SEQUENCE ENTROPY AND g-MEASURES

by

Paul Hulse

A thesis submitted for the degree of Doctor of Philosophy at the University of Warwick, Mathematics Institute, September, 1980.

CONTENTS

Acknowledgements

I would like to thank my supervisor, Dr. P. Walters, for his help and advice. I would also like to thank the Science Research Council for their financial support.

Declaration

The results in Chapter I, sections 3 and 4, and Theorem 1.13 in section 5 have appeared in the Journal of the London Mathematical Society, Vol.20 (1979) under the title 'On the sequence entropy of transformations with quasi-discrete spectrum'.

SUMMARY

This thesis presents some results on sequence entropy and g-measures.

Chapter I is concerned with the sequence entropy $h_A(T)$ of transformations with quasi-discrete spectrum. In [10], it was shown that T has discrete spectrum if and only if $h_A(T)$ is zero for all sequences A. This prompts the question: If T has quasi-discrete spectrum but not discrete spectrum, for which sequences is $h_A(T)$ positive? We first consider this problem for affine transformations on the torus and calculate the sequence entropies for certain types of sequence. In the general case, we obtain sufficient conditions on the sequence for zero and non-zero sequence entropy. With a suitable restriction imposed on the sequence, we get a necessary and sufficient condition for zero sequence entropy. Next, we determine a class of sequences for which $h_A(T)$ is infinite whenever T has quasi-discrete spectrum but not discrete spectrum and a larger class for which $h_A(T)$ is infinite whenever T has quasi-eigenfunctions of arbitrarily large order. An example is given to show that this last result does not characterize such transformations.

In [13] and unpublished work by Walters, $\sup_A h_A(T)$ was calculated for ergodic transformations. In Chapter II, §1, using the same method, we extend this result to show that the supremum is always attained. We then deduce necessary and sufficient conditions for weak-mixing and strong-mixing in terms of sequence entropy, strengthening similar results in [15]. In §2, we use a construction in [5] to construct sets W of arbitrarily small measure such that $\{T^{n_i}W\}_1^\infty$ generates the full σ-algebra, where $\{T^{n_i}\}_1^\infty$ converges weakly to the identity. By combining this with the results of §1, we deduce the existence of a transformation T and a sequence A such that $h_A(T)$ is infinite and there exist subsequence generators for A with arbitrarily small entropy. This contrasts with the case $A = \{n\}$, where the existence of a generator with finite entropy implies the entropy of T is finite and is the infimum of the entropies of the generators.

In Chapter III, we consider the uniqueness problem for g-measures. It is not known if g-measures are unique in general. However, a sufficient condition for uniqueness in terms of the variation of log g has been given in [18]. We construct examples to show this condition is not necessary for uniqueness.

INTRODUCTION

This thesis is concerned with two areas of ergodic theory, namely sequence entropy and g-measures.

The sequence entropy $h_A(T)$ of an automorphism T of a Lebesque space with respect to a sequence of integers A was introduced as an isomorphism invariant by Kushnirenko in [10]. If A is the sequence $\{n\}$, then $h_A(T)$ is just the usual entropy of T. In [10], Kushnirenko showed that sequence entropy did not reduce to the already known invariants of spectral isomorphism and entropy by giving an example of two spectrally isomorphic transformations with zero entropy but with different sequence entropies for the sequence $\{2^n\}$. In fact, it turns out that sequence entropy is only useful as an invariant when dealing with transformations with zero entropy, since Newton and Krug [12] showed that if the entropy is non-zero, then the sequence entropy depends only on the sequence and the entropy. Kushnirenko also showed that T has discrete spectrum if and only if $h_A(T)$ is zero for all sequences A. The concepts of eigenfunction and discrete spectrum can be generalized to those of quasi-eigenfunction and quasi-discrete spectrum. Abramov [1] has classified transformations with quasi-discrete spectrum and shown that they have zero entropy. In view of Kushnirenko's result on discrete spectrum and sequence entropy, it is natural to ask which sequences give positive sequence entropy when T has quasi-discrete spectrum but not discrete spectrum. In Chapter I, we first consider this question when T is an affine transformation on the torus and calculate the sequence entropies for certain types of sequences. This extends a result in [10] where the sequence entropy was calculated

for the sequence $\{2^n\}$. When considering the general case, we use Abramov's classification of transformations with quasi- discrete spectrum as affine transformations of compact connected abelian groups. This enables a sequence of factor automorphisms on finite dimensional tori to be constructed and by considering these factor automorphisms we obtain sufficient conditions on the sequence for zero and non-zero sequence entropy. With a suitable restriction imposed on the sequence, we get a necessary and sufficient condition for zero sequence entropy which depends only on the sequence. Next, we determine a class of sequences for which $h_A(T)$ is infinite whenever T has quasi-discrete spectrum but not discrete spectrum. Then we determine a class of sequences for which $h_A(T)$ is infinite whenever T has quasi-eigenfunctions of arbitrarily large order. This result does not characterize such transformations and we give an example where the quasi-eigenfunctions of T are all of order either one or two, but $h_A(T)$ is infinite for this class of sequences.

In Chapter II, §1, we consider connections between sequence entropy and the spectrum of an automorphism. Kushnirenko's result on sequence entropy and discrete spectrum has been extended by Pickel [13] and independently in unpublished work by Walters to show that if T is ergodic then $\sup_A h_A(T)$ is log k for some integer k or infinite. Saleski [15] has shown that T is weak-mixing if and only if $\sup_A h_A(T, \xi)$ is $H(\xi)$ for all partitions ξ with finite entropy, and strong-mixing if and only if $\sup_{B \subset A} h_B(T, \xi)$ is $H(\xi)$ for all partitions ξ with finite entropy and all increasing sequences A. We extend these results, using the

same method as Pickel and Walters, to show that the supremums are always attained. In §2, we consider subsequence generators, that is, partitions ξ such that $\{T^n\xi\}_{n\in N}$ generates the full σ-algebra $\mathcal{B}$, where N is a given subsequence of the integers. Ellis and Friedman [5] have shown that for any ergodic translation on a compact abelian group and any infinite subsequence of integers N there is a set W of arbitrarily small measure such that $\{T^nW\}_{n\in N}$ generates the full σ-algebra. Their method applies naturally to the setting of aperiodic automorphisms such that T^{n_i} converges weakly to the identity for some subsequence $\{n_i\}$, and in this case we use their method to prove the existence of arbitrarily small generators for the sequence $\{n_i\}$. By combining this result with the results of §1, we deduce the existence of an automorphism T and a sequence A such that $h_A(T)$ is infinite and there exist subsequence generators for A with arbitrarily small entropy. This result is in marked contrast to the situation when A = {n}. In this case, the existence of a generator with finite entropy implies the entropy of T is finite and is the infimum of the entropies of the generators.

In Chapter III, we turn our attention to g-measures, which were studied by Keane in [9]. We will consider the case where T is a one-sided subshift of finite type. The notion of g-measure occurs in statistical mechanics and probability theory, g being a positive, continuous function such that $\sum_{y\in T^{-1}x} g(y) = 1$ for all x. In the context of statistical mechanics, a g-measure is an equilibrium state for log g, that is, a T-invariant probability measure μ which

maximizes the quantity $h_\mu(T) + \int \log g \, d\mu$. Alternatively, if $\mathcal{E}$ is the partition into points, $T^{-1}\mathcal{E}$ is smaller than $\mathcal{E}$ and each set in $T^{-1}\mathcal{E}$ consists of a finite number of points. Given any Borel probability measure μ, there is an essentially unique canonical system of conditional measures $\{\mu_C\}_{C\in T^{-1}\mathcal{E}}$ (see [14, p.6]). These measures have the properties that each μ_C is a probability measure on C and they define a measurable function $g_\mu(x) = \mu_{T^{-1}Tx}(x)$ such that

$$\int_X f \, d\mu = \int_X \sum_{y\in T^{-1}x} g_\mu(y) f(y) \, d\mu \quad \text{for all } f \in L^1(\mu).$$

A T-invariant probability measure μ is a g-measure if $g_\mu = g$. Ledrappier [11] has shown these definitions are equivalent. It is natural to ask whether or not g-measures are unique. Non-unique g-measures correspond to phase-transition in statistical mechanics and in the context of probability theory, it is equivalent to asking whether or not a canonical system of conditional measures for $T^{-1}\mathcal{E}$ defined by a continuous function uniquely determines the corresponding measure. In general, the answer is not known. However, Walters [18] has given a sufficient condition for uniqueness in terms of the variation of log g. We show this condition is not necessary for uniqueness by constructing a class of functions which have unique g-measures but do not satisfy the condition.

CHAPTER I

ON THE SEQUENCE ENTROPY OF TRANSFORMATIONS WITH QUASI-DISCRETE SPECTRUM

Throughout this chapter, T is an automorphism, that is, an invertible measure-preserving transformation, of a Lebesgue space $(X,\mathcal{B},m)$. Whenever X is a compact abelian group, we will take it for granted that m is Haar measure.

§1. Sequence entropy

The notations and properties of entropy used can be found in [14]. In particular, $\mathcal{E}$ is the partition of a Lebesgue space into points and Z denotes the set of measurable partitions with finite entropy. Z is a complete, separable metric space with metric ρ defined by

$$\rho(\xi,\eta) = H(\xi|\eta) + H(\eta|\xi) .$$

Definition 1.1

If $A = \{t_n\}_{n=1}^{\infty}$ is a sequence of integers, the sequence entropy $h_A(T)$ of T with respect to A is defined as follows.

Let $\xi \in Z$. Then $h_A(T,\xi) = \limsup_{n\to\infty} \frac{1}{n} H(\bigvee_{i=1}^{n} T^{t_i}\xi)$

and $h_A(T) = \sup_{\xi\in Z} h_A(T,\xi)$.

Lemma 1.2 (see [10])

For any $\xi, \eta \in Z$ and sequence A

$$|h_A(T,\xi) - h_A(T,\eta)| \leq \rho(\xi,\eta) \quad .$$

Lemma 1.3 (see [10])

Let $\{\xi_k\}$ be an increasing sequence of partitions in Z such that $\bigvee_k \xi_k = \mathcal{E}$. Then

$$h_A(T) = \lim_{k\to\infty} h_A(T,\xi_k)$$

for any sequence A.

Let $\{\xi_n\}$ be an increasing sequence of T-invariant partitions and $\{T_n\}$ the corresponding factor automorphisms of T. If $\xi_n \nearrow \mathcal{E}$ then we write $T_n \nearrow T$.

Lemma 1.4

Let A be any sequence and let T' be a factor automorphism of T. Then $h_A(T') \leq h_A(T)$.

If $\{T_n\}$ is a sequence of factor automorphisms of T such that $T_n \nearrow T$, then $h_A(T_n) \nearrow h_A(T)$.

The proof of Lemma 1.4 is the same as that given in [14] for the case $A = \{n\}$.

If $A = \{t_n\}_{n=1}^{\infty}$ is a sequence of integers, let $\bar{t}_n = \sup_{1\leq j\leq n} |t_j|$.

If N is a subsequence of the natural numbers $\mathbb{N}$, N_n will denote the set $N \cap \{1,\ldots,n\}$.

The upper density of N is $\limsup_{n\to\infty} \frac{|N_n|}{n}$, and the density is $\lim_{n\to\infty} \frac{|N_n|}{n}$ if this limit exists.

§2. Quasi-discrete spectrum

In this section, we give a brief account of the theory of automorphisms with quasi-discrete spectrum. For a more detailed account and proofs of results stated here, see [1].

Definitions 1.5

1. The group G of quasi-eigenfunctions of T is defined inductively as follows.

   G_1 is the set of normalized eigenfunctions of T, that is,

   $$\{f \in L^2(X) : \|f\| = 1 \ , \ fT = \lambda f \quad \text{for some } \lambda \in \mathbb{C}\}.$$

   Having defined G_n, let

   $$G_{n+1} = \{f \in L^2(X) : \|f\| = 1 \ , \ fT = gf \text{ for some } g \in G_n\}.$$

   Then $G = \bigcup_{n \geq 1} G_n$.

2. An automorphism T has quasi- discrete spectrum if T is totally ergodic and G spans $L^2(X)$.

It can be easily proved by induction that each G_n is a group under pointwise multiplication, as is G, and that $G_n \subseteq G_{n+1}$ for all n. Also, it follows from the above definition that either $G_n = G_{n+1} = \ldots = G$ for some n, or $G_{n+1} \setminus G_n$ is non-empty for all n. In the first case, $n(T)$ will denote the smallest integer n for which $G_n = G_{n+1}$, otherwise we put $n(T) = \infty$. The order of a quasi-eigenfunction f is the least integer n such that $f \in G_n$. If $fT = gf$ then g is the quasi-eigenvalue corresponding to f. The condition that T is totally ergodic is equivalent to the condition

that G has no elements of finite order, apart from 1, which in turn is equivalent to the condition that quasi-eigenfunctions corresponding to different quasi-eigenvalues are orthogonal. Quasi-eigenfunctions corresponding to the same quasi-eigenvalue are multiples of each other, so if T has quasi-discrete spectrum then G contains an orthonormal basis for $L^2(X)$. Automorphisms with quasi-discrete spectrum can be characterized as follows.

<u>Theorem 1.6</u> (see [1])

Let T be an automorphism with quasi-discrete spectrum. Then T is isomorphic to an affine transformation T' of a compact connected abelian group X', given by $T(x) = S(x) + \alpha$ where S is a group automorphism of X' and $\alpha \in X'$. Moreover, under the isomorphism, $G = K \times \Gamma$ where K is the unit circle in $\mathbb{C}$ and Γ is the character group of X'. If $\gamma \in \Gamma$ then $\gamma \in G_n$ if and only if $\gamma_0(S - I)^n = 1$.

