L(j, t+1)} \cdot H(t, s)$$

where $0 \leq H(t, s) \leq L(t, t+1) \cdot L(C)$. $H(t, s)$ is due to the occurrence of C in positions which overlap two adjacent blocks $B(t, k)$ and $B(t, k')$. $L(C)$ is the length of C as a word built on the symbol set S_j.

Using the fact that $L(j, t+1) = L(j, t)L(t, t+1)$ (see (7.6)) and the definition of frequency of C in a word, we obtain

(7.22) $\alpha(C, B(t+1, s)) = \displaystyle\sum_{s'=0}^{n_t-1} \alpha(C, B(t, s')) \cdot \alpha(B(t, s'), B(t+1, s)) + \dfrac{1}{L(j, t)} \cdot H'(t, s)$

where $0 \leq H'(t, s) \leq L(C)$.

We now use (4.5) to replace the $\alpha(B(t, s'), B(t+1, s))$ by their approximate values $1/n_t$ to obtain

(7.23) $\alpha(C, B(t+1, s)) = \dfrac{1}{n_t} \displaystyle\sum_{s'=0}^{n_t-1} \alpha(C, B(t, s')) + G(t, s) + \dfrac{1}{L(j, t)} \cdot H'(t, s)$

where $0 \leq G(t, s) \leq \eta_t$; $0 \leq H'(t, s) \leq L(C)$.

Recalling from (7.18) the fact that

(7.24) $f_i(t+1) = \alpha(C, B(t+1, s))$ when $(\omega^{t+1})_i = B(t+1, s)$

and using (7.23), we obtain

(7.25) $\left| \dfrac{1}{n} \displaystyle\sum_{i=m}^{m+n-1} f_i(t+1) - \dfrac{1}{nn_t} \displaystyle\sum_{i=mn_t}^{mn_t+nn_t-1} f_i(t) \right| < \eta_t + \dfrac{L(C)}{L(j, t)} \leq \eta_t + \dfrac{L(C)}{(t-1)^2}.$

The last estimate follows from (4.16), which asserts that $L(j, t) \geq k_j$ and (4.7) which implies $k_j \geq (t-1)^2$.

Taking limits on n in (7.25), we obtain

$$(7.26) \qquad |d^j_{t+1}(C) - d^j_t(C)| \leq \eta_t + \frac{L(C)}{(t-1)^2}.$$

Since $\sum_{t=j}^{\infty} (\eta_t + L(C)/(t-1)^2) < +\infty$, our lemma is complete.

THEOREM 7.27. *If C is a word occurring in ω^j, then C has uniform density $d^j(C)$ in ω^j. That is, $\lim_t \alpha(C, D^j(s, s+t)) = d^j(C)$ uniformly in s (see Lemma 7.20).*

Proof. Let $\varepsilon > 0$ be arbitrary. We must show that

$$(7.28) \qquad |\alpha(C, D^j(s, s+t)) - d^j(C)| < \varepsilon$$

independently of s for all sufficiently large t.

From Lemma 7.20 we see that it suffices to show that

$$(7.29) \qquad |\alpha(C, D^j(s, s+t)) - d^j_v(C)| < \varepsilon/2$$

independently of s for all sufficiently large t and a fixed v, $v \geq j$, for which $|d^j_v(C) - d^j(C)| < \varepsilon/2$ and $L(C)/L(j, v) < \varepsilon/8$.

From Lemma 7.19 we see that it suffices to show that

$$(7.30) \qquad \left| \alpha(C, D^j(s, s+t)) - \frac{1}{n} \sum_{i=m}^{m+n-1} f_i(v) \right| < \varepsilon/4$$

independently of s for all sufficiently large t and all n larger than a fixed N_0 for which

$$(7.31) \qquad \left| \frac{1}{n} \sum_{i=m}^{m+n-1} f_i(v) - d^j_v(C) \right| < \varepsilon/4.$$

We will prove (7.30) much in the same manner as we proved Theorem 7.14.