In general, if T is an affine transformation of a compact connected abelian group, then $G \subsetneq K \times \Gamma$ and if γ is a character then $\gamma \in G_n$ if and only if $\gamma_0(S - I)^n = 1$ where S is the group automorphism part of T. Since the linear subspace generated by Γ is dense in $L^2(X)$, in order to show that T has quasi-discrete spectrum, it is enough to show that T is totally ergodic and for each $\gamma \in \Gamma$ there exists n such that $\gamma_0(S - I)^n = 1$.

The following two well-known results will come in useful for calculating sequence entropies.

Lemma 1.7

Let T be an automorphism with quasi-discrete spectrum. There exists a sequence of factor automorphisms $\{T_n\}$ of T such that $T_n \nearrow T$ where T_n is an affine transformation of a finite dimensional torus of the form $T_n(x) = S_n(x) + \alpha_n$ where α_n is constant and S_n is represented by an integer matrix (a_{ij}) where $a_{ij} = 0$ for $i < j$ and $a_{ii} = 1$.

Proof. By Theorem 1.6, we may assume that T is an affine transformation of a compact connected abelian group X given by $T(x) = S(x) + \alpha$ where S is a group automorphism of X. Let $\Gamma = \{\gamma_1, \gamma_2, \ldots\}$ be the character group of X. Then S induces a map $S^*: \Gamma \to \Gamma$ given by $S^*(\gamma) = \gamma \circ S$. For each n, define Γ_n to be the smallest subgroup of Γ containing $\{\gamma_1, \ldots, \gamma_n\}$ which is completely invariant under S^*. Since for all $\gamma \in \Gamma$ there exists an r such that $\gamma \circ (S - I)^r = 1$, it follows that Γ_n is finitely generated. Let H_n be the largest subgroup of X such that $\Gamma_n(H_n) = 1$. Then Γ_n is the character group of X/H_n in an obvious way and hence X/H_n is a finite dimensional torus. Also, $SH_n = H_n$. Let $\xi(H_n)$ denote the partition of X into H_n-cosets and let T_n be the automorphism of X/H_n given by $T_n(x + H_n) = S(x) + \alpha + H_n$. Then T_n is a factor of T. Since $\Gamma_n \nearrow \Gamma$, it follows that $\xi(H_n) \nearrow \mathcal{E}$, and hence $T_n \nearrow T$. The fact that the automorphisms T_n are of the required form follows from Lemma 2 in [17].

Lemma 1.8

Let T be an automorphism with quasi-discrete spectrum such that $n(T) \geqslant n$. Then the transformation on $\mathbb{R}^n/\mathbb{Z}^n$ given by $(x_1, x_2, \ldots, x_n) \longmapsto (x_1+\alpha,\ x_2+x_1, \ldots,\ x_n+x_{n-1}) \bmod 1$ is a factor of T for some $\alpha \in [0,1)$.

Proof. Since T is totally ergodic, all eigenvalues have modulus 1 and are not roots of unity, and all quasi-eigenfunctions map X into K, the unit circle in $\mathbb{C}$. Choose $\lambda \in K$ and $f_1, \ldots, f_n \in L^2(X)$ such that $f_1 T = \lambda f_1$ and $f_i T = f_{i-1} f_i$ for $2 \leqslant i \leqslant n$. Define a map $\phi : X \longrightarrow K^n$ by $\phi(x) = (f_1(x), \ldots, f_n(x))$. Since the functions f_i are measurable, ϕ is measurable. If $T' : K^n \longrightarrow K^n$ is the map given by $T'(z_1, \ldots, z_n) = (\lambda z_1,\ z_1 z_2, \ldots,\ z_{n-1} z_n)$, it is clear that

$$T'\phi = \phi T \qquad (1)$$

The map ϕ induces a measure μ on K^n given by $\mu(A) = m(\phi^{-1}A)$. It follows from (1) that μ is T'-invariant, but since λ is not a root of unity, T' is uniquely ergodic and so μ is Lebesgue measure. Hence, ϕ is measure-preserving and so T' is a factor of T.

§3. Transformations on the torus

Examples of automorphisms with quasi-discrete spectrum are affine transformations on the 2-torus given by the formula $(x,y) \longmapsto (x + \alpha, rx + y + \beta) \bmod 1$ where $r \in \mathbb{Z}\setminus\{0\}$, α and β are constants and α is irrational. In this case, $n(T) = 2$.

Theorem 1.9

Let T be an affine transformation of $\mathbb{R}^2/\mathbb{Z}^2$ given by $T(x,y) = (x + \alpha, rx + y + \beta) \bmod 1$ $(r \in \mathbb{Z}\setminus\{0\}$, α and β constants), and let $A = \{t_n\}_{n=1}^{\infty}$ be a sequence of integers.

(i) If $\bar{t}_n / \bar{t}_{n-1}$ is bounded then $h_A(T) = \limsup_{n\to\infty} \frac{1}{n} \log \bar{t}_n$.

(ii) If there exists a subsequence N of $\mathbb{N}$ with positive upper density such that $\lim_{\substack{n\to\infty \\ n\in N}} \bar{t}_n / \bar{t}_{n-1} = \infty$, then $h_A(T) = \infty$.

Proof. Let ξ_k $(k \geqslant 2)$ denote the partition of $\mathbb{R}^2/\mathbb{Z}^2$ into k^2 squares defined by the cycles $\{x = j/k\}$, $\{y = j/k\}$ $(0 \leqslant j \leqslant k-1)$. The boundaries of sets in $\bigvee_{i=1}^{n} T^{t_i}\xi_k$ are formed by cycles which are the images of $\{x = j/k\}$, $\{y = j/k\}$ under T^{t_i} for $1 \leqslant i \leqslant n$. If two such cycles are the images of T^{t_i} and T^{t_j} where $|t_i| \leqslant |t_j|$ then the number of times the cycles intersect is at most $2|rt_j|$. Since there are no more than 2kn cycles, the number of intersections on any one

cycle is at most $4kn|r|\bar{t}_n$. Therefore, the number of edges of sets in $\bigvee_{i=1}^{n} T^{t_i}\xi_k$ lying on any one cycle is at most $4kn|r|\bar{t}_n$, so the total number of edges is at most $8k^2n^2|r|\bar{t}_n$. Every set in $\bigvee_{i=1}^{n} T^{t_i}\xi_k$ has at least one edge and each edge belongs to two sets, so the number of sets in $\bigvee_{i=1}^{n} T^{t_i}\xi_k$ is at most $16k^2n^2|r|\bar{t}_n$. Therefore,

$$H(\bigvee_{i=1}^{n} T^{t_i}\xi_k) \leqslant \log 16k^2n^2|r|\bar{t}_n$$

and so,

$$h_A(T,\xi_k) \leqslant \limsup_{n\to\infty} \frac{1}{n} \log \bar{t}_n$$

for all $k \geqslant 2$. Since $\bigvee_{k=2}^{\infty} \xi_k = \varepsilon$,

$$h_A(T) \leqslant \limsup_{n\to\infty} \frac{1}{n} \log \bar{t}_n \quad . \qquad (2)$$

For a measurable set C in $\mathbb{R}^2/\mathbb{Z}^2$, let C_y denote the set $\{x \in \mathbb{R}/\mathbb{Z} : (x,y) \in C\}$.

If $C \in \xi_k \vee (\bigvee_{i=1}^{n} T^{t_i}\xi_k)$ and $y \in \mathbb{R}/\mathbb{Z}$, then C_y is contained in at most

$$|r| \prod_{i=1}^{n} \left[\frac{|t_i|/\ \bar{t}_{i-1} + 1 + k}{k}\right]$$

intervals of length

$$\frac{1}{k|r|\bar{t}_n} \qquad (\text{taking } \bar{t}_0 = 1). \quad (3)$$

(Here and later, $[x]$ denotes the integer part of x).

To prove this, note that (3) is true for $n = 0$ (taking the empty product to be 1). Assume it is true for $n-1$.

cycle is at most $4kn|r|\bar{t}_n$. Therefore, the number of edges of sets in $\bigvee_{i=1}^{n} T^{t_i}\xi_k$ lying on any one cycle is at most $4kn|r|\bar{t}_n$, so the total number of edges is at most $8k^2n^2|r|\bar{t}_n$. Every set in $\bigvee_{i=1}^{n} T^{t_i}\xi_k$ has at least one edge and each edge belongs to two sets, so the number of sets in $\bigvee_{i=1}^{n} T^{t_i}\xi_k$ is at most $16k^2n^2|r|\bar{t}_n$. Therefore,

$$H(\bigvee_{i=1}^{n} T^{t_i}\xi_k) \leqslant \log 16k^2n^2|r|\bar{t}_n$$

and so,

$$h_A(T,\xi_k) \leqslant \limsup_{n\to\infty} \frac{1}{n} \log \bar{t}_n$$

for all $k \geqslant 2$. Since $\bigvee_{k=2}^{\infty} \xi_k = \varepsilon$,

$$h_A(T) \leqslant \limsup_{n\to\infty} \frac{1}{n} \log \bar{t}_n \quad . \qquad (2)$$

For a measurable set C in $\mathbb{R}^2/\mathbb{Z}^2$, let C_y denote the set $\{x \in \mathbb{R}/\mathbb{Z} : (x,y) \in C\}$.

If $C \in \xi_k \vee (\bigvee_{i=1}^{n} T^{t_i}\xi_k)$ and $y \in \mathbb{R}/\mathbb{Z}$, then C_y is contained in at most

$$|r| \prod_{i=1}^{n} \left[\frac{|t_i|/\bar{t}_{i-1} + 1 + k}{k}\right]$$

intervals of length

$$\frac{1}{k|r|\bar{t}_n} \qquad (\text{taking } \bar{t}_0 = 1). \quad (3)$$

(Here and later, [x] denotes the integer part of x).

To prove this, note that (3) is true for n = 0 (taking the empty product to be 1). Assume it is true for n-1.

We have $C = T^{t_n}F \cap B$ for some $F \in \xi_k$ and $B \in \xi_k \vee (\bigvee_{i=1}^{n} T^{t_i}\xi_k)$.

Let I be one of the $|r| \prod_{i=1}^{n-1} \left[\frac{|t_i|/\bar{t}_{i-1} + 1 + k}{k}\right]$

intervals containing B_y. The set $T^{t_n}F$ is bounded by two

cycles $\{(x,y) \in \mathbb{R}^2/\mathbb{Z}^2 : y = rt_n x + \lambda_1\}$ and

$\{(x,y) \in \mathbb{R}^2/\mathbb{Z}^2 : y = rt_n x + \frac{1}{k} + \lambda_1\}$ for some constant λ_1.

Therefore, $(T^{t_n}F)_y$ is contained in a set of the form

$\{x \in \mathbb{R}/\mathbb{Z} : \lambda_2 \leq x \leq \frac{1}{k|rt_n|} + \lambda_2\}$ where λ_2 is a constant.

This set consists of intervals of length $\frac{1}{k|rt_n|}$ and the

distance between intervals is $\frac{k-1}{k|rt_n|}$. Therefore, if s is

the number of these intervals which intersect I then

$$\frac{1}{k|rt_n|}(s + (s-1)(k-1) - 2) \leq \frac{1}{k|r|\bar{t}_{n-1}}.$$

Hence,

$$s \leq \left[\frac{|t_n|/\bar{t}_{n-1} + 1 + k}{k}\right]$$

Therefore, $(T^{t_n}F)_y \cap I$ is contained in at most

$\left[\frac{|t_n|/\bar{t}_{n-1} + 1 + k}{k}\right]$ intervals of length $\frac{1}{k|r|\bar{t}_n}$.

Since there are at most $|r| \prod_{i=1}^{n-1} \left[\frac{|t_i|/\bar{t}_{i-1} + 1 + k}{k}\right]$

such I, (3) is proved. It follows from (3) that

$$m(C) = \int_0^1 \lambda(C_y)\, dy \quad (\lambda \text{ is Lebesgue measure on } [0,1]).$$

$$\leq \frac{1}{k\bar{t}_n} \prod_{i=1}^{n} \left[\frac{|t_i|/\bar{t}_{i-1} + 1 + k}{k}\right]$$

Therefore,

$$H(\xi_k \vee (\bigvee_{i=1}^{n} T^{t_i}\xi_k)) \geq -\log \frac{1}{k\bar{t}_n} \prod_{i=1}^{n} \left[\frac{|t_i|/\bar{t}_{i-1} + 1 + k}{k}\right] \tag{4}$$

If $\bar{t}_n/\bar{t}_{n-1}$ is bounded, choose k such that

$$\left[\frac{|t_i|/\bar{t}_{i-1} + 1 + k}{k}\right] = 1 \quad \text{for all } i.$$

Then

$$h_A(T,\xi_k) = \limsup_{n\to\infty} \frac{1}{n} H(\xi_k \vee (\bigvee_{i=1}^{n} T^{t_i}\xi_k))$$

$$\geq \limsup_{n\to\infty} \frac{1}{n} \log \bar{t}_n .$$

Therefore, by (2),

$$h_A(T) = \limsup_{n\to\infty} \frac{1}{n} \log \bar{t}_n .$$

Alternatively, if N is a sequence with upper density $\delta > 0$ such that $\lim_{n\in N} \bar{t}_n/\bar{t}_{n-1} = \infty$, it follows from (4) that

$$h_A(T,\xi_k) \geq \limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \log\left(\frac{k\,\bar{t}_i/\bar{t}_{i-1}}{\bar{t}_i/\bar{t}_{i-1} + 1 + k}\right)$$

$$= \log k + \limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \log\left(\frac{k\,\bar{t}_i/\bar{t}_{i-1}}{\bar{t}_i/\bar{t}_{i-1} + 1 + k}\right)$$

For all i,

$$\frac{\bar{t}_i/\bar{t}_{i-1}}{\bar{t}_i/\bar{t}_{i-1} + 1 + k} \geq \frac{1}{k+3}$$

and

$$\lim_{\substack{i\to\infty \\ i\in N}} \left(\frac{\bar{t}_i / \bar{t}_{i-1}}{\bar{t}_i / \bar{t}_{i-1} + 1 + k} \right) = 1 .$$

Therefore,

$$h_A(T,\xi_k) \geqslant \log k + \limsup_{n\to\infty} \frac{n - |N_n|}{n} \log \frac{1}{k+3}$$

$$= \log \frac{k}{k+3} + \delta \log(k+3) .$$

Taking the supremum over k gives $h_A(T) = \infty$.