Let $s \in Z^+$ and $t \in Z^+$, $t \neq 0$. Let $L = L(j, v) = L_j(B(v, w))$. There exist integers r and u for which

$$(r-1)L < s \leq rL < (r+1)L < \cdots < (r+u)L < s+t \leq (r+u+1)L.$$

We observe that

$$(7.32) \qquad \begin{aligned} N(C, D^j(s, s+t)) = L \sum_{i=0}^{u-1} \frac{1}{L} N(C, D^j((r+i)L, (r+i+1)L)) \\ + H(s, t) + G(s, t) \end{aligned}$$

where $0 \leq H(s, t) \leq 2L$ is the part obtained from the intervals $[s, rL)$ and $[(r+u)L, s+t)$. And $0 \leq G(s, t) \leq u \cdot L(C)$, where $G(s, t)$ comes from the occurrences of C which overlap adjacent intervals $[(r+i)L, (r+i+1)L)$ and $[(r+i+1)L, (r+i+2)L)$. Dividing (7.32) through by t and referring to (7.18), we obtain

$$(7.33) \qquad \begin{aligned} \alpha(C, D^j(s, s+t)) = \frac{1}{u} \sum_{i=r}^{r+u-1} f_i(v) + \left(\frac{Lu-t}{t}\right) \frac{1}{u} \sum_{i=r}^{r+u-1} f_i(v) \\ + \frac{1}{t} H(s, t) + \frac{1}{t} G(s, t). \end{aligned}$$

We observe that

$$0 \leq \frac{1}{u} \sum_{i=r}^{r+u-1} f_i(v) \leq 1 \quad \text{and} \quad \frac{1}{t} G(s, t) = \frac{uL}{t} \cdot \frac{1}{uL} G(s, t)$$

so

$$0 \leq \frac{1}{t} G(s, t) \leq \frac{uL}{t} \cdot \frac{L(C)}{L} < \frac{uL}{t} \cdot \frac{\varepsilon}{8}.$$

Since t is asymptotic to uL, we see that (7.30) holds (where we replace m by r and n by u) independently in s for all sufficiently large t.

THEOREM 7.34. *For each $j \in \mathbf{Z}^+$ the system $(O(\omega^j)^-, T_j^+)$ is uniquely ergodic.*

Proof. The previous theorem tells us that each word C which occurs in ω^j has a uniform density $d^j(C)$ (see Definition 7.3). Theorem 3.9 then applies, giving us our conclusion.

8. Entropy of ω^j. In this section we will show that the entropy of the almost periodic uniquely ergodic systems $(O(\omega^j)^-, T_j^+)$ lies between $\log n_j(1 - \delta)$ and $\log n_j$, where δ is the constant fixed at the beginning of our construction of ω^j and is subject only to the condition $0 < \delta < 1$. We shall briefly recall the definition of the Kolmogoroff-Sinai invariant (or entropy). For more details the reader is referred to [4, p. 241 et seq.].

Let $(\Omega, \mathfrak{B}, \mu)$ be a probability space and let T be a measure preserving transformation of Ω (e.g., $\mu(T^{-1}E) = \mu(E)$ for all $E \in \mathfrak{B}$). If $\mathfrak{F}$ is a finite Borel subfield of $\mathfrak{B}$ generated by the atoms $F_1 \cdots F_n$, the entropy of $\mathfrak{F}$ is defined by

$$H(\mathfrak{F}) = \sum_{i=1}^{n} -\mu(F_i) \log \mu(F_i).$$

A simple convexity argument shows that $H(\mathfrak{F}) \leq \log n$ and equality holds when $\mu(F_i) = 1/n$, $i = 1, 2, \ldots, n$, [4, p. 242]. It is also known that the following limit exists, [4, p. 269],

$$\bar{H}(\mathfrak{F}) = \lim_n \frac{1}{n} H(\mathfrak{F} \vee T^{-1}\mathfrak{F} \vee \cdots \vee T^{-n+1}\mathfrak{F})$$

where $\mathfrak{A} \vee \mathfrak{C}$ is used to designate the Borel field generated by the two Borel fields $\mathfrak{A}$ and $\mathfrak{C}$. The Kolmogoroff-Sinai invariant or the entropy of T with respect to the measure μ is defined as follows:

(8.1) $\hat{H}(T, \mu) = \sup \{\bar{H}(\mathfrak{F}): \mathfrak{F} \subset \mathfrak{B}, \mathfrak{F} \text{ a finite Borel field}\}.$

We will be interested in estimating the entropy of the uniquely ergodic systems $(O(\omega^j)^-, T_j^+)$.

We let

μ_j = unique invariant probability measure of $(O(\omega^j)^-, T_j^+)$,

(8.2) $\hat{H}_j = \hat{H}(T_j^+, \mu_j),$

$\mathfrak{B}_j$ = Borel field of $O(\omega_j)^-$.

We let $\mathfrak{F}_j^n$ be the finite Borel field generated by sets of the form

$$C(B, I^+(n)) \cap O(\omega^j)^-$$

(see (1.3), (1.4), (1.5)), where B is any word of length n built on the symbol set S_j. It is known, [4, p. 279], that

(8.3) $$\hat{H}_j = \lim_n \frac{1}{n} H(\mathfrak{F}_j^n).$$

LEMMA 8.4. *The entropy T_j^+ is bounded above by the combinatorial entropy of ω^j. That is, $\hat{H}_j \le K(\omega^j)$ (in this respect see [9, Theorem 6]).*

Proof. Since ω^j is almost periodic (§5), we see that $C(B, I^+(n)) \cap O(\omega^j)^-$ is not empty if and only if B occurs in ω^j (Lemma 2.6). Using the properties of $H(\mathfrak{F})$ and the definition of $\theta(n, \omega^j)$ (§6, second paragraph), we have

$$H(\mathfrak{F}_j^n) \le \log \theta(n, \omega^j).$$

Dividing by n and taking limits on n we obtain our result.

Our next few lemmas are needed to bound $\hat{H}_j$ from below. To do this, we will need to explicitly determine the μ_j measure of certain sets. We will use Lemma 3.10 for this, and we now make some preparatory remarks. Let $j \in Z^+$ and let S'_{j+1} be the set of *all* words of length r_{j+1} which are built on the symbol set S_j. We see that $S'_{j+1} \supset S_{j+1}$ (see (4.5)). We define the association map

$$\pi_{j,j+1} : \Omega^+(S_j) \to \Omega^+(S'_{j+1})$$

by defining the ith coordinate of $\pi_{j,j+1}\omega$,

(8.5) $(\pi_{j,j+1}\omega)_i$ = that word of length r_{j+1} which occurs in ω at the ir_{j+1}th place
                     $= x(\omega, [ir_{j+1}, (i+1)r_{j+1}))$.

We see easily that $\pi_{j,j+1}$ is continuous and that $\pi_{j,j+1} \circ (T_j^+)^{r_{j+1}} = T_{j+1}^+ \circ \pi_{j,j+1}$. If $s \ge j$, we define $\pi_{j,s}$ as iterates of mappings of the previous type and observe that

(8.6) $$\pi_{j,s} \circ (T_j^+)^q = T_s^+ \circ \pi_{j,s}$$

where $q = r_{j+1} \cdots r_s$.

From our construction of the sequences ω^j and the continuity of $\pi_{j,s}$ we see that $\pi_{j,s}(O(\omega^j)^-) = O(\omega^s)^-$.