§4. Zero sequence entropy

In this section we consider arbitrary transformations with quasi-discrete spectrum and give conditions on the sequence for zero and non-zero sequence entropy.

Theorem 1.10

Let T be an automorphism with quasi-discrete spectrum but not discrete spectrum and let $A = \{t_n\}$ be a sequence of integers. If there exists $\lambda > 1$ such that $\bar{t}_n / \bar{t}_{n-1} \geqslant \lambda$ on a subsequence with positive upper density, then $h_A(T) > 0$.

Proof. Using Lemma 1.8, it is sufficient to prove the result when T is a transformation on $\mathbb{R}^2/\mathbb{Z}^2$ given by $T(x,y) = (x + \alpha, x + y) \bmod 1$. Fix $k \geqslant 4$ and let ξ_k be the partition of $\mathbb{R}^2/\mathbb{Z}^2$ defined by the cycles $\{x = j/k\}$, $\{y = j/k\}$ $(0 \leqslant j \leqslant k-1)$. Using (4) in the proof of Theorem 1.9, we have

$$H(\xi_k \vee (\bigvee_{i=1}^{n} T^{t_i}\xi_k)) \geqslant -\log \frac{1}{k\bar{t}_n} \prod_{i=1}^{n} \left[\frac{|t_i|/\bar{t}_{i-1} + 1 + k}{k}\right]$$

$$= \log k + \sum_{i=1}^{n} \left(\log \bar{t}_i/\bar{t}_{i-1} - \log \left[\frac{|t_i|/\bar{t}_{i-1} + 1 + k}{k}\right]\right)$$

If $\bar{t}_i/\bar{t}_{i-1} < k - 1$, then

$$\left[\frac{|t_i|/\bar{t}_{i-1} + 1 + k}{k}\right] = 1$$

and if $\bar{t}_i/\bar{t}_{i-1} \geqslant k - 1$, then

$$\left[\frac{|t_i|/\bar{t}_{i-1} + 1 + k}{k}\right] \leq \frac{3\,\bar{t}_i/\bar{t}_{i-1}}{k} .$$

Therefore, if $\bar{t}_n/\bar{t}_{n-1} \geq \lambda$ on a subsequence with upper density $\delta > 0$, we have

$$\begin{aligned} h_A(T,\xi_k) &\geq \limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \log\left(\min\{\bar{t}_i/\bar{t}_{i-1}, \tfrac{k}{3}\}\right) \\ &\geq \delta \log\left(\min\{\lambda, \tfrac{k}{3}\}\right) \\ &> 0 \end{aligned}$$

Therefore, $h_A(T) > 0$.

<u>Theorem 1.11</u>

Let T be an automorphism with quasi-discrete spectrum and let $A = \{t_n\}$ be a sequence of integers. If $\limsup_{n\to\infty} \frac{1}{n}\log|t_n| = 0$, then $h_A(T) = 0$.

<u>Proof</u>. In view of Lemmas 1.4 and 1.7 it is sufficient to prove the result when T is a transformation of the m-torus given by $T(x) = S(x) + \alpha \bmod 1$ where S is represented by the integer matrix (a_{ij}) such that $a_{ij} = 0$ for $i < j$ and $a_{ii} = 1$ for all i. Consider the map $\tilde{T}: \mathbb{R}^m \to \mathbb{R}^m$ given by $\tilde{T}(x) = \tilde{S}(x) + \alpha$ where $\tilde{S}: \mathbb{R}^m \to \mathbb{R}^m$ is represented by the matrix (a_{ij}). The map $\tilde{S}$ is invertible and $\tilde{S}^{-1}$ is of the same form as $\tilde{S}$, that is, if $\tilde{S}^{-1} = (\bar{a}_{ij})$ then the numbers $\bar{a}_{ij}$ are integers and $\bar{a}_{ij} = 0$ for $i < j$ and $\bar{a}_{ii} = 1$ for all i. Put $M = \max_{i,j}\{|a_{ij}|, |\bar{a}_{ij}|\}$.

For $n \geqslant 1$, S^n is represented by the matrix

$$\left(\sum_{r_1=1}^{m} \cdots \sum_{r_{n-1}=1}^{m} a_{ir_1} a_{r_1 r_2} \cdots a_{r_{n-1} j} \right)_{i,j}$$

Since $a_{pq} = 0$ for $p < q$,

$$\sum_{r_1=1}^{m} \cdots \sum_{r_{n-1}=1}^{m} a_{ir_1} \cdots a_{r_{n-1} j} = 0$$

for $i < j$. For $i \geqslant j$,

$$\sum_{r_1=1}^{m} \cdots \sum_{r_{n-1}=1}^{m} a_{ir_1} \cdots a_{r_{n-1} j} = \sum_{i \geqslant r_1 \geqslant \cdots \geqslant r_{n-1} \geqslant j} a_{ir_1} \cdots a_{r_{n-1} j} \tag{5}$$

Since $a_{ii} = 1$ for all i, any term in (5) can be written as a product of s terms $a_{r_k r_{k+1}}$ where $r_k > r_{k+1}$ and $0 \leqslant s \leqslant \min\{n, i-j\}$. (If $s = 0$, then $i = j$ and $a_{ir_1} \cdots a_{r_{n-1} j} = 1$). The number of ways of choosing s values of k such that $r_k > r_{k+1}$ is $\dfrac{n!}{(n-s)!\,s!}$ and for each such k there are less than m^2 possible choices of $a_{r_k r_{k+1}}$. Therefore, if $n > i-j$,

$$\left| \sum_{i \geqslant r_1 \geqslant \cdots \geqslant r_{n-1} \geqslant j} a_{ir_1} \cdots a_{r_{n-1} j} \right| \leqslant \sum_{s=0}^{i-j} \frac{n!}{(n-s)!\,s!} m^{2s} M^s$$

$$\leqslant \frac{n(n-1)\cdots(i-j+1)}{(n-i+j)} \sum_{s=0}^{i-j} \frac{(i-j)!}{(i-j-s)!\,s!} (m^2 M)^s$$

$$\leqslant n^{i-j}(1 + m^2 M)^{i-j} .$$

If $n \leq i-j$, then

$$\left| \sum_{i \geq r_1 \geq \cdots \geq r_{n-1} \geq j} a_{ir_1} \cdots a_{r_{n-1}j} \right| \leq \sum_{s=0}^{n} \frac{n!}{(n-s)!\,s!} (m^2M)^s$$

$$= (1 + m^2M)^n \quad .$$

Therefore, if $\tilde{S}^n = (b_{ij})$ $(n \geq 1)$, then

$$|b_{ij}| \leq n^m(1 + m^2M)^m \quad \text{for all } i,j.$$

Similarily, if $\tilde{S}^{-n} = (\bar{b}_{ij})$, then

$$|\bar{b}_{ij}| \leq n^m(1 + m^2M)^m \quad \text{for all } i,j.$$

For $n \geq 1$, let $C_n = 2mn^m(1 + m^2M)^m$.

For $n \geq 1$, we have

$$\tilde{T}^n(x) = \tilde{S}^n(x) + \tilde{S}^{n-1}(\alpha) + \ldots + \alpha$$

and

$$\tilde{T}^{-n}(x) = \tilde{S}^{-n}(x) - \tilde{S}^{-n}(\alpha) - \ldots - \tilde{S}^{-1}(\alpha) \quad .$$

Therefore, if E is the unit cube $\{x \in \mathbb{R}^m : 0 \leq x_j \leq 1\}$ and n is any integer, then $\tilde{T}^nE$ is contained in a cube of the form

$$\prod_{1 \leq j \leq m} \left[y_j, \; y_j + C_{|n|} \right] \tag{6}$$

Consider the partition $\mathcal{E}_k$ $(k \geq 2)$ of $\mathbb{R}^m/\mathbb{Z}^m$ defined by the planes $P_{ij} = \{x_i = j/k\}$, $0 \leq j \leq k-1$, $1 \leq i \leq m$. The sets

in $\bigvee_{i=1}^{n} T^{t_i}\xi_k$ are m-dimensional volumes and their boundaries are formed by the images of the planes P_{ij} under T^{t_i} $(1 \leq i \leq n)$. Consider two such planes, $T^{t_i}P_{pq}$ and $T^{t_j}P_{rs}$ say. We have

$$T^{t_i}(P_{pq}) = \pi\tilde{T}^{t_i}i(P_{pq})$$

where $\pi: \mathbb{R}^m \longrightarrow \mathbb{R}^m/\mathbb{Z}^m$ is the natural projection and $i: \mathbb{R}^m/\mathbb{Z}^m \longrightarrow \mathbb{R}^m$ is the natural embedding. Similarily,

$$T^{t_j}(P_{rs}) = \pi\tilde{T}^{t_j}i(P_{rs})$$

Hence, it follows from (6) that the number of times $T^{t_i}P_{pq}$ and $T^{t_j}P_{rs}$ intersect is at most $(C_{|t_i|}C_{|t_j|})^m \leq (C_{\bar{t}_n})^{2m}$. Therefore, the number of intersections on any one plane is at most $m^2k^2n^2(C_{\bar{t}_n})^{2m}$. Since every m-dimensional volume needs at least one (m-1)-dimensional edge and each edge belongs to two volumes, the number of sets in $\bigvee_{i=1}^{n} T^{t_i}\xi_k$ is at most $2m^2k^2n^2(C_{\bar{t}_n})^{2m}$. Therefore,

$$h_A(T,\xi_k) \leq \limsup_{n\to\infty} \frac{1}{n} \log 2m^2k^2n^2(C_{\bar{t}_n})^{2m}$$

$$= \limsup_{n\to\infty} \frac{2m^2}{n} \log \bar{t}_n \quad .$$

However,

$$\limsup_{n\to\infty} \frac{1}{n} \log |t_n| = 0 \quad ,$$

and so,

$$\limsup_{n\to\infty} \frac{1}{n} \log \bar{t}_n = 0 \quad .$$

Therefore, $h_A(T,\xi_k) = 0$. Since $\bigvee_{k=2}^{\infty} \xi_k = \mathcal{E}$, $h_A(T) = 0$.

<u>Theorem 1.12</u>

Let T be an automorphism with quasi-discrete spectrum but not discrete spectrum, and let $A = \{t_n\}$ be a sequence of integers such that $\bar{t}_n / \bar{t}_{n-1}$ is bounded. Then $h_A(T) = 0$ if and only if $\lim\limits_{\substack{n\to\infty \\ n\notin N}} \bar{t}_n / \bar{t}_{n-1} = 1$ for some subsequence N with zero density.

<u>Proof</u>. We know that $\bar{t}_n / \bar{t}_{n-1} \geqslant 1$ for all n, and so, $\log \bar{t}_n / \bar{t}_{n-1}$ is a bounded sequence of non-negative numbers. Therefore,

$$\limsup_{n\to\infty} \frac{1}{n} \log \bar{t}_n = \limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \log \bar{t}_i / \bar{t}_{i-1} = 0$$

if and only if $\lim\limits_{\substack{n\to\infty \\ n\notin N}} \bar{t}_n / \bar{t}_{n-1} = 1$ for some subsequence N with zero density. Let $T' : \mathbb{R}^2/\mathbb{Z}^2 \longrightarrow \mathbb{R}^2/\mathbb{Z}^2$ be a factor of T as in Lemma 1.8. If $h_A(T) = 0$, then $h_A(T') = 0$, and so, $\limsup\limits_{n\to\infty} \frac{1}{n} \log \bar{t}_n = 0$ by Theorem 1.9(i). Conversely, if $\limsup\limits_{n\to\infty} \frac{1}{n} \log \bar{t}_n = 0$, then $h_A(T) = 0$ by Theorem 1.11.

§5. Infinite sequence entropy

The next result follows immediately from Theorem 1.9(ii) and Lemma 1.8.

Theorem 1.13

If T is an automorphism with quasi-discrete spectrum but not discrete spectrum, and $A = \{t_n\}$ is a sequence of integers such that $\lim_{\substack{n\to\infty \\ n\in N}} \bar{t}_n / \bar{t}_{n-1} = \infty$ for some subsequence N with positive upper density, then $h_A(T) = \infty$.

Theorem 1.14

Let T be an automorphism with quasi-discrete spectrum such that $n(T) = \infty$, and let $A = \{t_n\}$ be a sequence of integers such that $\bar{t}_n / \bar{t}_{n-1} \geqslant \lambda$ for some $\lambda > 1$ on a subsequence with positive upper density. Then $h_A(T) = \infty$.