If $s \ge j$, then $B(s, k)$, $0 \le k < n_s$, is a word of length $L_j(B(s, k))$ occurring in ω^j and is also a word of length one occurring in ω^s. We let

(8.7) $\tilde{B}_j(s, k) = C(B(s, k), [0, L_j(B(s, k)))) \cap O(\omega^j)^- \ne \varnothing$ (see (1.3))

where $B(s, k)$ is considered as a word of length $L_j(B(s, k))$ built on the symbols S_j.

$$\tilde{B}_s(s, k) = C(B(s, k), \{0\}) \cap O(\omega^s)^- \ne \varnothing \qquad \text{(see (1.3))}$$

where $B(s, k)$ is considered as a word of length one built on the symbols S_s. In words this means that $\tilde{B}_j(s, k)$ is the set of all sequences of ω^j which begin with the

word $B(s, k)$, and $\tilde{B}_s(s, k)$ is the set of all sequences of ω^s which begin with the symbol $B(s, k)$.

LEMMA 8.8.

$$\mu_j(\tilde{B}_j(s, k)) = d(s, k)$$

(see (8.2) and Lemma 7.11).

Proof. We see that $\pi_{js}^{-1}(\tilde{B}_s(s, k)) = \tilde{B}_j(s, k)$. By Lemma 3.10 we have

$$\mu_j(\pi_{js}^{-1}(\tilde{B}_s(s, k))) = \mu_s(\tilde{B}_s(s, k))$$

and thus it suffices to show

$$\mu_s(\tilde{B}_s(s, k)) = d(s, k).$$

This is the immediate consequence of Lemma 7.14 which asserts that $B(s, k)$ has uniform density $d(s, k)$ in ω^s and we then apply Theorem 3.9.

LEMMA 8.9. $|1 - n_s d(s, k)| < (1/n_s) \cdot 1/2^{s-3}$ for $0 \leq k < n_s$.

Proof. To prove this statement, we use the inequalities of (4.5b) and (7.13) to obtain (see (7.7) for definition of $\alpha(s, k; s+1, 0)$)

$$\left| \frac{1}{n_s} - d(s, k) \right| \leq \left| \frac{1}{n_s} - \alpha(s, k; s+1, 0) \right| + |\alpha(s, k; s+1, 0) - d(s, k)|$$

$$< \frac{\eta_s}{n_s} + 3 \sum_{i=s}^{\infty} \eta_i.$$

We multiply by n_s and apply (4.6) and (4.7) to obtain

$$|1 - n_s d(s, k)| < \eta_s + 3n_s \sum_{i=s}^{\infty} \eta_i$$

$$< \frac{1}{n_s^2} \cdot \frac{1}{2^s} + 3 \sum_{i=s}^{\infty} \frac{1}{n_i} \cdot \frac{1}{2^i}$$

$$< \frac{1}{n_s 2^{s-3}}.$$

LEMMA 8.10. *Let $\varepsilon > 0$ be arbitrary. Then there is an s_0 for which*

$$|n_s d(s, k) \log d(s, k) - \log (1/n_s)| < \varepsilon$$

for all $s > s_0$ and $0 \leq k < n_s$.

Proof. Let $n_s \cdot d(s, k) = \gamma(s, k)$, then $\log d(s, k) - \log (1/n_s) = \log \gamma(s, k)$. Moreover, by Lemma 8.9

$$|1 - \gamma(s, k)| < (1/n_s) \cdot 1/2^{s-3} \qquad \text{for } 0 \leq k < n_s;$$

thus

$$n_s d(s, k) \log d(s, k) - \log \frac{1}{n_s} = \gamma(s, k) \log \gamma(s, k) + (\gamma(s, k) - 1) \log \frac{1}{n_s}.$$

Thus

$$\left| n_s d(s, k) \log d(s, k) - \log \frac{1}{n_s} \right| \leq |\gamma(s, k) \log \gamma(s, k)| + \frac{1}{n_s} \cdot \frac{1}{2^{s-3}} \cdot \left| \log \frac{1}{n_s} \right|.$$

The function $x \log x$ is uniformly continuous for $0 \leq x \leq 2$. For large s, $\gamma(s, k)$ is uniformly close to 1 for $0 \leq k < n_s$ and $(1/n_s) \cdot (1/2^{s-3}) \cdot \log(1/n_s)$ is also small. These observations conclude our lemma.