Proof. Let $T_m : \mathbb{R}^m/\mathbb{Z}^m \longrightarrow \mathbb{R}^m/\mathbb{Z}^m$ be a factor of T as in Lemma 1.8. That is, $T_m(x) = S(x) + \alpha$ where $\alpha \in \mathbb{R}^m/\mathbb{Z}^m$ and S is given by the integer matrix (a_{ij}) where $a_{i+1,i} = a_{ii} = 1$ for $1 \leqslant i \leqslant m$ and $a_{ij} = 0$ otherwise. Then S^{-1} is represented by the matrix $(\bar{a}_{ij})$ where $\bar{a}_{i+1,i} = -1$, $\bar{a}_{ii} = 1$ and $\bar{a}_{ij} = 0$ otherwise. For $n \geqslant 1$, S^n is represented by the matrix

$$(S^n_{ij}) = \left(\sum_{r_1=1}^{m} \cdots \sum_{r_{n-1}=1}^{m} a_{ir_1} \cdots a_{r_{n-1}j}\right)_{i,j}$$

and if $n \leqslant -1$, the same holds with the a_{ij}'s replaced by $\bar{a}_{ij}$.

In either case $S^n_{ij} = 0$ for $i < j$, $S^n_{ii} = 1$ and $S^n_{i+1,i} = n$

(7)

Fix $k \geqslant 9$. Let ξ_j $(1 \leqslant j \leqslant m)$ be the partition of $\mathbb{R}^m/\mathbb{Z}^m$ defined by the planes $\{x_j = i/k\}$ $(0 \leqslant i \leqslant k-1)$ and let $\xi = \bigvee_{j=1}^{m} \xi_j$. Let $C \in \xi \vee (\bigvee_{i=1}^{n} T_m^{t_i}\xi)$. Then $C = \bigcap_{j=1}^{m} C_j$ where $C_j \in \xi \vee (\bigvee_{i=1}^{n} T_m^{t_i}\xi_j)$. If $C \subseteq B$ where $B \in \xi$, note that on B, χ_{C_j} depends only on coordinates x_i where $i \leqslant j$. If $y \in \mathbb{R}^{j-2}/\mathbb{Z}^{j-2}$ $(j \geqslant 2)$ and $W \in \mathcal{B}$, let $W(y)$ denote the subset $\Pi_j W \cap \{x : x_i = y_i,\ 1 \leqslant i \leqslant j-2\}$ where $\Pi_j : \mathbb{R}^m/\mathbb{Z}^m \longrightarrow \mathbb{R}^j/\mathbb{Z}^j$ denotes projection onto the first j coordinates. Then $C_j(y)$ is contained in at most

$$\prod_{i=1}^{n} \left[\frac{2|t_i|/\bar{t}_{i-1} + 2 + k}{k}\right]$$

rectangles of the form

$$B(y) \cap \{b \leqslant x_{j-1} \leqslant b + \frac{2}{k\bar{t}_n}\} \qquad (8)$$

To see this, note that (8) is true for $n = 0$ (taking $\bar{t}_0 = 1$). Assume it is true for $n-1$. We have $C_j = T_m^{t_n}F \cap E$ for some $F \in \xi_j$ and $E \in \xi \vee (\bigvee_{i=1}^{n-1} T_m^{t_i}\xi_j)$. It follows from (7) that $T_m^{t_n}F$ is bounded by the two planes

$$\{x \in \mathbb{R}^m/\mathbb{Z}^m : x_j = t_n x_{j-1} + \sum_{r=1}^{j-2} S^{t_n}_{jr} x_r + \lambda_1\}$$

and

$$\{x \in \mathbb{R}^m/\mathbb{Z}^m : x_j = t_n x_{j-1} + \sum_{r=1}^{j-2} S^{t_n}_{jr} x_r + \lambda_1\}$$

where λ_1 is some constant. Therefore, $(T_m^{t_n}F)(y)$ is bounded by two cycles

$$\{x_j = t_n x_{j-1} + \lambda_2 \bmod 1\}$$

and

$$\{x_j = t_n x_{j-1} + \lambda_2 + \frac{1}{k} \bmod 1\}$$

for some constant λ_2. Hence $(T_m^{t_n}F)(y)$ consists of strips whose width in the x_j-direction is $\frac{1}{k}$ and width in the x_{j-1}-direction is $\frac{1}{k|t_n|}$ and the distance in the x_{j-1}-direction between strips is $\frac{k-1}{k|t_n|}$. Let R be one of the rectangles containing E. Then R has width $\frac{2}{k\bar{t}_{n-1}}$, so if s is the number of strips of $(T_m^{t_n}F)(y)$ which intersect R, then

$$\frac{1}{k|t_n|}((s-1)k+1) \leq \frac{2}{k\bar{t}_{n-1}} + \frac{3}{k|t_n|}$$

and so

$$s \leq \left[\frac{2|t_n|/\bar{t}_{n-1} + 2 + k}{k}\right]$$

Moreover, the intersection of R and any strip of $(T_m^{t_n}F)(y)$ is contained in a rectangle of the form

$$B(y) \cap \{b \leq x_{j-1} \leq b + \frac{2}{k\bar{t}_n}\} .$$

Therefore, since there are at most $\prod_{i=1}^{n-1}\left[\frac{2|t_i|/\bar{t}_{i-1} + 2 + k}{k}\right]$ such R, (8) is proved.

It follows from (8) that $(\pi_{j-1}^{-1}\pi_{j-1}C_j)(y)$ is contained in

$$\prod_{i=1}^{n}\left[\frac{2|t_i|/\bar{t}_{i-1} + 2 + k}{k}\right]$$

rectangles of the form

$$\{b \leq x_{j-1} \leq b + \frac{2}{k\bar{t}_n}\} \quad .$$

Moreover, $\pi_{j-1}^{-1}\pi_{j-1}C_j$ depends only on coordinates x_i, $i \leq j-1$, and $C_j \subseteq \pi_{j-1}^{-1}\pi_{j-1}C_j$. Therefore, by repeated application of Fubini's Theorem,

$$\begin{aligned} m(C) &= \int_B \prod_{j=1}^{m} \chi_{C_j}\, dx_1 \ldots dx_m \\ &\leq \int_B \prod_{j=2}^{m} \chi_{(\pi_{j-1}^{-1}\pi_{j-1}C_j)}\, dx_1 \ldots dx_m \\ &\leq \left(\frac{2}{k\bar{t}_n}\prod_{i=1}^{n}\left[\frac{2|t_i|/\bar{t}_{i-1} + 2 + k}{k}\right]\right)^{m-1} \end{aligned}$$

Therefore,

$$H(\xi \vee (\bigvee_{i=1}^{n} T^{t_i}\xi)) \geq -(m-1)\log\frac{2}{k\bar{t}_n}\prod_{i=1}^{n}\left[\frac{2|t_i|/\bar{t}_{i-1} + 2 + k}{k}\right]$$

$$= (m-1)\left(\log\frac{k}{2} + \sum_{i=1}^{n}\left(\log \bar{t}_i/\bar{t}_{i-1} - \log\left[\frac{2|t_i|/\bar{t}_{i-1} + 2 + k}{k}\right]\right)\right) .$$

If $2\bar{t}_i/\bar{t}_{i-1} < k-2$, then

$$\left[\frac{2|t_i|/\bar{t}_{i-1} + 2 + k}{k}\right] = 1$$

and if $2\bar{t}_i/\bar{t}_{i-1} \geq k-2$, then

$$\left[\frac{2|t_i|/\bar{t}_{i-1} + 2 + k}{k}\right] \leq \frac{8\,\bar{t}_i/\bar{t}_{i-1}}{k} .$$

Therefore, if $\bar{t}_n/\bar{t}_{n-1} \geqslant \lambda$ on a subsequence with positive upper density $\delta > 0$, we have

$$h_A(T_m,\xi) \geqslant \limsup_{n\to\infty} \frac{m-1}{n} \sum_{i=1}^{n} \log\left(\min\left\{\bar{t}_i/\bar{t}_{i-1}, \frac{k}{8}\right\}\right)$$

$$\geqslant (m-1)\delta \log\left(\min\left\{\lambda, \frac{k}{8}\right\}\right)$$

Therefore, $h_A(T_m) \geqslant (m-1)\delta \log\left(\min\left\{\lambda, \frac{k}{8}\right\}\right)$ and taking the supremum over m gives $h_A(T) = \infty$.

Suppose T is an automorphism with quasi-discrete spectrum but not discrete spectrum and $A = \{t_n\}$ is a sequence of integers such that $\bar{t}_n/\bar{t}_{n-1}$ is bounded and $\bar{t}_n/\bar{t}_{n-1} \geqslant \lambda$ for some $\lambda > 1$ on a subsequence with positive upper density (otherwise $h_A(T) = 0$ for <u>all</u> such T). If $n(T) = \infty$, then $h_A(T) = \infty$ by Theorem 1.14. On the other hand, if T is an affine transformation of a finite dimensional torus, then n(T) is finite and $h_A(T)$ is finite by the proof of Theorem 1.11. However, in general, the converse does not hold, as the following example shows.

<u>Proposition 1.15</u>

Let T be the affine transformation of the infinite dimensional torus $\prod_{i=1}^{\infty} \mathbb{R}/\mathbb{Z}$ such that $T(x) = S(x) + \alpha$ where α_1 is irrational and $\alpha_i = 0$ for $i \geqslant 2$, and S is defined

as follows.

$$(Sx)_1 = x_1 \ , \quad (Sx)_2 = x_2 \ ,$$

and
$$(Sx)_{2i+1} = x_{2i+1} - x_{2i} - x_{2i-1} \ ,$$
$$(Sx)_{2i+2} = x_{2i+2} + x_{2i} + x_{2i-1} \qquad \text{for } i \geqslant 1.$$

Then T has quasi-discrete spectrum and $n(T) = 2$. Moreover, if $A = \{t_n\}$ satisfies the conditions of Theorem 1.14, then $h_A(T) = \infty$.

<u>Proof</u>. It is easily checked that T is totally ergodic and $(S - I)^2 = 0$. Hence, T has quasi-discrete spectrum and $n(T) = 2$. Let T_m be the factor automorphism of T on $\mathbb{R}^{2m+2}/\mathbb{Z}^{2m+2}$ defined by $T_m(x) = S_m(x) + \alpha_m$ where

$$\alpha_m = \begin{pmatrix} \alpha_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

and S_m is represented by the matrix

$$\begin{pmatrix} 1 & & & & & & & & & \\ 0 & 1 & & & & & & & & \\ -1 & -1 & 1 & & & & & & & \\ 1 & 1 & 0 & 1 & & & & & & \\ & & -1 & -1 & 1 & & & & & \\ & & 1 & 1 & 0 & 1 & \ddots & & & \\ & & & & & & \ddots & \ddots & & \\ & & & & & & & -1 & -1 & 1 \\ & & & & & & & 1 & 1 & 0 & 1 \end{pmatrix}$$

Clearly, $(S_m - I)^2 = 0$, so

$$S_m^n = ((S_m - I) + I)^n$$
$$= I + n(S_m - I) \qquad (9)$$

Fix $k \geqslant 9$. We use the same notation as in the proof of

as follows.

$$(Sx)_1 = x_1 \ , \quad (Sx)_2 = x_2 \ ,$$

and
$$(Sx)_{2i+1} = x_{2i+1} - x_{2i} - x_{2i-1} \ ,$$
$$(Sx)_{2i+2} = x_{2i+2} + x_{2i} + x_{2i-1} \qquad \text{for } i \geqslant 1.$$

Then T has quasi-discrete spectrum and $n(T) = 2$. Moreover, if $A = \{t_n\}$ satisfies the conditions of Theorem 1.14, then $h_A(T) = \infty$.

<u>Proof</u>. It is easily checked that T is totally ergodic and $(S - I)^2 = 0$. Hence, T has quasi-discrete spectrum and $n(T) = 2$. Let T_m be the factor automorphism of T on $\mathbb{R}^{2m+2}/\mathbb{Z}^{2m+2}$ defined by $T_m(x) = S_m(x) + \alpha_m$ where

$$\alpha_m = \begin{pmatrix} \alpha_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

and S_m is represented by the matrix

$$\begin{pmatrix} 1 & & & & & & & & & \\ 0 & 1 & & & & & & & & \\ -1 & -1 & 1 & & & & & & & \\ 1 & 1 & 0 & 1 & & & & & & \\ & & -1 & -1 & 1 & & & & & \\ & & 1 & 1 & 0 & 1 & & & & \\ & & & & & & \ddots & & & \\ & & & & & & & -1 & -1 & 1 \\ & & & & & & & 1 & 1 & 0 \; 1 \end{pmatrix}$$

Clearly, $(S_m - I)^2 = 0$, so

$$S_m^n = ((S_m - I) + I)^n$$
$$= I + n(S_m - I) \qquad (9)$$

Fix $k \geqslant 9$. We use the same notation as in the proof of

Theorem 1.14 (with m replaced by 2m+2). Let $C \in \xi \vee (\bigvee_{i=1}^{n} T_m^{t_i} \xi)$ and let $B \in \xi$ be such that $C \subseteq B$. If $j = 2i+1$ for some $1 \leq i \leq m$, then using (9) we can show in exactly the same way as in the proof of Theorem 1.14 that $(\pi_{j-1}^{-1} \pi_{j-1} C_j)(y)$ is contained in at most

$$\prod_{i=1}^{n} \left[\frac{2|t_i|/\bar{t}_{i-1} + 2 + k}{k} \right]$$

rectangles of the form

$$\{b \leq x_{j-1} \leq b + \frac{2}{k\bar{t}_n}\} .$$

Therefore,

$$m(C) \leq \int_B \prod_{i=1}^{m} \chi_{(\pi_{2i}^{-1} \pi_{2i} C_{2i+1})} \, dx_1 \ldots dx_{2m+2}$$

$$\leq \frac{2}{k\bar{t}_n} \prod_{i=1}^{n} \left[\frac{2|t_i|/\bar{t}_{i-1} + 2 + k}{k} \right]$$

If $\bar{t}_n/\bar{t}_{n-1} \geq \lambda > 1$ on a subsequence with positive upper density $\delta > 0$, we may deduce, as in the proof of Theorem 1.14, that

$$h_A(T_m) \geq h_A(T_m, \xi)$$

$$\geq m\delta \log (\min \{\lambda, \tfrac{k}{8}\})$$

Taking the supremum over m gives $h_A(T) = \infty$.

Finally in this chapter, we note that the results of this section and section 3 give examples of spectrally isomorphic automorphisms with the same entropy but different sequence entropies. For example, let $\alpha \in (0,1)$ be irrational and let $T_1 : \mathbb{R}^2/\mathbb{Z}^2 \longrightarrow \mathbb{R}^2/\mathbb{Z}^2$ and $T_2 : \mathbb{R}^3/\mathbb{Z}^3 \longrightarrow \mathbb{R}^3/\mathbb{Z}^3$ be defined by

$$T_1\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \alpha \\ 0 \end{pmatrix} \mod 1$$

and

$$T_2\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} \alpha \\ 0 \\ 0 \end{pmatrix} \mod 1$$

Both T_1 and T_2 have quasi-discrete spectrum but not discrete spectrum and all such transformations have countable Lebesgue spectrum in the orthogonal complement of the subspace generated by the eigenfunctions (see [1]). Moreover, T_1 and T_2 have the same eigenvalues and so they are spectrally isomorphic. Also, they both have zero entropy, but if $A = \{2^n\}$ then $h_A(T_1) = \log 2$ by Theorem 1.9(i), whereas it follows from the proof of Theorem 1.14 that $h_A(T_2) \geqslant 2 \log 2$.

CHAPTER II

SEQUENCE ENTROPY AND SUBSEQUENCE GENERATORS

Throughout this chapter, T is an automorphism of a Lebesgue space $(X,\mathcal{B},m)$. We will use the definitions, results and notation of Chapter I, §1.

§1. Sequence entropy and spectrum

Let $\gamma(T)$ be the T-invariant partition of X generated by the eigenfunctions of T. In other words, $\gamma(T)$ is the maximum partition such that the factor automorphism $T_{\gamma(T)}$ on $X/\gamma(T)$ has discrete spectrum. The next result strengthens a similar result proved by Pickel [13] and independently by Walters. The proof uses the same method.

Theorem 2.1

There exists an increasing sequence of natural numbers $A = \{t_n\}_{n=1}^{\infty}$ such that

$$h_A(T,\xi) = H(\xi)$$

for all $\xi \in Z$ which are independent of $\gamma(T)$.

Proof. Let Z' denote the set of partitions in Z which are independent of $\gamma(T)$. Then Z' is separable and so we may choose a countable, dense subset $\{\xi_k\}_{k=1}^{\infty}$ of finite partitions in Z'. If $E \in \xi_k$, then

$$\chi_E - m(E) \in L^2(X) \ominus L^2(\gamma(T)) \quad .$$

Therefore, for all $B \in \mathcal{B}$,

$$\lim_{\substack{n\to\infty \\ n\notin N}} m(T^n E \cap B) = m(E)m(B) \qquad (1)$$

where N is a subsequence of zero density depending on E and B. We define an increasing sequence $A = \{t_n\}_{n=1}^{\infty}$ such that

$$H(T^{t_n}\xi_k | \bigvee_{i=1}^{n-1} T^{t_i}\xi_k) \geqslant H(\xi_k) - 2^{-n}$$

for all $n \geqslant 2$, $1 \leqslant k \leqslant n$. Let t_1 be any positive integer and suppose $t_1,\ldots,t_n$ have been defined. Since ξ_k is finite for each k, using (1) we may choose t_{n+1} such that $t_{n+1} > t_n$ and

$$|\log m(T^{t_{n+1}}E \cap B) - \log m(E)m(B)| < 2^{-(n+1)}$$

for all $E \in \xi_k$, $B \in \bigvee_{i=1}^{n} T^{t_i}\xi_k$ $(1 \leqslant k \leqslant n+1)$. Then

$$H(T^{t_{n+1}}\xi_k | \bigvee_{i=1}^{n} T^{t_i}\xi_k) = \sum_{E,B} -m(T^{t_{n+1}}E \cap B) \log \frac{m(T^{t_{n+1}}E \cap B)}{m(B)}$$

$$= \sum_{E,B} m(T^{t_{n+1}}E \cap B)(\log m(B) - \log m(T^{t_{n+1}}E \cap B))$$

$$\geqslant \sum_{E,B} m(T^{t_{n+1}}E \cap B)(\log m(B) - \log m(E)m(B) - 2^{-(n+1)})$$

$$= \Big(\sum_{E,B} -m(T^{t_{n+1}}E \cap B) \log m(E)\Big) - 2^{-(n+1)}$$

$$= \Big(\sum_{E} -m(E) \log m(E)\Big) - 2^{-(n+1)}$$

$$= H(\xi_k) - 2^{-(n+1)}$$

Fix k. Then for $n > k$,

$$H(\bigvee_{i=1}^{n} T^{t_i}\xi_k) = H(\bigvee_{i=1}^{k} T^{t_i}\xi_k) + \sum_{j=k+1}^{n} H(T^{t_i}\xi_k | \bigvee_{j=i}^{i-1} T^{t_j}\xi_k)$$

$$\geqslant H(\bigvee_{i=1}^{k} T^{t_i}\xi_k) + (n-k)H(\xi_k) - \sum_{i=k+1}^{\infty} 2^{-i} .$$

Therefore,

$$h_A(T,\xi_k) \geqslant \limsup_{n\to\infty} \left(\frac{1}{n} H(\bigvee_{i=1}^{k} T^{t_i}\xi_k) + \frac{(n-k)}{n} H(\xi_k) - \frac{1}{n} \sum_{i=k+1}^{\infty} 2^{-i}\right)$$

$$= H(\xi_k) .$$

Now let ξ be any partition in Z'. Given $\delta > 0$, choose ξ_k such that $\rho(\xi,\xi_k) \leqslant \delta$. Then, by Lemma 1.2,

$$h_A(T,\xi) \geqslant h_A(T,\xi_k) - \delta$$

$$\geqslant H(\xi) - 2\delta$$

Since δ is arbitrary, $h_A(T,\xi) \geqslant H(\xi)$ and since $h_A(T,\xi) \leqslant H(\xi)$ for any $\xi \in Z$, it follows that

$$h_A(T,\xi) = H(\xi) \quad \text{for all } \xi \in Z'.$$

Let T be ergodic. Then we may write T as a skew product $(x,y) \longrightarrow (T_{\gamma(T)}x, T_x y)$ on $X/\gamma(T) \times M$ (see [2]). Since T is ergodic, either M consists of k atoms of measure $\frac{1}{k}$, or M is continuous. Pickel [13] and Walters have shown that $\sup_A h_A(T)$ is log k in the first case and infinite in the second. When M consists of atoms, it is clear from their proof that the supremum is always attained, so we will only

consider the case when M is continuous and use Theorem 2.1 to show the supremum is attained in this case also.

Theorem 2.2

If T is ergodic and M is continuous, there exists an increasing sequence of natural numbers A such that $h_A(T) = \infty$.

Proof. By Theorem 2.1, there exists an increasing sequence of natural numbers A such that $h_A(T,\xi) = H(\xi)$ for all $\xi \in Z$ independent of $\gamma(T)$. Since M is continuous, for any k we may choose a partition ζ of M consisting of k sets of measure $\frac{1}{k}$. Put $\xi = X/\gamma(T) \times \zeta$. Then ξ is independent of $\gamma(T)$ and $H(\xi) = \log k$. Therefore, $h_A(T,\xi) = \log k$. Taking the supremum over k gives $h_A(T) = \infty$.

By definition, T is weak-mixing if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} |m(T^{-i}B \cap C) - m(B)m(C)| = 0$$

for all $B, C \in \mathcal{B}$. The next result extends Saleski's result on weak-mixing and sequence entropy [15].

Theorem 2.3

T is weak-mixing if and only if there exists an increasing sequence of natural numbers A such that $h_A(T,\xi) = H(\xi)$ for all $\xi \in Z$.

Proof. If T is weak-mixing, $\gamma(T)$ is trivial. Hence, all partitions are independent of $\gamma(T)$ and the result follows

from Theorem 2.1. Conversely, if T is not weak-mixing, there exists a non-trivial $\xi \in Z$ such that $\xi \in \mathcal{S}(T)$. Then $H(\xi) > 0$, but it follows from Kushnirenko's result on sequence entropy and discrete spectrum ([10], Theorem 4) that $h_A(T,\xi) = 0$ for all A.

By definition, T is strong-mixing if

$$\lim_{n\to\infty} m(T^{-n}B \cap C) = m(B)m(C) \quad \text{for all } B, C \in \mathcal{B} \tag{1'}$$

Saleski [15] has shown that T is strong-mixing if and only if for every increasing sequence A and $\xi \in Z$,

$$\sup_{B\subseteq A} h_B(T,\xi) = H(\xi) .$$

If T is strong-mixing, we can replace (1) in the proof of Theorem 2.1 by the stronger condition (1'). Then, given any infinite sequence A, we may construct in a similar way a subsequence $B \subseteq A$ such that $h_B(T,\xi) = H(\xi)$ for all $\xi \in Z$. Conversely, if T is not strong-mixing, it follows immediately from Saleski's result that such a subsequence does not exist for all A. Hence, we have the following improved version of Saleski's result.

<u>Theorem 2.4</u>

T is strong-mixing if and only if given any infinite sequence A, there exists a subsequence $B \subseteq A$ such that

$$h_B(T,\xi) = H(\xi) \quad \text{for all } \xi \in Z.$$

§2. Subsequence generators

Definition 2.5

If N is an infinite set of integers, then a partition ξ is an N-generator of T if $\bigvee_{n\in N} T^n\xi = \mathcal{E}$ or, equivalently, $\bigcup_{n\in N} T^n\xi$ generates $\mathcal{B}$.

If we take $N = \mathbb{Z}$, we have the usual definition of a two-sided generator.

Theorem 2.6 (see [4])

Let T be an aperiodic automorphism. For every infinite set of integers N and $\varepsilon > 0$, there exists a set $B \in \mathcal{B}$ such that $m(B) < \varepsilon$ and $\bigcup_{n\in N} T^n B = X$.

We will be interested in automorphisms T such that T^{n_i} converges to the identity in the space of automorphisms of $(X,\mathcal{B},m)$ with the weak topology for some infinite sequence of integers $\{n_i\}$. In other words, $m(T^{n_i}B \Delta B) \longrightarrow 0$ as $n \to \infty$ for all $B \in \mathcal{B}$. ($B \Delta C$ denotes the symmetric difference $(B \cup C) \setminus (B \cap C)$). Such transformations were studied in [16]. Note that if we identify sets whose symmetric difference is zero, $\mathcal{B}$ is a complete metric space with metric d defined by $d(B,C) = m(B \Delta C)$ and if $\mathcal{A}$ is a dense subset of $\mathcal{B}$ then $\mathcal{A}$ generates $\mathcal{B}$.

Lemma 2.7

If T^{n_i} converges weakly to the identity, there exists a countable, dense, T-invariant sub-algebra in $\mathcal{B}$.

<u>Proof</u>. It is well-known that if $T^{n_i} \to I$ weakly, then $h(T) = 0$ (see [16]) and therefore, T has a two-sided generator ξ with finite entropy. Hence, ξ is countable (mod 0) and so, $\bigcup_{n=0}^{\infty} (\bigvee_{i=-n}^{n} T^i \xi)$ is a countable, dense, T-invariant sub-algebra.

<u>Lemma 2.8</u>

Let $\mathcal{A}$ be a countable, dense, T-invariant sub-algebra of $\mathcal{B}$ and let $\{n_i\}$ be a sequence of integers such that

$$\lim_{k\to\infty} m\left(\bigcup_{i=k}^{\infty} T^{n_i}B \setminus \bigcap_{i=k}^{\infty} T^{n_i}B\right) = 0$$

for all $B \in \mathcal{A}$. Then given $B \in \mathcal{A}$ and $\varepsilon > 0$, there exists $D \in \mathcal{A}$ such that $m(D) < \varepsilon$ and

$$m\left(B \,\Delta\, \bigcup_{i=1}^{\infty} T^{n_i}D\right) < \varepsilon .$$

<u>Proof</u>. Given $B \in \mathcal{A}$ and $\varepsilon > 0$, choose k such that

$$m\left(\bigcup_{i=k}^{\infty} T^{n_i}B \setminus \bigcap_{i=k}^{\infty} T^{n_i}B\right) < \frac{\varepsilon}{4}$$

Choose $F \in \mathcal{A}$ such that $m(F) < \frac{\varepsilon}{8k}$ and $m\left(\bigcup_{i=1}^{\infty} T^{n_i}F\right) > 1 - \varepsilon$.

Put $C = B \cap F$. Then $\bigcup_{i=k}^{\infty} T^{n_i}C \subseteq \bigcup_{i=k}^{\infty} T^{n_i}B$.