THEOREM 8.11. *For each $j \in Z^+$ we have*

$$\log n_j (1 - \delta) \leq \hat{H}_j < \log n_j$$

where δ is the constant used in §4 to construct ω^j and is subject only to the condition $0 < \delta < 1$.

Proof. The fact that $\hat{H}_j < \log n_j$ follows from Lemma 8.4 and Lemma 6.5 which state

$$\hat{H}_j \leq K(\omega_j) < \log n_j.$$

To obtain the other half of the inequality we proceed as follows: let $s > j$ and let $q_s = L_j(B(s, k))$. $\mathfrak{F}_j^{q_s}$ is the Borel field defined in (8.2). We have

$$\hat{H}_j \geq \lim_s \frac{1}{q_s} H(\mathfrak{F}_j^{q_s}).$$

If $\varepsilon > 0$ is arbitrary, we have (Lemma 8.10)

(8.12)
$$\left| \sum_{k=0}^{n_s-1} d(s, k) \log d(s, k) - \frac{1}{n_s} \log \frac{1}{n_s} \right|$$

$$\leq \frac{1}{n_s} \sum_{k=0}^{n_s-1} \left| n_s d(s, k) \log d(s, k) - \log \frac{1}{n_s} \right| < \varepsilon$$

for all sufficiently large s.

Thus using Lemma 8.8 and (8.12),

$$H(\mathfrak{F}_j^{q_s}) \geq \sum_{k=0}^{n_s-1} -\mu_j(\tilde{B}(s, k)) \log \mu_j(\tilde{B}(s, k))$$

(8.13)
$$= \sum_{k=0}^{n_s-1} -d(s, k) \log d(s, k)$$

$$> \left(\sum_{k=0}^{n_s-1} -\frac{1}{n_s} \log \frac{1}{n_s} \right) - \varepsilon$$

$$= (\log n_s) - \varepsilon.$$

In (6.10) we saw that

$$n_s \geq (n_j \cdot (1 - \delta))^{r_s \cdots r_{j+1}}.$$

We also have (4.16)

$$q_s = L_j(B(s, k)) = r_s \cdots r_{j+1}.$$

Using these and (8.13) we obtain

$$(1/q_s)H(\mathfrak{F}_j^{q_s}) \geqq (\log n_j(1-\delta)) - \varepsilon/q_s.$$

We let $s \to +\infty$ and we get

$$\hat{H}_j \geqq \log n_j(1-\delta)$$

which concludes our theorem.

COROLLARY 8.14. *Let* $0 \leqq M < +\infty$. *There exists a compact metric space* X *and a continuous mapping* $T: X \to X$ *such that the following conditions hold.*

(a) *For each* $x \in X$ *we have* $X = $ *closure* $\{T^j x : j \in \mathbf{Z}^+\}$. *That is,* (X, T) *is minimal,* [2].

(b) T *is almost periodic at each* $x \in X$.

(c) (X, T) *has a unique* T *invariant probability measure.*

(d) *The entropy satisfies the bounds* $M < \hat{H}(T) < +\infty$.

We remark here that our construction of §4 allows us considerable freedom in the invariant measure μ_j. We have chosen them so that $\mu_j(\tilde{B}(j, k)) = d(j, k)$ is close to $1/n_j$. We actually only need this for all sufficiently large j. Thus, for instance, we could have chosen a vastly different distribution of the $B(0, t)$ in the $B(1, s)$'s. However, in order to make the entropy high, eventually the $d(j, k)$ must be close to $1/n_j$.

9. **Case of bisequences.**   A slight modification of our construction of §4 allows us to carry over our results to bisequences. We modify (4.3) and (4.5) as follows:

(4.3′) A_{j-1} is the initial word of $B(j, 0)$ and the terminal word of $B(j, n_j - 1)$.

(4.5′) For $j > 0$, S_j is the set of all distinct words on the symbol set S_{j-1} whose length in terms of these symbols is $r_j = n_{j-1}k_j$ and which have the following properties:

(a) $A_{j-1}A_{j-1}$ occurs in each $B(j, k) \in S_j$,

(b) this remains the same as (4.5b).

Our inductive construction remains valid and (4.16) holds.

As in (4.17), we let $\omega^j = \bigcup_{i=j}^{\infty} A_i \in \Omega^+(S_j)$. Let $\Omega^*(S_j)$ be the set of all functions on the negative integers $\mathbf{Z}^* = \{-1, -2, -3, \ldots\}$ whose values are in S_j. We now define $\tilde{\omega}^j \in \Omega^*(S_j)$ as follows: If $i \in \mathbf{Z}^*$, choose any $m \geqq j$ for which $L_j(A_m) \geqq i$. We let $(\tilde{\omega}^j)_i$ be the symbol which occurs in the $(L_j(A_m) + i)$th place of A_m. (4.3′) assures us that $\tilde{\omega}^j$ is unambiguously defined.

We define $\dot{\omega}^j \in \Omega(S_j)$ as follows: Let $i \in \mathbf{Z}$ and set

$$\dot{\omega}_i^j = \tilde{\omega}_i^j \quad \text{if } i = -1, -2, -3, \ldots,$$

$$\dot{\omega}_i^j = \omega_i^j \quad \text{if } i = 0, 1, 2, \ldots.$$

The computations of §§5, 6, 7, and 8 are valid and we obtain

THEOREM 9.6. *For each $j \in Z^+$ the following statements are true*:

(a) $\dot{\omega}^j$ *is an almost periodic sequence*,

(b) $\log n_j(1 - \delta) \leqq K(\dot{\omega}^j) < \log n_j$,

(c) $(O(\dot{\omega}^j)^-, T_j)$ *is uniquely ergodic*,

(d) $\log n_j(1 - \delta) \leqq \hat{H}(T_j) \leqq K(\dot{\omega}^j) < \log n_j$, *where* δ *is subject only to the condition* $0 < \delta < 1$.

Thus we have

COROLLARY 9.7. *Let* $0 < M < +\infty$. *There is a compact metric space* X *and a homeomorphism* T *of* X *for which the following statements hold*:

(a) (X, T) *is minimal*. (*In* [2] *it is shown that this implies that* T *is almost periodic at each* $x \in X$.)

(b) (X, T) *is uniquely ergodic*.

(c) *The entropy satisfies the bounds* $M < \hat{H}(T) < +\infty$.

BIBLIOGRAPHY

1. R. Adler, A. Konheim and M. McAndrew, *Topological entropy*, Trans. Amer. Math. Soc. **114** (1965), 309–319.

2. H. Furstenberg, *Strict ergodicity and transformations of the torus*, Amer. J. Math. **83** (1961), 573–601.

3. W. Gottschalk and G. Hedlund, *Topological dynamics*, Amer. Math. Soc. Colloq. Publ., Vol. 36, Amer. Math. Soc., Providence, R. I., 1955.

4. K. Jacobs, *Lecture notes on ergodic theory*, Parts I and II, Mat. Inst. Aarhus Univ., 1962–1963.

5. M. Loève, *Probability theory*, 3rd ed., Van Nostrand, Princeton, N. J., 1963.

6. M. Morse and G. Hedlund, *Symbolic dynamics*, Amer. J. Math. **60** (1938), 815–866.

7. ———, *Symbolic dynamics*. II, Amer. J. Math. **62** (1940), 1–42.

8. J. Oxtoby, *Ergodic sets*, Bull. Amer. Math. Soc. **58** (1952), 116–136.

9. W. Parry, *Intrinsic Markov chains*, Trans. Amer. Math. Soc. **112** (1964), 55–66.

YALE UNIVERSITY,
        HEW HAVEN, CONNECTICUT
UNIVERSITY OF CALIFORNIA,
        BERKELEY, CALIFORNIA
STANFORD UNIVERSITY,
        STANFORD, CALIFORNIA