Therefore,

$$m\left(\bigcup_{i=k}^{\infty} T^{n_i}C \setminus T^{n_k}B\right) < \frac{\varepsilon}{4}$$

and so,

$$m\left(\bigcup_{i=1}^{\infty} T^{n_i}C \setminus T^{n_k}B\right) < \frac{\varepsilon}{2} \qquad (2)$$

Also,

$$\left(\bigcup_{i=1}^{\infty} T^{n_i}C\right)^c \subseteq \left(\bigcup_{i=1}^{\infty} T^{n_i}(F\setminus C)\right) \cup \left(\bigcup_{i=1}^{\infty} T^{n_i}F\right)^c$$

$$= \left(\bigcup_{i=1}^{k-1} T^{n_i}(F\setminus C)\right) \cup \left(\bigcup_{i=k}^{\infty} T^{n_i}(F\setminus C)\right) \cup \left(\bigcup_{i=1}^{\infty} T^{n_i}F\right)^c$$

$$\subseteq \bigcup_{i=1}^{k-1} T^{n_i}F \cup \left(\bigcup_{i=k}^{\infty} T^{n_i}B^c\right) \cup \left(\bigcup_{i=1}^{\infty} T^{n_i}F\right)^c$$

Now

$$m\left(\bigcup_{i=1}^{k-1} T^{n_i}F\right) < \frac{\varepsilon}{8} ,$$

$$m\left(\left(\bigcup_{i=1}^{\infty} T^{n_i}F\right)^c\right) < \frac{\varepsilon}{8}$$

and

$$m\left(T^{n_k}B \cap \bigcup_{i=k}^{\infty} T^{n_i}B^c\right) < \frac{\varepsilon}{4}$$

Therefore,

$$m\left(T^{n_k}B \setminus \bigcup_{i=1}^{\infty} T^{n_i}C\right) < \frac{\varepsilon}{2} \qquad (3)$$

Combining (2) and (3) gives

$$m\left(T^{n_k}B \,\Delta\, \bigcup_{i=1}^{\infty} T^{n_i}C\right) < \varepsilon .$$

Hence, $T^{-n_k}C$ is the required set.

<u>Theorem 2.9</u>

Let $\{n_i\}_{i=1}^{\infty}$ be an infinite sequence of integers such that T^{n_i} converges weakly to the identity. For every $\varepsilon > 0$, there exists $W \in \mathcal{B}$ such that $m(W) < \varepsilon$ and $\{T^{n_i}W\}_{i=1}^{\infty}$ generates $\mathcal{B}$.

Also,

$$(\bigcup_{i=1}^{\infty} T^{n_i}C)^c \subseteq (\bigcup_{i=1}^{\infty} T^{n_i}(F\setminus C)) \cup (\bigcup_{i=1}^{\infty} T^{n_i}F)^c$$

$$= (\bigcup_{i=1}^{k-1} T^{n_i}(F\setminus C)) \cup (\bigcup_{i=k}^{\infty} T^{n_i}(F\setminus C)) \cup (\bigcup_{i=1}^{\infty} T^{n_i}F)^c$$

$$\subseteq \bigcup_{i=1}^{k-1} T^{n_i}F \cup (\bigcup_{i=k}^{\infty} T^{n_i}B^c) \cup (\bigcup_{i=1}^{\infty} T^{n_i}F)^c$$

Now

$$m(\bigcup_{i=1}^{k-1} T^{n_i}F) < \frac{\varepsilon}{8} ,$$

$$m((\bigcup_{i=1}^{\infty} T^{n_i}F)^c) < \frac{\varepsilon}{8}$$

and

$$m(T^{n_k}B \cap \bigcup_{i=k}^{\infty} T^{n_i}B^c) < \frac{\varepsilon}{4}$$

Therefore,

$$m(T^{n_k}B \setminus \bigcup_{i=1}^{\infty} T^{n_i}C) < \frac{\varepsilon}{2} \qquad (3)$$

Combining (2) and (3) gives

$$m(T^{n_k}B \,\Delta\, \bigcup_{i=1}^{\infty} T^{n_i}C) < \varepsilon .$$

Hence, $T^{-n_k}C$ is the required set.

<u>Theorem 2.9</u>

Let $\{n_i\}_{i=1}^{\infty}$ be an infinite sequence of integers such that T^{n_i} converges weakly to the identity. For every $\varepsilon > 0$, there exists $W \in \mathcal{B}$ such that $m(W) < \varepsilon$ and $\{T^{n_i}W\}_{i=1}^{\infty}$ generates $\mathcal{B}$.

<u>Proof</u>. Let $\mathcal{A} = \{B_r\}_{r=1}^{\infty}$ be a countable, dense, T-invariant sub-algebra of $\mathcal{B}$. For each k, choose i_k such that

$$m(T^{n_i}B_r \Delta T^{n_j}B_r) < 2^{-k}$$

for all $i,j \geqslant i_k$, $1 \leqslant r \leqslant k$. Put $t_k = n_{i_k}$. Then

$$m(T^{t_i}B_r \Delta T^{t_j}B_r) < 2^{-k} \tag{4}$$

for all $i,j \geqslant k$, $1 \leqslant r \leqslant k$. Now fix B_r. For $k \geqslant r$,

$$m(\bigcup_{i=k}^{\infty} T^{t_i}B_r) \leqslant m(B_r) + \sum_{i=k+1}^{\infty} 2^{-i}$$

$$= m(B_r) + 2^{-k}$$

and

$$m(\bigcap_{i=k}^{\infty} T^{t_i}B_r) \geqslant m(B_r) - \sum_{i=k+1}^{\infty} 2^{-i}$$

$$= m(B_r) - 2^{-k}$$

Hence, for each B_r,

$$m(\bigcup_{i=k}^{\infty} T^{t_i}B_r \Delta \bigcap_{i=k}^{\infty} T^{t_i}B_r) \leqslant 2^{-k+1} \quad \text{for all } k \geqslant r.$$

Therefore,

$$\lim_{k\to\infty} m(\bigcup_{i=k}^{\infty} T^{t_i}B \Delta \bigcap_{i=k}^{\infty} T^{t_i}B) = 0 \quad \text{for all } B \in \mathcal{A}. \tag{5}$$

Let $\{A_i\}_{i=1}^{\infty}$ be an indexing of elements of $\mathcal{A}$ such that each element occurs infinitely often. Fix $\varepsilon > 0$. Define inductively sequences $\{C_k\}_{k=1}^{\infty}$, $\{D_k\}_{k=1}^{\infty}$, $\{E_k\}_{k=1}^{\infty} \subseteq \mathcal{A}$ and sequences $\{p_k\}_{k=1}^{\infty}$, $\{q_k\}_{k=1}^{\infty}$, $\{r_k\}_{k=1}^{\infty}$ of integers as follows. Let $C_1 = \emptyset$.

Given C_k, using (5) and Lemma 2.8, choose $p_k \geqslant r_{k-1}$ (or $p_k > 0$ if $k = 1$) and $D_k \in \mathcal{A}$ such that

$$m\left(\bigcup_{i=p_k}^{\infty} T^{t_i} C_k \setminus \bigcap_{i=p_k}^{\infty} T^{t_i} C_k\right) < 2^{-k}\varepsilon \qquad (6),$$

$$m(D_k) < 2^{-k-1} p_k^{-1} \varepsilon \qquad (7)$$

and

$$m\left(A_k \,\Delta\, \bigcup_{i=1}^{\infty} T^{t_i} D_k\right) < 2^{-k-1}\varepsilon \,. \qquad (8).$$

Using Lemma 2.8 again, choose $q_k > p_k$ such that

$$m\left(A_k \,\Delta\, \bigcup_{i=p_k}^{q_k} T^{t_i} D_k\right) < 2^{-k}\varepsilon \qquad (9)$$

and

$$m\left(\bigcup_{i=q_k}^{\infty} T^{t_i} D_k \setminus \bigcap_{i=q_k}^{\infty} T^{t_i} D_k\right) < 2^{-k}\varepsilon \qquad (10)$$

Choose $E_k \in \mathcal{A}$ such that

$$m(E_k) < 2^{-k-1} q_k^{-1} \varepsilon \qquad (11)$$

and

$$m\left(A_k \,\Delta\, \bigcup_{i=1}^{\infty} T^{t_i} E_k\right) < 2^{-k-1}\varepsilon \qquad (12)$$

Choose $r_k > q_k$ such that

$$m\left(A_k \,\Delta\, \bigcup_{i=q_k}^{r_k} T^{t_i} E_k\right) < 2^{-k}\varepsilon \qquad (13)$$

Put $C_{k+1} = (C_k \cup D_k) \setminus E_k$.

Let $W = \bigcup_{j=1}^{\infty} \left(\bigcap_{k=j}^{\infty} C_k \right)$. Then

$$m(W) \leq \sum_{k=1}^{\infty} m(D_k) < \sum_{k=1}^{\infty} 2^{-k-1}\varepsilon < \frac{\varepsilon}{2}$$

and $\{T^{t_i}W\}_{i=1}^{\infty}$ generates $\mathcal{A}$. To see this, let

$$F_k = \bigcup_{i=p_k}^{q_k} T^{t_i}W \setminus \bigcap_{i=q_k}^{r_k} T^{t_i}W$$

Note that

$$\bigcup_{i=p_k}^{q_k} T^{t_i}D_k \setminus \bigcup_{j=k}^{\infty} \left(\bigcup_{i=p_k}^{q_k} T^{t_i}E_j \right) \subseteq \bigcup_{i=p_k}^{q_k} T^{t_i}W$$

and

$$m\left(\bigcup_{j=k}^{\infty} \left(\bigcup_{i=p_k}^{q_k} T^{t_i}E_j \right) \right) < q_k \sum_{j=k}^{\infty} 2^{-j-1} q_j^{-1} \varepsilon$$

$$< 2^{-k}\varepsilon$$

by (11). Therefore,

$$m\left(A_k \setminus \bigcup_{i=p_k}^{q_k} T^{t_i}W\right) < m\left(A_k \setminus \bigcup_{i=p_k}^{q_k} T^{t_i}D_k\right) + 2^{-k}\varepsilon$$

$$< 2^{-k+1}\varepsilon \qquad (14)$$

Also, $E_k \setminus \bigcup_{j=k+1}^{\infty} D_j \subseteq W^c$. Therefore, (7) and (11) imply

$$m(A_k) \cap \left(\bigcap_{i=q_k}^{r_k} T^{t_i}W \right) \leq m(A_k) - m\left(A_k \cap \left(\bigcup_{i=q_k}^{r_k} T^{t_i}\left(E_k \setminus \bigcup_{j=k+1}^{\infty} D_j\right)\right)\right)$$

$$\leq m(A_k) + m\left(\bigcup_{i=q_k}^{r_k} \left(\bigcup_{j=k+1}^{\infty} T^{t_i}D_j \right) \right) - m\left(A_k \cap \left(\bigcup_{i=q_k}^{r_k} T^{t_i}E_k \right)\right)$$

$$< m(A_k) + r_k \sum_{j=k+1}^{\infty} 2^{-j-1} p_j^{-1} \varepsilon - m(A_k) + 2^k \varepsilon$$

$$< 2^{-k+1} \varepsilon \qquad (15)$$

It follows from (14) and (15) that

$$m(A_k \setminus F_k) \leq m(A_k \setminus \bigcup_{i=p_k}^{q_k} T^{t_i} W) + m(A_k \cap (\bigcap_{i=q_k}^{r_k} T^{t_i} W))$$

$$< 2^{-k+2} \varepsilon \qquad (16)$$

We have $W \subseteq C_k \cup D_k \cup (\bigcup_{j=k+1}^{\infty} D_j)$. Therefore, it follows that

$$m(F_k \setminus A_k) \leq m((\bigcup_{i=p_k}^{q_k} T^{t_i} C_k \setminus \bigcap_{i=q_k}^{r_k} T^{t_i} W) \setminus A_k) + m((\bigcup_{i=p_k}^{q_k} T^{t_i} D_k) \setminus A_k)$$

$$+ m((\bigcup_{i=p_k}^{q_k} (\bigcup_{j=k+1}^{\infty} T^{t_i} D_j)) \setminus A_k)$$

We have

$$m((\bigcup_{i=p_k}^{q_k} T^{t_i} D_k) \setminus A_k) < 2^{-k} \varepsilon$$

by (9) and

$$m((\bigcup_{i=p_k}^{q_k} (\bigcup_{j=k+1}^{\infty} T^{t_i} D_j)) \setminus A_k) < q_k \sum_{j=k+1}^{\infty} 2^{-j-1} p_j^{-1} \varepsilon$$

$$< 2^{-k-1} \varepsilon$$

by (7). Also, since $C_k \setminus \bigcup_{j=k}^{\infty} E_j \subseteq W$, it follows from (6), (11) and (13) that

$$m((\bigcup_{i=p_k}^{q_k} T^{t_i}C_k \setminus \bigcap_{i=q_k}^{r_k} T^{t_i}W) \setminus A_k)$$

$$\leq \quad m(\bigcup_{i=p_k}^{q_k} T^{t_i}C_k \setminus \bigcap_{i=q_k}^{r_k} T^{t_i}C_k) \quad + \quad m((\bigcap_{i=q_k}^{r_k} T^{t_i}C_k \setminus \bigcap_{i=q_k}^{r_k} T^{t_i}W) \setminus A_k)$$

$$< \quad 2^{-k}\varepsilon \quad + \quad m((\bigcup_{i=q_k}^{r_k} (\bigcup_{j=k}^{\infty} T^{t_i}E_j)) \setminus A_k)$$

$$< \quad 2^{-k}\varepsilon \quad + \quad m(\bigcup_{i=q_k}^{r_k} T^{t_i}E_k \setminus A_k) \quad + \quad m(\bigcup_{i=q_k}^{r_k} (\bigcup_{j=k+1}^{\infty} T^{t_i}E_j))$$

$$< \quad 2^{-k}\varepsilon \quad + \quad 2^{-k}\varepsilon \quad + \quad r_k \sum_{j=k+1}^{\infty} 2^{-j-1}q_j^{-1}\varepsilon$$

$$< \quad 3.2^{-k}\varepsilon$$

Hence,

$$m(F_k \setminus A_k) \quad < \quad 5.2^{-k}\varepsilon$$

Therefore, using (16), we have

$$m(A_k \Delta F_k) \quad < \quad 9.2^{-k}\varepsilon \qquad (17)$$

Since each $B \in \mathcal{A}$ occurs infinitely often in the sequence $\{A_k\}$, it follows from (17) that each $B \in \mathcal{A}$ can be approximated arbitrarily closely by a set generated by $\{T^{t_i}W\}_{i=1}^{\infty}$.

Therefore, since $\mathcal{A}$ is dense in $\mathcal{B}$, it follows that $\{T^{t_i}W\}_{i=1}^{\infty}$ generates $\mathcal{B}$; hence, so does $\{T^{n_i}W\}_{i=1}^{\infty}$.

It is well known that if there exists a (two-sided) generator $\xi \in Z$ of T, then h(T) is finite and $h(T) = h(T,\xi)$. Moreover, h(T) is the infimum of the entropies of all generators of T. The next result shows that none of this holds for arbitrary sequences.

Theorem 2.10

There exists a weak-mixing automorphism T and an increasing sequence of natural numbers A such that $h_A(T) = \infty$, $h_A(T,\xi) = H(\xi)$ for all $\xi \in Z$ and, given any $\varepsilon > 0$, there exists an A-generator ζ such that $H(\zeta) < \varepsilon$.

Proof. Let T be a weak-mixing automorphism such that T^{n_i} converges weakly to the identity for some increasing sequence of natural numbers $N = \{n_i\}_{i=1}^{\infty}$. (Such automorphisms exist; see [16]). By Theorem 2.3, there exists an increasing sequence of natural numbers $J = \{j_i\}_{i=1}^{\infty}$ such that

$$h_J(T,\xi) = H(\xi) \qquad \text{for all } \xi \in Z.$$

Now choose a subsequence $K = \{k_i\}$ of N so that $k_i > j_{i2}$. Put $A = J \cup K$ and index the elements so that A is an increasing sequence. Note that $J(n) \subseteq A(n+[n^{1/2}])$ where J(n) denotes the first n terms of the sequence J. If $W \in \mathcal{B}$, let $\xi_W = \{W, W^c\}$. Then for any $\varepsilon > 0$, there exists $\delta > 0$ such that

$$m(W) < \delta \implies H(\xi_W) < \varepsilon .$$

Since $K \subseteq N$, T^{k_i} converges weakly to the identity. Therefore, by Theorem 2.9, there exists $W \in \mathcal{B}$ such that $H(\xi_W) < \varepsilon$ and

ξ_W is a K-generator of T and, therefore, an A-generator of T. Since $J(n) \subsetneq A(n+[n^{1/2}])$,

$$\frac{1}{n+[n^{1/2}]} H\Big(\bigvee_{i\in A(n+[n^{1/2}])} T^i\xi\Big) \geqslant \frac{1}{n+[n^{1/2}]} H\Big(\bigvee_{i\in J(n)} T^i\xi\Big)$$

for any $\xi \in Z$ and so,

$$\begin{aligned} h_A(T,\xi) &\geqslant \limsup_{n\to\infty} \frac{1}{n+[n^{1/2}]} H\Big(\bigvee_{i\in J(n)} T^i\xi\Big) \\ &= h_J(T,\xi) \\ &= H(\xi) \quad . \end{aligned}$$

Therefore,

$$h_A(T,\xi) = H(\xi) \qquad \text{for all } \xi \in Z.$$

Since T is weak-mixing, m is a continuous measure and so we may choose $\xi \in Z$ with arbitrarily large entropy. Therefore, it follows that $h_A(T) = \infty$.

ξ_W is a K-generator of T and, therefore, an A-generator of T. Since $J(n) \subseteq A(n+[n^{1/2}])$,

$$\frac{1}{n+[n^{1/2}]} H\Big(\bigvee_{i\in A(n+[n^{1/2}])} T^i\xi\Big) \geqslant \frac{1}{n+[n^{1/2}]} H\Big(\bigvee_{i\in J(n)} T^i\xi\Big)$$

for any $\xi \in Z$ and so,

$$\begin{aligned} h_A(T,\xi) &\geqslant \limsup_{n\to\infty} \frac{1}{n+[n^{1/2}]} H\Big(\bigvee_{i\in J(n)} T^i\xi\Big) \\ &= h_J(T,\xi) \\ &= H(\xi) \quad . \end{aligned}$$

Therefore,

$$h_A(T,\xi) = H(\xi) \qquad \text{for all } \xi \in Z.$$

Since T is weak-mixing, m is a continuous measure and so we may choose $\xi \in Z$ with arbitrarily large entropy. Therefore, it follows that $h_A(T) = \infty$.

CHAPTER III

g-MEASURES

§1. General results and definitions

In this chapter, we consider the following set up.

Let S be a finite set $\{0,\ldots,n-1\}$ with the discrete topology. Let $X = \prod_{i=0}^{\infty} S$. X is given the product topology. With this topology, X is a compact metric space with metric d defined by $d(x,y) = 1/(k+1)$ where k is the largest integer such that $x_i = y_i$ for all $i < k$. Let T be the shift on X, that is, $(Tx)_i = x_{i+1}$. Then T is continuous. If A is an $n \times n$ matrix whose entries are zeros and ones, then X_A denotes the closed, T-invariant subset $\{x \in X : A_{x_i x_{i+1}} = 1 \text{ for all } i\}$. If T_A denotes T restricted to X_A, then T_A is said to be the one-sided subshift of finite type defined by the transition matrix A. We assume that no row of A is zero or else we could consider a smaller set S. By definition, T_A is topologically mixing if for all non-empty open sets U,V there exists N such that for all $i \geq N$, $T_A^{-i}U \cap V \neq \emptyset$. Clearly, this is equivalent to the condition A^M has all entries positive for some $M > 0$.

From now on, we will write T,X instead of T_A, X_A .

If $f \in C(X)$, let $\mathrm{var}_k(f) = \sup\{|f(x)-f(y)| : d(x,y) \leq \frac{1}{k+1}\}$

Note that $\mathrm{var}_k(f) \longrightarrow 0$ as $k \to \infty$. Let $\mathcal{B}$ denote the Borel σ-algebra of X, M(X) the set of Borel probability measures on X and M(T) the T-invariant members of M(X). Then M(X) is a convex, compact set in the weak *-topology and M(T) is a

closed, convex subset of $M(X)$.

Definition 3.1

Let $\phi \in C(X)$. Then $\mu \in M(T)$ is an equilibrium state for ϕ if

$$h_\mu(T) + \mu(\phi) = \sup_{m \in M(T)} \{h_m(T) + m(\phi)\}$$

where $h_\mu(T)$ denotes the entropy of T relative to μ, and $\mu(\phi)$ denotes $\int_X \phi \, d\mu$.

Let $P_T: C(X) \to \mathbb{R} \cup \{+\infty\}$ denote the pressure of T, which was defined in the most general context for continuous transformations on compact metric spaces by Walters in [19]. The pressure satisfies the following variational principle.

$$P_T(\phi) = \sup_{m \in M(T)} \{h_m(T) + m(\phi)\}$$

This gives an alternative definition of an equilibrium state for ϕ as a measure $\mu \in M(T)$ such that

$$P_T(\phi) = h_\mu(T) + \mu(\phi)$$

Note that when T is a one-sided subshift of finite type, the supremum in Definition 3.1, and therefore $P_T(\phi)$, is finite for all ϕ, since $h_m(T)$ is bounded by the topological entropy which is at most $\log|S|$.

Definition 3.2

If $\phi \in C(X)$, the Ruelle operator $\mathcal{L}_\phi$ on $C(X)$ is defined by

$$(\mathcal{L}_\phi f)(x) = \sum_{y \in T^{-1}x} e^{\phi(y)} f(y) \qquad (f \in C(X)).$$

Let $\mathcal{G} = \{g \in C(X) : g > 0 \text{ and } \sum_{y \in T^{-1}x} g(y) = 1 \text{ for all } x\}$

<u>Theorem 3.3</u> (see [11])

Let $g \in \mathcal{G}$ and $\mu \in M(X)$. The following are equivalent.

(i) μ is an equilibrium state for $\log g$.

(ii) $\mu \in M(T)$ and

$$E_\mu(f|T^{-1}\mathcal{B})(x) = \sum_{y \in T^{-1}Tx} g(y)f(y) \quad \text{a.e. } (\mu)$$

for all $f \in L^1(\mu)$.

(iii) $\mathcal{L}^*_{\log g}\mu = \mu$ ($\mathcal{L}^*_{\log g}$ denotes the dual of $\mathcal{L}_{\log g}$).

<u>Definition 3.4</u>

Let $g \in \mathcal{G}$. Then $\mu \in M(X)$ is a g-measure if μ satisfies the equivalent conditions in Theorem 3.3.

Note that if $g \in \mathcal{G}$, then $\|\mathcal{L}_{\log g} f\| \leq \|f\|$ for all $f \in C(X)$. Also, $\mathcal{L}_{\log g} 1 = 1$ and so $\mathcal{L}^*_{\log g}$ maps $M(X)$ into itself. Therefore, it is an easy consequence of the Shauder-Tychonoff Theorem [3, p.456] and (iii) of Theorem 3.3 that each $g \in \mathcal{G}$ has at least one g-measure. The following result gives a sufficient condition for the g-measure to be unique.

<u>Theorem 3.5</u> (see [18])

Let T be a topologically mixing one-sided subshift of finite type and let $g \in \mathcal{G}$. If $\sum_{k=0}^{\infty} \text{var}_k(\log g) < \infty$, then g has a unique g-measure.

§2. Examples of unique g-measures

In this section, we construct a class of functions which have unique g-measures, but do not satisfy the hypothesis of Theorem 3.5. The construction is similar to that of the Fisher model [7] and also that used by Hofbauer in [8].

We put $X = \prod_{i=0}^{\infty} \{0,1\}$ and take T to be the full shift. For $n \geqslant 1$, let $X_n = \{x \in X : x_i = 0 \text{ for } 0 \leqslant i \leqslant n \text{ and } x_n = 1\}$ and $X_0 = \{x \in X : x_0 = 1\}$. The X_n's together with the point (000...) form a partition of X.

Define $g \in \mathcal{G}$ as follows. Let $\{a_n\}_{n=1}^{\infty}$ be any decreasing sequence of real numbers such that $0 < a_n < 1$ for all n and $a_n \longrightarrow a$ as $n \to \infty$ where $0 < a < 1$. If $x_i = 0$ for all i, then $g(x) = a$. If $x \in X_n$ for some $n \geqslant 1$, then $g(x) = a_n$ and if $x \in X_0$, then $g(x) = 1 - g(0x_1x_2\ldots)$.

<u>Theorem 3.6</u>

Let $g \in \mathcal{G}$ be defined as above. Then g has a unique g-measure.

<u>Proof</u>. First, we introduce some notation. If $y_0 \ldots y_{n-1}$ is a finite sequence of zeros and ones and $x \in X$, then $(y_0 \ldots y_{n-1}x)$ denotes $z \in X$ where $z_i = y_i$ for $0 \leqslant i \leqslant n-1$ and $z_i = x_{i-n}$ for $i \geqslant n$. If $y \in \{0,1\}$, y^k denotes the finite sequence consisting of y repeated k times. We write $\mathcal{L}$ instead of $\mathcal{L}_{\log g}$.

It is enough to prove that for all $f \in C(X)$,

$$\mathcal{L}^n f(x) \longrightarrow c \quad \text{as } n \to \infty \quad \text{for all } x \in X \qquad (1)$$

where c is a constant depending on f. To see this, let μ

be any g-measure. Then $\mu(\mathcal{L}^n f) = \mu(f)$ by (iii) of Theorem 3.5. So, using (1) and the Dominated Convergence Theorem, we have $\mu(f) = c$ and hence μ is uniquely defined.

To prove (1), we may assume f is the characteristic function of a cylinder set

$$[w_0 \ldots w_{k-1}] = \{x \in X : x_i = w_i \ , \ 0 \leq i \leq k-1\}$$

since linear combinations of such functions are dense in C(X) and $\mathcal{L}$ is a bounded linear operator. Then

$$\mathcal{L}^{n+k} f(x) = \sum_{y \in T^{-(n+k)}x} g(y)g(Ty)\ldots g(T^{n+k-1}y)f(y)$$

$$= \sum_{y_0,\ldots,y_{n-1}} \{g(y_0 \ldots y_{n-1}x) \ldots g(y_{n-1}x)g(w_0 \ldots w_{k-1}y_0 \ldots y_{n-1}x) \ldots$$
$$\ldots g(w_{k-1}y_0 \ldots y_{n-1}x)\}$$

Suppose $w_i = 1$ for some i. Choose i, $0 \leq i \leq k-1$, such that $w_i = 1$ and $w_j = 0$ for $i < j \leq k-1$. Then if $0 \leq j < i$, $g(w_j \ldots w_{k-1}y_0 \ldots y_{n-1}x)$ does not depend on $y_0, \ldots, y_{n-1}$ or x. Put $\prod_{j=0}^{i-1} g(w_j \ldots w_{k-1}y_0 \ldots y_{n-1}x) = \lambda$. (If i = 0, take the empty product to be 1). Then

$$\mathcal{L}^{n+k} f(x)$$

$$= \lambda \sum_{y_0,\ldots,y_{n-1}} g(y_0 \ldots y_{n-1}x) \ldots g(y_{n-1}x) \prod_{j=i}^{k-1} g(w_j \ldots w_{k-1}y_0 \ldots y_{n-1}x)$$

$$= \lambda \sum_{y_0,\ldots,y_{n-1}} \{g(y_0 \ldots y_{n-1}x) \ldots g(y_{n-1}x)(1 - g(0^{k-i}y_0 \ldots y_{n-1}x)) \cdot$$
$$\cdot \prod_{j=1}^{k-i-1} g(0^j y_0 \ldots y_{n-1}x)\}$$

$$= \lambda(\mathcal{L}^{n+k-i-1}\chi_1(x) - \mathcal{L}^{n+k-i}\chi_2(x))$$

where χ_1 is the characteristic function of $[0^{k-i-1}]$, or 1 if $i = k-1$, and χ_2 is the characteristic function of $[0^{k-i}]$. It follows, therefore, that it is sufficient to consider the case when f is the characteristic function of $[0^k]$ $(k \geqslant 1)$.

Note that if $x,y \in X_k$ $(k \geqslant 0)$, then $\mathcal{L}^n f(x) = \mathcal{L}^n f(y)$ for all $n \geqslant 1$. Bearing this in mind, we will write $\mathcal{L}^n f(1)$, $\mathcal{L}^n f(0^k 1)$ instead of $\mathcal{L}^n f(1x)$, $\mathcal{L}^n f(0^k 1x)$. For $n \geqslant 1$ and $x \in X$,

$$\begin{aligned}\mathcal{L}^{n+k}f(x) &= (\prod_{i=1}^{n} g(0^i x))(\mathcal{L}^k f(0^n x) - \mathcal{L}^k f(1)) \\ &+ \sum_{i=1}^{n-1}(\prod_{j=1}^{n-i} g(0^j x))(\mathcal{L}^{k+i-1}f(1) - \mathcal{L}^{k+i}f(1)) \\ &+ \mathcal{L}^{n+k-1}f(1) \qquad (2)\end{aligned}$$

To see this, consider first the case $n = 1$. Then

$$\begin{aligned}\mathcal{L}^{k+1}f(x) &= g(0x)\mathcal{L}^k f(0x) + (1 - g(0x))\mathcal{L}^k f(1) \\ &= g(0x)(\mathcal{L}^k f(0x) - \mathcal{L}^k f(1)) + \mathcal{L}^k f(1) .\end{aligned}$$

So (2) holds for $n = 1$. Now suppose (2) holds for $n-1$.

$$\begin{aligned}\mathcal{L}^{n+k}f(x) &= g(0x)\mathcal{L}^{n+k-1}f(0x) + (1 - g(0x))\mathcal{L}^{n+k-1}f(1) \\ &= g(0x)(\mathcal{L}^{n+k-1}f(0x) - \mathcal{L}^{n+k-1}f(1)) + \mathcal{L}^{n+k-1}f(1) \\ &= g(0x)\Big\{\prod_{i=1}^{n-1} g(0^{i+1}x)(\mathcal{L}^k f(0^n x) - \mathcal{L}^k f(1)) \\ &\quad + \sum_{i=1}^{n-2}(\prod_{j=1}^{n-1-i} g(0^{j+1}x))(\mathcal{L}^{k+i-1}f(1) - \mathcal{L}^{k+i}f(1)) \\ &\quad + \mathcal{L}^{n+k-2}f(1) - \mathcal{L}^{n+k-1}f(1)\Big\} + \mathcal{L}^{n+k-1}f(1)\end{aligned}$$

$$= \prod_{i=1}^{n} g(0^i x)(\mathcal{L}^k f(0^n x) - \mathcal{L}^k f(1))$$

$$+ \sum_{i=1}^{n-2} \left(\prod_{j=1}^{n-i} g(0^j x)\right)(\mathcal{L}^{k+i-1} f(1) - \mathcal{L}^{k+i} f(1))$$

$$+ g(0x)(\mathcal{L}^{n+k-2} f(1) - \mathcal{L}^{n+k-1} f(1)) + \mathcal{L}^{n+k-1} f(1)$$

$$= \prod_{i=1}^{n} g(0^i x)(\mathcal{L}^k f(0^n x) - \mathcal{L}^k f(1))$$

$$+ \sum_{i=1}^{n-1} \left(\prod_{j=1}^{n-i} g(0^j x)\right)(\mathcal{L}^{k+i-1} f(1) - \mathcal{L}^{k+i} f(1)) + \mathcal{L}^{n+k-1} f(1)$$

This proves (2). Choose δ such that $0 < \delta < 1$ and $g \leq \delta$. Then

$$\left|\left(\prod_{i=1}^{n} g(0^i x)\right)(\mathcal{L}^k f(0^n x) - \mathcal{L}^k f(1))\right| \leq \delta^n$$

$$\longrightarrow 0 \quad \text{as } n \to \infty.$$

Therefore, it is enough to show $\mathcal{L}^{n+k} f(1)$ converges as $n \to \infty$, since then, given any $\varepsilon > 0$, we may choose N such that

$$n \geq N \implies \sum_{i=n}^{\infty} \delta^i < \varepsilon$$

$$\text{and} \quad |\mathcal{L}^{k+n-1} f(1) - \mathcal{L}^{k+n} f(1)| < \varepsilon .$$

Then if $n \geq 2N$,

$$\left|\sum_{i=1}^{n-1} \left(\prod_{j=1}^{n-i} g(0^j x)\right)(\mathcal{L}^{k+i-1} f(1) - \mathcal{L}^{k+i} f(1))\right|$$

$$\leq \sum_{i=1}^{N} \delta^{n-i} |\mathcal{L}^{k+i-1} f(1) - \mathcal{L}^{k+i} f(1)|$$

$$+ \sum_{i=N+1}^{\infty} \delta^{n-i} |\mathcal{L}^{k+i-1} f(1) - \mathcal{L}^{k+i} f(1)|$$

$$= \prod_{i=1}^{n} g(O^i x)(\mathcal{L}^k f(O^n x) - \mathcal{L}^k f(1))$$

$$+ \sum_{i=1}^{n-2} \left(\prod_{j=1}^{n-i} g(O^j x)\right)(\mathcal{L}^{k+i-1} f(1) - \mathcal{L}^{k+i} f(1))$$

$$+ g(Ox)(\mathcal{L}^{n+k-2} f(1) - \mathcal{L}^{n+k-1} f(1)) + \mathcal{L}^{n+k-1} f(1)$$

$$= \prod_{i=1}^{n} g(O^i x)(\mathcal{L}^k f(O^n x) - \mathcal{L}^k f(1))$$

$$+ \sum_{i=1}^{n-1} \left(\prod_{j=1}^{n-i} g(O^j x)\right)(\mathcal{L}^{k+i-1} f(1) - \mathcal{L}^{k+i} f(1)) + \mathcal{L}^{n+k-1} f(1)$$

This proves (2). Choose δ such that $0 < \delta < 1$ and $g \leq \delta$. Then

$$\left|\left(\prod_{i=1}^{n} g(O^i x)\right)(\mathcal{L}^k f(O^n x) - \mathcal{L}^k f(1))\right| \leq \delta^n$$

$$\to 0 \quad \text{as } n \to \infty.$$

Therefore, it is enough to show $\mathcal{L}^{n+k} f(1)$ converges as $n \to \infty$, since then, given any $\varepsilon > 0$, we may choose N such that

$$n \geq N \implies \sum_{i=n}^{\infty} \delta^i < \varepsilon$$

$$\text{and} \quad |\mathcal{L}^{k+n-1} f(1) - \mathcal{L}^{k+n} f(1)| < \varepsilon .$$

Then if $n \geq 2N$,

$$\left|\sum_{i=1}^{n-1} \left(\prod_{j=1}^{n-i} g(O^j x)\right)(\mathcal{L}^{k+i-1} f(1) - \mathcal{L}^{k+i} f(1))\right|$$

$$\leq \sum_{i=1}^{N} \delta^{n-i} |\mathcal{L}^{k+i-1} f(1) - \mathcal{L}^{k+i} f(1)|$$

$$+ \sum_{i=N+1}^{\infty} \delta^{n-i} |\mathcal{L}^{k+i-1} f(1) - \mathcal{L}^{k+i} f(1)|$$

$$\leq \sum_{i=N}^{\infty} \delta^i + \varepsilon \sum_{i=1}^{\infty} \delta^i$$

$$\leq (1 + \sum_{i=1}^{\infty} \delta^i)\varepsilon$$

and so

$$|\mathcal{L}^{n+k}f(x) - \mathcal{L}^{n+k}f(1)| \longrightarrow 0 \quad \text{as } n \to \infty \text{ for all } x \in X.$$

Using (2), we have

$$\mathcal{L}^{n+k}f(1) = \prod_{i=1}^{n} a_i(\mathcal{L}^k f(0^n 1) - \mathcal{L}^k f(1))$$

$$+ \sum_{i=1}^{n-1} \prod_{j=1}^{n-i} a_j(\mathcal{L}^{k+i-1}f(1) - \mathcal{L}^{k+i}f(1)) + \mathcal{L}^{n+k-1}f(1)$$

$$= \prod_{i=1}^{n} a_i \mathcal{L}^k f(0^n 1) + (\prod_{i=1}^{n-1} a_i - \prod_{i=1}^{n} a_i)\mathcal{L}^k f(1)$$

$$+ (\prod_{i=1}^{n-2} a_i - \prod_{i=1}^{n-1} a_i)\mathcal{L}^{k+1}f(1) + \ldots + (1-a_1)\mathcal{L}^{n+k-1}f(1)$$

If we put $\alpha_1 = (1-a_1)$, $\alpha_n = \prod_{i=1}^{n-1} a_i - \prod_{i=1}^{n} a_i$ $(n \geq 2)$ and

$\beta_n = \prod_{i=1}^{n} a_i \mathcal{L}^k f(0^n 1)$, then $\alpha_n > 0$, $\beta_n > 0$ $(n \geq 1)$,

$\sum_{n=1}^{\infty} \alpha_n = 1$, $\sum_{n=1}^{\infty} \beta_n < \infty$ and

$$\mathcal{L}^{n+k}f(1)$$

$$= \beta_n + \alpha_1 \mathcal{L}^{n+k-1}f(1) + \alpha_2 \mathcal{L}^{n+k-2}f(1) + \ldots + \alpha_n \mathcal{L}^k f(1)$$

Therefore, by the Renewal Theorem [5, p.330], $\mathcal{L}^{n+k}f(1)$ converges. This completes the proof.

Note that $\mathrm{var}_k(\log g) = \log a_k/a$ for $k \geqslant 1$. So, clearly we may choose $\{a_n\}$ so that $\sum_{k=0}^{\infty} \mathrm{var}_k(\log g)$ is infinite, but g has a unique g-measure. Indeed, we may choose $\{a_n\}$ so that $\mathrm{var}_k(\log g)$ converges as slowly as we like, so, clearly, the uniqueness of g-measures does not depend solely on the rate of convergence of $\mathrm{var}_k(\log g)$.

REFERENCES

1. L. M. Abramov, 'Metric automorphisms with quasi-discrete spectrum', A.M.S. Translations Series 2, Vol.39, 37-56.

2. L. M. Abramov and V. A. Rokhlin, 'The entropy of a skew-product of measure-preserving transformations', A.M.S. Translations Series 2, Vol.48, 255-265.

3. N. Dunford and J. T. Schwartz, 'Linear Operators I', Interscience, New York.

4. M. H. Ellis and N. A. Friedman, 'Sweeping out on a set of integers', Proc. A.M.S., Vol.72, No.3, 509-512.

5. M. H. Ellis and N. A. Friedman, 'Subsequence generators for ergodic group translations', Israel Journal of Maths., Vol.31 (1978), No.2, 115-121.

6. W. Feller, 'An Introduction to Probability Theory and Its Applications', Vol.1, 3rd. edition, Wiley, 1968.

7. M. Fisher, Physica 3 (1967), 255-283.

8. F. Hofbauer, 'Examples for the non-uniqueness of the equilibrium state', Trans. A.M.S., Vol.228 (1977), 223-241.

9. M. Keane, 'Strongly mixing g-measures', Invent. Math. Vol.16 (1972), 309-324.

10. A. G. Kushnirenko, 'On metric invariants of entropy type', Russian Math. Surveys, Vol.22 (1967), No.5, 53-61.

11. F. Ledrappier, 'Principe variationel et systèmes dynamiques symboliques', Z. Wahrscheinlichkeitstheorie verw. Gebiete, Vol.30 (1974), 185-202.

12. D. Newton and E. Krug, 'On sequence entropy of automorphisms of a Lebesgue space', Z. Wahrscheinlichkeitstheorie verw. Gebiete, Vol.24 (1972), 211-214.

13. B. S. Pickel, 'Some properties of A-entropy', Mat. Zametki, Vol.5 (1969), 327-334.

14. V. A. Rokhlin, 'Lectures on the entropy theory of transformations with invariant measure', Russian Math. Surveys, Vol.22 (1967), No.5, 1-52.

15. A. Saleski, 'Sequence entropy and mixing', Journal of Math. Analysis and Applications, Vol.60 (1977), 58-66.

16. P. Walters, 'Some invariant σ-algebras for measure-preserving transformations', Trans. A.M.S., Vol.163 (1972), 357-367.

17. P. Walters, 'Affine transformations and coalescence', Math. Systems Theory, Vol.8 (1974), No.1, 33-44.

18. P. Walters, 'Ruelle's operator theorem and g-measures', Trans. A.M.S., Vol.214 (1975), 375-387.

19. P. Walters, 'A variational principle for the pressure of continuous transformations', Amer. J. Math., Vol.97 (1975), 937-971.